# Measurement of the Single-Photon Tunneling Time

A. M. Steinberg, P. G. Kwiat, and R. Y. Chiao
*Department of Physics, University of California, Berkeley, California 94720*
(Received 5 January 1993)

Using a two-photon interferometer, we have measured the time delay for a photon to tunnel across a barrier consisting of a 1.1-$\mu$m-thick 1D photonic band-gap material. The peak of the photon wave packet appears on the far side of the barrier $1.47 \pm 0.21$ fs *earlier* than it would if it were to travel at the vacuum speed of light $c$. Although the apparent tunneling velocity $(1.7 \pm 0.2)c$ is superluminal, this is not a genuine signal velocity, and Einstein causality is not violated. The measured tunneling time is consistent with the group delay ("phase time"), but not with the semiclassical time.

PACS numbers: 42.50.Wm, 03.65.Bz, 73.40.Gk

Tunneling is one of the most striking consequences of quantum mechanics. The Josephson effect in solid state physics, fusion in nuclear physics, and instantons in high energy physics are all manifestations of this phenomenon. Every quantum mechanics text treats the calculation of the tunneling probability. And yet, the issue of how much *time* it takes a particle to tunnel through a barrier, a problem first addressed in the 1930s, remains controversial to the present day. The question arises because the momentum in the barrier region is imaginary. The first answer, the "phase time," i.e., the group delay as calculated by the method of stationary phase, can in certain limits be paradoxically small, implying barrier traversal at a speed greater than that of light in vacuum [1,2]. It has generally been assumed that such velocities cannot be physical [3], but in the case of tunneling no resolution has been universally accepted. This apparent violation of Einstein causality does not arise from the use of the nonrelativistic Schrödinger equation, since it also arises in solutions of Maxwell's equations, which are fully relativistic. As a result of developments in solid state physics, such as tunneling in heterostructure devices, this issue has acquired a new sense of urgency in the past decade, leading to much conflicting theoretical work [4,5]. In the past few years, several experimental papers presenting more or less indirect measurements of barrier traversal times have appeared. Some seem to agree with the "semiclassical time" of Büttiker and Landauer [6,7], while others [8] seem to agree with the "phase time." One experiment [9] using *classical* microwave fields in waveguides beyond cutoff has recently confirmed that the effective group velocity for evanescent waves may exceed $c$, but prior to the present Letter no direct time measurement had been presented, nor have results at the single-particle level. More experiments are needed in order to clarify the meanings and ranges of validity of the different tunneling times.

We recently proposed [10] an experiment which offers a relatively direct measurement of the time *delay* in tunneling. It employs a two-photon source in which pairs of photons are emitted essentially simultaneously. The advantage of using these "conjugate" particles is that after one particle traverses a tunnel barrier its time of arrival can be compared with that of its twin (which encounters no barrier), thus offering a clear operational definition of the tunneling time. The magnitude of this time is so small as to be inaccessible to electronic measurement, but a two-photon interference effect [11] can be used to study the overlap of the two photons' wave packets when they are brought together at a beam splitter, with subfemtosecond resolution. We have used this effect to confirm that single photons in glass travel at the group velocity [12]. Since this technique relies on coincidence detection, the particle aspect of tunneling can be clearly observed: Each coincidence detection corresponds to a single tunneling event.

In our current apparatus, the tunnel barrier is a multilayer dielectric mirror. Such mirrors are composed of quarter-wave layers of alternating high- and low-index materials, and hence possess a one-dimensional "photonic band gap" [13], i.e., a range of frequencies which correspond to pure imaginary values of the wave vector. They are optical realizations of the Kronig-Penney model of solid state physics, and thus analogous to crystalline solids possessing band gaps, as well as to superlattices. Our dielectric mirror has an $(HL)^5H$ structure, where $H$ represents titanium oxide (with an index of 2.22) and $L$ represents fused silica (with an index of 1.41). Its total thickness $d$ is 1.1 $\mu$m, implying a traversal time of $d/c = 3.6$ fs if a particle were to travel at $c$. Its band gap extends approximately from 600 to 800 nm, and the transmission reaches a minimum of 1% at 692 nm. This function is shown in Fig. 1, along with the group delay, the semiclassical time, and Büttiker's Larmor time. The group delay is the derivative with respect to angular frequency of the phase of the transmission amplitude. The semiclassical time is calculated from the group velocity which would hold inside an infinite periodic medium (i.e., neglecting reflections at the extremities of the barrier). As the wave vector becomes pure imaginary for frequencies within the band gap, so does the semiclassical time; in order to extend it into the band-gap region, we simply drop the factor of $i$, in analogy with the interaction time of Büttiker and Landauer [6]. The Larmor time [14] is a measure of the amount of Larmor rotation a tunneling electron would experience in an infinitesimal magnetic


