Theory and Simulation of Scanning Gate Microscopy Applied to the Investigation of Transport in Quantum Point Contacts

Wojciech Szewc

UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE DE PHYSIQUE ET CHIMIE-PHYSIQUE

INSTITUT DE PHYSIQUE ET CHIMIE DES MATÉRIAUX DE STRASBOURG

THÈSE présentée par :

Wojciech SZEWC

soutenue le : **18 septembre 2013**

pour obtenir le grade de : **Docteur de l'Université de Strasbourg**

Discipline/ Spécialité : Physique

Theory and Simulation of Scanning Gate Microscopy

Applied to the Investigation of Transport in Quantum Point Contacts

THÈSE dirigée par :

Prof.	**Rodolfo A. JALABERT**	Université de Strasbourg
Dr.	**Dietmar WEINMANN**	Université de Strasbourg

RAPPORTEURS :

Prof.	**Jean-Louis PICHARD**	CEA-Saclay
Dr.	**Marc SANQUER**	CEA-Grenoble/INAC

AUTRES MEMBRES DU JURY :

Prof.	**Bernard DOUDIN**	Université de Strasbourg

Summary

This work is concerned with the theoretical description of the Scanning Gate Microscopy (SGM) in general and with solving particular models of the quantum point contact (QPC) nanostructure, analytically and by numerical simulations.

Scanning Gate Microscopy is a local probe technique, which provides much more detailed information on the electron motion inside a nanostructure than traditional transport measurements. The SGM experiments on quantum point contacts and other structures reveal many interesting features; the SGM signal close to the QPC forms wide lobes, narrow branches away from it, and interference fringing patterns superimposed on both of these. A generally applicable theory, allowing for unambiguous interpretation of the results, is still missing.

Working within the scattering approach of quantum conductance and the Lippmann-Schwinger perturbation theory, we have developed a systematic theory, describing the response from the SGM system with a non-invasive probe (tip). The non-invasiveness is understood as posing a weak and localised obstacle for the moving electrons, as compared to their Fermi energy and Fermi wavelength, respectively. Additional insight can be obtained by simulating the SGM signal numerically, for which we use the Recursive Green Function algorithm, and also by numerically evaluating the SGM images from the analytical solutions obtained within specific models. The latter results allow also going beyond the non-invasive regime.

From the formulae we present, it follows that the SGM signal images the matrix elements of the probe's potential taken with two counter-propagating eigenstates of the transmission operator. It cannot be a freely chosen basis of the eigenstates, however, but one complying with the requirements set by entire scattering matrix. The SGM signal can be related with the densities of current and of charge at the Fermi level only in some special, symmetric cases. If the structure is time-reversal invariant (TRI) and has a quantised conductance curve, then the SGM signal in the lowest conductance plateau region can be related to the charge density at the Fermi energy. If it addi-

tionally has a central (four-fold) spatial symmetry and the leads accessing it are wide, then the signal away from the structure's exit is also related with the current density. Moreover, if the SGM probe is sufficiently local, then the plateau-region SGM signal from any TRI system can be decomposed as a product of the charge densities of the scattering states incoming from one lead and the scattering states incoming from the opposite one, both taken at the Fermi energy (one of these sets will be completely unoccupied).

The importance of the quantum point contact structure comes from its conductance quantisation. Through the expressions we derived, we predict that the SGM images obtained in the quantised conductance plateau region will be of second order in the probe potential strength, and for a sufficiently adiabatic structure will have no interference fringes, even with ideal coherence. It can have, as we will see from the computer simulation results, a chequerboard-like interference pattern close to the scatterer structure, when the tip potential is strong enough or when the structure is non-adiabatic. The non-adiabatic structures give irregular patterns, unless the system, apart from the central scatterer-structure, has no lateral boundaries confining it. It can be conjectured, that the same could apply in a branched signal, even at the lack of the confining walls.

The SGM signal in the step region of the conductance, or for a non-quantised conductance curve, will be of first order in the probe's potential and will have an interference fringing pattern, oscillating around zero.

Our theory is applicable to the Scanning Gate Microscopy of any coherent structure, the magnetic field, finite temperature effects and the influence of the disorder potential can be included in this theoretical framework, allowing for analytical, or at least numeric solutions. Neither the theory, nor the simulations, does not take into account the electron-electron interactions, whose inclusion into any of them remains a challenge for the future.

This dissertation is organised as follows. After defining and briefly reviewing the field of our interest in Chapter 1, we introduce the scattering theory of quantum conductance in Chapter 2. Special attention is paid to the issues related to the S-matrix and the bases in which we later perform our calculations and present the results.

In Chapter 3, we derive perturbative formulae for the conductance change, of the first and the second order in the tip's potential. This is followed by a detailed discussion of the obtained expressions and their basic consequences. The evaluation of certain integrals appearing during the derivation is explained in Appendix A, while certain simple manipulations for determining the current density formulae are placed in Appendix B.

In Chapter 4, we review the most important models for analytically solv-

ing the transport problem for the quantum point contacts. Based on the solutions of two of them, with quite opposite properties – the ideally adiabatic saddle-point model and the highly non-adiabatic abrupt hard-wall model of the QPC, we evaluate and compare the SGM response and the local densities of current and Fermi-energy charge. Finally in Chapter 5 we investigate a few hard-wall models of the QPC. We plot the resulting transmission vs. energy curves and SGM maps and find a few typical patterns of the SGM signal behaviour, together with the conditions in which they occur.

A concise statement of conclusions and perspectives is placed in the closing Chapter 6.

Related publications

- C. Gorini, R. A. Jalabert, W. Szewc, S. Tomsovic, and D. Weinmann, *Theory of scanning gate microscopy,* Phys. Rev. B, 88:035406, Jul, 2013

- R. A. Jalabert, W. Szewc, S. Tomsovic, and D. Weinmann, *What is Measured in the Scanning Gate Microscopy of a Quantum Point Contact?* Phys. Rev. Lett., 105:166802, Oct, 2010

Acknowledgements

I would like to thank everyone who helped me in creating this thesis and in the studies which led to it. I would like to thank my thesis advisors, Rodolfo Jalabert and Dietmar Weinmann, for everything I could learn while working in the Mesoscopic Physics Team of IPCMS, for their guidance and for their enormous patience.

I want to thank our collaborators from other research groups, notably to Jean-Louis Pichard, Gabriel Lemarié and Adel Abbout from CEA Saclay, I am also grateful to Ferdinand Evers from KIT, Germany, for giving me the opportunity to make my first external seminar, and to Dr. Maya Lucas, also from KIT, for an interesting discussion about the experimental aspects of the matter I investigated theoretically. I also thank my colleagues from IPCMS, especially to Guillaume Weick and Christian Mény. At the same time, I would like to thank Élisabeth Olory from the Faculty of Physics and Engineering, who supervised my work as a *moniteur*, and all my colleagues with whom I worked there.

I would like to express my gratitude to the members of jury for spending their time and effort to assess and evaluate my thesis work: to Marc Sanquer, Bernard Doudin, and, once more, to Jean-Louis Pichard.

I gratefully acknowledge the financing from University of Strasbourg, which I enjoyed throughout the greater part of my studies.

I feel tremendously indebted to my friends who supported me so cheerfully, and to my girlfriend, Marta, for her love and patience, which helped me so much during my long work. Finally, I would like to thank my family, especially my parents for their patience when they had not seen me at home for long months.

Contents

Chapter 1

Introduction

The second half of the twentieth century saw an enormous development in technology and science, an important part of which was driven by the everlasting trend towards miniaturisation in electronics, optics and materials' engineering and physics. As a fruit of this process came, starting in the 1980s, the techniques of manipulating and structuring matter at a nanometre scale, known under the name of Nanotechnology. Quantum effects can already become significant at these sizes, such as quantum confinement, quantum interference or single electron charging effects. The sub-field of Condensed Matter Physics investigating these phenomena is called Mesoscopic Physics.

Whereas the domain of Nanotechnology is usually defined from the engineering (fabrication) perspective, as embracing all artificial structures whose extent in at least one spatial dimension ranges from hundreds of nanometres down to the size of single atoms, the mesoscopic scale is determined with regard to the relevant physics, to be:

- small enough to observe the quantum-mechanical effects, i.e. on the order of electronic phase coherence length or smaller; and
- large enough for disregarding the details of the individual microscopic building-blocks and describing their behaviour statistically – which means that a sufficiently large number of atoms should be incorporated.

The progress in both these fields was tightly interwoven, with the scientific research being crucial for further advancement of the technology, and the technological side furnishing science with improved tools and even new, interesting systems to investigate. Downscaling in commercial electronics has already reached the nano-scale, but it cannot continue into the mesoscopic regime without a serious reshaping of the technology. (Transverse trends

have also started – most notably the wide use of the multi-core processors since 2000s.) Naturally, this reshaping must be based on a thorough fundamental understanding of the mesoscopic systems operation. For this, as well as for the design and characterisation of future devices, suitable imaging techniques will be indispensable.

In this work, we study one of the most promising among such techniques, the Scanning Gate Microscopy (SGM). Its particular advantages follow from the fact that it brings the spatially-resolved information, yielded by the Atomic Force Microscopy (AFM), into the domain of transport measurements. Based on scattering approach to quantum conductance, we develop a systematic theory describing the response from a non-invasive Scanning Gate Microscope (SGM). Further, a special attention will be paid to one particular nanostructure, the Quantum Point Contact (QPC). Patterned from a layer of 2-Dimensional Electron Gas (2DEG), it is one of the simplest conceivable structures and also the one to have been most extensively investigated with the means of Scanning Gate Microscopy. Its conductance characteristics and SGM maps reveal some very interesting phenomena, which we concisely review further in this chapter. We begin by introducing the technique of Scanning Gate Microscopy.

1.1 Scanning Gate Microscopy

Scanning Gate Microscopy is an experimental technique in which a charged probe is held in some place above the investigated sample, while conductance between source and drain (Ohmic contacts) is measured and recorded [1] – see fig. 1.1 for schema. The values obtained with many different positions of the probe are then gathered to create a spatial map of the conductance $G(\mathbf{r}_0)$ – with $\mathbf{r}_0$ being the probe position – or of the conductance change $\Delta G(\mathbf{r}_0)$, if the value with no probe present is used for reference.

As the probe, Scanning Gate Microscopy uses a metallised AFM tip [2]. The tip couples capacitively to the mobile electrons of the investigated structure (the electrons are usually buried inside the structure), so its influence can be modelled simply as a potential. Its role is to act as an additional, movable gate, which, by perturbing the system locally, influences a global property, the conductance. This influence can be described in terms of backscattering of the electron waves, which prevents a certain fraction of the electron flux from making its way to the drain electrode, leading to a negative conductance change. The reverse process is also possible, leading to a (weak) positive change in conductance when the backscattered waves interfere to cancel some other reflected electron wave – see below.

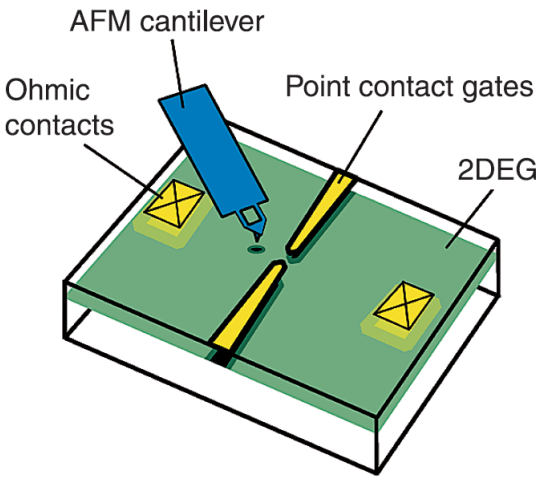
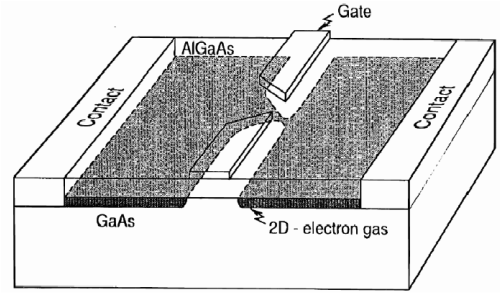


Figure 1.1: Schematic depiction of an SGM setup for investigating a 2DEG-structure containing a quantum point contact. Figure reproduced from [1].

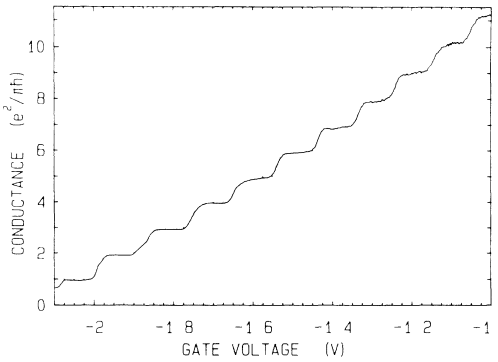
Although we work solely with the linear-response conductance here, other transport coefficients, like finite-bias conductance [1] or Hall resistance [3], can be mapped in the same way. The measurements are performed at Helium temperatures, so that the basic requirement of the coherence length exceeding the sample size could be easily satisfied. Depending on the desired regime of operation of the investigated structure, other characteristic lengths must be also controlled, like the mean free path or Fermi energy. These are influenced not only by the temperature, but also by the material properties and external fields.

Since its introduction in the mid-1990s [4], Scanning Gate Microscopy has been used to investigate many hallmark systems of the mesoscopic physics (for a brief review please refer to [5]), such as:

- quantum point contacts defined in two-dimensional electron gases [1, 6–18];
- carbon nanotubes [19];
- quantum wires [20];
- quantum dots defined in various structures, like
  - semiconductor 2DEGs [21, 22],
  - carbon nanotubes [23],
  - semiconductor nanowires [24, 25],
  - graphene [16];



(a) A schema of a QPC defined in the plane of GaAs/AlGaAs junction by a split-gate which controllably depletes the 2DEG underneath. Source: [33].



(b) The conductance quantisation in a QPC. The gate voltage parametrizes QPC's width, so the curve is equivalent to $G(W_{\mathrm{QPC}})$. Source: [33].

Figure 1.2

– Aharonov-Bohm rings [26–28];
– Hall bars [3];
– quantum Hall effect edge states [29];
– quantum billiards [30];
– bilayer graphene [31].

Out of the many, we place a special focus on the quantum point contacts.

1.2 Quantum Point Contacts

A quantum point contact is a short and narrow passage, of width on the order of the electron's Fermi wavelength, joining two wide electrically conducting regions [32].

It can be realised in various systems, among which by far the most important is the configuration in which the QPC is defined within the plane of a semiconductor heterojunction containing two-dimensional electron gas, see fig. 1.2(a) for a typical example. Owing to the low disorder rates, modulation-doped GaAs/AlGaAs is the standard material of choice.[1] A crucial property of the quantum point contacts is their conductance quantisation, discovered

[1]Experiments on InGaAs/InAlAs QPCs were also conducted [11]. In addition, outside the semiconductor realm, QPCs were obtained in metallic break-junctions [34].

at the end of the 1980s [33, 35] – when the QPC is operated in the ballistic transport regime, its linear-response conductance changes in a step-like way when its width W_{QPC}, or Fermi energy E_F, is varied:

$$G = G_0\, N(W, E_F)\,, \qquad N \in \mathbb{Z}\,,\ G_0 = const\ , \tag{1.1}$$

giving rise to a series of plateaus and ramps in the $G(W_{\mathrm{QPC}})$ and $G(E_F)$ curves, see fig. 1.2(b). Each rise of the conductance to a new plateau is caused by populating a new transverse mode inside the QPC's constriction, which since then acts as a perfectly transmitting, independent transport channel [36, 37] (see Ch. 4).

For an excellent review of the pre-SGM quantum point contacts research, see [38]. Now, we would like to recapitulate some of the findings from the Scanning Gate Microscopy of QPC transport available in the literature.

1.3 Most important experimental findings

1.3.1 Lobes in the angular structure

The most basic feature of the SGM signal, found already in the early experiments [1], is that it exits from the QPC in the form of wide, smoothly opening lobes, whose angular shape resembles the shape expected from the electronic wave functions protruding from a narrow constriction, see fig. 1.3, upper panel.[2] The images are taken with the system set to the first, second and third plateau of the quantised conductance.

Since the N-th plateau contains contributions due to the N lowest constriction modes, it is a natural postulate to subtract the previous, $(N-1)$-th, plateau's signal, as to isolate the signal from the highest open mode [1]. The single-mode SGM signal components obtained with this method are shown in the lower panel of figure 1.3, supplemented with simulated shapes of the corresponding mode wave functions in the middle-parts. There is a good qualitative agreement between the two. The correctness of this approach is also supported, with some minor restrictions, by the theory we introduce in Chapter 3 (see Ch. 3, subsec. 3.4.6).

1.3.2 Branching

In the regions further away from the QPC, the SGM signal no longer shapes as smooth lobes, but starts splitting into narrow branches instead [6] – see

[2] We do not provide the scale in fig. 1.3. The values of Δg are all negative, with the dark regions corresponding to zero signal.

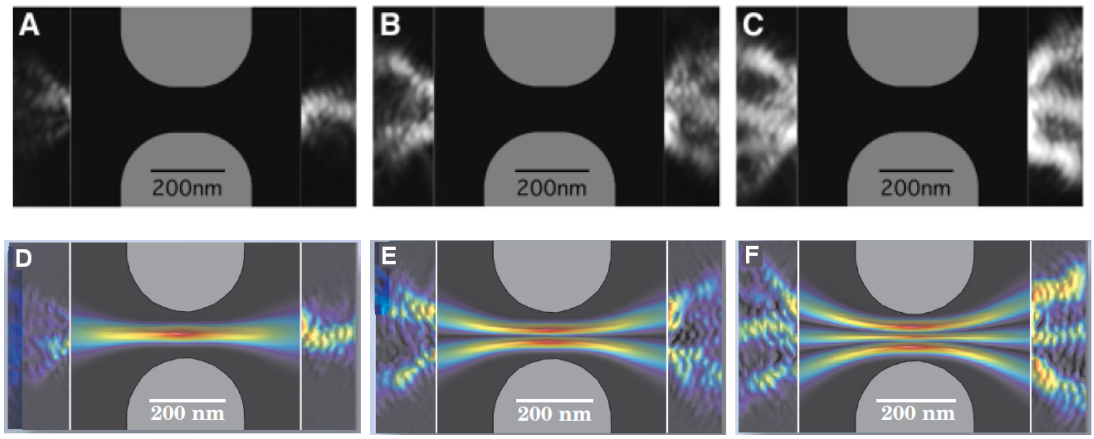


Figure 1.3:
Upper panel: Measured SGM signal with the QPC set to the first (A), second (B) and third (C) conductance plateau. The middle parts are not accessible to the tip-scan in the split-gate quantum point contacts, and show the shape of the spit-gate only. (Figures from [1].)
Lower panel: The first-mode (D), second-mode (E) and third-mode (F) SGM signal, isolated from the plateau-signals by subtracting the lower-plateau image from each one (except first). The middle parts show simulated shape of the modulus-squared of the constriction wave functions. (Figures from [39].)

figure 1.4. This effect is attributed to the disorder potential, or, more precisely, to its smooth component (due to the charged dopant atoms out of the junction plane) which, by small-angle scattering, can either disperse or focus the electron wave functions [6]. The focussing gives rise to the high-signal segments between two neighbouring branching events, by analogy to optics called caustics [13].

The typical length of such a segment shows a clear dependence on the electronic mean free path l_{mfp}, constituting, however, only a fraction of it [13] (fig. 1.4). The branching pattern is also quite robust against laterally shifting the QPC opening through asymmetric biasing of the two halves of the split-gate – the intensity of branches is prone to change, but their location varies very little [13]. The whole pattern, including the fringing (see below), is even more resistant to: thermal cycling of the sample, change of the AFM tip, and to change of the overall split-gate voltage, as long as the conductance stays on the same plateau [18]. This shows the importance of the medium through which the electrons travel for the shape of the SGM signal. When the conductance passes to a higher plateau, on the other hand, then the

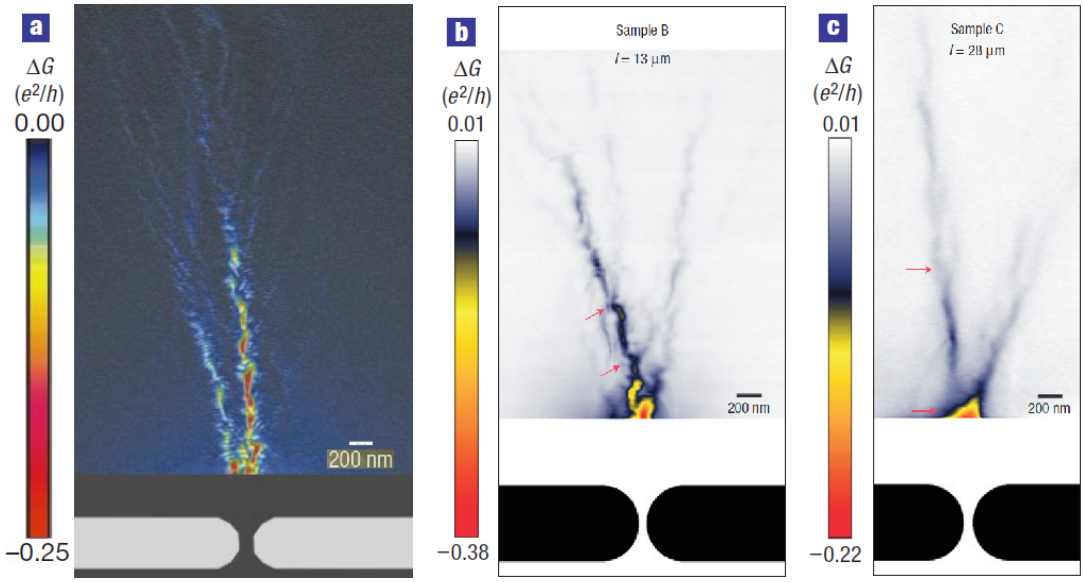


Figure 1.4: Maps of the SGM signal in a wide region neighbouring a QPC, taken with the system tuned to the first conductance plateau. The images clearly show branching patterns. The (b) and (c) images were taken for samples with increased electron mobility, and are added in order to illustrate two points: the dependence of the length of signal segments between neighbouring branching events (red arrows point to examples of such) on the mean free path, here denoted by l; and the disappearance of fringing pattern in conductance-plateau signal, which takes place in high-mobility samples (see below). (Image (a) is adapted from [6], images (b) and (c) – from [13].)

image develops new branches on top of the already existing ones [18].

Even when the branching is present, the average signal intensity scales with distance from the QPC opening, r_0 (we take the QPC opening as the reference position), as r_0^{-2} [7], which agrees well with the electron wave backscattering explanation of the SGM response.

The evolution of the signal from the lobed pattern to the branched one as the tip is moved away from the QPC shows the interplay of the influences on the signal: of the QPC structure from which it is "ejected", and of the environment to which it is "injected".

1.3.3 Fringes

A very important and surprisingly diverse attribute of the SGM signal is its fringing pattern. Most SGM images, whether branched or lobed, are, like

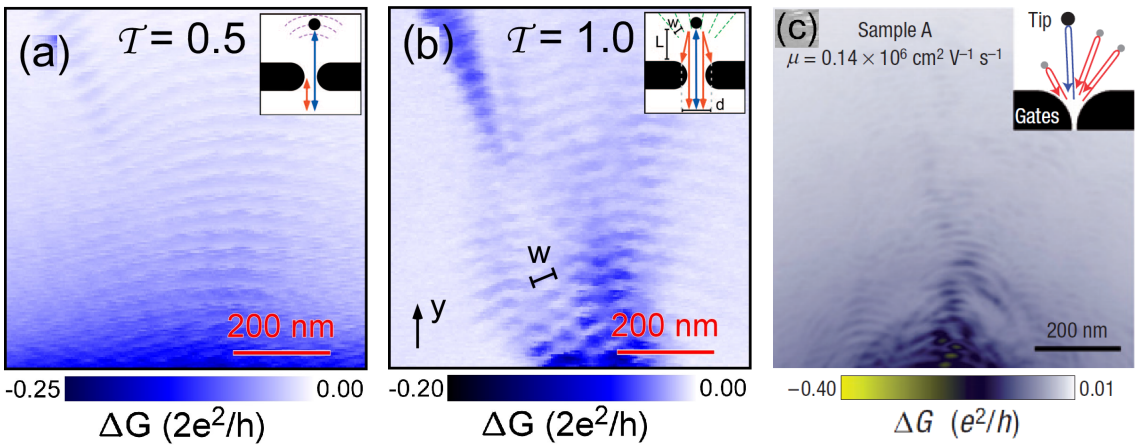


Figure 1.5: Example SGM maps with different types of fringing pattern:

(a) fringes due to backscattering from the tip and the QPC – fringes have a regular circular shape (high-mobility sample at 0.35K temperature, adapted from [14] and its EPAPS supplement);

(b) fringes due to backscattering from the tip once and twice – chequerboard pattern visible, with lateral spacing denoted by w (high-mobility sample at 0.35K temperature, adapted from [14] and its EPAPS supplement);

(c) fringes due to backscattering from the tip and an impurity – fringes have an irregular shape, cross each other, disappear and re-appear, precursors of branching can be seen (low-mobility sample at 4.2K temperature, adapted from [13]).

The insets illustrate the mentioned mechanisms of fringes formation.

the ones in the figs. 1.3 (A-C) and 1.4 (a), decorated with fringes, stretching perpendicularly to the presumed direction of electron motion and spaced by half of the local Fermi wavelength. Below, we review the experimentally observed fringing types, listing them by mechanism of their creation. The fringes always result from interference of the electron waves back-scattered by the tip towards the source electrode with the waves scattered at least once by some other object.

Backscattering from the tip and QPC

In the ramp region of the quantised conductance, where one of the constriction modes is only partially open, one interference path is provided by a wave partially reflected by the QPC, while another one is formed by its transmitted part, once it gets backscattered by the tip and re-transmitted by the QPC, see fig.1.5 (a), inset.

A measured pattern, attributed to this kind of mechanism, is shown in fig. 1.5 (a). The fringes created by such an interference have a circular shape. Ideally, their amplitude falls with distance like the overall signal, as r_0^{-2}. The pattern can be spoilt by the disorder potential, which can scatter the waves – introducing many different interferences, or deflect them – distorting the constant phase difference lines, and also by thermal dephasing. Hence, this type of fringes is visible only in high-purity samples in very low temperatures, and fades out within the thermal coherence length l_{th} from the QPC opening.

This mechanism is capable of producing a positive conductance change, by destructive interference between the two paths. The fact that it is not seen in figure 1.5 (a) indicates that other processes must be present.

Backscattering from the tip once and twice; chequerboard pattern

Still, in most experiments to date, it is the plateau-signal which has been probed.

In this case, there is no partial reflection at the QPC and first of the previously described paths does not form. The role of the base interference path is now overtaken by the second of these paths, i.e. source-QPC-tip-QPC-source, while the higher-order ones are created by the waves making at least one additional trip from the tip to the QPC-defining constriction walls and back to the tip. The lowest of those dominates over the rest, since each round trip of this kind diminishes the flux contained in the interfering wave by r_0^{-1} at least, r_0^{-2} at most.[3] There are, actually, two such paths, since each QPC wall offers one. Both interfere, superimposing their constant phase difference lines[4] on the circular ones of the base path, thus creating a *chequerboard-like pattern* in the SGM image. A relevant SGM map is shown in fig. 1.5 (b), and the mechanism is depicted in the inset.

Like the previous one, this fringing is limited by the thermal length, but due to the involvement of second-order paths, its decay is faster.

[3] Depending on the curvature of the constriction wall at the reflection point. Fitting of the fringes visibility decay in [14] indicates a value of r_0^{-2}.

[4] A family of hyperbolas with foci in the reflection points (green dashed lines in the fig. 1.5 (b), inset).

In principle, this pattern is present in the ramp-region signal as well, but is very feeble as compared to the basic one (cf. weak angular structure in fig. 1.5 (a)).

Backscattering from the tip and impurity

The third mechanism, see inset to fig. 1.5 (c), requires the presence of some hard scattering centres, like lattice defects or ionized dopant atoms close to the plane of the 2DEG. Each of such *impurities* introduces an interference path of its own, source-QPC-impurity-QPC-source, with the phase accumulation independent of the tip position. Their interference with the usual source-QPC-tip-QPC-source path gives a fringing pattern whose shape, normally circular, is hugely distorted by the influence of disorder (this includes the already discussed branching effect). Due to the measurement conditions, it is this type of fringes which is seen in all the early experimental results [1, 6–9], for example in fig. 1.4 (a). The mechanism itself has been studied in detail by adding an artificial, mirror-like scatterer, to create a strong and controllable interference path [10].

Because of the thermal dephasing, an impurity contributes to the interference pattern only if its distance from the QPC opening r_{imp} differs from the tip-opening distance, r_0, by no more than the thermal length: $r_0 - l_{\mathrm{th}} < r_{\mathrm{imp}} < r_0 + l_{\mathrm{th}}$. Therefore, the fringes in SGM images persist even at distances from the QPC beyond the thermal coherence length [6, 40], but disappear in high-purity samples (compare fig. 1.4 (a) vs. (b-c)) [13] unless the temperature is very low [14]. As the scanning tip is moved away from the QPC, the set of "active" impurities changes, which can lead to variation of phase and amplitude of the fringes. An example pattern can be found in fig. 1.5 (c).

Like in the first of the fringing mechanisms above, this one is, at least in principle, capable of producing positive SGM signal.

Fringes spacing, ring pattern

In each of these mechanisms, the fringes spacing should equal half of the local Fermi wavelength, which reflects the simple fact that the phase-path is twice the tip-QPC distance. It has been even proposed to use the spacing as a measure of the local electron density [8].

However, recent experiments in high-purity samples have found an interesting modulation of the fringes spacing as a function of the distance from the QPC [18], which escapes the simple picture presented above.

1.4 Interpretation of the SGM signal

1.4.1 Standard interpretation

As we mentioned before, the change of conductance, which constitutes the SGM signal, is caused by scattering of the electrons from the source-drain flux and turning them back towards the source. Based on the plausible argument that the greater the original flux at $\mathbf{r}_0$, the greater this back-scattering should be, it was proposed to interpret the SGM maps as images of electron flow [1]. This interpretation drew a great interest towards the SGM technique, since accessing the electron flow experimentally would mean touching upon the very fundamentals of quantum science, the relations between position and momentum of microscopic particles (not to mention the immediate utility for electronic devices).

A number of simulation results, both classical [41] and quantum [10, 12, 41–43], tend to support it – the conductance change and current density maps they yielded resemble each other to a reasonable degree. Still, certain experiments investigating small Aharonov-Bohm rings, combined with simulations, suggested a closer connection of SGM signal to the local density of states (LDOS) [27,28]. A generally applicable theory allowing for unambiguous interpretation of the SGM signal with arbitrary tip voltage strength is still missing.

Two conceptual issues about this standard, electron-flow interpretation, are worth mentioning. First, the above arguments are essentially semiclassical. The ubiquitous fringing pattern, obviously not shared by the electron current density, is explained within this picture by "dressing" of the expected electron paths with phase and considering their possible interferences (see previous section). Second, it tacitly assumes that, before hitting the tip, the electrons move in their usual fashion, as if the tip was not present. This point of view is local, which is not necessarily compatible with a fully quantum-mechanical picture, where the electron motion is described through the wave function – a global field, after all. Whether the SGM signal is a local or non-local quantity is one of the issues recently researched (see [44]).

1.4.2 The perturbative theory of SGM

As a first step towards the general theory fully describing the SGM response, a perturbative approach has been proposed [44, 45], leading to a clear-cut relationship between the infinitesimal conductance change due to a mild tip potential and the original (unperturbed) scattering properties and wave functions of the investigated system. Expressions up to second order in the tip

potential have been given. Presenting and discussing this theory is the main topic of this work.

A natural concern arises, whether the perturbative approach, operating with the notion of weak, non-invasive probe, can be useful in the description of the real-life experiments, where the tip's potential is strong enough to create the 2DEG depletion disc of a size larger than the Fermi wavelength. We will address this question analytically and also by the means of numeric simulations, which can reproduce the tip potential of arbitrary strength. It seems that the experimentally used tips do not meet the perturbation-theory criteria, making instead, due to their very large size, more of a classical obstacle for the electrons. Moreover, it is not impossible to employ milder [31] or spatially smaller tips [25] in the experimental measurements. Developing a truly non-invasive probe would be important for the sake of observing a completely quantum behaviour.

Naturally, the idealisations we use to set up our theoretical framework do not restrict to the choice of perturbation approach. We would like to mention the most crucial ones.

First of all, we work within the independent-particle approximation. Thus, the very interesting questions of electron-electron interactions (see [15,46,47]) are not discussed in here. We expect, nevertheless, that our theoretical considerations can lay a reliable basis for possible interacting models in the future, or at least for identifying which features of the experimental results could be traced back to the possible influence of electronic interactions.

Secondly, we work with no external fields. Our theory can be generalised to cater for weak magnetic fields, if appropriate treatment of completeness and orthonormality relations are established. The electric field (i.e. finite bias) could be incorporated in a simplistic way – see Ch. 2, eq. 2.52 – which would not make justice to all relevant processes, including the electron-electron scattering.

All our results are for zero temperature. Finite temperature effects are not targeted at with our theory, neither with the simulations, but are, in principle, tractable – see e.g. eq. (2.55) of Ch. 2 – and give a wealth of interesting phenomena. For a treatment of the thermal phenomena in a similar physical context see [48, 49]. Because of their large experimental relevance, we will occasionally comment on the influence of the finite temperature. Finally, the electronic disorder has been given very limited room throughout this work, nevertheless we mention its influence, as it plays an important role in the general context of interpreting the SGM images.

Chapter 2

Theory of quantum conductance

In this chapter we introduce the scattering approach to quantum conductance and build up our theoretical paradigm. We do it step by step, as shown schematically in figure 2.1(a-c): starting from the wave-mechanical description of electron motion in a simple two-dimensional wire and successively adding the elements needed for describing the conductance through a QPC, or any other coherent structure. We first introduce the electron states suitable for the description in each of these situations: the wire states and the scattering states (secs. 2.1-2.2). A particular attention will be paid to their mutual relations, expressed in terms of the transmission and reflection amplitudes which constitute the S-matrix (sec. 2.3), an entity of vivid importance throughout all of this work. Then we fill the states with non-interacting electrons (sec. 2.4) and determine the conductance within what is known as the *Landauer approach* (sec. 2.5). We will eventually work in the linear response approximation, so in the end solely the states at the Fermi energy will be of relevance. In the last section (2.6), we briefly discuss the way of calculating the S-matrix. The final object of our interest, the signal from an SGM system – fig. 2.1(d) – will be introduced only in the next chapter.

2.1 Electron states in a perfect wire

Let us consider the simplest conceivable 2D structure: an ideal wire – a piece of disorder-free material stretching infinitely along the x-coordinate ('longitudinal direction'), bounded along the y-coordinate ('transverse direction') and symmetric under translation in x. Such a wire will serve us as a model for a lead joining the proper structure with a reservoir, but it could also be

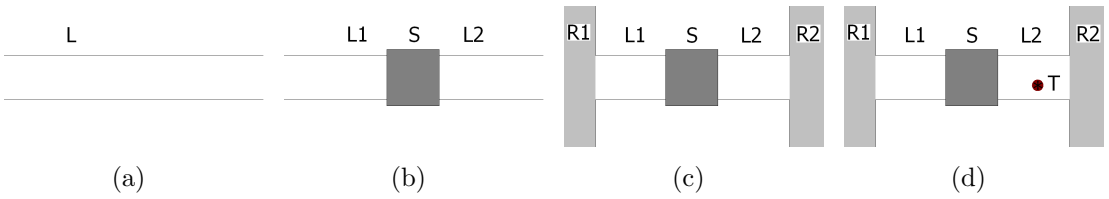


Figure 2.1: The subsequent levels of description of the SGM system:
(a) lead modelled as a regular wire,
(b) a nanostructure accessed by two leads,
(c) a nanostructure connected by the leads with two reservoirs,
(d) a nanostructure with leads, reservoirs, and the SGM tip.
Symbols:
L – lead(s), S – scatterer, R – reservoirs, T – the tip.

used as a simplistic model for the narrow part of the QPC (Ch. 4).

2.1.1 Hamiltonian and eigenfunctions

The effective mass Hamiltonian for an electron in such a wire can be written as

$$H_W = -\frac{\hbar^2}{2M_e}(\partial_x^2 + \partial_y^2) + U(y) \tag{2.1}$$

with the potential energy U depending only on the transverse variable. Its eigenfunctions read

$$\varphi(\mathbf{r}) = \frac{\exp(iskx)}{\sqrt{2\pi\hbar^2 k/M_e}}\,\phi(y)\ , \qquad s \in \{1, -1\}\ . \tag{2.2}$$

and the eigenenergies are denoted shortly by ε. The parameter k, that we identify with the longitudinal wave number, is associated with free propagation and therefore runs across a continuous set. Since the direction of propagation is marked explicitly with $s = \pm$, the wave number k has to be positive.[1] The $\phi(y)$ factor is the solution to the transverse part of the time-independent Schrödinger equation:

$$\left[-\frac{\hbar^2}{2M_e}\partial_y^2 + U(y)\right]\phi(y) = \left[\varepsilon - \frac{\hbar^2 k^2}{2M_e}\right]\phi(y)\ . \tag{2.3}$$

[1] In this context, we know that k is real, for the presence of an imaginary part of k would yield a real exponential in the wave function which then would become unphysical, growing indefinitely in one direction. We also exclude $k = 0$ states – they will no longer be solutions when we add the central structure (sec. 2.2) and for the integrals that we treat in here the apparent $1/\sqrt{k}$ singularity is always removable, hence the exclusion of the $k = 0$ point is irrelevant.

Due to the y-confinement, the energy spectrum of $H_W^{(\perp)} \equiv -\frac{\hbar^2}{2M_e}\partial_y^2 + U(y)$ is discrete and accordingly we label its eigenstates, called transverse modes, and their energies, with a mode index a running through the positive integer numbers (the order being defined by ascending transverse energies):

$$\phi_a(y), \quad \varepsilon_a^{(\perp)}, \quad a = 1, 2, \dots \ . \tag{2.4}$$

It is clear from the above, that the stationary states in a wire have energies

$$\varepsilon = \varepsilon_a^{(\perp)} + \frac{\hbar^2 k^2}{2M_e} \ . \tag{2.5}$$

Such states can in general be degenerate, their number being denoted by $N(\varepsilon)$, and to avoid the ambiguity we shall always write the longitudinal wave number with the corresponding transverse mode index:

$$\varepsilon = \varepsilon_a^{(\perp)} + \frac{\hbar^2 k_a^2}{2M_e} \ . \tag{2.6}$$

We see that we can label the wire eigenstates uniquely either by (k_a, a, s) or by (ε, a, s). Most times we will use the latter, and write: $\varphi_{\varepsilon a s}(\mathbf{r})$.

The eq. (2.6) has been plotted (for multiple a) in the figure 2.2. It gives the dispersion relation $\varepsilon(k_a, a)$ for the waves (2.2). A particle described by such a wave function moves with the velocity

$$v_a = \frac{1}{\hbar}\frac{\partial \varepsilon}{\partial k_a} = \frac{\hbar k_a}{M_e} \ . \tag{2.7}$$

In addition, the plot in 2.2 serves as a handy way of depicting on an ε-k plane the presence of the electron states. It is easily read from the figure, that the sum over all states can be expressed in equivalent forms as:

$$\sum_{s\in\{1,-1\}} \int_{\varepsilon_1^{(\perp)}}^{\infty} \mathrm{d}\varepsilon \sum_{a=1}^{N(\varepsilon)} \Big(\dots \Big) \ = \ \sum_{s\in\{1,-1\}} \sum_{a=1}^{\infty} \int_{\varepsilon_a^{(\perp)}}^{\infty} \mathrm{d}\varepsilon \Big(\dots \Big) \ . \tag{2.8}$$

2.1.2 Orthogonality and normalization

The transverse solutions (2.4) describe bound states, so they can be normalized in the usual particle-density sense. As eigenfunctions of a Hermitian operator they are also orthogonal [50], so we can write:

$$\int_{-\infty}^{\infty} \mathrm{d}y \, \phi_a^*(y)\phi_{a'}(y) = \delta_{aa'} \ , \tag{2.9}$$

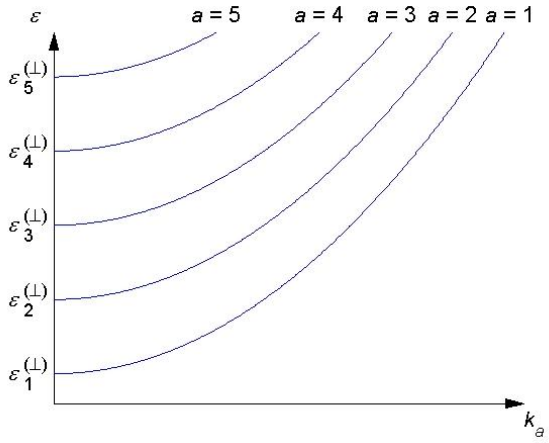


Figure 2.2: Dispersion relation for the wire states (2.2).

with the Kronecker delta $\delta_{aa'}$. They can be chosen real, hence the complex conjugate in (2.9) above could be dropped, as well as in (2.17) below.

Similarly, we can establish the orthonormality of the full wave functions, but the normalization here is set by the flux and not by the particle density. This is easily verified by the application of the current density operator ($\hat{\mathbf{r}}$ and $\hat{\mathbf{p}}$ are the position and momentum operators)

$$\hat{\mathbf{j}}(\mathbf{r}) = \frac{e}{2M_e}\left[\hat{\mathbf{p}}\,\delta(\hat{\mathbf{r}}-\mathbf{r}) + \delta(\hat{\mathbf{r}}-\mathbf{r})\hat{\mathbf{p}}\right] \ , \tag{2.10}$$

whose diagonal matrix elements $\langle\varphi_{\varepsilon as}\,|\,\hat{\mathbf{j}}(\mathbf{r})|\,\varphi_{\varepsilon as}\rangle$ yield the current densities of the wire states. In position representation they read:

$$\begin{aligned}\mathbf{j}\left\{\varphi_{\varepsilon as};\mathbf{r}\right\} &\equiv \langle\varphi_{\varepsilon as}\,|\,\hat{\mathbf{j}}(\mathbf{r})|\,\varphi_{\varepsilon as}\rangle = \\ &= \frac{e\hbar}{2iM_e}\left[\varphi^*_{\varepsilon as}(\mathbf{r})\nabla\varphi_{\varepsilon as}(\mathbf{r}) - \varphi_{\varepsilon as}(\mathbf{r})\nabla\varphi^*_{\varepsilon as}(\mathbf{r})\right] \\ &= \frac{e\hbar}{M_e}\,\mathrm{Im}\left[\varphi^*_{\varepsilon as}(\mathbf{r})\nabla\varphi_{\varepsilon as}(\mathbf{r})\right] \ . \end{aligned} \tag{2.11}$$

Only their x-components are of interest (ϕ_a carry no current anyway):

$$\begin{aligned} j^{(x)}\left\{\varphi_{\varepsilon as};\mathbf{r}\right\} &= \frac{e\hbar}{M_e}\,\mathrm{Im}\left[\frac{(\exp(isk_ax))^*\,isk_a\exp(isk_ax)}{2\pi\hbar^2k_a/M_e}\,\phi^*_a(y)\phi_a(y)\right] \\ &= s\,\frac{e}{h}\,|\phi_a(y)|^2 \ . \end{aligned} \tag{2.12}$$

Once we integrate over y to get the total current along the wire, with (2.9), we see that each (ε, a, s) state is normalized to carry the current

$$J^{(x)}\left\{\varphi_{\varepsilon as};x\right\} \equiv \int \mathrm{d}y\; j^{(x)}\left\{\varphi_{\varepsilon as};\mathbf{r}\right\} = s\frac{e}{h} \tag{2.13}$$

(per unit energy – cf. next paragraph). Revoking the definition contained in eq. (2.11), we can formally treat $J^{(x)}$ as an operator of the longitudinal current, and the expression $J^{(x)}\{.;x\}$ as its diagonal matrix element in the position-representation in a basis of our choice.

The normalization that we use gives the meaning to the wave functions $\varphi_{\varepsilon as}(\mathbf{r})$ as the amplitudes of probability density not only in space, but also in energy. In other words, the wave functions we work with already incorporate the densities of states in energy.[2] This is very convenient, as we work with continua of states – each a-branch of (2.6) is such a continuum – and it allows us to write sums over states in the energy domain in a simple form, without referring to the density of states.

It also means that the wire states are orthogonal. The orthogonality is demonstrated by taking the product of two arbitrary wave functions:

$$\begin{aligned}\int \mathrm{d}\mathbf{r}\, \varphi^*_{\varepsilon as}(\mathbf{r})\, \varphi_{\varepsilon' a' s'}(\mathbf{r}) = \\ &= \frac{M_e}{2\pi\hbar^2} \iint \mathrm{d}x\mathrm{d}y\, \frac{1}{\sqrt{k_a k'_{a'}}}\, \exp\left[-i(sk_a - s'k'_{a'})x\right] \phi^*_a(y)\phi_{a'}(y) \\ &= \frac{M_e}{\hbar^2\sqrt{k_a k'_{a'}}}\, \delta_{aa'}\delta(sk_a - s'k'_{a'}) \\ &= \frac{M_e}{\hbar^2 k'_a}\, \delta_{aa'}\delta_{ss'}\delta(k_a - k'_a) \\ &= \delta_{aa'}\delta_{ss'}\delta(\varepsilon - \varepsilon')\ , \end{aligned} \tag{2.14}$$

where we have used the Fourier decomposition of the Dirac delta function

$$\delta(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}x\, \exp(\pm ikx) \tag{2.15}$$

and the simple result derived from the properties of the Dirac's delta (gathered for example in [51]):

$$\begin{aligned}\delta_{aa'}\delta(\varepsilon - \varepsilon') &= \delta_{aa'}\delta\left(\varepsilon_a^{(\perp)} + \frac{\hbar^2 k_a^2}{2M_e} - \varepsilon_a^{(\perp)} - \frac{\hbar^2 k_a'^2}{2M_e}\right) \\ &= \frac{M_e}{\hbar^2 k'_a}\delta_{aa'}\left[\delta(k_a - k'_a) + \delta(k_a + k'_a)\right] \\ &= \frac{M_e}{\hbar^2 k'_a}\delta_{aa'}\delta(k_a - k'_a)\ , \end{aligned} \tag{2.16}$$

granted that $k_a, k'_a > 0$. The k'_a is defined through the eq. (2.6) as $k'_a \equiv k_a(\varepsilon')$.

[2] Equivalently, we could say that the density of these flux-normalized states in energy is identically equal to one state per unit energy. In this case the units of the wave function would have to be different, and the density of states $D_a(\varepsilon) = 1[\mathrm{J}^{-1}]$ would appear in the current formulae just for the unit conversion.

2.1.3 Completeness

Also from the Hermiticity of the $H_W^{(\perp)}$ operator over the space of bound functions, it follows that the transverse states constitute a complete set in their appropriate subspace [52, 53], which is often expressed by the means of the closure relation

$$\sum_{a=1}^{\infty} \phi_a(y)\phi_a^*(y') = \delta(y-y') \ . \tag{2.17}$$

In order to show the closure for the full wave functions (2.2), we need to take the sum of $\varphi(\mathbf{r})\varphi^*(\mathbf{r}')$ over all the possible (ε, a, s) states (cf. (2.8)):

$$\begin{aligned}
&\sum_{s\in\{1,-1\}} \sum_{a=1}^{\infty} \int_{\varepsilon_a^{(\perp)}}^{\infty} \mathrm{d}\varepsilon \, \varphi_{\varepsilon a s}(\mathbf{r}) \, \varphi_{\varepsilon a s}^*(\mathbf{r}') = \\
&\qquad = \sum_{s=\pm} \sum_{a=1}^{\infty} \int_{\varepsilon_a^{(\perp)}}^{\infty} \mathrm{d}\varepsilon \, \frac{\exp\left[isk_a(x-x')\right]}{2\pi\hbar^2 k_a/M_e} \, \phi_a(y)\phi_a^*(y') \\
&\qquad = \frac{1}{2\pi} \sum_{s=\pm} \sum_{a=1}^{\infty} \phi_a(y)\phi_a^*(y') \int_0^{\infty} \frac{\hbar^2 k_a}{M_e} \mathrm{d}k_a \, \frac{\exp\left[isk_a(x-x')\right]}{\hbar^2 k_a/M_e} \\
&\qquad = \frac{1}{2\pi} \sum_{a=1}^{\infty} \phi_a(y)\phi_a^*(y') \Big\{ \int_0^{\infty} \mathrm{d}k_a \, \exp\left[ik_a(x-x')\right] \\
&\qquad\qquad\qquad + \int_0^{\infty} \mathrm{d}k_a \, \exp\left[-ik_a(x-x')\right] \Big\} \\
&\qquad = \frac{1}{2\pi} \sum_{a=1}^{\infty} \phi_a(y)\phi_a^*(y') \int_{-\infty}^{\infty} \mathrm{d}k_a \, \exp\left[ik_a(x-x')\right] \\
&\qquad = \delta(x-x')\delta(y-y') \ .
\end{aligned} \tag{2.18}$$

In the third line the variable of integration has been changed to k_a according to the dispersion relation (2.6), and in the second integral on the fourth line we "flip" the integration variable, ie. change it to $-k_a$ and then rename this one as k_a in order to complete the first integral. The final passage is done by first applying (2.15) and then (2.17).

Completeness is a crucial property. The wire states we have introduced, being complete and linearly independent, form a basis for the description of the electron motion through a wire. Obviously, a travelling electron does not have to be in one of them. But any complicated state in which it might exist, can always be written as a linear combination of the basis states. Such a linear combination includes the continuous case, where the linear coefficients are given by an envelope function, ie. the creation of a wave packet. The

exact shapes of the wave packets, however, will not be relevant in the scope of this work.

2.1.4 Examples of transverse wave functions

Before closing this preliminary section, we would like to mention the form of transverse wave functions for two simple cases of the potential $U(y)$:

1) the hard-walled square well of width W,

$$U(y) = \begin{cases} \infty\,, & y \geq W/2 \\ 0\,, & |y| < W/2 \\ \infty\,, & y \leq -W/2 \end{cases}\,, \tag{2.19}$$

with the solutions

$$\phi_a(y) = \sqrt{\frac{2}{W}}\,\sin\left[\frac{\pi a}{W}\left(y + \frac{W}{2}\right)\right]\,, \tag{2.20}$$

and their associated energies,

$$\varepsilon_a^{(\perp)} = \frac{\hbar^2}{2M_e}\frac{\pi^2 a^2}{W^2}\,, \tag{2.21}$$

2) the harmonic potential,

$$U(y) = \frac{1}{2}M_e\omega_y^2 y^2\,, \tag{2.22}$$

whose solutions are given by

$$\phi_a(y) = \frac{1}{\sqrt{2^{a-1}(a-1)!}}\left(\frac{M_e\omega_y}{\pi\hbar}\right)^{1/4}\exp\left(\frac{-M_e\omega_y}{2\hbar}y^2\right)\,H_{a-1}\left(\sqrt{\frac{M_e\omega_y}{\hbar}}y\right) \tag{2.23}$$

with the Hermite polynomials H_{a-1}, and the energies

$$\varepsilon_a^{(\perp)} = \hbar\omega_y\left(a - \frac{1}{2}\right)\,. \tag{2.24}$$

The above solutions will be useful when we apply our theoretical results of chapter 3 to particular models of a quantum point contact in chapter 4.

2.2 Electrons in the leads, scattering states

Let us investigate a system that consists of the structure of our interest, whose Hamiltonian we denote by H_S, accessed from the left and right by a pair of leads, of which we assume that they are well approximated by the ideal wires defined in the previous section, with Hamiltonian H_W. (The structure-leads coupling we include into H_S.) At a reasonable distance from the structure the electrons can still be well described by the wire states, in this context also called *lead states*, or *lead modes*. But the presence of the structure at the end of the wire imposes some additional boundary conditions, and it turns out that another basis, constructed upon the states of the previous one, is more convenient for the problem of electron motion.

2.2.1 Scattering states

As long as the structure in question is coherent, we are able to describe a moving electron with one wave function stretching across the entire system. If we inject from the left an electron in a wire-basis state, its wave function will consist of:

- the incident wave (2.2) in the left lead, propagating to the right and reaching the structure,
- some, possibly complicated, form inside the structure, and
- the reflected and transmitted wave parts in the left and right lead, travelling away from the structure.

All are of the same energy, since we assume no inelastic processes. The transmitted and reflected parts away from the central structure can be resolved into linear combinations of the wire states. So the asymptotic form of such a scattering state can be written as

$$\begin{cases} \varphi_{\varepsilon a+}(\mathbf{r}) + \sum_{b=1}^{N(\varepsilon)} r_{ba}\varphi_{\varepsilon b-}(\mathbf{r}), & x \to -\infty \\ \sum_{b=1}^{N(\varepsilon)} t_{ba}\varphi_{\varepsilon b+}(\mathbf{r}), & x \to +\infty \end{cases}$$

In principle, we should also have included the non-propagating modes ($b > N$) for the basis expansion to be correct – the lead-wire ends at the structure, so decaying exponentials for the longitudinal part of the wave function are now in place. However, these terms fall very quickly with the distance from the central structure and in the asymptotic form can be neglected. Generally speaking, the N's can be different in different leads, but for simplicity we will incorporate the assumption of identical leads.

We would like to mark the items according to whether they propagate towards or away from the central structure. For this purpose, we introduce a superscript: (+) for the outgoing part of the wave and (−) for the incoming part of the wave, and use them instead of the left-right motion label $s = \pm$. Also, a lead label l is added: 1 for the left lead and 2 for the right lead. Thus, we have the wave function

$$\Psi_{1\varepsilon a}(\mathbf{r}) = \begin{cases} \varphi^{(-)}_{1\varepsilon a}(\mathbf{r}) + \sum_{b=1}^{N} r_{ba}\varphi^{(+)}_{1\varepsilon b}(\mathbf{r}), & x \ll -L/2 \\ \sum_{b=1}^{N} t_{ba}\varphi^{(+)}_{2\varepsilon b}(\mathbf{r}), & x \gg L/2 \end{cases} \tag{2.25a}$$

for a scattering state originating on the left, and

$$\Psi_{2\varepsilon a}(\mathbf{r}) = \begin{cases} \sum_{b=1}^{N} t'_{ba}\varphi^{(+)}_{1\varepsilon b}(\mathbf{r}), & x \ll -L/2 \\ \varphi^{(-)}_{2\varepsilon a}(\mathbf{r}) + \sum_{b=1}^{N} r'_{ba}\varphi^{(+)}_{2\varepsilon b}(\mathbf{r}), & x \gg L/2 \end{cases} \tag{2.25b}$$

for the scattering state from the right. L is the longitudinal extension of the structure. Each of these wave functions is labelled by the set of indices of the lead mode which provides the incoming part for the scattering state. We say that the scattering state $\Psi_{l\varepsilon a}$ is generated by the lead state $\varphi^{(-)}_{l\varepsilon a}$.

2.2.2 Current of a scattering state

Once we calculate the longitudinal component $j^{(x)}\{\Psi_{l\varepsilon a}; \mathbf{r}\}$ of the current density of a single scattering state,

$$\mathbf{j}\{\Psi_{l\varepsilon a}; \mathbf{r}\} = \langle \Psi_{l\varepsilon a} | \hat{\mathbf{j}}(\mathbf{r}) | \Psi_{l\varepsilon a} \rangle \tag{2.26}$$

– the explicit expression of which is given in Appendix B, eq. (3.17)[3] – and integrate it over dy to make use of the orthogonality of transverse modes (2.9), we obtain the current of each scattering state on the left and on the right in the simple forms:

$$J^{(x)}\{\Psi_{1\varepsilon a}; x\} = \frac{e}{h}\Big[1 - \big(r^\dagger r\big)_{aa}\Big], \qquad x \ll -L/2\,, \tag{2.27a}$$

$$J^{(x)}\{\Psi_{1\varepsilon a}; x\} = \frac{e}{h}\big(t^\dagger t\big)_{aa}\,, \qquad x \gg L/2\,, \tag{2.27b}$$

$$J^{(x)}\{\Psi_{2\varepsilon a}; x\} = \frac{e}{h}\big(t'^\dagger t'\big)_{aa}\,, \qquad x \ll -L/2\,, \tag{2.27c}$$

$$J^{(x)}\{\Psi_{2\varepsilon a}; x\} = \frac{e}{h}\Big[1 - \big(r'^\dagger r'\big)_{aa}\Big], \qquad x \gg L/2\,. \tag{2.27d}$$

[3]For one case only – in the region to the right of the scatterer, while electrons are injected from the lead $l = 1$.

It is important to remember that, due to our particular definition of the electron states, these currents retain the character of *densities* when viewed in the energy domain.

From the above considerations we see that out of each $\frac{e}{h}$ injected into the left lead in the incident lead mode (ε, a), we get $\frac{e}{h}[t^\dagger t]_{aa}$ on the right, while $\frac{e}{h}[r^\dagger r]_{aa}$ gets back to the left lead and diminishes the net (x-oriented) current inside it. Similarly for the right-lead injection. The Hermitian matrices $t^{(\prime)\dagger}t^{(\prime)}$ and $r^{(\prime)\dagger}r^{(\prime)}$ represent transmission and reflection operators, and their diagonal elements $T_a^{(\prime)} \equiv [t^{(\prime)\dagger}t^{(\prime)}]_{aa}$, $R_a^{(\prime)} \equiv [r^{(\prime)\dagger}r^{(\prime)}]_{aa}$ are transmission and reflection probabilities for a single mode. These are essential in the calculation of the structure's conductance. However, we still do not know the rate of injection of the electrons into our system, which will be considered in sec. 2.4.

2.2.3 Scattering states as a basis

It can be easily seen, that it is possible to reverse the relations (2.25a-2.25b) and obtain any of the wire-basis states $\varphi^{(\pm)}_{l\varepsilon a}$ in the appropriate asymptotic region by taking a suitable linear combination of the scattering states $\Psi_{l\varepsilon b}$:

– for the left lead, as $x \to -\infty$,

$$\varphi^{(-)}_{1\varepsilon a}(\mathbf{r}) = \Psi_{1\varepsilon a}(\mathbf{r}) + \sum_b (-t'^{-1}r)_{ba}\Psi_{2\varepsilon b}(\mathbf{r}) \ , \tag{2.28a}$$

$$\varphi^{(+)}_{1\varepsilon a}(\mathbf{r}) = \sum_b t'^{-1}_{ba}\Psi_{2\varepsilon b}(\mathbf{r}) \ , \tag{2.28b}$$

– for the right lead, as $x \to +\infty$,

$$\varphi^{(-)}_{2\varepsilon a}(\mathbf{r}) = \Psi_{2\varepsilon a}(\mathbf{r}) + \sum_b (-t^{-1}r')_{ba}\Psi_{1\varepsilon b}(\mathbf{r}) \ , \tag{2.29a}$$

$$\varphi^{(+)}_{2\varepsilon a}(\mathbf{r}) = \sum_b t^{-1}_{ba}\Psi_{1\varepsilon b}(\mathbf{r}) \ . \tag{2.29b}$$

The required invertibility of the t and t' matrices can be safely assumed, since vanishing of its determinant, would mean that at least one mode of the transmission eigenbasis gets fully stopped (see sec. 2.3.3), which is the case for an impenetrable barrier only – in all other cases there is at least a tunnelling transmission – and such a case is of no interest to us.

The above construction (2.28–2.29) shows that the set of the scattering states is complete inside each individual lead. They are not complete, however, in the joint space of both leads. Such a space is a tensor product of the

individual leads' spaces, its dimension is therefore greater than the dimension of the set of scattering states. Nonetheless, what we need is the space of the wave functions defined in both leads *and* complying with the fact that the leads are connected by the scatterer. In that space, the scattering states' completeness holds, as long as they are consistent with the "connecting" Hamiltonian, H_S.[4] It is in this sense, that we treat the presence of the scatterer as a boundary condition, disregarding, for now, what exactly happens inside it.

