Quantum Point Contacts

The quantization of ballistic electron transport through a constriction demonstrates that conduction is transmission.

Published in abbreviated form in Physics Today, July 1996, page 22.

Henk van Houten & Carlo Beenakker

Punctuated equilibrium, the notion that evolution in nature is stepwise rather than continuous, sometimes applies to evolution in science as well. It happens that the seed of a scientific breakthrough slumbers for a decade or even longer, without generating much interest. The seed can be a theoretical concept without clear predictions to test experimentally, or an intriguing but confusing experiment without a lucid interpretation. When the seed finally germinates, an entire field of science can reach maturity in a few years.

In hindsight, this is what happened ten years ago, when the authors (newly hired PhD’s at Philips Research in Eindhoven) ventured into the field of quantum ballistic transport. Together with Bart van Wees, then a graduate student at Delft University of Technology, we were confronted with some pretty vague challenges. On the experimental side, there was the search for a quantum-size effect on the conductance, which would reveal in a clear-cut way the one-dimensional density of states of electrons confined to a narrow wire. Experiments on narrow silicon transistors (at Yale University and AT&T Bell Labs., Holmdel) had come close, but suffered from irregularities due to disorder. (These irregularities would become known as “universal conductance fluctuations”, see Physics Today, December 1988, page 36.) We anticipated that the electron motion should be ballistic, i.e. without scattering by impurities. Moty Heiblum (IBM, Yorktown Heights) had demonstrated ballistic transport of hot electrons, high above the Fermi level. For a quantum-size effect one needs ballistic motion at the Fermi energy. Our colleague Thomas Foxon from Philips Research in Redhill (UK) could provide us with heterojunctions of GaAs and AlGaAs, containing at the interface a thin layer of highly mobile electrons. Such a “two-dimensional electron gas” seemed an ideal system for ballistic transport.

On the theoretical side, there was the debate whether a wire without impurities could have any resistance at all. Ultimately, the question was: “What is measured when you measure a resistance?” The conventional point of view (held in the classical Drude-Sommerfeld or the quantum mechanical Kubo theories) is that conduction is the flow of current in response to an electric field. An alternative point of view was put forward in 1957 by Rolf Landauer (IBM, Yorktown Heights), who proposed that “conduction is transmission” Landauer’s formula, a relationship between conductance and transmission probability, had evolved into two versions. One gave infinite conductance (= zero resistance) in the absence of impurity scattering, while the other gave a finite answer. Although the origin of the difference between the two versions was understood by at least one of the theorists involved in the debate, the experimental implications remained unclear.

Looking back ten years later, we find that the seed planted by Landauer in the fifties has developed into a sophisticated theory, at the basis of the entire field of quantum ballistic transport. The breakthrough can be traced back to experiments on an elementary conductor: a point contact. In this article we present a brief account of these developments. For a more comprehensive and detailed discussion, we direct the reader to the reviews in the bibliography.

Quantized conductance

The history of ballistic transport goes back to 1965, when Yuri Sharvin (Moscow) used a pair of point contacts to inject and detect a beam of electrons in a single-crystalline metal. In such experiments the quantum mechanical wave character of the electrons does not play an essential role, because the Fermi wave length (\( \lambda_F \approx 0.5 \text{ nm} \)) is much smaller than the opening of the point contact. The two-dimensional (2D) electron gas in a GaAs–AlGaAs heterojunction has a Fermi wave length which is a hundred times larger than in a metal. This makes it possible to study a constriction with an opening comparable to the wave length (and much smaller than the mean free path for impurity scattering). Such a constriction is called a quantum point contact.

In a metal a point contact is fabricated simply by pressing two wedge- or needle-shaped pieces of material together. A quantum point contact requires a more complicated strategy, since the 2D electron gas is confined at the GaAs–AlGaAs interface in the interior of the het-
A point contact of adjustable width can be created in this system using the split-gate technique developed in the groups of Michael Pepper (Cambridge) and Daniel Tsui (Princeton). The gate is a negatively charged electrode on top of the heterojunction, which depletes the electron gas beneath it. (See figure 1.) In 1988, the Delft-Philips and Cambridge groups reported the discovery of a sequence of steps in the conductance of a constriction in a 2D electron gas, as its width $W$ was varied by means of the voltage on the gate. (See Physics Today, November 1988, page 21.) As shown in figure 2, the steps are near integer multiples of $2e^2/h \approx 1/13 \text{kΩ}$ (after correction for a small gate-voltage independent series resistance).

