FINAL REPORT OF THE
NSF RESEARCH INITIATION AWARD ENTITLED:

SEMICONDUCTOR QUANTUM WAVE DEVICES
ECS-8909971

Award Dates: August 1989 — July 1992

Submitted to
the National Science Foundation
by the Georgia Institute of Technology

PRINCIPAL INVESTIGATOR
Elias N. Glytsis
School of Electrical Engineering and
Microelectronics Research Center

November 1992
SUMMARY

The efforts of the semiconductor industry to scale down the size of devices to achieve higher integration levels and higher speeds have been confronted by the major problem of the "troublesome" quantum-mechanical wave interference effects. These "troublesome" effects start to dominate device characteristics at dimensions of less than about one-quarter of a micron, thus, making further size reduction impossible. This limits the level of integration and the ultimate speed of integrated circuits. Further advances require fundamentally new device concepts and designs. The constructive use of the "troublesome" quantum effects was the basic premise of the proposed "Semiconductor Quantum Wave Optics" research. The long-term scientific objective of the completed research was to generate the necessary knowledge of nanometer-scale semiconductor quantum wave devices that would allow the practical development and use of these structures. Utilization of these quantum devices would be both for ultra-small high-speed versions of present-day devices (such as higher speed switches, oscillators, et.) and for future guided electron wave integrated circuits. This objective of this activity was to bring together understanding of the underlying physics, device modeling, development of software tools, device design methodology, evaluation of existing quantum interference experiments, crystal growth, device fabrication, and device testing. The work completed during the period of the award can be summarized into eight categories: (a) generalization of the analogies between electromagnetic waves in general dielectrics and electron waves in semiconductors, (including analogies with a generalized Hamiltonian) (b) theory and design of narrow-band semiconductor superlattice filter/emitters, (c) negative-differential-resistivity quantum interference devices, (d) theory and design of electron-wave impedance transformers, (e) space-charge effects in quantum wave devices, (f) time-dependent characteristics of the electron-wave interference filters and filter/emitters, (g) electron wave slab waveguides, (h) electron wave quantum interference grating devices, (i) a quantum interference effects in semiconductors bibliography, (j) a proposed innovative testing technique of quantum interference structures using ballistic electron emission microscopy, and (k) three related patents and four patent applications. The aforementioned research results are summarized in the following report.
MAIN RESULTS DURING THE AWARD PERIOD


A comprehensive set of analogies was established between electromagnetic propagation in general isotropic dielectrics (differing permeability and permittivity) and electron wave propagation in semiconductors. First, the electromagnetic results for propagation in non-magnetic dielectrics were generalized to describe propagation, reflection, and refraction in general dielectrics through the definition of separate phase and amplitude refractive indices. Through the analogous definition of electron wave phase and amplitude refractive indices, the expressions for electron wave propagation, reflection, and refraction were shown to have the same functional form as the electromagnetic results. Using these results, the reflectivity characteristics such as total internal reflection (critical angle) and zero reflectivity (Brewster angle) were analyzed as a function of material parameters for both general dielectrics and semiconductor materials. The critical angle and Brewster angle results were then applied to electron wave propagation in the Ga$_{1-x}$Al$_x$As material system, where it is shown that all interfaces in this material will have both a critical angle and a Brewster angle. This was the first prediction of an electron wave Brewster angle in semiconductors.


It was shown that exact, quantitative electromagnetic analogies exist for all forms of the general Hamiltonian of Morrow and Brownstein [Phys. Rev. B 30, 678 (1984)] which applies to single-band effective-mass electron wave propagation in semiconductors. It was further shown that these analogies are valid for propagation in the bulk, propagation past abrupt interfaces between materials, and propagation within two-dimensionally inhomogeneous materials. These results indicate that the correct form of the single-band effective-mass Hamiltonian can be determined through appropriate wavefunction-amplitude-sensitive experiments. Wavefunction-phase-sensitive experiments (such as the measurement of electron wave refraction directions) are not adequate to specify completely the Hamiltonian.
3. Theory and design of narrow-band semiconductor superlattice filter/emitters

A voltage-biased semiconductor superlattice structure was designed to operate as a continuously voltage-tunable, electron-wave interference filter and as an electron emitter. Using the analogies between electromagnetic waves and electron-waves, a systematic procedure was developed for designing the quantum wells and barriers comprising the electron-wave filter/emitter superlattice. A generalized procedure was also developed for analyzing the electron-current transmittance and reflectance spectral responses of these superlattice structures. The constraints required to have thicknesses that are integer multiples of the monolayer thickness and to minimize intervalley scattering were also included in the design. The filter/emitter was shown to operate over a wide tunable energy range. The full width at half maximum (FWHM) of the filter/emitter was shown to decrease as the number of the filter layers was increased. A sensitivity analysis of the device characteristics in the presence of fabrication errors revealed a very stable device response.


The transmission and current-voltage characteristics of semiconductor superlattice electron-wave quantum-interference filter/emitter negative differential resistance devices were analyzed with and without the self-consistency requirement. For the non-self-consistent calculation the single-band effective-mass time-independent Schroedinger equation was solved. For the self-consistent calculation, Schroedinger and Poisson equations were solved iteratively until a self-consistent electron potential energy and electron density were obtained. It was shown that suitably designed electron-wave quantum-interference filter/emitters can exhibit strong negative differential resistance in the current-voltage characteristics, similar to those of resonant tunneling diodes. For low to moderate (2-30 meV) Fermi energies in the conduction band of Ga$_{1-x}$Al$_x$As (doping concentration less or equal to $2 \times 10^{18}$ cm$^3$), and temperatures near 30 K (in the ballistic transport regime), it was shown that the space-charge effect is relatively small and results in a shift of the current-voltage and transmission characteristics toward higher bias voltages. In a fashion similar to that occurring in resonant tunneling diodes, the self-consistent field in electron-wave filter/emitter negative differential resistance devices effectively acts to screen the
positive applied bias. Designs of Ga$_{1-x}$Al$_x$As devices were presented. Resonant devices with current peak-to-valley ratios of 50 as well as nonresonant (not exhibiting negative differential resistance) devices were analyzed. The corresponding electron charge density distributions were also presented. Superlattice electron-wave filter/emitter negative differential resistance devices can be used as high-speed switches and oscillators and as monoenergetic emitters in electroluminescent devices and photodetectors.

5. Theory and design of electron wave impedance transformers

In the construction of semiconductor quantum devices and guided electron-wave integrated circuits it will be necessary to connect semiconductor materials with differing electron energy-band structure. In such a configuration, detrimental reflections will occur at the energy-band discontinuity between materials. These reflections can be eliminated or minimized using impedance transformers. These type of impedance transformers for ballistic (collisionless) electron-waves traveling between dissimilar energy-gap semiconductors were designed as a series of quarter (electron) wavelength layers in the form of a compositional superlattice, using the quantitative analogies between electromagnetic and electron waves. For the design energy, the electron-wave could be totally transmitted and the structure can be analogous to an antireflection coating in electromagnetic optics. Procedures for designing narrow-band, maximally flat (Butterworth), and equal-ripple (Chebyshev) impedance transformers of arbitrary spectral bandwidth were developed.

6. Space Charge Effects in Quantum Wave Devices

Theoretical techniques for studying the properties of electron wave superlattice structures have been developed. Specifically, we have included self-consistent effects in the design of superlattice electron-wave filters through the simultaneous solution of the Poisson and Schroedinger Equations. It was found that at doping concentrations less than or equal to $2 \times 10^{18}$ cm$^3$ and temperatures near 30 K (ballistic regime) that space-charge effects are relatively small in superlattice electron-wave filters. Nevertheless, self-consistent effects act in a similar manner as in resonant tunneling diodes; the current-voltage shifts towards higher voltages. In addition, we have solved the time-dependent Schroedinger Equation in order to determine the temporal response of electron-wave superlattice filters. It was found that the frequency response is much greater in these devices than in existing resonant tunneling devices offering the possibility of a new ultra-high frequency (THz) oscillator.
7. Time-Dependent Characteristics of Semiconductor Resonant Structures

Double-barrier tunneling structures operate based on quantum mechanical tunneling through two barriers. Quantum electron wave structures operate based on traveling-wave propagation above all conduction band edges. These are the fundamental structures proposed to realize ballistic electron transport devices in semiconductors. The time-dependent behavior of resonant tunneling structures has been discussed extensively in the literature, but no such analysis has been performed on quantum wave structures. A numerical solution of the time-dependent effective mass equation was used to calculate the traversal time of a Gaussian packet and the percentage of the packet transmitted for resonant tunneling and quantum wave structures.

8. Electron Waveguiding in Quantum Wells, Voltage-Induced Quantum Wells, and Quantum Barriers

Recent experiments have produced ballistic electron transport over micron lengths in semiconductor two-dimensional electron gas (2DEG) systems. This has made possible the demonstration of electron devices that exhibit impressive optical-like behavior. In these devices, the quantum well at the 2DEG interface acts as a slab waveguide for ballistic electron waves. In this work, we showed how finite-potential heterostructure wells, homostructure voltage-induced wells, and heterostructure barriers can act as electron slab waveguides. We found that the waveguiding in all of these structures is described by a single dispersion relation and can occur at energies above all band edges. The guided mode cutoffs, electron velocity, effective mass, density of states, and ballistic current density were determined. A multiple layer theory was developed to analyze wells and barriers with arbitrary potential energy profiles. The maximum ballistic guided current flowing in a given direction for a 10 monolayer Ga$_{0.75}$Al$_{0.25}$As/GaAs/-Ga$_{0.9}$Al$_{0.1}$As waveguide was found to be 2.3 mA per micron of waveguide width. This relatively large value suggests that interconnecting multiple ballistic electron devices through a single slab waveguide may be feasible.


Due to recent advances in the growth and fabrication of nanostructure electronic devices, it has been demonstrated that ballistic electron waves can be reflected, refracted,
interfered, waveguided, and diffracted in a manner analogous to electromagnetic optics. This has provided a surge of interest in the new field of semiconductor electron wave optics. In this paper, ballistic electron grating diffraction by a grating with an arbitrary effective mass and/or potential energy profile was analyzed using a rigorous coupled-wave analysis (RCWA). These results are related to electromagnetic diffraction by a permittivity grating. It was shown that electron diffraction by a kinetic energy grating was exactly analogous to TE electromagnetic diffraction and that electron diffraction by an effective mass grating was exactly analogous to TM electromagnetic diffraction. Approximate solutions to the RCWA equations were derived that are equivalent to Bragg regime and Raman-Nath diffraction. Using these results, sample electron wave diffractive switches and multiplexers were designed using achievable device configurations. The angular and energy selectivity of these devices were examined.

10. Quantum Interference Effects in Semiconductors: A Bibliography

Refinements in growth techniques such as molecular beam epitaxy (MBE) have produced materials with ballistic (collisionless) electron transport lengths of over a micron. Coupled with nanolithography it is now possible to fabricate structures with both lateral and vertical dimensions on the order of the deBroglie wavelength of a ballistic electron. In these structures quantum interference effects can dominate the electronic behavior. In view of the rapidly expanding interest and activity in this area, the following bibliography has been compiled as an introduction and study guide to this field. The papers listed describe the extensive theoretical and experimental results that have been obtained on quantum interference effects as well as discuss possible application areas. Works of a fundamental nature concerning phenomena that are basic to all semiconductor behavior have not been included. Articles on the properties and band structure of semiconductors, which are essential to a complete understanding of quantum interference effects, have not been included. Conference papers, though frequently very important, have not been included to conserve space. The papers are listed alphabetically according to the first author's surname. As in the compilation of any bibliography, numerous valuable and pertinent articles have probably been inadvertently omitted.


Ballistic electron emission microscopy (BEEM) has recently been developed to study
the electrical properties of buried interfaces where ballistic electrons are injected into a sample using a scanning tunneling microscope (STM). In this paper, a method was proposed that uses the BEEM technique to observe electron wave optical properties of ballistic transport in semiconductors. This method provides a three-terminal configuration for characterizing electron wave devices that overcomes many of the problems encountered in traditional two- and three-terminal techniques. Specifically, the method provides a highly collimated beam of ballistic carriers with a precisely controlled energy distribution. These carriers probe the quantum transmittance of a voltage-tunable electron wave interference device with minimal impurity scattering. A general procedure was presented for analyzing this experimental configuration based on a combination of the models used to describe BEEM and ballistic electron transport in semiconductors.

Using this procedure, a BEEM analysis of an electron wave energy filter was modeled, showing clear electron wave interference effects. This BEEM configuration allows for the precise characterization of a wide range of ballistic electron transport effects, such as quantum reflections from interfaces and electron wave interference effects, that are currently of wide interest.


Solid state quantum mechanical electron or hole wave devices which are analogous to optical thin-film devices provide among other things, energy selectivity for substantially ballistic electron or hole wave propagation in superlattice structures at energies above the superlattice potential energy barriers. Further, in accordance with this patent, the inventive devices may be designed by transforming existing optical thin-film design methods and existing optical interference filter designs into inventive semiconductor devices. This transformation from existing optical design methods and existing optical interference filter designs into semiconductor devices is performed for electron devices by mapping the optical phase index of refraction into a first solid state index of refraction for phase quantities which is proportional to the square root of the product of the electron kinetic energy and the electron effective mass and by mapping the optical amplitude index of refraction into a second solid state index of refraction for amplitude quantities which is proportional to the square root of the electron kinetic energy divided by the electron effective mass.

This patent pertains to solid state quantum mechanical electron and hole wave devices
and method for fabricating them and, in particular, to solid state quantum mechanical electron and hole wave devices such as, without limitation, low pass filters, high pass filters, narrow band and wide band notch filters, narrow band and wide band bandpass filters, impedance transformers, resonant electron and hole emitters, and so forth and method for fabricating them.

Recent progress in semiconductor growth technologies, particularly in molecular beam epitaxy (MBE) and metal organic chemical vapor deposition (MOCVD), enable the growth of multilayered superlattice structures with precise monolayer compositional control. For example, successively grown layers of narrow and wide band gap semiconductor materials such as GaAs and Ga\(_{1-x}\)Al\(_x\)As have been produced and widely used to provide multiple quantum well structures. In fact, there are many references in the prior art which are concerned with the use of these superlattice structures in resonant tunneling superlattice/multiple quantum well devices. Specifically, in such devices, a superlattice is formed by growing successive layers of narrow and wide band gap semiconductor material epitaxially and the materials and the widths of the layers in these devices are chosen so that quantum states which arise from spatial quantization effects in adjacent wells become coupled. Further, in such devices, the interaction of these coupled states leads to the formation of minibands of allowed energies through which carriers can tunnel.

The first embodiments of the inventive solid state electron wave devices comprise solid state analogs of Fabry-Perot optical interference filters which may be fabricated from alloys of GaAlAs and GaAs.


Semiconductor, quantum well, electron and hole slab waveguides include a substrate, semiconductor layer, a film semiconductor layer, and a cover semiconductor layer, wherein the semiconductor layers provide substantially ballistic transport for electrons and wherein the thicknesses and compositions of the semiconductor layers are determined in accordance with this patent to provide a waveguide.

Embodiments of the patent solve the above-identified need in the art by providing semiconductor, quantum well, electron and hole slab waveguides. Specifically, an electron slab waveguide is comprised of a substrate semiconductor layer, a film semiconductor layer,
and a cover semiconductor layer, wherein the semiconductor layers provide substantially ballistic transport for electrons and wherein the thicknesses and compositions of the semiconductor layers are determined in accordance with the inventive method which is set forth in detail below to provide a potential well.

In particular, in accordance with this patent, electron waveguide modes exist for electron energies in the well and for electron energies above one or both of the potential energy barriers of the substrate layer and the cover layer, respectively. Further, in contrast to the behavior of electromagnetic guided waves which only have a lower-energy cutoff due to dispersion, each electron waveguide mode also has an upper-energy cutoff wherein an electron wave is refracted into the substrate layer and/or the cover layer.

Doping of semiconductors is not required for embodiments of the present invention, however, it is preferred that doping not be done within the active region of the device in order to avoid scattering within the materials. This provides a further advantage for the inventive waveguide devices because the absence of doping makes them easier to fabricate.

Note that semiconductor electron slab waveguides can perform as described provided that ballistic transport can be achieved over sufficient distances and that the density of electrons is small enough to make electron-electron interactions negligible.

The inventive electron waveguides should be useful in high-speed electronic circuitry and as a central component in electron guided wave integrated circuits.


Continuously tunable, biased, semiconductor superlattice electron interference filter/emitter which can serve, for example, as a hot electron emitter in a ballistic transistor, provides energy selectivity for substantially ballistic electron wave propagation at electron energies above the superlattice potential barriers. The layers of the biased superlattice have alternatively high and low electron refractive indices wherein each layer is a quarter or half of an electron wavelength in thickness and wherein the quantum well barrier widths are adjusted in the direction of emission to provide the desired energy selectivity.

Embodiments of this patent solve the above-identified need in the art by providing biased, semiconductor superlattice tunable electron interference filters/emitters which can
serve, for example, as a hot electron emitter in a ballistic transistor. In particular, embodiments of the present invention comprise biased superlattice filter/emitters which provide energy selectivity for substantially ballistic electron wave propagation at electron energies above the superlattice potential barriers. Further, the layers of the biased superlattice comprise alternately high and low electron refractive indices wherein each layer is a multiple of a quarter or half of an electron wavelength in thickness and wherein the quantum well barrier widths are adjusted in the direction of emission to provide the desired energy selectivity.

Specifically, embodiments of the present invention comprise biased, semiconductor superlattice filter/emitters which are designed, in accordance with the inventive method, by transforming optical, thin-film interference filter designs which are designed in accordance with existing optical interference filter design methods into inventive semiconductor devices. In particular, the transformation from an optical interference filter design is performed by mapping the optical phase index of refraction into a solid state index of refraction for phase quantities which is proportional to the square root of the product of electron kinetic energy and electron effective mass and by mapping the optical amplitude index of refraction into a solid state index of refraction for amplitude quantities which is proportional to the square root of the electron kinetic energy divided by the electron effective mass. That is, the mapping makes an exact analogy between an electromagnetic optical wave and a quantum mechanical electron wave by using the electron wavevector above in place of the electromagnetic optical wavevector and by using the electron wave amplitude refractive index in expressions for reflectivity and transmissivity at a boundary, which expressions are well known from electromagnetic design.

The efficacy of the above-defined mapping between electromagnetic optical waves and quantum mechanical electron waves depends on the existence of ballistic electron transport in the solid state materials, i.e., where electrons travel through the solid state materials without being scattered by deviations from crystalline perfection. In the inventive filter/emitter devices, the ballistic electrons have energies above the potential barriers in the solid state materials and exhibit quantum mechanical plane wave behavior. Further, since these plane waves maintain their phase through the device, these coherent waves will refract, reflect, interfere, and diffract in a manner which is analogous to the behavior of electromagnetic waves traveling through dielectrics.

A solid state, quantum mechanical electron/hole wave device in the form of a switch or a multiplexor includes a layer of semiconductor material supporting substantially ballistic electron/hole transport and a periodic grating structure formed in the layer of semiconductor material, with the grating structure comprising a modulation in electron/hole potential energy and/or effective mass. Preferably, means are provided for applying and varying the grating modulation. By constructing the device to divide the input substantially completely into two output beams (0 and +1) (to operate in the Bragg regime), a useful switch is provided. Likewise, by constructing the device to divide the input into a selected number of three or more output beams (±1, 0, ±2) (to operate in the Raman-Nath regime), a useful multiplexor (broadcasting device) is provided.

HUMAN RESOURCES DEVELOPMENT

Three graduate students have been involved in the research work described previously. These are: (a) Mr. Gregory N. Henderson (Ph.D. expected December 1992) who has been working on analogies between electron wave and electromagnetic wave propagation, electron wave grating interference devices, and ballistic electron emission testing of quantum heterostructures; (b) Mr. Daniel W. Wilson (Ph.D. candidate) who has been working on electron wave slab waveguides as part of his thesis work; (c) Mr. David B. Walker (Ph.D. candidate) who has been working on time-dependent simulations of vertical transport devices. In addition, a post doctoral fellow, Dr. Karim Diff, was partially supported for a six month period working also on time-dependent solutions of the Schrödinger equation. Two undergraduate students have recently been involved in this research work. Mr. Daniel Guthrie (EE senior) has been working on simulations of quantum interference vertical transport biased and unbiased structures and calculations of density of states, and Ms. Patricia Dantzscher (EE senior) who has been working on software and hardware development for the scanning tunneling microscope.

JOURNAL PUBLICATIONS RESULTING FROM THE NSF AWARD


CONFERENCE PAPERS RESULTING FROM THE NSF AWARD


**PATENTS AND PATENT APPLICATIONS RESULTING FROM THE NSF AWARD**


8. G. N. Henderson, T. K. Gaylord, and E. N. Glytsis, “Quantum-mechanical semicon-
APPENDIX
Copies of most important Journal Publications
Semiconductor superlattice interference filter design

T. K. Gaylord, E. N. Glytsis, and K. F. Brennan
School of Electrical Engineering and Microelectronics Research Center, Georgia Institute of Technology,
Atlanta, Georgia 30332

(Received 25 October 1988; accepted for publication 21 November 1988)

The quantitative analogies that have been previously established [J. Appl. Phys. 65, 814 (1989)] between electron wave propagation in semiconductors and optical wave propagation in dielectrics may be used to translate thin-film optical device designs into semiconductor superlattice device designs. The procedure for this direct mapping is also described in the above reference. The resulting designs, however, have compositions that are not constrained to be within a usable compositional range and they have layer thicknesses that are not constrained to be integer multiples of a monolayer thickness. In the present work, a systematic design procedure is presented that includes these required practical constraints. This procedure is then applied to the design of Ga_{1-x}Al_{x}As superlattice narrow interference filters. For pass kinetic energies in the range of 0.14-0.20 eV, compositions (values of x) and numbers of monolayer thicknesses needed to produce quarter-wavelength layers are calculated. The detailed design of an example narrow bandpass (15.4 meV) filter with a pass electron energy of 0.20 eV is presented.

I. BACKGROUND AND MOTIVATION

The development and refinement of molecular-beam epitaxy (MBE) and metalorganic chemical vapor deposition (MOCVD) has enabled semiconductor multilayered superlattices to be grown with precise monolayer compositional control. Successively grown layers of narrow and wide band-gap semiconductor materials such as GaAs and Ga_{1-x}Al_{x}As have produced the well-known and widely used multiple quantum-well structures. Furthermore, improvements in the quality of materials grown have simultaneously produced ballistic electron transport. That is, electrons travel through the device without being scattered (by deviations from crystalline perfection). The resulting ballistic electrons exhibit clear quantum mechanical plane-wave behavior. Since they maintain their phase throughout the device, these coherent waves will refract, reflect, interfere, and diffract in a manner analogous to electromagnetic waves. Quantum interference effects have been observed experimentally in single-well double-barrier GaAs/Ga_{1-x}Al_{x}As and In_{1-x}Ga_{x}As/In_{1-y}Al_{y}As potential barrier structures. These measurements were made for single-well, double-barrier structures. Of course, electron interference effects are not limited to these single-well structures. In fact, the use of multilayer superlattices and electron energies above the potential barriers allows generalized semiconductor superlattice interference filters to be designed that are directly analogous to thin-film optical interference filters. These superlattice interference filters may exhibit very narrow passbands with full widths at half maximum (FWHM) on the order of only several meV. Monolithically integrated into solid-state devices, these structures have potential use as monoenerygetic emitters for electroluminescent devices, photodetectors, and fast ballistic transistors. In addition, these devices could assist in the control of free-space electron beams in fields such as electron spectroscopy, electron beam lithography, and electron diffraction analysis of crystals.

Electron wave propagation at energies above the potential barriers can be described by the analogies that exist between electron waves in semiconductors and electromagnetic waves in dielectrics. However, semiconductor superlattice interference filter designs cannot simply be copies of thin-film optical filter designs. In the design of semiconductor superlattice interference filters, the thicknesses of the layers are restricted to be integer multiples of the monolayer thickness. In addition, there is typically only a limited usable composition range available. The purpose of this paper is to present a systematic design procedure for semiconductor superlattice interference energy filters subject to these constraints and then to apply this procedure to the example case of designing Ga_{1-x}Al_{x}As superlattice filters.

II. SEMICONDUCTOR ELECTRON WAVE OPTICS

Quantum mechanical electron waves in semiconductors and electromagnetic optical waves in dielectrics exhibit transmission, reflection, interference, and diffraction characteristics that are analogous to each other. At an infinitely thick potential energy barrier, for example, an electron wave with energy above the barrier will be totally internally reflected if the angle of incidence is greater than the electron wave critical angle. Likewise, in semiconductor superlattices, an electron wave may be totally transmitted or totally reflected depending on its energy and angle of incidence. These characteristics are analogous to the corresponding optical thin-film devices (an antireflection coating and a dielectric coated mirror). Quantum interference effects have been analyzed for single potential energy boundaries, for normal incidence, and for the general case of any number of boundaries and any angle of incidence. From these results, a mapping has been established between electromagnetic and quantum mechanical quantities. As a consequence, existing optical device designs now have electron wave device counterparts.

To describe phase effects such as in wave interference
for either electron or optical waves, the corresponding wave vector is used. For an electron wave, the wave-vector magnitude is
\[ k = \left[ 2m^* (E - V) \right]^{1/2}. \] (1)
where \( m^* \) is the electron effective mass, \( E \) is the total electron energy, \( V \) is the electron potential energy, and \( \hbar \) is Planck's constant divided by \( 2\pi \). Thus, the electron wave phase refractive index, \( n_e \) (phase), is proportional to the square root of the product of the effective mass and the kinetic energy. That is,
\[ n_e (\text{phase}) \propto \left[ (E - V)/m^* \right]^{1/2}. \] (2)
Amplitude effects such as transmissivity and reflectivity may be described in terms of the wave-function amplitude for an electron wave or in terms of the electric field amplitude for an optical wave incident upon a boundary. Continuity of the wave function across a potential energy boundary is analogous to the continuity of the tangential component of the electric field across a boundary between dielectrics. Similarly, conservation of electron probability current normal to a potential energy boundary is analogous to conservation of power flow normal to a boundary between dielectrics. Using these, the electron wave and optical wave reflectivities and transmissivities are made equivalent by introducing, in quantum mechanics, an analogous index of refraction or characteristic impedance.\(^{18-22}\) Thus, the electron wave amplitude refractive index \( n_e \) (amplitude) is proportional to the square root of the ratio of the kinetic energy to the effective mass,\(^5\) and so
\[ n_e (\text{amplitude}) \propto \left[ (E - V)/m^* \right]^{1/2}. \] (3)
In the expressions for transmissivity and reflectivity, only dimensionless ratios of the electron wave amplitude refractive indices occur. Both types of electron wave refractive indices exhibit normal dispersion. That is, they increase with decreasing wavelength.

III. THIN-FILM OPTICAL INTERFERENCE FILTERS

Superlattice electron wave interference filters share common characteristics with thin-film optical interference filters. Therefore it is useful initially to review some of the primary properties of these optical filters.\(^{18-22}\) This is particularly appropriate since the procedure for designing superlattice electron wave filters is an extension of that used for thin-film optical filters.

The simplest type of narrow bandpass optical interference filter is the Fabry–Perot filter. It consists of a half-wavelength layer (frequently called a “spacer” in optical literature) sandwiched between reflectors. In the case of an all-dielectric Fabry–Perot filter, the reflectors are stacks of high index (designated \( H \)) and low index (designated \( L \)) quarter-wavelength layers. Only this elementary type of interference filter will be treated in this paper. The FWHM of the bandpass in this type of filter can be reduced by increasing the reflectivity at the boundaries between the layers. This may be accomplished by increasing the ratio of the high index \( (n_H) \) to the low index \( (n_L) \). Furthermore, for a given number of layers, the higher reflectance occurs with high index \( (H) \) layers on the outside boundaries of the filter. The half-wavelength (resonant) layer at the center of the filter may be of high index \( (n_H) \) or low index \( (n_L) \) material. Thus there are two basic types of all-dielectric Fabry–Perot interference filters. In the optical literature, these are symbolically represented as \([HL]^{N} [HH] [HL]^{N} \) and \([HL]^{N} LL [HL]^{N} H\) where \( H \) and \( L \) represent quarter-wavelength layers of high and low index materials, respectively, and \( N \) represents the number of repetitions of the layer-pair type indicated in square brackets.

Other important characteristics of the all-dielectric Fabry–Perot interference filters are as follows: (1) The maximum transmittance of the filter is 100%. (2) The maximum transmittance occurs at the wavelength for which the spacer layer is a half-wavelength thick (as measured in that material) and the reflector layers are a quarter-wavelength thick (as measured in those materials). This wavelength will be called the pass wavelength in this paper. (3) The FWHM decreases as the number of quarter-wavelength layers is increased. (4) The transmittance characteristic is symmetric about the pass wavelength when the characteristic is plotted as a function of the reciprocal of the wavelength (as measured in the material surrounding the filter). (5) A proportional change in the thicknesses of all the layers produces a simple displacement of the transmittance characteristic plotted as a function of the reciprocal wavelength. (6) If the thicknesses of all layers are increased by an odd integer factor (the parameter \( q \), in Sec. V), a passband will occur at the original pass wavelength and it will have a decreased FWHM. (7) As the angle of incidence upon the filter is increased, the pass wavelength is tuned to shorter wavelengths. (8) The transmittance characteristics are relatively insensitive to variations in the reflectivities and thicknesses of the layers. (9) Normal dispersion causes a narrowing of the FWHM. (10) The filter is effective over only a limited range since sidebands necessarily occur on either side of the passband. The wavelength range around the pass wavelength over which the transmission is low is called the rejection range. The range from the nearest passband peak below the pass wavelength to the nearest passband peak above the pass wavelength is called the free spectral range (FSR). Other terminology commonly employed in describing these filters include the finesse \( (\text{FSR}/\text{FWHM}) \) and the resolving power (pass wavelength/\( 4\text{FWHM} \)).

IV. SEMICONDUCTOR SUPERLATTICE INTERFERENCE FILTERS

Using the above electron wave refractive indices established from considering a single potential energy boundary, the characteristics of a many-boundary semiconductor superlattice system can be found.\(^5\) This may be done by applying the chain matrix approach commonly employed in electromagnetics\(^{23,24}\) but substituting the electron wave vector as given by Eq. (1) for the optical wavevector and using the electron wave amplitude refractive index as given by Eq. (3) in the expressions for the reflectivity and transmissivity at a boundary.\(^5\)
In layer \( m - 1 \) the amplitude of the electron wave incident (traveling in the positive direction) upon layer \( m \) is designated \( \psi_{r,m} \). In layer \( m - 1 \) the amplitude of the electron wave reflected (traveling in the negative direction) from layer \( m \) is designated \( \psi_{t,m} \). These complex amplitudes may be expressed in terms of the corresponding amplitudes \( \psi_{r,0} \) and \( \psi_{t,0} \) incident upon and reflected from the boundary between layer \( m \) and layer \( m + 1 \) as

\[
\begin{bmatrix}
\psi_{r,m} \\
\psi_{t,m}
\end{bmatrix} = \frac{1}{r_{r,m}} \begin{bmatrix} 1 & r_{r,m} \\ 0 & 1 \end{bmatrix} \exp(jk_{r,m}d_m \cos \theta_m) \begin{bmatrix}
\psi_{r,m-1} \\
\psi_{t,m-1}
\end{bmatrix} + \frac{1}{r_{t,m}} \begin{bmatrix} 1 & r_{t,m} \\ 0 & 1 \end{bmatrix} \exp(-jk_{t,m}d_m \cos \theta_m) \begin{bmatrix}
\psi_{r,m+1} \\
\psi_{t,m+1}
\end{bmatrix},
\]

where \( r_{r,m} \) is the amplitude transmissivity of the interface between layer \( m - 1 \) and layer \( m \), \( r_{t,m} \) is the amplitude reflectivity at the same interface, \( k_{r,m} \) is the magnitude of the electron wavevector in layer \( m \) as given by Eq. (1), \( d_m \) is the thickness of layer \( m \), and \( \theta_m \) is the angle of the wavevector direction in layer \( m \). For a stack of \( M \) layers, the total normalized transmitted electron wave amplitude \( \psi_{r,M+1} \) (in region \( M +1 \)) and the total normalized reflected electron wave amplitude \( \psi_{t,0} \) (in region 0) are obtained by chain multiplying a total of \( M +1 \) versions of Eq. (4) together, one for each of the \( M \) layers and one for the output region. The result is

\[
\begin{bmatrix}
\psi_{r,M+1} \\
\psi_{t,0}
\end{bmatrix} = \prod_{m=1}^{M} \begin{bmatrix} 1 & r_{r,m} \\ 0 & 1 \end{bmatrix} \exp(jk_{r,m}d_m \cos \theta_m) \begin{bmatrix}
\psi_{r,M+1} \\
\psi_{t,0}
\end{bmatrix}
\]

and this can be solved directly for the amplitude transmissivity \( \psi_{r,M+1} \) and the amplitude reflectivity \( \psi_{t,0} \).

The procedure for directly translating a thin-film optical filter design into a superlattice electron filter design has been described in Ref. 5. In that reference, the procedure was applied to translate an 11-layer thin-film optical filter into a corresponding superlattice filter. The resulting design, however, has compositions that are not constrained to fit a usable compositional range and it has layer thicknesses that are not integer multiples of a monolayer thickness. In the next section, a systematic design procedure that includes these practical constraints is presented.

V. SEMICONDUCTOR SUPERLATTICE INTERFERENCE FILTER DESIGN

A semiconductor compositional superlattice is depicted in Fig. 1. The given material system is taken to form a continuous set of alloys of the type \( F_{1-x}G_xH \). The range of usable compositions is \( 0 \leq x \leq x_{max} \) (for example, due to a possible transition at \( x_{max} \) from a direct to an indirect energy gap as in the case of \( GaAlAs \)). The materials surrounding the filter \( (i = 0 \) regions) have compositions \( x_0 \). The superlattice itself consists of two compositions of material. The composition in the high (electron wave amplitude) refractive index regions is \( x_1 \) \( (i = 1 \) regions). The composition in the low (electron wave amplitude) refractive index regions is \( x_2 \) \( (i = 2 \) regions). The monolayer thicknesses are \( r_1 \) \( (i = 1 \) region) and \( r_2 \) \( (i = 2 \) region). The electron potential energy in the three types of regions is given by

\[
V_i = \Delta E_x = A x_i, \quad i = 0, 1, 2,
\]

where \( \Delta E_x \) is the change in the energy of the conduction-band edge and \( A \) is a constant. The electron effective mass in the three types of regions is given by

\[
m^* = (B + C x_i) m_0, \quad i = 0, 1, 2,
\]

where \( B \) and \( C \) are constants and \( m_0 \) is the free-electron mass. The electron kinetic energy in the \( i \)th region is

\[
E = V_i = \hbar^2 / 2m^* \lambda_i^2, \quad i = 0, 1, 2,
\]

where \( \lambda_i \) is the pass wavelength as measured in the \( i \)th region. The overall pass kinetic energy of the filter (as measured in the material surrounding the filter) is simply \( E_o = V_0 \). This is the pass kinetic energy that will be specified by the user and is thus the starting point in the design procedure. Substituting Eqs. (6) and (7) into Eq. (8) and solving
for the pass wavelength gives
\[ (\lambda_p) = \frac{h}{(2m_0) \left( ACx^2 + (CE_p - AB)x + BE_p \right)^{1/2}}, \]
\[ i = 0, 1, 2. \]  
(9)

The thicknesses of the superlattice layers are designated \( d_i \), \( i = 1, 2 \). These thicknesses must be integer multiples of the monolayer thicknesses \( r_i \). Furthermore, these thicknesses must also be odd multiples of a quarter wavelength as measured in these regions. These constraints may be expressed as
\[ d_i = p_ir_i = (2q_i - 1)(\lambda_p)/4, \quad i = 1, 2, \]
(10)

where \( p_i \) is the integer number of monolayers for the \( i \)-th region and \( q_i \) is a positive integer \( (q_i = 1, 2, 3, ...) \). Eliminating \( (\lambda_p) \), between Eqs. (9) and (10) gives the following quadratic equation in the composition \( x_i \):
\[ ACx^2 + (AB - CE_p)x + (h^2/32m_0) \times [(2q_i^2 - 1)^2/p_i^2r_i^2] - BE_p = 0. \]  
(11)
The solution for the composition \( x_i \) is
\[ x_i = \left[ -b \pm (b^2 - 4ac)^{1/2} \right]/2a, \]
(12)
where
\[ a = AC, \quad b = AB - CE_p, \]
and
\[ c_i = (h^2/32m_0) \times [(2q_i^2 - 1)^2/p_i^2r_i^2] - BE_p. \]

In order to design a superlattice interference filter, at least two solutions for \( x_i \) must be found in the range \( 0 < x_i < x_{\text{max}} \). The smallest value of \( x_i \) within this range will become \( x_1 \), the composition of the high index material. The value of \( p_i \) that produces \( x_1 \) becomes \( p_1 \), the number of monolayers of type 1 material used to make a quarter-wavelength layer. Similarly, the largest value of \( x_i \) within this range will become \( x_2 \), the composition of the low index material. The value of \( p_i \) that produces \( x_2 \) becomes \( p_2 \), the number of monolayers of type 2 material used to make a quarter-wavelength layer.

To allow the broadest range of solutions, \( V_0 \) is set equal to \( V_{\text{max}} \). For a specified pass kinetic energy \( (E_p - V_0) \), the value of \( E_p \) is then determined. Furthermore, to minimize the total thickness of the filter, \( q_i \) is initially set equal to unity. Then Eq. (12) is repetitively evaluated for \( p_i = 1, 2, 3, \ldots \) until all of the positive real roots are found in the range \( 0 < x_i < x_{\text{max}} \). For the present case there are two roots. They are \( x_1 = 0.2063 \) (corresponding to \( p_1 = 6 \) and \( x_2 = 0.3984 \) (corresponding to \( p_2 = 7 \)). The smaller value of \( x_i \) is designated \( x_1 \) and the larger value \( x_2 \). A plot of the positive real root from Eq. (12) is shown as a function of \( p_i \) (taken as a continuous variable) in Fig. 2. The thickness of the \( \text{Ga}_{1-x}\text{Al}_x\text{As} \) quarter-wavelength layer is \( d_1 = p_1r_1 = 16.9599 \) \( \AA \) (6 monolayers). The thickness of the \( \text{Ga}_{0.60}\text{Al}_{0.40}\text{As} \) quarter-wavelength layer is \( d_2 = p_2r_2 = 19.7866 \) \( \AA \) (7 monolayers). The electron effective masses in the three regions are calculated from Eq. (7) and are \( m_1^e = 0.10435m_0, \quad m_2^e = 0.084123m_0, \quad m_3^e = 0.10007m_0 \). The pass kinetic energies in the three regions are calculated from Eq. (8) and are \( E_1 - V_0 = 0.2000 \) eV, \( E_2 - V_1 = 0.3884 \) eV, \( E_3 - V_2 = 0.2399 \) eV.

The electron wave amplitude refractive index for each type of region can be calculated from
\[ n_i(\text{amplitude}) \propto \left[ (E - V_i)/m_i^e \right]^{1/2}, \]
(13)
for the pass energy \( E_p \) or for any other energy. Similarly, the electron wave-vector magnitude in the 4th region is
\[ k_i = \left[ 2m_i^e(E - V_i) \right]^{1/2}/\hbar. \]  
(14)

VI. EXAMPLE \( \text{Ga}_{1-x}\text{Al}_x\text{As} \) SUPERLATTICE INTERFERENCE FILTER DESIGNS

To illustrate the design procedure of the previous section, superlattice interference filters fabricated from the \( \text{Ga}_{1-x}\text{Al}_x\text{As} \) material system will be treated here. This system is perhaps the most technologically advanced of current semiconductor superlattice material systems. For these alloys, all compositions are lattice matched. For growth along the \([100] \) direction, the monolayer thickness is
\[ r_1 = r_2 = 2.8266 \] \( \AA \). The material is a direct gap semiconductor for \( 0 < x < 0.45 \) and this represents the usable composition range. Furthermore, for \( \text{Ga}_{1-x}\text{Al}_x\text{As}, \)
\[ A = 0.77314 \text{ eV}, B = 0.067, \quad C = 0.083. \]

As an example, to design a \( \text{Ga}_{1-x}\text{Al}_x\text{As} \) superlattice interference filter with a pass kinetic energy of 0.20 eV (such as for an emitter in a high-speed ballistic transistor), the following calculations are performed. Let \( x_0 = x_{\text{max}} = 0.45 \) and \( E_0 = 0.347913 \) eV from Eq. (6). Since \( E_p - V_0 = 0.20 \) eV, then \( E_p = 0.547913 \) eV. Letting \( q_i = 1 \), the composition \( x_i \) is evaluated for \( p_i = 1, 2, 3, \ldots \) using Eq. (12) until all of the positive real roots are found in the range \( 0 < x < 0.45 \). For the present case there are two roots. They are \( x_1 = 0.2063 \) (corresponding to \( p_1 = 6 \) and \( x_2 = 0.3984 \) (corresponding to \( p_2 = 7 \)). The smaller value of \( x_i \) is designated \( x_1 \) and the larger value \( x_2 \). A plot of the positive real root from Eq. (12) is shown as a function of \( p_i \) (taken as a continuous variable) in Fig. 2. The thickness of the \( \text{Ga}_{0.70}\text{Al}_{0.30}\text{As} \) quarter-wavelength layer is \( d_1 = p_1r_1 = 16.9599 \) \( \AA \) (6 monolayers). The thickness of the \( \text{Ga}_{0.60}\text{Al}_{0.40}\text{As} \) quarter-wavelength layer is \( d_2 = p_2r_2 = 19.7866 \) \( \AA \) (7 monolayers). The electron effective masses in the three regions are calculated from Eq. (7) and are \( m_1^e = 0.10435m_0, \quad m_2^e = 0.084123m_0, \quad m_3^e = 0.10007m_0 \). The pass kinetic energies in the three regions are calculated from Eq. (8) and are \( E_1 - V_0 = 0.2000 \) eV, \( E_2 - V_1 = 0.3884 \) eV, \( E_3 - V_2 = 0.2399 \) eV. The electron wave amplitude refractive indices normalized to the surrounding \( \text{Ga}_{1-x}\text{Al}_x\text{As} \) are calculated from Eq. (13) and are \( n_1(\text{amplitude}) = 1.00000, \quad n_2(\text{amplitude}) = 1.552027, \quad n_3(\text{amplitude}) = 1.118372. \) For an

![Figure 2](image-url)
elementary 13-layer Fabry–Perot interference filter of the form \([HL]^6[HH][LH]^6\), these calculated material characteristics produce the transmission characteristic shown in Fig. 3. The pass energy is indeed 0.20 eV and the FWHM of the filter is 15.4 meV.

Repeating the above procedure, Ga\(_{1-x}\)Al\(_x\)As superlattice interference filters were designed for pass kinetic energies from 0.14 eV up through 0.20 eV, the range of energies potentially most useful in ballistic transistors. The positive real roots as given by Eq. (12) are shown in Table I for 6–10 monolayer thicknesses. Roots, of course, must be in the range \(0 < x < 0.45\). At the 0.14 eV low end of this energy range, there are essentially four roots. At the 0.20 eV high end of the energy range, there are two roots. The Ga\(_{1-x}\)Al\(_x\)As superlattice interference filter described in Ref. 10 (quarter-wavelength layers produced by 6 monolayers of GaAs and by 9 monolayers of Ga\(_{0.35}\)Al\(_{0.65}\)As) essentially corresponds to the 0.14-eV pass energy case given in Table I.

### VII. SUMMARY

The quantitative analogies that have been established between electron wave propagation in semiconductors and optical wave propagation in dielectrics may be used to translate thin-film optical device designs into semiconductor superlattice device designs. The procedure for this direct mapping has previously been described and applied to the case of semiconductor superlattice narrow bandpass interference filters in Ref. 5. The resulting designs, however, have compositions that are not constrained to be within a usable compositional range and have layer thicknesses that are not integer multiples of a monolayer thickness. In the present work, a systematic design procedure that includes these practical constraints is presented. This method is applicable to any material system. For illustration, this procedure is applied to the case of designing Ga\(_{1-x}\)Al\(_x\)As superlattice interference filters. For pass kinetic energies in the range investigated (0.14–0.20 eV), compositions and numbers of monolayer thicknesses needed to produce one \((q, = 1)\) quarter-wavelength layers are presented (Table I). Like their thin-film optical counterparts, semiconductor superlattice interference filters would be relatively insensitive to variations about the design composition values. Furthermore, the effects of nonparabolic and anisotropic energy band structure, if present, may be incorporated into the design process by using an energy-dependent anisotropic effective mass.

There is considerable flexibility in the design of semiconductor superlattice interference filters. For example, other odd multiples of a quarter wavelength may be used \((q, = 2, 3, 4, \ldots)\), the surrounding material can be changed \((V_o)\), and other crystallographic growth directions may be used \((\alpha, r, \eta)\). Beyond the simple Fabry–Perot interference filters treated here, there is a wide variety of more sophisticated thin-film designs that incorporate two or more half-wavelength resonant layers (“spacers”) that have been developed for optical applications. For example, the superlattice filter described in Ref. 5 contains two half-wavelength layers.

A wide variety of semiconductor superlattice electron wave filters are possible by analogy to thin-film optical filters. These include low-pass filters, high-pass filters, notch filters (narrow band and wide band), bandpass filters (narrow band and wide band), impedance transformers (antireflection coatings), and high reflectance surfaces (dielectric mirrors). In addition to being incorporated in semiconductor devices such as ballistic transistors, semiconductor superlattice filters could be used to control bremsstrahlung electron beams in applications such as electron spectroscopy, electron diffraction analysis of materials, and electron-beam lithography. Active devices that have electrically tunable characteristics are also possible.

### ACKNOWLEDGMENTS

This research was sponsored in part by a grant from the Joint Services Electronics Program under Contract No DAAL-03-87-K-0059. One of us (K.F.B.) was supported in...
part by a Presidential Young Investigator award from the National Science Foundation.