0031-9007/93/71(5)/708(4)$06.00

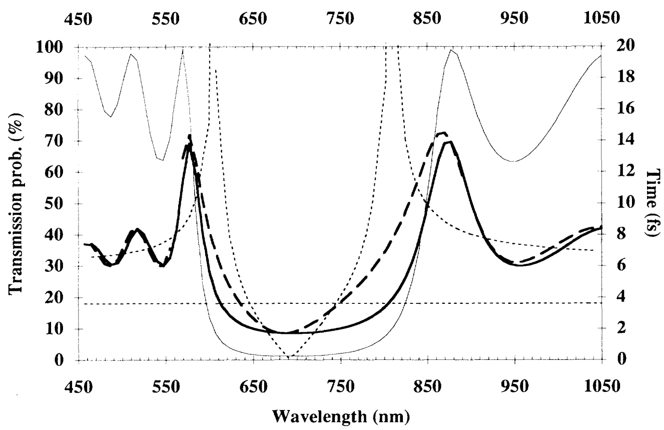


FIG. 1. Theoretical curves, where the light solid curve shows the transmission probability (left axis) of our multilayer coating, as a function of incident wavelength. The heavy solid curve shows the group delay, the heavy dashed curve shows the Büttiker-Larmor traversal time, and the light dotted curve the semiclassical time (right axis). The horizontal dotted line at 3.6 fs represents the "causality limit" $d/c$.

field confined to the barrier region. All three times are seen to dip below $d/c=3.6$ fs, although their detailed behaviors are quite different. Over most of the band gap, the group delay is less than 3.6 fs, and remains relatively constant near 1.7 fs. The semiclassical time, on the other hand, dips below 3.6 fs only over a narrower range of frequencies, and actually reaches zero at the center of the gap. The Larmor time approaches the group delay far from the band gap as well as at its center, but differs from it at intermediate points.

Our apparatus is shown in Fig. 2. A crystal with a $\chi^{(2)}$ nonlinearity (KDP) is pumped by a cw uv laser at 351 nm. Conjugate pairs of photons are emitted simultaneously in the process of spontaneous parametric down-conversion. By the use of 3-mm irises about 75 cm from the crystal, we select out nearly degenerate pairs centered at 702 nm, with an rms bandwidth of approximately 6 nm. One photon of each pair travels through air, while the conjugate photon impinges on our sample. This consists of a 7-mm-thick, 25-mm-diam etalon substrate of fused silica, which is coated over half of one face with the 1.1 $\mu$m coating described above, and uncoated on the other half of that face. The entire opposite face is antireflection coated. This sample is mounted on a precision translation stage, and can be placed in either of two positions separated by 6 mm: In one of these positions, the photon must tunnel through the 1.1 $\mu$m coating in order to be transmitted, while in the other position, it travels through 1.1 $\mu$m of air. In both positions, it traverses the same thickness of substrate. The two conjugate photons are brought back together by means of mirrors, so that they impinge simultaneously on the surface of a 50/50 beam splitter. A coincidence count is recorded when detectors (EG&G single-photon counting modules

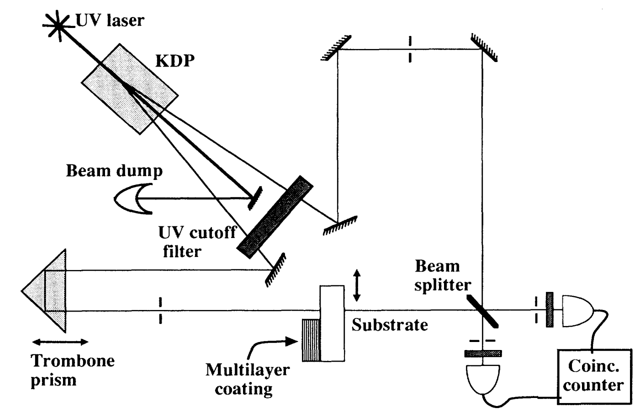


FIG. 2. Apparatus for measuring the single-photon tunneling time.

with 75% quantum efficiencies [15]) placed at the two output ports of the beam splitter register counts within 500 ps of one another.

