PHYSICAL REVIEW LETTERS

VOLUME 74 | 27 MARCH 1995 | NUMBER 13

How Much Time Does a Tunneling Particle Spend in the Barrier Region?

Aephraim M. Steinberg*
Department of Physics, University of California at Berkeley, Berkeley, California 94720
(Received 9 June 1994)

The question in the title may be answered by considering the outcome of a "weak measurement" in the sense of Aharonov, Albert, and Vaidman. Various properties of the resulting time are discussed, including its close relation to the Larmor times. It is a universal description of a broad class of measurement interactions, and its physical implications are unambiguous.

PACS numbers: 03.65.Bz, 73.40.Gk

The question posed in the title has remained controversial since the early days of quantum theory [1–5]. One commonly cited reason for this is the nonexistence of a quantum-mechanical time operator. However, it is quite possible to construct an operator Θ_B which measures whether a particle is in the barrier region or not. Such a projection operator is Hermitian, and may correspond to a physical observable. It has eigenvalues 0 and 1, and its expectation value simply measures the integrated probability density over the region of interest—it is this expectation value divided by the incident flux which is referred to as the "dwell time" [6]. Thus the central problem is not the absence of an appropriate Hermitian operator, but rather the absence of well-defined histories (or trajectories) in standard quantum theory. The dwell time measures a property of a wave function with both transmitted and reflected portions, and does not display a unique decomposition into portions corresponding to these individual scattering channels. Some workers consider the expectation value for a particular outgoing channel rather than a particular incident state [7]; this discards information about early times rather than about late times. Other related approaches follow phase space trajectories [7], Bohm trajectories [4], or Feynman paths [8]. No consensus has been reached as to the validity and the relationship of these various approaches. Ideally, transmission and reflection times τ_T and τ_R would, when weighted by the transmissivity and reflectivity $|t|^2$ and $|r|^2$, yield the dwell time τ_d:

$$|t|^2\tau_T + |r|^2\tau_R = \tau_d\,; \tag{1}$$

this relation has served as one of the main criteria in a broad review of tunneling times [3], but has also been criticized [5]. In this paper I present an approach to tunneling times which adheres strictly to the standard formulation of quantum mechanics, defining the "dwell" time by explicitly considering a general von Neumann-style measurement interaction. The mean value of the measurement outcome may then be calculated for transmitted and reflected particles individually, and I show that it automatically satisfies the above equation. This time is in general complex, and is closely related to the Larmor times and to the complex time of Sokolovski and Connor [8]; the consideration of a measuring apparatus, however, makes interpretation of its real and imaginary parts straightforward, as will be demonstrated by examining a generalization of the Larmor time.

Before presenting the new theory, let us recap several existing approaches. Relation (1) is violated by many of these approaches, including perhaps the simplest: the group delay or extrapolated "phase time." This time describes the appearance of a wave packet peak on the far side of the barrier. It has been pointed out that there is no fundamental reason in quantum theory to associate a "delay" time of this sort with the duration of the interaction itself [2]. For thick enough barriers the wave packet may appear to have traveled faster than the vacuum speed of light c; this has recently been verified for several different manifestations of tunneling [9]. Since such effects occur in the limit of very small transmission probability, the overall "center of mass" of the wave function (i.e., the expectation value of x) does not travel faster than light.

The anomalies occur only when a *subensemble* is defined both by preparation of a state incident on the barrier from the left and by postselection of particles emergent on the right. In the "standard" interpretation of quantum mechanics, this postselection comprises a collapse of the wave function, which occurs instantaneously [10], and the apparent superluminal motion of those particles which are transmitted is due entirely to the instantaneity of this uncontrollable collapse event, which is useless for signaling purposes.

Another class of approaches to tunneling relies instead on an external "clock." The Larmor time [6,11], in which a magnetic field confined to the barrier region causes the spin of a tunneling particle to precess, is the prototypical theory of this sort. As Büttiker observed, this time has two different components, the precession of the spin in the plane perpendicular to the applied field and the tendency of the spin to align itself with the field, and so far the choice of one of these components or their vector sum seems to be merely a matter of opinion. Several experiments related to "clocks" have been performed [5].