Semiconductor biased superlattice tunable electron interference filter/emitter

E. N. Glytsis, T. K. Gaylord, and K. F. Brennan
School of Electrical Engineering and Microelectronics Research Center, Georgia Institute of Technology, Atlanta, Georgia 30332

(Received 13 January 1989, accepted for publication 25 March 1989)

It is shown that a voltage-biased semiconductor superlattice structure can serve simultaneously as a tunable electron-wave interference filter and electron emitter. A systematic design procedure for selecting the quantum well and barrier widths to be alternately high and low electron refractive indices and a quarter (or a half) of an electron wavelength in thickness is developed. A practical narrow-band filter/emitter consisting of layers of Ga\(_{1-x}\)Al\(_x\)As and designed to emit 0.20-eV electrons is presented and analyzed. Such a structure would serve well as a tunable hot-electron emitter in ballistic transistors, and in future guided electron-wave integrated circuits.

Molecular-beam epitaxy and metalorganic chemical vapor deposition have allowed semiconductor multilayered superlattices to be grown with precise monolayer compositional control. Refinement of these methods has produced devices in which ballistic (collisionless) electron transport has been observed. Ballistic electrons are quantum mechanical deBroglie waves, and thus they can be refracted, reflected, diffracted, guided, and interfered in a manner analogous to electromagnetic waves. Quantum interference effects have been observed experimentally in single-well double-barrier GaAs/Ga\(_{1-x}\)Al\(_x\)As and In\(_{1-x}\)Ga\(_x\)As/In\(_{1-y}\)Al\(_y\)As potential barrier structures.

Phase effects such as path differences and wave interference for electron waves may be described using the wave-vector magnitude \(k = \frac{\sqrt{2m^*(E - V)}}{\hbar}\), where \(m^*\) is the electron effective mass, \(E\) is the total electron energy, \(V\) is the electron potential energy, and \(\hbar\) is Planck's constant divided by 2\(\pi\). Thus, the electron wave phase refractive index \(n_e\) (phase) is proportional to the square root of the product of the effective mass and the kinetic energy. Amplitude effects
such as transmissivity and reflectivity may be described in terms of the electron wave amplitude refractive index $n_{\text{e}}$ (amplitude) which is proportional to the square root of the ratio of the kinetic energy to the effective mass. Using these electron wave refractive indices, the characteristics of an unbiased or biased (this work) many boundary semiconductor superlattice system can be found. The simplest type of narrow bandpass interference filter consists of a half-wavelength layer sandwiched between reflectors which are stacks of high index (designated $H$) and low index (designated $L$) quarter-wavelength layers. The full width at half maximum (FWHM) of the bandpass in this type of filter is reduced by increasing the reflectivity at the boundaries between the layers. This is accomplished by increasing the ratio of the high index to the low index of the bounding materials. In this work, an electron emitter that incorporates a narrow-band interference filter is designed.

The electron potential energy of a quantum-well superlattice emitter/filter with a voltage bias applied is shown in Fig. 1. It consists of $M$ layers surrounded by bulk semiconductor material. When the design value of voltage is applied, electrons in a narrow spectral band around the pass energy $E_p$ traverse the device and are emitted with an output kinetic energy of $(KE)_{\text{out}}$. The $j$th quantum well or barrier has a thickness $d_j$, and at zero bias a potential energy of $V_j$. The surrounding regions are taken to have the same zero-bias potential energy of $V_0$. The layers have alternately low potential energy (high electron refractive index) and high potential energy (low refractive index). When the design value of bias potential energy $V_{\text{bias}}$ (product of electronic charge and design voltage) is applied, each reflector layer is exactly a quarter of an electron wavelength (as measured in that layer) in thickness, and the resonant layer is exactly one half of an electron wavelength. The given material system is taken to form a continuous set of alloys of the type Ga$_1-x$Al$_x$As. The range of usable compositions is $0<x<x_{\text{max}}$ (for example, due to a possible transition at $x_{\text{max}}$ from a direct to an indirect energy gap as in the case of Ga$_{1-x}$Al$_x$As). The materials surrounding the filter have compositions $x_0$. The electron potential energy is given by $V_j = V_{\text{bias}} + V_j$, where $V_j$ is the change in the energy of the conduction-band edge and $A$ is a constant. Thus, the corresponding range of potential energies is $0<V_j<V_{\text{max}}$. In addition, for an electron emitter/filter to be realizable, the thickness $d_j$ of each layer must be an integer multiple $p_j$ of the monolayer thickness $r$. For the $j$th layer to be a quarter of an electron wavelength in thickness at the pass energy $E_p$, the phase difference between the input boundary at $z_{j-1}$ and the output boundary at $z_j$ must be an odd multiple of $\pi/2$. That is,

$$\int_{z_{j-1}}^{z_j} k_j \, dz = \int_{z_{j-1}}^{z_j} (1/\hbar) \{2m_{\text{e}}^* \times [E_p - V_j(z)] \}^{1/2} \, dz = (2q_j - 1)(\pi/2),$$

where the potential energy in the $j$th layer with bias applied is given by $V_j(z) = V_{\text{bias}}(1 - z/L) + V_j$, $L$ is the total length of the superlattice, and $q_j$ is a positive integer. The pass energy may be expressed as $E_p = E_{\text{bias}} + V_0 + (KE)_{\text{in}}$, where $(KE)_{\text{in}}$ is the pass kinetic energy in the leftmost (input) region. The electron effective mass is given by $m_{\text{e}}^* = (B + Cx_{\text{bias}})m_0$, where $B$ and $C$ are constants and $m_0$ is the free-electron mass. Using $V_j = A x_j$ and the integral identity $\int [(a + bx)^{1/2} \, dz = (2/3b)(a + bx)^{3/2}$, the quarter electron wavelength condition may be expressed as

$${2L \left[2m_{\text{e}}(B + Cx_{\text{bias}})\right]^{1/2} / 3V_{\text{bias}}} \cdot \left[1 - \frac{1}{(2q_j - 1)^{1/2}} \right] = (2q_j - 1)(\pi/2).$$

This equation will be solved for the composition $x_j$ in the $j$th layer following the procedure described below. For the half-wavelength resonant section in the center of the filter, the phase difference in Eqs. (1) and (2) should be $\pi$ rather than $\pi/2$. The monolayers of the device will be numbered with the index $i$. The rightmost monolayer in the $j$th region is $i_j$. The total number of monolayers in the emitter/filter is designated $i_M$. Therefore, $z_{j-1} = (i_{j-1} - i_M)L$, $z_j = (i_j - i_M)L$, the total thickness is $L = i_Mr$, and the thickness of the $j$th layer is $d_j = p_j r$, where $p_j = i_j - i_{j-1}$ is the number of monolayers in the $j$th region.

A systematic procedure for designing a biased superlattice emitter/filter for a given output kinetic energy $[(KE)_{\text{out}} = V_{\text{bias}} + (KE)_{\text{in}}]$ using Eq. (2) is now described. The material system of the superlattice is taken to be Ga$_{1-x}$Al$_x$As with the surrounding regions being the same and having $V_0 = V_{\text{max}} = 0.3479$ eV, corresponding to an aluminum composition of $x_0 = 0.45$. The characteristics of (unbiased) superlattice electron wave interference filters may be used as a starting point in estimating the number of regions $M$ and the number of monolayers $i_M$ to be used in the emitter/filter. Following the example presented in Ref. 5, a nine-layer filter ($M = 9$) will be designed. The estimated number of monolayers needed is $i_M = 72$ (from Ref. 5). The device is to emit electrons with a kinetic energy of $0.20$ eV $[(KE)_{\text{out}} = 0.20$ eV]. The input kinetic energy is taken to be $(KE)_{\text{in}} = 0.10$ eV, and thus the bias potential energy $V_{\text{bias}} = (KE)_{\text{out}} - (KE)_{\text{in}} = 0.10$ eV. The parameters $r$ and $d_j$ are set equal to unity. Starting with the first layer, $i_{j-1} = 0$, the value of the monolayer index $i_j$ is incremented $1, 2, 3, \ldots$, and Eq. (2) solved for $x_j$ for each value of $i_j$. Since the $j = 1$ layer is to be a high electron refractive index layer, the resulting positive real value of $x_j$ closest to zero is selected. The corre-
TABLE I. Design parameters of electron interference filters/emitters consisting of nine layers in a GaAs superlattice.

<table>
<thead>
<tr>
<th>Layer number (j)</th>
<th>Layer type</th>
<th>Starting monolayer number (i_{j-1})</th>
<th>Ending monolayer number (i_j)</th>
<th>Number of monolayers thick (p_j)</th>
<th>Aluminum composition (x_j)</th>
<th>Unbiased electron potential energy (V_{bias})</th>
<th>Normalized effective mass (m_j^*/m_0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>H</td>
<td>0</td>
<td>7</td>
<td>7</td>
<td>0.2222</td>
<td>0.1718</td>
<td>0.0854</td>
</tr>
<tr>
<td>2</td>
<td>L</td>
<td>7</td>
<td>16</td>
<td>9</td>
<td>0.4151</td>
<td>0.3209</td>
<td>0.1015</td>
</tr>
<tr>
<td>3</td>
<td>H</td>
<td>16</td>
<td>23</td>
<td>7</td>
<td>0.2663</td>
<td>0.2059</td>
<td>0.0891</td>
</tr>
<tr>
<td>4</td>
<td>L</td>
<td>23</td>
<td>32</td>
<td>9</td>
<td>0.4493</td>
<td>0.3473</td>
<td>0.1043</td>
</tr>
<tr>
<td>5</td>
<td>HH</td>
<td>32</td>
<td>44</td>
<td>12</td>
<td>0.0639</td>
<td>0.0494</td>
<td>0.0723</td>
</tr>
<tr>
<td>6</td>
<td>L</td>
<td>44</td>
<td>52</td>
<td>8</td>
<td>0.4364</td>
<td>0.3374</td>
<td>0.1032</td>
</tr>
<tr>
<td>7</td>
<td>H</td>
<td>52</td>
<td>58</td>
<td>8</td>
<td>0.1442</td>
<td>0.1115</td>
<td>0.0790</td>
</tr>
<tr>
<td>8</td>
<td>L</td>
<td>58</td>
<td>65</td>
<td>7</td>
<td>0.3748</td>
<td>0.2898</td>
<td>0.0981</td>
</tr>
<tr>
<td>9</td>
<td>H</td>
<td>65</td>
<td>71</td>
<td>6</td>
<td>0.1951</td>
<td>0.1508</td>
<td>0.0832</td>
</tr>
</tbody>
</table>

FIG. 2. Transmission characteristic of a nine-layer Ga_{0.55}Al_{0.45}As superlattice interference filter/emitter as a function of the output kinetic energy for V_{bias} = 0.05, 0.10, and 0.15 eV. The tuning with bias is apparent.

K_1 = [(2m_j^*/\hbar^2)^{1/3}(E - V_0 - V_{bias})/M_j^*]^{1/2}(L/V_{bias})^{1/3},
K_2 = [(2m_j^*/\hbar^2)^{1/3}(E - V_0)/M_j^*]^{1/2}(L/V_{bias})^{1/3},
and ' indicates a derivative. These equations can be solved directly for the amplitude transmittance \psi_{LM+1} and the amplitude reflectance \psi_{LO}. The electron current transmittance is T = |\psi_{LM+1}|^2, where R = [(E - V_0)/M_j^*]^{1/2}/[(E - V_0 - V_{bias})/M_j^*]^{1/2}, and this is shown in Fig. 2 for the biased superlattice design described in Table I. At the design bias of V_{bias} = 0.10 eV, the device emits \psi_{LM+1} electrons into the output region. The full width at half maximum (FWHM) is 30.7 meV or 15.4% of the center energy. Furthermore, the output kinetic energy is tunable by changing the bias potential energy. Since the HH resonant layer is at the center (as measured in electron optical path length) of the device, the change in the output kinetic energy \Delta(KE)_{out} is one half of the change in the bias potential energy \Delta V_{bias}. Thus, as shown in Fig. 2, for \Delta V_{bias} = ±50 meV, the output kinetic energy is changed by \Delta(KE)_{out} = ±25 meV. Furthermore, the FWHM of the electron current transmittance can be decreased as required by incorporating additional quarter-wavelength thick superlattices.
lattice layers. The transmittance of the corresponding 13-layer filter/emitter is shown in Fig. 3. The overall characteristics are the same as those of the nine-layer filter/emitter of Fig. 2, except that the FWHM is reduced to 13.4 meV (6.7% of center energy).

The procedures for designing and analyzing narrow-band semiconductor superlattice interference filter/emitters have been presented in this communication. Such structures may be used as hot-electron emitters in ballistic transistors, electroluminescent devices, and photodetectors. They could also be used as tunable electron sources in the emerging new area of guided electron wave integrated circuits. Such tunable narrow-band emitters could provide numerous simultaneously energy multiplexed data channels in the same guided electron wave integrated circuit in a manner analogous to that used in electromagnetic optical integrated circuits. However, the corresponding electron wavelengths would be typically 100 times smaller.

This research was sponsored in part by a grant from the Joint Services Electronics Program under Contract No. DAAL-03-87-K-0059. One of us (K.F.B.) was supported in part by a Presidential Young Investigator Award from the National Science Foundation.
Semiconductor electron-wave slab waveguides

T. K. Gaylord, E. N. Glytsis, and K. F. Brennan

School of Electrical Engineering and Microelectronics Research Center, Georgia Institute of Technology, Atlanta, Georgia 30332

(Received 14 December 1988; accepted for publication 11 April 1989)

A one-dimensional semiconductor quantum well can act as a waveguide for ballistic electrons owing to the quantum mechanical wave behavior of these electrons. The allowed modes in an asymmetric quantum well slab waveguide are quantified. Electron waveguiding can occur for energies above one or both of the potential barriers. Due to dispersion, each electron waveguide mode has an upper-energy cutoff as well as a lower-energy cutoff. An example waveguide consisting of Ga\textsubscript{0.45}Al\textsubscript{0.55}As (substrate), GaAs (film), and Ga\textsubscript{0.70}Al\textsubscript{0.30}As (cover) is treated. This structure is a single-mode electron waveguide for [100] GaAs thicknesses of from 6 through 31 monolayers.

Semiconductor growth technologies such as molecular-beam epitaxy have enabled structures to be fabricated with precise monolayer compositional control. Improvements in these techniques have produced devices in which ballistic (collisionless) electron transport has been observed. Ballistic electrons are quantum mechanical de Broglie waves and thus they can be refracted, reflected, diffracted, interfered, and guided in a manner analogous to electromagnetic waves.

An asymmetric potential energy well is shown in Fig. 1. Using waveguide terminology, the three regions are denoted substrate (s), film (f), and cover (c), and the direction perpendicular to the waveguide surfaces is \( x_w \). The electron potential energy at the bottom of the quantum well (film) is \( E_f \). The potential energy barrier heights associated with the substrate and cover are \( V_s \) and \( V_c \), respectively. This quantum well can act as an asymmetric slab waveguide. The angle of incidence of the two plane-wave components that constitute the guided wave is the zigzag angle \( \theta \).

The magnitude of the electron wave vector in any of the three regions is \( k_i = [2m_i^* (E - V_i)]^{1/2}/\hbar \), where \( i = s, f, c \) and \( m_i^* \) is the electron effective mass, \( E \) is the total electron energy, and \( V_i \) is the electron potential energy. Thus, the electron-wave phase refractive index \( n_i \) (phase) is proportional to the square root of the product of the effective mass and the kinetic energy. The onset of total internal reflection occurs when \( \theta \) is equal to the critical angle \( \theta_c \) given by:

\[
\theta_c = \sin^{-1} \left[ m_s^* (E - V_s)/2m_f^* (E - V_f) \right]^{1/2},
\]

where \( i = s \) for the cover-film critical angle, \( i = f \) for the substrate-film critical angle, and \( E_f = (m_f^* V_f - m_s^* V_s)/(m_f^* - m_s^*) \). For an electron wave incident at an angle greater than \( \theta_c \), the wave is totally internally reflected for an infinitely thick barrier. At steady state, all of the electron current is reflected back into the film. The electron-wave function decays exponentially into the cover. If the kinetic energy \( E - V_f < 0 \), then total internal reflection occurs for any angle of incidence including normal incidence. For a guided mode, the zigzag angle \( \theta \) shown in Fig. 1 must be greater than both \( \theta_c \) and \( \theta' \). That is, \( \theta > \max(\theta_c, \theta') \) is a necessary condition for a guided mode.

A representation of the electron propagation constant versus total electron energy for the Ga\textsubscript{0.45}Al\textsubscript{0.55}As/GaAs/Ga\textsubscript{0.70}Al\textsubscript{0.30}As example is given in Fig. 2. For an infinite medium, the electron propagation constant is \( \beta = [2m_i^* (E - V_i)]^{1/2}/\hbar \), where \( i = c, f, s \). These three propagation constants are plotted in Fig. 2. For a given energy \( E \), the propagation constant of a guided mode can be no longer than \( \beta_f \). The region to the left of this curve corresponds to evanescent (nonphysical) modes. Thus the allowed guided modes must lie to the right of this curve in Fig. 2. However, in order to satisfy the \( \theta > \max(\theta_c, \theta') \) condition, the guided modes must also lie left of \( \beta_f \). An electron-guided wave mode can become cutoff by decreasing the electron energy to the lower-energy cutoff. The zigzag angle decreases with decreasing energy and this cutoff occurs when \( \theta = 0 \). The propagation constants \( \beta_v \) of the \( v \)th guided mode \( M_v \), where \( v = 0, 1, 2, ..., \) is given by:

\[
\beta_v = [2m_i^* (E - V_f)/\hbar^2]^{1/2} \sin \theta_v,
\]

where \( \theta_v \) is the waveguide mode angle given by:

\[
\theta_v = \sin^{-1} \left[ m_s^* (E - V_s)/2m_f^* (E - V_f) \right]^{1/2},
\]

where \( i = c \) for the cover-film critical angle, \( i = s \) for the substrate-film critical angle, and \( E_f = (m_f^* V_f - m_s^* V_s)/(m_f^* - m_s^*) \). For an electron wave incident at an angle greater than \( \theta_c \), the wave is totally internally reflected for an infinitely thick barrier. At steady state, all of the electron current is reflected back into the film. The electron-wave function decays exponentially into the cover. If the kinetic energy \( E - V_f < 0 \), then total internal reflection occurs for any angle of incidence including normal incidence. For a guided mode, the zigzag angle \( \theta \) shown in Fig. 1 must be greater than both \( \theta_c \) and \( \theta' \). That is, \( \theta > \max(\theta_c, \theta') \) is a necessary condition for a guided mode.

A representation of the electron propagation constant versus total electron energy for the Ga\textsubscript{0.45}Al\textsubscript{0.55}As/GaAs/Ga\textsubscript{0.70}Al\textsubscript{0.30}As example is given in Fig. 2. For an infinite medium, the electron propagation constant is \( \beta = [2m_i^* (E - V_i)]^{1/2}/\hbar \), where \( i = c, f, s \). These three propagation constants are plotted in Fig. 2. For a given energy \( E \), the propagation constant of a guided mode can be no longer than \( \beta_f \). The region to the left of this curve corresponds to evanescent (nonphysical) modes. Thus the allowed guided modes must lie to the right of this curve in Fig. 2. However, in order to satisfy the \( \theta > \max(\theta_c, \theta') \) condition, the guided modes must also lie left of \( \beta_f \). An electron-guided wave mode can become cutoff by decreasing the electron energy to the lower-energy cutoff. The zigzag angle decreases with decreasing energy and this cutoff occurs when \( \theta = 0 \). The propagation constants \( \beta_v \) of the \( v \)th guided mode \( M_v \), where \( v = 0, 1, 2, ... \), is given by:

\[
\beta_v = [2m_i^* (E - V_f)/\hbar^2]^{1/2} \sin \theta_v,
\]
FIG. 2. Electron-guided-mode propagation constant as a function of total electron energy showing the regions of evanescent modes, guided modes, substrate modes, and radiation modes for the example quantum well waveguide composed of GaAs/AlxAs (substrate), GaAs (film), and Ga0.7Al0.3As (cover). The mode dispersion curves for the fundamental mode $M_0$ are shown for film thicknesses of 10, 20, and 30 monolayers.

and so $\beta_c = 0$ at the lower-energy cutoff. The wave function is sinusoidal in the film and exponentially decaying in the substrate and cover. In this sense, the lower-energy cutoff is like the cutoff in an electromagnetic hollow metallic waveguide with finite conductivity walls.

As the electron energy of a guided mode is increased, an upper-energy cutoff will also occur. The upper-energy cutoff can be of three types: (1) cutoff to a substrate mode which is like the cutoff in an electromagnetic asymmetric dielectric waveguide with the substrate index higher than the cover index; (2) cutoff to a radiation mode which is like the cutoff in an electromagnetic symmetric dielectric waveguide (equal substrate and cover indices); and (3) cutoff to a cover mode which is like the cutoff in an electromagnetic asymmetric dielectric waveguide with the cover index higher than the substrate index. The type of upper-energy cutoff that occurs depends on the material parameters. For the example case depicted in Fig. 2, the upper-energy cutoff will be to substrate modes since $\beta_s$ occurs at lower energy, in general, than does $\beta_c$.

For a two-dimensional ($x_w, z_w$) quantum-well-guided electron wave, the wave function has a sinusoidal dependence in the $z_w$ direction and can be expressed as

$$\psi_c(x_w, z_w) = \psi_c(x_w) \exp(j \beta_z z_w),$$

where $\beta_z$ is the guided mode propagation constant. Using this, the Schrödinger time-independent wave equation becomes

$$d^2 \psi_c^2(x_w)/dx_w^2 + \left\{\left(2m^*/\hbar^2\right)[E - V(x_w)] - \beta_z^2\right\} \psi_c(x_w) = 0.$$

Thus for a guided mode, the wave-function amplitude in the substrate may be expressed as

$$\psi_s(x_w) = A_s \exp(\gamma_s x_w),$$

and in the film as

$$\psi_f(x_w) = A_f \exp(\gamma_f x_w) + A_f \exp(-\gamma_f x_w),$$

and in the cover as

$$\psi_c(x_w) = A_c \exp[-\gamma_c (x_w - d)].$$

where

$$\gamma_s^2 = \beta_z^2 - \left[\left(2m^*/\hbar^2\right)(E - V_f)\right],$$

$$\kappa_f^2 = \left[\left(2m_f^*/\hbar^2\right)(E - V_f)\right] - \beta_z^2,$$

and

$$\gamma_c^2 = \beta_c^2 - \left[\left(2m_c^*/\hbar^2\right)(E - V_c)\right].$$

By applying the boundary conditions that $\psi$ and $(1/m^*) (d\psi/dx)$ must be continuous across the cover-film and substrate-film boundaries, the dispersion equation is found to be

$$\kappa_f^2 - \tan^{-1}\left[\left(\gamma_s^2/m_s^2\right)/\left(\gamma_f^2/m_f^2\right)\right] = \nu \tau,$$

For $V_f = 0$ and normal incidence ($\beta_z = 0$), this dispersion equation predicts, in the limit as $V_c$ and $V_s$ approach infinity, that

$$E = E_s = (\nu + 1)\hbar^2 \tau^2/2m_f^* d^2.$$
in agreement with the well-known one-dimensional infinite potential well results.

Cutoff for the guided \( M_x \) modes can occur as the electron energy is decreased and the guided wave propagation constant goes to zero. The electron energy at which the lower-energy cutoff occurs is designated \( E_{\text{Loo}} \). The condition for this type of cutoff is obtained by substituting \( \beta_y = 0 \) into the dispersion equation (3). The cutoff condition is

\[
\left[ 2m_y (E_{\text{Loo}} - V_f) \right]^{1/2} / \hbar - \tan^{-1} \left[ m_y (V_s - E_{\text{Loo}}) / m_y (E_{\text{Loo}} - V_f) \right]^{1/2} = \tan^{-1} \left[ m_y (V_s - E_{\text{Loo}}) / m_y (E_{\text{Loo}} - V_f) \right]^{1/2} = \pi. 
\]

For given substrate, film, and cover materials, thickness, and integer mode number \( v \), this transcendental equation may be solved for the lower cutoff energy \( E_{\text{Loo}} \) corresponding to that mode. Cutoff for the guided \( M_y \) modes can also occur as the electron energy is increased and total internal reflection no longer occurs at the substrate-film boundary. The electron energy at which the upper-energy cutoff occurs is designated \( E_{\text{Uoo}} \). The condition for the occurrence of this type of cutoff is \( \gamma_z = 0 \), and thus the mode “leaks” into the substrate. This is analogous to the cutoff of an electromagnetic guided mode in an asymmetric dielectric slab waveguide. This type of cutoff occurs when the zigzag angle is equal to the substrate-film critical angle. Substituting \( \gamma_z = 0 \) into the dispersion equation (3), the condition for cutoff to substrate modes is thus

\[
\left[ 2 \left[ m_y V_s - m_y V_f - (m_y - m_y) E_{\text{Loo}} \right] \right]^{1/2} / \hbar \cdot \tan^{-1} \left[ m_y (V_s - m_y V_f - (m_y - m_y) E_{\text{Loo}}) / m_y (V_s - m_y V_f) \right]^{1/2} = \pi. 
\]

For a given set of material parameters, as the waveguide thickness is increased, a guided mode \( M_x \) first starts to propagate at an energy \( E = V_s \). This corresponds to the highest possible value of the cutoff energy for the lower-energy type cutoff as well as the lowest possible value of the cutoff energy for the higher-energy type cutoff. Substituting \( E = V_s \) into Eq. (3) gives the thickness at which the \( M_x \) mode starts to propagate as

\[
d = \left\{ \hbar / [2m_y (V_s - V_f)] \right\}^{1/2} \times \tan^{-1} \left[ m_y (V_s - V_f) / m_y (V_s - V_f) \right]^{1/2} + \nabla \eta. 
\]

For the quantum well waveguide consisting of \( \text{GaAs}_{0.30} \text{Al}_{0.70} \text{As} \) (substrate), \( \text{GaAs} \) (film), and \( \text{GaAs}_{0.70} \text{Al}_{0.30} \text{As} \) (cover), the potential energies are

<table>
<thead>
<tr>
<th>Substrate</th>
<th>GaAs (film)</th>
<th>Cover</th>
</tr>
</thead>
<tbody>
<tr>
<td>( E_{\text{Loo}} )</td>
<td>( E_{\text{Uoo}} )</td>
<td>( E_{\text{Uoo}} )</td>
</tr>
<tr>
<td>( 0.115 ) eV</td>
<td>( 0.115 ) eV</td>
<td>( 0.115 ) eV</td>
</tr>
<tr>
<td>( V_s = 0.231 ) 942 eV</td>
<td>( V_f = 0.231 ) 942 eV</td>
<td>( V_f = 0.231 ) 942 eV</td>
</tr>
<tr>
<td>( V_f = 0.231 ) 942 eV</td>
<td>( V_f = 0.231 ) 942 eV</td>
<td>( V_f = 0.231 ) 942 eV</td>
</tr>
<tr>
<td>( V_f = 0.231 ) 942 eV</td>
<td>( V_f = 0.231 ) 942 eV</td>
<td>( V_f = 0.231 ) 942 eV</td>
</tr>
<tr>
<td>( V_f = 0.231 ) 942 eV</td>
<td>( V_f = 0.231 ) 942 eV</td>
<td>( V_f = 0.231 ) 942 eV</td>
</tr>
<tr>
<td>( V_f = 0.231 ) 942 eV</td>
<td>( V_f = 0.231 ) 942 eV</td>
<td>( V_f = 0.231 ) 942 eV</td>
</tr>
<tr>
<td>( V_f = 0.231 ) 942 eV</td>
<td>( V_f = 0.231 ) 942 eV</td>
<td>( V_f = 0.231 ) 942 eV</td>
</tr>
<tr>
<td>( V_f = 0.231 ) 942 eV</td>
<td>( V_f = 0.231 ) 942 eV</td>
<td>( V_f = 0.231 ) 942 eV</td>
</tr>
</tbody>
</table>

The electron effective masses are \( m_y = 0.079 \) 45 \( m_0 \), \( m_y = 0.067 \) \( m_0 \), and \( m_y = 0.0919 \) \( m_0 \). Growth is taken to be along the [100] direction and thus each monolayer corre-
Semiconductor quantum wells as electron wave slab waveguides

T. K. Gaylord, E. N. Glytsis, and K. F. Brennan

School of Electrical Engineering and Microelectronics Research Center, Georgia Institute of Technology, Atlanta, Georgia 30332

(Received 6 December 1988; accepted for publication 21 April 1989)

A quantum well in a semiconductor can act as a slab waveguide for electron waves in a manner analogous to the way a layered dielectric can act as a slab waveguide for electromagnetic waves (e.g., as commonly employed in integrated optics). In this work, the case of a general electron asymmetric slab waveguide (a quantum well comprised of three materials each with a different potential energy and a different effective mass) is analyzed and the conditions for electron waveguiding are quantified. Electron waveguide modes exist for electron energies in the well and for electron energies above one or both of the potential energy barriers. Furthermore, due to dispersion, each electron waveguide mode has an upper-energy cutoff as well as a lower-energy cutoff. This is in contrast to electromagnetic guided modes which typically have only lower-energy (low-frequency) cutoffs. At the upper-energy cutoff the electron wave is refracted into the substrate and/or cover. An example quantum well waveguide consisting of Ga$_{0.80}$Al$_{0.20}$As (substrate), GaAs (film), Ga$_{0.55}$Al$_{0.45}$As (cover) is analyzed. This structure is a single-mode electron waveguide for GaAs thicknesses of from 5 (1.413 nm) to 26 monolayers (7.349 nm).

I. BACKGROUND AND MOTIVATION

Interest in quantum well devices has increased in recent years owing to progress in semiconductor growth technologies such as molecular beam epitaxy (MBE) and metal organic chemical vapor deposition (MOCVD) that enable the fabrication of structures with precise monolayer compositional control. Furthermore, improvements in the crystalline quality of these grown materials have simultaneously produced devices in which ballistic electron transport has been observed. That is, conduction electrons move through the material without being scattered and behave like waves (quantum mechanical deBroglie waves). Even in the presence of elastic scattering, electrons retain their phase coherence and can thus refract, reflect, diffract, and interfere in a manner analogous to electromagnetic waves. Also, electron waves may be guided by structures having characteristic dimensions on the order of an electron wavelength in a manner analogous to electromagnetic waves in a dielectric waveguide. Furthermore, the analogies between electron waves in semiconductors and electromagnetic waves in dielectrics provide a basis for designing semiconductor quantum wave devices such as electron waveguides.

The behavior of an electron in a single quantum well has been understood since the beginning of quantum mechanics. In the usual situation, the electron is normally incident upon the two potential energy barriers. This corresponds to a zigzag angle of $\theta = 0$ in the waveguide terminology used in subsequent sections. Furthermore, as shown in this paper, these solutions correspond to the lower-energy cutoff that occurs for guided modes of energies within the quantum well. For $\theta \neq 0$, a quantum well is shown to act like a waveguide. Electron waveguide modes exist for electron energies in the well and for electron energies above one or both of the potential energy barriers.

The purposes of this paper are as follows: (1) Present the conditions under which a quantum well acts like an electron waveguide; (2) quantify and label as $M_0$, $M_1$, $M_2$, ... the guided modes; (3) describe the lower-energy cutoff (at which the propagation constant $\beta_x = 0$) in which the electron waveguide is analogous to an electromagnetic hollow metallic waveguide with finite conductivity walls; (4) describe the upper-energy cutoff which may be a transition from a guided mode to a substrate "mode" or a radiation "mode" or a cover "mode"; (5) develop the dispersion relation governing electron guided modes; (6) derive the cutoff...
II. ELECTRON WAVEGUIDING IN QUANTUM WELLS

A. Configuration

An asymmetric potential energy well and its depiction as a slab waveguide are shown in Fig. 1. Using standard waveguide terminology, the three regions are denoted substrate (s), film (f), and cover (c), and the direction perpendicular to the waveguide surfaces is \( x_e \). The electron potential energy at the bottom of the quantum well (film) is \( V_f \). The potential energy barrier heights associated with the substrate and cover are \( V_s \) and \( V_c \), respectively. Given a material system of the type \( F_{1-x} Ga_x As \), the compositions of the substrate, film, and cover are \( x_s \), \( x_f \), and \( x_c \), respectively. This quantum well can act as an asymmetric slab waveguide. The waveguide coordinate system is also shown in Fig. 1. The direction of guided mode propagation is \( d \). The thickness of the waveguide is \( d \). The angle of incidence of the two plane-wave components that constitute the guided wave is the zig-zag angle \( \theta \).

B. Critical angles

The magnitude of the electron wave vector in any of the three regions may be expressed as

\[
k = \frac{2m^*(E - V)}{\hbar^2},
\]

where \( m^* \) is the electron effective mass, \( E \) is the total electron energy, and \( V \) is the electron potential energy. Thus, the electron wave phase refractive index, \( n_p \) (phase), is proportional to the square root of the product of the effective mass and the kinetic energy. The phase of the transmitted and reflected electron waves along the boundary must be identical to that of the incident electron wave. This phase-matching requirement means that the component of the wave vector parallel to the boundary must be the same before and after reflection and refraction. This gives rise to the equivalent of Snell's law for electrons. For incidence upon the film-cover interface, as shown in Fig. 1, it may be written as

\[
\sin \theta_s = \sin \theta_f \left( \frac{m^*(E - V_f)}{m^*(E - V_s)} \right)^{1/2}.
\]

The onset of total internal reflection occurs when \( \theta = 90^\circ \). This happens when the angle of incidence is equal to the critical angle \( \theta' \). Thus from Eq. (2), the cover-film critical angle is

\[
\theta'_c = \sin^{-1} \left( \frac{m^*(E - V_f)}{m^*(E - V_c)} \right)^{1/2},
\]

where \( E_{cf} \) is defined in Sec. II C. For an electron wave incident at an angle greater than \( \theta'_c \), the wave is totally internally reflected for an infinitely thick barrier. At steady state, all of the electron current is reflected back into the film. The electron wave function decays exponentially into the cover. If the kinetic energy \( E - V < 0 \), then total internal reflection occurs for any angle of incidence including normal incidence. This is in contrast to the electromagnetic case where total internal reflection can never occur at normal incidence due to the nonzero value of the refractive index. Similarly, the substrate-film critical angle is

\[
\theta'_s = \sin^{-1} \left( \frac{m^*(E - V_s)}{m^*(E - V_f)} \right)^{1/2},
\]

where \( E_{sf} \) is defined in Sec. II C. For a guided mode, the zig-zag angle \( \theta \) shown in Fig. 1 must be greater than both \( \theta'_c \) and \( \theta'_s \). That is,

\[
\theta > \max(\theta'_c, \theta'_s)
\]

are necessary conditions for a guided mode. For the example quantum well waveguide of \( Ga_{0.8}Al_{0.2}As \) (substrate), GaAs (film), \( Ga_{0.25}Al_{0.75}As \) (cover) treated in this paper, the critical angles \( \theta'_c \) and \( \theta'_s \) are shown in Fig. 2 on a plot of zig-zag angle versus total energy.

C. Types of guided modes and unguided "modes"

The range of zig-zag angles and total energies in Fig. 2 that satisfy (5) corresponds to the region of allowed guided modes. As the energy increases in a guided mode, the zig-zag angle \( \theta \) also increases. When \( \theta \) reaches \( \theta'_s \), the guided wave starts to refract into the substrate (rather than exponentially decaying). For energies greater than this energy, the electron wave is propagating in the substrate as well as in the film. This is called a substrate "mode." Quotation marks are used to emphasize that this is not a true guided mode. As the energy is further increased, \( \theta \) reaches \( \theta'_c \), and the electron wave starts to refract into the cover as well as into the substrate. The electron wave is now a propagating wave in all three regions, and this is called a radiation "mode." For a different set of material parameters, as the electron energy is increased, it is possible for the electron wave to be refracted into the cover. This situation is similarly called a cover "mode."

FIG. 2. The critical angles of \( \theta'_c \) and \( \theta'_s \) on a plot of zig-zag angle vs total energy for the example quantum well waveguide composed of \( Ga_{0.8}Al_{0.2}As \) (substrate), GaAs (film), \( Ga_{0.25}Al_{0.75}As \) (cover). The regions of guided modes, substrate modes, and radiation modes are shown.
D. Types of cutoff

An electron guided wave mode can become cutoff by decreasing the electron energy. This will be called the lower-energy cutoff. The zig-zag angle decreases with decreasing energy and this cutoff occurs when \( \theta = 0 \). The guided wave may be decomposed into two plane-wave components each having an angle of incidence of \( \theta \) with respect to the waveguide walls. The propagation constant, \( \beta_v \), of the vth guided mode, where \( v \) is an integer is

\[
\beta_v = 2\pi/\Lambda_v, \tag{6}
\]

where \( \Lambda_v \) is the period of the interference pattern produced by the two plane waves that constitute the vth guided mode as measured in the guided mode direction of propagation \( (z_v) \). The propagation constant \( \beta_v \) is the same as the component of the wave vector parallel to the boundary. In conventional solid state notation, \( \beta_v = k_\parallel \). Therefore it is given by

\[
\beta_v = [2m^*_v(E - V_f)/\hbar]^1/2 \sin \theta, \tag{7}
\]

and so \( \beta_v = 0 \) at the lower-energy cutoff. The wave function is sinusoidal in the film and exponentially decaying in the substrate and cover. In this sense, the lower-energy cutoff is like the cutoff in an electromagnetic hollow metallic waveguide with finite conductivity walls.

A representation of the electron propagation constant versus total electron energy for the Ga\(_{0.80}\)Al\(_{0.20}\)As/GaAs/ Ga\(_{0.55}\)Al\(_{0.45}\)As example is given in Fig. 3. For an infinite medium having the properties of the substrate, the electron propagation constant is

\[
\beta_v = [2m^*_v(E - V_s)/\hbar]^1/2, \tag{8}
\]

Similarly, for an infinite medium having the properties of the film, it is

\[
\beta_f = [2m^*_f(E - V_f)/\hbar]^1/2, \tag{9}
\]

and for an infinite cover medium, it is

\[
\beta_c = [2m^*_c(E - V_c)/\hbar]^1/2. \tag{10}
\]

These three propagation constant relationships are plotted in Fig. 3. For a given energy \( E \), the propagation constant of a guided mode can be no larger than that given by Eq. (9). The region to the left of this curve corresponds to evanescent (nonphysical) modes. Thus the allowed guided modes must lie to the right of this curve in Fig. 3.

As the electron energy of a guided mode is increased, an upper-energy cutoff will occur. The upper-energy cutoff can be of three types: (1) cutoff to a substrate "mode" which is like the cutoff in an electromagnetic asymmetric dielectric waveguide with the substrate index higher than the cover index, (2) cutoff to a radiation "mode" which is like the cutoff in an electromagnetic symmetric dielectric waveguide (equal substrate and cover indices), (3) cutoff to a cover "mode" which is like the cutoff in an electromagnetic symmetric dielectric waveguide with the cover index higher than the substrate index. The type of upper-energy cutoff that occurs depends on the material parameters. For the example case depicted in Fig. 4, the upper-energy cutoff will be to substrate "modes" since \( \beta_v \) occurs at lower energy, in general, than does \( \beta_c \).

The intersection of \( \beta_f \) and \( \beta_v \) occurs at an energy \( E_{cf} \) which is given by

\[
E_{cf} = (m^*_vV_f - m^*_fV_f)/(m^*_c - m^*_f). \tag{11}
\]

At this energy, the electron wave phase refractive indices for the film and the substrate are equal. When this energy is reached, the waveguide can no longer guide an electron wave even if it is at grazing incidence along the walls of the waveguide. Thus this energy is equivalent to the substrate-film critical angle \( \theta'_{cf} = 90' \). Similarly, the intersection of \( \beta_f \) and \( \beta_c \) occurs at an energy \( E_{cf} \) and this energy is given by

\[
E_{cf} = (m^*_cV_c - m^*_fV_f)/(m^*_c - m^*_f). \tag{12}
\]

At this energy, the electron wave phase refractive indices for the film and the cover are equal. This energy is equivalent to the cover-film critical angle \( \theta'_{cf} = 90' \). Similarly, the intersection of \( \beta_v \) and \( \beta_c \) occurs at an energy \( E_{cv} \) and this energy is given by

\[
E_{cv} = (m^*_vV_v - m^*_cV_v)/(m^*_c - m^*_v). \tag{13}
\]

At this energy, the electron wave phase refractive indices for the substrate and the cover are equal.

III. QUANTUM WELL ELECTRON SLAB WAVEGUIDES

A. Dispersion relation for guided modes

The analysis of a two-dimensional \((x_w, z_w)\) quantum well guided electron wave is similar to the conventional one-dimensional \((x_w)\) quantum well but with the addition of the guided mode dependence in the \( z_w \) direction and differing effective masses in the three regions. The wave function has a sinusoidal dependence in the \( z_w \) direction and can be expressed in the form

\[
\psi_v(x_w, z_w) = \psi_v(x_w) \exp (+j\beta_v z_w), \tag{14}
\]
where $\beta_v$ is the guided mode propagation constant. The Schrödinger time-independent wave equation becomes

$$\frac{d^2\psi_v(x_w)}{dx_w^2} + \left(\frac{2m^*}{\hbar^2} [E_v - V(x_w)] - \beta_v^2\right)\psi_v(x_w) = 0.$$  \hspace{1cm} (15)

For a guided mode, the wave function amplitude in the substrate may now be expressed as $\psi_v(x_w) = A_v \exp(\gamma_v x_w)$, in the film as

$$\psi_v(x_w) = A_{f1} \exp(j\kappa x_w) + A_{f2} \exp(-j\kappa x_w),$$

and in the cover as

$$\psi_v(x_w) = A_v \exp(-\gamma_v (x_w - d)).$$

where

$$\gamma_v^2 = \beta_v^2 - \left[(2m^*/\hbar^2)(E_v - V_f)\right],$$

$$\kappa^2 = \left[(2m^*/\hbar^2)(E_v - V_f)\right] - \beta_v^2,$$

and

$$\gamma_v^2 = \beta_v^2 - \left[(2m^*/\hbar^2)(E_v - V_f)\right].$$

Applying the boundary condition that the wave function amplitude and $1/(m^*) (d\psi_v/dx)$ be continuous at the substrate film and at the cover-film boundaries gives a set of four linear homogeneous equations in the four unknowns ($A_v, A_{f1}, A_{f2}, A_c$). For nontrivial solutions, the determinant of coefficients must vanish. This gives

$$\tan(\kappa d) = \frac{(\gamma_v/m^*) [\gamma_v (m^*) + \gamma_v (m^*)]}{(k_f/m^*) - (\gamma_v (m^*) (\gamma_v (m^*))}.$$  \hspace{1cm} (16)

Using the identity

$$\tan(x + y) = \frac{(\tan x + \tan y)/(1 - \tan x \tan y)}{1}$$

and then taking the arctangent of both sides of this equation gives

$$\kappa d = \tan^{-1}\left(\frac{\gamma_v (m^*)}{\kappa_f/m^*}\right) - \tan^{-1}\left(\frac{\gamma_v (m^*)}{\kappa_f/m^*}\right) = \pi n.$$  \hspace{1cm} (17)

For $V_f = 0$ and normal incidence ($\beta_v = 0$), this dispersion equation predicts, in the limit as $V_c$ and $V_f$ approach infinity, that

$$E_v = (n + 1)^2\hbar^2/2m^* d^2$$

in agreement with the quantum mechanical one-dimensional infinite potential well results.

B. Lower-energy cutoff (cutoff at normal incidence)

Cutoff for the guided $M_v$ modes can occur as the electron energy is decreased and the guided wave propagation constant goes to zero. This can occur only for an electron energy below the lower barrier energy ($E < V_1$). When $\beta_v = 0$, the mode is no longer propagating. This is analogous to the cutoff of an electromagnetic guided mode in a hollow metallic waveguide with finite conductivity walls. The plane-wave components of the guided wave are reflecting back and forth at normal incidence to the waveguide boundaries. The electron energy at which the lower-energy cutoff occurs is designated $E_{Leo}$. The condition for this type of cutoff is obtained by substituting $\beta_v = 0$ into the dispersion Eq. (17). The cutoff condition is

$$[2m^* (E_{Leo} - V_f)]^{1/2}/\hbar - \tan^{-1}\left(\frac{m^* (V_c - E_{Leo})}{m^* (E_{Leo} - V_f)}\right)^{1/2} = \pi n.$$  \hspace{1cm} (18)

For given substrate, film, and cover materials, thickness, and integer mode number $n$, this transcendental equation may be solved for the $\beta_v = 0$ cutoff energy $E_{Leo}$ corresponding to that mode.

C. Upper-energy cutoff (cutoff to substrate “mode”)

Cutoff for the guided $M_v$ modes can occur as the electron energy is increased and total internal reflection no longer occurs at the substrate-film boundary. In electromagnetic waveguides there is typically only a lower-energy (lower-frequency) cutoff, and no higher-energy cutoff exists. In the present situation, however, as the electron energy is increased, the electron wave is refracted into the substrate. As the energy is increased through this cutoff transition, the electron wave function amplitude in the substrate changes from being evanescent (exponentially decaying) to propagating (sinusoidal). This can occur only for an electron energy above the lower barrier energy ($E > V_1$). The electron energy at which the upper-energy cutoff occurs is designated $E_{Leo}$. The condition for the occurrence of this type of cutoff is $\gamma_v = 0$, and thus the mode “leaks” into the substrate. This is analogous to the cutoff of an electromagnetic guided mode in an asymmetric dielectric slab waveguide. This type of cutoff occurs when the zig-zag angle is equal to the substrate-film critical angle. The condition for this type of cutoff is thus obtained by substituting $\gamma_v = 0$ into the dispersion Eq. (17). The condition for cutoff to substrate “modes” is thus

$$\{2 \left[ (m^* V_c - m^* V_f - (m^* - m^*) E_{Leo}) \right]^{1/2}/\hbar - \tan^{-1}\left(\frac{m^* V_c - m^* V_f - (m^* - m^*) E_{Leo}}{m^* V_c - m^* V_f - (m^* - m^*) E_{Leo}}\right)^{1/2} = \pi n.$$  \hspace{1cm} (19)

This transcendental equation may be solved for the cutoff to substrate “mode” energy $E_{Leo}$ corresponding to the $M_v$ mode.

