Universal signature of ballistic transport in nanoscale field effect transistors

A. Schliemann
Technische Physik,
Universitat Wurzburg
L. Worschech
Technische Physik,
Universitat Wurzburg
A. Forchel
Technische Physik,
Universitat Wurzburg
G.Curatola
Dipartimento di Ingegneria dell'Informazione: Elettronica, Informatica, Telecomunicazioni,
Università di Pisa
Giuseppe Iannaccone
Dipartimento di Ingegneria dell'Informazione: Elettronica, Informatica, Telecomunicazioni,
Università di Pisa

A. Schliemann, L. Worschech, A. Forchel, G. Curatola, G. Iannaccone, Universal signature of ballistic transport in nanoscale field effect transistors , International Electron Device Meeting 2004, pp. 1039-1042, San Francisco 2004.

Universal signature of ballistic transport in nanoscale field effect transistors

A. Schliemann[1], L. Worschech[1], A. Forchel[1], G. Curatola[2], G. Iannaccone[2]
[1]Technische Physik, Universität Würzburg, Am Hubland, D-97074 Würzburg, Germany
[2]Dipartimento di Ingegneria dell'Informazione, Università degli Studi di Pisa, Via Caruso 56122, Pisa, Italy

Abstract

We present a universal signature of ballistic transport in field effect transistors with channel lengths down to 25 nm and widths up to 100 μm. By measuring with a two and four-probe setup the derivative of the differential conductance with respect to the gate voltage for finite drain bias a splitting is obtained, which is shown to be a clear indication of electrons traversing the channel conserving their initial energy.

Introduction

Ballistic transport has been observed for the first time in semiconductors with specially devised spectrometers [1–3]. In one-dimensional quantum wires ballistic transport results in the quantization of conductance in steps of $2e^2/h$ [4, 5]. When narrow constrictions are occupied by only a few modes coherent electron flow allows the observation of interference effects similar to wave phenomena well known for coherent light [6–8]. But ballistic transport can occur in more common devices operating far from the conductance quantization regime, such as nanoscale FETs. In this work, we present a universal signature of ballistic transport in multimode devices. This signature is observed in AlGaAs/GaAs heterostructure FETs (HFETs) with channel lengths down to 25 nm. We demonstrate that ballistic transistors show a peak splitting of the derivative of the four-probe differential conductance with respect to the gate voltage for finite drain bias.

Results and discussion

We have realized ultra-short FETs on modulation doped GaAs/AlGaAs heterostructures with a two-dimensional electron gas 120nm below the sample surface.

The mobility of the 2DEG has been measured to be 2.5×10^6 cm^2 V^{-1} s^{-1} at 4.2 K, with the electron concentration of 3-4 $\times 10^{11}$ cm^{-2}. Such parameters lead to an average scattering length at Fermi energy of 31-36 μm, which is a factor thousand larger than the gate length, and allows us to exclude the diffusive transport regime for the FETs.

A scheme of the FET is shown in Fig.1A. The gate electrode separates two wider regions serving as source and drain reservoirs. The scanning electron micrograph in Fig.1A shows a partial view of the gate section at which a tapered gate is connected with a narrow top gate. The gate length of the present structure was 25 nm. A sketch of the corresponding conduction band profile is depicted in Fig.1B. A voltage difference $V_{DS} = (\mu_S - \mu_D)/e$ applied between the drain and the source, with μ_D and μ_S the electrochemical potentials of the drain and the source, respectively, and e the electron charge, leads to an efficient current as long as the barrier maximum E_{MAX} is smaller than μ_S. The current depends on the voltage V_{DS} and the barrier height, which is controlled by the gate.

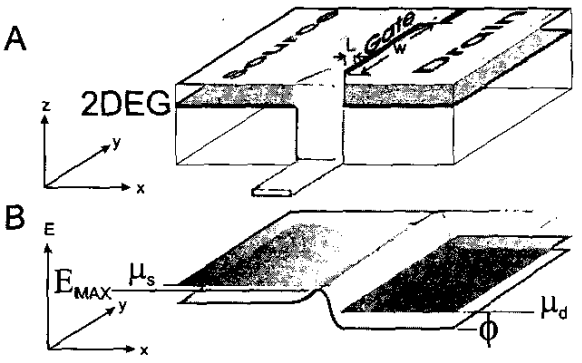

FIG. 1: **A**: Schematic representations of the field effect transistors with a scanning electron micrograph showing the gate section, at which the tapered gate is connected with the narrow, top gate. **B**: The conduction band energies and electrochemical potentials across the channel.

