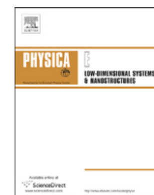
Physica E 44 (2011) 322–326

Contents lists available at SciVerse ScienceDirect

Physica E

journal homepage: www.elsevier.com/locate/physe



Optimal design of a semiconductor heterostructure tunnel diode with linear current–voltage characteristic

Kelly C. Magruder, A.F.J. Levi *

Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089-2533, USA

A R T I C L E I N F O

Article history:
Received 14 June 2011
Received in revised form
25 August 2011
Accepted 30 August 2011
Available online 6 September 2011

A B S T R A C T

Atomic layer precision in design of heterostructure tunnel diodes can be used to mimic an Ohmic (linear) current–voltage characteristic over a range of voltage bias. This is achieved by manipulating low energy electron transmission resonances. Calculations demonstrate how greater than 65 dB dynamic range can be obtained in an optimized heterostructure tunnel device that is only 17 nm thick.
© 2011 Elsevier B.V. All rights reserved.

1. Introduction

High levels of control over the input–output characteristics of electronic devices should be possible by simultaneously exploiting the wave nature of the electron on a length scale less than a characteristic inelastic mean free path, λ_{in}, and atomic layer precision that is available in fabrication of heterostructure devices. Because of the importance of Ohmic (linear) current–voltage behavior in analog and mixed signal circuits, we explore the extent to which it is possible to achieve such behavior in a semiconductor heterostructure tunnel diode design with non-equilibrium electron transport through an active region just a few nm thick.

The mechanisms responsible for linear current–voltage behavior in optimally designed tunnel diodes is different from those of conventional resistors and nanowires. Conventional resistors rely on near-equilibrium electron transport in relatively large structures ($\gg \lambda_{in}$) and dissipative processes involving the emission of phonons [1]. Nanowires rely on dimensional confinement and non-equilibrium electron transport with no scattering in the active region of the device giving a maximum electron conductance per spin of $G_n = e^2/2\pi\hbar$ [2]. Nano wires and carbon nanotubes are, by their very nature, small and so of interest for scaled electronic devices. However, inability to create low impedance contacts to nanowires has limited their utility and application to electronic devices. Depositing metal contacts directly onto semiconductor nanowires effectively forms a Schottky barrier [3] and failure to lower the contact resistance of single-walled carbon nanotubes below tens of kΩ has driven the research community to consider use of larger, and less controlled, multiple-walled and bundled nanotubes [4,5].

The design of optimized tunnel diodes considered here does not suffer from these difficulties. We use elastic electron scattering from a precisely defined potential to control the current–voltage characteristic of the device. The use of a single-crystal semiconductor heterostructure allows atomic layer control of composition and hence unique access to a wide range of conduction band potential profiles. Our work builds on previous studies of optimized electron transmission [6,7] and current density [8] by self-consistently solving the Schrödinger and Poisson equations to obtain the potential. Here we seek a proof-of-principle calculation of achievable levels of control using optimal design. In addition, physical mechanisms limiting feasibility of design are explored and we show how material parameter choices impact linearity (dynamic range).

2. Nonlinear current–voltage characteristic through single-barrier semiconductor tunnel diodes

In general, non-equilibrium electron transport through a nanoscale active region results in a nonlinear relation between current and voltage bias. As an example, consider a single-crystal semiconductor heterostructure tunnel diode for which the material consists of an intrinsic AlAs layer of thickness L_0 grown between two GaAs contacts each having n-type impurity concentration of $n = 1 \times 10^{18}$ cm^{-3}. At temperature $T = 300$ K an atomic monolayer of the [001]-oriented crystal is $\delta = 0.2827$ nm thick and the conduction band offset between AlAs and GaAs is 1.04 eV [9].