D. Energy of first appearance of modes

For a given set of material parameters, as the waveguide thickness is increased, a guided mode $M_v$ first starts to propagate at an energy $E = V_1$. This corresponds to the highest possible value of the cutoff energy for the lower-energy type cutoff as well as the lowest possible value of the cutoff energy for the higher-energy type cutoff. For $E = V_1$, Eqs. (18) and (19) are the same. Namely,

$$\left(2m^* (V_c - V_f)\right)^{1/2}/\hbar - \tan^{-1}\left(\frac{m^* (V_c - V_f)}{m^* (V_c - V_f)}\right)^{1/2} = \pi n.$$  \hspace{1cm} (20)

From Eq. (20), the $M_v$ first starts propagating as the thickness is increased to the value
the plane-wave components of the guided wave are reflecting barrier energy. This can occur only for an electron energy below the lower cutoff, again occurs as the energy is decreased and the propagation constant goes to zero. The lower-energy cutoff again occurs as the energy is decreased and the guided mode becomes evanescent to propagating. This can occur only for an electron energy above the barrier energy.

Also from Eq. (20), the range of thicknesses that will produce a waveguide that supports only the lowest-order ($\nu = 0$) mode ($M_0$) is obtained and is

\[
\frac{\hbar}{2m_s^*(V_s - V_f)} \left[ \tan^{-1} \left( \frac{m_s^*(V_c - V_s)}{m_s^*(V_f - V_s)} \right) \right]^{1/2} + n \pi. \tag{21}
\]

As with electromagnetic asymmetric slab waveguides, there is a minimum thickness required for any modes to propagate.

E. Symmetric waveguides

If the waveguide is symmetric, $V_c = V_s$ and the substrate and cover dispersion curves coincide. In this situation the lower-energy cutoff again occurs as the energy is decreased and the propagation constant goes to zero ($\beta_s = 0$). This can occur only for an electron energy below the lower barrier energy ($E < V_c = V_s$). When $\beta_s = 0$, also $\theta = 0$ and the plane-wave components of the guided wave are reflecting back and forth at normal incidence to the waveguide boundaries.

For a symmetric waveguide, the upper-energy cutoff occurs as the energy is increased and the guided mode becomes a radiation "mode." At the cutoff energy, total internal reflection occurs neither at the substrate-film boundary nor at the cover-film boundary. The electron wave is then refracted both into the substrate and the cover. As the electron energy is increased through this cutoff transition, the electron wave function amplitude in the substrate and cover change from being evanescent to propagating. This can occur only for an electron energy above the barrier energy ($E > V_c = V_s$).

When $\gamma_s = \gamma_c = 0$, the mode "leaks" into the substrate and cover. This is analogous to the cutoff of an electromagnetic guided mode in a symmetric dielectric slab waveguide. This type of cutoff occurs when the zig-zag angle becomes simultaneously equal to the substrate-film critical angle and the cover-film critical angle.

The first appearance of the $M_0$ again occurs when the electron energy $E = V_s$ ($= V_c$). The $M_0$ first starts propagating as the thickness is increased to the value

\[
d = \frac{\hbar}{2m_s^*(V_s - V_f)} \left[ \tan^{-1} \left( \frac{m_s^*(V_c - V_s)}{m_s^*(V_f - V_s)} \right) \right]^{1/2} + n \pi. \tag{22}
\]

IV. EXAMPLE Ga$_{1-x}$Al$_x$As ELECTRON WAVEGUIDE

To illustrate the analysis presented in Sec. III, the example quantum well waveguide consisting of a substrate of Ga$_{0.40}$Al$_{0.60}$As, a film of GaAs, and a cover of Ga$_{0.35}$Al$_{0.65}$As is further treated in this section. The electron potential energies in the three regions of the quantum well are given by the conduction band edge as

\[
V_i = A x_i, \quad i = s, f, c, \tag{23}
\]

where $x_i = 0.2$ for the substrate, $x_f = 0$ for the film, and $x_c = 0.45$ for the cover, and $d = 0.7731$ eV (in which the conduction band discontinuity has been taken to be 60% of the energy gap change). Therefore $V_s = 0.1546$ eV, $V_f = 0.0000$ eV, and $V_c = 0.3479$ eV. The electron effective mass in the three regions is given by

\[
m_s^* = (B + C x_i) m_0, \quad i = s, f, c, \tag{24}
\]

where $m_0$ is the free electron mass, $B = 0.067$, and $C = 0.083$. Therefore $m_s^* = 0.0836 m_0$, $m_f^* = 0.067 m_0$, and $m_c^* = 0.10435 m_0$.

Numerically solving the waveguide dispersion Eq. (17), the waveguide modes were quantified as a function of the GaAs waveguide (film) thickness. Growth is taken to be along the [100] direction and thus each monolayer corresponds to a thickness of 0.28267 nm. Mode dispersion curves for the lowest-order mode, $M_0$, are shown in Fig. 4 for thicknesses of 10 monolayers ($d = 2.8267$ nm) and 20 monolayers ($d = 5.6533$ nm). The modes first appear at $E = V_s$ and $\beta = 0$, the lower-right-hand corner of the guided mode region in a $\beta$ vs $E$ plot such as Fig. 4. As the thickness is further increased, the mode dispersion curves move to the left and upward. In the limit of very large thickness the mode lines approach the left-hand $\beta$ curve given by Eq. (9). For the composition treated, there are no modes at all for a thick-

![FIG. 4. Propagation constant of the $M_0$ electron guided mode as a function of total electron energy (mode dispersion curve) for the example quantum well waveguide composed of Ga$_{0.40}$Al$_{0.60}$As (substrate), GaAs (film), Ga$_{0.35}$Al$_{0.65}$As (cover) with film thicknesses of 10 monolayers and 20 monolayers. The lower-energy cutoffs ($\beta = 0$) and upper-energy cutoffs are apparent.](https://via.placeholder.com/150)
TABLE I. Upper- and lower-cutoff energies and range of energies for the allowed waveguide modes for various film thicknesses in a Ga<sub>0.45</sub>Al<sub>0.55</sub>As (substrate), GaAs (film), Ga<sub>1.05</sub>Al<sub>0.10</sub>As (cover) quantum well waveguide.

<table>
<thead>
<tr>
<th>Waveguide film (GaAs) thickness</th>
<th>1.1307 (4)</th>
<th>1.4133 (5)</th>
<th>2.8267 (10)</th>
<th>5.6533 (20)</th>
<th>7.3493 (26)</th>
<th>7.6320 (27)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Propagating or cutoff</td>
<td>Propagating</td>
<td>Cutoff</td>
<td>Propagating</td>
<td>Cutoff</td>
<td>Propagating</td>
<td>Cutoff</td>
</tr>
<tr>
<td>Upper-cutoff energy, E&lt;sub&gt;uc&lt;/sub&gt; (eV)</td>
<td>0.1546</td>
<td>0.1190</td>
<td>0.0623</td>
<td>0.0451</td>
<td>0.0430</td>
<td>0.1866</td>
</tr>
<tr>
<td>Lower-cutoff energy, E&lt;sub&gt;lc&lt;/sub&gt; (eV)</td>
<td>0.1635</td>
<td>0.5657</td>
<td>0.7017</td>
<td>0.7268</td>
<td>0.7297</td>
<td>0.0334</td>
</tr>
<tr>
<td>Propagation energy range, ΔE (eV)</td>
<td>0.0089</td>
<td>0.4467</td>
<td>0.6394</td>
<td>0.6817</td>
<td>0.6867</td>
<td>0.1532</td>
</tr>
</tbody>
</table>

FIG. 5. The wave function amplitude (normalized to the maximum value in the film region) as a function of position for various energies for a quantum well waveguide with thickness of 10 monolayers (d = 2.8267 nm). The wave function is shown for an electron energy at the lower-energy cutoff (E<sub>y</sub> = 0.1190 eV), for an energy above the substrate barrier but below the cover barrier (E<sub>y</sub> = 0.25 eV), for an energy above both barriers (E<sub>y</sub> = 0.40 eV), for an energy at the upper-energy cutoff (E<sub>y</sub> = 0.5657 eV), and for a substrate “mode” at an energy above the upper-energy cutoff (E<sub>y</sub> = 0.70 eV).

Plots of the wave function ψ<sub>ψ</sub> for the M<sub>0</sub> mode for GaAs thickness of 10 monolayers are given in Fig. 5: electron energies of E = E<sub>uc</sub> (B = 0 cutoff), V<sub>f</sub> < E < E<sub>uc</sub> to illustrate guided mode behavior in these ranges of electron energy. The wave function is also shown for an energy above the upper-cutoff energy (E > E<sub>uc</sub>) to illustrate substrate “mode” behavior.
V. SUMMARY AND DISCUSSION

The use of a semiconductor quantum well as an electron slab waveguide has been analyzed. The dispersion relation, Eq. (17), governing the guided modes has been developed. The guided modes have been labeled \( M_v \), where \( v \) is the integer mode number \((v = 0, 1, 2, \ldots)\). A lower-energy cutoff \((\beta_v = 0)\) has been quantified in which the electron waveguide is analogous to an electromagnetic hollow metallic waveguide with finite conductivity walls. This cutoff simply corresponds to the allowed energies of the one-dimensional potential energy well. Unlike electromagnetic waveguides, an upper-energy cutoff is also found to be present. The upper-energy cutoff may be a transition to a substrate "mode," a transition to a radiation "mode," or a transition to a cover "mode." These cutoffs are due to dispersion effects. The effective electron phase refractive indices become more nearly equal at higher electron energies (see Fig. 3) and thus the index of the film is not sufficiently large with respect to the substrate and/or cover to support a guided mode. That is, electron total internal reflection [Eqs. (3) and (4)] no longer occurs. The upper cutoff energies occur above the potential barriers and thus electron waveguiding has been shown to be possible at energies well above the potential energy barriers.

Electron waveguides are potentially useful in high-speed electronic circuitry and as a central component in future electron guided-wave integrated circuits which could perform "optical-like" processing that is analogous to present-day electromagnetic optical processing that is done with guided-wave integrated optical circuits.

ACKNOWLEDGMENTS

This research was sponsored in part by a grant from the Joint Services Electronics Program under Contract No. DAAL-03-87-K-0059. One of us (K.F.B) was supported in part by a Presidential Young Investigator award from the National Science Foundation.

Theory and design of semiconductor electron-wave interference filter/emitters

E. N. Glytsis, T. K. Gaylord, and K. F. Brennan
School of Electrical Engineering and Microelectronics Research Center, Georgia Institute of Technology, Atlanta, Georgia 30332-0230

(Received 15 June 1989; accepted for publication 30 August 1989)

A voltage-biased semiconductor superlattice structure is designed to operate simultaneously as a continuously voltage-tunable, electron interference filter and as an electron emitter. Using the analogies between electromagnetic waves and electron de Broglie waves, a systematic procedure for designing the quantum wells and barriers comprising the electron-wave filter/emitter superlattice is developed. A generalized procedure for analyzing the electron-current transmittance and reflectance spectral responses of these superlattice structures is then presented. A practical, continuously tunable filter/emitter consisting of multiple layers of GaAs/AlAs (compositional superlattice) is designed to emit nearly monoenergetic 0.20-eV electrons by appropriate selection of the layer compositions and thicknesses. The constraints required to have thicknesses that are integer multiples of the monolayer thickness and to avoid phonon scattering of electrons into the L band are included. The filter/emitter is shown to have a wide tunable energy range. A sensitivity analysis of the device characteristics in the presence of fabrication errors reveals a very stable device response. Such quantum electron-wave devices could serve as continuously tunable hot-electron emitters in ballistic transistors and in future guided electron-wave integrated circuits.

I. INTRODUCTION

Quantum-mechanical interference effects have recently been observed in devices such as GaAs/AlGaAs and InGaAs/InAlAs single-well double-barrier structures. This indicates that present-day semiconductor devices are rapidly approaching a fundamental size limitation. At dimensions of less than about 0.3 μm, potentially “troublesome” quantum-wave effects start to dominate device characteristics, making further reduction in their sizes undesirable in some applications, thereby limiting the ultimate speed of the devices. This holds true regardless of the material (silicon, gallium arsenide, etc.). However, these quantum-wave effects may also potentially be used in the design of novel devices.

Starting from fundamental principles, quantitative analogies between quantum-mechanical electron waves in semiconductor materials and electromagnetic optical waves in dielectrics have recently been developed. With these analogies, existing electromagnetic optical analysis and design techniques can be used for the analysis and design of new semiconductor quantum-wave devices. The possibility of realizing these device structures in practice has become more likely due to the rapid progress and relative maturity in semiconductor growth technologies such as molecular-beam epitaxy (MBE) and metalorganic chemical vapor deposition (MOCVD) that enable the fabrication of structures with precise monolayer compositional control. Possible devices include narrow-band superlattice interference filters and filter/emitters (present work). These structures could be incorporated into present-day ballistic transistors to produce subpicosecond switching times. Beyond improving the speed of existing devices, however, the totally new concept of guided electron-wave integrated circuits has recently been proposed. This next generation of integrated circuits would be comprised of many semiconductor quantum-wave devices interconnected by electron waveguides.

In these ultrasmall superlattice interference filters or filter/emitters, electron waves can travel through the device maintaining their quantum-mechanical phase coherence. Thus these waves can interfere, reflect, refract, and diffract in a manner analogous to the electromagnetic plane waves in dielectric media. Even though quantum interference effects have been observed experimentally in single-well double-barrier structures, these effects can also occur in multiple-layer superlattices and at energies above the potential barriers. Furthermore, exploiting the electron-wave interference, novel electron-wave interference filter/emitters can be designed that are analogous to thin-film optical interference filters. These superlattice interference filter/emitters can exhibit very narrow electron kinetic energy passbands and can be integrated into solid-state devices for potential use as monoenergetic emitters for electroluminescent devices, photodetectors, and fast ballistic transistors.

Although semiconductor superlattice interference filter designs can be visualized directly from the optical interference filters, their designs cannot simply be copies of thin-film optical designs. The reason for this is that due to the ultrasmall dimensions of each superlattice layer (on the order of a nanometer), the thickness quantization has to be taken into account (layer thicknesses must be an integer multiple of the monolayer thickness). Another constraint is the limited usable composition range that is available. Furthermore, in the case of an applied bias voltage, the potential energy along the undoped superlattice changes linearly with the device length. This is analogous to a varying optical refractive index in a thin-film filter. Presently, there are no available designs for this linearly varying index case. Another practical constraint is that the desired electron-energy states should be sufficiently below the L-band minimum in order to avoid phonon scattering which can reduce the electron coherence.
length. The purpose of this paper is to present a systematic procedure for the design of continuously tunable semiconductor electron interference filter/emitters subject to the above constraints. An example design using the Ga\textsubscript{1-x}Al\textsubscript{x}As alloy system is presented, and the filter/emitter is shown to have a broad tunable range. Moreover, a sensitivity analysis of the device performance in the presence of fabrication variations indicates stable device characteristics.

II. SEMICONDUCTOR QUANTUM ELECTRON-WAVE OPTICS

Quantum-mechanical electron waves in semiconductors and electromagnetic optical waves in dielectrics exhibit transmission, reflection, interference, and diffraction characteristics that are analogous to each other since they satisfy analogous wave equations and boundary conditions. Quantum interference effects have been analyzed for single potential energy boundaries,\textsuperscript{6,7} for normal incidence,\textsuperscript{8} and for the general case of any number of boundaries and any angle of incidence.\textsuperscript{9} From these results a quantitative mapping between electromagnetic optical waves and quantum electron waves has been established.\textsuperscript{4} Thus, using these quantitative analogies, existing optical device designs can now have electron-wave device counterparts.

The plane-wave solutions of the Schrödinger's equation contain a wave-vector quantity $k$ of magnitude $k = [2m^* (E - V)]^{1/2}/\hbar$, where $m^*$ is the electron effective mass, $E$ is the total electron energy, $V$ is the electron potential energy, and $\hbar$ is Planck's constant divided by $2\pi$. All plane-wave phase effects (for plane-wave interference) are described by the wave vector $k$. As a consequence, an electron-wave phase refractive index $n_e$ (phase), can be introduced that is proportional to the square root of the product of the effective mass and the kinetic energy.\textsuperscript{4} That is,

$$n_e (\text{phase}) \propto \left[ m^* (E - V) \right]^{1/2}. \quad (1)$$

Furthermore, amplitude effects such as electron transmissivity and reflectivity may be described in terms of the wavefunction amplitude for an electron wave or in terms of the electric field amplitude for TE-polarized optical wave incident upon a boundary. Continuity of the wave function across a potential energy boundary and conservation of electron probability current normal to a potential energy boundary are analogous to the continuity of the tangential component of the electric field across a boundary between dielectrics and to the conservation of power flow normal to a boundary between dielectrics, respectively. Using these analogies, an electron-wave amplitude refractive index $n_e$ (amplitude) can be introduced\textsuperscript{6} that makes the optical and electron-wave transmissivities and reflectivities equivalent. That is,

$$n_e (\text{amplitude}) \propto \left[ (E - V)/m^* \right]^{1/2}. \quad (2)$$

Since the phase and amplitude effects are mathematically decoupled in the equations describing electromagnetic optical designs, the same designs have quantum electron-wave counterparts that are described by the two electron-wave refractive indices [Eqs. (1) and (2)]. Both types of these refractive indices exhibit normal dispersion; i.e., they increase with decreasing wavelength.

III. THIN-FILM OPTICAL INTERFERENCE FILTERS

Due to the analogies between the electromagnetic optical waves and the quantum-mechanical electron waves, the thin-film optical interference filters and the electron-wave interference superlattice filters have many common characteristics. For completeness some of the more relevant properties of thin-film optical filters\textsuperscript{20-22} are summarized below.

A simple type of narrow-bandpass optical interference filter is the all-dielectric Fabry–Perot filter. It consists of a half-wavelength layer sandwiched between quarter-wavelength layers of high refractive index (designated $H$) and of low refractive index (designated $L$). The combination of quarter-wavelength layers constitutes a reflector. The electron-wave analog of this elementary type of interference optical filter is treated in this work. The full width at half maximum (FWHM) of the bandpass of this filter can be reduced by increasing the reflectivity of the reflectors, i.e., by increasing the ratio of high to low refractive indices. In the optical literature there are two basic types of all-dielectric Fabry–Perot interference filters which are symbolically represented as $[HL]^n[HH]^n$ and $H[LL]^n[HL]^nH$, where $H$ and $L$ represent quarter-wavelength layers of high and low refractive index, respectively, and $n$ represents the number of repetitions of the layer-pair type indicated in brackets. Some other important characteristics of the all-dielectric Fabry–Perot interference filters, that are related with the electron-wave interference filter designs, are the following:

1. The maximum transmittance of the filter is 100%.
2. The maximum transmittance occurs at the wavelength for which the central layer is a half-wavelength thick (as measured in that material) and the surrounding layers are quarter-wavelength layers (as measured in those materials).
3. The FWHM decreases as the number of the surrounding layers increases (as $n$ increases).
4. The transmittance characteristics are relatively insensitive to variations in the reflectivities and thicknesses of the layers.

IV. DESIGN OF SEMICONDUCTOR SUPERLATTICE INTERFERENCE FILTER/EMITTERS

A. Computation of the layer thicknesses and compositions for a given bias voltage and electron energy

When a voltage bias is applied to a superlattice structure, the resulting device can serve as an electron-wave interference filter and electron emitter. Using the previously presented analogies between the quantum electron waves and the electromagnetic optical waves, the quantum well and barrier widths and heights in the direction of emission can be systematically designed to comprise an interference filter which is embedded in the emitter. The optical analog of this device does not exist since the equivalent phase and amplitude refractive indices of the electron waves vary along the emission direction of the structure due to the applied bias potential energy. Thus the optical counterpart would be comprised of inhomogeneous regions where the correspond-
ing optical refractive indices would vary as functions of the propagation distance within the optical filter. However, the same concepts that are used for the design of homogeneous optical interference filters can be adopted in the case of the electron-wave interference filter/emitter. Moreover, the successive quantum wells and barriers can act as the high (H) and low (L) electron refractive index materials that comprise the reflectors of the interference filter/emitter.

The electron potential energy of a quantum-well superlattice interference filter/emitter with an applied bias voltage is shown in Fig. 1. The structure consists of M layers surrounded by bulk semiconductor material. The filter/emitter is designed to be a simple narrow-bandpass interference filter (the optical counterpart would be an all-dielectric inhomogeneous Fabry–Perot interference filter). The filter/emitter consists of M layers, where M is an odd integer, and the central layer is a half-wavelength layer sandwiched between quarter-wavelength layers of high (H) and low (L) equivalent electron-wave refractive index. When the design voltage is applied, electrons in a narrow spectral band around the prespecified pass energy \( E_p \) (Fig. 1) traverse the filter/emitter and are emitted with an output energy \( K_{E_{out}} \).

At the design voltage, with corresponding bias potential energy \( V_{bias} \), each reflector layer [of high (H) or low (L) average electron-wave refractive index] is exactly a quarter of an electron wavelength (as measured in that layer) in the thickness, and the resonant central layer is exactly one-half of an electron wavelength (as measured in that layer). In the case of the biased filter/emitter, the kinetic electron energy varies linearly within each individual layer due to the applied voltage. Thus the electron wavelength changes continuously inside each layer. However, a quarter or a half of an electron wavelength is defined as the required layer thickness to produce a total phase shift of the traveling electron-wave function inside the layer of \( \pi/2 \) or \( \pi \), respectively (or odd multiples of these phase shifts). Although only an average electron wavelength can be defined within each layer, the terminology of “quarter” or “half electron wavelength” are used in order to be analogous to the corresponding quantities appearing in the design of optical thin-film interference filters. The jth quantum well or barrier (Fig. 1) has a thickness \( d_j \) and at zero bias a potential energy of \( V_j \). The layers surrounding the emitter/filter regions are taken to have the same zero-bias potential energy of \( V_{o} \). The material system in which the filter/emitter can be implemented is taken to form a continuous set of alloys of the type \( \text{G}_1 \text{..H}_x \text{K} \) where G, H, and K are chemical elements, and \( x \) is the atomic percentage of the H chemical element. For example, a practical material system is \( \text{Ga}_1 \text{..Al}_x \text{As} \).

The electron potential energy is given by

\[
V_j = \Delta E_c = A x_j, \tag{3}
\]

where \( \Delta E_c \) is the change in the energy of the conduction-band edge, and \( \Delta \) is a constant. Due to restrictions on usable composition range, the corresponding range of potential energies is \( 0 < V_j < V_{\text{max}} = A x_{\text{max}} \). The layers surrounding the filter/emitter regions have a composition of \( x_0 \). An additional design constraint for the electron-wave filter/emitter is that the layer thicknesses should be an integer multiple of the monolayer thickness. Thus the thickness of the jth layer \( d_j \) should be an integer multiple \( p_j \) of the monolayer thickness \( r_j \). For jth layer to be a quarter of an electron wavelength in thickness at the design pass energy \( E_{p} \), the phase difference between the input boundary \( z_{j-1} \) and the output boundary \( z_{j} \) (Fig. 1) must an odd integer multiple of \( \pi/2 \). That is,

\[
\begin{align*}
\int_{z_{j-1}}^{z_{j}} k_j(z)dz &= \int_{z_{j-1}}^{z_{j}} \left(1 - \frac{1}{r_j}\right)[2m_{e}\left(E_{p} - V_j(z)\right)]^{1/2}dz \\
&= (2q_j - 1)\pi/2, \tag{4}
\end{align*}
\]

where the potential energy in the jth layer with bias applied is given by

\[
V_j(z) = V_{bias} \left(1 - \frac{z}{L}\right) + V_j. \tag{5}
\]

\( L \) is the total length of the superlattice (Fig. 1), and \( q_j = 1, 2, \ldots \). The electron pass energy may be expressed as \( E_{p} = V_{bias} + V_{o} + K_{E_{in}} \) where \( K_{E_{in}} \) is the pass kinetic energy in the input region (leftmost region in Fig. 1). The electron effective mass is given by \( m^*_{e} = (B+C_{x_j})m_{0} \) where \( m_{0} \) is the free-electron mass, and \( B \) and \( C \) are material-system-dependent constants. Using Eqs. (3), (5), and the integral identity \( \int (a + bx)^{1/2}dz = (2/3b)(a + bx)^{3/2} \), Eq. (4) can be rewritten

\[
2L \left[2m_{0}(B+C_{x_j})\right]^{1/2} \times \left[\left(V_{o} + K_{E_{in}} - A x_j + V_{bias} \frac{z_j}{L}\right)^{3/2} - \left(V_{o} + K_{E_{in}} - A x_j + V_{bias} \frac{z_{j-1}}{L}\right)^{3/2}\right] = (2q_j - 1)\pi/2, \tag{6}
\]

The above equation has to be solved for the composition \( x_j \) of the jth layer with \( V_{bias}, V_{o}, \) and \( K_{E_{in}} \) as design parameters. Equation (6) is valid for all the layers except the center reso-

---

Fig. 1. Schematic representation of a biased semiconductor superlattice electron-wave interference filter/emitter. At the design potential energy bias \( V_{bias} \), and input kinetic energy \( K_{E_{in}} \), the layers have a thickness of a quarter (or a half for the resonant central layer) of an electron wavelength as measured in that layer.
nant layer. In this case the \( \pi/2 \) term of Eq. (6) is replaced by \( \pi \). The monolayers of the device will be numbered with the index \( i \). The rightmost monolayer of the \( j \)-th region is \( i_j \). The total number of monolayers in the \( j \)-th region is \( i_j \). Therefore, the number of monolayers of the \( j \)-th layer is \( i_j = i_j - i_{j-1} \), and the corresponding thickness of this layer is \( d_i = p_i r_i = z_i - x_{j-1} \). The total thickness of the structure is \( L = \sum_{j=1}^{M} p_j r_j \) (if the monolayer thicknesses are the same in each layer, \( r_1 = r_2 = \ldots = r_M = r \), then \( L = i_j r \)). The boundaries between the various layers which are designated as \( z_i \) (Fig. 1) are given by \( z_i = \sum_{j=1}^{i-1} p_j r_j \) (for equal monolayer thicknesses, \( z_i = (i_j/i_M)L \)). It is worth mentioning that the solution of Eq. (6) is not trivial since Eq. (6) for the \( j \)-th layer depends on the thicknesses of the other layers. Furthermore, Eq. (6) is valid only if \( V_{\text{bias}} \neq 0 \). In the case of zero applied bias, \( V_{\text{bias}} = 0 \), Eq. (4) leads to Eq. (11) of Ref. 8, which is much simpler to solve.

A systematic procedure for designing a biased continuously tunable superlattice interference filter/emitter for a given output kinetic energy (\( KE_{\text{out}} = KE_{\text{in}} + V_{\text{bias}} \)) and given compositions \( (x_0) \) of the surrounding filter/emitter regions, using Eq. (6), is now described. The design parameters \( V_0 \) (or \( x_0 \)), \( KE_{\text{in}} \), and \( V_{\text{bias}} \) have to be specified at the beginning. Next the characteristics of the unbiased superlattice electron-wave interference filter \( (x_0) \) are used as a starting point in estimating the number of regions \( M \) and the number of monolayers \( p_j \) \((j = 1,2,\ldots,M)\) to be used in the filter/emitter. The initial values of the number of monolayers in each layer are designed as \( p_j \), the total initial thickness is designated as \( L^0 \), and the total number of monolayers is designated as \( i^0 \). It is important to keep the length of the structure as small as possible in order to be less than the electron coherence length. Consequently, it is more practical to set the parameter \( q_j = 1 \) in order to find the minimum thickness device. Starting from the first layer \( (j = 1) \), Eq. (6) is solved for \( x_j \). Since an iterative scheme is used, the solution of Eq. (6) for \( x_j \) is designated as \( x_j^{n} \), where \( n_1 \) is a counting index for the iterations of the \( j \)-th layer. In the next step it is determined if the solution \( x_j^{n} \) lies within the range of usable composition values \((0 < x_j^{n} < x_{\max}) \). If \( x_j^{n} \) is outside the usable range, then the number of monolayers is changed as \( p_j^{n} = p_j^{n-1} + \Delta p_j^{n} \), where \( \Delta p_j^{n} \) is the change in the number of monolayers of region \( j \). The procedure is repeated until a value for \( x_j^{n} \) is found within the usable range. For the filter to have a narrower response, it is important to increase the reflectivity between the low \( (L) \) and high \( (H) \) electron refractive index layers. For the low index layers the composition should be close to \( x_{\max} \) as possible. Similarly, for the high index layers the composition should be close to zero as possible. As a result, some additional refining of the calculated solution for the \( j \)-th layer can be found by changing the value of \( p_j^{n} \) by \( \pm 1 \) \((\Delta p_j^{n} = \pm 1) \). This is continued until the solution \( x_j^{n} \) is the maximum possible (within the usable composition range) for a low \( (L) \) index layer or the minimum possible (within the usable composition range) for a high \( (H) \) index layer. This procedure is repeated for all the layers \((j = 1,2,\ldots,M)\). However, the solution for the last layer may result in a new value of \( i_M \) (or equivalently \( L^0 \)). In this case the resultant value of \( L \) is set equal to \( L^0 \) and the design procedure is repeated until the final thickness \( L \) is consistent with the initially assumed thickness \( L^0 \), i.e., \( L = L^0 \). The design procedure is summarized in the algorithm presented in Fig. 2.

**B. Computation of the designed filter/emitter response**

In order to evaluate the response of the interference filter/emitter designed with the above-described procedure, the electron-wave function has to be calculated in the output region of the structure. Using the one-electron-wave function approximation, and neglecting the electron-electron and the exchange interaction which partially compensates one another, the electron-wave function in each layer satisfies the Schrödinger's equation. That is,

\[
\frac{\hbar^2}{2m^*} \frac{d^2 \psi_j(z)}{dz^2} + [E - V_j(z)] \psi_j(z) = 0, \tag{7}
\]

where \( \psi_j(z) \) is the electron-wave function in region \( j \) \((j = 1,2,\ldots,M)\), and \( V_j(z) \) is given by Eq. (5). The electron-wave function in any layer of the biased superlattice can be expressed as a linear combination of Airy functions \( Ai(\rho) \) and complimentary Airy functions \( Bi(\rho) \). By defining within each layer a new variable \( \rho_j = \rho_j(z) = (2m^* V_{\text{bias}}/\hbar^2 L)^{1/3} \left[z + (E - V_{\text{bias}} - V_j) L/V_{\text{bias}} \right] \), Eq. (7) can be transformed into the form \( d^2 \psi_j/ \rho_j^2 + \rho_j \psi_j = 0 \), which has solution of the form

\[
\psi_j(\rho_j) = C_j Ai[- \rho_j(z)] + D_j Bi[ - \rho_j(z)], \tag{8}
\]

where \( C_j \) and \( D_j \) are unknown amplitude constants. The solution of Schrödinger's equation in the two external regions (leftmost and rightmost regions in Fig. 1) is straightforward since the potential energy does not vary in these regions. Specifically, the wave function in the input region is given by \( \psi_0(z) = \exp(jk_0 z) + \exp(-jk_0 z) \), while the wave function for the output region is given by \( \psi_{M+1}(z) = \exp(jk_{M+1} z) + \exp(-jk_{M+1} z) \). The \( k_0 \) and \( k_{M+1} \) are the electron-wave vectors in the input and output regions, respectively. Using the boundary condition for the wave function, \( \psi_j(z_j) = \psi_{j+1}(z_j) \), and \( (1/m^*_j)d\psi_j(z_j)/dz = (1/m^*_{j+1})d\psi_{j+1}(z_j)/dz \), the reflected and transmitted amplitude of the wave function \((r \text{ and } t)\) may be expressed as...
The symmetry of the electron Fabry—Perot. For positional
being assumed). In the previously described design process,
the material composition (alloys of the type G, "H, K are
L-valley potential energy in the jth layer,
additional constraints can be imposed in the design process
mum, then the resulting phonon scattering in the device may
decrease the electron coherence length. As a result the
phonon scattering can deteriorate the performance of an
electron-wave interference filter/emitter. However, some
constraint:

\[ V_0 + \Delta E_1 < E < V_{L0} - \Delta E_2, \]  

where \( V_{L0} \) is the L-valley electron potential energy in the
output region \( (V_{L0} = D + Fx_0) \). It is implied that the bias
potential energy is non-negative as shown in Fig.
1 \( (V_{bias} > 0) \). Inequalities (13) impose the following condi-
tion for the allowable bias potential energy:

\[ V_{bias} < D - (A - F)x_0 - \Delta E_1 - \Delta E_2 \]  

Inequality (14) limits the maximum value of the applied bias
potential energy to \( (V_{bias})_{max} = D - (A - F) \)
\( x_0 - \Delta E_1 - \Delta E_2 \). The maximum applied bias potential energy
depends on the compositions of the input and output re-
gions. Constraints similar to inequality (13) can be applied
for each layer inside the device. That is,

\[ V_j (z) < E < V_{Lj}, \]  

for \( z_{j-1} < z < z_j \),

which imposes on the composition \( x_j \) the following con-
straint:

\[ \frac{1}{A} \left( Ax_0 + \Delta E_1 + V_{bias} \frac{z_j}{L} - D \right) \]  

\[ < x_j \frac{1}{F} \left( Ax_0 + V_{bias} \frac{z_j-1}{L} \right). \]  

Inequality (15) can be even tighter if energy margins \( \Delta E_{1j} \)
and \( \Delta E_{2j} \) are nearly equal to \( \Delta E_1 \) and \( \Delta E_2 \). However, for
practical purposes the inequalities (13) and (15) can be con-
sidered sufficient.

The design procedure described in Sec. IV A remains
basically the same with the exception that \( V_{bias} \) has to satisfy
inequality (14) and the layer compositions should lie in the
range expressed by inequality (16).

**V. Ga\(_{1-x}\)Al\(_x\)As EXAMPLE CASES**

A practical material system to be used in fabricating the
above-described electron-wave interference filter/emitters is
Ga\(_{1-x}\)Al\(_x\)As. For this material system the maximum com-
position in Al is \( x_{max} = 0.45 \) in order to avoid the direct/indirect
band-gap transition. Other parameters of this mate-
rial system include \( A = 0.773 \) eV, \( B = 0.067, C = 0.083, D = 0.284 \) eV,
and \( F = 0.168 \). The \( (100) \) monolayer thickness for any usable composition remains the same
\( r_j = 0.282 \) 665 nm (lattice-matched material system).

At first, some designs that are not restricted by the L-
valley energy are presented. For these filters/emitters the

\[ \left( \frac{1}{r} \right) = \frac{1}{2j_{K_1}} \left( \frac{1}{j_{K_2}} \right) \left[ S_j(z_{j0}) \right]^{-1} \left[ S_j(z_{j1}) \right]^{-1} \left[ S_j(z_{j2}) \right]^{-1} \left[ S_j(z_{j3}) \right]^{-1} \cdots \left[ S_j(z_{jM-1}) \right]^{-1} \]

where

\[ S_j(z) = \left( \frac{Ai[- \rho_j(z)]}{(1/\sqrt{m_j})^{2/3} \frac{Ai'}{- \rho_j(z)}} \right) \left( \frac{Bi[- \rho_j(z)]}{(1/\sqrt{m_j})^{2/3} \frac{Bi'}{- \rho_j(z)}} \right) \]

and where \( m_j^* = m_j^*/m_0, K_1 = [(2m_0/\hbar^2)^{1/3}(E - V_0 - V_{bias})/(L/\sqrt{m_0})]^{1/3}, \)
and \( K_2 = [(2m_0/\hbar^2)^{1/3}(E - V_0)/(L/\sqrt{m_0})]^{1/3} \)
and \( Ai' \) and \( Bi' \) are the first derivatives of the Airy and
complimentary Airy functions, respectively. Equations (9)
can be solved directly for the amplitude transmittance \( t \) and
the amplitude reflectance \( r \). The electron current trans-
mitance \( T_e \) is given by

\[ T_e = \left( \frac{E - V_0 - V_{bias}}{E - V_0 - V_{bias}} \right)^{1/2} \left( \frac{E - V_0 - V_{bias}}{E - V_0 - V_{bias}} \right)^{1/2} \]

Using Eq. (11), the response of the designed electron inter-
ference filter/emitter can be evaluated.

C. Design of a filter/emitter with L-valley electron
potential energy constraint

If the electron energy is near the \((111)\) L-valley mini-
num, then the resulting phonon scattering in the device may
decrease the electron coherence length. As a result the
phonon scattering can deteriorate the performance of an
electron-wave interference filter/emitter. However, some
additional constraints can be imposed in the design process
of the device that can minimize the phonon scattering. The
L-valley potential energy in the \( j \)th layer, \( V_{Lj} \), is given by

\[ V_{Lj} = D + Fx_j \]

where \( D \) and \( F \) are material-dependent constants, and \( x_j \) is
the material composition (alloys of the type \( Ga_{1-x}Al_x \) are
being assumed). In the previously described design process,
the composition of the input and the output regions of the
device were taken to be the same \( (x_0) \). This assumption does
not imply any restriction in the device design, but only re-

\[ m_j^* = m_j^*/m_0, K_1 = [(2m_0/\hbar^2)^{1/3}(E - V_0 - V_{bias})/(L/\sqrt{m_0})]^{1/3}, \]
and \( K_2 = [(2m_0/\hbar^2)^{1/3}(E - V_0)/(L/\sqrt{m_0})]^{1/3} \)
and \( Ai' \) and \( Bi' \) are the first derivatives of the Airy and
complimentary Airy functions, respectively. Equations (9)
can be solved directly for the amplitude transmittance \( t \) and
the amplitude reflectance \( r \). The electron current trans-
mitance \( T_e \) is given by

\[ T_e = \left( \frac{E - V_0 - V_{bias}}{E - V_0 - V_{bias}} \right)^{1/2} \left( \frac{E - V_0 - V_{bias}}{E - V_0 - V_{bias}} \right)^{1/2} \]

Using Eq. (11), the response of the designed electron inter-
ference filter/emitter can be evaluated.

C. Design of a filter/emitter with L-valley electron
potential energy constraint

If the electron energy is near the \((111)\) L-valley mini-
num, then the resulting phonon scattering in the device may
decrease the electron coherence length. As a result the
phonon scattering can deteriorate the performance of an
electron-wave interference filter/emitter. However, some
additional constraints can be imposed in the design process
of the device that can minimize the phonon scattering. The
L-valley potential energy in the \( j \)th layer, \( V_{Lj} \), is given by

\[ V_{Lj} = D + Fx_j \]

where \( D \) and \( F \) are material-dependent constants, and \( x_j \) is
the material composition (alloys of the type \( Ga_{1-x}Al_x \) are
being assumed). In the previously described design process,
the composition of the input and the output regions of the
device were taken to be the same \( (x_0) \). This assumption does
not imply any restriction in the device design, but only re-

\[ m_j^* = m_j^*/m_0, K_1 = [(2m_0/\hbar^2)^{1/3}(E - V_0 - V_{bias})/(L/\sqrt{m_0})]^{1/3}, \]
and \( K_2 = [(2m_0/\hbar^2)^{1/3}(E - V_0)/(L/\sqrt{m_0})]^{1/3} \)
and \( Ai' \) and \( Bi' \) are the first derivatives of the Airy and
complimentary Airy functions, respectively. Equations (9)
can be solved directly for the amplitude transmittance \( t \) and
the amplitude reflectance \( r \). The electron current trans-
mitance \( T_e \) is given by

\[ T_e = \left( \frac{E - V_0 - V_{bias}}{E - V_0 - V_{bias}} \right)^{1/2} \left( \frac{E - V_0 - V_{bias}}{E - V_0 - V_{bias}} \right)^{1/2} \]

Using Eq. (11), the response of the designed electron inter-
ference filter/emitter can be evaluated.
Design Parameters

\[ V_{\text{bias}}, K_{\text{E}_m}, V_0 \]

Set \( q = 1 \)

\[ j = 0 \]

\[ n_i = 1 \]

\[ j = j + 1 \]

Solve Eq. (6) for \( x_{n_i}^{(j)} \) with \( p_{j_0}^{(0)} \) as initial value

\[ p_{n_i}^{(j)} = p_{n_i}^{(j-1)} + \Delta p_{j_0}^{(j)} \]

\[ 0 \leq x_{n_i}^{(j)} \leq x_{\text{max}} \]

Is \( x_{n_i}^{(j)} \) the max or min possible value for \( L \) or \( H \) layers

\[ I = M \]

\[ L = \sum_{j=1}^{M} p_{j}^{(N)} f_j \]

Design is Completed

\[ x_i = x_{n_i}^{(j)} \]

\[ p_i = p_{n_i}^{(j)} \]

TABLE I. Design parameters of an electron-wave interference filter/emitter with the \((HL)^{N-1} HH(LH)^{N-1} [N = (M-1)/2]\) configuration. The surrounding input and output regions consist of GaAs/GaAlAs, and the filter/emitter is designed to emit 0.20-eV electrons when biased at 0.10 eV.

<table>
<thead>
<tr>
<th>Layer</th>
<th>( M = 9 )</th>
<th>( M = 13 )</th>
<th>( M = 17 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_j )</td>
<td>( x_j )</td>
<td>( p_j )</td>
<td>( x_j )</td>
</tr>
<tr>
<td>1</td>
<td>7</td>
<td>0.2222</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>9</td>
<td>0.4151</td>
<td>9</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>0.2663</td>
<td>7</td>
</tr>
<tr>
<td>4</td>
<td>9</td>
<td>0.4493</td>
<td>9</td>
</tr>
<tr>
<td>5</td>
<td>12</td>
<td>0.0639</td>
<td>7</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>0.4364</td>
<td>8</td>
</tr>
<tr>
<td>7</td>
<td>6</td>
<td>0.1442</td>
<td>6</td>
</tr>
<tr>
<td>8</td>
<td>7</td>
<td>0.3748</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>6</td>
<td>0.1951</td>
<td>6</td>
</tr>
<tr>
<td>10</td>
<td>8</td>
<td>0.3828</td>
<td>8</td>
</tr>
<tr>
<td>11</td>
<td>6</td>
<td>0.1626</td>
<td>6</td>
</tr>
<tr>
<td>12</td>
<td>7</td>
<td>0.3815</td>
<td>7</td>
</tr>
<tr>
<td>13</td>
<td>6</td>
<td>0.1982</td>
<td>6</td>
</tr>
<tr>
<td>14</td>
<td>7</td>
<td>0.3762</td>
<td>7</td>
</tr>
<tr>
<td>15</td>
<td>6</td>
<td>0.1729</td>
<td>6</td>
</tr>
<tr>
<td>16</td>
<td>7</td>
<td>0.3815</td>
<td>7</td>
</tr>
<tr>
<td>17</td>
<td>6</td>
<td>0.1982</td>
<td>6</td>
</tr>
</tbody>
</table>

FIG. 2. Flow chart of the calculational procedure used for the computation of the filter/emitter layer compositions and thicknesses.

The design parameters \( V_{\text{bias}} \) and \( K_{E_m} \) are chosen both to be 0.10 eV each. Using the procedure described in Sec. IV A (algorithm of Fig. 2), the layer thicknesses (expressed in number of monolayers) and the layer compositions can be calculated. These results are summarized in Table I for various number of layers \( M = 9, 13, \) and 17. The total thickness of the design structures is 20.1 nm (71 monolayers), 28 nm (99 monolayers), and 36.2 nm (128 monolayers) for \( M = 9, 13, \) and 17, respectively, which is within the desired electron coherence length. Experimental measurements in ballistic hot-electron devices on GaAs/GaAlAs and InGaAs/InAlAs indicate that the electron coherence length lies roughly between 10 and 100 nm. However, the electron coherence length is a statistical quantity. Therefore, the experimental data suggest that at least some measurable fraction of electrons exhibit a coherence length within the design for the filter/emitter requirements. The spectral response (transmittance versus output kinetic energy) of the designed structure can be computed using the analysis described in Sec. IV B. The electron-current transmittance \( T_e \) [Eq. (11)] is shown in Figs. 3(a), 3(b), and 3(c) as a function of the output kinetic energy \( E_{\text{out}} \) for \( M = 9, 13, \) and 17, respectively, with the applied bias potential energy \( V_{\text{bias}} \) as a parameter. From these figures it is shown that for \( V_{\text{bias}} = 0.10 \text{ eV} \) the filter/emitter has a peak in the electron-current transmittance (of about 100%) for the design output kinetic energy of 0.20 eV. The spectral response of the filter/emitter has a narrow passband around the designed output kinetic energy, and the FWHM decreases as the number of layers increases. For the cases shown in Figs. 3(a), 3(b), and 3(c), the FWHM values are 30.7, 13.4, and 5.9 meV for \( M = 9, 13, \) and 17, respectively. However, by increasing the number of the layers \( M \), (1) the total thickness of the device increases (thickness should be less than the electron coherence length), and (2) some secondary peaks in the electron-current transmittance appear as shown in Figs. 3(b) and 3(c). Thus there is a trade-off between the FWHM and the device thickness.
FIG. 3. Electron-current transmission of $M$-layer Ga$_{1-x}$Al$_x$As superlattice interference filter/emitter as a function of the electron output kinetic energy for $V_{\text{bias}} = 0.05, 0.10, \text{and} 0.15 \text{eV}$ for (a) $M = 9$, (b) $M = 13$, and (c) $M = 17$. The spectral tuning with the bias is apparent. Furthermore, the spectral widths become narrower as the number of the layers $M$ increases.

