

Chapter 1

Introduction

1.1. Introduction and preliminary warning

Matter stability and the way in which rigid crystalline or amorphous arrays of atoms can be formed are ruled by two pillars of physics: electromagnetism and quantum mechanics; nothing else, provided that we admit the existence of elementary constituents such as atom nuclei without having to derive their internal structure from the first principles (then we need to add nuclear forces to our bunch of tools). The postulates and basic equations of these two theories can be written on a couple of pages, and everything can be derived from them¹. If the world was ruled by classical mechanics, it would simply be impossible to obtain stable atoms² or stable chemical bonding to ensure the existence of matter as we all experience it in our everyday life. Thus, it is something of a misnomer to say that we are going to study quantum devices as opposed to devices which would not be quantum. Everything is ruled by quantum mechanics, from the insulating or conducting character to the color of any piece of matter or object that you can see inside the room where you are now reading this introduction (see also Figure 1.1). To understand our macroscopic world, we often feel that once we admit the existence of stable matter, we can content ourselves with using the second Newton's law of motion and classical gravitational forces. An aeronautics engineer does not put too much quantum mechanics in his calculations, but this is certainly no longer the case

1 Of course with a substantial amount of hard work and mathematics, and adding some thermodynamics. Note also that if quantum mechanical predictions can be verified with an astonishingly high precision, their interpretation was (and is) the source of thousands of scientific articles and books.

2 Classical electrons accelerated over orbits radiate electromagnetic waves and thus lose energy. Thus, bound electrons would collapse onto the atoms.

if we want to justify the way in which electrons and therefore the electrical current behaves in a bulk semiconductor. Without a periodic atomic lattice and quantum mechanics, we could not find free electrons able to carry a current in a p - n junction, or in the channel of transistors which form the integrated circuits inside our computers. Thus, the reason why the devices under study in this book are called quantum is that we can straightforwardly apply to them the basic quantum effects that students are accustomed to calculating in an introductory quantum mechanics course.

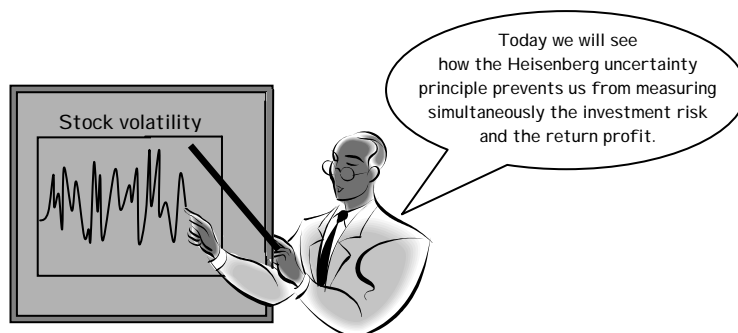


Figure 1.1. *The ubiquitous character of quantum mechanics*

In nanostructures, electrons can be confined in potential wells narrow enough to obtain energy quantization along the confining direction. Their dimension is small enough for probing the dual wave-particle nature of the electron in a straightforward manner, because the electron wave function phase can be kept coherent over the whole device length. Thus, it becomes possible to observe wave interference effects just by measuring the average current which can be passed through such components, and particle-like properties from current noise data. As once stated by the physicist Esaki, this looks like some kind of “do-it-yourself” quantum mechanics: you are not required to become a specialist in group theory and irreducible representations, or of field-theoretic methods to get in touch with the essence of the topic (see also Figure 1.2). In addition, other specific effects, although not quantum-mechanical, are also due to reduced dimensions: if you can inject a few electrons into a nanostructure and if the capacitance between this nanostructure and the rest of the world is very small, we can probe effects which are due to charge granularity (we cannot divide the electron charge), and which are known as Coulomb blockade. Such effects are the subject of intensive research in R&D laboratories, because many people hope to put them to good use to produce new types of memories and devices that are smaller, faster and require a smaller amount of operating power. The aim of this book is to give an introduction to the basic concepts which govern the conduction mechanisms taking place in such small devices.

