Emitter quantization and double hysteresis in resonant-tunneling structures: A nonlinear model of charge oscillation and current bistability

F. A. Buot
Naval Research Laboratory, Washington, D.C. 20375

P. Zhao and H. L. Cui
Stevens Institute of Technology, Hoboken, New Jersey 07030

D. L. Woolard
Army Research Office, Research Triangle Park, North Carolina 27709

K. L. Jensen and C. M. Krowne
Naval Research Laboratory, Washington, D.C. 20375

(Received 6 May 1999; revised manuscript received 14 October 1999)

The effects of emitter quantization on the current-voltage (I-V) characteristics of conventional double-barrier resonant tunneling structures (RTS's) are investigated by numerical, graphical, and analytical methods. Different stability and degrees of emitter quantization can lead to a host of different I–V characteristics in the negative differential resistance (NDR) region. Among these are simple NDR, NDR with a rising plateaulike region and well-separated double hysteresis, and NDR with a falling plateaulike region and well-separated double hysteresis. The ratio of the main hysteresis width to the secondary hysteresis width can vary between 1 and ∞. The use of large enough spacer layers can eliminate the hysteresis and plateaulike behavior. Our numerical results for RTS's are analyzed by employing graphical (based on simulated quantum-well charge) and analytical methods, and compared with experiments. We introduce a nonlinear physical model which is solved analytically for the limit cycle solution. The limit cycle predicts a rising average current, whereas the nonoscillatory solution predicts a falling current in the plateau region as a function of bias. The limit cycle also predicts a monotonically decreasing amplitude of the current oscillation as a function of bias in the plateau region. The fundamental frequency increases, reaches a maximum, and sharply decreases to zero as a function of bias in the plateau region. These analytical results agree with experiments and numerical simulations. The origin of inductive delay in RTS's is further clarified. We believe we have resolved in fine detail the controversy about the I-V characteristics of conventional RTS's. A prescription for this structure to operate as an all solid-state THz source is also given.

I. INTRODUCTION

It is well known that numerical quantum transport simulations and experiments of most resonant tunneling structures reveal the characteristic peak-to-plateau-to-valley behavior of the current as a function of applied voltage.[1–4] Double hysteresis behavior is also seen with forward and backward voltage sweep.[3,4] Moreover, time-dependent simulation results of different groups[5,6] reveal intrinsic high-frequency current oscillations in the plateau region of the current-voltage (I-V) characteristics. The purpose of this paper is to shed light on these phenomena. We have performed numerical quantum transport simulations of resonant tunneling structures (RTS's) by independently varying various device parameters. The results are explained as due to the presence of unstable two-dimensional (2D) quantization in the emitter through strong ''ripple effects'' on the electron-density distribution. First we use graphical methods to explain our results and other RTS results for the gross I-V characteristics. To explain the finer features, here we also introduce a nonlinear analytical model of charge oscillation and bistability, and obtain a limit cycle solution for the oscillations in the plateau region.

The only way to explain the current peak during forward bias is to postulate the passage of the quantum-well (QW) discrete energy level into the forbidden energy region of the emitter. This forbidden region does not necessarily correspond to the energy gap between the emitter conduction and valence bands. This forbidden region may be created above the emitter conduction band edge by virtue of the quantization of the supply electronic states by the emitter quantum well (EQW). This is illustrated in Fig. 1(a), which shows a triangular EQW. In what follows we will show that the QW energy-level passage into the forbidden region above the conduction band edge of the emitter, depicted in Fig. 1(a), is responsible for the high-frequency current oscillation and plateaulike behavior of the I-V characteristic of RTD's.

The passage of the QW energy level into the forbidden region of the emitter creates a sudden drop of the current across the device, producing a characteristic sharp current peak. However, the subsequent and inherent feedback mechanism prevents the average post-peak current from assuming the valley-current value. As the electrons are built up in the emitter, the interference of the reflected electrons and the incoming electrons creates a significant ''ripple effect'' on the electron-density distribution which effectively broad-

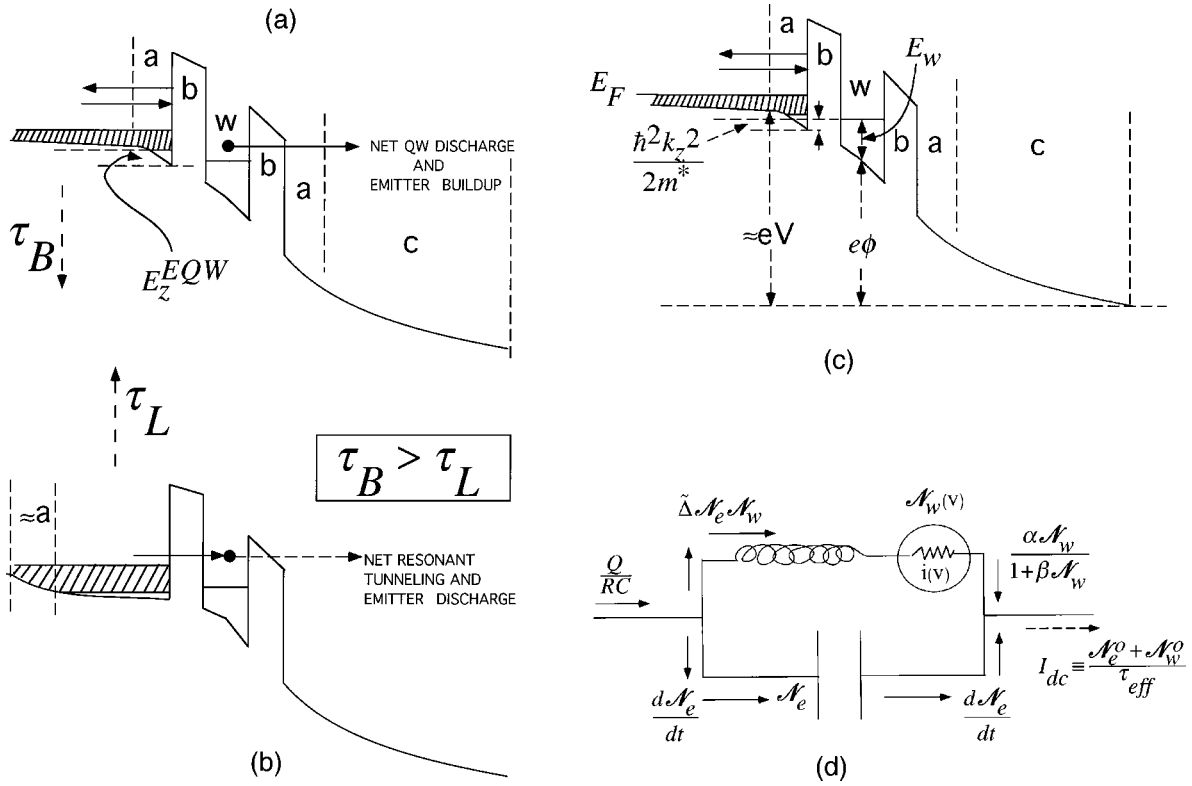

FIG. 1. (a) Emitter 2D quantization results in premature alignment of the QW energy level with the forbidden region of the emitter. The electrons accumulate in the emitter with density structure and a broadening of the EQW caused by interference of reflected and incoming electrons. The buildup time takes τ_B to achieve the condition in (b). (b) Realignment with occupied states causes rapid depletion of the builtup charge in time τ_L, after which condition (a) is restored. (c) Device parameters used in the calculations. (d) Equivalent circuit model, where $i(v)$ $[\mathcal{N}_w(v)]$ is the current (charge) of an ideal RTS for a voltage drop v across the negative resistor.

ens the EQW by virtue of the self-consistency of the charge and potential. The broadening is due to the redistribution of the electrons, with some regions becoming positive (deficit of electrons) and some regions negative (excess of electrons) in the emitter. This is clearly indicated by our numerical results.

This EQW renormalization yields a much broader EQW, and the lowering of the allowed discrete energy level in the emitter toward the conduction-band edge. With this EQW broadening, it is also plausible for a change from discrete quantization to the continuum energy levels down to the conduction-band edge of the emitter. In either case, the alignment of the QW discrete energy level with occupied states in the emitter will be restored, yielding high transmission coefficients and larger currents. This is depicted in Fig. 1(b). This feedback is basically a catalytic process, since the quantum-well charge, through the self-consistent potential, helps in restoring the QW energy-level alignment with the occupied states in the emitter. The resonant tunneling from the emitter (emitter discharge) and the self-consistency of the charge and potential lead to a restoration of the emitter potential profile which produces 2D quantization, and passage of the QW energy level through the occupied states into the forbidden region, i.e., back to the situation depicted by Fig. 1(a). The process therefore oscillates between that of Figs. 1(a) and 1(b), and the average is responsible for the plateau-like behavior above the valley-current minimum. From the steady-state point of view, it is as if the QW energy level maintains its alignment with some fraction of the occupied states in the emitter.[7]

It is thus clear that the driving source of this oscillatory condition is the buildup and redistribution of charge (significant ripple effect) due to the interference of the reflected and incoming electrons. In our analytical treatment of the problem, we identify this driving source in terms of the total charge buildup, Q, at the emitter in time duration RC, where R is the access resistance and C is the RTS capacitance. Therefore, Q/RC measures the maximum rate of buildup of supply electrons at the emitter in the absence of tunneling to the QW. In the presence of tunneling, the proper coupled rate equations are derived in Sec. III, where we introduce our nonlinear physical model.

We expect that the amplitude of oscillation is largest in the plateau region just after the current peak, as depicted in Fig. 1. This is because there is a considerable broadening or renormalization of the EQW in going from Fig. 1(a) to Fig. 1(b), i.e., in bringing the allowed EQW-allowed states in line with the QW energy level. On the other hand, well within the plateau region, there is only a further broadening of the EQW, and hence the amplitude of the oscillation will become smaller as the drain bias is further increased. This is indicated in Fig. 2. Indeed, this is what was found in various time-dependent numerical simulations of RTS's.[5,6] This conjecture is also confirmed by our analysis in Sec. VI.

We should point out that although the interference of reflected electrons with incoming electrons also occurs to some

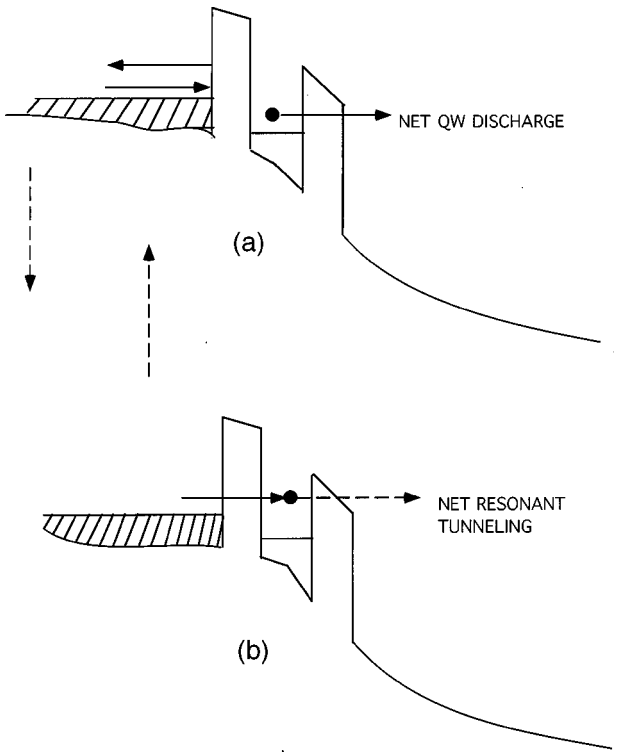

FIG. 2. A smaller amplitude of the oscillation occurs within the plateau region as the QW energy level continues to shift downward with an increase in applied bias. Hence the EQW broadening charge Q in Q/RC also decreases with an increase in applied voltage in the plateau region.

degree in the valley and prepeak region of the I-V plot, the accompanying interference-induced broad potential is only transient in nature as the bias is suddenly changed (transient ripple effect), and is not present in the simulation when the steady state is reached (the stationary transmission coefficient). The strong ripple effect on the electron-density distribution in the emitter is only sustained if the transmission coefficient is highly oscillating, such as what is occurring in the current-plateau region. This phenomenon is analogous to strong ripple on the surface of a water pool when the discharge valve is turned on and off periodically. For stationary discharge outside the plateau region, quantum mechanics still creates a ripple due to interference of reflected and incoming electrons, but not one significant enough to broaden the EQW and realign the QW energy level with occupied states in the emitter for applied bias beyond the plateau range.

Indeed, a deeper understanding of the oscillation and bistability in RTS's can be attributed to the unstable 2D quantization in the triangularlike EQW just after the current peak, as suggested by our recent steady-state simulation on resonant tunneling structures of different barrier widths, simulation box (simulated device) lengths, barrier heights, right-barrier widths, QW widths, spacer-layer widths, and doping densities. As indicated in Fig. 1(a), it is the passage of the QW energy level into the forbidden region of the emitter which marks the onset of the negative differential resistance (NDR) region of the I-V characteristics. Since the occupied states are above the conduction-band edge of the emitter, this premature onset is expected to occur at lower drain voltage than that when the QW energy level is aligned with the bottom of the conduction band of the emitter; this is where the current peak would have occurred if there was no forbidden 2D states formed in the emitter. We should also point out that the more accurate use of a longer simulation box length (longer simulated device) strengthens our physical arguments for two reasons: (a) it assures the natural formation of broadened EQW and low-lying allowed states, or even the formation of continuum states in the emitter, for the situation depicted in Fig. 1(b), while (b) providing for a more accurate approximation of the ''inflow'' boundary condition, typically used in the numerical simulation.[5,6] It also yields large enough series resistance for the oscillation to set in. Thus, in agreement with previous analysis,[1,2] large access resistance in longer device indeed helps create the oscillatory behavior.

The above proposed physical mechanism is particularly supported by the five salient features of the steady-state computer simulation of symmetrical RTS's of different barrier widths, holding other dimensions fixed (all variations reported here are in the sense of partial derivatives). These are as follows.

(1) The width of the plateau increases with the decrease of barrier widths (also with decrease of device lengths, barrier heights, right-barrier widths, QW widths, and spacer-layer widths).

(2) The ratio of the width of the primary hysteresis to the width of the plateau increases with the barrier width (also with increase in device lengths, barrier heights, right-barrier widths, QW widths, and spacer-layer widths).

(3) The average slope of the plateau decreases with increase in barrier widths (also with increase in device lengths, barrier heights, right-barrier widths, QW widths, and spacer-layer widths).

(4) There is a tendency to form secondary hysteresis between the plateau and peak current.

(5) The rising plateaulike region, which indicates the presence of time-dependent oscillatory current behavior and the hysteresis, disappears for large spacer-layer widths.

The fourth salient feature in particular has never been explained in the literature; however, this feature is usually seen in experiments[3,4] as well as in the time-dependent simulations of Jensen and Buot[5] and Biegel and Plummer.[6] We should also point out another very important information, namely, that the average stored charge in the quantum well calculated by Jensen and Buot[5] has the same shape as the calculated I-V characteristic. This observation is the key to explaining the fourth salient feature above, and is crucial to our graphical analysis presented in Sec. II. This information was completely missed in graphical analyses of RTS's by other authors. Sheard and Toombs[8] were the first to employ graphical methods to show the presence of hysteresis and bistability in RTS's. Although their result seems feasible, this does not resemble any of the experimentally measured I-V characteristics of Refs. 3 and 4, by virtue of failing to take into account the quantization in the emitter. In particular, the presence of double hysteresis in the experimental results[3,4] was not explained at all. Using similar model as that of Sheard and Toombs, the authors of Ref. 9 also employed graphical analysis in examining the effect of the discretization of charge in small area RTS's.

A more recent graphical work of Sheard and Toombs[10]

took into account the quantization of the accumulation layer in the emitter. However, their graphical analysis was entirely time independent, and failed to account the oscillatory behavior which intermittently renormalizes the emitter conduction-band edge after the current peak, causing an instability of the emitter quantization, a phenomenon which is clearly indicated in our numerical results. Moreover, they assumed that the QW energy level lies above the EQW energy level due to screening of the applied voltage by the charge in the QW. Thus, they claimed that the electric field across the emitter barrier and hence the charge in the accumulation layer remain almost constant with bias, which is contrary to our numerical results. Again no attempt was made to explain the double hysteresis and other salient features of the experiment of Refs. 3 and 4.

II. EFFECTS OF UNSTABLE EMITTER QUANTIZATION IN THE NDR REGION

The salient features of our simulation mentioned above can be explained if we invoke the presence of unstable discretization of the EQW energy levels in the NDR operation. The quantization of EQW energy levels is expected to drastically change the supply electrons from the emitter as ''seen'' by the QW energy level. The peak of the current will no longer correspond to the alignment of the QW energy level with the bottom of the conduction band of the emitter. The peak current will correspond to the alignment between the EQW-allowed subband minimum and QW energy level, at a lower current-peak bias. The supply electrons are thus reduced because the allowed energy level is raised from the bottom of the conduction band of the emitter. This does not imply that the observed current is necessarily decreased; on the contrary, the observed current usually increases since this condition occurs for smaller barrier widths and higher fields in the emitter. Moreover, because of the access-resistance voltage drop, larger currents will shift the I-V curve to higher-bias range, as shown by our simulation for smaller barrier width RTS's.