Calculation of current through the structure may be achieved by self-consistently solving the Schrödinger and Poisson equations. The Schrödinger equation is solved using the propagation matrix method [10] and the Poisson equation is solved using

* Corresponding author. Tel.: +1 213 740 7318.
E-mail address: alevi@usc.edu (A.F.J. Levi).

1386-9477/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.physe.2011.08.034

a predictor–corrector scheme [11]. Iterations between the two equations are continued until the relative error in calculated current density is less than $\gamma = 10^{-3}$, where current density is calculated using the Tsu–Esaki formula [12]. Inelastic scattering into accumulation region quasi-bound states is necessary for generating charge within this region. We assume these bound states to be in thermal equilibrium with the contact electrode [13]. To prevent depletion of the emitter contact in the absence of scattering a drifted Fermi distribution is used [14].

For simplicity we use a conduction band effective electron mass $m = 0.07 \times m_0$, where m_0 is the bare electron mass, in all heterostructure layers. We note that our model captures the essential physics of electron transport through a few nm of $Al_\xi Ga_{1-\xi}As$ heterostructure material. Inclusion of other details such as conduction band non-parabolicity [15] or attempts to include inelastic scattering mechanisms [16] when total thickness of the active region is less than λ_{in} can add significantly to the complexity of the model. More complex models introduce new ways in which current can flow through the device. However, so long as the change in current is slowly varying with voltage, it should be controllable by adjusting the conduction band potential profile of the active region and hence have little effect on the achievable degree of optimality.

Results of calculating the current–voltage characteristic of a single-barrier AlAs/GaAs tunnel diode are shown in Fig. 1 for the indicated values of L_0. For small bias voltage $|V_{bias}| < \mu$, increasing bias opens channels for current flow, causing an increase in current. At temperature $T = 300$ K, the chemical potential of electrons in the GaAs contacts is $\mu = 38.9$ meV. When $|V_{bias}| < \mu$ electron transmission probability through the tunnel barrier is weakly dependent on V_{bias}. However, once the voltage bias begins to enter the regime in which $|V_{bias}| > \mu$ the majority of current carrying states have been accessed. The slope $\partial J/\partial V_{bias}$ then decreases as changing current is now caused primarily by V_{bias} lowering the potential energy of the AlAs barrier relative to the energy of conduction band electrons in the cathode. Eventually, with increasing voltage bias and $|V_{bias}| \gg \mu$, the slope $\partial J/\partial V_{bias}$ increases *exponentially*. This change in physical mechanisms controlling electron transport should be difficult to linearize, providing a suitable test of achievable optimality.

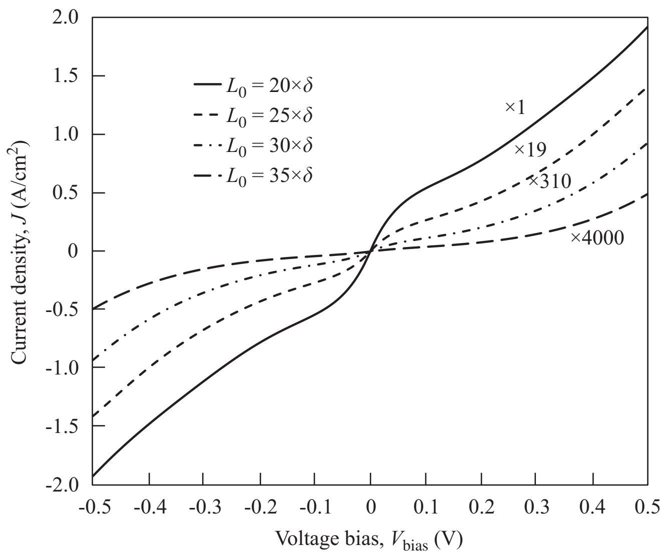


Fig. 1. Calculated current density through a single-barrier AlAs–GaAs tunnel diode of thickness L_0. Current is exponential for $|V_{bias}| > \mu$, the chemical potential of electrons in the electrode. Simulation parameters are electron effective mass $m = 0.07 \times m_0$, temperature $T = 300$ K, and n-type impurity concentration $n = 1 \times 10^{18}$ cm^{-3}. Vertical scale is multiplied by the indicated value for each curve.