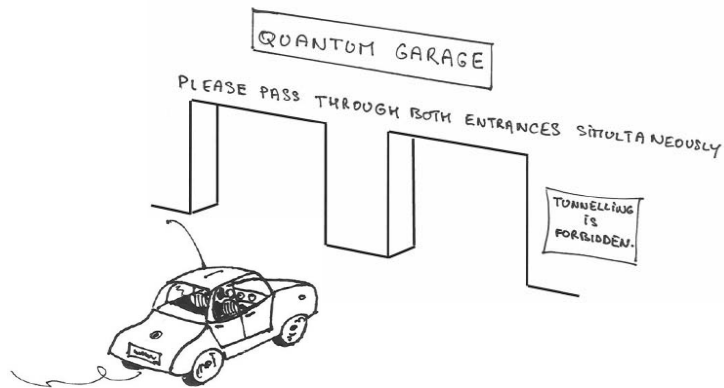


Figure 1.2. *The quantum garage*

Many (not to say most) of the phenomena described in this book usually take place at quite low temperatures, or in devices not yet (and for some of them never) used in the industry. The physics described here is not useful for understanding how industrial semiconductor devices behave in most applications right now, with the notable exception of resonant tunneling. Nevertheless, “today’s” silicon (*Si*) Metal-Oxide-Semiconductor Field-Effect Transistors (MOSFETs) definitely exhibit non-stationary and ballistic transport effects. Explaining these effects requires us to use some of the concepts developed in this book, even the high electric fields involved in MOSFET operation make the application of such concepts much more complicated than what is described in this introduction. At room temperature, the electron mean free path in silicon is in the 5-10 nm range, not far from the 45 nm channel length of the current CMOS technology, and integrated chips using a 32 nm process technology have already been demonstrated by the INTEL corporation in 2007. Figure 1.3 shows the picture of a 20 nm channel length prototype MOSFET produced in 2006 by LETI-CEA. Thus, even at room temperature some commercial electronic devices are close to the ballistic regime. Those *industrial* MOSFETs are fabricated with an incredibly high reproducibility in order to form extremely complex integrated circuits (and as a side note such precision and reproducibility are actually far from being achieved in most research laboratories working in the realm of mesoscopic physics and nanostructures, or with semiconductors more exotic and physically more appealing than silicon). Device-modeling based on ballistic properties has thus become an active research field, even in the case of silicon devices (see, e.g., [NAT 94] for one of the pioneering *Si* papers).

In addition, mesoscopic effects are important in four respects:

(i) they are often of great physical significance, and give a deep and straightforward insight into some of the most striking implications of quantum mechanics (for instance, they provide unambiguous and clear demonstrations of the dual electron nature, particle and wave);

(ii) although often obtained at low temperatures or high magnetic fields they are very useful for extracting physical parameters dealing with (nano)structures actually used in applications;

(iii) some of the effects are already used in (e.g. resonant tunneling) or potentially useful for (e.g. Coulomb blockade) applications;

(iv) although still difficult to engineer, devices made from graphene or carbon nanotubes exhibit truly ballistic and *quantum-coherent* effects even at room temperature. Thus, it is quite possible that not only ballistic, but also quantum-coherent effects may be present in electronic applications in the near future.

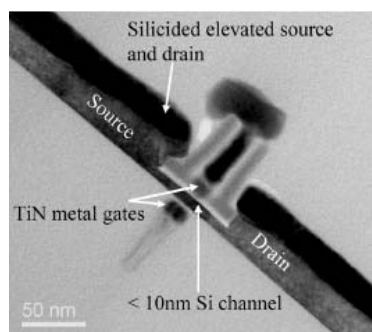


Figure 1.3. A transmission electron microscope view of a planar double-gate MOSFET fabricated by LETI-CEA with a 20 nm channel length; reproduced by permission after J. Widiez et al., *IEEE Transactions on nanotechnology*, vol. 5, p. 643 (2006), copyright ©2006IEEE ([WID 06])

As a consequence, in most of the largest semiconductor companies, and in a very large number of university labs, intensive research work is devoted to such structures. Scarcely applied though it may seem at first sight, this field of activity is in fact the leading edge of semiconductor research.

This book is designed to be accessible to the independent reader, and to students not having a strong background in solid-state physics (e.g. issued from engineering disciplines). As a matter of fact, this book is an attempt to answer the following question: what must be taught to students starting from scratch to make them

understand the bases of electron transport in mesoscopic devices? A professor placed in such a situation soon realizes that a good deal of solid-state physics and quantum mechanics is required. This explains the incorporation of chapters which are usually absent from the more specialized, already-existing books, and marks the difference between them and this. In addition, to follow the classification once given by J.M. Ziman, this book does not fall into the category of a “treatise” but into that of a “textbook”, with the purpose of introducing and explaining concepts. The text has been written with the aim of being as self-contained as possible, and is based on an oral course delivered at an international European master’s degree course involving three technical universities (GrenobleINP, EPFLausanne and Polit’oTorino). It is a deliberate choice of the author to keep in the book the spirit of the oral course, and this is the reason why the reader should not be surprised to be sometimes interpellated or hailed in a somewhat familiar way³.

Assimilating the quantum-mechanical rules summarized at the very beginning of the book suffices to derive any subsequent result, but should by no means be considered as enough to master quantum mechanics itself. Hence, and despite the fact that the text remains at an introductory level, a complete understanding of the course probably requires a minimum prior knowledge and self-maturation of the basic quantum-mechanical concepts. A reader not acquainted with this field will certainly feel the need to consult more authoritative manuals, due to the innumerable number of questions, either technical or fundamental, that a concise and *incomplete* presentation of quantum mechanics must arouse in any normally constituted mind. Some knowledge of solid-state and semiconductor physics certainly help as well, but all concepts useful for understanding the book can in principle be found in the book itself, and since this book is an introduction dedicated to a broad audience, maybe some of you are probably already acquainted with the required solid-state physics notions. For those who are experienced in solid-state physics it is possible to simply skip most of the reminders which make up Chapter 2. Besides, many of those reminders are not always quite rigorously demonstrated. All undemonstrated or heuristically-derived quantum-mechanical formulae can be found and are rigorously derived in a self-contained, encyclopedic textbook: [COH 77]. Solid-state physics has its self-contained book too: [ASH 76]. For *bulk* semiconductor physics and transport, an advanced and quite remarkable and complete textbook was written by [RID 82], but it is not essential for understanding this book. Eventually, we can find books specifically devoted to mesoscopic electron transport, which can be of great support for a better understanding or for gaining more information (the list below is not exhaustive): [BEE 91], [KEL 95], [DAT 95] and [FER 97]. The book which is the closest in spirit to this course is the one by Datta. It includes many exercises and also contains more advanced formalisms (e.g. Green’s functions) and discussions,

³ As you may have already noticed, the familiar way of addressing the reader began in the very first lines of this introduction.

which are not necessarily required at this introductory level. The book by Kelly presents a very large amount of data and also deals with aspects which are either more related to technological aspects or closer to the applications.