Meanwhile, Aharonov, Albert, and Vaidman have developed a theory of "weak measurements" precisely to deal with questions about subensembles such as those with which we are concerned [12]. Suppose we wish to measure the expectation value of some operator A (for instance, Θ_B) which acts on the particle under study. Following von Neumann [13], let us postulate a measuring apparatus or "pointer" whose position and conjugate momentum we term Q and P. A measurement results from a time-dependent interaction $H_{\text{int}} = g(t)PA$, since P is the generator of translations for the pointer, the mean position of the pointer after the interaction will have shifted by an amount proportional to the expectation value $\langle A\rangle$. In an ideal measurement, the relative shifts corresponding to different eigenvalues of A are large compared with the initial uncertainty in the pointer's position, and the resulting lack of overlap between the final states leads to the effective decoherence (or irreversible "collapse") between different eigenstates of A. In [12], this approach is modified in that the initial position of the pointer has a large uncertainty, so that the overlap between the pointer states remains close to unity, and hence that the measurement does not constitute a collapse. Seen another way, this means that the pointer momentum P may be very well defined, and therefore need not impart an uncertain "kick" to the particle; the measurement is "weak" in that it disturbs the state of the particle as little as possible between the state preparation and the postselection. One can give up measurement precision in exchange for delicacy. What Aharonov, Albert, and Vaidman calculate is the average shift of the pointer, according to standard quantum theory, for those of the particles prepared in state $|i\rangle$ which are later found to be in state $|f\rangle$. The derivation can be found in [12], but the same result can be loosely understood from the point of view of probability theory. Consider the conditional probability for the particle to be found in eigenstate $|A_i\rangle$ (with eigenvalue a_i) *given* that it is subsequently found to be in state $|f\rangle$, according to standard Bayesian theory, $P(A_i|f) = P(A_i\&f)/P(f)$. The probability $P(f)$ is represented in quantum theory by the expectation value of the projection operator $\text{Proj}(f) \equiv |f\rangle\langle f|$; we replace the logical *and* with multiplication since these projectors are binary valued. (It should, however, be noted that the product of two projectors need not be Hermitian, so these joint "probabilities" are not, in general, real. Detailed discussion of "extended" probabilities can be found in [14].) Thus we rewrite the relation

$$P(A_i|f) = \frac{\langle \text{Proj}(f)\text{Proj}(A_i)\rangle}{\langle \text{Proj}(f)\rangle}. \tag{2}$$

Summing over i, we find the weak (i.e., conditional) expectation value

$$\langle A\rangle_f = \sum_i a_i \frac{\langle \text{Proj}\ (f)\text{Proj}\ (A_i)\rangle}{\langle \text{Proj}\ (f)\rangle} = \frac{\langle |f\rangle\langle f|A\rangle}{\langle |f\rangle\langle f|\rangle}. \tag{3}$$

Evaluating the expectation values for a particle prepared in $|i\rangle$, we find

$$\langle A\rangle_{fi} = \frac{\langle i|\ |f\rangle\langle f|A\ |i\rangle}{\langle i|\ |f\rangle\langle f|\ |i\rangle} = \frac{\langle f|A|i\rangle}{\langle f|i\rangle}. \tag{4}$$

