Noise characteristics of sequential tunneling through double-barrier junctions

L. Y. Chen* and C. S. Ting
Department of Physics, University of Houston, Houston, Texas 77204
and Texas Center for Superconductivity, University of Houston, Houston, Texas 77204
(Received 17 July 1991; revised manuscript received 27 April 1992)

The shot noise and I-V characteristics of multilevel resonant-tunneling systems are investigated in a stochastic description including strong time correlation derived from the Pauli exclusion of tunneling electrons. In the first resonance regime, one tunneling electron occupying a quantum-well quasibound state blocks one channel for the follow-up electrons. This correlation suppresses the shot noise. In the second (and higher) resonance regime, the tunneling electrons in the quantum well, when experiencing strong inelastic scatterings, fall immediately to occupy the low-lying off-resonance subband states and thus do not block the tunneling channels of others. There the noise power density may approach the full shot-noise level and the tunneling current can be enhanced.

The transport characteristics of double-barrier junctions have received intense experimental and theoretical investigations.[1–4] In a recent experiment, the unique feature of shot-noise suppression has been found.[5] In the coherent tunneling limit,[1,3] when electrons are free from any scattering while traversing the junction, all the tunneling channels are independent of each other. So the Pauli exclusion among consecutive tunneling electrons via each same-channel state yields a strong time correlation and suppresses the shot noise.[6] Equivalently, the transport process can be regarded as elastic scattering of electrons with the double-barrier junction and the same suppression factor can be derived from the Landauer-Buttiker approach.[7] In the opposite sequential tunneling limit,[8] tunneling electrons experience immediate inelastic (typically electron-phonon) scatterings when they bounce back and forth in the quantum well (WQ).[8–11] In this paper, we present an analytical stochastic approach to examining the noise and I-V characteristics of multilevel resonant-tunneling systems in the sequential tunneling limit. In the first resonance regime, a tunneling electron occupying a QW quasibound state blocks one channel for the follow-up electrons. This correlation suppresses the shot noise in an identical way as in the coherent tunneling limit.[6] In the second (and higher) resonance regime and subject to certain condition as discussed in the context before Eq. (5), the tunneling electrons in the QW, experiencing inelastic scatterings immediately, fall to occupy the low-lying off-resonance subband states and thus do not block the tunneling channels of others. There the noise power density may approach the full shot-noise level. For simplicity and clarity, we shall work at zero temperature and assume that the inelastic scattering rate $1/\tau_{\text{in}}$ is much larger than the tunneling broadening of resonance levels but much smaller than spacing between the resonance levels so that quality resonance features survive.

Since the traversal time[12] for an electron to tunnel through each single barrier is very small compared with the lifetime of the resonant state in typical double-barrier structures, we can approximately regard the events of traversing either the emitter or the collector barrier as instantaneous. The overall tunneling time, then, is nearly determined by the inverse of the tunneling broadening of the resonance levels. The other relevant times are the inverse of the inelastic scattering rate of electrons in the QW, which is also much shorter than the resonance lifetime in the sequential tunneling case. In this case, the perpendicular transport through a double-barrier junction can be characterized as a stochastic process. In an infinitesimal time period Δt, an electron tunnels through the emitter barrier into a vacant QW state with probability $\gamma_1 \Delta t$. It then spends some finite time there and tunnels through the collector barrier with probability $\gamma_2 \Delta t$. γ_1 and γ_2 are, respectively, the tunneling rates of the emitter and the collector barriers. Since most inelastic scatterings occur in the quantum well (rather than within individual barriers), γ_1 and γ_2 are approximately identical as the ones for the coherent process[6,8] and nearly independent of the inelastic scattering rate $1/\tau_{\text{in}}$.