This book is an introduction and as such a number of important aspects have been omitted, mainly those which imply the use of mathematical concepts too involved to be developed in front of an audience new to the field. In particular, the reader will not find here a rigorous description of Green's function formalism, which is necessary to include electron-electron interactions in transport modeling. A general discussion and study of many-body effects is also absent, which would be mandatory to understand a physical phenomenon such as the fractional quantum Hall effect, metal-based mesoscopic devices, carbon nanotubes operating in the 1D form of a Luttinger liquid and many others. Justice has not been done to the electron spin and its possible applications. This book could thus be given a second title: how far can we go using only independent electrons and the exclusion Pauli principle (see also Figure 1.4)? Surprising though it may seem, a good deal of nanostructure physics can still be grasped that way, but the reader will not find in this book a wealth of phenomena associated with electron-electron interactions. If they are not discouraged by this introductory text their study should constitute the next step, to be achieved by studying more specialized treatises and articles. Thus, if after studying the various chapters the student decides to read further and deeper, the main objective of this book will have been fulfilled. In the same spirit, we shall skip some difficult demonstrations which would be required for a rigorous derivation of some important solid-state physics results⁴. However, even if difficult theoretical techniques have been deliberately banished from the text, "the language of physics is mathematics", and none of the chapters escape from the rule.



Figure 1.4. *The quantum society and Pauli's exclusion principle*

⁴ Whenever this occurs, the unsatisfied reader will always be left with the possibility of consulting the more advanced textbooks or specialized articles mentioned in the bibliography.

Most exercises proposed at the end of each chapter are easy and their purpose is to provide the reader with a means of checking that they have correctly assimilated the chapter content and concepts. However, some of them require more time, and have been inserted to complete points not detailed in the main text.

Not all the sections were dealt with during the original oral course. I have put indicators at the beginning of each section:

- This section is a reminder. Thus it can be skipped if the reader is already familiar with the corresponding field.
- This section is essential to the book (and, quite accessorially, it may be helpful to prepare an exam). Some reminders belong to this category.
- ⊕ This section is not a reminder, but is not considered as essential to understand the other parts.
- This section can be skipped at first reading.

1.2. Bibliography

- [ASH 76] ASHCROFT N.W. and MERMIN N.D., *Solid State Physics*, Wiley, New York, 1976.
- [BEE 91] BEENAKER C.W.J. and VAN HOUTEN H., *Quantum Transport in Semiconductor Nanostructures*, Solid State Physics 44, Academic Press, 1991.
- [COH 77] COHEN-TANNOUDJI C., DIU B. and LALOË F., *Quantum Mechanics*, Wiley, New York, 1977.
- [DAT 95] DATTA S., *Electronic Transport in Mesoscopic Systems*, Cambridge University Press, 1995.
- [FER 97] FERRY D.K. and GOODNICK S.M., *Transport in Nanostructures*, Cambridge University Press, 1997.
- [KEL 95] KELLY M.J., *Low-dimensional Semiconductors*, Oxford University Press, 1995.
- [NAT 94] NATORI K., “Ballistic metal-oxide-semiconductor field effect transistor”, *Journal of Applied Physics*, vol. 76, no. 8, 1994, p. 4879-4890.
- [RID 82] RIDLEY, B.K. *Quantum Processes in Semiconductors*, Clarendon Press, Oxford, 1982.
- [WID 06] WIDIEZ J., POIROUX T., VINET M., MOUIS M., DELEONIBUS S., “Experimental comparison between sub-0.1 μm ultrathin SOI single- and double-gate MOSFETs: performance and mobility”, *IEEE Transactions on Nanotechnology*, vol. 5, no. 6, 2006, p. 643-648.

Chapter 2

Some Useful Concepts and Reminders

2.1. Quantum mechanics and the Schrödinger equation ○

2.1.1. *A more than brief introduction*

The following is only a summary which includes the basic quantum-mechanical (QM) equations required for understanding the book. It is by no means a rigorous introduction to the topics, and if you want to go further, a wise thing to do would be to immerse yourself, e.g., in the introductory textbook by R.P. Feynman [FEY 65], and then in the book by Cohen-Tannoudji *et al.* [COH 77] for a while¹. Besides, several formulations can be used to describe quantum mechanics, and here we shall not really make the effort of differentiating them from one another. A concise description of those different formulations can be found in [STY 02].

