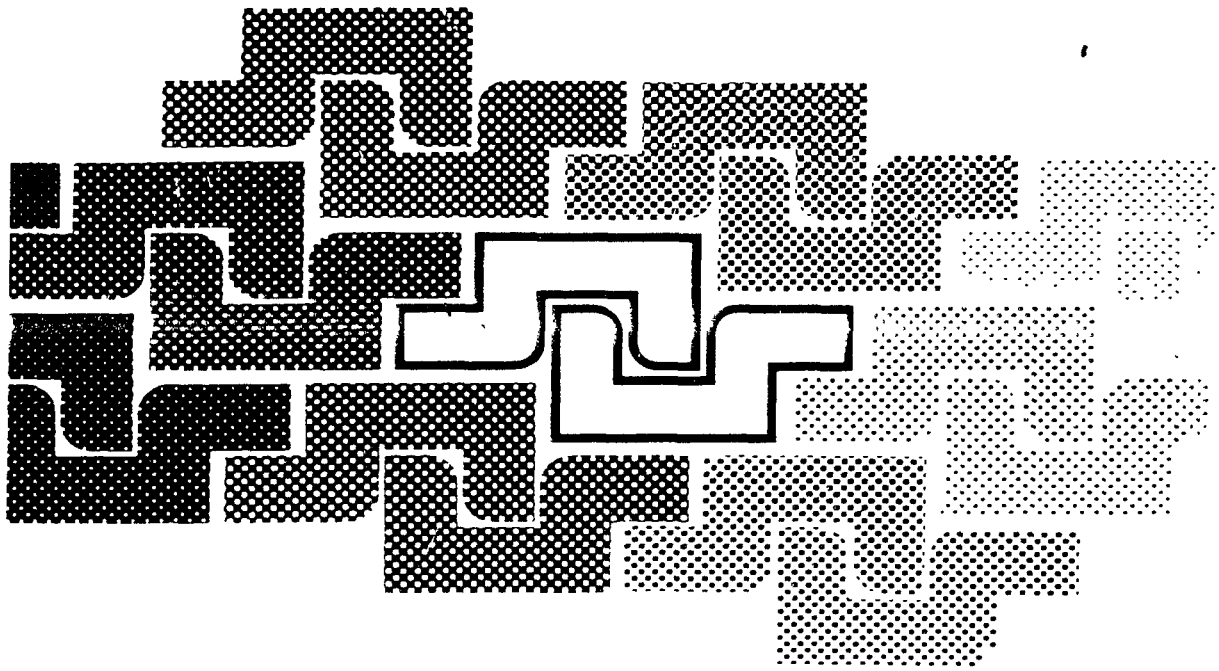
Principles of solid state electron optics

Beenakker, C.W.J.; Houten, H. van



Confined Electrons and Photons
New Physics and Applications

Edited by

Elias Burstein and
Claude Weisbuch

NATO ASI Series

Series B: Physics Vol. 340

PRINCIPLES OF SOLID STATE ELECTRON OPTICS

H. van Houten
Philips Research Laboratories
Prof. Holstlaan 4
5656 AA Eindhoven, The Netherlands

C.W.J. Beenakker
Instituut-Lorentz
University of Leiden
2300 RA Leiden, The Netherlands

1. INTRODUCTION

The science of vacuum electron optics has benefitted tremendously from the close analogy with light optics. This analogy exists on the level of classical motion (geometrical optics), as well as on the level of quantum mechanical motion (wave optics). The last two decades have witnessed a surge of interest in transport phenomena in low-dimensional semiconductor systems. Examples are the study of weak localization and conductance fluctuations in two-dimensional (2D) electron gases, resonant tunneling through confined states in quantum wells, transport through mini-bands in superlattices, and quantum ballistic transport through quantum point contacts. All of these phenomena have an optical analogue, and may be classified as manifestations of *solid state electron optics.*

In section 2 of this paper, we present the similarities in the fundamentals of optics and electron optics in vacuum, to prepare the ground for a discussion of the principles of solid state electron optics in section 3. Examples are discussed in section 4 and 5, which deal with ballistic transport through a quantum point contact, and with 2D refraction and (resonant) tunneling, respectively. The optical analogues of these phenomena are discussed as well. We chose these particular examples because of their relative simplicity, and because we wished to demonstrate how the quantum unit of conductance, e^2/h, appears in seemingly quite different transport phenomena (quantum ballistic transport and resonant tunneling). The common origin is the unit transmission probability of a single open scattering channel. The analogue for light scattering differs because e^2/h has no counterpart in optics. More precisely, the optical analogue of the conductance is the transmission cross section, which cannot be measured in units of fundamental constants (the velocity of light being the only one available).

Since this article is intended as a tutorial introduction, we have chosen to give a limited number of references to the original literature. A guide to the literature is provided in section 6

Confined Electrons and Photons
Edited by E. Burstein and C. Weisbuch, Plenum Press, New York, 1995

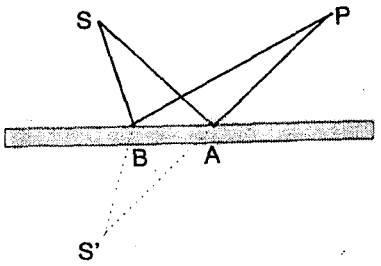


Figure 1. Specular reflection of a ray of light obeys a principle of least path length. The trajectory SAP, with equal angle of incidence and reflection, has the minimum length, as may be seen from the geometrical construction in this figure.

2. FUNDAMENTALS OF OPTICS AND ELECTRON OPTICS

2.1 Principles of Least Time and Action

A ray of light, propagating in a medium with a spatially varying index of refraction, or reflected by mirrors, may often be treated to sufficient accuracy by the laws of geometrical optics, which ignore the wave nature of light. This is analogous to the way in which classical mechanics is often a sufficiently accurate description of the motion of material particles in spatially varying potentials, or scattering elastically off a hard wall, even though the wave nature of matter is not taken into consideration. The search for the mathematical principles underlying the propagation of light and matter has intrigued scientists since classical antiquity. In those times, when calculus had yet to be invented, it may not have been as natural as it seems today to look for a principle governing the *local* dynamics of objects (as Newton succeeded in finding for material particles). Especially for light, it must have been quite natural to look for a principle governing the path traced as a whole. This is what was done by Hero of Alexandria, who wished to find an explanation for the equality of the angles of incidence and reflection for light incident on a mirror surface. In considering the possible paths that might be taken by a ray of light coming from a source at S, reflected at a mirror, and arriving at a point P, he hypothesized that the path actually taken is the shortest possible one. This *principle of least path length* indeed implies that the angle of incidence θ_i equals that of reflection θ_r, as may be proven by a simple geometrical construction (see Fig. 1).

Unfortunately, the minimum path length principle could not explain the refraction of a ray of light at the interface between two media of different optical density (such as air and water, see Fig. 2). This difficulty was removed in 1657, when Fermat introduced his famous *principle of least time*, which dictates that the actual path traced out by a ray of light is the one which takes the least time to complete. Since the velocity of light at position $\mathbf{r}$ in a medium with refractive index $n(\mathbf{r})$ is given by $v(\mathbf{r}) = c/n(\mathbf{r})$, Fermat's principle of least time may also be formulated as a principle of *least optical path length*

$$\int_S^P n(\mathbf{r})d\mathbf{r} = \text{minimum}$$

Since there exist situations where the actual optical path has a maximum rather than

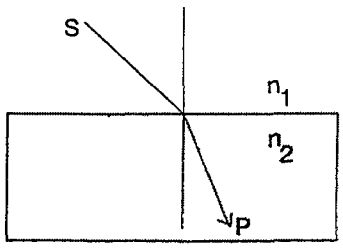


Figure 2: Refraction of light can be understood in terms of Fermat's principle of least time

a minimum length, it is more precise to express Fermat's principle as a variational one

$$\text{var} \int_S^P n(\mathbf{r}) d\mathbf{r} = 0, \tag{1}$$

stating that the optical path length is an extremum. Following common usage, we will still refer to Fermat's principle as the principle of least time. A derivation of Snell's law from Fermat's principle may be found in textbooks on optics [1].

In 1831, Hamilton formulated a principle of similar generality as Fermat's principle, but now for the mechanical motion of material particles in spatially varying potentials $V(\mathbf{r})$. Hamilton's *principle of least action* is the basis for formal treatments of classical mechanics. Each system is characterized by a function $L(\mathbf{r}, \mathbf{r}, t)$ called the Lagrangian. The general form of the Lagrangian can be constructed by considering the symmetries of the system[2]. Imagine a motion starting at t_1 and ending at t_2. One defines the *action* of the motion as

$$S = \int_{t_1}^{t_2} L dt. \tag{2}$$

According to Hamilton's principle, the path actually taken is the one which puts S at an extremum, so that $\text{var}S = 0$. From this variational principle one may derive Newton's equations of motion, describing the local dynamics of the system.

The analogy between Hamilton's principle (involving an integral over time) and Fermat's principle (involving an integral over space) may be made more explicit if one considers a single material particle with momentum $\mathbf{p}$ and kinetic energy $T = \frac{1}{2}\mathbf{p} \cdot \dot{\mathbf{r}}$ in a potential $V(\mathbf{r})$, for which the total energy $T + V(\mathbf{r})$ is a constant of the motion. The Lagrangian for this system is $L = T - V(\mathbf{r})$, so that

$$\text{var}S = \text{var} \int_{t_1}^{t_2} 2T dt = \text{var} \int_{\mathbf{r}_1}^{\mathbf{r}_2} \mathbf{p} \cdot d\mathbf{r},$$

where the integral is a line integral from $\mathbf{r}_1 = \mathbf{r}(t_1)$ to $\mathbf{r}_2 = \mathbf{r}(t_2)$. Hamilton's principle may thus be expressed as

$$\text{var} \int_{\mathbf{r}_1}^{\mathbf{r}_2} \mathbf{p} \cdot d\mathbf{r} = 0. \tag{3}$$

A comparison with Eq. (1) tells us that the path taken by a beam of classical particles in a potential $V(\mathbf{r})$ is analogous to that of a geometrical ray of light in a medium with refractive index $n(\mathbf{r})$, with the momentum playing the role of the refractive index. This analogy inspired Busch in 1925 to provide the first description of the focusing effects of electric and magnetic fields on a beam of electrons in optical terms[3]. Soon afterwards the electron microscope was invented, followed by other electron-optical instruments.

2.2 Huygens' Principle and Feynman Paths

The foundations of quantum mechanics were being completed just in time to support electron optics in its development as a successful branch of [illegible] physics [illegible] 1927. De Broglie[4] introduced his particle wave length h/mv, and in 1925 Schrödinger[5] presented his differential equation for the complex wave function, which describes the state of a non-relativistic particle at each instant of time. Also in 1925, it was suggested that the wave nature of particles might be demonstrated by studying the interaction of

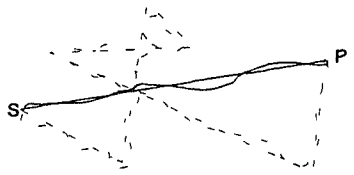


Figure 3. If the classically allowed path from S to P is the straight (full) line, neighboring paths have nearly the same classical action, so that they have little phase difference. For non-classical neighboring paths (dash-dotted lines) the action (and thus the phase) may differ strongly.

a beam of electrons with a single crystal[6]. Two years later Davisson and Germer[7] discovered (quite accidentally!) electron diffraction, and showed that the data were in agreement with the new theory.

In view of the analogy between geometrical optics and classical mechanics discussed above, it is natural to inquire whether a mathematical basis exists as well for the analogy between wave optics and quantum mechanics. In fact, such questions inspired the founding fathers of quantum mechanics to an extent that is perhaps not sufficiently appreciated today. The analogy was pushed furthest by Feynman, in his article on a "space-time approach to non-relativistic quantum mechanics"[8] This approach is related to Schrödinger's wave equation in a similar way as Hamilton's principle of least action is related to Newton's equations. Feynman introduced a complex probability amplitude $\phi[\mathbf{r}(t)]$, associated with a completely specified motion as a function of time $\mathbf{r}(t)$. This concept should be contrasted with Schrödinger's wave function $\Psi(t)$, which describes the state of a particle at a given instant of time. In Schrödinger's approach, the probability P to find the particle in a given region of space V at time t is given by $P = \int_V \mid \Psi(t) \mid^2 d\mathbf{r}$. Feynman's prescription treats space and time on equal footing:

to calculate the probability that a particle has a path lying in a given region of space-time R one should sum the amplitudes $\phi[\mathbf{r}(t)]$ for *all* paths in that region, and take the absolute square

$$P = |\sum_{\text{paths}\in R} \phi[\mathbf{r}(t)]|^2 . \tag{4}$$

As proposed earlier by Dirac[9], the amplitudes are of the form

$$\phi[\mathbf{r}(t)] = \text{constant} \times \exp\left(\frac{i}{\hbar} S_{cl}[\mathbf{r}(t)]\right) \tag{5}$$

where S_{cl} is the classical action for the path $\mathbf{r}(t)$

$$S_{\text{cl}}[\mathbf{r}(t)] = \int_{\text{path}} L dt. \tag{6}$$

Thus, the amplitudes carry equal weight, and differ in their phase only. Feynman's principle forms a wave-mechanical analogue of Huygens' principle in optics. The role of the phase $(\omega/c)\int n(\mathbf{r})d\mathbf{r}$ in optics is played by $(1/\hbar)S_{\text{cl}}[\mathbf{r}(t)]$ in mechanics.