There is yet another boundary condition that we have to mention, implicit in our construction of the scattering states. It is the (asymptotic) time evolution of a particle described by them. A wave packet built up from these states:

$$\int \mathrm{d}\varepsilon \sum_a C_{\varepsilon a} \exp(-i\varepsilon t/\hbar)\Psi_{l\varepsilon a}(\mathbf{r})$$

will follow the evolution expected from a "projectile" – in the distant past it describes a particle moving freely inside a wire towards the central structure and the waveform then contains no information of the structure, later it gets scattered and heads away in a complicated state determined by the influence of the scatterer. Such scattering states are referred to as the *outgoing scattering states* [54, 55], and are usually written with a (+) superscript which we omit for conciseness.

The relation between the scattering states and the lead states can also be described by means of the Lippmann-Schwinger equation with the retarded Green function of the clean wire, the choice of the retarded function being dictated by the just-described boundary condition in time [56].[5] Here, the Hamiltonian H_S of the central structure is treated as a perturbation (not necessarily weak) on the Hamiltonian H_W of the clean wire, and the scattering states are the solutions to the full Hamiltonian $H_W + H_S$. We shall occupy ourselves with this formalism when we add to our system yet another ingredient – the SGM tip. At this moment, we invoke only one of its results, whose proof can be found in [56] (sec. 4-2b, p. 298): the scattering states satisfy the same orthonormality relations as the unperturbed states they were made of,

$$\langle \Psi_{l\varepsilon a} | \Psi_{l'\varepsilon' a'} \rangle = \left\langle \varphi^{(-)}_{l\varepsilon a} | \varphi^{(-)}_{l'\varepsilon' a'} \right\rangle \ . \tag{2.30}$$

This, together with the completeness, grants that the set of all the scattering states is an *orthogonal basis*. At a given energy, we can decompose

[4] This is ensured by the Fisher-Lee relations, sec. 2.6.

[5] Note that in [56] the terminology is the reverse of what we use – the states of our interest are called *incoming scattering states*.

an arbitrary wave allowed by our system not only into the lead states in the asymptotic regions:

$$\Psi_\varepsilon = \sum_{la} C_{la}^{(-)} \varphi_{l\varepsilon a}^{(-)} + \sum_{la} C_{la}^{(+)} \varphi_{l\varepsilon a}^{(+)} \ , \tag{2.31}$$

but also into the scattering states in the entire conducting system:

$$\Psi_\varepsilon = \sum_{la} C_{la}^{(-)} \Psi_{l\varepsilon a} \ . \tag{2.32}$$

We stress that the coefficients $C_{la}^{(+)}$ are not independent. In order to comply with the form of the scattering states (2.25a–2.25b), they have to be given by a matrix equation

$$\begin{pmatrix} C_{1\cdot}^{(+)} \\ C_{2\cdot}^{(+)} \end{pmatrix} = \begin{pmatrix} r & t' \\ t & r' \end{pmatrix} \begin{pmatrix} C_{1\cdot}^{(-)} \\ C_{2\cdot}^{(-)} \end{pmatrix} \ , \tag{2.33}$$

where the reflection and transmission amplitudes of (2.25a) have been arranged as matrices $r^{(\prime)}$, $t^{(\prime)}$, whereas $C_{l\cdot}^{(-)}$ and $C_{l\cdot}^{(+)}$ should be understood as column vectors of the expansion coefficients of, respectively, the incoming and outgoing parts of the wave Ψ_ε.

The orthogonality of the basis formed from scattering states will also allow us for a very convenient eigendecomposition of the Green function in sec. 3.1, chapter 3.

2.3 Scattering matrix

The transformation matrix of (2.33),

$$\begin{pmatrix} r & t' \\ t & r' \end{pmatrix} \equiv \mathbf{S} \ , \tag{2.34}$$

is of great importance to our formalism, because it encapsulates all the relevant information about the scatterer. It is called the scattering matrix, or shortly S-matrix, and relates the amplitudes of the incoming and outgoing lead states. In general, the S-matrix depends on energy, but we do not display it in our notation.

2.3.1 Unitarity

Application of the current density operator to a wave incoming from the left, ie. $\Psi^{(in)}_{1\varepsilon} = \sum_a C^{(-)}_{1a}\varphi^{(-)}_{1\varepsilon a}$, followed by d$y$ integration and the use of orthonormality (2.9), yields a total longitudinal current of

$$J^{(x)}\left\{\Psi^{(in)}_{1\varepsilon}; x\right\} = \frac{e}{h}\sum_a |C^{(-)}_{1a}|^2 \ . \tag{2.35}$$

By adding to the above an analogous result for the wave coming from the right, we get the entire current towards the structure, in the units of e/h:

$$\frac{1}{e/h}J^{(x,in)} = \sum_{la}\left|C^{(-)}_{la}\right|^2 = \begin{pmatrix} C^{(-)}_{1\cdot} \\ C^{(-)}_{2\cdot}\end{pmatrix}^\dagger \begin{pmatrix} C^{(-)}_{1\cdot} \\ C^{(-)}_{2\cdot}\end{pmatrix}$$

and similarly for the current flowing outwards:

$$\frac{1}{e/h}J^{(x,out)} = \sum\left|C^{(+)}_{la}\right|^2 = \begin{pmatrix} C^{(-)}_{1\cdot} \\ C^{(-)}_{2\cdot}\end{pmatrix}^\dagger \mathbf{S}^\dagger\mathbf{S} \begin{pmatrix} C^{(-)}_{1\cdot} \\ C^{(-)}_{2\cdot}\end{pmatrix} \ .$$

Hence, for the sake of current conservation we demand that the scattering matrix be unitary:

$$\mathbf{S}^\dagger\mathbf{S} = \mathbf{S}\mathbf{S}^\dagger = \mathbf{1} \ . \tag{2.36}$$

This implies, in particular, that

$$r^\dagger r + t^\dagger t = \mathbf{1} \tag{2.37a}$$

and

$$r^\dagger t' + t^\dagger r' = \mathbf{0} \tag{2.37b}$$

in the appropriate subspaces.

2.3.2 Symmetry

Let us write in the lead-state basis a complex conjugate of some wave function allowed by our system:

$$\begin{aligned}\Psi^*_\varepsilon &= \sum_{la} C^{(-)*}_{la}\varphi^{(-)*}_{l\varepsilon a} + \sum_{la} C^{(+)*}_{la}\varphi^{(+)*}_{l\varepsilon a} \\ &= \sum_{la} C^{(-)*}_{la}\varphi^{(+)}_{l\varepsilon a} + \sum_{la} C^{(+)*}_{la}\varphi^{(-)}_{l\varepsilon a} \ .\end{aligned}$$

It is a spatial part of a time-dependent wave function with the reversed time. As we assume zero magnetic field, the system is time-reversal invariant, so its

time-dependent Schrödinger equation is satisfied by this state even with the non-reversed-time temporal part. Hence, the incoming and outgoing parts of the wave Ψ_ε^* have to obey the transformation (2.33), that is:

$$\begin{pmatrix} C_{1\cdot}^{(-)} \\ C_{2\cdot}^{(-)} \end{pmatrix} = \mathbf{S} \begin{pmatrix} C_{1\cdot}^{(+)} \\ C_{2\cdot}^{(+)} \end{pmatrix} . \tag{2.38}$$

Substituting to the above the relation (2.33) for the original state Ψ_ε, we obtain

$$\mathbf{S}^*\mathbf{S} = \mathbf{S}\mathbf{S}^* = \mathbf{1} \tag{2.39}$$

which, combined with the unitarity (2.36), tells us that the S-matrix must be symmetric, $\mathbf{S} = \mathbf{S}^{\mathrm{T}}$.

2.3.3 Polar representation of the S-matrix

By performing singular value decompositions of the transmission and reflection amplitude matrices , $t^{(\prime)}$ and $r^{(\prime)}$, and using the above condition that $\mathbf{S}$ be unitarity, we can obtain a particularly convenient form of the scattering matrix [57–59]:

$$\mathbf{S} = \begin{pmatrix} u_3 & 0 \\ 0 & u_4 \end{pmatrix} \begin{pmatrix} -\mathcal{R} & \mathcal{T} \\ \mathcal{T} & \mathcal{R} \end{pmatrix} \begin{pmatrix} u_1 & 0 \\ 0 & u_2 \end{pmatrix} . \tag{2.40}$$

The auxiliary matrices u_1, u_2, u_3, u_4 are all unitary. In the time-reversal invariant case, we can further simplify them by applying the S-matrix symmetry requirement, getting

$$u_3 = u_1^{\mathrm{T}} , \quad u_4 = u_2^{\mathrm{T}} . \tag{2.41}$$

The remaining $\mathcal{T}$ and $\mathcal{R}$ matrices are diagonal and contain the singular values of the $t^{(\prime)}$ and $r^{(\prime)}$ matrices, i.e. the square roots of the eigenvalues of the transmission and reflection operators $T^{(\prime)}$ and $R^{(\prime)}$. They can be parametrized in the following way:

$$\mathcal{R}_m \equiv \mathcal{R}_{mm} = \left(\frac{\lambda_m}{1+\lambda_m} \right)^{1/2} , \tag{2.42a}$$

$$\mathcal{T}_m \equiv \mathcal{T}_{mm} = \left(\frac{1}{1+\lambda_m} \right)^{1/2} , \tag{2.42b}$$

with real positive λ_m, $m = 1, 2, \ldots, N$.

More directly, we could write (2.40-2.41) as

$$t = u_2^{\mathrm{T}} \mathcal{T} u_1 = t'^{\,\mathrm{T}} , \qquad r = -u_1^{\mathrm{T}} \mathcal{R} u_1 , \qquad r' = u_2^{\mathrm{T}} \mathcal{R} u_2 . \tag{2.43}$$

It follows immediately, that the matrices u_1 and u_2 diagonalize the transmission operators $t^{(\prime)\dagger}t^{(\prime)}$:

$$t^\dagger t = u_1^{-1}\mathcal{T}^2 u_1 \ , \tag{2.44a}$$

$$t'^\dagger t' = u_2^{-1}\mathcal{T}^2 u_2 \ , \tag{2.44b}$$

and, consequently, that the wave functions

$$\overset{\approx}{\varphi}{}^{(-)}_{l\varepsilon m}(\mathbf{r}) = \sum_a [u_l]^\dagger_{am}\ \varphi^{(-)}_{l\varepsilon a}(\mathbf{r}) \tag{2.45}$$

are *transmission eigenmodes* with the corresponding eigenvalues $\mathcal{T}_m^2$, for both values of l. It is worthwhile to remark here, that by a transmission eigenmode we understand an eigenstate of the transmission operator acting over the space of the modes of the appropriate lead, $t^\dagger t$ for $l = 1$ and $t'^\dagger t'$ for $l = 2$. An eigenmode therefore exists in one lead only ($x < -L/2$ or $x > L/2$).

What we will also need are *the scattering eigenstates*, i.e. the scattering states generated by the transmission eigenmodes:

$$\overset{\approx}{\Psi}_{l\varepsilon m}(\mathbf{r}) = \sum_a [u_l]^\dagger_{am}\ \Psi_{l\varepsilon a}(\mathbf{r}) \ . \tag{2.46}$$

They, too, are eigenstates of the transmission operators $t^{(\prime)\dagger}t^{(\prime)}$, but in a different space (the space of functions over the entire position-space, cf. sec. 2.2.3). Their explicit asymptotic form in the basis of the transmission eigenmodes is

$$\overset{\approx}{\Psi}_{1\varepsilon m}(\mathbf{r}) = \begin{cases} \overset{\approx}{\varphi}{}^{(-)}_{1\varepsilon m}(\mathbf{r}) - \mathcal{R}_m \overset{\approx}{\varphi}{}^{(+)}_{1\varepsilon m}(\mathbf{r}), & x \ll -L/2 \\ \mathcal{T}_m \overset{\approx}{\varphi}{}^{(+)}_{2\varepsilon m}(\mathbf{r}), & x \gg \ \ L/2 \end{cases} , \tag{2.47a}$$

$$\overset{\approx}{\Psi}_{2\varepsilon m}(\mathbf{r}) = \begin{cases} \mathcal{T}_m \overset{\approx}{\varphi}{}^{(+)}_{1\varepsilon m}(\mathbf{r}), & x \ll -L/2 \\ \overset{\approx}{\varphi}{}^{(-)}_{2\varepsilon m}(\mathbf{r}) + \mathcal{R}_m \overset{\approx}{\varphi}{}^{(+)}_{2\varepsilon m}(\mathbf{r}), & x \gg \ \ L/2 \end{cases} , \tag{2.47b}$$

where the outgoing, (+)-superscripted, eigenmodes are defined simply as the eigenmodes with the reversed direction of propagation: $\overset{\approx}{\varphi}{}^{(+)}_{l\varepsilon m}(\mathbf{r}) \equiv \left[\overset{\approx}{\varphi}{}^{(-)}_{l\varepsilon m}(\mathbf{r})\right]^*$.

Despite the fact that an overall phase factor could be freely introduced to either of the u_l matrices in (2.44), we see from the relations (2.43) that compatibility with the $t^{(\prime)}$ and $r^{(\prime)}$ matrices, whose phases are fixed by the definition of the scattering states (2.25) and of the lead modes (sec. 2.1), does not allow for such a choice to be made when writing the S-matrix in

the polar form (2.40). It is therefore important to note that we work with one particular set of eigenmodes and with one particular set of the scattering eigenstates, both uniquely determined by the mentioned definitions of the scattering states and the lead modes. Each time we speak of *the transmission eigenmodes basis* or of *the scattering eigenstates basis*, we do not mean arbitrary bases, but precisely these sets.

Owing to the Hermiticity of $t^{(\prime)\dagger}t^{(\prime)}$, the transmission eigenmodes form orthogonal bases in their spaces, and so do the scattering eigenstates in their space. The currents of the eigenmodes can be evaluated from eq. (2.35). Unitarity of the u_l-matrices ensures that the normalization of the eigenmodes (also of the outgoing ones) is the same as of the ordinary lead states.

If we use (2.35) on the outgoing transmitted parts of the scattering eigenstates, we get

$$J^{(x)}\left\{\mathcal{T}_m \overset{\approx}{\varphi}{}^{(+)}_{l\varepsilon m}; x\right\} = \frac{e}{h}\mathcal{T}_m^2 \,, \tag{2.48}$$

for both l, which explicitly shows that the transmission eigenvalues $\mathcal{T}_m^2$ give the transmission probability for each eigenmode, as expected. And since the determinant of the transmission matrix, $\det(t^\dagger t) = \det(t)^* \det(t)$, is at the same time equal to the product of its eigenvalues, we see that indeed the vanishing of the determinant of the transmission amplitude matrix t, would eventually imply exactly zero transmission of at least one of such modes. Naturally, the same is true for t'.

The scattering states make the transport channels of our system. For the purposes of further chapters, we would like to classify the scattering eigenstates according to whether they are open or closed. We denote the set of all the eigenstates (or, equivalently, all the indices thereof) which are:

- completely open, $\mathcal{T}_m \approx 1$ – by $\underline{\mathcal{M}}$,
- completely closed, $\mathcal{T}_m \approx 0$ (this is limited by tunnelling) – by $\overline{\mathcal{M}}$,
- partially open – by $\not{\mathcal{M}}$,

while the space of all the active channels will be denoted by $\mathcal{M}$. Their respective numbers will simply be $N_{\underline{\mathcal{M}}}$, $N_{\overline{\mathcal{M}}}$, $N_{\not{\mathcal{M}}}$ and $N_{\mathcal{M}}$.

2.4 The reservoirs

In order to determine the current and the conductance through a nanostructure within the scattering formalism, it is necessary to supplement the system we have described so far with one important element – the reservoirs. [60] We

imagine that at the "loose" end of each lead there is a macroscopic reservoir (fig. 2.1(c)), held in thermal equilibrium at all times. These entities model the influence of all the rest of the macroscopic circuit that is normally present when measurements are made.

The reservoirs are characterized by two parameters each: the temperature ϑ_l and the chemical potential μ_l, where $l = 1$ for the left reservoir, $l = 2$ for the right one. The thermal equilibrium requires that the distribution of their electrons in energy is given by the Fermi-Dirac function

$$f(E) = \left(1 + \exp \frac{E - \mu_l}{k_B \vartheta_l}\right)^{-1} . \tag{2.49}$$

Since we do not investigate thermoelectric effects, we set both reservoirs to the same temperature: $\vartheta_1 = \vartheta_2 = \vartheta$. The applied bias voltage, measured in the reservoirs, is expressed in terms of the chemical potentials as $V_b = (\mu_1 - \mu_2)/e$.

We assume the electrons should be able to exit from the lead into the corresponding reservoir without suffering reflection [37]. We need this assumption in order to establish the distribution of the electrons in the leads – it grants us that any electron inside the lead *travelling towards the scatterer* has its origin in the reservoir attached to that lead.

We state that a reservoir injects into its corresponding lead as many carriers as the lead can cater for, the only two constraints being that:

- the electrons entering the lead from the reservoir retain the energy distribution they had in the reservoir (2.49),
- for the left lead, only the electrons with positive k can enter the lead from the reservoir (negative k for the right lead), so we are left with just a half of the Fermi-Dirac ensemble therein[6].

We remark, that the first point is an *assumption*, and its justification is not completely obvious. After all, there is a potential drop when passing from the reservoir to the lead (see [60] and references therein) and it is not a priori known whether the probability of passing depends on the energy or not. To support this assumption we would like to invoke a work by Yosefin and Kaveh [61], where an explicit treatment was presented for the leads that broaden as the distance from the scatterer grows. In that case, there is a smooth crossover between the leads and the reservoirs, and the local ensemble of the electrons in the vicinity of the quantum point contact complies with the statement we made.

[6] If account is made for the left-travelling electrons: the reflected ones (R_a) and the ones transmitted from the other side (T'_a), then the ensemble is something yet more complicated – see e.g. [37].

2.5 Conductance

Once we know both, the population of the incoming states and the current due to each of them, we can easily evaluate the overall current and conductance.

2.5.1 Current calculation

It follows from the discussion in the previous section, that the average occupancy of the scattering states coming from the left (2.25a) is given by the Fermi-Dirac function f_1 of the eq. (2.49). It has also been asserted before (subsec. 2.2.2), that each of those states carries $e/h\,(t^\dagger t)_{aa}$ into the right lead, and all of this reaches the right reservoir. The overall current from the left to the right is then given by the sum of all $(1\varepsilon a)$-state currents multiplied by the occupation probabilities,[7]

$$\begin{aligned} I_{1\to 2} &= \sum_{\{(1\varepsilon a)\}} f_1(\varepsilon)\, J^{(x)}\left\{\Psi_{1\varepsilon a}; x\right\} = 2\int_{\varepsilon_1^{(\perp)}}^{\infty} \mathrm{d}\varepsilon \sum_{a=1}^{N(\varepsilon)} f_1(\varepsilon)\, \frac{e}{h}\left[t^\dagger(\varepsilon)\, t(\varepsilon)\right]_{aa} \\ &= \frac{2e}{h}\int_{\varepsilon_1^{(\perp)}}^{\infty} \mathrm{d}\varepsilon\, f_1\, \mathrm{tr}\left(t^\dagger t\right) \ , \end{aligned} \tag{2.50}$$

where the energy of the first transverse mode in the left lead $\varepsilon_1^{(\perp)}$ has been used as the "band bottom" and the energy arguments on the second line are understood (we will omit them whenever it does not cause ambiguity). The additional factor of 2 in the expansion of the sum takes into account the spin degeneracy at zero magnetic field.

The sum is taken incoherently, since the electrons come from a thermal source, and – although travelling coherently through the leads and scatterer – they are not coherent with each other. Naturally, the electrons might not arrive exactly in $\Psi_{1\varepsilon a}$ states, but in some linear combinations thereof, $\Psi'_{1\varepsilon a} = \sum_{a'} c_{aa'}\Psi_{1\varepsilon a'}$. But this set of states has to form a basis as well. And since the trace of an operator is basis-invariant, the sum of currents over the states of this general set is exactly the same as 2.50.

The same calculation applies to the current from the right to the left, yielding

$$I_{2\to 1} = \frac{2e}{h}\int_{\varepsilon_1^{(\perp)}}^{\infty} \mathrm{d}\varepsilon\, f_2\, \mathrm{tr}\left(t^\dagger t\right) \tag{2.51}$$

[7] As we have discussed before (sec. 2.1), the number of electrons within each $\mathrm{d}\varepsilon_a$ is already accounted for by the normalization of the wave functions, so the density of states does not enter the integral.

From the form (2.44) we see that the transmission probability does not depend on the side from which the electrons are originating, as $\mathrm{tr}(t'^{\dagger}t') = \mathrm{tr}(\mathcal{T}^2) = \mathrm{tr}(t^{\dagger}t)$. The total current through the structure is therefore given by

$$I = I_{1\to 2} - I_{2\to 1} = \frac{2e}{h}\int_{\varepsilon_1^{(\perp)}}^{\infty} \mathrm{d}\varepsilon\ (f_1 - f_2)\ \mathrm{tr}\left(t^{\dagger}t\right)\ . \tag{2.52}$$

2.5.2 Linear response

If so it happens that $\mathrm{tr}(t^{\dagger}t)$ is constant across the energy interval (μ_2, μ_1) and within a few $k_B\vartheta$ around it, then we can place this term outside the integral. We will also assume that μ_1 and μ_2 are both securely greater than $\varepsilon_1^{(\perp)}$. Then we have

$$\begin{aligned} I &\approx \frac{2e}{h}\mathrm{tr}(t^{\dagger}t)\int_{\varepsilon_1^{(\perp)}}^{\infty}\mathrm{d}\varepsilon\ (f_1 - f_2) \approx \frac{2e}{h}\mathrm{tr}(t^{\dagger}t)\int_{-\infty}^{\infty}\mathrm{d}\varepsilon\ (f_1 - f_2) \\ &= \frac{2e}{h}\mathrm{tr}(t^{\dagger}t)\,(\mu_1 - \mu_2) = \frac{2e^2}{h}\mathrm{tr}(t^{\dagger}t)\,V_b\,, \end{aligned} \tag{2.53}$$

and the response to the applied voltage is linear.

If, on the other hand, $\mathrm{tr}(t^{\dagger}t)$ varies sharply, but the bias is much smaller then the thermal energy, $\mu_1 - \mu_2 \ll k_B\vartheta$, we can employ the differential approximation

$$f_1 - f_2 \approx (\mu_1 - \mu_2)\ \frac{\partial f_1}{\partial \mu_1}\bigg|_{\mu_1=\mu_2}\ , \tag{2.54}$$

and still get the linear response:

$$I \cong \frac{2e^2}{h}V_b\int_{-\infty}^{\infty}\mathrm{d}\varepsilon\, F_{th}\mathrm{tr}(t^{\dagger}t)\,, \tag{2.55}$$

Here, F_{th} is the thermal broadening function (for small bias $\mu_1 \approx \mu_2 \approx E_F$):

$$F_{th}(\varepsilon) = -\frac{\partial f_l}{\partial \varepsilon} = \frac{\partial f_l}{\partial \mu_l} = \frac{1/k_B\vartheta}{\left[\exp\frac{\varepsilon - E_F}{2k_B\vartheta} + \exp\left(-\frac{\varepsilon - E_F}{2k_B\vartheta}\right)\right]^2}\ . \tag{2.56}$$

In the limit of zero temperature F_{th} becomes the Dirac delta $\delta(\varepsilon - E_F)$ and (2.53) is retrieved. In conclusion, the conditions where

$$eV_b \ll k_B\vartheta + \Delta\varepsilon_c \tag{2.57}$$

with $\Delta\varepsilon_c$ being the scale of variation (ie. the correlation energy) of the transmission, we will refer to as the linear response regime.

Landauer-Büttiker conductance formula

The linear response conductance obtained in this way is

$$G^{(0)} = \frac{\partial I}{\partial V_b} = \frac{2e^2}{h}\mathrm{tr}(t^\dagger t)\,. \tag{2.58}$$

This is the (two-terminal) Landauer-Büttiker formula. The transmission amplitudes should be evaluated at the Fermi energy.

The dimensionless quantity $T \equiv \mathrm{tr}(t^\dagger t)$ is the total transmission of the structure and is of central importance in the method we have developed in here. It is the sum of the transmission probabilities over the lead modes, but any other normalized basis states are in place here, in particular $T = \sum_{m=1}^{N} \mathcal{T}_m^2$ (cf. (2.44)). Each electron state active at the Fermi energy, of whichever prescribed basis, constitutes an independent transport channel. A maximum conductance of such a channel is given by the constant prefactor $2e^2/h$. It is a ubiquitous and conceptually very important quantity, called the conductance quantum.

Using the total transmission T and the dimensionless conductance expressed in the units of the conductance quantum, $g^{(0)} \equiv G^{(0)}/\frac{2e^2}{h}$, we can write the Landauer-Büttiker formula in the laconic form:

$$g^{(0)} = T\ . \tag{2.59}$$

The point of view we have just described differs significantly from the classical one, because instead of the presence of the external electric field in the entire structure, giving a certain drift velocity to all of the free carriers, we only have to consider the effect the field has on the populations of travelling states on both sides of the scatterer, producing a net current carried only by the electrons at the Fermi energy.

2.6 Calculating the S-matrix, Fisher-Lee relation

How do we evaluate the transmission and reflection amplitudes $t_{ba}^{(\prime)}$, $r_{ba}^{(\prime)}$? An explicit way to do it would be by solving the Schrödinger equation for the allowed scattering states, and decomposing them inside each asymptotic region in the basis of the outgoing lead states, as is done for example in [62].

A more convenient approach, which we will employ, uses the retarded Green function:

$$\mathcal{G}^{(0)}(\mathbf{r},\mathbf{r}',E) = \left\langle \mathbf{r}\left|(E^+ - H_0)^{-1}\right|\mathbf{r}'\right\rangle\ , \tag{2.60}$$

where H_0 is the Hamiltonian of the entire system, acting in the space over $\mathbf{r}$, and E is the energy variable; the plus superscript indicates that an infinitesimal positive imaginary part has been added: $E^+ = E + i\eta,\ \eta \to 0^+$. In our case $H_0 = H_S + H_W$.

The transmission and reflection amplitudes are related to the Green function by what is known as the Fisher-Lee formulae [63]:

$$t_{ba} = i\hbar(v_a v_b)^{1/2} \exp\left[-i(k_b x - k_a x')\right] \int_{\mathcal{S}_x} \mathrm{d}y \int_{\mathcal{S}_{x'}} \mathrm{d}y'\ \phi_b^*(y)\ \mathcal{G}^{(0)}(\mathbf{r}, \mathbf{r}', \varepsilon)\ \phi_a(y')\ , \tag{2.61a}$$

$$\begin{aligned} r_{ba} = &-\delta_{ab} \exp\left[i(k_b x + k_a x')\right] \exp\left[ik_b|x - x'|\right] \\ &+i\hbar(v_a v_b)^{1/2} \exp\left[-i(k_b x + k_a x')\right] \int_{\mathcal{S}_x} \mathrm{d}y \int_{\mathcal{S}_{x'}} \mathrm{d}y'\ \phi_b^*(y)\ \mathcal{G}^{(0)}(\mathbf{r}, \mathbf{r}', \varepsilon)\ \phi_a(y')\ . \end{aligned} \tag{2.61b}$$

$\mathcal{S}_{x'}$, $\mathcal{S}_x$ denote the cross sections for integrations. Their longitudinal position is chosen:

- always in the left lead for x',
- for x, inside the right lead in the transmission amplitude expression, and inside the left lead for the reflection amplitude expression.

Chapter 3

Theory of Scanning Gate Microscopy, Perturbative Approach

In this chapter we develop the theoretical description of Scanning Gate Microscopy. We take the system described so far, consisting of a coherent nanostructure connected to the reservoirs with perfect leads, and add the last missing element – the SGM tip (fig.2.1(d) on page 14). We include it in our description by the means of perturbation theory, where the influence of the tip is treated as a weak potential perturbing the rest of the system.

The basic question we would like to answer is how such an additional potential changes the linear-response conductance of the structure. In order to do it, we will first consider the form of the wave functions of the perturbed scattering states, then their current densities and currents, and finally the conductance, constructed from the perturbed current expressions in the same way as we constructed the conductance $g^{(0)}$ from the currents of the non-perturbed system in the previous chapter.

All the quantities appearing in the calculation will be expressed in terms of the unperturbed wave functions and system properties. In particular, the transmission-reflection amplitudes and the Green functions will pertain to the unperturbed system, and we will not mark it with any special, additional notation.

A similar task has been tackled in a one-dimensional case by Gasparian *et al.* [64]. The approach there took a slightly different path of arriving at the conductance change, namely by calculating the perturbed Green function by the means of the Dyson equation, then applying the Fisher-Lee relations to get the perturbed transmission and reflection amplitudes, and finally obtaining the conductance from them through the Landauer-Büttiker equation.

The one-dimensional case is considerably simpler as compared to the one we treat in here, since in 1D the scattering states always coincide with the scattering eigenstates of transmission, whereas in 2D such a coincidence is true only in special cases (cf. notion of adiabaticity in Ch. 4).

The basic form of the main results, ie. first and second order conductance change in an SGM system, have already been published in [45].[1] Here, we will explain the underlying calculations and discuss the results, with their origins, form, and consequences.

3.1 SGM tip influence on the wave functions

The Hamiltonian of the complete system can be written as a sum

$$H = H_0 + V_{\mathrm{T}} \tag{3.1}$$

where $H_0 = H_W + H_S$, called the unperturbed Hamiltonian, describes the previously treated leads-plus-scatterer system, and V_{T} is the electrostatic potential of the tip.

The eigenfunctions of the total Hamiltonian H, denoted by $\chi_{l\varepsilon a}(\mathbf{r})$, obey the Lippmann-Schwinger equation

$$\chi_{l\varepsilon a}(\mathbf{r}) = \Psi_{l\varepsilon a}(\mathbf{r}) + \int \mathrm{d}\mathbf{r}'\; \mathcal{G}^{(0)}(\mathbf{r},\mathbf{r}',\varepsilon) V_{\mathrm{T}}(\mathbf{r}') \chi_{l\varepsilon a}(\mathbf{r}') \;, \tag{3.2}$$

$\mathrm{d}\mathbf{r}'$ being the volume element. Here, we use the retarded Green function $\mathcal{G}^{(0)}$ and the outgoing scattering states $\Psi_{l\varepsilon a}$, both of the unperturbed Hamiltonian H_0. The $\chi_{l\varepsilon a}$ therefore represent the *outgoing* perturbed scattering states. Like before, the usual plus superscript is omitted.

As can be verified by direct substitution to the defining equation (2.60), the Green function decomposes in the basis of the scattering states $\Psi_{l\varepsilon a}$ in the following way:

$$\mathcal{G}^{(0)}(\mathbf{r},\mathbf{r}',E) = \sum_{l=1}^{2} \int_{\varepsilon_1^{(\perp)}}^{\infty} \frac{\mathrm{d}\varepsilon}{E^+ - \varepsilon} \sum_{a=1}^{N(\varepsilon)} \Psi_{l\varepsilon a}(\mathbf{r}) \Psi^*_{l\varepsilon a}(\mathbf{r}') \;. \tag{3.3}$$

The equation (3.2) gives $\chi_{l\varepsilon a}(\mathbf{r})$ in an open form. Substituting it recursively into the integral on the right-hand side gives the Born series:

$$\chi_{l\varepsilon a}(\mathbf{r}) = \chi^{(0)}_{l\varepsilon a}(\mathbf{r}) + \chi^{(1)}_{l\varepsilon a}(\mathbf{r}) + \chi^{(2)}_{l\varepsilon a}(\mathbf{r}) + \ldots \tag{3.4}$$

[1] The second-order expression in [45] contained a mistake, cf. (3.42).

with the subsequent terms given by:

$$\begin{aligned}
\chi^{(0)}_{l\varepsilon a}(\mathbf{r}) &= \Psi_{l\varepsilon a}(\mathbf{r}) \ , \\
\chi^{(1)}_{l\varepsilon a}(\mathbf{r}) &= \int \mathrm{d}\mathbf{r}' \ \mathcal{G}^{(0)}(\mathbf{r},\mathbf{r}',\varepsilon) V_{\mathrm{T}}(\mathbf{r}') \Psi_{l\varepsilon a}(\mathbf{r}') \ , \\
\chi^{(2)}_{l\varepsilon a}(\mathbf{r}) &= \int \mathrm{d}\mathbf{r}' \mathrm{d}\mathbf{r}'' \ \mathcal{G}^{(0)}(\mathbf{r},\mathbf{r}',\varepsilon) V_{\mathrm{T}}(\mathbf{r}') \mathcal{G}^{(0)}(\mathbf{r}',\mathbf{r}'',\varepsilon) V_{\mathrm{T}}(\mathbf{r}'') \Psi_{l\varepsilon a}(\mathbf{r}'') \ , \\
&\text{etc.}
\end{aligned} \tag{3.5}$$

or, in a general, recursive form:

$$\chi^{(n+1)}_{l\varepsilon a}(\mathbf{r}) = \int \mathrm{d}\mathbf{r}' \ \mathcal{G}^{(0)}(\mathbf{r},\mathbf{r}',\varepsilon) V_{\mathrm{T}}(\mathbf{r}') \chi^{(n)}_{l\varepsilon a}(\mathbf{r}') \, , \qquad n = 0, 1, \ldots \ . \tag{3.6}$$

The convergence of this series is generally not easy to establish. We will comment it in somewhat more detail in subsection 3.4.3. For the moment, we will restrict ourselves to the simple statement that for modest tip potentials V_{T} the series should converge. Truncating it at the n-th term gives approximate expressions for the wave functions $\chi_{l\varepsilon a}(\mathbf{r})$ of the corresponding order in the tip potential. Each $\chi^{(n)}_{l\varepsilon a}(\mathbf{r})$ is then the n-th order perturbation correction to the scattering state wave function.

Using the eigenfunction decomposition (3.3) of the Green function, we bring the first order correction to the wave function to a form

$$\begin{aligned}
\chi^{(1)}_{l\varepsilon a}(\mathbf{r}) &= \int \mathrm{d}\mathbf{r}' \sum_{l'=1}^{2} \int_{\varepsilon_1^{(\perp)}}^{\infty} \frac{\mathrm{d}\varepsilon'}{\varepsilon^+ - \varepsilon'} \sum_{a'=1}^{N(\varepsilon')} \Psi_{l'\varepsilon'a'}(\mathbf{r}) \, \Psi^*_{l'\varepsilon'a'}(\mathbf{r}') V_{\mathrm{T}}(\mathbf{r}') \Psi_{l\varepsilon a}(\mathbf{r}') \\
&= \sum_{l'=1}^{2} \int_{\varepsilon_1^{(\perp)}}^{\infty} \frac{\mathrm{d}\varepsilon'}{\varepsilon^+ - \varepsilon'} \sum_{a'=1}^{N(\varepsilon')} \Psi_{l'\varepsilon'a'}(\mathbf{r}) \ [V_{\mathrm{T}}]^{l'l}_{a'a}(\varepsilon', \varepsilon) \ ,
\end{aligned} \tag{3.7}$$

where

$$[V_{\mathrm{T}}]^{l'l}_{a'a}(\varepsilon', \varepsilon) = \int \mathrm{d}\mathbf{r}' \ \Psi^*_{l'\varepsilon'a'}(\mathbf{r}') \, V_{\mathrm{T}}(\mathbf{r}') \, \Psi_{l\varepsilon a}(\mathbf{r}') \tag{3.8}$$

is the matrix element of the tip potential in the scattering wave function basis. Analogically, for the second order we obtain

$$\begin{aligned}
\chi^{(2)}_{l\varepsilon a}(\mathbf{r}) = \sum_{l'=1}^{2} \sum_{l''=1}^{2} \int_{\varepsilon_1^{(\perp)}}^{\infty} \frac{\mathrm{d}\varepsilon'}{\varepsilon^+ - \varepsilon'} \int_{\varepsilon_1^{(\perp)}}^{\infty} \frac{\mathrm{d}\varepsilon''}{\varepsilon^+ - \varepsilon''} \\
\sum_{a'=1}^{N(\varepsilon')} \sum_{a''=1}^{N(\varepsilon'')} \Psi_{l'\varepsilon'a'}(\mathbf{r}) \ [V_{\mathrm{T}}]^{l'l''}_{a'a''}(\varepsilon', \varepsilon'') \ [V_{\mathrm{T}}]^{l''l}_{a''a}(\varepsilon'', \varepsilon) \ ,
\end{aligned} \tag{3.9}$$

and the higher order terms will follow this pattern, each next one getting one more matrix element and one more summation over states (see 3.50 below).

3.2 First order conductance correction

In order to calculate the conductance of the full SGM system, we have to repeat the reasoning described in the previous chapter, only with the new, perturbed, scattering states. Within the linear response approach, the conductance is constructed in a simple way from the currents carried by the states at the Fermi level (see sec. 2.5, Ch. 2). The linearity of this operation grants us that the perturbation correction to the conductance is made up from the state current corrections in the very same way.

We first investigate the perturbative corrections of the first order, so we will use as the scattering states

$$\chi_{l\varepsilon a}(\mathbf{r}) \approx \chi^{(0)}_{l\varepsilon a}(\mathbf{r}) + \chi^{(1)}_{l\varepsilon a}(\mathbf{r}) \ . \tag{3.10}$$

3.2.1 Current density of a scattering state

The object of our interest in this subsection will be the first order correction to the current density of a single scattering state $(l\varepsilon a)$, denoted by $\delta^{(1)}\mathbf{j}_{l\varepsilon a}(\mathbf{r})$. The current density of the perturbed states (3.10) reads:

$$\begin{aligned} &\mathbf{j}\left\{\chi^{(0)}_{l\varepsilon a} + \chi^{(1)}_{l\varepsilon a}\right\} = \\ &\qquad = \frac{e\hbar}{M_e}\,\mathrm{Im}\left(\chi^{(0)*}_{l\varepsilon a}\nabla\chi^{(0)}_{l\varepsilon a} + \chi^{(1)*}_{l\varepsilon a}\nabla\chi^{(0)}_{l\varepsilon a} + \chi^{(0)*}_{l\varepsilon a}\nabla\chi^{(1)}_{l\varepsilon a} + \chi^{(1)*}_{l\varepsilon a}\nabla\chi^{(1)}_{l\varepsilon a}\right), \end{aligned} \tag{3.11}$$

(the argument $\mathbf{r}$ is understood). After removing the first addend on the right-hand side, which is just the unperturbed current density $\mathbf{j}\left\{\chi^{(0)}_{l\varepsilon a}\right\}$, and the last addend, which is of second order and will enter the equations of sec. 3.3.1, we complex-conjugate the $\chi^{(1)*}_{l\varepsilon a}\nabla\chi^{(0)}_{l\varepsilon a}$ term and substitute the form (3.7) for the wave function correction, to obtain

$$\begin{aligned} &\delta^{(1)}\mathbf{j}_{l\varepsilon a} = \\ &= \frac{e\hbar}{M_e}\,\mathrm{Im}\left[\sum_{l'=1}^{2}\int_{\varepsilon_1^{(\perp)}}^{\infty}\frac{\mathrm{d}\varepsilon'}{\varepsilon^+ - \varepsilon'}\sum_{a'=1}^{N(\varepsilon')}\left(\Psi^*_{l\varepsilon a}\nabla\Psi_{l'\varepsilon' a'} - \Psi_{l'\varepsilon' a'}\nabla\Psi^*_{l\varepsilon a}\right)\,\left[V_{\mathrm{T}}\right]^{l'l}_{a'a}(\varepsilon',\varepsilon)\right], \end{aligned} \tag{3.12}$$

where we recognise the matrix elements of the current density operator in the unperturbed scattering state basis $\langle\Psi_{l\varepsilon a}\,|\,\hat{\mathbf{j}}(\mathbf{r})|\,\Psi_{l'\varepsilon' a'}\rangle$ – arranging the indices according to the convention we used for the potential matrix elements (3.8), we will write them as $[\mathbf{j}(\mathbf{r})]^{l\,l'}_{aa'}(\varepsilon,\varepsilon')$. We therefore have:

$$\delta^{(1)}\mathbf{j}_{l\varepsilon a} = 2\,\mathrm{Re}\left[\sum_{l'=1}^{2}\int_{\varepsilon_1^{(\perp)}}^{\infty}\frac{\mathrm{d}\varepsilon'}{\varepsilon^+ - \varepsilon'}\sum_{a'=1}^{N(\varepsilon')}[\mathbf{j}(\mathbf{r})]^{l\,l'}_{aa'}(\varepsilon,\varepsilon')\,\left[V_{\mathrm{T}}\right]^{l'l}_{a'a}(\varepsilon',\varepsilon)\right], \tag{3.13}$$

with

$$\left[\mathbf{j}(\mathbf{r})\right]^{ll'}_{aa'}(\varepsilon,\varepsilon') = \frac{e\hbar}{2iM_e}\left[\Psi^*_{l\varepsilon a}(\mathbf{r})\nabla\Psi_{l'\varepsilon' a'}(\mathbf{r}) - \Psi_{l'\varepsilon' a'}(\mathbf{r})\nabla\Psi^*_{l\varepsilon a}(\mathbf{r})\right]. \tag{3.14}$$

The current density matrix elements (3.14) contain all the spatial dependence of the $\delta^{(1)}\mathbf{j}_{l\varepsilon a}$ expression.[2] Its elements diagonal in all indices ($l\varepsilon a = l'\varepsilon' a'$) give the currents of each individual unperturbed scattering state.

Within the Landauer approach, we need the current injected from one side only, so we set the lead of origin as the left one, $l = 1$. Moreover, due to the current conservation, the current does not depend on which intersecting surface $\mathcal{S}$ we choose for the current density integration. We choose it in the right lead for definiteness, and we choose it to be a planar section $\mathcal{S}_x$ perpendicular to the lead's axis, so that we need not bother with the transverse current density components. Later we will also specify its position x to go to infinity.

By substituting the explicit form of the unperturbed scattering state wave functions – the eqs. (2.25a) and (2.2) – into the above, we obtain the current density matrix elements in terms of the basic properties of the scatterer (transmission and reflection amplitudes) and of the wire-lead (transverse wave functions, dispersion relations). Their longitudinal components are:

$$\left[j^{(x)}(\mathbf{r})\right]^{11}_{aa'}(\varepsilon,\varepsilon') = \tag{3.15a}$$
$$= \frac{e}{2h}\sum_{b=1}^{N(\varepsilon)}\sum_{b'=1}^{N(\varepsilon')}\left(\sqrt{\frac{k'_{b'}}{k_b}} + \sqrt{\frac{k_b}{k'_{b'}}}\right) t^\dagger_{ab}(\varepsilon)t_{b'a'}(\varepsilon')\,\phi_b(y)\phi_{b'}(y)\exp\left[i(k'_{b'} - k_b)x\right],$$

$$\left[j^{(x)}(\mathbf{r})\right]^{12}_{aa'}(\varepsilon,\varepsilon') = \tag{3.15b}$$
$$= \frac{e}{2h}\left\{\sum_{b=1}^{N(\varepsilon)}\left(\sqrt{\frac{k_b}{k'_{a'}}} - \sqrt{\frac{k'_{a'}}{k_b}}\right) t^\dagger_{ab}(\varepsilon)\,\phi_b(y)\phi_{a'}(y)\exp\left[-i(k'_{a'} + k_b)x\right]\right.$$
$$\left. + \sum_{b=1}^{N(\varepsilon)}\sum_{b'=1}^{N(\varepsilon')}\left(\sqrt{\frac{k'_{b'}}{k_b}} + \sqrt{\frac{k_b}{k'_{b'}}}\right) t^\dagger_{ab}(\varepsilon)r'_{b'a'}(\varepsilon')\,\phi_b(y)\phi_{b'}(y)\exp\left[i(k'_{b'} - k_b)x\right]\right\}.$$

As usual, the primed wave numbers are defined through the dispersion relation (2.6): $k'_{\cdot} \equiv k_{\cdot}(\varepsilon')$.

[2] Not to be confused with the dependence on the SGM tip's position, which is contained solely in the tip potential matrix elements (3.8).

3.2.2 Current of a scattering state

For the longitudinal current correction we need to integrate its density (3.13) over the cross section, getting an analogous expression:

$$\delta^{(1)}J^{(x)}_{1\varepsilon a} = \int_{\mathcal{S}_x} \mathrm{d}y\, \delta^{(1)}\mathbf{j}_{1\varepsilon a} = \tag{3.16}$$
$$= 2\mathrm{Re}\left[\sum_{l'=1}^{2}\int_{\varepsilon_1^{(\perp)}}^{\infty}\frac{\mathrm{d}\varepsilon'}{\varepsilon^+-\varepsilon'}\sum_{a'=1}^{N(\varepsilon')}\left[J^{(x)}\right]^{1\,l'}_{aa'}(\varepsilon,\varepsilon')\ \left[V_{\mathrm{T}}\right]^{l'1}_{a'a}(\varepsilon',\varepsilon)\right]$$

with the y-integrated current density matrix elements

$$\left[J^{(x)}\right]^{1\,1}_{aa'}(\varepsilon,\varepsilon') = \frac{e}{2h}\sum_{b=1}^{\bar{N}}\left(\sqrt{\frac{k'_b}{k_b}}+\sqrt{\frac{k_b}{k'_b}}\right)t^{\dagger}_{ab}(\varepsilon)t_{ba'}(\varepsilon')\exp\left[i(k'_b-k_b)x\right]\,,$$

$$\left[J^{(x)}\right]^{1\,2}_{aa'}(\varepsilon,\varepsilon') = \frac{e}{2h}\left\{\left(\sqrt{\frac{k_{a'}}{k'_{a'}}}-\sqrt{\frac{k'_{a'}}{k_{a'}}}\right)t^{\dagger}_{aa'}(\varepsilon)\exp\left[-i(k'_{a'}+k_{a'})x\right]\right. \tag{3.17}$$
$$\left.+\sum_{b=1}^{\bar{N}}\left(\sqrt{\frac{k'_b}{k_b}}+\sqrt{\frac{k_b}{k'_b}}\right)t^{\dagger}_{ab}(\varepsilon)r'_{ba'}(\varepsilon')\exp\left[i(k'_b-k_b)x\right]\right\},$$

where the orthonormality of the transverse modes (eq. (2.9)) has been used, and $\bar{N}\equiv\min\{N(\varepsilon),N(\varepsilon')\}$.

Next, we take the limit $x\to\infty$ and use the oscillating $\exp(ikx)$ terms of the $[J^{(x)}]$-elements together with $1/(\varepsilon^+-\varepsilon')$ factors in order to facilitate the $\mathrm{d}\varepsilon'$ integration. We also assume that the variation of the elements of $[V_{\mathrm{T}}]$ and the transmission and reflection matrices is smooth with respect to the appropriate wave numbers. The details of this calculation are explained in the Appendix A. Denoting those submatrices of the entire $[V_{\mathrm{T}}]$-matrix, which are diagonal in energy, by $\mathcal{V}$:

$$\mathcal{V}^{ll'}_{\cdot\cdot} \equiv \left[V_{\mathrm{T}}\right]^{ll'}_{\cdot\cdot}(\varepsilon,\varepsilon)\,, \tag{3.18}$$

we write the resulting first order correction to the current of a single mode as

$$\delta^{(1)}J^{(x)}_{1\varepsilon a} = 4\pi\frac{e}{h}\sum_{a'=1}^{N(\varepsilon)}\mathrm{Im}\Big[(t^{\dagger}t)_{aa'}\,\mathcal{V}^{1\,1}_{a'a}+(t^{\dagger}r')_{aa'}\,\mathcal{V}^{2\,1}_{a'a}\Big]$$
$$= 4\pi\frac{e}{h}\,\mathrm{Im}\Big[\left(t^{\dagger}t\,\mathcal{V}^{1\,1}+t^{\dagger}r'\,\mathcal{V}^{2\,1}\right)_{aa}\Big]\,. \tag{3.19}$$

3.2.3 Total current

Just like we did in the section 2.5, we calculate the first order correction to the total current, $\delta^{(1)}I$, by summing over all the states (including the degenerate spin-up and spin-down states). We disregard the thermal broadening for simplicity. The only relevant states are the left-originating scattering states within the "Fermi window", i.e. between μ_1 and μ_2 in energy:

$$\delta^{(1)}I = 2\int_{\mu_2}^{\mu_1} \mathrm{d}\varepsilon \sum_a \delta^{(1)}J^{(x)}_{1\varepsilon a} \,. \tag{3.20}$$

The intermediate sum, $\delta^{(1)}J^{(x)}_{1\varepsilon} \equiv \sum_a \delta^{(1)}J^{(x)}_{1\varepsilon a}$, is the first order total current correction per unit energy and per spin. It reads

$$\delta^{(1)}J^{(x)}_{1\varepsilon} = 4\pi\frac{e}{h}\mathrm{Im}\Big[\mathrm{tr}\left(t^\dagger t\,\mathcal{V}^{11} + t^\dagger r'\,\mathcal{V}^{21}\right)\Big] \,. \tag{3.21}$$

Recalling the definition of the matrices $\mathcal{V}^{ll'}$, eqs. (3.9) and (3.18), we note that for $l = l'$ they are hermitian. Using this, together with the invariance of trace operation under cyclic permutation of factors and under transposition, we easily prove that the first addend in the above is purely real:

$$\mathrm{tr}\left(t^\dagger t\,\mathcal{V}^{11}\right) = \mathrm{tr}\left(\mathcal{V}^{11}\,t^\dagger t\right) = \mathrm{tr}\left(\mathcal{V}^{11}\,t^\dagger t\right)^{\mathrm{T}} = \mathrm{tr}\left[(t^\dagger t)^*\,\mathcal{V}^{11*}\right] = \mathrm{tr}\left(t^\dagger t\,\mathcal{V}^{11}\right)^*,$$

and consequently disappears from the expression, leaving

$$\delta^{(1)}J^{(x)}_{1\varepsilon} = 4\pi\frac{e}{h}\mathrm{Im}\Big[\mathrm{tr}\left(t^\dagger r'\,\mathcal{V}^{21}\right)\Big] \,. \tag{3.22}$$

3.2.4 Conductance

The above calculated current change gives rise to the first order correction to the linear response conductance (cf. (2.58))

$$\begin{aligned} G^{(1)} = \frac{2e^2}{h}g^{(1)} &= \frac{\partial}{\partial V_b}\left(\delta^{(1)}I\right) = 2e\,\frac{\partial}{\partial(\mu_1-\mu_2)}\int_{\mu_2}^{\mu_1}\mathrm{d}\varepsilon\,\delta^{(1)}J^{(x)}_{1\varepsilon} \\ &= 4\pi\frac{2e^2}{h}\mathrm{tr}\Big[\mathrm{Im}\left(t^\dagger r'\,\mathcal{V}^{21}\right)\Big]_{\varepsilon=E_F} \,. \end{aligned} \tag{3.23}$$

By (2.37b), an equivalent form is:

$$g^{(1)} = -4\pi\mathrm{tr}\Big[\mathrm{Im}\left(r^\dagger t'\,\mathcal{V}^{21}\right)\Big]_{\varepsilon=E_F} \,. \tag{3.24}$$

This expression approximates the SGM signal for small tip potential values – up to the first order in the tip potential strength. It consists of the $t^\dagger r'$ (or

$r^\dagger t'$) factors, depending solely on the unperturbed structure, and the matrix elements $\mathcal{V}$ which describe the tip's influence. The latter also depend on the structure through the scattering states $\Psi_{l\varepsilon a}$ which they contain. All the terms are evaluated at the Fermi energy, which we will tacitly assume throughout the rest of this work.

Scattering eigenstates basis

It turns out, that we can express this result in a way which provides for a more immediate and simple interpretation. With the aid of the S-matrix polar decomposition (sec. 2.3.3), we can render the conductance correction into the transmission eigenstates basis $\left\{\overset{\approx}{\Psi}_{l\varepsilon m}\right\}$ of eq. (B.12):

$$g^{(1)} = 4\pi \sum_{m=1}^{N} \mathcal{R}_m \mathcal{T}_m \operatorname{Im}\left\{\mathcal{U}^{21}_{mm}\right\} \ . \tag{3.25}$$

Here, $\mathcal{U}$ are the tip potential matrix elements in the new basis:

$$\mathcal{U}^{ll'} = \int \mathrm{d}\mathbf{r}\ \overset{\approx}{\Psi}{}^{*}_{l\varepsilon.}\, V_{\mathrm{T}}\, \overset{\approx}{\Psi}_{l'\varepsilon.} = u_l\, \mathcal{V}^{ll'} u^\dagger_{l'} \ . \tag{3.26}$$

Transmission eigenmodes basis

If we choose the tip potential to be localized entirely outside the scatterer region (i.e. entirely inside one of the leads), we can go one step further and use the basis of the transmission eigenmodes (B.11). According to (2.47a–2.47b), we have for the tip placed to the right of the scatterer:

$$\mathcal{U}^{21}_{mm'} = \mathcal{T}_m \left\langle \overset{\approx}{\varphi}{}^{(+)}_{1\varepsilon m} \middle| V_{\mathrm{T}} \middle| \overset{\approx}{\varphi}{}^{(-)}_{1\varepsilon m'} \right\rangle - \mathcal{R}_{m'} \mathcal{T}_m \left\langle \overset{\approx}{\varphi}{}^{(+)}_{1\varepsilon m} \middle| V_{\mathrm{T}} \middle| \overset{\approx}{\varphi}{}^{(+)}_{1\varepsilon m'} \right\rangle \ , \tag{3.27a}$$

and for the tip placed to the left of it:

$$\mathcal{U}^{21}_{mm'} = \mathcal{T}_{m'} \left\langle \overset{\approx}{\varphi}{}^{(-)}_{2\varepsilon m} \middle| V_{\mathrm{T}} \middle| \overset{\approx}{\varphi}{}^{(+)}_{2\varepsilon m'} \right\rangle + \mathcal{R}_m \mathcal{T}_{m'} \left\langle \overset{\approx}{\varphi}{}^{(+)}_{2\varepsilon m} \middle| V_{\mathrm{T}} \middle| \overset{\approx}{\varphi}{}^{(+)}_{2\varepsilon m'} \right\rangle \ , \tag{3.27b}$$

For $m = m'$, it is easy to see that the rightmost terms in both of the above lines have no imaginary part. The conductance correction $g^{(1)}$ then assumes the form

$$\begin{aligned} g^{(1)} &= 4\pi \sum_{m=1}^{N} \mathcal{R}_m \mathcal{T}_m^2 \operatorname{Im}\left\{ \left\langle \overset{\approx}{\varphi}{}^{(+)}_{1\varepsilon m} \middle| V_{\mathrm{T}} \middle| \overset{\approx}{\varphi}{}^{(-)}_{1\varepsilon m} \right\rangle \right\} \qquad (3.28a) \\ &= 4\pi \sum_{m=1}^{N} \mathcal{R}_m \mathcal{T}_m^2 \operatorname{Im}\left\{ \int \mathrm{d}\mathbf{r}\, V_{\mathrm{T}}(\mathbf{r})\, \overset{\approx}{\varphi}{}^{(-)\,2}_{1\varepsilon m}(\mathbf{r}) \right\} \ , \end{aligned}$$

with the tip inside the left lead, and

$$g^{(1)} = 4\pi \sum_{m=1}^{N} \mathcal{R}_m \mathcal{T}_m^2 \, \mathrm{Im} \left\{ \left\langle \overset{\approx}{\varphi}{}^{(+)}_{1\varepsilon m} \middle| V_{\mathrm{T}} \middle| \overset{\approx}{\varphi}{}^{(-)}_{1\varepsilon m} \right\rangle \right\} \tag{3.28b}$$
$$= 4\pi \sum_{m=1}^{N} \mathcal{R}_m \mathcal{T}_m^2 \, \mathrm{Im} \left\{ \int \mathrm{d}\mathbf{r} \, V_{\mathrm{T}}(\mathbf{r}) \, \overset{\approx}{\varphi}{}^{(+)\,2}_{2\varepsilon m}(\mathbf{r}) \right\} \, ,$$

with the tip inside the right one. This form of the conductance correction result will be useful at the time of drawing conclusions about the shape of the SGM signal in the leads region, or the broad regions which the leads could model.

The previous form (3.25), on the other hand, is very important in the general context, because it addresses directly the transmission process, with its amplitudes and eigenstates. Each term of the sum expresses the contribution from one individual transmission eigenstate. We can see, that whenever an eigenmode from one lead is:

– fully transferred by the structure, $\mathcal{T}_m \cong 1$, $\mathcal{R}_m = \sqrt{1 - \mathcal{T}_m^2} \cong 0$,

– or fully stopped by it, $\mathcal{T}_m \cong 0$, $\mathcal{R}_m \cong 1$,

its contribution to $g^{(1)}$ vanishes, regardless of the tip position. It can, therefore, occur that none of the channels contributes,

$$\mathcal{R}\mathcal{T} = \mathbf{0} \, , \tag{3.29}$$

in which case the first order formula is not enough for the correct perturbative description of the SGM signal. Thus, before discussing the results (3.23-3.28) in more detail, we turn to the derivation of the second order correction to the conductance.