Another important feature of the filter/emitter spectral response is that using the same design (of $V_{\text{bias}} = 0.10 \text{eV}$ and $K_{E\text{out}} = 0.10 \text{eV}$), and applying other bias potential energies, the narrow-band characteristic of the filter/emitter remains almost unchanged, but shifted in output kinetic energy. For example, the spectral response of the designed filter/emitter is shown in Figs. 3(a), 3(b), and 3(c) for $V_{\text{bias}} = 0.05, 0.10, \text{and} 0.15 \text{eV}$. It is observed that there is no significant change in the electron-current transmittance peak and the FWHM, but there is a shift in the output kinetic energy for which the peak electron-current transmittance occurs. The change in the output kinetic energy $\Delta K_{E\text{out}}$ is approximately half of the change in the applied bias potential energy $\Delta V_{\text{bias}}$. This can be explained by the fact that the resonant central layer $(HH)$ is almost at the middle of the device and consequently experiences half of the change of the bias potential energy. Furthermore, due to this property of the filter/emitter, the device can be continuously tuned over a range of output kinetic energies, around the designed value, by varying the applied voltage. The tunability of the device is further illustrated in Fig. 4 (for $M = 13$) for a range of bias potential energies from 0.0 to 0.2 eV. It is observed that the output kinetic energy at the peak electron-current transmittance varies linearly with the applied bias potential energy in the range between $\sim 0.15$ and $\sim 0.25 \text{eV}$ (the slope of the straight line in Fig. 4 is 0.496). Moreover, the peak of the electron-current transmittance, $T_{\text{peak}}$, varies smoothly around the design value of $V_{\text{bias}}$ (0.10 eV) in the range between 92.1$\%$ and $\sim 100\%$.

The electron-current transmittance of the filter/emitter as a function of the applied bias potential energy with the electron output kinetic energy $K_{E\text{out}}$ as a parameter is shown in Fig. 5 for $V_{\text{bias}}$, varying between 0.0 and 0.3 eV with $K_{E\text{out}} = 0.10, 0.15, 0.20$ (design value), 0.25, and 0.30 eV. The design parameters of this filter/emitter are given in Table I for $M = 9$. Figure 5 is equivalent to Fig. 3(a), but with the roles of $K_{E\text{out}}$ and $V_{\text{bias}}$ interchanged. It is observed that for $K_{E\text{out}} = 0.15, 0.20, 0.25,$ and $0.30 \text{eV}$ there is a peak in the electron-current transmittance. For $K_{E\text{out}} = 0.10 \text{eV}$ the transmittance is low as it can also be seen from Fig. 3(a).

FIG. 4. Tunability characteristics of a filter/emitter for $M = 13$. The design values are $V_{\text{bias}} = 0.10 \text{eV}$ and $K_{E\text{out}} = 0.10 \text{eV}$. The electron output kinetic energy at peak electron-current transmittance and the value of the peak electron-current transmittance are shown as functions of the bias potential energy $V_{\text{bias}}$. 
The peak value of the transmittance decreases as the output kinetic energy deviates from the design value (of 0.20 eV at $V_{\text{bias}} = 0.10$ eV).

Another important practical factor is the sensitivity of the designed electron-wave interference filter/emitters to fabrication process variations. Since these devices are ultrasmall, the layer thicknesses should ideally be the design number of monolayers. Furthermore, the Al composition usually can be controlled to within 1%-2%. In order to evaluate the sensitivity of the electron-wave filter/emitter, the layer thicknesses and compositions were varied from the design values, and the spectral response of the resulting devices were calculated using Eq. (11). Three cases were considered along with the design case (case 1, $V_{\text{bias}} = 0.10$ eV, $KE_{\text{in}} = 0.10$ eV, and $M = 9$). These are summarized in Table II: Case 2 corresponds to fabrication variations of the Al compositions ($x_i$'s), while cases 3a/3b and 4 correspond to fabrication variations of the central (resonant) layer thickness and of the surrounding layer thicknesses, respectively. The spectral responses that correspond to the filter/emitters of cases 1 and 2 are shown in Fig. 6(a). It is observed that even if the layer compositions vary between 0.5%-2.2% (possible fabrication variation) from the design values the filter/emitter response is not substantially altered. When the thicknesses of the layers deviate from the optimal design values (cases 3a/3b and 4), the filter emitter response is shown in Fig. 6(b). From Figs. 6(a) and 6(b) it is observed that deviation from the design thicknesses is much more important than the deviation from the design compositions. Especially, the variation of the thickness of the resonant layer is most important. However, from the same figures [6(a) and 6(b)] it is shown that even though the designs that correspond to cases 2, 3a/3b, and 4 are nonoptimal, the filter/emitter characteristics remain basically unaffected. Shifts of the peak electron-current transmittance are observed, while the peak transmittance value and the corresponding FWHM remain approximately the same. Even if the filter/emitter response is shifted due to fabrication process variations, it is possible to shift the response back to its designed position by
changing the applied voltage.

Finally, an additional consideration in the design of electron-wave interference filter/emitters is the restriction that the design electron energy be sufficiently below the L-valley potential energy to minimize phonon scattering. Selecting \( \Delta E_1 = KE_{in} = \Delta E_2 = 0.05 \text{ eV} \) and the input/output regions compositions 0.10 and 0.20 \((x_0 = 0.10 \text{ or } 0.20)\), two additional filter/emitters were designed. As is described by inequalities (14) and (16), there is a restriction in the maximum applied bias potential energy and the layer compositions. For \( x_0 = 0.10 \) and 0.20 the corresponding maximum values of the applied bias potential energy are 0.124 and 0.063 eV, respectively. If the layer compositions are less than \( x_0 \) \((x_j < x_0 \text{ for } j = 1,2,...,M)\), then inequality (16) is also satisfied. Using \( V_{\text{bias}} = 0.12 \) and 0.06 eV as the design parameters for \( x_0 = 0.10 \) and 0.20, respectively, and \( M = 9 \) layers, two additional designs were developed that take into account this electron-energy constraint. The characteristics of these two designs are summarized in Table III. The spectral (energy) response of the corresponding two filter/emitters is shown in Figs. 7(a) \((x_0 = 0.10, V_{\text{bias}} = 0.12 \text{ and } 0.07 \text{ eV})\) and 7(b) \((x_0 = 0.20, V_{\text{bias}} = 0.06 \text{ and } 0.03 \text{ eV})\). From these figures it is again observed that the electron-current transmittance has a peak at the design KE_{out} and \( V_{\text{bias}} \) values and, in addition, has the same tunability properties that were described previously. However, the design that corresponds to \( x_0 = 0.10 \) does not have a particularly good rejection band (since the transmittance value is not very small). This is due to the reduced reflectivity at layer boundaries. The reduced reflectivity is a result of the small differences in the layer compositions, which in turn correspond to small differences in the electron-wave refractive indices. Especially at larger electron energies, the corresponding electron-wave refractive indices become nearly equal, and this degrades the performance of the filter/emitter at energies above the pass energy. However, when \( x_0 = 0.20 \) there is a significant variation of the electron-wave refractive index, which results in an increased reflectivity between the filter/emitter layers. For this reason the response that corresponds to \( x_0 = 0.20 \) has much narrower energy passband and lower rejection band transmittance.

VI. DISCUSSION AND SUMMARY

In all the electron-wave interference filter/emitters presented, the semiconductor materials had parabolic band structure. However, for sufficiently high electron energies, the band structure is more appropriately treated as nonparabolic. Furthermore, the band structure may vary with the direction of the electron-wave propagation. Both of these effects can be incorporated by using an energy-dependent anisotropic effective mass. In this case the solutions of Schrödinger's equation described in Sec. III B needs to be modified; however, the design procedure of the filter/emitter remains unchanged.

The response of the electron-wave Fabry-Perot filter/emitter has been shown to have a narrow passband around the design value of the output kinetic energy. However, additional passbands in kinetic energy exist beyond the rejection band. At large electron energies the corresponding electron-wave refractive indices of the filter/emitter layers become

<table>
<thead>
<tr>
<th>Layer</th>
<th>( V_{\text{bias}} = 0.12 \text{ eV}, x_0 = 0.10)</th>
<th>( V_{\text{bias}} = 0.06 \text{ eV}, x_0 = 0.20)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( J )</td>
<td>( p_j )</td>
<td>( x_j )</td>
</tr>
<tr>
<td>1</td>
<td>12</td>
<td>0.0194</td>
</tr>
<tr>
<td>2</td>
<td>13</td>
<td>0.0677</td>
</tr>
<tr>
<td>3</td>
<td>12</td>
<td>0.0278</td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>0.0831</td>
</tr>
<tr>
<td>5</td>
<td>20</td>
<td>0.0307</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>0.0602</td>
</tr>
<tr>
<td>7</td>
<td>9</td>
<td>0.0077</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>0.0965</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>0.0489</td>
</tr>
</tbody>
</table>
detectors, and in future guided electron-wave integrated circuits.

ACKNOWLEDGMENTS

This research was sponsored in part by a grant from the Joint Services Electronics Program under Contract No. DAAL-03-87-K-0059. One of us (K.F.B.) was supported in part by a Presidential Young Investigator award from the National Science Foundation and another (E.N.G.) by a Research Initiation award from the National Science Foundation.
Electron-wave quarter-wavelength quantum well impedance transformers between differing energy-gap semiconductors

T. K. Gaylord, E. N. Glytsis, and K. F. Brennan
School of Electrical Engineering and Microelectronics Research Center, Georgia Institute of Technology, Atlanta, Georgia 30332

(Received 21 July 1989; accepted for publication 13 November 1989)

Impedance transformers for ballistic (collisionless) electron waves traveling between dissimilar energy-gap semiconductors are designed as a series of quarter (electron) wavelength layers in the form of a compositional superlattice. The quantitative analogies that have been previously established [J. Appl. Phys. 65, 814 (1989)] between electron-wave propagation in semiconductors and electromagnetic-wave propagation in dielectrics are used. For the design energy, the electron wave would be totally transmitted and the structure is analogous to an antireflection coating in electromagnetic optics. Practical constraints on the impedance transformer layers are (1) their compositions must be within the usable compositional range and (2) their thicknesses must be integer multiples of a monolayer thickness. These constraints are included in the design process. Procedures for designing narrow-band, maximally flat (Butterworth), and equal-ripple (Chebyshev) impedance transformers of arbitrary spectral bandwidth are presented. Example practical single-layer and three-layer transformers for connecting GaAs and Ga$_{0.4}$Al$_{0.6}$As are presented.

I. BACKGROUND

Molecular-beam epitaxy (MBE) and metalorganic chemical vapor deposition (MOCVD) have been developed and refined to the point where single monolayers can be grown with precise compositional control. These technologies have allowed important multiple quantum well devices composed of wide and narrow band-gap semiconductors such as GaAs and Ga$_{1-x}$Al$_x$As to be produced. With further improvements in the quality of the materials grown, ballistic transport has been observed. That is, electrons travel through the device without being scattered by defects from crystalline perfection. Even with the addition of elastic scattering, the electrons exhibit clear quantum mechanical plane-wave behavior. Since these coherent waves maintain their phase throughout the device, they can be refracted, reflected, interfered, guided, and diffracted in a manner quantitatively analogous to electromagnetic waves. Quantum interference effects have been observed experimentally for electron energies below the potential energy barriers in resonant tunneling structures and for electron energies above the barriers in negative differential resistance and potential barrier devices.

In the construction of semiconductor quantum devices and guided electron-wave integrated circuits it will be necessary to connect semiconductor materials with differing electron energy-band structures. This occurs, for example, in the electron flow from the GaAs base of a bipolar junction transistor to the Ga$_{1-x}$Al$_x$As collector where Ga$_{1-x}$Al$_x$As is used to increase the breakdown voltage in the high electric field collector. In such a configuration, detrimental reflections will occur at the energy-band discontinuity between materials. However, for a given electron energy or a given band of electron energies, these reflections can be completely eliminated and total transmission achieved by fabricating an impedance transformer between the two materials. The type of impedance transformer directly amenable to semiconductor technology consists of a series of layers that are each a quarter of an electron wavelength thick. Transformers of arbitrary bandwidth can be designed using the quantitative analogies between electron-wave propagation in semiconductors and electromagnetic waves in dielectrics and the theory of impedance transformers as developed for microwave transmission lines and waveguides. In addition, practical constraints require that the impedance transformer layers have (1) compositions that are within the usable compositional range and (2) thicknesses that are an integer multiple of a monolayer thickness. The purpose of this paper is to present systematic design procedures for narrow-band, maximally flat (Butterworth), and equal-ripple (Chebyshev) impedance transformers subject to the above constraints and then apply these procedures to example impedance transformers between GaAs and Ga$_{0.4}$Al$_{0.6}$As.

II. SEMICONDUCTOR ELECTRON-WAVE OPTICS

The transmission and reflection characteristics of quantum mechanical electron waves in semiconductors have been analyzed for single potential energy boundaries, for normal incidence, and for the general case of any number of boundaries and any angle of incidence. From these results, a mapping has been established between electromagnetic and quantum mechanical quantities. As a consequence, existing electromagnetic optical device designs now have electron-wave device counterparts.

Electron-wave phase effects such as the phase thicknesses of the impedance transformer layers are described by the electron wave-vector magnitude which is

\[ k = \left[ 2m^* (E - V) / \hbar \right]^{1/2}, \]

where \( m^* \) is the electron effective mass, \( E \) is the total electron energy, \( V \) is the electron potential energy, and \( \hbar \) is Planck's constant divided by \( 2\pi \). Thus, the electron-wave phase refractive index \( n_e \) is proportional to the square root of the
product of the effective mass and the kinetic energy. That is,

\[ n_s \propto \left[ \frac{m^*}{(E-V)} \right]^{1/2}. \] (2)

Amplitude effects such as transmissivity and reflectivity may be described in terms of the wave-function amplitude for an electron wave or in terms of the electric field amplitude for a transverse electric (TE) electromagnetic optical wave incident upon a boundary between dielectrics. Continuity of the wave function across a potential energy boundary is analogous to the continuity of the tangential component of the electric field across a boundary between dielectrics. Similarly, conservation of electron probability current normal to a potential energy boundary is analogous to conservation of power flow normal to a boundary between dielectrics. Using these, the electron-wave and electromagnetic-wave reflectivities and transmissivities are made equivalent by introducing, in quantum mechanics, an analogous index of refraction or characteristic impedance. Thus, the electron-wave amplitude refractive index \( n_s \) is proportional to the square root of the ratio of the kinetic energy to the effective mass, and so

\[ n_s \propto \left[ \frac{(E-V)}{m^*} \right]^{1/2}. \] (3)

The amplitude transmissivity \( t \) of an electron wave normally incident upon a boundary between semiconductors 1 and 2 is given by

\[ t = \frac{2n_{a1}}{(n_{a1} + n_{a2})}. \] (4)

The amplitude reflectivity \( r \) is similarly given by

\[ r = \frac{(n_{a1} - n_{a2})}{(n_{a1} + n_{a2})}. \] (5)

The electron current density in quantum mechanics is analogous to Poynting vector power in electromagnetics. Thus, the electron current transmitted \( T \) is given by

\[ T = \frac{(n_{a2}/n_{a1})}{2} t^2. \] (6)

and the electron current reflected \( R \) is given by

\[ R = r^2. \] (7)

In the above expressions, only dimensionless ratios of the electron-wave amplitude refractive indices occur. Both types of electron-wave refractive indices exhibit normal dispersion. That is, they increase with decreasing wavelength.

III. IMPEDANCE TRANSFORMERS

A. General

In microwave engineering, impedance transformation between transmission lines and waveguiding systems of differing characteristic impedance can be achieved with a variety of devices including quarter-wave transformers, double-stub tuners, triple-stub tuners, slot lines, and \( E \)-plane-H-plane tuners. Quarter-wave transformers may also be used for unguided electromagnetic plane waves and unguided electron plane waves. In electromagnetic optics they are simply thin-film antireflection coatings. Each layer is a dielectric with a thickness of quarter of a wavelength (as measured in that material) at the pass frequency. In electron-wave optics in semiconductors, quarter-wave transformers are merely layers of semiconductor material with thicknesses that are a quarter of an electron wavelength (as measured in that material) at the pass frequency. Such a semicon-

![FIG. 1. Electron-wave quarter-wavelength impedance transformer shown in Fig. 1 in terms of its conduction-band edge (labeled \( \Gamma \) for the case of the conduction-band minimum occurring at the center of the Brillouin zone). If impedance matching is required only over a narrow band of energies, then a single section (layer) impedance transformer may be used. In this situation the transformer acts as a narrow-band filter. If impedance matching is required over a broad range of energies, then multiple layers will be needed. The energy range covered increases with increasing number of layers. The impedance transformer of Fig. 1 is a general quarter-wave device consisting of \( M \) layers. The wave-function reflectivity of an electron wave in region 0 incident upon a boundary with a semi-infinite region 1 is designated \( r_0 \). Similarly the reflectivity of an electron wave in region \( i \) incident upon a boundary with a semi-infinite region \( i + 1 \) is designated \( r_i \). The physical thickness of the \( i \)th layer is denoted \( d_i \). In practical impedance transformers, the reflectivities \( r_i \) are relatively small. In this situation, the total electron-impedance reflection coefficient \( r \) for the entire impedance transformer is given by the sum of the first-order reflected waves. Thus

\[ r = r_0 + r_1 \exp(-2jkd_1) + r_2 \exp(-2jkd_2) + \ldots. \] (8)

Since all sections (layers) have the same electron optical thickness, then \( k_d = k_d = \cdots = k_d = \theta \). For quarter-wavelength transformers, this means \( \theta = \pi/2 \) at the pass energy. Furthermore, for a symmetrical transformer, \( r_0 = r_M, r_1 = r_{M-1}, \ldots \). For this common case, Eq. (8) may be rewritten as

\[ r = 2 \exp(-jM\theta) \left[ r_0 \cos(M\theta) + r_1 \cos((M-2)\theta) + \ldots + r_M \cos((M-2M)\theta) \right]. \] (9)

By appropriate selection of the reflectivities \( r_0, r_1, \ldots, r_M \), a wide variety of passband characteristics can be obtained.

B. Maximally flat transformers

A maximally flat passband is characterized by the first \( M - 1 \) derivatives of the total electron amplitude transmission coefficient \( t \) with respect to the electron energy being
zero at the pass energy. Maximally flat transformers are also called Butterworth or binomial transformers. The maximally flat characteristic is obtained when

$$r = 2^{-M} n_{a0} - n_{aM+1} \sum_{j=0}^{M} C_j^M \exp(-j2\theta),$$

(10)

where $C_j^M$ are the binomial coefficients given by

$$C_j^M = M!/[j!(M-j)!].$$

(11)

Comparing Eq. (10) with Eq. (8) yields

$$r_i = 2^{-M} \left[ (n_{a0} - n_{aM+1})/(n_{a0} + n_{aM+1}) \right] C_j^M.$$

(12)

Since the reflectivities $r_i = (n_{a0} - n_{aM+1})/(n_{a0} + n_{aM+1})$ are small, then $\ln(n_{a0}/n_{aM+1}) \approx 2r_i$. Thus

$$\ln(n_{a0}/n_{aM+1}) = 2^{-M} C_j^M \ln(n_{a0}/n_{aM+1}),$$

(13)

providing a relationship between the refractive indices of the various layers.

C. Equal-ripple transformer

In an equal-ripple impedance transformer, the electron amplitude reflectivity varies in an oscillatory fashion between 0 and $r_m$, where $r_m$ is a specified maximum reflectivity in the passband. This is obtained when the reflection coefficient is given by

$$r = r_m \exp(-jM\theta) T_M \sec \theta_m \cos \theta,$$

(14)

where $T_M(x)$ are Chebyshev polynomials and are given by

$$T_1(x) = x, \quad T_2(x) = 2x^2 - 1, \quad T_3(x) = 4x^3 - 3x, \quad T_4(x) = 8x^4 - 8x^2 + 1, \ldots, T_M(x) = 2xT_{M-1}(x) - T_{M-2}(x).$$

The quantity $\theta_m$ is the normalized half-bandwidth parameter given by $\theta_m = (\pi/2) (E_m/E_p)$, where $E_m$ is the lowest electron energy at which the reflectivity equals $r_m$, and $E_p$ is the pass energy. For a specified value of the maximum reflectivity $r_m$, the quantity $\theta_m$ may be determined by solving the polynomial equation

$$T_M \sec \theta_m = (1/r_m) (n_{a0} - n_{aM+1})/(n_{a0} + n_{aM+1}).$$

(15)

The electron amplitude refractive indices may then be determined by equating Eqs. (9) and (14). This gives a series of equations by then equating like terms in $\cos(i\theta)$. Each of these equations gives one of the reflectivities $r_{m1}, r_{m2}, \ldots$, and thus allows the amplitude refractive indices to be calculated as illustrated in Sec. IV.

IV. OPTIMUM SEMICONDUCTOR IMPEDANCE TRANSFORMERS

A. Configuration of semiconductors

The interconnection of GaAs and Ga$_{0.8}$Al$_{0.2}$As with several types of impedance transformers will be treated for illustration. This might correspond to electron flow from the base (GaAs) to the collector (Ga$_{0.8}$Al$_{0.2}$As) in a bipolar junction transistor as mentioned previously. The Ga$_{1-x}$Al$_x$As material system is perhaps the most advanced of current semiconductor superlattice material systems. For these alloys, all compositions are lattice matched. The material is a direct gap semiconductor for 0 < x < 0.45 with the conduction-band minimum occurring at the center of the Brillouin zone (Γ point). The electron potential energy of the conduction band $V_i$ is given by

$$V_i = A x_i, \quad i = 0, 1, \ldots, M + 1,$$

(16)

and the electron mass is given by

$$m^* = (B + C x_i)m_0, \quad i = 0, 1, \ldots, M + 1,$$

(17)

where $A$, $B$, and $C$ are constants and $m_0$ is the free-electron mass. For Ga$_{1-x}$Al$_x$As, $A = 773.14$ meV (using a 62% conduction-band offset), $B = 0.067$, and $C = 0.083$. In Fig. 1, region 0 would be GaAs and region $M + 1$ would be Ga$_{0.8}$Al$_{0.2}$As. As also shown in Fig. 1, the conduction-band potential energy edge increases monotonically from GaAs to Ga$_{0.8}$Al$_{0.2}$As through the transformer.

The next higher conduction-band minimum in Ga$_{1-x}$Al$_x$As for this range of compositions occurs along the [111] directions at the edge of the Brillouin zone (L point). The electron potential energy at the bottom of the L band, $V_{Li}$, is given by

$$V_{Li} = D + F x_i, \quad i = 0, 1, \ldots, M + 1,$$

(18)

relative to the Γ conduction-band edge for GaAs ($x = 0$). The quantities $D$ and $F$ are constants given by $D = 284$ meV and $F = 168.14$ meV for Ga$_{1-x}$Al$_x$As. The corresponding curve appears in Fig. 1 and is labeled L. To support ballistic transport by ensuring that intervalley scattering does not occur through phonon absorption, the pass electron energy will be arbitrarily set to be 50 meV below the L band of GaAs. The pass electron energy is thus $E_p = 234$ meV. As the aluminum concentration increases through the transformer, the energy separation $V_{Li} - E_p$ further increases until it is 83.6 meV in the Ga$_{0.8}$Al$_{0.2}$As as depicted in Fig. 1.

Since the reference for the electron amplitude refractive index is arbitrary, it will be taken to be unity in the Ga$_{0.8}$Al$_{0.2}$As since this is the low index side of the impedance transformer. That is $n_{aM+1} = 1$. The amplitude index of any other region is then given by

$$n_{ai} = [(E_p - V_i)m_{aM+1}^* / (E_p - V_{Mi})m_{aM}^*]^{1/2} n_{aM+1},$$

(19)

Therefore in the $i = 0$ GaAs region, the amplitude index is $n_{a0} = 1.917 962$.

The electron wavelength in any of the regions determines the electron optical phase thickness of that region. The electron wavelength for the pass energy in the $i$th region is given by

$$\lambda_p = h/[2m_{aM}^*(E_p - V_{Mi})]^{1/2},$$

(20)

where $\lambda_p$ is the pass wavelength as measured in the $i$th region and $h$ is Planck's constant.

B. Optimum single-layer transformer

For the single-layer narrow-band impedance transformer, maximally flat and equal-ripple design procedures both
give the same result. It is the well-known single-layer antireflection coating result that the amplitude index should be

\[ n_{a1} = (n_{o0} n_{aM+1})^{1/2}. \]  

(21)

Thus, for the present case \( n_{a1} = 1.384\,905 \). The composition of this layer may be determined by substituting Eqs. (16) and (17) into Eq. (19) and solving for \( x_i \). The result is

\[ x_i = (E_p - B G m_0) / (A + C G m_0), \]

(22)

where \( G = [(E_p - V_{M+1}) / m_{M+1}^*] (n_{o0} / n_{aM+1})^2 \). With the composition known, the potential energy \( V_i \) and the effective mass \( m_i^* \) are determined through Eqs. (16) and (17), respectively. The required physical thickness is then

\[ d_i = \lambda_\mu / 4 = h / [32 m_i^* (E_p - V_i)]^{1/2}, \]

(23)
to produce a quarter-wavelength thickness. For the present case, Eqs. (22), (16), (17), and (23) yield \( x_1 = 0.1212 \), \( V_1 = 93.682 \) meV, \( m_1^* = 0.07706 m_0 \), and \( d_1 = 2.9482 \) nm, respectively. The resulting overall electron current transmittance \( T \) for this optimum single-layer impedance transformer is shown in Fig. 2. The transmittance was calculated by the procedure described in Ref. 6. At the pass kinetic energy \( K_E = 79.372 \) meV, the electron current transmittance is unity as it should be.

C. Optimum three-layer maximally flat transformer

The design of a three-layer maximally flat impedance transformer to connect GaAs and \( Ge/A 1_{0.2}As \) for a broader range of electron energies near the passband is presented in this section. For a three-layer device, \( M = 3 \). For \( M = 3 \) and \( i = 0 \), Eq. (13) becomes \( \ln(n_{o0} / n_{a1}) = \ln(n_{o0} / n_{aM+1}) \) and so

\[ n_{a1} = n_{o0}^{3/4} n_{aM+1}^{1/4}. \]

(24)

For the present case \( n_{o0} = 1.917\,962 \) and \( n_{aM+1} = 1 \) and so \( n_{a1} = 1.768\,011 \). The required composition is obtained from

\[ (E_p - B G m_0) / (A + C G m_0), \]

using Eq. (25) this may be rewritten as

\[ n_{a2} = n_{o0}^{3/4} n_{a3}^{1/4}. \]

(25)

Thus the center section of a maximally flat transformer is the same as that for the single-layer transformer [see Eq. (21)]. Therefore \( n_{a2} = 1.384\,905 \), \( x_2 = 0.1212 \), and \( d_2 = 2.9482 \) nm. For \( M = 3 \) and \( i = 2 \), Eq. (13) gives \( n_{a3} = n_{o0}^{3/4} n_{a2}^{1/4} \). Using Eq. (25) this may be written as

\[ n_{a3} = n_{o0}^{3/4} n_{aM+1}^{1/4}. \]

(26)

For the present case, \( n_{a3} = 1.084\,813 \). The required composition and physical thickness are \( x_3 = 0.1838 \) and \( d_3 = 3.5259 \) nm. A summary of the parameters of this impedance transformer is given in Table I. The overall electron current transmittance \( T \) for this optimum three-layer maximally flat impedance transformer is shown in Fig. 2. Again at the pass kinetic energy \( K_E = 79.372 \) eV, the electron current transmittance \( T \) is unity. The passband, however, is much broader than that for the one-layer transformer (Fig. 2). The maximally flat character of the passband is also evident in the transmittance shown in Fig. 3.

D. Optimum three-layer equal-ripple transformer

The design of a three-layer equal-ripple impedance transformer to connect GaAs and \( Ge/A 1_{0.2}As \) is presented in this section. For a maximum electron current reflectivity within the passband of 2%, the amplitude reflectivity is \( r_m = (0.02)^{1/2} = 0.1414 \). For a three-section device \( M = 3 \). Using the polynomial representation for \( T_3(x) \), Eq. (15) may be written as

\[ 4x^2 - 3x = (1/r_m)(n_{o0} - n_{aM+1}) / (n_{o0} + n_{aM+1}) = 2.224\,488, \]

(27)

TABLE I. Design parameters of optimum three-layer maximally flat and three-layer equal-ripple electron impedance transformers for connecting GaAs to Ga\(_{0.8}\)Al\(_{0.2}\)As.

<table>
<thead>
<tr>
<th>Layer</th>
<th>Amplitude refractive index</th>
<th>Aluminum composition</th>
<th>Thickness ( d_i ) (nm)</th>
<th>Number of monolayers ( p_i )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.9180</td>
<td>0.0000</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>1</td>
<td>1.7680</td>
<td>0.0345</td>
<td>2.5472</td>
<td>9.0113</td>
</tr>
<tr>
<td>2</td>
<td>1.3849</td>
<td>0.1212</td>
<td>2.9482</td>
<td>10.4300</td>
</tr>
<tr>
<td>3</td>
<td>1.0848</td>
<td>0.1838</td>
<td>3.5259</td>
<td>12.4739</td>
</tr>
<tr>
<td>4</td>
<td>1.0000</td>
<td>0.2000</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

Maximally flat

<table>
<thead>
<tr>
<th>Layer</th>
<th>Amplitude refractive index</th>
<th>Aluminum composition</th>
<th>Thickness ( d_i ) (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.9180</td>
<td>0.0000</td>
<td>...</td>
</tr>
<tr>
<td>1</td>
<td>1.7680</td>
<td>0.0345</td>
<td>2.5472</td>
</tr>
<tr>
<td>2</td>
<td>1.3849</td>
<td>0.1212</td>
<td>2.9482</td>
</tr>
<tr>
<td>3</td>
<td>1.2440</td>
<td>0.1515</td>
<td>3.1786</td>
</tr>
<tr>
<td>4</td>
<td>1.0000</td>
<td>0.2000</td>
<td>...</td>
</tr>
</tbody>
</table>

Equal ripple

<table>
<thead>
<tr>
<th>Layer</th>
<th>Amplitude refractive index</th>
<th>Thickness ( d_i ) (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.9180</td>
<td>...</td>
</tr>
<tr>
<td>1</td>
<td>1.5738</td>
<td>2.7180</td>
</tr>
<tr>
<td>2</td>
<td>1.3992</td>
<td>2.9280</td>
</tr>
<tr>
<td>3</td>
<td>1.2440</td>
<td>3.1786</td>
</tr>
<tr>
<td>4</td>
<td>1.0000</td>
<td>2.0000</td>
</tr>
</tbody>
</table>
for a transformer between GaAs and Ga$_{0.4}$Al$_{0.6}$As, where $x = \sec \theta_m$. Solving this cubic equation gives $x = 1.117068$ and $x = -0.558534 \pm 0.431139$. Only the real solution is meaningful in this design procedure. Equating Eqs. (9) and (14) for the case of $M = 3$ gives

$$2(r_0 \cos 3\theta + r_1 \cos \theta) = r_m [\sec^3 \theta_m (\cos 3\theta + 3 \cos \theta) - 3 \sec \theta_m \cos \theta].$$

Equating like terms in $\cos 3\theta$ produces $2r_0 = r_m \sec^3 \theta_m$. Solving this gives $r_0 = 0.098565$. Equating like terms in $\cos \theta$ similarly produces $2r_1 = 3r_m (\sec^3 \theta_m - \sec \theta_m)$ and solving this yields $r_1 = 0.058730$. From Eq. (5), the amplitude refractive index in region $i + 1$ is

$$n_{a,i+1} = \left( \frac{1 - r_i}{1 + r_i} \right) n_{a,i}.$$  

For the present configuration of materials, $n_{a,0} = 1.917962$ and so for $i = 0$, Eq. (29) becomes

$$n_{a,1} = \left[ \frac{(1 - r_0)/(1 + r_0)}{n_{a,0}} \right] n_{a,0} = 1.573796.$$  

The required composition is obtained from Eq. (22) and is $x_1 = 0.0789$. The needed physical thickness is found from Eq. (23) and is $d_1 = 2.717966$ nm. For $i = 1$, Eq. (29) gives

$$n_{a,2} = \left( \frac{1 - r_1}{1 + r_1} \right) n_{a,1} = 1.399193.$$  

The corresponding composition is $x_2 = 0.1180$ and the corresponding thickness is $d_2 = 2.9280$. Since the transformer is symmetrical, then $r_2 = r_1$ and thus the third layer has

$$n_{a,3} = \left( \frac{1 - r_2}{1 + r_2} \right) n_{a,2} = 1.243961.$$  

The corresponding composition and thickness are $x_3 = 0.1515$ and $d_3 = 3.1786$ nm. A summary of the parameters of this impedance transformer is also given in Table I. The resulting electron current transmittance for this optimum Chebyshev impedance transformer is shown in Fig. 4. The equal-ripple character of the passband is evident. The passband of a multiple-section equal-ripple transformer is broader than that of the corresponding multiple-section maximally flat transformer. Furthermore, both multiple-layer transformers have much broader bandwidths than that of the narrow-band one-layer transformer (Fig. 2). Upon detailed inspection, the center layer in the Chebyshev transformer should have been $n_{a,2} = (n_{a,0} n_{a,M-1})^{1/2}$. That is, it should have been 1.3849 rather than 1.3992 as calculated. This is due to minor approximations that are commonly made in the design of equal-ripple transformers. An exact, but more complicated, procedure is available but the additional effort required to implement it produces only a negligible change in the resultant design.

V. MONOLAYER-CONSTRAINED SEMICONDUCTOR IMPEDANCE TRANSFORMER

A. Monolayer thickness constraint

The thicknesses of the layers in the optimum impedance transformers as designed in Sec. IV can have any of a continuum of values. In reality the thicknesses of all layers must be integer multiples of the monolayer thickness for that region. For lattice-matched Ga$_{1-x}$Al$_x$As the monolayer thickness $s$ is the same for any composition $x$. For this material system $s = 0.282665$ nm.

To construct a quarter-wavelength impedance transformer, the thicknesses of the layers must also be an odd multiple of a quarter wavelength as measured in that layer for the pass energy $E_p$. That is

$$d_i = p_i s = (2q_i - 1) \lambda_p/4, \quad i = 1, 2, ..., (30)$$

where $p_i$ is the integer number of monolayers for the $i$th region and $q_i$ is a positive integer ($q_i = 1, 2, 3, ...$). Substituting Eqs. (16) and (17) into Eq. (20) and then substituting the resulting $\lambda_p$ into Eq. (30) gives the following quadratic
The solution for the composition \( x_i \) is

\[
x_i = \left[ -b \pm \left( b^2 - 4ac \right)^{1/2} \right]/2a,
\]

where \( a = AC \), \( b = AB - CE_p \), and \( c_i = (h^2/32m_0) \left[ (2q_i - 1)^2/p_i^2 \right] - BE_p \).

In order to design an impedance transformer, solutions \( x_i \) must be found in the usable composition range that correspond closely to the optimum values calculated by the procedures given in Sec. IV. To minimize the total thickness of the filter, \( q_i \) is initially set equal to unity. Then Eq. (32) is repetitively evaluated for \( p_i = 1, 2, 3 \ldots \) until all of the positive real roots in the range \( 0 < x_i < x_{\text{max}} \) are found. Plots of aluminum composition \( x \) as a function of the number of monolayers are shown for various values of the pass kinetic energy \( KE_p \). For large values of \( KE_p \), there are only a few values of composition that correspond to thicknesses that are integer multiples of the monolayers. For smaller values of \( KE_p \), there are a greater number of thicknesses available that simultaneously satisfy the quarter-wavelength and monolayer constraints.

### B. Monolayer-constrained single-layer transformer

The monolayer thickness constraint is now applied to the design impedance transformers for a transition from GaAs to \( \text{Ga}_{0.45} \text{Al}_{0.55} \text{As} \). Using \( E_p = 234 \text{ meV} \), setting \( q_i = 1 \), and repetitively evaluating Eq. (32) for integer values of \( p_i \), yields within the range \( 0 < x_i < x_{\text{max}} \) the following aluminum concentrations: \( x_i = 0.0335 \) for \( p_i = 9 \), \( x_i = 0.1008 \) for \( p_i = 11 \), \( x_i = 0.1728 \) for \( p_i = 12 \), and \( x_i = 0.1944 \) for \( p_i = 13 \).

For a monolayer thickness constrained single-layer narrow-band impedance transformer, the middle value of aluminum composition may be used to approximate the optimum single-layer transformer. Therefore \( x_i = 0.1433 \). With the composition \( x_i \) and number of monolayers \( p_i \) known, the potential energy \( V_i \), effective mass \( m_i^* \), the amplitude refractive index \( n_i \), the thickness of the layer \( d_i \), and wavelength in the layer \( \lambda_i \) can be determined from Eqs. (16), (17), (19), (30), and (23), respectively. For the present monolayer-thickness-constrained \( (p_i = 11) \) single-layer impedance transformer, these quantities are \( V_i = 110.78 \text{ meV}, \)

\( m_i^* = 0.07889 \eta_{\infty} \), \( d_i = 3.1093 \eta_{\infty}, \)

\( n_i = 1.2826, \) and \( \lambda_i = 12.437 \eta_{\infty} \). The resulting overall electron current transmittance \( T \) for this monolayer thickness constrained impedance transformer is shown in Fig. 2 along with the transmittance for the optimum (without monolayer constraint) single-layer transformer. The passband characteristics have similar functional forms. However, at the pass kinetic energy \( KE_p = 79.4 \text{ meV} \), the electron current transmittance is reduced from the optimum 100% to 99.41%.

### C. Monolayer-constrained three-layer maximally flat transformer

For an optimum three-layer maximally flat impedance transformer as described in Sec. IV C and summarized in Table I, the needed values of aluminum concentration are near the lower end, near the middle, and near the upper end of the range \( 0 < x < x_{\text{max}} \). The optimum three-layer maximally flat impedance transformer may be approximated by selecting the aluminum compositions of \( x_1 = 0.0335 \) for \( p_1 = 9 \), \( x_2 = 0.1433 \) for \( p_2 = 11 \), and \( x_3 = 0.1944 \) for \( p_3 = 13 \). As before, a knowledge of the composition \( x_i \) and number of monolayers \( p_i \) allows the potential energy \( V_i \), effective mass \( m_i^* \), amplitude refractive index \( n_i \), thickness of the layer \( d_i \), and wavelength in the layer \( \lambda_i \) to be calculated. For the present monolayer-thickness-constrained \( (p_1 = 9, p_2 = 11, p_3 = 13) \) three-layer maximally flat impedance...
TABLE II. Design parameters of monolayer-constrained three-layer maximally flat and three-layer equal-ripple electron impedance transformers for connecting GaAs to Ga$_{0.8}$Al$_{0.2}$As. The design pass energy is $E_p = 234$ meV and the design output kinetic energy is $KE_p = 79.372$ meV.

<table>
<thead>
<tr>
<th>Layer $i$</th>
<th>Amplitude refractive index $n_{ai}$</th>
<th>Aluminum composition $x_i$</th>
<th>Thickness $d_i$ (nm)</th>
<th>Number of monolayers $p_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximal flat</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1.9180</td>
<td>0.0000</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>1</td>
<td>1.7723</td>
<td>0.0335</td>
<td>2.5440</td>
<td>9</td>
</tr>
<tr>
<td>2</td>
<td>1.2826</td>
<td>0.1433</td>
<td>3.1093</td>
<td>11</td>
</tr>
<tr>
<td>3</td>
<td>1.0299</td>
<td>0.1944</td>
<td>3.6746</td>
<td>13</td>
</tr>
<tr>
<td>4</td>
<td>1.0000</td>
<td>0.2000</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Equal ripple</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1.9180</td>
<td>0.0000</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>1</td>
<td>1.4769</td>
<td>0.1088</td>
<td>2.8267</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>1.2826</td>
<td>0.1433</td>
<td>3.1093</td>
<td>11</td>
</tr>
<tr>
<td>3</td>
<td>1.1403</td>
<td>0.1728</td>
<td>3.3920</td>
<td>12</td>
</tr>
<tr>
<td>4</td>
<td>1.0000</td>
<td>0.2000</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

The quantitative analogies that have been previously established between electron-wave propagation in semiconductors and electromagnetic-wave propagation in dielectrics have been used to design impedance transformers for ballistic (collisionless) electron waves traveling between dissimilar energy-gap semiconductors. These devices have been designed as a series of quarter (electron) wavelength quantum well layers in the form of compositional superlattices. Procedures for designing maximally flat (Butterworth) and equal-ripple (Chebyshev) impedance transformers of arbitrary spectral bandwidth have been presented. An alternate method to design maximally flat and equal-ripple transformers is through the use of precalculated tables of refractive indices' for the tables contain the ratio of beginning and ending refractive indices that is needed.

Additional constraints on the impedance transformer quantum well layers in practice are that (1) their compositions must be within the usable compositional range and that (2) their thicknesses must be integer multiples of a monolayer thickness. These constraints have been incorporated into the design procedure producing designs that approximate the optimum designs. However, even with these additional constraints, it has been shown that the general characteristics of the optimum design can be retained. While the transmittances presented in Figs. 2-4 have been calculated for electrons moving perpendicular to the boundaries, these results accurately apply to electrons with angles of incidence beyond $10^\circ$. Thus the results presented are rather insensitive to the electron angle of incidence.

The electron potential energy shown in Fig. 1 is idealized and does not include the effects of space charge in these undoped layers. The space charge which is a function of the current, will alter the potential energy (as calculated by the Poisson equation). This new potential energy will, in turn, alter the wave function (as calculated by the Schrödinger equation). Thus, in general, a self-consistent solution of these equations is needed, particularly at high current levels.

The designs of practical single-layer and three-layer transformers for connecting GaAs and Ga$_{0.8}$Al$_{0.2}$As have been presented. Although this well-developed, lattice-matched material system was used as an example, the procedures presented would also apply to other material systems. Dispersion effects in electromagnetic (e.g., microwave) impedance transformers are relatively small compared to the dispersion effects in semiconductor electron-wave devices. For the present case of impedance transformers this causes a decrease in the electron current transmittance at energies...
below the pass energy and an increase in the transmittance at energies above the pass energy.

ACKNOWLEDGMENTS

This research was sponsored in part by a grant from the Joint Services Electronics Program under Contract No. DAAL-03-87-K-0059. One of us (K.F.B.) was supported in part by a Presidential Young Investigator award from the National Science Foundation and another (E.N.G.) by a Research Initiation award from the National Science Foundation.

21. L. Young, IRE Trans. Microwave Theory Tech. 8, 243 (1960).
Electron wave diffraction by semiconductor gratings: Rigorous analysis and design parameters

Gregory N. Henderson, Elias N. Glytsis, and Thomas K. Gaylord
School of Electrical Engineering and Microelectronics Research Center, Georgia Institute of Technology, Atlanta, Georgia 30332-0250

(Received 21 March 1991; accepted for publication 22 April 1991)

An exact rigorous coupled-wave analysis has been developed to model ballistic electron wave diffraction by gratings with periodic effective mass and/or potential energy variations. Design expressions have been derived to calculate diffracted angles, to identify evanescent orders, and to identify the Bragg condition. Design expressions for Bragg regime (up to 100% diffraction efficiency in a single order) and Raman–Nath regime (high diffraction efficiency divided among multiple orders) diffraction are presented along with example Ga$_1-x$Al$_x$As grating designs. Design procedures for ballistic electron switches, multiplexers, spectrometers, and electron waveguide couplers are described.

Semiconductor growth techniques such as molecular beam epitaxy and metalorganic chemical vapor deposition together with nanolithography have been refined so that semiconductor structures can be fabricated with device dimensions on the order of electron wavelengths. Using these techniques, devices exhibiting ballistic (collisionless) electron transport have been fabricated. These ballistic electrons are quantum-mechanical deBroglie waves and can be reflected, refracted, diffracted, guided, and interfered like optical waves in dielectrics.

The growth procedures mentioned above have been used to produce periodic structures such as superlattices, arrays of quantum wires, and periodic gate structures, whose periods are on the order of electron wavelengths in semiconductors. These periodic structures can be used as ballistic electron wave diffractive devices analogous to gratings used in integrated optics. Previous analyses of electron wave diffraction in semiconductors have been limited to the use of a two-mode amplitude transmittance analysis (to model diffraction from a symmetric rectangular potential-energy grating), and a Fraunhofer analysis (to model single-slit diffraction). The rigorous coupled-wave analysis (RCWA) presented in this letter can model ballistic electron wave diffraction from a general effective mass and/or potential energy grating. The Floquet theorem (Bloch theorem) for waves in a periodic medium states that the electron wave inside the grating can be described as

$$\psi = \sum_{m} U_m(z) \exp(-i \mathbf{k}_m \cdot \mathbf{r}),$$

where $\psi$ is the electron wave amplitude, $\mathbf{k}$ is Planck's constant divided by $2\pi$, and $m$ and $V$ are the periodic effective mass and potential energy in the grating (periodic in the grating vector $\mathbf{K}$). The Floquet theorem (Bloch theorem) for waves in a periodic medium states that the electron wave inside the grating can be described as

$$\psi = \sum_{m} U_m(z) \exp(-i \mathbf{k}_m \cdot \mathbf{r}),$$

Likewise, the electron wave in the input region is given as

$$\psi = \sum_{m} U_m(z) \exp(-i \mathbf{k}_m \cdot \mathbf{r}),$$

$\mathbf{k}_m$ and $\mathbf{R}_m$ are the wave vector and amplitude of the $m$th backward-diffracted order, respectively. The electron wave in the output region is

$$\psi = \sum_{m} U_m(z) \exp(-i \mathbf{k}_m \cdot \mathbf{r}),$$

$\mathbf{k}_m$ and $\mathbf{T}_m$ are the wave vector and amplitude of the $m$th forward-diffracted order, respectively.

An electron wave diffraction grating is shown in Fig. 1 with grating vector $\mathbf{K}$, period $A (|\mathbf{K}| = 2\pi/A)$, slant angle $\phi$, and thickness $d$. The input and output regions are described by effective masses $m^*$ and $m^*_H$, and potential energies $V_1$ and $V_{mil}$, respectively. An electron wave of energy $E$ and wave vector $\mathbf{k'}$ is incident from the input region at an angle $\theta'$ and is diffracted into forward- and backward-diffracted orders. The Hamiltonian used for the electron wave inside the grating is given as

$$\frac{-\hbar^2}{2m^*(r)} \nabla^2 \psi(r) + V_{mil}(r) \psi(r) = (E - V(r)) \psi(r).$$

FIG. 1. Schematic diagram of an electron wave grating. A ballistic electron wave is incident upon the grating from region 1 and is diffracted into backward- and forward-diffracted orders in regions 1 and 3, respectively.
By employing phase matching of the input and output waves to the space harmonics in the grating region \([U(z)]\), one finds the grating equations to be

\[
(2m^2_p(E - V_i))^{1/2} \sin \theta - (2\pi h/A) \sin \phi
= r_n (2m^2_p(E - V_a))^{1/2} \sin \theta,
\]

where \(n = 1, 3, r_1 = -1, r_3 = 1\), and \(\theta_i, \theta_0, \theta_1, \theta_2, \theta_3\) are the angles of propagation of the \(n\)th, \(r_1\)th, \(r_2\)th, \(r_3\)th, and \(r_3\)th backward- and forward-diffracted orders, respectively. Those space harmonics that cannot phase match to propagating orders are evanescent (\(k, \text{and} \theta\), are imaginary) or "cutoff" orders. The directions of the propagating diffracted orders are shown schematically in Fig. 1.

