An electron wave directional coupler and its analysis

Nadir Dagli, Gregory Snider, Jonathan Waldman, and Evelyn Hu
Electrical and Computer Engineering Department, University of California, Santa Barbara, California 93106

(Received 4 September 1990; accepted for publication 9 October 1990)

In this work a specific quantum interference device, which is two coupled-electron waveguides in close proximity, is described and analyzed. After an initial analysis outlining the principle of operation and the constraints on the structure dimensions and material parameters, a specific structure is simulated using a self-consistent Poisson and Schrödinger solver. Results indicate that this device is feasible using current fabrication technology.

I. INTRODUCTION

The possibility of fabrication at lateral dimensions from a few hundred to a few thousand angstroms, together with the submonolayer resolution of heterostructures grown by molecular-beam epitaxy (MBE) or metalorganic chemical vapor deposition (MOCVD) make accessible new classes of "quantum" devices. For example, devices at these dimensions may utilize the coherent, wavelike properties of electrons, and new device ideas very much like guided-wave optical and microwave devices could be possible. Several different structures which are analogs of optical interferometers,[1] single stub tuners,[2] and directional couplers[3,4] have already been proposed. In this work a specific structure, which is the analog of an optical directional coupler, is described and analyzed. This structure consists of two electron waveguides in close proximity to one another, and is represented schematically in Fig. 1. If the shortest separation between the waveguides, S, allows sufficient overlap between coherent electron waves traveling through either branch A or B of the coupler, of width W, then a periodic transfer of signal from A to B can take place, with a spatial period of L. We shall initially develop the analysis for the electron-wave directional coupler. Using the insights gained from the analysis, we then describe a particular realization of the structure, and discuss the constraints on the structure dimensions and other parameters to assess the feasibility of this approach.

II. GENERAL ANALYSIS

We assume that the dimensions of the coupler shown in Fig. 1 are sufficient to serve as waveguides for coherent electron wavefunctions, and that the separation, S, between the waveguides allows overlap and coupling of the wavefunctions. The effect of this coupling is to split the energy eigenstates in the original electron waveguide into a symmetric and an antisymmetric state. Hence, it is best to describe the electronic states in the coupled region in terms of the superposition of a symmetric and an antisymmetric wavefunction as depicted in Fig. 1. Then assuming that electrons behave as coherent electron waves in the coupled section and that, by proper design only the lowest lying symmetric and antisymmetric eigenstates are occupied, the total wavefunction in the coupled region is

$$\psi(x,y,z) = a_s\phi_s(x,y)e^{ik_sz} + a_a\phi_a(x,y)e^{ik_az}, \quad (1)$$

where subscripts s and a refer to symmetric and antisymmetric eigenstates. $\phi_s(x,y)$ and $\phi_a(x,y)$ are normalized wavefunctions for the lowest symmetric and antisymmetric eigenstates such that

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \phi_s^*(x,y)\phi_s(x,y)dx\,dy = \langle\phi_s|\phi_s\rangle = 1, \quad (2)$$

$$\langle\phi_a|\phi_a\rangle = 1, \quad (3)$$

$$\langle\phi_s|\phi_a\rangle = 0. \quad (4)$$

Now suppose that electrons of one particular energy E are injected into the coupled system through the waveguide A. This incident electron wave, which can be represented with a wavefunction

$$\psi(x,y,z) = \Psi(x,y)e^{ikz}, \quad (5)$$

will excite the symmetric and antisymmetric eigenstates at the beginning of the coupled section, i.e., at $z = 0$, in such a way that their superposition will be equivalent to the incident wave, as shown in Fig. 1. The amplitudes of the symmetric and antisymmetric wavefunctions will depend on the overlaps of these wavefunctions with the overlap of the incident wavefunction, i.e.,

$$a_s = \langle\phi_s|\Psi\rangle \text{ and } a_a = \langle\phi_a|\Psi\rangle. \quad (6)$$