3.3 Second order current correction

3.3.1 Current density and Current – single mode

The calculation of the current density correction of the second order, $\delta^{(2)}\mathbf{j}_{l\varepsilon a}$, follows the same path as of the first order one. We write the current density of the approximate state $\chi^{(0)}_{l\varepsilon a} + \chi^{(1)}_{l\varepsilon a} + \chi^{(2)}_{l\varepsilon a}$ and dispose of all the terms whose order is different than two (i.e. zeroth, first, third and fourth):

$$\delta^{(2)}\mathbf{j}_{1\varepsilon a} = \frac{e\hbar}{M_e} \, \mathrm{Im} \left(\chi^{(0)*}_{1\varepsilon a} \nabla \chi^{(2)}_{1\varepsilon a} + \chi^{(1)*}_{1\varepsilon a} \nabla \chi^{(1)}_{1\varepsilon a} + \chi^{(2)*}_{1\varepsilon a} \nabla \chi^{(0)}_{1\varepsilon a} \right) \tag{3.30}$$

Like before, all are functions of $\mathbf{r}$. This expression is then subdivided into two types of terms,

$$\delta^{(2)}\mathbf{j}_{1\varepsilon a} = \delta^{(2)}\mathbf{j}^{\alpha}_{1\varepsilon a} + \delta^{(2)}\mathbf{j}^{\beta}_{1\varepsilon a}\,, \tag{3.31}$$

in the following way:

$$\delta^{(2)}\mathbf{j}^{\alpha}_{1\varepsilon a} = \frac{e\hbar}{M_e}\,\mathrm{Im}\left(\chi^{(0)*}_{1\varepsilon a}\nabla\chi^{(2)}_{1\varepsilon a} - \chi^{(2)}_{1\varepsilon a}\nabla\chi^{(0)*}_{1\varepsilon a}\right)\,, \tag{3.32a}$$

$$\delta^{(2)}\mathbf{j}^{\beta}_{1\varepsilon a} = \frac{e\hbar}{M_e}\,\mathrm{Im}\left(\chi^{(1)*}_{1\varepsilon a}\nabla\chi^{(1)}_{1\varepsilon a}\right)\,. \tag{3.32b}$$

The form of the perturbed states is now taken from both, eq. (3.7) and eq. (3.9); the integration of the above expressions over a cross section $\mathcal{S}_x$ to the right of the scatterer yields the current correction with

$$\delta^{(2)}J^{(x)\alpha}_{1\varepsilon a} = 2\mathrm{Re}\left[\sum_{l',l''=1}^{2}\int_{\varepsilon_1^{(\perp)}}^{\infty}\frac{\mathrm{d}\varepsilon'}{\varepsilon^+-\varepsilon'}\int_{\varepsilon_1^{(\perp)}}^{\infty}\frac{\mathrm{d}\varepsilon''}{\varepsilon^+-\varepsilon''}\right. \tag{3.33a}$$
$$\left.\sum_{a'=1}^{N(\varepsilon')}\sum_{a''=1}^{N(\varepsilon'')}\left[J^{(x)}\right]^{1\,l'}_{aa'}(\varepsilon,\varepsilon')\;[V_{\mathrm{T}}]^{l'l''}_{a'a''}(\varepsilon',\varepsilon'')\;[V_{\mathrm{T}}]^{l''1}_{a''a}(\varepsilon'',\varepsilon)\right]$$

and

$$\delta^{(2)}J^{(x)\beta}_{1\varepsilon a} = 2\mathrm{Re}\left[\sum_{l',l''=1}^{2}\int_{\varepsilon_1^{(\perp)}}^{\infty}\frac{\mathrm{d}\varepsilon'}{\varepsilon^--\varepsilon'}\int_{\varepsilon_1^{(\perp)}}^{\infty}\frac{\mathrm{d}\varepsilon''}{\varepsilon^+-\varepsilon''}\right. \tag{3.33b}$$
$$\left.\sum_{a'=1}^{N(\varepsilon')}\sum_{a''=1}^{N(\varepsilon'')}[V_{\mathrm{T}}]^{1l'}_{aa'}(\varepsilon,\varepsilon')\;[X]^{l'\,l''}_{a'a''}(\varepsilon',\varepsilon'')\;[V_{\mathrm{T}}]^{l''1}_{a''a}(\varepsilon'',\varepsilon)\right]\,.$$

The $\left[J^{(x)}\right]$-factors have been given in (3.17) above; and the $[X]$-factors are the y-integrated longitudinal components of the matrix elements

$$\left\langle\Psi_{l\varepsilon a}\left|\,\hat{\mathbf{j}}_{\mathbf{1/2}}(\mathbf{r})\right|\Psi_{l'\varepsilon' a'}\right\rangle \tag{3.34}$$

of the operator

$$\hat{\mathbf{j}}_{\mathbf{1/2}}(\mathbf{r}) = \frac{e}{2M_e}\delta(\hat{\mathbf{r}}-\mathbf{r})\hat{\mathbf{p}}\,, \tag{3.35}$$

which could be described as a "half" of the current density operator. Their specific form is

$$[X]^{1\,1}_{a'a''}(\varepsilon',\varepsilon'') = \frac{e}{2h}\sum_{b=1}^{\bar{N}}\sqrt{\frac{k''_b}{k'_b}}\;t^{\dagger}_{a'b}(\varepsilon')t_{ba''}(\varepsilon'')\;\exp\left[i(k''_b-k'_b)x\right]\,, \tag{3.36a}$$

$$[X]^{1\,2}_{a'a''}(\varepsilon',\varepsilon'') = \frac{e}{2h}\left\{-\sqrt{\frac{k''_{a''}}{k'_{a''}}}\; t^{\dagger}_{a'a''}(\varepsilon')\;\exp\left[-i(k''_{a''}+k'_{a''})x\right]\right. \tag{3.36b}$$

$$\left. + \sum_{b=1}^{\bar{N}} \sqrt{\frac{k''_b}{k'_b}}\; t^{\dagger}_{a'b}(\varepsilon') r'_{ba''}(\varepsilon'')\;\exp\left[i(k''_b-k'_b)x\right]\right\},$$

$$[X]^{2\,1}_{a'a''}(\varepsilon',\varepsilon'') = \frac{e}{2h}\left\{\sqrt{\frac{k''_{a'}}{k'_{a'}}}\; t_{a'a''}(\varepsilon'')\;\exp\left[i(k''_{a'}+k'_{a'})x\right]\right. \tag{3.36c}$$

$$\left. + \sum_{b=1}^{\bar{N}} \sqrt{\frac{k''_b}{k'_b}}\; r'^{\dagger}_{a'b}(\varepsilon') t_{ba''}(\varepsilon'')\;\exp\left[i(k''_b-k'_b)x\right]\right\},$$

$$[X]^{2\,2}_{a'a''}(\varepsilon',\varepsilon'') = \frac{e}{2h}\left\{-\delta_{a'a''}\;\sqrt{\frac{k''_{a'}}{k'_{a'}}}\;\exp\left[-i(k''_{a''}-k'_{a'})x\right]\right. \tag{3.36d}$$

$$\begin{aligned}
&+\;\sqrt{\frac{k''_{a'}}{k'_{a'}}}\; r'_{a'a''}(\varepsilon'')\exp\left[i(k''_{a'}+k'_{a'})x\right]\\
&-\;\sqrt{\frac{k''_{a''}}{k'_{a''}}}\; r'^{\dagger}_{a'a''}(\varepsilon')\exp\left[-i(k''_{a''}+k'_{a''})x\right]\\
&+\;\left.\sum_{b=1}^{\bar{N}}\sqrt{\frac{k''_b}{k'_b}}\; r'^{\dagger}_{a'b}(\varepsilon') r'_{ba''}(\varepsilon'')\;\exp\left[i(k''_b-k'_b)x\right]\right\}\;.
\end{aligned}$$

After the energy integrations in (3.33a) and (3.33b), conducted as described in the Appendix A, point 3 and point 2, we obtain for the α-terms:

$$\begin{aligned}
\delta^{(2)}J^{(x)\alpha}_{1\varepsilon a} = &-4\pi^2\,\frac{e}{h}\,\mathrm{Re}\Big\{\Big[t^{\dagger}t\left(\mathcal{V}^{1\,1}\mathcal{V}^{1\,1}+\mathcal{V}^{1\,2}\mathcal{V}^{2\,1}\right) \qquad (3.37)\\
&\qquad + t^{\dagger}r'\left(\mathcal{V}^{2\,1}\mathcal{V}^{1\,1}+\mathcal{V}^{2\,2}\mathcal{V}^{2\,1}\right)\Big]_{aa}\Big\}\\
&+4\pi\frac{e}{h}\sum_{l''=1}^{2}\mathrm{Im}\Big\{\mathcal{P}\int_{\varepsilon_1^{(\perp)}}^{\infty}\frac{\mathrm{d}\varepsilon''}{\varepsilon-\varepsilon''}\Big[t^{\dagger}(\varepsilon)t(\varepsilon)\;[V_{\mathrm{T}}]^{1l''}(\varepsilon,\varepsilon'')\;[V_{\mathrm{T}}]^{l''1}(\varepsilon'',\varepsilon)\\
&\qquad + t^{\dagger}(\varepsilon)r'(\varepsilon)\;[V_{\mathrm{T}}]^{2l''}(\varepsilon,\varepsilon'')\;[V_{\mathrm{T}}]^{l''1}(\varepsilon'',\varepsilon)\Big]_{aa}\Big\}\;,
\end{aligned}$$

where $\mathcal{P}$ denotes the Cauchy principal value, and the energy arguments have been given explicitly in the second part to distinguish the integration variable from the fixed one.

Likewise, integration of the β-terms yields:

$$\delta^{(2)} J_{1\varepsilon a}^{(x)\beta} = 4\pi^2 \frac{e}{h} \mathrm{Re}\Big[\big(\mathcal{V}^{1\,1}\, t^\dagger t\, \mathcal{V}^{1\,1} + \mathcal{V}^{1\,1}\, t^\dagger r'\, \mathcal{V}^{2\,1} + \mathcal{V}^{1\,2}\, r'^\dagger t\, \mathcal{V}^{1\,1} + \mathcal{V}^{1\,2}\, r'^\dagger r'\, \mathcal{V}^{2\,1}\big)_{aa}\Big] . \tag{3.38}$$

With the invariance of the form $\mathrm{Re}(\ldots)_{aa}$ under hermitian conjugation, the third addend turns out to be equal to the second one and the formula takes the form

$$\delta^{(2)} J_{1\varepsilon a}^{(x)\beta} = 4\pi^2 \frac{e}{h} \mathrm{Re}\Big[\big(\mathcal{V}^{1\,1} t^\dagger t\, \mathcal{V}^{1\,1} + 2\mathcal{V}^{1\,1} t^\dagger r'\, \mathcal{V}^{2\,1} + \mathcal{V}^{1\,2} r'^\dagger r'\, \mathcal{V}^{2\,1}\big)_{aa}\Big] . \tag{3.39}$$

3.3.2 Total Current

To obtain the second order current correction, we add up the α- and β-terms, and sum over the scattering states. Using cyclic rearrangements of terms, transposition and complex conjugation, we verify that:

- the first terms of the α- and β-parts cancel each other,
- the third term of the α-part cancels with the factor of two from the middle term of the β-part,
- the second term of the α- and the last term of the β-part are purely real, so the $\mathrm{Re}(\ldots)$ symbol can be dropped,
- the first term (both l''-addends) of the principal value in (3.37) is real and therefore vanishes when the imaginary part is taken.

Thus, we arrive at the current correction per unit energy:

$$\begin{aligned}\delta^{(2)} J_{1\varepsilon}^{(x)} = 4\pi^2 \frac{e}{h} \mathrm{tr}\Big\{ & r'^\dagger r'\, \mathcal{V}^{2\,1}\mathcal{V}^{1\,2} - t^\dagger t\, \mathcal{V}^{1\,2}\mathcal{V}^{2\,1} \\ & + \mathrm{Re}\big[t^\dagger r'\, \big(\mathcal{V}^{2\,1}\mathcal{V}^{1\,1} - \mathcal{V}^{2\,2}\mathcal{V}^{2\,1}\big)\big]\Big\} \\ + 4\pi \frac{e}{h} \mathrm{tr}\Big\{ & \mathcal{P}\int_{\varepsilon_1^{(\perp)}}^{\infty} \frac{\mathrm{d}\varepsilon''}{\varepsilon - \varepsilon''} \sum_{l''=1}^{2} \mathrm{Im}\big[t^\dagger r'\, [V_\mathrm{T}]^{2\,l''}(\varepsilon, \varepsilon'')\, [V_\mathrm{T}]^{l''\,1}(\varepsilon'', \varepsilon)\big]\Big\}\end{aligned} \tag{3.40}$$

to which the second order correction to the total current is connected by the usual relation:

$$\delta^{(2)} I = \int_{\mu_2}^{\mu_1} \mathrm{d}\varepsilon\, \delta^{(2)} J_{1\varepsilon}^{(x)} . \tag{3.41}$$

3.3.3 Conductance

Differentiating the above with respect to the bias voltage according to the usual linear response procedure, we get the second order conductance correction in its full form:

$$\begin{aligned}\frac{1}{2e^2/h}G^{(2)} = g^{(2)} = \\ = 4\pi^2 \operatorname{tr}\Big\{ r'^{\dagger} r' \, \mathcal{V}^{2\,1}\mathcal{V}^{1\,2} - t^{\dagger}t \, \mathcal{V}^{1\,2}\mathcal{V}^{2\,1} + \operatorname{Re}\big[t^{\dagger} r' \left(\mathcal{V}^{2\,1}\mathcal{V}^{1\,1} - \mathcal{V}^{2\,2}\mathcal{V}^{2\,1} \right) \big] \Big\}_{\varepsilon = E_F} \\ + 4\pi \operatorname{tr}\left\{ \mathcal{P}\int_{\varepsilon_1^{(\perp)}}^{\infty} \frac{\mathrm{d}\varepsilon''}{E_F - \varepsilon''} \sum_{l''=1}^{2} \operatorname{Im}\big[t^{\dagger} r' \, [V_{\mathrm{T}}]^{2\,l''}(E_F, \varepsilon'') \, [V_{\mathrm{T}}]^{l''\,1}(\varepsilon'', E_F) \big] \right\}.\end{aligned} \tag{3.42}$$

If we take into account that the second order correction could dominate the first order one only when the multiplier $t^{\dagger}r' = u_1^{\dagger}\mathcal{R}\mathcal{T}u_2$ vanishes (see (3.29)), we can immediately simplify this expression to

$$g^{(2)} = 4\pi^2 \operatorname{tr}\Big\{ r'^{\dagger} r' \, \mathcal{V}^{2\,1}\mathcal{V}^{1\,2} - t^{\dagger}t \, \mathcal{V}^{1\,2}\mathcal{V}^{2\,1} \Big\}. \tag{3.43}$$

Scattering eigenstates basis

After transforming the previous equation into the scattering eigenstates basis (B.12), we have:

$$g^{(2)} = 4\pi^2 \sum_{m} \left[\mathcal{R}_m^2 \left(\mathcal{U}^{2\,1}\mathcal{U}^{1\,2} \right)_{mm} - \mathcal{T}_m^2 \left(\mathcal{U}^{1\,2}\mathcal{U}^{2\,1} \right)_{mm} \right]. \tag{3.44}$$

Once again, we are going to use the fact that this formula is relevant when the condition (3.29) is met.

We now substitute the approximate one-or-zero values for $\mathcal{R}_m$ and $\mathcal{T}_m$, and write the matrix multiplications explicitly. Then, we swap the index names within the second term ($m \leftrightarrow m'$), and carry out the subtraction:

$$\begin{aligned} g^{(2)} &= 4\pi^2 \sum_{\substack{m\in\overline{\mathcal{M}} \\ m'\in\underline{\mathcal{M}}\cup\overline{\mathcal{M}}}} \mathcal{U}^{2\,1}_{mm'}\,\mathcal{U}^{1\,2}_{m'm} - 4\pi^2 \sum_{\substack{m\in\underline{\mathcal{M}} \\ m'\in\underline{\mathcal{M}}\cup\overline{\mathcal{M}}}} \mathcal{U}^{1\,2}_{mm'}\,\mathcal{U}^{2\,1}_{m'm} \\ &= 4\pi^2 \sum_{\substack{m\in\overline{\mathcal{M}} \\ m'\in\underline{\mathcal{M}}\cup\overline{\mathcal{M}}}} \mathcal{U}^{2\,1}_{mm'}\,\mathcal{U}^{1\,2}_{m'm} - 4\pi^2 \sum_{\substack{m\in\underline{\mathcal{M}}\cup\overline{\mathcal{M}} \\ m'\in\underline{\mathcal{M}}}} \mathcal{U}^{2\,1}_{mm'}\,\mathcal{U}^{1\,2}_{m'm} \\ &= 4\pi^2 \sum_{\substack{m\in\overline{\mathcal{M}} \\ m'\in\overline{\mathcal{M}}}} \mathcal{U}^{2\,1}_{mm'}\,\mathcal{U}^{1\,2}_{m'm} - 4\pi^2 \sum_{\substack{m\in\underline{\mathcal{M}} \\ m'\in\underline{\mathcal{M}}}} \mathcal{U}^{2\,1}_{mm'}\,\mathcal{U}^{1\,2}_{m'm}\,. \end{aligned} \tag{3.45}$$

This expression is valid for all tip positions, but once we decide to place the tip inside a lead, the first term will be suppressed: by the presence of the $\overset{\approx}{\Psi}_{2\varepsilon m}$ factors when the tip is in the left lead, and by the presence of the $\overset{\approx}{\Psi}_{1\varepsilon m'}$ factors when the tip is in the right one. The second order correction is then written as

$$\begin{aligned} g^{(2)} &= -4\pi^2 \sum_{m,m' \in \underline{\mathcal{M}}} \left| \mathcal{U}^{2\,1}_{mm'} \right|^2 \\ &= -4\pi^2 \sum_{m,m' \in \underline{\mathcal{M}}} \left| \int \mathrm{d}\mathbf{r}\, \overset{\approx}{\Psi}_{2\varepsilon m} V_{\mathrm{T}} \overset{\approx}{\Psi}_{1\varepsilon m'} \right|^2 , \end{aligned} \tag{3.46}$$

where we have used $\mathcal{U}^{ll'\,\dagger} = \mathcal{U}^{l'l}$, following from the definition (B.20).

Transmission eigenmodes basis

In order to pass to the eigenmode basis, we need to recall the form of $\mathcal{U}^{2\,1}_{mm'}$ given in the eqs. (3.27). Once again the rightmost terms will be cancelled, this time through the $\mathcal{R}$-multipliers, which for $m, m' \in \mathcal{M}$ are zero. We therefore have the matrix elements relevant for the second order correction:

$$\mathcal{U}^{2\,1}_{mm'} = \begin{cases} \left\langle \overset{\approx}{\varphi}{}^{(+)}_{1\varepsilon m} \middle| V_{\mathrm{T}} \middle| \overset{\approx}{\varphi}{}^{(-)}_{1\varepsilon m'} \right\rangle \\ \\ \left\langle \overset{\approx}{\varphi}{}^{(-)}_{2\varepsilon m} \middle| V_{\mathrm{T}} \middle| \overset{\approx}{\varphi}{}^{(+)}_{2\varepsilon m'} \right\rangle \end{cases} , \tag{3.47}$$

with the tip held inside the left and the right lead, respectively. We substitute them into (3.46) and in the right-lead form of the matrix element we apply complex-conjugation and change of the summation indices $m \leftrightarrow m'$. Altogether, this brings us to

$$\begin{aligned} g^{(2)} &= -4\pi^2 \sum_{\substack{m \in \underline{\mathcal{M}} \\ m' \in \underline{\mathcal{M}}}} \left| \left\langle \overset{\approx}{\varphi}{}^{(-)}_{l\varepsilon m} \middle| V_{\mathrm{T}} \middle| \overset{\approx}{\varphi}{}^{(+)}_{l\varepsilon m'} \right\rangle \right|^2 \\ &= -4\pi^2 \sum_{\substack{m \in \underline{\mathcal{M}} \\ m' \in \underline{\mathcal{M}}}} \left| \int \mathrm{d}\mathbf{r}\; V_{\mathrm{T}}(\mathbf{r})\, \overset{\approx}{\varphi}{}^{(+)}_{l\varepsilon m}(\mathbf{r})\, \overset{\approx}{\varphi}{}^{(+)}_{l\varepsilon m'}(\mathbf{r}) \right|^2 . \end{aligned} \tag{3.48}$$

Here, the lead index l should agree with the lead in which the tip is placed.

As we can see, the ideal second order correction is always negative, except inside the structure, and involves only the open modes. Deviations of the second order response with respect to (3.48) can arise due to non-perfect

closing of the closed modes ($\mathcal{T}_m \neq 0$) and non-perfect transmission in the open modes ($\mathcal{R}_m \neq 0$). Although they are, to some degree, inevitable, there is no use investigating their effect on the results – the corrections due to such imperfections will be of greater order in their causing factors, $\mathcal{T}_m$ or $\mathcal{R}_m$, then the first-order conductance correction (3.25) itself.

3.4 SGM signal – discussion

The theoretical approach presented in this chapter brought us to the two conductance correction expressions, (3.24) and (3.42), which describe the signal from a Scanning Gate Microscope for weak, non-invasive probes. Both of these formulae are given in terms of trace operations over the space of the lead modes, which ensures the basis invariance. Moreover, it can be verified that if the nanostructure probed by the SGM has a left-right or up-down symmetry, the SGM map as described by our formulae will have the respective symmetry as well [44, 45].

A particularly convenient version of these expressions is given in the scattering eigenstates basis – eqs. (3.25) and (3.44), and in the transmission eigenmodes basis, for the regions inside the leads – eqs. (3.28) and (3.48). These simple forms pertain to one eigenstates' basis only and to one eigenmodes' basis only – see subsec. 3.4.9 of Ch. 2 – but, as we see directly from the expressions, the results for $g^{(1)}$ and $g^{(2)}$ would not change if an overall m-dependent (but l-independent) phase factor were added, it is therefore enough to evaluate the $\overset{\approx}{\Psi}_{l\varepsilon m}$-functions only up to such a phase. Another important form, for a perfectly local tip potential, will be given below, in subsec. 3.4.9.

3.4.1 Overview

The conductance corrections incorporate two kinds of items:

- structure-dependent – transmission and reflection amplitudes, expressed in terms of the products like $r^\dagger t'$, $t^\dagger t$, $\mathcal{R}\mathcal{T}$; they depend on the investigated nanostructure only, and can be further subdivided into magnitude-controlling quantities $\mathcal{R}_m$, $\mathcal{T}_m$, and the matrices $[u_l]_{ma}$ which control the mode-composition and phases acquired on passing through the scatterer;

- tip-dependent – the matrix elements of the tip potential, in the appropriate bases; these describe the tip-scattering of the electron waves and,

as such, they hold all the corrections' dependence on the tip potential, its shape, strength, and, most importantly, the position where it is applied; they *also* depend on the structure through the wave functions of the corresponding basis.

Passing between the different bases that we have featured amounts to shifting some of the structure-dependent terms from one of the groups to the other. By the choice of a suitable basis we can emphasise either the relation of the SGM signal with the leads, i.e. the environment in which the electron waves move, or with the scatterer, which in this context could be called the "emitter" or "injector", cf. Ch. 1.

3.4.2 Comments on the derivation; role of the leads

We would like to remark that, despite the use of the lead-specific form of the scattering state wave functions in the matrix elements of the current density (3.14), and of the "half" current density (3.34), the derivation did not require placing the tip inside the lead. It is valid for any tip position, including the interior of the scatterer structure. What it requires, is that the cross section $\mathcal{S}_x$ for the current integration in (3.16), be inside the (right) lead, and since its longitudinal coordinate is specified as the infinity ($\mathcal{S}_x : x \to \infty$) this requirement only amounts to stating that the leads must exist somewhere, while we are left with quite much freedom in the choice of their position and width.

This means that, although our system cannot be infinite in both dimensions, we can approximate such a situation by taking the leads to be very broad. Taking the infinity limit of the leads width takes us to a "fully" two-dimensional space. In particular, we never have a two-dimensional travelling wave in our ("finite") system, but in the limit passage an appropriate linear combination of many quasi-one-dimensional terms (each consisting of a 1D travelling wave in longitudinal direction and the stationary transverse mode) will compose such a wave. Before taking the limit, the correctness of this quasi-1D approximation of a 2D-wave holds only up to a place where a substantial part of the signal hits the lead border.

If, for a correctness check, we decide to situate the integration cross section in the other lead ($\mathcal{S}_x : x \to -\infty$), we get results equivalent to those obtained above; their equality is guaranteed by the S-matrix properties described in sec. 2.3. Also reversing the current direction, which amounts to putting $l = 2$ as the lead of origin in subsecs. 3.2.1 and 3.3.1, yields the same results.

3.4.3 Applicability of the perturbation theory

Certainly, the condition

$$\left|\chi_{l\varepsilon a}^{(n+1)}\right| \ll \left|\chi_{l\varepsilon a}^{(n)}\right|, \qquad n = 0, 1, 2, \ldots \,, \tag{3.49}$$

if met, would guarantee the convergence. Furthermore, from the Lippmann-Schwinger equation it follows that

$$\chi_{l\varepsilon a}^{(n+1)} = \sum_{l'=1}^{2} \int_{\varepsilon_1^{(\perp)}}^{\infty} \frac{\mathrm{d}\varepsilon'}{\varepsilon^+ - \varepsilon'} \sum_{a'=1}^{N(\varepsilon')} \chi_{l'\varepsilon' a'}^{(n)} \; [V_{\mathrm{T}}]_{a'a}^{l'l}(\varepsilon', \varepsilon) \,, \tag{3.50}$$

which can be read off from the equations (3.7–3.9). From this expression, we see that it is the matrix elements $[V_{\mathrm{T}}]$, especially the ones diagonal in energy, $\mathcal{V}$, which should be small for the Born series to converge.

This can be achieved by keeping the tip potential low in strength and in extent. If we denote by Δr_0 the linear size of the region in which the tip potential is significant, and by V_0 the tip potential's maximum value, then the condition for the applicability of the Born approximation in a one-dimensional space can be expressed as (after [65], §45, p. 161):

$$V_0 \Delta r_0 \ll \hbar v_F^{(1\mathrm{D})} \,, \tag{3.51a}$$

where $v_F^{(1\mathrm{D})} \equiv \hbar k_F^{(1\mathrm{D})}/M_e$ is the free Fermi velocity. The condition can be rewritten in a more pictorial form as

$$V_0 \Delta r_0 \ll E_F^{(1\mathrm{D})} \lambda_F^{(1\mathrm{D})} \,. \tag{3.51b}$$

An estimate for a fully 2-dimensional case is not available (because of a logarithmic divergence – see [65], §45, p. 161, first footnote), but the 1D-condition is reasonable for our situation, because the form of our formulae, such as (3.50), is essentially quasi-one-dimensional.[3] They contain many one-dimensional terms weighted by the transverse modes and added together – and this can only favour the convergence. When the tip is placed in the broad region away from the Quantum Point Contact, the unperturbed wave functions from the $[V_{\mathrm{T}}]$ matrix elements, being normalised with respect to the extent of the transverse cross-section, have relatively small amplitudes. The addition of the many terms in (3.50) cannot restore the value of the sum to the magnitude it would have if just one mode existed, since every term that is added has a different phase factor coming from the longitudinal part

[3] We just need to bear in mind that the (1D)-superscripted quantities of these expressions will be the propagation (longitudinal) quantities in our quasi-1D system.

of the wave function. The quantitative assessment of this effect, however, might not be straightforward and is beyond the scope of this work. At the same time we would like to remark, that this effect might loose some of its importance when the disorder causes the coherent branching of the SGM signal, since this inevitably requires some kind of greater localisation in the transverse direction of the wave functions of the travelling electrons.

We see from the condition (3.51a) that our theory gives good description for weak or very tightly localized tip potentials. This is how we understand the non-invasiveness of the SGM probe.

The question of how relevant the perturbation theory results are for the SGM signal with finite tip strengths, will also be addressed, along with other tasks, through the means of numerical simulations in Chapter 5.

3.4.4 Utility of different bases

The standard basis, in which we derived our results, uses the advantage of being closely related to the states of the lead, whose form is usually known and simple (cf. the examples in sec. 2.1.4). This allows, for example, for an easy numeric implementation in a tight-binding scheme and easy interpretation of its outcomes. However, it forces us to consider all the modes existing in the leads, even though only few eigenstates might actually be transmitted. Besides, this basis makes it somewhat difficult to draw general conclusions from the formulae (3.23) and (3.42-3.43).

This is largely remedied by the conversion to the scattering eigenstates basis. Here, we do not have to work with more states than necessary: the terms participating in the transmission are readily extracted, and the imaginary/real or positive/negative parts are easy to obtain. The basis of the scattering eigenstates gives us better means of interpreting the results for $g^{(1)}$ and $g^{(2)}$ and of exploring the conditions in which each of them is relevant (eq. (3.29)). Due to the emphasis on the structure and not the leads, it is especially advantageous for setups where the leads are very broad or away from the interesting regions. The cost of these simplifications is the fact that we no longer have an easy way for determining the form of the basis wave functions, as evaluation of the u_l matrices might be a demanding task.

In both of the bases described above, the wave functions, and hence also the matrix elements of the tip potential, contain the transmission and reflection amplitudes. By passing to the transmission eigenmodes basis, we push the $\mathcal{R}$ and $\mathcal{T}$ magnitudes completely out of matrix elements, which then gain a simpler relation to the lead modes. They still retain the information on the transmission process, expressed through the u_l matrix coefficients, responsible for the shape of the eigenmodes, but not for their magnitude on

either side of the structure. In other words, this basis achieves the separation between the magnitude-controlling factors and the shape of the transmitted wave functions. Naturally, the transmission eigenmodes basis is useful only when the region of our interest is inside the leads, even if they are very wide.

3.4.5 Relevance of the 2nd order correction

The first order conductance correction will be sufficient for the perturbative approximation of the SGM signal, provided it is not equal to zero.[4] Therefore, once the tip potential matrix elements are small enough for the perturbation theory to be valid, the condition of relevance for the second order correction $g^{(2)}$ is set by the vanishing of $g^{(1)}$.

Out of the two categories of terms specified in 3.4.1, only the structure-dependent ones will realise this (see below). The condition for nullifying $g^{(1)}$ through the transmission and reflection factors has been concisely expressed in the equation (3.29):

$$\mathcal{R}\mathcal{T} = \mathbf{0} \ .$$

Since the $\mathcal{R}_m$ and $\mathcal{T}_m$ values complement each other (cf. (2.42)), this requirement means that all active modes have to be *fully closed or fully open*, which is reflected in the form of the eq. (3.45). Also the form (3.25) of the first order correction could have been restricted, by stating that $m \in \not{\mathcal{M}}$, so as to express the fact that only the *partially open* modes contribute to $g^{(1)}$.

From the above, it follows that the applicability of $g^{(2)}$ necessarily implies the quantisation of the unperturbed structure's conductance, because

$$g^{(0)} = \sum_{m \in \mathcal{M}} \mathcal{T}_m^2 = N_{\underline{\mathcal{M}}} \tag{3.52}$$

(eq. (2.59)). If we know that the eigenstates open one by one, like it is in the case of well-chosen QPCs (see Ch. 4), and possibly close one by one, we can state that the reverse implication is also true. Then, the condition of applicability of the second order correction is identical to the conductance quantisation (3.52). Each time we need to be specific, we will call such conductance quantisation a regular one. In general, however, it is possible to imagine a situation, where two partially open modes add up to unity, giving rise to conductance quantisation and not suppressing $g^{(1)}$ anyway. A somewhat exaggerated example is shown in fig. 3.1. We will not consider such cases in here, viewing them as incidental. More realistic are instances of such "false" quantisation which happen in a point-like way, but these ones – as we discuss below – bear no practical significance.

[4] What we mean by this in the present context, is the presence of factors other than V_0, tending to zero faster than V_0.

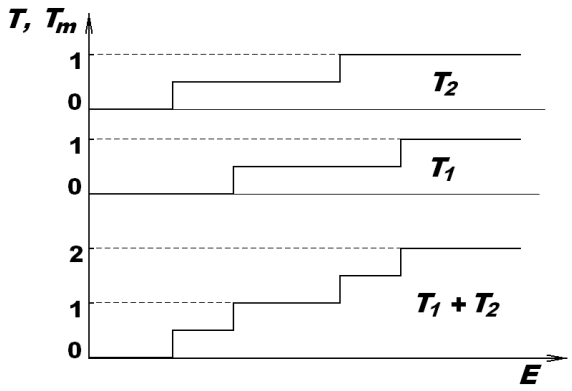


Figure 3.1: An example of quantized transmission which does not cause vanishing of the first order conductance correction when $T = 1$. This is due to overlapping of the regions where the two modes are only partially open.

In order to see a quantized conductance curve when a measurement is taken, the above mentioned requirement (3.29) has to be met for certain ranges, or intervals, of the unperturbed system parameters varied in the measurement, and not just for isolated points. Each such a range will referred to as a *conductance plateau* region. Complementarily, the intervals when at least one mode is partially open, we will call *conductance step* or *ramp* region. The single-point-only compliance also would alter the SGM signal, but has no practical importance, as it would demand very precise tuning of the system parameters. Moreover, a conductance curve could be quantised with its plateaus at non-integer multiples of the conductance quantum. Quite obviously, on such plateaus we will not find the second order expression useful, and therefore we are not going to treat them on the same footing as the regular plateaus of the previous paragraph.

As to the tip-dependent part, the matrix elements appearing in (3.28) do get suppressed, become purely real, or, with a particular set of $\mathcal{R}_m\mathcal{T}_m^2$ coefficients, cancel each other. But these can happen only for certain positions of the tip and cannot be sustained in entire regions (granted that the tip potential is sufficiently local[5]), therefore cannot lead to significant changes in the SGM signal.

Finishing this subsection, we would like to point out that the first and second order expressions suffice for the perturbative description of the SGM.

[5]In contrast, when the basis wave functions oscillate very quickly on the scale of variation of the tip potential, it might occur that the matrix elements vanish for any tip position. This pertains to highly energetic electrons and spatially extended tip potentials, in which case a completely classical theory should apply, see 3.4.8.

The only situation where both, $g^{(1)}$ and $g^{(2)}$, vanish simultaneously (not counting the interior of the scatterer region) is when all modes are fully closed.

3.4.6 Basic interpretation

Based on our first and second order conductance correction expressions, we would like to interpret the SGM signal in the context of two kinds of measurements that can be made:

- SGM-mapping of the structure with all the non tip-related parameters held constant, and

- taking the conductance curve with respect to some of the system parameters, with or without perturbing the system with the tip.

As follows from the previous section, the SGM maps obtained with the non-invasive tip will differ substantially when taken on the conductance plateau, from the ones taken in the conductance step region or for a structure revealing no quantisation at all.

Maps – step region

First, we would like to look at the conductance step region. Here, the SGM signal is of the first order in the tip potential. According to (3.25), what we probe when we move the tip over the system, is the spatial structure of the imaginary part of the tip potential matrix elements $\mathcal{U}^{2\,1}_{mm}$, taken between scattering eigenstates originating on the right and the corresponding, i.e. same-index, eigenstates originating on the left. The map is a superposition of such signals coming from all the states which under the measurement conditions remain partially open. The set of real positive coefficients, $\mathcal{R}_m\mathcal{T}_m$, by which the individual signals are weighted, is constant for each map, and can be understood as measuring how far the channels are from being fully open or fully closed.

In general, the presence of many intertwined matrix elements can make the map complicated and difficult to interpret. This need not be the case, if there are eigenstates which open within a well-defined region in the energy domain, not intersecting with the opening regions of other channels. Throughout the entire step region corresponding to such an opening, the SGM technique maps one channel only. With the regular conductance quantisation, all the conductance steps have this plausible feature.

Maps – plateau region

The second situation we will feature is the plateau region of a quantised conductance, where the SGM response to a non-invasive probe is dominated by the second order conductance correction. Here, the SGM scan maps the spatial variation of the squared modulus of the tip potential matrix elements $\mathcal{U}^{2\,1}_{mm'}$. The map is created by adding signals from all the pairs of open states and all the pairs of closed states; first state of the pair has to originate on the right, second one – on the left. As we have pointed out in the derivation, sec. 3.3.3, and discuss in the next section, the closed modes can be neglected for most purposes, since their influence is significant only inside the scatter region, and only for certain types of scatterers.

Now, even if the conductance is regularly quantised, we have no general way of isolating these individual components. The only simple map is the one obtained at the lowest conductance plateau (with $g^{(0)} = 1$, or, in terms of the total transmission, $T = 1$), with only one channel open. The higher the plateau, the more complicated the SGM image will get. An attempt at separating the partial signals might be realised by subtracting a lower plateau's map from the given, n-th one, and considering $g^{(2)}\big|_{T=n} - g^{(2)}\big|_{T=n-1}$ (see Ch. 1, sec. 1.3). This brings in two sorts of issues. Firstly, by doing so, we do not dispose of the cross-terms between the highest open state ($m = n$) and all the other open states ($m' \in \underline{\mathcal{M}} - \{n\}$). This problem disappears when the transmission is quasi-adiabatic, which we introduce and discuss in subsec. 3.4.10. Secondly, as the conductance plateaus do not overlap, the subtracted $(n-1)$-st map must have been taken at a different Fermi energy than the n-th one. This is alleviated by the fact that the transverse parts of the lead modes, upon which the scattering eigenstates are constructed, depend solely on the mode index. Thus, we can probe the transverse shape of the matrix elements, provided that we obtain information on the longitudinal parts, i.e. on the propagation wave numbers and the u_l matrices,[6] from some other sources. And here, the just mentioned quasi-adiabatic approximation will also offer us some help.

Conductance curves

Conductance curves of the unperturbed system serve the purposes of determining whether the conductance quantisation takes place and telling the values of the parameters, for which the conductance plateaus and steps occur. Thus, they inform us about the conditions at which the SGM map is or should be taken. In the case of the regular quantisation, the information

[6] Or, if we average it out like in [1].

from the conductance curve translates immediately into the knowledge of $\mathcal{T}_m$ (and hence $\mathcal{R}_m$ too).

The varied parameter is often chosen in such a way that its variation is, at least theoretically, equivalent to varying the Fermi energy. For the sake of simplicity, we will talk here about the conductance dependence on the Fermi energy or the Fermi wave number, and not on other parameters, like constriction width, shape or the potential inside it. In addition, we will interchangeably treat the conductance and the total transmission, which, in the view of the Landauer-Büttiker relation introduced in Ch. 2, is fully justified.

The idea of the conductance curve can be stretched onto a measurement at the presence of the SGM probe. Thus, the arising conductance correction, i.e. the difference between the perturbed and unperturbed curves $\Delta g(E_F) = g(E_F) - g^{(0)}(E_F)$, can give a "spectroscopic" information about the matrix elements $\mathcal{U}^{2\,1}$ for one position of the tip.

In the step regions, the signal would be modulated by the $\mathcal{R}_m(E_F)\mathcal{T}_m(E_F)$ function, whose values are known also from the reference curve. Naturally, it will provide us with the information on the imaginary part of the matrix elements. A full information should be available after use of the Kramers-Kronig relations, but we do not follow this thread in this work.

On the plateaus, on the other hand, even if we have the means to extract the much smaller second order signal, the conductance correction curves might be not easy to interpret, except for the lowest plateau, due to impossibility of deconvoluting the individual $|\mathcal{U}^{2\,1}_{mm'}|$ signals.

3.4.7 Remarks on the form of the expressions

The fact that all the conductance correction formulae are based on the $\mathcal{U}^{2\,1}$ matrix elements is a consequence of the scattering states' basis that we used all the way through the derivation. With the current impinging from the left, like we assumed it, only a scattering of electrons from the left-originating scattering state to the right-originating one can describe a change in conductance, since only that one carries a net current back towards the left.

Taking for the first order correction only the diagonal elements of $\mathcal{U}^{2\,1}_{\cdot\cdot}$ is reminiscent of the trace operation. The double sum in the second order expressions introduces the mixed (mm') terms, but does not mix between closed and open states. This is understandable, because:

a) there is always one lead, in which the closed state has zero wave function and hence cannot participate in scattering;

b) in the other lead:

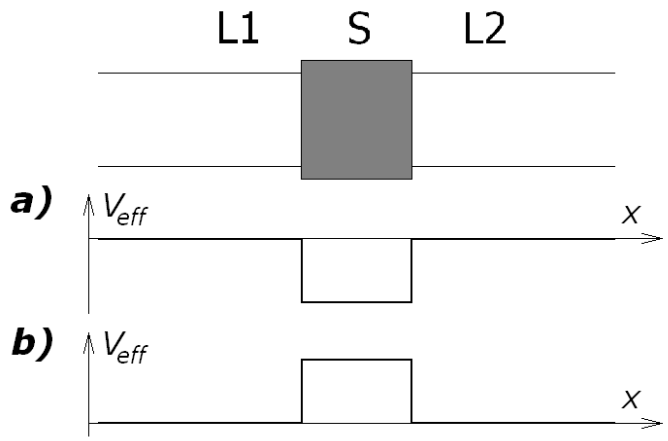


Figure 3.2: A schematic depiction of the difference between the "trapping" (a)) and "repulsive" (b)) scatterer. a) When the effective potential is smaller then the background potential in the leads, the scattering state has a sizeable part inside the structure, which is attributed to neither the transmitted, nor the reflected part. b) The reverse situation happens when the structure creates a convex potential landscape – the scattering state's amplitude is small here.

- the electron scattered into the closed state has no chance to get back towards the left (this applies when the tip is in the right lead), or
- the electron which scatters from the closed state to an open one, would have gotten back even without the scattering event (this applies in the left lead).

We also see (point a)), that the closed-closed scattering can be important only when the tip is inside the scatterer region. Furthermore, its influence will be more pronounced for the structures which tend to "trap" the wave function inside them, and less pronounced for the structures which push the wave functions out, like in the case of the QPC. The difference is schematically explained in the fig. 3.2.

3.4.8 Role of the tip potential

The presented formulae apply to the SGM probes of arbitrary potential shape, as long as they meet the conditions described in sec. 3.4.3. It is, however, desirable to limit our considerations to localized tip potentials, as

this is always the case in the SGM technique. In fact, we will restrict ourselves even further, to the idealised case of the point-like tip.

We observe that the matrix elements of any given potential $V_{\mathrm{T}}(\mathbf{r})$ can be expressed in terms of the delta-potential matrix elements:

$$\begin{aligned}\mathcal{U}^{ll'}_{mm'} &= \int \mathrm{d}\mathbf{r}\ \tilde{\tilde{\Psi}}^{*}_{l\varepsilon m}(\mathbf{r})\, V_{\mathrm{T}}(\mathbf{r})\, \tilde{\tilde{\Psi}}_{l'\varepsilon m'}(\mathbf{r}) \\ &= \int \mathrm{d}\mathbf{r}\ \tilde{\tilde{\Psi}}^{*}_{l\varepsilon m}(\mathbf{r}) \left[\int \mathrm{d}\mathbf{r}'\ V_{\mathrm{T}}(\mathbf{r}')\delta(\mathbf{r}-\mathbf{r}')\right] \tilde{\tilde{\Psi}}_{l'\varepsilon m'}(\mathbf{r}) \\ &= \frac{1}{U_0}\int \mathrm{d}\mathbf{r}'\ V_{\mathrm{T}}(\mathbf{r}') \left[\int \mathrm{d}\mathbf{r}\ \tilde{\tilde{\Psi}}^{*}_{l\varepsilon m}(\mathbf{r})\, U_0\delta(\mathbf{r}-\mathbf{r}')\, \tilde{\tilde{\Psi}}_{l'\varepsilon m'}(\mathbf{r})\right] .\end{aligned} \tag{3.53}$$

Here, U_0 is an arbitrary constant, but, in order to keep the correct units for the matrix elements, its dimension should be energy times length squared. In the context of the ideal point-like tip potential

$$V^{(\delta)}_{\mathrm{T}}(\mathbf{r}) = U_0\, \delta(\mathbf{r}-\mathbf{r}_0)\ , \tag{3.54}$$

(with $\mathbf{r}_0$ denoting the tip position), the constant U_0 has a clear significance as the space-integrated value of the potential; we shall call it the *tip strength* and occasionally use it interchangeably with the tip voltage. With appropriately defined linear extent of the tip potential, Δr_0, we can write $U_0 = V_0 \Delta r_0^2$.

Thanks to its linearity, the first order correction $g^{(1)}$ shares the delta-decomposition property (3.53). If we denote by $g^{(1,\delta)}(\mathbf{r}_0)$ the first order signal due to the delta tip potential $V^{(\delta)}_{\mathrm{T}}$ with the tip held at $\mathbf{r}_0$, then the response to the general potential is

$$g^{(1)}(\mathbf{r}_0) = \frac{1}{U_0}\int \mathrm{d}\mathbf{r}'_0\ V_{\mathrm{T}}(\mathbf{r}'_0)\, g^{(1,\delta)}(\mathbf{r}'_0)\ . \tag{3.55}$$

This tells us not only that knowing the delta-tip signal we can reconstruct the signal of any other tip, but also, more importantly, that the only effect the size of a non-invasive tip has on the signal is to blur it, by taking the values of the idealized signal from the points around $\mathbf{r}_0$ and adding them, with a certain weight, to the signal at $\mathbf{r}_0$. We can, therefore, say that the maximum information that could be obtained from a first order SGM map is contained in the delta-tip signal. This is why throughout the rest of this work we will focus on this shape of tip potential. In the rest of the formulae, we will abandon the (δ)-superscript.

It follows from the above, in particular, that a non-invasive finite-size tip potential should blur out the fringes in the signal completely, when its effective radius becomes larger than their spacing. We know that this does

not hold in the case of the standard, invasive-tip experiments, like [1, 6], because the observed fringes constitute a variation of signal on a scale smaller than the size of the 2DEG-depletion disc under the tip. It is an additional confirmation that the depletion disc under the invasive tip acts essentially like a classical object. And the classical explanation for the unexpectedly high resolution is that the fraction of the tip-backscattered flux which ultimately returns to the QPC and gets transmitted through it, must have been reflected from a small and relatively well-defined region (circular segment) on the disc's circumference (the so-called "glint effect" [39]).

We cannot write an analogue of eq. (3.55) for the second order correction, as this one is not linear. Since the matrix elements get smoothed out when the tip potential has finite size, we can surely say that the $g^{(2)}$-signal will also get smoothed, but additional effects occur as well – most importantly the impossibility of separating the summations as we do it in (3.57) below, which is important for the interpretation, see subsec. 3.4.12.

3.4.9 Conductance correction for the delta-tip potential

First order With the delta tip potential, the first order conductance correction (3.25) can be rewritten as

$$g^{(1)} = 4\pi U_0 \sum_{m=1}^{N} \mathcal{R}_m \mathcal{T}_m \operatorname{Im}\left\{ \overset{\approx}{\Psi}{}^{*}_{2\varepsilon m}(\mathbf{r}_0)\, \overset{\approx}{\Psi}_{1\varepsilon m}(\mathbf{r}_0) \right\} , \tag{3.56a}$$

and the lead-specific version (3.28) as

$$g^{(1)} = 4\pi\, U_0 \sum_{m\in\mathcal{M}} \mathcal{R}_m \mathcal{T}_m^2 \operatorname{Im}\left\{ \overset{\approx}{\varphi}{}^{(\pm)\,2}_{l\varepsilon m}(\mathbf{r}_0) \right\} , \tag{3.56b}$$

with l indexing the lead in which we place the tip, while the $(+)$ superscript should be chosen if $l = 2$, and the $(-)$ superscript – if $l = 1$.

Second order Similarly, the second order correction (3.46) reads:

$$\begin{aligned} g^{(2)} &= -4\pi^2 U_0^2 \sum_{m,m'\in\underline{\mathcal{M}}} \left| \overset{\approx}{\Psi}{}^{*}_{2\varepsilon m}(\mathbf{r}_0)\, \overset{\approx}{\Psi}_{1\varepsilon m'}(\mathbf{r}_0) \right|^2 \\ &= -4\pi^2 U_0^2 \left[\sum_{m\in\underline{\mathcal{M}}} \left| \overset{\approx}{\Psi}_{2\varepsilon m}(\mathbf{r}_0) \right|^2 \right] \left[\sum_{m\in\underline{\mathcal{M}}} \left| \overset{\approx}{\Psi}_{1\varepsilon m}(\mathbf{r}_0) \right|^2 \right] . \end{aligned} \tag{3.57}$$

This is nothing else but a product of two *partial local densities of states*, one for the open states incoming from the left, and one for the open states incoming from the right. Such form points out to the physical origin of the correction, which is the scattering of the electrons by the tip potential from any of the open $(1\varepsilon m)$-states to any of the back-propagating open states. It does not include the occupation factors, but the states are fully occupied and fully empty anyway.

The first plateau case If we apply the above formula in a time-reversal invariant system tuned to the first plateau of the quantised conductance, an even simpler interpretation can be obtained.

For a fully open channel, allowing no reflected wave components, the time-reversal invariance (see subsec.2.3.2, Ch. 2) guarantees that the complex conjugate of a right-originating eigenstate can be decomposed in the following way: $\overset{\approx}{\Psi}{}^{*}_{2\varepsilon m}(\mathbf{r}) = \sum_{m'\in\underline{\mathcal{M}}} c^{(m)}_{m'}\, \overset{\approx}{\Psi}_{1\varepsilon m}(\mathbf{r})$. The sum contains fully open left-originating states only. On the first conductance plateau the sum can run over one index only, $\underline{\mathcal{M}} = \{1\}$, and the constant multiplier is a pure phase. The second order conductance correction then takes the form

$$g^{(2)} = -4\pi^2 U_0^2 \left| \overset{\approx}{\Psi}_{1\varepsilon 1}(\mathbf{r}_0) \right|^4 . \tag{3.58a}$$

Likewise, the eigenmode-basis form (eq. (3.48)) is

$$g^{(2)} = -4\pi^2\, U_0^2 \left| \overset{\approx}{\varphi}{}^{(\pm)}_{l\varepsilon 1}(\mathbf{r}_0) \right|^4 \tag{3.58b}$$

inside the l-th lead. The choice between the $(\pm)$ superscripts is arbitrary.

We see that there is a direct and very simple relation between the $g^{(2)}$ correction and the local density of the only occupied open state, which, in turn, is related to the charge density at the Fermi energy. We will discuss this briefly in subsec. 3.4.12 below.

3.4.10 General features of the signal inside the leads

Even in the simplest situations, with only one partially open channel when the step-signal is considered, or only one fully open channel for the plateau-signal, we are faced with a superposition of many oscillations, since the basic constituents of the $g^{(1)}$ and $g^{(2)}$ expressions are not simple oscillations, but linear combinations thereof, namely $\overset{\approx}{\varphi}{}^{(+)}_{l\varepsilon m}(\mathbf{r}) = \sum_a [u_l]_{ma}\, \varphi^{(+)}_{l\varepsilon a}(\mathbf{r})$.

In order to see how the signal decomposes to those simple parts, we take the lead-specific formulae (3.56) and (3.57) back to the basis of lead modes (keeping the $m^{(\prime)}$-sums):

$$g^{(1)} = 4\pi\, U_0 \sum_{m\in\cancel{\mathcal{M}}} \mathcal{R}_m \mathcal{T}_m^2 \,\mathrm{Im}\left\{ \sum_{aa'} [u_l]_{ma}\, [u_l]_{ma'}\, \varphi^{(+)}_{l\varepsilon a}(\mathbf{r}_0)\, \varphi^{(+)}_{l\varepsilon a'}(\mathbf{r}_0) \right\}, \quad (3.59a)$$

$$g^{(2)} = -4\pi^2\, U_0^2 \sum_{m,m'\in\underline{\mathcal{M}}} \left| \sum_{aa'} [u_l]_{ma}\, [u_l]_{m'a'}\, \varphi^{(+)}_{l\varepsilon a}(\mathbf{r}_0)\, \varphi^{(+)}_{l\varepsilon a'}(\mathbf{r}_0) \right|^2 . \quad (3.59b)$$

First order

The fact that the eigenmodes get squared in the first order formula (3.56) gives rise to the oscillating terms with increased longitudinal wave numbers,

$$k_{aa'} = k_a + k_{a'} \ , \quad (3.60)$$

which lie between $2k_N$ and $2k_1$, and to a mix of increased and decreased transverse wave numbers – for the simple hard-walled leads (ex. 1, sec. 2.1.4) each transverse mode splits into two harmonics, with[7]

$$k^{(\perp)}_{aa'} = \left| k^{(\perp)}_a \pm k^{(\perp)}_{a'} \right| . \quad (3.61)$$

This produces a myriad of oscillating terms with the effective (total) wave numbers $\left\| \left(k^{(\perp)}_{aa'}; k_{aa'}\right) \right\| = \left(k^{(\perp)\,2}_{aa'} + k^2_{aa'} \right)^{1/2}$, ranging from $2k_N$ for $a = a' = N$, which can be nearly zero, to maximally $2k_F$.

Second order

At the same time, in the second order formula (3.57) we have a similar mix of simple oscillations in the expression under the modulus sign, only differing by that the off-diagonal mm'-pairs are also allowed, which affects the amplitudes u_l but not the form of individual oscillating terms. After factoring out of the sum the term associated with the middle wave number, $\exp[i(k_1 + k_N)x]$, and cancelling it at the modulus, we are left with the set of terms whose propagation wave numbers are differences of the original ones,

$$k_{aa'} = k_a - k_{a'} \ , \quad (3.62)$$

[7] As

$$\phi_a \phi_{a'} = (2/W) \sin(k^{(\perp)}_a y) \sin(k^{(\perp)}_{a'} y) = (1/W)\left\{ \cos[(k^{(\perp)}_a - k^{(\perp)}_{a'})y] - \cos[(k^{(\perp)}_a + k^{(\perp)}_{a'})y] \right\} .$$

spanning a symmetric set between $-(k_1 - k_N)$ and $k_1 - k_N$ ($< k_F$) which includes zero wave numbers, i.e. constant terms, absent in the first-order case. Squaring the modulus-expression expands this set to reach twice larger bounds, while the overall set of longitudinal periodicities stretches from 0 to $2(k_1 - k_N)$, since each negative wave number has its positive counterpart. This set would coincide with what we had for the first order signal, if we shifted it by $2k_N$ towards the lower wave numbers.

The transverse wave numbers, on the other hand, are given by (3.61) before the modulus and the square are taken. After this is done,[8] the effective magnitudes of wave numbers of the oscillations found in the second order signal will lie between 0 and $2\sqrt{\left(k_1 - k_N\right)^2 + \left(k_1^{(\perp)} + k_N^{(\perp)}\right)^2}$. (We stress that the oscillations are not circular waves, so the periodicities implied by these wave numbers can be found only in certain directions.)

We see that the second order case can produce an equally complicated structure as the first order, with slightly slower oscillations, but richer structure in the transverse direction (compare footnote 8 with footnote 7).

The experimentally observed fringes of regular periodicity, half the local Fermi wavelength, can be explained through our formulae only if there are some regularities encoded in the u_l matrices, other than just unitarity. Now, we would like to discuss the simplest possible of such regularities, namely when only one of the terms on each column is appreciably different from zero.

3.4.11 Quasi-adiabatic case

The special case we would like to highlight here, amounts to assuming that, for at least one lead l, the column vector of coefficients $[u_l]_{m\cdot}$ has only one non-zero element, that is

$$[u_l]_{ma} = \delta_{\tilde{a}a} \exp(i\alpha_m^{(l)}) \tag{3.63}$$

for some index $\tilde{a}$ depending on m. The quantity $\alpha_m^{(l)}$ represents (part of) the phase acquired on getting transmitted.

[8] Here, we use

$$\sin^2(k_a^{(\perp)}y)\sin^2(k_{a'}^{(\perp)}y) = \frac{1}{4}\Big\{1 - \cos\left(2k_a^{(\perp)}y\right) - \cos\left(2k_{a'}^{(\perp)}y\right) + \frac{1}{2}\cos\left[2\left(k_a^{(\perp)} - k_{a'}^{(\perp)}\right)y\right] + \frac{1}{2}\cos\left[2\left(k_a^{(\perp)} + k_{a'}^{(\perp)}\right)y\right]\Big\}.$$

The above equality basically means that the m-th scattering eigenmode coincides within the lead l with one of the lead modes, up to a phase,[9] which is very similar to requiring a (globally) adiabatic transmission of the lead mode, i.e. such where that mode would be transmitted into exactly one, corresponding, lead mode on the other side. For this reason, we shall call eq. (3.63) the *quasi-adiabatic* condition. The *adiabatic* transmission of a lead mode implies such identification not only between the given lead mode and some incoming transmission eigenmode in one lead, but also between its other-side correspondent and some outgoing eigenmode in the other lead, so we could say that the adiabaticity is nothing else but a double-sided quasi-adiabaticity. As we shall discuss in Ch. 4, in the QPCs the quasi-adiabaticity is not as unusual as it might seem.

Thanks to the quasi-adiabatic assumption, we can easily evaluate the conductance correction expressions (3.56) and (3.57) in the scattering eigenstate and lead eigenmode bases.

First order

The first order signal from a single partially open state is then

$$\begin{aligned} g^{(1)} &= 4\pi\, U_0 \sum_{m\in\not{\mathcal{M}}} \mathcal{R}_m \mathcal{T}_m^2\, \mathrm{Im}\left\{ \varphi_{l\varepsilon\tilde{a}}^{(\pm)\,2}(\mathbf{r}_0) \exp(2i\alpha_m^{(l)}) \right\} \\ &= 4\pi U_0\, \mathcal{R}_m \mathcal{T}_m^2\, \frac{\phi^2_{\tilde{a}(m)}(y_0)}{h v_{\tilde{a}}} \sin\left(2k_{\tilde{a}(m)} x_0 + 2\alpha_m^{(l)} \right) . \end{aligned} \tag{3.64}$$

This is a regular, x-y separable wave. It reproduces the experimentally reported fringes of the positive- and negative-correction stripes alternating in the direction of propagation with the periodicity of half the local Fermi wavelength.[10] Based on this and the analysis from the previous subsection, we can conclude that the regular fringing pattern, containing one periodicity only, as reported in the many experiments is the signature of not only the coherence, which is a necessary condition, but also of a certain degree of (quasi-) adiabaticity – meaning that the vector $[u_l]_{m\cdot}$, even if not delta-like as in (3.63), is at least peaked around one of its indices $\tilde{a}$.

What it cannot reproduce, is the decay of the fringes with the distance from the structure. This is due to the straight, non-widening leads, and only

[9] This is true for both, the incoming and the outgoing m-th eigenmode, but only in one lead.

[10] In this context, the *local* Fermi wavelength should be understood as the propagation wavelength at Fermi energy, i.e. the spatial periodicity based only on the kinetic part of the electron's energy.

passing to the infinite leads width, which we do not perform in here, would tell us something about how an ideal quasi-adiabatic signal should decay in a fully two-dimensional QPC-containing system. It is very interesting to note, however, that the results do not depend on whether the "ray" of signal is kept together by the hard walls or by the prior focussing. This suggests, that in a branched pattern of SGM signal, a single branch could perhaps behave like a separate lead. (The fact that each aa'-term in (3.59a) has a periodicity-direction of its own, means actually that this should work also for the non-quasi-adiabatic case, because all the parts of signal with the "wrong" periodicity would simply leave such an imaginary lead through its "walls" by moving in a different direction, and only one wavelength would remain in the branch. Still, such a scenario would imply the existence of branches with different periodicities, which is contrary to the experimental findings.)

Another price we pay for the simplicity of this solution, is the inability to satisfy a natural requirement on the signal in a usual (abrupt) QPC system, namely the requirement that the signal is concentrated near the exit from the narrow constriction and spreads wider as the tip is moved further from the exit. This can be reproduced only by adding many terms with different periodicities, like we had in the previous subsection. This is another way of justifying the well-known fact that the abrupt structures are necessarily non-adiabatic, complementary to the one we will present in Ch. 4. And conversely – we see that the quasi-adiabatic model can serve as a reasonable approximation for the QPCs which open gradually and smoothly, so that the wave functions inside them manage to "straighten" before they enter the lead.

Second order

The quasi-adiabatic expression for the second order reads

$$\begin{aligned} g^{(2)} &= -4\pi^2\, U_0^2 \sum_{m,m'\in\underline{\mathcal{M}}} \left|\varphi^{(+)}_{l\varepsilon\tilde{a}(m)}(\mathbf{r}_0)\,\varphi^{(+)}_{l\varepsilon\tilde{a}(m')}(\mathbf{r}_0)\right|^2 \qquad (3.65)\\ &= -4\pi^2 U_0^2 \sum_{m,m'\in\underline{\mathcal{M}}} \frac{\phi^2_{\tilde{a}(m)}(y_0)\;\phi^2_{\tilde{a}(m')}(y_0)}{h^2\, v_{\tilde{a}(m)} v_{\tilde{a}(m')}}\\ &= -4\pi^2 U_0^2 \left|\sum_{m\in\underline{\mathcal{M}}} \frac{\phi^2_{\tilde{a}(m)}(y_0)}{h\, v_{\tilde{a}(m)}}\right|^2, \end{aligned}$$

which does not vary in the longitudinal dimension at all – in particular, it shows no fringes. On the other hand, for plateaus beyond the first one,

the transverse structure of the signal is considerably richer here than in the first order case, due to the presence of the mixed terms $\tilde{a}(m)\tilde{a}(m')$. (The non-diagonal mixed terms, $m \neq m'$, for the low plateaus, like second and third, will have relatively small overlaps, which might be at the origin of the qualitative success of the mode subtracting in [1].)

The lack of fringes in the plateau signal is quite remarkable, because it does not follow from the absence of QPC-backscattering ($\mathcal{T}_m$ can be lower than one), but from the quasi-adiabaticity of it. As we can see from (3.59b), any departure from the quasi-adiabatic model will lead to a revival of the fringes. Also, owing to the effectively higher powers of transverse modes in that expression, the transverse structure of this "imperfect" (non-quasi-adiabatic) second order signal will be richer than of the one of its first order analogue. In conjunction with the possible small fringing, this could give a weak chequerboard pattern (cf. simulation results in Ch. 5). We remind that these conclusions hold inside the leads. As we will see in Ch. 4, outside leads even a fully adiabatic structure can have a fringed SGM signal.

3.4.12 Comparison with local densities

Closing this chapter, we would like to compare the obtained SGM conductance corrections to the unperturbed local densities: density of current and density of charge.