In classical mechanics the elementary constituents of matter are massive point particles whose movement is controlled by electromagnetic or gravitational forces. At any instant we can precisely define the particle position and, provided that at a time t we are given the position and velocity of all the system particles, we can calculate everything at any other time, and obtain well-defined trajectories (with a powerful enough computer if the particles are numerous, etc.), even if the system remains isolated. Thus, the whole picture is *in principle* perfectly deterministic. In quantum mechanics the situation is far more subtle. Experimentally, it appears that if we let a system evolve isolated for a while, the maximum information concerning

¹ Of course these are not the only useful introductory QM textbooks, and the reader can also consult, e.g., [DIR 58], [SCH 68], [MES 62], [BOH 54], [MER 70], [LAN 65], among which many present a historical interest in addition to their scientific value, and there are many others.

this system that is physically accessible to human knowledge does not allow us to predict in a deterministic and unique way the result which will be obtained once we act on this system to measure some of its properties.

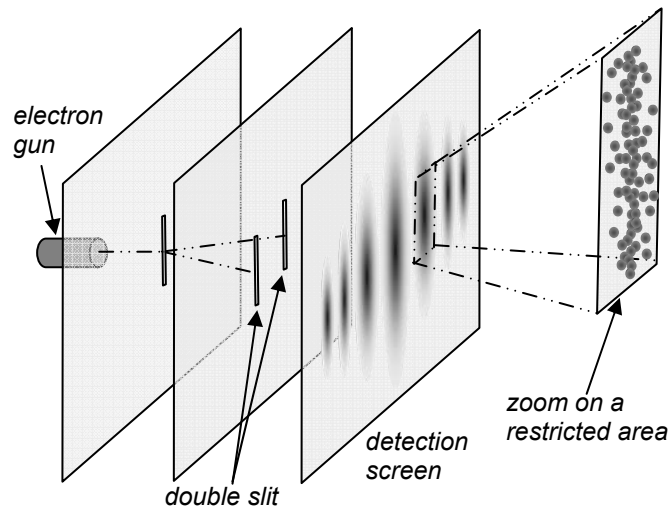


Figure 2.1. *Quantum-mechanical interference experiment illustrating the dual wave-particle nature of the electron*

The celebrated double-slit interference experiment is probably one of the more striking and meaningful illustrations of the quantum nature of matter. This effect figures in due place in almost any introduction chapter on quantum mechanics, and we shall respect this very justified habit. Interference experiments such as that illustrated by Figure 2.1 reveal that it is no longer possible to consider an entity such as an electron or a proton as a particle, and that it is not possible to consider it as a pure wave either [FEY 65]. “Identically prepared” electrons propagating through double slits exhibit interference patterns like waves [JON 61], but if we put a screen behind the plane of those slits we always obtain localized spots, as for particles [TON 89]. It is the statistical collection of a large number of such individual events which forms the interference pattern. Thus, in quantum mechanics (and in the real world) we have to assign a dual nature to electrons, whose behavior can be modeled only as a combination of both a particle and a wave. Suppress one slit and we lose the interference pattern. The wave really passes through both slits. Try to detect the electron at one of the two slits and we also lose the pattern, because the particle-like detection at one slit instantaneously reduces the extended propagating wave.

In some textbooks it is stated that quantum mechanics does not allow us, *even in principle*, to calculate any trajectory, and that it is a probabilistic theory *in essence*. This is not a correct statement, because one interpretation, known as the de Broglie-Bohm theory, gives a perfectly deterministic picture of quantum mechanics (at least for massive particles). In such an interpretation both a wave and a particle co-exist. The wave guides the particle, and in Bohm's version the *only* guiding rule states that the particle momentum is equal to the phase gradient of the complex wave obeying the Schrödinger equation multiplied by \hbar , a physical quantity known as the Planck constant [HOL 93]. In such a picture we can calculate well-defined trajectories (which are quite weird compared to classical ones, due to the action of the guiding wave). The unknown parameters, or "hidden variables", which make experiments exhibit a statistical aspect are nothing but the initial particle coordinates with respect to the wave. Thus, it is not possible to say from quantum mechanics that the basic facts of nature are undeterministic *in principle*. However, since this interpretation does not provide any new prediction with respect to the usual quantum rules, and exhibits the moral drawback of being obviously non-local, it has not attracted the favor of most physicists².

The "orthodox" interpretation of the quantum-mechanical formalism is that in between measurements we cannot precisely define a thing such as a particle; the experimental indeterminacy obtained when repeating the same experiment a large number of times with identically prepared systems results from the indeterminacy of nature itself, and not from a difference in system preparation which would be unknown to the observer. This was quite an incredible statement when quantum mechanics emerged, but it has now become a common "philosophical" view among scientists. In this book we shall not enter into those considerations any longer. We shall just use the quantum-mechanical rules, which up to now have always been experimentally validated with a numerical precision unprecedented by any other physical theory.

² Non-locality, or the possibility of instantaneous action-at-a-distance, does not agree with another pillar of modern physics, special relativity, so physicists are reluctant to accept it in other theories. In Bohm's interpretation experiments conducted on space-separated but correlated particles make this non-locality quite explicit. Note that in private most scientists would admit that quantum mechanics is deeply non-local, whatever the interpretation, but in public they see it as an unforgivable flaw, as soon as this non-locality is no longer buried in the mysteries of the "Copenhagen's school" or "orthodox" interpretation, and is given a straightforward meaning. Note that Bohm's interpretation also has its "weaknesses". For instance, if at some time t the particle distribution is given by the square modulus of the wave function, this implies that it will be true forever; however, there is nothing which rigorously justifies when and how it began to be so.