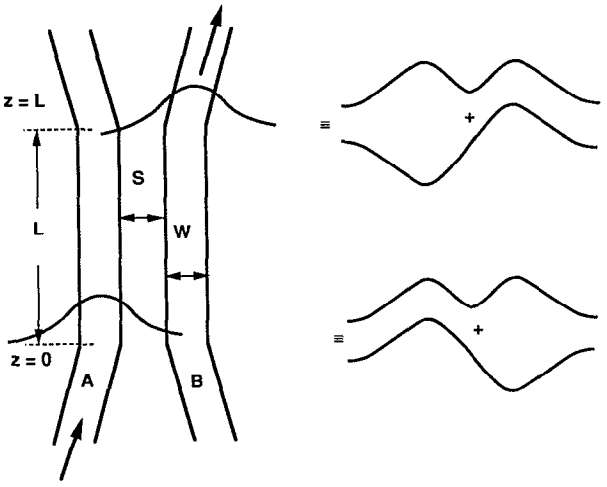

FIG. 1. Two coupled electron waveguides and the electron wave functions at various points along the coupled region when electrons are incident on it through waveguide A.

 0021-8979/91/021047-05$03.00 © 1991 American Institute of Physics

Downloaded 10 Mar 2005 to 128.111.56.182. Redistribution subject to AIP license or copyright, see http://jap.aip.org/jap/copyright.jsp

If the individual waveguides in the coupler are identical and the coupling between them is not very strong, $a_s \approx a_a$. Both the symmetric and antisymmetric eigenstates will have the same total energy, E, which is the energy of incident electrons. Then

$$E = E_s + \frac{\hbar^2 k_s^2}{2m^*} = E_a + \frac{\hbar^2 k_a^2}{2m^*}, \tag{7}$$

where m^* is the electron effective mass, E_s and E_a are the lowest symmetric and antisymmetric energy eigenvalues and k_s and k_a are the corresponding symmetric and antisymmetric wave vectors of the coupled system. As a result of this condition the symmetric and antisymmetric wavefunctions will travel along the length of the coupler at different speeds, since they have different wave vectors in the direction of propagation. Therefore, there will be a phase difference between these two wavefunctions as they propagate along the length of the coupled region. At the end of the coupled region i.e., at $z = L$, the electron wavefunction will be

$$\psi(x,y,z) = \left[a_s\phi_s(x,y) + a_a\phi_a(x,y)e^{-i(k_s - k_a)L}\right]e^{ik_sL}. \tag{8}$$

If L is chosen such that $(k_s - k_a)L = \pi$ the resulting electron distribution at $z = L$ will be in the second waveguide as shown in Fig. 1. If the coupling is terminated at this point, an electron wave initially traveling in z direction in waveguide A can entirely be transferred to waveguide B. This spatial transfer of charge from one channel to the other is periodic in the z direction and the electron wave essentially meanders between these two waveguides. Hence, if the length over which the waveguides are coupled is made just right, so that $(k_s - k_a)L = (2n + 1)\pi$, charge that enters waveguide A will be ultimately leave through waveguide B. On the other hand, one can make $(k_s - k_a)L = 2\pi$ by destroying the synchronism between the waveguides, for example by making them no longer identical. This way at $z = L$ electron distribution will be in waveguide A. This switching of the electron wavefunction from one waveguide to the other will externally manifest itself as current switching from one terminal to another, hence this structure will form an active switching device.

