Nature of "Superluminal" Barrier Tunneling

Herbert G. Winful
Department of Electrical Engineering and Computer Science, University of Michigan, 1301 Beal Avenue, Ann Arbor, Michigan 48109-2122
(Received 27 August 2002; published 17 January 2003)

We show that the distortionless tunneling of electromagnetic pulses through a barrier is a quasistatic process in which the slowly varying envelope of the incident pulse modulates the amplitude of a standing wave. For pulses longer than the barrier width, the barrier acts as a lumped element with respect to the pulse envelope. The envelopes of the transmitted and reflected fields can adiabatically follow the incident pulse with only a small delay that originates from energy storage. The theory presented here provides a physical explanation of the tunneling process and resolves the mystery of apparent superluminality.

DOI: 10.1103/PhysRevLett.90.023901 PACS numbers: 42.25.Bs, 03.65.Ta, 42.70.Qs

It has been more than 70 years since MacColl suggested that there is no appreciable delay in the transmission of a tunneling wave packet through a potential barrier [1]. Yet, to this day, the dynamics of the tunneling process remains shrouded in mystery and mired in controversy as the idea of instantaneous propagation violates all our cherished notions of causality [2]. Experimental observations of apparent superluminal tunneling are usually presented with a disclaimer that no violation of causality has occurred and that a mere reshaping of the wave packet is sufficient to explain the observed phenomena [2–5]. Notwithstanding these explanations, there remains a deep unease regarding the lack of a truly physical description of the tunneling process that explains all the paradoxical effects. This is especially true in those experiments where no reshaping of the transmitted wave packet is apparent to the naked eye.

In this paper, we make the following key point: True tunneling without distortion is a *quasistatic* process in which the output intensity adiabatically follows the input intensity with a small delay due to energy storage in the barrier. Quasistatic conditions obtain whenever the pulse envelope varies slowly compared to the transit time of light across the barrier. Thus, the barrier acts as a *lumped circuit* with respect to the pulse envelope [6]. Indeed, in all the experiments in which tunneling without distortion has been observed, the pulse lengths have been several times larger than the barrier width. Such hitherto mystifying phenomena as the "Hartman effect" [7], in which the tunneling time becomes independent of length for an opaque barrier, are easily explained in the context of quasistatics and energy storage [8]. While we present results for electromagnetic tunneling, the well-known analogy between classical tunneling and quantum-mechanical tunneling means that the tunneling of particles can be understood in the same manner [2,9].

To be specific, we consider propagation through a one-dimensional periodic dielectric structure (photonic band gap structure) such as has been used in several tunneling experiments [3–5]. The results, however, apply to other tunneling situations such as in constricted waveguides [10] and in frustrated total internal reflection. The refractive index of the periodic structure is of the form

$$n(z) = n_0 + n_1 \cos(2\beta_0 z), \tag{1}$$

where $n_1 \ll n_0$, $\beta_0 = n_0\omega_0/c$ is the Bragg wave number, and ω_0 is the carrier angular frequency that satisfies the Bragg condition for the structure. The periodic structure extends from $z = 0$ to $z = L$ and is embedded in a homogeneous region of refractive index n_0. The complex electric and magnetic fields within the structure are given by

$$E(z,t) = E_+(z,t)e^{i(\beta_0 z-\omega_0 t)} + E_-(z,t)e^{-i(\beta_0 z+\omega_0 t)}, \tag{2}$$

$$H(z,t) = (1/\eta)[E_+(z,t)e^{i(\beta_0 z-\omega_0 t)} - E_-(z,t)e^{-i(\beta_0 z+\omega_0 t)}], \tag{3}$$

where E_+ and E_- are the forward and backward components of the field envelopes, and $\eta = \sqrt{\mu/\varepsilon}$ is the intrinsic impedance of the unperturbed medium. Within the slowly varying envelope approximation, use of Eqs. (1)–(3) in Maxwell's equations leads to the following coupled-mode equations for the forward and backward fields [11]:

$$\frac{\partial E_\pm}{\partial z} \pm \frac{1}{v}\frac{\partial E_\pm}{\partial t} = \pm i\kappa E_\mp. \tag{4}$$

Here $\kappa = \pi n_1/\lambda_0$ describes the coupling between forward and backward waves and $v = c/n_0$ is the group and phase velocity in the unperturbed medium, assumed dispersionless. It should be noted that the carrier waves $\exp[i(\beta_0 z \pm \omega_0 t)]$ are solutions of the homogeneous wave equation for the lossless unperturbed medium and are always propagating modes. The envelope functions modulate the amplitude of these carrier waves and determine how much of the carrier will be transmitted by the barrier.