5.8. Fano resonance ●

Consider an electron prepared in a discrete state coupled to a continuum as in the previous section. Add to this system a bound state to which are coupled both the continuum and the discrete state. It turns out that the excitation spectrum resulting from the discrete state dilution into the continuum is the product of the interference between the two escape paths that can be followed by the electron, either through direct coupling or through an indirect coupling mediated by the bound state. The analytical shape of the spectrum is specific to the effect and thus constitutes a clear signature of its occurrence. For this reason Fano interference is without doubt one of the few essential landmarks with respect to quantum interference experiments ([FAN 35] and [FAN 61]), to be put on the same footing as the double-slit interference set-up, the Aharonov-Bohm effect or the weak localization phenomenon. In the double slit experiment, interference is evidenced through the spatially-dependent collected signal. In the Fano effect, interference is evidenced through the energy-dependent collected signal. It is worth noting that the discrete state need not be bound, so that the overall transmission through a device which incorporates in parallel both an open channel and a channel resulting from a resonant-tunneling state is a typical Fano resonance example. This explains why Fano profiles have in fact been obtained in a number of mesoscopic physics experiments, ranging from Aharonov-Bohm rings with a quantum dot in one arm to carbon nanotubes.

Consider a system similar to the discrete state coupled to a continuum as in section 5.7, but add to this another discrete state, which can be coupled either to the first discrete state with a matrix element

$$\langle \xi | W | \varphi_i \rangle = w_i \quad (5.68)$$

and to the continuum with a matrix element

$$\langle \xi | W | k \rangle = w \quad (5.69)$$

The problem can in fact be reduced to studying the transition probability between the discrete state $|\xi\rangle$ and the “new” continuum formed by the states resulting from the dilution of the discrete state $|\varphi_i\rangle$ into the continuum $|k\rangle$. Thus, we have to apply Fermi’s golden rule between the initial state $|\xi\rangle$ and the final state $|\psi_f\rangle$. From equation (5.67) the transition probability from the discrete state $|\xi\rangle$ to the continuum is proportional to the squared modulus of the matrix element

$$\langle \xi | W | \psi_f \rangle = \sum_k \langle k | \psi_f \rangle \langle \xi | W | k \rangle + \langle \varphi_i | \psi_f \rangle \langle \xi | W | \varphi_i \rangle \quad (5.70)$$

where the right-hand side has been obtained by using the decomposition of this final state over $|\varphi_i\rangle$, the $|k\rangle$ s and $|\xi\rangle$, and where we assumed as before that $\langle \xi | W | \xi \rangle = 0$. Equation (5.70) shows that the probability amplitude is now the sum of two complex amplitudes corresponding to the continuum and to the discrete state, respectively. By inserting the expressions given by equations (5.50) and (5.52) into equation (5.70), and taking equation (5.44) into account, we obtain

$$\langle \xi | W | \psi_F \rangle = \frac{w_i v + w(E_f - E_i)}{\sqrt{v^2 + \Gamma^2/4 + (E_f - E_i)^2}} \quad (5.71)$$

It is common practice to re-write this formula with reduced variables defined as

$$\varepsilon = \frac{2(E_f - E_i)}{\Gamma} \quad \text{and} \quad q = \frac{\delta E}{\pi v} \frac{w_i}{w} \quad (5.72)$$

where q is a term which measures the asymmetry degree between the coupling of the discrete state $|\xi\rangle$ to the evanescent bound state $|\varphi_i\rangle$ and to the continuum $|k\rangle$, respectively. This enables us to turn the square of equation (5.71) into

$$\left| \langle \xi | W | \psi_F \rangle \right|^2 \cong \frac{|\varepsilon + q|^2}{1 + \varepsilon^2}, \quad (5.73)$$

in which we neglected the term v^2 in the denominator of equation (5.71). Let us now examine this last expression in further detail: the magnitude of the numerator depends on the respective signs of the level energy and q , which may add or subtract, depending on whether the energy is below or above the resonance and the sign of q is positive or negative. The excitation profiles into the continuum are thus expected to reflect this asymmetry, and some particular cases are illustrated in Figure 5.20.