3. Optimal design of heterostructure diodes for linear objective

Control necessary to linearize the current–voltage characteristic shown in Fig. 1 may be achieved by introducing $Al_\xi Ga_{1-\xi}As$ layers with varying composition. We choose a device structure consisting of one AlAs layer of thickness L_0 followed by an integer number, N, of $Al_\xi Ga_{1-\xi}As$ layers, each of thickness $L_\delta = 4 \times \delta$. Total thickness of the active region is $L = L_0 + NL_\delta$. The GaAs contacts have n-type impurity concentration $n = 1 \times 10^{18}$ cm^{-3}. In this arrangement, the AlAs layer serves as the primary current-limiter and the $Al_\xi Ga_{1-\xi}As$ layers modify the current through the AlAs layer. The relative permittivity of each $Al_\xi Ga_{1-\xi}As$ layer is assumed to vary as $\varepsilon_r = 13.2 - 3.1 \times \xi$ [17].

The optimization parameters consist of L_0 and the set of $Al_\xi Ga_{1-\xi}As$ conduction band offsets, $\{U_j\}$, where the potential energy of the jth $Al_\xi Ga_{1-\xi}As$ layer is $U_j = 0.8355 \times \xi$ and the Al alloy concentration is restricted to $0 \le \xi \le 0.42$. The cost function is defined as

$$C = \sum_{i=1}^{v} w_i (J_{obj}(V^i_{bias}) - J_{sim}(V^i_{bias}, p))^2, \quad (1)$$

where design parameters L_0 and $\{U_j\}$ are contained in vector p and w_i is a weighting factor. Eq. (1) is a least squares measure of the difference between simulated current density, J_{sim}, and the objective function, J_{obj}. This form of cost function guarantees a continuous derivative at local minima, enabling use of gradient-based methods for optimization. The linear objective function is $J_{obj} = V_{bias}/r$, where r defines the slope of the objective function with units of $\Omega \times \mu m^2$. For a device of cross-sectional area A characterized by current density $J = V_{bias}/r$, the resistance is $R = r/A$. The optimization is limited to $0 \le V_{bias} \le 0.5$ V, and auxiliary parameters for optimization are r and the number of $Al_\xi Ga_{1-\xi}As$ layers N.

Since higher optimality is associated with lower cost, the optimization is formulated as a minimization problem

$$\min_p \quad C,$$

subject to physical constraints (2)

$$0 \le U_j \le 0.35 \text{ eV}, \quad (3)$$

$$L_0 = n \times \delta, \quad n \in N^+. \quad (4)$$

Sufficient conditions for an optimal set of design parameters, p^*, are $\nabla_p C|_{p=p^*} = 0$, $\mathbf{H}_p(C)|_{p=p^*} > 0$, and $p^* \in \Omega$, where Ω is the feasible design space and $\mathbf{H}_p$ is the Hessian matrix, a matrix of second derivatives with respect to design parameters p.

Initial device configurations consist of a random set $\{U_j\}$ and an initial guess for L_0 that ensures $J(V_{bias} = 0.5\text{ V})$ is of the same order of magnitude as the objective function. Minimization of Eq. (1) is completed using the Newton–Raphson method to find a local minimum satisfying the above criteria, which requires evaluation of the gradient $\nabla_p C$. The partial derivatives $\partial C/\partial U_j$ and $\partial C/\partial L_0$ are efficiently calculated using the adjoint method [8,18]. Solution space is non-convex and so the best of multiple optimizations was selected for each objective.