As the QW energy level passes a sufficiently high-enough EQW energy level during a forward-bias sweep, oscillatory charging and discharging of the QW result by virtue of the self-consistency and feedback between the charge and potential, with intermittent broadening of the confining potential in the emitter. In effect, as will be analytically demonstrated below the unstable alignment of the EQW and QW energy levels converges into a limit cycle, and hence in the oscillation in the QW charges and corresponding currents.[1] These oscillations persist until the QW energy level is below the conduction-band edge of the emitter; the current drastically drops to the valley-current value, since any broadening of the EQW is no longer effective in realigning the pertinent energy levels. The average values essentially map the steady-state current plateau, given by the steady-state simulation.

All the salient features listed above can readily be explained by using a graphical analysis in the manner used by Buot and Rajagopal[11–13] and Buot.[14] The graphical analysis involves the simultaneous solution of two different expressions for the QW charge, one derived from the Poisson equation and another from the transport-rate equation. Here we have to make provision, based on the numerical result of Jensen and Buot,[5] for the changes in the average supply electrons from the emitter due to the formation of an oscillating forbidden region above the emitter conduction-band edge. This was not accounted for in graphical analyses by other authors. As usual, the current can then be approximated by Q_w/τ_c,[11–13] where Q_w is QW charge and τ_c is the decay time to the collector.

From the Poisson equation, we have, for Q_w, an expression,

$$Q_w=\frac{(2+\zeta)}{e'}\Xi\left\{\frac{\hbar^2k_z^2}{2m^*}-E_w+\frac{eV'}{(2+\zeta)}\right\}, \tag{1}$$

where $\zeta=c/(a+b+w/2)$, $\Xi=\epsilon/(a+b+w/2+c)$, $e'=e\{1-\Xi/[\epsilon/a(1+\zeta)]\}$ and $V'=V[1-\zeta a/c]$. Here ϵ is the dielectric constant. In deriving Eq. (1), we simply use the Poisson equation[9] $E_R-E_L=Q_w/\epsilon$, where E_R is the electric field in the right region $a+b+w/2+c$ and E_L is the electric field in the left region $a+b+w/2$, indicated in Fig. 1(c). Note that the ''capacitance'' parameter Ξ and the ratio ζ vary inversely with the barrier width b, i.e., the slope of Q_w vs k_z^2 increases with the decrease in b. We shall soon see that this causes the ratio of the hysteresis width to the plateau width to increase with increase in barrier width.

From the transport-rate equation,[11–13] we have another expression for the QW charge given by

$$\begin{aligned}Q_w=&\frac{e\ m^*}{\pi\hbar^2\beta}\ln\left\{1+\exp\beta\left(E_F-\frac{\hbar^2k_z^2}{2m^*}\right)\right\}\\&\times\left(\frac{\tau_d}{\tau_e}\right)\theta\left(\frac{\hbar^2k_z^2}{2m^*}-E_z^{\mathrm{EQW}}\right)\\&+\frac{F\ e\ m^*}{\pi\hbar^2\beta}\ln\left\{1+\exp\beta\left(E_F-\frac{\hbar^2k_z^2}{2m^*}\right)\right\}\\&\times\left(\frac{\tau_d}{\tau_e}\right)\theta(k_z^2)\,\theta\left(E_z^{\mathrm{EQW}}-\frac{\hbar^2k_z^2}{2m^*}\right),\end{aligned} \tag{2}$$

where E_z^{EQW} is the energy level formed in the emitter, $1/\tau_d=1/\tau_e'+1/\tau_c$, $1/\tau_c$ is the effective rate of decay of Q_w into unoccupied collector states, and $1/\tau_e$ is the effective rate of decay of Q_w into unoccupied emitter states which may be assumed to be equal to the rate of supply of electrons from the emitter to the quantum well, $1/\tau_e'$. In a highly nonequilibrium situation we can put $\tau_d\approx\tau_c$. In Eq. (2) we set the average value of the oscillatory quantum well charge at the plateau ($0<\hbar^2k_z^2/2m^*<E_z^{\mathrm{EQW}}$) as equal to a fraction F of Q_w obtained if there were no forbidden 2D states in the emitter. The first term of Eq. (2) comes from the electrons in the allowed states above E_z^{EQW}, whereas the second term, which is further clarified in Sec. VI, describes either the effective averaged of the oscillating quantum well charge or the stationary value in the plateau after the QW energy level passes the E_z^{EQW}. Note that we do not exclude the possibility of nonoscillatory behavior in the plateau, as indicated by Fig. 3. However, our analytical results in Secs. V and VI strongly suggest that this can occur only for very weak quantizing field in the emitter, and furthermore this is characterized by a

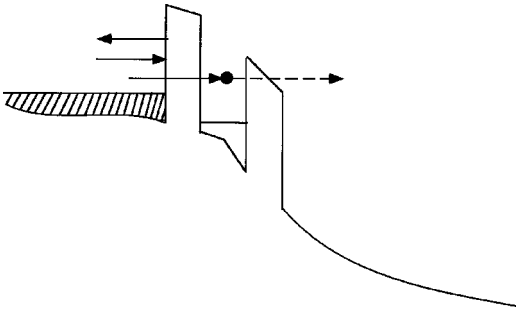

FIG. 3. Stable stationary condition can occur if the buildup rate in the emitter is balanced by the resonant tunneling rate. This condition is estimated to be readily attained in the plateau region for a very weak quantizing field in the emitter that causes the premature current peak.

decreasing plateau current as a function of drain bias, suggesting that F is a decreasing function of bias for nonoscillatory condition.

For the oscillatory case, F is expected to be an increasing function of barrier width b, since the amplitude of quantum-well charge oscillation should be small when the difference of bias is small between the alignment of the QW energy level with the bottom of the conduction band and with the EQW energy level. This holds for the larger barrier width b, and hence a weaker quantizing field in the emitter for the situation depicted in Fig. 1(a). More will be said relating to F in the discussion of the analytical solution in Sec. VI. E_z^{EQW} is expected to lie higher above the bottom of the emitter conduction-band edge for smaller barrier width than for larger barrier width, holding other RTS dimensions fixed, by virtue of increased localized band bending at the emitter for smaller barrier width. This translates to a larger plateau width for smaller barrier width than for larger barrier width. Other factors also enter, such as the spacer layer widths whose effect will be discussed later.

Based on the simulated quantum-well charge[5] and in the light of the typical experimental and simulated I-V curves of RTS's *in the absence of a current plateau*, Fig. 4 shows the empirical plot of Eq. (2), together with the approximation of Eq. (1) as straight dotted lines for different barrier widths. For simplicity we use a constant value of F for each plateau.

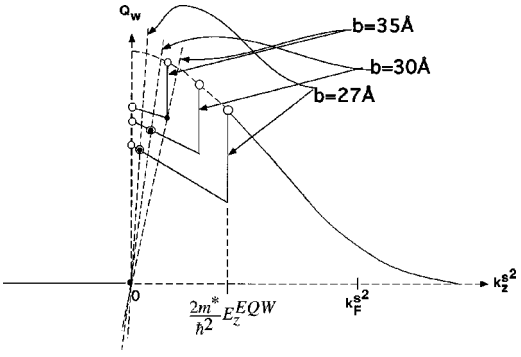

FIG. 4. Schematic diagram comparing the plots of Eq. (1) (straight lines) and Eq. (2) for RTS's with different barrier widths. Solid circles stand for reverse-bias sweep (progresses to the right) solutions, and open circles for forward-bias sweep (progresses to the left) solutions.

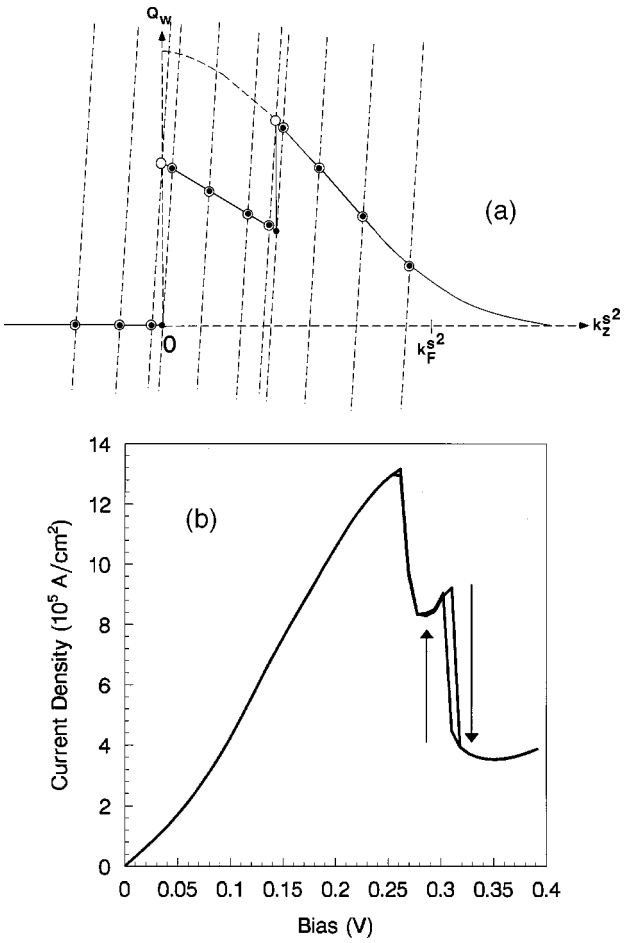

FIG. 5. (a) Graphical solutions of Eqs. (1) and (2) for a barrier width of 27 Å. Solid circles stand for reverse-bias sweep solutions, and open circle for forward-bias sweep solutions. (b) Corresponding simulation results for the current density.

However, note that for the 35-Å barrier width one should use a decreasing F with bias in light of our numerical and analytical results. In all these figures the *solid circle* represents the simultaneous solution of Eqs. (1) and (2) for a *reverse-bias sweep*, and the *open circle* corresponds to the solution for a *forward-bias sweep*. According to Eq. (1), an increase in drain bias will result in the displacement of the dotted lines to the left, and a decrease in drain bias will result in displacement to the right.

The current is approximately given by Q_w/τ_c.[11–13] In Fig. 4, we see that the slope and width of the plateau, as a function of bias (or k_z^2), is also a decreasing function of the barrier width, in agreement with our numerical simulation. Moreover, because of the dependence of Ξ and ζ of Eq. (1) on the barrier width b, the slope of the dotted straight lines is inversely proportional to the barrier width. This results in a decreasing fraction of the hysteresis width to the plateau width, as well as an increased separation between the two hysteresis, as the barrier width is decreased. The hysteresis is defined by four solution points consisting of two diagonally opposite solid and open circles and two diagonally opposite points, each consisting of a solid circle inside an open circle representing overlapping solutions to the both reverse and forward-bias sweeps. Figures 5(a)–7(a) show the solution points for reverse and forward-bias sweeps for different bar-

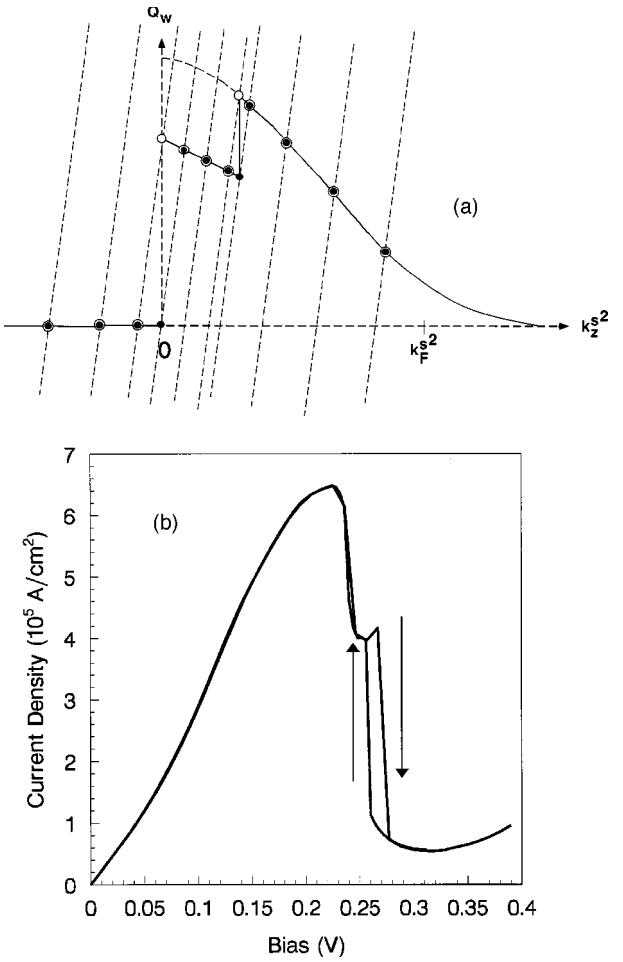

FIG. 6. (a) Graphical solutions of Eqs. (1) and (2) for a barrier width of 30 Å. Solid circles stand for reverse-bias sweep solutions, and open circle for forward-bias sweep solutions. (b) Corresponding simulation results for the current density.

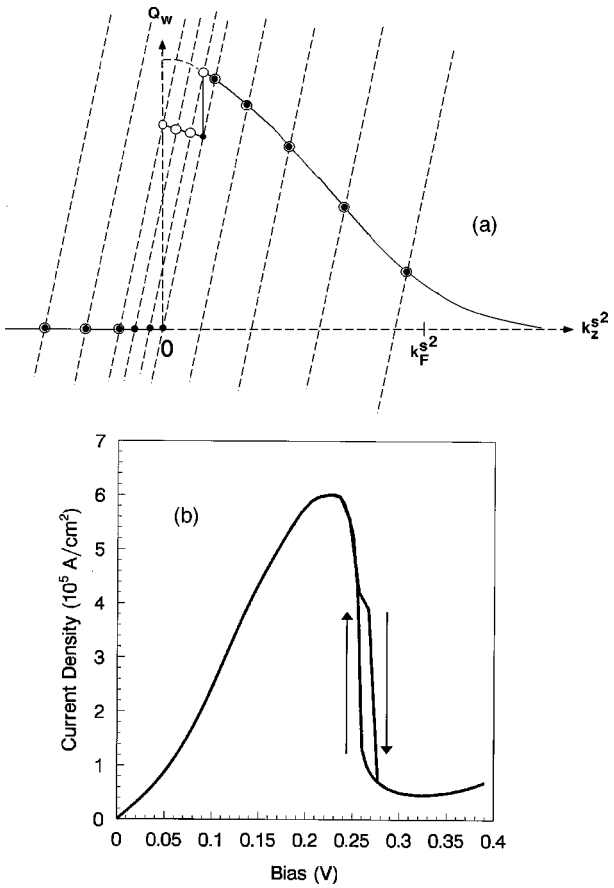

FIG. 7. (a) Graphical solutions of Eqs. (1) and (2) for a barrier width of 35 Å. Solid circles stand for reverse-bias sweep solutions, and open circle for forward-bias sweep solutions. (b) Corresponding simulation results for the current density.

rier widths, from which one can observe the salient features listed above. In particular the tendency to form double hysteresis, not very well accounted for in the literature, is quite apparent from these figures. Figures 5, 6, and 7 demonstrate that the graphical solution has the same prediction as the extensive numerical simulation concerning the dependence of the I-V characteristics as a function of the barrier width.

In the graphical solution of Fig. 5(a) for a barrier width of 27 Å, note that the primary lower hysteresis width and secondary upper hysteresis width are small and approximately equal, and these two hysteresis are very well separated in the plateaulike region. These salient features agree with our actual numerical quantum transport simulation shown in Fig. 5(b). In both our graphical solution [Fig. 6(a)] and in the actual numerical simulation [Fig. 6(b)] for a barrier width of 30 Å, the primary lower hysteresis width is clearly larger than the secondary upper hysteresis width, and their separation in the plateau region is less than those in Figs. 5(a) and 5(b). In our graphical solution of Fig. 7(a) for a barrier width of 35 Å, the upper secondary hysteresis width is finally masked by the large width of the primary lower hysteresis which occupies the whole plateau region, i.e., the two hysteresis widths overlapped. There could only be one visible hysteresis occupying the whole plateau region, although the simulated I-V characteristics for large doping levels, discussed below, closely follow the graphical solution of Fig. 7(a). Hence our graphical results agree with that of the actual numerical quantum transport simulation shown in Fig. 7(b).

