Resonant tunneling diodes for switching applications

S. K. Diamond, E. Özbay, M. J. W. Rodwell, and D. M. Bloom
Edward L. Ginzton Laboratory, Stanford University, Stanford, California 94305-4085

Y. C. Pao and J. S. Harris
Department of Electrical Engineering, Stanford University, Stanford, California 94305-4055

(Received 8 August 1988; accepted for publication 25 October 1988)

Rise times for simple pulse-forming circuits are presented. Switching times for present best devices are in the range of 5–15 ps. An equivalent circuit model for resonant tunneling diodes inclusive of space-charge effects and transit time effects in the depletion region is presented. From these models it is shown that switching times are limited by the device RC time constants and are relatively unaffected by the resonant state lifetime or depletion layer transit times. Appropriate figures of merit for switching applications are the device capacitance and peak current density. Less emphasis should be placed on improving the peak-to-valley ratio. Optimally designed devices which maximize the current density should be capable of switching in under 5 ps.

With maximum frequency of oscillation in the millimeter-wave range, resonant tunneling diodes (RTDs) have been extensively studied for their potential as oscillators and mixers.[1] Their application to high-speed switching and pulse forming has also been examined.[2] Key to the optimization of such high-speed circuits is an accurate and physical equivalent circuit model of the RTD. This letter presents a RTD circuit model and discusses its implication for device switching times. We find that the importance of the peak-to-valley ratio (PVR) diminishes above 2:1 and that the current density and device capacitance are more suitable figures of merit.

To analyze a typical RTD pulse generator circuit [Fig. 1(a)], we replace the RTD with its equivalent circuit and the external circuit with its Thevenin equivalent [Fig. 1(b)]. While the circuit output voltage is V_{out}, the transition time of the voltage V_1 internal to the series resistance of the diode is more readily calculated:

$$V_{out} = V_1 \frac{R_l}{R_l + R_s} + V_{in} \frac{R_s}{R_l + R_s}. \tag{1}$$

The resistance R_s of ohmic contacts and substrate[3] determines f_{max} (relevant in oscillator applications) and degrades switching waveforms by coupling a fraction of the input signal V_{in} into the output signal V_{out}.

The switching circuit load line is shown in Fig. 2. Bistable switching requires an average value of negative resistance less than the equivalent source impedance, $R_n < R_l = Z_0/2$. The device is initially biased at the peak current, and as V_{in} increases from V_{in1} to V_{in2}, the diode voltage switches from V_p to V_f, resulting in a positive rising voltage transition. Similarly, a negative transition is formed when the source voltage is lowered and the device is reset to the initial state.

The transient behavior of V_1 is given by $C(V_1)(dV_1/dt) = I_l(V_1) - I_d(V_1)$, where $I_l(V_1) = (V_s - V_1)/R_l'$ is the current supplied by the source, $R_l' = R_l + R_s$, and $I_d(V_1)$ is the current through the well region. The 10–90% rise time is

$$T_{rise} = \int_{V_p + 0.1(V_f - V_p)}^{V_f - 0.1(V_f - V_p)} \frac{C(v)}{I_l(v) - I_d(v)}\, dv. \tag{2}$$

Ignoring the variation of capacitance with voltage and taking maximum and minimum values for $I_l(v)$ and $I_d(v)$, $I_l(v) < I_p$, $I_d(v) > I_v$, a lower bound on the rise time is

$$T_{rise} > \frac{0.8(V_f - V_p)C}{I_p - I_v} = \Gamma \frac{0.8(V_f - V_p)}{1 - 1/\mathrm{PVR}}, \tag{3}$$

where $\Gamma = C/I_p$ (dimensionally Γ has the units of ps/V) and PVR = peak-to-valley ratio $= I_p/I_v$. Note that if the PVR increases from 2:1 to ∞:1, the switching time only decreases by a factor of 2. PVR may be useful for evaluation of epitaxial growth quality, but is of minor significance to switching circuits. Since the rise time is proportional to the ratio Γ of capacitance per unit area to current density, Γ is a more appropriate figure of merit.

Equation (3) demonstrates the primary parameters for switching applications, but significantly underestimates the rise time. For a more accurate estimate of the switching time, the $I_d(v)$ in Eq. (2) is approximated by two linear segments (Fig. 2) of slope $(1/R_n)$ in the negative resistance region and slope $(1/R_d)$ in the region past resonance. The rise time can then be calculated by applying Eq. (2) to the linear segments separately:

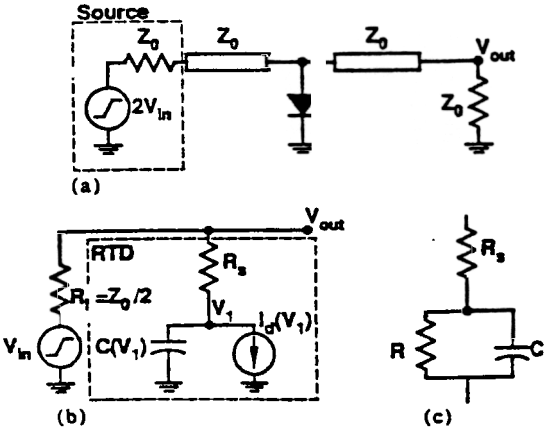

FIG. 1. High-speed RTD pulse-forming circuit. The RTD is shunted to ground across a 50 Ω transmission line. In (b) the matched transmission lines and source have been replaced by their Thevenin equivalent, and the equivalent circuit for the RTD is used. For small-signal applications the RTD equivalent circuit in (c) is used.

