Extended Abstracts of the 20th (1988 International) Conference on Solid State Devices and Materials, Tokyo, 1988, pp. 491-494 **S-IIIA-1**

Invited

Quantum Interference Transistors

Supriyo DATTA

School of Electrical Engineering
Purdue University
West Lafayette, IN 47907

We discuss theoretically the possibility of quantum transistors that utilize interference effects to control the drain current. Low switching voltages compared to conventional transistors are predicted. The possibility of implementing programmable quantum networks with an intriguing similarity to neural networks is also pointed out.

Consider first the most well-known quantum device: the resonant tunneling diode.[1] To calculate the current (I) for a given voltage (V) we start from the expression

$$I = \frac{e}{\pi\hbar}\int dE\,[f(E)-f(E+eV)]\sum_{\mathbf{k}_t} |t(E,\mathbf{k}_t)|^2 \qquad (1)$$

The transmission coefficients $t(E,\mathbf{k}_t)$ for electrons incident from the left contact with energy E and transverse wavevector $\mathbf{k}_t=(k_x,k_y)$ are evaluated from the (effective mass) Schrödinger equation.

$$-\frac{\hbar^2}{2m^*}\nabla^2\Psi + E_C(\mathbf{r})\Psi = E\Psi \qquad (2)$$

$E_C(\mathbf{r})$ is the bottom of the conduction band and m^* is the effective mass. Usually in theoretical calculations all scattering processes within the device are neglected assuming that it is short compared to a mean free path. Elastic scattering can be included straightforwardly by adding the corresponding scattering potential $V_s(\mathbf{r})$ to $E_c(\mathbf{r})$ in eq. (2). It is more difficult (even conceptually) to include dissipative inelastic processes within the device, since such processes are irreversible while the Schrödinger equation is reversible. Eq. (1) assumes that all the dissipation takes place in the contacts and not in the device. This assumption is usually fairly justified for a useful quantum device and we will use eq. (1) for our discussions in this paper.

The resonant tunneling diode has found widespread use[2] though one of its limitations is that it is basically a *two-terminal* rather than a *three-terminal* device. The latter is preferred from an applied point of view and resonant tunneling diodes have been adapted to three-terminal configurations.[3] The objective of this paper is to discuss theoretically the possibility of 'quantum transistors' where a gate voltage V_G is used to control the transmission coefficients $t(E,\mathbf{k}_t)$ in eq. (1) through various interference effects and thereby modulate the current I for a fixed source-drain voltage V.[4] This requires structures with multiple paths from the left contact to the right contact.

---

1) R. Tsu, L. Esaki: Appl. Phys. Lett. 22 (1973) 562.

2) E. R. Brown et.al.: Appl. Phys. Lett. 50 (1987) 83.

3) F. Capasso et.al.: IEEE Electron Dev. Lett. ED-7 (1986) 573.

4) S. Datta et. al.: Second International Conference on Modulated Semiconductor Structures, Kyoto, Sept. 1985 (Surf. Sci. 174 (1986) 439); A. B. Fowler, Patent No. 45503320 (1985).

For example, consider the structure in Fig. 1a which is basically an ordinary Field Effect Transistor (FET) with a barrier in the middle of the channel.[5] One way to fabricate this structure may be to grow a barrier layer over the entire film, interrupt the growth to etch away the barrier where it is not needed and then continue the growth process. The major challenge lies in ensuring the quality of the regrown interfaces. Neglecting multiple reflections[6] we can write

$$t(E,\mathbf{k}_t) \simeq t^{(1)}(E,\mathbf{k}_t) + t^{(2)}(E,\mathbf{k}_t) \tag{3}$$

where the superscripts 1 and 2 refer to the two channels. Now,

$$t^{(1)} \sim e^{ik_1L} \tag{4a}$$

$$t^{(2)} \sim e^{ik_2L} \tag{4b}$$

where k_1 and k_2 are the wavenumbers in the x-direction in the two channels respectively.

$$E = \epsilon_1 + \frac{\hbar^2 k_t^2}{2m^*} + \frac{\hbar^2 k_1^2}{2m^*} \tag{5a}$$

$$E = \epsilon_2 + \frac{\hbar^2 k_t^2}{2m^*} + \frac{\hbar^2 k_2^2}{2m^*} \tag{5b}$$