To calculate the amplitudes of the diffracted waves, a RCWA similar to that employed in electromagnetism is used. The effective mass and potential energy in the grating region are expanded as Fourier series in \(k\) and substituted into Eq. (1). In order to allow for discontinuities in the effective mass (as in GaAlAs superlattices), Eq. (1) is separated into two first-order vector equations where the derivatives of the effective mass have been eliminated. These equations are solved using the state-variable approach. The number of orders is chosen to ensure convergence of the diffraction efficiencies. Through the application of the boundary conditions (continuity of \(\psi\) and \(\nabla \psi \cdot \hat{z} / m^2_p\)), the \(R_i\)'s and \(T_i\)'s are calculated. By analogy to electromagnetics, the diffraction efficiency for electron wave gratings is defined as the ratio of the \(z\) component of the probability current carried by the \(n\)th diffracted order to the \(z\) component of the probability current of the incident wave. Using the probability current \(J = \hat{z} (\nabla \psi^* - \psi^* \nabla \psi) / 2m^2_p\), one finds the diffraction efficiencies in regions I and III to be

\[
D_E^{(1)} = \left( \cos \theta_0 / \cos \theta \right) R_1^2, \quad \text{and} \quad D_E^{(III)} = \left( \sqrt{m^2_p(E - V_{III})} \cos \theta_0 / \sqrt{m^2_p(E - V_I)} \cos \theta \right) T_1^3,
\]

respectively.

The RCWA is a robust procedure that can be applied to arbitrary grating profiles. For design, however, it is useful to have analytical expressions that describe the diffraction efficiencies in terms of the grating parameters. In electromagnetics, the Bragg regime and the Raman-Nath regime are often used for grating design. In the Bragg regime, the angle of incidence is adjusted so that the first diffracted wave vector inside the grating, \(\theta_i\), lies on the allowed wave vector surface inside the grating, yielding

\[
\sqrt{2m^2_p(E - V_{III})} = \pi h / A \cos (\theta_0 - \phi),
\]

where \(m^2_p\) and \(V_{III}\) are the average effective mass and average potential energy inside the grating, and \(\theta\) is the refracted wave vector angle inside the grating. When an electron wave is incident at the Bragg condition \([\text{i.e., it satisfies Eq. (3)}]\), the diffraction is in the Bragg regime if the grating thickness parameter \(Q\) and the grating strength parameter \(\gamma\) satisfy the relations \(\rho_b = Q / 2\gamma > 1\), where \(\rho_b\) is the Bragg regime parameter. \(Q\) is the effective thickness of the grating given by \(Q = 4\pi h / \sqrt{2m^2_p(E - V_{III})} A^2 \cos \theta\). \(\gamma\) is the coefficient that couples the power between the \(i = 0\) and the \(i = +1\) order when only two waves are included in the coupled-wave analysis. For a sinusoidal effective mass variation \(\gamma = (m^2_p(E - V_I) + m^2_p(E - V_{III}) \cos (2\theta) d / \sqrt{2m^2_p(E - V_{III})} \pi h / A \cos \theta)\), where \(m^2_p\) and \(V_e\) are the amplitudes of the sinusoidal effective mass and potential energy variation. The diffraction efficiency of the first diffracted order for a sinusoidal grating in the Bragg regime is found to be \(D_E^{(III)} = \sin^2(\gamma)\). From this expression for \(D_E^{(III)}\), it is apparent that 100% diffraction efficiency is possible in the Bragg regime. Figure 2(a) shows both the approximate analysis and the RCWA for a Bragg regime grating constructed from a periodic aluminum composition variation in GaAlAs. For this grating, \(\rho_b = 2.60\). As \(\rho_b\) increases, the approximate analysis approaches the exact analysis, becoming indistinguishable around \(\rho_b = 20.0\).

In the Raman-Nath regime, power is diffracted into multiple orders, and can be distributed almost evenly among many orders. This occurs when the Raman-Nath regime parameter \(\rho_{RN} = Q\gamma < 1\). In this regime, the approximate solution to the coupled-wave equations is \(D_E^{(III)} = J_1(2\gamma)\), where \(J_1\) is an \(n\)th order ordinary Bessel function of the first kind. Figure 2(b) shows the approximate analysis and the RCWA (for \(i = \pm 1\)) for a Raman-Nath regime grating. This grating leaves the Raman-Nath regime at \(d = 18\) nm, which explains the failure of the approximate analysis for \(d > 20\) nm.

The electrons incident upon the grating will have a
distribution in both angle of incidence and energy. Therefore, it is important to understand the factors that affect the angular and energy selectivity of gratings. The angular (energy) selectivity describes the variation in the diffraction efficiency as the angle of incidence (energy) is changed. Strong angular (energy) selectivity refers to the case when the diffraction efficiency is appreciable for only a narrow corridor of angles of incidence (energies). It has been shown in electromagnetics that the angular and energy selectivity of a grating is a function of the ratio \( d/A \). A "thick" grating \( (d/A > 10) \) exhibits strong angular and energy selectivity. As a grating becomes "thinner" \( (d/A \) is decreased), the grating angular and energy selectivity become weaker. This behavior is demonstrated in Figs. 3(a) and 3(b) for three grating thicknesses. The suppression of the angular selectivity side lobes for angles less than 6.4° \( [\text{Fig. } 3(a)] \) is due to the fact that the \( i = +1 \) order becomes evanescent or "cutoff" at this angle of incidence \( [\text{see Eq. } (2)] \). Similarly, the rapid drop in diffraction efficiency for kinetic energies around \( E - V_1 = 0.025 \text{ eV} \) \( [\text{Fig. } 3(b)] \) is due to the fact that the \( i = +1 \) order becomes evanescent at this energy. The sharp discontinuity in the diffraction efficiency for \( E - V_1 = 0.102 \text{ eV} \) \( [\text{Fig. } 3(b)] \) is analogous to a Wood's anomaly in electromagnetic optics and is due to the "cutting on" of the \( i = +2 \) order at this energy.

With this information, one can design electron wave gratings for many applications such as switches, energy multiplexers, energy spectrometers, and electron waveguide couplers. These devices could be invaluable in the development of future guided-wave electron devices. If one desires a device with nearly 100\% diffraction efficiency in a given order \( (\text{e.g., a } 1 \times 1 \text{ switch}) \) then the following design steps should be followed. First, the grating should be designed such that the electron wave is incident at the Bragg condition \( [\text{Eq. } (3)] \). Then, the grating parameters should be chosen such that \( \rho_{\text{eff}} > 1 \). Finally, the thickness and grating modulation should be chosen such that \( \gamma \) is a half-integer multiple of \( \pi \) so that \( \text{DE}_{\gamma} \approx 100\% \). For this case, the grating parameters can also be adjusted such that all orders other than \( i = 0 \) and \( i = +1 \) are cutoff \( [\text{as in Fig. } 2(b)] \). If one desires a grating with a large diffraction efficiency divided among multiple orders \( (\text{e.g., a } 1 \times N \text{ switch}) \), the following design steps should be followed.

First, the grating should be designed \( [\text{Eq. } (2)] \) to have many propagating forward-diffracted orders. Then, the grazing parameters should be chosen such that \( \rho_{\text{eff}} < 1 \). Finally, the argument of the Bessel functions \( (2\gamma) \) should be chosen to allow for a large diffraction efficiency in the desired orders. If one desires a grating with a large angular and energy selectivity \( (\text{e.g., an energy multiplexer}) \), then \( d/A \) should be made large \( (> 10) \). If one desires a grating with a low angular and energy selectivity \( (\text{e.g., a } 1 \times 1 \text{ switch with high diffracted current}) \) then \( d/A \) should be made small \( (< 1) \). After the grating has been designed using the above procedures, the diffraction can be determined using the RCWA to test the performance and fine-tune the grating design.

This research was supported in part by a grant from the Joint Services Electronics Program under contract No. DAAL-03-90-C-0004. One of us (G.N.H.) was supported by an Office of Naval Research Graduate Research Fellowship, and one of us (E.N.G.) was supported by a Research Initiation Award from the National Science Foundation.

Ballistic Electron Transport in Semiconductor Heterostructures and Its Analogies in Electromagnetic Propagation in General Dielectrics

GREGORY N. HENDERSON, STUDENT MEMBER, IEEE, THOMAS K. GAYLORD, FELLOW, IEEE, AND ELIAS N. GLYTSIS, SENIOR MEMBER, IEEE

Developments in growth and fabrication technology have produced semiconductor structures with ballistic (collisionless) transport lengths of over a micron. In this paper, a comprehensive set of analogies is established between ballistic electron wave propagation in semiconductors (arbitrary kinetic energy and effective mass) and electromagnetic propagation in general dielectrics (arbitrary permittivity and permeability). First, electromagnetic results for propagation in nonmagnetic dielectrics are generalized to describe propagation, reflection, and refraction in general dielectrics through the definition of separate phase and amplitude refractive indexes. The expressions for electron wave propagation, reflection, and refraction are then developed and are shown to have the same functional form as in electromagnetics, if analogous definitions of electron wave phase and amplitude refractive indexes are used. The reflectivity characteristics such as total internal reflection (critical angle) and zero reflectivity (Brewster angle) are analyzed as a function of material parameters for both general dielectrics and semiconductor materials. The critical angle and Brewster angle results are then applied to electron wave propagation in Ga$_{1-x}$Al$_x$As where it is shown that all interfaces in this material will have both a critical angle and a Brewster angle due to differing effective masses across the interface. This is the first prediction of an electron wave Brewster angle in semiconductors.

1. INTRODUCTION

Semiconductor growth techniques such as molecular beam epitaxy (MBE) and metalorganic chemical vapor deposition (MOCVD) have been refined so that semiconductor structures can be grown with precise monolayer and compositional control [1]. This has produced semiconductor materials in which ballistic electron transport has been observed [2]-[4]. That is, the electrons traverse the sample as quantum mechanical plane waves experiencing no elastic or inelastic scattering events. Ballistic electrons can account for more than 50% of the current in small devices [3] and have been shown to be prominent even in the presence of elastic scattering [5]. Since ballistic electrons are quantum mechanical deBroglie waves, they can be reflected, refracted, diffracted, and interfered in a manner analogous to optical waves in dielectrics [6], [7]. Electron wave interference effects have been observed experimentally for electron energies below the barriers in double-barrier and multibarrier resonant tunneling devices [8]-[12] and for electron energies above the conduction band edges in GaAs/Ga$_{1-x}$Al$_x$As and In$_{1-x}$Ga$_x$As/In$_{1-x}$Al$_x$As heterostructures [13]-[16]. In addition, electron wave refraction has been experimentally demonstrated through the fabrication of electrostatic lenses [17], [18] and prisms [19] in a two-dimensional GaAs/Ga$_{1-x}$Al$_x$As electron gas.

In recent years, there has been a surge of interest in the development of analogies between electron wave (EW) propagation in semiconductors and electromagnetic wave (EMW) propagation in dielectrics [20]. Due to the fundamental differences between electrons and photons, however, it would seem difficult to draw quantitative, exact analogies between them. Specifically, as indicated in Fig. 1, electrons and photons differ in rest mass, charge, spin, particle specie (Fermion or Boson), and velocity. However, in the propagation of an EW in a charge-neutral crystal, where the electron is either in a single nondegenerate conduction band or in a degenerate conduction band with no zero-field spin splitting, the EW propagation is well described by the single-electron effective mass equation [21]. This equation is analogous to the Helmholtz equation for EMW propagation in dielectrics. Under these conditions, often referred to as the envelope-function approximation or the effective-mass approximation, exact, quantitative analogies can be drawn between EW and EMW propaga-
A number of conceptual and quantitative analogies have previously been developed. An analogy has been established between the nonmagnetic-dielectric Helmholtz equation and the time-independent Schrödinger equation for constant effective-mass conduction electrons [21], [22]. In addition, optical ray tracing methods for dielectrics have been used to describe ballistic electron refraction [17]–[19] and a Schrödinger equation has been derived for photons to relate constant effective-mass EW propagation to EMW propagation [23]. All of these analogies have been drawn to EMW propagation in nonmagnetic dielectrics. Such analogies are equivalent to neglecting effective mass changes across material interfaces. Separate EW phase and amplitude refractive indexes have previously been defined and used with the electromagnetic chain matrix method to describe EW propagation in semiconductor heterostructures with effective mass differences [6]. It is shown in the present work that the requirement of two EW refractive indexes is exactly analogous to EMW propagation in general dielectrics. Furthermore, recent experimental work has demonstrated that effective mass differences can have propagating effects on the propagation of electrons across material interfaces [24]–[26]. Ohno, Mendez, and Wang [24] have demonstrated that effective mass differences between the emitter and well of a resonant tunneling diode have dramatic effects on the resulting current-voltage characteristic of the diode. These effects were well explained by the authors as being due to variations in the number of electrons that can phase match to the state in the well as a function of the effective masses. Similar results have been demonstrated through the ultrahigh resolution of Ballistic-Electron-Emission Microscopy (BEEM) [25], [26] of surfaces. In these experiments, ballistic electrons tunnel through an air gap and probe the electrical properties of a metal-semiconductor interface. Bell and Kaiser have explained the ultrahigh resolution in these experiments (on the order of 1 nm [25]) in terms of the total internal reflection of the nonnormal incidence electrons at the metal-semiconductor interface [26]. The existence of such a small critical angle (≈ 3 deg for GaAs [26]) is due to the large effective mass difference between the metal and the semiconductor. It is straightforward to show that without the effective mass difference, the critical angle would be ≈ 52 deg and the resolution would be about 25 nm. These experimental results of Ohno et al. [24] and Bell et al. [25] are verification of the importance of including effective mass effects in the representation of ballistic electron transport.

In order to include these important effective mass effects, analogies in this paper are drawn to propagation in general dielectrics with arbitrary permittivity and permeability. A comprehensive set of analogies is established between EW propagation in semiconductor heterostructures and EMW propagation in general dielectrics. First, starting with Maxwell's equations, the expressions for EMW reflection and refraction in lossless nonmagnetic dielectrics are generalized to include lossless magnetic materials (with arbitrary permittivity and permeability). It is shown that the inclusion of lossless magnetic materials (general dielectrics) requires the definition of separate electromagnetic phase and amplitude refractive indexes. The phase refractive index is required to describe phase effects such as phase matching and total internal reflection at an interface between dielectrics. The amplitude refractive index is required to describe amplitude effects such as the reflectivity at an interface and the Brewster angle. In addition, by considering the reflectivity at interfaces between these general dielectrics, inherent analogies between the reflectivity of TE and TM polarizations are clarified. Then, starting with the single-electron effective-mass equation for conduction electrons, EW propagation in a semiconductor is treated. An EW phase refractive index is defined analogously to the electromagnetic phase refractive index. The EW phase refractive index describes phase matching at an interface between two semiconductor materials (with differing potential energy and effective mass) and total internal reflection from a semiconductor interface. By considering the reflectivity of an EW incident upon an interface between semiconductor materials an EW amplitude refractive index is defined analogously to the electromagnetic amplitude refractive index. Through the appropriate definition of these EW refractive indexes, it is shown that the electromagnetic expressions for calculating the critical angle, the reflectivity and transmissivity at an interface, and the Brewster angle can be used to calculate analogous quantities for EW propagation at interfaces between differing semiconductor materials. These analogies allow for the design of novel EW heterostructure devices (such as monoelectronic electron filters [27], [28]) using standard optical design techniques. The ranges of material parameters that allow for the existence of total internal reflection and for a Brewster angle are investigated for both optical and semiconductor materials. These results are applied to a boundary between two semiconductors in the GaAs–AlAs material system to investigate the reflectivity characteristics as a function of material composition. These analogies lead to a number of new results about reflection and refraction.

**Fig. 1.** Comparison of electrons and photons in terms of rest mass, charge, particle species (Fermion or Boson), spin, and velocity. In spite of their differences, a quantitative set of exact analogies can be drawn between EMW propagation and EW propagation for the propagation of an EW in a charge-neutral crystal, where the electron is either in a single nondegenerate conduction band or in a degenerate conduction band with no zero-field spin splitting.
First, by defining appropriate reference regions for the EW phase refractive index, inherent analogies are drawn between the electromagnetic constitutive parameters (permittivity and permeability) and the electron wave propagation parameters (kinetic energy and effective mass). These analogies allow the direct application of phase matching results from electromagnetics to the phase matching of electron waves. These analogies are used to calculate the angles of reflection and refraction at an interface between semiconductors and the conditions for total internal reflection at the interface.

Second, by considering the reflectivity of a dielectric interface between general dielectrics, an inherent analogy is developed between the reflectivity and transmissivity for TE polarized EMW's, for TM polarized EMW's, and for EW's. In the electromagnetic case, by defining a separate amplitude refractive index for both TE and TM polarizations, it is shown that a single set of expressions can be used to describe the reflectivity and transmissivity for both polarizations. Through the definition of analogous amplitude refractive indexes for an EW incident upon a boundary between two semiconductor materials, the reflectivity and transmissivity of the EW are shown to have the same form as the reflectivity of the polarization component of the EMW parallel to the interface (the electric field for TE polarization and the magnetic field for TM polarization). Thus a single set of reflectivity equations can be used for all three cases. In general dielectrics, the EMW phase (amplitude) refractive index is proportional to the square root of the product (ratio) of the relative permittivity and permeability of the material. In semiconductor materials, the EW phase (amplitude) refractive index is proportional to the square root of the product (ratio) of the kinetic energy and effective mass of the EW.

Third, an investigation of the material parameters that satisfy the total internal reflection condition shows that for both general dielectrics and semiconductor materials, there are a wide range of material parameters that can produce a given critical angle. In general dielectrics, the critical angle can be adjusted by varying the relative permittivity and permeability of the materials adjoining the interface. In semiconductors, the critical angle can be adjusted by varying the effective mass and the band offset of the semiconductor materials adjoining the interface. It is shown that for a Ga1-x1Alx1As/Ga1-x2Alx2As interface the aluminum compositions x1 and x2 can be adjusted to achieve a given critical angle. By including these effective mass contributions it is shown that an EW at a Ga1-x1Alx1As/Ga1-x2Alx2As interface can experience total internal reflection from both a potential rise and a potential drop.

Fourth, it is shown that since the reflectivity has the same form for TE and TM polarized EMW's, then one, both, or neither of the polarizations can have a Brewster angle. This result is not found in the standard nonmagnetic optical case (in which TM polarization always has a Brewster angle and TE polarization never has a Brewster angle). These results for the electromagnetic Brewster angle in general dielectrics are applied to the EW case. It is shown that when the kinetic energy and effective mass differ on the two sides of the interface, the EW will have a Brewster angle. If the effective mass difference were not included in the analysis, the EW results would follow the TE nonmagnetic case and a Brewster angle would not occur. It is shown that a Brewster angle will occur for a Ga1-x1Alx1As/Ga1-x2Alx2As interface for all compositional combinations that maintain direct gap semiconductors. This is the first prediction of a Brewster angle for ballistic electron transport in semiconductors.

II. PHASE REFRACTIVE INDEX

A. Electromagnetic Phase Refractive Index

Sinusoidal steady-state propagation of electromagnetic radiation in a lossless source-free medium is described by the sinusoidal steady-state form of Maxwell's equations where the time variation of the fields is taken to be \( \exp(-j\omega t) \). A minus sign is used in the exponent in order to be consistent with the quantum mechanical formulation (see Section II-B). For an isotropic, homogeneous medium, Maxwell's equations reduce to Helmholtz equations for wave propagation,

\[
\nabla^2 \vec{E} = -\omega^2 \mu \vec{H}, \quad \nabla^2 \vec{H} = -\omega^2 \varepsilon \vec{E},
\]

where \( \vec{E} \) and \( \vec{H} \) are the complex electric and magnetic fields, \( \omega \) is the radian frequency of the wave, and \( \varepsilon \) and \( \mu \) are permittivity and permeability of the medium, respectively. The forward propagating solution to (1) is a plane wave of the form: \( \vec{E} = \vec{E}_0 \exp(\pm j k \vec{r} \cdot \vec{r} - j \omega t) \) or \( \vec{H} = \vec{H}_0 \exp(\pm j k \vec{r} \cdot \vec{r} - j \omega t) \) where \( k \) is the wavevector for the forward propagating wave (which has a positive sign due to the choice of \( \exp(-j\omega t) \)). \( \vec{E}_0 \) and \( \vec{H}_0 \) are the vector wave amplitudes, and \( j = \sqrt{-1} \). The isotropic dispersion relation for the wavevector of the forward propagating wave is found by inserting the plane wave solution into the Helmholtz equation, yielding the dispersion equation \( |k| = \omega \sqrt{\mu \varepsilon} \).

In electromagnetics, a phase refractive index is commonly defined to relate the magnitude of the wavevector in the medium of propagation, \( |k| \), to the magnitude of the wavevector of the same frequency wave propagating in a reference medium, \( |k_{ref}| \), by

\[
\eta_{ph} \triangleq \frac{|k_{ph}|}{|k_{ref}|} = \frac{\sqrt{\mu \varepsilon}}{\sqrt{\mu_{ref} \varepsilon_{ref}}},
\]

where the wavevector of the reference region is related to the reference permittivity, \( \varepsilon_{ref} \), and permeability, \( \mu_{ref} \), through the isotropic dispersion relationship: \( |k_{ref}| = \omega \sqrt{\mu_{ref} \varepsilon_{ref}} \). This definition of the phase refractive index given in (2) is a function of the material parameters of both the region of propagation and the reference region. It would be useful to normalize the refractive index such that it is only a function of quantities in the region of propagation. This normalization can be performed if the reference region...
is dispersionless. That is, $\varepsilon_{\text{ref}}$ and $\mu_{\text{ref}}$ are not functions of the frequency of the EMW.

With a dispersionless reference region the phase refractive index can be defined in terms of the relative permittivity $\varepsilon_r \triangleq \varepsilon / \varepsilon_{\text{ref}}$ and relative permeability $\mu_r \triangleq \mu / \mu_{\text{ref}}$, yielding

$$\eta_{\text{ref}}^{\text{EM}} = \sqrt{\mu_r \varepsilon_r}.$$  \hspace{1cm} (3)

In electromagnetics, the reference region is usually taken to be the vacuum. Using the vacuum as a reference region, (3) gives the definition of the electromagnetic phase refractive index that is commonly found in optics textbooks [29], [30]. In this case, the relative permittivity and permeability are normalized with respect to the vacuum quantities: $\varepsilon_{\text{ref}} = \varepsilon_0 = 8.85 \times 10^{-12} \text{F/m}$ and $\mu_{\text{ref}} = \mu_0 = 4\pi \times 10^{-7} \text{H/m}$.

B. Electron Wave Phase Refractive Index

The propagation of single parabolic-band conduction electrons in a semiconductor crystal is described by the time-dependent effective-mass Schrödinger equation [21]. If the energy of the electron is constant (not a function of time or position) and if the propagation is limited to a region of constant potential energy, the time-dependent effective-mass equation reduces to the time-independent effective-mass Schrödinger equation for constant potential energy given as

$$\nabla^2 \psi = -\frac{2m^* (E - V)}{\hbar^2} \psi = -\frac{2m^*(KE)}{\hbar^2} \psi \tag{4}$$

where $\psi$ is the EW amplitude, $E$ is the total EW energy, $V$ is the EW potential energy (band edge), $(KE)$ is the EW kinetic energy, $m^*$ is the electron effective mass, and $\hbar$ is Planck's constant divided by $2\pi$. This constant energy formulation of the Schrödinger equation is analogous to the constant frequency (steady-state) formulation of Maxwell's equations where $\omega$ is replaced by the quantum mechanical radian frequency $\omega = E/\hbar$. From this construction, it is apparent that the differential equation describing propagation of an EW in a region of constant potential energy (see (4)) is analogous to the Helmholtz equation describing propagation of an EMW in a homogeneous general dielectric (see (1)), where the EW amplitude is analogous to either the electric or the magnetic field. The solution to the EW propagation equation (see (4)) is also a plane wave of the form $\psi = \psi_0 \exp (+jK_{\text{EW}} \cdot r)$, where, in the EW case, the dispersion equation for the wavevector is $|K_{\text{EW}}| = \sqrt{2m^*(KE)/\hbar^2}$.

As in electromagnetics, one can define a phase refractive index as the ratio of the magnitude of the wavevector in the region of propagation, $|K_{\text{EW}}|$, to the magnitude of the wavevector of the same total energy $(E)$ wave propagating in a reference medium, $|K_{\text{EW}}|_{\text{ref}}$, yielding

$$\eta_{\text{EW}}^{\text{EW}} \triangleq \frac{|K_{\text{EW}}|}{|K_{\text{EW}}|_{\text{ref}}} = \frac{\sqrt{m^*(KE)}}{\sqrt{m_{\text{ref}}^*(KE)_{\text{ref}}}} \tag{5}$$

where $|K_{\text{EW}}| = \sqrt{2m_{\text{ref}}^*(KE)_{\text{ref}}/\hbar^2}$. This is analogous to the definition of the phase refractive index in the electromagnetic case (see (2)).

As in the electromagnetic case, the dispersion in the EW phase refractive index needs to be proportional to the dispersion of the medium of propagation. This requires the use of a dispersionless EW reference region where the kinetic energy $(KE)_{\text{ref}}$ and effective mass $m_{\text{ref}}^*$ are not functions of the total energy of the EW. The fact that this region may not exist in a real semiconductor is of no consequence. All physical calculations use ratios of refractive indexes and $(KE)_{\text{ref}}$ and $m_{\text{ref}}^*$ will cancel.

Using a dispersionless reference region, the EW phase refractive index can be written in terms of a relative kinetic energy, $(KE)_{\text{ref}}$, and a relative effective mass, $m_{\text{ref}}^*$, as

$$\eta_{\text{EW}}^{\text{EW}} = \sqrt{m_{\text{ref}}^*(KE)_{\text{ref}}}.$$

Since the choice of a dispersionless reference region does not affect the physical calculations, $m_{\text{ref}}^*$ and $(KE)_{\text{ref}}$ will be defined to be unity such that the relative kinetic energies and effective masses are equal to the actual values in the region of propagation.

Equation (6) is the expression for the EW phase refractive index given a dispersionless reference region. It is interesting to note that the EW phase refractive index is expressed as the square root of the product of two material parameters, the same form as is the EMW phase refractive index. In real semiconductor material systems, such as Ga1-xAlxAs, the kinetic energy and effective mass differ for each composition. Thus it is important that effects described by the EW phase refractive index, such as phase matching (Section III) and total internal reflection (Section V), be analyzed with respect to the variation of both material parameters. It is for this reason that the analogies in this paper are drawn between EMW propagation in general dielectrics and EW propagation in semiconductors. An analogy to electromagnetic propagation in nonmagnetic dielectrics ($\mu_r = 1$) is not complete since the EMW phase refractive index is expressed as the square root of one parameter, the relative permittivity of the medium.

III. PHASE MATCHING UPON REFLECTION AND REFRACTION

When an EMW is incident upon a lossless dielectric boundary, Maxwell's equations must be satisfied on both sides of the boundary, and the solutions must match at the boundary. By applying the appropriate boundary conditions, the angles of the reflected and transmitted waves can be calculated. An EW incident upon a boundary between two semiconductor materials will obey the same type of relationships. In the EW case, the EW must satisfy the time-independent Schrödinger equation on both sides of the boundary, and the solutions must match at the boundary.

A. Electromagnetic Phase Matching

Consider an electromagnetic plane wave incident upon a boundary between two charge-free general dielectric materials as shown in Fig. 2. The incident, reflected, and transmitted waves are propagating in the $x-z$ plane, and...
Fig. 2. Electromagnetic reflection and transmission from a general dielectric boundary for (a) TE and (b) TM polarizations.

the interface is located at $z = 0$. The dielectric to the left (right) of the interface, region 1 (2), is described by the relative permittivity $\varepsilon_1$ ($\varepsilon_2$) and the relative permeability $\mu_1$ ($\mu_2$). The expressions for the incident ($E_1, H_1$), reflected ($E_r, H_r$), and transmitted ($E_t, H_t$) electric and magnetic fields are given as

$$
W_1 = W_1 \exp(+j\omega_1 \cdot \mathbf{r}) \\
W_r = W_r \exp(+j\omega_r \cdot \mathbf{r}) \\
W_t = W_t \exp(+j\omega_t \cdot \mathbf{r})
$$

where $\omega_i = \omega_{ri} = \omega_{rt}$, $\omega_{rt} = \omega_{rt}$, and $\omega_{rt}$ are complex wave amplitudes and $F_{EM}^1, F_{EM}^r, F_{EM}^t$ are the corresponding wavevectors that satisfy the isotropic dispersion relationship in regions 1 (for $i$ and $r$) and 2 (for $t$).

Maxwell's equations require the continuity of tangential component of the electric field, $E_1 \cdot \hat{\mathbf{z}} = 0$, and the tangential component of the magnetic field, $H_1 \cdot \hat{\mathbf{z}} = 0$, across the boundary. By substituting the field expressions (see (7)) into either of these two boundary conditions, it is found that the boundary conditions can be satisfied only if two phase matching conditions are satisfied, namely,

$$
\theta_1 = \theta_r = \theta_1, \quad \text{and} \quad n_{ph,1} \sin \theta_1 = n_{ph,2} \sin \theta_2
$$

where $\theta_1$ and $\theta_2 = \theta_1$ are the angles between the incident (reflected) and transmitted wavevectors and the surface normal, measured counterclockwise as is shown in Fig. 2. The second expression of (8) is commonly referred to as Snell's law. If the transmitted wave cannot satisfy the phase matching condition, total internal reflection occurs, which is described in Section V.

B. Electron Wave Phase Matching

The reflection and refraction of an EW incident upon an interface between two semiconductor materials can be analyzed in a manner analogous to the reflection and refraction of an EMW at an interface between dielectrics. As was the case in electromagnetics, expressions for the reflected and transmitted angles can be calculated by considering phase matching across the interface.

Consider an EW incident upon a boundary between two semiconductor materials as shown in Fig. 3. The EW propagation in region 1 (2) is characterized by the relative kinetic energy, $(KE)_1$, $(KE)_2$, and the relative effective mass, $m_i^* (m_2^*)$. For the interface shown in Fig. 3, the material in region 2 has a higher potential energy (a lower band gap) than the material in region 1. The results, however, are equally applicable to propagation into a region of lower potential energy. The expression for the incident, reflected, and transmitted, EW's are given as

$$
\psi_i = \psi_{ir} \exp(+jF_{EM}^{ir} \cdot \mathbf{r}) \\
\psi_r = \psi_{or} \exp(+jF_{EM}^{or} \cdot \mathbf{r}) \\
\psi_t = \psi_{ot} \exp(+jF_{EM}^{ot} \cdot \mathbf{r})
$$

where $\psi_{ir}, \psi_{or},$ and $\psi_{ot}$ are complex wave amplitudes and $F_{EM}^{ir}, F_{EM}^{or}, F_{EM}^{ot}$ are the wavevectors that satisfy the EW isotropic dispersion relationship in regions 1 (for $i$ and $r$) and 2 (for $t$). In a manner analogous to the EMW, the EW amplitude, $\psi$, and the transmitted component of the product of the reciprocal of the effective mass and the gradient of the EW amplitude, $(1/m^*)\nabla \psi \cdot \hat{\mathbf{z}}$, must be continuous across the boundary.

These EW boundary conditions have a direct analogy in the boundary conditions given above for an EMW incident upon a general dielectric interface. The electromagnetic boundary conditions are the matching of the tangential components of vector fields ($E$ and $H$). The EW boundary conditions involve the matching of a scalar field and the transmitted component of the gradient of the scalar field. In order to investigate the form of the analogies between the EMW boundary conditions and the EW boundary conditions, an analogy must be drawn between the EW amplitude and one of the electromagnetic field components $E$ or $H$. The choice of field component will not affect the form of the results. For this analysis, an analogy will be drawn between the EW amplitude and the electric field amplitude. The same results would be obtained if an analogy were drawn between the magnetic field amplitude and the EW amplitude.

Using the analogy between the electric field $E$ and the electron wave amplitude $\psi$, the EMW boundary condition matching the continuity of the tangential component of the electric field has the same functional form as the EW boundary condition matching the continuity of the wave amplitude across the boundary. The EMW boundary condition matching the continuity of the magnetic field should then have the same form as the EW boundary condition matching the continuity of the gradient of the wave amplitude. This analogy between the magnetic field
and the gradient of the electron wavefunction can be seen if the magnetic field is rewritten in terms of the electric field using Maxwell's curl relation, yielding \( \mathbf{H} = (1/j \omega \mu_e) \mathbf{E} \). Hence, the EMW boundary condition matching the continuity of the magnetic field can be expressed as a boundary condition matching the continuity of the tangential component of the curl \( \mathbf{H} \) of the electric field. The two electromagnetic boundary conditions are the matching of the tangential component of a vector field and the tangential component of its vector derivative. The EW boundary conditions involve the matching of a scalar field and the transmitted component of its vector derivative (gradient).

Since the boundary conditions and field expressions are similar for the EMW and the EW, it would be expected that the EW phase matching conditions have the same form as the EMW phase matching conditions. By substituting the EW amplitude expressions (see (9)) into either of the two electron wave boundary conditions, the EW phase matching conditions are found to be identical to the EMW phase matching conditions given in (8), where the EW phase refractive index, \( n_{EW}^E \), is used in place of the EMW phase refractive index, \( n_{EM}^E \). Since these equations describing the directions of propagation of the reflected and transmitted EW's are identical to the EMW results, standard electromagnetic results for describing phase propagation effects in general dielectrics, such as reflection, refraction, and total internal reflection (Section V) can be used to describe ballistic electron propagation in semiconductor heterostructures.

IV. AMPLITUDE REFRACTIVE INDEX

A. Electromagnetic Amplitude Refractive Index

Phase matching of the boundary conditions for an EMW incident upon a dielectric boundary was used to calculate expressions for the propagation directions of the reflected and transmitted waves. The relative amplitudes of the reflected and transmitted waves can be calculated by considering the conditions placed on the field amplitudes by the boundary conditions. The reflectivity and transmissivity can be calculated by considering any two of the following three boundary conditions: continuity of the tangential component of the electric field, continuity of the tangential component of the magnetic field, and conservation of normal power flow across the boundary, where the power flow is given as \( 1/2 \text{Re} (\mathbf{E} \times \mathbf{H}^*) \) and \( * \) denotes complex conjugate. Since any arbitrary wave polarization can be decomposed into a sum of the eigenmodes of the interface (TE and TM), and since the TE and TM modes are uncoupled (isotropic media), one can analyze TE and TM polarizations separately. Then, the reflection of any arbitrary polarization can be calculated from a vector superposition of the results for TE and TM polarizations.

1) TE Polarization: For a TE polarized plane wave (Fig. 2(a)), the polarization in which the electric field \( \mathbf{E} \) is polarized parallel to the interface, the three phase matched boundary conditions reduce to

$$ E_0 + E_{or} = E_{ot} \quad (10) $$

$$ (E_0/Z_1) \cos \theta_1 - (E_{or}/Z_1) \cos \theta_1 = (E_{ot}/Z_2) \cos \theta_2 \quad (11) $$

$$ (E_0^2/Z_1) \cos \theta_1 - (E_{or}^2/Z_1) \cos \theta_1 = (E_{ot}^2/Z_2) \cos \theta_2 \quad (12) $$

where \( Z_i \) is the characteristic impedance of the bulk medium \( i \) for \( i = 1 \) and \( i = 2 \). The characteristic impedance is \( Z_i = |E_i|/|H_i| = \sqrt{\mu_i \varepsilon_i} \) where \( \varepsilon_i \) and \( \mu_i \) are the relative permittivity and relative permeability of region \( i \). By using any two of these three boundary conditions (see (10)–(12)), the reflectivity and transmissivity of the electric field can be calculated in terms of the reflected and transmitted angles and the characteristic impedances of the two dielectrics as

$$ r_{TE}^E = \left( \frac{E_{or}}{E_0} \right)_E = \frac{Z_2 \cos \theta_1 - Z_1 \cos \theta_2}{Z_2 \cos \theta_1 + Z_1 \cos \theta_2} \quad (13) $$

$$ t_{TE}^E = \left( \frac{E_{ot}}{E_0} \right)_E = \frac{2Z_2 \cos \theta_1}{Z_2 \cos \theta_1 + Z_1 \cos \theta_2} \quad (14) $$

As was the case for the wavevector, it is convenient to rewrite these equations in terms of quantities normalized to a reference region. This normalization is performed through the use of an amplitude refractive index defined as

$$ n_{am}^E = \frac{Z_{ref}}{Z} = \sqrt{\varepsilon_r/\mu_r} \quad (15) $$

where the vacuum was used as a dispersionless reference region.

Using this expression for the amplitude refractive index (see (15)) the equations for the reflectivity and transmissivity can be rewritten in terms of the amplitude refractive indexes as

$$ r_{TE}^E = \left( \frac{E_{or}}{E_0} \right)_E = \frac{n_{am,1}^E \cos \theta_1 - n_{am,2}^E \cos \theta_2}{n_{am,1}^E \cos \theta_1 + n_{am,2}^E \cos \theta_2} \quad (16) $$

$$ t_{TE}^E = \left( \frac{E_{ot}}{E_0} \right)_E = \frac{2n_{am,1}^E \cos \theta_1}{n_{am,1}^E \cos \theta_1 + n_{am,2}^E \cos \theta_2} \quad (17) $$
Since these expressions for the reflectivity and transmissivity (see (16)–(17)) contain ratios of the amplitude refractive indexes, the reference dependence is removed. By using Maxwell’s curl equations the expressions for the reflectivity and transmissivity of the magnetic field can be calculated in terms of the reflectivity and transmissivity of the electric field as
\[
\begin{align*}
\tau_E^{\text{TM}} &= \left(\frac{H_{or}}{H_o}\right)_{\text{TM}} = \tau_E^\text{TE} \\
t_E^{\text{TM}} &= \left(\frac{H_{ot}}{H_o}\right)_{\text{TM}} = \frac{\eta_{\text{am},2}^{\text{TE}} \cdot t_E^{\text{TE}}}{\eta_{\text{am},1}^{\text{TE}}}. 
\end{align*}
\]
Equations (16)–(19) are the generalized Fresnel equations for the reflectivity and transmissivity of a TE polarized plane wave. These equations require the definition of an amplitude refractive index that in general differs from the phase refractive index. Most lossless dielectric materials, however, have a relative permeability (\(\mu_r\)) of unity. In this case of nonmagnetic dielectrics, the amplitude and phase refractive indexes are equal (as can be seen from (3) and (15)). Thus the distinction between the amplitude and phase refractive indexes is usually ignored. This is why the Fresnel equations are often written in terms of the phase refractive indexes in optics texts [29, 30]. It is important to note, however, that in general one must use the phase refractive index for phase matching, and a differing amplitude refractive index for calculation of the reflectivity and transmissivity of a dielectric interface.

2) TM Polarization: For a TM polarized plane wave shown in Fig. 2(b), the polarization in which the magnetic field (\(H\)) is parallel to the interface, the phase matched boundary conditions reduce to
\[
\begin{align*}
(H_1/Y_1) \cos \theta_1 - (H_2/Y_2) \cos \theta_1 &= (H_{ot}/Y_2) \cos \theta_2 \\
H_1 + H_{ot} &= H_{ot} \\
(H_2^2/Y_1) \cos \theta_1 - (H_1^2/Y_2) \cos \theta_1 &= (H_{ot}/Y_2) \cos \theta_2
\end{align*}
\]
where \(Y_i\) is the characteristic admittance of the bulk medium \(i\) for \(i = 1\) and \(i = 2\). The characteristic admittance is defined as the ratio of the magnetic field amplitude to the electric field amplitude and is the reciprocal of the characteristic impedance, yielding \(Y_i = |H_i|/|E_i| = \sqrt{\varepsilon_i\mu_i}/\sqrt{\varepsilon_i\mu_i} = 1/2i\). By combining any two of the above three phase matched boundary conditions, the equations for the reflectivity and transmissivity of the magnetic field can be expressed as a function of the angles of propagation and the dielectric admittances:
\[
\begin{align*}
\tau_H^{\text{TM}} &= \left(\frac{H_{or}}{H_o}\right)_{\text{TM}} = \frac{Y_2 \cos \theta_1 - Y_1 \cos \theta_2}{Y_2 \cos \theta_1 + Y_1 \cos \theta_2} \\
t_H^{\text{TM}} &= \left(\frac{H_{ot}}{H_o}\right)_{\text{TM}} = \frac{2Y_2 \cos \theta_1}{Y_2 \cos \theta_1 + Y_1 \cos \theta_2}.
\end{align*}
\]
As was performed for TE polarization, the reflectivity and transmissivity of the interface can be rewritten in terms of normalized admittances through the use of an amplitude refractive index. The TM amplitude refractive index is defined in an analogous manner to the TE amplitude refractive index as
\[
n_{\text{am}}^{\text{TM}} = \frac{Y_{\text{ref}}}{Y} = \sqrt{\mu_r/\varepsilon_r} = 1/n_{\text{am}}^{\text{TE}}(25)
\]
where the vacuum is chosen as the reference region.

Using this definition for the TM amplitude refractive index (see (25)), the equations for the reflectivity and transmissivity of the magnetic field can be written in terms of the amplitude refractive indexes as
\[
\begin{align*}
\tau_H^{\text{TM}} &= \left(\frac{H_{or}}{H_o}\right)_{\text{TM}} = \frac{n_{\text{am},1}^{\text{TM}} \cos \theta_1 - n_{\text{am},2}^{\text{TM}} \cos \theta_2}{n_{\text{am},1}^{\text{TM}} \cos \theta_1 + n_{\text{am},2}^{\text{TM}} \cos \theta_2} \\
t_H^{\text{TM}} &= \left(\frac{H_{ot}}{H_o}\right)_{\text{TM}} = \frac{2n_{\text{am},1}^{\text{TM}} \cos \theta_1}{n_{\text{am},1}^{\text{TM}} \cos \theta_1 + n_{\text{am},2}^{\text{TM}} \cos \theta_2}.
\end{align*}
\]

As was the case for TE polarization, the equations for the reflectivity and transmissivity of the electric field can be obtained using Maxwell’s curl equations, and are
\[
\begin{align*}
\tau_E^{\text{TM}} &= \left(\frac{E_{or}}{E_o}\right)_{\text{TM}} = \tau_H^{\text{TM}} \\
t_E^{\text{TM}} &= \left(\frac{E_{ot}}{E_o}\right)_{\text{TM}} = \frac{Y_1}{Y_2} t_H^{\text{TM}} = \frac{n_{\text{am},2}^{\text{TM}}}{n_{\text{am},1}^{\text{TM}}}.
\end{align*}
\]

Equations (26)–(29) are the generalized Fresnel equations for the reflectivity and transmissivity of TM polarized light. The amplitude refractive index in these equations is the reciprocal of the amplitude refractive index for TE polarization. In an analogous manner to TE polarization, there is a special case, \(\varepsilon_r = 1\), in which the TM amplitude refractive index and the phase refractive index are identical. In this case, the phase refractive index can be used to calculate the reflectivity (transmissivity) at the interface. Although this case is less common for lossless dielectric materials, and thus not generally used in optics, it is important to note that there is nothing “special” about TE polarization that allows the amplitude refractive index to be identical to the phase refractive index. It just happens that most lossless dielectric materials have a relative permeability of unity, leading to identical TE amplitude and phase refractive indexes. For general dielectrics, three separate refractive indexes are required in order to use the unified results given in this paper: one for phase matching, one for TE reflectivity (transmissivity), and one for TM reflectivity (transmissivity). These refractive indexes are summarized in Table 1.
3) Duality between TE and TM:

The reflectivity and transmissivity results for TE and TM polarized light are summarized in Table 2. The reflectivity and transmissivity for the electric (magnetic) field of a TE (TM) polarized wave are the same as those for the magnetic (electric) field of a TM (TE) polarized wave, where the TE amplitude refractive index ($n_{mTE}$) is used for TE polarization and the TM amplitude refractive index ($n_{mTM}$) is used for TM polarization. The similarity between the TE and TM reflectivity (transmissivity) results can be clarified by considering the polarization directions for both cases. For TE polarized light the electric field is parallel to the interface (tangential) and the magnetic field has a component normal to the interface (nontangential). For TM polarized light, the magnetic field is parallel to the interface (tangential) and the electric field has a component normal to the interface (nontangential). Thus there is one equation for the reflectivity (transmissivity) of the tangential field component ($E$ for TE and $H$ for TM) and another equation for the reflectivity (transmissivity) of the nontangential field component ($H$ for TE and $E$ for TM).

The duality between the reflectivity and transmissivity of TE and TM polarized light could alternatively be derived through the application of Babinet’s principle [31] which states that if $E_{a,n}$ and $H_{a,n}$ are solutions to Maxwell’s wave equation where $n = \text{i.r.t}$ for incident, reflected, and transmitted waves respectively, then the fields $E_{b,n} = \pm Z_n H_{a,n}$ and $H_{b,n} = \pm X_n E_{a,n}$ are also solutions to Maxwell’s equations. If field $a$ describes TE polarization and satisfies the boundary conditions, the application of Babinet’s transformation shows that field $b$ describes TM polarization and also satisfies the boundary conditions. If the transformations are applied to the boundary conditions for TE polarization (field $a$) given in (10)–(12), the resultant equations are identical to the phase matched boundary conditions for TM polarization (field $b$) given in (20)–(22). Thus the reflectivity (transmissivity) for TM polarized light could be calculated through the application of Babinet’s transformation to the boundary conditions for the dual TE problem. The fact that the equations describing TE and TM polarizations have the same functional form is a manifestation of this duality principle and is required by Maxwell’s equations.