 0003-6951/89/020153-03$01.00 © 1989 American Institute of Physics

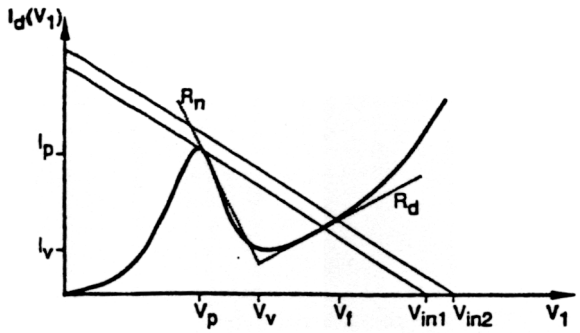

FIG. 2. As the supply voltage increases, the RTD voltage switches from V_p to V_f. The dotted lines represent the approximations to the I-V curve in the piecewise linear model.

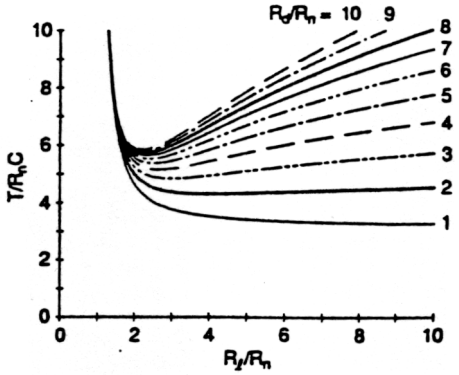

FIG. 3. Normalized rise time is plotted as a function of different ratios of load resistance to device negative resistance for various shapes of current voltage curves. Minimum obtainable rise time is typically 4–5 R_nC.

$$\frac{T_{rise}}{|R_n|C} = \left(\frac{x}{x-1}\right)\ln\left(\frac{10(x+y)}{x(x-y)}\right) + \left(\frac{xy}{x-y}\right)\ln\left(\frac{10(x+y)y}{x(x-y)}\right), \tag{4}$$

where $y \equiv R_d/|R_n|$ and $x \equiv R_l'/|R_n|$. The product $|R_n|C$ limits the device switching speed. Figure 3 plots the normalized rise time ($T_{rise}/|R_n|C$) as a function of R_l' for several $R_d/|R_n|$. As R_l' approaches the $|R_n|$ ($x \to 1$), the rise time increases dramatically because of a decrease in the current available for charging up the RTD capacitance. As the load impedance increases ($x \to \infty$), ($V_f - V_p$) increases, increasing the rise time. Table I gives the rise times, calculated from Eq. (4) with R_l' optimized for minimum rise time, for several RTDs discussed in the literature. The capacitance was estimated from published growth parameters at the bias voltage corresponding to the peak current.

For device optimization RTD circuit models are developed. Conceptually, the RTD is split up into three regions: ohmic contacts and substrate, quantum well region, and depletion region. The ohmic contacts and substrate are modeled by a resistance $R_s = \rho_s/A$, where A is the active area.

In the well region there is conduction current due to tunneling through the barriers and displacement current from the parallel capacitance through the structure. The small-signal impedance for the well region is thus $Z_w = R_w(V_{well}) + 1/j\omega C_w$, where $C_w = \epsilon_s A/L_{wb}$ (L_{wb} is the combined thickness of the well and barriers). $R_w(V_{well})$ can be obtained by Taylor series expansion of static I-V expressions as listed in Ref. 3.

The impedance Z_d of the RTD depletion region was calculated by Brown *et al.*[3] using a technique previously applied to barrier injection diodes[4]:

$$Z_d = \frac{1}{j\omega C_d}\left(1 - \gamma\frac{(1-e^{-j\theta})}{j\theta}\right), \tag{5}$$

where $C_d = \epsilon_s A/L_d$ (L_d is the thickness of the depletion layer), $\theta = \omega L_d/v_{sat}$ is the transit angle where v_{sat} is the saturation velocity, and γ is the complex ratio of the conduction current density to the total current (displacement and conduction) which enters the depletion region. Equation (5) includes both space-charge and transit time effects. From the expression for the well impedance, $\gamma = 1/(1 + j\omega R_w C_w)$.

