New Journal of Physics
The open access journal at the forefront of physics
Deutsche Physikalische Gesellschaft DPG IOP Institute of Physics

PAPER • OPEN ACCESS

The relativistic tunneling flight time may be superluminal, but it does not imply superluminal signaling

To cite this article: Randall S Dumont *et al* 2020 *New J. Phys.* **22** 093060

New Journal of Physics
The open access journal at the forefront of physics
Deutsche Physikalische Gesellschaft DPG
IOP Institute of Physics
Published in partnership with: Deutsche Physikalische Gesellschaft and the Institute of Physics

CrossMark
OPEN ACCESS
RECEIVED
31 May 2020
REVISED
27 August 2020
ACCEPTED FOR PUBLICATION
3 September 2020
PUBLISHED
18 September 2020

PAPER

The relativistic tunneling flight time may be superluminal, but it does not imply superluminal signaling

Randall S Dumont[1], **Tom Rivlin**[2] and **Eli Pollak**[2,*]

[1] Department of Chemistry and Chemical Biology, McMaster University, Hamilton, Ontario, Canada L8S 4M1
[2] Chemical and Biological Physics Department, Weizmann Institute of Science, 76100 Rehovot, Israel
[*] Author to whom any correspondence should be addressed.

E-mail: dumontr@mcmaster.ca, tom.rivlin@weizmann.ac.il and eli.pollak@weizmann.ac.il

Keywords: tunneling time, superluminality, time dependent Dirac equation, phase time, momentum filtering

Abstract

Wavepacket tunneling, in the relativistic limit, is studied via solutions to the Dirac equation for a square barrier potential. Specifically, the arrival time distribution (the time-dependent flux) is computed for wavepackets initiated far away from the barrier, and whose momentum is well below the threshold for above-barrier transmission. The resulting distributions exhibit peaks at shorter times than those of photons with the same initial wavepacket transmitting through a vacuum. However, this apparent superluminality in time is accompanied by very low transmission probabilities. We discuss these observations, and related observations by other authors, in the context of published objections to the notion that tunneling can be superluminal in time. We find that many of these objections are not consistent with our observations, and conclude that post-selected (for transmission) distributions of arrival times can be superluminal. However, the low probability of tunneling means a photon will most likely be seen first and therefore the superluminality does not imply superluminal signaling.

1. Introduction

The phenomenon of quantum tunneling was discovered early on by Hund [1], who considered the spectra of optical isomers, and independently by Mandelstam and Leontowitsch [2], who studied scattering through a barrier potential. Their work inspired Gamow [3], who realized that the decay of heavy nuclei and the associated emission of alpha particles occurs due to this tunneling [4].

Quantum tunneling is arguably one of the strangest physical phenomena. Even today, it remains puzzling, especially when considering the question 'how much time does it take a particle to tunnel through a barrier?' This question has received many different answers [6, 7], all indicating that there remains a puzzle. On the one hand, the question is easily posed, but when considering it deeply, the answer is not entirely clear. In fact, the question can be met with another: what do we mean when we talk about time?

The consensus opinion is that an observable, which may be measured experimentally, is an operator and its value is determined by its expectation value. To consider time, one needs to know the time operator. Pauli [5] noted that such an operator cannot be conjugate to the Hamiltonian operator since this would imply an infinite energy spectrum, whereas it is well understood that for almost all physical systems, the energy is bounded from below.

This in itself does not imply that operators with the dimension of time do not exist. Many suggestions have been made over the years, and many of these may be found in the reviews of the topic presented by Muga and coworkers [6–8]. More recent work may be found in references [9, 10].

The definition of a time operator by itself is not sufficient for answering the question, however. It is necessary to show that the suggested operator may be measured experimentally. This is not readily obvious when considering the so called dwell time [6], for instance, as that depends on the density of the scattering wavefunction in the classically forbidden region for a tunneling problem.

Many researchers have suggested coupling the system to an external measurement apparatus which exerts a force on the particle only in the classically disallowed region [11–13]. This strategy has recently been implemented in an experiment by Steinberg and coworkers [14], who deduce a finite time for the tunneling of a Bose–Einstein condensate. In their experiment, the coupling to the external spin system is weak, and therefore hardly changes the tunneling particle itself. Such experiments measure the density of the particle in the tunneling region, but do not directly reveal the time of flight of the scattering event.

One should also note that a flight time should be defined completely within quantum theory without the necessity of recourse to classical mechanical concepts such as the classically allowed or disallowed region of the potential energy surface. Various investigators have shown that such 'Larmor-type' experiments also provide information on what is known as the imaginary time [15] associated with tunneling [12, 16, 17].

What then is the 'real' time it takes a particle to tunnel? Wigner introduced the concept of a (real) time delay as the imaginary part of the logarithmic energy derivative of the scattering amplitude weighted by $\hbar$ [18]. One may use the complex scattering amplitude at the scattering energy E, $T(E) \equiv |T(E)|\exp(\mathrm{i}W(E)/\hbar)$ to define a complex time:

$$\tau = \mathrm{Re}\ \tau + \mathrm{i}\ \mathrm{Im}\ \tau = -\mathrm{i}\hbar \frac{\partial\ \ln[T(E)]}{\partial E} \tag{1.1}$$

from which one readily see that the energy derivative of the phase $W(E)$ gives the real time while the logarithmic derivative of the probability $|T(E)|^2$ gives the imaginary time. We note in passing that with this definition the real and imaginary times are related by a Kramers–Kronig relationship.

Semiclassically, the imaginary time is just the classical period it takes the particle to traverse the barrier on the upside-down potential. The complex time is well defined for a given energy, but it is well understood that if one considers a wavepacket incident on a potential whose energy spectrum is narrow, then the scattered wavepacket when measured at a 'screen' in the transmitted or reflected region of the potential will have a time dependent intensity whose peak appears at a time which is related to the real time—Wigner's time delay. The imaginary part is responsible for the momentum filtering effect by which, especially in the tunneling regime, the exponential dependence of the probability on the energy leads to a predominant filtering of the high energy components of the incident wavepacket as discussed further below.

Many authors have studied this complex time over the years, both theoretically [15, 16, 19–26] and experimentally [27]. A variety of different interpretations have been proposed by authors of these studies for the physical significance of the real and imaginary parts of the time, which we do not discuss in detail here.

The Wigner phase time is related to another problem, namely superluminality. In a relatively unappreciated publication, MacColl, already in 1932 [28] used a steepest descent analysis to show that the appearance time of the peak of the transmitted distribution is independent of both the barrier width and barrier height. This was rediscovered and reaffirmed thirty years later by Hartman [29], who, using Wigner's time delay, also reached the conclusion that the appearance of the transmitted peak on the 'screen' will occur at a time which is independent of the barrier width. But if so, the barrier can be made arbitrarily broad leading to a superluminal effect, that the particle may traverse a distance faster than the speed of light!