If we count the numbers of these two kinds of events $n_1(t)$ and $n_2(t)$ which occur in the time interval $[0,t]$, their time derivatives give, respectively, the emitter and the collector barrier currents, $i_1(t)=\partial_t n_1(t)$ and $i_2(t)=\partial_t n_2(t)$. The terminal current, due to the Ramo-Shockley theorem[13]

$$i(t)=[i_1(t)+i_2(t)]/2$$

when the junction is biased by a voltage source. For convenience, we work with two random variables $[N(t),n(t)]$: the total number of events

$$N(t)=n_1(t)+n_2(t)$$

and the number of electrons staying in the QW

$$n(t)=n_1(t)-n_2(t)+n_0$$

with initial value n_0. In this, the measured current

$$I=\langle i(t)\rangle=\lim_{t\to\infty} e\langle N(t)\rangle/2t \ , \tag{1}$$

and the low-frequency noise current power density[14]

$$S_i(\Omega\sim 0)=\lim_{t\to\infty} e^2\langle \delta N(t)^2\rangle/2t \ . \tag{2}$$

The evolution of the random variables $[N(t),n(t)]$ is governed by the following stochastic equation:

$$\begin{bmatrix} N(t+\Delta t) \\ n(t+\Delta t)\end{bmatrix}=\left\{\begin{matrix}\begin{bmatrix} N(t)+1 \\ n(t)+1\end{bmatrix} \\ \begin{bmatrix} N(t)+1 \\ n(t)-1\end{bmatrix} \\ \begin{bmatrix} N(t) \\ n(t)\end{bmatrix}\end{matrix}\right\} \text{ with probability } \left\{\begin{matrix} p\Delta t \\ q\Delta t \\ 1-(p+q)\Delta t\end{matrix}\right\} . \tag{3}$$

When the system is biased on resonance to the first QW level, one two-dimensional (2D) subband of QW states is below the emitter Fermi level. The number of participating channels N_V is determined by matching the 2D QW states to the 3D emitter states within the Fermi sphere,[8] which is obviously dependent upon the bias voltage V. Although strong inelastic scattering (in the sequential case) makes all the electrons in the QW fall immediately to occupy only the low-energy portion of the N_V channel states, each electron in the well blocks one channel for the follow-up tunneling electrons. Therefore, the probability coefficients p and q in the first resonance region are

$$p=[N_V-n(t)]\gamma_1, \quad q=n(t)\gamma_2 \ , \tag{4}$$

in which the Pauli exclusion effect is represented by the factor $[N_V-n(t)]$. When the system is biased on resonance to the second (or higher) QW level, two (or more) 2D subbands of QW states are below the emitter Fermi level and the lower levels are off resonance. Then only N'_V on-resonance channels are responsible for the emitter current, which are via the on-resonance 2D subband QW states matching those within the emitter Fermi sphere. Electrons tunneling through the emitter barrier into the QW are subject to immediate inelastic scattering so that they fall to occupy the low-energy states in the lower subbands. When the number of the low-lying off-resonance states is large enough, the N'_V on-resonance channel states will be virtually empty and always available for tunneling electrons to come in. Consider, for example, a three-dimensional system whose first subband states $[\varepsilon_1(V)+\varepsilon_{q\parallel}]$ ($q_\parallel$ is the parallel lateral momentum which is conserved in tunneling through one barrier) are off resonance and below the emitter band bottom, i.e., $\varepsilon_1(V)<0$. The second subband level $[\varepsilon_2(V)+\varepsilon_{q\parallel}]$ is on resonance, $0<\varepsilon_2(V)<\mu_e$ (emitter Fermi energy). The abovementioned condition is $\gamma_2[\varepsilon_2(V)-\varepsilon_1(V)] \gg \gamma'_1[\mu_{er}-\varepsilon_2(V)]$ which obviously involves the bias voltage V. Therefore, the QW electrons occupying only the off-resonance first subband states, do not block the participating channels, and the Pauli exclusion effect will not show up in this situation. Here and in what follows, γ'_1 and γ'_2 denote, respectively, the emitter and collector barrier tunneling rates at the second (or higher) resonance level. Then

$$p=N'_V\gamma'_1, \quad q=n(t)\gamma_2 \ , \tag{5}$$

so that larger shot noise is expected.