An elementary explanation of the quantization views the constriction as an electron wave guide, through which a small integer number $N \approx 2W/\lambda_F$ of transverse modes can propagate at the Fermi level. The wide regions at opposite sides of the constriction are reservoirs of electrons in local equilibrium. A voltage difference $V$ between the reservoirs induces a current $I$ through the constriction, equally distributed among the $N$ modes. This equipartition rule is not immediately obvious, because electrons at the Fermi level in each mode have different group velocities $v_n$. However, the difference in group velocity is canceled by the difference in density of states $\rho_n = 1/h v_n$. As a result, each mode carries the same current $I_n = V e^2 \rho_n v_n = V e^2 / h$. Summing over all modes in the wave guide, one obtains the conductance $G = I/V = Ne^2/h$. The experimental step size is twice $e^2/h$ because spin-up and spin-down modes are degenerate.

The electron wave guide has a non-zero resistance even though there are no impurities, because of the reflections occurring when a small number of propagating modes in the wave guide is matched to a larger number of modes in the reservoirs. A thorough understanding of this mode-matching problem is now available, thanks to the efforts of many investigators.

The quantized conductance of a point contact provides firm experimental support for the Landauer formula,

$$G = \frac{2e^2}{h} \sum_n t_n,$$

for the conductance of a disordered metal between two electron reservoirs. The numbers $t_n$ between 0 and 1 are the eigenvalues of the product $tt^\dagger$ of the transmission matrix $t$ and its Hermitian conjugate. For an “ideal” quantum point contact $N$ eigenvalues are equal to 1 and all others are equal to 0. Deviations from exact quantization in a realistic geometry are about 1%. This can be contrasted with the quantization of the Hall conductance in strong magnetic fields, where an accuracy better than 1 part in $10^7$ is obtained routinely. One reason why a similar accuracy can not be achieved in zero magnetic field is the series resistance from the wide regions, whose magnitude can not be determined precisely. Another source of excess resistance is backscattering at the entrance and exit of the constriction, due to the abrupt widening of the geometry. A magnetic field suppresses this backscattering, improving the accuracy of the quan-
Contact of atomic dimensions is so large that the con-

The propagating modes in the quantum Hall effect are the magnetic Lan-

ductance steps are visible at room temperature. Nicolás Garcia and his group at the Autonomous University of Madrid have made use of this property to develop a classroom experiment of quantized conductance. (See Physics Today, February 1996, page 9.)

Photons and Cooper pairs

The interpretation of conduction as transmission of electrons at the Fermi level suggests an analogy with the transmission of monochromatic light. The analogue of the conductance is the transmission cross-section $\sigma$, defined as the transmitted power divided by the incident flux. Figure 4 shows the transmission cross-section of a slit of variable width, measured by Edwin Montie and collaborators from Philips. Steps of equal height occur whenever the slit width $W$ equals half the wave length $\lambda = 1.55 \mu m$ of the light. Because $\sigma$ equals $W$ for large slit widths, the step height is also equal to $\lambda/2$. Two-dimensional isotropic illumination was achieved by passing the light through a random array of glass fibres parallel to the slit. The isotropy of the illumination mimics
the reservoirs in the electronic case, and is crucial for the effect. The two-dimensionality is not essential, but was chosen because a diaphragm of variable area of the order of $\lambda^2$ is difficult to fabricate. (For a diaphragm, the steps in $\sigma$ are $\lambda^2/2\pi$.)