These two issues are the motivation for the present paper. Instead of worrying about time operators, one may consider time as a parameter in the time dependent Schrodinger equation. An initial wavepacket scattered on a barrier potential will be either transmitted or reflected and will appear as a function of time at a suitably placed screen in the transmitted or reflected region of the potential. The flux distribution of times may be experimentally measured, at least in principle [30]. One should therefore use these transmitted and reflected time distributions to elucidate the tunneling time. Such an approach is the idea behind the so called 'presence time distribution' [31], with one caveat. In an experiment, the wavepacket, or equivalently the particle would be released at some initial time and a synchronized clock at the screen would measure the arrival time, or in different words, the flight time. In contrast to the presence time distribution which is defined on the time interval $[-\infty, \infty]$ one should consider the flight time distribution which is defined on the time interval $[0, \infty]$. This approach has been recently implemented using a non-relativistic theory for the density flight time distribution [32, 33]. It was shown that when the incident wavepacket is suitably defined, the mean time of arrival of the transmitted and reflected particles, is the same. This was interpreted as indicating that the tunneling time vanishes, since the path of transmitted particles would be expected to be longer than the reflected ones.

But what to make of the supposed superluminality? One possibility is to note that the theories mentioned thus far were all non-relativistic. In a non-relativistic theory it is possible to travel at arbitrary speeds. Could there be a fundamental difference between the tunneling time in a non-relativistic and relativistic theory? This question was addressed by De Leo and Rotelli [34] and separately again a few years later by De Leo [35] using the time-dependent Dirac equation for an electron scattered by a symmetric step potential. Their conclusion was that superluminal tunneling times exist also in this relativistic limit. The same conclusion was presented by del Barco and Gasparian [36].

What about causality? Would it be violated? De Leo and Rotelli left the question open, considering that perhaps one would reach a different conclusion in a field-theoretical study of the scattering. However, their model leaves open some additional questions. They estimated a tunneling time using the Wigner time and then used the ratio of the barrier length to this time to define a velocity. It was this velocity that was then found to be greater than the speed of light. But this is not directly related to any measurement, and as already argued in reference [32], the conditions needed to assure that the transmitted distribution reflects only tunneling would mask any such superluminal measurement. The incident wavepacket has to be very narrow in energy, and would therefore be extremely broad in space, masking any effect of the barrier.

In this paper, we consider a somewhat different approach based on the scattering of wavepackets and the measured time-dependent flux distribution of scattered particles. In a similar manner to De Leo, we consider a relativistic electron scattered on a square barrier and follow the time evolution through the time-dependent Dirac equation. To assure that the observed transmitted wavepacket is due only to tunneling, we place stringent conditions on the incident wavepacket. The probability for energies greater than the barrier are kept much smaller than the below barrier transmission probability, so that the signal is not contaminated by above-barrier transmission.

Secondly, to prevent any misinterpretation, one has to assure that initially, the incident wavepacket does not 'leak' into the transmitted region. This means that it has to be initiated at a distance from the barrier which is much larger than the spatial width of the initial distribution. Our 'screen' is placed close to the edge of the barrier in the transmitted region (in analogy to near-field detection in optics). We then compare the time-dependence of the transmitted flux distribution with the time-dependent flux distribution of a free particle described initially by the same wavepacket but travelling at the speed of light (or arbitrarily close to it).

We find that the peak of the tunneling particle arrives at the screen earlier than the free particle does. In other words, the tunneling particle is superluminal in time. It takes less time for it to arrive at the screen than a free particle traveling at the speed of light. We do not follow references [34, 35] and reinterpret the time distribution in terms of a speed, since one does not know where in space a specific particle originated. In this sense, superluminality may occur even when the quantum dynamics is relativistic. Furthermore, we find that the resulting time distribution agrees well with conclusions based on the quantum phase time delay. Does this superluminality in time violate causality? We will show that this is not necessarily the case; it does not lead inexorably to the conclusion that information can be transmitted faster than the speed of light.

In section 2 we review relativistic time-dependent Dirac theory for an electron scattered upon a barrier. The numerical results showing unequivocal superluminality in time are presented in section 3. The fact that this does not necessarily contradict the special theory of relativity is considered in some detail in section 4. We end with a discussion of further extensions and possible implications.

2. Theory

2.1. The time-dependent Dirac equation

To explore the possibility of superluminal tunneling in the relativistic regime we consider electrons with quantum dynamics governed by the time-dependent Dirac equation, which has the following form in $1+1$ dimensions [37]:

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix} = \hat{H}\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix}, \tag{2.1}$$

where

$$\hat{H} = mc^2\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} + \mathrm{i}c\hbar\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\frac{\partial}{\partial z} + V(z), \tag{2.2}$$

and hats denote operators. This follows from the $3+1$ dimensional Dirac equation when the potential depends only on z. With these simplifications, the x and y degrees of freedom separate. The flux in the z direction is given by

$$\begin{aligned} J &= -\frac{\mathrm{i}}{\hbar}\left[H, \Theta_{(z_1,\infty)}\right] = -\frac{\mathrm{i}}{\hbar}\left[\mathrm{i}c\hbar\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\frac{\partial}{\partial z}, \Theta_{(z_1,\infty)}\right] \\ &= c\left[\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\frac{\partial}{\partial z}, \Theta_{(z_1,\infty)}\right] = c\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\delta(z-z_1). \end{aligned} \tag{2.3}$$

We will consider the temporal properties of an initial Gaussian wavepacket of an electron tunneling through a square barrier, located in the region $z_1 \leqslant z \leqslant z_2$. At a given energy, E, the Dirac equation has (stationary) scattering eigenfunctions denoted as $\psi_E(z)$, such that the time-dependent wavepacket may be expanded in the energy representation as:

$$\psi\left(z,t\right) = \int_{-\infty}^{\infty} \mathrm{d}E\ b_E\ \exp\left(\frac{-\mathrm{i}Et}{\hbar}\right)\psi_E\left(z\right), \tag{2.4}$$

where

$$b_E = \int_{-\infty}^{\infty} \mathrm{d}z\ \psi_E^{\dagger}\left(z\right)\psi\left(z,0\right) \tag{2.5}$$

are the expansion coefficients of the initial wavefunction with the scattering eigenfunctions.

Our objective is to study the tunneling phenomenon. As such, we must assure that the initial wavepacket is chosen such that initially it does not leak into the transmitted region, and such that it has negligible above-barrier energy components. The momentum representation is taken to be:

$$\langle p|\psi\rangle \equiv \psi(p,z;t=0) = C\ \exp\left(-\frac{\left(p-p_0\right)^2}{2\Gamma} + \frac{\mathrm{i}}{\hbar}\left[p\left(z-z_0\right)\right]\right)\mathbf{u}^{(1)}\left(p\right), \tag{2.6}$$

where C is a normalization factor, z_0 and p_0 are the centers of the wavepacket in coordinate and momentum space respectively, and

$$\mathbf{u}^{(1)}\left(p\right) = \begin{pmatrix} 1 \\ \dfrac{cp}{E\left(p\right)+mc^2}\end{pmatrix} \tag{2.7}$$

is the two-component Dirac spinor of a spin-up, positive energy eigenstate that is slowly varying in p. The two-component spinor is sufficient, since spin cannot flip in a one-dimensional potential. The normalization is chosen to best facilitate localization of the wavepacket to the left of the barrier. It is not the usual covariant normalization. Since the goal is a space-localized superposition, that is not necessary here.