B. Electron Wave Amplitude Refractive Index

The directions of propagation of the reflected and transmitted EW’s from a semiconductor interface (Fig. 3) were calculated by using phase matching considerations. The relative amplitudes of the reflected and transmitted EW’s can be calculated from the phase matched boundary conditions, as they were in the electromagnetic case. The reflectivity and transmissivity of the EW amplitude can be calculated by considering any two of the following three boundary conditions: continuity of the EW amplitude across the boundary, continuity of the transmitted component of the gradient of the EW amplitude, and conservation of the transmitted component of the probability current, $\mathbf{J} \cdot \hat{z}$, where the probability current is defined as $J = \hbar / 2 jm^* [\nabla \psi^* \psi - \psi^* \psi / m^*]$ [32]. The analogy between the continuity of the wavefunction (gradient of the wavefunction) and the electromagnetic boundary conditions matching the tangential field component (nontangential field component) were developed in Section III. The third boundary condition, the conservation of probability current across the interface, is analogous to the conservation of power flow across the boundary in electromagnetics, since both are expressions of the conservation of the transmitted wave intensity.

In order to calculate the reflectivity (transmissivity) of the EW, a characteristic amplitude constant, $X$, will be defined in the same manner as was the characteristic impedance for TE polarization and the characteristic admittance for TM polarization, as

$$X = \frac{|\psi|}{|\nabla \psi| / m^*} = \sqrt{\frac{\hbar^2 m^*}{2(K E)}}.$$  (30)

This characteristic amplitude constant is analogous to the characteristic impedance for TE polarization, in which case $|E|$ is analogous ($\sim$) to $|\psi|$ and $|H| \sim (1/m^*)|\nabla \psi|$, and to the characteristic admittance for TM polarization, in which case $|H| \sim |\psi|$ and $|E| \sim (1/m^*)|\nabla \psi|$. This characteristic amplitude constant has the same form as the characteristic admittance defined in a transmission line description of resonant tunneling processes in semiconductors (where $|E| \sim |\psi|$) [33]. Using the characteristic amplitude constant, the equations for the reflectivity and transmissivity of the material interface can be calculated as

$$r_{EW} = \frac{\psi_{in}}{\psi_{in}} = \frac{X_2 \cos \theta_1 - X_1 \cos \theta_2}{X_2 \cos \theta_1 + X_1 \cos \theta_2}$$  (31)

$$t_{EW} = \frac{\psi_{in}}{\psi_{in}} = \frac{2X_2 \cos \theta_1}{X_2 \cos \theta_1 + X_1 \cos \theta_2}.$$  (32)

By inspection of the above equations, it is apparent that the reflectivity (transmissivity) of an EW incident upon a potential barrier have the same form as the equations describing the reflectivity (transmissivity) of the electric field for TE polarization [$r^E_{AB}$, (13)] and the reflectivity (transmissivity) of the magnetic field for TM polarization [$r^TM_{BA}$, (23)]. Thus the reflectivity (transmissivity) of the EW is analogous to the reflectivity (transmissivity) of the tangential component of the EMW. As would be expected, the equations describing the reflectivity (transmissivity) of the gradient of the EW amplitude have the same form as the equations describing the reflectivity (transmissivity) of the nontangential component of the EMW ($r^E_{AB}$ and $r^TM_{BA}$). That is,

$$r^{EW}_{AB/m} = \left( \frac{|\nabla \psi_{in}| / m^*}{|\nabla \psi_{in}| / m^*} \right) = r^{EW}_{in}$$  (33)

$$t^{EW}_{AB/m} = \left( \frac{|\nabla \psi_{in}| / m^*}{|\nabla \psi_{in}| / m^*} \right) = \frac{X_1}{X_2} t^{EW}_{in}.$$  (34)

The analogy between the reflectivity (transmissivity) of the EW amplitude and the reflectivity (transmissivity) of
the tangential component of the EMW can be understood as follows. The EW boundary conditions describing the EW continuity at the boundary are scalar equations. The EMW boundary conditions describing the continuity of the tangential field component have the same form as the EW scalar equations, since the field is tangential to the interface. Therefore, the EW amplitude reflectivity (transmissivity) equation has the same form as the tangential EMW component reflectivity (transmissivity) equation. The gradient of the EW amplitude, on the other hand, is a vector that points in the direction of propagation of the EW. Thus the boundary condition describing the continuity of the transmitted component of the gradient of the EW amplitude are vector equations that require resolution of the gradient in the direction normal to the interface. The EMW boundary conditions for the nontangential field require a resolution of the field component in the direction tangential to the interface. Therefore, the reflectivity (transmissivity) of the gradient of the EW amplitude has the same form as the reflectivity (transmissivity) of the nontangential EMW component.

These analogies show that there is no inherent analogy between one of the EMW field amplitudes (E or H) and the EW amplitude. For bulk propagation, the EW amplitude is analogous to either E or H. For propagation past an interface, the EW amplitude is analogous to the tangential component of the electromagnetic field.

By defining the EW amplitude refractive index, the equations for the EW reflectivity and transmissivity can be expressed in terms of normalized quantities as were the EMW equations. The amplitude refractive index is defined as the ratio of the characteristic amplitude constant of a reference medium to the characteristic amplitude constant of the medium of propagation [(15) and (25)],

$$n_{am}^{EW} \triangleq \frac{X_{ref}}{X} = \sqrt{(KE)_{ref}/m}$$

(35)

where the characteristic amplitude constant of the reference region is found through (30).

With this expression for the amplitude refractive index the reflectivity and transmissivity of the EW amplitude can be expressed as

$$n_{am}^{EW} = \frac{r_{TE}^{EW}}{r_{TE}^{EW}}$$

$$n_{am}^{EW} = \frac{t_{TE}^{EW}}{t_{TE}^{EW}}$$

$$n_{am}^{EW} = \frac{r_{TM}^{EW}}{r_{TM}^{EW}}$$

$$n_{am}^{EW} = \frac{t_{TM}^{EW}}{t_{TM}^{EW}}$$

(36)

Through the definition of the EW amplitude refractive index, the expressions for the reflectivity and transmissivity at a semiconductor material boundary are identical to the expression for the reflectivity and transmissivity of the tangential component of the electromagnetic field. A general set of equations that can be used to describe the reflectivity and transmissivity for TE polarization, TM polarization, and EW's is given in Table 2. As is the case in electromagnetics, there is a special case in which the electron wave phase refractive index is equal to the electron wave amplitude refractive index. For EW's this occurs when the effective masses are equal on the two sides of the boundary. This constant effective mass case is analogous to the nonmagnetic dielectric case ($\mu_r = 1$) for TE polarization and the $e_r = 1$ case for TM polarization. In many material systems (such as Ga$_{1-x}$Al$_x$As), however, the effective mass can change by more than 50% across a material interface. In these material systems, the dependence of the effective mass in the amplitude refractive index is important, requiring the use of the EW amplitude refractive index in the reflectivity (transmissivity) expressions for general dielectrics. For this reason, it is instructive to analyze the reflection (transmission) properties of general dielectric interfaces. This analysis is performed in the next two sections, and the results are applied to the Ga$_{1-x}$Al$_x$As material system in Section VII.

V. TOTAL INTERNAL REFLECTION

In Section III, phase matching considerations were used to calculate an expression for the refracted angle for both an EMW incident upon a boundary between dielectrics and an EW incident upon a boundary between semiconductor materials, yielding $n_{ph,1} \sin \theta_1 = n_{ph,2} \sin \theta_2$ (see (8)). This phase matching condition cannot be satisfied if

$$n_{ph,1} > n_{ph,2} \text{ and } \theta_1 > \theta_c = \sin^{-1}(n_{ph,2}/n_{ph,1})$$

(37)

where $\theta_c$ is referred to as the critical angle. When the phase matching condition is not satisfied, there is no transmitted wave. This condition corresponds to total internal reflection (TIR). In order to have a real critical angle, the parameters of the materials adjoining the interface must satisfy the first

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Electromagnetic</th>
<th>Electron</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reflectivity</td>
<td>$r_{TE}^{TE}$</td>
<td>$r_{TM}^{TM}$</td>
<td>$r_{EW}^{EW}$</td>
</tr>
<tr>
<td>Reflectivity</td>
<td>$r_{TE}^{TE}$</td>
<td>$r_{TM}^{TM}$</td>
<td>$r_{EW}^{EW}$</td>
</tr>
<tr>
<td>Transmissivity</td>
<td>$t_{TE}^{TE}$</td>
<td>$t_{TM}^{TM}$</td>
<td>$t_{EW}^{EW}$</td>
</tr>
<tr>
<td>Transmissivity</td>
<td>$t_{TE}^{TE}$</td>
<td>$t_{TM}^{TM}$</td>
<td>$t_{EW}^{EW}$</td>
</tr>
</tbody>
</table>

Table 2: Reflectivity and Transmissivity Expressions for Electromagnetic Waves (TE and TM) and Electron Waves
part of (38). In order to have TIR, the incident angle must satisfy the second part of (38).

For optical nonmagnetic dielectric interfaces, (38), shows that TIR can occur if \( \epsilon_1 > \epsilon_2 \) where the critical angle is given as \( \theta_c = \sin^{-1}(\sqrt{\epsilon_2/\epsilon_1}) \). For general dielectrics and semiconductor interfaces, the range of material combinations that allow a critical angle is shown as the shaded region in Fig. 4, where \( \Omega = \mu_1/\mu_2 \) and \( \Gamma = \epsilon_1/\epsilon_2 \) for EMW's, and \( \Omega = m_1^*/m_2^* \) and \( \Gamma = (KE)_1/(KE)_2 \) for electron waves. The hyperbola \( \Omega = 1 \) \( (n_{ph,1} = n_{ph,2}) \) represents the onset of TIR, where the critical angle \( \theta_c \) is equal to 90 deg. For material combinations below this curve \( (\Omega < 1) \), the phase refractive index in region 2 is larger than the phase refractive index in region 1 and total internal reflection will not occur for any angle of incidence. For material combinations above this curve \( (\Omega > 1) \), there is a real critical angle for the interface. A given critical angle \( \theta_c \) is represented on the graph by the hyperbola \( \Omega \Gamma = 1/\sin^2 \theta_c \). As the product \( \Omega \Gamma \) is increased from unity, the value of the critical angle decreases, approaching normal incidence as the product \( \Omega \Gamma \) approaches infinity. Plots of the material combinations that yield a constant critical angle are shown for \( \theta_c = 90, 66.5, 45, \) and 22.5 deg.

The above conditions for TIR reduce to the standard expressions for the nonmagnetic dielectric case in which \( \mu_1 = \mu_2 = 1 \). This nonmagnetic dielectric condition is represented in Fig. 4 by the line \( \Omega = 1 \). Considering only the line \( \Omega = 1 \) gives the condition for a critical angle as \( \Gamma > 1 \), which reduces to \( \epsilon_1 > \epsilon_2 \), the familiar nonmagnetic dielectric condition for TIR. In the electron wave case, the line \( \Omega = 1 \) corresponds to the effective masses being equal, \( m_1^* = m_2^* \), in which case the condition for having total internal reflection is \( (KE)_1 > (KE)_2 \).

A given critical angle is represented by the curve \( \Omega \Gamma = 1/\sin^2 \theta_c \), there is an infinite set of relative material combinations that can be used to achieve a given critical angle. In electromagnetics, it is generally not possible to obtain lossless dielectrics with arbitrary permittivity and permeability. Thus it is difficult to use regions in Fig. 4 other than the line \( \Omega = 1 \). In this case, the critical angle is limited to the intersection of the line \( \Omega = 1 \) and the curve of constant critical angle. In semiconductor materials, since the kinetic energy and effective mass often differ on the two sides of an interface, there is more flexibility in the design of TIR devices such as EMW waveguides [34]. This effective mass dependence on the critical angle is manifested in the Ga1-xAlxAs material system in which TIR can occur from a potential drop due to the effective mass difference on the two sides of the boundary. The critical angle results discussed in this section are applied to the Ga1-xAlxAs material system in Section VII.

VI. BREWSTER ANGLE

In electromagnetics, the Brewster angle is defined as the angle of incidence that produces zero reflectivity at the interface between two dielectrics. At this angle, the incident wave is totally transmitted. By solving the equation for zero reflectivity (for either an EMW incident upon a boundary between two dielectrics or an EM wave upon an interface between semiconductor materials) the equation for the Brewster, \( \theta_B \), is found to be given as

\[
\theta_B = \sin^{-1}\left(\frac{1}{\sqrt{(1/A \Lambda)/((1-A^2))}}\right)
\]

where \( A \) and \( \Lambda \) are functions of the material parameters. Since the amplitude refractive indexes differ for TE polarization, TM polarization, and electron waves, the definitions of \( A \) and \( \Lambda \) also differ. For TE polarization, \( A = \mu_1/\mu_2 \) and \( \Lambda = \epsilon_1/\epsilon_2 \); for TM polarization, \( A = \mu_1/\mu_2 \) and \( \Lambda = \epsilon_1/\epsilon_2 \); and for electron waves, \( A = m_1^*/m_2^* \) and \( \Lambda = (KE)_1/(KE)_2 \). In order for the interface to have a real Brewster angle, the argument of the square root in (39) must lie between 0 and 1. The ranges of material parameters that yield a real \( \theta_B \) are shown as the shaded regions in Fig. 5. The material combinations that lie in these regions have a Brewster angle given by (39). These regions are bounded by the curve \( A\Lambda = 1 \), which is the condition \( n_{ph,1} = n_{ph,2} \), and by the curve \( A = \Lambda \), which is the condition \( n_{am,1} = n_{am,2} \). The material combinations that lie along the boundary curve \( A\Lambda = 1 \) do not have a Brewster angle. For the material combinations along this curve, the reflectivity is a constant and is not a function of the angle of the incident wave! The material combinations that lie on the line \( A = \Lambda \) have a Brewster angle at normal incidence. In other words, there is no reflected wave at normal incidence for these material combinations! For the material parameters in between these two boundary curves the Brewster angle is given by (39). Contour plots o
material combinations for five constant Brewster angles (0, 22.5, 45, 66.5, and 90 deg) are shown in Fig. 5. For small Brewster angles, the material combinations stay close to the line \( \Lambda = \Gamma \), for small values of \( \Gamma \), but approach the curve \( A = 1 \) as \( \Gamma \) approaches infinity. As the value of the Brewster angle is increased, the material combinations move away from the line \( \Lambda = \Gamma \) and toward the curve \( A = 1 \) at a faster rate. As the Brewster angle approaches grazing incidence, the material parameters rise sharply for small values of \( \Gamma \) and rapidly approach the curve \( A = 1 \). It should be noted that all Brewster angles pass through the point \( \Lambda = \Gamma = 1 \). At this point, the two materials are the same and there is no interface, thus all of the incident wave is transmitted for every angle of incidence.

All of the above results are valid for both an EMW incident upon an interface between two general dielectrics and an EMW incident upon a boundary between two semiconductor materials. In addition, the results are valid for both TE and TM polarizations in the electromagnetic case. Thus for general dielectric, there are ranges of material parameters that allow for a Brewster angle for TE polarization as well as for TM polarization. This result might seem to contradict the common optical understanding that a Brewster angle always occurs for TM polarization, and never occurs for TE polarization. This contradiction can be resolved by considering the case for nonmagnetic dielectric materials \((\mu_1 = \mu_2)\), in which the graph in Fig. 5 is reduced to the line \( \Lambda = 1 \) (for TE polarization). It is apparent that this line does not intersect either region that allows for a Brewster angle. Thus there is no TE Brewster angle for nonmagnetic dielectric materials. The restriction to nonmagnetic dielectric materials reduces the TM polarization graph to the line \( T = 1 \). This line is completely contained in the regions that have a Brewster angle. Thus there is a TM Brewster angle for all combinations of nonmagnetic dielectric materials. In general, however, one, both, or neither of the polarizations can have a Brewster angle. For most material combinations, either the TE polarized wave or the TM polarized wave will have a Brewster angle. For those material combinations that lie along the line \( \Lambda T = 1 \), neither of the polarizations will have a Brewster angle; the reflectivity is independent of angle. For those material combinations that lie along the curve \( \Lambda = \Gamma \), both polarizations will have a Brewster angle at normal incidence (TE and TM polarization are identical at normal incidence). For all other material combinations, either one or the other of the polarizations will have a Brewster angle.

For material interfaces in which the effective mass is the same on the two sides of the boundary, no Brewster angle will occur. This result can be obtained directly from Fig. 5. The constant effective mass case \((\mu_1 = \mu_2)\) reduces Fig. 5 to the line \( \Lambda = 1 \). It is apparent from the graph that the line does not pass through the regions that have a Brewster angle. This condition is analogous to the nonmagnetic dielectric case \((\mu_1 = 1)\) for TE polarization and the \( \epsilon_r = 1 \) case for TM polarization. Thus an effective mass difference (across the interface) is required for an EW to have a Brewster angle. These results are applied to the \( \text{Ga}_{1-x}\text{Al}_x\text{As} \) material system in the next section where it is shown that all interfaces created in this material system will have a Brewster angle.

VII. ELECTRON WAVE REFLECTION AND REFRACTION IN \( \text{Ga}_{1-x}\text{Al}_x\text{As} \)

The results that were derived for the reflectivity and transmissivity of an EW at a material interface can be directly applied to heterostructures in the \( \text{Ga}_{1-x}\text{Al}_x\text{As} \) material system. By inserting the appropriate expressions for the kinetic energy and effective mass into the expressions defining the phase and amplitude refractive indexes (see Tables 1 and 2) the reflectivity characteristics of a \( \text{Ga}_{1-x}\text{Al}_x\text{As}/\text{Ga}_{1-y}\text{Al}_y\text{As} \) interface can be analyzed as a function of EW energy. For this material system, the potential energy and effective mass are a function of total electron energy \( (E) \) and aluminum composition \( (x) \). It is instructive to analyze the ranges of EW energies and boundary compositions that allow for TIR (critical angle, \( \theta_c \)) and zero reflectivity (Brewster angle, \( \theta_B \)). The reflectivity of the boundary will be considered for two cases: \( x_1 < x_2 \) which is a potential rise shown in Fig. 6 and \( x_2 > x_1 \) which is a potential drop shown in Fig. 7.

For interfaces constructed using the \( \text{Ga}_{1-x}\text{Al}_x\text{As} \) material system, the range of usable compositions that maintain direct gap semiconductors is given as \( 0 \leq x_1 \leq 0.45 \). The electron kinetic energy and effective mass for electrons in
the Γ valley in each of the two regions are given as [35]

\[(KE)_i = E - Ax_i, \quad m^*_i = m_o(B + Cx_i), \quad i = 1, 2\] 

where \(A\), \(B\), and \(C\) are constants, \(m_o\) is the free electron mass, and the bottom of the conduction band in pure GaAs is used as the zero of total energy. For Ga\(_{1-x}\)Al\(_x\)As, \(A = 0.77314\) eV (assuming a 62% conduction band offset [35]-[38]), \(B = 0.067\), and \(C = 0.083\) [35]. By using (40) for the kinetic energy and effective mass, the existence of both a critical angle and a Brewster angle can be investigated as a function of compositions \((x_1\) and \(x_2)\) and total electron energy \((E)\).

For this material system, the ratio of effective masses in the two regions is limited by the range of usable compositions given above. For interfaces with \(x_1 < x_2\), the range of effective mass ratios \((\Omega \text{ for TIR and } \Lambda \text{ for Brewster angle})\) is limited to the region \(B/(B + 0.48C) \leq m^*_1/m^*_2 \leq 1\). For interfaces with \(x_1 > x_2\), the range of effective mass ratios is limited to the region \(1 \leq m^*_1/m^*_2 \leq (B + 0.48C)/B\). These regions are shown graphically as the dotted regions of Fig. 8. In the case of the potential rise \((x_1 < x_2)\), when \((KE)_2\) is equal to zero, the ratio of kinetic energies \((\Gamma \text{ for TIR and } \Lambda \text{ for Brewster angle})\) is equal to infinity. As the kinetic energy in the output region is increased, the ratio of kinetic energies decreases, approaching unity as \((KE)_2\) approaches infinity. The compositional difference of the rise gives the effective mass ratio \((E)\), fixing the ordinate of the graph. The energy dependence of the critical and Brewster angles is given as a vertical line at this ordinate. The points of intersection between the curves of constant critical (Brewster) angle and this line give the values of the EW kinetic energy that yield a given critical (Brewster) angle for the chosen compositional difference. An example line is shown in Fig. 8, where \(x_1 = 0\) and \(x_2 = 0.3\). The kinetic energy dependence of the critical and Brewster angle of a potential drop behaves in a similar manner to the potential rise. When \((KE)_1\) is equal to zero, the ratio of kinetic energies is equal to zero. As the kinetic energy in the input region increases, the ratio of kinetic energies increases, approaching unity as \((KE)_1\) approaches infinity. Again, the compositional difference fixes the ordinate on the graph. The ratio of kinetic energies for a given critical (Brewster) angle is given by the intersection of the curves of constant critical (Brewster) angle with a vertical line at the ordinate fixed by the effective mass ratio. A sample line is shown in Fig. 8 for a potential drop of \(x_1 = 0.3\) and \(x_2 = 0.0\). These energy dependencies of the critical and Brewster angle can be used to show the regions of EW input energies that allow a critical angle and/or a Brewster angle. These regions are discussed as follows.

For the potential rise (Fig. 6), if the electron energy in region 1 is less than the size of the potential rise, i.e., if \((KE)_1 < A(x_2 - x_1)\), then the EW will experience TIR from the potential barrier for all angles of incidence. For these electron energies, the critical angle \(\theta_c\) occurs at normal incidence \((\theta_c = 0)\). As \((KE)_1\) is increased past \(A(x_2 - x_1)\), the EW will be partially transmitted at normal
with energies less than this critical value, there is no Brewster angle. As the EW energy is increased, critical angle. For EW energies below the critical energy, there is a Brewster angle that exists for all interfaces in this material system. In

As (KE) is further increased, the critical angle increases from normal incidence, until the critical angle reaches grazing incidence at the electron energy \( E_c = A(x_1 + x_2) + AB/C \). For all electron energies larger than \( E_c \), the EW will not experience TIR for any angle of incidence. At these energies, a fraction of the EW amplitude will be transmitted past the interface. As the electron energy is increased infinitesimally above \( E_c \), TIR occurs at grazing incidence. As the energy is further increased above \( E_c \), the critical angle asymptotically approaches the value \( \theta_{\text{B}} \) given by:

\[
\theta_{\text{B}} = \sin^{-1} \left( \frac{1 - \epsilon_x \epsilon_y}{1 + \epsilon_x \epsilon_y} \right)
\]

As the EW energy is further increased above \( E_c \), the Brewster angle decreases asymptotically to the value \( \theta_{\text{B}} = \sin^{-1} \left( \frac{1 - \epsilon_x \epsilon_y}{1 + \epsilon_x \epsilon_y} \right) \) as the energy approaches infinity. Thus for electron energies below the critical value \( E_c \), a fraction of the electron probability current will always be reflected. For electron energies above the critical value \( E_c \), there is an angle of incidence for each electron energy for which the electron probability current will be totally transmitted. As the EW energy approaches infinity, the angle of incidence that allows for total transmission asymptotically approaches the value \( \theta_{\text{B}} \). It is interesting to note that the asymptotic value of the Brewster angle is not normal incidence. This is because the reflectivity at normal incidence does not go to zero for high energies due to the effective mass difference. At high EW energies, the kinetic energy difference becomes negligible but the effective mass difference remains.

The energy dependence of the critical angle and the Brewster angle for a potential drop (Fig. 7) are similar to the results discussed above for the potential rise. In fact, these results predict a region of EW energies for which an EW can be totally internally reflected from a potential drop. This TIR occurs for energies at which the relative difference in effective masses is larger than the relative change in kinetic energies. Thus TIR will occur at a potential drop for electron energies whose input kinetic energy is significantly larger than the magnitude of the potential drop. The onset of TIR for a potential drop occurs at the same critical energy \( E_c \) at which the TIR disappeared for the potential rise. For electron energies below \( E_c \), there is no critical angle. At these energies, all electron waves will be partially transmitted past the interface. As the electron energy is increased infinitesimally above \( E_c \), TIR occurs at grazing incidence. As the energy is further increased above \( E_c \), the critical angle asymptotically approaches the value \( \theta_{\text{B}} \) given by:

\[
\theta_{\text{B}} = \sin^{-1} \left( \frac{1 - \epsilon_x \epsilon_y}{1 + \epsilon_x \epsilon_y} \right)
\]

As the EW energy is further increased above \( E_c \), the Brewster angle decreases asymptotically to the value \( \theta_{\text{B}} = \sin^{-1} \left( \frac{1 - \epsilon_x \epsilon_y}{1 + \epsilon_x \epsilon_y} \right) \) as the energy approaches infinity. Thus for electron energies below the critical value \( E_c \), a fraction of the electron probability current will always be reflected. For electron energies above the critical value \( E_c \), there is an angle of incidence for each electron energy for which the electron probability current will be totally transmitted. As the EW energy approaches infinity, the angle of incidence that allows for total transmission asymptotically approaches the value \( \theta_{\text{B}} \). It is interesting to note that the asymptotic value of the Brewster angle is not normal incidence. This is because the reflectivity at normal incidence does not go to zero for high energies due to the effective mass difference. At high EW energies, the kinetic energy difference becomes negligible but the effective mass difference remains.

The energy dependence of the critical angle and the Brewster angle for a potential drop (Fig. 7) are similar to the results discussed above for the potential rise. In fact, these results predict a region of EW energies for which an EW can be totally internally reflected from a potential drop. This TIR occurs for energies at which the relative difference in effective masses is larger than the relative change in kinetic energies. Thus TIR will occur at a potential drop for electron energies whose input kinetic energy is significantly larger than the magnitude of the potential drop. The onset of TIR for a potential drop occurs at the same critical energy \( E_c \) at which the TIR disappeared for the potential rise. For electron energies below \( E_c \), there is no critical angle. At these energies, all electron waves will be partially transmitted past the interface. As the electron energy is increased infinitesimally above \( E_c \), TIR occurs at grazing incidence. As the energy is further increased above \( E_c \), the critical angle asymptotically approaches the value \( \theta_{\text{B}} \) given by:

\[
\theta_{\text{B}} = \sin^{-1} \left( \frac{1 - \epsilon_x \epsilon_y}{1 + \epsilon_x \epsilon_y} \right)
\]

As the EW energy is further increased above \( E_c \), the Brewster angle decreases asymptotically to the value \( \theta_{\text{B}} = \sin^{-1} \left( \frac{1 - \epsilon_x \epsilon_y}{1 + \epsilon_x \epsilon_y} \right) \) as the energy approaches infinity. Thus for electron energies below the critical value \( E_c \), a fraction of the electron probability current will always be reflected. For electron energies above the critical value \( E_c \), there is an angle of incidence for each electron energy for which the electron probability current will be totally transmitted. As the EW energy approaches infinity, the angle of incidence that allows for total transmission asymptotically approaches the value \( \theta_{\text{B}} \). It is interesting to note that the asymptotic value of the Brewster angle is not normal incidence. This is because the reflectivity at normal incidence does not go to zero for high energies due to the effective mass difference. At high EW energies, the kinetic energy difference becomes negligible but the effective mass difference remains.

The energy dependence of the critical angle and the Brewster angle for a potential drop (Fig. 7) are similar to the results discussed above for the potential rise. In fact, these results predict a region of EW energies for which an EW can be totally internally reflected from a potential drop. This TIR occurs for energies at which the relative difference in effective masses is larger than the relative change in kinetic energies. Thus TIR will occur at a potential drop for electron energies whose input kinetic energy is significantly larger than the magnitude of the potential drop. The onset of TIR for a potential drop occurs at the same critical energy \( E_c \) at which the TIR disappeared for the potential rise. For electron energies below \( E_c \), there is no critical angle. At these energies, all electron waves will be partially transmitted past the interface. As the electron energy is increased infinitesimally above \( E_c \), TIR occurs at grazing incidence. As the energy is further increased above \( E_c \), the critical angle asymptotically approaches the value \( \theta_{\text{B}} \) given by:

\[
\theta_{\text{B}} = \sin^{-1} \left( \frac{1 - \epsilon_x \epsilon_y}{1 + \epsilon_x \epsilon_y} \right)
\]
addition, these figures illustrate the important role of the effective mass in electronic propagation in semiconductors. Merely viewing the band (potential energy) diagrams does not give complete information about the quantum reflectivity of a material interface between two semiconductors.

VIII. SUMMARY AND DISCUSSION

A comprehensive set of analogies have been presented between electronic wave propagation in semiconductors and electromagnetic propagation in general dielectrics. It has been shown that separate phase and amplitude refractive indexes are required for both cases. The phase refractive index is required for describing phase effects such as phase matching and total internal reflection, and the amplitude refractive index is required for describing amplitude effects such as the reflectivity of a material interface. For the electron wave, the phase (amplitude) refractive index is proportional to the square root of the product (ratio) of the kinetic energy and the effective mass of the medium of propagation. For the EMW, the phase (amplitude) refractive index is proportional to the square root of the product (ratio) of the permittivity and permeability of the medium of propagation.

By considering the phase matching and amplitude conditions of the electromagnetic boundary conditions, the reflection and refraction of an EMW incident upon an interface between two general dielectrics has been analyzed. It has been shown that one set of equations can be used to calculate the angles and the amplitudes of the reflected and transmitted waves for both TE and TM polarizations through the definition of different amplitude refractive indexes for each polarization. In addition, by applying the analogies discussed above, it has been demonstrated that the same set of expressions can be used to describe electron wave reflection and refraction at an interface between semiconductor materials through the definition of electron wave phase and amplitude refractive indexes. The behavior of the phase matching and reflectivity expressions have been determined as a function of material parameters. Through this analysis, it has been shown that for general dielectrics there is an infinite set of parameter values that produce a given critical angle. In addition, it has been shown that for general dielectrics a Brewster angle can occur for TE polarization, TM polarization, both TE and TM polarization, or neither TE nor TM polarization depending on the material parameters.

Using analogies between the electromagnetic and electronic wave propagation indexes, the critical angle and Brewster angle for a semiconductor material interface have also been determined as a function of kinetic energy and effective mass of the electron wave on the two sides of the material interface. Analogous to the electromagnetic case, it has been shown that for general material systems (with arbitrary kinetic energy and effective mass), there are an infinite number of material compositions that produce a given critical angle. In addition, it has been shown that an electron wave can encounter a Brewster angle if the effective mass differs on the two sides of the boundary. For general material systems, this Brewster angle is a function of the energy of the electron wave and can occur at normal incidence for appropriate material combinations. These results for electron wave propagation have been applied to interfaces in the Ga1_x Al_x As material system to determine the reflectivity of an interface as a function of electron energy and material composition. By applying this analysis, it has been shown that electron waves in the Γ band in this material system can be totally internally reflected from both a potential rise and a potential drop. In addition, it has been shown that a Brewster angle will exist for all interfaces in this material system. It should be noted that the Brewster angle calculated here for Γ band electrons in Ga1-x Al_x As occurs at energies above the L band satellite minima. The analysis method, however, also applies to other material systems whose range of energies for Brewster angles may fall above or below those of any minima, depending on the material system parameters.

The present formalism can be used to analyze electron wave optical structures, such as the impressive experimental devices of Sivan et al. [17] and Spector et al. [18–19]. The phase matching relationships of the present work (in the limiting case of a constant effective mass throughout the device) reduce to the same expressions that they used to explain successfully the results of these experiments. In their work, a complete ray tracing analysis was performed, using the electron wave phase indexes of refraction, to predict the device response. In order to demonstrate the procedure of applying the present formalism to the analysis of electron wave optical devices, the electron wave prism switch [19] is treated. The results of this analysis are shown in Fig. 9. In this device, constructed in the two-dimensional electron gas configuration, an incident electron wave is refracted by a potential energy prism and is switched to collector A, B, or C. The potential energy prism is created by applying a gate voltage to the electrode labeled n_E'(n_E''). The electron wave inside the prism (index n_E'(n_E'')) is incident upon the prism surface at angle θ_1 and is refracted into the output region (index n_E''(n_E')) at angle θ_2. By using the current-voltage and electron density data given in [19], the index ratios (n_E''(n_E')) that switch the electron into collectors A, B, and C are calculated to be 1.4, 1.3, and 1.0, respectively. These index ratios would result in angles of refraction (θ_2) of 38.2°, 52.7°, and 67.0° for collector A, B, and C, respectively. From the geometry of the device, the ranges of incident angles that can be accepted by each collector are approximately 30 to 39° for collector A, 46 to 53° for collector B, and 60 to 68° for collector C. Thus the calculated angles of refraction are consistent with these angular ranges. In the limit of an index ratio of n_E''/n_E' = ∞, the electron wave would be refracted normal to the prism surface. For an index ratio of n_E''/n_E' = 0.92, the electron wave would be refracted parallel to the prism surface (onset of total internal reflection). For this index ratio, the angle of incidence θ_1 (≈ 67°) is equal to the critical angle as given by (38). The
calculated. which the reflectivity (transmissivity) characteristics can be
with direct gap Ga\textsubscript{1-x}Al\textsubscript{x}As. Furthermore, in addition to analyzing phase effects (refraction angles), the present formalism can be used to extend these results to include amplitude effects such as reflectivity and transmissivity at boundaries. Furthermore, the present methods can be used to include simultaneously both phase and amplitude effects, incorporating important quantum interference physical optics effects [7H16] so as to calculate the total reflectance and transmittance of multiple-boundary, multiple-layer, and multiple-component electron optical devices.

Even though parabolic band, direct gap Ga\textsubscript{1-x}Al\textsubscript{x}As has been treated in this work, it should be emphasized that the methods can be applied to other material systems. It is conceptually straightforward to include nonparabolicity effects (often represented as \( E(1 + \alpha E + \beta E^2 + \ldots) = \langle \Delta E^2 \rangle / 2m^* \) in the analysis. Similarly, anisotropic effects can be incorporated through the use of a tensor effective mass and potential energy. This anisotropic case could be used if the lowest energy band consisted of ellipsoidal X minima energy surfaces (e.g., Si, AlAs, etc.) or ellipsoidal L minima energy surfaces (e.g., Ge) rather than the spherical \( \Gamma \) minimum energy surface associated with direct gap Ga\textsubscript{1-x}Al\textsubscript{x}As.

Although the form of the potential energy has been taken to be \( V = A x \), since this linear relationship applies to direct gap Ga\textsubscript{1-x}Al\textsubscript{x}As and numerous other material systems, the analysis presented can be straightforwardly extended to general potentials given as \( V = f(x) \). For example, for indirect gap Ga\textsubscript{1-x}Al\textsubscript{x}As (0.45 \( \leq x \leq 1 \)) the potential \( V = 1.900 + 0.125 x + 0.193 x^2 \) [35], [36] can be used to calculate the phase and amplitude refractive indexes, from which the reflectivity (transmissivity) characteristics can be calculated.

The analysis presented in this paper is based on the single-band effective-mass equation. Such an analysis is valid for conduction electrons in type I superlattices constructed from semiconductors such as Ga\textsubscript{1-x}Al\textsubscript{x}As [40]. For the conduction of holes in the valence band of type I superlattices (Ga\textsubscript{1-x}Al\textsubscript{x}As [21]), the conduction of electrons in material systems with zero-field spin splitting in the conduction band (such as Ga\textsubscript{1-x}In\textsubscript{x}As [39]), or the conduction of electrons in type II superlattices [40], a multiband effective mass equation is required. In this case, the effective mass equation is a \( N \)-dimensional matrix equation rather than a scalar equation. The electron-wave amplitude \( \psi \) for single-band bulk propagation (the case treated in this paper) is analogous to either the electric or the magnetic field. For propagation across an interface, \( \psi \) is analogous to the component of the electromagnetic field that is normal to the plane of incidence. In the case of the two-band effective mass equation (lack of spin degeneracy), the resulting electromagnetic analogy differs from that treated in this paper. Bulk propagation in the two-band case is analogous to EM propagation in an anisotropic media. The two bands represent a two-sheeted wavevector surface, analogous to the two-sheeted wavevector surface used to describe electromagnetic propagation in anisotropic materials [30].

In the electromagnetic case, the direction of the electric field polarization determines the wave decomposition into the two wavevector surfaces. In the electron wave case, the orientation of the electron spinor determines the decomposition of the electron wave into the two bands [21], [39]. In this manner, the electron spin is analogous to the electromagnetic field polarization. However, the electron wave eigen-solutions for the two-band model (spinors) are two-dimensional. Thus the two-band model describes a two-dimensional wavevector surface as opposed to the three-dimensional wavevector surface in electromagnetic optics [30]. In the \( N \)-band case, the \( N \) bands describe an \( N \)-sheeted wavevector surface in direct extension to the standard two-sheeted surface in anisotropic optics. The orientation of the \( N \)-dimensional vector describing the electron wave amplitude determines the wave decomposition into the wavevector surfaces. These analogies allow the description of multiband transport in terms of EM propagation in an anisotropic crystals.

An important result of this work is the demonstration that the analogies from the optics of nonmagnetic dielectrics are not applicable to electron wave propagation in semiconductors due to kinetic energy and effective mass differences across the interface. The optical expressions must first be generalized to include nonmagnetic dielectrics and then the results can be applied to the design of ballistic electron devices.

If the effective mass difference across the interface were not included in this analysis, many of the results would differ dramatically. By including the effective mass difference across a boundary, it was shown that an electron wave could be totally internally reflected from a potential drop. In addition, the existence of an electron wave Brewster angle is dependent upon the inclusion of the effective mass.
dependence. In fact, when the effective mass is included in the analysis, the reflectivity of an interface for an arbitrary angle of incidence does not go to zero as the energy of the electron wave goes to infinity. Thus there is always a component of the electron wave that is reflected due to the effective mass difference on the two sides of the interface. These reflections will be eliminated if the electron waves are incident at the Brewster angle. For these material systems that have effective mass changes, the existence of an electron wave Brewster angle could become useful in the design of antireflection interfaces in ballistic electron devices [41]. As was discussed in Section I, recent experiments have demonstrated the importance of including these effective mass effects [24]-[26].

With the comprehensive set of quantitative analogies presented in this work, it is clear that a wide variety of electron wave optical devices are possible using electron wave propagation and that these devices can be designed starting with existing optical designs.

Using the results of the present formalism, electron wave optical components can be constructed including a wide variety of electron interference filters [27], [28] that have high-pass, low-pass, band-pass (narrow-band or wide-band), or notch (narrow-band and wide-band) characteristics. Impedance transformers (antireflection coatings) to reduce unwanted electron reflections at boundaries can be designed [41]. These could be used, for example, in the electron flow from the GaAs base of a bipolar junction transistor to the Ga_{1-x}Al_{x}As collector where Ga_{1-x}Al_{x}As is used to increase the breakdown voltage in the high electric field collector. One-dimensional periodic structures (diffraction gratings) can be designed to diffract electron waves [42]-[44], allowing a variety of beam-splitting and transistor-like switching functions to be performed.

In addition, nearly monoelectronic electron emitters can be designed that are voltage tunable [28], [46]. Electron emitters can be incorporated as hot electron injectors in the design of electroluminescent devices [46]-[47], photodetectors [46], and fast ballistic transistors [46]. Negative resistivity devices can also be designed utilizing linearly varying period (chipped) superlattices [48], [49] or Fabry-Perot emitter/filter structures [28], [45], [50]. All of these devices could be used in the construction of ultrahigh speed switches and high-frequency oscillators.

The quantitative analogies presented here may be used in the design of slab waveguides (quantum wells with a two-dimensional electron gas) [34] and channel waveguides (quantum wires with a one-dimensional electron gas). Electron waveguides must be used in nanometer-scale high-speed integrated circuits. They are also required in future electron guided-wave integrated circuits for implementing high-speed integrated switches, modulators, and logic devices for computers and communications. Such guided wave circuitry could also perform "optical-like" parallel data processing, in a manner analogous to present-day electromagnetic optical data processing which is done with guided-wave integrated optical circuits (with a wavelength approximately 100 times smaller than the wavelength in the electromagnetic optical devices). In addition, the interface to conventional digital electronic circuitry could be contained within the same semiconductor chip, thus greatly simplifying the "optics-electronics" interface.

REFERENCES


Gregory N. Henderson (Student Member, IEEE), for a photograph and biography please see page 1180 of the August 1991 issue.

Thomas K. Gaylord (Fellow, IEEE), for a photograph and biography please see page 1180 of the August 1991 issue.

Elias N. Glytsis (Senior Member, IEEE), for a photograph and biography please see page 1180 of the August 1991 issue.
Ballistic current-voltage characteristics of semiconductor superlattice electron-wave quantum-interference filter/emitter negative differential resistance devices

Elias N. Gytsis, Thomas K. Gaylord, and Kevin F. Brennan
School of Electrical Engineering and Microelectronics Research Center, Georgia Institute of Technology, Atlanta, Georgia 30332

(Received 2 October 1990; accepted for publication 27 June 1991)

The transmission and current-voltage characteristics of Ga_{1-x}Al_{x}As superlattice electron-wave quantum-interference filter/emitter negative differential resistance devices are analyzed with and without the self-consistency requirement. The analysis neglects the scattering effects within the superlattice since it is assumed that it is very thin (less than the electron coherence length). Thus, the calculated characteristics correspond to the ballistic component of the current. For the non-self-consistent calculation the single-band effective-mass time-independent Schrödinger equation is solved. For the self-consistent calculation, the Schrödinger and Poisson equations are solved iteratively until a self-consistent electron potential energy and electron density are obtained. It is shown that suitably designed electron-wave quantum-interference filter/emitters can exhibit strong negative differential resistance in the current-voltage characteristics, similar to those of resonant tunneling diodes. For low-to-moderate (2-30 meV) Fermi energies in the conduction band of Ga_{1-x}Al_{x}As (Si doping concentration less or equal to 2 \times 10^{18} \text{ cm}^{-3}) and temperatures near 30 K (in the ballistic transport regime), it is shown that space-charge effects are relatively small and result in a slight shift of the current-voltage and transmission characteristics toward higher bias voltages. In a fashion similar to that occurring in resonant tunneling diodes, the self-consistent field in electron-wave filter/emitter negative differential resistance devices partially screens the positive applied bias. Designs of Ga_{1-x}Al_{x}As resonant devices with current peak-to-valley ratios of ~50 as well as nonresonant (not exhibiting negative differential resistance) devices are analyzed. The corresponding electron charge density distributions are also presented. Superlattice electron-wave filter/emitter negative differential resistance devices can be used as high-speed switches, oscillators, and as monoenergetic emitters in electroluminescent devices and photodetectors.

I. INTRODUCTION

Nanostructure devices like multiple quantum wells and superlattices have been rapidly developed due to recent advances of microfabrication technology, particularly metalorganic chemical vapor deposition (MOCVD) and molecular-beam epitaxy (MBE). These fabrication technologies have been refined to the point where single monolayers can be grown with precise compositional control. At these spatial dimensions, quantization effects largely dominate device characteristics, making further reduction in their sizes undesirable in some applications, thereby limiting the ultimate speed of the devices. However, quantum interference effects may also potentially be used in an advantageous way in novel devices. Quantum-mechanical interference effects have been experimentally observed for electron energies below the potential energy barriers in double-barrier and multibarrier resonant tunneling devices and for electron energies above the barriers in single- and double-barrier structures. With further improvements in the quality of the materials grown, ballistic transport has also been experimentally observed. That is, electrons travel through the device without being scattered by deviations from crystalline perfection. Even with the addition of elastic scattering, the electrons exhibit clear quantum-mechanical plane-wave behavior. These coherent waves maintain their phase throughout the device and thus can interfere, reflect, refract, and diffract in a manner analogous to the electromagnetic plane waves in dielectric media. Impressive experimental evidence of optical-like electron behavior in GaAs semiconductors has recently been reported.

Quantitative analogies between quantum-mechanical electron waves in semiconductor materials and electromagnetic optical waves in dielectrics have already been developed. With these analogies, existing electromagnetic optical analysis and design techniques can be used for the analysis and design of novel semiconductor quantum-wave devices. Possible devices include narrow-band superlattice interference filters and filter/emitters which can exhibit very narrow electron kinetic energy passbands and can be integrated into solid-state devices for potential use as monoenergetic emitters for electroluminescent devices, photodetectors, and subpicosecond ballistic transistors. Beyond improving the speed of existing devices, however, the totally new concept of guided electron-wave integrated circuits has been proposed.
potentially next generation of integrated circuits could be comprised by many semiconductor quantum electron-wave devices interconnected by electron waveguides.24-29

Semiconductor superlattice interference filter and filter/emitter designs can be visualized from the optical interference filters counterparts but their designs cannot be simple copies of thin-film optical designs due to the constraints that are imposed by the ultrasmall dimensions, the usable composition range, and the applied bias voltage. The design of such quantum interference filters and filter/emitters has been presented in Refs. 17-20. In the present work the current-voltage (I-V) characteristics of these filter/emitters are calculated self-consistently. It is shown that these devices can exhibit negative differential resitivity similar to that occurring in resonant tunneling diodes. In addition, space-charge effects due to electron localization that have been shown to be very important in semiconductor device analysis and design are quantified for these structures. Several analytical techniques have been used and can be classified into two major categories: the classical approaches that include the solution of the non-linear Poisson equation using Boltzmann or Fermi–Dirac statistics for the electron and hole concentrations,30-35 and the quantum-mechanical approaches in which both the Schroedinger and Poisson equations are solved self-consistently through an iterative algorithm until a self-consistent electrostatic potential and electron density are achieved.36-47 Most of these analyses have been applied to the resonant tunneling diodes37,39-41,44 in which the space-charge effects have been shown to be significant. In this paper, for the first time, a complete quantum-mechanical self-consistent analysis of the semiconductor superlattice filter/emitter is presented. These devices inherently differ from the resonant tunneling diodes in that they are designed to operate at electron energies above all barriers and wells of the semiconductor superlattice. Thus, electron localization is purely due to the quantum-mechanical interference mechanism.