This is the result proved in [12] for the mean shift in the pointer position. The symmetry of this expression is related to the time reversibility of quantum evolution in the absence of collapse. Note that this expression may, in general, be complex. Unlike the case for complex times found in the past for tunneling, however, the physical significance of the real and imaginary parts of a weak value is clear. If the initial state of the pointer is $\exp[-Q^2/4\sigma^2]$, then after the measurement it will, for suitably normalized $g(t)$, be in the state $\exp[-(Q - \langle A\rangle_{fi})^2/4\sigma^2]$. The real part of $\langle A\rangle_{fi}$ corresponds to the mean shift in the pointer position, while the imaginary part constitutes a shift in the pointer *momentum*. This latter effect is a reflection of the backaction of a measurement on the particle. It does have physical significance, but, since it does not correspond to a spatial translation of the pointer, should *not* be thought of as part of the measurement outcome. Furthermore, unlike the spatial translation $\Delta Q = \text{Re}\langle A\rangle_{fi}$, this effect is sensitive to the initial state of the pointer: $\Delta P = \text{Im}\langle A\rangle_{fi}/2\sigma^2$. As σ becomes large, the measurement becomes weak, and the momentum shift of the pointer (like the backaction on the particle) vanishes, while the spatial shift remains constant.

Consider these weak values for a complete orthonormal set of final states $|f_j\rangle$. We take the weighted average

$$\begin{aligned}\overline{\langle A\rangle_{fi}} &\equiv \sum_j |\langle f_j|i\rangle|^2 \langle A\rangle_{f_j i} = \sum_j \langle i|f_j\rangle\langle f_j|i\rangle \frac{\langle f_j|A|i\rangle}{\langle f_j|i\rangle} \\ &= \sum_j \langle i|f_j\rangle\langle f_j|A|i\rangle = \langle i|A|i\rangle = \langle A\rangle;\end{aligned} \tag{5}$$

thus relations in the form of Eq. (1) are satisfied automatically if τ_T and τ_R are defined in terms of weak values. This is because the weak values represent average measurement outcomes for various subensembles, which must clearly reproduce the average for the ensemble as a whole when weighted properly.

Let us examine a particle tunneling through a rectangular barrier extending from $x = -d/2$ to $x = +d/2$. As explained in the first paragraph, the operator A we wish to measure is simply the projector onto this region, $\Theta_B \equiv \Theta(x + d/2) - \Theta(x - d/2)$. Suppose as above that we allow a pointer to interact with the particle so that its position will reflect the expectation value of this operator. We prepare the particle in a stationary state $|i\rangle$ defined such that to the right of the barrier it contains only right-going components; a superposition of several such states with a range of energies would describe a wave packet incident from the left at early times. Inside the barrier, we have $\psi_i(-d/2 < x < d/2) = Be^{-\kappa x} + Ce^{\kappa x}$. At a later time, we may find the particle in the transmitted state $|t\rangle$ (the time-dependent case is treated explicitly in [15], and in [16] for the Larmor clock). In this event, we observe the pointer position, which constitutes a measurement of how much time the transmitted particle spent in the barrier region (compare the definition of dwell time in [6]). Over many trials, the average position is given by Eq. (4), where $|f\rangle$ represents the transmitted state $|t\rangle$; since $|t\rangle$ contains no left-going components to the left of the barrier, it is related to $|i\rangle$ by a parity flip combined with a time reversal: $\psi_t(x) = \psi_i^*(-x)$. Similarly, $\psi_r(x) = \psi_i^*(x) = \psi_t(-x)$. The transition amplitudes $\langle t|i\rangle$ and $\langle r|i\rangle$ are simply t and r. Thus, defining $\tau_f = \langle\Theta_B\rangle_{fi}/J_{\rm in}$, where $J_{\rm in} = \hbar k/m$, we have

$$\tau_T = \frac{m}{\hbar k}\frac{1}{t}\int_{-d/2}^{d/2} dx(Be^{\kappa x} + Ce^{-\kappa x})(Be^{-\kappa x} + Ce^{\kappa x}),$$

$$\tau_R = \frac{m}{\hbar k}\frac{1}{r}\int_{-d/2}^{d/2} dx(Be^{-\kappa x} + Ce^{\kappa x})(Be^{-\kappa x} + Ce^{\kappa x}),$$

$$\tau_d = \frac{m}{\hbar k}\int_{-d/2}^{d/2} dx|Be^{-\kappa x} + Ce^{\kappa x}|^2, \qquad (6)$$

which are trivial to evaluate.