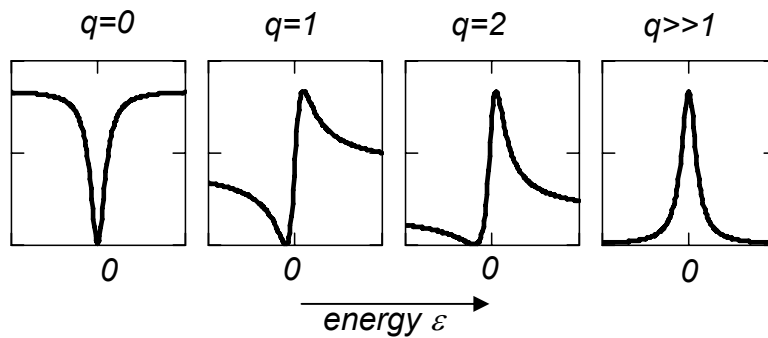


Figure 5.20. Fano profiles for different values of the coupling parameter q

For some limiting cases the shape of the spectrum can be either inverted (anti-resonance for $q=0$) or rendered symmetric ($q \gg 1$). This physical situation thus leads to a much larger variety of resonance spectra than the simpler phenomenon of resonant tunneling studied in section 5.5. It must also be stressed that an evanescent bound state can have an appreciable lifetime, so that observing such an effect in mesoscopic devices usually requires very low temperatures, to keep the wave function coherent over times longer than this lifetime.

5.9. Fano resonance in a quantum-coherent device ●

Here we are going to illustrate by a simple analytical example how a Fano resonance may occur in a quantum device. To make the analysis easier we shall restrict ourselves to the case of one input channel. A first ingredient is of course to define a structure in which the electron waves can follow two differentiated paths, and a second one is to include a resonant tunneling part. Consider a structure as in Figure 5.21, which incorporates both aspects. The electrons issued from the left contact can be transferred to the output either through the upper arm (which gives us a resonance equivalent to that of the bound state communicating with the continuum in section 5.8), or through the lower arm. First we have to define the S -matrix of the tunneling resonant structure. A typical matrix is⁴

$$S_u = \frac{1}{E + i\Gamma} \begin{bmatrix} E & i\Gamma \\ i\Gamma & E \end{bmatrix}. \quad (5.74)$$

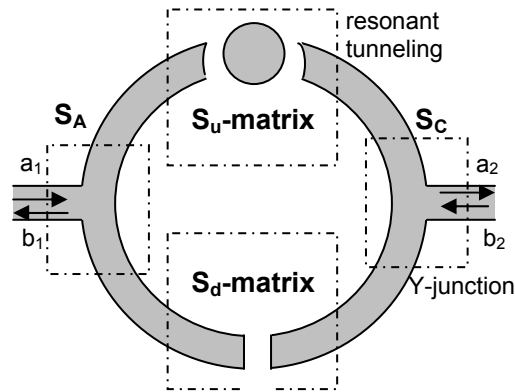


Figure 5.21. Quantum ring exhibiting a Fano-like resonance

⁴ Each time we introduce a new type of S -matrix in this section we should check that it is unitary and convince ourselves that it does correspond to what we are looking for. For instance here check that the transmission probability is formally similar to equation (5.23).

Note that here we use the symbol Γ to define an energy rather than an emission rate, so as to simplify the notations. The lower arm could be just a line inducing wave dephasing, the corresponding S -matrix being obviously of the form

$$S_d = \begin{bmatrix} 1 & e^{-i\theta} \\ e^{i\theta} & 1 \end{bmatrix} \quad (5.75)$$

However, it is more reasonable to also expect some wave attenuation in the lower arm, so that a possible S -matrix is

$$S_d = \begin{bmatrix} \sqrt{1-t_d^2} & it_d \\ it_d & \sqrt{1-t_d^2} \end{bmatrix} = \begin{bmatrix} r_d & it_d \\ it_d & r_d \end{bmatrix}. \quad (5.76)$$

where t_d is a number that is positive, real and lower than 1 (note that we did not choose an arbitrary phase as in equation (5.75), because the final result critically depends on it; this point is discussed later on). The S -matrices of the Y -junctions are given by equation (4.20). To calculate the overall transmission we can just apply the results that are thoroughly expounded in section 4.4, and in particular use the analytical transmission coefficients given by equation (4.25) or equation (4.26). First we voluntarily choose a Y -junction with a small transmission t , so that the full transmission can possibly correspond to the sum over two different paths, as explained in section 4.4, so that it is conceptually close to the principle of the Fano resonance. The various factors which enter into equation (4.24) are easily deduced from the matrix coefficients appearing in equations (5.74) and (5.76):

$$\alpha_u = \frac{4E}{E+i\Gamma}, \quad s_u = \frac{E-i\Gamma}{E+i\Gamma}, \quad \alpha_d = 2(1+r_d), \quad s_d = 1. \quad (5.77)$$

However, we see that factor α_u defined by equation (4.24) gives us a quantity which vanishes at the resonance. Thus the zeroth order approximation cannot be used, and we have to calculate the transmission using equation (4.26) instead of the simpler formula equation (4.25)! After inserting the expressions above into equation (4.26) and a few lines of algebra, we find an overall transmission coefficient

$$t_T = \frac{it^2t_d}{2(1+r_d)(1-t^2)} \frac{E + \frac{\Gamma(1+r_d)}{2t_d}}{E + i\frac{\Gamma t^2}{4(1-t^2)}}, \quad (5.78)$$

so that we obtain a transmission probability

$$|t_T|^2 = \lambda^2 \frac{|\varepsilon + q|^2}{1 + \varepsilon^2}, \quad (5.79)$$

where we have defined the reduced energy ε , the constant λ and the factor q as

$$\varepsilon = E \left(\frac{\Gamma t^2}{4(1-t^2)} \right)^{-1}, \quad \lambda = \frac{t^2 t_d}{2(1+r_d)(1-t^2)}, \quad q = \frac{2(1+r_d)(1-t^2)}{t_d t^2}. \quad (5.80)$$