 0031-9007/03/90(2)/023901(4)$20.00

To elucidate the physics of the tunneling process, we first assume an input sinusoidal envelope modulation of the form $E_+(0, t) = E_0 \cos\Omega t$, with $\Omega \ll \omega_0$, as befits a modulation, and impose the condition $E_-(L, t) = 0$. With these boundary conditions, the sinusoidal steady state solutions (i.e., after transients have died out) of Eq. (4) are

$$E_+(z, t) = \frac{E_0 \cos\phi}{\cosh\gamma L}\left\{\cosh\gamma(z - L)\cos(\Omega t - \phi) + \frac{\Omega}{\gamma v}\sinh\gamma(z - L)\sin(\Omega t - \phi)\right\}, \tag{5a}$$

$$E_-(z, t) = -\left[\frac{\kappa E_0 \cos\phi}{\gamma \cosh\gamma L}\right]\sinh\gamma(z - L)\sin(\Omega t - \phi), \tag{5b}$$

where $\gamma = \sqrt{\kappa^2 - (\Omega/v)^2}$ and the phase shift is

$$\phi = \tan^{-1}[(\Omega/\gamma v)\tanh\gamma L]. \tag{6}$$

Here we have assumed $\Omega < \kappa v$ so that the envelope modulation is below cutoff. This implies that the period of the modulation must satisfy the requirement $T > (2\pi/\kappa L)T_{tr}$, where $T_{tr} = L/v$ is the transit time through the unperturbed medium. For "opaque" barriers with reasonable transmission ($1 < \kappa L \leq 6$), true tunneling thus requires that the envelope period (and, hence, pulse length) exceed the transit time. This is what defines lumped-circuit behavior and the domain of quasistatics [6].

These driven solutions correspond to envelope standing waves oscillating at the same frequency as the input but with phase lags of ϕ and $\phi + \pi/2$. The cosh and sinh functions are normal modes of the finite barrier, and as a result they move up and down in their entirety in response to the input modulation. In doing so, they instantaneously modulate the carrier waves throughout the barrier. The phase of each mode is determined by the coupling and by the boundary conditions and is established during the transient stage of duration $t \sim 2\pi/\kappa v$, a time related to the Q of the barrier. Since the envelope does not propagate in the sinusoidal steady state, the phase shifts seen here are not due to propagation but are a result of the reactive nature of the coupling and of any impedance mismatch at the boundaries of the barrier. As a result of these phase shifts, the peaks of the forward and backward envelope modulations are delayed in time with respect to the driving modulation. This delay is the group delay or envelope delay (also referred to as the phase time) and is given by $d\phi/d\Omega$. For the transmitted $[E_F(L, t)]$ and reflected $[E_B(0, t)]$ fields, we find the same group delay:

$$\tau_d = \frac{d\phi}{d\Omega}\bigg|_{\Omega=0} = \frac{\tanh\kappa L}{\kappa v}. \tag{7}$$

We emphasize that this is not a propagation delay since the envelope does not propagate when the modulation frequency is below cutoff. The behavior is similar to that of an RC (resistance-capacitance) high-pass filter with a time constant τ_d. The barrier is acting as a lumped capacitor with coupling to the outside world providing an effective dissipation.

In the limit $L \to \infty$, Eqs. (5) reduce to

$$E_+(z, t) = E_0 e^{-\gamma z}\cos\Omega t, \tag{8a}$$

$$E_-(z, t) = E_0 e^{-\gamma z}\sin(\Omega t - \phi). \tag{8b}$$

The normal modes of the infinite structure are pure exponential functions which do not depend on length since they never see the exit. Inside the barrier, the forward envelope simply oscillates up and down in phase with the input modulation. Since there is no dissipation, the driving force E_0 does no work and, hence, in the sinusoidal steady state, the forward envelope is everywhere in step with the input modulation, including at the distant reaches of this very long barrier. The backward wave, on the other hand, is driven not directly by the input modulation but by scattering from the forward wave and is related to it through spatial and temporal derivatives. It therefore oscillates with a phase shift relative to the input. This is a coupling induced phase shift that depends on the spatial distribution of the forward wave.