Current density

As we discussed in Ch. 2, the net current across the investigated system is carried only by the states in the "Fermi window" of energies, that is, between the chemical potentials μ_1, μ_2 of the two reservoirs, as all currents below μ_2 ($< \mu_1$) cancel. All of those originate in the left lead (reservoir), and in Appendix B, we show that this is also the case for current densities. Thus, we only need to consider here $\mathbf{j}_{1\varepsilon}$, the current density due to the ensemble of left-originating scattering states, at the Fermi energy (i.e. between μ_1 and $\mu_2 = \mu_1 - \mathrm{d}E_F$).

The lack of dependence on the right-originating states might be seen as an opportunity to rule out immediately the possible relation between the current density and the SGM conductance corrections, which contain wave functions of both, left- and right-originating states. We see, however, by the virtue of the same reasoning, that there is a relation between left-originating current and the right-originating states ($\mathbf{j}_{1\varepsilon} = -\mathbf{j}_{2\varepsilon}$). We will therefore seek comparison between the current density and the SGM response by writing both quantities explicitly.

The current density components, longitudinal $j_{1\varepsilon}^{(x)}$ and transverse $j_{1\varepsilon}^{(y)}$, are given in eqs. (B.2a-B.2b) of the appendix. We rewrite them as:

$$j_{1\varepsilon}^{(x)} = \frac{e\hbar}{M_e} \sum_{m\in\mathcal{M}} \mathcal{T}_m^2 \,\mathrm{Im}\left\{ \sum_{aa'} [u_2^*]_{ma}\, [u_2]_{ma'}\; \varphi_{2\varepsilon a}^{(-)}(\mathbf{r})\varphi_{2\varepsilon a'}^{(+)}(\mathbf{r})\; ik_{a'} \right\} ,$$

$$j_{1\varepsilon}^{(y)} = \frac{e\hbar}{M_e} \sum_{m\in\mathcal{M}} \mathcal{T}_m^2 \,\mathrm{Im}\left\{ \sum_{aa'} [u_2^*]_{ma}\, [u_2]_{ma'}\; \varphi_{2\varepsilon a}^{(-)}(\mathbf{r})\varphi_{2\varepsilon a'}^{(+)}(\mathbf{r})\; \frac{1}{\phi_{a'}} \frac{\partial \phi_{a'}}{\partial y} \right\} .$$

This form is convenient for confronting it with the delta-tip conductance corrections expressed in eqs. (3.59a-3.59b) of subsec. 3.4.10. We need to bare in mind that what should be relevant about the conductance correction is its magnitude

$$\|\mathbf{j}_{1\varepsilon}\| = \sqrt{j_{1\varepsilon}^{(x)^2} + j_{1\varepsilon}^{(y)^2}} \,, \tag{3.66}$$

which complicates tremendously the analysis. Hence, we will compare only the most general features of the expressions.

Eigenmode-composition The first discrepancy we notice, is the difference of the eigenmode summation range. The current density expression includes contributions from all the channels, except the fully closed ones – due to the $\mathcal{T}_m^2$-factor. We can make the eigenmode-composition match that of the conductance corrections by a suitable choice of the Fermi energy:

– in the first step region of the quantised conductance – the range matches with that of $g^{(1)}$,

– in any of the plateau regions – the range of $g^{(2)}$ is matched.

With a non-quantised conductance curve, where only $g^{(1)}$ is relevant, the ranges will only agree if none of the channels if perfectly open. In all other cases there is no way to represent all the contributions to the current density through the SGM signal.

In addition, the current density's single sum over eigenmodes can be turned into a double sum, similar to the one in second order conductance correction, when the squares of the $j_{1\varepsilon}^{(x)}$ and $j_{1\varepsilon}^{(y)}$ components are taken. The expressions become complicated after squaring and we will not trace them exactly, contenting ourselves with considering only the simplest, quasi-adiabatic case – see below.

Lead-mode-composition The most important point, however, is the absence of any factors in the SGM conductance corrections, that could give them the dependence on the wave vector, and hence on the velocity, of electrons, which are present in the expressions for $\mathbf{j}_{1\varepsilon}$.

The wave-vector-related factors in the current density, $k_{a'}$ and $\frac{1}{\phi_{a'}}\frac{\partial \phi_{a'}}{\partial y}$, distort the relative magnitudes of terms in the sum over the lead modes (a, a'), with respect to what there is in the g-corrections. The second one of them also distorts the spatial variation. However, in the limit of infinitely wide leads ($W \to \infty$), this effect of distorting the lead-mode composition can get cancelled by adding the two squared components in (3.66). The electron waves sufficiently far from the scatterer nanostructure become separable in polar coordinates, so the current density and charge density are proportional, the proportionality coefficient containing the Fermi wave vector k_F. Then, relating the current density and the SGM response is still possible.

Phase-composition Owing to the complex conjugation of the u_2 matrix and the presence of one incoming and one outgoing lead mode, the relative phases of the terms in the aa'-sum are also distorted, but only with respect to what we find in $g^{(1)}$-correction, while they can be made to agree with the diagonal, $m = m'$, elements of the second order correction.[11] In particular, on the first plateau – with only one possible choice for m and m' – the phase-compositions of $g^{(2)}$ and $\mathbf{j}_{1\varepsilon}$ agree.

Quasi-adiabatic case Using the quasi-adiabatic approximation, (3.63), we can retrieve the simple expressions for the current density,

$$j_{1\varepsilon}^{(x)} = \frac{e\hbar}{M_e} \sum_{m\in\mathcal{M}} \mathcal{T}_m^2 \, k_{\tilde{a}} \left|\varphi_{2\varepsilon\tilde{a}(m)}^{(\pm)}(\mathbf{r})\right|^2 = \frac{e}{h} \sum_{m\in\mathcal{M}} \mathcal{T}_m^2 \, \phi_{\tilde{a}(m)}^2(y)$$

and

$$j_{1\varepsilon}^{(y)} = 0 \ .$$

(cf. (2.12) and (2.52)), which show explicitly the scaling by mode wave numbers (see above), with the lack of it in the quasi-adiabatic SGM signal (3.64) and (3.65). Apart from this, the first order correction also differs from the current density by the fringes-generating factor $\sin(2k_{\tilde{a}}x_0)$.

[11] Thanks to the complex modulus sign, the second order correction can be re-written as

$$g^{(2)} = -4\pi^2 U_0^2 \sum_{m,m'\in\underline{\mathcal{M}}} \left|\sum_{aa'} [u_l^*]_{ma} \, [u_l]_{m'a'} \, \varphi_{l\varepsilon a}^{(-)}(\mathbf{r}_0) \, \varphi_{l\varepsilon a'}^{(+)}(\mathbf{r}_0)\right|^2 \tag{3.67}$$

Charge density

While it cannot be directly linked to the unperturbed current flux,[12] the non-invasive SGM response in some cases finds a clear relation to the structure's charge density at the Fermi level.

The total density of charge at the Fermi energy, or in the "Fermi window", in the region to the right of the QPC is given by

$$\varrho_\varepsilon = e \sum_{m\in\mathcal{M}} \left| \overset{\approx}{\Psi}_{1\varepsilon m}(\mathbf{r}_0) \right|^2 = e \sum_{m\in\mathcal{M}} \mathcal{T}_m^2 \left| \overset{\approx}{\varphi}{}^{(\pm)}_{2\varepsilon m}(\mathbf{r}) \right|^2 . \tag{3.68}$$

($\varepsilon = E_F$) and is the same, up to a constant, as the local density of the open left-incoming states at E_F. It includes the left-incoming states only, since the right-incoming ones, both open and closed, remain empty.

Considering the result (B.23), we see that the second order correction in the first plateau region of the quantised conductance is proportional to the square of the Fermi-energy charge density.

Apart from the time-reversal invariance, two things are crucial for this simple relation. The first one is the locality of the tip potential. We remind that the splitting of summations in (3.57) would not have been possible for tip potentials other then the assumed delta-potential. This could be a serious restriction from the practical point of view, since the tip potential width is comparable to the tip's radius of curvature and cannot be controlled by other factors. The finite width will introduce not only blurring of the SGM signal, like in the linear case ($g^{(1)}$) shown explicitly in subsec. 3.4.8, but also an additional variation caused by the different evolution of the phase in $\overset{\approx}{\Psi}_{1\varepsilon m}$ and $\overset{\approx}{\Psi}_{2\varepsilon m}$ as we move to a neighbouring point of space.

The second essential requirement is the good quality of the conductance plateau. Any departures from $\mathcal{T}_m = 1$ for the open channel and from $\mathcal{T}_m = 0$ for the closed ones will change the character of $g^{(2)}$ by introducing the $r't$-related terms (see (3.42), subsec. 3.3.3) and diminish its importance in favour of $g^{(1)}$. We will revisit the matter of the imperfect plateau conductance in the context of numerically simulated maps in Ch. 5.

We cannot establish analogous relations the plateaus higher than the first one, unless the investigated system has additional symmetries. It has been shown in [44] that if it has a central (four-fold) symmetry, then the SGM response on all plateaus is proportional to both, the square of the current density and the square of the Fermi-energy density of the *mobile* charge. This relation holds only at sufficient distances away from the QPC opening

[12] Except in the wide-lead limit – see below.

(or from the exit of another small structure), $k_F r_0 \gg 1$, and only in the limit of infinitely wide leads.

The first order correction $g^{(1)}$ does not share the same direct relation to the charge density as $g^{(2)}$. One situation, where it bears some resemblance to the charge or current density, is when we consider a quasi-adiabatically transmitting structure and look at the signal inside the (finite-width) leads, far enough from the structure. There, the oscillation given by (3.64) should have an amplitude which is related to the charge density. This situation, however, seems to have a limited practical importance.

Chapter 4

Application to Quantum Point Contacts

We would now like to apply the theory developed so far to the case of the quantum point contacts, whose most prominent feature is the conductance quantization. As we discussed in the previous chapter, focussing on this simple case allows for some serious simplifications to be made in the SGM conductance corrections formulae.

In order to obtain the SGM maps according to our formulae from the Chapter 3, a specific model for the QPC geometry has to be chosen. In the first part of the present chapter, we review the most important models describing the QPC conduction. Some of them are particularly suitable for the numerical simulations – and these will be the subject of our investigations in Ch. 5, while some other ones, owing to the analytical solutions they provide, can be used directly within our theory – this is the subject of the second part of this brief chapter.

A special attention will be paid to the notion of adiabaticity, since, as we have seen in Ch. 3, it can play an important role in explaining the simple shape of the SGM signal observed in the experiments.

4.1 Conductance quantization in the QPC

4.1.1 Wire model

A quick-and-simple way to understand the quantization of conductance in steps of the height $2e^2/h$, occurring in Quantum Point Contacts in the ballistic regime of transport, is supplied by the Landauer-Büttiker formula, applied to the simplest conceivable case – with no scatterer at all, as in the fig. 4.1.

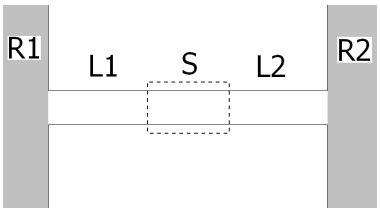


Figure 4.1: The simplest approximation for a QPC: the perfect thin and long wire (represented here by L1, S and L2 all together) joined in a reflectionless way with the reservoirs (R1 and R2).

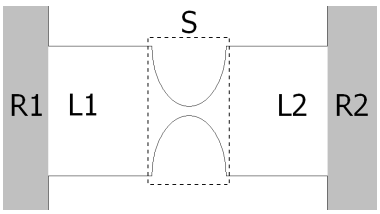


Figure 4.2: QPC as a part of the scatterer-leads-reservoirs system.

The lead, an ideal wire perfectly coupled to the reservoirs, does not disturb the passing electrons in any way, so $t_{ab} = \delta_{ab}$. The only restriction on the transmission is that at a given Fermi energy E_F the lead of the effective width W supports only a limited number of the transverse modes, $N(E_F)$, which enters the formula as the dimension of the matrix t.

The Landauer conductance (2.58) is then:

$$g^{(0)} = N(E_F) \ . \tag{4.1}$$

This produces a perfect step-like structure for g when we vary the Fermi energy. The subsequent steps rise at the thresholds given by the transverse eigenenergies $\varepsilon_a^{(\perp)}$ of (2.4), Ch. 2, so a variation of the wire's effective width produces the same effect. From the example (2.19-2.21) we gave in sec. 2.1, a simple geometric corollary can be drawn for the hard-walled structure: a new constriction mode is opened when the quantity $2W/\lambda_F$ trespasses an integer number, ie. when an additional half of Fermi wavelength fits inside the constriction.

Each of the conductance steps is associated with a *single* open mode. Furthermore, there are no inter-mode processes, so the modes behave like independent quasi-one-dimensional channels, each of them with a constant background potential $\varepsilon_a^{(\perp)}$ and its own conductance $G_0 = 2e^2/h$. The physical origins of the fact that, despite the lack of scattering, a mode does not conduct more than this, are the following:

– a single mode can host no more than a certain number (certain density) of electrons[1] at a time, and any excessive electrons fed by the reservoir

[1]Or, more correctly, a certain total probability of finding an electron – as we do not work with many-electron states.

would not enter – a phenomenon known as the contact resistance [37, 66],

– all the electrons in a given mode move with a fixed velocity, v_a.

The cancellation of the mode's velocity and the electron density in the current calculation, leading to the equal conductance for all modes, is an important conclusion from the Landauer theory.

Because of our assumption about the perfect channel filling and voiding at the reservoirs (sec. 2.4 of Ch. 2), this simple model of the QPC neglects *a priori* any kind of reflection of the modes that can exist inside it. On the other hand, the assumption about long leads (secs. 2.1-2.2) prevents it from including the tunnelling transmission of those modes which cannot. Thus, we do not know anything about the accuracy of the quantization that can be achieved in a real structure. It is interesting to note, that the experimental discovery of the $2e^2/h$ steps in the conductance curve came unexpectedly, even though the Landauer formalism was already known at that time. The reason was the lack of estimates of the steps' quality and a common belief that it was unlikely to have the quality good enough for the steps to be observed [67]. In addition, a natural question arises whether this kind of description could be valid for a very short constriction, when transverse states in the narrowest part might be not well defined [68].

4.1.2 Separable potential

Having said this, we would like to get back to modelling the QPC as the central structure accessed by two leads, as in fig. 4.2.

The QPC itself can be described by the potential $U(x, y)$ that the split-gate exerts inside the simple wire. We have to consider a slightly more general Hamiltonian as compared to the one of Chapter 2 (eq. (2.1)),

$$H_{\mathrm{QPC}} = -\frac{\hbar^2}{2M_e}(\partial_x^2 + \partial_y^2) + U(x, y) \ , \tag{4.2}$$

but we assume $U(x, y)$ remains separable and binding in the y-dimension. So, the description in terms of the transverse modes and transverse energies is still valid, only with the allowance for the x-coordinate appearing as a parameter:

$$\phi_a(y; x), \quad \varepsilon_a^{(\perp)}(x), \quad a = 1, 2, \dots \ . \tag{4.3}$$

We can decompose the wave function of an electron propagating through the constriction using the completeness of the transverse modes; at any x we

have

$$\varphi(\mathbf{r}) = \sum_{a=1}^{\infty} \xi_a(x)\, \phi_a(y;x) \ , \tag{4.4}$$

which, for mathematical correctness, should include all the modes: the open ones, the ones which get closed inside the narrow region only and the ones which are evanescent even inside the leads. The last category can be neglected due to the long-leads assumption. Under such a decomposition, the time-independent Schrödinger equation has [2, 69]:

– the transverse part

$$\left[-\frac{\hbar^2}{2M_e}\partial_y^2 + U(x,y) \right] \phi_a(y;x) = \varepsilon_a^{(\perp)}(x)\, \phi_a(y;x) \ , \tag{4.5}$$

– and the longitudinal part

$$\sum_{a=1}^{\infty} \left\{ -\frac{\hbar^2}{2M_e}\xi_a''(x)\, \phi_a(y;x) + \left[\varepsilon_a^{(\perp)}(x) - \varepsilon\right] \xi_a(x)\, \phi_a(y;x) \right\} = \sum_{a=1}^{\infty} \left[-\frac{\hbar^2}{M_e}\xi_a'(x)\, \partial_x \phi_a(y;x) - \frac{\hbar^2}{2M_e}\xi_a(x)\, \partial_x^2 \phi_a(y;x) \right] \tag{4.6}$$

(where prime denotes the ordinary derivation).

Projecting the latter on some arbitrary transverse mode we get a more tractable form:

$$\left[-\frac{\hbar^2}{2M_e}\partial_x^2 + \varepsilon_a^{(\perp)}(x) - \varepsilon \right] \xi_a(x) = \sum_{b=1}^{\infty} A_{ab}\xi_b(x) \ , \tag{4.7}$$

where the transverse energies $\varepsilon_a^{(\perp)}$ play the role of an effective potential, and the operator

$$A_{ab} = -\frac{\hbar^2}{M_e} \left\langle \phi_a \left| \partial_x \right| \phi_b \right\rangle \partial_x \ - \frac{\hbar^2}{2M_e} \left\langle \phi_a \left| \partial_x^2 \right| \phi_b \right\rangle \tag{4.8}$$

couples different modes. (The bra-ket multiplication is understood to take place in the y-space only.)

In the approach described above, the travelling electrons can be viewed as passing through a series of potential wells. The potential $U(x,y)$ carves out a certain region accessible for the electrons at a given Fermi energy, thus defining the lateral constriction: *soft-walled* (SW) – in the ordinary case, and *hard-walled* (HW) – if the transverse potential wells are closed by

infinite values of the potential. Both these cases are well described by the above analysis. The latter type of constrictions involves imposing appropriate additional boundary conditions, but this makes it particularly suitable for computer treatment. It is the hard-walled model that we will use for our simulations in Chapter 5.

4.1.3 Adiabatic approximation

The form of the operator A, with the first and second derivative of the transverse mode wave functions along the longitudinal direction, suggests how the set of equations (4.7) could be simplified [2, 69].

The transverse modes, at each x, are fully determined by the transverse potential in the equations (4.5), and when the QPC potential varies slowly and smoothly with x as compared to its corresponding changes in y, so will do the transverse wave functions $\phi_a(y;x)$. For such potentials, the right-hand side terms of (4.7) are small and we could neglect them, hence decoupling the equations. The electron's evolution when passing through the structure is then *globally adiabatic* – it stays all the time in the same transverse state.[2]

Globally adiabatic hard-wall model

The globally adiabatic approximation was applied by Glazman *et al.* [68], to the hard-walled structures with zero potential at the bottom, like in the example (2.19) of the sec. 2.1. The constriction is then fully described by its width as a function of the longitudinal position, $W(x)$. A formula to describe the transmission was given, which includes the previously neglected effects, partial reflection in the open channels and the tunnelling in the closed ones,

$$T_a = \frac{1}{1+\exp\left[-\frac{2\pi}{\sqrt{W_{\mathrm{QPC}}W''_{QPC}}}\left(k_F W_{\mathrm{QPC}} - \pi a\right)\right]} \,. \tag{4.9}$$

Here W_{QPC} is the constriction width in its narrowest point, and by W''_{QPC} we mean second derivative taken in that point. For the validity of the adiabatic approximation in this system, the boundary's radius of curvature should exceed the Fermi wavelength: $W''^{-1}_{QPC} \gg \lambda_F$. The formula (4.9) naturally sets the quantity $2\pi\sqrt{W_{\mathrm{QPC}}/\,W''_{\mathrm{QPC}}}$ in the role of a measure of the conductance steps' quality, called the adiabatic parameter. An important conclusion is that the ideally adiabatic transmission depends only on the constriction properties in its narrowest point.

[2] More precisely: the electron stays in the mode *corresponding* to the one it started in.

Non-perfect adiabaticity, local adiabaticity

Analysis of the adiabaticity conditions and implications [70, 71] shows that even a small, but finite, boundary curvature builds up its effect as the electrons propagate through the constriction, leading necessarily to breaking the adiabaticity at a certain distance $x_{ad} \sim W''^{-1}_{QPC}$ away from the centre of the QPC. The adiabaticity is therefore kept only *locally*, while globally we can only speak of greater or smaller degree of adiabaticity.

Beyond that distance, an electron propagation state previously described by the single transverse mode (of the central part of the constriction) starts incorporating other transverse modes, which manifests through non-zero off-diagonal, i.e. mode-mixing, terms in the transmission amplitude matrix in the lead-mode basis. In the basis of scattering eigenstates, on the other hand, there is by definition no mode-mixing, and the non-adiabaticity is described through the non-diagonal elements of the auxiliary matrix u_l in the considered lead.[3] The scattering eigenstates, at least the propagating ones, can still be chosen to contain single transverse modes in constriction's central part. This shows that there is a direct correspondence between constriction (central region) states and the transmission eigenstates of the entire system. (This correspondence has been confirmed also in the abrupt hard-wall model of the QPC [44], which lacks the adiabaticity completely.) From this perspective, the non-adiabaticity can be understood as a mismatch, or an incompatibility, between the injector-determined electron motion and the environment-determined one.

The effect of small departures from adiabaticity on the structure's conductance was studied by Yacoby and Imry in [69]. They introduced a method for iterative solution of the eq. (4.7), by which they demonstrated that the inter-mode reflection and transmission terms have the same character as the diagonal reflection terms, i.e. exponentially falling with the growth of the constriction smoothness parameter. Furthermore, they showed that for the adiabatic shape of the conductance curve it is sufficient, if the constriction opens smoothly in a short region close to the narrowest point, regardless of what happens further. Thus the local adiabaticity can successfully replace the global one.

As we have seen in Ch. 3, adiabaticity also has an important influence on the shape of the SGM signal, but not on its interpretation. It seems that the SGM-mapping gives us better means of discerning the effects of the adiabaticity or its lack than simply the conductance curves. We will qualitatively consider these effects in the simulated images in Ch. 5.

[3] More precisely – through having more then one non-zero element in a column, i.e. not satisfying (3.63).

Non-adiabaticity, abrupt hard-walled structures

The opposite case, with rectangular constriction borders, was investigated by Szafer and Stone [72]. This type of geometry is expected to be highly mode-mixing. Their results show that a good quantization is still attainable. The shape of transmission curve is no longer of exponential character, as in (4.9), but is much sharper than it would follow from solving a 1D rectangular-potential problem. An additional mechanism to spoil the plateaus was reported. The electron waves reflected at the entrance and at the exit of the constriction can now interfere with each other, giving Fabry-Perot-type resonant structures in the conductance curves. This effect requires, apart from substantial reflection at the exit, an appreciable phase difference between those two, so it will be visible for elongated structures (see Ch. 5). Its experimental relevance was discussed in somewhat more detail in [73]. Later, this kind of feature in the conductance plateaus was experimentally observed by van Wees *et al.* [74].

In shorter structures, on the other hand, the previously mentioned below-barrier tunnelling is large, which also deteriorates the accuracy of quantization.

4.1.4 Büttiker saddle-point model

Since the electrostatically defined constriction must have a smooth potential everywhere, it is of interest to see what happens in a soft-walled structure created by a potential with no sharp edges nor instantaneous rises. The key insights can be extracted from the double-harmonic model of the QPC introduced by Büttiker [36].

Here, the QPC potential is expanded in a Taylor series up to the second order around its saddle point:

$$U(x,y) \cong V_{\mathrm{QPC}} - \frac{1}{2}M_e\omega_x^2 x^2 + \frac{1}{2}M_e\omega_y^2 y^2 \ . \tag{4.10}$$

The coefficients $\omega_{x(y)}$ are the curvatures of the QPC potential, and hence they fully describe the constriction's geometry in this model.

The transverse wave functions, as evaluated from (4.5), are the 1D harmonic oscillator solutions, given by the expression (2.23) of the sec. 2.1. They are independent of x, and completely determined by the mode index a and the scale parameter $\sqrt{M_e\omega_y/\hbar}$. The effective energies $\varepsilon_a^{(\perp)}$ are then (cf. (2.3-2.4) and (2.24)):

$$\varepsilon_a^{(\perp)}(x) = V_{\mathrm{QPC}} - \frac{1}{2}M_e\omega_x^2 x^2 + \hbar\omega_y\left(a - \frac{1}{2}\right) \ . \tag{4.11}$$

The coupling terms A_{ab} vanish and the longitudinal Schrödinger equation can be written in terms of the dimensionless energy and position

$$\mathcal{E}_a = 2\,\frac{\varepsilon - V_{\mathrm{QPC}} - \hbar\omega_y\left(a-\frac{1}{2}\right)}{\hbar\omega_x} \qquad X = \sqrt{\frac{M_e\omega_x}{\hbar}}\,x \tag{4.12}$$

as

$$\left(\partial_X^2 + \mathcal{E}_a + X^2\right)\xi(X) = 0\ . \tag{4.13}$$

This is a form of Weber's differential equation [75]. The associated transmission problem was studied by Connor [76] and in the context of two-dimensional electron motion by Fertig and Helperin [62]. They gave the wave functions satisfying (4.13) in terms of the confluent hypergeometric functions (see 4.2.2) and also the semiclassical solutions valid for large $|X|$. From these, the reflection and transmission amplitudes are deduced:

$$t_{ab} = \exp\left(-i\alpha_a\right)\frac{1}{\sqrt{1+\exp\left(-\pi\mathcal{E}_a\right)}}\,\delta_{ab}\,, \tag{4.14a}$$

$$r_{ab} = -i\exp\left(-i\alpha_a\right)\frac{\exp\left(-\pi\mathcal{E}_a/2\right)}{\sqrt{1+\exp\left(-\pi\mathcal{E}_a\right)}}\,\delta_{ab}\,, \tag{4.14b}$$

where the phase factor α_a is a function of energy:

$$\alpha_a = \alpha(\mathcal{E}_a) = \frac{1}{2}\mathcal{E}_a + \arg\,\Gamma\left(\frac{1+i\mathcal{E}_a}{2}\right) - \frac{1}{2}\mathcal{E}_a\ln\frac{1}{2}\left|\mathcal{E}_a\right|$$

and has been deduced from the semiclassical connection formulae [76].

The transmission calculated from (4.14) has the familiar exponential (Fermi-Dirac) shape with a small modification with respect to (4.9), as the energy is used instead of the wave number:

$$T_a = (t^\dagger t)_{aa} = \frac{1}{1+\exp(-\pi\mathcal{E}_a)}\ . \tag{4.15a}$$

The role of adiabatic parameter is now played by $2\pi/\hbar\omega_x$, but once we consider the energy spacing between the steps' rising points, $\hbar\omega_y$, we see that the good quantization condition is actually $\omega_y \gg \omega_x$. Complementary to the above, the total reflection probability for a single mode reads

$$R_a = (r^\dagger r)_{aa} = \frac{\exp(-\pi\mathcal{E}_a)}{1+\exp(-\pi\mathcal{E}_a)}\ . \tag{4.15b}$$

When it comes to calculating the conductance, these expressions have to be evaluated at the Fermi energy and summed up over all the modes, which gives

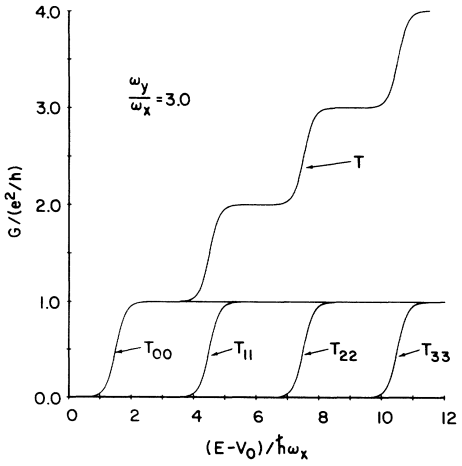


Figure 4.3: The conductance through a QPC modelled by the saddle-point potential, as a function of a dimensionless energy variable. The individual-mode transmission curves T_{aa}, which add up to create it, are also shown. (Figure reproduced from [36].)

the rounded step-like structure in the plot of the conductance as a function of the Fermi energy – see fig. 4.3 – or as a function of the constriction parameters $\omega_{x(y)}$.

The Büttiker model can match a real potential only in the region close to the saddle-point. It therefore replaces the global problem of the electron's passage through the structure, by the local scattering problem. The validity of this approach follows from the above-mentioned fact that the most important contribution to the transmission in a smoothly varying system, i.e. locally adiabatic, comes exactly from that small central region. Accuracy is the greatest when the electron energies are close to the saddle point energy V_{QPC}, so that the region they pass through is indeed close to the saddle point.

An important advantage of this model is that it provides, as sort of a by-product, the exact wave functions of the system [62]. We will use them in the next section to plot the expected SGM signal, first- and second-order, as well as the appropriate local densities.

4.2 SGM signal

Once we choose the model of the QPC and specify the constriction geometry within that model, we can solve it to obtain the quantities necessary for the evaluation of the SGM signal: the transmission and reflection amplitudes, and the scattering state wave functions, from which we build up the tip-potential matrix elements $\mathcal{V}^{21}$ or $\mathcal{U}^{21}$, as well as the local densities ϱ and $\mathbf{j}$ to compare with.

4.2.1 Wire

The case of the clean wire (sec. 4.1.1), even though quite trivial, can serve as a useful reference point when we compare signals of different structures.

The scattering states here are just the wire states (sec. 2.1), based on which we easily evaluate the SGM conductance corrections:

$$g^{(1)} = 0 \ , \qquad g^{(2)} = -4\pi^2 \, U_0^2 \ \left[\sum_{a=1}^{N} \left| \varphi_{\varepsilon a}^{(+)}(\mathbf{r}_0) \right|^2 \right]^2 \tag{4.16}$$

and the relevant local densities:

$$\varrho = e \sum_{a=1}^{N} \left| \varphi_{\varepsilon a}^{(+)}(\mathbf{r}) \right|^2 \ , \qquad \mathbf{j} = \frac{e\hbar}{M_e} \ \sum_{a=1}^{N} k_a \left| \varphi_{\varepsilon a}^{(+)}(\mathbf{r}) \right|^2 \hat{\mathbf{x}}$$

(the lead index l, meaningless in this case, has been omitted). The vanishing of $g^{(1)}$ is determined by the zero-width character of the step regions – see sec. 5.2.1, Ch. 5.

4.2.2 Büttiker model

The saddle-point model of the QPC is very convenient, as it provides analytical expressions for all the quantities required to create the perturbative SGM maps according to the formulae of Chapter 3.

Basic ingredients

Owing to the adiabaticity, the scattering states generated by the transverse modes a are identical to the transmission eigenstates m. Hence, from (4.15) the transmission and reflection singular values can be readily established:

$$\mathcal{T}_m = \frac{1}{\sqrt{1 + \exp(-\pi \mathcal{E}_m)}} \ , \qquad \mathcal{R}_m = \frac{\exp\left(-\pi \mathcal{E}_m / 2\right)}{\sqrt{1 + \exp(-\pi \mathcal{E}_m)}} \ , \tag{4.17}$$

and the scattering eigenstates are described by the wave functions:

$$\overset{\approx}{\Psi}_{l\varepsilon m}(\mathbf{r}) = \xi_m(x) \, \phi_m(y) \ . \tag{4.18}$$

As we mentioned, the transverse functions ϕ_m are the harmonic oscillator solutions (2.23), while the longitudinal parts ξ_m are given by the solutions to the Weber equation (4.13), called the parabolic cylindrical functions. They

can be written in terms of the confluent hypergeometric functions $F(.|.|.)$ as [62]

$$\xi_m(x) = A \exp\left(-iX^2/2\right)\ F\left(\frac{1+i\mathcal{E}_m}{4}\,\middle|\,\frac{1}{2}\,\middle|\,iX^2\right) + BX \exp\left(-iX^2/2\right)\ F\left(\frac{3+i\mathcal{E}_m}{4}\,\middle|\,\frac{3}{2}\,\middle|\,iX^2\right)\ , \tag{4.19}$$

with two arbitrary constants $A, B \in \mathbb{C}$. This form has to be supplemented by the appropriate boundary conditions on ξ_m, so that we could have the incoming-wave component only on the left, if the scattering state's lead of origin is specified as the left one, $l = 1$, or the incoming-wave component only on the right, if $l = 2$.

In order to determine the values of A and B complying with these conditions, the asymptotic expansion of the function $F(.|.|.)$ is used. An explicit form of (4.19) at large values of the position argument, $X \to \pm\infty$, is [62]

$$\xi_m(x) = A\left(\mathcal{F}_1 + \mathcal{F}_1^*\right) + B\left(\mathcal{F}_3 + \mathcal{F}_3^*\right)\ , \tag{4.20}$$

with the short-hand notation

$$\mathcal{F}_1 = \frac{\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{1+i\mathcal{E}_m}{4}\right)} \exp\left(-i\pi\frac{1-i\mathcal{E}_m}{8}\right) |X|^{-\left(\frac{1-i\mathcal{E}_m}{2}\right)} \exp\left(iX^2/2\right)\ ,$$
$$\mathcal{F}_3 = \frac{\Gamma\left(\frac{3}{2}\right)}{\Gamma\left(\frac{3+i\mathcal{E}_m}{4}\right)} \exp\left(-i\pi\frac{3-i\mathcal{E}_m}{8}\right) |X|^{-\left(\frac{3-i\mathcal{E}_m}{2}\right)} X \exp\left(iX^2/2\right)\ .$$

It is easy to see, that the functions $\mathcal{F}_1$ and $\mathcal{F}_3$ represent the outgoing-wave components, whereas their complex conjugates – the incoming-wave ones. Hence, the scattering eigenstates originating in the left lead, $\overset{\approx}{\Psi}_{1\varepsilon m}(\mathbf{r})$, can be described with the constants

$$A^{(l=1)} = \frac{\Gamma\left(\frac{1-i\mathcal{E}_m}{4}\right)}{\Gamma\left(\frac{1}{2}\right)} \exp\left(-i\frac{\pi}{8}\right)\ , \quad B^{(l=1)} = -\frac{\Gamma\left(\frac{3-i\mathcal{E}_m}{4}\right)}{\Gamma\left(\frac{3}{2}\right)} \exp\left(-i\frac{3\pi}{8}\right)\ ,$$

and the ones originating in the right lead, $\overset{\approx}{\Psi}_{2\varepsilon m}(\mathbf{r})$, with

$$A^{(l=2)} = \frac{\Gamma\left(\frac{1-i\mathcal{E}_m}{4}\right)}{\Gamma\left(\frac{1}{2}\right)} \exp\left(i\frac{3\pi}{8}\right)\ , \quad B^{(l=2)} = \frac{\Gamma\left(\frac{3-i\mathcal{E}_m}{4}\right)}{\Gamma\left(\frac{3}{2}\right)} \exp\left(i\frac{\pi}{8}\right)\ .$$

A phase factor can be added to the above coefficients, but it has to be common to all four of them, because the scattering eigenstates basis is formed by all the scattering eigenstates, including both values of l, and we work with a fixed basis – see the remarks in 2.3.3, Ch. 2.

In order to fit within our theoretical framework, the eigenstate wave functions we use for the calculations will be normalised to $\frac{e}{h}\mathcal{T}_m^2$ total flux.[4][5]

In addition, an asymptotic form equivalent to (4.20) can be obtained through a semiclassical analysis of the wave functions [76]. The advantage of this approach is that it provides an easier relation to the physics. It has been used in [45] to construct an analytical expression for the $\mathcal{V}^{2\,1}$ matrix element and thus determine the shape of the SGM signal in the regions away from the QPC.[6]

We would like to use the above wave functions to evaluate and plot the local densities and the SGM conductance corrections. Before we do it, however, it is important to know the ranges of step (ramp) and plateau regions.

Step and plateau regions

The step region is defined by the rising of the prefactor $\mathcal{R}_m\mathcal{T}_m$ significantly above zero. Naturally, the step and plateau regions can be discriminated by simply looking at the conductance curve. Due to the quick variation of the function $\mathcal{R}_m\mathcal{T}_m = \mathcal{T}_m\sqrt{(1-\mathcal{T}_m^2)}$, when the state transmission falls below one, this could be misleading, and evaluation of the prefactor each time gives a better measure. Within the saddle-point model, it reads

$$\mathcal{R}_m\mathcal{T}_m = \frac{1}{\exp(\pi\mathcal{E}_m/2)+\exp(-\pi\mathcal{E}_m/2)} = \frac{1}{2}\operatorname{sech}\left(\frac{\pi\mathcal{E}_m}{2}\right)\,, \tag{4.22}$$

and we plot in the figure 4.4 an example of its variation as a function of energy, for a few lowest m-indices. The peaks around $\mathcal{E}_m = 0$ mark the step regions. The peaks' width is $\hbar\omega_x\frac{\ln(2+\sqrt{3})}{\pi/2}$ (full width at half-maximum) and they are spaced by $\hbar\omega_y$. The regions in between, with zero $\mathcal{R}_m\mathcal{T}_m$, are the plateaus.

In the case of poor conductance quantisation, the peaks will overlap and the tunnelling transmission "tails" from higher m-indices will enter the lower ones' plateaus. Then, despite the total transmission values close to an integer, the SGM signal will not have a pure plateau-character. We will investigate the effect of this spurious process on the SGM maps in some more detail in the next chapter.

[4] Otherwise we would have to add to the calculations the density of particle-normalised states.

[5] We will present our results in arbitrary units, so the normalisation is important only for correctly determining the relative weights of the different modes' signals.

[6] In [45], however, the off-diagonal matrix elements have been (erroneously) neglected.

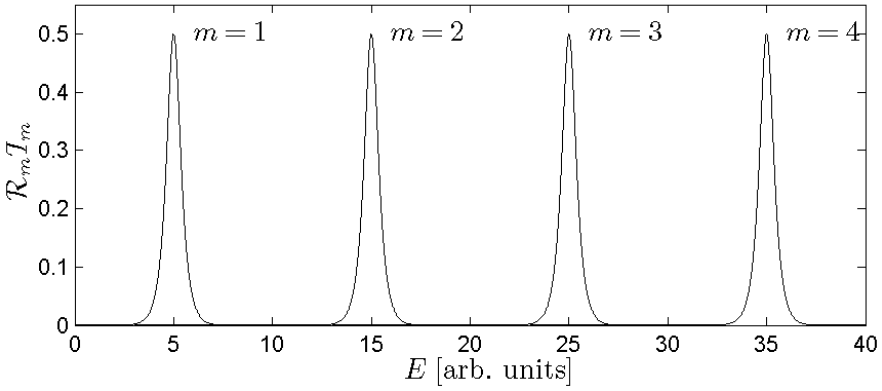


Figure 4.4: The $\mathcal{R}_m\mathcal{T}_m$-prefactor as a function of energy, showing the extent of the step/plateau regions. The constriction parameters are $\omega_y/\omega_x = 2$ and $V_{\mathrm{QPC}} = 0$.

Calculation of the conductance corrections and local densities

A short MATLAB code has been set up to evaluate the wave functions according to what we have described in this section and calculate from them the quantities of our interest:

- first and second order conductance corrections – according to (3.25) and (3.44),
- the charge density at the Fermi energy – (3.68),
- the current density – (2.26).

The values of the longitudinal wave functions ξ_m are obtained by directly summing the confluent hypergeometric series [52] for moderate ranges of the dimensionless position (an external procedure is used for this, [77]). In the regions far from the structure's centre, where the numerical summation becomes troublesome, the asymptotic form (4.20) could be used.

As we discussed in Ch. 3, the legitimate use of the SGM correction formulae requires having straight, non-broadening leads attached to our structure (even soft-walled). This is not the case in the saddle-point model, but – since the contribution to the scattering process is negligible away from the centre of the structure – we can assume that the wave functions will not change significantly in the region of our interest, when we attach very broad leads with an appropriate parabolic potential far away from the centre. The formulae (3.28) and (3.48) are, naturally, of no use in this situation.

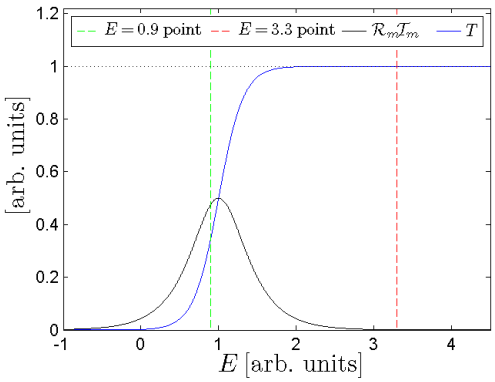


Figure 4.5: Plot of the transmission and the $\mathcal{R}_m\mathcal{T}_m$-prefactor as functions of energy, for the lowest step and plateau. Example step and plateau points, for which the maps of the following figure were taken, are marked with dashed horizontal lines.

Maps of the conductance corrections and local densities

In fig. 4.6 on page 85 we have plotted the local densities and the SGM signals due to the lowest state, $m = 1$, for two example energy-space points: the left-hand panel images pertain to the Fermi energy value chosen in the first step (ramp) region, while the right-hand panel ones have their Fermi energy in the first plateau region. Both these points are marked in the transmission vs. energy curve, shown in fig. 4.5 together with the $\mathcal{R}_m\mathcal{T}_m$ function. The saddle-point potential parameters are chosen such that $\omega_y/\omega_x = 2$ and $V_{\mathrm{QPC}} = 0$. In the maps, two white contour lines have been added to mark the boundaries of the classically forbidden regions. The first row of images shows the charge density at the Fermi energy ϱ_{E_F}, equivalent to the local density of the left-originating states, the second row shows the structure's current density, and the third one – the SGM response. For the step-region energy, only the first order SGM correction is displayed, since the second-order one is irrelevant here. In the plateau region, conversely, the second order conductance correction is depicted, because the linear one vanishes.

The analytical character of the solutions gives us a possibility of isolating the signals of individual m-states, which we use for a more convenient presentation of the results. In particular, we have pushed the first plateau-region energy well into the second eigenstate's step region (cf. fig. 4.5 with fig. 4.7 below) – as to diminish $\mathcal{R}_m$ and thus obtain a good representation of the plateau-signal by $g^{(2)}$ – without including the $m = 2$ (or any higher) eigenstate contribution, as this would obscure our images. We did the same for the second plateau point, where the third eigenstate has been omitted.

In the step region, the Fermi-energy charge density is asymmetric due to the electron wave reflection at the QPC potential. This feature works to spoil the relation between the charge density and the SGM corrections, which was established under the assumption of equality of ϱ_{E_F} and its right-

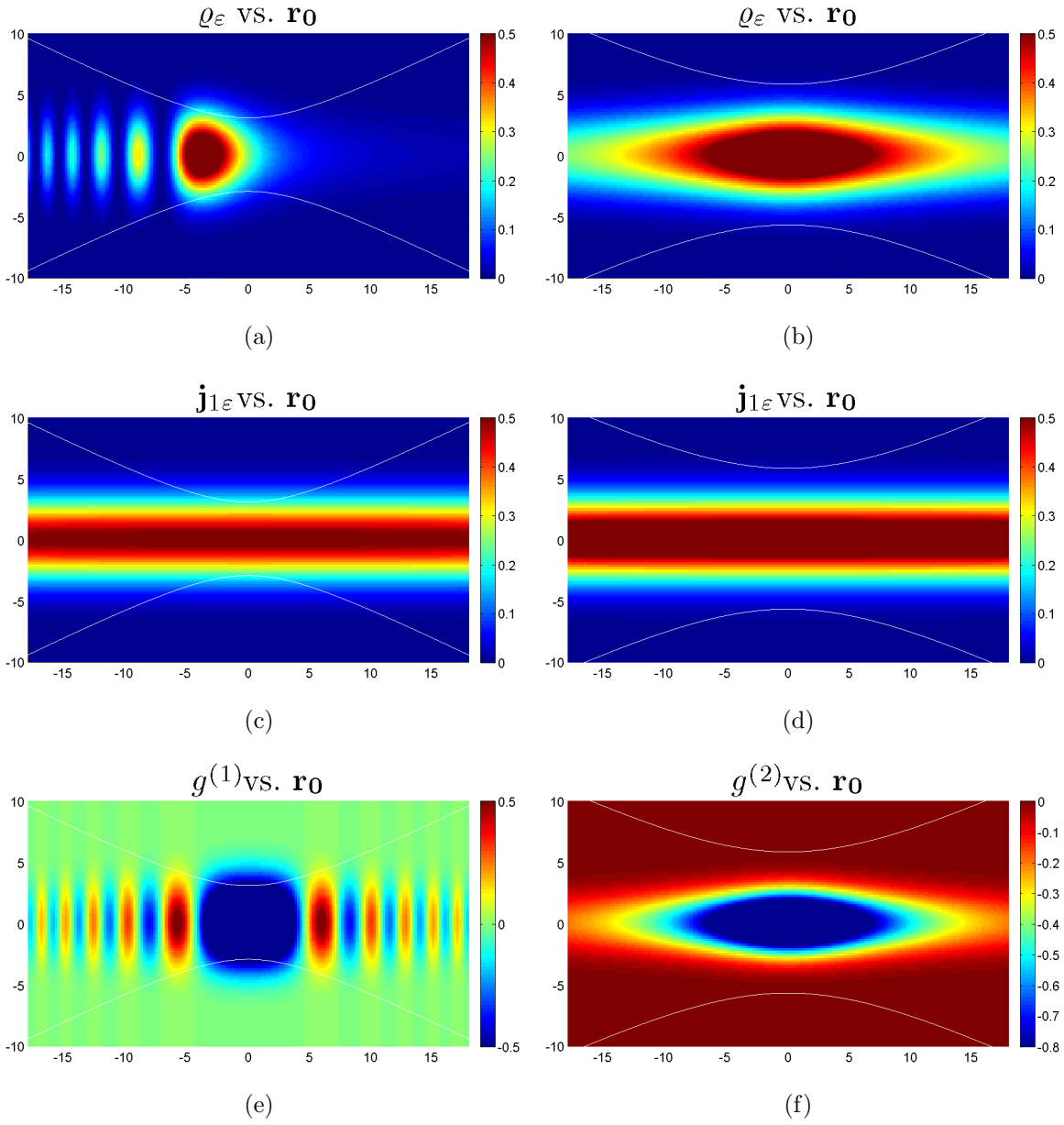


Figure 4.6: Saddle-point model of the QPC, first step and first plateau regions. Plots of local densities:
(a), (b) – charge density at the Fermi energy,
(c), (d) – current density,
and of the SGM signal, given by the:
(e) – first order conductance correction,
(f) – second order conductance correction.
Left panel maps (a,c,e) are taken in the step region, at $E = 0.9$, and the right panel ones (b,d,f) are taken in the plateau region, at $E = 3.3$ – see also fig. 4.5. All values are given in arbitrary units, scaled for better visibility. The white lines show boundaries of the classically forbidden region.

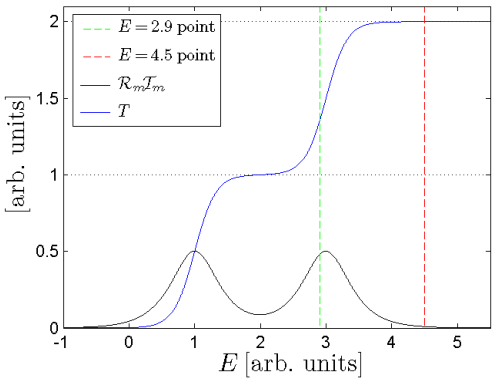


Figure 4.7: Plot of the transmission and the $\mathcal{R}_m\mathcal{T}_m$-prefactor as functions of energy. for the two lowest steps and plateaus. Example step and plateau points, for which the maps of the following figure were taken, are marked with dashed horizontal lines.

originating counterpart (sec. 3.4.9). Apart from this, the fringing patterns in both these signals have the same spatial periodicity, but the SGM signal oscillates around zero, with positive and negative crests – hence it has twice more fringes. Also, the oscillating patterns are shifted in phase: the maxima of the charge density fall in the places where the SGM signal has its nodes. The spatial period decreases as we move away from the centre, because the QPC potential falls, leaving more and more energy for the propagation.

In the plateau region, on the other hand, with the reflection coefficient very small, the asymmetry is no longer visible and both signals, the SGM correction and the charge density, can indeed be related. The signal does not oscillate, which is in agreement with our predictions about the SGM maps in the adiabatic case. It does not, however, allow relating it to the current density.

The current density, plotted in parts (c) and (d) of the fig. 4.6 has the shape of a straight ray, non-broadening and not fading, which follows from the bounded, and hence real, character of the transverse wave functions. Together with the fading charge density, this means that the electrons move faster as they are further from the centre. As we mentioned in subsec. 3.4.12, the lack of scaling with the electron velocity is another important difference between the non-invasive SGM signal and the current density.

Analogous images, but for the second step- and second plateau-region points, have been plotted in fig. 4.8 on p. 87. The relevant total transmission curve, the $\mathcal{R}_m\mathcal{T}_m$-function and the chosen energy-points have been shown in fig. 4.7. All the quantities are now created by the contributions from two states, $m = 1$ and $m = 2$, with the notable exception for the SGM correction in the step region (part (e)). This one is not affected by the $m = 1$-state and therefore has a nodal line along the axis of the QPC. From the SGM correction in the plateau region (p. (f)) and the Fermi-energy charge density

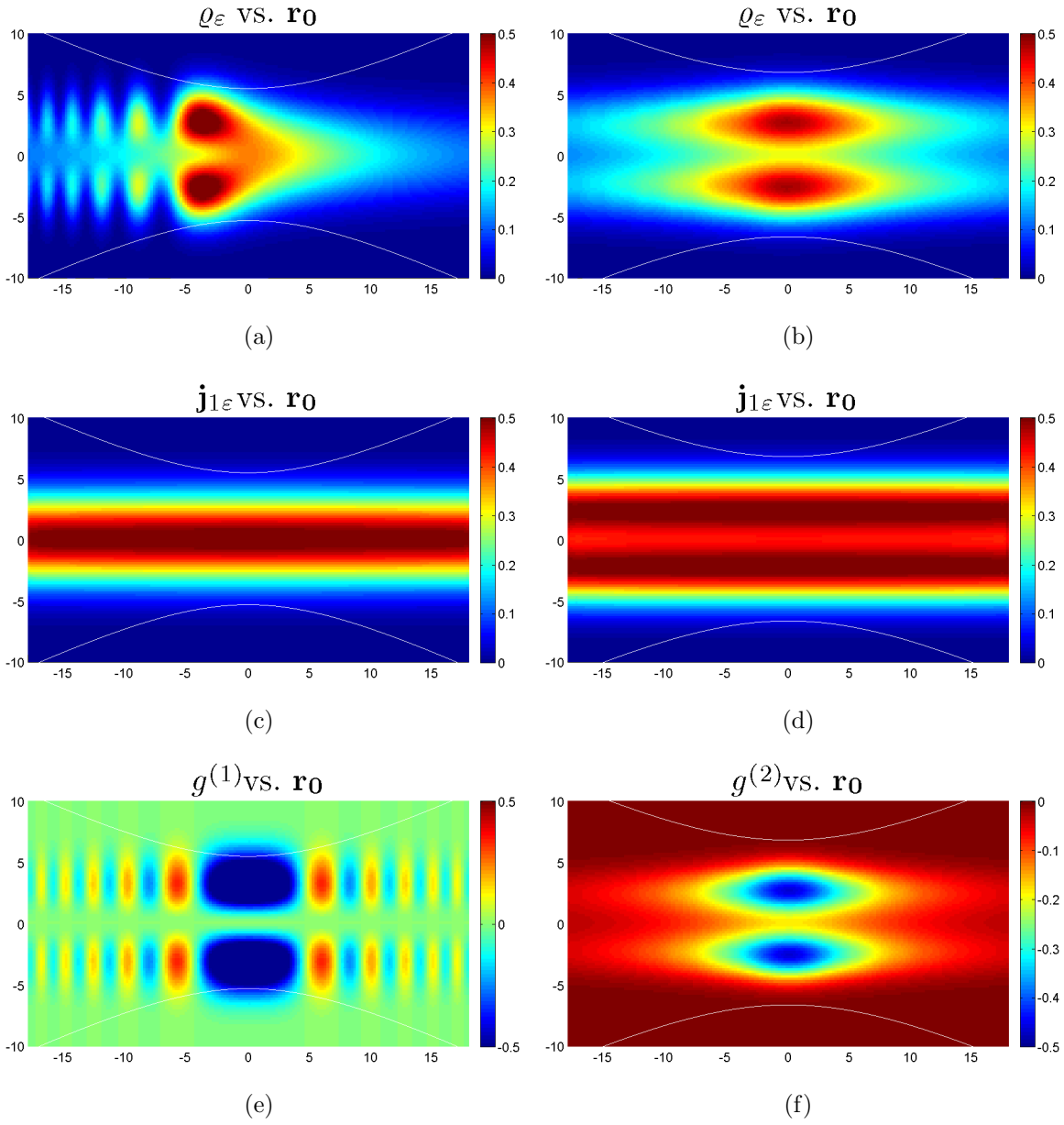


Figure 4.8: Saddle-point model of the QPC, second step and second plateau regions. Plots of local densities:
(a), (b) – charge density at the Fermi energy,
(c), (d) – current density,
and of the SGM signal, given by the:
(e) – first order conductance correction,
(f) – second order conductance correction.
Left panel maps (a,c,e) are taken in the step region, at $E = 2.9$, and the right panel ones (b,d,f) are taken in the plateau region, at $E = 4.5$ – see also fig. 4.7. Boundaries of classically forbidden region are marked with white lines.

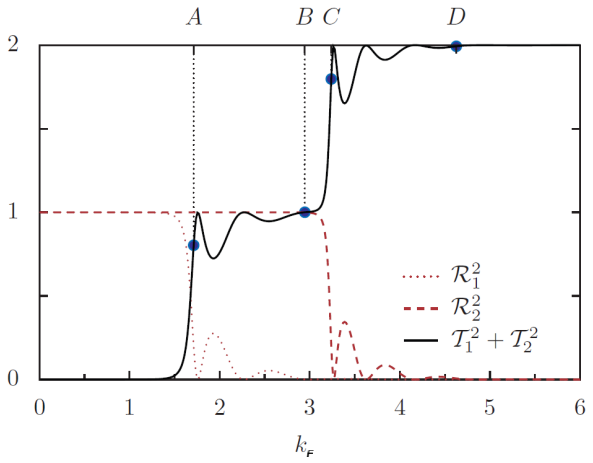


Figure 4.9: Plot of the total transmission and individual reflections from the two lowest eigenstates as functions of the Fermi wave number, given in the units of W_{QPC}^{-1}. The letters A-D mark the example energy points for which the maps of two following figures were taken.

in both, step and plateau (parts (a-b)), we see that the second eigenstate's contribution dominates – this is caused by the wave function normalisation factors $1/\sqrt{k_m}$, which are much greater for the states just having been open than for all the states lying below. The plateau-region current density, which has this factor removed by the velocity-multiplication, shows a significant value on the longitudinal axis (p. (d)). In the step region, on the other hand, the $m = 2$-state contributes very little to the current, because of the relatively small value of $\mathcal{T}_m^2$ ($\sim$0.12).

4.2.3 Abrupt QPC model

Not only the saddle-point model can offer us analytical solutions. We would like to mention very briefly an analytical model for the opposite type of structures – abrupt and hard-walled, which are highly mode-mixing.

The already-mentioned work by Szafer and Stone [72] introduced the so-called Mean Field Approximation, thanks to which the transmission problem in such a structure can be solved analytically. It does not, however, provide the access to the scattering state wave functions. Based on a similar approach, the wave functions have been obtained in [44] through a less restrictive approximation, the Smooth Field Approximation (SFA). Both approximations require the leads to be much wider than the constricted region ($W/W_{\mathrm{QPC}} \to \infty$).

An important point which SFA shows, is that even in this non-adiabatic geometry a direct one-to-one correspondence between the constriction modes and the transmission eigenstates is kept (see eq. (52) in [44]).

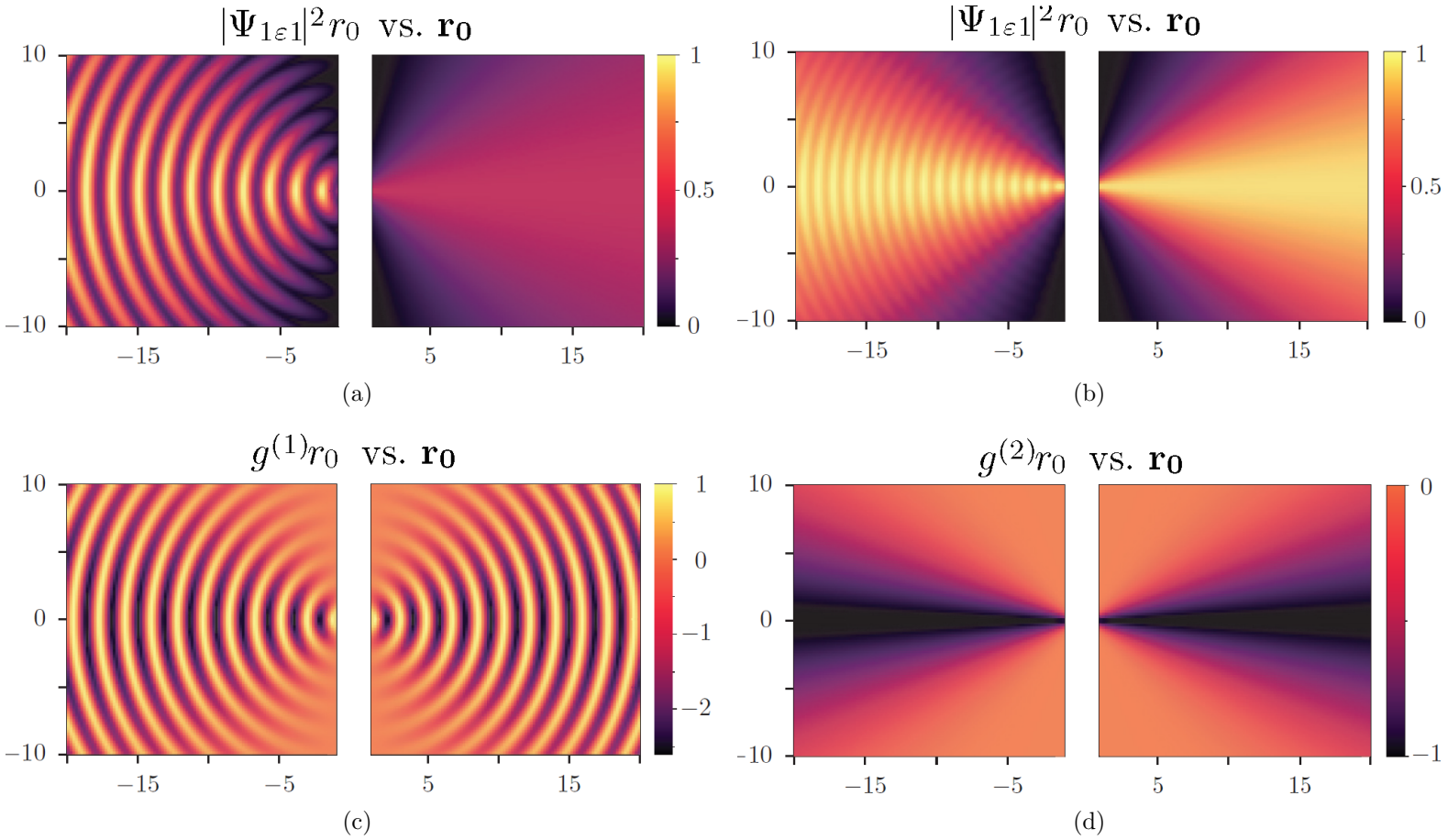


Figure 4.10: Abrupt QPC, SFA-model. First conductance step and plateau regions. Images of the:
(a) – squared amplitude of the first eigenstate wave function in the step region (energy point A of the fig. 4.9),
(b) – charge density at the Fermi energy in the plateau region (energy point B of the fig. 4.9),
(c) – SGM signal in the step region (point A), represented by the first-order conductance correction,
(d) – SGM signal in the plateau region (point B), represented by the second-order conductance correction.
All maps are scaled by the tip distance from the QPC exit to compensate the decay of the signal. Adapted from [44].

At large enough distances from the QPC opening, such that

$$k_F r_0 \gg 1 \ , \tag{4.23}$$

the obtained scattering eigenstate wave functions can be separated in polar coordinates into the radial parts (incoming and outgoing), whose variation is given by the ordinary circular waves $\frac{1}{\sqrt{k_F r_0}} \exp(\pm i k_F r_0)$, and the angular part. The latter is peaked at those polar angles $\theta^{(m)}$, for which the transverse wave number in the wide region matches the transverse wave number of the constriction mode: $|k_F \sin\theta^{(m)}| \approx k_m^{(\perp)}$.