Equation (5.79) has the form of a Fano resonance. However, with small transmission coefficients t and t_d , the factor q is clearly much larger than 1, and we do not expect a clear Fano-like resonance shape (see the case $q \gg 1$ in Figure 5.20). For this example it is therefore preferable to make the calculation for a large transmission factor t . This is however achieved at the price of more involved analytical results, and a higher conceptual difficulty, because we know that for a large t value the final transmission is the result of multiple scattering inside the two arms, which are not clearly separated. However, in this case we are going to see that it is indeed easier to obtain a clearly asymmetric resonance line shape. Inserting the parameters given by equation (5.77) into the general formula equation (4.23) leads after some tedious but easy algebra to

$$t_T = \frac{2it^2(2t_d E + \Gamma(1+r_d))}{\left((1-t^2)(5+3r_d) + 3\sqrt{1-2t^2} \right) E + \Gamma \left(t_d \left(1-t^2 - \sqrt{1-2t^2} \right) + i \times 2t^2(1+r_d) \right)} \quad (5.81)$$

Since there is an energy offset between the real parts appearing in the numerator and the denominator, we do expect a Fano resonance. This is readily observable in Figure 5.22, for which the transmission is plotted versus energy, with the

transmission in the lower arm as a parameter, and taking for the Y -junction transmission the maximum admissible value $t=1/2^{1/2}$.

Although it exhibits a Fano-like resonance, the case illustrated by Figure 5.22 is somewhat more complicated than the archetype expounded in section 5.8, due to the multiple scattering which occurs between both arms. Let us examine if by choosing a more reasonable matrix, i.e. putting some attenuation in the upper arm as well, we can clearly separate the transmission through both arms before realizing a quantum interference at the output and still obtain an outcome which exhibits a Fano-like resonance, even for small transmissions. Let us combine in series a resonant tunneling matrix as given by equation (5.74) and an “attenuation” matrix as given by equation (5.76). After some easy algebra, the usual combination rules (equation (4.15)) lead us to a new upper S -matrix of the form

$$S_u = \begin{bmatrix} \frac{E + i\gamma r_a}{E + i\gamma} & \frac{i\gamma t_a}{E - i\gamma} \\ \frac{i\gamma t_a}{E + i\gamma} & \frac{E - i\gamma r_a}{E - i\gamma} \end{bmatrix}, \quad (5.82)$$

where r_a and t_a are the reflection and transmission coefficients corresponding to an attenuation, and γ is equal to

$$\gamma = \frac{\Gamma}{1 - r_a}. \quad (5.83)$$

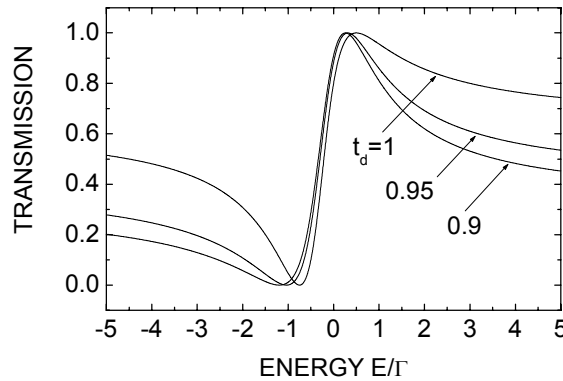


Figure 5.22. Fano-like resonance in a quantum ring with resonant tunneling in the upper arm, calculated according to equation (5.82) and with $t=2^{-1/2}$

The transmission is still resonant, but now the resonance magnitude is limited to t_a , instead of being equal to I . If we calculate the factor α_u defined by equation (4.24) from the coefficients appearing in S_u (equation (5.82)), we now find that it is equal to

$$\alpha_u = 2 \frac{2E^2 + \gamma^2(1 + r_a)}{E^2 + \gamma^2} \quad (5.84)$$

Provided that the attenuation is important (*i.e.* r_a is close to I), we see that $\alpha_u \approx 4$. Thus, the inverse quantity never exhibits a pathological divergence, and we can apply the zeroth order approximation found for the quantum ring transmission coefficient, to find

$$t_T \cong t^2 \left(\frac{t_u}{4} + \frac{t_d}{\alpha_d} \right) \quad (5.85)$$

Choose for the lower arm an attenuation S -matrix as given by equation (5.76). If the lower arm transmission is weak we have $\alpha_d \approx 4$ and we obtain a transmission

$$t_T \cong \frac{t^2}{4} \left(\frac{i(\gamma t_a + Et_d) - \gamma t_d}{E + i\gamma} \right). \quad (5.86)$$

Since in the numerator the imaginary part is shifted in energy we do find a Fano-like resonance component, to be added to a conventional resonant tunneling part. This can be re-written as

$$|t_T|^2 \cong \frac{t^4 t_d^2}{16} \left(\frac{1 + |q + \varepsilon|^2}{1 + \varepsilon^2} \right). \quad (5.87)$$

where the energy $\varepsilon = E/\gamma$ is normalized and the coefficient q is defined as $q = t_a/t_d$. This form is quite close to equation (5.73). It is instructive to see that here the q factor is really physically equivalent to the one found in the derivation given in section 5.8, since it is the ratio between the transmissions corresponding to the two interfering paths that can be followed by the electrons. This factor thus reflects the asymmetry between these two paths. In addition, by adjusting this transmission ratio, Fano resonance line shapes are clearly no longer restricted to large

transmission values. For this particular example, making $q=0$ leads to a constant transmission and not to an anti-resonance.