The difference may be clarified as follows[10]. In optics, the frequency ω is a constant along any trajectory, so that the phase shift may be written as

$$\int_{\text{path}} \mathbf{k}\ \ d\mathbf{r} = (\omega/c)\int_{\text{path}} n(\mathbf{r})d\mathbf{r}, \tag{7}$$

with $k = \omega n/c$ the wave vector. In quantum mechanics, the energy E need not be a constant along the (possibly non-classical) path, so that the phase shift is

$$\frac{1}{\hbar}\int_{\text{path}} \mathbf{p}\cdot d\mathbf{r} - \frac{1}{\hbar}\int_{\text{path}} E dt = \frac{1}{\hbar}\int_{\text{path}} (\mathbf{p}\cdot\dot{\mathbf{r}} - E)dt. \tag{8}$$

For non-relativistic motion. and in Cartesian coordinates, one has $\mathbf{p}\cdot\dot{\mathbf{r}} - E = 2T - (T + V) = L$, so that the phase shift along the path equals $(1/\hbar)\int_{\text{path}} L dt = S_{\text{cl}}/\hbar$.

One of the most appealing aspects of Feynman's path integral formulation is that it gives insight in the connection between classical mechanics and quantum mechanics. In the classical limit $\hbar \to 0$, so that the phase factors of neighboring trajectories differ wildly — except for the classical paths, for which var $S_{\text{cl}} = 0$. Thus, one may imagine a classical path (obeying Hamilton's principle of least action) to be the result of the constructive interference between neighboring trajectories of constant phase, whereas non-classical paths are suppressed because of destructive interference (see Fig. 3). A similar connection exists between Fermat's principle of least time for geometrical optics and Huygens' principle in wave optics.

Feynman's path integral formulation is completely equivalent to the Schrödinger equation. The connection may be established by defining the wave function $\Psi(\mathbf{r}, t)$ with initial condition $\Psi(\mathbf{r}, 0) = \delta(\mathbf{r})$ as the sum or path integral of the complex amplitudes $\phi[\mathbf{r}(t)]$ over all paths with $\mathbf{r}(0) = 0$ and $\mathbf{r}(t) = \mathbf{r}$,

$$\Psi(\mathbf{r}, t) = \sum_{\mathbf{r}(0)=0}^{\mathbf{r}(t)=\mathbf{r}} \exp\left(\frac{i}{\hbar} S_{\text{cl}}[\mathbf{r}(t)]\right). \tag{9}$$

2.3 Wave Equations

Let us finally examine the analogy between optics and electron optics from the wave equation point of view The Schrodinger equation for an electron in a potential $V(\mathbf{r})$ reads

$$-\frac{\hbar^2}{2m}\nabla^2\Psi(\mathbf{r},t) + V(\mathbf{r})\Psi(\mathbf{r},t) = \imath\hbar\frac{\partial\Psi(\mathbf{r},t)}{\partial t}, \tag{10}$$

which reduces to a stationary wave equation on substituting a mono energetic wave $\Psi(\mathbf{r},t) = \Psi_0(\mathbf{r})\exp(-\imath Et/\hbar)$

$$\nabla^2\Psi_0(\mathbf{r}) = -\frac{2m}{\hbar^2}\left(E - V(\mathbf{r})\right)\Psi_0(\mathbf{r}) \tag{11}$$

This implies a *quadratic* dispersion relation

$$E - V = \frac{\hbar^2 k^2}{2m} \tag{12}$$

for a plane wave $\Psi_0 \propto \mathrm{e}^{\imath\mathbf{k}\,\mathbf{r}}$ in the case of a slowly varying potential

In contrast, the wave equation for the electric field $\mathcal{E}(\mathbf{r},t)$ of an optical wave in a medium with refractive index $n(\mathbf{r})$ is second order in time [1]

$$\nabla^2\mathcal{E}(\mathbf{r},t) = \frac{1}{\epsilon_0 c^2}\frac{\partial^2}{\partial t^2}\left(P(\mathbf{r},t) + \epsilon_0\mathcal{E}(\mathbf{r},t)\right) \tag{13}$$

For a monochromatic wave $\mathcal{E}(\mathbf{r}\ t) = \mathcal{E}_0(\mathbf{r})\exp(-\imath\omega t)$ in a linear medium with polariza tion $P(\mathbf{r},t) = (n(\mathbf{r})^2 - 1)\epsilon_0\mathcal{E}(\mathbf{r},t)$ this reduces to the Helmholtz equation

$$\nabla^2\mathcal{E}_0(\mathbf{r}) = -\left(\frac{n(\mathbf{r})\omega}{c}\right)^2\mathcal{E}_0(\mathbf{r}) \tag{14}$$

which implies a *linear* dispersion relation

$$\omega = \frac{ck}{n} \tag{15}$$

Electron waves and light waves thus obey similar stationary wave equations, (11) and (14) A comparison of these equations tells us again that electron wave optics is similar to light wave optics if we treat the momentum $(2m(E-V))^{1/2}$ as the refractive index

2.4 Limitations of the Analogy

The analogy between electron and wave optics is not a perfect one In this sub-section we briefly discuss some fundamental limitations of the analogy

The different dispersion relations (12) and (15) imply that an electron has a wave length $\lambda = 2\pi/k$ that is inversely proportional to its velocity ($v = \mathrm{d}E/\hbar\mathrm{d}k = h/m\lambda$, whereas a photon has a wavelength that is directly proportional to its velocity $v = c/n = \lambda\omega/2\pi$ As a result, Snell's law, expressed in terms of phase velocities, reads differently for electrons and photons (see section 5)

Further differences were discussed by Ehrenfest[11] He pointed out that the electron wave function Ψ is not an observable quantity, whereas the electric field $\mathcal{E}$ or the magnetic field B is An implication is that Ψ is an intrinsically complex quantity, but

[1] In deriving Eq (13) from Maxwell's equations a term containing a spatial derivative of $n(\mathbf{r})$ has been neglected This is justified if n varies more slowly than $\mathcal{E}$, i e slow on the scale of the wave length

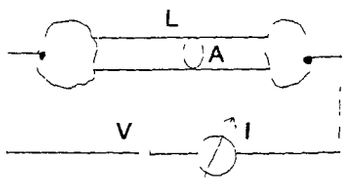


Figure 4. A set-up for the measurement of the two-terminal conductance of a wire of length L and cross-sectional area A.

$\mathcal{E}$ is real. (When one writes $\mathcal{E} = \mathcal{E}_0 e^{i\mathbf{k}\cdot\mathbf{r}}$ one really means $\mathcal{E} = \mathrm{Re}[\mathcal{E}_0 e^{i\mathbf{k}\cdot\mathbf{r}}]$.) For light one could just as well have opted to work exclusively with real quantities, but this is impossible for electrons. A second difference mentioned by Ehrenfest is that $\mid \Psi \mid^2$ is a probability density, but $\mid \mathcal{E}^2 + B^2 \mid$ an energy density. If there is only a single energy in the problem, then the energy density is also a probability density (apart from a proportionality factor). However, for non-monochromatic light, it is impossible to obtain from $\mid \mathcal{E}^2 + B^2 \mid$ the probability that one may find a photon in a given region of space-time.

Feynman has pointed out[8] that –whereas his formulation of quantum mechanics is exact– Huygens' principle is not. The reason is that the optical wave equation is second order in time. In an exact theory of optical waves it is necessary to specify the derivative of the wavefunction (in addition to its amplitude and phase) on a given wavefront, to be able to predict its further evolution in space and time. This is known as Kirchoff's modification of Huygens' principle.

It is possible to construct an approximate wave equation for light which is more closely analogous to Schrödinger's equation because it is also first order in time. This may be done using the slowly-varying-envelope approximation[12]. There are some interesting analogies that may be fruitfully discussed in terms of this Schrödinger equation for light, one example being the analogy between Andreev reflection of electrons at normal metal-superconductor interfaces and optical phase conjugation[13, 14].

Additional differences exist, such as the different statistics for electrons and photons, but these require a discussion beyond the level of the Maxwell equations.

3. PRINCIPLES OF SOLID STATE ELECTRON OPTICS

The main theme of this paper is the idea that transport of conduction electrons in the solid state can in many different regimes be treated as a form of electron optics In this section, we discuss the basic principles which justify such a treatment.

3.1 Electrical Conduction in Linear Response

An elementary electrical circuit consists of a conductor connected via a pair of contacts and leads to a voltage source (Fig. 4). A current I flows through the conductor in response to the application of a voltage difference V between the two contacts. For small applied voltages, I depends linearly on V. This is the regime of linear response. The coefficient of proportionality between current and voltage is the *conductance* $G = \lim_{V\to 0} I/V$. The conductance of a macroscopic and homogeneous conducting wire in zero magnetic field is proportional to its cross-section A and inversely proportional

to its length L. The coefficient of proportionality is the *conductivity* σ, defined by $G = (A/L)\sigma$. The conductivity relates the current density $j = I/A$ to the electric field $\mathcal{E} = V/L$ by $j = \sigma\mathcal{E}$.

We emphasize the difference between conductance and conductivity. The conductivity is a property of the material, while the conductance is a property of a specific sample (including contacts and leads). If the conductor is too small or not homogeneous, then its conductance does not scale with the dimensions of the sample and can not be obtained from the conductivity of the material. Since conductance is a concept of a more general experimental significance than conductivity, one needs a theoretical framework which deals with the conductance explicitly.

The most important property of the linear response regime is the possibility to relate the conductance and conductivity to Fermi level properties of the conductor. The Einstein relation is one such relation, the Landauer formula another. The Einstein relation expresses the conductivity tensor as the product of density of states and diffusion coefficient, both evaluated at the Fermi level E_F. The Landauer formula relates the conductance to the transmission probability at E_F. Since the Landauer formula involves the conductance, rather than the conductivity, it is more generally applicable than the Einstein relation. The Einstein relation is the more familiar of the two, so we discuss it first. We restrict ourselves in this article to non-interacting electrons.

3.2 Einstein Relation

The Einstein relation follows from the thermodynamic rule that the current density is zero if the electrochemical potential μ is uniform throughout the sample. The electrochemical potential μ is the sum of the electrostatic potential energy $-eV$ and the chemical potential (or Fermi energy) E_F. A difference in electrochemical potential between two regions in the sample means that energy is gained (or lost) on transporting an electron from one region to the other. The system is thus not in equilibrium. Electrons will drift from the high μ region to the low μ region, until the electrochemical potentials are equalized.

The conduction electrons in a semiconductor or in a metal form an *electron gas*, moving randomly through the crystal lattice. Through the interaction with the periodic electrostatic potential due to the lattice, the quantum states accessible to these electrons are Bloch states organized in bands, with dispersion relation $E_n(k)$ and density of states $\rho_n(E)$. For our purposes it is sufficient to consider only a parabolic conduction band, for which $E(k) = \hbar^2 k^2/2m$, with m the effective mass (which is typically less than the free electron mass). The electrons occupy the available states according to the Fermi-Dirac distribution function

$$f(E - E_F) = [1 + \exp(E - E_F)/kT]^{-1}\,. \tag{16}$$

The density of electrons in the partially filled conduction band is thus given by

$$n = \int_0^\infty \rho(E) f(E - E_F) dE. \tag{17}$$

When E_F is more than a few kT below the bottom of the conduction band E_c(i.e. in the band gap, where no states are available), the electron gas density is very low, and the Fermi-Dirac distribution may be approximated by the classical Maxwell-Boltzmann distribution. One then speaks of a non-degenerate electron gas. In this article, however, we are concerned with the opposite limit of a degenerate electron gas, where $E_F - E_c \gg kT$. We will now derive the Einstein relation for such a degenerate electron gas,

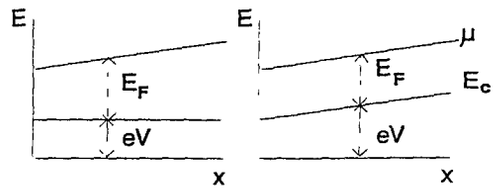


Figure 5. A gradient in the electrochemical potential $\mu = -eV + E_F$ can be caused by an electric field $\mathcal{E} = -\nabla V$, or by a density gradient $\nabla n = \rho(E_F)\nabla E_F$.

considering for simplicity the limit of zero temperature (the generalization to a finite temperature is straightforward).