Since $|B/C| = \exp[\kappa d]$, the dominant contributions to (6) are the B^2 terms, which decay exponentially with x for τ_R and τ_d, reflecting the fact that little of the wave penetrates far into the barrier. By contrast, the B^2 term is constant across the entire barrier for τ_T. In a sense, this suggests that the transmitted particles differ from the others in that they sample the entire barrier thickness. However, τ_T is predominantly imaginary, and its *real part* stems predominantly from the two extremes of the barrier, becoming independent of thickness along with the other times. This feature bears an interesting resemblance to the intrabarrier group delay time, which remains nearly constant through most of the barrier (since $B\exp[-\kappa x]$ accumulates no phase and $|B| \gg |C|$) and grows at a rate of approximately $m/\hbar k$ for the *last* exponential decay length ($1/\kappa$).

It is straightforward (but unenlightening) to demonstrate that $\mathrm{Re}[(B^2 + C^2)/t] = \mathrm{Re}[2BC/r] = 2\mathrm{Re}[B^*C]$ and $\mathrm{Re}[2BC/t] = \mathrm{Re}[(B^2 + C^2)/r] = |B|^2 + |C|^2$, which immediately imply the equality of the real parts (that is, the actual measurable shift of the pointer) for all three times, and thus trivially of relation (1). This holds for all values of k and d, including the classical limit as well as the tunneling regime. The equality is *not* a general property of weak values, but rather a consequence of (1) and the symmetry of the barrier. Since the integrands in (6) have different spatial dependences, we see that the conditional dwell times in very small regions will not be the same for the different scattering channels, although when integrated over the entire barrier they are both equal to the "unconditional" dwell time. The dwell time in regions outside the barrier depends even more strikingly on which channel is considered.

The expression for τ_T is the same one found by Sokolovski *et al.* by performing an appropriately weighted average over Feynman paths [8]: its real part is equal to the dwell time, as is the in-plane portion of the Larmor time. Its imaginary part is equal to minus the out-of-plane portion of the Larmor time, and its magnitude is thus equal to Büttiker's proposed Larmor time. Because of the emphasis on real measuring devices, however, the different significance of these portions becomes clear in the present context: the real part is the mean shift in the pointer's position, while the imaginary part is instead a measure of the backaction on the particle due to the measurement interaction itself. In fact, as discussed earlier, this latter effect is sensitive to the details of the measurement apparatus (in particular, to its initial uncertainty in momentum), unlike the position shift.

Although τ_T calculated in this way is a general description of *any* von Neumann-style measurement, these results can be particularly well understood in the context of the Larmor time, which is defined by the interaction between a magnetic field confined to the barrier region and the particle's spin. $H_{\rm int} = -\gamma \mathbf{S}\cdot\mathbf{B} = -\gamma S_z B_0 \Theta_B$ is nothing more than a particular realization of the von Neumann interaction. S_z plays the role of the pointer momentum; as it is the difference between the occupation *numbers* of the spin-up and spin-down states, the conjugate pointer position is the difference between the *phases* of the spin-up and spin-down components [17]. This relative phase determines the in-plane rotation angle, which is why the Larmor precession serves as a measure of the dwell in the barrier region. The present approach shows that *any* interaction which is designed to measure Θ_B will yield τ_T; the Larmor result is thus perfectly general, and the natural extension of standard quantum measurements to experiments which include postselection.