Note that the dephasing terms appearing in the S -matrix coefficients are essential to obtain a Fano-like resonance, which is of course not surprising since this phenomenon results from an interference effect. For instance, just consider a lower arm S -matrix S_d of the kind

$$S_d = \begin{bmatrix} \sqrt{1-t_d^2} & -t_d \\ t_d & \sqrt{1-t_d^2} \end{bmatrix} = \begin{bmatrix} r_d & -t_d \\ t_d & r_d \end{bmatrix}, \quad (5.88)$$

which only differs by transmission phase factors from equation (5.76), along with the upper matrix S_u defined by equation (5.74). As in the first case, we have to calculate the transmission coefficient using equation (4.26), and here we just quote the final outcome,

$$|t_T|^2 = t_u^2 \beta^2 \frac{\Gamma^2 + \alpha^2 E^2}{\Gamma^2 \beta^2 + E^2}, \quad (5.89)$$

where parameters α and β are defined as

$$\alpha = 1 + \frac{2t_d}{t_u(1+r_d)} \quad \beta = \frac{t^2}{4(1-t^2)}. \quad (5.90)$$

This transmission is always an even function of energy, and there is simply no Fano resonance at all! This explains why experimental data of Fano resonance are often obtained by varying a magnetic field through the ring, because the magnetic field value allows us to adjust the dephasing between the two arms.

5.10. Fano resonance in the real world \oplus

Fano resonance has now been observed in a number of quantum devices (see, e.g., [GOR 00], [KOB 02], [KOB 03], [KOB 04]), and here we reproduce data obtained from quantum rings incorporating a dot in one arm and purposely designed to exhibit and control a Fano resonance effect.

A scanning electron micrograph of the full device is shown in Figure 5.23. The conductance curves reproduced in the same figure very clearly exhibit asymmetric resonance peaks, and a Fano-like signature is obtained only when the upper arm transmission is not reduced down to zero, demonstrating thereby that an asymmetric shape is induced by the interference between the two arms. The shape of these peaks can be simply fitted by adjusting the value of the q factor. This shape varies qualitatively by passing from small to high q values. In addition, Fano resonance only occurs at the lowest temperatures, which are required to produce a large enough coherence length (see the conductance curves at the bottom of equation (5.28)). Note that the oscillation periodicity is due to the Coulomb blockade effect, which is the subject of a full chapter of this book.

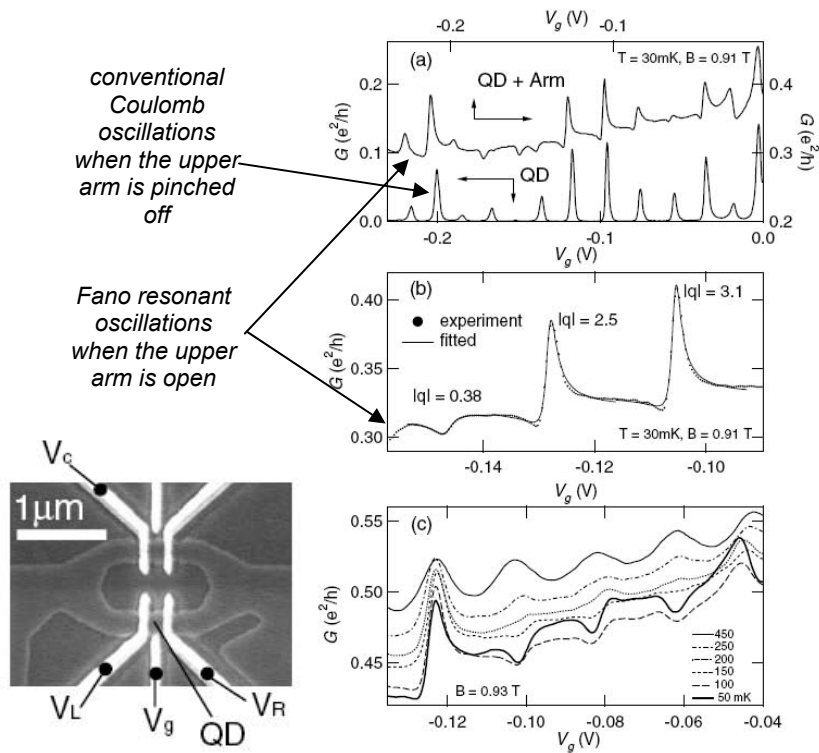


Figure 5.23. Quantum ring with a dot exhibiting a Fano-like resonance. V_c is used to control the upper arm transmission; reproduced with permission from K. Kobayashi et al., *Phys. Rev. Lett.* 88, 256801 (2002), copyright (2002) by the American Physical Society