Figures 8–11 demonstrate the significant renormalization of the emitter conduction, band edge at NDR, often ignored in previous analyses of RTS's. Figure 8, for the forward-bias sweep, and Fig. 9, for the backward-bias sweep, demonstrate the degree of broadening of the emitter region as a function of the barrier widths. These figures indicate the increase of EQW-confining electric fields as the barrier width decreases. At the largest simulated barrier width of 35 Å, it is estimated that the oscillatory behavior in the plateau is not maintained due to a very weak quantizing field, and evolves into a stationary behavior. Figures 10 and 11 show the corresponding electron-density distribution for the forward- and backward-bias sweeps, respectively. These figures again show large changes in the structure of the emitter region for smaller barrier widths than for larger barrier widths, indicating that the EQW-allowed states lie higher from the conduction-band edge of the emitter for smaller barrier widths.

Thus the width of the plateaulike region in the I-V characteristics is proportional to the value of the 2D energy level of the EQW measured from the bottom of the emitter conduction-band edge, denoted as E_z^{EQW} in Fig. 1(a). This value is in turn proportional to the band-edge bending in the emitter region at the barrier edge. We are thus led to similar

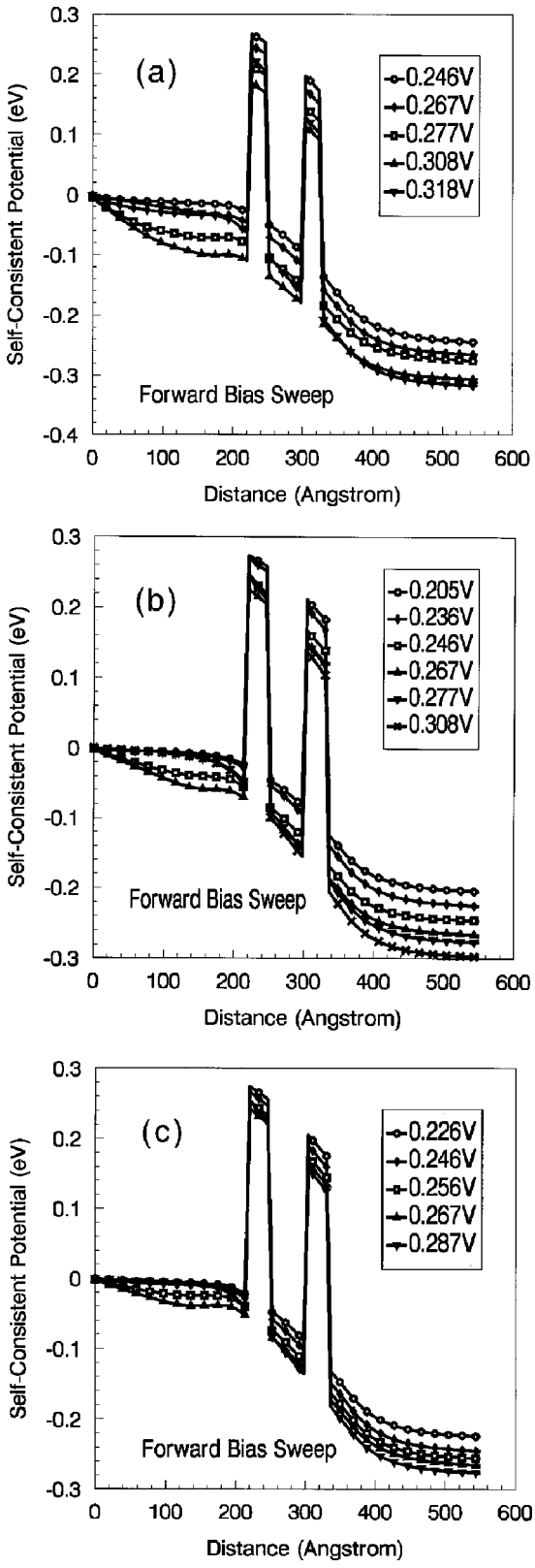

FIG. 8. Simulated forward-bias self-consistent potential for barrier widths of (a) 27 Å, (b) 30 Å, and (c) 35 Å.

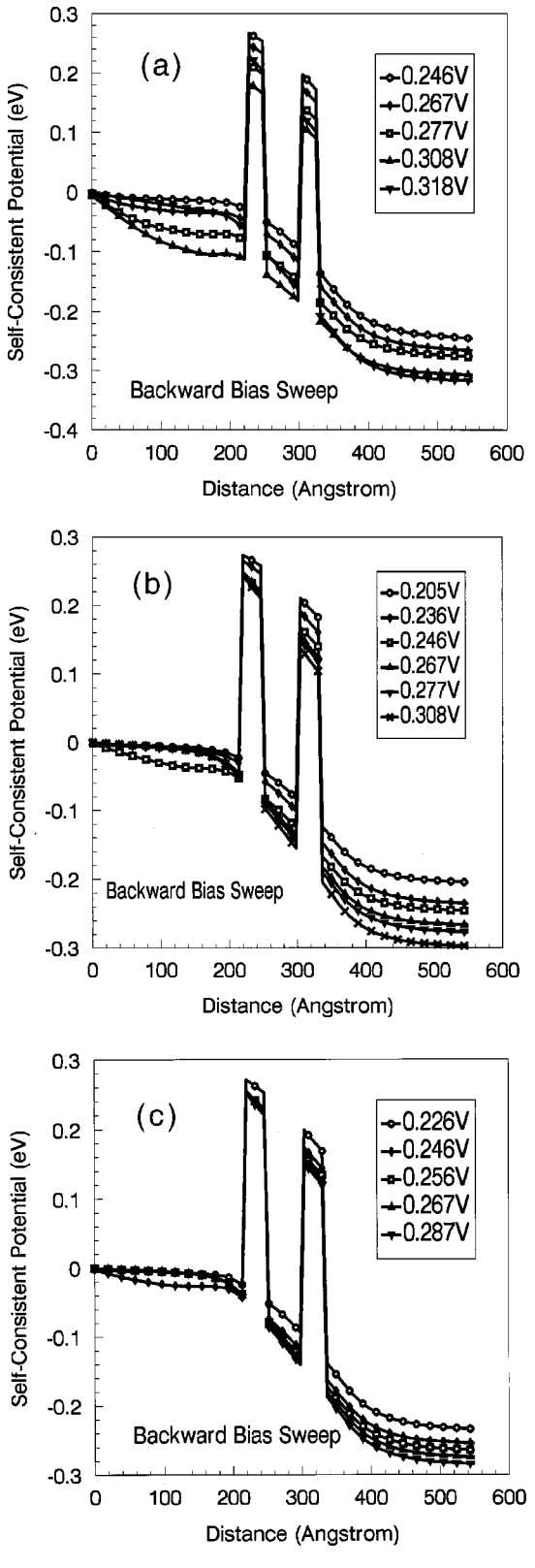

FIG. 9. Simulated backward-bias self-consistent potential for barrier widths of (a) 27 Å, (b) 30 Å, and (c) 35 Å.

explanations for the variation of the features of the I-V characteristic in the NDR with respect to device lengths (Fig. 12), heights of the barrier (Fig. 13), right-barrier widths (Fig. 14), quantum-well widths (Fig. 15), spacer-layer widths (Fig. 16), temperatures (Fig. 17), and doping density (Fig. 18).

In our simulation of the actual I-V characteristic as functions of the device lengths, we found that the use of a smaller device length produces a result which mimics the use of a smaller barrier width [Fig. 5(a)], in the sense of having very small hysteresis widths (shown in Fig. 12 to be zero for both the primary and secondary hysteresis) and very well sepa-

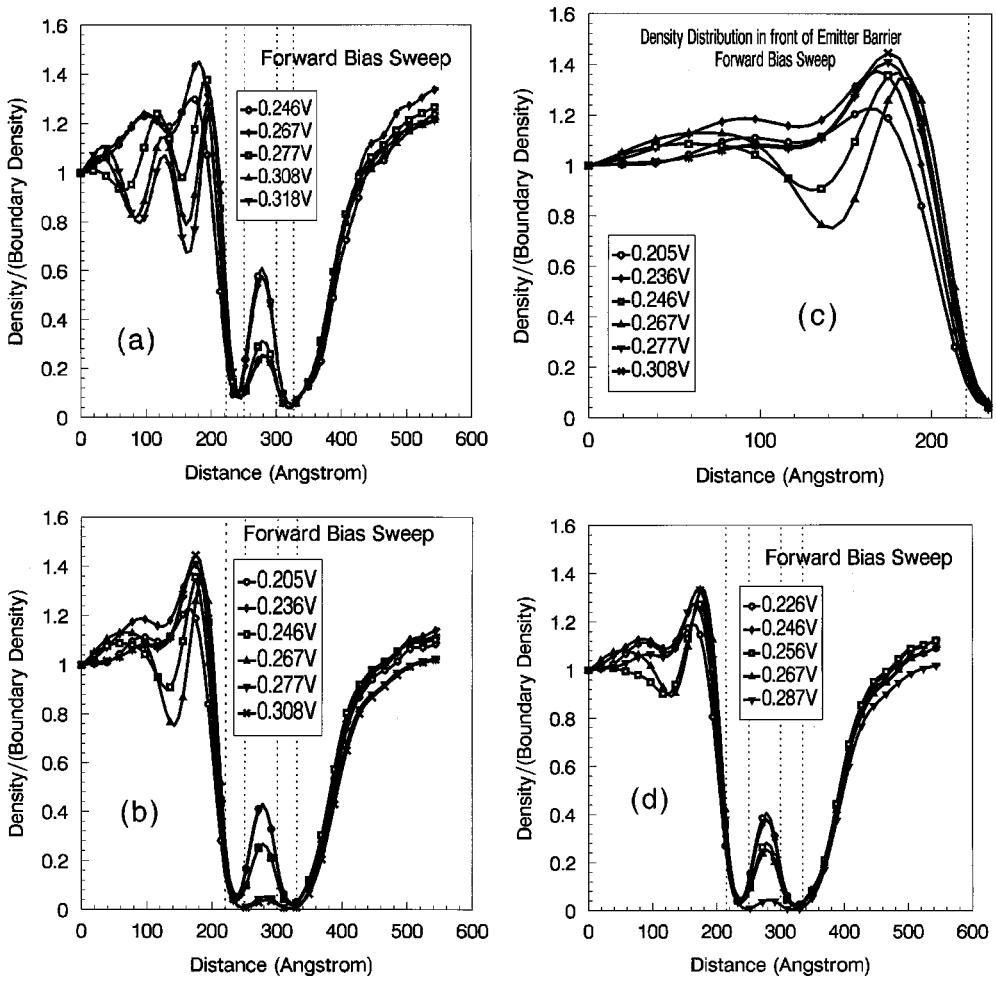

FIG. 10. Simulated forward-bias electron-density distribution for a barrier width of (a) 27 Å and (b) 30 Å. (c) Details of the emitter structure for 30 Å and (d) 35 Å.

rated in the plateau region. This is because the condition of a higher quantizing field at the emitter and a larger value of the ''capacitance'' parameter Ξ of Eq. (1) are achieved for smaller lengths than for larger device lengths. Indeed, a similar trend is obtained in going from smaller to larger simulation box lengths (Fig. 12), as in the use of increasing barrier widths (Figs. 5–7). Note, however, that in Fig. 12 all the plateau currents are rising as a function of the applied bias, indicating the presence of significant oscillations in all three cases in Fig. 12, as opposed to Fig. 7(b), which exhibits a falling current in the plateau and where oscillation does not seem to be present in the light of the discussions given in Secs. V and VI. The reason for this maintained oscillation is that longer simulation box length enhances oscillatory behavior in a manner discussed in the Sec. I, thereby compensating for the weakened confining field at the emitter and maintaining the driving source.

Approximately similar trend as in Figs. 5–7 is also obtained in our numerical simulation as one goes from smaller barrier heights to larger barrier heights shown in Fig. 13. The explanation again hangs on the two important parameters that were used in our graphical analysis, namely, the quantizing field in the emitter and the ''capacitance'' parameter. The variation with respect to the barrier heights is more subtle than the dimensional effects on the confining potential at the emitter and the capacitance parameter Ξ. It is important and helpful to realize that a decrease in barrier heights will actually lead to greater band bending at the emitter in the NDR operation. The reason for this is that smaller barrier heights yield larger values of the QW energy level measured from the bottom of the conduction-band edge of the QW, denoted as E_w in Fig. 1(c), due to a lesser confining QW potential. This means that the NDR for these devices occurs at larger values of the drain bias, which results in stronger confining fields at the emitter in the NDR. Hence we see that the width of the plateau also increases with a decrease in the barrier heights. Moreover, smaller barrier heights will result in larger currents and a larger charging of the quantum well. This translates into a larger capacitance with a decrease in barrier heights. The large slope of the plateaulike region in the NDR for a smaller simulated barrier height of 0.20 eV suggests that the large charging of the QW eventually renders the stable disappearance of 2D states at the emitter in a manner of a decreasing amplitude of oscillation with bias in this region. Indeed, as can be seen from Fig. 13, the peak point of the increasing ''plateau'' nicely connects through the first peak point interpolating the I-V characteristics in the absence of oscillation and plateau. The absence of hysteresis is mainly due to the presence of large capacitance.

In our numerical simulation for different right-barrier widths (Fig. 14), the results approximately mimicked those

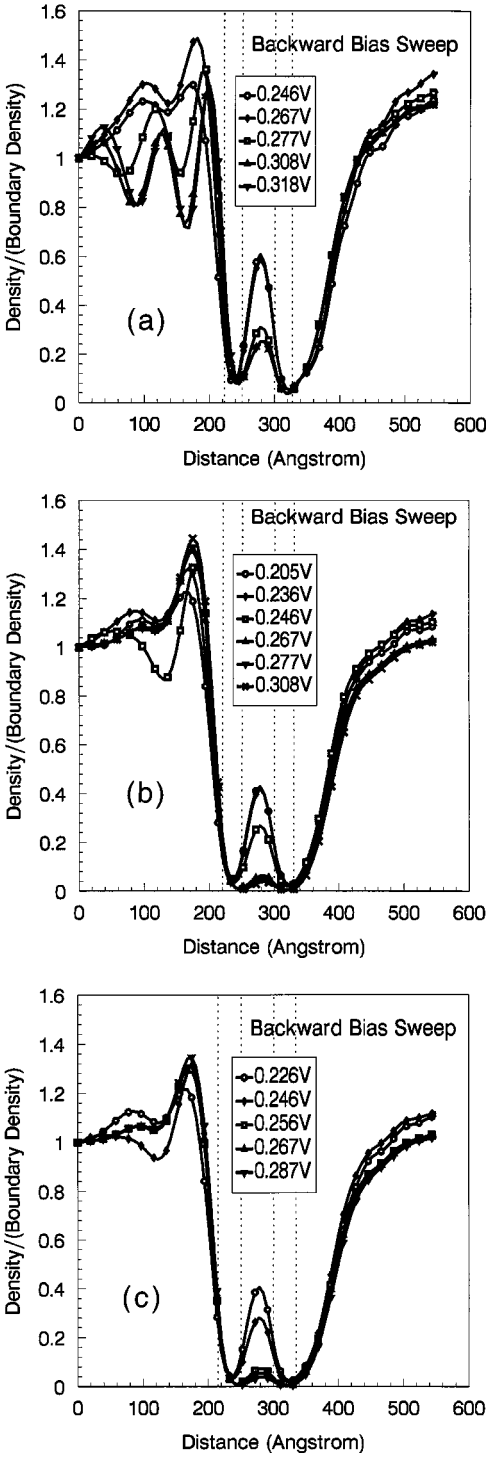

FIG. 11. Simulated backward-bias electron-density distribution for barrier widths of (a) 27 Å, (b) 30 Å, and (c) 35 Å.

of Figs. 5–7, where the right and left barriers are symmetrically varied, thus indicating the principal role of the right barrier in changing the quantizing field at the emitter. The numerical simulation for different QW widths also produces similar trend as one goes from smaller to wider QW widths, as shown by Fig. 15. Basically, the smaller QW width, with its higher QW resonant energy level, produces a stronger quantizing field at the emitter in the NDR operation than those for wider QW widths. Note the higher peak bias in Fig. 15 for the smallest QW width. Note also that the use of a

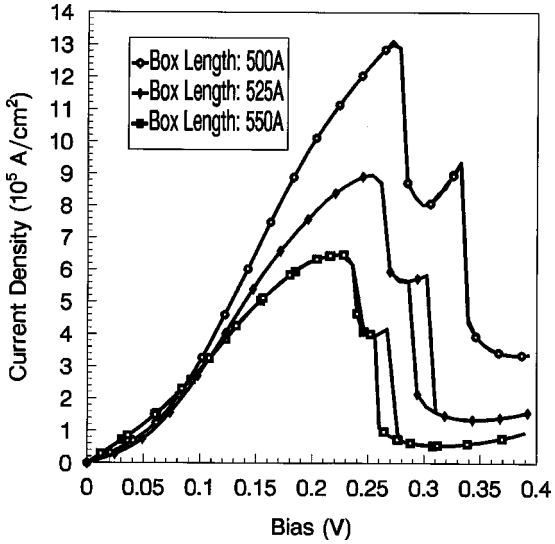

FIG. 12. Simulated current density as a function of simulation box lengths for RTS's with $b=30$ Å, $w=50$ Å, and a spacer-layer width of 30 Å. The doping density is 2.0×10^{18} cm^{-3}.

larger QW width of 60 Å in our simulation virtually eliminates the oscillatory behavior and hence the increasing plateau behavior, as shown in Fig. 15. This last remark can be understood in the light of the analytical treatment of Secs. V and VI.