The phase shifts and delays seen here can be related to energy storage within the barrier. Figure 1 shows the electric field pattern over a few optical cycles for a steady state envelope in a finite length barrier. Interference between forward and backward carrier waves sets up a standing wave within the barrier. This standing wave is a complete standing wave in the front part of the barrier where the backward wave amplitude is nearly equal to that of the forward wave. As with standing waves on a transmission line, the spatial patterns of the electric and magnetic fields are in phase quadrature. Energy also sloshes back and forth in time between electric and

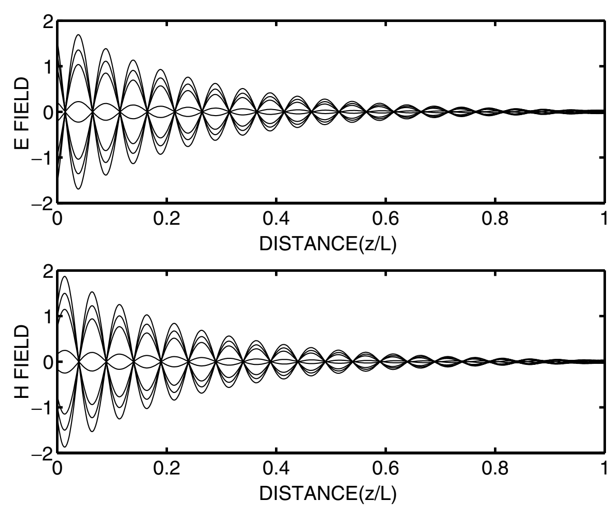

FIG. 1. The electric and magnetic fields in the barrier over a few cycles of the optical carrier wave. Here $\beta_0 L = 20\pi$ and $\kappa L = 4$. The envelopes are in steady state.

magnetic forms but does not propagate. Near the exit of the barrier, however, the backward carrier wave, since it starts with zero amplitude at $z = L$, is much weaker than the forward wave, and therefore the latter can exist as a traveling wave in this region. In a finite length barrier, the coupling to the unperturbed region at the exit provides an effective dissipation that makes it possible for real power to be transmitted. This alters the ratio of stored reactive power to real power transport and thereby introduces a phase shift in the transmitted forward wave.

Poynting's theorem relates the time rate of change of the total instantaneous stored energy U to the incident (P_i), reflected (P_r), and transmitted (P_t) power:

$$\frac{dU}{dt} = P_i(t) - P_r(t) - P_t(t). \tag{9}$$

By using Eqs. (5) in (9), we find that the peaks of the total stored energy lag the peaks of the incident intensity by $\tau_d/2$. The delay τ_d between an incident peak and a reflected or transmitted peak relates to the time it takes for energy to be stored and then released. For a very long barrier, an integration over the length L and area A yields the time-averaged stored energy

$$\langle U \rangle = \tfrac{1}{2}\varepsilon E_0^2 A(1 - e^{-2\gamma L})/\gamma, \tag{10}$$

which, for $L \rightarrow \infty$, reduces to $\frac{1}{2}\varepsilon E_0^2 A L_{\mathrm{eff}}$, where $L_{\mathrm{eff}} = 1/\gamma$ is the effective length of the barrier, the distance at which the field drops to $1/e$ of its initial value. The Hartman effect is precisely a result of the fact that the stored energy (which is proportional to the group delay) becomes independent of length, residing primarily in the region $z < L_{\mathrm{eff}}$ [8]. Note that for a very slow modulation such that $\Omega \ll \kappa v$, the delay approaches $1/\kappa v$, which is the inverse of the cutoff angular frequency of the barrier.

For narrowband pulses, the tunneling process is essentially a quasisteady state phenomenon in which the field envelope throughout the barrier can follow the slow variations of the input envelope with little phase lag (after the initial transient). In this quasistatic limit, we can obtain approximate solutions to the coupled-mode equations for arbitrary input pulse profiles by expanding the complex amplitudes of the sinusoidal solutions to first order in the frequency parameter $\Omega/\kappa v$ and performing an inverse Fourier transform. The resulting solutions are

$$E_+(z, t) = \frac{\cosh\kappa(z - L)}{\cosh\kappa L} \times \left\{ A(t) - \frac{1}{\kappa v}[\tanh\kappa L + \tanh\kappa(z - L)]A'(t) \right\}, \tag{11a}$$

$$E_-(z, t) = -i\frac{\sinh\kappa(z - L)}{\cosh\kappa L}\left\{ A(t) - \frac{\tanh\kappa L}{\kappa v}A'(t) \right\}. \tag{11b}$$

Here $A(t)$ is the envelope of the incident pulse as measured at $z = 0$ and the primes denote derivatives with respect to time. These quasistatic solutions are in excellent agreement with the numerical solutions of the coupled-mode equations.