In figures 4.10-4.11, we display the calculated [44, 78] maps of the left-incoming $m = 1$ and $m = 2$ eigenstates' charge densities, i.e. the squared amplitude of their wave functions, and the maps of the SGM signal, represented by the first or second order corrections. The transverse and longitudinal position axes are scaled with respect to the QPC opening width, W_{QPC}. The maps are multiplied by the distance between the tip and the QPC opening, r_0, for better visibility of the signal. The general features of the signal, including the non-fringed character of the plateau-type correction, are very similar to the adiabatic signal features. The main difference is that now the signal spreads angularly, which was not the case in the ideally adiabatic model.

The great regularity of the SGM for this abrupt constriction case is owed to the infinite leads' width limit – which essentially means the lack of lateral boundaries – and it will not be shared by the simulated maps of Ch. 5.

Another important conclusion from the SFA-model is that the non-invasive SGM signal is proportional to the current density of the unperturbed structure [44] – and at the same time to the Fermi-energy charge density – if the system is tuned to the conductance plateau and only for structures with a strict four-fold spatial symmetry. Naturally, the constriction shape of this model always has this symmetry, but the addition of a disorder potential will break it, and hence also the relation between the signal and the current density. Also, a structure with laterally constricted wide region will not allow for such proportionality.

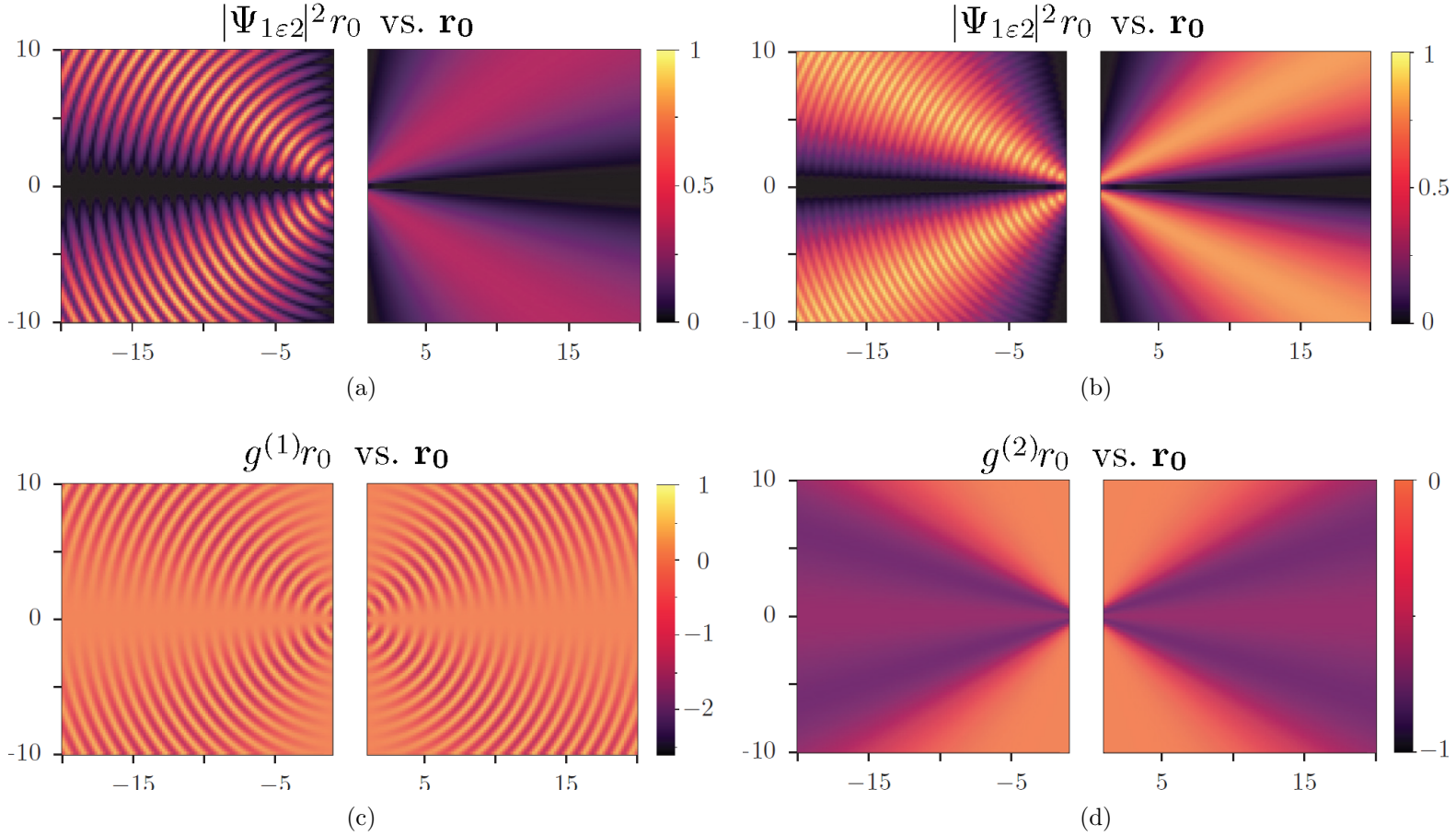


Figure 4.11: Abrupt QPC, SFA-model. Second conductance step and plateau regions. Images of the:
(a) – partial charge density, due to $m = 2$ eigenstate, at the Fermi energy, step region (point C of the fig. 4.9),
(b) – partial charge density, due to $m = 2$ eigenstate, at the Fermi energy, plateau region (point D of the fig. 4.9),
(c) – SGM signal in the step region (point C), represented by the first-order conductance correction,
(d) – SGM signal in the plateau region (point D), represented by the second-order conductance correction.
All maps are scaled by the tip distance from the QPC exit to compensate the decay of the signal. Adapted from [44].

Chapter 5

Simulation of SGM experiments in QPCs

In this chapter, we present and discuss the results of numerical simulations of the SGM signal. The simulations give us the direct access to the details of the signal in the concrete systems, whereas our theory only sets the frames for it. They can also serve as an additional test, confirming the predictions of the theory (as in [45]), and as a way to go beyond it by including arbitrarily strong tip potentials.

We consider not only the SGM maps, but also the simulated conductance curves. Apart from informing about the step and plateau region ranges, they also tell us about the achieved quality of the conductance quantisation – both these things are crucial to the nature and to the interpretation of the non-invasive SGM signal. As a natural generalisation of such a study, we will also plot the conductance curves taken in the presence of the SGM tip, which, compared to the original ones, provide a simple method for investigating the tip influence on the structure, before the SGM maps are taken. Certain insight can be extracted by analysing such curves in terms of the possible electron interference paths, like it is often done in the case of experimentally obtained SGM maps (see Ch. 1), but the scope of this approach is limited. As we have mentioned in Ch. 3, the conductance correction curves are complementary to the maps, in the sense that they provide information on energetic and spatial variation of the tip potential matrix elements $\mathcal{U}^{21}$ introduced in Ch. 3.

Finally, we will present the simulated SGM maps pertaining to both, small and large tip strengths. The character of step- and plateau-regime signals will be discussed, as well as additional factors which influence the shape of the SGM images. In this context, a special attention will be paid to the adiabaticity of the considered geometries, and to the processes of tunnelling

and Fabry-Perot resonances.

5.1 The method

In order to simulate the Scanning Gate Microscopy signal, we have to calculate the conductance through our lead-structure-lead system (fig. 2, Ch. 2.1) for two cases: with and without the probe, and subtract them. The method should allow for variation of the energy of the travelling electrons and the tip position, so that we could produce the transmission curves and SGM maps for Fermi energies of our choice.

5.1.1 Calculation scheme

Conductance from Green function

As we described it in Ch. 2, the conductance is equivalent to the total transmission, which can be evaluated by taking an incoherent sum of the squares of the transmission amplitudes, eq. (2.58). The amplitudes are, in turn, obtained from the retarded Green function linking one side of the structure, i.e. one lead-structure interface, with the other one. This is done using the Fisher-Lee relation, eq. (2.61a). Finally, to calculate the Green function, we need to invert the wave operator of the Schrödinger equation of our system, $E^+ - H_0$.

Discretisation

For the computer solution of this task, we employ a tight-binding (T-B) model where the structure (see e.g. fig. 4.2, Ch 4) is represented by a set of points on a finite rectangular M-by-L lattice. The lattice spacing, denoted by a_{TB}, becomes now the natural unit of length, and the position is expressed by a pair of integers (i, j), the first one being the row index (i.e. the y-coordinate) and the second one being the column index (x). The natural unit of energy is set by the hopping term of the tight-binding Hamiltonian $t_{\mathrm{TB}} = \hbar^2/(2M_e a_{\mathrm{TB}}^2)$. The system of units generated by these two is called the Anderson units; we will use it across this chapter without further mentioning.

Due to its discreteness, the T-B scheme is less accurate for faster spatial variations, and not capable at all of reproducing oscillations of wavenumber greater than π, where the T-B band starts closing (the entire band range is from 0 to 2π). As a rule of thumb, we shall try not to exceed $k = 0.5$, for which the relative error in energy estimation is about 2%.

For a correct description of the transmission process, the T-B Hamiltonian has to take into account the boundary conditions, set by the hard-wall closing of the structure in the transverse direction and by the perfectly open interfaces with the leads in the longitudinal direction. The latter introduce complex self-energy terms to the Hamiltonian. (As a by-product of this, we can now neglect the infinitesimal imaginary part in the wave operator $E^+ - H_0$.)

RGF algorithm

Instead of inverting the entire wave operator $E - H_0$ at once, we use a considerably faster method, the Recursive Green Function (RGF) algorithm [72,79]. Here, the structure is subdivided into a collection of L vertical stripes (i.e. stretching in the transverse direction), and each of the stripes, considered as decoupled from all the rest, has its individual green function calculated by inverting its own wave operator. Then, the Green function across the structure is built up, step by step, by attaching subsequent individual stripes' Green functions to the leftmost one. Likewise, the analytically known Green function of the lead is attached to the first stripe's one in the beginning, and to the entire structure's function in the end. This ensures the open lead-structure interface.

The computational complexity of this procedure scales like M^3L for a single transmission calculation. An important advantage of this algorithm is its numerical stability. It does not provide, however, the access to the local densities, which we had in the previous chapter.

5.1.2 Implementation

An existing FORTRAN program, performing the computation described above, has been adopted and developed, so that in addition to the transmission curves $T(k_F)$ it could also produce the SGM maps $\Delta T(\mathbf{r}_0)$. The output data and plotting are handled using MATLAB computation environment. The parameters determining the system state while performing the calculation are: lattice size, constriction shape, lattice potential, tip potential and (unused) magnetic field. The particle energy E_F is treated as a dependent variable linked to the independent one – the wavenumber – by the discrete dispersion relation.

Since, for a full SGM scan, the tip will be placed on all the lattice sites (or a fraction, if the scanning range is restricted), the overall complexity will be M^4L^2.

5.1.3 Parameter space and preliminary choices

QPC geometry

The procedure described above can be performed with an arbitrary potential on the discretisation lattice. For simplicity, however, we will model our quantum point contact with a zero-potential hard-wall model, where the constriction is implemented by narrowing of the lateral boundaries. (The results do not differ much with respect to the ones obtained with the QPC defined by a potential.) We chose three specific geometries, denoted concisely as HC, MC and SC, and defined in the following way:

– **HC geometry** – based on the shape of level lines of a saddle-point potential, the boundaries are hyperbolas $y(x)$, parametrised by x_{sc} and y_{sc}:

$$\pm\, y = \sqrt{y_{sc}^2 + (x/x_{sc})^2} + 1 \ ; \tag{5.1}$$

we will describe the HC constriction by giving its minimum width

$$W_{\mathrm{QPC}} = 2(y_{sc} + 1) \ , \tag{5.2}$$

and its characteristic length, defined as the position at which the hyperbola's asymptote meets the lateral lattice edge,

$$\eta_{\mathrm{QPC}} = (M + 1)\, x_{sc} \ ; \tag{5.3}$$

– **MC geometry** – whose boundaries are given by exponential curves:

$$\pm\, y = y_{sc} \exp\left[(x/x_{sc})^2\right] + 1 \ ; \tag{5.4}$$

with the width

$$W_{\mathrm{QPC}} = 2(y_{sc} + 1) \ , \tag{5.5}$$

and the characteristic length

$$\eta_{\mathrm{QPC}} = x_{sc} \ ; \tag{5.6}$$

– **SC geometry** – with the abrupt, rectangular boundaries:

$$\pm\, y = y_{sc} + 1 \ , \quad \text{for} \quad |x| \leq x_{sc} \ ; \tag{5.7}$$

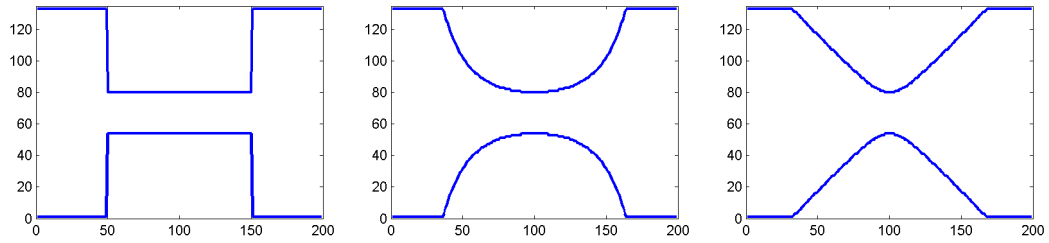
the constriction width is

$$W_{\mathrm{QPC}} = 2(y_{sc} + 1) \ , \tag{5.8}$$

while its length is given by

$$\eta_{\mathrm{QPC}} = 2x_{sc} + 1 \ . \tag{5.9}$$

The lateral position will usually be shifted so that the all the points inside the constriction can be described with positive indices. We plot an example of each of these geometries in the fig. 5.1.



(a) SC geometry (b) MC geometry (c) HC geometry

Figure 5.1: Example constrictions of the three QPC geometries considered in this chapter.

Tip potential

Also the tip effect is introduced into our calculation by adding its potential $V_{\mathrm{T}}(\mathbf{r})$ on the appropriate lattice site or sites. Here, we have the choice over the tip's integrated strength U_0, shape $\frac{1}{U_0}V_{\mathrm{T}}(\mathbf{r}-\mathbf{r}_0)$, and, naturally, on the range of the tip scan.

Shape Following the conclusions of sec. 3.4.8, Ch. 3, we will use only one tip potential shape, the discrete delta function, which puts its energy V_0 on one lattice point only. The spatial extent of the tip is then just one lattice spacing, $\Delta r_0 = 1$, which means that V_{T} and the integrated strength $U_0 = V_0\,\Delta r_0^2$ can be used interchangeably.

Strength From the point of view of interpreting the signal, it will be important to distinguish between the weak, non-invasive tip potentials, and the strong ones. We will therefore perform our simulations in two different tip strength modes.

Certain way of laying a border in between the two regimes is through the eq. (3.51) of subsec. 3.4.3. As we will see, the propagation wavelength as inferred from the fringes of the signal does not exceed a few tens of lattice spacings, suggesting the condition $V_0 \ll 10^{-2}$. This estimate, however, is not completely satisfactory, since, among the many wide-region modes, the one which just has been opened will be very slow, setting a much more restrictive requirement (but pertaining to only a small contribution towards the entire process).

In the numerical simulations we will rather attempt at achieving the non-invasiveness by simply taking the tip strength as low as possible, while still keeping away from the numerical noise level. Thus, we will consider the

infinitesimal differences of the transmission and represent them by numerical derivatives with respect to the tip strength, evaluated at $V_0 = 0$, with a possibly small value of the potential step ϵ (see next subsection). The linear and quadratic regions in the transmission vs. tip strength curves are where the non-invasiveness is certainly achieved.[1]

5.1.4 Differential signal

The first and second order differential signal from the SGM are calculated by taking the right-side three-point numerical derivatives:

$$\left.\frac{\partial T}{\partial V_0}\right|_{V_0=0^+} \cong \frac{-T(2\epsilon) + 4T(\epsilon) - 3T(0)}{2\epsilon} , \tag{5.10a}$$

$$\left.\frac{\partial^2 T}{\partial V_0^2}\right|_{V_0=0^+} \cong \frac{T(2\epsilon) - 2T(\epsilon) + T(0)}{\epsilon^2} . \tag{5.10b}$$

The accuracy of this calculation is of the order $O(\epsilon^2)$ in the case of the first derivative, and $O(\epsilon)$ in the case of the second one. We will usually set the infinitesimal element of the tip potential to $\epsilon = 10^{-5}$.

In order to relate these signals to the theoretical results of Ch. 3, we use the fact that

$$\frac{g^{(1)}}{U_0} = \frac{\partial g}{\partial U_0} = \frac{1}{\Delta r_0^2}\frac{\partial g}{\partial V_0} , \qquad \frac{g^{(2)}}{U_0^2} = \frac{1}{2}\frac{\partial^2 g}{\partial U_0^2} = \frac{1}{2\Delta r_0^4}\frac{\partial^2 g}{\partial V_0^2} , \tag{5.11}$$

and we write the delta-tip expressions (3.56a) and (3.57), in the form

$$\frac{\partial T}{\partial V_0} = 4\pi \sum_{m=1}^{N} \mathcal{R}_m \mathcal{T}_m \,\mathrm{Im}\left\{ \overset{\approx}{\Psi}{}^{*}_{2\varepsilon m}(\mathbf{r}_0)\, \overset{\approx}{\Psi}_{1\varepsilon m}(\mathbf{r}_0) \right\} , \tag{5.12a}$$

$$\frac{\partial^2 T}{\partial V_0^2} = -8\pi^2 \sum_{\substack{m\in\underline{\mathcal{M}} \\ m'\in\underline{\mathcal{M}}}} \left| \overset{\approx}{\Psi}{}^{*}_{2\varepsilon m}(\mathbf{r}_0)\, \overset{\approx}{\Psi}_{1\varepsilon m'}(\mathbf{r}_0) \right|^2 . \tag{5.12b}$$

where the product $\overset{\approx}{\Psi}{}^{*}_{2\varepsilon m}(\mathbf{r}_0)\, \overset{\approx}{\Psi}_{1\varepsilon m'}(\mathbf{r}_0)$ is nothing else but a normalised and local version of the matrix element $\mathcal{U}^{2\,1}_{mm'}$. If we were to consider the tips of finite spatial extent, the integral would have to be retrieved, but also the factors Δr_0^2 and Δr_0^4 would have to be added in the right-hand side of the first and second of these equations.

[1] Because if the finite difference is well described by the (n-th) differential, then no higher terms ($> n$) give a significant contribution.

Our theory of Ch. 3 operates in the regime of local and weak tip potentials. Comparing the non-invasive differential signal with the difference signal in the strong-tip regime can provide a way of assessing the relevance of our theory in a wider context.

5.2 Transmission curves

The curves of the transmission as a function of energy calculated with various methods, including the RGF, were already published some time ago [72, 73]. However, we shall pay some time to investigating them, since they are an indispensable basis for what we do in the following sections. As we saw in Chapter 4, different the behaviours of the SGM signal in the plateau and step regions follow from the vanishing of the $\mathcal{R}_m\mathcal{T}_m$ factors in the perturbative expressions. So, any deviations from the integer transmission values on the plateaus will mix the two behaviours. Though it is only for small tip voltages that this discrimination follows from our theory, we will see later that the results with a finite tip potential reveal a similar behaviour.

5.2.1 Wire

First, we consider results for the oversimplified case: the ideal wire, described previously in sec. 4.1. They serve us as a simple check on the correctness of our programs and as a reference for the QPC transmission curves which will appear further. The total transmission T as a function of the Fermi wavenumber k_F is shown in the fig. 5.2. The curves are for three different widths of the wire: $W = 4, 14, 34$. (These correspond to $M = 3, 13, 33$ lattice sites in the transverse direction.) The wavenumber covers the entire band range, $k_F \in (0, 2\pi)$.

We see that the perfect steps, (ch. 4), are reproduced. The leads' width is set to the same value as the one of the wire's, so we do not get any features in the curve: the plateaus are exactly flat (up to the penultimate digit of the machine-represented number) and the steps rise over a single k-point interval, so we can securely say that there is no ramp region at all. The steps rise exactly at the positions $k_a^{(\perp)} = \pi a/W$, in accord with the eq. (2.21) and the preliminary discussion of sec. 4.1.

5.2.2 Quantum point contacts

We now pass to quantum point contacts. We would like to discuss the shape of their transmission curves, how they deviate from the perfectly quantised

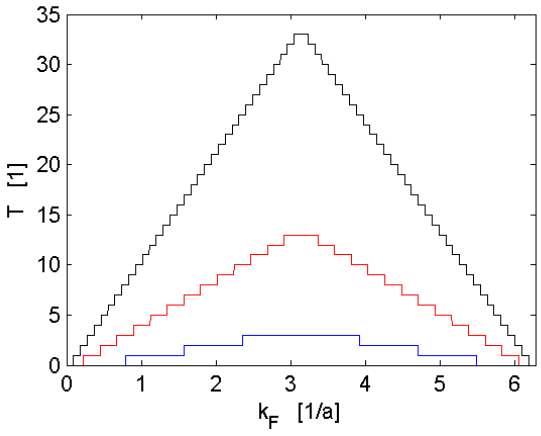


Figure 5.2: The $T(k_F)$ curves for the ideal wire structure for three different widths. Perfectly quantised steps and band closing are seen.

shape we had in the case of the wire, and the factors that influence the quality of its plateaus and steps.

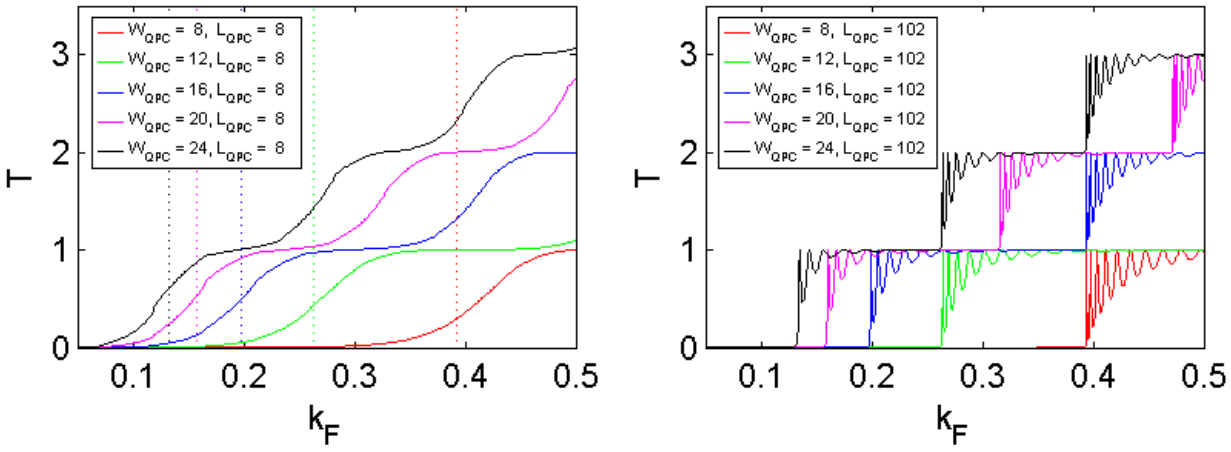
The discretisation board contains $M = 133$ (i.e. $W = 200$) lattice sites in transverse dimension and $L = 199$ in the longitudinal one, except when the HC-constriction length η_{QPC} surpasses 199, in which case we take $L = \eta_{\mathrm{QPC}} + 3$.

SC constriction

Transmission curves of an SC Quantum Point Contact are plotted in the fig. 5.3(a) and in the fig. 5.3(b). As expected, the position of the steps is controlled by the constriction width, the step thresholds are at around

$$k_m^{(\perp)} = \frac{\pi m}{W_{\mathrm{QPC}}} \ . \tag{5.13}$$

These positions are marked for the first step of each curve by the dotted lines of the corresponding colour (fig. 5.3(a)). Still, these positions do not mark the middle of each step region and the transmission there does not assume the value of $T = 0.5$ as might be expected from the adiabatic model, eq. (4.9). Instead, we have more tunnelling in the wider structures, which pushes the transmission curve towards the left with respect to the threshold, and more reflection in the narrower structures, causing the curve to move towards the right. In the case of longer structures, fig. 5.3(b) the dotted threshold lines have been omitted, but the threshold positions $k_m^{(\perp)}$ are always at the very beginning of the ramp regions, with low T values. The tunnelling now is



(a) (b)

Figure 5.3: Total transmission through an SC-QPC as a function of the Fermi wavelength for various constriction widths, W_{QPC}, and lengths, L_{QPC}. (a) The short (small aspect ratio) QPCs show rounded steps because of the tunnelling through the closed channels. (b) The elongated QPCs show sharp steps, but with a Fabry-Perot-type resonant structure.

negligible, but the ramp's steepness is spoilt by another mechanism, the Fabry-Perot (F-P) resonances. We will now discuss these two, for it is their interplay that decides about the quality of the conductance quantisation.

The essential difference between the figures 5.3(a) and 5.3(b) is the length of the structures they feature. It is, however, not the absolute length that makes this classification, but the aspect ratio $L_{\mathrm{QPC}}/W_{\mathrm{QPC}}$.

Fabry-Perot resonances As mentioned in [72,73], the entrance-exit interference resonances in m-th eigenstate transmission start appearing whenever the phase change on passing twice the length of the constriction can reach at least π; for the plateau associated with the m-th constriction state this is

$$2k_m\, L_{\mathrm{QPC}} \geq \pi\ . \tag{5.14}$$

Following from the resonant interference condition on the base path ($2k_m L_{\mathrm{QPC}} = \pi n\,,\ n \in \mathbb{N}$), the resonant features appear at the positions

$$k_F^{(\mathrm{FP})} = \sqrt{\left(\frac{\pi m}{W}\right)^2 + \left(\frac{\pi n}{2L}\right)^2}\ , \tag{5.15}$$

in k-space, giving peaks when n is even and valleys when n is odd. The above equation is satisfied only approximately, nevertheless it gives a useful

estimate. Since k_m is the longitudinal wavelength inside the constriction, $k_m = (k_F^2 - k_m^{(\perp)2})^{1/2}$, the condition (5.14) boils down to:

$$\frac{L_{\mathrm{QPC}}}{W_{\mathrm{QPC}}} \geq \frac{1}{\sqrt{2m+1}} \ . \tag{5.16}$$

The resonances are of Fabry-Perot type and their number on the plateau grows with growing aspect ratio. All their maxima reach the plateau level exactly. In addition, their minima for varied constriction lengths lie on a single curve[2], as long as the width is kept constant. This means that if we could strip off the transmission curves from the F-P effect, the long structures' transmission would not depend on the length. As we move on the plateau in the positive k_F direction, the resonances fade and become broader, but when we consider the transmission as a function of k_m, which is not shown in the figures, then the period is approximately constant and close to the theoretical value $\Delta k_m = \pi / L_{\mathrm{QPC}}$.

Tunnelling On the other hand, when the aspect ratio is smaller than the value from the criterion (5.16), the tunnelling through the first evanescent constriction mode becomes important [72], as it depends on

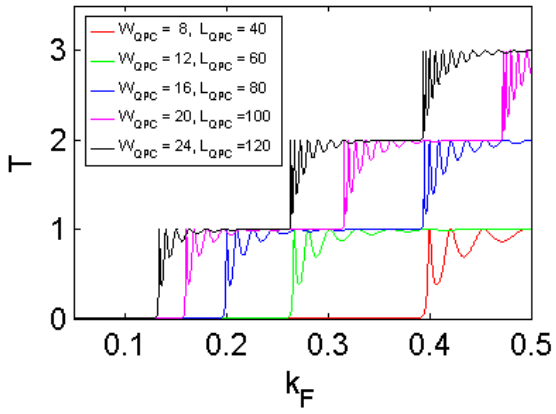
$$\exp[-(\varepsilon_{m+1}^{(\perp)} - E_F)^{1/2} L_{\mathrm{QPC}}]$$

and spoils the plateau quality by giving it a visible slope. Both of these effects can be seen from the fig. 5.3.

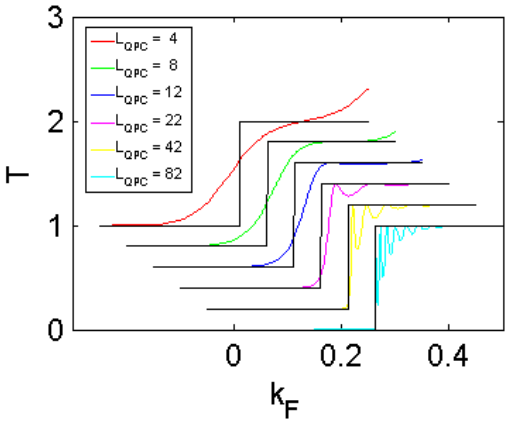
Quality of the plateaus Curves for different structure dimensions but a fixed aspect ratio, such as the ones displayed in fig. 5.4(a), match each other quite well, when scaled in terms of the variable $k_F W_{\mathrm{QPC}} - \pi m$. This additionally confirms that it is indeed the aspect ratio that parametrizes the plateaus' quality. We can see the crossover from the short-QPC to the long-QPC behaviour in the fig. 5.4(b), where the SC transmission curves for single constriction width ($W_{\mathrm{QPC}} = 12$) but multiple lengths are plotted. The criterion (5.16) is approximately satisfied, though the F-P features just having entered the plateau from the right are poorly visible.

Clearly, seeking a good quantisation of conductance requires trading-off between the rounded shape of the transmission curve due to the tunnelling and the Fabry-Perot resonant structure in it. The latter poses only a small problem for the computer simulation of the SGM signal maps, since we are able to fine-tune the transmission to an F-P peak and make the $\mathcal{R}_m \mathcal{T}_m$ vanish.

[2]This curve is well fitted by an exponential of the type $1 - \exp(-Ak_m + B)$.



(a) Total transmission through an SC-QPC as a function of the Fermi wavelength for a fixed aspect ratio $AR = 5$.



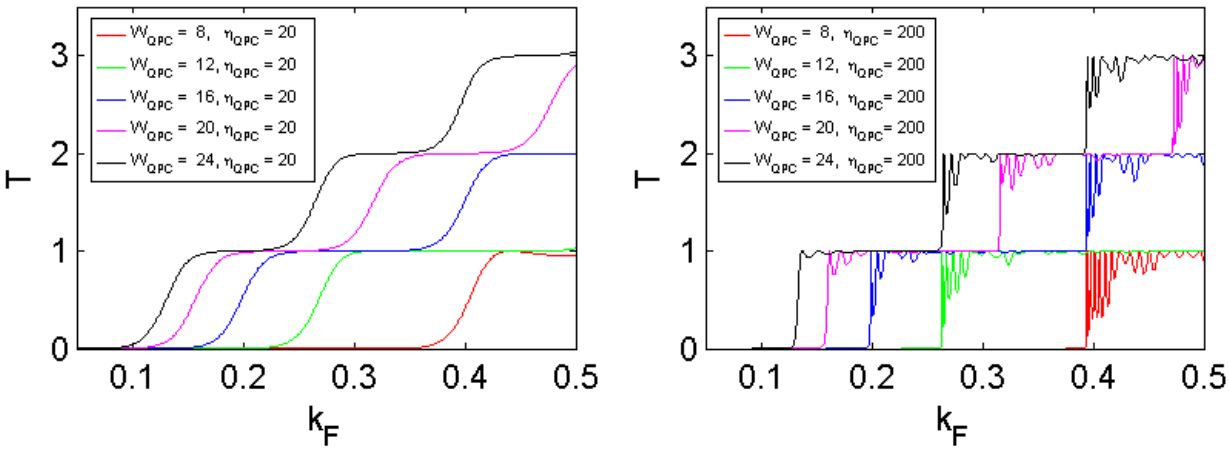
(b) Total transmission through an SC-QPC as a function of the Fermi wavelength for $W_{\mathrm{QPC}} = 12$ and various constriction lengths. The reference lines (black) are the transmission curves for an ideal wire of the same width. The curves are offset for better visibility and the axes annotation applies to the last one ($L_{\mathrm{QPC}} = 82$).

Figure 5.4

In experiment, however, the finite temperature averaging will smooth out the peaks, rounding off the sharp step shape and the F-P peak shape, as well as pulling the transmission values below the plateau. Furthermore, the difficulty in precise estimation of the series resistance that needs to be subtracted when measuring the one of QPC's and the presence of other irregularities, such as residual disorder to which the samples are very sensitive at low temperatures, can make the fine tuning troublesome (see e.g. [80, 81]).

MC and HC geometries

In the case of MC and HC quantum point contacts, where we have no clear-cut constriction length, we use the characteristic length η_{QPC} instead (see subsec. 5.1.3). In the figure 5.5 we present the transmission curves for MC structures, short (a) and elongated (b). With a smooth boundary the position where reflections occur is now blurred, which gives rise to smoothing of possible Fabry-Perot resonances. At larger aspect ratios we still observe the resonant structure in the plateaus, but the trade-off we just mentioned in the previous paragraph becomes easier when we deal with MC QPCs. This can be seen explicitly from the fig. 5.6(a), where the MC-curve has not only



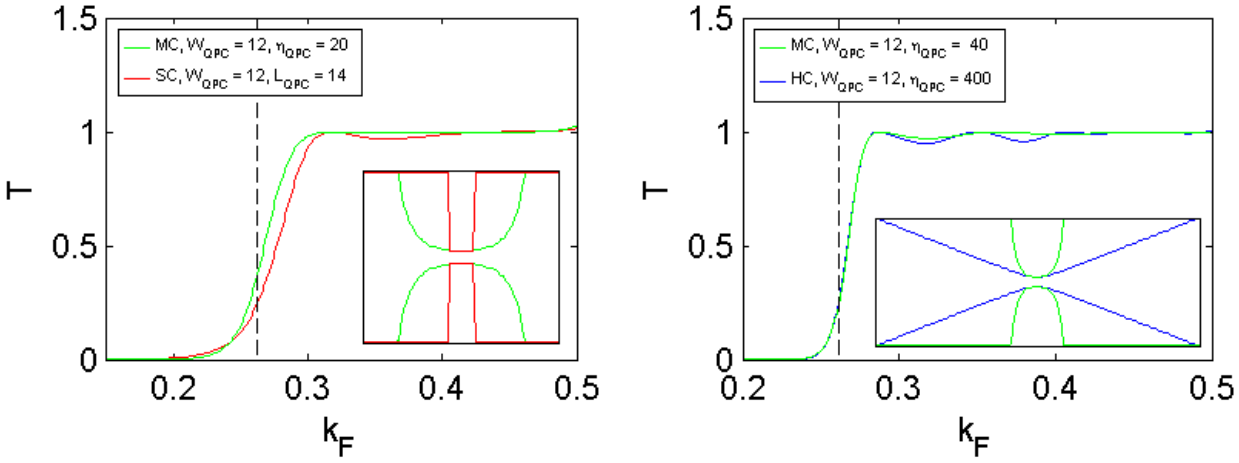
(a) (b)

Figure 5.5: Total transmission through an MC-QPC as a function of the Fermi wavelength for various constriction widths, W_{QPC}, and lengths, η_{QPC}. (a) The short QPCs. (b) The elongated QPCs – we can see the beatings of the Fabry-Perot resonances.

better quality of quantisation on the plateau, but also takes smaller extent along the k-axis to rise at the step. For large aspect ratios, in 5.5(b), we see additional irregularities, which are due to the non-abrupt character of the constriction opening – the wave parts reflected at different longitudinal positions interfere with each other, giving rise to the beating-like patterns in the curves.

The transmission curves for the HC constrictions can follow the step shape for corresponding[3] MC curves quite closely, which was not the case with SC geometry. The conductance quantisation accuracy on the plateau is slightly better in the MC case – see fig. 5.6(b). We therefore conclude that attempts at making the structure very smooth by tapering do not yield better quantisation, and that it is indeed the *local* adiabaticity that counts, as long as the the transmission curve shape is concerned. The MC geometry also has an added advantage from the technical point of view, because it does not require as large structure lengths as the HC one.

[3]This correspondence is, in general, not easy to establish. Because of the resonances, the plateau quality is not a good measure. We tried to match the curves according to their slope in the middle of their steps (the ramp regions). An estimate might be obtained from the adiabatic parameter of Glazman's formula (see below), but we do not use it as an *a priori* measure for the equivalence.



(a) (b)

Figure 5.6: A comparison of the total transmission vs. Fermi wavelength curves between: (a) SC- and MC-QPCs; (b) MC- and HC-QPCs. The insets compare schematically the geometries of the constrictions. The dashed lines show the step positions for an ideal wire of the same width.

Adiabaticity Having the knowledge of the structure's shape we can evaluate the adiabatic parameter of the Glazman formula (subsec. 4.1.3) and compare the RGF-computed transmission curves against the globally adiabatic model. The agreement is good for the structures, which are not too short, both HC and MC. An example with the HC structure, first conductance step, is given in the fig. 5.7 (a-c). When we shorten the structures, making them more abrupt, the agreement between the analytic curves and the ones calculated for the same constriction diminishes. On the other hand, with the SC structure, where the adiabatic parameter cannot be assessed, and which is highly non-adiabatic, it is possible to imitate the adiabatic transmission curve quite well by just choosing a suitable structure length, see fig. 5.7 (d).

It is also worth noting, that even when the agreement is quite good, the globally adiabatic model systematically underestimates the ramps' steepness. At first sight this seems counter-intuitive, because we expected the adiabatic transmission to yield the best quality conductance quantisation. It seems that the influence of the mode mixing not only does not spoil the quality, but can even slightly enhance it. This happens because in the step region the off-diagonal r and t terms follow a similar variation as the diagonal ones, cf. [69].

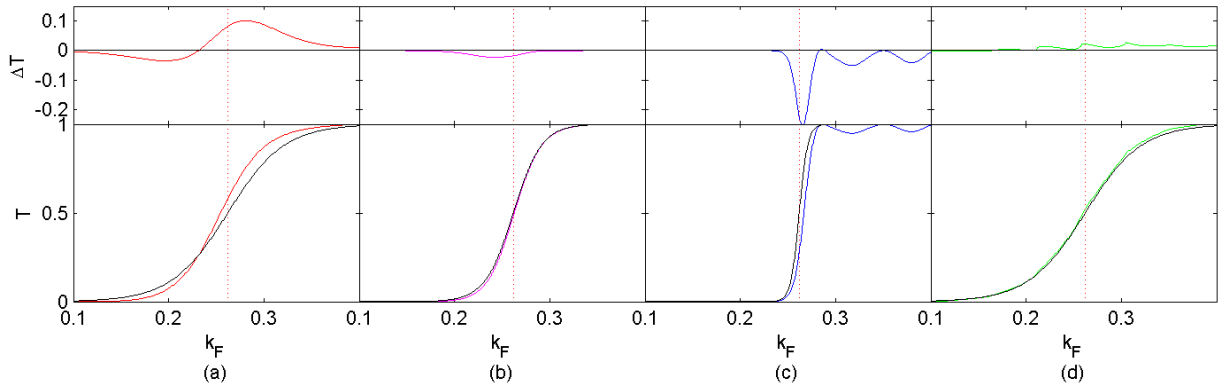


Figure 5.7: Total transmission vs. Fermi wavelength $T(k_F)$ compared with the adiabatic model prediction $T_{adiab.}(k_F)$ (black reference curves) for a HC structure: (a) short: $W_{\mathrm{QPC}} = 12$, $\eta_{\mathrm{QPC}} = 60$; (b) intermediate: $W_{\mathrm{QPC}} = 12$, $\eta_{\mathrm{QPC}} = 120$; (c) elongated: $W_{\mathrm{QPC}} = 12$, $\eta_{\mathrm{QPC}} = 400$; and for: (d) a short SC structure, $W_{\mathrm{QPC}} = 12$, $L_{\mathrm{QPC}} = 5$, with the adiabatic reference curve taken from (a). The upper insets show the difference $T(k_F) - T_{adiab.}(k_F)$. The dashed lines show the constriction mode threshold $k_m^{(\perp)}$, $m = 1$.

We therefore see that, although there is some relation between the shape of conductance curves and adiabaticity, we cannot use the curves to assess the degree of the adiabaticity of a structure.

5.3 Influence of the SGM tip

We introduce now the charged tip to our system, hence obtaining a complete SGM setup. Before simulating the SGM signal maps, it is of a natural interest to see how the tip's presence affects the transmission curves that we just discussed. We will plot the transmission curves $T(k_F)$ for a few different tip potential values and placements, and the difference curves $\Delta T(k_F)$, resulting from subtracting $T(k_F)$ and a reference curve obtained without the tip ($V_t = 0$). This is nothing else than the SGM response in the wavenumber domain.

We begin by describing the simple model of the SGM process based on considering interference paths covered by the back-scattered electrons and the patterns they give rise to. This kind of analysis has been used to explain many features found in the experimental SGM results – see Ch. 1. It will prove useful in this section, where we investigate the tip influence in the energy space. On the other hand, it is of little help when the SGM maps in space are considered.

The interference paths model

The essential approximation here is that the electron waves move along well-defined paths, on which they keep the wavenumber k_F (instead of obeying the stationary transverse modes – i.e. negligence of the quasi-1D character of the system). A few most important examples of the paths have been shown in the fig. 5.8. Each path is described by:

- its amplitude, which depends on its length l_{path}, order, tip-backscattering amplitude and the probability of getting re-transmitted by the QPC, and
- its phase accumulation, given by $k_F l_{path}$ and the tip scattering phase shift,

and gives rise to a family of constant phase lines. The interference features, like fringing or chequerboard pattern, will lie on the crossings of these lines, for which the phase difference meets the usual interference condition. They will be weighted by the respective amplitudes.

This can produce a very complicated pattern, with a multitude of features packed closely in the area of a simulated SGM map. Furthermore, the features created by different pairs of paths will overlap, often concealing each other, so the observed simple patterns of the simulated images might be extremely difficult to explain correctly by finding the appropriate interference paths. We note that the truly two-dimensional systems are much less affected by this issue, since their lack of lateral boundaries seriously limits the number of the possible interference paths.

The analysis is made somewhat easier and more robust by passing to the k-space, where interference features due to the paths of different lengths get decoupled, even if they lie in the same position in the real space. The difference in lengths for a pair of interfering paths can be estimated from the position of the k-space peaks thanks to the reciprocity relation $k^{(peak)} = 2\pi/\Delta l_{path}$.

Tip scattering amplitudes

The tip scattering amplitudes and phase shifts are determined by the tip potential value V_0 and spatial extent Δr_0, already set as 1, (due to the local character, the exact shape of the tip potential should not be of great importance), scaled against the electron local kinetic energy and wavelength.

For a mode that has just been opened and has a very low kinetic energy, the scattering amplitude will be large, hence the SGM correction will be the greatest close to the new mode opening. Once the Fermi wavenumber k_F

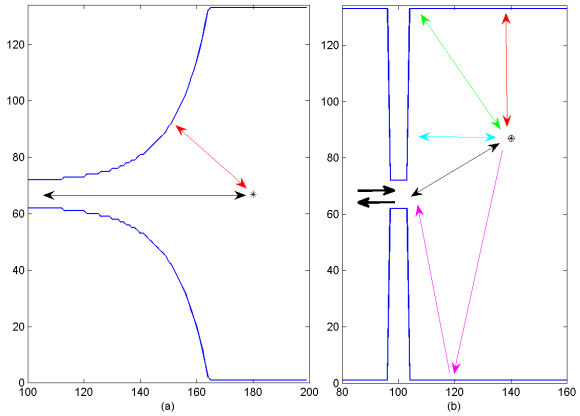


Figure 5.8: A few possible electron interference paths for:
(a) long MC structure;
(b) short SC structure.
The featured paths are of zero order (thick black), first order (thin black, magenta) and second order (red+black, cyan+black, green+black) in the tip potential.

grows, raising also the k_m value, the influence of the tip will diminish. It is best visible in the case of an ideal wire, but works quite generally – see below.

In a constricted geometry, different propagating states will have different phase shifts, which can give additional distortions to the interference patterns predicted by the simple path-model described above.

5.3.1 Wire

As a preliminary insight we present the results for the simple wire. Because of the perfect quantisation of transmission in this case, we can view these results as an idealisation of the plateau behaviour, stripped off from the influence of tunnelling and F-P resonances.

Fig. 5.9 shows the transmission curves $T(k_F)$ for various tip voltages. The upper fields of the figure show the difference signal $\Delta T(k_F)$. The discretisation board is the usual size, $M = 133$ by $L = 199$ points so the $(i_0, j_0) = (67, 132)$ tip position means that it is placed just on the longitudinal axis of the wire. The column index is of no relevance here, because of the translational symmetry of this structure[4]. The transversally centred position makes the tip couple well to the odd modes, lowering the odd plateaus transmission, while exerting no effect on the even ones. The influence is slightly smaller for each subsequent odd mode (i.e. decays more quickly). The situation is different when the tip is placed aside, at $(i_0, j_0) = (87, 132)$. There, the second plateau is also affected, but the effect on the third one is small.

Considering the k-dependence of the transmission, we see that the mode just being opened, with its longitudinal wavenumber close to zero, is com-

[4]Only the numerical noise will depend on it.

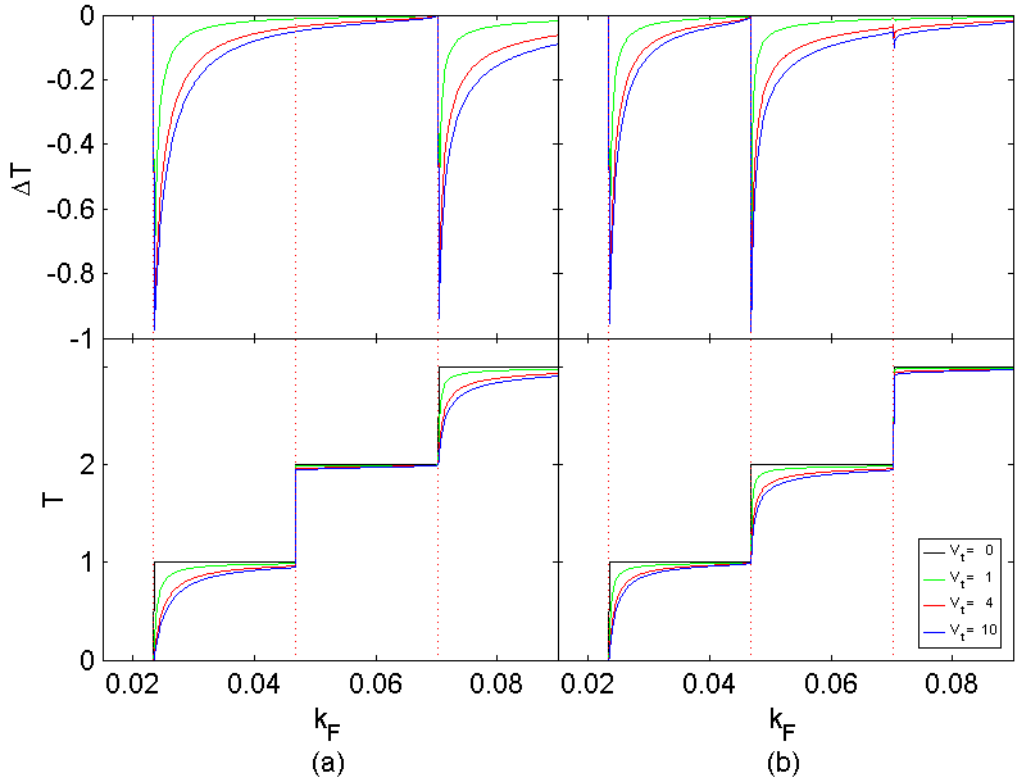


Figure 5.9: The transmission correction ΔT (upper fields) and the transmission T (lower fields) through an ideal wire as a function of the Fermi wavelength k_F, for a few different values of the SGM tip potential $V_t = 0, 1, 4, 10$. The red dotted lines show the channel thresholds $k_a^{(\perp)}$, $a = 1, 2, 3$. Two tip positions are featured:

(a) tip placed in the wire's axis, at the lattice site (67,132).

(b) tip placed off the axis, at the lattice site (87,132).

pletely stopped ($\Delta T = -1$) by the tip's presence – regardless of the tip potential. Then, along the plateaus the transmission curve exhibits monotonous growth, i.e. the tip's influence regularly falls as the wavenumber increases, just like we argued above. No interference peaks or dips are seen.

The dependence of the curves on the tip potential suggests that the SGM response saturates with growing V_T. Putting a very large potential value at the lattice site occupied by the tip is effectively equivalent to excluding this site from the lattice, so the saturation must indeed take place. The curves of the transmission as a function of the tip voltage for two different k-points plotted in fig. 5.10 show that the saturation takes place at rather large tip

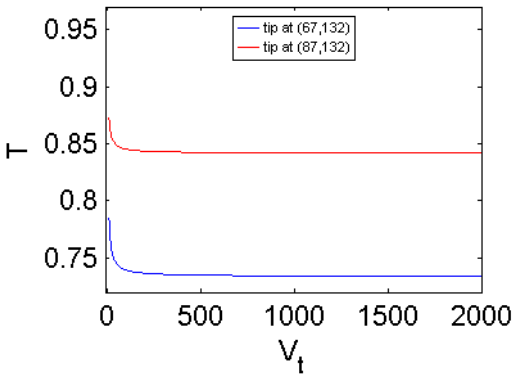


Figure 5.10: Total transmission through an ideal wire as a function of the tip potential V_t, at $k = 0.03$, for two different tip placements: symmetric, $(i_0, j_0) = (67, 132)$; and non-symmetric, $(i_0, j_0) = (87, 132)$.

potential values and we will not use such in our SGM scans. (A few control-scans have been run with $V_{\mathrm{T}} = 1000$ on the QPC-containing structures and they did not show any essential differences with respect to the ones performed with the standard tip strength, $V_{\mathrm{T}} = 10$.)

The saturation tip-strength values will certainly be smaller if we place the tip in the narrow channel of a QPC, but should be of comparable magnitude when we place it in the wide region outside the channel, as it is done in the SGM experiments.

5.3.2 QPCs, tip inside the constriction

We might expect a similar influence on the transmission curves, when we have a QPC structure with a rectangular (i.e. SC) constriction with the tip placed inside the narrow channel.

In the fig. 5.11 we plot the transmission curves and the transmission correction curves for such a case. The featured structure is a short SC Quantum Point Contact ($W_{\mathrm{QPC}} = 12$, $L_{\mathrm{QPC}} = 30$). The influence of the below-threshold tunnelling leads to a non-vanishing signal in the step region. The signal grows in magnitude to create a negative peak in ΔT around the beginning of the plateau, while on the plateau it behaves in a similar way as we saw for the wire. The plateau position for the modes affected by the tip's presence is shifted towards greater wavenumbers, since larger electron energies are needed to break though the constriction with an additional obstacle. It should be mentioned, that the results for higher wavenumbers ($k_F \gtrsim 0.5$) are burdened with uncertainty due to the computational model.

The situation becomes slightly more complex when we put the tip inside the constricted region of a long structure, as we have it in the fig. 5.12. The below-threshold signal is once again suppressed, like it was in the wire case, but the Fabry-Perot structure on the plateau introduces oscillatory features

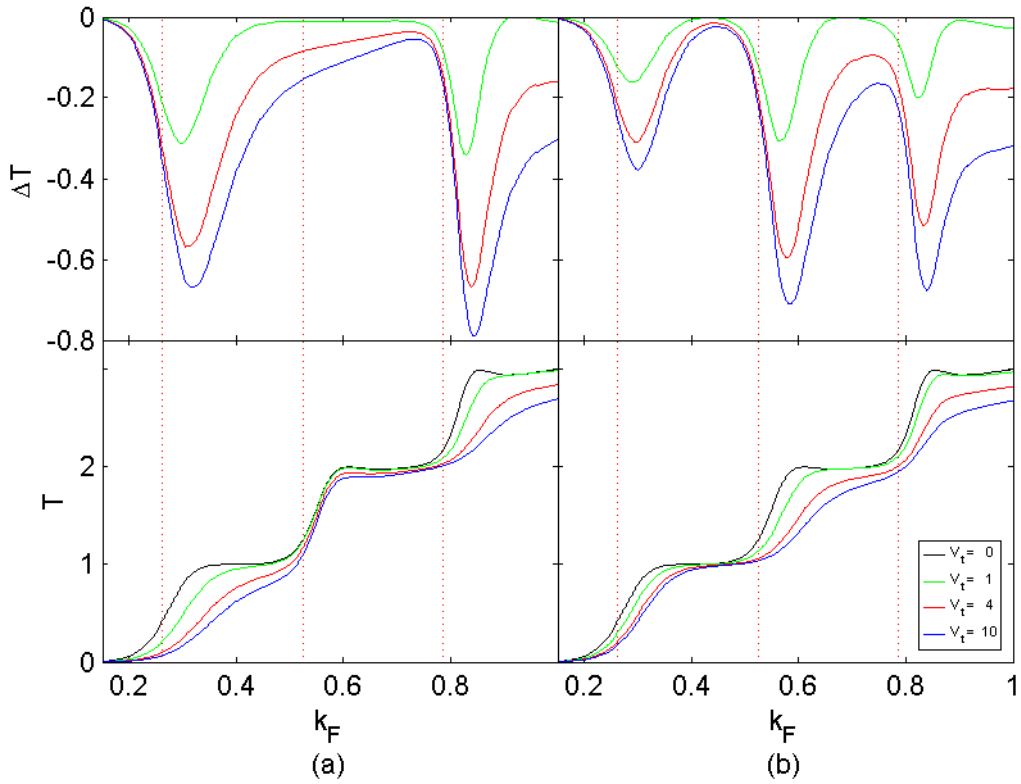


Figure 5.11: The transmission correction ΔT (upper fields) and the transmission T (lower fields) through a short SC-QPC as a function of the Fermi wavelength k_F, for a few different values of the SGM tip potential $V_t = 0, 1, 4, 10$. The red dotted lines show the channel thresholds $k_a^{(\perp)}$, $a = 1, 2, 3$. Two tip positions are featured:
(a) tip placed symmetrically in the middle of the channel, $(i_0, j_0) = (67, 100)$.
(b) tip placed non-symmetrically, $(i_0, j_0) = (70, 100)$.

to the difference curves. These are only partially due to shifting the $T(k_F)$ signal to the right. The resonating cavity is now split in half by the tip (part (a) of the figure), so the spacing of the F-P resonances in terms of k_a gets doubled, and this additional oscillation is seen in the difference curve (upper part). In part (b) the tip is placed in about five sixths of the channel length, which brings in two additional oscillation patterns, with the periods in k_a-space equal to $\frac{6}{5}\Delta k_a$ and $6\Delta k_a$. The second pattern has its first peak far to the right and is therefore barely visible. The oscillations do not bring the transmission above the plateau level and the "envelope" variation is still similar to the idealised case of the wire.

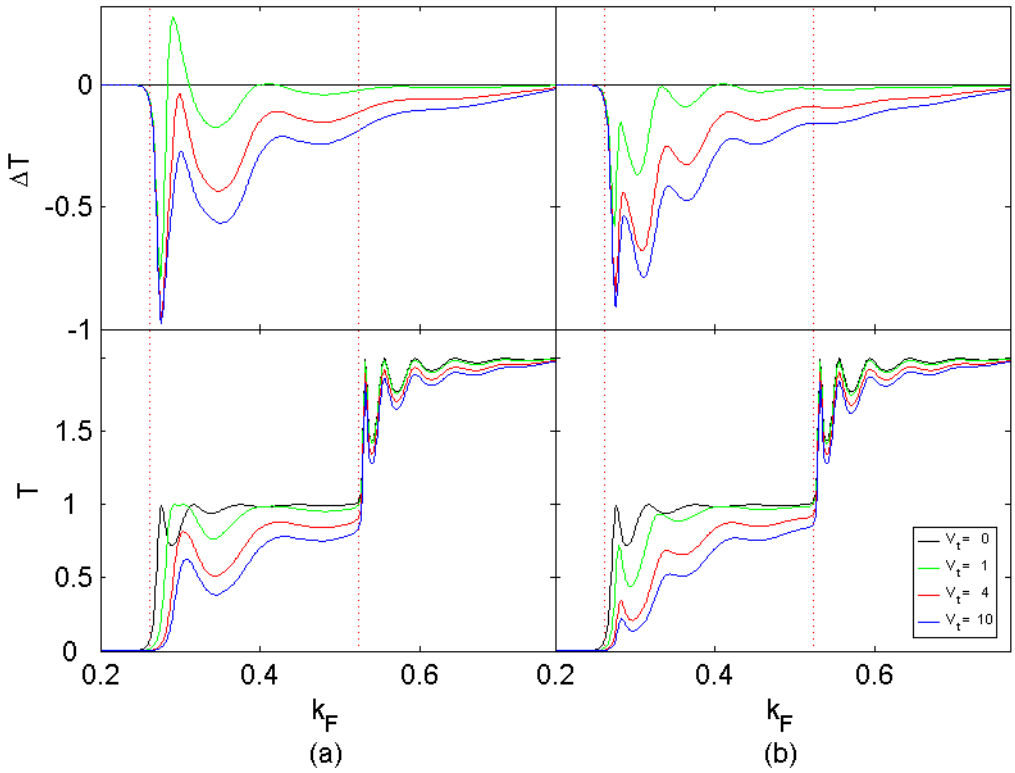


Figure 5.12: The transmission correction ΔT (upper fields) and the transmission T (lower fields) through a long SC-QPC as a function of the Fermi wavelength k_F, for a few different values of the SGM tip potential $V_t = 0, 1, 4, 10$. The red dotted lines show the channel thresholds $k_a^{(\perp)}$, $a = 1, 2$. Two tip positions are featured:
(a) tip placed symmetrically in the middle of the channel, $(i_0, j_0) = (67, 100)$.
(b) tip placed symmetrically, $(i_0, j_0) = (67, 110)$.

5.3.3 QPCs, tip in the wide region

Now we proceed to placing the tip in the wide region, like it is done in the experiments. The k-space structure is much richer now, because many interference paths for the electron waves can be created when the tip-reflection is involved.