Here ϵ_1 and ϵ_2 are the energies at the bottom of the subband in channels 1 and 2 respectively; it is assumed that only the lowest subband is occupied everywhere in the device. From eqs. (3) and (4) we have,

$$|t(E,\mathbf{k}_t)|^2 = 4\,|t^{(1)}(E,\mathbf{k}_t)|^2 \cos^2(\theta/2) \tag{6a}$$

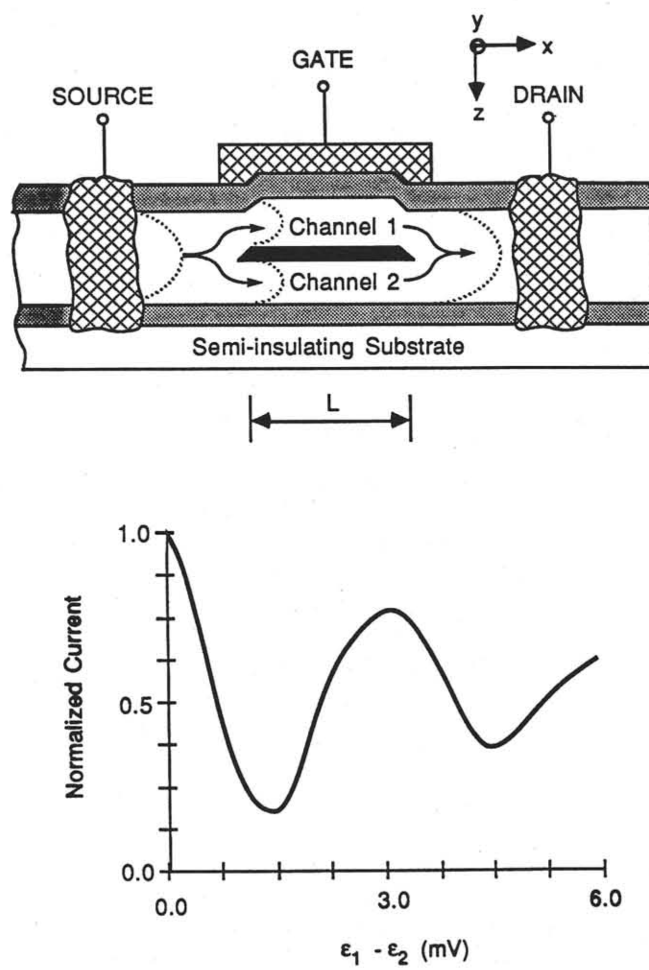
$$\text{where} \quad \theta = (k_1 - k_2)L \tag{6b}$$

It is evident from eq. (6) that the transmission coefficient $|t|^2$ can be modulated through $(k_1 - k_2)$. From eqs. (5a,b),

$$\theta \equiv (k_1 - k_2)L = (\epsilon_2 - \epsilon_1)L/\hbar v \tag{7a}$$

$$\text{where} \quad v = \hbar\,(k_1 + k_2)/2m^* \tag{7b}$$

Eq. (7a) shows that $(k_1 - k_2)$ can be changed by changing $\epsilon_2 - \epsilon_1$, the energy difference between the subbands in channels 1 and 2. This is achieved readily with an applied gate voltage. Fig. 1b shows the normalized current calculated as a function of $\epsilon_1 - \epsilon_2$. The current can be turned off with a fairly small potential $\sim$ mV. It seems that the speed of operation of this device should ultimately be limited by the transit time across the two-channel region of length L. However, it should be noted that as L is decreased the potential difference $(\epsilon_2 - \epsilon_1)$ required to switch the device is increased and becomes comparable to the potential required to deplete the channel if L is comparable to the DeBroglie wavelength of the electrons ($kL \sim 2\pi$).



(b)

Fig. 1: A proposed Quantum Interference Transistor (after Ref. 5)

(a) The structure consists of a conducting channel with a barrier in the middle.

(b) Calculated current versus gate voltage for a small source-to-drain voltage.