To see that the tendency of the particle's spin to align with the applied field (which Büttiker showed dominated over the in-plane rotation, in the opaque limit) should not be considered an additional term for the traversal time, we must let the measurement become more weak by increasing the initial uncertainty σ in the pointer position, as discussed above after Eq. (4). It might at first seem that it would suffice to consider particles of spin $\gg 1/2$, in order to approach the correspondence principle limit. Büttiker found, however, that the Larmor time was the same for higher-spin particles. The various S_z eigenstates remain split by $\hbar\omega_L \equiv \hbar\gamma B_0$, but if the pointer is prepared in the maximum-S_x eigenstate (the case typically considered, in order to follow rotation in the x-y plane) the width of the initial distribution in S_z is $\sqrt{S/2}$. For small ω_L, the transmission varies linearly with S_z, so the shift it introduces in $\langle S_z \rangle$ scales as the square of the width (near the peak, treat the distribution as quadratic: $[1 - (S_z/\Delta S)^2][1 + \xi S_z] \sim 1 - [(S_z - \xi \Delta S^2/2)/\Delta S]^2$). The alignment angle goes as S_z/S and is thus independent of S, as is the Larmor frequency; the time defined by their ratio is hence unchanged by going to $S > 1/2$. However, it is possible to prepare the particle in a "spin-squeezed state" [18,19], in which ΔS_z may be much smaller than $\sqrt{S/2}$, while $\langle \mathbf{S} \rangle$ still points along $\mathbf{x}$. The uncertainty-principle trade-off is that the angle in the x-y plane can no longer be determined as precisely. The mean rotation angle, however, is determined by the relative phase difference between successive S_z states, and this is not affected by the weighting of the different components. As the Larmor measurement is made weaker by using a highly squeezed initial spin state, the in-plane rotation (due to the real part of τ_T) remains constant, while the out-of-plane rotation (due to the imaginary part of τ_T) vanishes. Thus while the former time scale is indeed a property of the tunneling process, regardless of how it is observed, the latter is merely a measure of the backaction provoked by the measurement, sensitive to the initial state of the measurement apparatus.

It follows from the above argument that $\mathrm{Re}\,\tau_T$ is in essence a derivative of the transmission phase with respect to potential energy, while $\mathrm{Im}\,\tau_T$ is a derivative of the transmission amplitude. It is well known that the phase shifts for reflection and for transmission, given a spatially symmetric barrier, differ only by a constant, leading to equal group delay times for the two scattering channels [16,20]. This is also why $\mathrm{Re}\,\tau_R = \mathrm{Re}\,\tau_T$, and hence [taking (1) into account] why both times are equal to τ_d. Similarly, it is the fact that $|r|^2 + |t|^2 = 1$ which determines the behavior of the derivatives of the amplitudes, ensuring that $|t|^2 \mathrm{Im}\,\tau_T = -|r|^2 \mathrm{Im}\,\tau_R$.

It might be asked whether $\mathrm{Re}\,\tau_T$ alone represents the entire duration of the tunneling interaction, or whether "self-interference delays" prior to the barrier constitute an additional contribution. When the three times discussed in this paper are generalized to treat regions other than the barrier, all three do contain interference terms, but the resulting spatial oscillations continue indefinitely, leading to no overall contribution when integrated over all the space to either side of the barrier. It is interesting to note that, while the dwell time displays such oscillations only to the left of the barrier, the other times oscillate everywhere outside the barrier. In particular, the "reflection dwell time" at point x to the right of the barrier $d\tau_R(x > d/2) \propto \sin(2kx + \arg t)\,dx$. It may seem odd that the reflected particles display a nonzero dwell time on the far side of the barrier, but the physical interpretation is clear. Any interaction which can detect the presence of a particle (including, but not restricted to, an applied magnetic field) has some probability of reflecting the particle. In the Larmor case, the spin-up and spin-down particles will be reflected with opposite phase by a small magnetic field at some point beyond the barrier. Some of these particles are transmitted back to the left side of the barrier, and the phase difference means that on average the spin angle of particles emerging on the left now depends on the strength of the applied field on the far side of the barrier. One may be tempted to argue that the effect comes about due to particles which would not have been reflected in the absence of a measurement, and hence is not truly indicative of the behavior of reflected particles in general; in particular, [16] and [21] conclude that such times are unphysical. If, however, we follow the Copenhagen interpretation, we have no choice but to define our question in terms of conceivable measurements. This leads automatically to τ_T and τ_R as found in this paper, including their odd properties beyond the barrier. While the backaction on the particle may vary with measurement technique, the delays calculated in this fashion do not. As discussed in [15], for example, the momentum transfer due to the Coulomb interaction between a proton and an electron may become repulsive. Whether or not one describes this as a negative interaction time, it is a counterintuitive and physically testable consequence of these conditional expectation values.