At $T = 0$, E_F is the energy of the highest occupied energy level, measured relative to the conduction band bottom. As illustrated in Fig. 5, a gradient in the electrochemical potential $\mu = -eV + E_F$ can be caused by an electric field $\mathcal{E} = -\nabla V$, or by a density gradient $\nabla n = \rho(E_F)\nabla E_F$

$$\nabla\mu = e\mathcal{E} + \rho(E_F)^{-1}\nabla n. \tag{18}$$

An electric field induces a current density $j_{\text{drift}} = \sigma\mathcal{E}$. A density gradient induces a current density $j_{\text{diffusion}} = eD\nabla n$, with D the diffusion constant. If $\nabla\mu = 0$ we have from Eq. (18) that $\nabla n = -e\rho(E_F)\mathcal{E}$. Hence, the total current density is

$$\begin{aligned} j &= j_{\text{drift}} + j_{\text{diffusion}} \\ &= [\sigma - e^2\rho(E_F)D]\mathcal{E}, \text{when} \nabla\mu = 0. \end{aligned} \tag{19}$$

The requirement $j = 0$ when $\nabla\mu = 0$ (for arbitrary $\mathcal{E}$) yields the Einstein relation for a degenerate electron gas at $T = 0$

$$\sigma = e^2\rho(E_F)D. \tag{20}$$

Because of the Einstein relation we can write

$$\begin{aligned} j &= eD\nabla n + \sigma\mathcal{E} \\ &= \sigma\nabla E_F/e - \sigma\nabla V \\ &= \sigma\nabla\mu/e. \end{aligned} \tag{21}$$

This relation expresses the fact that the fundamental driving force for the current in a system out of thermal equilibrium is $\nabla\mu$.[2]

For small ∇n, and at low temperatures, only states near E_F contribute to $j_{\text{diffusion}}$. The diffusion coefficient D is thus by definition a Fermi level property. The current j_{drift} caused by an electric field in general contains contributions from all states below E_F. The different distribution over energies of drift and diffusion currents arises because $e\mathcal{E}$ is a force which enters in the equations of motion, and hence acts on all electrons, while ∇n is a "thermodynamic" force, which only affects the occupation of states near the Fermi level. The importance of the Einstein relation (20) is that it shows that, although

[2]Thus, a "voltmeter" actually measures μ, not V, and a "voltage source" maintains a constant difference in electrochemical potential between its two terminals.

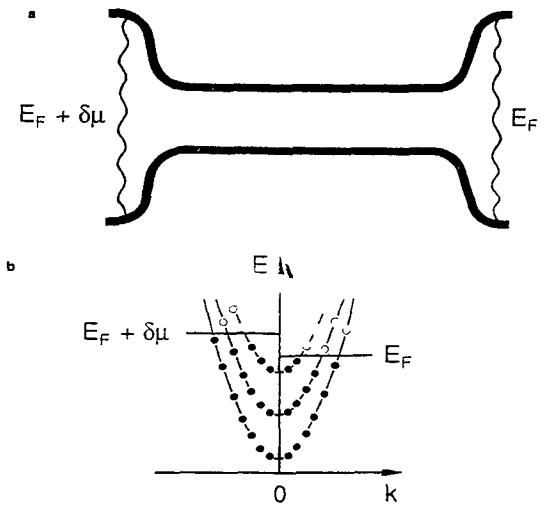


Figure 6 a An ideal electron waveguide connected through ideal leads (the gradually widening regions) to reservoirs at different electro-chemical potentials b Plot of the dispersion relation for the lowest three one-dimensional subbands in the waveguide The combination of a reservoir and an ideal lead ensures complete filling of all available states up to the electro-chemical potential of the reservoirs (at zero temperature)

σ is not manifestly a Fermi level property of the sample, it can nevertheless be expressed entirely in terms of Fermi level properties This fact is at the heart of solid state electron optics, because it allows us to treat electrical conduction as a transmission problem of (nearly) monochromatic particles

The Einstein relation for the conductivity may be generalized to a relation for the conductance Imagine two wide electron gas reservoirs having a slight difference δn in electron density, which are brought into contact by means of a narrow channel, as in Fig 6a A current I will flow in the channel, carried by electrons with energies between the Fermi energies E_F and $E_F + \delta\mu$ in the low and high-density regions At zero temperature, and for small δn, one has $\delta\mu = \delta n/\rho(E_F)$ The *diffusance* $\mathcal{D}$ is defined by $I = e\mathcal{D}\delta n$, and is related to the conductance G by

$$G = e^2\rho(E_F)\mathcal{D}, \tag{22}$$

which implies

$$I = G\delta\mu/e \tag{23}$$

Eq (22) is a generalization to the conductance G and diffusance $\mathcal{D}$ of the Einstein relation (20), and is derived in a completely analogous way The implication is that one may express the conductance in terms of the properties of the quantum states at the Fermi level

3.3 Electron Waveguide

The prototypical conductor is an electron waveguide An ideal waveguide has completely smooth boundaries, a uniform electron density along the channel, and an absence of defects that might scatter electrons Some elementary properties of the propagating modes in an electron waveguide, which we will need to determine its conductance are derived here We consider a cylindrical conducting channel along the $x-$axis, defined by

a lateral confining potential $V(y,z)$. For such a potential, the motion in the $x-$direction is separable. The Hamiltonian has the form (for a single spin component)

$$\mathcal{H} = \frac{p^2}{2m} + V(y,z) \tag{24}$$

with $p^2 = p_x^2 + p_y^2 + p_z^2$. Because the momentum $p_x = -\imath\hbar\partial/\partial x$ along the channel commutes with $\mathcal{H}$, the eigenfunctions of $\mathcal{H}$ can be chosen to be also eigenfunctions of p_x. The wavefunction

$$\langle \mathbf{r} \mid n,k\rangle = \Psi_{n,k}(y,z)\mathrm{e}^{\imath kx} \tag{25}$$

is an eigenfunction of p_x with eigenvalue $\hbar k$. It is also an eigenfunction of $\mathcal{H}$ with eigenvalue $E_n(k)$ if Ψ satisfies

$$\left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial y^2} - \frac{\hbar^2}{2m}\frac{\partial^2}{\partial z^2} + \frac{\hbar^2 k^2}{2m} + V(y,z)\right]\Psi_{n,k}(y,z) = E_n(k)\Psi_{n,k}(y,z). \tag{26}$$

Eq. (26) is the Schrödinger equation for motion in the $y-z$ plane in the effective potential

$$V_{\mathrm{eff}}(y,z,k) = V(y,z) + \frac{\hbar^2 k^2}{2m}. \tag{27}$$

Because the motion is bounded, Eq. (26) has for each k a discrete set of eigenvalues $E_n(k)$, $n = 1,2,\ldots$. It should be emphasized that V_{eff}, since it depends on k, is not a true electrostatic potential (which should only depend on the coordinates). The eigenvalues $E_n(k)$ depend quadratically on k,

$$E_n(k) = E_n^{\min} + \frac{\hbar^2 k^2}{2m}. \tag{28}$$

The conventional terminology in solid state physics refers to the collection of states for a given value of n as a one-dimensional subband. In a waveguide terminology, the index n labels the modes, and the dependence of the energy $E_n(k)$ on the wavenumber k is called the *dispersion relation* of the $n-$th mode. The dispersion relation (28) is illustrated in Fig. 6b for the lowest 3 subbands in an electron waveguide. The lowest energy on the curve $E_n(k)$ is the *cutoff energy* $E_n^{\min}$ of the $n-$th mode. The *propagating* modes at energy E are those for which $E_n^{\min} < E$ so that the equation $E_n(k) = E$ has a solution for a real value of k. The wavefunction (25) is then a non-decaying plane wave along the channel. The modes with $E_n^{\min} > E$ do not propagate at energy E. A localized perturbation (such as the quantum point contact considered in section 4) can excite such *evanescent* modes, but they then decay exponentially along the channel. For a propagating mode, one can define the *group velocity*

$$v_n(k) = \frac{1}{\hbar}\frac{dE_n(k)}{dk}. \tag{29}$$

In view of Eq. (28), one has $v_n(k) = \hbar k/m$. Note, however, that the group velocity differs from the velocity $\hbar k/m$ derived from the wavenumber (the *phase velocity*) if one places the waveguide in an external magnetic field[22].

To calculate the conductance, we need to know the number of states in a given energy interval, in addition to their velocity. The number of states for the $n-$th mode in an energy interval $[E_n(k), E_n(k) + dE_n(k)]$ in a waveguide section of length L may be written as $\rho_n(E)LdE_n(k)$, where $\rho_n(E)$ is the density of states per unit length. This

number is equal to $4Ldk/2\pi$ (the factor of 4 contains a factor of 2 from the spin-degeneracy, and another factor of 2 from the two velocity directions) We thus find for the density of states of a single subband

$$\rho_n(E) = \frac{2}{\pi}\left(\frac{dE_n(k)}{dk}^{-1}\right) = \frac{2m}{\pi\hbar^2}\left(\frac{\hbar^2}{2m(E-E_n^{\min})}\right)^{1/2} \tag{30}$$

Comparison with Eq (29) shows that the density of states of a waveguide mode is inversely proportional to its group velocity,

$$\rho_n(E) = (\pi\hbar v_n(k)/2)^{-1} \tag{31}$$

It is useful to define also the density ρ_n^+ of *positive velocity states*, which is just one half ρ_n, $\rho_n(E)^+ = (\pi\hbar v_n(k))^{-1}$. The density of states for a multi-mode electron waveguide in a 2D electron gas with a hard-wall confining potential is shown in Fig. 7.

3.4 Conductance of an Ideal Electron Waveguide

To calculate the conductance of an electron waveguide, we adopt Landauer's viewpoint,

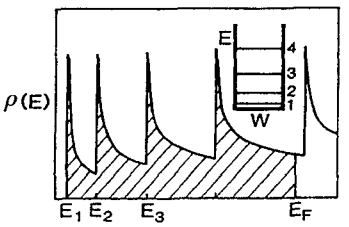


Figure 7 Density of states of a multi-mode electron waveguide in a 2D electron gas, with a hard-wall lateral confining potential.

which is to treat transport as a transmission problem. This point of view is justified by the following considerations. The inelastic scattering length at low temperatures can be quite long (on the order of 10 μm), exceeding the length of micron or sub-micron sized conductors, typically used for the study of quantum transport It is then reasonable to ignore inelastic scattering in the conductor entirely, and to assume that it occurs in the contacts exclusively. Ideal contacts function as electron *reservoirs* A *source* reservoir at electrochemical potential $E_F + \delta\mu$ feeds the conductor with an incoherent flux of incident electrons, a second reservoir at electrochemical potential E_F is a *drain* for the electrons that have traversed the conductor. The conductance can thus be expressed in terms of the transmission probability from source to drain. Elastic scattering in the conductor reduces the transmission probability, because some electrons are reflected back into the source contact

In this section we consider the case of an *ideal* electron waveguide between two ideal contacts. An ideal contact can be formed by inserting a smoothly widening region

(an ideal *lead*) between the reservoirs and the waveguide proper (see section 4 for a discussion of the role of a smooth region). Because of the assumed absence of scattering processes in the waveguide, an electron occupying a certain quantum state $|\, n, k\rangle$ at one point in the waveguide will occupy the same state further downstream. The assumption of ideal contacts implies that within the waveguide the right-moving states are occupied up to $E_F + \delta\mu$ (the electro-chemical potential of the left reservoir), while the left-moving states are occupied up to E_F (cf. Fig. 6b). We write $I = \sum_{n=1}^{N} I_n$, with I_n the current in mode n and N the number of propagating modes. The current I_n is carried by the occupied states in mode n with energy between E_F and $E_F + \delta\mu$. States below E_F give no net contribution to the current, because the contribution of each positive velocity state cancels against that of the corresponding negative velocity state. The amount of current dI_n carried by states of mode n in the infinitesimal interval $[E_n(k), E_n(k) + dE_n(k)]$ is given by the product of the charge e, the number of positive velocity states in that interval $\rho_n^+(E)dE_n(k)$, and the group velocity $v_n(k)$. This yields simply

$$dI_n = \frac{2e}{h} dE_n(k), \tag{32}$$

because the group velocity cancels against the density of states, cf. Eq. (31). Again, we assume a two-fold spin degeneracy of the energy levels, hence the prefactor of 2. The total current I_n in mode n follows on integration from E_F to $E_F + \delta\mu$,

$$I_n = \frac{2e}{h} \int_{E_F}^{E_F + \delta\mu} dE_n(k) = \frac{2e}{h} \delta\mu. \tag{33}$$

Remarkably, for an ideal electron waveguide, the current I_n induced in mode n by a difference $\delta\mu$ in Fermi energies between the ideal contacts, equals $(2e/h)\delta\mu$ *independent of mode index or Fermi energy.* The current in the channel is shared equally ("equipartitioned") among the N propagating modes at the Fermi level, because of the cancellation of group velocity and density of states (cf. Eqs. (29) and (30)). Since $G = I/(\delta\mu/e)$, this equipartition rule implies that the conductance of an ideal electron waveguide is quantized in units of $2e^2/h$:

$$G = \frac{2e^2}{h} N, \tag{34}$$

with N the number of propagating modes in the waveguide.