Interestingly enough, the effect of varying the spacer-layer widths in our numerical simulation in Fig. 16 follows the same trend as that of Fig. 15. The only difference is in the significant change in the magnitude and shift of the current peak as function of applied bias in Fig. 15 as compared to Fig. 16. Whereas in Fig. 15 the resonant energy level is also affected as one varies the QW well widths, this is not the case as one varies the spacer-layer widths in Fig. 16. The loss of the oscillatory behavior and significant reduction in the hysteresis for a spacer-layer width of 40 Å in our numerical simulation clearly indicates that for large enough spacer-layer widths, the plateaulike behavior and hysteresis can be

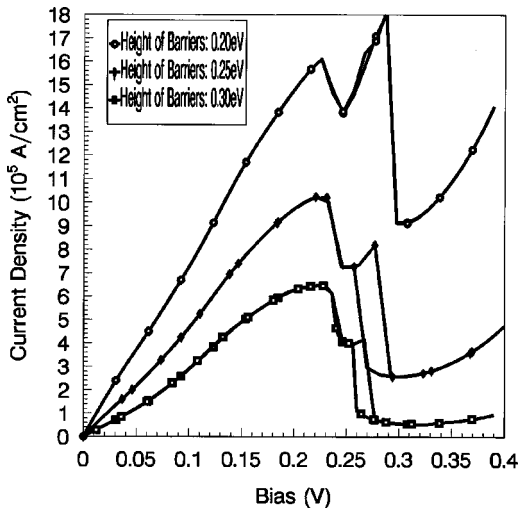

FIG. 13. Simulated current density as a function of barrier heights for RTS's with $b=30$ Å, $w=50$ Å, and a spacer-layer width of 30 Å. The doping density is 2.0×10^{18} cm^{-3}, and the simulation box length is 600 Å.

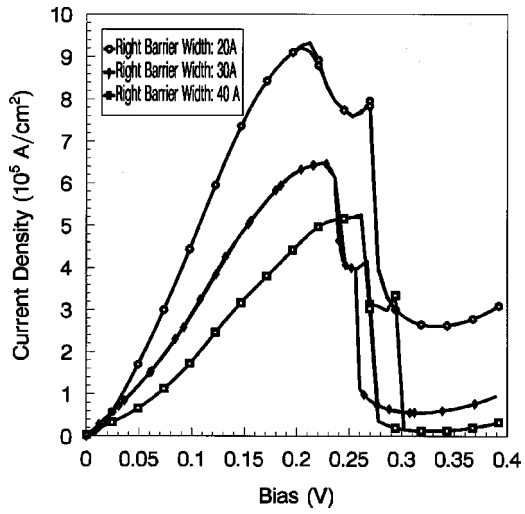

FIG. 14. Simulated current density as a function of right barrier widths for RTS's with $w=50$ Å, and a spacer-layer width of 30 Å. The doping density is $2.0\times10^{18}\,\mathrm{cm}^{-3}$, and the simulation box length is 600 Å.

completely eliminated. We have given a definite device parameter criteria for the RTS's to exhibit the ideal resonant I-V characteristic without the plateaulike behavior and hysteresis at very low temperature.

The variation of the I-V characteristics as a function of the spacer-layer widths probes the role of unstable 2D states in the emitter in the formation of the plateaulike region and hysteresis. In particular, one expects that the localization of the confining potential at the emitter is a strong function of the spacer-layer widths. For zero spacer-layer width, with a not too large doping density, the confining potential at the emitter is expected to be highly localized by virtue of the efficient screening of the doped emitter region, resulting in a larger value of the 2D energy level and hence in a larger current plateau width. At zero spacer-layer width the capacitance is also increased, resulting in a very small ratio of the hysteresis width to the plateau width, in ageement with our simulation results.

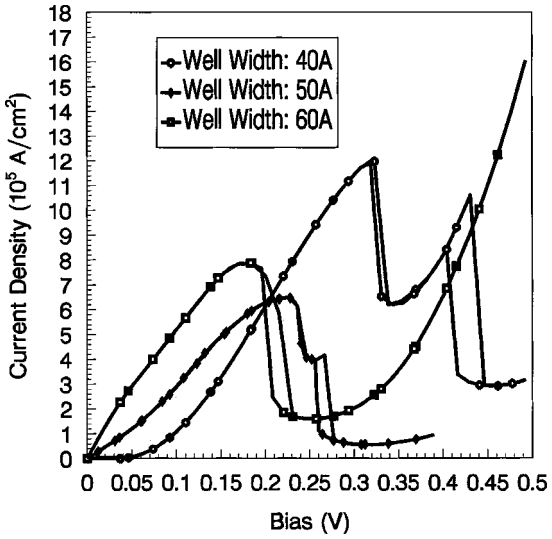

FIG. 15. Simulated current density as a function of quantum-well widths for RTS's with $b=30$ Å, and a spacer-layer width of 30 Å. The doping density is $2.0\times10^{18}\,\mathrm{cm}^{-3}$, and the simulation box length is 600 Å.

Figure 17 shows the dependence of the I-V characteristic as a function of temperature. The disappearance of plateau-like behavior and hysteresis at room temperature is expected from the analysis given in this paper, which is based on the existence of sharp energy levels. Effects on the confinement potential of temperature almost resemble the effects of spacer-layer widths (Figs. 16 and 17).

Figure 18 shows the numerical simulation of the I-V characteristic as a function of the doping density. A larger doping density creates a larger current and a smaller access resistance, such that the accompanying voltage drop in the access resistor is practically constant. Thus there is no observable shift of the current peak between the I-V characteristic of different doping densities in Fig. 18. The decreasing current in the plateau region for the highest doping density of $3\times10^{18}\,\mathrm{cm}^{-3}$ that was simulated indicates that the strong screening effect dampens the oscillation while producing a very weak penetration of the electric field at the emitter (weak quantization). Because of the interplay of the depletion and QW capacitance in a series, the capacitance parameter is probably slowly varying between different doping densities. However, the quantizing field at the emitter is rapidly varying, resulting in corresponding different widths of the plateau region with different doping densities. Thus a trend similar to that exhibited in Figs. 5–7 is obtained in going from a small doping density of $1.0\times10^{18}\,\mathrm{cm}^{-3}$ to a larger value of $3.0\times10^{18}\,\mathrm{cm}^{-3}$, where the I-V characteristics more closely follow the graphical solution of Fig. 7(a).

III. NONLINEAR MODEL OF THE OSCILLATORY DYNAMICS

Several nonlinear models were attempted in the literature to explain the oscillatory behavior of RTS's. A nonlinear model for RTS's by Abe[15] only focused on the self-consistent electron charge in the QW and its effect on the tunneling probability, i.e., only on the catalytic feedback process due to the self-consistent QW potential mentioned in Sec. I. The dynamical coupling with the emitter (reservoir) was completely ignored. Highly mathematical treatments of a similar model were given in Refs. 16 and 17. A Schrödinger equation with nonlinearities concentrated in the double-barrier region was proposed. The major assumption is that the emitter region is left undisturbed, which is unrealistic in light of all our numerical results. No comparison with salient features of the experimental results[3–4] were attempted. To our knowledge the present paper is the first to explain all the different salient features between the experimental results of Refs. 3 and 4.

A satisfactory analysis of the current voltage characteristics needs to explain not only the plateaulike behavior as a function of bias and double hysteresis of the average currents[3,4] but also the oscillatory behavior[4] found in the time-dependent simulation of the I-V characteristic of RTS's.[5,6] As seen in Sec. II, double hysteresis is a consequence of the plateaulike behavior of the average quantum-well charge. We should also explain the two different behaviors of the current in the plateau, namely, a decreasing

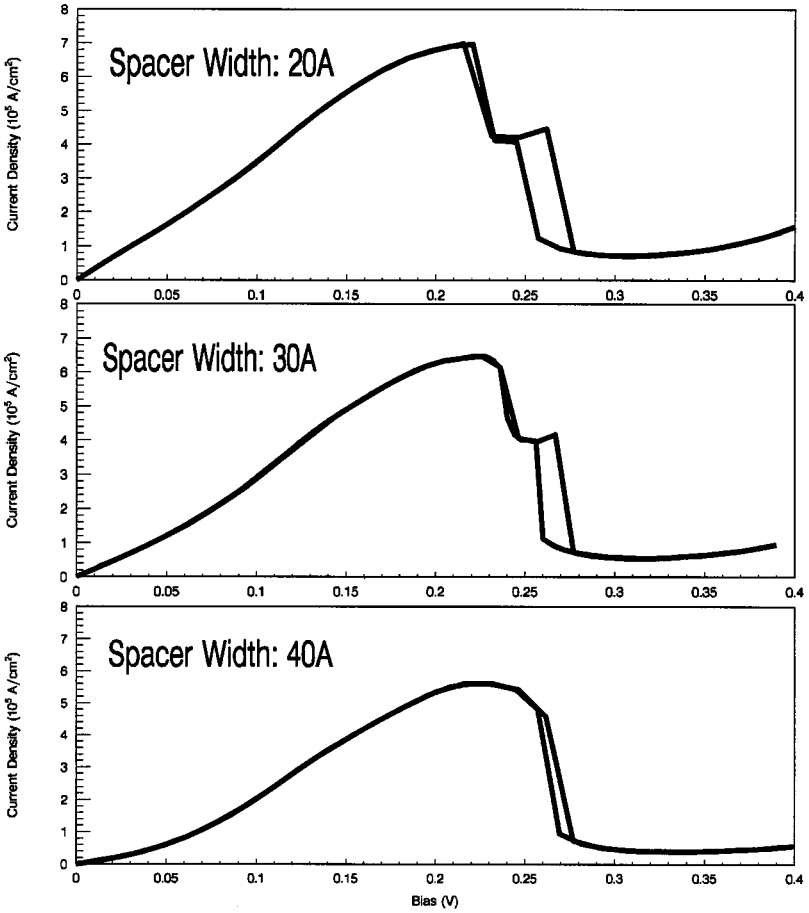

FIG. 16. Simulated current density as a function of spacer-layer widths for RTS's with $b=30$ Å, and $w=50$ Å. The doping density is 2.0×10^{18} cm^{-3}, and the simulation box length is 600 Å.

current with smaller plateau width[3] and an increasing current[4] with larger plateau width as a function of bias. Here we introduce a physical model for the oscillatory dynamics. This model does not take into account the steady leakage current from the emitter (source) to the collector (drain) while the QW energy level is above the conduction-band edge of the emitter. This dc component will be treated simply as a background dc current in the NDR, assumed to be negligible for strong emitter quantization.

Let $\mathcal{N}_e$ be the number of supply electrons at the emitter that are generated by the EQW broadening and participate in tunneling to the QW, and let $\mathcal{N}_w$ be the number of corresponding electrons generated in the quantum well. In the NDR, the frequency $\omega_{\rm eq}$ of tunneling from the emitter to the quantum well is given by $\omega_{\rm eq}=\tilde{\Delta}\mathcal{N}_e\mathcal{N}_w$, where $\tilde{\Delta}$ is the tunneling probability factor which takes into account the dependence of tunneling coefficient on the barrier height and width, taking into consideration the longitudinal quantization of the supply electrons in the emitter. $\tilde{\Delta}$ therefore could be dependent on the driving source Q/RC, which affects the existence of limit cycle solution for very small values of Q/RC. The explicit dependence of $\omega_{\rm eq}$ on $\mathcal{N}_w$ is explained as follows. In the NDR, the frequency of tunneling from the emitter to the quantum well is enhanced by the presence of $\mathcal{N}_w$. This is because the realignment of the allowed energy levels in the emitter and the QW discrete energy level is enhanced by the presence of $\mathcal{N}_w$, by virtue of the feedback due to the self-consistency of the potential. This represents the catalytic feedback process.

We can now write the ''effective'' generation rate of supply electrons at the emitter as

$$\frac{\partial}{\partial t}\mathcal{N}_e=\frac{Q}{RC}-\frac{\mathcal{N}_e}{\tau_{\rm eq}}, \tag{3}$$

where $\tau_{\rm eq}=2\pi/\omega_{\rm eq}$. As mentioned in Sec. I, Q/RC is the driving source term, Q is the electron buildup at the emitter in time constant RC to produce enough broadening of the EQW, R is the series resistance, and C is the double-barrier capacitance. Similarly, the effective generation rate of electrons in the quantum well is

$$\frac{\partial}{\partial t}\mathcal{N}_w=\frac{\mathcal{N}_e}{\tau_{\rm eq}}-\frac{\mathcal{N}_w}{\tau_c}. \tag{4}$$

Equations (3) and (4) describe the situation depicted in Fig. 1(a), namely, at $\mathcal{N}_e=0$ the generation rate of $\mathcal{N}_e$ is at its maximum while the generation rate of $\mathcal{N}_w$ is negative, i.e., $\mathcal{N}_w$ is actually decaying. An equivalent-circuit model closely describing Eqs. (3) and (4) is shown in Fig. 1(d). This is the same equivalent-circuit model for RTS's with inductive delay as introduced by Buot and Jensen[1] to explain the presence of intrinsic high-frequency oscillations in their numerical results.[5] This was discussed in more detail by Buot and Rajagopal.[11] The coupled rate equations above are similar to the one used to describe an interband-tunnel high-frequency oscillator introduced by Buot,[14] and discussed in more detail by Buot and Krowne.[18] The maximum generation rate of $\mathcal{N}_e$, given by Q/RC, is the parameter of our ''dual'' theory

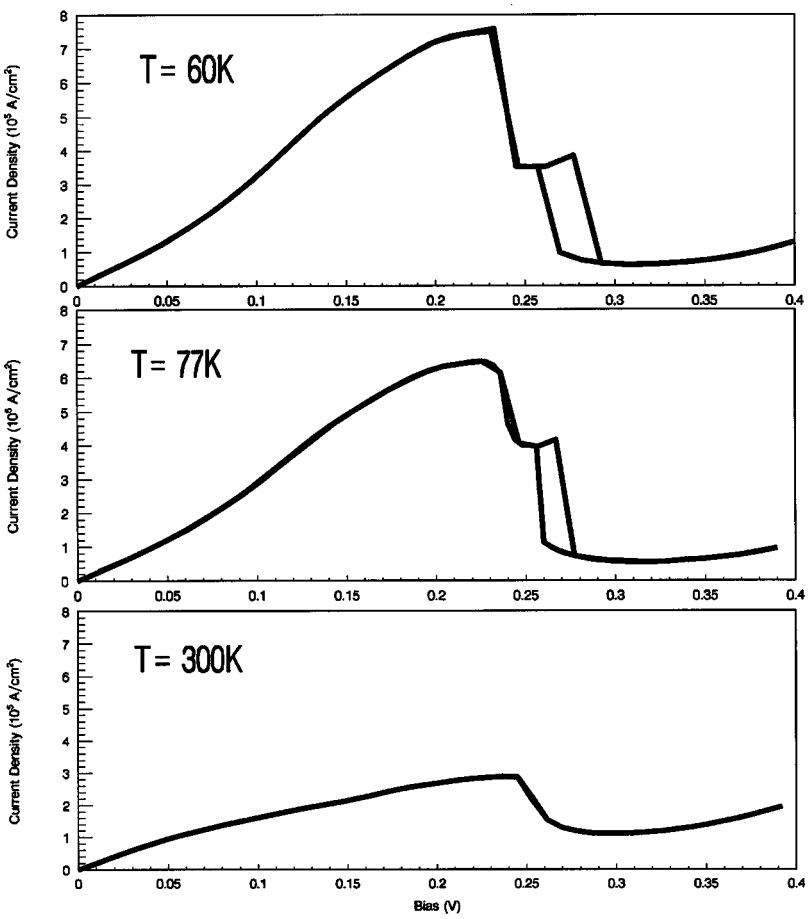

FIG. 17. Simulated current density as a function of temperature for RTS's with $b=30$ Å, $w=50$ Å, and a spacer-layer width of 30 Å. The doping density is $2.0\times10^{18}\,\mathrm{cm}^{-3}$, and the simulation box length is 600 Å.

here. This is expected to depend on the confining electric field at the emitter for a given bias, when the QW energy level passes into the forbidden region of the emitter. In the last equation, τ_c is the characteristic time for the decay of $\mathcal{N}_w$ due to tunneling to the collector.