4. Results

Device structures having $N = 5$, 7, 10, 12, 15, and 17 layers of $Al_\xi Ga_{1-\xi}As$ and values of r ranging from $r = 10\ \text{k}\Omega \times \mu m^2$ to $r = 100\ \text{M}\Omega \times \mu m^2$ by factors of 10 were studied. Designs for $r = 1\ \text{k}\Omega \times \mu m^2$ also yielded linear (Ohmic) behavior, but the 2–3 monolayer thicknesses of the AlAs layer in the design violated the

assumption of an equilibrium distribution of electrons in the accumulation region quasi-bound states [13].

The potential profile as a function of position x and relative error in r for a design is shown in Fig. 2(a). Fig. 2(b) shows current–voltage characteristics of the most optimal design for each r. The currents have been scaled such that the objective functions are equal and the curves overlap. It can be seen that device designs spanning four orders of magnitude in r display essentially linear behavior over the optimization range.

Linearity of the optimized designs is determined by calculating total harmonic distortion (THD) and spectral intensity at frequencies $2\omega_0$ and $3\omega_0$, where ω_0 is the frequency of the input signal. The input signal applied to the diode is a sinusoidal voltage of frequency ω_0 having constant offset voltage $V_{bias} = 0.25$ V and peak amplitude V_p ($V_{pp} = 2V_p$). Fig. 3 shows the calculated distortion for the diode in Fig. 2(a). Suppression of spectral intensity at frequency $2\omega_0$ for $V_p \leq 0.1$ V results from odd symmetry in the differential resistance about $V_{bias} = 0.25$ V.