We have measured the current I_{DS} through the transistor for varying voltages V_{DS} applied between the drain and the source of the FETs and V_{GS} applied between the gate and the source. The samples were immersed in liquid He at a temperature of 4.2 K. In the upper part of Fig.2A the differential conductance $G = dI_{DS}/dV_{DS}$ of an FET with 50 nm gate length is plotted versus the gate voltage V_{GS} for $V_{DS} = 1$, 20 and 40 mV. For $V_{DS} = 1$ mV the FET channel is getting conductive at $V_{GS} > -0.62$ V. G increases from 0 to a maximum value of 7.4×10^{-4} AV^{-1} with increasing gate voltage up to -0.52 V. A further increase in the gate voltage does not change the conductance. For $V_{DS} = 20$ mV the onset of conductance is shifted by about 50 mV to lower gate voltages compared to $V_{DS} = 1$ mV. Conversely, the gate voltage at which G saturates is shifted 50 mV to larger values. Starting from the threshold gate voltage with increasing V_{GS} a similar slope of the conduc-

0-7803-8684-1/04/$20.00 ©2004 IEEE

tance dG/dV_{GS} compared to V_{DS} = 1 mV is found up to $G = 4.4 \times 10^{-4}$ AV^{-1}. A further increase of the gate voltage for -0.62 V $< V_{GS} < -0.55$ V results in a reduction of dG/dV_{GS}. The larger dG/dV_{GS} is recovered for the gate sweep region -0.55 V $< V_{GS}$ until G remains constant for increasing V_{GS}. When V_{DS} is increased to 40 mV the FET is getting conductive for a smaller gate voltage. It is interesting to note that here as well the gate voltage above which G remains constant is shifted to a larger value compared to $G(V_{GS})$ traces detected for smaller drain voltages. As shown in the lower part of Fig.2A, the derivative dG/dV_{GS} of the differential conductance with respect to the gate voltage shows a narrow peak for small bias voltages. The peak splits into two well-resolved components when the drain to source voltage is increased as shown for V_{DS} = 1 to 10 mV in steps of 1 mV and V_{DS} = 20 and 40 mV in the lower part of Fig. 2A. We found that for a constant drain to source voltage the peak splitting is getting larger the smaller the gate length of the FET. Fig.2B shows a gray-scale plot of dG/dV_{GS} as a function of gate voltage and drain bias voltage for an FET with 25 nm gate length. Bright contrast reflects a large derivative of the differential conductance and dark contrast small values. For small V_{DS} a clear peak is found for $V_{GS} = -6.5$ V. The peak splits linearly with increasing drain to source bias voltage up to 3 V for V_{DS} = 40 mV.

We interpret the peak splitting of dG/dV_{GS} with increasing drain to source bias voltage by taking into account the ballistic transport of electrons through the channel. As electrons in the channel of the FET are strongly confined in the vertical direction, their density of states is given by contributing of two-dimensional subbands. Typically only the lowest subband is populated, and the subband edge has the profile sketched in the left part of Fig. 3. E_{MAX} is the maximum of the subband in the channel, which is controlled by the gate to source voltage, V_{GS}. For a constant non-zero drain to source voltage V_{DS} the conduction band in the channel decreases with gradually increasing gate to source voltage V_{GS}, and E_{MAX} decreases correspondingly. The two-dimensional subband edge profile shown in the left part of Fig. 3 (i) corresponds to a low V_{GS}. When E_{MAX} is larger than μ_S, the differential conductance is negligible. As E_{MAX} aligns with μ_S (Fig. 3 (ii)) the differential conductance increases abruptly, because of the sharpness of the Fermi-Dirac distribution. Then, as long as $\mu_S > E_{MAX} > \mu_D$, the ballistic drain current is only due to electrons with energy in the x direction between E_{MAX} and μ_S and is proportional to $F_{1/2}[(\mu_S - E_{MAX})/kT]$, where $F_{1/2}$ is the Fermi-Dirac integral of order 1/2 [9]. Therefore, the current depends on V_{GS} and V_{DS} only through E_{MAX}. To first order, we can write

$$E_{MAX} = E_{MAX0} - e\alpha(V_{GS} - V_{GS0}) - e\beta(V_{DS} - V_{DS0}), \quad (1)$$

where E_{MAX0} is the value of E_{MAX} for the gate-to-source bias V_{GS0} and drain-to-source bias V_{DS0}.