It is instructive to consider the special case of a channel in a 2D electron gas, defined by a square-well confining potential. In this case, the equipartitioning of the current among the modes can be understood graphically, see Fig. 8. This diagram shows the Fermi circle of an unbounded 2D electron gas in $k-$space. The right-moving states in the energy interval $(E_F, E_F + \delta\mu)$ are shaded. The modes in the channel (of width W) correspond to the pairs of horizontal lines at $k_y = \pm n\pi/W, n = 1, 2, \ldots$. The number of propagating modes at the Fermi level is $N = \mathrm{Int}[k_F W/\pi]$. Each mode can be characterized by an angle ϕ_n (indicated in Fig. 8), such that $n = (k_F W/\pi)\sin\phi_n$. The group velocity $v_n = \hbar k_x/m$ is proportional to $\cos\phi_n$, and thus decreases with increasing n. However, the decrease in v_n is compensated by an increase in the number of states in the shaded region in Fig. 8. This number is proportional to the length of the horizontal lines within the shaded region, and hence to $1/\cos\phi_n$. The current I_n in mode n is proportional to the product of group velocity and number of states (per unit channel length), and hence the dependence on the mode index n drops out. Each mode carries the same amount of current.

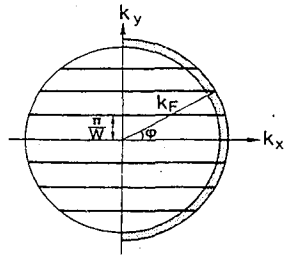


Figure 3. Semi-classical illustration of the equipartitioning of current among 1D subbands in an ideal electron waveguide. The net current is carried by the shaded region in $k-$space. In an ideal electron waveguide the allowed states lie on the horizontal lines. These correspond to quantized values for $k_y = \pm n\pi/W$, and continuous values for k_x.

3.5 Landauer Formula

We need to take just one more step to arrive at the Landauer formula, for a conductor with scattering (for example due to impurities). Scattering causes partial reflection of the injected current back into the source reservoir. If a fraction T_n of the current I_n injected by the source reservoir is transmitted to the drain reservoir, then the total current through the conductor becomes $I = (2e/h)\delta\mu \sum_{n=1}^{N} T_n$. Using $G = I/(\delta\mu/e)$ one obtains the Landauer formula

$$G = \frac{2e^2}{h} \sum_{n=1}^{N} T_n. \tag{35}$$

Eq. (35) may also be written in the form

$$G = \frac{2e^2}{h} \sum_{n,m=1}^{N} \mid t_{mn} \mid^2 = \frac{2e^2}{h} \mathrm{Trtt}^\dagger, \tag{36}$$

where $T_n = \sum_{m=1}^{N} \mid t_{mn} \mid^2$ has been expressed in terms of the matrix t (with elements t_{mn}) of transmission probability amplitudes from an incident mode n to a transmitted mode m.

So far, we have treated the case of zero temperature, where only electrons at the Fermi level have to be considered. This may be expressed in Landauer's formula by making the energy dependence of the transmission probability explicit

$$G = \frac{2e^2}{h} \sum_{n} T_n(E_\mathrm{F}) \equiv \frac{2e^2}{h} T(E_\mathrm{F}). \tag{37}$$

At finite temperatures, energies within a few kT from E_F have to be taken into account. The current I may now be written as the difference $I_{12} - I_{21}$ of the current from source to drain

$$I_{12} = \frac{2e}{h} \int f(E - (E_\mathrm{F} + \delta\mu)) T(E) dE$$

and the current from drain to source

$$I_{21} = \frac{2e}{h} \int f(E - E_F) T(E) dE.$$

For small $\delta\mu$ one has $f(E - (E_F + \delta\mu)) = f(E - E_F) - (\partial f/\partial E)\delta\mu$, so that

$$I = \frac{2e}{h} \int_0^\infty \left(-\frac{\partial f}{\partial E}\right) T(E) dE \delta\mu,$$

or, in view of Eq (23)

$$G = \frac{2e^2}{h} \int_0^\infty \left(-\frac{\partial f}{\partial E}\right) T(E) dE$$

This is the finite temperature generalization of the Landauer formula. The effect of a finite temperature is to average $T(E)$ near E_F over a range of energies of a few kT in width.

4. CONDUCTANCE QUANTIZATION AND TRANSMISSION STEPS

4.1 Quantum Point Contacts

In the previous section we have shown that the conductance of an ideal electron waveguide, attached to ideal contacts, is quantized in units of $2e^2/h$ (for a two-fold spin-degeneracy),

$$G = \frac{2e^2}{h} N, \tag{39}$$

with N the number of propagating modes at the Fermi level. In 1988, it was discovered [15, 16] that the conductance of a *quantum point contact* obeys Eq (39) to a quite reasonable accuracy (better than 1 %). A quantum point contact is a constriction in a 2D electron gas, defined electrostatically by means of a split gate on top of the heterostructure (a schematical view is given in Fig 9). In the experiment[15], the width is continuously variable from 0 to 250 nm, or from 0 to about 7 times the Fermi wave length of the electrons in the 2D electron gas. The length is much less than the mean free path, so that transport through the point contact is ballistic. The conductance of a quantum point contact is shown in Fig. 10. Each step reflects an increase in the number of propagating modes by one due to the increase of the point contact width. This effect is a manifestation of the equipartition of current among an integer number of propagating modes in the constriction, each mode carrying a current of $2e^2/h$ times the applied voltage V, as in an ideal electron waveguide.

It remains to be explained, of course, *why* the quantum point contact behaves as an ideal electron waveguide, since diffraction at the entrance and exit of the constriction might be expected to induce large deviations from precise quantization. To analyze such deviations it is necessary to solve the Schrödinger equation in the narrow point contact and the adjacent wide regions, with plane wave boundary conditions at infinity. The resulting transmission coefficients determine the conductance via the Landauer formula (36). This scattering problem has been solved numerically for point contacts of a variety of shapes and analytically in special geometries. When considering the mode coupling at the entrance and exit of the constriction it is important to distinguish between the case of a gradual (*adiabatic*) and of an *abrupt* transition from wide to narrow regions.

If the constriction width $W(x)$ changes sufficiently gradually, the transport through the constriction is adiabatic, i.e. without intersubband scattering[18]. The transmission

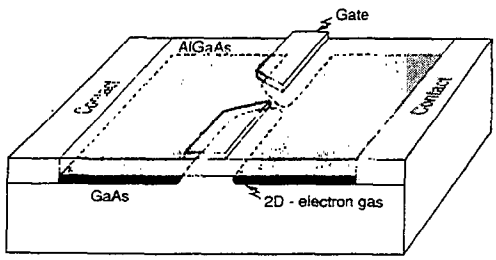


Figure 9. Schematic cross-sectional view of a quantum point contact, defined in a high mobility 2D electron gas at the interface of a GaAs-AlGaAs heterostructure. The point contact is formed when a negative voltage is applied to the gate electrodes on top of AlGaAs layer.

fined in a high
ure The point
odes on top of

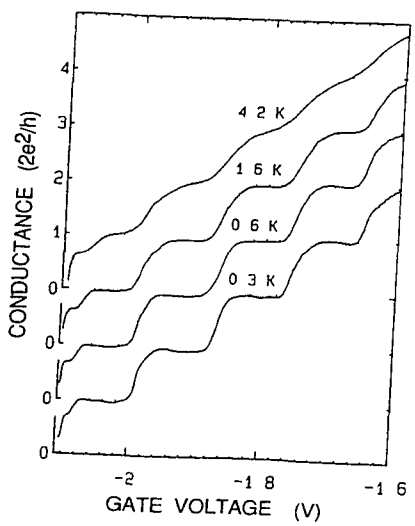


Figure 10 Point contact conductance as a function of gate voltage at 0.3-4 K demonstrating the conductance quantization in units of $2e^2/h$. The conductance is obtained from a two-terminal measurement (after subtraction of a small background resistance). The constriction width increases with increasing voltage on the gate. (From Van Wees et al. [15]).

coefficients then vanish, $\mid t_{nm} \mid^2 = 0$, unless $n = m \leq N_{\rm min}$, with $N_{\rm min}$ the smallest number of propagating modes in the constriction. The conductance quantization (39), with N replaced by $N_{\rm min}$, then follows immediately from Eq. (36). The criterion for adiabatic transport is $dW/dx \lesssim 1/N(x)$, with $N(x) \sim k_{\rm F}W(x)/\pi$ the local number of subbands. As the constriction widens, $N(x)$ increases and adiabaticity is preserved only if $W(x)$ increases more and more slowly. In practice, adiabaticity breaks down at a width $W_{\rm max}$ which is at most a factor of two larger than the minimum width $W_{\rm min}$. This does not affect the conductance of the constriction, however, if the breakdown of adiabaticity results in a mixing of the subbands without causing reflection back through the constriction. If such is the case, the total transmission probability through the constriction remains the same as in the hypothetical case of fully adiabatic transport. As pointed out by Yacoby and Imry[17], a relatively small adiabatic increase in width from $W_{\rm min}$ to $W_{\rm max}$ is sufficient to ensure a drastic suppression of reflections at $W_{\rm max}$. The reason is that the subbands with the largest reflection probability are close to cut-off, i.e. they have subband index close to $N_{\rm max}$, the number of subbands occupied at $W_{\rm max}$. Because the transport is adiabatic from $W_{\rm min}$ to $W_{\rm max}$, only the $N_{\rm min}$ subbands with the smallest n arrive at $W_{\rm max}$, and these subbands have a small reflection probability. In the language of waveguide transmission, one has impedance matched the constriction to the wide regions. The filtering of subbands by a gradually widening constriction restricts the emission cone of electrons injected through it into the wide regions. This *horn collimation effect*[19] has been observed experimentally[20]. It allows one to perform solid state electron optical experiments using a quantum point contact as injector of a collimated electron beam (cf. section 5).

An adiabatic constriction improves the accuracy of the conductance quantization, but is not required to observe the stepwise increase of the conductance. Calculations have shown that well-defined conductance plateaux persist for *abrupt* constrictions, although transmission resonances lead to periodic dips in the conductance below the quantized plateau value[21]. Further details and references to the literature may be found in ref. [22].

The results described above do not only explain the conductance quantization of a quantum point contact, but they also show that equipartitioning of the current over the waveguide modes inside the constriction is approximately valid regardless of the detailed shape of the connection to the wide 2D electron gas. This provides some justification for the use of the concept of a reservoir and an ideal lead, and thus for the use of the Landauer formula in practical cases.

4.2 Steps in the Optical Transmission through a Slit

The unexpected discovery of the conductance quantization of a quantum point contact has led to a search for its optical analogue. A considerable literature exists on the coupling of light into fibers, or microwaves into waveguides, but the optical analogue was not noticed previously. At the basis for the analogy are three facts.