As mentioned, the initial conditions must be chosen appropriately, such that $|\langle z_1|\psi\rangle|^2 \ll |T\left(p_0\right)|^2$, where the transmission amplitude is [34]

$$T\left(p\right) = \frac{\exp\left(-\mathrm{i}pl/\hbar\right)}{\cosh\left(ql/\hbar\right) + \frac{1+\alpha^2}{2\alpha}\sinh\left(ql/\hbar\right)}, \tag{2.8}$$

with

$$\alpha = \mathrm{i}\frac{q}{p}\frac{E+m}{E-V+m}, \tag{2.9}$$

where $p = \sqrt{E^2/c^2 - m^2c^2}$, $q = \sqrt{m^2c^2 - (E-V)^2/c^2}$, and $l = z_2 - z_1$ is the width of the barrier. Therefore, the initial location of the wavepacket must be such that $z_1 \gg \hbar/\sqrt{\Gamma}$. The second condition—that the probability for energy components above the barrier energy (V) must be much smaller than the transmission probability—implies that $\Gamma/m \ll V - E$. In other words, the momentum width must be narrow, the coordinate width must be very broad, and the initial spatial center of the packet must be far removed from the left side of the barrier.

With these preliminaries, we observe the time-dependent transmitted wavepacket at the right of the barrier, where the spatial form of the eigenstate of the Dirac Hamiltonian is $T\exp(\mathrm{i}pz)$, and we may write down the time-dependent coordinate representation of the transmitted wavepacket as

$$\psi\left(z,t\right) = C\int_{-\infty}^{\infty}\mathrm{d}p\ \exp\left(-\frac{\left(p-p_0\right)^2}{2\Gamma}\right)T\left(p\right)\mathbf{u}^{(1)}\left(p\right)\exp\left(\mathrm{i}\frac{p\left(z-z_0\right)-E\left(p\right)t}{\hbar}\right). \tag{2.10}$$

Note also that the relativistic relation between the momentum and the energy is given by

$$E\left(p\right) = \sqrt{m^2c^4 + c^2p^2}. \tag{2.11}$$

In what follows, we will determine the flux time distribution at the right edge (z_2) of the barrier, which is by definition

$$\tilde{P}\left(t\right) = c\begin{pmatrix}\psi_0^*\left(z_2,t\right) & \psi_3^*\left(z_2,t\right)\end{pmatrix}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}\psi_0\left(z_2,t\right)\\ \psi_3\left(z_2,t\right)\end{pmatrix}. \tag{2.12}$$

As defined here, $\tilde{P}(t)$ is positive and integrates over all t to the total transmission probability. The normalized flux time distribution ($P(t)$) is then given by dividing $\tilde{P}$ by the transmission probability.

2.2. The steepest descent approximation

Below, we will evaluate the integral in equation (2.10) with numerical accuracy. However, it is useful when analyzing the results to consider a steepest descent evaluation of the time-dependent wavefunction, which is especially justified since the width parameter, Γ, is chosen to be small to assure that we only consider tunneling. The steepest descent approximation reveals the MacColl–Hartman effect: we expect the tunneling time to be independent of barrier width. With this in mind, it is convenient to write the transmission amplitude in the form

$$T\left(p\right) = \exp\left(\mathrm{i}W\left(p\right)/\hbar\right), \tag{2.13}$$

where

$$\begin{aligned} W\left(p\right) &= \left(\mathrm{i}q - p\right)l + \mathrm{i}\hbar\ \ln\left(\left(\frac{1}{2} + \frac{1+\alpha^2}{4\alpha}\right) + \left(\frac{1}{2} - \frac{1+\alpha^2}{4\alpha}\right)\exp\left(-2ql/\hbar\right)\right) \\ &\equiv \left(\mathrm{i}q - p\right)l + \mathrm{i}\hbar\ \ln\ \sigma\left(p\right), \end{aligned} \tag{2.14}$$

and α has been defined in equation (2.9).

We denote the exponent appearing in equation (2.10) as

$$F\left(p,t\right) = \mathrm{i}\hbar\frac{\left(p-p_0\right)^2}{2\Gamma} + W\left(p\right) + p\left(z_2 - z_0\right) - E\left(p\right)t \tag{2.15}$$

and the momentum dependent velocity as:

$$v\left(p\right) = \frac{\mathrm{d}E}{\mathrm{d}p} = \frac{cp}{\sqrt{p^2 + m^2c^2}}. \tag{2.16}$$

The steepest descent point $p^\sharp\left(t\right)$ of the integral is the solution of the steepest descent condition

$$\frac{\partial F\left(p,t\right)}{\partial p} = 0 = \mathrm{i}\hbar\frac{p-p_0}{\Gamma} + \frac{\partial W}{\partial p} + \left(z_2 - z_0\right) - \frac{\partial E}{\partial p}t. \tag{2.17}$$

The steepest descent approximation for the time dependent wavefunction at the location of the screen is then readily found to be:

$$\psi\left(z_2,t\right) \simeq C\frac{\left(2\pi\right)^{1/2}}{\left|\frac{\partial^2 F}{\partial p^2}\left(p^\sharp,t\right)\right|^{1/2}}\mathbf{u}^{(1)}\left(p^\sharp\right)\exp\left(\frac{\mathrm{i}}{\hbar}F\left(p^\sharp,t\right)\right). \tag{2.18}$$

The second derivative term in the prefactor is given by

$$\frac{\partial^2 F\left(p,t\right)}{\partial p^2} = \mathrm{i}\frac{\hbar}{\Gamma} - \frac{\mathrm{d}^2\ \ln\ W}{\mathrm{d}p^2} - \frac{t}{m\gamma^2}, \tag{2.19}$$

since

$$\frac{\mathrm{d}^2E}{\mathrm{d}p^2} = \frac{\mathrm{d}v}{\mathrm{d}p} = \frac{1}{m\gamma^2}, \tag{2.20}$$

where $\gamma = \left(1 - v^2/c^2\right)^{-1/2}$ is the ubiquitous Lorentz factor of relativity.

In the narrow-in-energy initial wavepacket limit, the two components of the spinor wavefunction are proportional, such that

$$\tilde{P}\left(t\right) \cong \tilde{P}_{\mathrm{sd}}\left(t\right) = \frac{2\pi|C|^2}{\left|\frac{\partial^2 F}{\partial p^2}\left(p^\sharp,t\right)\right|}\ \exp\left(-2\ \mathrm{Im}\ \frac{F\left(p^\sharp,t\right)}{\hbar}\right)\mathrm{Re}\ \frac{c^2p^\sharp}{E^\sharp + mc^2}, \tag{2.21}$$

where $|C|^2$ is a normalization constant. The most probable tunneling time corresponds to a maximum in the argument of the exponential, with respect to the varying of the time:

$$\frac{\mathrm{d}}{\mathrm{d}t}\ \mathrm{Im}\left(\mathrm{i}\hbar\frac{\left(p^\sharp - p_0\right)^2}{2\Gamma} + W\left(p^\sharp\right) + p^\sharp\left(z_2 - z_0\right) - E^\sharp t\right) = 0. \tag{2.22}$$

This reduces to

$$\mathrm{Im}\left(E^{\sharp}\right)=0, \tag{2.23}$$

on account of equation (2.17). The most probable tunneling time thus occurs when $E^{\sharp}$ is real, or equivalently, when $p^{\sharp}$ is real.