In this work, the effects of the self-consistent potential on quantum-interference filter/emitter negative differential resistance devices are isolated and analyzed. The details of the model used are described in Sec. II. Several example cases are presented in Sec. III for the GaAs/Ga0.75Al0.25As alloy system. Resonant (exhibiting negative differential resistance) and nonresonant (not exhibiting negative differential resistance) filter/emitter devices are analyzed. The effects of the Fermi energy (doping) and temperature on the filter/emitter transmittance and current-voltage characteristics are included. Electron densities in the superlattice are also presented for resonant and nonresonant filter/emitters. It is found that the space-charge effects on the transmission and I-V characteristics of the filter/emitter devices can be significant for relatively high Fermi energies (above 30 meV) and generally lead to shifting of the device characteristics toward higher bias voltages. Finally, in Sec. IV, a summary and some discussion are presented.

II. MODEL DESCRIPTION

A voltage-biased semiconductor superlattice can serve simultaneously as an electron-wave interference filter and electron emitter. The theory and the design of such electron-wave filter/emitters have been presented in Refs. 19 and 20 using a non-self-consistent (linear electrostatic potential) model. The present self-consistent model is similar to the one used for the resonant tunneling devices.36-45 The model formulation is based on the following given conditions:

(a) The semiconductor superlattice consists of an intrinsic region where quantum-mechanical analysis applies. Thus, the length of the device is smaller than the electron coherence length. Experimental measurements in ballistic hot-electron devices16,11,14,15,48,49 (on GaAs/GaAlAs and InGaAs/InAlAs structures) indicate that the electron coherence length lies roughly between 10 and 100 nm. Of course, the electron coherence length depends strongly on the temperature (electron coherence length of the order of several micrometers was recently reported at very low temperatures14,45) and it is a statistical quantity. Therefore, the experimental data suggest that a measurable fraction of the electrons will exhibit coherent behavior within the filter/emitter. The transport of electrons through the filter/emitter negative differential resistance device is assumed to be strictly collisionless.

(b) For the filter/emitter region, the one-electron single-band effective mass time-independent Schroedinger equation is used. This is a valid approximation for electron energies near the \( \Gamma \) point in accordance with the filter/emitter design requirements.20 The boundary conditions used at the heterojunction interfaces of the filter/emitter region (intrinsic region) are the continuity of the envelope wave function, \( \psi \), and the continuity of the normal component of the probability current density or equivalently the continuity of the envelope wave function first derivative weighted by the inverse effective mass, \( (1/m^*)\psi \) (where \( \psi \) denotes a first derivative). These two boundary conditions work sufficiently well for the III-V semiconductor family as long as the same band edge is considered on both sides of the discontinuity.30 (The conduction-band direct-gap minimum for the electron-wave filter/emitters).

(c) The applied voltage bias appears across the filter/emitter region. The voltage drop in the contacts is assumed negligible.

(d) The contact regions outside the filter/emitter are treated as having constant Fermi energy levels and equilibrium Fermi–Dirac statistics apply.

The electron potential energy of the electron-wave filter/emitter negative differential resistance device with an applied bias voltage is shown in Fig. 1. In this figure the electrostatic potential energy is assumed to be linear which is the first estimate used in order to start the iterative solution of the Schroedinger and Poisson equations. The filter/emitter consists of \( M \) layers, where \( M \) is an odd integer. The center layer is the resonant layer while the surrounding layers are quarter electron-wavelength layers (for the design pass energy of the filter/emitter).19,20 When the design voltage is applied, electrons in a narrow energy
FIG. 1. Schematic representation of a biased semiconductor superlattice electron-wave interference resonance filter/emitter. At the design potential energy bias, \( V_{\text{bias}} \), and input kinetic energy, \( KE_{\text{in}} \), the layers have a thickness of a quarter (or a half for the resonant central layer) of an electron wavelength as measured in that layer. In this figure a linear (non-self-consistent) electrostatic potential is assumed.

FIG. 2. Part of the resonant filter/emitter region between points \( z_{i-1} \) and \( z_i \), where the stair-step representation is used for the self-consistent potential \( V(z) \).

One-electron envelope function time-independent Schroedinger equation for the \( i \)th subregion can be written

\[
\frac{\hbar^2}{2m_i^*} \frac{d^2 \psi}{dz^2} + (E - V_i - v_i) \psi_i = 0, \quad z_{i-1} < z < z_i;
\]

(1)

where \( \hbar \) is Planck's constant divided by \( 2\pi \) and \( E \) is the total electron energy. In order to solve Eq. (1), knowledge of the electrostatic potential energy \( V(z) \) is required. Assuming that \( V(z) \) is known, the solution for \( \psi_i \) can be written in the form

\[
\psi_i(z) = C_i \exp\left( + jk_i(z - z_{i-1}) \right) + D_i \exp\left( - jk_i(z - z_{i-1}) \right)
\]

(2)

for \( z_{i-1} < z < z_i \) and with the electron wave vector \( k_i \) = \([2m_i^*(E - V_i - v_i)/\hbar^2]^{1/2} \). The coefficients \( C_i \) and \( D_i \) of each subregion \( i \) can be found from the boundary conditions for the envelope wave functions and assuming that \( \psi_0(z) = \exp(\pm jk_0z) + r_e \exp(\mp jk_0z) \) and \( \psi_{N+1}(z) = r_e \exp(\pm jk_0z) \), where \( r_e \) and \( t_e \) are the amplitude reflection and transmission coefficients of the wave function. Using a transfer matrix approach, \( r_e \) and \( t_e \) can be calculated first, and then all \( C_i \) and \( D_i \) can be computed recursively. Thus, once the overall transmission and reflection coefficients are found, the steady-state wavefunction \( \psi(z) \) for \( z_{i-1} < z < z_i \) \((i = 1,2,...,N) \) can be determined at each incident energy \( E \) over the full range of interest.

After the wave function \( \psi(z) \) is known throughout the entire structure, the electron density, \( n(z) \), can be calculated by \( n(z) = \langle \psi(z) \psi^*(z) \rangle \), where \( \langle \cdot \rangle \) denotes summation (or equivalently integration) over all energies.\(^{36}\) Furthermore, assuming Fermi–Dirac statistics the previous equation becomes

\[
n(z) = \langle \psi(z) \psi^*(z) f(E,E_f) \rangle,
\]

(3)

where \( f(E,E_f) \) is the Fermi distribution function and \( E_f \) is the Fermi energy level. The distribution function is taken as being in equilibrium, in spite of the fact that a current...
flows through the filter/emitter when a bias voltage is applied. However, the distribution function \( f(E,E_f) \) = \( 1/(1 + \exp[(E - E_f)/k_B T]) \), where \( k_B \) is Boltzmann's constant and \( T \) is the absolute temperature, is reasonably accurate since the \( xy \) system is essentially decoupled from the quantization direction \( z \), and can be considered in a quasi-equilibrium state.\(^{32-43}\) Evaluation of Eq. (3) yields to the following approximate expression for \( n(z) \):

\[
n(z) \approx \frac{k_B T (m^*_n)^{3/2}}{2 \sqrt{\pi}} \frac{1}{E_z} \int_0^\infty |\psi(z)|^2 \ln[1 + \exp(E_f - E_z)/k_B T] \, dE, \tag{4}
\]

where \( E_z \) is the longitudinal component of the energy \( E \). For the above computation, parabolic band structure and independence of \( |\psi(z)|^2 \) on the transverse energy component (momentum) were assumed. None of the two previous assumptions is strictly valid, but for low energies near the bottom of the conduction band (\( \Gamma \) valley) these are very good approximations and they greatly simplify the computational procedure for the evaluation of the electron density.\(^{37-43}\) In addition, Eq. (4) accounts for only the left-to-right incident electrons. In order to be strictly correct, the electrons impinging from right to left should also be taken into account,\(^{39-41}\) and the total electron density should be the sum of these two streams of electrons. However, the right-to-left incident stream of electrons has a negligible effect on the current density of the device for the design values of the applied voltage and temperature. The charge density corresponding to the spatial electron distribution given by Eq. (4) can be calculated by \(-en(z)\) (where \( e \) is the electron charge) and then substituted into the one-dimensional Poisson equation,

\[
\frac{d}{dx} \left( e_0 \varepsilon_r(z) \frac{d\phi}{dz} \right) = -en(z), \tag{5}
\]

where \( e_0 \) is the permittivity of free space, \( \varepsilon_r \) is the relative permittivity along the device, and \( \phi(z) \) is the electrostatic potential. Defining the electrostatic potential energy as before by \( V(z) = -e\phi(z) \), Eq. (5) can be written as

\[
\frac{d^2V(z)}{dz^2} = -\frac{e_n(z)}{e_0 \varepsilon_r}, \tag{6}
\]

where \( \varepsilon_r \) is the relative permittivity of GaAs (assumed constant throughout the filter/emitter for simplicity). Solving Eq. (6) numerically, using the method of finite differences, the electrostatic potential energy \( V(z) \) can be found and used in the Schroedinger equation [Eq. (1)].

The computation algorithm starts with the solution of Eq. (1) assuming linear electrostatic potential energy (Fig. 1). After computation of the wave function, the electron density can be calculated from Eq. (4). Then Eq. (6) can be solved for a new estimate of the electrostatic potential energy \( V(z) \). This procedure is iterated until convergence of the electrostatic potential energy \( V(z) \) is obtained.

Then, the current density, \( J \), can be calculated at each applied bias using the equation\(^{34,53}\)

\[
J = -\frac{em^*_n}{2\pi^2} \int_0^\infty T(E_z) \times \ln \left( \frac{1 + \exp[(E_f - E_z)/k_B T]}{1 + \exp[(E_f - E_z - V_{bias})/k_B T]} \right) dE, \tag{7}
\]

where \( T(E_z) \) is the transmission coefficient of the device, calculated using the self-consistent electrostatic potential, and is given by

\[
T(E_z) = \left( \frac{(E - V_0)/m^*_n}{(E - V_0 - V_{bias})/m^*_n} \right)^{1/2} |t_z|^2. \tag{8}
\]

For the derivation of Eq. (7), parabolic band structure and independence of \( T(E_z) \) on transverse energy component were assumed, similar to the assumptions for the validity of Eq. (4). In the following section several results of resonant (exhibiting negative differential resistance) and nonresonant (not exhibiting negative differential resistance) devices in the Ga\(_{1-x}\)Al\(_x\)As alloy family are presented.

### III. RESULTS OF THE CALCULATIONS FOR Ga\(_{1-x}\)Al\(_x\)As ALLOYS

A practical material system to be used in fabricating electron-wave interference devices is Ga\(_{1-x}\)Al\(_x\)As. For this material system the maximum allowable composition in aluminum is \( x_{max} = 0.45 \) in order to avoid the direct/indirect band-gap transition. The electron potential energy of the \( j \)th layer is given by \( V_j = \Delta E_j \,(A = 0.773 \text{ eV})\), using the \(-60/40\) rule conduction-to-valence band-edge discontinuity,\(^{54,55}\) and its effective mass is given by \( m^*_j = (B + Cx_j)m_0 \,(B = 0.067, C = 0.083) \), where \( m_0 \) is the free-electron mass. The \(100\) monolayer thickness for any usable composition remains the same and equals \( 0.2827 \text{ nm} \) (lattice-matched material system). Two resonant and one nonresonant example designs will be presented. These designs will take into account the minimization of the inter-valley scattering (\( \Gamma-L \) band transition) according to the analysis presented in Ref. 20. The two resonant devices will be referred to as resonant filter/emitters A and B respectively.

#### A. Resonant filter/emitter A design criteria

Resonant filter/emitter A was designed with the algorithm of Ref. 20 for a bias potential energy \( V_{bias} = 0.10 \text{ eV} \) and an output kinetic energy at resonance \( K \_E_{out} = 0.13 \text{ eV} \), assuming a Fermi energy level at the input region of \( 30 \text{ meV} \) (measured from the bottom of the conduction band). In order to reduce intervalley transitions for the applied resonance bias and kinetic energy, the maximum allowable aluminum composition had to be reduced\(^{20} \) to \( 0.2 \), which was also selected to be the input and output region compositions (Fig. 1). The operating temperature was assumed to be \( 25 \text{ K} \). The characteristics of the nine-layer \((M = 9)\) resonant filter/emitter A (exact design) are given in Table I. In this table, the thickness of each layer in
TABLE I. Design parameters of an electron-wave interference resonant filter/emitter consisting of nine layers surrounded by GaAs and AlAs and designed to emit 0.13-eV electrons for the exact design and 0.125-eV electrons for the constrained-composition design when biased at 0.10 eV.

<table>
<thead>
<tr>
<th>Layer No.</th>
<th>Exact design</th>
<th>Resonant filter/emitter A design</th>
<th>Constrained-composition design</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number monolayers thick</td>
<td>Aluminum composition $x_j$</td>
<td>Unbiased electron potential energy $V_j$ (eV)</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>0.0249</td>
<td>0.0193</td>
</tr>
<tr>
<td>2</td>
<td>12</td>
<td>0.1225</td>
<td>0.0947</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>0.0630</td>
<td>0.0487</td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>0.1557</td>
<td>0.1204</td>
</tr>
<tr>
<td>5</td>
<td>18</td>
<td>0.0402</td>
<td>0.0311</td>
</tr>
<tr>
<td>6</td>
<td>11</td>
<td>0.1712</td>
<td>0.1324</td>
</tr>
<tr>
<td>7</td>
<td>9</td>
<td>0.0855</td>
<td>0.0661</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>0.1622</td>
<td>0.1254</td>
</tr>
<tr>
<td>9</td>
<td>8</td>
<td>0.0142</td>
<td>0.0109</td>
</tr>
</tbody>
</table>

monolayers, the aluminum composition $x_j$, the unbiased electron potential energy $V_j$, and the normalized effective mass $m^*_j/m_0$ are given. The total thickness of the filter/emitter A is 100 monolayers (28.3 nm), which is within the electron coherence length [assumption (a) of Sec. II] at 25 K. An alternate design which takes into account the practical constraint of having a finite number of composition levels is also shown in Table I (constrained-composition design). This version of resonant filter/emitter A restricts the aluminum compositions to four distinct levels (0.0, 0.07, 0.13, and 0.20 as is required by typical MBE fabrication systems). The transmission characteristics (calculated using a non-self-consistent analysis) of these exact and constrained-composition designs are shown in Fig. 3(a) as functions of the output kinetic energy $KE_{out}$ (Fig. 1) for an applied bias potential energy of 0.10 eV (resonance design value). The constrained-composition design has its peak electron transmission at $KE_{out} = 0.125$ eV instead of the 0.13-eV design value due to the four distinct aluminum composition levels. Interestingly, the transmission (transmission of 99.8% at 0.125 eV) peak of the constrained-composition design is larger than the peak of the exact design (transmission of 94.4% at 0.13 eV). This is due to the more symmetric configuration of the constrained composition design. However, the exact design remains the optimal design for $KE_{out} = 0.13$ eV and for $V_{bias} = 0.10$ eV.

B. Resonant filter/emitter B design criteria

Resonant filter/emitter B was designed using the same procedure as resonant filter/emitter A. The Fermi energy level of resonant filter/emitter B was calculated to be $E_F = 2.26$ meV using a donor concentration of $N_D = 2 \times 10^{18}$ cm$^{-3}$, a donor ionization energy of 5.8 meV (for silicon dopant), and an operating temperature of 25 K. The calculational procedure was based on the numerical solution of the charge neutrality equation at flat-band equilibrium using Fermi–Dirac statistics. This is a conservative calculation which does not take into account the
broadening of the donor ionization energies and the band tailing due to the heavy doping. To restrict intervalley scattering, the maximum allowable aluminum composition was chosen again to be 0.2, which is also equal to the compositions of the input and output regions. In this design, due to the lower Fermi energy level, the design bias potential energy had to be chosen lower than previously and was selected to be 0.05 eV. The output kinetic energy of resonant filter/emitter B at its resonance was selected to be 0.052 eV in order to be near the Fermi level of the input region. The design characteristics of the nine-layer ($M = 9$) resonant filter/emitter B (exact design) are shown in Table II. For this design the total thickness of the filter/emitter is 136 monolayers (38.4 nm), which is again within the electron coherence length at 25 K. The thickness of resonant filter/emitter B is larger than that of resonant filter/emitter A due to the lower resonance energy and applied voltage bias. A constrained-composition design is also included in Table II with three only-aluminum composition levels (the 0.13 composition level is not necessary in this case). In this constrained-composition design the center layer thickness was increased by 1 monolayer (21 instead of 20) to shift the transmission peak closer to the design transmission peak and at lower output kinetic energies (which is desirable due to the low Fermi energy level). The transmission characteristics of the exact and the constrained-composition design compositions are shown in Fig. 3(b) as a function of the output kinetic energy for a bias potential energy of 0.05 eV (non-self-consistent calculation). For both of the resonant filter/emitter B designs the peak transmissions are smaller than those of resonant filter/emitter A (37.9% for the exact design and 34.4% for the constrained-composition design) for the corresponding design energy and bias potential. The smaller peak transmission is due to the lower input kinetic energies and applied resonant bias voltages. In the remainder of this paper, the constrained-composition designs for both resonant filter/emitters A and B will be analyzed, since these designs are practical from a fabrication point of view.

C. Self-consistent and non-self-consistent analyses

Using the algorithm presented in Sec. II, constrained-composition resonant filter/emitters A and B were analyzed with both the non-self-consistent (linear electrostatic potential) and the self-consistent approach. The stair-step representation used for the self-consistent computation consisted of steps of monolayer thickness, i.e., $\Delta z = 0.2827$ nm and $N = 100$ for resonant filter/emitter A and $N = 137$ for resonant filter/emitter B. Comparing the stair-step results with those of the exact analysis (using Airy and complimentary Airy functions for non-self-consistent potential) excellent agreement was found (error was smaller than $\pm 10^{-5}$ for the wave-function amplitude). Thus, the above selection of $\Delta z$ was also considered more than adequate for the self-consistent analysis.

The electron potential energy across the filter/emitter is shown as a function of the longitudinal distance ($z$) for the non-self-consistent (linear) and the self-consistent electrostatic potential in Figs. 4(a) (resonant filter/emitter A for $V_{\text{bias}} = 0.10$ eV) and 4(b) (resonant filter/emitter B for $V_{\text{bias}} = 0.05$ eV). It is observed that the difference between the self-consistent and the linear potential is larger for resonant filter/emitter A (with a Fermi energy $E_F = 30$ meV) than for filter/emitter B (with a Fermi energy $E_F = 2.26$ meV). This difference is at least an order of magnitude smaller for resonant filter/emitter B. The difference between the self-consistent and the linear electrostatic potential energy, $\Delta V = V_{\text{SC}}(z) - V_L(z)$, is shown in Figs. 5(a) (resonant filter/emitter A for $V_{\text{bias}} = 0.10$ eV) and 5(b) (resonant filter/emitter B for $V_{\text{bias}} = 0.05$ eV). Since resonant filter/emitter A has more carriers available and a larger transmission peak at the resonant energy, stronger electron localization is expected, which raises the potential energy profile for the transmitted electrons higher than that of filter/emitter B. The wave-function square amplitude, $|\psi(z)|^2$, as function of the longitudinal ($z$) direction, is shown for both the

<table>
<thead>
<tr>
<th>Layer No.</th>
<th>Number monolayers thick</th>
<th>Aluminum composition $x_i$</th>
<th>Unbiased electron potential energy $V_i$ (eV)</th>
<th>Normalized effective mass $m^*/m_0$</th>
<th>Number monolayers thick</th>
<th>Aluminum composition $x_i$</th>
<th>Unbiased electron potential energy $V_i$ (eV)</th>
<th>Normalized effective mass $m^*/m_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>11</td>
<td>0.0226</td>
<td>0.0175</td>
<td>0.0689</td>
<td>11</td>
<td>0.00</td>
<td>0.0000</td>
<td>0.0670</td>
</tr>
<tr>
<td>2</td>
<td>21</td>
<td>0.1702</td>
<td>0.1316</td>
<td>0.0811</td>
<td>21</td>
<td>0.20</td>
<td>0.1546</td>
<td>0.0836</td>
</tr>
<tr>
<td>3</td>
<td>11</td>
<td>0.0419</td>
<td>0.0324</td>
<td>0.0705</td>
<td>11</td>
<td>0.07</td>
<td>0.0541</td>
<td>0.0728</td>
</tr>
<tr>
<td>4</td>
<td>20</td>
<td>0.1814</td>
<td>0.1402</td>
<td>0.0821</td>
<td>20</td>
<td>0.20</td>
<td>0.1546</td>
<td>0.0836</td>
</tr>
<tr>
<td>5</td>
<td>20</td>
<td>0.0140</td>
<td>0.0109</td>
<td>0.0682</td>
<td>20</td>
<td>0.00</td>
<td>0.0000</td>
<td>0.0670</td>
</tr>
<tr>
<td>6</td>
<td>17</td>
<td>0.1819</td>
<td>0.1407</td>
<td>0.0821</td>
<td>17</td>
<td>0.20</td>
<td>0.1546</td>
<td>0.0836</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td>0.0348</td>
<td>0.0269</td>
<td>0.0699</td>
<td>10</td>
<td>0.07</td>
<td>0.0541</td>
<td>0.0728</td>
</tr>
<tr>
<td>8</td>
<td>16</td>
<td>0.1866</td>
<td>0.1443</td>
<td>0.0825</td>
<td>16</td>
<td>0.20</td>
<td>0.1546</td>
<td>0.0836</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>0.0513</td>
<td>0.0397</td>
<td>0.0712</td>
<td>10</td>
<td>0.00</td>
<td>0.0000</td>
<td>0.0670</td>
</tr>
</tbody>
</table>
non-self-consistent and the self-consistent approaches in Figs. 6(a) (resonant filter/emitter A) and 6(b) (resonant filter/emitter B). In both cases the square amplitude of the wave function was calculated at the resonant energy of the non-self-consistent design. Consequently for resonant filter/emitter A the wavefunction was computed for $KE_{\text{out}} = 0.0125$ eV (or total electron energy $E = V_0 + KE_{\text{out}} = 0.28$ eV) and $V_{\text{bias}} = 0.10$ eV, while for filter/emitter B it is computed for $KE_{\text{out}} = 0.0536$ eV (or $E = 0.21$ eV) and $V_{\text{bias}} = 0.05$ eV. From Fig. 6(a) it is observed that the square of the wave-function amplitude peaks at the boundaries between layers of differing compositions. In addition, the highest peaks appear at the boundaries of the central (resonant) layer. The square of the wave-function amplitude is not exactly symmetric around the central (resonant) layer due to the asymmetry of the filter/emitter in layer thicknesses. For the self-consistent calculation it can be seen from Figs. 6(a) and 6(b) that the wave functions corresponding to the self-consistent potentials have lower and slightly shifted peaks [Fig. 6(a)] since the self-consistent electrostatic potential is expected to shift the filter/emitter characteristics to higher applied voltages. The self-consistent effect is more pronounced in the case of resonant filter/emitter A since for this design a higher Fermi energy level ($E_F = 30$ meV) was assumed, and consequently the screening effect was stronger. Comparing the square of the wave-function amplitudes of these filter/emitters with those of resonant tunneling diodes, a fundamental difference is observed. The peaks of the square of the wave-function amplitude in the resonant tunneling diodes occur inside the well region (resonant layer), while in the resonant filter/emitter cases the peaks occur at the boundaries of the central (resonant) layer with the surrounding layers. Thus, electron localization is expected near and at the interfaces between differing aluminum composition layers and especially near the interfaces of the central layer. This may be due to higher electron velocities inside the well regions than inside the barrier regions, re-
resulting in an electron localization near the interfaces between the differing electron velocity regions.

The electron-wave transmission characteristics of the resonant filter/emitters can be calculated using Eqs. (1), (2), and (8) for both the non-self-consistent and the self-consistent electrostatic potential. The results of the calculation are shown in Figs. 7(a) (resonant filter/emitter A for $V_{bias} = 0.10$ eV) and 7(b) (resonant filter/emitter B for $V_{bias} = 0.05$ eV). In both self-consistent electrostatic potential cases there is a slight shift of the transmittance curve toward higher output kinetic energies (or equivalently input kinetic energies) from its corresponding linear electrostatic potential case. This was expected due to the increased electron potential energy. The effect is again more pronounced for resonant filter/emitter A due to its higher Fermi energy.

The electron density, $n(z)$, can be calculated from Eq. (4) using the self-consistent electrostatic potential. The electron densities of the two resonant filter/emitters are shown in Figs. 8(a) (resonant filter/emitter A for $V_{bias} = 0.10$ eV) and 8(b) (resonant filter/emitter B for $V_{bias} = 0.05$ eV) as functions of the longitudinal distance ($z$). Comparisons of Figs. 6 and 8 reveal some similarity between the $|\psi(z)|^2$ and $n(z)$. This is expected since $n(z)$ is an energy average over all electrons using Fermi–Dirac statistics [Eq. (3)], and to this average, electrons with energies near the resonant energy (Figs. 6) mainly contribute.
Finally the current-voltage characteristics of the resonant filter/emitters can be calculated using Eq. (7) with either a linear or a self-consistent electrostatic potential. The $I$-$V$ characteristics of resonant filter/emitters A and B are shown in Figs. 9(a) and 9(b), respectively. The effect of the self-consistent electrostatic potential is again stronger for resonant filter/emitter A. In both cases the self-consistent characteristics are shifted (in the case of resonant filter/emitter B this shift is small) to higher applied bias voltages. This is in agreement with the $I$-$V$ characteristics of the resonant tunneling diodes.\textsuperscript{37-43} From the $I$-$V$ characteristics a negative differential resistivity (NDR) is clearly observed. For resonant filter/emitter A the peak-to-valley ratio is 5.74 (with peak current density $J = 3.22 \times 10^6$ A/cm$^2$ at 0.155 V) for the self-consistent calculation, and 5.78 for the non-self-consistent calculation (with peak current density $J = 3.20 \times 10^6$ A/cm$^2$ at 0.135 V). For resonant filter/emitter B the peak-to-valley ratio is 46.1 (with peak current density $J = 1.09 \times 10^6$ A/cm$^2$ at $-0.0525$ V) for the self-consistent calculation, and 47.2 with peak current density $J = 1.11 \times 10^3$ A/cm$^2$ at 0.0525 V) for the non-self-consistent calculation. It is observed that filter/emitter B has about three orders of magnitude smaller current density even if its Fermi energy is only one order of magnitude smaller than that of resonant filter/emitter A. However, resonant filter/emitter B exhibits an order of magnitude larger peak-to-valley ratio. All the above calculations were performed for $T = 25$ K and for a relative dielectric permittivity of GaAs of 13.18.

D. Effects of temperature variation

The temperature dependence of resonant filter/emitter B current-voltage characteristics is shown in Fig. 10 for $T = 25, 35$, and 50 K. The Fermi energy levels were computed\textsuperscript{36} as described previously for resonant filter/emitter B and are $E_F = 2.26, 3.63,$ and 5.26 meV, respectively. The self-consistent as well as the non-self-consistent $I$-$V$ characteristics are shown for the three selected temperatures. For higher temperatures the electron coherence
length is expected to be smaller than the dimensions of the filter/emitter, thus significantly degrading the quantum interference behavior of the device due to scattering mechanisms.

E. Analysis of nonresonant filter/emitters

The two resonant filter/emitters A and B described previously have been designed to exhibit electron transmission peaks for a specified input (or output) kinetic energy and applied bias potential. A question that now arises is how a nonresonant filter/emitter would perform. In order to answer this question, resonant filter/emitter B is redesigned so that the central layer is a quarter-electron wavelength layer instead of a half-electron wavelength layer. The resulting nonresonant filter/emitter C (Table III) has a 10-monolayer-thick central layer (instead of a 21-monolayer-thick central layer for the resonant filter/emitter B). Consequently, the nonresonant filter/emitter C is not expected to exhibit any negative differential resistance behavior. The rest of the layers of nonresonant filter/emitter C have the characteristics shown in Table II for the resonant filter/emitter B. The electron potential energy profile of the nonresonant filter/emitter C is shown in Fig. 11(a) for both the non-self-consistent and self-consistent computation. The difference $\Delta V$ between the self-consistent and the linear electrostatic potential energy is shown in Fig. 11(b). It is observed that this difference is approximately two orders of magnitude smaller than that of the resonant filter/emitter B. Thus, the space-charge effects are much less significant in the nonresonant design. The square of the wave-function amplitude, $|\psi(z)|^2$, for an output kinetic energy of 0.0536 eV (resonant energy of the resonant filter/emitter B) and for an applied bias potential energy of 0.05 eV (resonant bias potential energy of the resonant filter/emitter B) is shown in Fig. 11(c). Comparing Fig. 11(c) with Fig. 6(b) a significant difference can be observed. In the nonresonant filter/emitter C the wave-function square amplitude drops fast along the longitudinal direction of the device, revealing very small electron localization inside the nonresonant filter/emitter. The difference between the self-consistent and the non-self-consistent calculation is insignificant for the nonresonant device. The electron transmission characteristics of the nonresonant filter/emitter C are shown in Fig. 11(f). No negative differential resistance is observed in this case as was expected [compare with Fig. 9(b) of the resonant filter/emitter B]. The electron density distribution, $n(z)$, is shown in Fig. 11(e). This distribution resembles the $|\psi(z)|^2$ distribution since for all the electron energies near the Fermi level the filter/emitter exhibits the same nonresonant electron transmission response. Finally, the $I-V$ characteristics of the nonresonant filter/emitter C are shown in Fig. 11(f). No negative differential resistance is observed in this case as was expected [compare with Fig. 9(b) of the resonant filter/emitter B]. In addition, the current density is much smaller (2–3 orders of magnitude) for voltages up to 0.10 V since most of the incident electrons are reflected due to the nonresonant behavior of the filter/emitter. All the computations for the nonresonant filter/emitter C were performed again for a temperature of 25 K.

IV. DISCUSSION AND SUMMARY

In all the example cases presented for both the resonant and the nonresonant filter/emitters, the step size $\Delta z$ of the stair-step approximation was selected to be one monolayer. Smaller step sizes (fraction of a monolayer) have also been tested but the final results were almost unaffected (observed changes were smaller than $10^{-6}$). Therefore, the 1-monolayer thickness was retained for savings and efficiency of computer time. The number of energy points (in momentum space) for the integral computation [Eqs. (4) and (7)] was varied between 400 and 1000 points depending on the electron transmission characteristics of

![FIG. 10. Temperature dependence of the I-V characteristics of resonant filter/emitter B for $T=25$, 35, and 50 K. Both self-consistent (solid lines) and non-self-consistent (dotted lines) characteristics are included.](image-url)
FIG. 11. Characteristics of the nonresonant filter/emitter C. (a) The electron potential energy profile; self-consistent potential (solid line) and non-self-consistent (dotted line) potential energy profile along the filter/emitter region for $V_{\text{bias}} = 0.05$ eV. (b) The difference $\Delta V$ between the self-consistent and non-self-consistent (linear) electron potential energy profiles along the filter/emitter region for $V_{\text{bias}} = 0.05$ eV. (c) The square of the wave-function amplitude distribution for $V_{\text{bias}} = 0.05$ eV and $KE_{\text{bias}} = 0.0536$ eV. The dashed lines represent boundaries between the layers of differing aluminum composition of the filter/emitter. (d) The electron transmission characteristics for both the self-consistent (solid line) and the non-self-consistent (dotted line) electron potential energy profile for $V_{\text{bias}} = 0.05$ eV. (e) The self-consistent electron density distribution along the nonresonant filter/emitter for $V_{\text{bias}} = 0.05$ eV. The dashed lines represent boundaries between different aluminum composition layers of the filter/emitter. (f) Current-voltage characteristics of the nonresonant filter/emitter using the self-consistent (solid line) and the non-self-consistent (dotted line) electron potential energy profile.
the filter/emitter. For example, resonant filter/emitter B
needed more energy points than resonant filter/emitter A
due to its sharper energy (momentum) space characteris-
tics. Convergence of the algorithm was obtained within
3–25 iterations for resonant filter/emitter A, and 2–6 itera-
tions for resonant filter/emitter B. The convergence cri-
terion selected was that the maximum absolute differ-
ence of the electrostatic potential energy between two successive
iterations should be less than a preselected value. This
value for the examples presented was set equal to $10^{-4}$. All
computations were performed in a CDC-Cyber 855 or 990
mainframe computer with CPU times varying between 400
and 45 000 s, with the most time-consuming cases being
the I-V characteristics computations (using 80 voltage val-
ues).

In the analysis presented in this paper, only states cor-
responding to propagating electrons are considered. In ac-
tual devices, it is possible for electrons injected from the
contacts to scatter inelastically into lower-energy states
that could be bound states for the superlattice structure.
However, due to the low temperature, the small longitudi-
dinal dimensions (compared to the electron coherence
length), and the absence of doping, inelastic scattering can
be neglected and only the ballistic component of the cur-
rent considered. Thus, in a ballistic analysis like the one
presented in this paper, the link between propagating and
bound states has been removed. However, the contribution
of the bound (localized) states could be significant in a
self-consistent solution since the electrostatic potential
would be altered by the bound electrons. For this reason
the exact eigenstates of the bound electrons were calculated
for both filter/emitters A and B. The bound electronic
state energies can be determined exactly, under the effec-
tive mass approximation, through the solution of the ap-
propriate eigenvalue equation for the superlattice. Using
this approach for the unbiased filter/emitter A the follow-
ing four eigenstates were found: $E_0 = 0.0530$, $E_1 = 0.0812$,
$E_2 = 0.1053$, and $E_3 = 0.1304$ eV, where all energies are
measured from the bottom of the conduction band of
GaAs. However, under the design operation bias of 0.10 V
the number of bound eigenstates of the superlattice reduces
to two: $E_0 = 0.1066$ and $E_1 = 0.1143$ eV. In order to deter-
mine the resonant strength of these localized states the
electron transmissivity was computed over the total range
of energies. It was found that under the design bias voltage
the contribution of the eigenenergy $E_0 = 0.1066$ eV eigen-
state was very small (transmissivity of less than 7%) and
of the eigenenergy $E_1 = 0.1143$ eV was not observable. The
small transmissivity reveals that the localized eigenwave
functions cannot be efficiently excited, thus making the
electron localization in these states very small and con-
sequently of negligible effect on the electrostatic potential
distribution and the self-consistent solution presented for
filter/emitter A. Similarly, for the unbiased filter/emitter B
the following five eigenstates were calculated: $E_0 = 0.0519$, $E_1 = 0.0925$, $E_2 = 0.0973$, $E_3 = 0.1259$, and $E_4 = 0.1353$ eV. Under the designed operation bias of 0.05 V,
the number of bound eigenstates reduces to four: $E_0 = 0.0749$, $E_1 = 0.1001$, $E_2 = 0.1396$, and $E_3 = 0.1457$ eV.

Again using the electron transmissivity calculation
method, the maximum transmissivity was found for the
$E_3 = 0.1457$ eV eigenstate and was less than 7%. The exci-
tations of the remainder of the localized eigenenergies were
negligible. Again this reveals that these localized states can
not be efficiently excited to produce any significant local-
ized space-charge effects due to the bound states. Thus the
presented self-consistent calculations are not altered. Sim-
ilar results were also calculated for the nonresonant filter/
emitter C. More detailed information about the calculation
of the bound eigenstates and their resonant strengths will
be presented in a future publication. If strong localized
states exist in designed filter/emitter devices, then the elec-
tron density of Eq. (4) would be replaced by

$$n(z) = \sum_i n_{bound,i}(z),$$

where $n_{prop}(z)$ is given by Eq. (4) and $n_{bound,i}(z)$ ($i$th
bound state contribution) for sufficiently long times
approaches $m^*(z)k_BT$ $\psi_i(z)|^2$

$$n_{bound,i}(z) = \frac{m^*(z)k_BT}{\pi} \times \ln|1 + \exp[(E_i - E_j)/k_BT]|,$$

where $E_i$ and $\psi_i(z)$ are the bound eigenenergy and the
corresponding eigenwave function, and $m^*(z) = m$, when
$z$ is within the $i$th layer. The above expression is approxi-
mate in that it uses Fermi-Dirac statistics inside the su-
perlattice, and these are valid only under equilibrium.
However, in cases where the bound states can be strongly
excited, Eq. (9) can give more accurate results than Eq.
(4). A remedy in the case of strong localized states would
be the inclusion of a depletion region on the collector side
so as to prevent electrons from being backscattered into the
structure. The presence of a sizable electric field produced
within the depletion region would serve to redirect the
carrier's momentum in the forward direction after a pos-
sible scattering, thus greatly reducing backscattered injec-
tion. This is similar to the action of the collector in a
junction transistor biased in the active mode. Of course,
bound states in the structure can be eliminated altogether
by selecting GaAs to be the material of the output region.
This does not affect the design of the filter/emitter. In this
case the bound states cannot exist in the superlattice.

The analysis presented describes a purely quantum-
mechanical transport and is valid for distances that are less
than the electron coherence length. In all the examples
analyzed, it was assumed that the semiconductor materials
had parabolic bandstructure. This was a good approxima-
tion since both filter/emitters have been designed to reso-
nate at low electron energies. However, for sufficiently high
applied bias voltages, the electron energies exceed the par-
obolic bandstructure regime, and the band structure is
more appropriately treated as nonparabolic. Furthermore,
the band structure may vary with the direction of the elec-
tron-wave propagation (anisotropy). Both of these effects
can be incorporated by using an energy-dependent aniso-
tropic effective mass. In the latter case the solutions of the Schrödinger equation have to be modified but the remainder of the analysis is still valid. The anisotropic case could be used when the lowest-energy band consisted of ellipsoidal $X$ minima energy surfaces (Si, AlAs, etc.) or ellipsoidal $L$ minima energy surfaces (Ge) rather than the spherical $\Gamma$ minimum energy surface associated with the direct gap Ga$_1$-$_x$Al$_x$As.

Finally, in the analysis and the examples presented, the effect of the space-charge field due to the contact-filter/emitter junctions was not taken into account. The present analysis has focused on the filter/emitter region design and function. The impinging electrons (from the left-hand side of Fig. 1) were assumed to originate from the quasi-equilibrium contact regions and are incident on the filter/emitter. The self-consistent potential will not be significantly altered by the contacts in shape or relative magnitude. However, the contacts will cause an elevation of the electron potential energy profile of the whole filter/emitter structure due to the inevitable alignment of the Fermi energy levels at equilibrium. For this reason a specialized transit region between the contacts and the filter/emitter will be required for the designed operation of the device. The design and fabrication of this specialized region and its effect in the function of the filter/emitter is under investigation.

In summary, for the first time to our knowledge, the transmission and the current-voltage characteristics of electron-wave quantum-interference resonant filter/emitters were studied using a non-self-consistent and a self-consistent analysis. It was shown that the resonant filter/emitters can exhibit negative differential resistivity similar to that in resonant tunneling diodes. Several example designs on Ga$_1$-$_x$Al$_x$As alloys were presented including both resonant and nonresonant filter/emitters. For low temperatures (in the ballistic transport regime) it was shown that the effect of the space-charge fields is insignificant for nonresonant filter/emitters, while it can be quite significant for resonant filter/emitters at high Fermi energy levels. The $I$-$V$ characteristics of the resonant filter/emitters are shifted to higher applied voltages due to the screening effect of the electron localization. However, the peak-to-valley ratios remained almost the same as those from the non-self-consistent computation. Importantly, in contrast to resonant tunneling diodes the valley current is lower and remains low over a larger voltage range ( "flat valley" ). The square of the wave-function amplitude distributions, electron potential energy profiles, and electron density profiles were also presented. Semiconductor superlattice electron-wave quantum-interference resonant filter/emitters have potential use as high-speed switches and oscillators and as monoelectronic emitters in electroluminescent devices and photodetectors.

This research was supported in part by a grant from the Joint Services Electronics Program under contract No. DAAL-03-90-C-0004. One of us (K.F.B.) was supported in part by a Presidential Young Investigator award from the National Science Foundation and another (E.N.G.) by a Research Initiation award from the National Science Foundation.

41. M. Cahay, M. McLennan, S. Datta, and M. S. Lundstrom, in Funda-

Quantum Interference Effects In Semiconductors: A Bibliography


Refinements in growth techniques such as molecular beam epitaxy (MBE) have produced materials with ballistic (collisionless) electron transport lengths of over a micron. Coupled with nanolithography, it is now possible to fabricate structures with both lateral and vertical dimensions on the order of the deBroglie wavelength of a ballistic electron. In these structures, quantum interference effects can dominate the electronic behavior. In view of the rapidly expanding interest and activity in this area, the following bibliography has been compiled as an introduction and study guide to this field. The papers listed describe the extensive theoretical and experimental results that have been obtained on quantum interference effects as well as discuss possible application areas. Works of a fundamental nature concerning phenomena that are basic to all semiconductor behavior have not been included. Articles on the properties and band structure of semiconductors, which are essential to a complete understanding of quantum interference effects, have not been included. Conference papers, though frequently very important, have not been included to conserve space. The papers are listed alphabetically according to the first author's surname. As in the compilation of any bibliography, numerous valuable and pertinent articles have probably been inadvertently omitted.

I. BIBLIOGRAPHY


Manuscript received January 16, 1991; revised July 8, 1991. This work was supported in part by a grant from the Joint Services Electronics Program under Contract DAAL-03-90-C-0004. G. N. Henderson was supported by an Office of Naval Research Graduate Research Fellowship and E. N. Glytsis was supported by a Research Initiation Award ECS-8909971 from the National Science Foundation. K. P. Martin was supported by the National Science Foundation under Grant ECS-8922512. The authors are with the School of Electrical Engineering and Microelectronics Research Center, Georgia Institute of Technology, Atlanta, GA 30332. IEEE Log Number 9101226.


PROCEEDINGS OF THE IEEE, VOL. 79, NO. 8, AUGUST 1991


Thomas K. Gaylord (Fellow, IEEE) received the B.S. degree in physics and the M.S. degree in electrical engineering from the University of Missouri-Rolla, Columbia. He received the Ph.D. degree in electrical engineering from Rice University, Houston, TX.

Presently he is with the Georgia Institute of Technology, Atlanta, where he is now Regents' Professor and Julius Brown Chair Professor of Electrical Engineering. He is the author of some 180 technical journal publications in the areas of solid state holography, semiconductors, grating diffraction, and optical computing.

Dr. Gaylord is a Fellow of the Optical Society of America. He is also the recipient of seven Sigma Xi research awards, the "Curtis W. McGraw Research Award" from the American Society for Engineering Education, the IEEE Centennial Medal, the "Georgia Tech Outstanding Teacher" award, and the "Georgia Tech Outstanding Doctoral Advisor" award. He received an honorary "Professional Degree of Electrical Engineer" from the University of Missouri-Rolla in 1985 and the "Engineer of the Year" award from Georgia Society of Professional Engineers in 1988.

Elias N. Glytsis (Member, IEEE) received the B.S. degree in electrical engineering and the M.Sc. degree from the National Technical University of Athens, Greece in 1982. He received the M.S. and Ph.D. degrees in electrical engineering from the Georgia Institute of Technology, Atlanta, in 1984 and 1987.

Since January 1988 he has been with the School of Electrical Engineering of Georgia Institute of Technology as an Assistant Professor. His current research interests are in electromagnetic theory of grating diffraction, optical computing, integrated-optics, thin films, electro-optic and optoelectronic devices, semiconductor quantum electron-wave devices, and effects of geomagnetic fields on power systems. He has published over 30 journal and conference publications and has been co-guest editor of a special issue of the Optical Society of America on grating diffraction.

Gregory N. Henderson (Student Member, IEEE) received the B.S.E.E. degree from Texas Tech University, Lubbock, in 1989 and the M.S. degree in electrical engineering from the Georgia Institute of Technology, Atlanta, in 1991. Currently, he is pursuing the Ph.D. degree in electrical engineering at the Georgia Institute of Technology with the support of an Office of Naval Research Graduate Research Fellowship.

His research interests include analysis, design, and experimental characterization of ballistic electron transport in semiconductor nanostructures. He has also worked in the development of optical computing processors using photorefractive materials.

Mr. Henderson is a member of the Optical Society of America, Sigma Xi, Eta Kappa Nu, and Tau Beta Pi.

Kevin P. Martin received the B.S., M.S., and Ph.D. degrees in physics from the Ohio State University, Columbus.

Presently, he is a senior research scientist in the Microelectronics Research Center at the Georgia Institute of Technology, Atlanta. His research interests include transport and defect properties in III-V semiconductors and devices. Most recently, his activities have focused on transport mechanisms in semiconductor quantum well devices. This includes tunneling processes under ideal (without scattering) and nonideal (with energy-momentum relaxation effects) conditions.

Dr. Martin is a member of the American Physical Society.

David B. Walker received the B.S.E.E. degree from Purdue University, West Lafayette, IN, in 1982.

After graduation, he went to work at the Polaroid Corporation as an electronics design engineer. He was recognized as the Outstanding Engineer of the Year at Polaroid in 1985. In 1990 he came to Georgia Institute of Technology, Atlanta, where, as a President’s Fellow, he is pursuing the Ph.D. degree in electrical engineering. His current research interests include ballistic electron transport devices and optical information processing.

Mr. Walker received the Col. Oscar P. Cleaver Award as the outstanding graduate student in 1991.

Kevin F. Brennan (Senior Member, IEEE) was born in Elizabeth, NJ, on October 18, 1956. He received the B.S. degree in physics from the Massachusetts Institute of Technology, Cambridge, in 1978 and the M.S. degree in physics and the Ph.D. degree in electrical engineering from the University of Illinois at Urbana-Champaign in 1984.

He is currently an Associate Professor with the School of Electrical Engineering at the Georgia Institute of Technology, Atlanta.

His research interests include modeling of avalanche photodiodes, high-field effects in compound semiconductors, and transport in submicrometer Si and GaAs devices.
Quantum well, voltage-induced quantum well, and quantum barrier electron waveguides: Mode characteristics and maximum current

Daniel W. Wilson, Elias N. Glytsis, and Thomas K. Gaylord
Department of Electrical Engineering and Microelectronics Research Center, Georgia Institute of Technology, Atlanta, Georgia 30332

(Received 18 April 1991; accepted for publication 22 July 1991)

It is shown that finite-potential heterostructure wells, homostructure voltage-induced wells, and homostructure barriers can act as waveguides for ballistic electrons and that waveguiding is described by a single dispersion relation and can occur at energies above all band edges. The guided mode cutoffs, electron velocity, effective mass, density of states, and ballistic current density (applicable to 2D electron gases) are presented. The maximum ballistic guided current flowing in a given direction for a 10 monolayer Ga0.75Al0.25As/GaAs/Ga0.53Al0.47As waveguide is found to be 2.3 mA per μm of waveguide width—allowing considerably greater currents than in single-mode quantum wires.