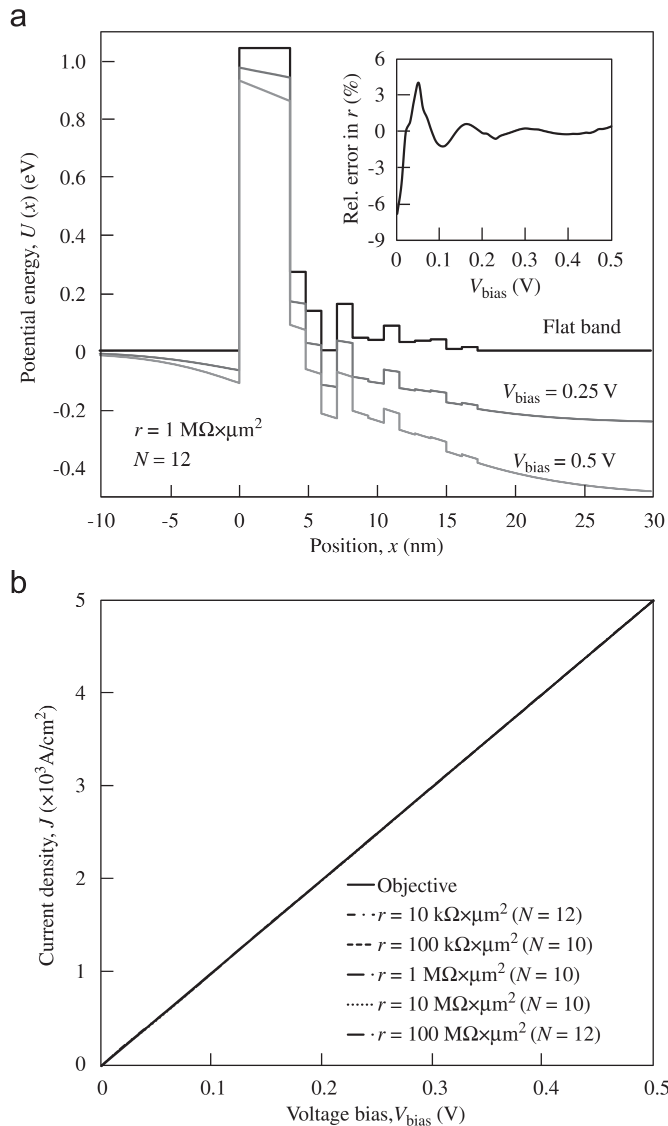


Fig. 2. (a) Optimized potential profile as a function of position x for a diode characterized by $r = 1$ M$\Omega \times \mu$m^2 and $N=12$. The inset shows error in differential resistance relative to r. (b) Overlay of simulated current through the most optimal design for each r, with corresponding N indicated. Currents for $r > 10$ k$\Omega \times \mu$m^2 have been scaled so that all curves overlap. Simulation parameters are electron effective mass $m = 0.07 \times m_0$, temperature $T=300$ K, and impurity concentration $n = 1 \times 10^{18}$ cm^{-3}.

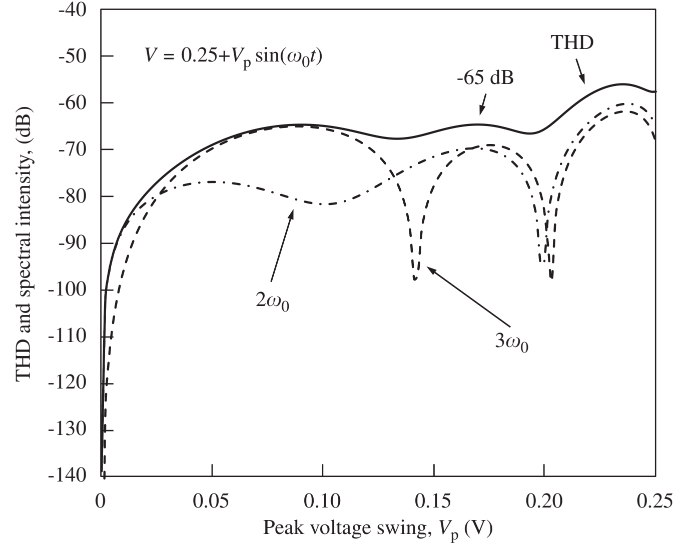


Fig. 3. Spectral intensities at $2\omega_0$ and $3\omega_0$ and total harmonic distortion (THD) for optimized diode shown in Fig. 2(a). The input voltage is a sinusoidal wave of frequency ω_0 and amplitude V_p with constant offset voltage $V_{bias} = 0.25$ V. Total harmonic distortion of less than -65 dB indicates greater than 10 bit dynamic range.

Table 1
Deviation from objective, Δ, of the most optimal design for each N.

N	5	7	10	12	15	17
Δ	−4.23	−5.09	−5.62	−5.58	−4.97	−4.72

Excellent linearity is achieved, with THD of less than -65 dB for $V_p \leq 0.2$ V indicating greater than 10 bit dynamic range for signals up to $0.4V_{pp}$. Decreasing the relative error for convergence to $\gamma = 10^{-5}$ produced similar results indicating the existence of limits to achievable optimality for the system.

To compare optimality independent of the magnitude of current, it is convenient to define the deviation from the objective as

$$\Delta = \log_{10}(1 - R_d^2), \tag{5}$$

where R_d^2 is the coefficient of determination between the objective and simulated current density, constrained to $0 \leq R_d^2 \leq 1$ with $R_d^2 = 1$ corresponding to $J_{sim} = J_{obj}$. The logarithmic nature of this metric coupled with $R_d^2 \approx 1$ for our designs makes this a very sensitive measure of optimality. Presented in Table 1 is Δ for the most optimal design for each value of N considered. As indicated in Fig. 2(b), $N=10$ and $N=12$ were the most successful values of auxiliary parameter N. The deviation from objective of the design of Fig. 2(a) is $\Delta = -5.58$.

5. Discussion

Control over device behavior is accessible through manipulation of quantum mechanical effects such as tunneling and resonant electron transmission. The extent to which these features can be accessed by thermally distributed electrons in the electrodes and controlled by the design parameters for the given range of V_{bias} and N determine feasibility of a given design.