It is remarkable that this optical phenomenon, with its distinctly nineteenth century flavour, was not noticed prior to the discovery of its electronic counterpart. There is an interesting parallel in the history of the discovery of the two phenomena. In the electronic case, the Landauer formula was already known before the quantized conductance of a point contact was discovered. Yoseph Imry (Weizmann Institute, Israel) had made the connection with Sharvin’s work on point contacts. The reason that the conductance quantization came as a surprise, was that the relation $\sum t_n = N$ for ballistic transport was regarded as an order-of-magnitude estimate. To have quantization, the relative error in this estimate must be smaller than $1/N$, which is not obvious. The equivalent of the Landauer formula for the transmission cross-section has long been familiar in optics but also in this field it was not noticed that $\sum t_n = N$ holds with better than $1/N$ relative accuracy.

One can speak of the optical analogue as a quantum point contact for photons. Can the analogue be extended towards a quantum point contact for Cooper pairs? The answer is “Yes”: The maximal supercurrent through a narrow and short, impurity-free constriction in a superconductor is an integer multiple of $e\Delta/h$, with $\Delta$ the energy gap of the bulk superconductor. A superconducting quantum point contact has been realized by Hideaki Takayanagi and collaborators (NTT, Japan) but the superconducting analogue of the quantized conductance remains to be observed experimentally.

**Thermal analogues**

The conductance is the coefficient of proportionality between current and voltage. The additional presence of a small temperature difference $\delta T$ across the point contact gives rise to a matrix of coefficients:

$$
\begin{pmatrix}
\text{electrical current} \\
\text{heat current}
\end{pmatrix}
= \begin{pmatrix}
G & L \\
L' & K
\end{pmatrix}
\begin{pmatrix}
-V \\
\delta T
\end{pmatrix}.
$$

The thermal conductance $K$ relates heat current to temperature difference. The thermo-electric cross-phenomena are described by coefficients $L$ and $L'$. As first deduced by Lord Kelvin, time-reversal symmetry requires that $L' = -LT$ (at a temperature $T$).

The two new transport coefficients $K$ and $L$ can be expressed in terms of the transmission probabilities, just like the electrical conductance $G$. (See sidebar.) Approximately, $K \propto t$ and $L \propto dt/dE_F$, where $t = \sum t_n$ is the total transmission probability at the Fermi energy $E_F$. (The proportionality of $K$ to $t$, and hence to $G$, is the Wiedemann-Franz law of solid-state physics.) The step-wise energy dependence of the transmission probability through a quantum point contact implies two types of quantum-size effects: steps in $K$ and peaks in $L$. Both effects have been observed by Laurens Molenkamp and collaborators from Philips.

The thermal conductance $K$ of a quantum point contact exhibits steps when the gate voltage is varied, aligned with the steps in the electrical conductance. Each step signals the appearance of a new mode at the Fermi level which can propagate through the constriction. A step in the transmission probability leads to a peak in the thermo-electric coefficient.

In figure 5 we show measurements of the thermopower $S = -L/G$ of a quantum point contact. (The thermopower is proportional to the voltage produced by a temperature difference for zero electrical current.) The coincidence of peaks in the thermopower with steps in the conductance (measured for the same point contact) is clearly visible. Joule heating was used to create a temperature difference across the point contact in this work. Local heating by means of a focused beam of far-infrared radiation has been used in a more recent experiment.

**Shot noise**

The electrical current through a point contact is not
constant in time, but fluctuates. The conductance determines only the time-averaged current. The noise power $P = 2 \int dt \langle \delta I(0) \delta I(t) \rangle \cos \omega t$ at frequency $\omega$ is the Fourier transform of the correlator of the time-dependent fluctuations $\delta I(t)$ in the current at a given voltage $V$ and temperature $T$. One distinguishes equilibrium thermal noise ($V = 0$, $T \neq 0$) and non-equilibrium shot noise ($V \neq 0$, $T = 0$). Both types of noise have a white power spectrum (i.e. the noise power does not depend on frequency over a very wide frequency range). Thermal noise is directly related to the conductance through the fluctuation-dissipation theorem ($P_{\text{thermal}} = 4kT G$). Therefore, the thermal noise of a quantum point contact does not give any new information.