- Firstly, as we have seen, linear response implies transport at $E_{\rm F}$, which is analogous to optical experiments with monochromatic light. The Helmholtz equation (14) for the electric field of monochromatic light in vacuum (polarized in the $z-$direction)

$$\nabla^2 \mathcal{E}_z = -(\omega/c)^2 \mathcal{E}_z \tag{40}$$

is similar to the Schrödinger equation. Secondly, the boundary condition at a metal surface parallel to z is that $\mathcal{E}_z$ vanishes, which corresponds to the vanishing of the electron wavefunction Ψ at an infinitely steep potential wall. Thirdly, the expression for

unless $n = m \leq N_{\rm min}$, with $N_{\rm min}$ th
onstriction. The conductance quantiz
s immediately from Eq. (36). The cr
), with $N(x) \sim k_F W(x)/\pi$ the local
$V(x)$ increases and adiabaticity is pres
ly. In practice, adiabaticity breaks
of two larger than the minimum wi
f the constriction, however, if the bre
bbands without causing reflection ba
he total transmission probability th
. ..theoretical use of fully adiabatic
), a relatively small adiabatic increas
re a drastic suppression of reflections
largest reflection probability are close
max, the number of subbands occupie
$W_{\rm min}$ to $W_{\rm max}$, only the $N_{\rm min}$ subband
bands have a small reflection probabil
has impedance matched the constric
by a gradually widening constrictio
through it into the wide regions.
experimentally[20]. It allows one t
using a quantum point contact as in

the accuracy of the conductance qu
se increase of the conductance. Ca
e plateaux persist for *abrupt* constri
periodic dips in the conductance
tails and references to the literatur

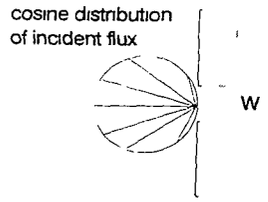
nly explain the conductance quantiz
w that equipartitioning of the curren
approximately valid regardless of th
electron gas. This provides some ju
nd an ideal lead, and thus for the

n through a Slit

nce quantization of a quantum poi
,ue. A considerable literature exis
es into waveguides, but the optical
or the analogy are three facts.
e implies transport at E_F, which is
c light. The Helmholtz equation (1
uum (polarized in the $z-$direction

$\omega/c)^2\mathcal{E}_z$

econdly, the boundary condition a
which corresponds to the vanishi
ep potential wall. Thirdly, the expr



re 11. An isotropic velocity distribution in a wide region of an electron gas in
nal equilibrium causes a $\cos\theta$ distribution of the flux incident on a narrow slit. θ
the angle with respect to normal incidence. The magnitude of the flux is indicated
straight lines for various angles of incidence.

ergy current density in optics is given by the Poynting vector

$$\mathbf{j} = \frac{1}{2}\epsilon_0 c^2 \mathrm{Re}(\mathcal{E} \times \mathbf{B}^*) = \frac{1}{2}\epsilon_0 \frac{c}{k}\mathrm{Re}(i\mathcal{E}_z \nabla \mathcal{E}_z^*) \tag{41}$$

tical, up to a numerical factor, to the quantum mechanical particle current density,

$$\mathbf{j} = \frac{\hbar}{m}\mathrm{Re}(i\Psi\nabla\Psi^*). \tag{42}$$

ows that the ratio of transmitted to incident power in the optical problem is the
as the ratio of the transmitted to incident current in its electronic counterpart.
et us now look at the ingredients needed in optics to mimick the conditions of
periment on the conductance quantization of a quantum point contact. In optics
ually studies the transmission of a single incident plane wave, as a function of
gle of incidence. In contrast, electrons are incident at a point contact with an
ic velocity distribution, or equivalently with a $\cos\theta$ distribution of the incident
ee Fig. 11). In optics such an angular distribution is known as a Lambertian
, and is provided by any diffusely scattering medium. The analogue of a point
t in a 2D electron gas is a slit in a metal screen illuminated diffusively in a plane
dicular to the slit, with $\mathcal{E}$ polarized parallel to the slit (and monochromatic).
he analogue of the conductance G of the point contact is the transmission cross
per unit length σ of the slit, defined by $\sigma = P/j_{\rm in}$, where P is the transmitted
per unit slit length, and $j_{\rm in}$ is the incident energy flux (so that σ has the dimension
ngth). If the shape of the slit is the same as the shape of the point contact, then
ws from the equivalence discussed above that G and σ are related by

$$\frac{h}{2e^2}G = \frac{2}{\lambda}\sigma, \tag{43}$$

the wavelength of the light. To see this, one should note that, because of the
er formula, $(h/2e^2)G = T$ equals the transmitted current divided by the incident
per mode. In the optical case, $j_{\rm in}\lambda/2$ is the incident current per mode, so that
$/2) = (2/\lambda)\sigma \equiv T'$ equals the transmitted power divided by the incident power

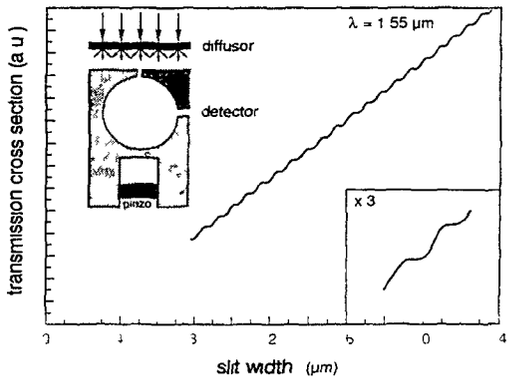


Figure 12 Experimental demonstration of equidistant steps in the transmission cross-section of a slit of adjustable width A 2D Lambertian monochromatic source is obtained by illuminating a diffusor consisting of a random array of parallel fibers by a diode laser beam An integrating sphere is used to obtain a detector signal proportional to the transmission crosssection [From Montie et al [23]]

per mode But in the previous paragraph we have proved that the two ratios are the same, hence $T = T'$ and $(h/2e^2)G = (2/\lambda)\sigma$ The dependence of transmission cross section σ on the slit width W should thus be a stair case, with steps separated by $\lambda/2$, and with a constant step height given also by $\lambda/2$ The role of the shape of the slit should be identical to the role of the shape of a hard wall confining potential in the case of a quantum point contact This is why one expects $\sigma = N\lambda/2$ to a good aproximation

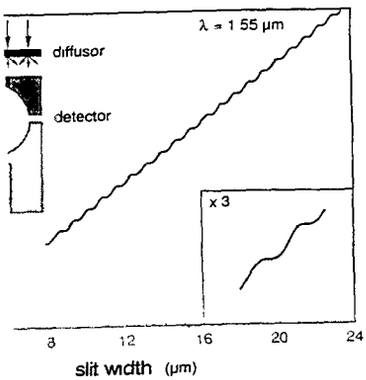
This prediction[25] has been verified experimentally by Montie et al [23] Their result is reproduced in Fig 12 The generalization of the optical analogue to the case where a dielectric fills the wide regions (but not the slit) is straightforward since (as explained in the next section) a negative step in refractive index is analogous to a positive step in the electrostatic potential (or in the local conduction band bottom) in the solid state electron optics case Such a step reduces the number of propagating modes in the constriction – but has no effect on the conductance quantization

Absorption at the slit boundaries gives rise to a rounding of the transmission steps[23] This effect has of course no counterpart in solid state electron optics We also note that, unlike in the electronic case, it is straightforward to generalize the optical experiment to transmission through an aperture (a hole in a screen) Although this experiment has not yet been performed, the theory[25] predicts $\sigma = N\lambda^2/2\pi$ for this case (assuming that the two independent polarizations of the modes in the aperture can be resolved)

5 REFRACTION AND TUNNELING

5.1 Snell's Law for Electrons and Photons

Consider a 2D electron gas, with Fermi energy E_F, containing a region of reduced electron density The local conduction band bottom is raised in such a region to a value



ation of equidistant steps in the transm
A 2D Lambertian monochromatic source
of a random array of parallel fibers by
ed to obtain a detector signal proport
ontie et al [23]]

agraph we have proved that the two r
$= (2/\lambda)\sigma$. The dependence of transm
d thus be a stair case, with steps separ
ven also by $\lambda/2$. The role of the shap
shape of a hard-wall confining potentia
why one expects $\sigma = N\lambda/2$ to a good a
rified experimentally by Montie et al.
generalization of the optical analogue
ons (but not the slit) is straightforwa
gative step in refractive index is ana
ential (or in the local conduction band

Such a step reduces the number of
no effect on the conductance quantizat
ies gives rise to a rounding of the
no counterpart in solid state electron
c case, it is straightforward to general
gh an aperture (a hole in a screen). A
ied, the theory[25] predicts $\sigma = N\lambda^2 s$
ent polarizations of the modes in the a

ELING

Photons

rmi energy E_F, containing a region
band bottom is raised in such a regio

n electron at the Fermi level, impinging on the region of reduced density, thus potential barrier of height E_c. Classically, there are just two possibilities. The n will be reflected specularly if its kinetic energy along the direction of normal ice is less than E_c, or

$$E_F \cos^2\theta_1 < E_c \rightarrow \text{reflection} \tag{44}$$

tory (1) in Fig. 13a). The electron will be refracted when

$$E_F \cos^2\theta_1 > E_c \rightarrow \text{refraction} \tag{45}$$

tory (2) in Fig. 13a). One may derive Snell's law for refraction of conduction ns by invoking conservation of tangential momentum

$$k_1 \sin\theta_1 = k_2 \sin\theta_2, \tag{46}$$

$k_1 = (2mE_F/\hbar^2)^{1/2}$, and $k_2 = (2m(E_F - E_c)/\hbar^2)^{1/2}$. This result is identical to law in optics. In terms of the velocity $v_i = \hbar k_i/m$, Snell's law for electrons reads

$$v_1 \sin\theta_1 = v_2 \sin\theta_2. \tag{47}$$

tices a difference with Snell's law in optics, $n_1 \sin\theta_1 = n_2 \sin\theta_2$, which corresponds (46), but which may be rewritten as

$$v_2 \sin\theta_1 = v_1 \sin\theta_2, \tag{48}$$

e velocity of light is $v_i = c/n_i = \omega/k_i$, i.e. *inversely* proportional to the wavenum- illustrated in Fig. 14 (see also Fig. 13b), this has the amusing consequence positive lens in solid state electron optics, constructed out of a region of reduced density (i.e. with reduced velocity) has a concave shape, in contrast to optics. positive lens made out of a material with reduced velocity (such as glass) is This difference is a consequence of the different dispersion laws for electrons tons (cf. Section 2).

ng a quantum point contact to inject an electron beam at the Fermi level in a 2D gas it has been possible to demonstrate total specular reflection of electrons at rostatic boundary and magnetic focusing[25], and focusing of an electron beam electrostatic lens[26, 27].

this section, we have discussed Snell's law for electrons and photons in terms ctories (or rays). Alternatively, one may derive Snell's law by matching the s of the wave equations for electrons or for light at the interface between two regions. Such a derivation adds to our understanding, but the result is the fraction being essentially a classical phenomenon. In the next section we discuss g, which may only be understood in terms of quantum mechanics, and which nalogue in geometrical optics.

neling of Electrons and Photons

sition from refraction to tunneling occurs when the potential barrier in the re- educed electron density is increased above the Fermi energy. The optical coun- f this phenomenon is known as frustrated total internal reflection (FTIR). One ly encounters treatments of FTIR[28, 29] as a (somewhat imperfect) analogue *imensional* electron tunneling. As we will show, a more satisfactory analogy

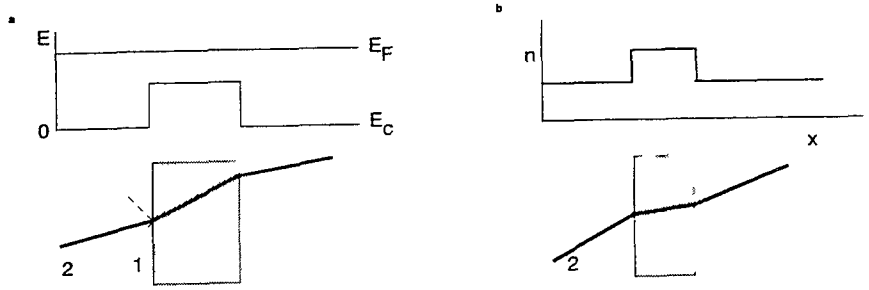
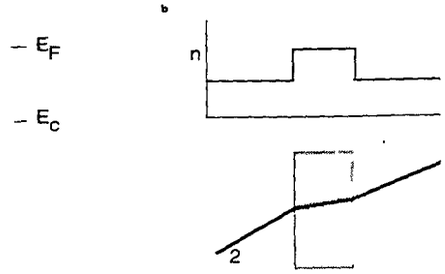
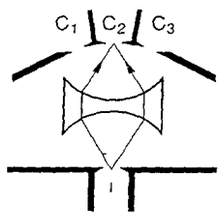


Figure 13 a Refraction (for small angle of incidence) and total reflection (for large angle of incidence) of a conduction electron at a potential barrier defined electrostatically in a 2D electron gas Note that the electron is refracted *away* from the surface normal on entering the barrier region b In optics, a ray is refracted *towards* the surface normal, at a positive step in refractive index