It can be seen from Fig. 1b that the current is not completely turned off as $(\epsilon_1 - \epsilon_2)$ is changed. The reason is that the current is determined by

5) S. Bandyopadhyay, S. Datta, M. R. Melloch: Superlattices and Microstructures.

6) S. Datta, S. Bandyopadhyay: Phys. Rev. Lett. 58 (1987) 717.

$|t(E,\mathbf{k}_t)|^2$ averaged over a range of energies (E) and wavevectors ($\mathbf{k}_t$) (Eq. (1)). The current is a minimum when $\theta \sim \pi$; but electrons with different energies E and wavevectors $\mathbf{k}_t$ have different velocities v and hence different values of θ for the same value of $(\epsilon_2 - \epsilon_1)$. The spread $\delta\theta$ in the phase-shift is proportional to the spread $\delta\tau_t$ in the transit times $\tau_t = L/v$.

$$\delta\theta = \frac{\epsilon_2 - \epsilon_1}{\hbar}\,\delta\tau_t \tag{8}$$

It is easy to see from eq. (8) why the oscillations in the conductance get smaller as $\epsilon_2 - \epsilon_1$ is increased (Fig. 1b); it is because for a fixed $\delta\tau_t$, $\delta\theta$ gets larger. Again, it will be noted that multiple elastic scattering processes increase the spread in the transit time for the electrons, causing the effect to wash out.[7] For this reason ballistic transport is desirable over the length L. This does not seem difficult to achieve using present-day technology. A more difficult requirement seems to be to ensure an adequate degree of symmetry between the channels.

Since the minimum in the current is limited basically by the spread $\delta\tau_t$ in the transit times, it should be possible to increase the percentage modulation by restricting the y-dimension of the device. This should reduce the number of allowed transverse wavevectors $\mathbf{k}_t$. In fact, in the limit when there is only a single allowed subband, we can ideally expect the current to go to zero at low temperatures when the spread in the energies E is small. For a carrier density of $n_s = 7 \times 10^{11}/cm^2$ in GaAs we need a width $\sim$ 150Å in the y-direction in order to achieve single-moded operation. This seems too small to define lithographically. However, recent experiments have shown that the actual channel width is significantly smaller than the physical width due to depletion layers.[8),9)] Using single-moded quantum wires it seems feasible to conceive of quantum networks analogous in principle to the microwave networks well-known in electromagnetics. The main difference is that with classical waves coherent monochromatic sources with very little spread in the wavelength are quite common. By contrast, electrons in solid state devices commonly have a large spread in energy, E (analogous to the frequency of classical waves) and also have a large spread in the transverse wavevector, $\mathbf{k}_t$ ('uncollimated'). The resultant spread in the wavelength tends to wash out quantum interference effects. Moreover, inelastic scattering events which randomize phases are present in solids. For these two reasons we normally do not worry about such effects in everyday devices. But with single-moded quantum wires at low temperatures, devices such as the ones shown in Figs. 2a and 2b may be practicable. A surprising feature of these devices is that the gate is not positioned between the source and the drain as we normally expect in electronic devices. From a microwave network point of view, however, it is quite well-known that in a multiport network the transmission between two ports is affected by the load at the other ports. This 'remote control' (or non-local effect[10),11)] as it is usually called) is an interesting aspect of quantum devices that may find useful applications. Another aspect that has not yet received much attention is the large spin-orbit coupling term[12)] in the conduction band of narrow-gap semiconductors like InAs. Electron spin plays a role analogous to the polarization of electromagnetic waves and it may be possible to exploit this degree of freedom too.

Finally, it should be noted that the true power and utility of quantum devices may eventually lie not in the implementation of conventional transistors with a source, a drain and a gate but in implementing programmable multiterminal resistor networks analogous to neural networks.[13)] For example, the four-terminal structure in Fig. 3a may be represented as a resistor network with four nodes as shown in Fig. 3b; the conductance G_{mn} connecting node m and node n is proportional to the transmission coefficient of an electron wave

7) S. Datta, M. Cahay, M. McLennan: Phys. Rev. B36 (1987) 5655.

8) G. Timp et. al.: Phys. Rev. Lett. 59 (1987) 732.

9) G. Timp et. al.: Phys. Rev. Lett. 58 (1987) 2814.