This work was supported by ONR Grant N00014-90-J1259. Discussions with R. Chiao, A. Elby, P. Kwiat, and M. Mitchell were helpful.

*Now at NIST, Phys A167, Gaithersburg, MD 20899.
Elecronic address: aephraim@physics.berkeley.edu

[1] E.U. Condon, Rev. Mod. Phys. **3**, 43 (1931); L.A. MacColl, Phys. Rev. **40**, 621 (1932); L. Eisenbud, Ph.D. Thesis, Princeton University, 1948; E.P. Wigner, Phys. Rev. **98**, 145 (1955); T.E. Hartman, J. Appl. Phys. **33**, 3427 (1962); K.W.H. Stevens, Europhys. J. Phys. **1**, 98 (1980).
[2] M. Büttiker and R. Landauer, Phys. Rev. Lett. **49**, 1739 (1982).
[3] S. Collins, D. Lowe, and J.R. Barker, J. Phys. C **20**, 6213 (1987); E.H. Hauge and J.A. Støvneng, Rev. Mod. Phys. **61**, 917 (1989).

[4] C.R. Leavens and G.C. Aers, in *Scanning Tunneling Microscopy III,* edited by R. Wiesendanger and H.-J. Güntherodt (Springer-Verlag, Berlin, 1993).
[5] R. Landauer and Th. Martin, Rev. Mod. Phys. **66**, 217 (1994).
[6] M. Büttiker, Phys. Rev. B **27**, 6178 (1983).
[7] B.A. van Tiggelen, A. Tip, and A. Lagendijk, J. Phys. A **26**, 1731 (1993); M. Sassoli de Bianchi, Helv. Phys. Acta **66**, 361 (1993); J.G. Muga, S. Brouard, and R. Sala, Phys. Lett. A **167**, 24 (1992).
[8] D. Sokolovski and J.N.L. Connor, Phys. Rev. A **47**, 4677 (1993); H.A. Fertig, Phys. Rev. Lett. **65**, 2321 (1990).
[9] A.M. Steinberg and R.Y. Chiao, Phys. Rev. A (to be published), and references therein.
[10] W. Heisenberg, *The Physical Principles of the Quantum Theory* (Dover, New York, 1930), p. 39.
[11] A.I. Baz', Sov. J. Nucl. Phys. **5**, 161 (1967); V.F. Rybachenko, Sov. J. Nucl. Phys. **5**, 635 (1967).
[12] Y. Aharonov, D.Z. Albert, and L. Vaidman, Phys. Rev. Lett. **60**, 1351 (1988); Y. Aharonov and L. Vaidman, Phys. Rev. A **41**, 11 (1990).
[13] J. von Neumann, in *Quantum Theory and Measurement,* edited by J.A. Wheeler and W.H. Zurek (Princeton University Press, Princeton, NJ, 1983).
[14] W. Mückenheim *et al.,* Phys. Rep. **133**, 337 (1986); S. Youssef, Mod. Phys. Lett. A **6**, 225 (1991); M.O. Scully *et al.,* Phys. Rev. A **49**, 1562 (1994).
[15] A.M. Steinberg (to be published).
[16] J.P. Falck and E.H. Hauge, Phys. Rev. B **38**, 3287 (1988).
[17] P.A.M. Dirac, Proc. R. Soc. London A **114**, 243 (1927).
[18] D.J. Wineland, J.J. Bollinger, W.M. Itano, F.L. Moore, and D.J. Heinzen, Phys. Rev. A **46**, R6797 (1992).
[19] M. Kitagawa and M. Ueda, Phys. Rev. A **47**, 5138 (1993).
[20] A.M. Steinberg and R.Y. Chiao, Phys. Rev. A **49**, 3283 (1994).
[21] C.R. Leavens and W.R. McKinnon, Phys. Lett. A **194**, 12 (1994).