14. Experimental device used to demonstrate focusing of a ballistic electron beam
gle of incidence) and total reflection (fo
n et al.[26] and by Spector et al.[27] A concave lens is positive, even though it
ı at a potential barrier defined electro
ed out of a region with reduced phase velocity.
tron is refracted *away* from the surfac
ics, a ray is refracted *towards* the surf

exists with *two-dimensional* electron tunneling The relevant geometries are depicted in Fig 15 [3]

Consider a monochromatic electro-magnetic wave, polarized linearly with $\mathcal{E}$ in the $z-$direction, propagating in the $x-y$ plane in a medium of refractive index nr The scalar wave equation (13) becomes

$$\nabla^2\mathcal{E} + \left(\frac{\omega}{c}\right)^2 n(x)^2\mathcal{E} = 0 \tag{49}$$

Let us now see what happens at a step in refractive index, from n_1 to the lower value n_2 We look for a plane wave solution in the $y-$direction,

$$\mathcal{E} = \Psi(x) \exp[i(\omega t - \beta y)] \tag{50}$$

Substitution of this wave in Eq (49) yields an equation for $\Psi(x)$,

$$\frac{\partial^2\Psi}{\partial x^2} + \left[\left(\frac{\omega}{c}\right)^2 n^2 - k^2\right]\Psi = 0 \tag{51}$$

For an incident plane wave at angle θ_1 with the $x-$axis, one has $k = (\omega n_1/c)\sin\theta_1$ Hence, in region 1 Eq (50) reduces to

$$\frac{\partial^2\Psi}{\partial x^2} + \left(\frac{\omega}{c}\right)^2 n_1^2\cos^2\theta_1\Psi = 0 \tag{52}$$

whereas in region 2 one has

$$\frac{\partial^2\Psi}{\partial x^2} + \left(\frac{\omega}{c}\right)^2 \left(n_2^2 - n_1^2\sin^2\theta_1\right)\Psi = 0 \tag{53}$$

Tunneling of light occurs when $n_2^2 - n_1^2\sin^2\theta_1 < 0$, so that Eq (53) does not have a propagating solution (Note that the frequency ω does not enter in this condition)

The Schrodinger equation for tunneling at the Fermi level through a planar potential barrier of height E_c in a 2D electron gas reads

$$\frac{\partial^2\Psi}{\partial x^2} + \frac{2m}{\hbar^2}E_F\cos^2\theta_1\Psi = 0, \tag{54}$$

in the 2D electron gas, and

$$\frac{\partial^2\Psi}{\partial x^2} + \frac{2m}{\hbar^2}\left(E_F\cos^2\theta_1 - E_c\right)\Psi = 0, \tag{55}$$

in the barrier region Tunneling thus occurs whenever $E_F\cos^2\theta_1 - E_c < 0$, a condition that depends explicitly on the energy of the electron, E_F In contrast to the optical case, tunneling at normal incidence ($\theta_1 = 0$) is possible in the electronic case if $E_c > E_F$ (The optical condition $n_2^2 - n_1^2\sin^2\theta_1 < 0$ has no solution for $\theta_1 < \arcsin n_2/n_1$) Apart from these differences, a comparison with Eqs (52) and (53) shows that 2D electron tunneling through a planar barrier is analogous to 2D photon tunneling (FTIR) through a region of reduced refractive index, with the following identifications

$$\frac{2m}{\hbar^2}E_F\cos^2\theta_1 \iff \left(\frac{\omega}{c}\right)^2 n_1^2\cos^2\theta_1 \tag{56}$$

$$\frac{2m}{\hbar^2}\left(E_F\cos^2\theta_1 - E_c\right) \iff \left(\frac{\omega}{c}\right)^2\left(n_1^2\cos^2\theta_1 - (n_1^2 - n_2^2)\right) \tag{57}$$

[3] FTIR is commonly studied by bringing two rectangular prisms in close proximity in such a way that they nearly form a cube consisting of the two prisms separated by a narrow air gap A plane wave incident perpendicularly on one of the rectangular faces of a prism is deflected by 90 degrees through total internal reflection On narrowing the air gap some of the light may be transmitted towards the second prism, because of tunneling The total internal reflection is then frustrated

·on tunneling. The relevant geometries ar

lectro-magnetic wave, polarized linearly
· — y plane in a medium of refractive in
s

$$+\left(\frac{\omega}{c}\right)^2 n(x)^2\mathcal{E} = 0.$$

step in refractive index, from n_1 to the
ion in the $y-$direction,

$$\ldots \exp[i(\ldots y - \omega t)]$$

9) yields an equation for $\Psi(x)$,

$$\left[\left(\frac{\omega}{c}\right)^2 n^2 - k^2\right]\Psi = 0.$$

le θ_1 with the $x-$axis, one has $k = (\omega$
to

$$\left(\frac{\omega}{c}\right)^2 n_1^2\cos^2\theta_1\Psi = 0,$$

$$\left(n_2^2 - n_1^2\sin^2\theta_1\right)\Psi = 0.$$

$- n_1^2\sin^2\theta_1 < 0$, so that Eq. (53) does
frequency ω does not enter in this co
neling at the Fermi level through a plan
gas reads

$$\frac{n}{2}E_F\cos^2\theta_1\Psi = 0,$$

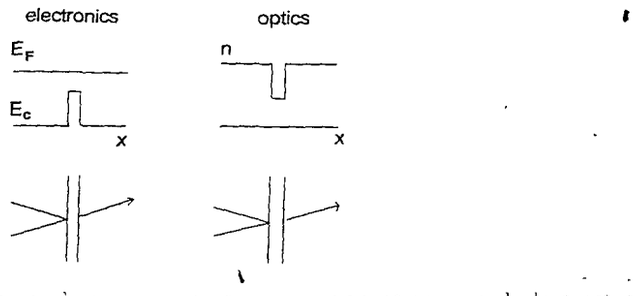
$$_F\cos^2\theta_1 - E_c)\,\Psi = 0,$$

occurs whenever $E_F\cos^2\theta_1 - E_c < 0$,
f the electron, E_F. In contrast to the c
is possible in the electronic case if E_c
is no solution for $\theta_1 < \arcsin n_2/n_1$.)
s. (52) and (53) shows that 2D electro
o 2D photon tunneling (FTIR) throu
llowing identifications

$$\Rightarrow \left(\frac{\omega}{c}\right)^2 n_1^2\cos^2\theta_1$$

$$\Rightarrow \left(\frac{\omega}{c}\right)^2\left(n_1^2\cos^2\theta_1 - (n_1^2 - n_2^2)\right).$$

vo rectangular prisms in close proximity, i
two prisms separated by a narrow air gap.
lar faces of a prism is deflected by 90 deg
gap, some of the light may be transmitted
nternal reflection is then frustrated.



15 a. Tunneling of an electron incident on a region of increased electrostatic
tial (reduced momentum) at an angle such that classically it would be totally
ed. b. Frustrated total reflection of light incident on a region of reduced refractive

rst relation (56) expresses nothing but the correspondence of the wavenumber of
cident electron wave at the Fermi level to that of the incident optical wave. The
l relation (57) may thus be rewritten as

$$\frac{2m}{\hbar^2}E_c \Longleftrightarrow \left(\frac{\omega}{c}\right)^2\left[n_1^2 - n_2^2\right], \tag{58}$$

expresses the fact that a change in electrostatic potential in the electronic problem
change in squared momentum) corresponds to a change in the square of the
ive index. This specific example illustrates the assertion of Sec. 2 that electron
is analogous to optics when one identifies the refractive index with the electron
ntum.

andauer Formula and Fermi's Golden Rule for Tunneling

andauer formula for the conductance in terms of the energy dependent transmis-
robability $T(E)$ (for one spin direction)

$$G = \frac{e^2}{h}\int_0^\infty T(E)\left(-\frac{\partial f}{\partial E}\right)\mathrm{d}E \tag{59}$$

applied straightforwardly to elastic tunneling. This approach is equivalent to
re traditional approach, based on Fermi's golden rule, as we now discuss for 1D
onsider a planar barrier across which a voltage V is maintained (see Fig. 16). The
n gas regions on each side are characterized by shifted Fermi-Dirac distribution
ns $f_1(E) = [1 + \exp(E - E_F)/kT]^{-1}$ and $f_2(E) = [1 + \exp(E + eV - E_F)/kT]^{-1}$.
ansverse momentum is conserved in the tunneling process, so that we can consider
ansverse momentum state separately. The following results are for just one such
or conduction channel. The tunnel rate for an electron approaching the barrier

from region 1 with energy E is

$$\begin{aligned}\Gamma_{12}(E) &= \frac{2\pi}{\hbar}\int dE_2\rho_2(E_2)|W_{12}(E)|^2\delta(E-E_2) \\ &= \frac{2\pi}{\hbar}\rho_2(E)|W_{12}(E)|^2 \end{aligned} \tag{60}$$

Here $|W_{12}(E)|$ is the tunnel matrix element, and $\rho_2(E)$ is the density of states in electron gas region 2 at energy E for the specified transverse momentum state Note that $\rho_2(E)$ depends on the applied voltage, due to the shift in conduction band bottom in region 2 (see Fig 16) To arrive at the current due to electrons moving from 1 to 2 we have to sum the tunnel rate times the electron charge over all occupied states in region 1 excluding the occupied states in region 2 (in view of the Pauli principle) The result is

$$\begin{aligned} I_{12} &= e\int \rho_1(E)f_1(E)\Gamma_{12}(E)[1-f_2(E)]dE \\ &= \frac{e}{h}\int 4\pi^2|W_{1\,2}|^2\rho_1(E)f_1(E)\rho_2(E)[1-f_2(E)]dE \end{aligned} \tag{61}$$

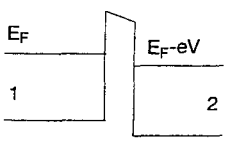


Figure 16 Planar potential barrier separating two degenerate electron gas regions of equal chemical potential, but with shifted Fermi levels because of the voltage V applied across the barrier

The term in the integrand containing the product of Fermi functions cancels on adding I_{21}, so that the net current $I = I_{12} - I_{21}$ is given by

$$I = \frac{e}{h}\int 4\pi^2|W_{1\,2}|^2\rho_1(E)\rho_2(E)[f_1(E)-f_2(E)]dE \tag{62}$$

For small applied voltage $f_2(E) \approx f(E) + eV\partial f/\partial E$ (the subscript 1 is now dropped) so that one finds a linear response conductance (for a single spin direction)

$$G = -\frac{e^2}{h}\int_0^\infty 4\pi^2|W_{1\,2}|^2\rho_1(E)\rho_2(E)\frac{\partial f}{\partial E}dE \tag{63}$$

$$= \frac{e^2}{kT}\int \Gamma_{12}(E)\rho_1(E)f(E)[1-f(E)]dE \tag{64}$$

To arrive at the final result we used the identity $-\partial f/\partial E = (kT)^{-1}f(1-f)$

This equation explicitly contains the density of states of the electron gas regions on either side of the tunnel barrier This may seem puzzling, because of the cancellation of group velocity and 1D density of states in the derivation of Landauer's formula (59)

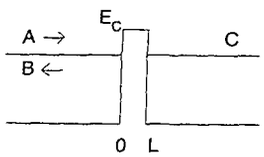


Figure 17. Rectangular planar barrier of height E_c, with incident wave of amplitude A, reflected wave of amplitude B, and transmitted wave of amplitude C.

To establish the equivalence of both results, we note that the tunnel rate from region 1 to region 2 may also be written as the product of an attempt frequency $\nu_{12}(E)$ and the transmission probability $T(E)$

$$\Gamma_{12}(E) = \nu_{12}(E)T(E) \,. \tag{65}$$

The attempt frequency equals the group velocity of the electron incident on the barrier, divided by twice the length L of electron gas region 1 or, equivalently,

$$\nu_{12}(E) = 1/h\rho_1(E) \,, \tag{66}$$

where we have used the relation $v_1 = 2L/h\rho_1$ between group velocity and density of states for one spin direction (cf. Eq. (31), which is for two spin directions). Consequently, one may write

$$T(E) = h\Gamma_{12}(E)\rho_1(E) \,. \tag{67}$$

One may thus express the transmission probability $T(E)$ in terms of the tunnel matrix element, according to

$$T(E) = 4\pi^2|W_{1,2}|^2\rho_1(E)\rho_2(E) \,. \tag{68}$$

This relation proves the equivalence of the standard result (63) for the conductance due to tunneling through a single barrier and the Landauer formula (59). The analysis given above closely follows the one given in 1970 in a textbook by Harrison[30]. However, at that time it was not obvious that the result (59) applies for any value of T: the equivalence to the Fermi golden rule formula holds only in the limit $T \ll 1$, since this rule is based on perturbation theory.