Shot noise is more interesting, because it contains information on the temporal correlation of the electrons which is not contained in the conductance. Maximal shot noise ($P_{\text{max}} = 2eI$) is observed when the stream of electrons is fully uncorrelated. A typical example is a tunnel diode. Correlations reduce $P$ below $P_{\text{max}}$. One source of correlations, operative even for non-interacting electrons, is the Pauli principle, which forbids multiple occupation of the same single-particle state. A typical example is a ballistic point contact in a metal, where the stream of electrons is completely correlated by the Pauli principle in the absence of impurity scattering.

A quantum point contact in a 2D electron gas has a different behavior. Using a Landauer-type formula (see sidebar), Gordey Lesovik (Moscow) has predicted peaks in the shot noise at the steps in the conductance. The peak height $P_{\text{peak}} = e I$ is half the maximal value for uncorrelated electrons. The shot noise vanishes in between the steps. Michael Reznikov and collaborators from the Weizmann Institute in Israel have recently presented a convincing demonstration of this quantum-size effect in the shot noise (See figure 6). By going to microwave frequencies (8–18 GHz) they could avoid the ubiquitous “1/f noise” at lower frequencies.

**Solid-state electron optics**

The effects discussed so far refer to properties of the quantum point contact itself. A wealth of new phenomena has been discovered using a quantum point contact as a spatially coherent point source and detector, and specially formed electrodes as mirror, prism, or lens.

The basic experiment, coherent electron focusing, is shown in figure 4. A point contact injects electrons with the Fermi momentum $p_F$ into the 2D electron gas, in the presence of a perpendicular magnetic field $B$. The electrons follow a “skipping orbit” along the boundary, consisting of circular arcs of cyclotron diameter $d = 2p_F/eB$. Some of the electrons are collected at a second point contact, at a separation $L$ from the first. The voltage measured at the collector is proportional to the transmission probability between the two point contacts. V. S. Tsoi (Moscow) first used this focusing technique in a metal. The magnetic field acts as a lens, bringing the divergent trajectories at the injector together at the collector. The collector is at a focal point of the lens when $L$ is a multiple of $d_c$, hence when $B$ is a multiple of $2p_F/eL$ (arrows in figure 7). For reverse magnetic fields the injected electrons are deflected away from the collector, so that no signal is generated. The observation of peaks at the expected positions demonstrates that a quantum point contact acts as a monochromatic point source of ballistic electrons, and that the reflections at the boundary of the 2D electron gas are specular. The fine-structure on the focusing peaks is due to quantum interference of trajectories between the two point contacts. Such fine-structure does not appear in metals. It demonstrates that the quantum point contact is a spatially coherent source and that the phase coherence is maintained over a distance of several microns to the col-
FIG. 7 Magnetic focusing in a 2D electron gas at 50 mK. The top panel shows the experimental arrangement. Electrons injected through one point contact (i) follow skipping orbits over a distance of 3.0 µm to a second point contact (c). The arrows indicate the positions of the focusing peaks expected when the point contact separation is a multiple of the cyclotron diameter. The fine-structure on the peaks is due to quantum interference. (Adapted from ref. 24.)

Magnetic focusing has been used by several groups to obtain information on the dynamics and scattering of quasiparticles in the 2D electron gas. An intriguing application in the regime of the fractional quantum Hall effect is the focusing of composite fermions, which can be thought of as electrons bound to an even number of flux quanta. In the regime of the integer quantum Hall effect, the geometry of figure 7 has been used to selectively populate and detect the magnetic edge states mentioned earlier. The observation of plateaus in the Hall conductance at anomalously quantized values provides support for the edge-state theory of the quantum Hall effect.

Electrostatic focusing, by means of the electric field produced by a lens-shaped electrode, provides an alternative technique to focus the beam of electrons injected by a point contact. Instead of focusing the beam, one can also deflect it — either by means of a magnetic field or by means of a prism-shaped electrode. The building blocks of electron optics in the solid state have by now all been realized.

Ultimate confinement

A quantum point contact which is nearly pinched off (so that its conductance is less than $2e^2/h$) is a tunnel barrier of adjustable height for electrons near the Fermi level. This property has been used to inject and detect electrons in a small confined region of a 2D electron gas, called a quantum dot. A quantum dot coupled to the outside by a pair of quantum point contacts has provided an ideal model system for the investigation of the effects of Coulomb repulsion on resonant tunneling. (See PHYSICS TODAY, January 1993, page 24.)