To derive an explicit expression for the last term of Eq. (4) from a more fundamental consideration, we need to formulate the process describing the decay of $\mathcal{N}_w$ due to tunneling from the quantum well to the collector or drain. Let **N** be the total number of matching states in the drain for the electrons in the quantum well to transition to. This is a finite number for real semiconductors. Let N_x be the number of matching states already occupied by virtue of electrons transitioning to these states. The production rate of N_x is proportional to the product of the available number of matching states and $\mathcal{N}_w$. Let λ be this proportionality constant. And let γ be the decay rate of N_x by virtue of electron drift in the depletion region followed by absorption at the metal contact. Then we can write the rate equation for N_x as

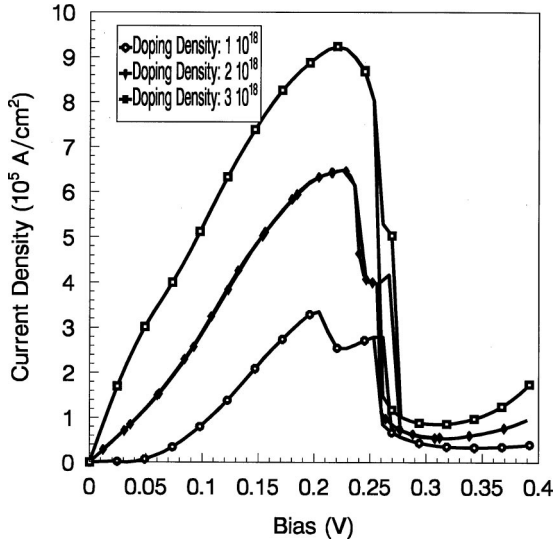

FIG. 18. Simulated current density as a function of doping density for RTS's with $b=30$ Å, $w=50$ Å, and a spacer-layer width of 30 Å. The simulation box length is 600 Å.

$$\frac{\partial}{\partial t}N_x=\lambda(N-N_x)\mathcal{N}_w-\gamma N_x, \tag{5}$$

where the first term is also the decay rate of $\mathcal{N}_w$. The process described by the last term of Eq. (5) is the fastest process in the problem, N_x is therefore expected to relax much faster than $\mathcal{N}_w$ and $\mathcal{N}_e$. Thus by adiabatic elimination of fast variables, we can let $\dot{N}_x\Rightarrow0$. Then we obtain

$$N-N_x=\frac{N}{1+\lambda\mathcal{N}_w/\gamma}. \tag{6}$$

Upon substituting the expression of Eq. (6) into the first term of Eq. (5), we obtain the decay rate of $\mathcal{N}_w$ given by $\lambda N\mathcal{N}_w/(1+\lambda\mathcal{N}_w/\gamma)$. Thus we can express the decay rate of the quantum-well electrons as

$$\mathcal{N}_w \text{ decay rate} = \frac{\alpha \mathcal{N}_w}{1+\beta \mathcal{N}_w}, \tag{7}$$

where $\alpha/\beta = \gamma N$, and $1/\beta$ is proportional to the sum of available states in the collector. Equation (7) is a typical decay rate for dynamical systems involving a fast relaxing ''sink,'' and is similar to the Michaelis-Menten decay law in chemical kinetics. The parameter $\alpha = \lambda N$ is the decay rate constant and α/β is the value of the saturated decay rate of $\mathcal{N}_w$. Therefore, we can also write the generation rate for $\mathcal{N}_w$ as

$$\frac{\partial}{\partial t}\mathcal{N}_w = \tilde{\Delta}\mathcal{N}_e^2 \mathcal{N}_w - \frac{\alpha \mathcal{N}_w}{1+\beta \mathcal{N}_w}, \tag{8}$$

where the first term is also the explicit expression for $\mathcal{N}_e/\tau_{\text{eq}}$ in Eq. (4). Comparing with Eq. (4), we obtained the following relation: $\tau_c = (1+\beta \mathcal{N}_w)/\alpha$. As seen in Eq. (10) below, the physical situation corresponds to $\alpha/\beta > Q/RC$. This means that the maximum discharging rate of the quantum well is larger than the buildup rate of supply electrons at the emitter. Indeed, we can estimate that $Q/\tau_B \approx Q/RC$ and $Q/\tau_L \leqslant \alpha/\beta$, where τ_B is the length of time to bring the device from the state of Fig. 1(a) to that of Fig. 1(b), and τ_L is the corresponding length of time for bringing back from state of Fig. 1(b) to Fig. 1(a). Therefore, the physical requirement that $\alpha/\beta > Q/RC$ implies that we may arrange for R to be large enough for a given capacitance C such that $\tau_B > \tau_L$. This is the situation we are considering in our analysis of the oscillation. Note that oscillatory behavior and hysteresis very much depends on these two characteristic times as discussed by Buot and Rajagopal.[11,13]

$\tau_B > \tau_L$ implies that at steady state one would only see the average values of the built-up charge in the emitter and broadened EQW, since the process of charge buildup takes longer compared to the time duration for discharging the emitter (which is roughly the leakage time of the QW charge to the collector). Moreover, the average QW charge will be a slowly varying function of the bias, i.e., resembling the current plateau plot. This behavior of the quantum-well charge mimicking the current plateau was indeed found by Jensen and Buot[5] in their quantum transport numerical simulation of RTS's.

IV. STABILITY ANALYSIS

The stationary solution to the coupled rate equations, (3) and (8) is given by

$$Q/RC = \tilde{\Delta}\mathcal{N}_3^2 \mathcal{N}_w = \frac{\alpha \mathcal{N}_w}{1+\beta \mathcal{N}_w}. \tag{9}$$

Thus the stationary values of $\mathcal{N}_w$ and $\mathcal{N}_e$, respectively, are given by

$$\mathcal{N}_e^0 = \left(\frac{\alpha - \beta Q/RC}{\tilde{\Delta}}\right)^{1/2},$$

$$\mathcal{N}_w^0 = \frac{Q/RC}{\alpha - \beta Q/RC}. \tag{10}$$

Note that as the drain bias approaches the plateau edge and into the valley-current value, Q also decreases, as indicated by Figs. 1(a) and 2. This situation is indicated in our simulation results by the decaying amplitude of the oscillatory behavior as the bias approaches the current-valley value.[5]

For the following stability and nonlinear analyses, it is convenient to simplify the fundamental rate equations and write them in dimensionless form as

$$\frac{\partial}{\partial \tau}\Pi = \Delta \mathcal{Q}^2 \Pi - \frac{\Pi}{1+\Pi}, \tag{11}$$

$$\frac{\partial}{\partial \tau}\mathcal{Q} = \mathcal{G} - \Delta \mathcal{Q}^2 \Pi, \tag{12}$$

where

$$\Pi = \beta \mathcal{N}_w,$$

$$\mathcal{Q} = \beta \mathcal{N}_e,$$

$$\Delta = \frac{\tilde{\Delta}(1/\beta)^2}{\alpha}, \tag{13}$$

$$\mathcal{G} = \frac{Q/RC}{(\alpha/\beta)},$$

$$\tau = \alpha t.$$

In the absence of any data, we can make a rough estimate for $\mathcal{G}$. Assume that $Q \approx 10^9\ \text{cm}^{-2}$. From Ref. 1, we can estimate $RC \approx 10^{-14}$ s. Then the driving rate (or electron-flux density source) $Q/RC \approx 10^{23}\ \text{cm}^{-2}\ \text{s}^{-1}$. From the product of $L|G| \approx 10^{-13}$ s, where L is the positive inductance of the double-barrier structure and G is the ideal negative conductance in the NDR region, determined from Ref. 1, we can also estimate $\alpha \approx 10^{13}\ \text{s}^{-1}$. We can approximate $1/\beta \approx 10^{11}\ \text{cm}^{-2}$ to represent the available states in the collector. Then $\mathcal{G} \approx 0.1$. Because of inductive delay, we estimate $\tilde{\Delta}\mathcal{N}_e^2 \ll \alpha$ using typical values of $\mathcal{N}_e$. Using these values, a reasonable value for Δ comes out to be about the same magnitude as $\mathcal{G}$, which can lead to the inequality $(1-\mathcal{G})^3/4 > \Delta$. We will see that this last inequality has a very important role in our limit cycle analysis in Sec. V. In what follows, we take $\mathcal{G} < 0.5$ to cover the physical range for $\mathcal{G}$ in the I-V plateau region.

In terms of these dimensionless variables, the stationary values of $\mathcal{Q}$ and Π are given by

$$\Pi^0 = \frac{\mathcal{G}}{1-\mathcal{G}} \equiv \beta \mathcal{N}_w^0, \tag{14}$$

$$\mathcal{Q}^0 = \left(\frac{1-\mathcal{G}}{\Delta}\right)^{1/2} \equiv \beta \mathcal{N}_e^0. \tag{15}$$

As mentioned above for a physical situation in the relevant RTS's, $0 < \mathcal{G} < 1$. The dynamical system cannot be sustained or becomes unbounded in the presence of a catalytic process, represented by the first term of Eq. (8), if $\mathcal{G} \geqslant 1.0$.

The question of whether there is a nonstationary solution to our fundamental rate equations can first be answered by examining the stability of the stationary point in $(\Pi, \mathcal{Q})$

space. This is done by examining the neighborhood of the stationary point. Let us denote the coordinates of this neighborhood by

$$\Pi = \Pi^o + p,$$

$$\mathcal{Q} = \mathcal{Q}^o + q. \tag{16}$$

Substituting these into the coupled rate equations (11) and (12), and retaining only linear terms in p and q, we have

$$\frac{\partial}{\partial \tau}\begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} \mathcal{G}(1-\mathcal{G}) & 2\mathcal{G}[\Delta/(1-\mathcal{G})]^{1/2} \\ -(1-\mathcal{G}) & -2\mathcal{G}[\Delta/(1-\mathcal{G})]^{1/2} \end{pmatrix}\begin{pmatrix} p \\ q \end{pmatrix}. \tag{17}$$

The solution for the trajectories in $(\Pi, \mathcal{Q})$ space about the equilibrium point is given by

$$\begin{pmatrix} p \\ q \end{pmatrix} = A_1 e^{\lambda_1 \tau}\begin{pmatrix} V_1^p \\ V_1^q \end{pmatrix} + A_2 e^{\lambda_2 \tau}\begin{pmatrix} V_2^p \\ V_2^q \end{pmatrix}, \tag{18}$$

where λ_1 and λ_2 are the eigenvalues of the matrix (M) defined by Eq. (17), and the corresponding eigenvectors are V_1 and V_2, respectively. The eigenvalues are given by

$$\lambda_{1,2} = \frac{\mathrm{Tr}(M)}{2} \pm \frac{1}{2}\sqrt{[\mathrm{Tr}(M)]^2 - 4\,\det(M)}. \tag{19}$$

The character of the stationary point can thus be determined with the help of the invariants of the matrix (M), namely, $\mathrm{Tr}(M)$, $\det(M)$, and $D(M) = [\mathrm{Tr}(M)]^2 - 4\det(M)$. The stationary point can not be a saddle point for physical reason since $\det(M) = 2\mathcal{G}(1-\mathcal{G})^{3/2}\sqrt{\Delta} > 0$. The physical processes depicted in Fig. 1 also suggest that the stationary point can only be any one of the following cases: stable focus $[\mathrm{Tr}(M) < 0]$, center focus $[\mathrm{Tr}(M) = 0]$, or unstable focus $[\mathrm{Tr}(M) > 0]$. These cases mean that $D(M) < 0$ or $4\det(M) > [\mathrm{Tr}(M)]^2$. Thus there are two out of three chances that the oscillating processes depicted in Figs. 1 and 2 are sustained depending on the value of Δ relative to $\mathcal{G}$. On physical grounds, we expect the limit cycle solution for uniqueness and structural stability. For the unstable focus we have to demonstrate that a limit cycle exists. The region in $(\mathcal{G}, \Delta)$ parameter space where the structurally stable limit cycle is possible lies in the area under the bifurcation curve [the locus of $\mathrm{Tr}(M) = 0$] in this space.

The trace of M is given by $\mathrm{Tr}(M) = \mathcal{G}[(1-\mathcal{G}) - 2\{\Delta/(1-\mathcal{G})\}^{1/2}]$. Thus $\mathrm{Tr}(M) > 0$ implies $(1-\mathcal{G})^3 > 4\Delta$. On the other hand, $D(M) < 0$ implies $(1-\mathcal{G})^3 < 4\Delta + 8\mathcal{G}^{-1}(1-\mathcal{G})^{5/2}\Delta^{1/2} + 4\sqrt{2}\mathcal{G}^{-1}(1-\mathcal{G})^{5/4}\Delta^{3/4}$. In Sec. V, we will employ a nonlinear perturbation technique using the method of multiple time scales with values of the parameter around $\mathrm{Tr}(M) = 0$. As we shall show in the following nonlinear analysis, the limit cycle indeed occurs at $\mathrm{Tr}(M) > 0$. The amplitude and frequency of oscillation is expected to depend on the actual values of the two parameters $\mathcal{G}$ and Δ in this region.

V. LIMIT CYCLE SOLUTION

Retaining nonlinear terms for p and q measured from the stationary point, the rate equation from Eqs. (11) and (12) becomes a matrix equation,

$$\frac{\partial}{\partial \tau}\begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} \mathcal{G}(1-\mathcal{G}) & 2\mathcal{G}[\Delta/(1-\mathcal{G})]^{1/2} \\ -(1-\mathcal{G}) & -2\mathcal{G}[\Delta/(1-\mathcal{G})]^{1/2} \end{pmatrix}\begin{pmatrix} p \\ q \end{pmatrix} + \begin{pmatrix} N^p \\ N^q \end{pmatrix}, \tag{20}$$

where

$$N^p = (1-\mathcal{G})^3 p^2 + 2[\Delta(1-\mathcal{G})]^{1/2} pq + \frac{\Delta\mathcal{G}}{(1-\mathcal{G})}q^2 + \Delta p q^2 + \sum_{n=3}^{\infty}(-1)^n (1-\mathcal{G})^{n+1} p^n, \tag{21}$$

$$N^q = -2[\Delta(1-\mathcal{G})]^{1/2} pq - \frac{\Delta\mathcal{G}}{(1-\mathcal{G})}q^2 - \Delta p q^2. \tag{22}$$

The perturbation technique employed in what follows essentially transforms the above nonlinear equation into a hierarchy of solvable and simpler equations, obtained by equating coefficients of powers of the smallness parameter. Near $\mathrm{Tr}(M) = 0$, we use as our smallness parameter the departure of $\mathrm{Tr}(M)$ from zero, i.e., the departure of Δ from Δ_c where $(1-\mathcal{G})^3 = 4\Delta_c$. Thus, let the smallness parameter be $\epsilon = \sqrt{\{\Delta - [(1-\mathcal{G})^3/4]\}/\mathcal{D}}$, where $\mathcal{D}$ is determined from the expansion of Δ in powers of ϵ. $\mathcal{D} \approx \Delta_2$ in the analysis that follows. $\mathcal{G}$ is assumed constant at fixed bias, i.e., a function only of the external bias. We make the following expansion:

$$\Delta = \sum_{j=0}^{\infty} \epsilon^j \Delta_j, \quad \text{where } \Delta_0 = \Delta_c. \tag{23}$$

We also expand the matrix M in powers of ϵ through direct Taylor expansion in powers of $\Delta - \Delta_c$ as

$$(M) = (M_c) + \epsilon\Delta_1\left(\left.\frac{\partial M(\Delta)}{\partial \Delta}\right|_{\Delta=\Delta_c}\right) + \frac{1}{2}\epsilon^2\left[\Delta_2\left(\left.\frac{\partial M(\Delta)}{\partial \Delta}\right|_{\Delta=\Delta_c}\right) + \Delta_1^2\left(\left.\frac{\partial^2 M(\Delta)}{\partial \Delta^2}\right|_{\Delta=\Delta_c}\right)\right] + O(\epsilon^3). \tag{24}$$

Using $(1-\mathcal{G})^3 = 4\Delta_c$, we obtain the following expressions:

$$(M_c) = \mathcal{G}(1-\mathcal{G})\begin{pmatrix} 1 & 1 \\ -\mathcal{G}^{-1} & -1 \end{pmatrix}, \tag{25}$$

$$(M_1) \equiv \left(\left.\frac{\partial (M)}{\partial \Delta}\right|_{\Delta=\Delta_c}\right) = 2\mathcal{G}(1-\mathcal{G})^{-2}\begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix}, \tag{26}$$

$$(M_2) \equiv \left(\left.\frac{1}{2}\frac{\partial^2 (M)}{\partial \Delta^2}\right|_{\Delta=\Delta_c}\right) = -2\mathcal{G}(1-\mathcal{G})^{-5}\begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix}, \tag{27}$$