Recent experiments have demonstrated micron-length ballistic (collisionless) electron transport and ballistic electron refraction in two-dimensional electron gas (2DEG) structures. In these devices, the quantum well at the 2DEG interface acts as a slab waveguide for electron waves. Other experiments have demonstrated that "one-dimensional (1D) wires" defined in GaAs-GaAlAs heterostructures can act as channel electron waveguides. In most analyses of 2D and 1D electron waveguides, it has been assumed that the waveguide has hard-wall boundaries (infinite potential), or that the confining potential does not couple the transverse and longitudinal electron wave vector components. While the hard-wall boundary may be reasonable for the gate-voltage defined sidewalls of one-dimensional wires, it is not valid for the boundaries formed by a heterostructure. The assumption that the transverse and longitudinal wave vector components are independent leads to a parabolic energy dispersion relation that does not give proper electron wave phase matching at the boundary. In this letter, 2D electron waveguides are analyzed without any of the above assumptions.

We first consider a ballistic electron wave with energy $E$ incident at an angle $\theta_1$ on a single potential energy step which can be either a rise ($V_1 < V_2$) or a drop ($V_1 > V_2$). The effective masses are taken to vary in the same manner as the potential energies, i.e., $m^*_{1,2}$ for $V_1 < V_2$ and vice versa. This holds explicitly for the Ga$_{1-x}$Al$_x$As system, in which the conduction band potential energy $V = Ax$ and the effective mass $m^* = m_0(Bx + C)$, where $A = 0.7731$ eV, $B = 0.067$, $C = 0.083$, and $m_0$ is the free electron mass. The time-independent Schrödinger equation for the envelope wave function $\psi$ in the effective mass approximation is

$$(-\hbar^2/2m^*)\nabla^2\psi(r) + V(r)\psi(r) = E\psi(r),$$

where $j$ is the region index ($j = 1,2$). The boundary conditions are that $\psi$ and $(1/m^*)\nabla\psi$ must be continuous. Snell's law for ballistic electrons is

$$k_1 \sin \theta_1 = k_2 \sin \theta_2,$$

where $k_j = \sqrt{(2m^*_j/E_j)}(E - V_j)$ ($j = 1,2$). The onset of total internal reflection (TIR) ($\theta_j = 90^\circ$) occurs at the critical angle $\theta_{cr} = \sin^{-1}(k_j/k_1)$. TIR can occur below the critical energy $E_{cr} = (m^*_1V_2 - m^*_2V_1)/(m^*_2 - m^*_1)$ for a potential rise and above the critical energy for a potential drop. The existence of TIR for a potential drop requires differing effective masses. When TIR occurs, the wave function in region 2 decays as exp$(-\gamma_2 x)$, and the electron wave in region 1 experiences a phase shift $\phi = -2\tan^{-1}(m^*_1/m^*_2)(\gamma_2/k_1)$, where $\gamma_2 = \sqrt{\beta^2 - k_1^2}$, $k_{1x} = \sqrt{k_1^2 - \beta^2}$, and the wave vector components are $k_{1x}$ (transverse) and $\beta$ (longitudinal).

The two-dimensional (slab) electron waveguide is shown in Fig. 1 and is composed of a cover, film, and substrate with potential energies $V_\alpha$, $V_f$, and $V_s$ and effective masses $m^*_\alpha$, $m^*_f$, and $m^*_s$, respectively. The potential energy diagrams for the three electron waveguide configurations are shown in insets in Fig. 2. They are the heterostructure well ($V_f < V_s < V_\alpha$, $m^*_s < m^*_f < m^*_\alpha$), the homostructure voltage-induced well ($V_f < V_s < V_\alpha$, $m^*_f = m^*_s$), and the heterostructure barrier ($V_f > V_s > V_\alpha$, $m^*_f > m^*_s$). The guiding region in all cases is the film which has thickness $d$, and the cover and substrate are taken to be infinitely thick. Both wells and barriers can act as electron waveguides since TIR can occur for both a potential rise and a potential drop. The heterostructure well and barrier may be formed by MBE growth, and the homostructure voltage-induced well may be formed by the gate-defined sidewalls of one-dimensional wires.

An electron injected into the film can be guided by TIR if its zigzag angle $\theta$ is greater than both the film-cover and film-substrate critical angles. However, only those electrons that constructively interfere with themselves as they reflect from the cover and substrate boundaries will be guided over significant distances. This requires that the sum of the phase shifts on TIR from the cover and substrate plus the phase shift incurred due to round-trip transverse propagation through the film must equal a multiple of $2\pi$. Thus the dispersion equation is

$$2k_f x + \phi_s + \phi_c = 2n\pi,$$

where $k_{fs}$ is the transverse propagation constant in the film, and $\phi_c$ and $\phi_s$ are

---

1855 Appl. Phys. Lett. 69 (15), 7 October 1991 0003-6951/91/401855-03$02.00 © 1991 American Institute of Physics 1855
The guided modes have propagation vector $\beta$ and zig-zag angle $\delta$.

$\phi$, are the phase shifts on TIR from the cover and substrate respectively. Thus

$$k_fd - \tan^{-1} \left( \frac{m_\gamma \gamma_c}{m_\delta \gamma_f} \right) - \tan^{-1} \left( \frac{m_\gamma \gamma_c}{m_\delta \gamma_f} \right) = \nu\pi,$$  \hspace{1cm} (1)

where $\gamma_{c,s} = \sqrt{\beta^2 - k_{c,s}^2}$, $k_{c,x} = \sqrt{\beta^2 - \beta_c^2}$, and $\nu = 0,1,2,\ldots$. This equation is valid for all three types of electron waveguides. Electron waves that satisfy Eq. (1) are guided modes and have wave functions of the form $\psi_e(r) = \psi_e(x)e^{i(\beta y + \nu \pi)}$, where the mode propagation vector $\beta = \beta_c + \beta_x$ and $\nu = \nu_c + \nu_x$. The transverse wave functions of the guided modes will be evanescent in the cover and substrate, and oscillatory in the film. Note that separation of the wave function into transverse and longitudinal parts does not lead to independent wave vector components.

Cutoff of the guided modes occurs when: $\beta = 0$ or $\gamma_c = 0$. When $\beta$ decreases to zero, the electron no longer propagates down the waveguide, it merely reflects back and forth between the cover and the substrate. When $\gamma_c$ decreases to zero, the wave function is no longer evanescent, and the electron is refracted into the substrate. The condition $\gamma_c = 0$ is sufficient for guided mode cutoff since $\gamma_c < \gamma_f$ by definition. Equations for the cutoff energies $E_{c,0}$ and $E_{c,0} = 0$ can be found by substituting $\beta = 0$ and $\gamma_c = 0$ into Eq. (1).  

Equation (1) and the cutoff equations can be numerically solved to yield the $E$ vs $\beta$ dispersion curve and the cutoff energies of each guided mode $M_v$. These results are shown in Fig. 2. The Ga$_{1-x}$Al$_x$As compositions of Fig. 2 were used for consistency among the three waveguide types. The heterostructure well [Fig. 2(a)] has both a $\beta = 0$ (lower energy) cutoff and a $\gamma_c = 0$ (upper energy) cutoff. The voltage-induced well [Fig. 2(b)] has only a $\beta = 0$ (lower energy) cutoff, and the heterostructure barrier [Fig. 2(c)] has only a $\gamma_c = 0$ (lower energy) cutoff. Clearly, waveguiding can occur at energies above the band edges of all three regions ($c,f,$). Further, all of the dispersion curves are nonparabolic [Eq. (1)]. This is in contrast to the commonly used parabolic subband approximation, $E(\beta) = E_{c,0} + \hbar^2 \beta^2/(2m^*_c)$. From Eq. (1), the dispersion curves are parabolic only in the limit of infinite confinement ($V_c = \infty, V_s = \infty$).

The guided electron velocity $v_B$ is the group velocity of an electron wave packet, $v_B = (1/\hbar) (\partial E/\partial \beta)$. By differentiating Eq. (1), it can be shown that $v_B = P_j \nu_c + P_j \nu_f + P_j \nu_p$ where $P_j$ is the probability of finding the electron in region $j$, and $v_j = \hbar \beta / m_j^*$ is the velocity an electron would have in bulk material like that of region $j$ ($j = c,f,s$). Example velocity curves for the $\nu = 1$ mode in each of the three waveguide types are plotted in Fig. 3. Note that at the $\gamma_c = 0$ cutoff, $v_B = v_c$. By differentiating the velocity expression, the guided mode effective mass $1/m^*_c = (1/\hbar^2)(\partial^2 E/\partial \beta^2)$ is obtained. From Fig. 3(b), it is seen that $m^*_c$ can differ significantly from $m^*_c$, and that the guided electron can become infinitely "heavy" when $\partial^2 E/\partial \beta^2 = 0$ (usually near the $\gamma_c = 0$ cutoff). The guided mode density of states per unit area $g_B(E)$ was derived by applying periodic (Born--von Karman) boundary conditions to the guided mode wave function. This yields the
The ballistic guided current density flowing in a given region, \( J_x \), is defined as
\[
J_x = -e \sum_{\nu=0}^{\infty} \int_0^\infty \int_{-\pi/2}^{\pi/2} \beta(E) \cos(\alpha) f_{\text{bal}}(E, \alpha) dE d\alpha,
\]
where \( f_{\text{bal}}(E, \alpha) \) is the distribution of ballistic electrons in energy and angle. The ballistic electron injector determines \( f_{\text{bal}}(E, \alpha) \) in much the same way a light source determines the wavelength- and angle-dependent distribution of photons launched into an optical waveguide. Equation (2) is valid for all three types of electron waveguides. As an example, we calculate the maximum current density for a single-mode 10 monolayer \( (d = 2.8267 \, \text{nm}) \) Ga\(_{0.75}\)Al\(_{0.25}\)As/GaAs/Ga\(_{0.9}\)Al\(_{0.1}\)As heterostructure well waveguide. For this waveguide, \( E_{\text{r},0} = 0.0743 \, \text{eV} \) and \( E_{\text{v},0} = 0.2885 \, \text{eV} \) (below the \( L \)-valley minima in GaAs to avoid intervalley scattering) which yields a mode propagation energy range \( \Delta E = 0.2142 \, \text{eV} \). For maximum current, all the allowed modes propagating in the forward direction, would be filled, i.e., \( f_{\text{bal}}(E, \alpha) = 1 \) for \( E_{\text{g},0} < E < E_{\text{r},0} - \pi/2 < \alpha < \pi/2 \), and \( f_{\text{bal}}(E, \alpha) = 0 \) for all other \( E, \alpha \). This case yields \( J_x^{\text{max}} \approx 375 \, \text{mA} \) per \( \mu \text{m} \) of waveguide width. This value can be roughly compared to that predicted by the known maximum current in a single-vertical-mode (subband) hard-wall 1D wire of width \( W \), \( J_{1D}^{\text{max}} = N e \Delta E / (\pi \hbar) \), where \( N = 163.5 \) and \( J_{1D}^{\text{max}} = 2.7 \, \text{mA} \) which is in close agreement with \( J_x^{\text{max}} \). The sizeable value of \( J_x^{\text{max}} \) makes it feasible to interconnect multiple ballistic electron devices with a single slab waveguide, as opposed to using multiple low-current-capacity 1D quantum wires.

Electron waveguides are potentially useful in high-speed electronic circuitry. They could also be a central component in future integrated electron guided-wave circuits similar to present-day integrated optical circuits. This research was sponsored by the Joint Services Electronics Program under Grant No. DAAL-03-90-C-0004 and by the National Science Foundation Grant No. ECS-9111866. One of us (E.N.G.) was supported by a Research Initiation Award from the National Science Foundation.

BALLISTIC ELECTRON EMISSION TESTING OF SEMICONDUCTOR HETEROSTRUCTURES

Gregory N. Henderson,* Thomas K. Gaylord,* Elias N. Glytsis,* and Phillip N. Furst,**
Microelectronics Research Center, Georgia Institute of Technology, Atlanta, Georgia 30332
and William J. Kaiser
Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109

(Received 11 July 1991 by R. H. Silsbee)
(In Revised Form September 19, 1991, by R.H. Silsbee)

In ballistic electron emission microscopy (BEEM) and spectroscopy, ballistic electrons are injected into a sample using a scanning tunneling microscope to probe the electrical properties of buried interfaces. In this communication, a method is proposed that uses the BEEM technique to observe the electron wave optical properties of semiconductor heterostructures. This method provides a three-terminal configuration for characterizing electron wave devices that overcomes many of the limitations encountered in other two- and three-terminal measurement techniques. Specifically, the method provides an injector, which is well isolated from the heterostructure, that injects a collimated beam of ballistic carriers with a precisely controlled energy distribution. These carriers accurately probe the quantum transmittance of a voltage-tunable electron wave interference structure, which can be designed with a light doping to minimize impurity and electron-electron scattering. A general procedure is presented for analyzing this experimental configuration based on a combination of the models used to describe BEEM and ballistic electron transport in semiconductors. Using this procedure, BEEM testing of an electron wave energy filter is modeled and clear quantum interference effects are predicted. This BEEM configuration should allow for the precise characterization of a wide range of ballistic electron transport effects such as quantum reflections from interfaces and electron wave interference effects, phenomena that are presently of wide interest.

Due to recent advances in nanostructure fabrication techniques such as molecular beam epitaxy and nanolithography, electron wave interference effects have been observed in semiconductor heterostructures for electron energies below the barriers in resonant tunneling structures [1-3] and for energies above the barriers in semiconductor heterostructures [2-7]. In all of these experiments, a degenerately doped emitter region was used as the source of ballistic electrons. The ballistic electrons were injected by tunneling through a barrier or by being emitted above a barrier. Interference effects were manifested as peaks in the current-voltage and conductance-voltage characteristics. Two of these structures [4,6] used a three-terminal configuration in order to tune the interference resonances without changing the distribution of injected electrons.

In spite of impressive experimental results that have been achieved in these electron wave interference experiments, all of these configurations suffer from three common limitations. First, since electron wave interference effects are dominant only at low temperatures (due to phonon scattering), the emitter must be degenerately doped in order to provide sufficient carriers. This degenerate doping causes increased impurity and electron-electron scattering, which have been shown to be the principal mechanisms that destroy quantum interference effects at low temperatures [8]. Also, as a consequence of heavy emitter doping, there is generally large band bending induced between the emitter and the interference structure [4]. This band bending alters the quantum transmittance of the structure, making quantitative modeling difficult. Second, it is difficult to make a third contact to the base region of the structure. The contact must be extremely small, on the order of 50-100 nm, and requires degenerate doping of the base layer [4,6], thus greatly increasing the
scattering. Third, collimation of the injected electrons has been limited in previous designs because the devices have relied on potential-induced acceleration as the collimating mechanism. Since bias voltages are constrained to be less than the threshold voltage for intervalley scattering (≈ 1 V), the degree of collimation is limited. A poorly collimated input distribution smears out the interference resonances, further obscuring the locations and amplitudes of the interference peaks.

In this communication it is proposed to use ballistic electron emission microscopy (BEEM) and spectroscopy as a diagnostic measurement technique for semiconductor electron wave heterostructures that can overcome the above limitations. This complements the previous uses of BEEM for studying buried interfaces [9,10] and scattering in the base electrode [11]. In the BEEM configuration ballistic electrons are injected with precise energy control from a scanning tunneling microscope (STM) into a metal base. These electrons propagate ballistically through the base (approximately 10-20 nm) to probe a base-semiconductor interface. By monitoring the number of electrons transmitted across the interface as a function of tip-base voltage, the electrical characteristics of the interface, such as the barrier height, can be determined. The method presented in the present communication is an extension of conventional BEEM where a multilayer electron wave structure is placed directly below the base-semiconductor interface, as is shown in Fig. 1. By monitoring the number of electrons that are transmitted through the structure as a function of tip-base voltage, the quantum transmittance of the structure can be determined. Interference resonances will appear as peaks in the conductance-voltage curves in a manner similar to other quantum interference measurements [1-7].

The BEEM diagnostic measurement technique described here is not restricted by the first two limitations of past quantum interference measurement methods. The first limitation, the degenerate doping of the semiconductor structure, is removed since the carriers are generated by the STM. This significantly reduces the impurity and electron-electron scattering, thus enhancing the quantum interference effects. A light doping (around $10^{14}$ cm$^{-3}$) of the structure can be used to control the band bending imposed by the Fermi level pinning at the base-semiconductor interface. If the device is made sufficiently thinner than the depletion width, there will be minimal band bending through the structure; the Fermi level will be pinned near mid-gap throughout. The second limitation, the difficulty of providing a base contact, is removed because the base is exposed and thus easily accessible in this configuration. Changing the base-collector voltage will allow precise tuning of the resonance energies, without affecting the energy of injection. The third limitation, poor collimation of the injected distribution, can be improved in the BEEM technique. Of the electrons incident at the base-semiconductor interface, only those within a few degrees of normal incidence are transmitted into the semiconductor, due to the small critical angle of the base-semiconductor interface [10,12]. The electrons within this small angular range are, however, refracted into the semiconductor into a broad distribution from 0 to 90° deg. This distribution is collimated in the forward direction due to the nonlinearity of the refraction process [12] and the high reflectivity for those electrons incident near the critical angle in the base [12]. If one assumes the worst case condition in which the electrons in the base are uniformly distributed in solid angle between normal incidence and the critical angle, then the normalized distribution in the semiconductor is given as $g(\theta) = [1 - r^2(\theta)] \cos \theta$, where $\theta$ is the angle of refraction (measured from the interface normal) inside the semiconductor, and $r$ is the reflectivity of the base-semiconductor interface for a given incident energy. This distribution is shown in Fig. 2 for an interface between gold and $Ga_{0.4}Al_{0.6}As$ with an energy of incidence of 100 meV above the conduction band in the $Ga_{0.4}Al_{0.6}As$. Although the electrons tunnel through the $AlAs$ layer and then directly enter the first $GaAs$ layer of the three-layer filter, the input distribution is

---

**Fig. 1** - Energy band configuration in BEEM testing of a quantum electron wave interference structure, where the semiconductor structure is directly below the gold base. Ballistic electrons injected from the STM tip are used to probe the quantum transmittance of the structure. The structure shown is an electron energy filter which is designed to transmit at 53 meV above the $\Gamma$ conduction band minimum in the output region. At this energy, the barrier and wells are quarter- and half-wavelength layers, respectively.
The collector current \( I_c \) of the structure can be calculated by integrating the product of the electron transmittance of the heterostructure \([T(E_s,E_i) = 1 - r^2(E_s,E_i)]\) and the injected electron distribution over the tip energies \( E = E_s + E_i \), where \( E_i \) is the energy associated with the component of the momentum normal to the interface, and \( E_s \) is the energy associated with the component of the momentum tangential to the interface. The electrons in the base are propagating at an angle \( \theta_i = \arctan(\sqrt{E_i/(E_s - V_i)}) \) (measured from the interface normal) which are then refracted into \( \theta_s \) in the output region, where \( V_i \) is the potential energy of the electrons in the base [12]. The integral of the collector current can be evaluated numerically as a function of tip-base voltage \( V_{tip-base} \) to calculate the current-voltage characteristic of the structure. By inserting \( T_s(E_s,E_i) \) into the collector current expression developed in Ref. [10], one achieves an expression for the collector current \( I_c \), normalized to the tip current \( I_t \),

\[
\frac{I_c}{I_t} = \frac{R \int_{E_{min}}^{E_{max}} D(E_s) \int_{E_i}^{E_{max}} T_s(E_s,E_i) f(E) dE_s dE_i}{\int_{E_{min}}^{E_{max}} D(E_s) \int_{E_i}^{E_{max}} [f(E) - f(E + eV_{tip-base})] dE_s dE_i}
\]  

(1)

where \( f(E) \) is the Fermi-Dirac distribution, \( E \) is the electron energy in the tip, \( D(E_s) \) is the tunneling probability, \( R \) is an energy-independent scattering term in the base, \( E_{min} \) is the minimum normal energy that can be transmitted to the collector due to the Schottky barrier \( \Psi_{\text{Schottky}} \) (Fig. 1), and \( E_{max} \) is the maximum energy (for a given \( E_s \)) that can be transmitted to the collector due to total internal reflection at the base-semiconductor interface [10]. The electron transmittance of the structure can be calculated using any of the methods employed to describe quantum transport such as the chain matrix extension of the envelope function approximation [14] or the Wigner function approach [15].

In order to demonstrate the utility of this method, an electron wave interference filter was designed for \( 1 \) minimum electrons in \( Ga_{0.8}Al_{0.2}As \) (Fig. 1) using the design method presented in Ref. [16]. This device is analogous to a thin-film Fabry-Perot filter in optics, where, at the design energy, the center layer is a half-wavelength resonant layer and the two adjacent layers are quarter-wavelength reflectors. The filter in Fig. 1 was designed to be resonant for a normal incidence (i.e. \( E_i = 0 \)) kinetic energy \( (KE) \) of 53 meV above the conduction band minimum in the output region, \( Ga_{0.8}Al_{0.2}As \). The resonant layer was
when illuminated by a distributed source \( \theta_2 = 0.0 \text{ deg} \) and an incidence of \( \theta_2 = 30.0 \text{ deg} \). At the design kinetic energy, \( KE = 53 \text{ meV} \), the normal-incidence transmittance of the device peaks at 80%, as is shown in Fig. 3. The transmittance at resonance is not unity due to reflections from the base-semiconductor boundary. As the angle of incidence is increased, the resonant peak shifts to higher energies (shorter wavelengths). The transmittance at \( \theta_2 = 30.0 \text{ deg} \) peaks at 70 meV which is consistent with the expected shift for this angle, as calculated from thin-film optics principles [12,17]. As the energy of injection is varied from the resonance energy, the transmittance drops. The full-width-at-half-maximum (FWHM) of the resonance peak is 16.6 meV.

By substituting the complete filter transmittance (for all angles of incidence) into Eq. (1), the current-voltage, conductance-voltage, and differential-conductance-voltage characteristics of the filter can be calculated. These results are shown in Figs. 4a, 4b, and 4c respectively for temperatures of 4.2 K and 77 K. The two peaks in the differential-conductance characteristic correspond to the two peaks in the transmittance \( T_\theta \), shown in Fig. 3. This similarity can be seen in Fig. 4c where the transmittance has been superimposed on the differential-conductance curves. In the experiment, it is likely that the amplitude of the second peak will be significantly reduced since the electrons can start entering the L minima in Ga\(_{0.8}\)Al\(_{0.2}\)As at \( V_{\text{tip-base}} - V_{\text{barrier}} = 0.16 \text{ V} \). The characteristics below this voltage, however, should closely resemble Fig. 4.

It should be noted that each differential-conductance peak occurs a voltage for which the Fermi energy in the tip is slightly above the energy of that transmittance peaks. This shift is a complicated function of the power-law dependence of the collector current near the threshold \( V_{\text{tip-base}} = V_{\text{barrier}} \) [10], the distribution of injected electrons \( g(\theta_2) \), and the transmittance of the filter. This shift is consistent, however, with the electromagnetic case in which the transmittance of an interference filter peaks at an energy that is displaced from the normal incidence resonant energy when illuminated by a distributed source [17]. For the first peak this shift amounts to 2 meV above the resonance energy, corresponding to a tip-base voltage of \( V_{\text{tip-base}} - V_{\text{barrier}} = 55 \text{ meV} \). The FWHM of the conductance peaks can be approximated as a sum of the width of the transmittance peak and the width of the derivative of the Fermi-Dirac distribution \( (\approx 2k_B T) \). For the first peak, this estimate gives a FWHM of 17.3 meV at 4.2 K and 31.2 meV at 77 K, which can be compared with the simulation values of 14.8 meV and 29.9 meV, respectively.

It has been demonstrated analytically that the primary effect of the application of a base-collector voltage to this filter structure is a linear shift of the resonance energy [18]. Thus, tuning of the base-collector voltage will shift the resonance peak in the conductance-voltage characteristic of the device. This tunability of the resonance should allow for great flexibility in the characterization of electron wave interference effects.

The analysis used to achieve the results of Fig. 4 were based on an ideal experimental configuration with no background noise current, no scattering at the base semiconductor boundary, and no scattering in the device, the three mechanisms that are most likely to reduce the magnitude of the interference peaks. Since the steps in the \( d(I_c/I_f)/dV \) curve shown in Fig. 4
Fig. 4 - Current-voltage (a), conductance-voltage (b), and differential-conductance-voltage (c) characteristics of the interference filter of Fig. 1. The peaks in the differential-conductance-voltage characteristic (c) correspond to the two peaks in the transmittance shown in Fig. 3. The differential conductance peaks occur at energies slightly higher than the transmittance peaks as discussed in the text. The first peak occurs at $V_{\text{ito-base}} - V_{\text{barrier}} = 55 \text{ meV}$, with a FWHM of 17.3 meV at 4.2 K and a FWHM of 31.2 meV at 77 K. The second peak lies above the $L$ minima in the output region and is therefore likely to have a significantly reduced amplitude.

are an order magnitude larger than previously resolved $dI/dV$ steps [10], it is likely that the signal peaks will be significantly larger than the background noise current fluctuations. The accurate agreement between the momentum-conservation theory and BEEM spectra for GaAs [10] indicate that the component of the current that scatters at the base-semiconductor interface will be small compared to the momentum conserving current. However, if there is significant scattering at the base-semiconductor interface, the primary effect is to change the distribution of electrons ($g(\theta, E)$) injected into the semiconductor. Simulations that incorporate various injected distributions show similar behavior to the results shown in Fig. 4. The locations and amplitudes of the interference peaks shift, but the interference resonances are still resolvable. The final contributor to the noise current is scattering in the interference structure. Since the structure is significantly thinner ($\approx 200 \text{ nm}$) and more lightly doped ($\approx 10^{18} \text{ cm}^{-3}$) than previous structures used to demonstrate quantum interference effects ($\approx 800 \text{ to } 1000 \text{ nm}$), it is likely that this component of the noise current will not significantly degrade the interference measurements.

The modeling of this simple ballistic electron interference filter clearly demonstrates the potential BEEM holds as a diagnostic tool for quantum structures. This method could be used in the precise analysis and characterization of other quantum interference structures such as resonant tunneling diodes [19], multilayer energy filters [16,18], impedance transformers [20], and quantum interference transistors [21]. In addition, this method could be used to investigate many other quantum transport effects that are presently of great interest. For instance, the technique could be used to analyze the quantum mechanical reflection and transmission past a single interface, demonstrating such effects as the dependence on effective mass differences across the interface and the existence of a Brewster angle [12,22]. Additionally, the method could be used to characterize the reflectance and resonance effects occurring in quarter- and half-wavelength material layers [4-6,16,18]. A complete understanding and characterization of these fundamental effects is extremely important in the design of future electron guided wave integrated circuits [12].

In conclusion, a method has been proposed that uses ballistic electron emission microscopy and spectroscopy to test and characterize quantum electron wave heterostructures. The BEEM technique provides a three-terminal testing configuration that overcomes many of the limitations of other two- and three-terminal testing configurations. A simulation of the use of this technique has been performed for an electron wave interference filter and the quantum interference
effects are shown to be clearly observable. This method can be used to characterize a variety of ballistic transport effects such as quantum reflections from material interfaces and electron wave interference effects. In addition, the method could be combined with other quantum transport measurement techniques such as magneto-transport [23] and optical [24] techniques in order analyze a wide range of quantum structures. In the future, the technique could be expanded to two-probe and multi-probe configurations in order to analyze complex ballistic transport effects such as propagation in electron waveguides [25] and ballistic electron diffraction from gratings [26].

Acknowledgements - This research was supported in part by grant no. DAAL-03-90-C-0004 from the Joint Services Electronics Program and by grant no ECS-9111866 from the National Science Foundation. One of us (G.N.H.) was supported by an office of Naval Research Graduate Research Fellowship, and one of us (E.N.G.) was supported by a Research Initiation Award from the National Science Foundation. The authors gratefully acknowledge the original suggestion of C. A. Kukkonen of the Jet Propulsion Laboratories that led to the present work.

References

Electromagnetic analogies to general-Hamiltonian effective-mass electron wave propagation in semiconductors with spatially varying effective mass and potential energy

Gregory N. Henderson, Thomas K. Gaylord, and Elias N. Glytsis
School of Electrical Engineering and Microelectronics Research Center, Georgia Institute of Technology, Atlanta, Georgia 30332

(Received 1 October 1991; revised manuscript received 10 December 1991)

It is shown that exact, quantitative electromagnetic analogies exist for all forms of the general Hamiltonian \( [R. A. Morrow and K. R. Brownstein, Phys. Rev. B 30, 678 (1984)] \), which applies to single-band effective-mass electron wave propagation in semiconductors. It is further shown that these analogies are valid for propagation in the bulk, propagation past abrupt interfaces between materials, and propagation within one- and two-dimensionally inhomogeneous materials. These results indicate that the correct form of the single-band effective-mass Hamiltonian can be determined through appropriate wave-function-amplitude-sensitive experiments. Wave-function-phase-sensitive experiments (such as the measurement of electron wave refraction directions) are not adequate to specify completely the Hamiltonian. The present analogies suggest many wave-function-amplitude-sensitive experiments that can be used to determine the correct form of the Hamiltonian. The results of the present analysis are broadly applicable to general effective-mass propagation, unlike other recent work that has treated specific cases.

Recent advances in nanostructure growth and fabrication techniques (such as molecular-beam epitaxy and nanolithography) have led to the development of semiconductor devices in which the device response is dominated by ballistic-electron (phase-coherent) transport.\(^1\)\(^\text{-}\)\(^5\) Such ballistic electrons have been reflected and refracted,\(^1\)\(^\text{-}\)\(^3\) and interfered\(^4\)\(^\text{-}\)\(^5\) in a manner analogous to electromagnetic waves in dielectrics. Based on these results, it has been shown analytically, that under the effective-mass approximation, exact, quantitative analogies can be drawn between ballistic (collisionless) electron transport in semiconductors and electromagnetic wave propagation in dielectrics.\(^6\)

These previous analogies were developed both for propagation in the bulk and for propagation past abrupt interfaces between materials.\(^6\) In developing these analogies, the electron wave boundary conditions at an abrupt interface between dissimilar semiconductors were assumed to be the conservation of the electron wave amplitude \( \psi \) and the conservation of the product of the inverse effective mass and the normal component of the gradient of the electron wave amplitude, \( \nabla \psi \cdot \hat{n} / m \). The choice of these boundary conditions is equivalent to choosing the Hamiltonian \( H \) such that

\[
H\psi = -\frac{\hbar^2}{2m(r)} \nabla \cdot \left[ \frac{\nabla \psi}{m(r)} \right] + V(r)\psi = E\psi,
\]

for electron wave propagation in a region of spatially varying effective mass \( m(r) \) and spatially varying potential energy \( V(r) \), where \( \hbar \) is Planck's constant divided by \( 2\pi \), and \( E \) is the total electron energy.\(^7\) This Hamiltonian is probably the most widely used form of the effective-mass Hamiltonian.\(^7\)\(^\text{-}\)\(^8\) There are, however, other Hermitian forms of the effective-mass Hamiltonian and each results in different boundary conditions.\(^9\)\(^,\)\(^10\) von Roos\(^9\) has suggested a Hermitian class of effective-mass Hamiltonian functions. Using this class of functions, Morrow and Brownstein\(^10\) have shown that only those Hamiltonians that lie within a subset of this class of functions have physical meaning when considering the matching of the boundary conditions across an abrupt interface. There is, however, significant disagreement as to the exact form of the Hamiltonian within this class, based on consideration of a number of specific cases.\(^11\)\(^\text{-}\)\(^16\) The purpose of the present paper, therefore, is to draw a set of exact, quantitative analogies between electromagnetic wave propagation in dielectrics and effective-mass electron wave propagation described by the complete class of Hamiltonians given by Morrow and Brownstein. These analogies will be drawn for propagation in the bulk, propagation past abrupt interfaces, and for propagation within one- and two-dimensionally inhomogeneous materials, and will be valid for whatever form of the general Hamiltonian is ultimately shown to be correct. In addition, these analogies present the specific types of experiments that can be performed to identify the correct form of the effective-mass Hamiltonian.

Morrow and Brownstein\(^10\) demonstrated that, of the general class of Hamiltonians \( (H) \) suggest by von Roos,\(^9\) only those that take the form

\[
H\psi = -\frac{\hbar^2}{2} (m(r)\partial^2 \psi / [m(r)\partial](m(r)\partial\psi)) + V(r)\psi = E\psi
\]

with the constraint

\[
2\alpha + \beta = -1
\]

have physical meaning, when considering propagation past an abrupt interface between dissimilar semiconductors. Using comparisons of more exact theories and the effective-mass theory, many authors have attempted to
deduce the values of $\alpha$ and $\beta$, resulting in a wide range of values from $\alpha = -\frac{1}{2}$ and $\beta = 0$ (Ref. 11) to $\alpha = 0$ and $\beta = -1$ (Refs. 12–14). As Morrow suggests, the determination of the correct values of $\alpha$ and $\beta$ will doubtlessly depend on experiments. Galbraith and Duggan have used photoluminescence data to show that $\alpha = 0$ and $\beta = -1$ for GaAs/Ga$_{1-x}$Al$_x$As quantum wells. Similar results have recently been reported for GaAs/Ga$_{1-x}$Al$_x$As quantum wells by Mojahedi and Osinski. However, these results are valid only for GaAs/Ga$_{1-x}$Al$_x$As heterostructures. Since the determination of $\alpha$ and $\beta$ is still an open problem in general, this paper will draw electromagnetic analogies to the general form of the Hamiltonian given in Eq. (2) for all values of $\alpha$ and $\beta$. The analogies that are drawn dictate the form experiments must take in order to determine the correct values of $\alpha$ and $\beta$.

For the Hamiltonian of Eq. (2), the boundary conditions for an electron wave at an interface are

$$m^a\psi \text{ continuous}$$

and

$$m^{a+\beta}\frac{\partial \psi}{\partial n} \text{ continuous},$$

where $\hat{n}$ is the unit vector normal to the interface. Analogously, the boundary conditions for an electromagnetic wave at an interface between two dielectrics require the continuity of the tangential component of the electric field ($\mathcal{E}$) and the continuity of the tangential component of the magnetic field ($\mathcal{H}$) across the interface. Based on this consideration, it is reasonable to look for analogies between $\Phi = m^a\psi$ and either $\mathcal{E}$ or $\mathcal{H}$. In the previous work based on Eq. (1), it was demonstrated that $\psi$ (not $\Phi$) was analogous to $\mathcal{E}$ for TE polarization and to $\mathcal{H}$ for TM polarization. The analogies of the present paper will be consistent with these analogies because Eq. (1) is the $\alpha = 0$ special case of Eq. (2) for which $\Phi = \psi$.

For bulk propagation in a homogeneous medium, an exact analogy can be drawn between $\Phi$ and both $\mathcal{E}$ and $\mathcal{H}$. In this case, the Hamiltonian for the electron wave equation [Eq. (2)] reduces to a Helmholtz equation of the form

$$\nabla^2 \Phi = -k^2 \Phi,$$

where $k^2 = 2m(E-V)/\hbar^2$. This wave equation [Eq. (6)] is exactly analogous to the Helmholtz equation for an electromagnetic wave propagating in a homogeneous dielectric of permittivity $\varepsilon$ and permeability $\mu$, where $\Phi$ is replaced by $\mathcal{E}$ for the electric-field equation and by $\mathcal{H}$ for the magnetic-field equation. In the electromagnetic case, $k^2 = \omega^2\mu\varepsilon$, where $\omega$ is the radian frequency of the wave. Since the electron wave Helmholtz equation has exactly the same form as both the electric-field Helmholtz equation and the magnetic-field Helmholtz equation, an exact analogy can be drawn between $\Phi$ and both $\mathcal{E}$ and $\mathcal{H}$. Using these analogies and the definitions given in Ref. 6, one can define a phase-refractive index for electron waves as

$$n_{ph}^E = m_r^{1/2}(E-V)^{1/2},$$

where $m_r = m / m_{ref}$ is the relative effective mass and

$$(E-V) = (E-V_1)/(E-V_{ref})$$

is the relative kinetic energy, where $m_{ref}$ and $V_{ref}$ are the effective mass and potential energy in a reference region. This electron wave phase-refractive index is analogous to the phase-refractive index for electromagnetic waves $n_{ph}^E = \sqrt{\mu \varepsilon}$, where $\mu$ is the relative permeability and $\varepsilon$ is the relative permittivity of the dielectric. With these results, phase-propagation effects, such as interference, can be analyzed using standard electromagnetic results where $\mathcal{E}$ (or $\mathcal{H}$) is replaced by $\Phi$ and $n_{ph}^E$ is replaced by $n_{ph}^E$. These results are valid for all the Hamiltonians given in Eq. (2).

The above analogies can be extended to describe electron wave propagation past an abrupt interface between materials 1 and 2 with effective masses $m_1$ and $m_2$ and potential energies $V_1$ and $V_2$, respectively. When a plane wave [the eigensolution to Eq. (6)] is incident upon such an interface, part of the wave is reflected back into region 1 and part of the wave is transmitted (refracted) into region 2. The boundary conditions [Eqs. (4) and (5)] are used to calculate the directions of propagation and the amplitudes of the reflected and transmitted waves. By substituting $\Phi = \exp(\pm ik_{1,r}r) + r \exp(\pm ik_{2,r}r)$ and $\Phi_r = t \exp(ik_{2,r}r)$ into the boundary conditions [Eqs. (4) and (5)], one finds that

$$t = \frac{2n_{amp,1} \cos \theta_1}{n_{amp,1} \cos \theta_1 + n_{amp,2} \cos \theta_2},$$

and

$$\theta_1 = \theta_2 = 0,$$

$$n_{ph,1} \sin \theta_1 = n_{ph,2} \sin \theta_2,$$

$$r = \frac{n_{amp,1} \cos \theta_1 - n_{amp,2} \cos \theta_2}{n_{amp,1} \cos \theta_1 + n_{amp,2} \cos \theta_2},$$

where the electron wave amplitude index of refraction is defined as

$$n_{amp,1} = m_r^{1/2}(E-V)^{1/2}$$

for region 1. These expressions [Eqs. (8)–(11)] are exactly the same as the analogous electromagnetic expressions for the reflection and refraction of an electromagnetic wave from an interface between dielectrics 1 and 2 with relative permittivities $\varepsilon_{r,1}$ and $\varepsilon_{r,2}$ and relative permeabilities $\mu_{r,1}$ and $\mu_{r,2}$ respectively. In the electromagnetic case, Eqs. (10) and (11) give the reflectivity and transmissivity of the electric field for TE polarization and of the magnetic field for TM polarization. Therefore, when considering propagation past an abrupt material interface, $\Phi$ is analogous to the electric field for TE polarization and to the magnetic field for TM polarization. In other words, $\Phi$ is analogous to the electromagnetic field quantity that is parallel to the interface. In the electromagnetic case, the amplitude index of refraction for region 1 has one value for TE polarization, $n_{amp,1} = \varepsilon_{r,1}^{1/2}/\mu_{r,1}^{1/2}$, and another value for TM polarization, $n_{amp,1} = \mu_{r,1}^{1/2}/\varepsilon_{r,1}^{1/2}$. Using the above results for the indices of refraction, one can construct a general set of analogies between electron wave propagation, TE-polarized electromagnetic wave propagation, and TM-
polarized electromagnetic wave propagation. This set of analogies is shown in Table I. These analogies, which have been developed for the general form of the effective-mass Hamiltonian given in Eq. (2), are valid for both propagation in the bulk and for propagation past abrupt interfaces between materials. In the case of an abrupt material interface, Eqs. (8)–(11) are valid for electron waves, TE-polarized electromagnetic waves, and TM-polarized electromagnetic waves, where the appropriate indices of refraction are used for each case.

Motivated by the results for abrupt interfaces, one can attempt to draw similar analogies for propagation within materials with general one- and two-dimensional inhomogeneities in effective mass and/or potential energy. Again, the analogy will be drawn between $\Phi$ and either $\mathcal{E}$ (for TE polarization) or $\mathcal{H}$ (for TM polarization), where the electromagnetic wave is propagating in a one- or two-dimensionally inhomogeneous dielectric. In this case, TE (TM) polarization is defined as the polarization in which the electric (magnetic) field is polarized normal to the plane containing the gradient of the inhomogeneity. In the case of such an inhomogeneity, the Hamiltonian for the electron wave [Eq. (2)] can be expanded as a wave equation for $\Phi$.

$$\nabla^2 \Phi - \left[ \frac{\nabla m_e^2(r) \cdot \nabla}{m_e^2(r)} \right] \Phi + k_0^2 m_e^2(r) [E - V(r)], \Phi = 0, \tag{13}$$

where $m_e^2(r) = m_e^2(r)/m_0^2$ is the varying relative effective mass, $(E - V(r)) = (E - V(r))/m_0$ is the varying relative kinetic energy, $m_0$ is the average effective mass, $(E - V_0)$ is the average kinetic energy, and $k_0 = \sqrt{2m_0^2(E - V_0)/\hbar^2}$ is the average wave vector of propagation in the medium. This wave equation [Eq. (13)] is exactly analogous to the wave equation for TE propagation in a one- or two-dimensionally inhomogeneous dielectric,

$$\nabla^2 \mathcal{E} - \left[ \frac{\nabla \mu_e^2(r) \cdot \nabla}{\mu_e^2(r)} \right] \mathcal{E} + k_0^2 \mu_e^2(r) \varepsilon_e(r) \mathcal{E} = 0, \tag{14}$$

where $\mu_e^2(r) = \mu_e^2(r)/\mu_0^2$ is the relative permeability modulation, $\varepsilon_e(r) = \varepsilon_e(r)/\varepsilon_0$ is the relative permittivity modulation, $\mu_0$ is the average permeability, $\varepsilon_0$ is the average permittivity, and $k_0 = (\omega^2 \mu_0 \varepsilon_0)^{1/2}$ is the average wave vector of propagation. By comparison of these wave equations [Eqs. (13) and (14)], one can see that the analogies between electron wave propagation within a one- or two-dimensionally inhomogeneous semiconductor and TE-polarized electromagnetic wave propagation within a one- or two-dimensionally inhomogeneous dielectric are the same analogies as those developed for propagation past abrupt material interfaces, which are shown in Table I. As one would expect, a similar analogy exists between electron wave propagation within a one- or two-dimensionally inhomogeneous semiconductor and TM-polarized electromagnetic wave propagation within a one- or two-dimensionally inhomogeneous dielectric,

$$\nabla^2 \mathcal{H} - \left[ \frac{\nabla \varepsilon_e(r) \cdot \nabla}{\varepsilon_e(r)} \right] \mathcal{H} + k_0^2 \mu_e^2(r) \varepsilon_e(r) \mathcal{H} = 0, \tag{15}$$

where the analogies are again given in Table I. Thus, the analogies of Table I are valid for propagation in the bulk, propagation past abrupt material interfaces, and propagation within one- and two-dimensionally inhomogeneous semiconductors. For all of these cases, standard electromagnetic analysis techniques can be used to analyze electron wave effects such as interference, propagation, reflection, refraction, and diffraction, where the analogies of Table I are used.

At this point, one might wonder whether such exact analogies exist for general three-dimensional inhomogeneities. In this case, the analogies do not hold. For general three-dimensional inhomogeneities, decoupled TE and TM polarization cannot be defined. Therefore, one cannot write scalar wave equations [like Eqs. (14) and (15)] for the electric and the magnetic field, but must use the curl equations. Since the vector field quantities are coupled, no exact analogy can be drawn between the vector electromagnetic fields and the scalar electron wave amplitude.

In conclusion, this work has shown that exact, quantitative analogies exist for all forms of the general Hamiltonian of Morrow and Brownstein. In addition, these analogies were developed for propagation in the bulk, propagation past abrupt interfaces between materials, and propagation within one- or two-dimensionally inhomogeneous materials. With these analogies, one can analyze a wide class of electron wave effects such as reflection and refraction, interference, and diffraction using well-understood electromagnetic analysis methods.

An understanding of these electron wave optical effects in semiconductors has become of increasing importance in the past few years. Recent experiments have verified that the electron wave phase index of refraction is proportional to the product of the square root of the kinetic energy and the square root of the effective mass. It is likely that, in the near future, similar experiments will be performed to verify the dependence of the amplitude index of refraction on kinetic energy and effective mass. Since the form of the amplitude refractive index is linked to the form of the effective-mass Hamiltonian (through $n_{amp} = m_0^{3/2}$), experiments that establish the power dependence of the effective mass in the amplitude index of refraction can be used to identify the correct form of the effective-mass Hamiltonian. The recent experiments on transition energies in GaAs/Ga_1-xAl_xAs...
quantum wells\textsuperscript{15,16} fit this category since the transition energies are strongly dependent on the reflectivity of the barriers (and thus strongly dependent on $\beta$).\textsuperscript{15} Due to the exact analogies to electromagnetics, it is easy to conceive of other numerous experiments (such as measuring interface reflectivity) to establish this dependence. However, regardless of the results of such experiments, the exact, quantitative analogies established in this paper remain valid. In addition, if the correct form of the Hamiltonian (for material systems other than GaAs/Ga$_{1-x}$Al$_x$As) is shown to be other than $\alpha=0$ and $\beta=-1$, the results of previous work based on this assumption (such as Refs. 6 and 17) can be simply modified using Table I, with the analysis methods remaining valid.

This research was supported in part by Grant No, DAAL-03-90-C-0004 from the Joint Services Electronics Program and by Grant No. ECS-9111866 from the National Science Foundation. One of us (G.N.H.) was supported in part by the Office of Naval Research, and one of us (E.N.G.) was supported in part by the National Science Foundation.

\textsuperscript{8}S. Datta, \textit{Quantum Phenomena} (Addison Wesley, Reading, MA, 1989), Chap. 6.