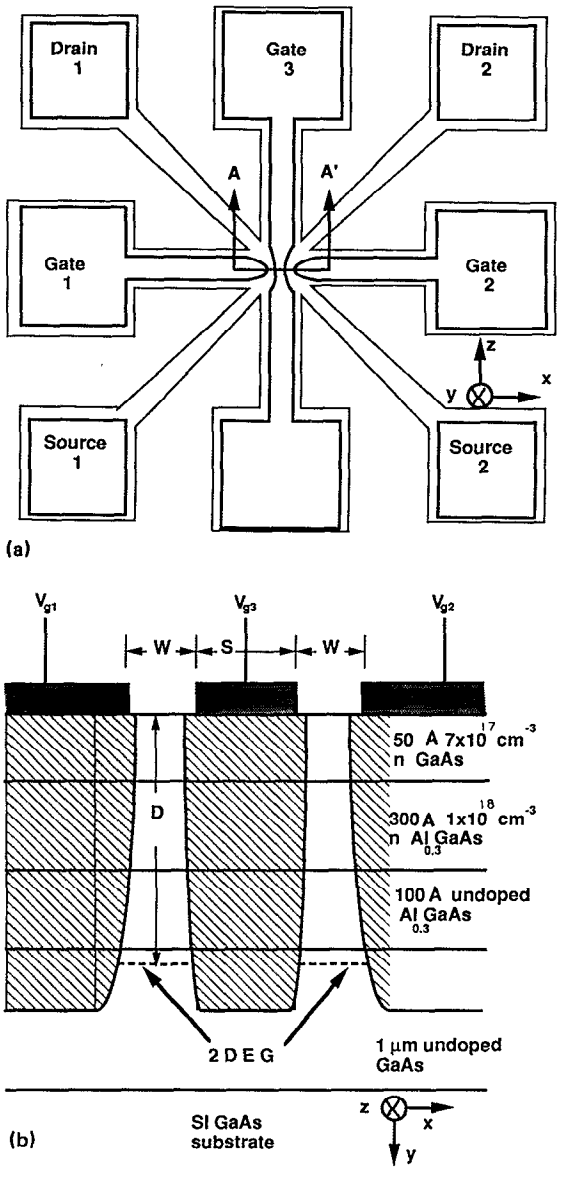

FIG. 2. (a) Top view of the proposed specific electron wave directional coupler. The enlarged view of the central part of the device in the vicinity of section AA' is the same as Fig. 1. (b) Cross-sectional profile of the coupled waveguides along AA'. The shaded areas in the semiconductor show the depletion regions.

III. A SPECIFIC ELECTRON WAVE DIRECTIONAL COUPLER: DESIGN CONSIDERATIONS

The specific structure proposed for the realization of the electron waveguide coupler is shown in Fig. 2. The starting material, is a modulation doped GaAs/AlGaAs heterostructure which provides a high-mobility 2DEG. Initially a mesa, whose boundary is defined by the thin solid line in Fig. 2(a), is etched all the way to the semi-insulating GaAs substrate. Then Schottky gate patterns and ohmic contacts are fabricated on the surface as shown in Fig. 2(a). The resulting structure is very much like two split-gate FETs in very close proximity, and its contacts will be called drain, source, and gate in accordance with the FET terminology as indicated in Fig. 2(a). The mesa etching provides isolation and makes drain 1 separate from drain 2 and source 1 separate from source 2. Under appropriate bias areas under the gate electrodes are also depleted of electrons and only two narrow channels of high-mobility electrons between the drains and sources exist. Figure 2(b) shows the cross-sectional profile of the structure along section AA' in Fig. 2(a). In the proposed structure the electron waveguides are created by confining a two-dimensional electron gas (2DEG) into a narrow channel using the depletion edge of a split Schottky barrier gate. Electron waveguides created this way have been successfully used by different groups to observe quantized conductance.[5-7] Therefore, the resulting structure is two separate waveguides decoupled everywhere except for a short length, where they are in close proximity, hence cou-

Downloaded 10 Mar 2005 to 128.111.56.182. Redistribution subject to AIP license or copyright, see http://jap.aip.org/jap/copyright.jsp

pled. Electrons launched by either source 1 or source 2 enter into the coupled waveguide region formed by gates 1, 2, and 3. The electrons are finally collected either at drain 1 or drain 2. Electron waveguiding and switching can be achieved by applying appropriate voltages on the gate, source, and drain electrodes. For example, by applying a small voltage between source 1 and drain 1 electrons are injected into waveguide A from the 2DEG reservoir near source 1. The same drain voltage should also be applied to drain 2 and source 2 should be terminated with an open circuit to prevent any electron injection into waveguide B from the 2DEG reservoir near source 2. Furthermore, the two channels should be made identical by applying proper gate voltages. Under these conditions with the proper choice of L, a charge injected into the coupled section from source 1 exits through drain 2. The channels can also be made different by application of a different gate voltage. This is equivalent to varying the width, hence the energy eigenvalues of the affected waveguide. In that case, the coupling between the two waveguides will be disturbed. If the energy eigenvalues of guide 2 are sufficiently modulated, so that $(k_s - k_a)L = 2\pi$, then the charge injected at source 1 will appear at drain 1. Therefore, it is possible to switch the current between drain 1 and 2 by modulating the voltage at gate 2. In this mode of operation the device is equivalent to a current steering logic gate. If the depth of modulation of either drain current is kept low by applying a small modulating voltage to either gate 1 or gate 2 the device will work as a linear modulator. It can also work as a multiplexer or demultiplexer if the voltage at gate 2 is properly synchronized to the repetition frequency of current pulses incident on drain 1.