HC geometry

We first present data for a long structure of the HC geometry. The difference curves $\Delta T(k_F)$ are shown in fig. 5.13, with the constriction width and

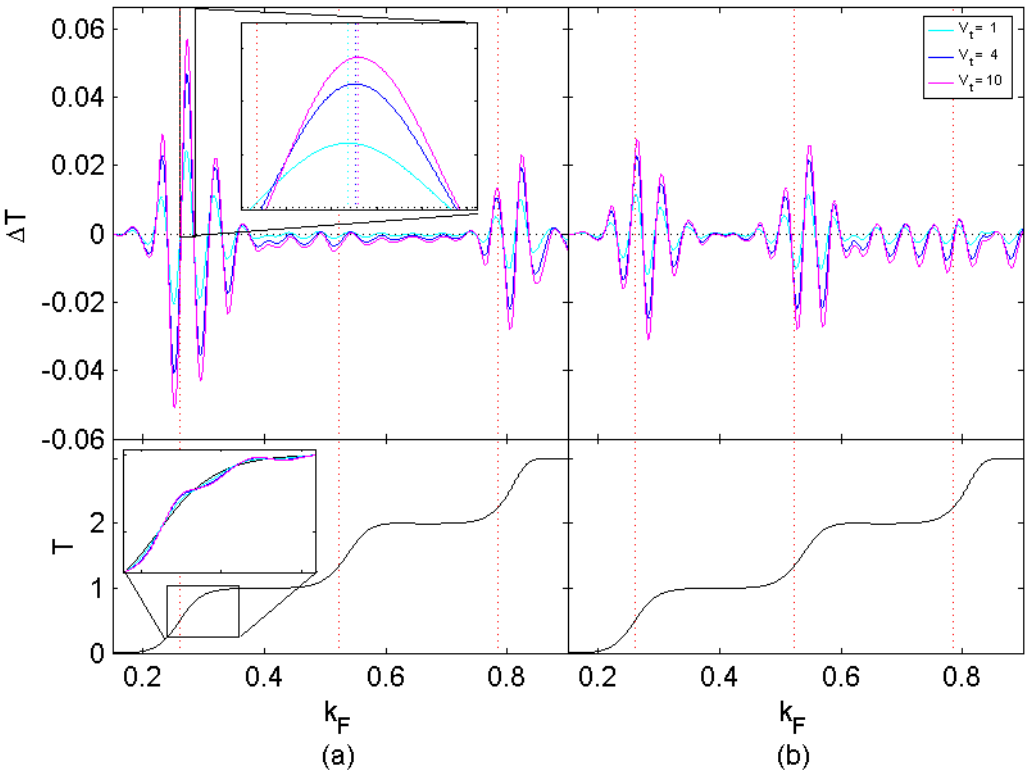


Figure 5.13: The correction ΔT to the total transmission through a long HC-QPC as a function of the Fermi wavelength k_F, for a few different values of the SGM tip potential $V_t = 1, 4, 10$. The lower plots show the transmission curves with no tip, for reference. Red dotted lines show the theoretical channel thresholds $k_a^{(\perp)}$, $a = 1, 2, 3$. Two tip positions are featured:
(a) tip placed symmetrically – $(i_0, j_0) = (67, 210)$.
(b) tip placed non-symmetrically – $(i_0, j_0) = (87, 210)$.
The inset in the upper plot shows the ΔT peaks with a stretched k_F-scale – we see that the peaks slightly change their positions when V_t changes. The inset in the lower plot contains a part of the $T(k_F)$ curves (not displayed in the lower pictures themselves) and is intended to give an overview of the scale of the tip influence.

length parameters of $W_{\mathrm{QPC}} = 12$ and $\eta_{\mathrm{QPC}} = 120$. The lower plots give the transmission curves, $T(k_F)$, for easier orientation.

The average magnitude of the ΔT signal follows a similar variation to what we saw earlier, with the tip placed inside the short constriction (cf. fig. 5.11 of the previous section). Naturally, now the scale of the correction is rela-

tively much smaller – see the inset in the lower field of the figure 5.13. On top of this envelope-variation, we have now the oscillatory pattern. This testifies that an important contribution to the SGM signal is made by the interferences between the electron waves travelling along different paths.

The period is not constant over the k_F-space, but the oscillation is otherwise regular, which shows that the interference paths different than the base and QPC-tip-QPC are either of small importance or of very similar effective lengths as the latter one.

It is only in the step region that the tip-induced oscillation can enhance the transmission to values above the reference, i.e. produce resonant tunnelling. On the plateau, as long as the tunnelling from the higher modes can be neglected, the tip correction does not bring the transmission above the plateau value, and consequently stays negative. Here, with the QPC reflection close to zero, the former base path disappears, and the oscillation decays with growing wavenumber.

The peaks of $\Delta T(k_F)$ change their position towards larger wavenumbers when the tip potential value is increased. The shift is very small and generally grows with growing k_F. At the origin of this lies the phase shift acquired by the electron wave when scattered off the tip, which generally should be small, but will scale with the tip strength and the propagation wavenumber. For example, the s-wave phase shift in a non-constricted geometry is $\delta_0 = k_F \Delta r_0 \left[\tanh(\sqrt{V_0})/V_0 - 1\right]$, (for a uniform disc-like scatterer, after [65], §132, Problem 2).

MC and SC geometries, multiple interference paths

The transmission correction curves become slightly more complex when we deal with MC and SC geometries. Example curves for those are plotted in fig. 5.14. The higher order interference paths will now introduce additional oscillation patterns.

In the MC structure, the strongest of such contributions will be given by the path drawn in red in the fig. 5.8 (a) above. Its effective length is comparable to the one of QPC-tip-QPC itinerary, but smaller. This results in amplitude-modulating of the basic oscillation of ΔT, seen in the (a)-part of the figure 5.14.

A different effect can be extracted from part (b) of the fig. 5.14. Here, narrow and sharp dips cut into the $\Delta T(k_F)$ curve's structure with a shorter repeat period than the periods of its other features. The narrow shape of these (anti-)resonances suggests a long path for the electron waves and indeed the period indicates about 60- to 70-sites' distance from an obstacle. This must engage a reflection from the lateral boundary of the discretisation grid.

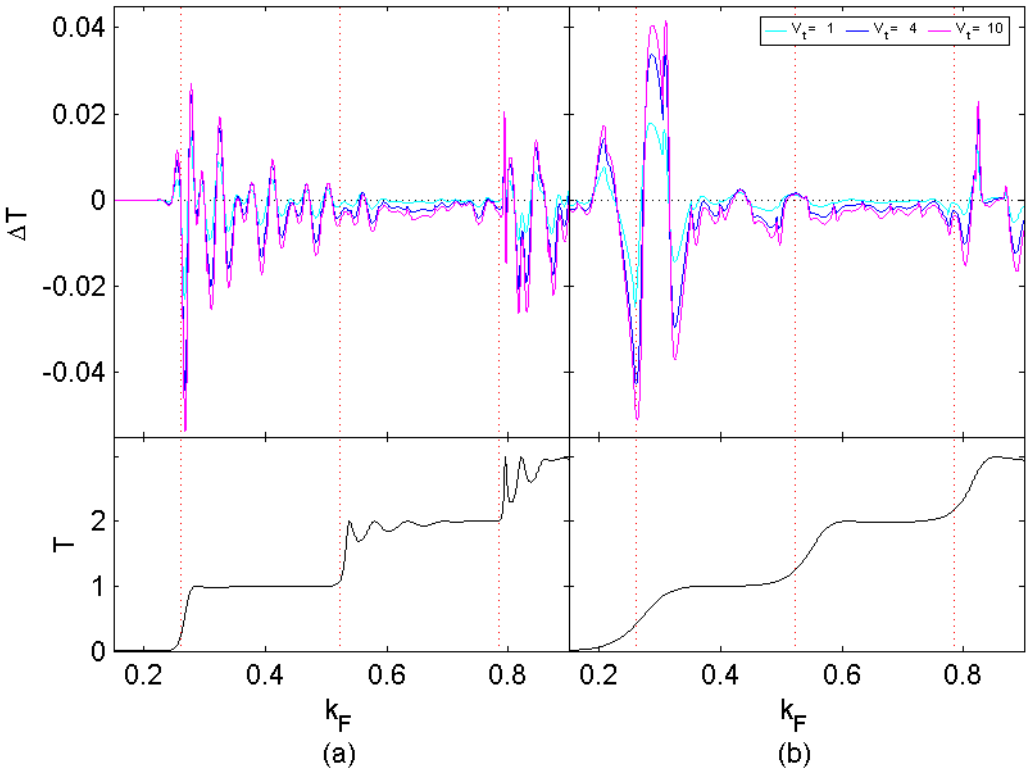


Figure 5.14: The correction ΔT to the total transmission as a function of the Fermi wavelength k_F for a few different values of the SGM tip potential $V_t = 1, 4, 10$. The lower plots show the transmission curves with no tip, for reference. Dotted red lines show the theoretical channel thresholds $k_a^{(\perp)}, a = 1, 2, 3$. Two structures are featured:

(a) long MC structure with tip placed symmetrically, $(i_0, j_0) = (67, 180)$.

(b) short SC structure with tip placed symmetrically, $(i_0, j_0) = (67, 132)$.

Two manifestations of additional interference paths are seen: modulation in (a), sharp dips in (b).

In fact, such boundary-related resonances appear in all structures (except the HC, in which case we just do not place the tip in regions allowing the lattice boundary reflections to take place), but seeing them often requires enlarging the scale.

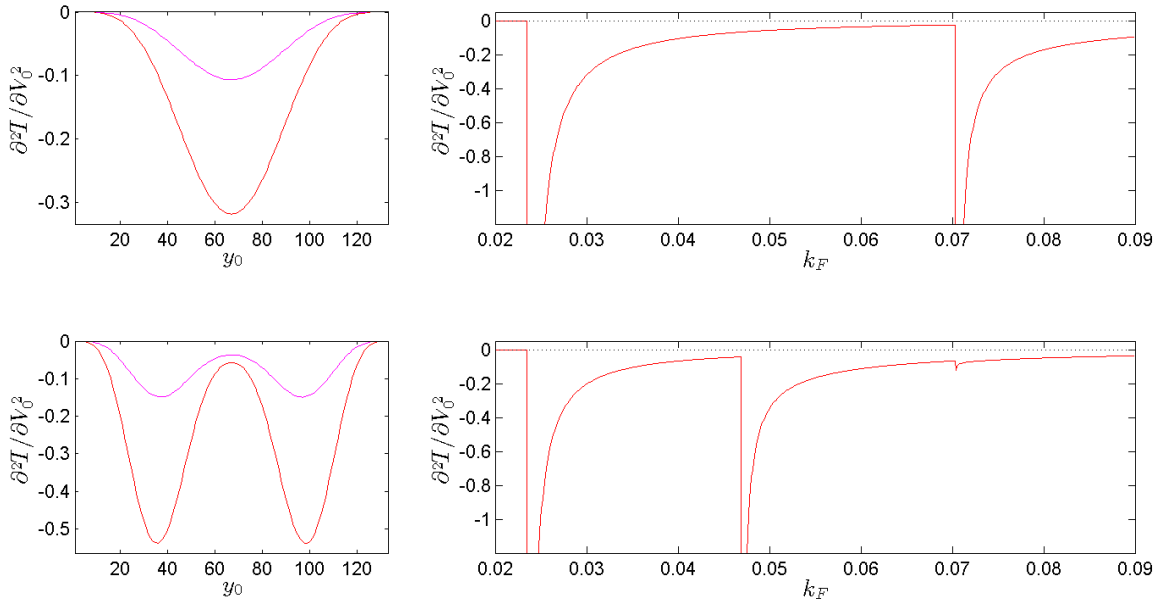


Figure 5.15: Second order differential signal in a wire.
Left panel: transverse variation of $\partial^2 T/\partial V_0^2$ for a few example k-points,
(a) in the first plateau region, $k = 0.03$ (red) and $k = 0.04$ (magenta),
(c) in the second plateau region, $k = 0.05$ (red) and $k = 0.06$ (magenta).
Right panel: k-space variation of $\partial^2 T/\partial V_0^2$ two example tip positions,
(b) tip placed at $(i_0, j_0) = (67, 132)$,
(d) tip placed at $(i_0, j_0) = (67, 132)$.

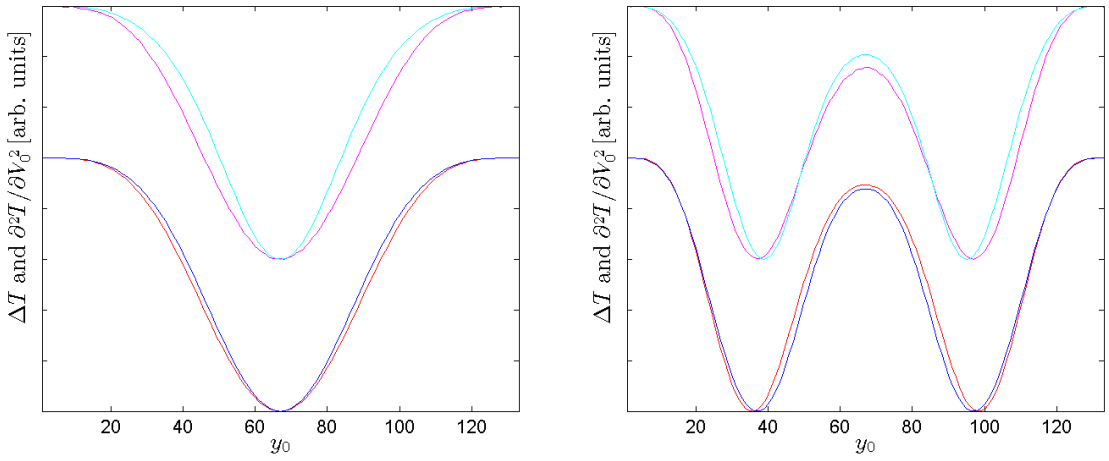
5.4 Wire signal as a basic check

The RGF-simulated differential signal in the wire strictly obeys the variation predicted for this simple geometry by our formulae ((4.16), Ch. 4), written in Anderson units as:

$$\partial T/\partial V_0 = 0 \ , \qquad \partial^2 T/\partial V_0^2 = -2 \left[\sum_{a=1}^{N} \frac{\sin^2 \left(k_a^{(\perp)} y \right)}{k_a (M+1)} \right]^2 \tag{5.17}$$

The magnitude and the variation agree in both, the position space and the k-space. Example simulated curves are plotted in fig. 5.15.

This is an additional confirmation that the analytically obtained formulae and the computation performed by the computer indeed pertain to *exactly* the same quantities.



(a) (b)

Figure 5.16: The transverse variation of the full tip strength SGM signal ΔT and the differential signal $\partial^2 T/\partial V_0^2$ in a straight wire, for a few different k-points. All curves are normalised to unit height. The curves are pairwise offset to compare the corresponding ones:
(a) $\partial^2 T/\partial V_0^2$ (magenta) with ΔT (cyan) at $k = 0.03$, and $\partial^2 T/\partial V_0^2$ (red) with ΔT (blue) at $k = 0.04$;
(b) $\partial^2 T/\partial V_0^2$ (magenta) with ΔT (cyan) at $k = 0.05$, and $\partial^2 T/\partial V_0^2$ (red) with ΔT (blue) at $k = 0.06$.

Numerical noise

The vanishing of the first order derivative can also serve as a method of assessing the levels of the numerical noise in the simulations, which, thus estimated, does not rise above the level of $10^{-8} - 10^{-9}$ amplitude and 10^{-9} root-mean-square variance. With $\epsilon = 10^{-5}$, this means that the noise level in the finite difference signal $T|_{V_0=\epsilon} - T|_{V_0=0}$ is certainly below 10^{-13}, and in the second order signal does not exceed 10^{-3}, which gives us at least one order of magnitude buffer in all the images that we present further.

The noise estimated from subtracting symmetric signals (not shown) is even slightly smaller.

Comparison with the finite difference signal

The variation of the full tip strength SGM response in a wire differs slightly from the differential response. Its features are generally sharper, so resolving it – or, better, its square root – in the basis of the squared transverse wave

functions would require adding some small contribution from the higher order terms. We plot the finite signal and the differential signal, normalised to unit height, in fig. 5.16.

5.5 SGM maps

We will now present the simulated maps of the first and second order differential signal in the SGM setup and compare them against the finite-difference signal ΔT obtained with the tip strength value $V_0 = 10$.

Checks performed on the $T(V_0)$-curves (not shown) confirm that with the potential step ϵ that we chose, the linear- and quadratic-differential regions of operation are indeed achieved. Also the potential value chosen for the finite difference signal is, as we estimate with the help of eq. (3.51), Ch.3, sufficient to probe the system in the invasive way.

Since the simulated maps have the expected symmetries, we plot only their halves.

5.5.1 Step-region and plateau-region signals

For an orderly presentation and systematic discussion of the results, we begin with the simplest case among the ones that we have investigated. In figure 5.17 we present the maps of the differential SGM signals for a moderately elongated ($W_{\mathrm{QPC}} = 12$, $\eta_{\mathrm{QPC}} = 40$) structure of the MC type. The maps were taken with the Fermi wavenumber in the step region ($k_F = 0.2668$) and in the plateau region ($k_F = 0.2881$) of the quantised transmission.

Differential signal

The elongated MC geometry optimises quite well the trade-off between the spurious effect of the tunnelling and the excess of the Fabry-Perot oscillation in the transmission curve, which is essential for having a well-pronounced difference between the step and the plateau signal. We immediately see this difference when we compare part (a) and part (b) of the figure, both showing the first order differential map. On passing from step – (a), to the plateau – (b), the signal keeps roughly the same shape, but diminishes substantially in magnitude – here, by three orders, but with fine tuning of the k-space position to the F-P peak, it can be brought down to the level of numerical noise. This confirms the formerly envisaged need for the second order correction, which we plot in part (c). (The second-order step-region signal is not featured, as,

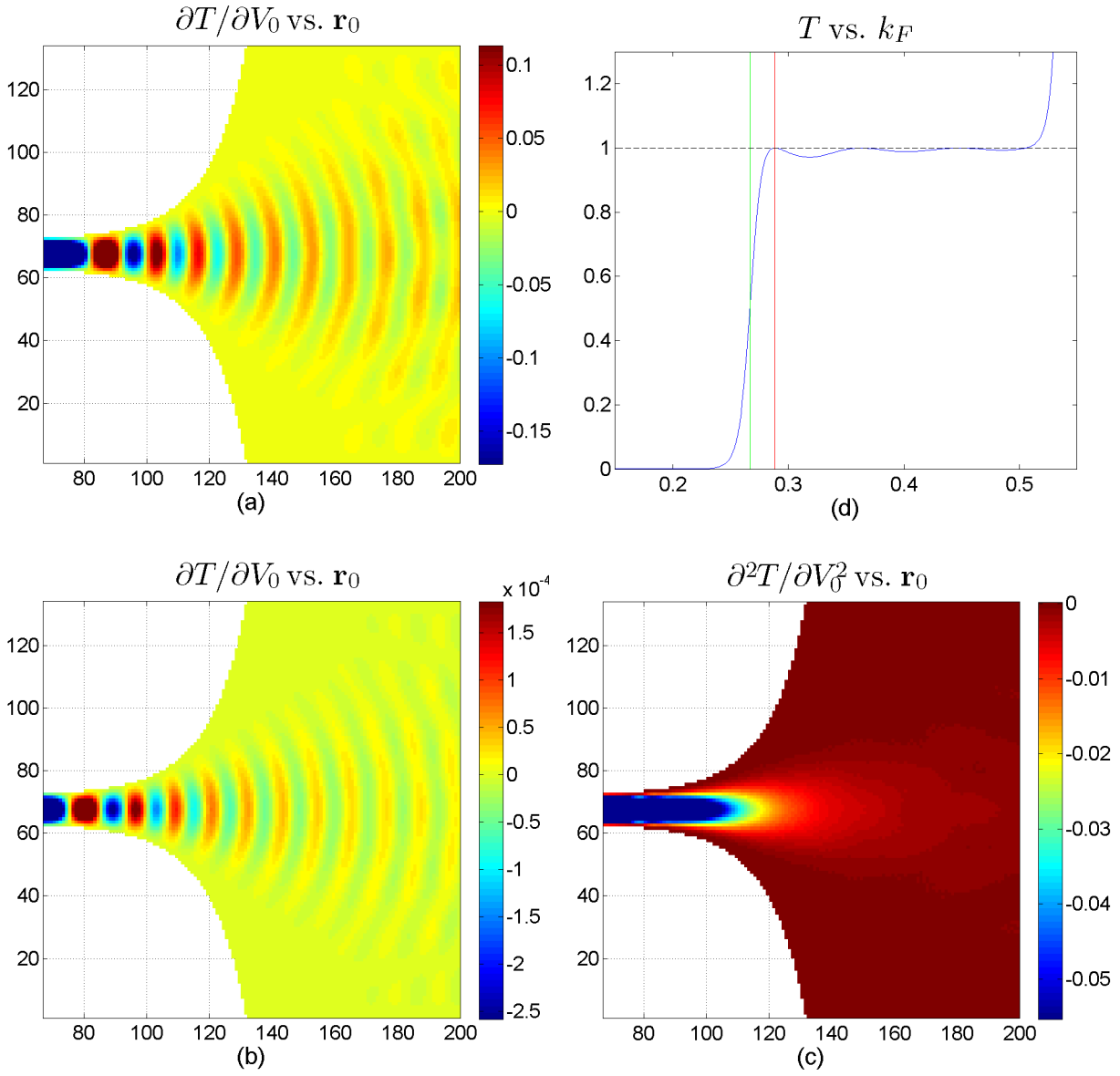


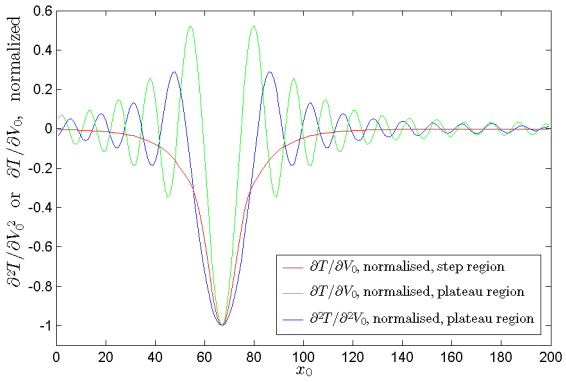
Figure 5.17: Spatial maps of the first- and second-order derivative of the total transmission with respect to the tip strength, $\partial T/\partial V_0$ and $\partial^2 T/\partial^2 V_0$, representing 1st- and 2nd-order components of the non-invasive tip SGM signal:

(a) 1st-order signal in the transmission step region,

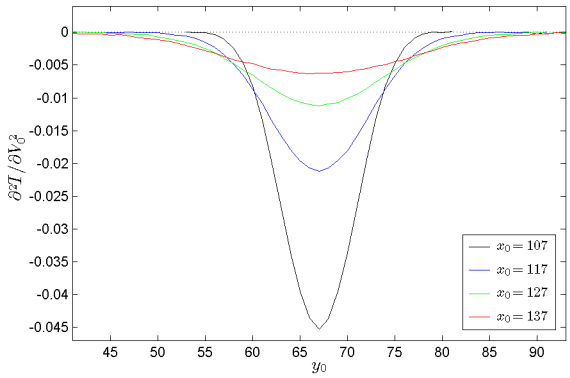
(b) 1st-order signal in the transmission plateau region,

(c) 2nd-order signal in the transmission plateau region.

The featured structure is a moderately long MC-structure, with $W_{\mathrm{QPC}} = 12$ and $\eta_{\mathrm{QPC}} = 40$. Part (d) contains the corresponding transmission curve $T(k_F)$, with vertical lines marking the k-space points for which the step (green, $k_F = 0.2668$) and plateau (red, $k_F = 0.2881$) maps have been taken.



(a) The longitudinal profiles along the structure's axis ($y_0 = W/2$) of the first-order SGM signal ($\partial T/\partial V_0$) in the transmission step region and the first- and second-order ($\partial^2 T/\partial V_0^2$) signals in the transmission plateau region. All are normalized to have a minimum value of -1.



(b) The transverse profiles of the second-order signal ($\partial^2 T/\partial V_0^2$) on the transmission plateau ($k_F = 0.2881$), taken at a few different longitudinal positions – the last one is already outside the constricted region, where the broadening (and decaying, cf. (a)) tendency continues.

Figure 5.18

with non-vanishing first order correction, it contributes only negligibly to the mild-tip SGM signal.)

Also the shapes of the signals fit into our general predictions from Ch. 3. The first order signal oscillates around zero, and we see that it is dominated by one simple oscillation, whose periodicity matches, up to the resolution restrictions,[5] the half of the local Fermi wavelength. In addition, the signal contains relatively little angular structure. In each of the figures we can see a regular "cone" of signal, which spreads as we move away from the QPC, until it hits the lateral borders. Both these facts are consequences of the relatively high adiabaticity of the long MC-constriction.

If we consider the map's profile along the longitudinal axis of the system, plotted in fig. 5.18(a) (green and blue curves), we find that the decay of the oscillation, being at the same time the decay of the signal itself, is between x_0^{-1} and x_0^{-2}.

The second order transmission correction has a completely different character. It is always negative and reveals no oscillation (it can oscillate only inside the constriction – which is caused by the formation of Fabry-Perot

[5] In the broad region the resolution makes it impossible to tell, from the longitudinal periodicity, which of the lowest modes ($a = 1, 2, 3$) is dominating, but it can be determined from the signal's transverse variation.

modes). In particular, its transverse variation, see fig. 5.18(b), is well described by the fourth power of the function $\phi_1(y)$ (the fitting is not displayed in the figure), but with an effective width W_{eff}, considerably smaller than the lead's width W. This feature agrees with eq. (3.65) of Ch. 3, – and, since we are at the first conductance plateau, with one eigenmode open, the presence of only one transverse mode in the signal additionally confirms the adiabaticity of the transmission.

The agreement with the idealised eq. (3.65) is only partial, for the signal does not reach a state in which it would be given by an x-y separable function. Instead, it falls towards zero as x_0 is increased, while its effective width grows, causing transverse spreading of the signal. The decay of the second order signal is much faster than of the first order one cf. (5.18(b)). The spreading, on the other hand, progresses more slowly than in the first order case, which we can see from the fig. 5.17 (Part (b) vs. Part (c)).

Full tip-strength signal

In fig. 5.19, we have plotted the finite transmission correction $\Delta T(\mathbf{r}_0) = T|_{V_0} - T|_0$, obtained with the tip potential value $V_0 = 10$, taken as standard for all the maps presented below.

The finite signal has the same character as the corresponding first (for the step k-points) and second order derivatives (for the plateau), plotted above, but has more sharp features. This is a general tendency, which we will see in all the finite tip-strength images featured below. In particular, it has more transverse variation. Its plateau-region form also acquires some irregular longitudinal variation, but it does not have the character of the first-order signal fringes, so we cannot say that the finite tip-strength influence is to mix the plateau and step signals.

Increasing the tip potential value (to $V_0 = 1000$, not shown) does not introduce any new features to the SGM maps, only enhances slightly the ones already existent in the $V_0 = 10$ signal.

5.5.2 Non-perfect plateau-region signals

Now, we would like to pay attention to how the differential signals get modified, when the system is non-perfectly tuned on the transmission plateau.

Two factors that we will take into account are:

– Fabry-Perot resonant structure, and

– tunnelling.

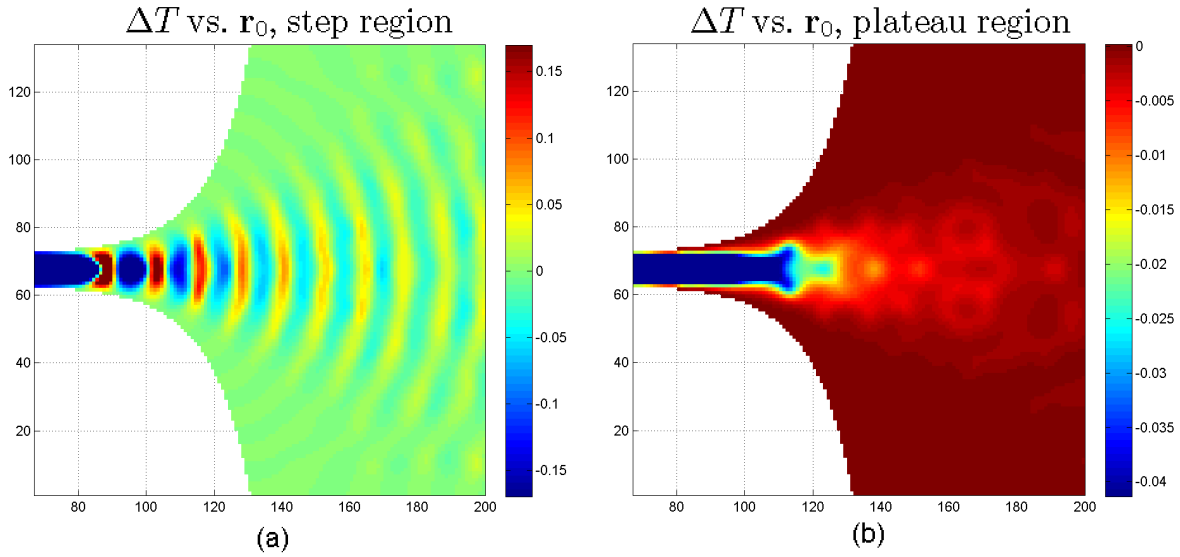
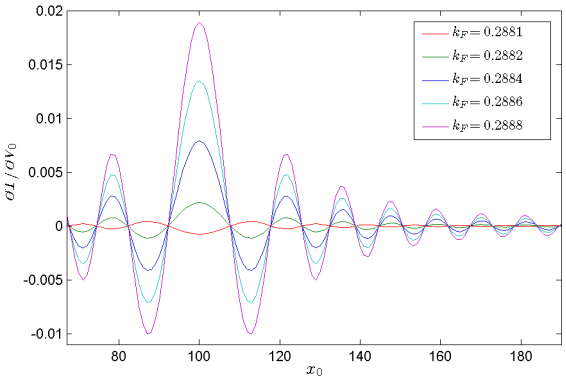


Figure 5.19: Spatial maps of the change of the total transmission ΔT which represents the invasive-tip SGM signal. The images were taken in:
(a) the step region, with $k_F = 0.2668$,
(b) the plateau region, with $k_F = 0.2881$,
and at the tip strength $V_0 = 10$. The featured structure, like before, is a moderately long MC-structure ($W_{\mathrm{QPC}} = 12$, $\eta_{\mathrm{QPC}} = 40$).

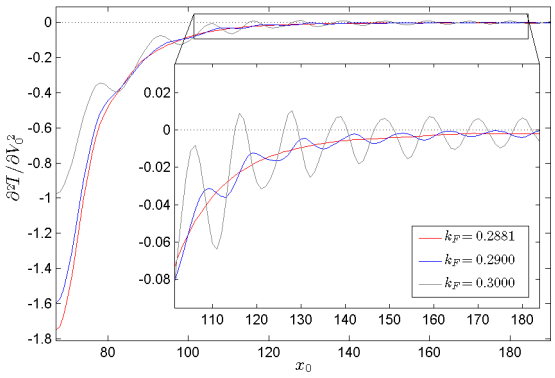
Fabry-Perot oscillation, long structures

The first of such situations takes place when the Fermi wavenumber belongs to the plateau region, but the transmission is lowered by the Fabry-Perot oscillation. The plateau-results in the previous subsection were taken with the system tuned close to the F-P peak, which guaranteed that $\mathcal{T}_m \approx 1$.

Now, as we tune the Fermi wavenumber off the F-P peak towards the F-P valley, the overall magnitude of the first-order correction grows quite quickly, as we can see from fig. 5.20(a) showing the longitudinal profiles of the first order signal for a few k-points lying very close to the peak. The sign flip between first two curves is caused by the fact that the $k_F = 0.2881$ value, chosen as representative for the plateau, is slightly lower than the exact the F-P peak wavenumber – once the maximum is trespassed, $g^{(1)}$ goes through zero and changes the sign. The growth of the signal's magnitude outside the constricted region should scale like the prefactor $F(k_F) \equiv \mathcal{T}_m^2(k_F)\sqrt{1-\mathcal{T}_m^2(k_F)}$ of eq. (3.56), Ch. 3. (It is not possible to verify this growth based on our simulated data, for the signal's magnitude at a given tip position depends on the oscillatory factors $\overset{\approx}{\varphi}{}^{(\pm)\,2}_{l\varepsilon m}$ of eq. (3.56) too.) Apart from this, the sig-



(a) The longitudinal profiles along the structure's axis ($y_0 = W/2$) of the first-order differential SGM signal for a few k-points close to the Fabry-Perot peak, including the previously introduced $k_F = 0.2881$ representing the Fabry-Perot peak maximum, in the transmission plateau region.



(b) The longitudinal profiles along the structure's axis ($y_0 = W/2$) of the second-order differential SGM signal in the plateau region – at the Fabry-Perot peak ($k_F = 0.2881$) and two other k-points in the vicinity (see Part (d) of fig. 5.17(b)).

Figure 5.20

nal keeps its shape, only changing slightly the fringes spacing due to the k_F change.

The second order, on the other hand, gradually looses its importance and develops certain traits of what is typical for the first order – despite staying mostly negative, it acquires a mild fringing pattern, which can at times bring it above zero. We can see it changing in figure 5.20(b), where the longitudinal profiles have been plotted for the F-P peak point ($k_F = 0.2881$) along with two other points in the vicinity, and also in fig. 5.21(a) which presents an example map of such a non-perfect second-order differential signal. The fringing pattern is caused by the $t^\dagger r'$-related terms in eq. (3.42), Ch. 3,[6], which in this situation are no longer zero. With the $F(k_F)$ function growing very quickly around its zero value, the departure from the ideal second-order shape already occurs as close to the F-P maximum as at $k_F = 0.2900$.

The fringes have the same periodicity and decay as the ones from the first order signal, but are shifted in phase by approximately π. This is important, because it lets us distinguish the signature of the first-order and the second-order patterns in the SGM signal for finite tip strengths. A comparison of the figures 5.21(a) and 5.21(b) shows clearly that the small fringes emerging

[6] The $r'^\dagger r'$-related terms will grow much more slowly, so they should not play role in here.

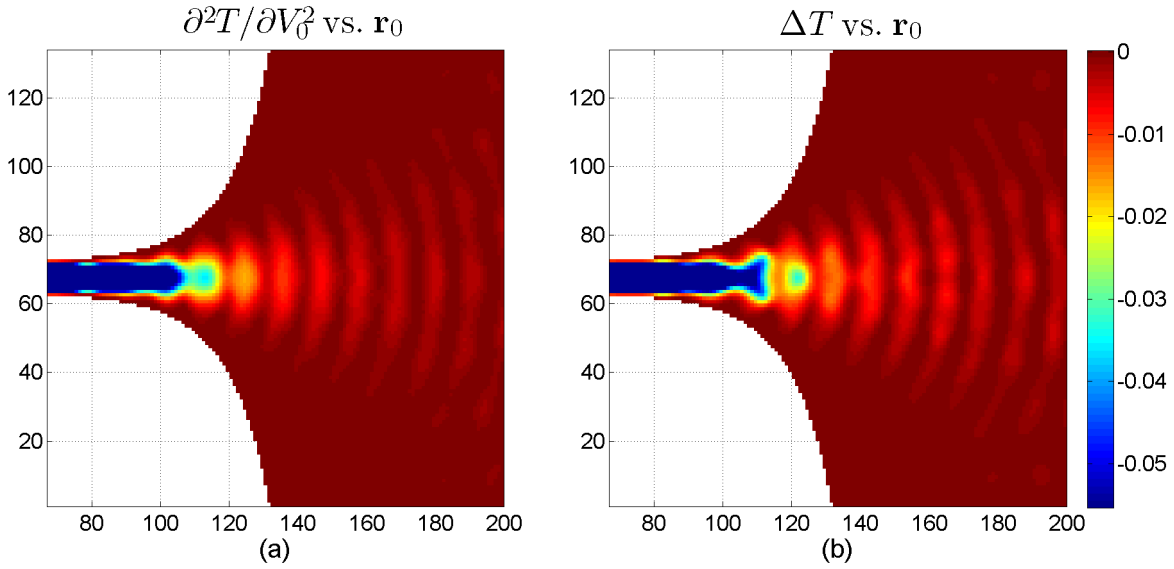


Figure 5.21: Maps of the:
(a) second-order differential SGM signal,
(b) finite tip strength SGM signal,
with the Fermi wavenumber tuned to $k_F = 0.2900$, which lies in the plateau region, slightly off the Fabry-Perot peak. The colour-axis is common for both images. The maps (a) and (b) have roughly the same character, but the fringes in (b) are shifted in phase with respect to (a) by approximately π, which is caused by the influence of the first-order differential signal.

in the finite-tip signal agree in phase with the first-order differential signal's fringing, and not the second-order one. This is not surprising, as the second-order fringes depend on the same factor F as the first-order signal, which is more important at the small tip strengths. And, indeed, no k-points have been found in which the finite-tip image would have clearly visible second-order type fringes.

In addition, we observe that the full-tip signal on the plateau, even in this situation, does not exceed zero. This is a natural consequence of the simple fact that the transmission of the first eigenmode is already very close to unity and cannot be risen more, while the second mode is not yet open and its tunnelling part is negligible.

Two important conclusions follow. Firstly, the longitudinal oscillation in the ordinary SGM signal in a structure of this geometry can only be caused by the first-order conductance correction. That is, the one typical for the step region. Secondly, because of the Fabry-Perot oscillations, throughout

a considerable part of the plateau the SGM signal will have a non-perfect character, mixed between the step-signal and the plateau-signal. Of course, its effect will be most visible in the beginning of the plateau, where the F-P oscillation has a large amplitude and varies sharply, and less important in plateau's end. Moreover, if the finite temperature effects are present, it might turn out that the ideal plateau signal cannot be observed at all – this will happen when the thermal energy $k_B\theta$ is comparable to the F-P peak width in energy.

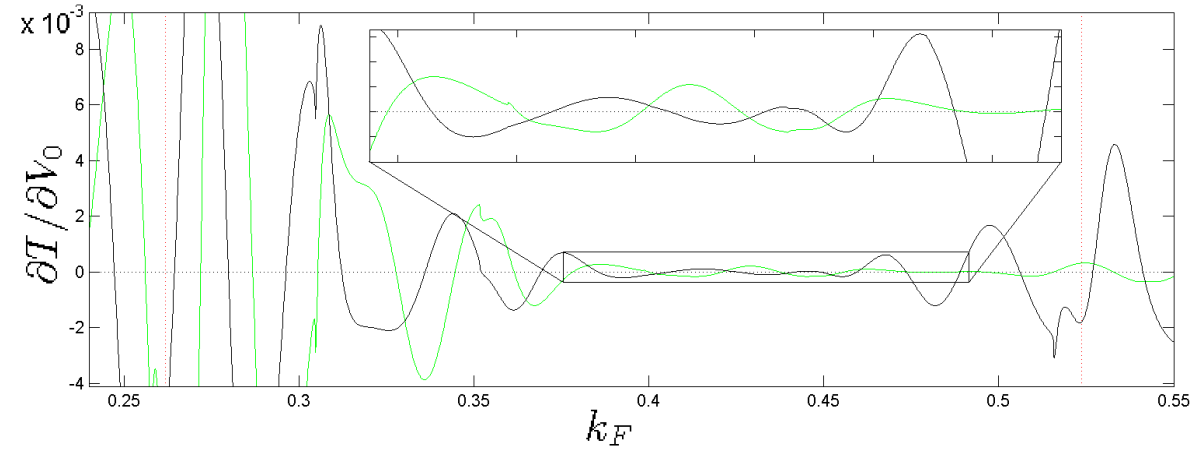
The situation is even more complicated on the higher plateaus, since for having a perfect plateau level, we need all the open modes to be either on their F-P peak, or far enough in k-space from their thresholds, so that the F-P oscillation has already died out (see subsec. 5.5.3)

Tunnelling, short structures

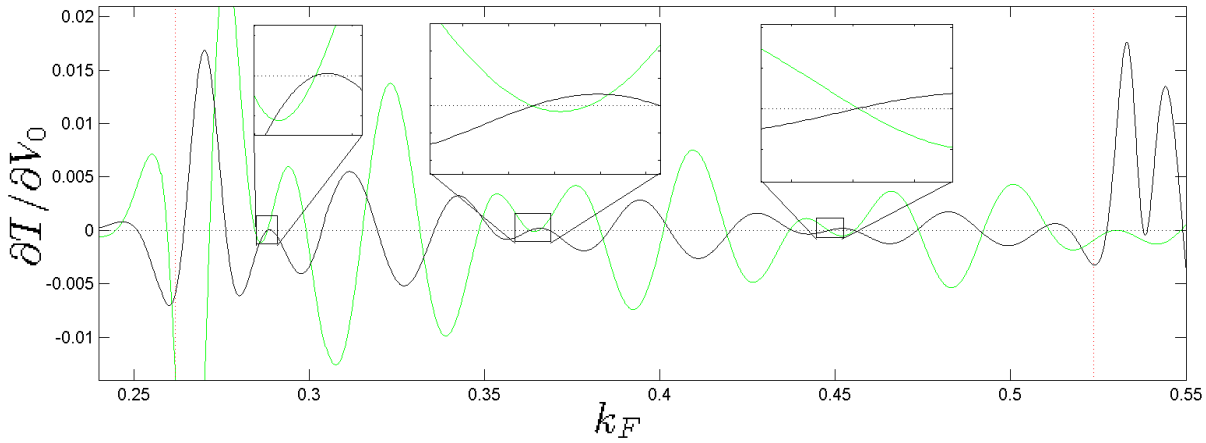
A natural question arises, whether one could avoid this imperfection by making the structure short and hence killing the Fabry-Perot oscillatory structure in the transmission curve. The price we have to pay for this is the entering, by tunnelling, of the higher QPC-eigenmode's signal into the given eigenmode's plateau region. In this situation, we expect a superposition of the previously described $m = 1$-related plateau signal, either perfect or non-perfect, and the signal arising from the tunnelling eigenmode, $m = 2$.

The $m = 1$-sate will keep loosing its spurious first-order component as k_F increases, while the latter gains in magnitude and can become significant in the right part of the plateau region. In the short structures, this complementarity prevents the SGM signal from achieving the idealised plateau-form – there is no intermediate region in which both contributions would be small at the same time. To see it directly, we plot the first order differential signal in the k-space domain for the short MC structure in fig. 5.22(a), for two different tip positions. The two curves have no common zeros, which means that it is not possible to have a k-point for which the first order signal would vanish in all space. We can contrast it with analogous curves for the previously featured elongated MC structure, plotted in the adjacent figure, 5.22(b). There, the common zeros appear in the positions of the Fabry-Perot peaks, which we saw in the fig. 5.17 (d) on p. 119.

The SGM images of the short MC quantum point contact (not plotted), throughout most of the plateau region, have a shape very similar to the shape of the elongated structures of the fig. 5.21. The only difference is that their angular structure is slightly richer, due to a lower adiabaticity of the short geometry.



(a)



(b)

Figure 5.22: First order differential SGM correction $\partial T/\partial V_0$ as a function of the Fermi wave number k_F for:

(a) the short MC structure, $W_{\mathrm{QPC}} = 12$, $\eta_{\mathrm{QPC}} = 12$,

(b) the moderately long MC structure, $W_{\mathrm{QPC}} = 12$, $\eta_{\mathrm{QPC}} = 40$,

each of them for two different tip positions – green curves for $(i_0, j_0) = (67, 180)$ and black curves for $(i_0, j_0) = (87, 180)$. The two dotted red lines mark the first and second constriction mode thresholds.

The tunnelling signal According to the $g^{(1)}$ and $g^{(2)}$ formulae (eqs. (3.23) and (3.42), Ch. 3), the tunnelling-related contribution to the SGM response is just a weak step-region signal, of the second constriction mode. However, due to the small propagation wave number and to the evanescent character

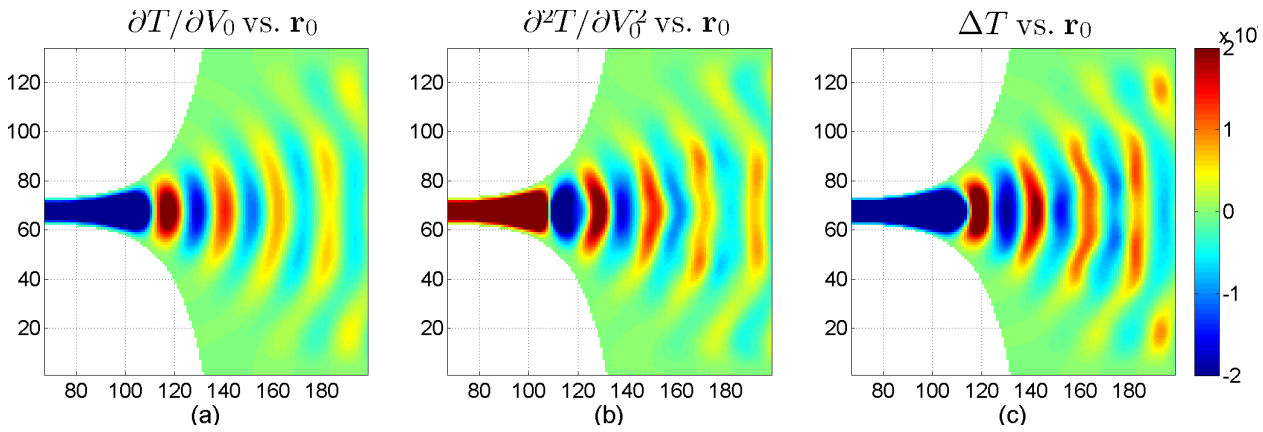


Figure 5.23: Maps of the:
(a) first-order differential SGM signal,
(b) second-order differential SGM signal,
(c) finite tip strength SGM signal,
with the Fermi wavenumber chosen at $k_F = 0.15$, below the first mode's threshold.

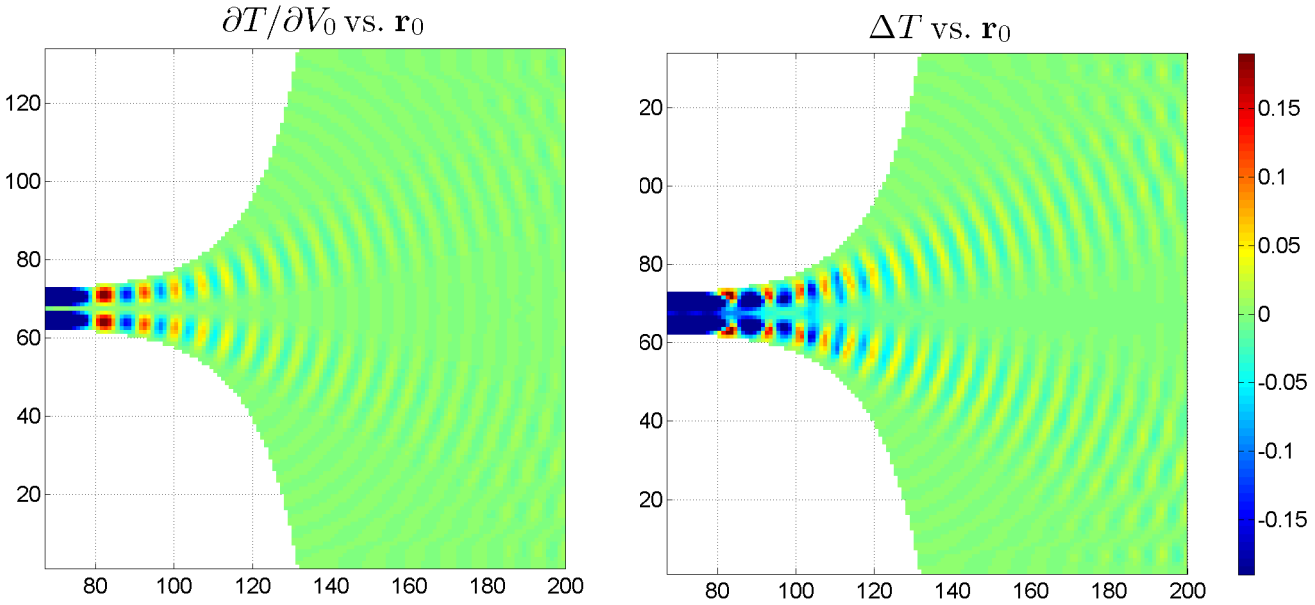
of the wave functions throughout a considerable part of the central region, it is worth comparing it to the ordinary step-region signal. We have no way of isolating it from the first plateau's signal, but we will extract the basic features from the images obtained below the first ramp, where the total signal is scarce and the first mode's tunnelling contribution is its only constituent. These are plotted in fig. 5.23.

Inside the constriction, the signal creates a non-oscillating "blob", Outside, it resumes the fringing, but the images show a more polygonal shape of the fringes pattern, which is directly related to the fact that structure's adiabaticity has a meaning only when compared to the propagation wavelength, so for the slowly propagating electron waves any structure is less adiabatic.[7]

5.5.3 Beyond the first plateau

In figure 5.24, we present the SGM response for the second step region. First order differential signal and full tip strength difference signal are plotted in parts (a) and (b), respectively. The general character of the maps does not depart from what we have described so far. The only difference is that now

[7] Another way of looking at this, is that the lower the number of active transverse modes, the more difficult it is to combine them into a smoothly rounded shape.



(a) (b)

Figure 5.24: Maps of the:
(a) first-order differential SGM signal,
(b) finite tip strength SGM signal,
with the Fermi wavenumber tuned to the second step region, $k_F = 0.531785$.

the second constriction mode decides of the images' angular variation, giving two wide lobes. The zero value along the axis testifies that there are indeed no contributions from the fully open first mode.

The situation becomes slightly more complex in the plateau region, where both constriction modes $m = 1, 2$ are important. Each of them, independently, can be burdened with the plateau-spoiling factors and it is impossible to tell from the total transmission curves which of the modes' contributions are nearing $\mathcal{T}_m = 1$ at the given k-point. In addition, even in the case of the Fabry-Perot structure, which for some geometries is more or less predictable – see the eq. (5.15) in subsec. 5.2.2, it is highly unlikely to have the second mode's F-P peak occurring at the same place as an F-P peak of the first mode's contribution.

As we move along the plateau, the signal will always have a non-perfect, mixed character, and will alternate between two shapes:

– when the $m = 2$-mode has a peak, the step-type component will have the $m = 1$ mode's shape, with one lobe – which is shown in the left panel of fig. 5.25,

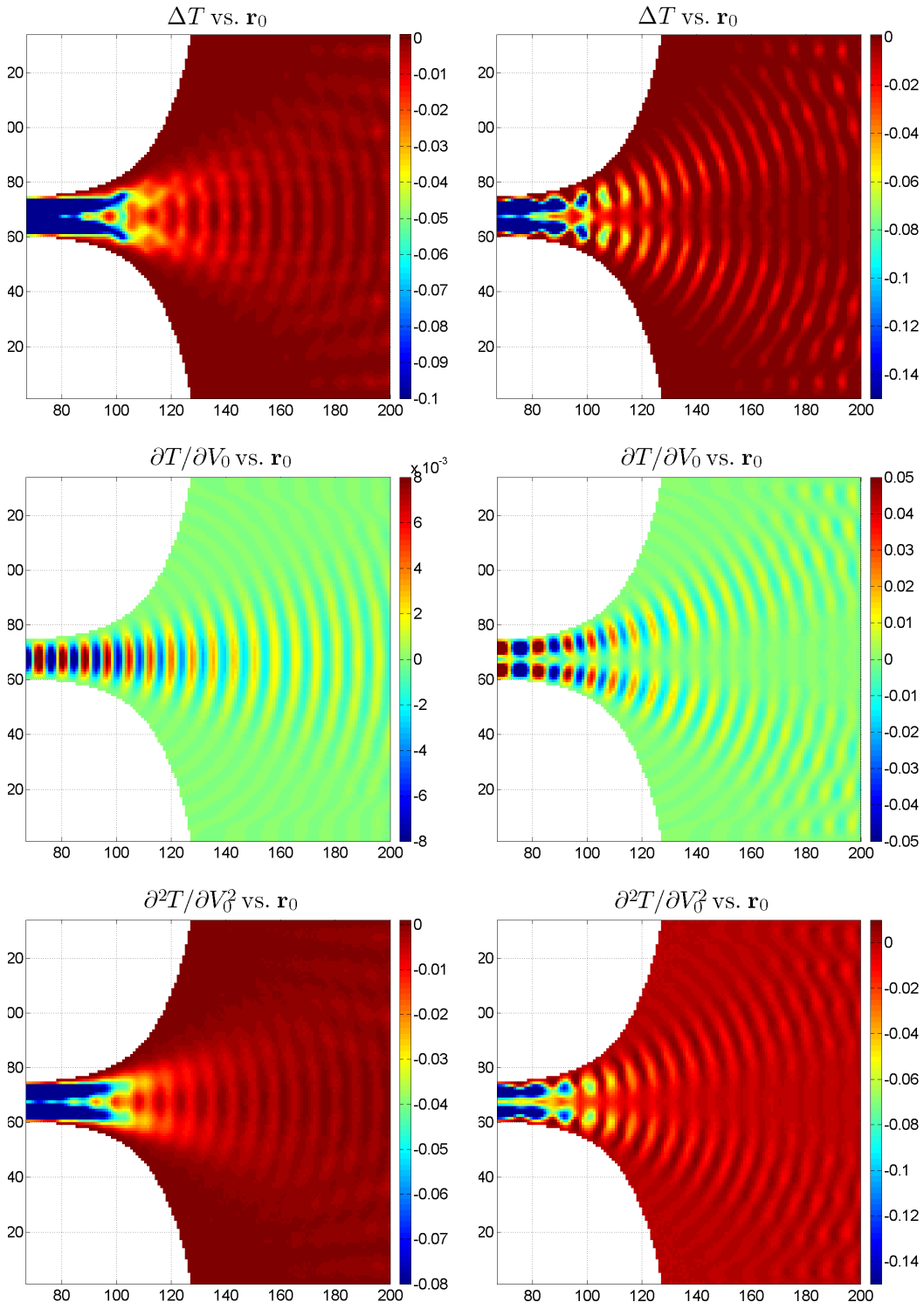


Figure 5.25: Maps of the finite tip strength (upper row of figures), first order differential (middle row) and second order differential (lower row) SGM signal on the second plateau of the quantised conductance. The left column maps are for $k_F = 0.4177$, where the signal imperfectness is caused by $m = 1$-mode influence. The right column maps are for $k_F = 0.44$, where the imperfectness due to $m = 2$-mode influence prevails.

– when the $m = 1$-mode has a peak, the step-type (fringed) component of signal will have the shape typical for the $m = 2$-mode, i.e. with two lobes – shown in the right panel of fig. 5.25.

In the intermediate regions, where both of the step-type signals are present, the $m = 2$ one will prevail, because it is weighted by a larger $\mathcal{R}_m\mathcal{T}_m^2$-prefactor. Also the normalisation factors in the wave functions will work to marginalise the influence of the lower mode. Therefore, the signal featured in the right panel of the figure will be more common. Differences in fringes spacing between these two will also occur, but only inside the constricted region.

5.5.4 Other geometries, non-adiabaticity

We have seen in sec. 5.2 that the lack of adiabaticity does not have to affect the transmission curves, if it follows only from the shape of the constriction and not from its length. This fits well into the theoretical results of [36, 69]. On the other hand, in subsec. 3.4.11 of Ch. 3 we found that in the structures constricted by the straight lateral borders – which we simulate here – the fact of having regular oscillations in the SGM signal can be explained only by the adiabaticity of the structure.

In order to see what influence the degree of adiabaticity can have on the SGM maps, we will plot the simulated SGM images for the two remaining geometries, HC and SC. Their adiabaticity is determined primarily by the curvature in the narrowest region and by their length, see Ch. 4. We therefore assess that the HC geometry has a considerably high level of adiabaticity, but somewhat lower than the MC geometry considered so far. The SC structure, on the other hand, is abrupt and thus has to be highly mode-mixing (see 4.1.3,Ch. 4 or the discussion in 3.4.11,Ch. 3).

The HC- and SC-geometry results are shown in fig. 5.26 for the step signal and fig. 5.27 for the plateau signal. As to investigate the clean step- and plateau-behaviours, we chose moderately elongated structures and, when on a plateau, tune the wave number closely to the unit-transmission points.

The most basic observation is that the mode-mixing SC geometry has much more transverse structure than the HC and MC ones. It quickly activates the higher order lead-modes and at a distance of a few Fermi wave lengths can already loose the memory of its original shape at the exit from the QPC. The mode-mixing introduces also a well visible chequerboard pattern and a mild fringing into the plateau-signal (fig.5.27 (a)), which for the other geometries remains plain. The chequerboard pattern in the adiabatic structures can be obtained by using an invasive, finite-potential probe.

The finite tip strength in the SC structure does not bring in new sharp

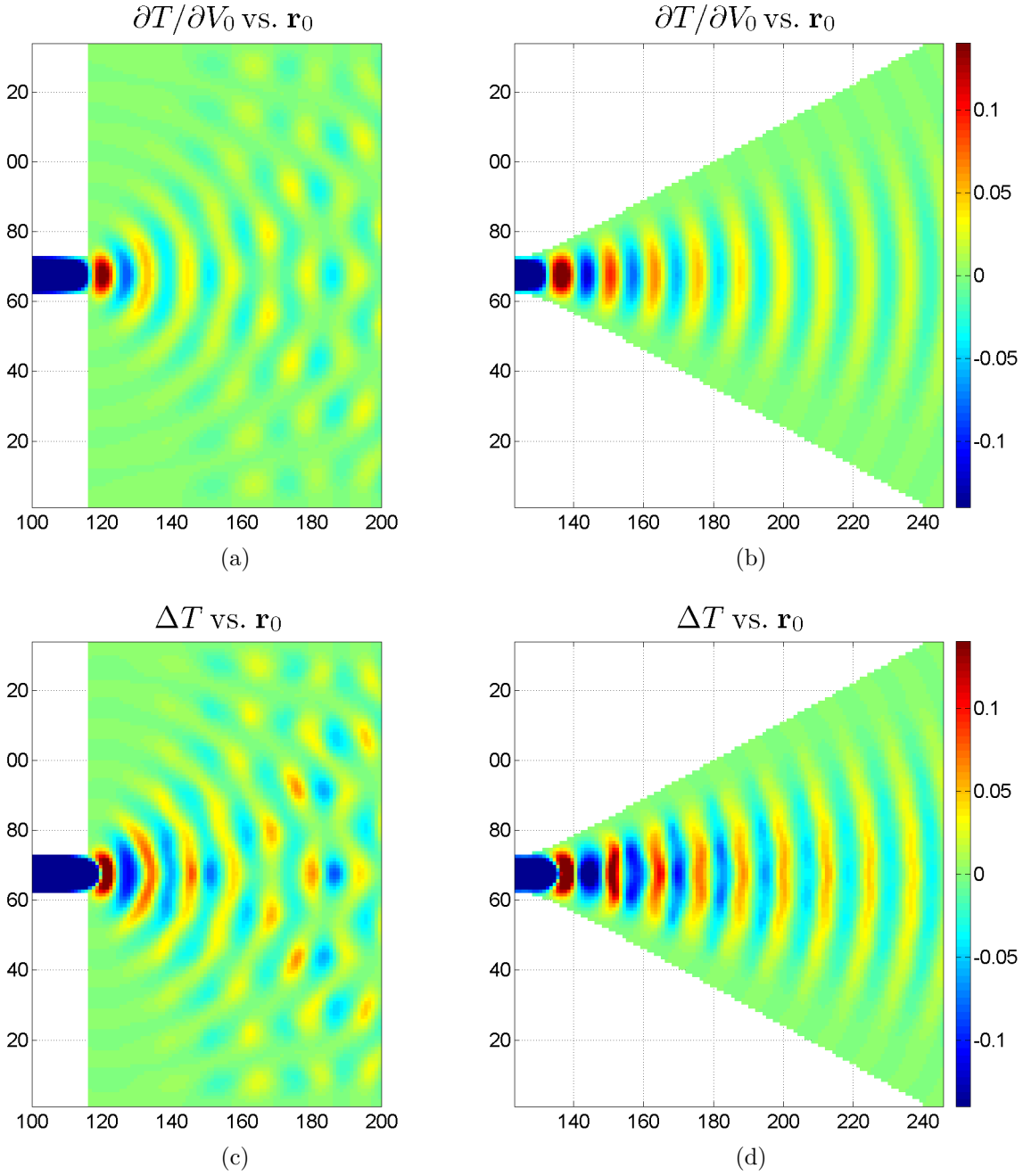


Figure 5.26: Maps of the first order differential SGM signal (a,b) and finite tip strength SGM signal (c,d) on the conductance step, for: (a,c) SC structure with $W_{\mathrm{QPC}} = 12$, $L_{\mathrm{QPC}} = 32$, at $k_F = 0.263$; (b,d) HC structure with $W_{\mathrm{QPC}} = 12$, $\eta_{\mathrm{QPC}} = 120$, at $k_F = 0.271$.