It is useful then to consider a complex time (the complex time of equation (1.1)),

$$\tau\left(p\right)=\frac{\mathrm{d}\left[W+p\left(z_2-z_0\right)\right]}{\mathrm{d}E}=\frac{z_1-z_0}{v\left(p\right)}-\mathrm{i}\frac{z_2-z_1}{u\left(p\right)}+\mathrm{i}\hbar\,\frac{\mathrm{d}\,\ln\,\sigma}{\mathrm{d}E}, \tag{2.24}$$

where

$$u\left(p\right)=-\frac{\mathrm{d}E}{\mathrm{d}q}=\frac{cq\left(p\right)}{\sqrt{m^2c^2-q^2\left(p\right)}} \tag{2.25}$$

gives the positive imaginary velocity, $\mathrm{i}u\left(p\right)$, under the barrier. As is evident from equation (2.24), the complex time $\tau(p)$ is associated with travel from z_0 to z_2, at energy, E. In equation (2.24) the first term on the right-hand side is the real, classical time to travel to z_1. The second term is the negative imaginary, under-barrier contribution to time, and the third term adds a quantum contribution to both the real and imaginary parts.

Returning then to the most probable time as determined from equation (2.22), the real and imaginary parts of the equation are separated to give the most probable time

$$t_{\mathrm{mp}}=\mathrm{Re}\;\tau\left(p^{\sharp}_{\mathrm{mp}}\right). \tag{2.26}$$

As we shall see, this leads to the MacColl–Hartman effect, and

$$p^{\sharp}_{\mathrm{mp}}=p_0-\frac{\Gamma}{\hbar}\;\mathrm{Im}\;\tau\left(p^{\sharp}_{\mathrm{mp}}\right)v\left(p^{\sharp}_{\mathrm{mp}}\right), \tag{2.27}$$

expresses a momentum filtering effect (as discussed below).

From the definition of $\tau(p)$ (equation (2.24)) and the transmission amplitude exponent (equation (2.14)), one readily finds that

$$\mathrm{Re}\;\tau^{\sharp}_{\mathrm{mp}}=\frac{z_1-z_0}{v^{\sharp}_{\mathrm{mp}}}-\hbar\;\frac{\mathrm{d}}{\mathrm{d}E}\;\mathrm{Im}\left[\ln\left(\left(\frac{1}{2}+\frac{1+\alpha^2}{4\alpha}\right)+\left(\frac{1}{2}-\frac{1+\alpha^2}{4\alpha}\right)\exp\left(-2q^{\sharp}_{\mathrm{mp}}l/\hbar\right)\right)\right], \tag{2.28}$$

where α is evaluated at $p^{\sharp}_{\mathrm{mp}}$. Here, the first term is the classical time it takes to reach the left edge of the barrier. The second term is much smaller, and in the limit of a long barrier it asymptotically tends towards a constant. This constant is independent of the barrier width, and reduces to

$$-\hbar\;\frac{\mathrm{d}}{\mathrm{d}E}\;\mathrm{Im}\;\ln\left(\frac{1}{2}+\frac{1+\alpha^2}{4\alpha}\right)\bigg|_{p=p^{\sharp}_{\mathrm{mp}}}=\frac{2\hbar}{v^{\sharp}_{\mathrm{mp}}}\left(-\frac{\mathrm{d}\left|\alpha\right|}{\mathrm{d}p}\right)_{p=p^{\sharp}_{\mathrm{mp}}}\frac{1}{1+\left|\alpha^{\sharp}_{\mathrm{mp}}\right|^2}. \tag{2.29}$$

This term is positive, since $|\alpha|$ is a decreasing function of p and independent of the barrier width l. This independence is the MacColl–Hartman effect. If the steepest descent estimate is indeed valid, then the tunneling time in the wide barrier limit becomes independent of the barrier width.

Equation (2.27) expresses the momentum filtering effect [38–42]. Since the transmission probability increases exponentially with increasing incident momentum, the transmitted wavepacket will filter preferentially the higher momentum components of the incident wavepacket. In our case, $\mathrm{Im}\;\tau^{\#}$ is negative, so the steepest descent momentum is larger than the incoming momentum, in proportion to $\left|\mathrm{Im}\;\tau^{\#}\right|$ and the incident wavepacket momentum width parameter Γ. This corresponds to the transmitted portion of the wavepacket traveling at higher momentum than the incoming wavepacket central momentum p_0. We do note that this effect is mitigated in the relativistic limit, where $v\to c$ while $p\to\infty$.

Finally, one may also use the steepest descent approximation to approximate the tunneling time probability distribution. One expands to second order the exponent $F\left(p,t\right)$ about the most probable time and associated saddle point momentum. The result—see the appendix A—is

$$\mathrm{Im}\ F\left(p^{\sharp},t\right)_{\mathrm{sd}}=\mathrm{Im}\ F\left(p^{\sharp}_{\mathrm{mp}},t_{\mathrm{mp}}\right)+\frac{\kappa\left(v^{\sharp}_{\mathrm{mp}}\right)^{2}}{2}\left(t-t_{\mathrm{mp}}\right)^{2}, \tag{2.30}$$

where

$$\kappa=\mathrm{Re}\ \frac{1}{\frac{\hbar}{\Gamma}-\mathrm{i}\frac{\mathrm{d}^{2}\ \ln\ W^{\sharp}_{\mathrm{mp}}}{\mathrm{d}p^{2}}+\mathrm{i}\frac{t_{\mathrm{mp}}}{m\gamma^{2}}}. \tag{2.31}$$

The steepest descent flux time distribution is (see also equation (2.21))

$$\tilde{P}_{\mathrm{sd}}\left(t\right)=\frac{v_{0}(\pi\Gamma)^{-1/2}\hbar^{1/2}}{\left|\frac{\partial^{2}F}{\partial p^{2}}\left(p^{\sharp}_{\mathrm{mp}},t_{\mathrm{mp}}\right)\right|}\ \exp\left(-\mathrm{Im}\ \frac{F\left(p^{\sharp},t\right)}{\hbar}\right)\frac{E_{0}+mc^{2}}{E^{\sharp}_{\mathrm{mp}}+mc^{2}}\frac{p^{\sharp}_{\mathrm{mp}}}{p_{0}}. \tag{2.32}$$