10) A. Benoit et. al.: Phys. Rev. Lett. 58 (1987) 2343; W. J. Skocpol et. al.: Phys. Rev. Lett. 58 (1987) 2347.

11) H. U. Baranger et. al.: Phys. Rev. B37 (1988).

12) Y. A. Bychkov, E. I. Rashba: J. Phys. C17 (1984) 6039.

13) S. Datta: Proceedings of the International Conference on Electronic Materials, Tokyo, 1988.

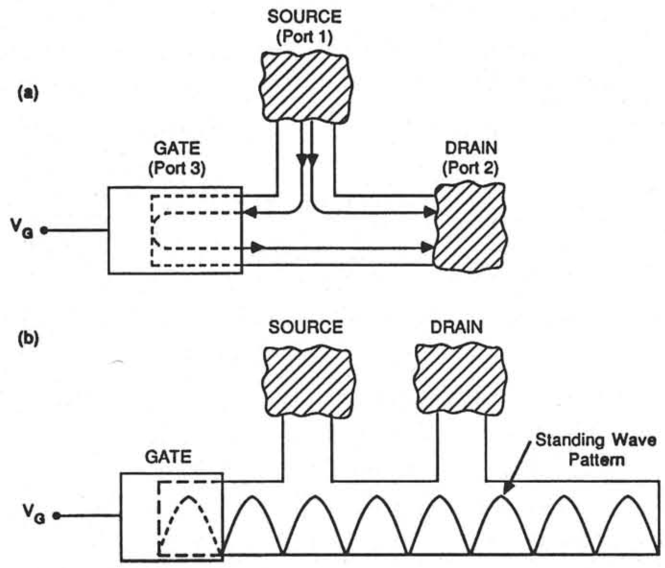


Fig. 2: Possible structures for quantum transistors with remote gates.

(a) The gate potential changes the phase-difference between the two primary paths between the source and the drain.

(b) The gate potential shifts the standing wave pattern in the main waveguide to which the source and drain are weakly coupled. A large source-drain current flows when peaks of the standing wave are aligned with the source and drain.

from node m to node n.[13,14] A gate at one end of the structure could be used to change the conductances G_{mn}. However, a lot of theoretical and experimental work remains to be done before it is clear whether optimal structures can indeed be designed to achieve large and controlled modula-

14) M. Büttiker: Phys. Rev. Lett. 57 (1986) 1761.

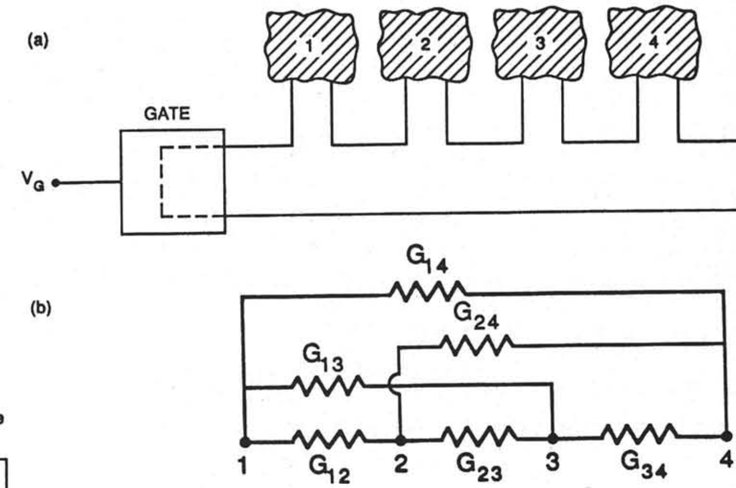


Fig. 3: A programmable multiterminal quantum network.

(a) Schematic diagram.

(b) Equivalent circuit.

tion of the resistor network.[15]

Acknowledgements

This work was supported by the National Science Foundation under Grant no. ECS-83-51-036, by the Semiconductor Research Corporation under Contract no. 87-SJ-089, and by the Office of Naval Research under Contract No. N00014-87-K-0693.

15) S. Datta: Quantum Interference Devices, chapter in Physics of Quantum Electron Devices edited by F. Capasso, to be published in 1989 by Springer-Verlag.