In the first on-resonance region, the master equation for the distribution function $P(N,n,t)$ can be derived from Eqs. (3) and (4) as

$$\begin{aligned} P(N,n,t+\Delta t)=&P(N,n,t)\{1-\Delta t\gamma_1[N_V-n(t)]-\Delta t\gamma_2 n(t)\} \\ &+P(N-1,n-1,t)\Delta t\gamma_1[N_V-(n-1)]+P(N-1,n+1,t)\Delta t\gamma_2(n+1) \ , \end{aligned} \tag{6}$$

or, equivalently,

$$\begin{aligned} \partial_t P(N,n,t)=&-[\gamma_1(N_V-n)+\gamma_2 n]P(N,n,t) \\ &+\gamma_1(N_V-n+1)P(N-1,n-1,t) \\ &+\gamma_2(n+1)P(N-1,n+1,t) \ . \end{aligned} \tag{7}$$

The initial distribution can be physically chosen as

$$P(N,n,t=0)=\delta_N P_0(n) \ , \tag{8}$$

where $\delta_{N=0}=1$ and $\delta_{N\neq 0}=0$. $P_0(n)$ is the distribution for the initial number n_0 of electrons in the QW and will be chosen as $\sum_N P(N,n,t\to\infty)$.

In the second (or higher) resonance region, Eqs. (3) and (5) yield a three-variable differential-difference equation for the distribution as

$$\begin{aligned}\partial_t P(N,n,t) = &-(\gamma_1' N_V' + \gamma_2 n) P(N,n,t) \\ &+ \gamma_1' N_V' P(N-1,n-1,t) \\ &+ \gamma_2 (n+1) P(N-1,n+1,t) ,\end{aligned} \tag{9}$$

subject to the same initial condition in Eq. (8). Time t is continuous while N and n are discrete.

To extract useful information from Eqs. (7) and (9), we introduce the following reduced distribution functions:

$$P(n,t) \equiv \sum_N P(N,n,t) , \tag{10}$$

$$Q(n,t) \equiv \sum_N N P(N,n,t) , \tag{11}$$

$$S(n,t) \equiv \sum_N N^2 P(N,n,t) , \tag{12}$$

and their generating functions

$$P(z,t) \equiv \sum_n z^n P(n,t) , \tag{13}$$

$$Q(z,t) \equiv \sum_n z^n Q(n,t) , \tag{14}$$

$$S(z,t) \equiv \sum_n z^n S(n,t) . \tag{15}$$

At $z=1$, these three generating functions give all the physical information we need to calculate the tunneling current and the low-frequency noise power density, i.e.,

$$P(z=1,t) = \sum_{N,n} P(N,n,t) = 1 , \tag{16}$$

$$Q(z=1,t) = \sum_{N,n} N P(N,n,t) = \langle N(t) \rangle , \tag{17}$$

$$S(z=1,t) = \sum_{N,n} N^2 P(N,n,t) = \langle N^2(t) \rangle , \tag{18}$$

in which the first corresponds to the unitarity of probability. The second and the third give, respectively, the physical mean and fluctuation.

In the first on-resonance region, the generating functions satisfy the following equations, which can be derived from Eq. (7) with Eq. (8):

$$[\partial_t + N_V \gamma_1 (1-z)] P(z,t) + (\gamma_1 z + \gamma_2)(z-1) \partial_z P(z,t) = 0 , \tag{19}$$

$$[\partial_t + N_V \gamma_1 (1-z)] Q(z,t) + (\gamma_1 z + \gamma_2)(z-1) \partial_z Q(z,t) = [\gamma_1 N_V z + (\gamma_2 - \gamma_1 z^2) \partial_z] P(z,t) , \tag{20}$$

$$[\partial_t + N_V \gamma_1 (1-z)] S(z,t) + (\gamma_1 z + \gamma_2)(z-1) \partial_z S(z,t) = [\gamma_1 N_V z + (\gamma_2 - \gamma_1 z^2) \partial_z] P(z,t) + 2 [\gamma_1 N_V z + (\gamma_2 - \gamma_1 z^2) \partial_z] Q(z,t) . \tag{21}$$