Due to the choice of a narrow-in-momentum wavepacket, this Gaussian steepest descent approximation, is, as in the non-relativistic case, rather accurate, as shall be demonstrated in the next section. For wavepackets narrow in momentum, the associated width in the time distribution scales linearly with the inverse of $\sqrt{\Gamma}$—it is approximately $\hbar(\Gamma)^{-1/2}\left(v^{\sharp}_{\mathrm{mp}}\right)^{-1}$. The normalization constant, $|C|^{2}$, is eliminated by requiring $\tilde{P}_{\mathrm{sd}}\left(t\right)$ to normalize to unity in the case of a free particle. The above flux distribution is normalized to the total transmission probability. Thus, the model also provides an approximate (in fact, very accurate) formula for the total transmission probability,

$$C_{\mathrm{trans,sd}}=\frac{v_{0}\Gamma^{-1/2}2^{1/2}\hbar}{v^{\sharp}_{\mathrm{mp}}\kappa^{1/2}\left|\frac{\partial^{2}F}{\partial p^{2}}\left(p^{\sharp}_{\mathrm{mp}},t_{\mathrm{mp}}\right)\right|}\ \exp\left(-\mathrm{Im}\ \frac{F\left(p^{\sharp}_{\mathrm{mp}},t_{\mathrm{mp}}\right)}{\hbar}\right)\frac{E_{0}+mc^{2}}{E^{\sharp}_{\mathrm{mp}}+mc^{2}}\frac{p^{\sharp}_{\mathrm{mp}}}{p_{0}}. \tag{2.33}$$

Also of interest here is the post-selected tunneling time distribution—the normalized (to one) distribution of tunneling times among realizations of the process where the particle did tunnel. The post-selected tunneling time distribution is $P\left(t\right)=\tilde{P}\left(t\right)/C_{\mathrm{trans}}$, while in the steepest descent approximation, the equivalent quantity is $P_{\mathrm{sd}}\left(t\right)=\tilde{P}_{\mathrm{sd}}\left(t\right)/C_{\mathrm{trans,sd}}$.

3. Numerical evidence for superluminal tunneling transit times

We first considered the scattering of an electron whose initial mean velocity is 0.99 c on a square barrier potential with barrier height 6.7 mc^2 with varying width. The momentum width of the wavepacket was $\hbar\Gamma^{-1/2}=15\ \lambda_{\mathrm{rc}}$ (where λ_{rc} is the reduced Compton wavelength, 3.86×10^{-13} m). The initial and final coordinates of the barrier were $z_1=120\ \lambda_{\mathrm{rc}}$ and $z_2=125$, 130 and 135 λ_{rc}. Initially, the wavepacket was centered at $z_0=0$–$8\ \hbar\Gamma^{-1/2}$ away from the leading edge of the barrier. The distance to the barrier ensures that the density is negligible at (and beyond) $z_1=120$. Specifically, it is less than 10^{-27} times its peak value. In the momentum representation, the wavepacket is initially centered at $p_0=7.0179\ mc$. At the threshold for above-barrier transmission, $p=7.6348\ mc$, the wavepacket density is less than 10^{-37} times its peak value.

Figure 1 shows the resulting flux time distributions of the tunneling particle, for the three barrier widths considered. The top panels show the flux probability distributions $\tilde{P}\left(t\right)$ on a logarithmic scale and the associated cumulative probability $\left(\int_0^t \mathrm{d}t'\tilde{P}\left(t'\right)\right)$. The bottom panels show the flux probability distribution on a linear scale, normalized to unity.

Also shown are the corresponding flux time probability distributions for photons in a vacuum with initial wavepackets with the same shape. The numerically exact flux time distributions are compared to their steepest descent estimates—the two are practically identical. Superluminal transmission is evident here, as the peak of the tunneled flux arrives earlier than that of light in a vacuum.

The wider the barrier, the larger the gap between the earlier arrival time of the tunneled particle and the free photon. Note, however, that the transmission time distribution for photons is, at all times, much larger than the equivalent for the tunneling electrons. This is the case for all parameter choices we have considered thus far. The likelihood of photon detection, especially at short (but also long) times, always exceeds the likelihood of tunneled electron detection, due to the small tunneling probability. This is related to the fact that the barrier width is sufficiently large such as to create a noticeable speedup of the tunneling electron.

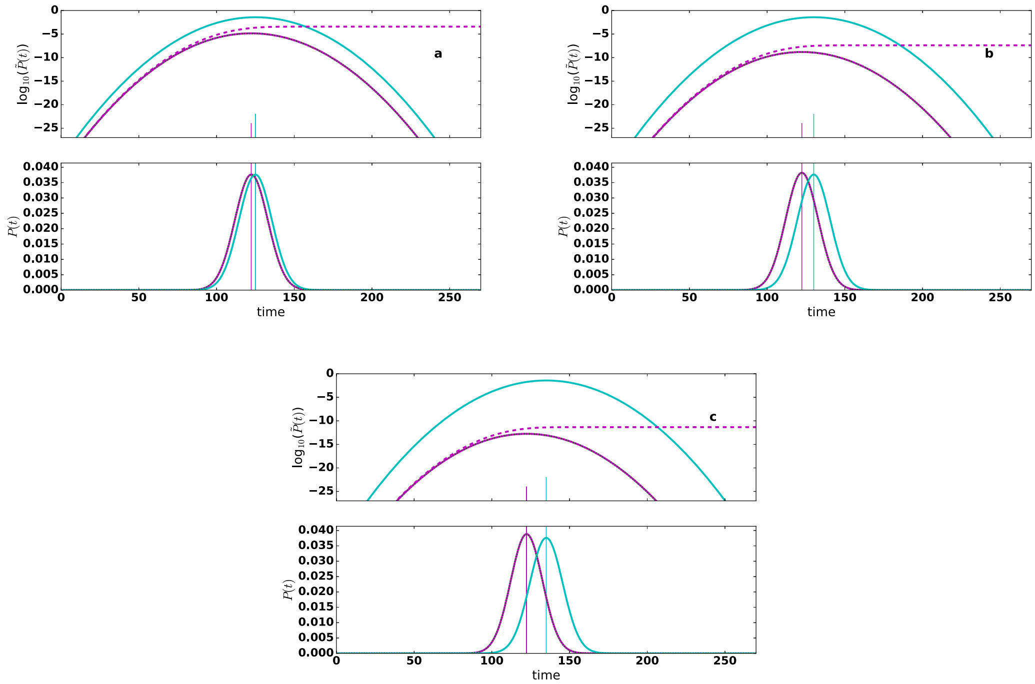

Figure 1. The tunneling flux time probability distributions are compared to the free photon distributions. The top panel of each subfigure shows the distributions normalized to the transmission probability, the bottom panel shows the same distributions normalized to unity. Tunneling distributions are in magenta, photon distributions are in cyan. The barrier widths in the three figures are 5 λ_{rc} for the graph labelled (a), 10 λ_{rc} for (b), and 15 λ_{rc} for (c). Also shown are the associated cumulative probability distributions (dashed magenta). The dotted green line shows the steepest descent estimate of the flux time probability distribution. It is almost identical to the numerical result. The tick marks above the time axis indicate the times at which the electron and photon fluxes are at their maxima. Time has units of $\lambda_{rc}/c = 1.2881 \times 10^{-21}$ s.

Another interesting aspect of the three figures is that the tunneled distributions are slightly narrower than the photon distribution, and that this narrowing also increases with barrier width. This feature is a result of the momentum filtering by the barrier. As the barrier width grows, momentum filtering reduces the contribution of lower momenta, thus slightly reshaping the time distribution. Decreasing the momentum width of the initial distribution would reduce this effect. Observations of apparent superluminality, such as those seen in figure 1, have been subject to much debate. We examine below discussions of this issue by other authors in the context of our computational results.