For frequencies ω less than the depletion layer transit time L_d/v_{sat} (i.e., $\omega < v_{sat}/L_d$), θ is less than 1 and Eq. (5) can be expanded accurately to second order in θ. Summing the quantum well region impedance with the depletion region impedance, we find a small-signal equivalent circuit as in Fig. 1(c), where

$$R = \left(\frac{L_d + L_{wb}}{L_{wb}}\right)(1+K)R_w, \tag{6}$$

$$C = \frac{\epsilon_s}{L_{wb} + L_d}\left(\frac{1}{1+K}\right), \tag{7}$$

and

$$K = \frac{L_d^2 L_{wb}}{2\epsilon_s v_{sat}(L_d + L_{wb})R_w}. \tag{8}$$

The equivalent circuit is valid for $\omega < v_{sat}/L_d$; hence switching times are correctly predicted if the depletion layer transit

TABLE I. Estimated switching time with piecewise linear model.

Research group	PVR	I_p (kA/cm²)	avg. R_n ($\mu\Omega$ cm²)	C (nF/cm²)	$\frac{R_d}{\lvert R_n\rvert} \equiv y$	T_{rise} (ps)
Lincoln Labs[a]	3.5	40	16	160	2.6	12
Texas Instruments[b]	1.6	80	4.2	340	1.0	4.6
Fujitsu[c]	14	23	13	340	1.5	13

Reference 1.
M. A. Reed, J. W. Lee, and H. L. Tsai, Appl. Phys. Lett. 49, 158 (1986).
T. Inata, S. Muto, Y. Nakata, S. Sasa, T. Fujii, and S. Hiyamizu, Jpn. J. Appl. Phys. 26, L1332 (1987)

time L_d/v_{sat} is less than the calculated rise time T_{rise}. Because R_w varies with V_{well}, the small-signal parameter R varies with V_1, where $V_1 = V_{well} + V_{depl}$ is the sum of the quantum well and depletion region voltages. The large-signal slope parameters R_n and R_d then correspond to fits to $R(V_1)$ in the regions $V_p < V_1 < V_v$ and $V_v < V_1 < V_f$, respectively. Note that the parameters $|R_n|C$ and R_d/R_n, which determine the pulse rise time, are independent of L_d to second order in θ.

By varying L_d the device impedance can be altered to match a source impedance without affecting switching speed. This is most effectively accomplished with an undoped buffer layer placed between the well region and a heavily doped N^+ anode. In this manner, L_d will be controlled in growth, L_d will be independent of voltage, and the device series resistance will be minimized.

To increase current densities, thinner barriers can be used to increase the transmission width at resonance. AlAs barriers four atomic layers thick will have a transmission width of 15 meV, roughly matching the energy distribution of the electron supply.

As the current density of a RTD with thin buffer and depletion layers is increased to $\sim 10^5$ A/cm^2, the negative resistance R_n will approach the series resistance of Au/Ge ohmic contacts. Switching transition times are minimized (Fig. 3) with $R_l \simeq 2|R_n|$. Hence, as $|R_n|$ decreases, $R_s/(R_l + R_s)$ increases, and switching waveforms are degraded by direct coupling [Eq. (1)] of V_{in} to V_{out} through R_s. In the negative resistance region, R_w is negative and thus there exists a L_d for which $|R_n|$ is maximized [Eqs. (6) and (7)]. Growing a $\sim$700–1200 Å undoped buffer layer increases $|R_n|$ (and thus increases R_l for minimum switching time) while keeping R_s constant. The buffer will thus decrease the coupling of the input waveform into the output waveform.

Three time constants for RTDs are discussed in the literature. Based on uncertainty relations, the intrinsic time constant of the electron in the well appears to be on the order of 100's of fs. For depletion lengths less than 1000 Å, the depletion layer transit times are also subpicosecond. As our analysis has shown, device switching speed is not being limited by quantum transport phenomenon, but is instead dominated by the more classical effects of current densities and capacitances. For typical best devices published to date, the R_wC_w time constant limits switching transition times to around 5–15 ps. By increasing doping levels in the cathode, using thinner barriers, and optimizing L_d through the use of an undoped buffer layer, transition times below 5 ps should be attainable.

This work was supported by Office of Naval Research contract N00014-86-K-0530. S. Diamond acknowledges an IBM fellowship. The authors wish to thank T. C. L. G. Sollner and E. Wolak for helpful discussion.

[1]E. R. Brown, T. C. L. G. Sollner, W. D. Goodhue, and C. D. Parker, Appl. Phys. Lett. 50, 83 (1987).
[2]H. C. Liu and D. D. Coon, Appl. Phys. Lett. 50, 1246 (1987).
[3]E. R. Brown, W. D. Goodhue, and T. C. L. G. Sollner, J. Appl. Phys. 64, 1519 (1988).
[4]S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed. (Wiley, New York, 1981), Sec. 10.7.2.