We let the solution depend on time τ in a combination $\tau_0 = \tau$ and $\tau_1 = (\Delta - \Delta_c)\tau$. Thus, instead of determining the solution in terms of τ, we seek the solution as a function of τ_0, τ_1, and ϵ. This method of doing the nonlinear perturbation analysis is well known, and is often referred to as the method of multiple time scales. This has the virtue that it separates the dependence of the solution into fast and slow

time scales. For limit cycle behavior, for example, we expect that the amplitude of the oscillation is only a function of the slow time scale. The left side of the rate equation can now be written as

$$\frac{\partial}{\partial\tau}\begin{pmatrix} p(\tau_o,\tau_1,\epsilon) \\ q(\tau_o,\tau_1,\epsilon)\end{pmatrix}=\left\{\frac{\partial}{\partial\tau_o}+(\Delta-\Delta_c)\frac{\partial}{\partial\tau_1}\right\}\begin{pmatrix} p(\tau_o,\tau_1,\epsilon) \\ q(\tau_o,\tau_1,\epsilon)\end{pmatrix}. \tag{28}$$

Since the last term in Eq. (20) represents the nonlinear term for the solution, we adopt the following expansion:

$$\begin{pmatrix} p \\ q\end{pmatrix}=\sum_{j=0}^{\infty}\epsilon^{j+1}\begin{pmatrix} p_j \\ q_j\end{pmatrix}. \tag{29}$$

Therefore, any finite solution that will be found in this analysis will invariably indicate that the limit cycle occurs for values of the parameter away from the critical point, $\mathrm{Tr}(M)=0$, i.e., away from the bifurcation point. This holds, for example, in our numerical nonlinear equivalent-circuit simulation for the limit cycle of conventional RTS's operating at the NDR region.[2] With Eq. (29), the nonlinear term in Eq. (20) acquires the following expansion in terms of the smallness parameter:

$$\begin{pmatrix} N^p \\ N^q\end{pmatrix}=\epsilon^2\begin{pmatrix} N_2^p \\ N_2^q\end{pmatrix}+\epsilon^3\begin{pmatrix} N_3^p \\ N_3^q\end{pmatrix}+O(\epsilon^4), \tag{30}$$

where

$$\begin{pmatrix} N_2^p \\ N_2^q\end{pmatrix}=(1-\mathcal{G})^2\begin{pmatrix} [p_0q_0+(\mathcal{G}/4)q_0^2+(1-\mathcal{G})p_0^2] \\ -[p_0q_0+(\mathcal{G}/4)q_0^2]\end{pmatrix}. \tag{31}$$

$$\begin{pmatrix} N_3^p \\ N_3^q\end{pmatrix}=(1-\mathcal{G})^2\begin{pmatrix} \left[p_0q_1+p_1q_0+(\mathcal{G}/2)q_1q_0+\frac{(1-\mathcal{G})}{4}p_0q_0^2+\frac{\Delta_1\mathcal{G}}{(1-\mathcal{G})^3}q_0^2 \right. \\ \left. +\left\{\frac{\Delta_2}{(1-\mathcal{G})^3}\right\}^{1/2}2p_0q_0+2(1-\mathcal{G})p_0p_1-(1-\mathcal{G})^2p_0^3\right] \\ -\left[p_0q_1+p_1q_0+(\mathcal{G}/2)q_1q_0+\frac{(1-\mathcal{G})}{4}p_0q_0^2+\frac{\Delta_1\mathcal{G}}{(1-\mathcal{G})^3}q_0^2 \right. \\ \left. +\left\{\frac{\Delta_2}{(1-\mathcal{G})^3}\right\}^{1/2}2p_0q_0\right]\end{pmatrix}. \tag{32}$$

We did not show nonlinear terms with fractional powers of ϵ in Eq. (30) associated with Δ_1 in Eqs. (21) and (22), since the left-hand side of the rate equation does not contain fractional powers of ϵ. To eliminate the occurrence of these fractional powers of ϵ, we have to make $\Delta_1\equiv 0$ in the expansion of Δ [Eq. (23)] and also in Eqs. (24) and (32).

Upon substituting all the expanded quantities in the nonlinear rate equation [Eq. (20)], we obtain a hierarchy of simpler equations. Those arising from the first up to the third powers of ϵ are given below:

$$\mathcal{L}_0\begin{pmatrix} p_0 \\ q_0\end{pmatrix}=0. \tag{33}$$

$$\mathcal{L}_0\begin{pmatrix} p_1 \\ q_1\end{pmatrix}=\begin{pmatrix} N_2^p(p_0,q_0) \\ N_2^p(p_0,q_0)\end{pmatrix}. \tag{34}$$

$$\mathcal{L}_0\begin{pmatrix} p_2 \\ q_2\end{pmatrix}+\Delta_2\mathcal{L}_1\begin{pmatrix} p_0 \\ q_0\end{pmatrix}=\begin{pmatrix} N_3^p(p_0,q_0,p_1,q_1) \\ N_3^q(p_0,q_0,p_1,q_1)\end{pmatrix}, \tag{35}$$

where

$$\mathcal{L}_0=\left(\frac{\partial}{\partial\tau_0}-(M_c)\right). \tag{36}$$

$$\mathcal{L}_1=\left(\frac{\partial}{\partial\tau_1}-(M_1)\right). \tag{37}$$

The first equation in the hierarchy turns out to be a simple eigenvalue problem, analogous to our linear-stability analysis before. The only difference is that the present eigenvalue problem has to be solved with values of the parameter at the critical point, where $\mathrm{Tr}(M)=0$, using the matrix (M_c). The solutions to Eqs. (33), (34), and (35) are given in the Appendix to second order in ϵ, where it is shown that limit cycle exists for $\Delta<\Delta_c$.

Thus, to second order in the smallness parameter, and by virtue of Eqs. (16) and (23), we have the limit cycle solution given as

$$\begin{pmatrix} \Pi \\ \mathcal{Q}\end{pmatrix}=\begin{pmatrix} \Pi^0 \\ \mathcal{Q}^0\end{pmatrix}+\left(\left|\frac{\Delta-\Delta_c}{\Delta_2}\right|\right)^{1/2}\begin{pmatrix} p_0 \\ q_0\end{pmatrix}+\left(\left|\frac{\Delta-\Delta_c}{\Delta_2}\right|\right)\begin{pmatrix} p_1 \\ q_1\end{pmatrix} +O\left[\left(\left|\frac{\Delta-\Delta_c}{\Delta_2}\right|\right)^{3/2}\right], \tag{38}$$

where from Eq. (A3) in the Appendix, we have

$$\begin{pmatrix} p_0 \\ q_0\end{pmatrix}=\frac{|\Theta(\infty)|}{\mathcal{G}^{1/2}}\begin{pmatrix} 2\cos\Omega\tau \\ -2\mathcal{G}^{1/2}\cos\Omega\tau-(1-\mathcal{G})^{1/2}2\sin\Omega\tau\end{pmatrix}, =\frac{2|\Theta(\infty)|}{\mathcal{G}^{1/2}}\begin{pmatrix} \cos\Omega\tau \\ -\sin(\Omega\tau+\Phi)\end{pmatrix}, \tag{39}$$

where

$$\tan\Phi=\left\{\frac{\mathcal{G}}{(1-\mathcal{G})}\right\}^{1/2}, \tag{40}$$

$$\Omega\equiv\sqrt{\mathcal{G}}(1-\mathcal{G})^{3/2}+\left\{\mathrm{Im}\,\eta(\Delta-\Delta_c)+\frac{\mathrm{Im}\,\sigma}{\Delta_2}|\Theta(\infty)|^2(\Delta-\Delta_c)\right\}. \tag{41}$$

From Eqs. (A7), (A10), and (A11) in the Appendix, we also have

$$\begin{pmatrix}p_1\\ q_1\end{pmatrix}=|\Theta(\infty)|^2\frac{(1-\mathcal{G})}{\mathcal{G}}\begin{pmatrix}2\mathcal{G}\\ -\frac{1}{2}\end{pmatrix}+|\Theta(\infty)|^2\begin{pmatrix}\frac{2}{6\mathcal{G}(1-\mathcal{G})}\{4(1-\mathcal{G})^3\cos2\Omega\tau-\omega(1+2\mathcal{G})\sin2\Omega\tau\}\\ \frac{2}{6\mathcal{G}(1-\mathcal{G})}\left\{\left(\frac{15}{2}\mathcal{G}-\frac{9}{2}-3\mathcal{G}^2\right)(1-\mathcal{G})\cos2\Omega\tau\right.\\ \left.+\omega(8-3\mathcal{G}-2\mathcal{G}^{-1})\sin2\Omega\tau\right\}\end{pmatrix}. \tag{42}$$

We note that Eq. (42) also contains a time-independent term, indicating a higher-order shift of the center of the limit cycle from the stationary point $(\mathcal{Q}^0,\Pi^0)$. Therefore, the average value of $\binom{\Pi}{\mathcal{Q}}$ is given by

$$\begin{pmatrix}\Pi\\ \mathcal{Q}\end{pmatrix}_{\text{average}}=\begin{pmatrix}\Pi^0\\ \mathcal{Q}^0\end{pmatrix}+\left|\frac{\Delta-\Delta_c}{\Delta_2}\right||\Theta(\infty)|^2\frac{(1-\mathcal{G})}{\mathcal{G}}\begin{pmatrix}2\mathcal{G}\\ -1/2\end{pmatrix}+O(\epsilon^3)=\begin{pmatrix}\Pi^0\\ \mathcal{Q}^0\end{pmatrix}+(\text{higher-order corrections}), \tag{43}$$

where the leading higher-order corrections come from the time-independent terms. Thus we have demonstrated that a unique limit cycle exists away from the bifurcation point $(\Delta<\Delta_c)$, and the average value is determined by the time-independent terms of Eq. (16).

In examining the dependence of various quantities on the driving source, $\mathcal{G}$, we make the assumption that $|\Delta-\Delta_c|$ is approximately a constant. The physical reason for this is that Δ is a measure of the inductive delay, indicated in Fig. 1(d). Δ is large for small inductive delay and small for large inductance. For a larger driving source, $\mathcal{G}$, which happens immediately after the current peak bias, Δ_c is smaller and we also expect larger inductive delay at this point, and hence Δ is also smaller. For larger values of bias in the plateau, meaning a weaker driving source $\mathcal{G}$, Δ_c is larger and we also expect the inductive delay to be smaller, meaning Δ is also larger. Therefore, $|\Delta-\Delta_c|$ is approximately constant. It is taken small enough such that the second term in Eq. (43) is only a very small correction to the first term; otherwise one has to include other higher-order terms.

VI. DISCUSSIONS

A. Average value of the current in the plateau range

We see that the limit cycle occurs within the range of values of the parameter Δ where the criterion for unstable focus $(1-\mathcal{G})^3>4\Delta$ [i.e., $\mathrm{Tr}(M)>0$] holds, supporting the results of our numerical calculation of the limit cycle of an $Al_xGa_{1-x}As/GaAs/Al_xGa_{1-x}As$ double-barrier heterostructure operating in the NDR region.[2] Because of strong correlation of the oscillatory charges $\mathcal{N}_w$ and $\mathcal{N}_e$, and the consequent current oscillation caused by the alternate buildup of $\mathcal{N}_w$ and $\mathcal{N}_e$, the average value of $\mathcal{N}_w$ alone does not determine the measurable averaged value of the current. The averaged measurable value of the current is determined by the sum of the stationary values of $\mathcal{N}_w$ and $\mathcal{N}_e$, more specifically, the average current $I_{\mathrm{dc}}\equiv(\mathcal{N}_e^0+\mathcal{N}_w^0)/\tau_{\mathrm{eff}}$, as indicated in the equivalent-circuit model of Fig. 1(c), where $\tau_{\mathrm{eff}}\approx2\tau_c$. From Eqs. (13), (14), and (15), the leading average value of this sum is determined by

$$\Sigma^0(\mathcal{G})=\Pi^0+\mathcal{Q}^0=\frac{\mathcal{G}}{(1-\mathcal{G})}+\left(\frac{1-\mathcal{G}}{\Delta}\right)^{1/2}. \tag{44}$$

We have

$$\frac{d\Sigma^0(\mathcal{G})}{d\mathcal{G}}=\frac{1}{(1-\mathcal{G})^2}-\left(\frac{1}{4\Delta(1-\mathcal{G})}\right)^{1/2}<0 \tag{45}$$

for $\Delta<\Delta_c$ at the limit cycle. This slope goes to zero at $\Delta=\Delta_c$ and becomes positive at $\Delta>\Delta_c$, which defines the absence of inherent self-oscillation. We conclude from Eq. (45) that the average of the oscillatory current increases as $\mathcal{G}$ decreases (or as bias increases), shown in Fig. 19. $\mathcal{G}$ is large right after the current peak, and decreases with applied bias in the plateau range. We expect the charging rate Q/RC to be proportional to the energy difference (which decreases with bias) between the QW energy level and the conduction-band edge of the emitter.

We estimated $\mathcal{G}$ just beyond the current peak to be <0.5, and it decreases as a function of bias in the plateau range. Indeed, our nonlinear model supports the increasing current in the plateau range as a function of bias in the presence of intrinsic oscillations, as indicated by Fig. 19. It is important to point out, however, that in the absence of intrinsic oscillatory behavior, our nonlinear model, as described by the equivalent circuit of Fig. 1(d), indicates that the stationary current is proportional to $\mathcal{N}_w^0$ alone, which increases with the charging rate in the emitter, $\mathcal{G}$, shown in Fig. 20. Thus in the

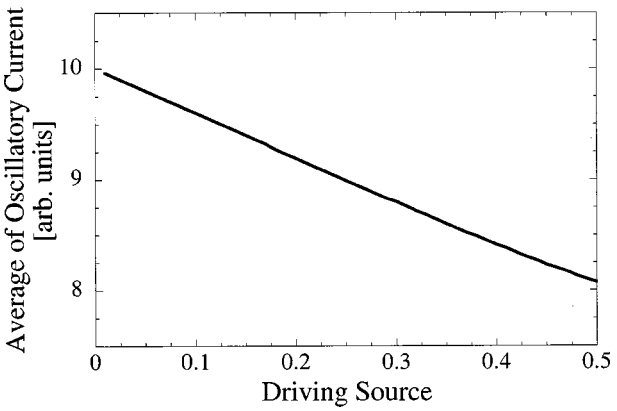

FIG. 19. Plot of the average of the oscillatory current as a function of the driving source $\mathcal{G}=(Q/RC)/(\alpha/\beta)$, for $\Delta<\Delta_c$. The increasing drain bias goes with decreasing values of $\mathcal{G}$.

absence of intrinsic oscillation, the current is expected to exhibit a decreasing behavior as a function of bias (increase of bias means a decrease of $\mathcal{G}$) in the plateau range. Indeed, results of our recent steady-state numerical simulation for a large barrier width, where the quantization field in the emitter is very weak, show a decreasing current as a function of bias in the plateau range, indicating the absence of noticeable oscillation. Moreover, the background current mentioned in Sec. III is expected to be larger for weak emitter quantization, i.e., less control of the total current by the driving source $\mathcal{G}$. Thus our nonlinear model is able to discriminate the presence or absence of oscillation in the plateaulike region, and is definitely a refinement of the graphical analysis presented in Sec. II.

B. Relation between amplitude and driving source, $\mathcal{G}$

The oscillatory current is proportional to $p/\tau_c+\partial q/\partial\tau$ [refer to Fig. 1(d)] expanded to second order using Eqs. (38)–(42). We now show that the amplitude of oscillation increases with $\mathcal{G}$. From Eqs. (38) and (39), this amplitude of the fundamental frequency component is typified by the expression

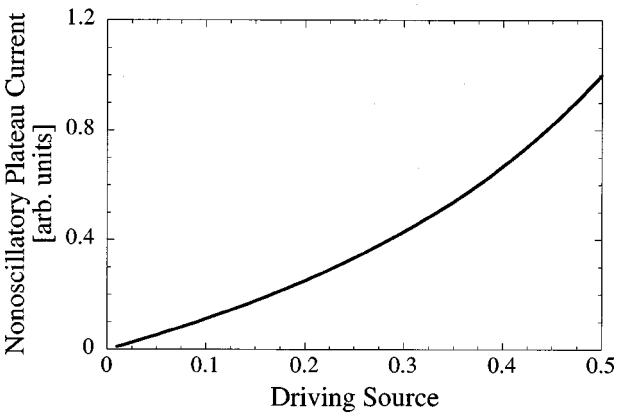

FIG. 20. Plot of the stable stationary values of the plateau current as a function of $\mathcal{G}$ for $\Delta>\Delta_c$. The increasing drain bias goes with decreasing values of $\mathcal{G}$.