Having established the superluminal effect, the question remains as to what the tunneling time is. For this purpose, in figure 2 we plot the mean transmission time as well as the peak arrival time (the two are almost identical) for the same initial wavepacket as considered in figure 1 as a function of the barrier width for the different barrier heights 6.5, 6.6 and 6.7 mc^2. The mean transmission time is defined as

$$\langle t\rangle = \int_0^\infty \mathrm{d}t\ tP(t)\,. \tag{3.1}$$

The dashed blue line at the bottom of the figure, is the time the maximum of the distribution appears at z_1. This is the time that would be observed if the barrier width vanishes. If the tunneling time vanishes, then one expects that this time would coincide with the mean tunneling time. As is evident from the figure, however, it does not. It is, in fact, slightly lower than the mean time. The difference between the two is thus the tunneling time—it is finite but small.

For each barrier width, this tunneling time is less than $c\,l$, the time needed to traverse the barrier at the speed of light. In accordance with the MacColl–Hartman effect, as the barrier width increases, the effective time approaches a constant that is independent of the barrier width.

The horizontal blue lines on the right give this constant time—evaluated using equations (2.28) and (2.29) with $p^\sharp_{\mathrm{sd}}$ for the barrier width, 19 λ_{rc}—for the three barrier heights, thus establishing that the tunneling time is indeed given by the phase time.

Note that for the barrier height of 6.5 mc^2, the mean and maximum times are still rising as the asymptote is approached. This is because the total probability associated with above-barrier transmission is 6.6×10^{-19} when the barrier height is 6.5 mc^2. This becomes non-negligible compared to the tunneling probability for the largest barrier widths considered. For barrier width 19 λ_{rc}, the total transmission probability is 3.3×10^{-15}.

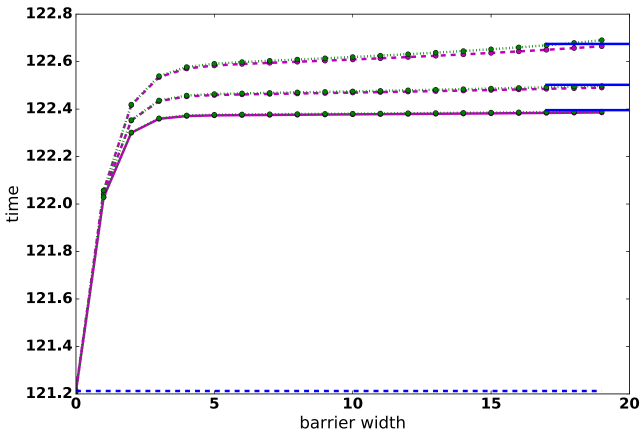

Figure 2. The barrier-width dependence of the mean tunneling time (magenta) and most probable time (green). The no-barrier (i.e., zero width) mean tunneling time is indicated with the dashed blue line. Initial electron velocity and initial wavepacket width are as in figure 1. The barrier heights (from lower curve to top curves) are 6.7, 6.6 and 6.5 mc^2 (the lower curves correspond to the case of figure 1). The short solid blue lines on the right indicate the tunneling time estimate using the phase, in the limit of large barrier width.

These results are qualitatively similar to those presented in figure 2(a) of reference [34]. However they are not the same, since De Leo and Rotelli considered only the time from one edge of the barrier to the other. Here, we considered the time from the initial mean position of the wavepacket z_0 to the right edge of the barrier, using parameters which ensure that initial leakage of the wavepacket to the transmitted side is negligible. To be sure that the time is superluminal one must initiate the wavepacket sufficiently far to the left of the barrier such that any density initially under or to the right of the barrier is negligible. It is then necessary to consider the total time to reach the right edge of the barrier (the target) from the initial position far to the left of the barrier, as shown in figure 1.

4. Discussion: interpretations of superluminality in the MacColl–Hartman effect

4.1. Tunneling and special relativity

Any result that claims to detect physical objects being transmitted at superluminal times must be subjected to close scrutiny. The MacColl–Hartman effect is no exception, and many authors have discussed these results and their implications over the years. It is important to explore the implications of the specific results in this paper in light of these discussions, especially since this paper explicitly works in a Lorentz-invariant framework.

Special relativity is often phrased as saying that no object can travel at a speed faster than that of light in a vacuum, but it is perhaps more accurate to say that *signals* cannot be sent superluminally [43]. There are many instances of seeming superluminality (e.g. phase velocities [43], entangled particles [44]), but in all these cases it can be shown that no information can be meaningfully transmitted at FTL speeds.

It can even be shown that under certain circumstances, classical *group velocities* can be greater than the speed of light [45]. One manifestation of this is a classical analogue of quantum tunneling where a classical wave 'tunnels' through a barrier [46]. The tunneling can be thought of as being due to evanescent modes of the wave [47]. (This has been described as a classical analogue of virtual photons [48].)

This classical tunneling also leads to a classical version of the MacColl–Hartman effect, which has been realized experimentally in the so-called 'double-prism' experiment [49]. As such, many authors have discussed the physical implications of the MacColl–Hartman effect in a classical setting over the years, and the possibility of superluminal transmission [45–52]. Many arguments made in the classical case are also adapted for the quantum case, and thus are worth reviewing together.

We identify three main interpretations of the MacColl–Hartman effect. Within each interpretation there are a number of separate claims about observable quantities.

4.2. Superluminal signals vs causality violation

Most authors dispute that there is any form of superluminal signalling, but we identified one author who claims this signalling does occur. Thus, the first interpretation, due to Nimtz, is that superluminal *signal transmission* is possible, but this does not violate *causality* [49].

In special relativity, it is usually stated that one always implies the other, thanks to Lorentz transformations, and the two terms are often used interchangeably. Nimtz, however, presents an argument that it is possible to have one without the other in the case of these tunneled wavepackets. For this reason, we have refrained from commenting on causality in this work, focusing only on the question of superluminal signals. This is only an academic distinction, however, as we currently believe that there is no superluminal signal transmission, despite Nimtz's claims to the contrary [52].

4.3. Wavepacket continuity

Why is it then that one observes superluminality, but it cannot be utilized for superluminal signalling? A number of interpretations have been suggested. One example is the interpretation that the wavepacket that emerges from the barrier is somehow causally disconnected from the one that entered it. There are two distinct empirical claims within this interpretation.

One common claim is that the wavepacket is re-shaped so much by the tunneling process that the object that emerges from the barrier must be thought of as a separate, distinct wavepacket from the one that entered. One author [45], for instance, describes the tunneling process as distorting the shape of the wavepacket, destroying the signal. A review by Nimtz [47] quotes another author as saying 'the initial packet gets completely deformed and unsuitable for perfect signal transmission during the course of propagation'.

In his article on the topic, Winful states: 'input and output peaks are not related by a simple causal translation' [46]. Sokolovski makes a similar point in his chapter in Muga *et al*'s monograph [6]. To illustrate the point, Winful's article [46] includes a cartoon of an elephant entering the tunnel and an ant exiting, emphasizing the vast difference in scale between the pre- and post-tunnel wavepackets, and implying they are not directly comparable.