In designing the particular structure of Fig. 2(a), the following considerations apply. The electron transport through the channel should be ballistic and the cross-sectional dimensions of the channel should be comparable to the Fermi wavelength of the electrons. Furthermore, the length of the channel should be short enough, shorter than the phase coherence length, so that the phase information of the electrons is preserved and electrons behave as coherent waves. To satisfy these requirements, while keeping the lateral dimensions of the structure as large as possible, a very high mobility of the 2DEG is needed. Furthermore, this 2DEG should be as close to the gate electrodes as possible. The shorter this distance, D, the sharper the potential step that is created at the 2DEG plane. Also, for small D, one can modify this potential step sufficiently with realistic gate voltages. On the other hand, there is a lower limit on D. In order to maintain high-electron mobility in the 2DEG the thickness of the undoped AlGaAs spacer should be greater than 100 Å. The thickness of the doped AlGaAs supply layer cannot be made very low either, in order to have sufficient number of carriers in the channel. Furthermore, a thin GaAs cap layer is desirable to protect AlGaAs from air contact, hence oxidation. One particular vertical design that satisfies all these constraints is shown in Fig. 2(b). For this structure, in our laboratory we routinely achieve sheet-carrier concentration in the channel of about 4×10^{11} cm^{-2}, and electron mobilities over 300 000 cm^2/V s at 4.2 °K without any parallel conduction paths in AlGaAs supply layer.

Transfer length of the coupler, L, should be less than the phase coherence length of electrons in order to preserve the phase information of the electronic wavefunctions, and to observe coherent wave coupling. Using data from previous experiments on ballistic transport and quantized conductance on similar structures, we judge that L should be less than 0.3 μm for the expected mobility value in our design.[7] The phase coherence of the electron wave function sets an upper bound on L. On the other hand, L must be long enough to allow transfer of the electron wave from one waveguide to the other, i.e., $L \geqslant \pi/(k_s - k_a)$.

The transfer length depends exponentially on the separation between the waveguides, which is S plus the width of the depletion layer on either side of the center electrode. To maintain a short value for the transfer length, S is desired to be as narrow as possible. Furthermore, S directly influences the strength of coupling, hence the energy splitting between the symmetric and antisymmetric energy eigenvalues. That energy splitting should be much larger than the thermal energy, kT, to eliminate the possibility of thermal transitions between the symmetric and antisymmetric states. In this way, S affects not only the transfer length, but also the upper temperature of operation. Hence, it is the most critical parameter for the successful operation of this device.

The width, W, of the waveguides is another important parameter. Coupled with the potential step that is created at the 2DEG plane, waveguide width determines the number of eigenstates or modes that will exist, the energy gap between them, and their confinement. The larger the value of W, the stronger the confinement of the modes and the smaller the coupling. As a result, the transfer length will have to be longer for full switching. Furthermore, for larger W, the number of modes will increase and the energy differences between them will decrease. The proper operation of this device requires that only the lowest symmetric and antisymmetric subbands be occupied. That means that electrons injected into the coupled region should populate the lowest symmetric and antisymmetric subbands and, even though higher order subbands can exist, the energy separation of those subbands with the lowest subbands should be many kT. But, this condition is almost always satisfied if the waveguide width is not very wide and S is narrow enough, so that the energy gap between the lowest symmetric and antisymmetric eigenstates is many kT. Therefore, width has a direct effect on the transfer length and some effect on the upper temperature of operation. The additional limit on the temperature of operation is the temperature dependence of the phase coherence length.