If the two photons' wave packets are made to overlap in time at the beam splitter, a destructive interference effect leads to a theoretical null in the coincidence detection rate. Thus as the path-length difference is changed by translating a "trombone" prism with a Burleigh Inchworm system (see Fig. 2), the coincidence rate exhibits a dip with an rms width of approximately 20 fs, which is the correlation time of the two photons (determined by their 6 nm bandwidths) [11,12,16]. The rate reaches a minimum when the two wave packets overlap perfectly at the beam splitter. For this reason, if an extra delay is inserted in one arm of this interferometer (i.e., by sliding the 1.1 $\mu$m coating into the beam), the prism will need to be translated in order to compensate for this delay and restore the coincidence minimum. In order to eliminate as far as possible any systematic errors, we conducted each of our data runs by slowly scanning the prism across the dip, while sliding the coating in and out of the beam periodically, so that at each prism position we had directly comparable data with and without the barrier.

In our experiment, we found that inserting the barrier into the beam caused the center of the dip to be shifted to a position in which the prism was located *farther* from the barrier (see Fig. 2). This determines the sign of the effect: The external delay had to be *lengthened*, implying that the mean delay time experienced by the photon inside the barrier was *less than* the delay time for propagating through the same distance in air. We performed twelve 1-h runs, alternating the direction in which we translated the prism in case the direction had a systematic effect on the result (although none was observed). In each run, the data with and without the barrier were fitted by separate Gaussians, and the relative shift between their centers was calculated, with an uncertainty on the order of 0.6 fs (about one-thirtieth of the rms width). As can be seen in Fig. 3, these fits match the data relatively well; a typical run yielded a $\chi^2$ of 86 for 107 points taken with the barrier, and a $\chi^2$ of 258 for 127 points with no barrier. (The latter $\chi^2$ is high because the uncertainties are sufficiently low in the absence of the barrier that laser fluctuations and deviations from a Gaussian shape become noticeable.)

By averaging the results of all these runs, we found that $\Delta t = -1.47 \pm 0.19$ fs, where $\Delta t$ is the transit time for the tunneling photons minus that for the nontunneling photons. We also noted that in some of the runs the "baseline" coincidence rates, taken far from the dip, changed somewhat over time, presumably due to laser pointing instability. Since a linear drift will shift the apparent location of a dip, we reanalyzed the same data after normalizing each of the rates to a straight line between the baseline measurements made before and after each run. The average value found for $\Delta t$ was not significantly different when we used this correction technique: $\Delta t = -1.57 \pm 0.16$ fs. However, it reduced the $\chi^2$ of our twelve runs from 25 to 15, which implies that it largely corrects for errors introduced by these drifts. We therefore used the same technique in analyzing a separate 4-h run, which then yielded $\Delta t = -1.1 \pm 0.3$ fs. The weighted average of these two numbers yields $-1.47 \pm 0.21$ fs, when estimated systematic errors are taken into account as well. This demonstrates that the tunneling delay time is *smaller* by 7 standard deviations than the time it takes to traverse the barrier width at $c$. It is

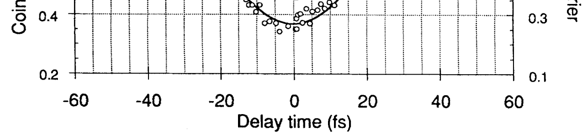


FIG. 3. Coincidence profiles with and without the tunnel barrier, taken by scanning the trombone prism (see Fig. 2) map out the single-photon wave packets. The upper profile (right axis) shows the coincidences with the barrier; this profile is shifted by $1.1 \pm 0.3$ fs to *negative* times relative to the one with no barrier (lower curve, left axis). The wave packet which tunnels through the barrier arrives *earlier* than one which travels the same distance in air. (For comparison, the arrow corresponds to the delay time one would expect from the optical path length of the coating divided by $c$.) Thirteen fits of this type yielded a relative delay of $-1.47 \pm 0.21$ fs, including estimated systematic errors.

within about 2 standard deviations of the group delay and the Larmor time, which both give $\Delta t = -1.9$ fs. It disagrees with the semiclassical prediction, $\Delta t = -3.0$ fs.