The Laplace transformation with respect to the t variable helps solve the above three equations to produce the physical mean and fluctuation,

$$\langle N(t) \rangle = \frac{2\gamma_1 \gamma_2 N_V}{\gamma_1 + \gamma_2} t + O(t^0) , \tag{22}$$

$$\langle N^2(t) \rangle = \langle N(t) \rangle^2 + \langle N(t) \rangle \left[1 + \frac{(\gamma_1 - \gamma_2)^2}{(\gamma_1 + \gamma_2)^2} \right] + O(t^0) . \tag{23}$$

In the above-mentioned derivation, Eqs. (16)–(18) have been used. Therefore, the tunneling current [Eq. (1)]

$$I = e \frac{\gamma_1 \gamma_2 N_V}{\gamma_1 + \gamma_2} \tag{24}$$

and the noise power density [Eq. (2)]

$$S_i = 2eI \left[1 - \frac{2\gamma_1 \gamma_2}{(\gamma_1 + \gamma_2)^2} \right] , \tag{25}$$

which represents a suppressed shot-noise process. The suppression factor depends mainly on the symmetry of the two barriers, which is an implicit function of the bias voltage V via the two tunneling rates. When a system of proper parameters is biased at such a point that $\gamma_1 = \gamma_2$, the shot noise is suppressed most. The noise power density is only at the half-full level. In this situation, the average time for a tunneling electron to spend in the QW $\tau_0 = 2/(\gamma_1 + \gamma_2)$ reaches its maximum over the average transport time $\tau = (\gamma_1 + \gamma_2)/\gamma_1 \gamma_2$. So the Pauli-exclusion correlation has the strongest effect. When the system is very asymmetrical, either $\gamma_1 \gg \gamma_2$ or $\gamma_1 \ll \gamma_2$, the ratio τ_0/τ is very small. Then the transport is similar to that through a single-barrier junction and nearly full shot noise is produced. This suppression feature is identical to that of the coherent tunneling[6] and has already been discovered experimentally by Li *et al.*[5]

In the second (or higher) on-resonance region, we can derive, from Eq. (9) together with initial condition Eq. (8), the following equations for the generating functions:

$$[\partial_t + N_V' \gamma_1' (1-z)] P(z,t) + \gamma_2 (z-1) \partial_z P(z,t) = 0 , \tag{26}$$

$$\begin{aligned}[\partial_t + N_V' \gamma_1' (1-z)] Q(z,t) &+ \gamma_2 (z-1) \partial_z Q(z,t) \\ &= [\gamma_1' N_V' z + \gamma_2 \partial_z](z,t) ,\end{aligned} \tag{27}$$

$$\begin{aligned}[\partial_t + N_V' \gamma_1' (1-z)] S(z,t) &+ \gamma_2 (z-1) \partial_z S(z,t) \\ &= [\gamma_1' N_V' z + \gamma_2 \partial_z] P(z,t) + 2 [\gamma_1' N_V' z + \gamma_2 \partial_z] Q(z,t) .\end{aligned} \tag{28}$$

With the help of a Laplace transformation with respect to t variable and of Eqs. (16)–(18), the above three equations can be solved to produce the physical mean and fluctuation

$$\langle N(t)\rangle = 2\gamma_1' N_V' t = O(t^0) , \tag{29}$$

$$\langle N^2(t)\rangle = \langle N(t)\rangle^2 + \langle N(t)\rangle + O(t^0) . \tag{30}$$

In this way, we arrive at the tunneling current [Eq. (1)]

$$I = e\gamma_1' N_V' \tag{31}$$

and noise power density [Eq. (2)]

$$S_i = 2eI \tag{32}$$

which stands for full shot noise. Comparing Eq. (31) with tunneling current

$$I_{\rm coh} = N_V' e\gamma_1'\gamma_2'/(\gamma_1'+\gamma_2')$$

in the absence of inelastic scattering shows that inelastic scattering tends to enhance tunneling current.