For example, optimal $Al_{\xi}Ga_{1-\xi}As$ region lengths of $NL_0 \approx 7-12$ nm restrict access to low-energy ($< \mu$) resonances at small V_{bias}. Low-energy resonances are necessary for control in the low bias regime. Absence of these resonances is apparent in Fig. 2(a) by the negative error in r for $V_{bias} < 15$ mV and increasing error with decreasing V_{bias}. This feature is common to all optimized designs. Access to these resonant states could be made available by including $In_{\xi}Ga_{1-\xi}As$ alloys and in this way allowing $U_j \leq 0$ eV.

Naively, one might expect that thicker devices, with more degrees of freedom and access to additional resonances in the electron transmission spectrum, would have more control and hence higher optimality. However, for large values of N this is not the case due to a reduced ability of V_{bias} to control resonant spectra with a necessarily smaller electric field. Differential resistance curves for optimized designs with $N=15$ and $N=17$ show that the amplitude of oscillations about the objective resistance are enhanced.

The potential profile of Fig. 2(a) shows decreasing average U_j along the device length. This feature is common to many optimized designs with $7 \leq N \leq 12$ and is selected by the optimization algorithm to set the background electron transmission spectrum. Fitting the average conduction band offset within the $Al_{\xi}Ga_{1-\xi}As$ region of the structure considered in Fig. 2(a) to $U(x)=0.25\exp(-(x-L_0)/16)$ eV, with x measured in nm, yielded $\Delta=-1.2$ relative to a linear fit. Potential well features in the exponentially decaying potential that are added by the optimization algorithm form electron transmission resonances that further linearize and fine-tune the current.