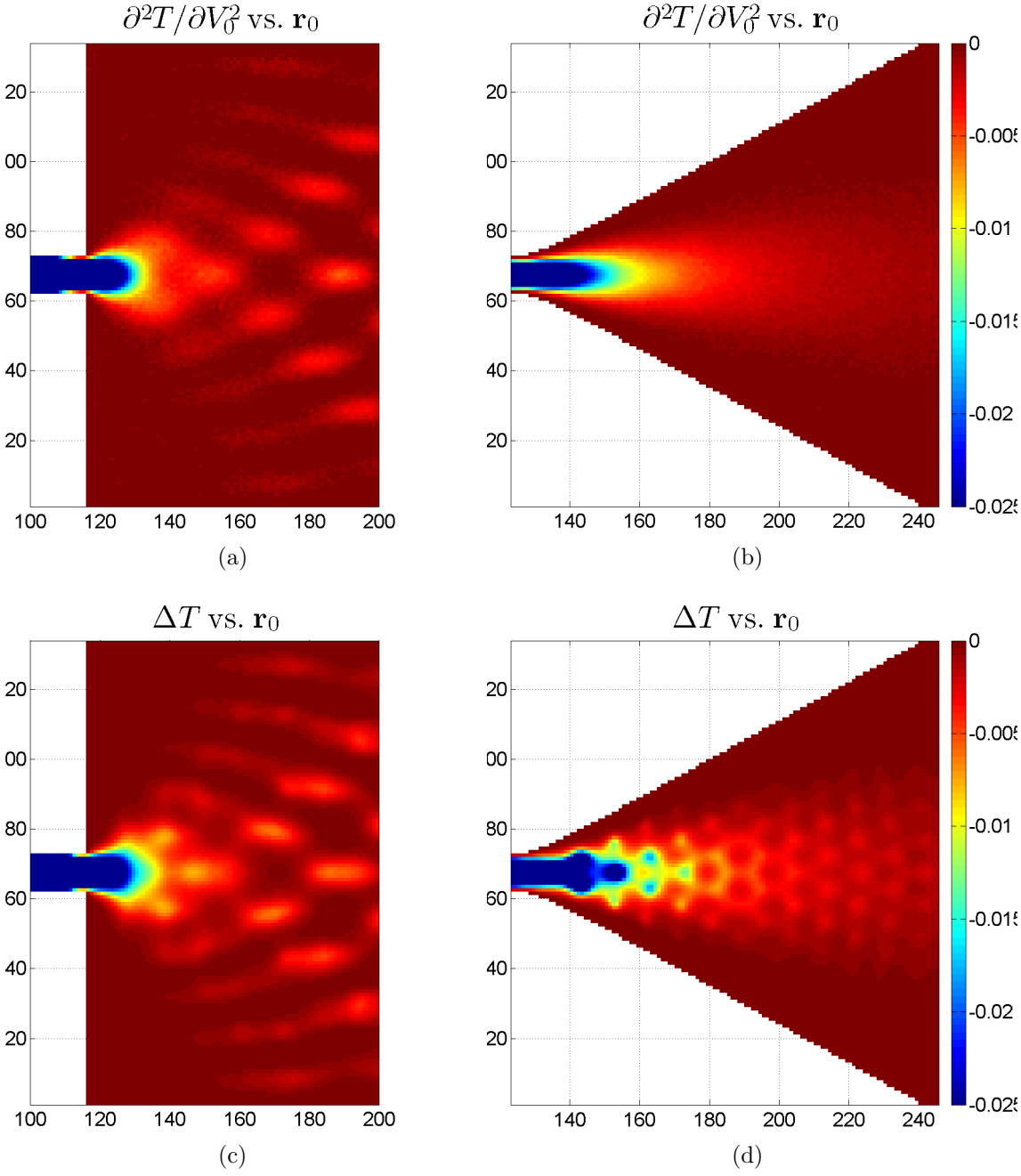


Figure 5.27: Maps of the second order differential SGM signal (a,b) and finite tip strength SGM signal (c,d) on the conductance plateau, for: (a,c) SC structure with $W_{\mathrm{QPC}} = 12$, $L_{\mathrm{QPC}} = 32$, at $k_F = 0.3773723$; (b,d) HC structure with $W_{\mathrm{QPC}} = 12$, $\eta_{\mathrm{QPC}} = 120$, at $k_F = 0.276363859$.

features, but only increases the intensity of the ones already present. On the other hand, the finite tip strength applied to the adiabatic structures, as we see in figs. 5.27 and 5.26 and also have seen in the fig. 5.19, works as though it was diminishing the adiabaticity. This is related to the fact that both these things tend to activate more lead modes and should manifest itself in the composition of the u_l-matrices of (2.40). This correspondence of finite tip-strength and non-adiabaticity suggests, perhaps counter-intuitively, that the non-invasive SGM signal should be easier to obtain in the non-adiabatic structures (in other words, they are more robust against the invasive action of the tip).

Furthermore, these results point out that beyond certain distances, there would be no fringing at all in the signal exiting from the non-adiabatic constrictions. The situation can, however, be very different if the lateral borders are not present. Then, the environment into which the electron waves are injected from the QPC (i.e. the free space) does not exert any influence upon them, so they can retain the shape determined at their exit from the opening. There, the information about the adiabaticity of the QPC constriction might be contained in the angular structure of the fringed signal.

In Ch. 1, we have cited some of the most important experimental SGM-results. The wide-lobed structure with the number of lobes corresponding to the constriction mode index is reproduced by our maps. We predict, however, that there could be points at the plateau, for which the lower number of the angular lobes are seen. This is probably limited by the temperature effects, which would mix such signals, relatively weak in amplitude, with the ones with the usual structure.

The fringes can be created by mechanisms of different type. The fringes we have in the simulated figures would not be able persist beyond the thermal length, like, for example, the ones in fig. 1.4 (a), as we do not include any disorder. On the other hand, their vanishing in cleaner samples does not follow from having a perfect plateau-signal, because lowering the temperature retrieves the fringes in both, step and plateau regions. The step-image fringes of fig. 1.5 (a) can be thought of as corresponding to the same situation as the simulated map in fig. 5.19, which, in turn, does not differ significantly from the non-invasive signal. Apart from fringing, all three contain a similar mild angular variation. The plateau map, fig. 1.5 (b), has the non-perfect plateau character, because it mixes the chequerboard pattern, occurring in the finite tip-strength plateau signal, and the fringing from the step-signal – and therefore should be compared with fig. 5.21 (b).

The effect of the finite temperature on the fringes will generally lead to blurring them, but, as pointed out by Abbout *et al.*[49], below certain temperature, the thermal averaging will cause mixing into the non-fringed

(or weakly fringed) plateau-signal, the fringed signal from the step – thus enhancing its fringes instead of suppressing them. The Fabry-Perot valleys, will give the same effect of fringes enhancement of fringes for the well-tuned plateau k-points, only of smaller extent.

We cannot try to reproduce the branching, because of the lack of disorder in our simulations.

Chapter 6

Closing remarks

6.1 Overview

In this work, we have presented the theory of non-invasive Scanning Gate Microscopy, its application for investigating the transport in quantum point contacts and numerical simulations of this process.

Working within the Landauer approach to the quantum transport, we employed the perturbation theory to describe the influence of the SGM probe on the conductance, thus obtaining differential conductance corrections, expressed in terms of the unperturbed system properties. A detailed derivation has been included. We have given estimates of the applicability of the perturbative formulae, providing in this way a definition of the non-invasive measurement. We addressed the question of the interpretation of the SGM signal, with particular attention to the conditions at which it can be related to the current density or the local density of states.

In Chapter 4, we have briefly reviewed the most important models of the quantum point contact transport and applied our conductance correction formulae to obtain SGM signal maps, as well as the images of current and charge density, within two of these models, essentially different from each other.

Another class of models has been chosen for the computer simulations of Chapter 5, performed with the use of Recursive Green Function algorithm. The character of ideal non-invasive SGM signal is shown and discussed, along with the mechanisms and conditions which contribute to impairing it. It is also compared against the signal obtained with invasive, strong tip potential.

Our theoretical results neglect the electron-electron interactions and external fields, and so do the numeric models. The effects of temperature and disorder, on the other hand, have not been treated in our calculations, but

can be included as a straightforward generalisation of our theory.

6.2 Conclusions

The central results of our theory are the formulae for the perturbation corrections:

- of the first order in the tip potential, $g^{(1)}$ – relevant for the description in the quantised conductance step region, or when the conductance curve is not quantised,

- of the second order in the tip potential, $g^{(2)}$ – relevant in the conductance plateau region.

They give the meaning to the SGM measurement as probing the tip potential matrix elements $\left\langle \approx\!\!\!\!\!\!\sim\!\!\!\!\!\!\sim \Psi_{2\varepsilon.} \middle| V_{\mathrm{T}} \middle| \approx\!\!\!\!\!\!\sim\!\!\!\!\!\!\sim \Psi_{1\varepsilon.} \right\rangle$, and through them the scattering eigenstate wave functions. The first order correction has the interesting feature of isolating the influence of a single pair of counter-propagating transmission eigenstates which are just being opened, while the second order corrections mixes the influence from all the open channels.

The fact of having the eigenstates incoming from the opposite leads prevents general simple interpretations. Only under special conditions can the non-invasive SGM response be related to the investigated system's current density or to the density of the mobile charge (ie. the local density of the open states) at the Fermi-energy.

When the delta function is taken as the tip potential, the $g^{(2)}$-formula can be factored to give a product of the Fermi-energy local densities of states, one for the left-incoming states, which are fully occupied, and the other one for the right-incoming states, which are empty.

If the system is time-reversal invariant and is tuned to the first plateau of the quantised conductance, then the SGM signal maps the square of the Fermi-energy charge density of the conducting state. If, on top of the time-reversal symmetry, a four-fold spatial symmetry is added and the leads can be taken as infinitely wide, then the SGM signal on any plateau can be related to the square of the Fermi-energy charge density and also to the current density, as these two are now proportional. Straying from the perfect plateau value will impair this plausible relations, and, as we have seen in the Chapters 4 and 5, even within the idealised models this is difficult to avoid. Presence of finite temperature and disorder potential will make it even harder.

In the general case, the non-invasive Scanning Gate Microscopy does not image the current density. Although it has certain common features, its dependence on the individual mode propagation velocities is different, which gives it different relative weights of the eigenstates contributions. The analytical expressions and the simulation results show that the SGM signal resembles the local density of states to a greater extent. In the special, symmetric, case mentioned above, the current density and SGM signal are related, but only outside the QPC or any other laterally constricted structure. Since, the simulated strong-tip SGM signal differs from the simulated non-invasive results by features of relatively small importance, we believe that also the invasive SGM technique does not image the current density, unless the tip is large enough to be treated as a purely classical object.

Non-invasiveness

In the discussion of the results in Ch. 3, we have defined the non-invasiveness by requiring that the tip potential is weak and/or tightly localised, $V_0 \Delta r_0 \ll E_F^{(1\mathrm{D})} \lambda_F^{(1\mathrm{D})}$ – which is a sufficient criterion.

It was satisfied in all the simulation results which we presented. On the other hand, in most of the experimental work on quantum point contacts published so far, this requirement is not met. As reported in these works, the large voltages on the tip created a divot of depletion in the free electron gas under the tip, whose size is comparable to the diameter (radius of curvature) of the tip ending [82] and definitely larger than the Fermi wavelength (otherwise, e.g., the so-called "glint effect" [39] would not be possible). Therefore, these experiments will probably be better described by some semiclassical approach [83] or even, perhaps, a completely classical one (see [6, 41]).

Nevertheless, it is not impossible to employ milder or smaller SGM probes, like it has been done in [22, 31] or in [27, 28] – where the dependence of the SGM signal on the probe voltage was tracked down to very small voltages. In fact, the latter two works also found a close relation between the SGM response and the local density of states. In addition, experiments on quantum dots were reported [25], in which the tip potential is enough localised as to only "dent" the wave functions inside the investigated system. For such probes, our theory should be appropriate. Moreover, in the quest for genuinely quantum behaviour, developing non-invasive probe techniques is certainly desired.

Shape and features of the SGM images

In Chapter 5, we have identified the typical shapes of the SGM maps in the step and plateau regions.

The basic property of the step-region signal is the regular fringing, with some more transverse variation if the strong tip is used instead of the non-invasive one. The non-invasive SGM signal in plateau-region has a smooth decay, with no oscillation and no fine features in the transverse direction. At the same time the ideal strong-tip SGM plateau-signal has a chequerboard pattern, but not the fringes. In each case, a non-perfect tuning of the system onto the plateau will mix each of these behaviours with the corresponding step-like behaviour.

The exact tuning of the system for ideal plateau-region character of the signal is not possible for the short structures, where the constriction mode through which the transmission takes place does not manage to open fully before the onset of the next mode. It is also difficult for the structures which are too elongated and therefore have very thin Fabry-Perot resonances. Small but finite temperature will tend to spoil the plateau-signal quality by first *enhancing* the fringing pattern, and only when it gets larger by blurring it.

We have compared the step-region signal and imperfect plateau-region signal from the invasive SGM simulation with the experimental results from [14] (see fig. 1.5, Ch. 1), obtained for a clean structure at low temperature, finding that the general character – the fringes and the chequerboard pattern – agrees very well.

Adiabaticity and mode-mixing

We have envisaged in Ch. 3 and confirmed in Ch. 5 that the non-invasive signal we simulate can have the simple form – with exactly one fringing pattern for the first order signal and no fringes for the second order – solely in adiabatic structures. In the non-adiabatic structures, many oscillations are superposed, hence the first order signal looses the regular fringing in favour of the sharper dot-like features already at a few wave lengths from the constriction exit. Even the non-invasive SGM signal ideally tuned to a plateau has at least a chequerboard pattern.

This kind of signal, of course, does not resemble the experimental images from fig. 1.5, Ch. 1. This fact can be viewed as a confirmation of the adiabatic character of the electrostatically defined experimental constrictions, but can also be viewed as an artefact of the boundary influence, which in non-adiabatic structures is higher than in the other ones. In that case, still, a conclusion can be drawn that the non-adiabatic structures direct much of the

signal towards the walls, while the adiabatic ones focus it to a larger degree along the axis, or the axes for higher constriction modes. The focussing is certainly not an artefact of the hard-wall-leaded structure, which we saw in Ch. 4 in the results from the ideally adiabatic saddle-point model, where the structure is soft-walled and constantly widening – see fig. 4.6.

Thus, the simulated non-invasive SGM maps allow for assessing the level of adiabaticity of a structure without problem. Could the experimental images serve the same purpose? The lead-mode composition in the quasi-one-dimensional geometry is controlled by the u_l matrices of eq. (2.40), Ch. 2. When passing to the limit of infinitely wide leads, the u_l's become functions controlling the angular shape of the signal. We expect that this function should contain the information about how the electron waves got injected into the 2D free space, and thus also about the injecting structure.

6.3 Perspectives

The non-invasive SGM theory can serve as a framework for investigating some interesting issues which have not yet been answered.

The above-described question about the structure's adiabaticity and the focussing of signal is one of such. The problem of the angular profile of the signal has been solved in [44] for one geometry, the abrupt (rectangular) QPC, and assuming small, point-like exit from the constriction. Naturally, the answer to the focussing problem would require a generic solution, capable of determining the angular spreading of the SGM signal for at least some classes of constriction shapes. A controlled injection of the electrons into the wide regions of the 2DEG might be of use for experimental setups and for potential future nanodevices.

Another topic which we did not discuss in the present work is the disorder. The structure's disorder, as we mentioned in Ch. 1, is very important from the experimental point of view, as it is virtually impossible to eliminate it the real structures. It is also a very interesting topic for the theoretical research. It is widely agreed that the branching behaviour of the SGM signal is caused by the smooth component of the structure's disorder potential. Is it possible to obtain a similar effect with an abrupt potential, of shorter correlation length? What are the relations between the correlation length of the disorder and the correlation length of the signal? Can we relate the average characteristics of the branching pattern to the correlation and strength of the potential? Those questions are most easily answered by running the numerical simulations. This requires, however, investigating structures larger

than the ones presented in this work.[1] It would also be interesting to see the deviations of the SGM signal, both, invasive and non-invasive, from a calculated local density of states or even the density of current. (A dedicated program could be written for that.)

Our perturbative formulae can describe the conductance corrections in a disordered system, but the disorder potential enters the corrections through both, the eigenfunctions (or u_l matrices if choose the lead-state basis) and the transmission and reflection amplitudes. It might turn out that the alternative formulation is more suitable for this purpose, in which the "perturbed" quantities are not the the wave functions, but the transmission and reflection amplitudes (see introduction to Ch. 3), to which the random matrix theory could be applied.

There are two important research subjects which our theory cannot treat, but perhaps can be used for developing an approach which would tackle them – the electron-electron interactions and the non-linear (finite bias) transport. Both are closely related, as the correct description of the latter cannot omit including the earlier. The SGM of quantum point contacts is considered a potentially very useful technique in looking for the electronic interactions. Even in the early experiments, attempts have been made to asses how the electrons dephase as they travel. These methods have later been refined, but the issues related to the electronic interactions and correlations remain the field of active research. It has been suggested that the interaction effects can reveal in the finite bias SGM measurements inside the the constriction [15], or close to it [46]. In this context, the theoretical and simulation methods we have developed might play two kinds of roles: first, compared against possible future experiments with non-invasive probes, they can help to identify which measured features are not explicable within non-interacting models, thus pointing to the electronic interactions as their potential origin; second, they can eventually be used within one of the theoretical schemes of including the electron-electron correlations and lay a basis for an interacting-electron SGM theory.

[1] Preliminary tests have shown that branching cannot be reproduced by quantum calculations at these sizes.

Appendix A

The evaluation of the energy integrals

In this appendix, we explain the way of calculating the energy integrals appearing in the formulae for $\delta^{(1)}J^{(x)}_{1\varepsilon a}$ and $\delta^{(2)}J^{(x)}_{1\varepsilon a}$, that is, in the equations (3.16) and (3.33a-3.33b) of chapter 3. We recollect them here for the ease of access:

$$\delta^{(1)}J^{(x)}_{1\varepsilon a} = 2\mathrm{Re}\left[\sum_{l'=1}^{2}\int_{\varepsilon_1^{(\perp)}}^{\infty}\frac{\mathrm{d}\varepsilon'}{\varepsilon^{+}-\varepsilon'}\sum_{a'=1}^{N(\varepsilon')}\left[J^{(x)}\right]^{1\,l'}_{aa'}(\varepsilon,\varepsilon')\;\left[V_{\mathrm{T}}\right]^{l'1}_{a'a}(\varepsilon',\varepsilon)\right]\,, \tag{A.1}$$

$$\delta^{(2)}J^{(x)}_{1\varepsilon a} = \delta^{(2)}J^{(x)\alpha}_{1\varepsilon a} + \delta^{(2)}J^{(x)\beta}_{1\varepsilon a}\,, \tag{A.2}$$

with

$$\delta^{(2)}J^{(x)\alpha}_{1\varepsilon a} = 2\mathrm{Re}\left[\sum_{l',l''=1}^{2}\int_{\varepsilon_1^{(\perp)}}^{\infty}\frac{\mathrm{d}\varepsilon'}{\varepsilon^{+}-\varepsilon'}\int_{\varepsilon_1^{(\perp)}}^{\infty}\frac{\mathrm{d}\varepsilon''}{\varepsilon^{+}-\varepsilon''}\right. \tag{A.3a}$$
$$\left.\sum_{a'=1}^{N(\varepsilon')}\sum_{a''=1}^{N(\varepsilon'')}\left[J^{(x)}\right]^{1\,l'}_{aa'}(\varepsilon,\varepsilon')\;\left[V_{\mathrm{T}}\right]^{l'l''}_{a'a''}(\varepsilon',\varepsilon'')\;\left[V_{\mathrm{T}}\right]^{l''1}_{a''a}(\varepsilon'',\varepsilon)\right]\,,$$

$$\delta^{(2)}J^{(x)\beta}_{1\varepsilon a} = 2\mathrm{Re}\left[\sum_{l',l''=1}^{2}\int_{\varepsilon_1^{(\perp)}}^{\infty}\frac{\mathrm{d}\varepsilon'}{\varepsilon^{-}-\varepsilon'}\int_{\varepsilon_1^{(\perp)}}^{\infty}\frac{\mathrm{d}\varepsilon''}{\varepsilon^{+}-\varepsilon''}\right. \tag{A.3b}$$
$$\left.\sum_{a'=1}^{N(\varepsilon')}\sum_{a''=1}^{N(\varepsilon'')}\left[V_{\mathrm{T}}\right]^{1l'}_{aa'}(\varepsilon,\varepsilon')\;\left[X\right]^{l'\,l''}_{a'a''}(\varepsilon',\varepsilon'')\;\left[V_{\mathrm{T}}\right]^{l''1}_{a''a}(\varepsilon'',\varepsilon)\right]\,,$$

and

$$\left[J^{(x)}\right]^{1\,1}_{aa'}(\varepsilon,\varepsilon') = \frac{e}{2h}\sum_{b=1}^{\bar{N}}\left(\sqrt{\frac{k'_b}{k_b}}+\sqrt{\frac{k_b}{k'_b}}\right)t^{\dagger}_{ab}(\varepsilon)t_{ba'}(\varepsilon')\exp\left[i(k'_b-k_b)x\right]\,,$$

$$\left[J^{(x)}\right]^{1\,2}_{aa'}(\varepsilon,\varepsilon') = \frac{e}{2h}\left\{\left(\sqrt{\frac{k_{a'}}{k'_{a'}}}-\sqrt{\frac{k'_{a'}}{k_{a'}}}\right)t^{\dagger}_{aa'}(\varepsilon)\exp\left[-i(k'_{a'}+k_{a'})x\right]\right.$$

$$\left.+\sum_{b=1}^{\bar{N}}\left(\sqrt{\frac{k'_b}{k_b}}+\sqrt{\frac{k_b}{k'_b}}\right)t^{\dagger}_{ab}(\varepsilon)r'_{ba'}(\varepsilon')\exp\left[i(k'_b-k_b)x\right]\right\},$$

$$[X]^{1\,1}_{a'a''}(\varepsilon',\varepsilon'') = \frac{e}{2h}\sum_{b=1}^{\bar{N}}\sqrt{\frac{k''_b}{k'_b}}\;t^{\dagger}_{a'b}(\varepsilon')t_{ba''}(\varepsilon'')\;\exp\left[i(k''_b-k'_b)x\right],\tag{A.5a}$$

$$[X]^{1\,2}_{a'a''}(\varepsilon',\varepsilon'') = \frac{e}{2h}\left\{-\sqrt{\frac{k''_{a''}}{k'_{a''}}}\;t^{\dagger}_{a'a''}(\varepsilon')\;\exp\left[-i(k''_{a''}+k'_{a''})x\right]\right.\tag{A.5b}$$

$$\left.+\sum_{b=1}^{\bar{N}}\sqrt{\frac{k''_b}{k'_b}}\;t^{\dagger}_{a'b}(\varepsilon')r'_{ba''}(\varepsilon'')\;\exp\left[i(k''_b-k'_b)x\right]\right\},$$

$$[X]^{2\,1}_{a'a''}(\varepsilon',\varepsilon'') = \frac{e}{2h}\left\{\sqrt{\frac{k''_{a'}}{k'_{a'}}}\;t_{a'a''}(\varepsilon'')\;\exp\left[i(k''_{a'}+k'_{a'})x\right]\right.\tag{A.5c}$$

$$\left.+\sum_{b=1}^{\bar{N}}\sqrt{\frac{k''_b}{k'_b}}\;r'^{\dagger}_{a'b}(\varepsilon')t_{ba''}(\varepsilon'')\;\exp\left[i(k''_b-k'_b)x\right]\right\},$$

$$[X]^{2\,2}_{a'a''}(\varepsilon',\varepsilon'') = \frac{e}{2h}\left\{-\delta_{a'a''}\;\sqrt{\frac{k''_{a'}}{k'_{a'}}}\;\exp\left[-i(k''_{a''}-k'_{a'})x\right]\right.\tag{A.5d}$$

$$+\sqrt{\frac{k''_{a'}}{k'_{a'}}}\; r'_{a'a''}(\varepsilon'')\exp\left[i(k''_{a'}+k'_{a'})x\right]$$
$$-\sqrt{\frac{k''_{a''}}{k'_{a''}}}\; r'^{\dagger}_{a'a''}(\varepsilon')\exp\left[-i(k''_{a''}+k'_{a''})x\right]$$
$$+\sum_{b=1}^{\bar{N}}\sqrt{\frac{k''_b}{k'_b}}\; r'^{\dagger}_{a'b}(\varepsilon')r'_{ba''}(\varepsilon'')\;\exp\left[i(k''_b-k'_b)x\right]\Bigg\}\;.$$

For our purpose, we will need the following equalities ([56], sec. 4-2b, p. 292):

$$\lim_{x\to-\infty}\lim_{\eta\to0^+}\frac{\exp(ikx)}{k-i\eta}=0\,, \tag{A.6a}$$
$$\lim_{x\to+\infty}\lim_{\eta\to0^+}\frac{\exp(ikx)}{k-i\eta}=2\pi i\delta(k)\,, \tag{A.6b}$$
$$\lim_{x\to-\infty}\lim_{\eta\to0^+}\frac{\exp(ikx)}{k+i\eta}=-2\pi i\delta(k)\,, \tag{A.6c}$$
$$\lim_{x\to+\infty}\lim_{\eta\to0^+}\frac{\exp(ikx)}{k+i\eta}=0\,, \tag{A.6d}$$

which should be understood in the distribution sense, ie. can be meaningful only when multiplied onto[1] some smooth function of k.

Below, we list six distinct types of terms, that appear in (A.1) and (A.2) once the $[J^{(x)}]$- and $[X]$-expressions are substituted, and outline the way of evaluating the energy integrals inside them, in the limit of $x\to\infty$. For the variable changes, we use the dispersion relation $\varepsilon(k_n,n)$, eq. (2.6):

$$\varepsilon=\frac{\hbar^2}{2M_e}k_n^2+\varepsilon_n^{\perp}\;.$$

The expressions that we replace by the dots inside the integrands – containing square roots of longitudinal wave numbers k_n, matrix elements of the tip potential $[V_{\mathrm{T}}]$, and the ingredients of the S-matrix – vary smoothly with k_n along its positive real axis. We assume that, with appropriate choice of branch-cut lines, they have analytic continuations at least in the interior of the first quadrant of the complex plane. These properties can be readily verified in the particular situations we consider in Ch. 4. In the case when one wave number is treated as a function of another wave number, care has to be taken about the ranges and mode indices, so that the function

$$k_m(k_n)=\sqrt{k_n^2+\bar{\kappa}_{nm}^2}$$

[1] We mean the scalar multiplication in an appropriate function space.

with

$$\bar{\kappa}_{nm} = \sqrt{\frac{2M_e}{\hbar^2}\left(\varepsilon_n^{(\perp)} - \varepsilon_m^{(\perp)}\right)}$$

remains analytic.

At the very end, in point 3, we treat the $\delta^{(2)} J_{1\varepsilon a}^{(x)\alpha}$ terms of the eq. (A.3a) ((3.33a), subsec. 3.3.1).

1. The single integrals of eq. (A.1) ((3.16) of sec. 3.2.1)

Here $\bar{N} \equiv \min\{N(\varepsilon),\, N(\varepsilon')\}$.

1.1 *Delta plus* terms

$$\int_{\varepsilon_1^{(\perp)}}^{\infty} \frac{\mathrm{d}\varepsilon'}{\varepsilon^+ - \varepsilon'} \sum_{a'=1}^{N(\varepsilon')} \tag{A.7}$$

$$\frac{e}{2h} \sum_{b=1}^{\bar{N}} \left(\sqrt{\frac{k_b'}{k_b}} + \sqrt{\frac{k_b}{k_b'}} \right) \left\{ \begin{array}{l} t_{ab}^{\dagger}(\varepsilon) t_{ba'}(\varepsilon')\ [V_{\mathrm{T}}]_{a'a}^{1\,1}(\varepsilon', \varepsilon) \\ t_{ab}^{\dagger}(\varepsilon) r_{ba'}'(\varepsilon')\ [V_{\mathrm{T}}]_{a'a}^{2\,1}(\varepsilon', \varepsilon) \end{array} \right\} \exp\left[i(k_b' - k_b)x\right] =$$

where the upper variant of the expression in braces corresponds to the $l' = 1$ addend of the equation (A.1), and the lower one, to the $l' = 2$ addend of the same equation; continuing:

$$= \sum_{b=1}^{N(\varepsilon)} \int_{\varepsilon_b^{(\perp)}}^{\infty} \frac{\mathrm{d}\varepsilon'}{\varepsilon - \varepsilon' + i\eta} \exp\left[i(k_b' - k_b)x\right] \sum_{a'=1}^{N(\varepsilon')} \Big(\dots \Big)$$

after a variable change $\varepsilon' \mapsto k_b'$ and bearing in mind that the same mapping exists for the ε parameter, $\varepsilon \mapsto k_b$:

$$= \sum_{b=1}^{N(\varepsilon)} \int_0^{\infty} \frac{-2k_b'\,\mathrm{d}k_b'}{k_b' + k_b + i\bar{\eta}} \frac{\exp\left[i(k_b' - k_b)x\right]}{k_b' - k_b - i\bar{\eta}} \sum_{a'=1}^{N(\varepsilon(k_b', b))} \Big(\dots \Big)$$

where $\bar{\eta} \to 0^+$ as $\eta \to 0^+$; then, based on (A.6b),

$$= \sum_{b=1}^{N(\varepsilon)} \int_0^{\infty} \frac{-2k_b'\,\mathrm{d}k_b'}{k_b' + k_b} 2\pi i \delta(k_b' - k_b) \sum_{a'=1}^{N(\varepsilon(k_b', b))} \Big(\dots \Big)$$

$$= -2\pi i \sum_{a'=1}^{N(\varepsilon)} \sum_{b=1}^{N(\varepsilon)} \Big(\ \dots \ \Big) \Big|_{k'_b = k_b}$$

$$= -2\pi i \frac{e}{h} \sum_{a'=1}^{N(\varepsilon)} \sum_{b=1}^{N(\varepsilon)} \begin{cases} t^\dagger_{ab}(\varepsilon) t_{ba'}(\varepsilon) \ [V_{\rm T}]^{1\,1}_{a'a}(\varepsilon, \varepsilon) \\ \\ t^\dagger_{ab}(\varepsilon) r'_{ba'}(\varepsilon) \ [V_{\rm T}]^{2\,1}_{a'a}(\varepsilon, \varepsilon) \end{cases} ,$$

or, in the shorthand notation (3.18), of ch. (3):

$$= -2\pi i \frac{e}{h} \begin{cases} \left(t^\dagger t \, \mathcal{V}^{1\,1} \right)_{aa} \\ \\ \left(t^\dagger r' \, \mathcal{V}^{2\,1} \right)_{aa} \end{cases} . \tag{A.8}$$

1.2 *Sigma minus* terms

These come from the $l' = 2$ addend of the eq. (A.1):

$$\int_{\varepsilon_1^{(\perp)}}^{\infty} \frac{{\rm d}\varepsilon'}{\varepsilon^+ - \varepsilon'} \sum_{a'=1}^{N(\varepsilon')} \tag{A.9}$$

$$\frac{e}{2h} \left(\sqrt{\frac{k_{a'}}{k'_{a'}}} - \sqrt{\frac{k'_{a'}}{k_{a'}}} \right) t^\dagger_{aa'}(\varepsilon) \ [V_{\rm T}]^{l'1}_{a'a}(\varepsilon', \varepsilon) \ \exp\left[-i(k'_{a'} + k_{a'})x \right] =$$

$$= \sum_{a'=1}^{\infty} \int_{\varepsilon_{a'}^{(\perp)}}^{\infty} \frac{{\rm d}\varepsilon'}{\varepsilon - \varepsilon' + i\eta} \exp\left[-i(k'_{a'} + k_{a'})x \right] \ \Big(\ \dots \ \Big)$$

$$= \sum_{a'=1}^{\infty} \exp\left(-2ik_{a'}x \right) \int_0^{\infty} \frac{-2k'_{a'} \, {\rm d}k'_{a'}}{k'_{a'} + k_{a'} + i\bar{\eta}} \ \frac{\exp\left[-i(k'_{a'} - k_{a'})x \right]}{k'_{a'} - k_{a'} - i\bar{\eta}} \ \Big(\ \dots \ \Big)$$

$$= 0 \ , \tag{A.10}$$

as follows from (A.6a) when $-x \to -\infty$. The limit is not endangered by the $\exp\left(-2ik_{a'}x \right)$ factor, since the independent variable $k_{a'}$ in the integration acts merely as a parameter.

2. The double integrals of (A.3a) and (A.3b)

Here $\bar{N} \equiv \min\{N(\varepsilon'), \, N(\varepsilon'')\}$.

2.1 *delta plus* terms

$$\int_{\varepsilon_1^{(\perp)}}^{\infty} \frac{\mathrm{d}\varepsilon'}{\varepsilon^- - \varepsilon'} \int_{\varepsilon_1^{(\perp)}}^{\infty} \frac{\mathrm{d}\varepsilon''}{\varepsilon^+ - \varepsilon''} \sum_{a'=1}^{N(\varepsilon')} \sum_{a''=1}^{N(\varepsilon'')} \tag{A.11}$$

$$\frac{e}{2h} \sum_{b=1}^{\bar{N}} \sqrt{\frac{k_b''}{k_b'}} \left\{ \begin{array}{l} [V_{\mathrm{T}}]^{1\,1}_{aa'}(\varepsilon,\varepsilon')\; t^{\dagger}_{a'b}(\varepsilon') t_{ba''}(\varepsilon'')\; [V_{\mathrm{T}}]^{1\,1}_{a''a}(\varepsilon'',\varepsilon) \\ [V_{\mathrm{T}}]^{1\,1}_{aa'}(\varepsilon,\varepsilon')\; t^{\dagger}_{a'b}(\varepsilon') r'_{ba''}(\varepsilon'')\; [V_{\mathrm{T}}]^{2\,1}_{a''a}(\varepsilon'',\varepsilon) \\ [V_{\mathrm{T}}]^{1\,2}_{aa'}(\varepsilon,\varepsilon')\; r'^{\dagger}_{a'b}(\varepsilon') t_{ba''}(\varepsilon'')\; [V_{\mathrm{T}}]^{1\,1}_{a''a}(\varepsilon'',\varepsilon) \\ [V_{\mathrm{T}}]^{1\,2}_{aa'}(\varepsilon,\varepsilon')\; r'^{\dagger}_{a'b}(\varepsilon') r'_{ba''}(\varepsilon'')\; [V_{\mathrm{T}}]^{2\,1}_{a''a}(\varepsilon'',\varepsilon) \end{array} \right\} \exp\left[i(k_b'' - k_b')x\right] =$$

where the factors in the braces are due to, respectively:

- the last term of the $[X]^{1\,1}_{a'a''}(\varepsilon',\varepsilon'')$ factor,
- the last term of the $[X]^{1\,2}_{a'a''}(\varepsilon',\varepsilon'')$ factor,
- the last term of the $[X]^{2\,1}_{a'a''}(\varepsilon',\varepsilon'')$ factor,
- the last term of the $[X]^{2\,2}_{a'a''}(\varepsilon',\varepsilon'')$ factor,

from the equation (A.3b),

$$= \sum_{b=1}^{\infty} \int_{\varepsilon_b^{(\perp)}}^{\infty} \frac{\mathrm{d}\varepsilon'}{\varepsilon - \varepsilon' - i\eta} \int_{\varepsilon_b^{(\perp)}}^{\infty} \frac{\mathrm{d}\varepsilon''}{\varepsilon - \varepsilon'' + i\eta}\; \exp\left[i(k_b'' - k_b')x\right] \sum_{a'=1}^{N(\varepsilon')} \sum_{a''=1}^{N(\varepsilon'')} \Big(\dots \Big)$$

$$= \sum_{b=1}^{\infty} \int_0^{\infty} \frac{-2k_b'\,\mathrm{d}k_b'}{k_b' + k_b + i\bar{\eta}}\; \frac{\exp\left[-i(k_b' - k_b)x\right]}{k_b' - k_b + i\bar{\eta}}$$

$$\int_0^{\infty} \frac{-2k_b''\,\mathrm{d}k_b''}{k_b'' + k_b + i\bar{\eta}}\; \frac{\exp\left[i(k_b'' - k_b)x\right]}{k_b'' - k_b - i\bar{\eta}} \sum_{a'=1}^{N(\varepsilon(k_b',b))} \sum_{a''=1}^{N(\varepsilon(k_b'',b))} \Big(\dots \Big)$$

we use (A.6c) and (A.6b):

$$= \sum_{b=1}^{\infty} \int_0^{\infty} \frac{-2k_b'\,\mathrm{d}k_b'}{k_b' + k_b}\; (-2\pi i \delta(k_b' - k_b))$$

$$\int_0^{\infty} \frac{-2k_b''\,\mathrm{d}k_b''}{k_b'' + k_b}\; 2\pi i \delta(k_b'' - k_b) \sum_{a'=1}^{N(\varepsilon(k_b',b))} \sum_{a''=1}^{N(\varepsilon(k_b'',b))} \Big(\dots \Big)$$

$$= 4\pi^2 \sum_{b=1}^{N(\varepsilon)} \sum_{a'=1}^{N(\varepsilon)} \sum_{a''=1}^{N(\varepsilon)} \Big(\dots \Big) \Big|_{k''_b = k'_b = k_b}$$

$$= 4\pi^2 \frac{e}{2h} \begin{cases} \left(\mathcal{V}^{1\,1}\ t^\dagger t\ \ \mathcal{V}^{1\,1} \right)_{aa} \\ \left(\mathcal{V}^{1\,1}\ t^\dagger r'\ \mathcal{V}^{2\,1} \right)_{aa} \\ \left(\mathcal{V}^{1\,2}\ r'^\dagger t\ \mathcal{V}^{1\,1} \right)_{aa} \\ \left(\mathcal{V}^{1\,2}\ r'^\dagger r'\ \mathcal{V}^{2\,1} \right)_{aa} \end{cases} .$$

The upper restriction on the b-sum in the penultimate line is dictated by the fact that for $b > N(\varepsilon)$ the points $k''_b = k'_b = k_b$ sifted by the delta functions are no longer inside the integration scope.

2.2 *sigma minus* terms
of eq. (A.3b), due to the first addend of $[X]^{1\,2}_{a'a''}(\varepsilon', \varepsilon'')$:

$$\int_{\varepsilon_1^{(\perp)}}^{\infty} \frac{\mathrm{d}\varepsilon'}{\varepsilon^- - \varepsilon'} \int_{\varepsilon_1^{(\perp)}}^{\infty} \frac{\mathrm{d}\varepsilon''}{\varepsilon^+ - \varepsilon''} \sum_{a'=1}^{N(\varepsilon')} \sum_{a''=1}^{N(\varepsilon'')} \tag{A.12}$$

$$- \frac{e}{2h} \sqrt{\frac{k''_{a''}}{k'_{a''}}}\ t^\dagger_{a'a''}(\varepsilon')\ \exp\left[-i(k''_{a''} + k'_{a''})x\right] =$$

$$= \sum_{a''=1}^{\infty} \int_{\varepsilon_{a''}^{(\perp)}}^{\infty} \frac{\mathrm{d}\varepsilon'}{\varepsilon - \varepsilon' - i\eta} \int_{\varepsilon_{a''}^{(\perp)}}^{\infty} \frac{\mathrm{d}\varepsilon''}{\varepsilon - \varepsilon'' + i\eta} \exp\left[-i(k''_{a''} + k'_{a''})x\right] \sum_{a'=1}^{N(\varepsilon')} \Big(\dots \Big)$$

$$= \sum_{a''=1}^{\infty} \exp\left(-2ik_{a''}x\right) \int_0^{\infty} \frac{-2k'_{a''}\,\mathrm{d}k'_{a''}}{k'_{a''} + k_{a''} + i\bar{\eta}}\ \frac{\exp\left[-i(k'_{a''} - k_{a''})x\right]}{k'_{a''} - k_{a''} + i\bar{\eta}}$$

$$\int_0^{\infty} \frac{-2k''_{a''}\,\mathrm{d}k''_{a''}}{k''_{a''} + k_{a''} + i\bar{\eta}}\ \frac{\exp\left[-i(k''_{a''} - k_{a''})x\right]}{k''_{a''} - k_{a''} - i\bar{\eta}} \sum_{a'=1}^{N(\varepsilon(k'_{a''}, a''))} \Big(\dots \Big)$$

$$= 0\ ,$$

after the use of (A.6c) and (A.6a).

2.3 *sigma plus* terms
of eq. (A.3b), due to the first addend of $[X]^{2\,1}_{a'a''}(\varepsilon',\varepsilon'')$:

$$\int_{\varepsilon_1^{(\perp)}}^{\infty}\frac{\mathrm{d}\varepsilon'}{\varepsilon^- - \varepsilon'}\int_{\varepsilon_1^{(\perp)}}^{\infty}\frac{\mathrm{d}\varepsilon''}{\varepsilon^+ - \varepsilon''}\sum_{a'=1}^{N(\varepsilon')}\sum_{a''=1}^{N(\varepsilon'')}\frac{e}{2h}\sqrt{\frac{k''_{a'}}{k'_{a'}}}\;t_{a'a''}(\varepsilon'')\;\exp\left[i(k''_{a'}+k'_{a'})x\right] \tag{A.13}$$

This case is exactly analogous to the previous one – the integral vanishes, only the use is made of (A.6d) and (A.6b) instead, and the leading sum is over a'.

2.4 *delta minus* terms
of eq. (A.3b), due to the first addend of $[X]^{2\,2}_{a'a''}(\varepsilon',\varepsilon'')$:

$$\int_{\varepsilon_1^{(\perp)}}^{\infty}\frac{\mathrm{d}\varepsilon'}{\varepsilon^- - \varepsilon'}\int_{\varepsilon_1^{(\perp)}}^{\infty}\frac{\mathrm{d}\varepsilon''}{\varepsilon^+ - \varepsilon''}\sum_{a=1}^{\bar{N}}\frac{e}{2h}\sqrt{\frac{k''_{a}}{k'_{a}}}\;\exp\left[-i(k''_{a}-k'_{a})x\right] = \tag{A.14}$$

$$\begin{aligned}
&= \sum_{a=1}^{\infty}\int_{\varepsilon_a^{(\perp)}}^{\infty}\frac{\mathrm{d}\varepsilon'}{\varepsilon-\varepsilon'-i\eta}\int_{\varepsilon_a^{(\perp)}}^{\infty}\frac{\mathrm{d}\varepsilon''}{\varepsilon-\varepsilon''+i\eta}\exp\left[-i(k''_a-k'_a)x\right]\Big(\;\ldots\;\Big)\\
&= \sum_{a=1}^{\infty}\int_{0}^{\infty}\frac{-2k'_a\,\mathrm{d}k'_a}{k'_a+k_a+i\bar{\eta}}\;\frac{\exp\left[i(k'_a-k_a)x\right]}{k'_a-k_a+i\bar{\eta}}\\
&\qquad\qquad\int_{0}^{\infty}\frac{-2k''_a\,\mathrm{d}k''_a}{k''_a+k_a+i\bar{\eta}}\;\frac{\exp\left[-i(k''_a-k_a)x\right]}{k''_a-k_a-i\bar{\eta}}\;\Big(\;\ldots\;\Big)\\
&= 0\;,
\end{aligned}$$

after the use of (A.6d) and (A.6a).

3. The α-terms

The equation (A.3a) for the α-terms of the second order current correction can be rearranged in the following way:

$$\begin{aligned}
\delta^{(2)}J^{(x)\alpha}_{1\varepsilon a} = 2\mathrm{Re}\Bigg\{&\int_{\varepsilon_1^{(\perp)}}^{\infty}\frac{\mathrm{d}\varepsilon'}{\varepsilon^+-\varepsilon'}\sum_{a'=1}^{N(\varepsilon')}\\
&\left[\left[J^{(x)}\right]^{1\,1}_{aa'}(\varepsilon,\varepsilon')\;[W]^{1\,1}_{a'a}(\varepsilon',\varepsilon)+\left[J^{(x)}\right]^{1\,2}_{aa'}(\varepsilon,\varepsilon')\;[W]^{2\,1}_{a'a}(\varepsilon',\varepsilon)\right]\Bigg\}
\end{aligned} \tag{A.15}$$

with

$$[W]^{1\,1}_{a'a}(\varepsilon',\varepsilon) = \int_{\varepsilon_1^{(\perp)}}^{\infty} \frac{\mathrm{d}\varepsilon''}{\varepsilon^+ - \varepsilon''} \sum_{a''=1}^{N(\varepsilon'')} \Big[[V_{\mathrm{T}}]^{1\,1}_{a'a''}(\varepsilon',\varepsilon'')\, [V_{\mathrm{T}}]^{1\,1}_{a''a}(\varepsilon'',\varepsilon) + [V_{\mathrm{T}}]^{1\,2}_{a'a''}(\varepsilon',\varepsilon'')\, [V_{\mathrm{T}}]^{2\,1}_{a''a}(\varepsilon'',\varepsilon) \Big] \tag{A.16a}$$

$$[W]^{2\,1}_{a'a}(\varepsilon',\varepsilon) = \int_{\varepsilon_1^{(\perp)}}^{\infty} \frac{\mathrm{d}\varepsilon''}{\varepsilon^+ - \varepsilon''} \sum_{a''=1}^{N(\varepsilon'')} \Big[[V_{\mathrm{T}}]^{2\,1}_{a'a''}(\varepsilon',\varepsilon'')\, [V_{\mathrm{T}}]^{1\,1}_{a''a}(\varepsilon'',\varepsilon) + [V_{\mathrm{T}}]^{2\,2}_{a'a''}(\varepsilon',\varepsilon'')\, [V_{\mathrm{T}}]^{2\,1}_{a''a}(\varepsilon'',\varepsilon) \Big] \tag{A.16b}$$

The expression (A.15) is a direct analogue of (A.1), so we can apply the methods for single integrals, described in point 1. These bring us to

$$\delta^{(2)} J^{(x)\alpha}_{1\varepsilon a} = 4\pi \frac{e}{h} \operatorname{Im}\left[\left(t^\dagger t\, \mathcal{W}^{1\,1} + t^\dagger r'\, \mathcal{W}^{2\,1} \right)_{aa} \right] , \tag{A.17}$$

with the notation

$$\mathcal{W}^{ll'}_{\cdot\cdot} \equiv [W]^{ll'}_{\cdot\cdot}(\varepsilon,\varepsilon) . \tag{A.18}$$

By Sochocki-Weierstrass theorem ([84], sec. 5.7), we can further write:

$$\mathcal{W}^{1\,1}_{a'a} = -i\pi \left(\mathcal{V}^{1\,1}\, \mathcal{V}^{1\,1} + \mathcal{V}^{1\,2}\, \mathcal{V}^{2\,1} \right)_{a'a} + \mathcal{P} \int_{\varepsilon_1^{(\perp)}}^{\infty} \frac{\mathrm{d}\varepsilon''}{\varepsilon - \varepsilon''} \sum_{a''=1}^{N(\varepsilon'')} \Big\{ [V_{\mathrm{T}}]^{1\,1}_{a'a''}(\varepsilon,\varepsilon'')\, [V_{\mathrm{T}}]^{1\,1}_{a''a}(\varepsilon'',\varepsilon) + [V_{\mathrm{T}}]^{1\,2}_{a'a''}(\varepsilon,\varepsilon'')\, [V_{\mathrm{T}}]^{2\,1}_{a''a}(\varepsilon'',\varepsilon) \Big\} , \tag{A.19a}$$

$$\mathcal{W}^{2\,1}_{a'a} = -i\pi \left(\mathcal{V}^{2\,1}\, \mathcal{V}^{1\,1} + \mathcal{V}^{2\,2}\, \mathcal{V}^{2\,1} \right)_{a'a} + \mathcal{P} \int_{\varepsilon_1^{(\perp)}}^{\infty} \frac{\mathrm{d}\varepsilon''}{\varepsilon - \varepsilon''} \sum_{a''=1}^{N(\varepsilon'')} \Big\{ [V_{\mathrm{T}}]^{2\,1}_{a'a''}(\varepsilon,\varepsilon'')\, [V_{\mathrm{T}}]^{1\,1}_{a''a}(\varepsilon'',\varepsilon) + [V_{\mathrm{T}}]^{2\,2}_{a'a''}(\varepsilon,\varepsilon'')\, [V_{\mathrm{T}}]^{2\,1}_{a''a}(\varepsilon'',\varepsilon) \Big\} . \tag{A.19b}$$

Appendix B

Current density

Current density

For definiteness, we will consider the region to the right of the scatterer structure.

The longitudinal and transverse components of the current density $\mathbf{j}\{\Psi_{l\varepsilon a};\mathbf{r}\}$ – see eq. (2.26), subsec. 2.2.2 of Ch. 2 – due to a single left-originating scattering state $(1\varepsilon a)$ are:

$$j^{(x)}\{\Psi_{1\varepsilon a};\mathbf{r}\} = \frac{e\hbar}{M_e}\sum_{b,b'\in\mathcal{M}} \mathrm{Im}\left(ik_{b'}\; t_{b'a}t^{\dagger}_{ab}\; \varphi^{(-)}_{2\varepsilon b}(\mathbf{r})\,\varphi^{(+)}_{2\varepsilon b'}(\mathbf{r})\right)\;, \tag{B.1a}$$

$$j^{(y)}\{\Psi_{1\varepsilon a};\mathbf{r}\} = \frac{e\hbar}{M_e}\sum_{b,b'\in\mathcal{M}} \mathrm{Im}\left(\frac{1}{\phi_{b'}}\frac{\partial\phi_{b'}}{\partial y}\; t_{b'a}t^{\dagger}_{ab}\; \varphi^{(-)}_{2\varepsilon b}(\mathbf{r})\,\varphi^{(+)}_{2\varepsilon b'}(\mathbf{r})\right)\;. \tag{B.1b}$$

The total x- and y-current densities at a given energy are obtained by taking an incoherent sum over all the states at that energy. We take only the states coming from the left lead, getting:

$$\begin{aligned} j^{(x)}_{1\varepsilon}(\mathbf{r}) &= \sum_a j^{(x)}\{\Psi_{1\varepsilon a};\mathbf{r}\} \\ &= \frac{e\hbar}{M_e}\,\mathrm{Im}\left\{ \sum_{b'b} ik_{b'}\;\left(tt^{\dagger}\right)_{b'b}\;\varphi^{(-)}_{2\varepsilon b}(\mathbf{r})\,\varphi^{(+)}_{2\varepsilon b'}(\mathbf{r})\right\}\;, \end{aligned} \tag{B.2a}$$

and

$$\begin{aligned} j^{(y)}_{1\varepsilon}(\mathbf{r}) &= \sum_a j^{(y)}\{\Psi_{1\varepsilon a};\mathbf{r}\} \\ &= \frac{e\hbar}{M_e}\,\mathrm{Im}\left\{ \sum_{b'b} \frac{1}{\phi_{b'}}\frac{\partial\phi_{b'}}{\partial y}\;\left(tt^{\dagger}\right)_{b'b}\;\varphi^{(-)}_{2\varepsilon.}(\mathbf{r})\,\varphi^{(+)}_{2\varepsilon.}(\mathbf{r})\right\}\;. \end{aligned} \tag{B.2b}$$

These two components make up the vector field of current density

$$\mathbf{j}_{1\varepsilon} = \hat{\mathbf{x}}\, j_{1\varepsilon}^{(x)} + \hat{\mathbf{y}}\, j_{1\varepsilon}^{(y)} \ . \tag{B.3}$$

With the aid of the S-matrix unitarity and symmetry properties (see subsecs. 2.3.1–2.3.2, Ch. 2), we can retrieve exactly the same expressions, but with the opposite signs, for the total current density of the scattering states originating in the right lead, $\mathbf{j}_{2\varepsilon}(\mathbf{r})$:

$$\mathbf{j}_{1\varepsilon}(\mathbf{r}) = -\mathbf{j}_{2\varepsilon}(\mathbf{r}) \ . \tag{B.4}$$

While in Ch. 2 we have shown that the left- and right-originating currents cancel each other at the energies below the right lead's chemical potential, μ_2, here we see that not only the currents, but also their spatial (and energetic) distributions cancel each other exactly, everywhere in space and for all the energies below μ_2.[1] Therefore, the current density $\mathbf{j}(\mathbf{r})$ in our system is made up solely by the left-originating states – the ones which remain occupied in the "Fermi window" of energies between the leads' chemical potentials μ_2 and $\mu_1 = \mu_2 + \mathrm{d}E_F$:

$$\mathbf{j}(\mathbf{r}) = \mathbf{j}_{1E_F}(\mathbf{r}) \ . \tag{B.5}$$

It is, still, not completely unrelated to the right-originating states, even though they stay empty and do not participate in creating it. The relation is given simply by the eq. (B.4).

[1] We remark that no spatial symmetry of the structure is required; satisfying the time-reversal invariance (TRI) and total flux conservation is sufficient. On the other hand, this reasoning is spoilt when magnetic field is present, as this breaks the TRI. Then, circulating currents appear, altering the current density, but not the total current.

Bibliography

[1] M. A. Topinka, B. J. LeRoy, S. E. J. Shaw, E. J. Heller, R. M. Westervelt, K. D. Maranowski, and A. C. Gossard. Imaging coherent electron flow from a quantum point contact. Science, 289(5488):2323–2326, 2000.

[2] D. K. Ferry, S. M. Goodnick, and J. Bird. Transport in Nanostructures. Cambridge University Press, 2009.

[3] A. Baumgartner, T. Ihn, K. Ensslin, K. Maranowski, and A. C. Gossard. Quantum hall effect transition in scanning gate experiments. Phys. Rev. B, 76:085316, Aug 2007.

[4] M. A. Eriksson, R. G. Beck, M. Topinka, J. A. Katine, R. M. Westervelt, K. L. Campman, and A. C. Gossard. Cryogenic scanning probe characterization of semiconductor nanostructures. Applied Physics Letters, 69(5):671–673, 1996.

[5] H. Sellier, B. Hackens, M. G. Pala, F. Martins, S. Baltazar, X. Wallart, L. Desplanque, V. Bayot, and S. Huant. On the imaging of electron transport in semiconductor quantum structures by scanning-gate microscopy: successes and limitations. Semiconductor Science and Technology, 26(6):064008, 2011.

[6] M. A. Topinka, B. J. LeRoy, R. M. Westervelt, S. E. J. Shaw, R. Fleischmann, E. J. Heller, K. D. Maranowski, and A. C. Gossard. Coherent branched flow in a two-dimensional electron gas. Nature, 410(6825):183–186, Mar 2001.

[7] M. A. Topinka, B. J. LeRoy, R. M. Westervelt, K. D. Maranowski, and A. C. Gossard. Imaging coherent electron wave flow in a two-dimensional electron gas. Physica E: Low-dimensional Systems and Nanostructures, 12(1-4):678–683, 2002. Proceedings of the Fourteenth International Conference on the Electronic Properties of Two-Dimensional Systems.

[8] B. J. LeRoy, M. A. Topinka, R. M. Westervelt, K. D. Maranowski, and A. C. Gossard. Imaging electron density in a two-dimensional electron gas. Applied Physics Letters, 80(23):4431–4433, 2002.

[9] B. J. LeRoy, M. A. Topinka, A. C. Bleszynski, R. M. Westervelt, S. E. J. Shaw, E. J. Heller, K. D. Maranowski, and A. C. Gossard. Imaging coherent electron flow in a two-dimensional electron gas. Applied Surface Science, 210(1-2):134–139, 2003. 5th International Conference on non-contact AFM in Montreal, Canada.

[10] B. J. LeRoy, A. C. Bleszynski, K. E. Aidala, R. M. Westervelt, A. Kalben, E. J. Heller, S. E. J. Shaw, K. D. Maranowski, and A. C. Gossard. Imaging electron interferometer. Phys. Rev. Lett., 94:126801, Apr 2005.

[11] N. Aoki, C. R. Da Cunha, R. Akis, D. K. Ferry, and Y. Ochiai. Scanning gate microscopy investigations on an InGaAs quantum point contact. Applied Physics Letters, 87(22):223501, 2005.

[12] K. E. Aidala, R. E. Parrott, T. Kramer, E. J. Heller, R. M. Westervelt, M. P. Hanson, and A. C. Gossard. Imaging magnetic focusing of coherent electron waves. Nat Phys, 3(7):464–468, Jul 2007.

[13] M. P. Jura, M. A. Topinka, L. Urban, A. Yazdani, H. Shtrikman, L. N. Pfeiffer, K. W. West, and D. Goldhaber-Gordon. Unexpected features of branched flow through high-mobility two-dimensional electron gases. Nature Physics, 3:841–845, 2007.

[14] M. P. Jura, M. A. Topinka, M. Grobis, L. N. Pfeiffer, K. W. West, and D. Goldhaber-Gordon. Electron interferometer formed with a scanning probe tip and quantum point contact. Phys. Rev. B, 80:041303, Jul 2009.

[15] M. P. Jura, M. Grobis, M. A. Topinka, L. N. Pfeiffer, K. W. West, and D. Goldhaber-Gordon. Spatially probed electron-electron scattering in a two-dimensional electron gas. Phys. Rev. B, 82:155328, Oct 2010.

[16] S. Schnez, C. Rössler, T. Ihn, K. Ensslin, C. Reichl, and W. Wegscheider. Imaging the lateral shift of a quantum point contact using scanning gate microscopy. Phys. Rev. B, 84:195322, Nov 2011.

[17] N. Aoki, R. Brunner, A. M. Burke, R. Akis, R. Meisels, D. K. Ferry, and Y. Ochiai. Direct imaging of electron states in open quantum dots. Phys. Rev. Lett., 108:136804, Mar 2012.

[18] A. A. Kozikov, C. Rössler, T. Ihn, K. Ensslin, C. Reichl, and W. Wegscheider. Interference of electrons in backscattering through a quantum point contact. New Journal of Physics, 15(1):013056, 2013.

[19] A. Bachtold, M. S. Fuhrer, S. Plyasunov, M. Forero, Erik H. Anderson, A. Zettl, and Paul L. McEuen. Scanned probe microscopy of electronic transport in carbon nanotubes. Phys. Rev. Lett., 84:6082–6085, Jun 2000.

[20] R. Crook, C. G. Smith, M. Y. Simmons, and D. A. Ritchie. Imaging electrostatic microconstrictions in long 1d wires. Physica E: Low-dimensional Systems and Nanostructures, 12(1-4):695–698, 2002. Proceedings of the Fourteenth International Conference on the Ele ctronic Properties of Two-Dimensional Systems.

[21] A. Pioda, S. Kičin, T. Ihn, M. Sigrist, A. Fuhrer, K. Ensslin, A. Weichselbaum, S. E. Ulloa, M. Reinwald, and W. Wegscheider. Spatially resolved manipulation of single electrons in quantum dots using a scanned probe. Phys. Rev. Lett., 93:216801, Nov 2004.

[22] P. Fallahi, A. C. Bleszynski, R. M. Westervelt, J. Huang, J. D. Walls, E. J. Heller, M. Hanson, and A. C. Gossard. Imaging a single-electron quantum dot. Nano Letters, 5(2):223–226, 2005.

[23] M. T. Woodside and P. L. McEuen. Scanned probe imaging of single-electron charge states in nanotube quantum dots. Science, 296(5570):1098–1101, 2002.

[24] A. C. Bleszynski, F. A. Zwanenburg, R. M. Westervelt, A. L. Roest, E. P. A. M. Bakkers, and L. P. Kouwenhoven. Scanned probe imaging of quantum dots inside InAs nanowires. Nano Letters, 7(9):2559–2562, 2007.

[25] E. E. Boyd and R. M. Westervelt. Extracting the density profile of an electronic wave function in a quantum dot. Phys. Rev. B, 84:205308, Nov 2011.

[26] B. Hackens, F. Martins, T. Ouisse, H. Sellier, S. Bollaert, X. Wallart, A. Cappy, J. Chevrier, V. Bayot, and S. Huant. Imaging and controlling electron transport inside a quantum ring. Nat Phys, 2(12):826–830, Dec 2006.

[27] F. Martins, B. Hackens, M. G. Pala, T. Ouisse, H. Sellier, X. Wallart, S. Bollaert, A. Cappy, J. Chevrier, V. Bayot, and S. Huant. Imaging

electron wave functions inside open quantum rings. Phys. Rev. Lett., 99:136807, Sep 2007.

[28] M. G. Pala, B. Hackens, F. Martins, H. Sellier, V. Bayot, S. Huant, and T. Ouisse. Local density of states in mesoscopic samples from scanning gate microscopy. Phys. Rev. B, 77:125310, Mar 2008.