To follow a pulse as it tunnels through a barrier, we numerically integrate Eq. (4) along forward and backward characteristics. The input pulse is Gaussian of the form

$$A(t) = E_+(0, t) = \exp[-(t - t_0)^2/2\tau_p^2], \tag{12}$$

with time measured in units of the transit time and with $\tau_p = 3$ chosen to satisfy the narrowband condition. Figure 2(a) shows the incident pulse, the transmitted (tunneled) pulse, and the reference pulse, i.e., a pulse that travels the distance L in a barrier-free region. It is seen that the pulse is transmitted with no distortion and that its peak at the exit is delayed with respect to the peak

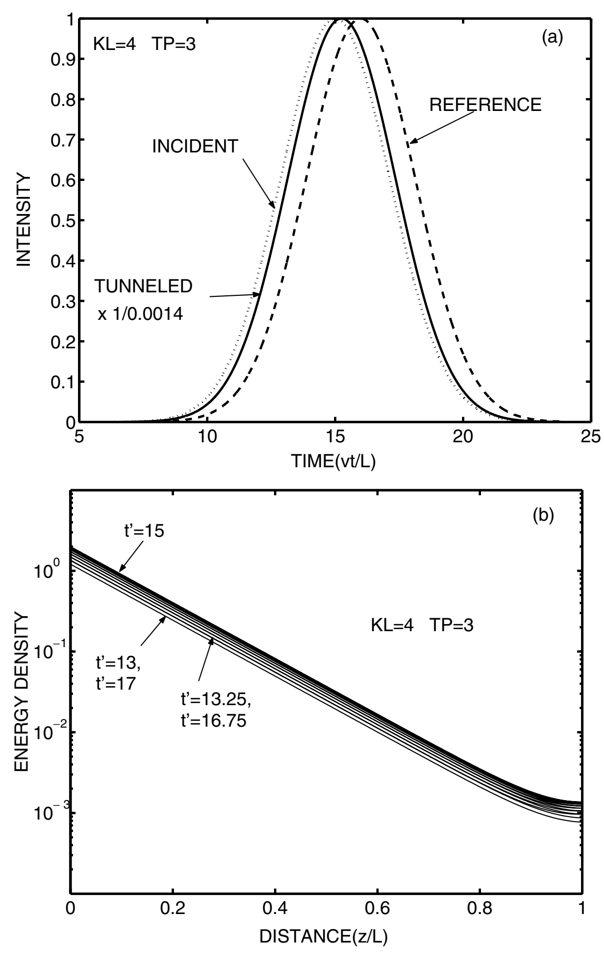

FIG. 2. (a) Incident (dotted), transmitted (solid line), and reference (dashed) pulses for the barrier of strength $\kappa L = 4$. The input narrowband pulse has normalized width $\tau_p = 3$ and a peak at $t_0 = 5\tau_p$. The tunneled pulse is undistorted, has a peak intensity of 1.4×10^{-3}, and is delayed by $t_d' = 0.25$. (b) Snapshots of the spatial distribution of energy density in the tunneling pulse. These are taken at time instants from $t' = 13$ to $t' = 17$ in steps of 0.25 around the peak of the incident pulse. Note that the entire distribution moves up and down in phase.