The normalized data from the 4-h run are presented in Fig. 3. Each data point with the barrier in the beam has been averaged for 70 s, and each point with no barrier for 7 s. The prism was translated approximately 12 nm every 7 s; the position data are interpolated from the 0.1-$\mu$m-resolution output of our encoder. Without the barrier, the singles rate was about 200000 $s^{-1}$, and the coincidence rate 1800 $s^{-1}$. Insertion of the barrier generally reduced the singles rate by a factor of 3 and the coincidence rate (including accidentals) by a factor of 50. Typical visibilities were 60%, with or without the barrier in place [16]. Since the transmission probability was nearly constant over the bandwidth of our photons, the *form* and *width* of the photon wave packets (and thus of the dip) were not appreciably changed by the insertion of the barrier, despite an overall reduction of the count rate. For this reason, there is physical significance to following the peak of the photon wave packets. Furthermore, unlike in the electronic case, the shift cannot be understood as arising from the higher speed of the preferentially transmitted energy components before they reach the barrier [17]. The transmission is a weak function of energy near the middle of the band gap, and photons of all energies travel at the same velocity, except in a few dispersive optical elements. Even the small dispersive effects of those elements, however, have been shown to cancel out in this type of experiment [12,16].

Systematic errors which could arise from the substrate's $< 1$ arcsec wedge and $\lambda/10$ flatness (both determined interferometrically prior to coating) should be about $\pm 0.15$ fs. The orientation of the substrate was observed to be stable to better than 1 mrad, implying that any extra delay due to angle was limited to less than $\pm 0.01$ fs. As a direct check of these effects, we performed several runs in which we measured times through two different positions on the uncoated half, separated by 6 mm, as before. The shift of $+0.15 \pm 0.09$ fs confirms that there was no strong systematic effect arising either from variations in the thickness of the substrate or from any mechanical effect of the translation itself. We also changed the position of the sample between several of the data runs, and rotated it by 180° about the vertical axis once; no significant differences were observed. We therefore estimate our systematic error to be $\pm 0.15$ fs, which we combine with our statistical uncertainty of $\pm 0.14$ fs to yield an overall uncertainty of $\pm 0.21$ fs.

Ideally, one would like to repeat the measurements at a second wavelength. Since this was not convenient, we "tuned" the barrier by rotating the sample about the vertical axis by 23° and by 45°, thus shifting the center of the band gap. For the $p$-polarized photons we employed, this also reduces the width of the band gap. At 45°, the upper edge shifted to 708 nm, close to our photons' wavelengths. The semiclassical time diverges at the edge, already exceeding $d/c$ by 12 fs at 702 nm (the transmission rises to 10%). As can be seen from Fig. 1, the group delay varies much more slowly, approximately equaling $d/c$ at the band edge, where the Larmor time exceeds $d/c$ by about 4 fs. At 23°, we found a delay time of $\Delta t = -1.0 \pm 0.4$ fs. At 45°, we found $\Delta t = +0.78 \pm 0.15$ fs. These results demonstrate that the delay time is a relatively weak function of angle, in contrast to the semiclassical time. While they seem to agree somewhat better with the group delay than with the Larmor time, the theories need to be extended to two-dimensional problems and more data may be necessary.

While the analogy between quantum mechanical tunneling and evanescent wave propagation in electromagnetism is well known [8,10,18], there is an important difference between classical wave propagation and single-particle tunneling. In classical optics, the existence of group velocities greater than $c$, and even negative ones under certain conditions, is known, and has been observed experimentally [19,20]. This phenomenon is understood as a "pulse reshaping" process, in which a medium preferentially attenuates the later parts of an incident pulse, in such a way that the output peak appears shifted towards earlier times. Einstein causality is not violated in this process; at all times, the output intensity is less in the presence of the medium than it would have been in its absence; this differs from superluminal propagation in an inverted medium [21]. In both effects, however, the output peak arises from the forward *tails* of the input pulse in a strictly causal manner, and no abrupt disturbance in the input pulse would travel faster than $c$ [3].