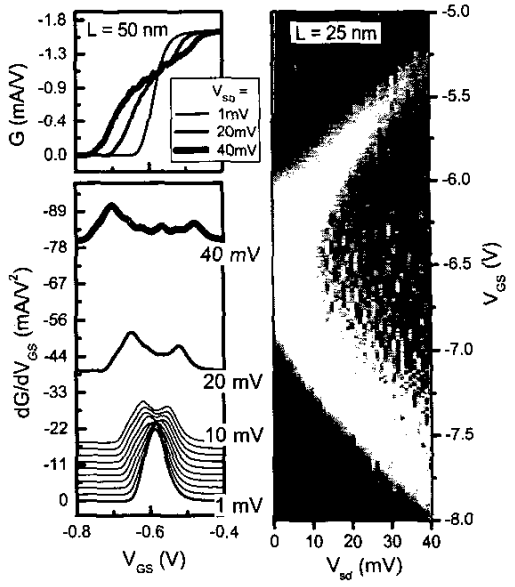

FIG. 2: Peak splitting of the derivative of the differential conductance for constant drain to source bias voltage V_{DS} of ballistic field effect transistors at T=4 K. **Top left**: Step like differential conductance for source drain bias voltages V_{DS} = 1, 20 and 40mV. **Bottom left**: Derivative of differential conductance for a 50 nm gate length FET. **Right hand side**: Grey scale plot of the derivative of the differential conductance for an FET with 25nm gate length. The splitting is clearly seen and increases linearly with the applied source to drain voltage V_{DS}.

It is the dependence of E_{MAX} on V_{DS}, due to the so-called Drain Induced Barrier Lowering (DIBL) [10], that is the key to understand the observed behavior of G. Indeed, G is proportional to the rate of increase of available propagating states with increasing V_{DS}, i.e., dE_{MAX}/dV_{DS}. From electrostatics, $dE_{MAX}/dV_{DS} \approx -e\beta$, with $0 < \beta < 1$. As V_{GS} is further increased, E_{MAX} goes below μ_D (left part of Fig. 3 (iii)), and electron states contributing to the current are those with energy between μ_S and μ_D. Again, G is proportional to the rate of increase of current-carrying states with increasing V_{DS}, that is $d\mu_D/dV_{DS} \propto -e$, i.e., a factor $\beta^{-1} > 1$ larger than before. As a consequence, we observe an abrupt increase of the differential conductance as $E_{MAX} = \mu_D$, which is better highlighted by the second peak of dG/dV_{GS}. As V_{GS} is increased further dG/dV_{GS} goes to zero due to the dependence of the number of transmitted modes on E_{MAX}. Therefore, the two peaks of dG/dV_{GS} correspond to the two cases $E_{MAX} = \mu_S$, and $E_{MAX} = \mu_D$. Their presence reveals that two different electrochemical

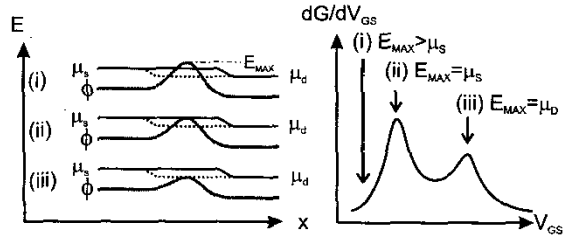

FIG. 3: Sketch of subband energies for interpretation of the experimental results: **(i)** Profile of the lowest two-dimensional subband in the HFET channel for low gate voltage. Since the subband maximum E_{MAX} is larger than both μ_{S} and μ_{D} the device is in cut-off. **(ii)** As V_{GS} is increased, the subband maximum E_{MAX} decreases. When E_{MAX} goes below μ_{S} the conductance increases abruptly. Therefore the first peak of dG/dV_{GS} occurs. **(iii)** As V_{GS} is further increased, E_{MAX} goes below μ_{D} and the conductance increases faster, since now the electron flux is limited only by the availability of empty states at the drain, and not by the channel barrier. For further increases of V_{GS} the derivative of the differential conductance goes to zero due to the dependence of the number of transmitted modes on E_{MAX}.

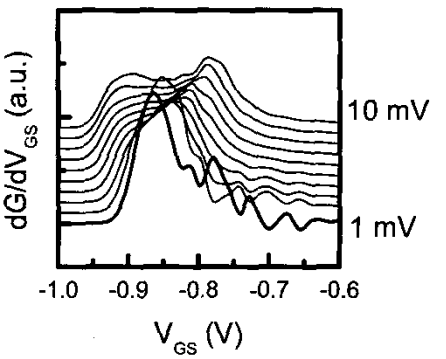

FIG. 4: : Calculated dG/dV_{GS} as a function of V_{GS} obtained from a numerical simulation of ballistic electron transport in an FET with a gate length of 50nm. Curves for different values of V_{DS} are shifted for clarity of presentation. The development of the two peaks with increasing V_{DS} is clearly reproduced.

potentials exist in the channel, since electrons, in their ballistic motion, conserve the electrochemical potential of the contact from which they are injected.