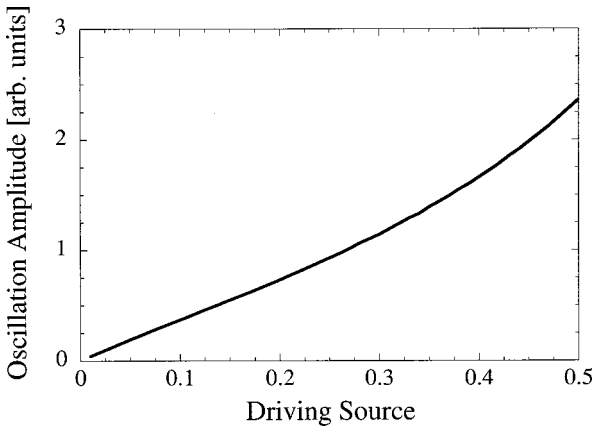

FIG. 21. Plot of the amplitude of the oscillation of an oscillatory current as a function of $\mathcal{G}$ for $\Delta<\Delta_c$. The increasing drain bias goes with decreasing values of $\mathcal{G}$.

$$\left(\left|\frac{\Delta-\Delta_c}{\Delta_2}\right|\right)^{1/2} 2\mathcal{G}^{-1/2}|\Theta(\infty)|\Omega$$

$$=\Omega\left(\left|\frac{\Delta-\Delta_c}{\Delta_2}\right|\right)^{1/2} 2\mathcal{G}^{-1/2}\left[-\frac{\mathrm{Re}\,\eta\Delta_2}{\mathrm{Re}\,\sigma}\right]^{1/2}, \tag{46}$$

where η is given by Eq. (A21) and σ is determined from Eq. (A27) in the Appendix. From Eqs. (A21) and (A27), we have the final expression for the representative amplitude, $\mathcal{A}_0$, given by

$$\mathcal{A}_0=\frac{\Omega}{\beta}(|\Delta-\Delta_c|)^{1/2}\left[\frac{64}{(1-\mathcal{G})^5\{4+19\mathcal{G}-8\mathcal{G}^2\}}\right]^{1/2}. \tag{47}$$

Figure 21 is a plot of this amplitude as a function of $\mathcal{G}$. Since $\mathcal{G}$ is a decreasing function of bias in the plateau range as discussed above, we conclude that the amplitude of oscillation is also a decreasing function of bias in the plateau range, as indicated by Fig. 21. The interference between the two cosine terms in the expression for the current further enhances this behavior in the physical range of $\mathcal{G}$. Indeed, all time-dependent numerical simulation of RTS's (Refs. 5 and 6) found the largest oscillation amplitude right after the current peak, and after which it decreases as a function of bias in the plateau range. These salient features, namely, the increase of the averaged current and the decrease in oscillation amplitude in the plateau range, have never been explained before, to our knowledge. It should be pointed out that the simulations of Biegel and Plummer[6] actually showed a small-amplitude oscillation at the far end of the plateau, which they interpreted as decaying to the steady state in the limit of large time.

Another important salient feature of I-V curves of the time-dependent simulation is that the oscillating current just after the current peak contains other harmonics, whereas at higher bias in the plateau it becomes purely harmonic. This is understandable since the largest amplitude occurs immediately after the current peak, with higher-order harmonics making contributions to the fundamental oscillating current.

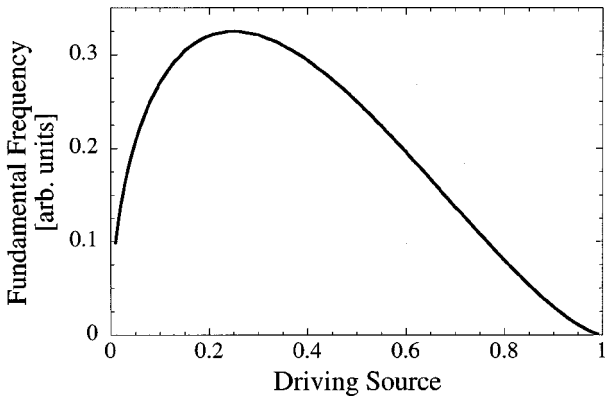

FIG. 22. Plot of $\sqrt{\mathcal{G}(1-\mathcal{G})^3}$ as a function of $\mathcal{G}$ for $\Delta<\Delta_c$.

C. Oscillation frequency and its dependence on charge buildup rate

The fundamental frequency of oscillation is given by $\omega_0 = \alpha\Omega$, where α becomes the conversion of τ to real time. We have

$$\omega_0 \equiv \alpha\sqrt{\mathcal{G}}(1-\mathcal{G})^{3/2} + \alpha\left\{ \mathrm{Im}\ \eta(\Delta-\Delta_c) + \frac{\mathrm{Im}\,\sigma}{\Delta_2}|\Theta(\infty)|^2(\Delta-\Delta_c)\right\}, \tag{48}$$

where $\alpha=\lambda N$ as defined in Sec. III. Thus oscillation is indeed driven by the rate of charging of the emitter, $\mathcal{G}$, in the presence of quantizing field causing premature current peak to occur. The scaling factor $\alpha=\lambda N$ is expected to bring the frequency values to the THz range in our nonlinear model, in agreement with our time-dependent numerical simulation of RTS's.[5] Using our estimates in Sec. III for $\alpha\approx 10^{13}\,\mathrm{s}^{-1}$ and $\mathcal{G}\approx 0.1$, for the leading term of Eq. (48) we obtain $\omega_0 \approx 2.5\,\mathrm{THz}$.

It is more revealing to examine the dependence of the fundamental frequency on the charge buildup rate at the emitter. From Eq. (48), the fundamental frequency is proportional to $\sqrt{\mathcal{G}(1-\mathcal{G})^3}$; this is plotted in Fig. 22. This figure shows that within the range of $\mathcal{G}$ considered in Figs. 19–21, the frequency temporarily increases as $\mathcal{G}$ decreases toward zero (with increase in bias) in the plateau. Comparing with Fig. 21, we see that while the amplitude of oscillation decreases, the frequency of oscillation increases at first before sharply decreasing to zero toward the end of the plateau. These behaviors were indeed noticeable in various numerical quantum transport simulations of RTS's,[1,5–6] clearly demonstrating the validity of our analytical model.

D. External sources and *I-V* measurements: resolution of controversy

Another interesting result which follows from our analysis is the case where Δ is relatively large, meaning the inductive delay is very small, or the inductance in the equivalent-circuit model of RTS's is very small. The intrinsic self-oscillation would then be absent since $\Delta>\Delta_c$; the stable stationary current is only determined by Π^0 of Eq. (14); this was discussed after Eq. (45). However, if the device is connected to an external circuit with significant lead inductance, then the whole system would still self-oscillate at the NDR by virtue of the two energy storage circuit elements; the RTS capacitance (in parallel with the NDR circuit element) and the dominant wiring inductance in series. For this particular case, the driving source $\mathcal{G}$ becomes externally induced. We can still apply Eqs. (44) and (45) for calculating the current in the plateau range. However, since $\Delta>\Delta_c$, the slope of the average of the oscillating current as a function of $\mathcal{G}$ is positive, and this translates to a decreasing average current in the plateau.

This result seems to occur in actual *I-V* measurement reported in Ref. 3, where the barrier width of 85 Å employed was in the range of size where our simulation would indicate that the quantizing field in the emitter is relatively weak when the QW energy level passes into the forbidden region of the emitter [refer to the decreasing current in the plateau of Fig. 7(b), where the simulated barrier width of 35 Å is even smaller than 85 Å]. It follows that the inductive delay is also small in the device of Ref. 3. Also their use of a ''biasing'' capacitor across the device has the effect of increasing the dimension c, and perhaps also a, in Eq. (1) compared to one without biasing capacitor, for their doping level of $2\times 10^{17}\,\mathrm{cm}^{-3}$. This will cause a decrease in the slope of Q_w as a function of k_z^2 in Eq. (1). In our graphical analysis, this would mean an increase in the widths of the two hystereses while also making the double hystereses approach each other in the plateau region, compared to their measurement without the biasing capacitor. This explains the experimental results of Ref. 3.

On the other hand, an experimental rising current in the plateau region was reported by Sollner,[4] where the device was also found to be oscillating as measured by spectrum analyzer. In both experiments, the double hystereses are clearly manifested in the *I-V* characteristics, with Sollner's device yielding an order of magnitude larger current than the device of Ref. 3. Although the device parameters were not given in Sollner's paper, we suspect that the measured device has enough intrinsic inductive delay to exhibit intrinsic oscillation. This is basically supported by the larger width of the plateau region of about $0.6V$ reported by Sollner,[4] compared to the width of $0.07V$ reported in Ref. 3, about an order of magnitude smaller. This observation is consistent with our simulation of the effect of barrier widths on the slope of the current and width of the plateau (Figs. 5–7). Therefore, only when the intrinsic inductance of the RTS's dominates the lead inductance do the measured *I-V* characteristics resemble the intrinsic one where oscillation is present. However, we should also point out that according to the discussion in Sec. VI A, hysteresis in the *I-V* characteristics can occur without intrinsic oscillation, characterized by a decreasing current in the plateau, by virtue of the possibility of stable stationary solutions for a weak quantizing field in the emitter.

VII. CONCLUDING REMARKS

We have demonstrated that the formation of unstable 2D states in the emitter under NDR operation plays a principal role in explaining a host of numerical and experimental results on RTS's. We have given a graphical explanation of our steady-state simulation results of RTS's with different

barrier widths, holding other dimensions fixed. We use a graphical method similar to that employed before[12–14] to explain other nonlinear effects in RTS's. Figures 5–7 specifically explain the following salient features of the simulation: (1) the width of the plateau increases with the decrease of barrier widths; (2) the ratio of the width of the primary hysteresis to the width of the plateau increases with the barrier width; (3) the average slope of the plateau decreases with increase in barrier widths; and (4) double hysteresis is a rule rather than an exception. These figures show, in accordance with simulation and experimental results, that the width of the secondary hysteresis is smaller than the primary one. We should emphasize that to our knowledge no other analysis can explain the presence of double hysteresis in the I-V characteristics so ubiquitous in the simulation[5] and experimental results[3,4] of resonant tunneling structures. Furthermore, the present explanations also account for the very sharp I-V peak found experimentally,[3–5] as indicated by Fig. 4.

Our graphical analysis also explained the following salient features in the NDR region of the I-V curves: (a) the width of the plateaulike region increases with a decrease in the simulation box lengths, barrier heights, right-barrier widths, QW widths, and spacer-layer widths; (b) the ratio of the width of the primary hysteresis to the width of the plateaulike region increases with an increase in the simulation box lengths, barrier heights, right-barrier widths, QW widths, and spacer-layer widths; (c) the average slope of the plateaulike region decreases with an increase in the simulation box lengths, barrier heights, right-barrier witdths, QW widths, and spacer-layer widths; (d) in almost all cases, there is a definite tendency to form a secondary hysteresis between the plateaulike region and the peak current; and (e) the plateaulike region and hysteresis disappear for large enough spacer-layer widths and higher temperatures.

We have also presented refinements of our graphical analysis by introducing a nonlinear physical model to describe the time-dependent oscillation in the rising plateau of the I–V characteristics. The analytical limit cycle solution supports the current oscillation in the plateau range found in various numerical quantum transport simulations of RTS's,[5,6] as well as in other analytical models employing the solution of the many-particle Schrödinger equation.[19] Specifically, our nonlinear model predicts a rising plateau current and a decreasing amplitude of current oscillation as a function of bias in the plateau range. In the current plateau, the frequency of oscillation is found to increase, reach a maximum, and then decrease sharply to zero as a function of bias. It also predicts large signal amplitude and the presence of higher harmonics just after the current peak. All these findings are in agreement with numerical simulations[5,6] and experiments.[3,4] Furthermore, in the absence of oscillation, our model predicts a decreasing current in the plateau. This can happen if there is a weak quantizing field in the emitter brought about, for example, by a large barrier width.[3]

Finally, we point out that a conventional RTS, under appropriate device parameters, has the potential for operating as an all solid-state THz source. Our analysis dictates that the device must be operated just after the resonant current peak to maximize the output power at THz frequencies. This THz source is expected to significantly extend the domain of application of the traditional impact ionization avalanche transit time (IMPATT) diode and Gunn effect microwave solid-state sources.

ACKNOWLEDGMENT

The authors are grateful for partial support from the Army Research Office and the Office of Naval Research.

APPENDIX

Using the matrix expression of Eq. (25) for (M_c), the eigenvalues for Eq. (33) are

$$\gamma_{1,2}=\pm i\sqrt{\mathcal{G}}(1-\mathcal{G})^{3/2}\equiv\pm i\omega, \tag{A1}$$

and the corresponding eigenvectors are

$$\mathbf{Y}_0^{1,2}=\begin{pmatrix} 1 \\ -1\pm\dfrac{i\omega}{\mathcal{G}(1-\mathcal{G})}\end{pmatrix}. \tag{A2}$$

The solution can be written in the form

$$\begin{aligned}\begin{pmatrix} p_0 \\ q_0\end{pmatrix}&=\Theta(\tau_1)\boldsymbol{\Psi}(\tau_0)+\text{c.c.}\\ &=\Theta(\tau_1)\left\{\exp(i\omega t_0)\begin{pmatrix} 1 \\ -1+\dfrac{i\omega}{\mathcal{G}(1-\mathcal{G})}\end{pmatrix}\right\}+\text{c.c.},\end{aligned} \tag{A3}$$

where the separation between the slow and fast time scales is explicitly written. Note that Eq. (35) determines the nature of the dependence of the solution on the slow time scale τ_1 by virtue of the presence of the operator $\boldsymbol{\mathcal{L}}_1$.

Next we obtain the solution for $\binom{p_1}{q_1}$ by solving Eq. (34). The right-hand side is now known since it is only a function of $\binom{p_0}{q_0}$. The solution can be obtained separately for terms involving fast and slow time scales, where the fast time scale occurs only in exponential terms. The right-hand side of Eq. (34) can be written as

$$\begin{aligned}\begin{pmatrix} N_2^p(p_0,q_0) \\ N_2^q(p_0,q_0)\end{pmatrix}&=\begin{pmatrix} \mathcal{A}^p \\ \mathcal{A}^q\end{pmatrix}|\Theta(\tau_1)|^2\\ &\quad+\left\{\begin{pmatrix} \mathcal{B}^p \\ \mathcal{B}^q\end{pmatrix}\Theta(\tau_1)^2\exp 2i\omega\tau_0+\text{c.c.}\right\},\end{aligned} \tag{A4}$$

where

$$\begin{pmatrix} \mathcal{A}^p \\ \mathcal{A}^q\end{pmatrix}=(1-\mathcal{G})^2\begin{pmatrix} 1/2-2\mathcal{G} \\ 3/2\end{pmatrix}, \tag{A5}$$

$$\begin{pmatrix} \mathcal{B}^p \\ \mathcal{B}^q\end{pmatrix}=(1-\mathcal{G})^2\begin{pmatrix} -((1+2\mathcal{G})/4)+\dfrac{i\omega(2-\mathcal{G})}{2\mathcal{G}(1-\mathcal{G})} \\ (5/4-\mathcal{G}/2)-\dfrac{i\omega(2-\mathcal{G})}{2\mathcal{G}(1-\mathcal{G})}\end{pmatrix}. \tag{A6}$$

If we write the solution for $\binom{p_1}{q_1}$ as

$$\begin{pmatrix} p_1 \\ q_1 \end{pmatrix} = \begin{pmatrix} \alpha^p \\ \alpha^q \end{pmatrix} |\Theta(\tau_1)|^2 + \left\{ \begin{pmatrix} \beta^p \\ \beta^q \end{pmatrix} \Theta(\tau_1)^2 \exp 2i\omega\tau_0 + \text{c.c.} \right\}; \tag{A7}$$

then we have to solve the following equations for the coefficients:

$$\begin{pmatrix} \alpha^p \\ \alpha^q \end{pmatrix} = -(M_c)^{-1} \begin{pmatrix} A^p \\ A^q \end{pmatrix}, \tag{A8}$$

$$\begin{pmatrix} \beta^p \\ \beta^q \end{pmatrix} = \begin{pmatrix} 2i\omega - \mathcal{G}(1-\mathcal{G}) & -\mathcal{G}(1-\mathcal{G}) \\ (1-\mathcal{G}) & 2i\omega + \mathcal{G}(1-\mathcal{G}) \end{pmatrix}^{-1} \begin{pmatrix} B^p \\ B^q \end{pmatrix}. \tag{A9}$$

Substituting the expressions given by Eqs. (A5) and (A6) into Eqs. (A8) and (A9), we obtained the expressions

$$\begin{pmatrix} \alpha^p \\ \alpha^q \end{pmatrix} = \frac{(1-\mathcal{G})}{\mathcal{G}} \begin{pmatrix} 2\mathcal{G} \\ -\frac{1}{2} \end{pmatrix}, \tag{A10}$$

$$\begin{pmatrix} \beta^p \\ \beta^q \end{pmatrix} = \frac{1}{6\mathcal{G}(1-\mathcal{G})} \begin{pmatrix} \{4(1-\mathcal{G})^3 + i\omega(1+2\mathcal{G})\} \\ \left\{ \left(\frac{15}{2}\mathcal{G} - \frac{9}{2} - 3\mathcal{G}^2 \right)(1-\mathcal{G}) - i\omega\left(8 - 3\mathcal{G} - 2\mathcal{G}^{-1}\right) \right\} \end{pmatrix}. \tag{A11}$$

and $\binom{p_1}{q_1}$ is thus determined by Eq. (A7).