5.4 Rectangular potential

To illustrate how $T(E)$ is calculated, we discuss the text-book example of one-dimensional tunneling through a rectangular potential barrier of height E_c, separating two regions of zero potential (see Fig. 17). The solutions $u(x)$ of the Schrödinger equation in the regions on either side of the barrier are plane waves with (positive or negative) wavevector $k = (2mE)^{1/2}/\hbar$. By reference to Fig. 17 it is clear that

$$x \leq 0 \,;\; u(x) \;=\; Ae^{ikx} + Be^{-ikx} \tag{69}$$

$$x \geq L \,;\; u(x) \;=\; Ce^{ikx} \,, \tag{70}$$

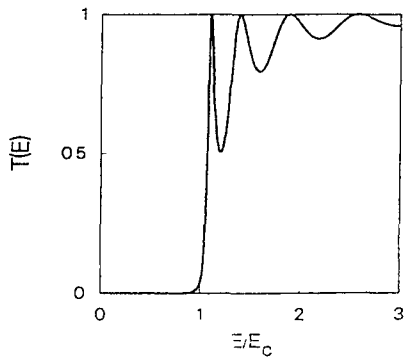


Figure 18 Transmission probability versus energy for a rectangular barrier of height E_c and thickness L, for the case that $\hbar^2/2mL^2 = 0.01E_c$. The curve has been calculated from Eqs (72) and (75)

whereas in the barrier region the solutions are plane waves with wavevector $\kappa = [2m(E - E_c)]^{1/2}/\hbar$

$$0 \leq x \leq L\ ,\ u(x) \ = \ Fe^{i\kappa x} + Ge^{-i\kappa x}\ , \tag{71}$$

For the tunneling problem $E \leq E_c$, so that κ is imaginary, $\kappa = i[2m(E_c - E)]^{1/2}/\hbar$. The transmission probability $T(E) \equiv |C|^2/|A|^2$ can be found by matching the propagating wavefunctions in the regions adjacent to the barrier to the the decaying wavefunction in the barrier. The matching conditions require that both $u(x)$ and $\partial u(x)/\partial x$ are continuous at $x = 0$ and $x = L$. The result for $E \leq E_c$ reads

$$T(E) = [1 + \frac{E_c^2 \sinh^2(\mid \kappa \mid L)}{4E(E_c - E)}]^{-1} \tag{72}$$

This general result has a number of interesting limits. If the barrier is high and thick, $\mid \kappa \mid L \gg 1$ (or equivalently $E_c - E \gg \hbar^2/2mL^2$), then

$$T(E) \approx 16\frac{E(E_c - E)}{E_c^2}\exp(-2 \mid \kappa \mid L) \tag{73}$$

The transmission probability due to tunneling is exponentially small for such a barrier.

Eq (72) reduces to the transmission probability for tunneling through a potential of the form $H\delta$r, if one takes the limit $\mid \kappa \mid L \ll 1$ and defines $H \equiv E_c L$,

$$\begin{aligned} T(E) \ &\approx \ [1 + Z^2]^{-1} \\ Z \ &\equiv \ H/\hbar(2E/m)^{1/2} \end{aligned} \tag{74}$$

The reflection probability at such a one-dimensional delta-scatterer is $R \equiv 1 - T = [1 + Z^{-2}]^{-1}$

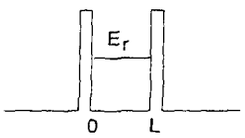


Figure 19 Double planar barrier forming a well with quasi-bound state at energy E_r.

The above approach is applicable as well to the transmission of electrons over a square barrier. Eqs. (69)–(71) still hold, but in this case κ is real. The result for $E \geq E_c$ is

$$T(E) = [1 + \frac{E_c^2 \sin^2(\kappa L)}{4E(E - E_c)}]^{-1} . \tag{75}$$

A plot of $T(E)$ is given in Fig. 18, for the case that $\hbar^2/2mL^2 = 0.01E_c$. The transmission resonances seen for $E > E_c$ correspond to virtual bound states above the barrier, occurring at energies for which κL is an integer times π.

The resonances due to "over-the-barrier" reflection are less pronounced if the potential barrier is rounded, as is often the case experimentally [31]. A similar suppression of transmission resonances occurs in the case of the conductance quantization of a quantum point contact, due to the rounding of the shape of the constriction near entrance and exit.

5.5 Resonant Tunneling

When two barriers are placed in series, the transmission probability $T(E)$ may show resonances due to tunneling through quasi-bound states in the well between the barriers (see Fig. 19). The double barrier is the analogue of the Fabry-Pérot resonator in optics. A theoretical study of resonant tunneling has been made by Breit and Wigner, in the context of resonant enhancement of the neutron capture cross section observed in nuclear physics [32]. Resonant tunneling has since become relevant for solid state physics as well, in particular because of the proposal by Tsu and Esaki [33] to build multiple barrier "superlattice" devices using semiconductor heterostructures. Evidence for resonant tunneling through a double barrier structure was first reported by Chang, Esaki and Tsu [34]. As in most of the subsequent experiments, they measured the current-voltage characteristic to detect the resonance as a negative differential resistance at finite bias. In this section we will discuss instead the transmission probability at zero or negligibly small bias, which determines the linear response conductance. It will be assumed that either the Fermi energy or the energy of the quasi-bound states in the well can be tuned by means of an external parameter (such as gate voltage or magnetic field).

Resonant tunneling through a double barrier may be treated by summing the complex amplitudes of all possible paths which finally lead to transmission. We denote the complex transmission amplitudes of the individual barriers by t_1 and t_2, and the complex reflection amplitudes by r_1 and r_2. These amplitudes are related to the transmission

and reflection probabilities T_i and R_i ($i = 1, 2$) by

$$r_i = R_i^{1/2} e^{i\Delta\phi} \tag{76}$$
$$T_i = 1 - R_i \tag{77}$$

In addition to the phase shifts $\Delta\phi_i$ incurred on reflection off a barrier, there is a phase shift $\phi_\pm$ corresponding to traversal of the well in the positive or negative $x-$direction The total transmission amplitude through the double barrier then is

$$t = t_1 e^{i\phi_+} t_2 + t_1 e^{i\phi_+} r_2 e^{i\phi_-} r_1 e^{i\phi_+} t_2 + \cdots$$
$$= \frac{t_1 t_2 e^{i(\phi_+ + \phi_-)}}{1 - r_1 r_2 e^{i(\phi_+ + \phi_-)}}$$

The transmission probability follows from $T = |t|^2$

$$T = \frac{T_1 T_2}{1 + R_1 R_2 - 2R_1^{1/2} R_2^{1/2} \cos\chi} , \tag{79}$$

where χ is the total phase shift for one round trip in the well

$$\chi = \phi_+ + \phi_- + \Delta\phi_1 + \Delta\phi_2 \tag{80}$$

The transmission probability T has a maximum whenever $\chi = n2\pi$ as a consequence of destructive interference of the backscattered partial waves Since this is precisely the condition for the existence of a quasi bound state in the well the resonance occurs when the energy of the incident electron coincides with the energy E_r of a quasi bound state The maximum and minimum transmission probabilities are given by

$$T_{\max} = \frac{T_1 T_2}{(1 - R_1^{1/2} R_2^{1/2})^2} \approx \frac{4T_1 T_2}{(T_1 + T_2)^2} \tag{81}$$
$$T_{\min} = \frac{T_1 T_2}{(1 + R_1^{1/2} R_2^{1/2})^2} \approx \frac{T_1 T_2}{4} , \tag{82}$$

where the approximate equalities hold only if $T_1 \ll 1$, and $T_2 \ll 1$ Note that if the double barrier is symmetric ($T_1 = T_2$), the maximum transmission probability is unity, regardless of the magnitude of the barrier transparencies The conductance then equals e^2/h (for one spin direction) A plot of T as a function of χ is given in Fig 20, for $T_1 = T_2 = 0\,8$ and $T_1 = T_2 = 0\,2$ The energy dependent transmission probability $T(E)$ may be obtained from Eq (79) provided the phase shift χ and the transmission probabilities of the individual barriers are known as a function of energy For planar rectangular barriers this may be done by the wavefunction matching method discussed in the previous subsection [36]

If the barriers are sufficiently high and thick, both $T_1 \ll 1$ and $T_2 \ll 1$, and $T(E)$ reduces to the Breit Wigner form for energies close to a resonance, as we will now discuss [35] The phase shifts incurred on reflection off the barrier are $\Delta\phi_i = -\pi/2$, independent of energy If the separation of the barriers is L, then the resonance condition $\chi = n2\pi$ reduces to the familiar Bohr Sommerfeld quantization condition $2L/\lambda = n + \frac{1}{2}$ (here $\lambda = 2\pi/k$ with $k = (2mE)^{1/2}/\hbar$) Consider one such state at energy E_r For energies close to E_r the round-trip phase shift χ is linear in $\epsilon_r \equiv E - E_r$,

$$\frac{d\chi}{dE} = \frac{d\chi}{dk}\frac{dk}{dE} = (2L)(\pi\rho/L) = 1/\hbar\nu , \tag{83}$$

(76)
(77)

here is a phase
ve $x-$direction.

(78)

(79)

(80)

2π, as a con-
Since this is
well, the reso-
e energy E_r of
ities are given

(81)

(82)

te that if the
bility is unity,
e then equals
n Fig. 20, for
n probability
transmission
. For planar
od discussed

I, and $T(E)$
l now discuss
independent
ion $\chi = n2\pi$
$n + \frac{1}{2}$ (here
For energies

(83)

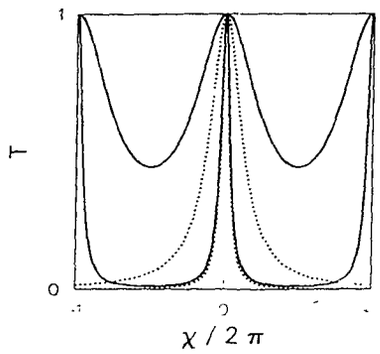


Figure 20. Transmission probability through a double barrier as a function of the round-trip phase shift χ, calculated from Eqs. (79) and (80), for $T_1 = T_2 = 0.8$ (upper full curve) and $T_1 = T_2 = 0.2$ (lower full curve). Also shown are the corresponding Breit-Wigner lineshapes for a single quasi-bound state, calculated from Eq. (85) assuming a linear dependence of χ on $E - E_r$ as in Eq. (84) (dashed curves).

with $\nu = 1/h\rho$ the attempt frequency and $\rho = (L/\pi)dk/dE$ the density of states in the well. Close to resonance we may thus write

$$\chi \approx 2\pi n + \epsilon_r/\hbar\nu \tag{84}$$

By expanding $\cos\chi \approx 1 - \frac{1}{2}(\epsilon_r/\hbar\nu)^2$ and $R_i^{1/2} \approx 1 - \frac{1}{2}T_i$ we then find from Eq. (79)

$$\begin{aligned} T &= \frac{T_1 T_2}{(T_1+T_2)^2/4 + (\epsilon_r/\hbar\nu)^2} \\ &= \frac{\Gamma_1\Gamma_2}{(\Gamma/2)^2 + (\epsilon_r/\hbar)^2}\,, \end{aligned} \tag{85}$$

where we have used $\Gamma_i = \nu T_i$ and $\Gamma \equiv \Gamma_1 + \Gamma_2$. Eq. (85), with its characteristic lorentzian lineshape, is known as the Breit-Wigner formula [32]. The Breit-Wigner formula is a good approximation of the lineshape close to a resonance, where $\epsilon \ll \hbar\nu$, provided $T_1, T_2 \ll 1$ (or, equivalently, $h\Gamma \ll \Delta E \equiv 1/\rho$). The deviations from the exact result (79) can be quite large if the barrier transparencies approach unity, as shown in Fig. 20.

The Breit-Wigner formula has a wider range of validity than to the one-dimensional tunneling problem considered here, where only a single conduction channel or 1D-subband in the electron gas regions adjacent to the potential well couples to the quasi-bound state in the well. As discussed by Büttiker [35], in the multi-channel case one simply has to replace the tunnel rates Γ_1 and Γ_2 by the sum of the tunnel rates of the localized state to each of the available conduction channels. This has the consequence that the maximum transmission probability due to resonant tunneling through a single localized state is unity, regardless of the number of scattering channels in the leads which couple to the localized state. This implies a maximum contribution of e^2/h to the conductance (for one spin direction) for each localized state in the well, the maximum

being realized if the tunneling rates through the two barriers are equal. This important result has been found following a different route by Kalmeyer and Laughlin [37], and by Xue and Lee [38].