The zero-dimensional quantum dot forms the logical end to the reduction of dimensionality of the two-dimensional electron gas. In this article we have reviewed the role played by the one-dimensional quantum point contact in the conceptual development started by Landauer four decades ago. The concept of electrical conductance was conceived in the nineteenth century, at a time when the electron was not even discovered. It is amusing that it required the sophisticated micro-electronics technology of the late twentieth century to demonstrate experimentally that “conduction is transmission”.

Landauer formulas

Landauer’s original 1957 formula:

$$G = \frac{2e^2}{h}rac{t}{1-t}.$$  

expresses the conductance of a one-dimensional system as the ratio of transmission and reflection probabilities. As explained by Imry, this formula gives infinity for unit transmission because it excludes the finite contact conductance contained in:

$$G = \frac{2e^2}{h}t.$$  

Extension to higher dimensions is achieved by replacing the transmission probability $t$ by the eigenvalue $t_n$ of the transmission matrix product $tt^\dagger$, and summing over $n$.

Generalizations of the Landauer formula have been found for a variety of other transport properties, besides the conductance. At zero temperature, these expressions are of the form:

$$\text{transport property} = A_0 \sum_n a(t_n),$$

- conductance $G$: $A_0 = 2e^2/h$, $a(t) = t$.
- shot-noise power $P$: $A_0 = 4e^2V/h$, $a(t) = t/(1-t)$.
- conductance $G_{NS}$ of a normal-metal–superconductor junction: $A_0 = 4e^2/h$, $a(t) = t^2(2-t)^{-2}$.
- supercurrent $I$ through a Josephson junction with phase difference $\phi$: $A_0 = e\Delta/h$, $a(t) = \frac{1}{t}t\sin\phi(1 - t\sin^2\phi/2)^{-1/2}$. 

The expressions for the thermo-electric coefficients involve an integration over energies around the Fermi energy $E_F$, weighted by the derivative $f' = df/dE$ of the Fermi-Dirac distribution function at temperature $T$:

$$\text{transport property} = -A_0 \int dE (E - E_F)^p f' \sum_n t_n,$$

- electrical conductance $G$: $p = 0$, $A_0 = 2e^2/h$.
- thermo-electric coefficient $L$: $p = 1$, $A_0 = 2e/h T$.
- thermal conductance $K$: $p = 2$, $A_0 = -2/h T$.

If the energy-dependence of the transmission eigenvalues involves an oscillatory dependence in the case of $G$, $G_{NS}$, $I(\phi)$, $K$, and an oscillatory dependence in the case of $P, L$.

### References

[1] For an historical account of the controversy, see R. Landauer, IBM J. Res. Dev. 32, 306 (1988); A. D. Stone, A. Szafer, ibidem, page 384. For a more recent review and a comprehensive list of references, see Y. Imry, Introduction to Mesoscopic Physics, Oxford University Press (1996).

[2] R. Landauer, IBM J. Res. Dev. 1, 223 (1957).

[3] Y. Imry, in: Directions in Condensed Matter Physics, eds. G. Grinstein and G. Mazenko, World Scientific, Singapore (1986).

[4] Yu. V. Sharvin, Sov. Phys. JETP 21, 655 (1965).

[5] T. J. Thornton, M. Pepper, H. Ahmed, D. Andrews, G. J. Davies, Phys. Rev. Lett. 56, 1198 (1986); H. Z. Zheng, H. P. Wei, D. C. Tsui, G. Weimann, Phys. Rev. B 34, 5635 (1986).

[6] B. J. van Wees, H. van Houten, C. W. J. Beenakker, J. G. Williamson, L. P. Kouwenhoven, D. van der Marel, C. T. Foxon, Phys. Rev. Lett. 60, 848 (1988); Phys. Rev. B 43, 12431 (1991).

[7] 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, G. A. C. Jones, J. Phys. C 21, L209 (1988).