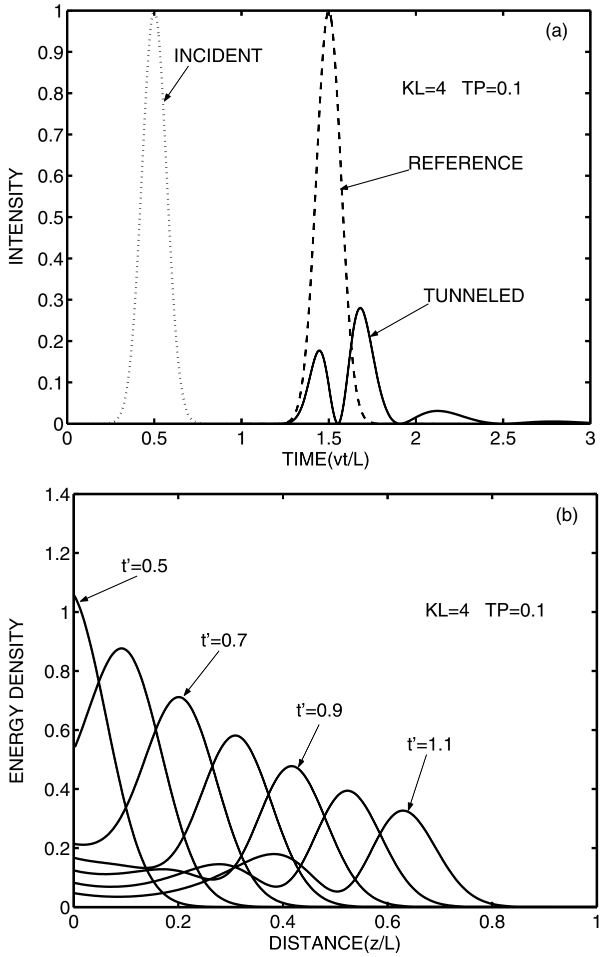

FIG. 3. (a) Incident (dotted), transmitted (solid line), and reference pulse (dashed) for a short pulse of width $\tau_p = 0.1$. (b) Snapshots of spatial distribution of energy density in the "tunneling" pulse of width $\tau_p = 0.1$.

at the input by a time $t_d \approx 0.25$, which is exactly the value predicted by the phase time.

The rather prompt appearance of the peak at the output should not be cause for concern. The incident peak *did not propagate* to the output. In fact, under these quasistatic conditions the pulse peak never even enters the barrier since the energy density has no peak in the interior of the interval $(0, L)$. We can check this by taking snapshots of the spatial distribution of the field energy density $u(z, t) \propto [|E_+(z, t)|^2 + |E_-(z, t)|^2]$ in the barrier at instants of time just before the peak, right at the peak, and just after the peak of the incident pulse. Figure 2(b) shows the energy density distributions taken at normalized times $t = 13$–17 in steps of 0.25, based on the numerical solutions of Eq. (4). The intrabarrier energy density decreases monotonically from the input to the output even as the incident pulse goes through a maximum. There is no sense in which the input peak travels to the output and, hence, input and output peaks are not connected by causal propagation as noted by Landauer [9]. The entire envelope moves up and down as a semirigid entity in response to the input modulation. Near the exit, the energy density deviates from its exponential decay as the fields adjust to the loading conditions at that end. Clearly, the duration of the tunneling process is just the length of the input pulse.

The behavior of narrowband pulses seen here is in marked contrast to that of broadband pulses that violate the adiabatic criterion. Figures 3(a) and 3(b) show the propagation of a pulse of normalized width $\tau_p = 0.1$ through the same structure. First, it is noted that the transmitted pulse suffers significant distortion because its spectrum extends over regions of frequency where the magnitude of the transmission function is not uniform. The overall propagation is luminal since it is dominated by frequency components that lie outside the stop band. Second, as shown in Fig. 3(b), snapshots taken at different instants show the peak in energy density actually traveling through the barrier. This pulse, however, is not tunneling but "flying over" the barrier because it has significant spectral content in the filter pass bands. The important point is that for broadband pulses one can track the propagation of a peak through the barrier, whereas the peak of a narrowband pulse does not even enter the barrier.

In conclusion, we have shown that the apparent superluminal tunneling of pulses is a quasistatic phenomenon in which the output envelope adiabatically follows the input. The incident peak does not actually propagate to the exit which means that the notion of a transit time is meaningless. The input field merely modulates the amplitude of a standing wave created through the interference between forward and backward waves. When properly interpreted in this context, no superluminal transport is seen.

This research was supported in part by the NSF under Grants No. PHY01-14336 and No. PHY99-07949.

[1] L. A. MacColl, Phys. Rev. **40**, 621 (1932).

[2] R. Y. Chiao and A. M. Steinberg, in *Progress in Optics*, edited by E. Wolf (Elsevier, Amsterdam, 1997), Vol. XXXVII, p. 345.

[3] A. M. Steinberg, P. G. Kwiat, and R. Y. Chiao, Phys. Rev. Lett. **71**, 708 (1993).

[4] Ch. Spielmann, R. Szipocs, A. Stingl, and F. Krausz, Phys. Rev. Lett. **73**, 2308 (1994).

[5] S. Longhi, M. Marano, P. Laporta, and M. Belmonte, Phys. Rev. E **64**, 055602 (2001).

[6] R. B. Adler, L. J. Chu, and R. M. Fano, *Electromagnetic Energy Transmission and Radiation* (Wiley, New York, 1960).

[7] T. E. Hartman, J. Appl. Phys. **33**, 3427 (1962).

[8] H. G. Winful, Opt. Express **10**, 1491 (2002).

[9] Th. Martin and R. Landauer, Phys. Rev. A **45**, 2611 (1992).

[10] A. Enders and G. Nimtz, J. Phys. I (France) **3**, 1089 (1993).

[11] H. G. Winful, Appl. Phys. Lett. **46**, 527 (1985).