We have verified such a behavior by performing a two-dimensional quantum simulation of the FET structures, taking into account both quantum confinement and ballistic transport in the channel. The subbands are obtained from a self-consistent solution of the Poisson and Schrödinger equations with density functional theory, according to a method described in Ref. [11]. We assume that the occupation factors of all propagating states obey Fermi-Dirac statistics with the electrochemical potential equal to that of the originating contact. This is consistent with the assumption that electrons injected into the device conserve the electrochemical potential of the emitting contact until they exit the device region through the source or the drain. Such an assumption, by itself, is sufficient to ensure current continuity per unit energy in each subband. In our simulation, while we consider fully ballistic transport in the channel, we have included two series resistances at the source and the drain contact. The derivative of G with respect to V_{GS} for the 50 nm HFET is shown in Fig.4. Results are in good agreement with the experiments, showing that the relevant physics of the phenomenon is well described by our interpretation.

In order to verify that the observed signature is due to ballistic transport we have modelled the effect of dissipative transport in the channel. Let us consider zero temperature with V_{GS} such that the differential conductance G for a given V_{DS} is zero. With increasing V_{GS} when the FET is getting conductive, we can write within the drift-diffusion model $I \approx -\mu n\, dE_F/dx$ where μ is the electron mobility, n is the electron concentration at the source, and dE_F/dx is the spatial derivative of the quasi-Fermi energy at the source. If we take into account only the lowest 2D subband $n \approx \rho_{2\mathrm{D}}(\mu_{\mathrm{S}} - E_{\mathrm{MAX}})$, where $\rho_{2\mathrm{D}}$ is the 2D electron density of states. In the case of saturation $dE_F/dx \approx -(\mu_{\mathrm{S}} - E_{\mathrm{MAX}})/L$, where L is the channel length. Therefore we have $I \approx \mu\rho_{2\mathrm{D}}(\mu_{\mathrm{S}} - E_{\mathrm{MAX}})^2/L$. Using (1) we can write $dG/dV_{\mathrm{GS}} = 2\mu\rho_{2\mathrm{D}}e^2\alpha\beta/L$. For larger V_{GS} the FET enters the ohmic region with $dE_F/dx \approx -(\mu_{\mathrm{S}} - \mu_{\mathrm{D}})/L$, and therefore $I \approx \mu\rho_{2\mathrm{D}}(\mu_{\mathrm{S}} - E_{\mathrm{MAX}})(\mu_{\mathrm{S}} - \mu_{\mathrm{D}})/L$, which gives us $dG/dV_{\mathrm{GS}} = 2\mu\rho_{2\mathrm{D}}e^2\alpha/L$, a factor β^{-1} larger than in the saturation region. Therefore, when the FET changes its transport regime with increasing V_{GS} for constant V_{DS} from cutoff to saturation and then from saturation to the ohmic region dG/dV_{GS} as a function of V_{GS} results in a staircase with two steps. Only the contact resistances at the source and the drain contacts lead to a descending curve superimposed to the staircase due to the reduced effective voltage drop on the channel with increasing current. Experimentally one can easily exclude the contact resistance by transport measurements in a 4 terminal (4T) configuration.

Therefore we have realized FETs with 2 further voltage probes separated from the source and the drain terminals, which allow the detection of the intrinic V_{DS} of the device by neglecting the voltage drops on the source and drain contacts. In the case when the contact resistance causes

the observed peak splitting, the peaks in the 4T dG/dV_{GS} curve should vanish and only a step like function should appear.

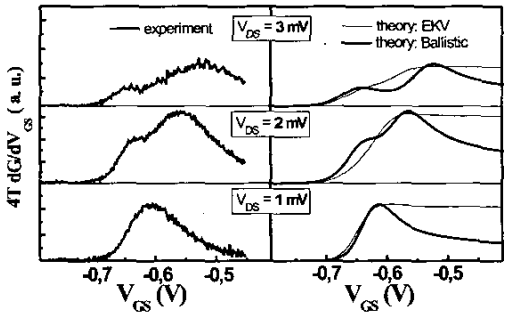

FIG. 5: **Left hand side**: Experimental four-probe dG/dV_{GS} as a function of V_{GS} obtained with a four-probe measurement for V_{DS} ranging from 1 to 3 mV for 50 nm gate length FET. **Right hand side**: Theoretical best fitting curve obtained with the analytic EKV model (thin solid line) and with an analytic model for ballistic transport (thick solid line).