A different claim is that the 'time of flight' that we have been using in this work is in some way not a meaningful metric. By this logic, it is meaningless to describe the time of flight of the wavepacket as being superluminal. In practice, this claim often appears in tandem with the previous one about the wavepacket, as is the case in Winful's article, for instance [46]. There he claims that the tunneling time in fact measures something other than a time of flight. Winful even rejects the idea that the time the particle spends inside the barrier is a meaningful quantity to discuss, for the same reasons.

We object to both of these claims. Our results clearly show that the tunneled wavepacket is almost identical in shape to the one that entered the barrier, differing only in absolute amplitude due to the low transmission probability. There is a momentum filtering effect due to the tunneling probability's dependence on the momentum, but for the systems studied here this effect is small in the non-relativistic case and even smaller in the relativistic one.

Consequently, we believe that it is possible to observe the arrival of the tunneled particle at a screen, and use that to infer an experimentally observable time of flight for the wavepacket that begins on one side of the barrier and ends on the other. In that regard, the MacColl–Hartman effect must be taken seriously as genuinely shortening this time of flight, such that superluminal time at least appears possible.

4.4. Superluminal times vs superluminal signals

The final interpretation we consider here is that there are superluminal times, but that the circumstances under which they arise also preclude them from being used for superluminal signalling. This is also our interpretation, although there are still claims made in this interpretation that we take issue with. Such claims pertain to the shape of the wavepacket in different ways.

One claim focuses on an attempt to divide the wavepacket into separate 'parts'. A popular science article written by authors who have studied this phenomenon claims that the transmitted pulse is 'front-loaded', meaning the 'parts' of the initial pulse that 'begin' closer to the barrier are the ones that are more likely to be transmitted, thus appearing to explain the superluminal times without needing to invoke superluminality [53]. In the words of the authors: 'only the leading edge of the incident pulse survives the tunnelling event without being severely attenuated' [53]. This, of course, has overlap with the previous interpretation, as it appears to imply that the wavepackets on either side of the barrier are incomparable in some meaningful ways.

Arguments such as these that make use of the shape of the wavepacket in this particular way appear to run into a certain quantum mechanical problem: they seem to rely on the ability to know which part of the initial wavepacket a final measurement originates from. Outside of Bohmian mechanics, this information does not exist in quantum mechanics. Within Bohmian mechanics, this information exists but is fundamentally unknowable.

In essence, we contend that this form of argument is not permitted by the postulates of quantum mechanics—it is impossible to assign a certain set of measurements to a specific part of the overall initial probability distribution, just as it is impossible to relate a single particle to its passage through one of two slits in the double-slit experiment.

Two further claims involve the ‘front’ of the wavepacket. The front plays a role in many explanations of this phenomenon, in both the classical and quantum cases [46], and other authors [54] have used discussions of the front to incite broader discussions about the Hartman effect as we have here.

One of these claims is that observing the front of the wavepacket allows one to use analytic continuation to infer the shape of the full wavepacket [45], without needing to consider at all the time of arrival of the peak of the wavepacket. The velocity of the front of the wavepacket is therefore equivalent to the velocity of the transfer of information. As one source phrases it, ‘there is no information contained in the peak of an analytic wavepacket which is not already present in its forward tail’ [50]. It is then asserted that since the front travels no faster than light, no superluminal signals occur.

Relatedly, a further claim is that ‘new information is contained in a signal only when it deviates from the analytic continuation, especially when it shows singularities like steps or cusps’ [45]. In other words, the signal must be ‘turned on’ at some point. Of course, one implication of a discontinuity in position space in the initial wavepacket is that there must be a large variance in momentum space, meaning there must be above-barrier transmission.

We do not agree with these two claims. Our results do not have any such discontinuities, all of our initial wavepackets are narrow in momentum space well below the barrier height, and they are all well-defined. Furthermore, Nimtz has repeatedly [47, 49, 51, 52] objected to trying to use the ‘front velocity’ to explain away superluminality, claiming that the ‘front velocity’ is simply the same as the group velocity, a front ‘has no physical meaning’, and that ‘a single photon has no front’.

This final point is key to our understanding of the flaws of these arguments. In quantum mechanics, the wavefunction gives us information on the probabilistic behavior of a large ensemble of scattered particles. But, as is well known from the double-slit experiment, a single particle is just one manifestation. In any scattering experiment of many particles we cannot know in advance that we are measuring the front of the wavepacket, as opposed to its back or its middle. Since the probability amplitude at the front is very small, one may expect generically that in any measurement one will first see the central peak rather than the front.

The claim about the fronts not traveling superluminally has been expanded in two analogies by Chiao, Kwiat, and Steinberg, and our rebuttal to these claims is well-illustrated through a detailed discussion of these analogies, along with a thought experiment of our own.

4.5. Analogies and thought experiments

The first analogy that Steinberg uses is one of a train losing cars. In a popular science article on the subject, Steinberg says ‘the stopwatch starts when the center of the train leaves the station, but the train leaves cars behind at each stop. So when the train arrives in New York, now comprising only two cars, its center has moved ahead, although the train itself hasn’t exceeded its reported speed’ [55].

This argument appears to imply that we are actually able to track individual components of the initial wavepacket (the ‘train cars’, in the analogy), and identify parts of the final wavepacket as having come from the front or back of the initial distribution (those same train cars). This implication is problematic, as it would only be possible to track the individual cars in this way with access to some hidden variables as in the case of Bohmian mechanics and the like. This is not the true spirit of the analogy, however. In fact, the essence of Steinberg’s train analogy is that the peak of the tunneled wavepacket being earlier is irrelevant because the fronts of the wavepackets still arrive at the same time [56].

This is more clearly stated in another analogy due to Steinberg, along with collaborators Chiao and Kwiat. In their ‘turtle race’ analogy presented in Scientific American [56], (inspired by experimental results they had recently obtained [57]), two turtles race, one going through the barrier. The ‘peak’ of the tunneled turtle’s shell arrives earlier, and thus by that metric it ‘wins the race’, but the heads of both tortoises still arrive at the same time. We claim that the results presented in figure 1 negate this argument. The head of the tunneling ‘tortoise’ arrives earlier than the one which did not tunnel.

As far as we understand the issue, the essential question which should be addressed is whether it is in principle possible to receive a signal from a tunneling particle before one would expect to receive it from a free particle travelling at the speed of light.

In that spirit we propose a simple thought experiment. Suppose that Bob is in trouble and wants to alert Alice, who is distant from him, in the quickest possible way. Is it preferable for him to send a particle through a barrier or without a barrier? As long as a single particle can be fired at and read by a detector, the possibility for FTL signaling seems to exist, and the tunneled peak being earlier makes this possibility sound quite plausible.

The upshot of all these discussions is straightforward. On the one hand, the results in the bottom panels of each subfigure of figure 1 show that the tunneling particle is superluminal in the sense that *if* it arrives, it does so earlier than the free particle. At the same time the results in the top panels of the subfigures will convince Bob to send a free particle, since the probability of the particle arriving at an earlier time is exponentially larger. Our results show that tunneling is another case of superluminality without superluminal signaling.