In the case of a single-particle wave packet in quantum mechanics, there is no meaning to the question of which "part" of a minimum-uncertainty wave packet gives rise to a given detection event. If the peak of such a wave packet is shifted forward in time, this means that the mean delay between the single-particle emission event and the corresponding detection event is smaller than $d/c$, *whenever the particle is transmitted.* The source of this anomaly is that the incident particles are transmitted with low probability. It has been shown by Aharonov and Vaidman [22] that when a "weak measurement" (one with a sufficiently large uncertainty as to leave the state on which the measurement is performed essentially unperturbed) is made on a subensemble defined both by state preparation and by a *postselection* of low probability, mean values can be obtained which would be strictly forbidden for any complete ensemble. Interpretation is discussed further in [23].

Our measurements indicate that the peak of the undistorted (but attenuated) single-photon wave packet appears on the far side of a tunnel barrier *earlier* than it would were it to propagate at $c$. There is, however, no genuine violation of Einstein causality, as explained above. The tunneling time does not appear to be a strong function of the angle of incidence, and the data indicate that in our experiment the group delay (or "phase time") gives a better description of the physically observable *delay* than does the "semiclassical" interaction time.

This work was supported by ONR Grant No. N000-14-90-J-1259. We would like to thank I. H. Deutsch, E. L. Bolda, J. C. Garrison, G. Kurizki, R. Landauer, A. Stern, E. Yablonovitch, A. Jacobson, and B. E. Johnson.

[1] L. A. MacColl, Phys. Rev. **40**, 621 (1932).
[2] E. P. Wigner, Phys. Rev. **98**, 145 (1955).
[3] L. Brillouin, *Wave Propagation and Group Velocity* (Academic, New York, 1960).
[4] E. H. Hauge and J. A. Støvneng, Rev. Mod. Phys. **61**, 917 (1989), and references therein.
[5] H. A. Fertig, Phys. Rev. Lett. **65**, 2321 (1990).
[6] M. Büttiker and R. Landauer, Phys. Rev. Lett. **49**, 1739 (1982).
[7] R. Landauer, Nature (London) **341**, 567 (1989).
[8] A. Ranfagni, D. Mugnai, P. Fabeni, G. P. Pazzi, G. Naletto, and C. Sozzi, Physica (Amsterdam) **175B**, 283 (1991), and references therein.
[9] A. Enders and G. Nimtz, J. Phys. I (France) **3**, 1089 (1993).
[10] R. Y. Chiao, P. G. Kwiat, and A. M. Steinberg, Physica (Amsterdam) **175B**, 257 (1991).
[11] C. K. Hong, Z. Y. Ou, and L. Mandel, Phys. Rev. Lett. **59**, 2044 (1987).
[12] A. M. Steinberg, P. G. Kwiat, and R. Y. Chiao, Phys. Rev. Lett. **68**, 2421 (1992).
[13] E. Yablonovitch and K. M. Leung, Physica (Amsterdam) **175B**, 81 (1991), and references therein.
[14] M. Büttiker, Phys. Rev. B **27**, 6178 (1983).
[15] P. G. Kwiat, A. M. Steinberg, R. Y. Chiao, P. Eberhard, and M. Petroff, Phys. Rev. A **48**, 867 (1993).
[16] A. M. Steinberg, P. G. Kwiat, and R. Y. Chiao, Phys. Rev. A **45**, 6659 (1992); J. Jeffers and S. M. Barnett, Phys. Rev. A **47**, 3291 (1993).
[17] R. Landauer and Th. Martin, Solid State Commun. **84**, 115 (1992); R. S. Dumont and T. L. Marchioro II, Phys. Rev. A **47**, 85 (1993).
[18] Th. Martin and R. Landauer, Phys. Rev. A **45**, 2611 (1992).
[19] C. G. B. Garrett and D. E. McCumber, Phys. Rev. A **1**, 305 (1970).
[20] S. Chu and S. Wong, Phys. Rev. Lett. **49**, 1293 (1982).
[21] R. Y. Chiao, Phys. Rev. A **48**, R34 (1993).
[22] Y. Aharonov and L. Vaidman, Phys. Rev. A **41**, 11 (1990).
[23] A. M. Steinberg, P. G. Kwiat, and R. Y. Chiao, in Proceedings of the XXVIII Rencontres de Moriond, edited by J. Tran Thanh Van *et al.* (to be published).