The above discussion should serve to highlight the critical importance of lateral dimension control for S and W. Fortunately the structure under consideration provides a certain degree of tunability. It is possible to move the depletion layer edges at the 2DEG plane by varying the gate voltages. This way one can control the widths of the waveguides as well as the spacing between them, i.e., the subband separation as well as the coupling length. Furthermore, any asymmetry in the fabrication process can be compensated for. There is one other advantage in addition to tunability. The potential created in the 2DEG plane is proportional to the

Downloaded 10 Mar 2005 to 128.111.56.182. Redistribution subject to AIP license or copyright, see http://jap.aip.org/jap/copyright.jsp

second integration of the charge distribution on the gate electrode. This charge distribution will exactly follow the profile of the gate electrode, hence, will have the same inevitable irregularities on the edges of the gate electrode. But due to double integration the potential profile will be a lot smoother. Therefore, this particular structure smooths out the effect of edge irregularities that will inevitably occur in the fabrication. Obviously, apart from these considerations, tunability is essential for the operation of an electron wave coupler in the indicated modes of operation.

IV. A SPECIFIC ELECTRON WAVE DIRECTIONAL COUPLER: ANALYSIS

Keeping all these design criteria in mind, we analyzed several different geometries with the help of simplified models. One of these designs was later simulated using a self-consistent Schrödinger and Poisson's solver.[8] For the design that was simulated $W = 1500$ Å and $S = 300$ Å. These lateral dimensions push the limits of today's technology, however the fabrication is still feasible. The resulting potential energy profile in the 2DEG plane for the bias values indicated, is shown in Fig. 3(a). For these particular bias values, the lowest symmetric and antisymmetric energy eigenvalues turn out to be $E_s = -3.585$ meV and $E_a = -1.701$ meV with respect to the Fermi level in 2DEG for which $E = 0$. The large energy difference between the lowest eigenstates and the potential energy minimum is due to the vertical confinement of the carriers at the 2DEG plane. The corresponding transfer length L can be calculated using the following equation,

$$L = \sqrt{\pi^2\hbar^2/2m^*}\left[1/\left(\sqrt{E - E_s} - \sqrt{E - E_a}\right)\right], \quad (9)$$

which is easily derived using Eq. (7) and $(k_s - k_a)L = \pi$.

Since electrons are incident from the 2DEG reservoir on either side, under small-drain to source biases, it is reasonable to take E as the Fermi energy of the 2DEG, which is 0 in this simulation. Then L for the case under discussion turns out to be 1267 Å, which is acceptable. Furthermore, at $T = 2$ K, $E_a - E_s \cong 10.9kT$ which implies that thermal activation between the symmetric and antisymmetric eigenstates will be low enough to allow proper device operation. The next higher symmetric and antisymmetric subbands are more than $16\,kT$ above the lowest subbands and $6\,kT$ above the Fermi level of the 2DEG. Hence, injected electrons will only populate the lowest symmetric and antisymmetric eigenstates and will not be thermally excited to higher subbands. Therefore, for this structure, with the indicated bias values at $T = 2$ K or lower, it should be possible to switch the current injected through drain 2 to source 1.

The other obvious question is whether or not it will be possible to switch the current back to drain 1 by changing the voltage at gate 2. The simulations indicate that if V_{g_2} is changed to -2.3 V, which is a realistic voltage and does not pinch off the channel, $(k_s - k_a)L = 6.32$ which is slightly larger than 2π. The potential energy profile at the 2DEG plane and the corresponding bias values are shown in Fig. 2(b). Hence it is possible to find a V_{g_2} value somewhere in the range between -2.1 and -2.3 V that will result in switching the current back to drain 1. Furthermore, for these bias values the energy gap between the two lowest energy eigenvalues is more than $13\,kT$ at $T = 2$ K. The next higher energy eigenvalue is more than $18\,kT$ above the Fermi level and $19\,kT$ above the two lowest energy eigenvalues. All these results indicate that such a device is feasible and is within the reach of today's technology.