[29] N. Paradiso, S. Heun, S. Roddaro, L.N. Pfeiffer, K.W. West, L. Sorba, G. Biasiol, and F. Beltram. Selective control of edge-channel trajectories by scanning gate microscopy. Physica E: Low-dimensional Systems and Nanostructures, 42(4):1038–1041, 2010. 18th International Conference on Electron Properties of Two-Dimensional Systems.

[30] R. Crook, C. G. Smith, A. C. Graham, I. Farrer, H. E. Beere, and D. A. Ritchie. Imaging fractal conductance fluctuations and scarred wave functions in a quantum billiard. Phys. Rev. Lett., 91:246803, Dec 2003.

[31] M. R. Connolly, R. K. Puddy, D. Logoteta, P. Marconcini, M. Roy, J. P. Griffiths, G. A. C. Jones, P. A. Maksym, M. Macucci, and C. G. Smith. Unraveling quantum Hall breakdown in bilayer graphene with scanning gate microscopy. Nano Letters, 12(11):5448–5454, 2012.

[32] Henk van Houten and Carlo Beenakker. Quantum point contacts. Physics Today, 49(7):22–27, 1996.

[33] B. J. van Wees, H. van Houten, C. W. J. Beenakker, J. G. Williamson, L. P. Kouwenhoven, D. van der Marel, and C. T. Foxon. Quantized conductance of point contacts in a two-dimensional electron gas. Phys. Rev. Lett., 60:848–850, Feb 1988.

[34] N. Agrait, A. Levy Yeyati, and J. M. van Ruitenbeek. Quantum properties of atomic-sized conductors. Physics Reports, 377(2-3):81–279, 2003.

[35] D. A. Wharam, T. J. Thornton, R. Newbury, M. Pepper, H. Ahmed, J. E. F. Frost, D. G. Hasko, D. C. Peacock, D. A. Ritchie, and G. A. C. Jones. One-dimensional transport and the quantisation of the ballistic resistance. Journal of Physics C: Solid State Physics, 21(8):L209, 1988.

[36] M. Büttiker. Quantized transmission of a saddle-point constriction. Phys. Rev. B, 41:7906–7909, Apr 1990.

[37] S. Datta. Electronic transport in mesoscopic systems. Cambridge University Press, Cambridge, 1997.

[38] H. van Houten, C.W.J. Beenakker, and B.J. van Wees. Chapter 2: Quantum point contacts. volume 35 of Semiconductors and Semimetals, pages 9–112. Elsevier, 1992.

[39] M. A. Topinka, R. M. Westervelt, and E. J. Heller. Imaging electron flow. Physics Today, 56(12):47–52, 2003.

[40] S. E. J. Shaw, R. Fleischmann, and E. J. Heller. Quantum coherence beyond the thermal length, 2001. `arXiv:cond-mat/0105354`.

[41] S. E. J. Shaw. Propagation in Smooth Random Potentials. PhD thesis, Harvard University, 2002.

[42] G. Metalidis and P. Bruno. Green's function technique for studying electron flow in two-dimensional mesoscopic samples. Phys. Rev. B, 72:235304, Dec 2005.

[43] A. Cresti. Microscopic current imaging in quantum point contact devices. Journal of Applied Physics, 100(5):053711, 2006.

[44] C. Gorini, R. A. Jalabert, W. Szewc, S. Tomsovic, and D. Weinmann. Theory of scanning gate microscopy. Phys. Rev. B, 88:035406, Jul 2013.

[45] R. A. Jalabert, W. Szewc, S. Tomsovic, and D. Weinmann. What is measured in the scanning gate microscopy of a quantum point contact? Phys. Rev. Lett., 105:166802, Oct 2010.

[46] A. Freyn, I. Kleftogiannis, and J.-L. Pichard. Scanning gate microscopy of a nanostructure where electrons interact. Phys. Rev. Lett., 100:226802, Jun 2008.

[47] D. Weinmann, R. A. Jalabert, A. Freyn, G.-L. Ingold, and J.-L. Pichard. Detection of interaction-induced nonlocal effects using perfectly transmitting nanostructures. The European Physical Journal B, 66(2):239–244, 2008.

[48] A. Abbout. Thermoelectric transport in quantum point contacts and chaotic cavities: Thermal effects and fluctuations. PhD thesis, Université de Paris 6, 2011.

[49] A. Abbout, G. Lemarié, and J.-L. Pichard. Thermal enhancement of interference effects in quantum point contacts. Phys. Rev. Lett., 106:156810, Apr 2011.

[50] L. I. Schiff. Quantum mechanics. McGraw-Hill, New York, 3rd edition, 1968.

[51] E. W. Weisstein. Delta function. From MathWorld – A Wolfram Web Resource. Last visited on 13.07.2012.

[52] P. M. Morse and H. Feshbach. Methods of Theoretical Physics, Part I and II. McGraw-Hill, New York, 1953.

[53] J. J. Sakurai. Modern Quantum Mechanics (Revised Edition). Addison Wesley, 1 edition, 1993.

[54] P.A. Mello and N. Kumar. Quantum Transport in Mesoscopic Systems: Complexity and Statistical Fluctuations. Mesoscopic Physics and Nanotechnology. Oxford University Press, USA, 2004.

[55] L. E. Ballentine. Quantum Mechanics: A Modern Development. World Scientific, may 1998.

[56] P. Roman. Advanced quantum theory: an outline of the fundamental ideas. Addison-Wesley series in advanced physics. Addison-Wesley Pub. Co., Reading, Massachusetts, 1965.

[57] P.A. Mello, P. Pereyra, and N. Kumar. Macroscopic approach to multichannel disordered conductors. Annals of Physics, 181(2):290–317, 1988.

[58] P. A. Mello and J.-L. Pichard. Symmetries and parametrization of the transfer matrix in electronic quantum transport theory. J. Phys. I France, 1:493–513, Apr 1991.

[59] R. A. Jalabert and J.-L. Pichard. Quantum Mesoscopic Scattering: Disordered Systems and Dyson Circular Ensembles. Journal de Physique I, 5:287–324, March 1995.

[60] R. Landauer. Conductance from transmission: common sense points. Physica Scripta, 1992(T42):110, 1992.

[61] M. Yosefin and M. Kaveh. Conduction in curvilinear constrictions: Generalization of the landauer formula. Phys. Rev. Lett., 64:2819–2822, Jun 1990.

[62] H. A. Fertig and B. I. Halperin. Transmission coefficient of an electron through a saddle-point potential in a magnetic field. Phys. Rev. B, 36:7969–7976, Nov 1987.

[63] D. S. Fisher and P. A. Lee. Relation between conductivity and transmission matrix. Phys. Rev. B, 23:6851–6854, Jun 1981.

[64] V. Gasparian, T. Christen, and M. Büttiker. Partial densities of states, scattering matrices, and Green's functions. Phys. Rev. A, 54:4022–4031, Nov 1996.

[65] Lifshitz E. M. Landau L. D. Quantum mechanics. Non-relativistic Theory. Pergamon, New York, 3ed. edition.

[66] Y. Imry. Introduction to mesoscopic physics. Oxford University Press, Oxford, 2nd edition, 2002.

[67] H. van Houten and C. Beenakker. Quantum point contacts. Physics Today, 49(7):22–27, 1996.

[68] R.I. Shekhter G.B. Lesovik, D.E. Khmelnitskii. Reflectionless quantum transport and fundamental ballistic-resistance steps in microscopic constrictions. JETP Letters, 48:238–241, 1988.

[69] A. Yacoby and Y. Imry. Quantization of the conductance of ballistic point contacts beyond the adiabatic approximation. Phys. Rev. B, 41:5341–5350, Mar 1990.

[70] L. I. Glazman and M. Jonson. Cross-over to global adiabatics in 2d ballistic transport. Journal of Physics: Condensed Matter, 1(32):5547, 1989.

[71] M. C. Payne. Adiabaticity in quantum transport. Journal of Physics: Condensed Matter, 1(30):4939, 1989.

[72] A. Szafer and A. D. Stone. Theory of Quantum Conduction through a Constriction. Physical Review Letters, 62(3):300–303, January 1989.

[73] G. Kirczenow. Resonant conduction in ballistic quantum channels. Phys. Rev. B, 39:10452–10455, May 1989.

[74] B. J. van Wees, L. P. Kouwenhoven, E. M. M. Willems, C. J. P. M. Harmans, J. E. Mooij, H. van Houten, C. W. J. Beenakker, J. G. Williamson, and C. T. Foxon. Quantum ballistic and adiabatic electron transport studied with quantum point contacts. Phys. Rev. B, 43:12431–12453, May 1991.

[75] E. W. Weisstein. Weber differential equations. From MathWorld – A Wolfram Web Resource. Last visited on 14.10.2012.

[76] J. N. L. Connor. On the analytical description of resonance tunnelling reactions. Molecular Physics, 15(1):37–46, 1968.

[77] Confluent hypergeometric function. From Matlab Central File Exchange, a web resource. Last visited on 3.08.2013.

[78] C. Gorini. private communication.

[79] P. A. Lee and D. S. Fisher. Anderson localization in two dimensions. Phys. Rev. Lett., 47:882–885, Sep 1981.

[80] C. W. J. Beenakker and H. van Houten. Quantum Transport in Semiconductor Nanostructures. eprint arXiv:cond-mat/0412664, December 2004.

[81] P. E. Lindelof and M. Aagesen. Measured deviations from the saddle potential description of clean quantum point contacts. Journal of Physics: Condensed Matter, 20(16):164207, 2008.

[82] M. A. Eriksson, R. G. Beck, M. A. Topinka, J. A. Katine, R. M. Westervelt, K. L. Campman, and A. C. Gossard. Effect of a charged scanned probe microscope tip on a subsurface electron gas. Superlattices and Microstructures, 20(4):435 – 440, 1996.

[83] J. Vaníček and E. J. Heller. Uniform semiclassical wave function for coherent two-dimensional electron flow. Phys. Rev. E, 67:016211, Jan 2003.

[84] V. S. Vladimirov. Equations of mathematical physics. Pure and applied mathematics. M. Dekker, 1971.

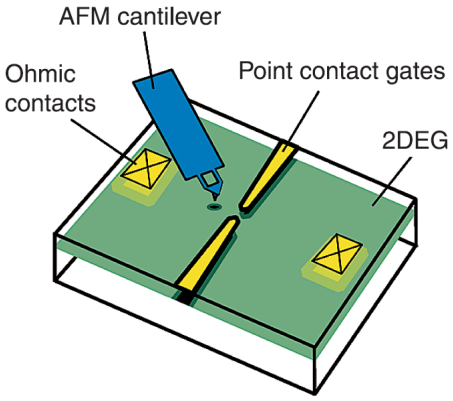
Résumé de la thèse (in French)

Introduction

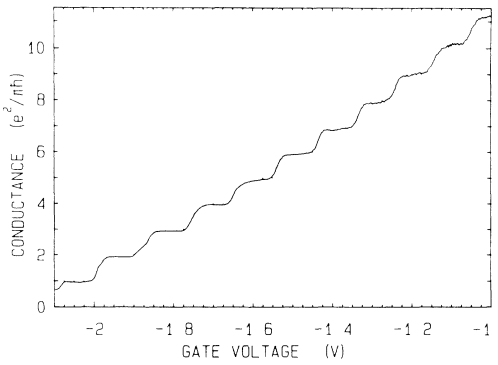
La recherche présentée dans cette thèse appartient aux domaines de la nanotechnologie et de la physique mésoscopique. La nanotechnologie concerne la manipulation de la matière à l'échelle allant de simples atomes jusqu'à des centaines de nanomètres dans au moins une dimension, tandis que la physique mésoscopique analyse les systèmes qui sont suffisamment petits pour produire des effets quantiques, mais assez grand pour ne pas avoir a considérer la structure microscopique.

L'objet de ce travail est la microscopie à grille locale (Scanning Gate Microscopy, SGM) et le transport électronique à travers le points quantiques (quantum point contacts, QPCs), étudié par cette technique. Dans une mesure SGM, la conductance à travers l'échantillon étudié est mesurée et enregistrée, tandis qu'une pointe de microscope à force atomique chargée est placée au-dessus de la structure, et produit un couplage capacitif [1]. Voir la figure B.1(a) pour le schéma. Le rôle de la pointe est de disperser les électrons de retour vers leur électrode d'origine et donc de changer la conductance. Les valeurs obtenues avec de nombreuses positions différentes de la pointe sont réunis pour créer une carte spatiale de la conductance $G(\mathbf{r}_0)$ - avec $\mathbf{r}_0$ la position de pointe - ou du changement de conductance $\Delta G(\mathbf{r}_0)$, si la valeur sans pointe actuelle est utilisée à titre de référence. Afin de maintenir la cohérence électronique, les mesures sont prises à basse température. D'autres paramètres, tels que le désordre de la nanostructure et la longueur d'onde de Fermi électronique sont également importants.

La technique SGM a été utilisé pour étudier divers systèmes mésoscopiques (nanotubes de carbone, points quantiques, anneaux de Aharonov-Bohm, bars de Hall, etc.). Parmi ceux-ci nous porterons une attention particulière aux contact quantiques définis dans un gaz d'électrons bidimensionnel de semiconducteur (two-dimensional electron gas, 2DEG). Un QPC est un passage étroit entre deux régions électriquement conductrices larges. La caractéristique la plus marquante de ces structures est la quantification de



(a) Représentation schématique du dispositif. Figure reproduite de [1].



(b) La quantification de la conductance dans un QPC. La tension de grille paramétrise la largeur W_{QPC} de QPC, de sorte que la courbe est équivalente à $G(W_{\mathrm{QPC}})$. Source : [33].

FIGURE B.1

la conductance, voir fig. B.1(b).

Les expériences SGM sur QPC révèlent de nombreuses caractéristiques intéressantes. A proximité de l'ouverture du QPC, le signal SGM forme de larges lobes lisses, dont le nombre correspond au nombre du plateau sur lequel la mesure est effectuée – voir fig. B.2 (a-b), tout en s'éloignant de l'ouverture *while away from the opening*, il se forme des branches étroites – fig. B.2 (c). Les images présentent des franges d'interférence, espacées de la moitié de la longueur d'onde de Fermi et d'étirement perpendiculairement aux branches ou lobes. L'origine des franges est l'interférence des ondes électroniques rétrodiffusés par la pointe, puis retransmises par le QPC vers la source, avec les ondes :

- réfléchie par la QPC – voir fig. B.3 (a),
- diffusée par la pointe deux fois – voir fig. B.3 (b),
- diffusée par une impureté – voir fig. B.3 (c).

Les franges crées par le dernier de ces mécanismes peuvent atteindre au-delà de la longueur thermique loin du QPC, mais disparaissent dans les échantillons à haute mobilité, tandis que les premier et second mécanismes donnent des franges visibles uniquement dans des échantillons de grande pureté à de basses températures. Le second mécanisme donne également un motif de damier dans les images.

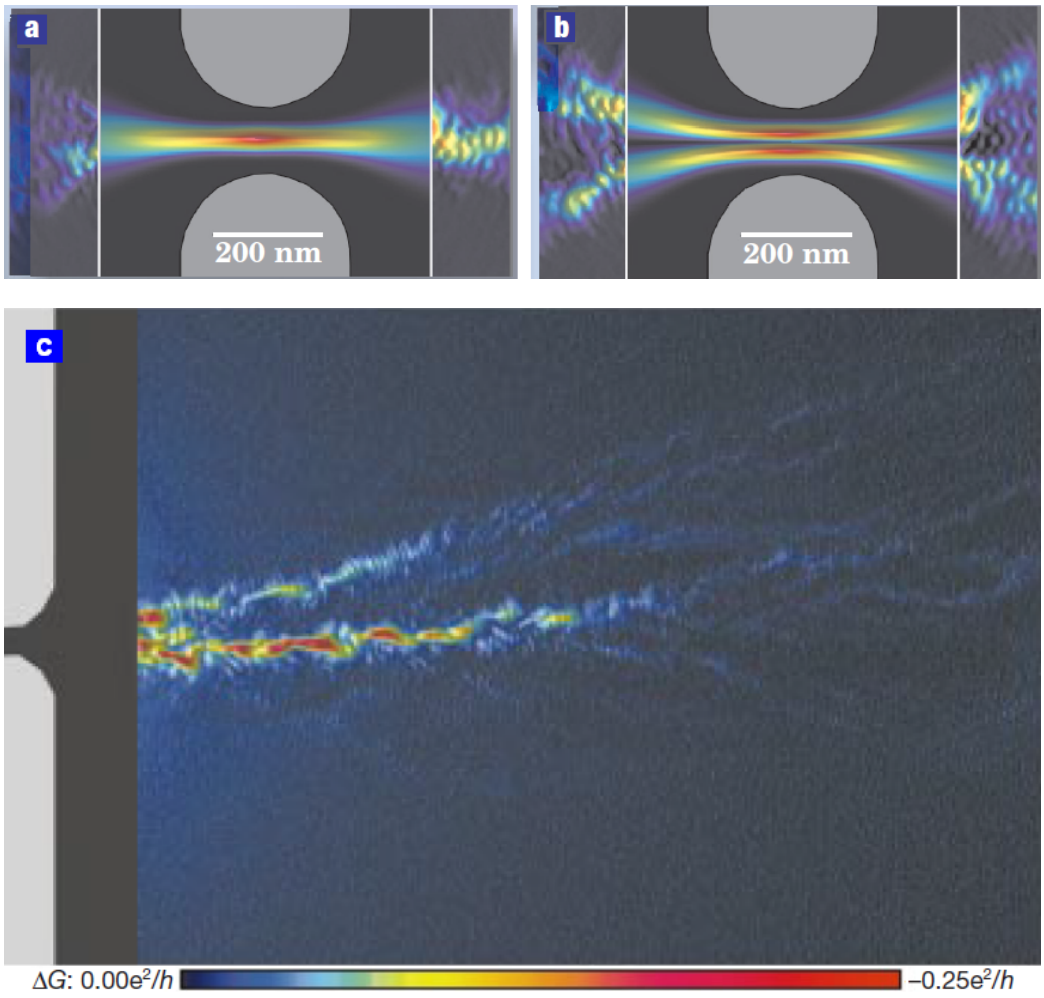


FIGURE B.2 – Images SGM proches de la QPC (a,b) et à des distances plus grandes de celle-ci (c). Les images (a,c) ont été prises sur le premier plateau de conductance. L'image (b) a été effectué sur le deuxième plateau, et le premier signal de commande soustrait de celle-ci. Les régions intérieures de (a) et (b), inaccessibles à la pointe de balayage, contiennent le carré absolu du premier (a) et seconde (b) fonctions d'onde du mode de constriction simulé. (De [39] et [6].)

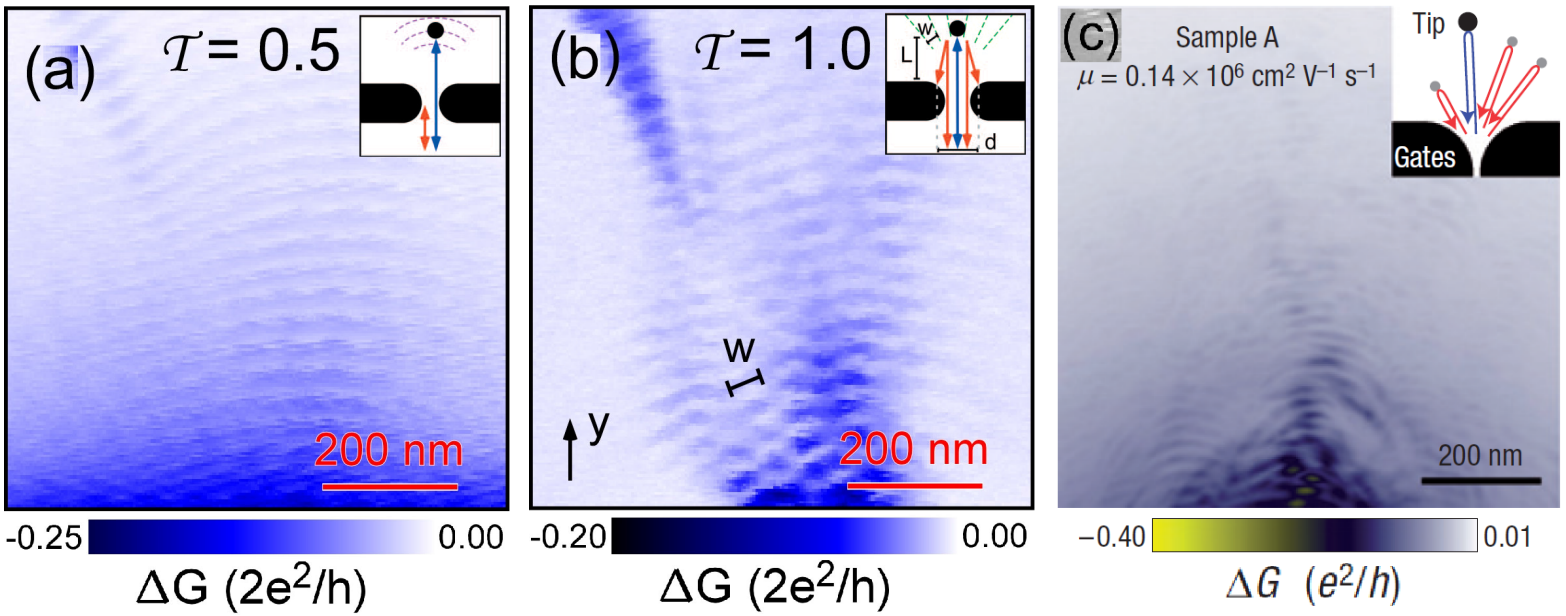


FIGURE B.3 – Cartes SGM avec différents types de motif de franges. Les encarts montrent le mécanisme de création de franges. (Adapté de [14] et [13]).

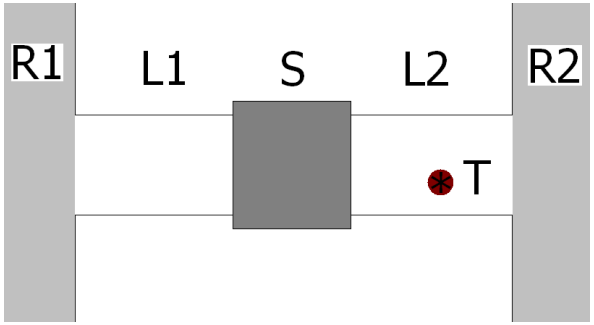


FIGURE B.4 – Les ingrédients d'un système SGM :
S – diffuseur ; L1, L2 – fils ; R1, R2 – réservoirs ; T – la pointe.
(La pointe peut être placé à l'intérieur du diffuseur aussi.)

Puisque le changement de conductance est proportionnel au flux incident la pointe de SGM, il a été proposé que la carte SGM est une image du flux d'électrons dans la structure [1]. Cette interprétation a été corroborée par un certain nombre de simulations [10, 12, 41–43]. D'autre part, certaines expériences sur de petits anneaux Aharonov-Bohm, également combinés avec des simulations, suggèrent un lien plus étroit du signal SGM à la densité d'états locale. Aucune théorie généralement applicable en donnant une interprétation univoque est disponible à ce jour. Dans une première étape vers une telle théorie, nous avons proposé une approche perturbative [44,45], pour fournir une interprétation des mesures de SGM non invasives, soit réalisée avec des potentiels de pointe faibles, en fonction des amplitudes de diffusion et des fonctions d'onde du système sans la pointe. Les expressions du premier et du second ordre dans le potentiel de la pointe ont été donnés. La condition de non-invasivité n'est très probablement pas atteinte dans les expériences SGM sur QPC publiés à ce jour, mais il n'est pas impossible d'employer des potentiels de pointe plus faibles [31], ou des pointes à l'extension spatiale réduite [25] dans les montages expérimentaux, et de développer une sonde véritablement non-invasive serait certainement désirable.

Le cadre théorique

Notre description théorique est basée sur le modèle de la masse effective et de particules indépendantes. Nous n'incluons pas les effets de température finie et le champ magnétique, mais une généralisation de ces cas est possible. Dans notre description, le système étudié est conceptuellement divisé en la nanostructure de notre intérêt, dénommé le diffuseur, idéales fils filiformes, les réservoirs et la pointe – voir fig. B.4.

Les fils sont confinés dans la direction transversale y, semi-infinie dans la direction longitudinale x et invariant à la translation le long de x. Les confinement selon y donne naissance à un ensemble de modes transversaux discrets avec des fonctions d'onde réelles $\phi_a(y)$ et énergies $\varepsilon_a^{(\perp)}$. Un électron à l'intérieur du fil gauche (étiqueté $l = 1$), qui propage en direction du diffuseur $(-)$ ou s'éloignant de celle-ci $(+)$, à l'énergie ε et en état transversal a, peut être décrite par la fonction d'onde stationnaire (mode de fil)

$$\varphi_{1\varepsilon a}^{(\pm)}(\mathbf{r}) = \frac{\exp(\mp i k_a x)}{\sqrt{2\pi\hbar^2 k_a/M_e}}\ \phi_a(y)\ , \tag{B.6}$$

où $\mathbf{r} \equiv (x, y)$, avec le nombre d'onde longitudinal $k_a > 0$ satisfaisant

$$\varepsilon = \varepsilon_a^{(\perp)} + \frac{\hbar^2 k_a^2}{2M_e}\ , \tag{B.7}$$

avec M_e la masse effective d'électron. Le nombre des modes propageant autorisées par le dessus est désigné par $N(\varepsilon)$. A l'intérieur du fil droit ($l = 2$), les signes $\mp$ dans l'exponentielle devraient être inversées. Les modes électroniques dans les fils portent le courant de $\frac{e}{h}$ par unité d'énergie chaque.

Tant que le diffuseur est une structure cohérente, un électron voyageant à travers notre système, sans pointe présente, peut être décrit avec une fonction d'onde qui s'étend à travers l'ensemble du système. Une base commode pour cette description est fournie par les états de diffusion sortants $\Psi_{1\varepsilon a}$, dont la forme asymptotique est donnée par

$$\Psi_{1\varepsilon a}(\mathbf{r}) = \begin{cases} \varphi^{(-)}_{1\varepsilon a}(\mathbf{r}) + \sum_{b=1}^{N} r_{ba}\varphi^{(+)}_{1\varepsilon b}(\mathbf{r}) \ , & x \ll -L/2 \\ \sum_{b=1}^{N} t_{ba}\varphi^{(+)}_{2\varepsilon b}(\mathbf{r}) \ , & x \gg \ L/2 \end{cases} \tag{B.8a}$$

pour l'électron provenant du réservoir gauche dans l'état (εa), et

$$\Psi_{2\varepsilon a}(\mathbf{r}) = \begin{cases} \sum_{b=1}^{N(\varepsilon)} t'_{ba}\varphi^{(+)}_{1\varepsilon b}(\mathbf{r}) \ , & x \ll -L/2 \\ \varphi^{(-)}_{2\varepsilon a}(\mathbf{r}) + \sum_{b=1}^{N(\varepsilon)} r'_{ba}\varphi^{(+)}_{2\varepsilon b}(\mathbf{r}) \ , & x \gg \ L/2 \end{cases} \tag{B.8b}$$

pour un électron originaire du réservoir droit dans l'état (εa). Ici, L est l'étendue longitudinale du diffuseur et $r^{(\prime)}$, $t^{(\prime)}$ sont les matrices d'amplitudes de réflection et de transmission au travers du diffuseur. Ensemble, ils forment la matrice de diffusion (S-matrice)

$$\mathbf{S} = \begin{pmatrix} r & t' \\ t & r' \end{pmatrix} \ . \tag{B.9}$$

La matrice S est unitaire et, dans les systèmes invariants par renversement du temps (TRI), symétrique. Elle peut être décomposée dans la représentation polaire

$$\mathbf{S} = \begin{pmatrix} u_1^{\mathrm{T}} & 0 \\ 0 & u_2^{\mathrm{T}} \end{pmatrix} \begin{pmatrix} -\mathcal{R} & \mathcal{T} \\ \mathcal{T} & \mathcal{R} \end{pmatrix} \begin{pmatrix} u_1 & 0 \\ 0 & u_2 \end{pmatrix} \ , \tag{B.10}$$

(cas TRI), où u_l sont unitaires, tandis que $\mathcal{R} = \mathrm{diag}(\mathcal{R}_m)$ et $\mathcal{T} = \mathrm{diag}(\mathcal{T}_m)$ contient les valeurs singulières de la matrice $r^{(\prime)}$ et $t^{(\prime)}$, respectivement. Nous allons utiliser les modes propres de transmission, entrants et sortants,

$$\overset{\approx}{\varphi}{}^{(-)}_{l\varepsilon m}(\mathbf{r}) = \sum_a [u_l]^{\dagger}_{am} \ \varphi^{(-)}_{l\varepsilon a}(\mathbf{r}) \ , \qquad \overset{\approx}{\varphi}{}^{(+)}_{l\varepsilon m}(\mathbf{r}) = \left[\overset{\approx}{\varphi}{}^{(+)}_{l\varepsilon m}(\mathbf{r})\right]^* \ , \tag{B.11}$$

et les *états de diffusion propres* (entrant seulement),

$$\overset{\approx}{\Psi}_{l\varepsilon m}(\mathbf{r}) = \sum_a [u_l]^{\dagger}_{am} \ \Psi_{l\varepsilon a}(\mathbf{r}) \ , \tag{B.12}$$

qui sont les vecteurs propres de la matrice de transmission $t^{(\prime)\dagger}t^{(\prime)}$.

Compte tenu des courants d'états de diffusion, nous pouvons voir que, sur chaque e/h porté par un mode électronique de fil ou un mode propre de transmission, on obtient, respectivement, $\frac{e}{h}\left[t^{(\prime)\dagger}t^{(\prime)}\right]_{aa}$ ou $\frac{e}{h}\mathcal{T}_m^2$, transmise vers le côté opposé du diffuseur. Tenant compte de la distribution énergétique des électrons provenant des réservoirs, nous pouvons écrire le courant total à travers la structure

$$I = \frac{2e}{h}\int_{\varepsilon_1^{(\perp)}}^{\infty} \mathrm{d}\varepsilon\ (f_1 - f_2)\ \mathrm{tr}\left(t^\dagger t\right)\ , \tag{B.13}$$

où f_l est la fonction de distribution de Fermi-Dirac dans le réservoir l. La conductance en réponse linéaire dans la forme sans dimension (soit dans les unités de $2e^2/h$) est alors

$$g^{(0)} = \mathrm{tr}(t^\dagger t)\ , \tag{B.14}$$

où les amplitudes de transmission doivent être évaluées à l'énergie de Fermi. Ce résultat, connu sous le nom de la formule de Landauer-Büttiker (à deux canaux), nous permet d'identifier la transmission totale et la conductance dans le régime de la réponse linéaire. Les amplitudes de transmission et de réflexion peuvent être évaluées grâce à la relation de Fisher-Lee [63] avec la fonction de Green retardée $\mathcal{G}_0$ correspondant au hamiltonien du système sans la pointe, H_0.

Théorie des perturbations de SGM

L'hamiltonien du système de SGM complet (fig. B.4) est une somme d'hamiltonien H_0 et potentiel de la pointe V_T, que nous allons traiter comme une petite perturbation. Les états de diffusion sortants du système complet $\chi_{l\varepsilon a}$ sont donnés sous une forme ouverte par l'équation de Lippmann-Schwinger

$$\chi_{l\varepsilon a}(\mathbf{r}) = \Psi_{l\varepsilon a}(\mathbf{r}) + \int \mathrm{d}\mathbf{r}'\ \mathcal{G}^{(0)}(\mathbf{r},\mathbf{r}',\varepsilon)V_\mathrm{T}(\mathbf{r}')\chi_{l\varepsilon a}(\mathbf{r}')\ , \tag{B.15}$$

dont l'application récursive conduit à la série de Born :

$$\chi_{l\varepsilon a}(\mathbf{r}) = \sum_{n=0}^{\infty} \chi_{l\varepsilon a}^{(n)}(\mathbf{r})\ , \tag{B.16}$$

où $\chi_{l\varepsilon a}^{(n)}(\mathbf{r})$ est la correction de l'état de diffusion à l'ordre n dans le potentiel de la pointe V_T. En utilisant de la décomposition en fonctions propres de la fonction de Green $\mathcal{G}_0$, on peut écrire chacune de ces corrections en termes des

états de diffusion non-perturbés $\Psi_{l\varepsilon a}$ et des éléments de matrice du potentiel de la pointe $[V_{\mathrm{T}}]^{l'l}_{a'a}(\varepsilon', \varepsilon)$.

De même, les densités de courant résultant en raison de chaque somme partielle de série (B.16) peuvent être liées aux éléments de matrice $[V_{\mathrm{T}}]^{l'l}_{a'a}(\varepsilon', \varepsilon)$ et les éléments de matrice de l'opérateur de densité de courant $[\mathbf{j}(\mathbf{r})]^{l\,l'}_{aa'}(\varepsilon, \varepsilon')$. Les différences de densité de courant par rapport à la densité de courant non-perturbée sont ensuite calculées. Intégrées sur la section transversale du fil et sommées sur toutes les modes disponibles à l'énergie de Fermi (et sur les deux états de spin possibles), elles donnent la correction au courant total et donc également la correction de la conductance, de l'ordre approprié dans le potentiel de la pointe, $g^{(n)}$. Des expressions jusqu'au deuxième ordre ont été obtenues.

Avec des potentiels de pointe suffisamment faibles (voir ci-dessous), la correction du premier ordre va dominer dans les régions où la conductance présente une marche, ou dans le cas où elle n'est pas du tout quantifié. Cette correction est donnée par l'expression

$$g^{(1)} = -4\pi \,\mathrm{tr}\Big[\mathrm{Im}\left(r^\dagger t'\, \mathcal{V}^{2\,1}\right)\Big] \,, \tag{B.17}$$

avec la notation abrégée $\mathcal{V}^{ll'}_{\cdot\cdot} \equiv [V_{\mathrm{T}}]^{ll'}_{\cdot\cdot}(\varepsilon, \varepsilon)$. Toutes les quantités sur le côté droit sont évalués à l'énergie de Fermi. Dans les régions des plateaux de la conductance quantifié, d'autre part, les termes $r^\dagger t'$ provoquent la correction $g^{(1)}$ à disparaître, donc la description perturbative complète nécessite la deuxième correction de l'ordre. Sa forme pertinente pour la région du plateau est

$$g^{(2)} = 4\pi^2 \,\mathrm{tr}\Big\{r'^\dagger r'\, \mathcal{V}^{2\,1}\mathcal{V}^{1\,2} \,- t^\dagger t\, \mathcal{V}^{1\,2}\mathcal{V}^{2\,1}\Big\} \,. \tag{B.18}$$

Ces deux formules peuvent être encore écrites dans la base des états de diffusion propres comme

$$g^{(1)} = 4\pi \sum_{m\in\not{\mathcal{M}}} \mathcal{R}_m \mathcal{T}_m \,\mathrm{Im}\left\{\mathcal{U}^{2\,1}_{mm}\right\} \,, \tag{B.19a}$$

$$g^{(2)} = -4\pi^2 \sum_{m,m'\in\underline{\mathcal{M}}} \left|\mathcal{U}^{2\,1}_{mm'}\right|^2 \,, \tag{B.19b}$$

avec

$$\mathcal{U}^{ll'} = \int \mathrm{d}\mathbf{r}\; \overset{\approx}{\Psi}{}^*_{l\varepsilon\cdot}(\mathbf{r})\, V_{\mathrm{T}}(\mathbf{r})\, \overset{\approx}{\Psi}_{l'\varepsilon\cdot}(\mathbf{r}) = u_l\, \mathcal{V}^{ll'} u^\dagger_{l'} \tag{B.20}$$

étant les éléments de matrice du potentiel de pointe dans la nouvelle base, et $\not{\mathcal{M}}$, $\underline{\mathcal{M}}$ désignent les ensembles d'indices de modes partiellement ouvertes et complètement ouvertes, respectivement. Si la courbe de conductance est

quantifiée, comme c'est le cas dans les contacts quantiques, $\not{\mathcal{M}}$ ne contient qu'un seul indice. Les expressions dans la base des modes propres de transmission sont également utiles, mais ne sont pertinents que dans les fils.

Une condition suffisante pour l'application des formules perturbatives est donnée par [65]

$$V_0 \Delta r_0 \ll E_F^{(1\mathrm{D})} \lambda_F^{(1\mathrm{D})} \ , \tag{B.21}$$

où V_0 est la valeur maximale du potentiel de pointe et Δr_0 est son diamètre effectif. Les quantités sur le côté droit, il faut comprendre comme l'énergie de propagation et la longueur d'onde de propagation dans un mode transversal donné. L'équation ci-dessus définit pour nous la définition du régime de non-invasivité.

Les expressions (B.19) montrent que la technique SGM dans le régime non-invasive sonde les éléments de matrice du potentiel de pointe entre les états de diffusion propres se propageant en sens inverse. Le signal sur le plateau contient la contribution de toutes les paires d'états propres ouverts, tandis que à la marche, le signal ne contient que la contribution des états venant de s'ouvrir.

Le maximum d'informations sur la structure des états de diffusion propres est contenue dans le signal SGM avec le potentiel de pointe parfaitement localisé

$$V_{\mathrm{T}}(\mathbf{r}) = U_0 \, \delta(\mathbf{r} - \mathbf{r}_0) \ , \tag{B.22}$$

où $\mathbf{r}_0$ est la position de la pointe. Le seul effet, que l'extension spatiale finie de la pointe a, est de bruiller le signal. Dans le cas de la localisation ideale, la réponse de SGM sur le premier plateau dans un système invariant par renversement du temps peut être écrite comme

$$g^{(2)} = -4\pi^2 U_0^2 \ \left[\sum_{m \in \underline{\mathcal{M}}} \left| \tilde{\tilde{\Psi}}_{l\varepsilon m}(\mathbf{r}_0) \right|^2 \right]^2 \ , \tag{B.23}$$

où le fil d'origine, l, est arbitraire. Cela signifie que les images SGM sur le premier plateau sont proportionnels au carré de la densité des particules à l'énergie de Fermi découlant des états ouverts seulement, ou de manière équivalente – de la densité du charge électrique dans une fenêtre entre deux potentiels chimiques des réservoirs opposés. Dans les systèmes réels, cette dépendance est modifiée par la localité non parfaite du potentiel de pointe et par la transmission non parfaite sur le plateau, c'est-à-dire $\mathcal{T}_m < 1$. De plus, si le système possède aussi une symétrie spatiale centrale et est accèdé par les contacts trés larges, son signal SGM en dehors de la nanostructure est proportionnel au même temps au carré de la densité de courant, sur chaque plateau, pas seulemnt le premier.

Le signal à l'intérieur des fils contient plusieurs oscillations, à moins que le diffuseur soit adiabatique. Dans ce cas, le signal du premier ordre (à la marche) contient une seule oscillation et le signal du deuxième ordre (sur le plateau) n'a pas du tout des oscillations. Les informations sur le degré de l'adiabaticité sont codées dans les matrices u_l, qui contrôlent la composition de l'état de diffusion propre en modes de fil.

L'application aux QPC

Plusieurs modèles de transport dans les QPC sont déjà disponibles dans la littérature et peuvent être adaptés pour être utilisés dans le cadre de notre théorie. Les modèles peuvent être classés en fonction du potentiel qu'ils utilisent pour décrire la constriction de QPC, qui peut être aux parois dures ou aux parois souples, et peuvent avoir différentes formes menant à différents niveaux de l'adiabaticité.

Le modèle point-selle de Büttiker, parfaitement adiabatique, donne la courbe de transmission composée des marches de forme des fonctions de distribution Fermi-Dirac, chaque marche correspondant à l'ouverture d'un mode électronique nouveau dans la constriction. Les états propres de transmission sont des fonctions x-y-séparables proposées par les produits des fonctions propres de l'oscillateur harmonique selon la direction transversale et les fonctions paraboliques de cylindre (fonctions de Weber) selon la direction longitudinale. Des conditions aux limites et les formules de connexion semi-classique permettent de déterminer les états propres originaires du fil gauche ou du fil droit. L'application de la formule (B.19) donne du signal SGM pour la conductance des points plateau-région étape et de la conductance, présentés dans la figure B.5.

Le modèle de QPC abrupt est non adiabatique. Il suppose que la constriction a les parois dures d'une forme rectangulaire. Le calcul des états propres de diffusion de ce modèle est rendu possible par l'introduction d'une nouvelle approximation, l'approximation du champ lisse (Smooth Field Approximation, SFA) [44]. Les cartes de SGM non invasives, évalués selon les hypothèses de fils larges et la distance entre l'ouverture de QPC et la pointe, r_0, dépassant au moins quelques longueurs d'onde de Fermi, sont affichés dans la figure B.6. Sous les mêmes hypothèses, on peut montrer que le signal de SGM sur le plateau, dans la QPC rectangulaire avec aucun désordre, peut être lié au carré de la densité de courant.

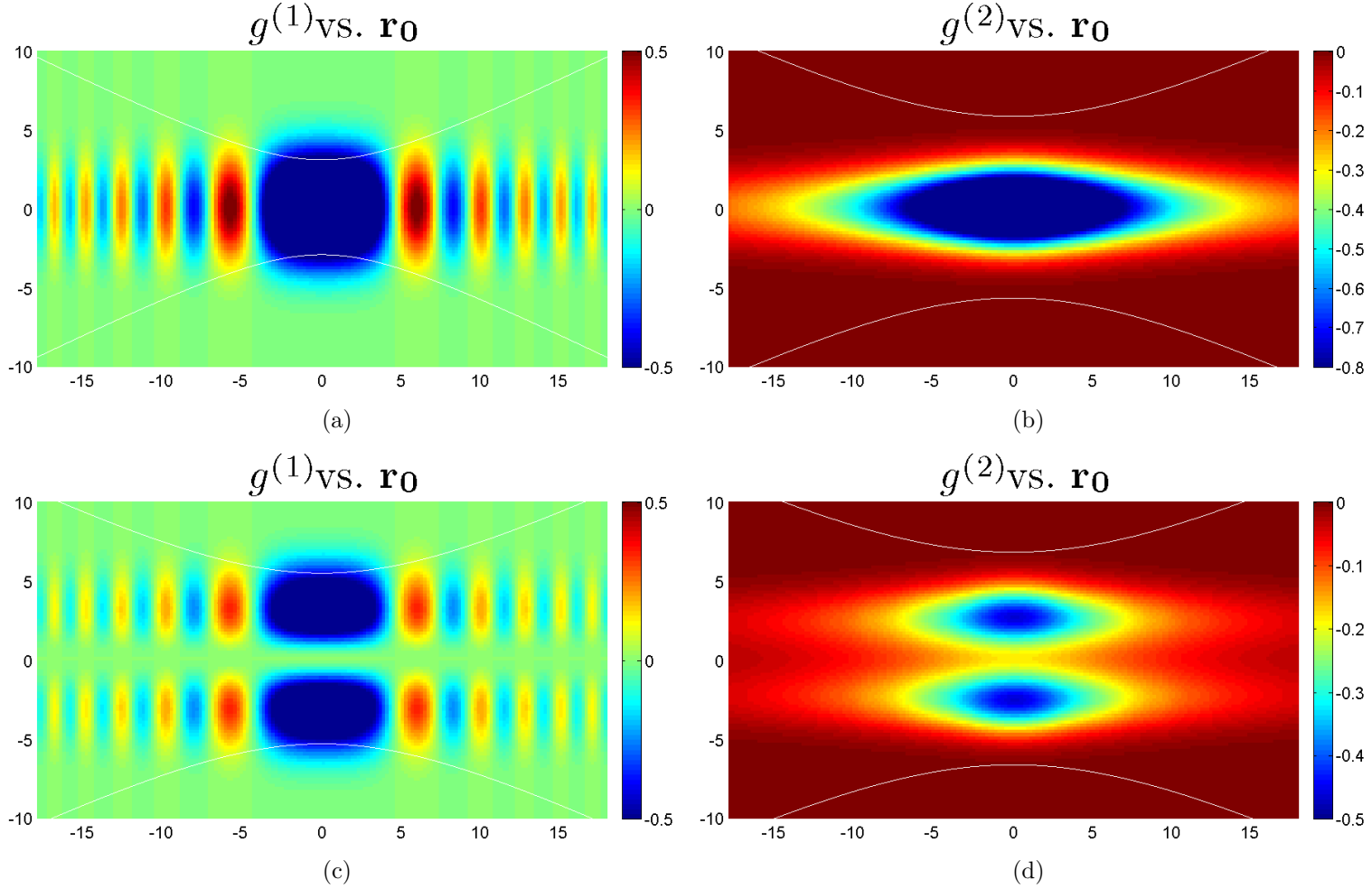


FIGURE B.5 – Corrections de conductance représentant le signal de SGM non-invasive :
(a) – point-selle QPC, première zone de l'étape de la conductance, (b) – point-selle QPC, première région du plateau de conductance, (c) – point-selle QPC, première zone de l'étape de la conductance, (d) – point-selle QPC, première région du plateau de conductance.

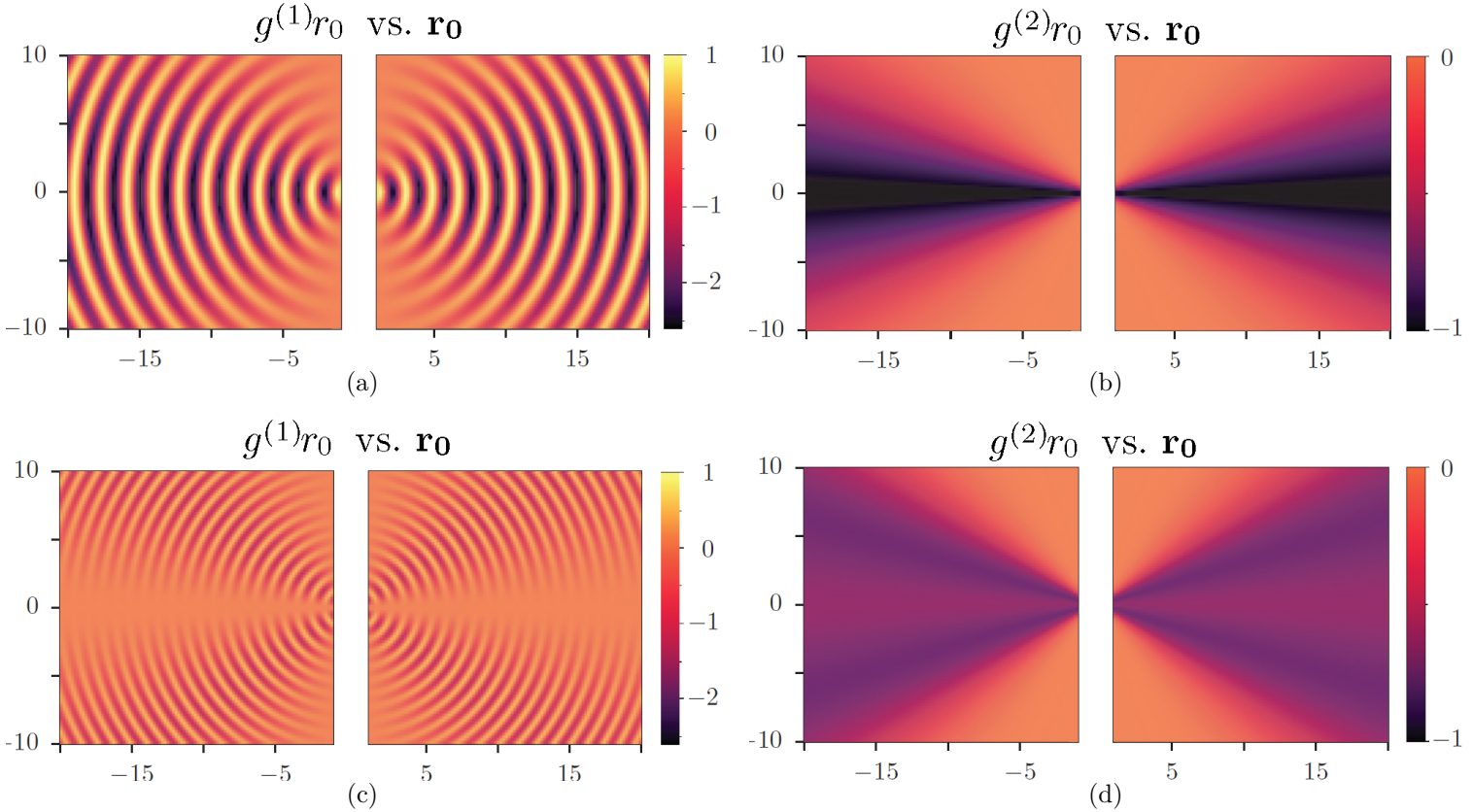


FIGURE B.6 – Corrections de conductance représentant le signal de SGM non-invasive :
(a) – QPC abrupt, première région d'étape de conductance (adapté de [44]), (b) – QPC abrupt, première région du plateau de conductance ([44]), (c) – QPC abrupt, deuxième région d'étape de conductance ([44]), (d) – QPC abrupt, deuxième région du plateau de conductance ([44]).

Simulations de SGM en QPC

Enfin, nous avons étudié quelques structures à parois dures de diverses adiabaticités par des simulations numériques. L'algorithme récursif de la fonction de Green sur un réseau serré de liaison a été utilisé pour évaluer la transmission avec et sans pointe.

Nous avons obtenu les courbes de transmission et de la correction SGM a la transmission, en fonction de l'énergie de Fermi, et les cartes spatiales SGM. La correction de conductance non invasive est représenté par les dérivés numériques de la transmission par rapport à la tension de pointe. Le signal d'intensité extrémité finie, d'autre part, représente la réponse SGM au-delà du régime perturbatif.

Nous avons constaté que l'effet de la pointe est le plus grand dans la région de marche de la conductance quantifiée et diminue comme le nombre d'onde de Fermi est augmenté le long du plateau. Les caractéristiques les plus importantes des cartes de SGM obtenus sont que le signal différentiel de premier ordre a un motif de franges régulier et s'annule lorsque le système est réglé sur le plateau de conductance, tandis que le signal différentiel de deuxième ordre n'oscille pas et est purement négatif – voir la figure B.7 (a-b). Deux effets diminuent la qualité de quantification de la conductance : l'effet tunnel et les vallées de résonance de Fabry-Perot. Le signal SGM sur le plateau non parfait possède un caractère mélangé, entre les signaux typiques du plateau et de la marche. Le désordre et la température finie diminuent aussi la qualité des plateaux de quantification. Si leur influence est modérée, ils peuvent améliorer le visibilité des franges.

Le principal effet de la force de la pointe fini (c'est-à-dire du régime invasif de SGM) est l'augmentation de la variation transversale du signal SGM – voir fig. B.7 (c-d). En particulier, il introduit le motif en damier au signal sur le plateau qui, dans le cas non-invasive, ne l'a pas. L'influence de la force de la pointe fini est beaucoup plus visible dans le signal de structures non-adiabatiques.

La non-adiabaticité manifeste très faiblement dans les courbes de transmission, mais a une influence significative sur les images SGM. Elle augmente leurs traits transversales d'une manière très similaire à celle de l'effet de la force de la pointe finie. Cela diminue la régularité des franges, et même peut, a quelques distances de l'ouverture de QPC, détruire le motif des franges.

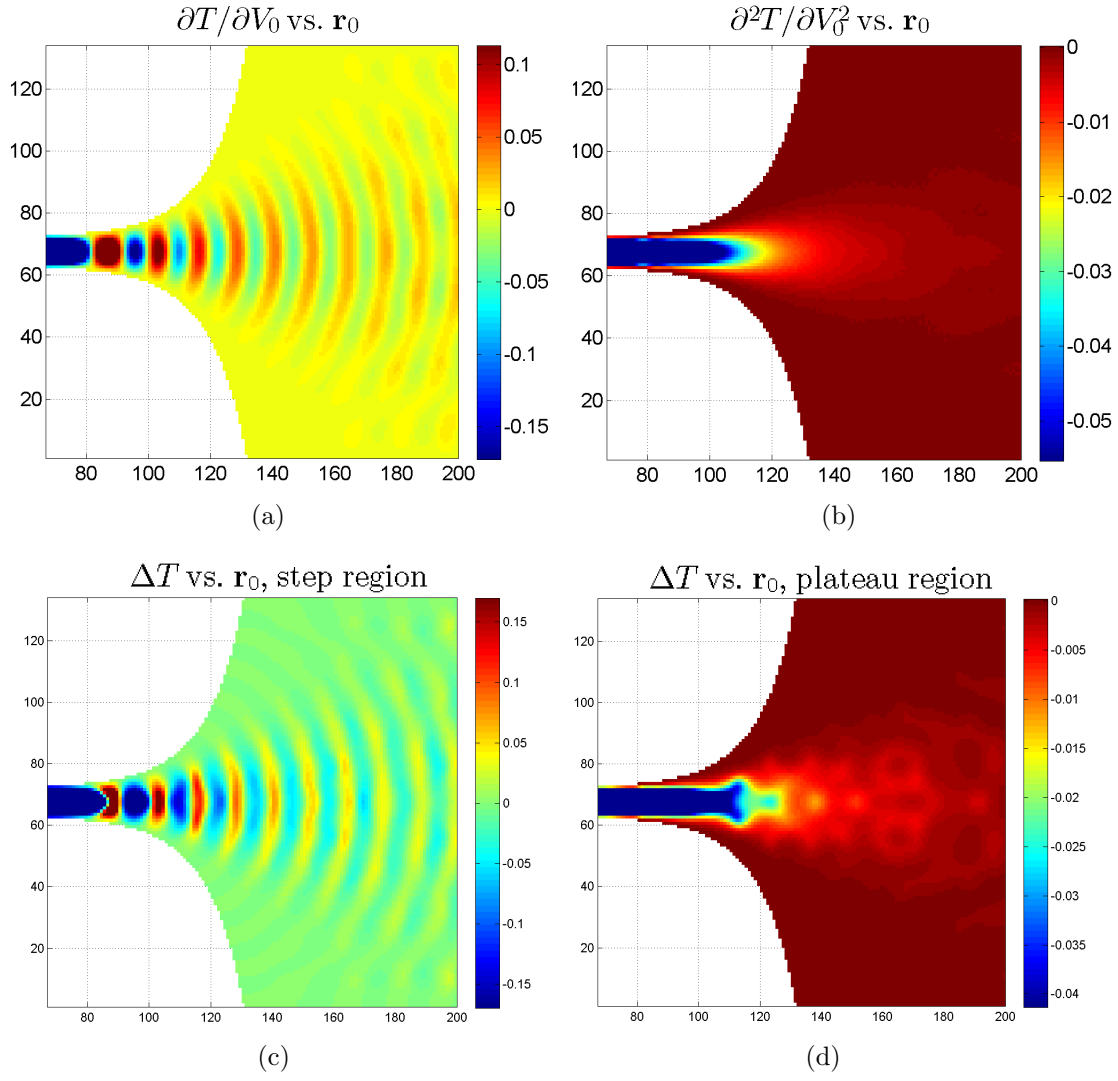


FIGURE B.7 – Les cartes SGM simulées :
(a) non-invasive, région d'étape de transmission ;
(b) non-invasive, région du plateau de transmission ;
(c) le potentiel de la pointe fini, région de l'étape de transmission ;
(d) le potentiel de la pointe fini, région du plateau de transmission.

Conclusions

Nous avons présenté une théorie de la microscopie à grille locale (SGM), applicable dans le régime non-invasive. Les corrections de conductance de première et de deuxième ordre dans le potentiel de pointe ont été données, pertinentes aux régions de la marche et du plateau de la conductance quantifié, respectivement. Le signal SGM ne peut pas être généralement lié aux quantités locales, comme la densité de courant ou la densité de charge. Une telle relation ne peut être trouvée que dans certaines situations idéalisées. Avec la qualité de plateau idéal et le potentiel de pointe parfaitement localisé, le signal sur les plateaux est lié à la densité de charge des états de diffusion ouverts à l'énergie de Fermi.

Nous avons obtenu et étudié les cartes spatiales, analytiques et numériques, du signal SGM dans les QPC. Dans les QPC avec les fils de connexion de largeur infinie, le signal sur les plateaux peut être lié à la densité de courant, et, au même temps, à la densité de charge. Nous avons trouvé les franges et motifs de damier. Les franges apparaissent dans le signal a la marche de conductance et aussi sur les plateaux, si ces-ci présentent une transmission non parfaite ($\mathcal{T}_m < 1$). Le motif de damier apparaît dans le signal sur les plateaux obtenu avec une pointe de tension finie. Nous avons également étudié les effets parasites introduites par l'effet tunnel et ;es résonances de Fabry-Pérot, et par le manque d'adiabaticité. Nous avons constaté que le manque d'adiabaticité est capable de détruire le motif de franges dans des structures ou des régions délimitées latéralement, aux distances de quelques longueurs d'onde de propagation de l'ouverture de QPC.

Notre théorie peut servir comme un bon cadre pour étudier les phénomènes induites par le désordre de nanostructure et induites par la forme de constriction. Aussi, elle peut servir comme un bon point de départ pour étudier le transport non-linéaire et les effets des interactions entre électrons.

Wojciech SZEWC

Theory and Simulation of Scanning Gate Microscopy

Résumé

Ce travail porte sur la description théorique de la microscopie à grille locale (SGM) et sur la résolution de modèles particuliers de contacts quantiques (QPC), analytiquement et numériquement. SGM est une technique expérimentale, qui mesure la conductance d'une nanostructure, lorsqu'une pointe de microscope a force atomique chargée balaye la surface, sans contacter cette derniere. Les images de SGM révèlent de nombreuses traits intéressants, tels que des lobes, des branches, des franges d'interférence et des motifs de damier. Aucune théorie généralement applicable, donnant une interprétation univoque, n'est disponible à ce jour. En utilisant la théorie de la diffusion de Lippman–Schwinger, nous avons développé une description perturbative de signal de SGM non-invasive. Les expressions du premier et du second ordre ont été données, se rapportant aux régions de marche et de plateau de la courbe de conductance. Dans les systèmes invariants par renversement du temps (TRI), adaptés au premier plateau de conductance, les images SGM sont liées à la densité de charge à l`énergie de Fermi. Dans un système TRI, avec une symétrie spatiale centrale et de très larges contacts, les images sont aussi liées à la densité de courant, quelque soit le plateau. Nous présentons et discutons les images calculées pour deux modèles analytiques de QPC et les images obtenues numériquement avec la méthode des fonctions de Green récursives, réproduisant certains motifs observés expérimentalement, et pointant les difficultés fondamentales a se bien positionner sur le plateau de conductance.

Mots-clés : physique mésoscopique, gaz bidimensionnel d'électrons, théorie du transport, conductance, microscopie à grille locale, contacts quantiques, théorie des perturbations, simulations numériques

Résumé en anglais

This work is concerned with the theoretical description of the Scanning Gate Microscopy (SGM) in general and with solving particular models of the quantum point contact (QPC) nanostructure, analytically and numerically. SGM is an experimental technique, which measures the conductance of a nanostructure, while a charged AFM tip is scanned above its surface. It gives many interesting results, such as lobed and branched images, interference fringes and a chequerboard pattern. A generally applicable theory, allowing for unambiguous interpretation of the results, is still missing. Using the Lippman-Schwinger scattering theory, we have developed a perturbative description of non-invasive SGM signal. First and second order expressions are given, pertaining to the ramp- and plateau-regions of the conductance curve. The maps of time-reversal invariant (TRI) systems, tuned to the lowest conductance plateau, are related to the Fermi-energy charge density. In a TRI system with a four-fold spatial symmetry and very wide leads, the map is also related to the current density, on any plateau. We present and discuss the maps calculated for two analytically solvable models of the QPC and maps obtained numerically, with Recursive Green Function method, pointing to the experimental features they reproduce and to the fundamental difficulties in obtaining good plateau-tuning which they reveal.

Keywords: mesoscopic physics, two-dimensional electron gas, transport theory, conductance, scanning gate microscopy, quantum point contacts, perturbation theory, numerical simulations