In Fig. 5 (left part) the 4T dG/dV_{GS} of an FET with a channel length of 50 nm at 4.2 K as a function of V_{GS} for V_{DS} ranging from 1 to 3 mV is shown. When V_{DS}=1 mV one clear peak in the 4T dG/dV_{GS} trace with increasing V_{GS} appears. We emphasize that within the drift-diffusion model a single staircase is predicted. However, in order to compare the experimental data we included in Fig. 5 (right part) analytically obtained results using the transistor model of Enz, Krummenacher and Vitoz (EKV), which assumes dissipative quasi-equilibrium transport in the channel [12], and analytical results obtained using a ballistic transport model described in Ref. [9]. The coefficients of both the EKV and the ballistic models have been chosen in order to provide the best fit with experiments. As anticipated, in the case of dissipative transport each change of operating region only provides a step in the curves of Fig. 5, and while in the case of ballistic transport we clearly see the two peaks. Notwithstanding the noise of the experimental results due to the numeric second derivative performed the agreement with the predictions of the ballistic transport model is apparent. Therefore, the double peak feature observed in the dG/dV_{GS} is a specific signature of ballistic transport in the device.

The dependence of the separation between the peaks of dG/dV_{GS} upon V_{DS} can provide information about the capability to control the two-dimensional subband in the channel through the gate voltage. This is particularly relevant in ballistic devices since the current only depends on E_{MAX} and on the efficiency of the emitting contact (source). From our considerations it follows that $|dE_{MAX}/qdV_{GS}|$ is close to $V_{DS}/\Delta V_{GS}$, where ΔV_{GS} is the peak splitting corresponding to V_{DS}. For example, in the case of the 50 nm device, shown in Fig. 5, it is pretty clear that the gate efficiency is rather poor, since the term $V_{DS}/\Delta V_{GS}$ is close to 1/40.

In conclusion, our experiments and simulations demonstrate that the four-probe differential conductance of field effect transistors gives direct access to information on ballistic transport in multimode devices far from the equilibrium, where conductance is not quantized.

Let us stress the fact that the double peak feature is a unique signature of ballistic transport in nanoscale FETs, where elastic diffusive transport can be easily ruled out. For devices in which such assumption cannot be made, for example because the device length is much larger then the typical elastic scattering length, one must consider that elastic diffusive transport could also provide a two-step electron distribution in the channel, as shown in Ref. 13 for polycrystalline metal wires. In such cases the double peak feature would only be a unique signature of *elastic* transport.

Finally, we would like to note that as long as the channel is shorter than the inelastic scattering length the reported effect should be observable at higher temperature.

This work was supported from the EU project NanoTCAD IST-1999-10828.

[1] M. Heiblum, M. I. Nathan, D. C. Thomas, and C. M. Knoedler, *Phys. Rev. Lett.* Vol. 55, pp. 2200-2203, 1985.
[2] A. Palevski, M. Heiblum, C. P. Umbach, C. M. Knoedler, A. N. Broers, and R. H. Koch, *Phys. Rev. Lett.* vol. 62, pp. 1776-1779, 1989.
[3] K. Hess and G. J. Iafrate, *Proc. IEEE* vol. 76, pp. 519-532, 1988.
[4] B. J. van Wees et *al.*, *Phys. Rev. Lett.* vol. 60, pp. 848-851, 1988.
[5] D. A. Wharam et *al. J. Phys. C*, vol. 21, L209-L214, 1988.
[6] S. Frank, P. Poncharal, Z. L. Wank, and W. A. D. Heer, *Science* vol. 280, pp. 1744-1746, 1988.
[7] M. A. Topinka et *al.*, *Science*, vol. 289, pp. 2323-2326, 2000.
[8] M. A. Topinka *et al.*, *Nature*, vol. 410, pp. 183-186, 2001.
[9] K. Natori, *J. Appl. Phys.* vol. 76, pp. 4879-4890, 1994.
[10] W. Liang, M. Bockrath, D. Bozovic, J. H. Hafner, M. Tinkham, and H. Park, *Nature*, vol. 411, pp. 665-669, 2001.
[11] G. Fiori and G. Iannaccone, *Appl. Phys. Lett.* vol. 81, pp. 3672-3674, 2002.
[12] C. C. Enz, F. Krummenacher, and E. A. Vittoz, *Analog Int. Circ. Sig. Proc.* vol. 8, pp. 83-114, 1995.
[13] H. Pothier, S. Guéron, and M. H. Devoret, *Phys. Rev. Lett.* vol. 79, pp. 3490-3492, 1997.