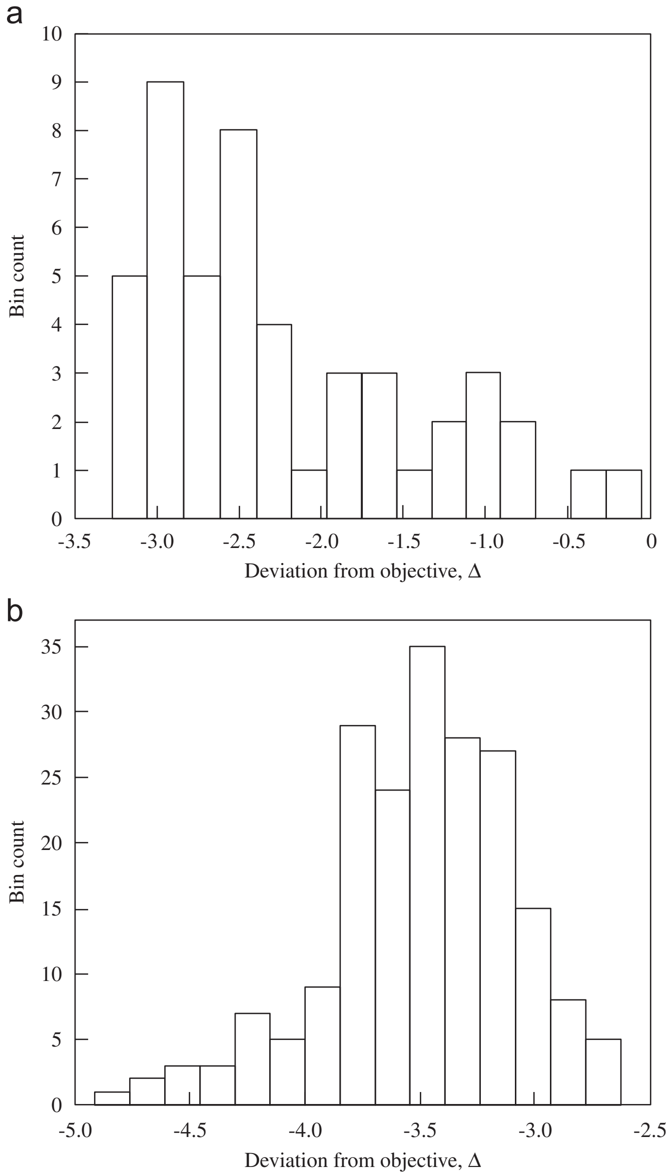


Fig. 4. (a) Histogram showing sensitivity of the optimal design of Fig. 2(a) to errors in $Al_{\xi}Ga_{1-\xi}As$ layer thickness, by calculating Δ with a single $Al_{\xi}Ga_{1-\xi}As$ layer perturbed by $\pm 1\times\delta$ or $\pm 2\times\delta$. (b) Histogram showing sensitivity of optimal design to alloy fluctuations and error in Al concentration. In each simulation all $Al_{\xi}Ga_{1-\xi}As$ layers were perturbed by a random potential energy uniformly distributed between ± 3 meV. Simulation parameters are electron effective mass $m=0.07\times m_0$, temperature $T=300$ K, and impurity concentration $n=1\times10^{18}$ cm^{-3}.

Sensitivity to errors in $Al_{\xi}Ga_{1-\xi}As$ layer thickness is shown in Fig. 4(a). The figure shows Δ calculated in the presence of $\pm 1\times\delta$ or $\pm 2\times\delta$ perturbations in the thickness of individual $Al_{\xi}Ga_{1-\xi}As$ layers for the design considered in Fig. 2(a). Skew in the distribution is because electron scattering strength of the first four layers from the AlAs layer is significant. These layers are responsible for all bin counts in which $\Delta > -2$. The third and fourth layers are most responsible for formation of transmission resonances because they are located near a quarter of the characteristic electron (Fermi) wavelength in the system. The two leftmost layers help the AlAs potential barrier set the scale of the current.

Sensitivity to alloy fluctuations and errors in Al concentration are simulated by perturbing U_j of all $Al_{\xi}Ga_{1-\xi}As$ layers by a random potential energy uniformly distributed within ± 3 meV. A histogram of 200 of these simulations for the design considered in Fig. 2(a) is shown in Fig. 4(b). Skew of this distribution toward smaller Δ may be a result of the strongest scattering occurring in 4 of the 12 $Al_{\xi}Ga_{1-\xi}As$ layers closest to the AlAs barrier.

Since the AlAs layer is the primary current limiter, the optimized design displays enhanced sensitivity to errors in L_0 thickness. Current through the design of Fig. 2(a) is scaled roughly by a factor of two for each monolayer change in L_0 between $\pm 2\times\delta$. Approximately linear behavior is preserved however, with a maximum deviation from objective of $\Delta=-2.65$ relative to a linear fit.

Changes in the Fermi–Dirac distribution with temperature will alter the current–voltage characteristic of the tunnel diode. Linearity of the current–voltage characteristic of the design in Fig. 2(a) as a function of temperature is evaluated by computing the deviation from objective with respect to a linear fit forced through the origin. Approximately linear current–voltage behavior over the temperature range $250\text{ K} \leq T \leq 350$ K is obtained with a maximum deviation from objective $\Delta=-2.95$ at temperature $T=350$ K. Variation of r over this temperature range is exponential with approximately $r(T)=(3.21\text{ M}\Omega\times\mu\text{m}^2\times\exp(-T/257.3)$.

Optimization with respect to temperature variation provides an interesting direction for future work. A more advanced model might incorporate additional scattering mechanisms, such as electron–electron scattering, in which scattering rates increase with temperature. It seems likely this would tend to increase the resistance with temperature and hence counter the trend of decreasing resistance for the model without temperature dependent inelastic scattering effects.