This example and the one discussed in section 4 illustrate how e^2/h may show up in seemingly unrelated contexts (quantum ballistic transport through a point contact, and resonant tunneling through a localized state). Further examples are the quantum Hall effect, and universal conductance fluctuations[22]. Solid state electron optics is the viewpoint that transport properties are in essence transmission properties of the modes (or quantum channels) in the conductor. A single open channel universally contributes e^2/h to the conductance, which explains why this quantity is ubiquitous in quantum transport.

6. READING GUIDE

Fundamentals of the Analogy

1. R.P. Feynman, *Space-Time Approach to Non-Relativistic Quantum Mechanics*, Rev.Mod.Phys. **20**, 367 (1948). A lucid and self-contained description of Feynman's path-integral formulation of quantum mechanics, that may also be regarded as the final building block of the fundamental analogy between non-relativistic mechanics and optics.
2. R.P. Feynman, R.B. Leighton, and M. Sands, *The Feynman Lectures on Physics*, vol. II, chapter 19: *The Principle of Least Action* (Addison-Wesley, New York, 1964). A personal account of Feynman's fascination with variational principles, which led to his formulation of quantum mechanics.
3. M. Born and E. Wolf, *Principles of Optics*, (Pergamon, Oxford, 1965). Appendix II describes the basics of the analogy with electron optics.

Principles of Solid State Electron Optics

1. W. van Haeringen and D. Lenstra, eds. *Analogies in Optics and Micro-Electronics*, (Kluwer Academic, Dordrecht, 1990) and also Physica B **175** (1990). Two collections of articles on analogies, which provide an excellent reading guide for those interested in this field.
2. We would also like to refer to the Chapters in this volume by Agranovich, John, and Yablonovitch, which deal with particular analogies between optics and solid state transport.
3. R. Landauer, *Spatial Variation of Currents and Fields Due to Localized Scatterers in Metallic Conduction*, IBM J.Res.Dev. **1**, 223 (1957). The main message of this paper was that more attention should be given to the role of the strongly inhomogeneous (self-consistent) electric field in some transport problems, but it is most often cited because it pioneered the idea of a relation between conductance and the transmission probability: transport is viewed as a consequence of a flux of carriers, incident from a reservoir, rather than as a result of an applied electric field.
4. R. Landauer, *Conductance from Transmission: Common Sense Points*, Physica Scripta **T42**, 110 (1992). A highly readable discussion of some objections that one might raise (or that some have in fact raised) against the transmission approach, complete with arguments why they are irrelevant in many cases.
5. M. Büttiker, *Four-Terminal Phase-Coherent Conductance*, Phys.Rev.Lett. **57**, 1761 (1986). Another pioneering paper, that provided the generalization of the Landauer formula (as we know it today) to multi-probe measurements of the conductance. A reciprocity relation is derived for the conductance (generalizing Onsager's reciprocity relation to one for the conductivity).
6. M. Büttiker, *Symmetry of Electrical Conduction*, IBM J.Res.Dev. **32**, 317 (1988). A detailed and more tutorial account of the Landauer-Büttiker formalism.
7. L. Esaki and R. Tsu, *Superlattice and Negative Differential Conductivity in Semiconductors*, IBM J.Res.Dev. **14**, 61 (1970). The first paper on the idea of a semiconductor superlattice, which has had a tremendous influence on semiconductor physics ever since.
8. Y. Imry, *Physics of Mesoscopic Systems*, in *Directions in Condensed Matter Physics*, G. Grinstein and G. Mazenko, eds. (World Scientific Press, Singapore, 1986). A review that has set the agenda for much of the subsequent developments in mesoscopic physics.
9. V.S. Tsoi, J. Bass, and P. Wyder, *Transverse Electron Focusing as a Way of Studying Surface Crystallography*, Advances in Physics **41**, 365 (1992). A review of pioneering work on ballistic electron transport in metals.

10. H. van Houten, C.W.J. Beenakker, J.G. Williamson, M E.I Broekaart, P.H.M. van Loosdrecht, B.J. van Wees, J.E. Mooij, C.T. Foxon, and J.J. Harris, *Coherent electron focusing with quantum point contacts in a two dimensional electron gas*, Phys. Rev. B. **39**, 8556 (1989). This paper established the idea of solid state electron optics using quantum point contacts as coherent point sources, potential walls as mirrors, and a magnetic field as a lens.

Conductance Quantization and Transmission Steps

1. C.W.J. Beenakker and H. van Houten, *Quantum Transport in Semiconducting Nanostructures*, Solid State Physics **44**, 1 (1991) A review dealing with the regimes of coherent diffusion, quantum ballistic transport, lateral tunneling, and adiabatic transport in the quantum Hall effect regime.
2. M A. Reed, ed. *Nanostructured Systems*, a volume of Semiconductors and Semimetals, (Academic Press, New York, 1992). This book contains 4 extended reviews entitled *Quantum Point Contacts* by H van Houten C W J Beenakker and B J van Wees [illegible] *The Quantum Hall Effect in Open Conductors*, by M Buttiker, and *Electrons in Laterally Periodic Nanostructures*, by W. Hansen, J.P Kotthaus, and U Merkt.

Refraction and Tunneling

1. E. Burstein and S Lundquist, eds. *Tunneling Phenomena in Solids*, (Plenum, New York, 1969). Still the best introduction to solid-state tunneling.
2 L. Solymar, *Superconductive Tunneling and Applications*, (Chapman and Hall, London 1972) This book also contains a chapter on normal metal tunneling, and has an extensive bibliography
3. H. Grabert and M. Devoret, eds *Single Charge Tunneling. Coulomb Blockade Phenomena in Nanostructures*, NATO ASI Series B: vol. 294 (Plenum, New York, 1992). This book is a collection of review articles of tunneling beyond the independent electron approximation, where the analogy with optics no longer holds.
4. M. Büttiker, *Coherent and Sequential Tunneling in Series Barriers*, IBM J Res.Dev **32**, 63 (1988) Contains a good introduction and guide to the literature on resonant tunneling

7. ACKNOWLEDGEMENT

The authors thank M.A.M. de Jong and G.W. 't Hooft for valuable comments and discussions

REFERENCES

1. E. Hecht and A. Zajac, *Optics*, (Addison-Wesley, New York, 1974).
2 L D. Landau and E. M. Lifshitz, *Mechanics*, (Pergamon, Oxford, 1976)
3. H. Busch. *Berechnung der Bahn von Kathodenstrahlen im axialsymmetrischen elektromagnetischen Felde*, Ann Phys. Leipzig **81**, 974 (1926).
4. L. de Broglie, *Recherches sur la Théorie des Quanta*, Annales de Physique **10**, 22 (1925)
5. E. Schrödinger, *Quantisierung als Eigenwertproblem*, Ann. Phys. **79**, 32 (1926)
6 W. Elsasser, *Bemerkungen zur Quantenmechanik freier Elektronen*, Naturwiss. **13**, 711 (1925)
7. C. Davisson and L. H. Germer, *Diffraction of Electrons by a Crystal of Nickel*, Phys. Rev. **30**, 705 (1927)
8. R. P. Feynman, *Space-Time Approach to Non-Relativistic Quantum Mechanics*, Rev Mod.Phys. **20** (1948), 367
9. P. A. M. Dirac, Physikalische Zeitschrift der Sowjetunion, **3**, No. 1 (1933).
10. D. B. Beard and G. B. Beard, *Quantum Mechanics with Applications*, (Allyn and Bacon, Boston, 1970).
11. P. Ehrenfest, *Einige die Quantenmechanik betreffende Erkundigungsfragen*, Zeits. f. Physik **78**, 555 (1932).
12. D. Lenstra and W van Haeringen, *Playing with Electrons and Photons in Rings*, in *Analogies in Optics and Microelectronics*, W. van Haeringen and D. Lenstra, eds., (Kluwer, Dordrecht, 1990), p.3-19.
13. D. Lenstra, *Wave-Particle Analogies in Optics and Microelectronics*, in: *Studies in Mathematical Physics. Proc. Huygens Symposium*, E. van Groenen and E.M. de Jager, eds (North-Holland Amsterdam, 1992).
14. H. van Houten and C. W. J. Beenakker, *Andreev Reflection and the Josephson Effect in a Quantum Point Contact: an Analogy with Phase-Conjugating Resonators*, Physica B**175**, 187 (1991)

15 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** (1988) 848 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 (1991)
16 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 Quantization of the Ballistic Resistance*, J Phys C**21** (1988) L209
17 A Yacoby and Y Imry, *Quantization of the Conductance of Ballistic Point Contacts Beyond the Adiabatic Approximation* Phys Rev B **41**, 5341 (1990)
18 L I Glazman G B Lesovick D E Khmel'nitskii and R I Shekhter *Reflectionless Quantum Transport and Fundamental Ballistic Resistance Steps in Microscopic Constrictions* **48** 218 (1988) [JETP Lett **48** 238 (1988)]
19 C W J Beenakker and H van Houten, *Magnetotransport and Nonadditivity of Point-Contact Resistances in Series*, Phys Rev B **39**, 10445 (1989)
20 L W Molenkamp, A A M Staring, C W J Beenakker R Eppenga, C E Timmering J G Williamson, C J P M Harmans, and C T Foxon, *Electron-Beam Collimation with Quantum Point Contacts*, Phys Rev B **41** 1274 (1990)
21 A Szafer and A D Stone, *Theory of Quantum Conduction through a Constriction* Phys Rev Lett **62**, 300 (1989)
22 For a review see C W J Beenakker and H van Houten, *Quantum transport in semiconductor nanostructures*, Solid State Physics **44** 1 (1990)
23 E A Montie, E C Cosman, G W 't Hooft, M B van der Mark, and C W J Beenakker *Observation of the Optical Analogue of Quantized Conductance of a Point Contact* Nature **350** 594 (1991), see also Physica **B 175**, 149 (1991)
24 H van Houten and C W J Beenakker, *Quantum Point Contacts and Coherent Electron Focusing* in *Analogies in Optics and Microelectronics*, W van Haeringen and D Lenstra eds (Kluwer Dordrecht, 1990)
25 H van Houten, B J van Wees, J E Mooij, C W J Beenakker, J G Williamson and C T Foxon *Coherent Electron Focusing in a Two-Dimensional Electron Gas*, Europhys Lett **5** 721 (1988)
26 E Sivan, M Heiblum, C P Umbach, and H Shtrikman, *Electrostatic Lens in the Ballistic Regime* Phys Rev B **41**, 7937 (1990)
27 J Spector, H L Stormer, K W Baldwin, L N Pfeiffer and K W West Appl Phys Lett *Electron Focusing in Two-Dimensional Systems by Means of an Electrostatic Lens* **56** 1290 (1990) J Spector, J S Weiner H L Stormer, K W Baldwin, L N Pfeiffer, and K W West *Ballistic Electron Optics*, Surf Sci **263** 240 (1992)
28 S Zhu, A W Yu, D Hawley, and R Roy *Frustrated Total Internal Reflection A Demonstration and Review* Am J Phys **54**, 601 (1986)
29 R Y Chiao, P G Kwiat, and A M Steinberg, *Analogies between Electron and Photon Tunneling* Physica B**175**, 257 (1991)
30 W A Harrison *Solid State Theory*, (McGraw Hill, New York 1970)
31 S Washburn, A B Fowler, H Schmid and D Kern, *Possible Observation of Transmission Resonances in GaAs-AlGaAs Transistors*, Phys Rev B **38**, 1554 (1988)
32 G Breit and E Wigner, *Capture of Slow Neutrons* Phys Rev **49**, 519 (1936)
33 R Tsu and L Esaki, *Tunneling in a Finite Superlattice*, Appl Phys Lett **22** 562 (1973)
34 L L Chang, L Esaki and R Tsu, *Resonant Tunneling in Semiconductor Double Barriers*, Appl Phys Lett **24**, 593 (1974)
35 M Buttiker, *Coherent and Sequential Tunneling in Series Barriers*, IBM J Res Dev **32** 63 (1988)
36 E Burstein and S Lundqvist, eds *Tunneling Phenomena in Solids*, (Plenum, New York 1969)
37 V Kalmeyer and R B Laughlin, *Differential Conductance in Three-Dimensional Resonant Tunneling*, Phys Rev B **35**, 9805 (1987)
38 W Xue and P A Lee *Two-Dimensional Resonant Tunneling*, Phys Rev B **38**, 3913 (1988)

Henk van Houten is department head of the experimental physics group of Philips Research Laboratories in Eindhoven and professor of solid state physics at Leiden University Apart from

his duties as an industrial research manager, he maintains a strong interest in mesoscopic transport phenomena

Carlo Beenakker is professor of theoretical physics at the Institute Lorentz of the University of Leiden where he conducts research on electronic properties of mesoscopic systems