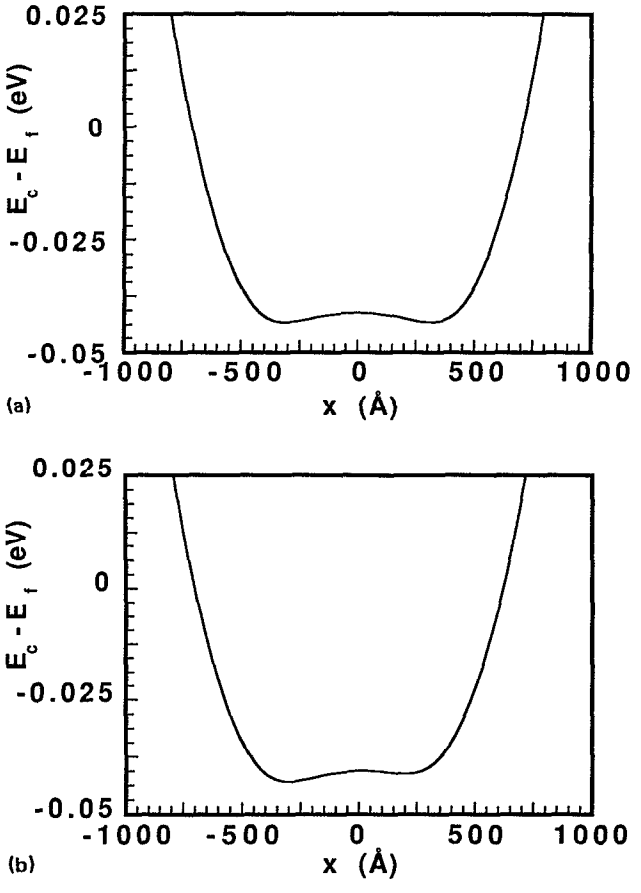

FIG. 3. (a) Potential energy profile in the 2DEG plane of the coupled waveguides for $V_{g_1} = V_{g_2} = -2.1$ V and $V_{g_c} = -0.1$ V. (b) Potential energy profile in the 2DEG plane of the coupled waveguides for $V_{g_1} = -2.1$ V, $V_{g_2} = -2.3$ V and $V_c = -0.1$ V.

V. CONCLUSIONS

In this work a specific electron-wave directional coupler which operates based on the interference of two coherent electron waves in a solid is described and analyzed. Results indicate that it should be possible to fabricate such a device with today's technology and in the future this novel device could be one of the basic building blocks of high-performance digital circuits.

ACKNOWLEDGMENTS

This work is supported by NSF through QUEST, the NSF Science and Technology Center on Quantized Electronic Structures at UC Santa Barbara. Authors also thank

Downloaded 10 Mar 2005 to 128.111.56.182. Redistribution subject to AIP license or copyright, see http://jap.aip.org/jap/copyright.jsp

Professor Arthur Gossard for valuable discussions and support.

[1] S. Datta, M. R. Melloch, S. Bandyopadhyay, and R. Noren, Phys. Rev. Lett. **55**, 2344 (1985).
[2] F. Sols, M. Macucci, U. Ravaioli, and K. Hess, Appl. Phys. Lett. **54**, 350 (1989).
[3] J. A. del Alamo and C. C. Eugster, Appl. Phys. Lett. **56**, 78 (1990).
[4] N. Tsukada, A. D Wieck, and K. Ploog, Appl. Phys. Lett. **56**, 2527 (1990).
[5] 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, Phys. Rev. Lett. **60**, 848 (1988).
[6] 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, J. Phys. C **21**, L209 (1988).
[7] G. Timp, R. E. Behringer, S. Sampare, J. E. Cunningham, and R. Howard, in *Proceedings of the International Symposium on Nanostructure Physics and Fabrication*, edited by W. P. Kirk and M. Reed (Academic, New York, 1989) 331.
[8] G. Snider, I.-H. Tan, and E. Hu, J. Appl. Phys. **68**, 2849 (1990).

Downloaded 10 Mar 2005 to 128.111.56.182. Redistribution subject to AIP license or copyright, see http://jap.aip.org/jap/copyright.jsp