We note that the high levels of linearity achieved in the semiconductor heterostructure tunnel diode current–voltage characteristics we have considered is possible because use of $Al_{\xi}Ga_{1-\xi}As$ layers provide the needed degrees of freedom in design of the conduction band potential profile. An alternative approach using a single AlAs tunnel barrier similar to that considered in Fig. 1, but with very high n-type carrier concentration in the GaAs electrodes, fails to attain the same level of performance. For example, with $3\times10^{19}\text{ cm}^{-3} \leq n \leq 1\times10^{20}\text{ cm}^{-3}$ and $L_0=12\times\delta$ one obtains $\Delta \geq -3.3$ and THD > -36 dB for $V_p \leq 0.2$ V indicating less than 6 bit dynamic range for signals up to $0.4V_{pp}$.

The physical model we have used to describe electron transport is highly simplified involving a single Γ-valley band structure in GaAs. Use of a more sophisticated band structure that includes conduction band non-parabolicity and the possibility of electron transport via subsidiary L- and X-minima in GaAs will change the calculated current–voltage characteristic. However, these changes should be controllable and hence compensated for by varying the $Al_{\xi}Ga_{1-\xi}As$ layers used to fine-tune the optimal design.

6. Conclusion

The design of electronic heterostructure devices can utilize atomic layer precision in semiconductor fabrication to mimic an Ohmic (linear) current–voltage characteristic over a limited, but useful, range of voltage bias values. Even though electron transport is dominated by tunneling, it is possible to manipulate low energy electron scattering and resonances to approach a desired linear current–voltage behavior. However, practical and materials constraints limit the degree of linearity that may be achieved. Here, we demonstrate how greater than 62 dB (10 bit) dynamic range may be obtained in design of an optimized semiconductor heterostructure tunnel device that is only 17 nm thick. This type of control, in which there is no inelastic electron scattering in the active region $(\lambda_{in} \gg L)$, is only possible because of the wave nature of electrons.

References

[1] J.M. Ziman, Principles of the Theory of Solids, second ed., Cambridge University Press, New York, 1972.
[2] R. Landauer, Philosophical Magazine 21 (1970) 863.
[3] W. Lu, P. Xie, C.M. Lieber, IEEE Transactions on Electron Devices 55 (2008) 2859.
[4] A. Nieuwoudt, Y. Massoud, IEEE Transactions on Electron Devices 55 (2008) 2097.
[5] J.J. Plombon, K.P. O'Brien, F. Gstrein, V.M. Dubin, Applied Physics Letters 90 (2007) 063106.
[6] J. Zhang, R.L. Kosut, Robust design of quantum potential profile for electron transmission in semiconductor nanodevices, in: European Control Conference 2007, July 2–5, 2007, Kos, Greece.
[7] P. Schmidt, S. Haas, A.F.J. Levi, Applied Physics Letters 88 (2006) 013502.
[8] K.C. Magruder, P. Seliger, A.F.J. Levi, in: S. Haas, A.F.J. Levi (Eds.), Optimal Device Design, Cambridge University Press, New York, 2010.
[9] I. Vurgaftman, J.R. Meyer, L.R. Ram-Mohan, Journal of Applied Physics 89 (2001) 5815.
[10] A.F.J. Levi, Applied Quantum Mechanics, second ed., Cambridge University Press, New York, 2006.
[11] A. Trellakis, A.T. Galick, A. Pacelli, U. Ravaioli, Journal of Applied Physics 81 (1997) 7880.
[12] R. Tsu, L. Esaki, Applied Physics Letters 22 (1973) 562.
[13] T. Fiig, A.P. Jauho, Applied Physics Letters 59 (1991) 2245.
[14] W. Pötz, Journal of Applied Physics 66 (1989) 2458.
[15] R.C. Bowen, G. Klimeck, R.K. Lake, W.R. Frensley, T. Moise, Journal of Applied Physics 81 (1997) 3207.
[16] P. Roblin, W. Liou, Physical Review B 47 (1993) 2146.
[17] S. Adachi, Journal of Applied Physics 58 (1985) R1.
[18] A.F.J. Levi, I.G. Rosen, SIAM Journal on Control and Optimization 48 (2010) 3191.