[8] An early theoretical paper is by L. I. Glazman, G. B. Lesovik, D. E. Khmel’nitskii, R. I. Shekhter, JETP Lett. 48, 238 (1988). More references are given in the review by C. W. J. Beenakker, H. van Houten, Sol. State Phys. 44, 1 (1991).

[9] M. E. Cage, in: The Quantum Hall Effect, eds. R. E. Prange and S. M. Girvin, Springer, New York (1990).

[10] M. Büttiker, Phys. Rev. B 38, 9375 (1988).

[11] J. K. Gimzewski, R. Möller, Phys. Rev. B 36, 1284 (1987).

[12] J. M. Kraus, J. M. van Ruitenbeek, V. V. Fisun, I. K. Yanson, L. J. de Jongh, Nature 375, 767 (1995).

[13] H. van Houten, C. W. J. Beenakker, in: Analogies in Optics and Micro Electronics, eds. W. van Haeringen and D. Lenstra, Kluwer, Dordrecht (1990).

[14] E. A. Montie, E. C. Cosman, G. W. ’t Hooft, M. B. van der Mark, C. W. J. Beenakker, Nature 43, 594 (1991).

[15] A. W. Snyder, C. Pask, J. Opt. Soc. Am. 63, 806 (1973).

[16] C. W. J. Beenakker, H. van Houten, Phys. Rev. Lett. 66, 3056 (1991); A. Furusaki, H. Takayanagi, M. Tsukada, Phys. Rev. B 45, 10563 (1992).

[17] H. Takayanagi, T. Akazaki, J. Nitta, Phys. Rev. Lett. 75, 3333 (1995).

[18] P. Stáeda, J. Phys. Condens. Matter 1, 1025 (1989).

[19] A review of thermo-electric effects in quantum point contacts is: H. van Houten, L. W. Molenkamp, C. W. J. Beenakker, C. T. Foxon, Semicond. Sci. Technol. 7, B215 (1992).

[20] L. W. Molenkamp, H. van Houten, C. W. J. Beenakker, R. Eppenga, C. T. Foxon, Phys. Rev. Lett. 65, 1052 (1990).

[21] R. A. Wyss, C. C. Eugster, J. A. del Alamo, Q. Hu, M. J. Rooks, M. R. Melloch, Appl. Phys. Lett. 66, 1144 (1995).

[22] G. B. Lesovik, JETP Lett. 49, 592 (1989).

[23] M. Reznikov, M. Heiblum, H. Shtrikman, D. Mahalu, Phys. Rev. Lett. 75, 3340 (1995).

[24] H. van Houten, B. J. van Wees, J. E. Mooij, C. W. J. Beenakker, J. G. Williamson, C. T. Foxon, Europhys. Lett. 5, 721 (1988); Phys. Rev. B 39, 8556 (1989).

[25] V. S. Tsai, JETP Lett. 19, 70 (1974).

[26] V. J. Goldman, B. Su, J. K. Jain, Phys. Rev. Lett. 72, 2065 (1994).

[27] U. Sivan, M. Heiblum, C. P. Umbach, H. Shtrikman, Phys. Rev. B 41, 7937 (1990).

[28] L. W. Molenkamp, A. M. Staring, C. W. J. Beenakker, R. Eppenga, C. E. Timmering, J. G. Williamson, C. J. P. M. Harmans, C. T. Foxon, Phys. Rev. B 41, 1274 (1990).

[29] J. Spector, H. L. Stormer, K. W. Baldwin, L. N. Pfeiffer, K. W. West, Appl. Phys. Lett. 56, 2433 (1990).

[30] E. N. Economou, C. M. Soukoulis, Phys. Rev. Lett. 46, 618 (1981).

[31] D. S. Fisher, P. A. Lee, Phys. Rev. B 23, 6851 (1981).

[32] The formulas for $K$ and $L$ are due to Uri Sivan and Imry [Phys. Rev. B 33, 551 (1986)], for $P$ to Büttiker [Phys. Rev. Lett. 65, 2901 (1990)], and for $G_{NS}$ and $I(\phi)$ to Beenakker [Phys. Rev. Lett. 67, 3836 (1991); Phys. Rev. B 46, 12841 (1992)].