At this stage of the calculation, we can expect that all the $\binom{p_j}{q_j}$'s contain, as factors, various powers of $\Theta(\tau_1)$, as well as powers of its absolute value, and their combinations. Therefore, in order to find out if a limit cycle exists, it is important to examine the τ_1 dependence of $\Theta(\tau_1)$, and thereby determine if a well-defined finite limit exists for $\Theta(\tau_1)$ as $t \to \infty$. Moreover, before we can calculate $\binom{p_2}{q_2}$, we need to know the τ_1 derivative of $\Theta(\tau_1)$ in Eq. (35). This information can be obtained by imposing the ''solvability'' condition. This condition makes use of the property of the solution to the adjoint of $\mathcal{L}_0$, denoted as $\mathcal{L}_0^\dagger$. Let $\mathcal{L}_0^\dagger \mathfrak{R} = 0$, then $\langle \mathfrak{R}, \mathcal{L}_0 \binom{p_2}{q_2} \rangle = \langle \mathcal{L}_0^\dagger \mathfrak{R}, \binom{p_2}{q_2} \rangle = 0$, where the scalar product is defined by $\langle \boldsymbol{\nu}, \boldsymbol{\mu} \rangle \equiv \lim_{T\to\infty} T^{-1} \int_0^T \boldsymbol{\nu}^* \cdot \boldsymbol{\mu}\, d\tau$. Therefore, from Eq. (35) we must have

$$\left\langle \mathfrak{R}, -\Delta_2 \mathcal{L}_1 \begin{pmatrix} p_0 \\ q_0 \end{pmatrix} + \begin{pmatrix} N_3^p(p_0, q_0, p_1, q_1) \\ N_3^q(p_0, q_0, p_1, q_1) \end{pmatrix} \right\rangle = 0. \tag{A12}$$

We refer the readers to Morse and Feshback,[20] in showing that the eigensolutions of $\mathcal{L}_0$ and $\mathcal{L}_0^\dagger$ from biorthogonal set of eigenvectors, where $\mathcal{L}_0^\dagger$ here is given by

$$\mathcal{L}_0^\dagger = \left(-\frac{\partial}{\partial \tau_0} - \mathcal{G}(1-\mathcal{G}) \begin{pmatrix} 1 & -\mathcal{G}^{-1} \\ 1 & -1 \end{pmatrix} \right). \tag{A13}$$

For example, the eigensolutions to $\mathcal{L}_0^\dagger \mathfrak{R} = 0$ with eigenvalues $\mu_1 = i\omega$ and $\mu_2 = -i\omega$ are given by $\mathfrak{R}^1 = \exp(-i\omega\tau_0) \binom{\mathfrak{R}_1^p}{\mathfrak{R}_1^q}$ and $\mathfrak{R}^2 = \mathfrak{R}^{1*}$, respectively, where

$$\begin{pmatrix} \mathfrak{R}_1^p \\ \mathfrak{R}_1^q \end{pmatrix} = \begin{pmatrix} 1 \\ \mathcal{G}\left(1 - \dfrac{i\omega}{\mathcal{G}(1-\mathcal{G})}\right) \end{pmatrix}. \tag{A14}$$

The eigensolution $\mathfrak{R}^1$ is orthogonal to the eigenvector of $\mathcal{L}_0$ for the same eigenvalue, i.e., $\langle \mathfrak{R}^1 | \exp(i\omega\tau_0) Y_0^1 \rangle = 0$, since $\mu_1^* = -i\omega = \gamma_2$ of $\mathcal{L}_0$, Eq. (36). Thus, we are led to the relation

$$\langle \mathfrak{R}^2 | \boldsymbol{\Psi}(\tau_0) \rangle = 2(1-\mathcal{G}) + \frac{2i\omega}{(1-\mathcal{G})}, \tag{A15}$$

where only $\mathbf{Y}_0^1$ in $\Psi(\tau_0)$ of Eq. (A3) contributes by virtue of the biorthogonality, i.e., the complex conjugate part of $\Psi(\tau_0)$ also does not contribute in Eq. (A15). Thus, with $\mathfrak{R}$ chosen to be equal to $\mathfrak{R}^2$, the scalar product in Eq. (A12) can be evaluated, and defines the differential equation for $\Theta(\tau_1)$. We obtain

$$\frac{\partial}{\partial \tau_1} \Theta(\tau_1) = \frac{\langle \mathfrak{R}, (M_1) \boldsymbol{\Psi}(\tau_0) \rangle}{\langle \mathfrak{R}, \boldsymbol{\Psi}(\tau_0) \rangle} \Theta(\tau_1) + \frac{\Delta_2^{-1}}{\langle \mathfrak{R}, \boldsymbol{\Psi}(\tau_0) \rangle} \left\langle \mathfrak{R}, \begin{pmatrix} N_3^p(p_0, q_0, p_1, q_1) \\ N_3^q(p_0, q_0, p_1, q_1) \end{pmatrix} \right\rangle. \tag{A16}$$

Note that in Eq. (A16) nonzero τ_0 integration comes only from $p_0 q_1$, $p_1 q_0$, $q_1 q_0$, $p_0 p_1$, p_0^3, and $p_0 q_0^2$ terms in

$$\begin{pmatrix} N_3^p(p_0, q_0, p_1, q_1) \\ N_3^q(p_0, q_0, p_1, q_1) \end{pmatrix}$$

[Eq. (32)], while the rest, including the complex conjugate part, do not contribute as $T \to \infty$ by virtue of the appearance of a δ function of the frequency sum after the τ_0 integration. Therefore, by taking into account only the contributing terms, we can write Eq. (A16) in a simpler form

$$\frac{\partial}{\partial\tau_1}\Theta(\tau_1)=\eta\Theta(\tau_1)+\frac{\sigma}{\Delta_2}\Theta(\tau_1)|\Theta(\tau_1)|^2, \tag{A17}$$

where

$$\eta=\frac{\langle\mathfrak{R},(M_1)\Psi(\tau_0)\rangle}{\langle\mathfrak{R},\Psi(\tau_0)\rangle} \tag{A18}$$

and

$$\sigma\Theta(\tau_1)|\Theta(\tau_1)|^2=\frac{(1-\mathcal{G})^2}{\langle\mathfrak{R},\Psi(\tau_0)\rangle}\left\langle\mathfrak{R},\begin{pmatrix}Q_3(p_0,q_0,p_1,q_1)+2(1-\mathcal{G})p_0p_1-(1-\mathcal{G})^2p_0^3\\-Q_3(p_0,q_0,p_1,q_1)\end{pmatrix}\right\rangle, \tag{A19}$$

with

$$Q_3(p_0,q_0,p_1,q_1)=p_0q_1+p_1q_0+(\mathcal{G}q_1q_0)/2+(1-\mathcal{G})p_0q_0^2/4. \tag{A20}$$

We thus obtain

$$\eta=-\mathcal{G}(1-\mathcal{G})^{-2}+i\omega(1-\mathcal{G})^{-3} \tag{A21}$$

and

$$\sigma\Theta(\tau_1)|\Theta(\tau_1)|^2=\lim_{T\to\infty}\frac{1}{T}\int_0^T\exp(-i\omega\tau_0)d\tau_0(1-\mathcal{G})\times\begin{Bmatrix}[(1-\mathcal{G})^2+i\omega]Q_3\\+2(1-\mathcal{G})^2p_0p_1-(1-\mathcal{G})^3p_0^3\end{Bmatrix}\times\left(2(1-\mathcal{G})+\frac{2i\omega}{(1-\mathcal{G})}\right)^{-1}. \tag{A22}$$

Upon performing the τ_0 integration in Eq. (A22), we obtain

$$\sigma\Theta(\tau_1)|\Theta(\tau_1)|^2=\frac{1}{4}\begin{Bmatrix}[(1-\mathcal{G})^2+i\omega]\hat{Q}_3+2(1-\mathcal{G})^2\hat{p}_0\hat{p}_1\\-(1-\mathcal{G})^3\hat{p}_0^3\end{Bmatrix}\times\left(2(1-\mathcal{G})-\frac{2i\omega}{(1-\mathcal{G})}\right), \tag{A23}$$

where

$$\hat{Q}_3=A_pC_q+A_p^*B_q+A_qC_p+A_q^*B_p+\frac{\mathcal{G}}{2}(A_qC_q+A_q^*B_q)-\frac{(1-\mathcal{G})}{4}(2A_p|A_q|^2+A_p^*A_q^2), \tag{A24}$$

$$\hat{p}_0^3=3A_p|A_p|^2,\quad \hat{p}_0\hat{p}_1\equiv A_pC_p+A_p^*B_p, \tag{A25}$$

$$A_p=\Theta(\tau_1),\quad A_q=\Theta(\tau_1)\{-1+i\omega/[\mathcal{G}(1-\mathcal{G})]\},$$

$$C_p=2(1-\mathcal{G})|\Theta(\tau_1)|^2,\quad C_q=-[(1-\mathcal{G})/2\mathcal{G}]|\Theta(\tau_1)|^2,$$

$$B_p=[4(1-\mathcal{G})^3+i\omega(1+2\mathcal{G})]\Theta(\tau_1)^2/[6\mathcal{G}(1-\mathcal{G})],$$

$$B_q=[(15\mathcal{G}/2-9/2-3\mathcal{G}^2)(1-\mathcal{G})-i\omega(8-3\mathcal{G}-2\mathcal{G}^{-1})]\Theta(\tau_1)^2/[6\mathcal{G}(1-\mathcal{G})]. \tag{A26}$$

Carrying out the operation in Eq. (A23), using Eqs. (A24)–(A26), we obtained

$$\sigma=-\frac{(1-\mathcal{G})^3}{16\mathcal{G}}\{4+19\mathcal{G}-8\mathcal{G}^2\}-\frac{i\omega(1-\mathcal{G})}{48\mathcal{G}^2}\{24\mathcal{G}^3+17\mathcal{G}^2-31\mathcal{G}+8\}\}. \tag{A27}$$

We note that for $0.0<\mathcal{G}<1.0$, Re $\eta<0.0$ and Re $\sigma<0$.

We solve for the absolute value and phase of $\Theta(\tau_1)$ by writing this in polar form and equating the real and imaginary parts on both sides of Eq. (A17). With $\Theta(\tau_1)\equiv|\Theta(\tau_1)|\exp i\phi(\tau_1)$, we obtained exactly solvable equations:

$$\frac{\partial}{\partial\tau_1}|\Theta(\tau_1)|=\mathrm{Re}\,\eta|\Theta(\tau_1)|+\frac{\mathrm{Re}\,\sigma}{\Delta_2}|\Theta(\tau_1)|^3, \tag{A28}$$

$$\frac{\partial}{\partial\tau_1}\phi(\tau_1)=\mathrm{Im}\,\eta+\frac{\mathrm{Im}\,\sigma}{\Delta_2}|\Theta(\tau_1)|^2. \tag{A29}$$

A solution of Eq. (A28) in which $|\Theta(0)|$ can be arbitrarily independent of the limiting value $|\Theta(\infty)|$, the possible limit cycle value, is of the form

$$|\Theta(\tau_1)|=\frac{|\Theta(0)||\Theta(\infty)|\exp[\mathrm{Re}\,\eta\tau_1]}{[|\Theta(\infty)|^2+\{[\exp(2\,\mathrm{Re}\,\eta\tau_1)-1]|\Theta(0)|^2\}]^{1/2}}, \tag{A30}$$

where $|\Theta(\infty)|=[-\mathrm{Re}\,\eta\Delta_2/\mathrm{Re}\,\sigma]^{1/2}$, which is a real value if $\Delta_2<0$ since Re $\eta<0$ and Re $\sigma<0$.

Indeed, in real time $|\Theta(\infty)|=\lim_{t\to\infty}|\Theta(\tau_1)|$ only if Re $\eta\tau_1\to\infty$ as $t\to\infty$; hence only if $\Delta-\Delta_c<0$ or $\Delta<\Delta_c$ in Eq. (A30), i.e., $\Delta_2<0$. This is consistent with the criteria for unstable focus in the linear analysis. Otherwise, $\lim_{t\to\infty}|\Theta(\tau_1)|=0$ if $\Delta-\Delta_c>0$ or $\Delta>\Delta_c$. Thus a well-defined limiting value of $|\Theta(\tau_1)|$ as $t\to\infty$ exists only for $\Delta<\Delta_c$. This is the limit cycle value of $|\Theta(\tau_1)|$. Substituting the now known functional form of $|\Theta(\tau_1)|$ into Eq. (A30), we can also integrate Eq. (A29). The result is

$$\phi(\tau_1)=\text{const}+\text{Im}\,\eta\tau_1+\frac{\text{Im}\,\sigma}{\Delta_2}\frac{|\Theta(\infty)|^2}{2\,\text{Re}\,\eta}\begin{Bmatrix}\ln[|\Theta(\infty)|^2+\{\exp 2\,\text{Re}\,\eta\tau_1-1\}|\Theta(0)|^2]\\ -\ln[|\Theta(\infty)|^2-|\Theta(0)|^2]\end{Bmatrix}. \tag{A31}$$

The limiting values of $\phi(\tau_1)$ are

$$\lim_{t\to\infty}\phi(\tau_1)=\begin{cases}\text{const}+\text{Im}\,\eta\tau_1+0 & \text{if } \Delta>\Delta_c,\\ \text{const}+\text{Im}\,\eta\tau_1+\dfrac{\text{Im}\,\sigma}{\Delta_2}|\Theta(\infty)|^2\tau_1+\ln\left[\dfrac{|\Theta(0)|^2}{|\Theta(\infty)|^2-|\Theta(0)|^2}\right] & \text{if } \Delta<\Delta_c.\end{cases} \tag{A32}$$

Taking the overall constant of integration equal to zero, we end up with the expression for the limit cycle value of $\Theta(\tau_1)$, given as

$$\lim_{t\to\infty}\Theta(\tau_1)=\left[-\frac{\text{Re}\,\eta\Delta_2}{\text{Re}\,\sigma}\right]^{1/2}\exp i\left\{\text{Im}\,\eta(\Delta-\Delta_c)+\frac{\text{Im}\,\sigma}{\Delta_2}|\Theta(\infty)|^2(\Delta-\Delta_c)\right\}\tau. \tag{A33}$$

With the limit cycle value of $\Theta(\tau_1)$ known, $\binom{p_0}{q_0}$ and $\binom{p_1}{q_1}$ at the limit cycle are completely determined. This yields the solution in the form given by Eq. (16) to second-order accuracy.

[1] F. A. Buot and K. L. Jensen, Int. J. Comput. Math. Elec. Electron. Eng., COMPEL **10**, 241 (1991).

[2] D. W. Woolard, F. A. Buot, D. L. Rhodes, X. L. Lu, R. A. Lux, and B. S. Perlman, J. Appl. Phys. **79**, 1515 (1996).

[3] V. J. Goldman, D. C. Tsui, and J. E. Cunningham, Phys. Rev. Lett. **58**, 1256 (1987).

[4] T. C. L. G. Sollner, Phys. Rev. Lett. **59**, 1622 (1987).

[5] K. L. Jensen and F. A. Buot, Phys. Rev. Lett. **66**, 1078 (1991).

[6] B. A. Biegel and J. D. Plummer, Phys. Rev. B **54**, 8070 (1996).

[7] P. Zhao, H. L. Cui, D. L. Woolard, K. L. Jensen, and F. A. Buot (unpublished).

[8] F. W. Sheard and G. A. Toombs, Appl. Phys. Lett. **52**, 1228 (1988).

[9] A. N. Korotkov, D. V. Averin, and K. K. Likharev, Physica B **165**, 927 (1990).

[10] F. W. Sheard and G. A. Toombs, Semicond. Sci. Technol. **7**, B460 (1992).

[11] F. A. Buot and A. K. Rajagopal, Phys. Rev. **48**, 17 217 (1993). The argument of this paper concerning alignment should be corrected; that is, in the absence of 2D states in the emitter, the peak or resonant current correspond to the alignment of the quantum-well energy level with the bottom of the conduction band of the emitter. In view of the present findings, oscillation occurs during passage across the boundary between allowed and forbidden 2D states, not across the bottom of the conduction band of the emitter.

[12] F. A. Buot and A. K. Rajagopal, J. Appl. Phys. **76**, 5552 (1994).

[13] F. A. Buot and A. K. Rajagopal, Mater. Sci. Eng., B **35**, 303 (1995).

[14] F. A. Buot, J. Phys. D **30**, 3016 (1997).

[15] Y. Abe, Semicond. Sci. Technol. **7**, B498 (1992).

[16] G. Jona-Lasino, C. Presilla, and J. Sjostrand, Ann. Phys. (N.Y.) **240**, 1 (1995).

[17] C. Presilla and J. Sjostrand, Phys. Rev. B **55**, 9310 (1997); J. Math. Phys. **37**, 4816 (1996).

[18] F. A. Buot and C. M. Krowne, J. Appl. Phys. **86**, 5215 (1999); see also J. Appl. Phys. (to be publiished).

[19] E. Kececioglu and M. C. Yalabik, Phys. Rev. B **59**, 2111 (1999).

[20] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953), pp. 884–886.