In summary, an analytical stochastic approach is developed to investigate the noise and I-V characteristics of multilevel resonant-tunneling structures. In the first resonance regime, sequential tunneling and coherent tunneling have idential shot noise suppression behavior. In the second and higher resonant regions, inelastic scattering tends to enhance tunneling current and has a rather strong effect on noise behavior. The coherent-tunneling process always leads to the Puali-exclusion-derived suppression, while the sequential process yields full shot noise when

$$\gamma_2[\varepsilon_2(V)-\varepsilon_1(V)] \gg \gamma_1'[\mu_2-\varepsilon_2(V)] .$$

In some systems under certain bias voltage,

$$\gamma_2[\varepsilon_2(V)-\varepsilon_1(V)] \le \gamma_1'[\mu_e-\varepsilon_2(V)] ,$$

the lower off-resonance QW states, even when they are fully occupied, cannot drain all of the electrons from the upper on-resonance QW states. Then the on-resonance channels can be partly blocked and the shot noise becomes suppressed but the suppression level is lower than that in the coherent limit. The noise power density lies in between the

$$2eI[1-2\gamma_1'\gamma_2'/(\gamma_1'+\gamma_2')^2]$$

and the full $2eI$ levels. In the extreme case where $\gamma_2=0$, i.e., when electrons in the lower off-resonance QW states cannot tunnel out into the collector, all these states will be fully occupied as soon as a steady transport state is established and cannot drain any electrons from the upper on-resonance QW states. In this limit, the Pauli-blockade effect is identical as in the first on-resonance region [Eq. (4)] and the shot-noise density is again given by Eq. (25).

Finally, varying the thickness and/or height of the emitter and collector barriers and thus tuning the tunneling broadening of the resonance levels, one can go gradually from the coherent to the sequential limit, i.e., from $\gamma_{1,2} \gg 1/\tau_{\rm in}$ to $1/\tau_{\rm in} \gg \gamma_{1,2}$. Therefore, to well understand experimental measurements of the shot noise, not only theoretical studies of the two limits, but also a more general one to deal with the situation in between them, are necessary.

This work has been supported by Grant No. MDA 972-88-G-0002 from the U.S. Defense Advanced Research Projects Agency and the State of Texas.

---

*Present address: Physics Department, Brown University, Providence, RI 02912.

[1] L. L. Chang, L. Esaki, and R. Tsu, Appl. Phys. Lett. **24**, 593 (1974).

[2] T. C. L. G. Sollner, W. D. Goodhue, P. E. Tannenwald, C. D. Parker, and D. D. Peck, Appl. Phys. Lett. **43**, 588 (1983).

[3] B. Ricco and M. Ya. Azbel, Phys. Rev. B **29**, 1970 (1984).

[4] For a recent review, see F. Capasso, F. Beltram, S. Sen, and A. Y. Cho, in *Physics of Quantum Electron Devices,* edited by F. Capasso (Springer-Verlag, Heidelberg, 1990).

[5] Y. P. Li, A. Zaslavsky, D. C. Tsui, M. Santos, and M. Shayegan, Phys. Rev. B **41**, 8388 (1990).

[6] L. Y. Chen and C. S. Ting, Phys. Rev. B **43**, 4534 (1991).

[7] M. Büttiker, Phys. Rev. B **45**, 3807 (1992); G. B. Lesovik, Pis'ma Zh. Eksp. Teor. Fiz. **49**, 594 (1989) [JETP Lett. **49**, 683 (1989)]; M. Büttiker, Phys. Rev. Lett. **65**, 2901 (1990); R. Landauer and T. Martin, Physica B **175**, 161 (1991).

[8] S. Luryi, Appl. Phys. Lett. **47**, 490 (1985).

[9] M. Büttiker, IBM J. Res. Dev. **32**, 63 (1988).

[10] V. J. Goldman, D. C. Tsui, and J. E. Cunningham, Solid-State Electron **31**, 731 (1988).

[11] T. K. Woodward, S. D. Chemla, I. Bar-Joseph, H. U. Baranger, D. L. Sivco, and A. Y. Cho, Phys. Rev. B **44**, 1353 (1991).

[12] M. Büttiker and R. Landauer, Phys. Rev. Lett. **49**, 1739 (1982).

[13] S. Ramo, Proc. IRE **27**, 584 (1939); W. Shockey, J. Appl. Phys. **9**, 639 (1938).

[14] F. N. H. Robinson, *Noise and Fluctuations in Electronic Devices and Circuits* (Clarendon, Oxford, 1974).