It is undeniable that the peak time of the tunneled distribution is shorter than that of the free particle distribution, and in this sense it is accurate to say that something is travelling 'superluminally'. However, figure 1 clearly shows that this phenomenon cannot be used to send a signal with a single, tunneled particle faster than one could send that same signal using a free particle. In other words: it will always be preferable to transmit a free particle traveling at or close to the speed of light in order to convey a signal, and not a tunneled particle traveling at the same speed. Sala, Brouard, and Muga make a similar argument in their 1995 paper [54].

One final caveat worth noting is that our results do not definitively prove that an *ensemble* of free particles will always be a preferable method of transmitting a signal to an ensemble of tunneled particles. It is our intention to include a discussion of this in a follow-up paper.

5. Final remarks

The fundamental issue is that the probability of seeing a tunneled particle is so small that the free particle will always be seen first. That is why signaling will be subluminal. However, if we consider the post-selected density (i.e. the density conditional on the electron having tunneled), then we see the opposite result—the tunneled electron arrives first. As for the shape and causal connection between the incident and transmitted wavepacket, as long as the incident wavepacket is a simple Gaussian far to the left of the barrier, and well below the above-barrier momentum threshold, the tunneled wavepacket has almost the same shape as the initial wavepacket—only it is almost uniformly diminished by the tunneling amplitude. There is a slight enhancement of the leading portion of the wavepacket, relative to the lagging portion, since momentum filtering favors the faster components of the initial wavepacket. However, this effect is small for narrow-in-momentum wavepackets. In the relativistic limit, it is very small, since momentum filtering is suppressed in this case. The result is an almost-Gaussian tunneling time distribution with width simply related to the width of the initial wavepacket.

Wavepacket reshaping may occur, for instance in wide-in-momentum wavepackets and higher order Hermite–Gaussian wavepackets [58], which have multiple momentum components. However, it plays a negligible role in our results. The tunneled flux arises from the entire initial wavepacket, not just the near-barrier portion, as claimed by some authors. A single particle detected at the screen cannot be attributed as coming from the front, center or back of the incident wavepacket.

Finally, we note that we have established, at least as far as the tunneling through a square barrier is concerned, that the tunneling time is finite, but short, and given by the phase time. In references [32, 33] it was shown in the non-relativistic case that the mean flight time of transmitted and reflected particle is the same in the limit of an infinitely narrow in momentum incident wavepacket. This was then interpreted as implying that the tunneling time vanishes, since the path of the transmitted particle is seemingly longer than that of the reflected one. The present results negate this claim, in fact the phase time associated with the reflected and transmitted particles is the same, so that the means are the same, yet the tunneling time does not vanish.

Acknowledgments

This work was supported by a Grant from the Israel Science Foundation. RSD thanks the Weizmann Institute of Science for a Visiting Faculty Fellowship, including a last-minute plane ticket to get back home before the Covid-19 lockdown, and McMaster University for a Professional Development Allowance.

Appendix A. The steepest descent approximation for the tunneling time distribution

The approximation is based on a second order expansion of the phase, $F(p, t)$ (equation (2.15)), about the most probable time and associated steepest descent momentum:

$$F\left(p,t\right)=F\left(p^{\sharp}_{\mathrm{mp}},t_{\mathrm{mp}}\right)+\begin{pmatrix}\frac{\partial F}{\partial p} & \frac{\partial F}{\partial t}\end{pmatrix}\begin{pmatrix}\delta p\\ \delta t\end{pmatrix}+\frac{1}{2}\left(\delta p\ \delta t\right)\begin{pmatrix}\frac{\partial^2 F}{\partial p^2} & \frac{\partial^2 F}{\partial p\partial t}\\ \frac{\partial^2 F}{\partial p\partial t} & \frac{\partial^2 F}{\partial t^2}\end{pmatrix}\begin{pmatrix}\delta p\\ \delta t\end{pmatrix}$$

$$=F\left(p^{\sharp}_{\mathrm{mp}},t_{\mathrm{mp}}\right)+\begin{pmatrix}\frac{\partial G}{\partial p}-vt, & -E\end{pmatrix}\begin{pmatrix}\delta p\\ \delta t\end{pmatrix}+\frac{1}{2}\left(\delta p\ \delta t\right)\begin{pmatrix}\frac{\partial^2 G}{\partial p^2}-\frac{t}{m\gamma^2} & -v\\ -v & 0\end{pmatrix}\begin{pmatrix}\delta p\\ \delta t\end{pmatrix}, \tag{A.1}$$

where

$$G\left(p\right)=\mathrm{i}\hbar\frac{\left(p-p_0\right)^2}{2\Gamma}+\tilde{W}\left(p\right) \tag{A.2}$$

If $p=p^{\sharp}$,

$$\frac{\partial G^{\#}}{\partial p}-v^{\#}t=0, \tag{A.3}$$

and

$$F\left(p^{\sharp},t\right)=F\left(p^{\sharp}_{\mathrm{mp}},t_{\mathrm{mp}}\right)-E^{\sharp}_{\mathrm{mp}}\delta t+\frac{1}{2}\left(\left(\frac{\partial^2 G^{\sharp}_{\mathrm{mp}}}{\partial p^2}-\frac{t_{\mathrm{mp}}}{m\gamma^2}\right)\left(\delta p^{\sharp}\right)^2-2v^{\sharp}_{\mathrm{mp}}\delta p^{\sharp}\delta t\right). \tag{A.4}$$

Expanding the former equation about the most probable time, to first order, gives

$$\left(\frac{\partial^2 G^{\sharp}_{\mathrm{mp}}}{\partial p^2}-\frac{t_{\mathrm{mp}}}{m\gamma^2}\right)\delta p^{\sharp}=v^{\sharp}_{\mathrm{mp}}\delta t, \tag{A.5}$$

which leads to equation (2.30).

ORCID iDs

Tom Rivlin https://orcid.org/0000-0002-9275-2917

References

[1] Hund F 1927 *Z. Phys.* **43** 805
[2] Mandelstam L and Leontowitsch M 1928 *Z. Phys.* **47** 131–6
[3] Tamm I E 1991 *Selected Papers* ed B M Bolotovskii and V Y Frankel (Berlin: Springer) p 273
[4] Gamow G 1928 *Z. Phys.* **51** 204
[5] Pauli W 1933 *Handbuch der Phys.* **24** 83
[6] Muga J G, Mayato R S and Egusquiza I 2008 *Time in Quantum Mechanics, Volume 1* (Lecture Notes in Physics) (Berlin: Springer) p 734
[7] Muga J G, Ruschaupt A and del Campo A 2009 *Time in Quantum Mechanics, Volume 2* (Lecture Notes in Physics) (Berlin: Springer) p 789
[8] del Campo A, García-Calderón G and Muga J G 2009 *Phys. Rep.* **476** 1–50
[9] Pablico D A L and Galapon E A 2020 *Phys. Rev.* A **101** 022103
[10] Maccone L and Sacha K 2020 *Phys. Rev. Lett.* **124** 110402
[11] Rybachenko V F 1967 *Yad. Fiz.* **5** 895
[12] Buettiker M and Landauer R 1982 *Phys. Rev. Lett.* **49** 1739
[13] Leavens C R and Aers G C 1989 *Phys. Rev.* B **39** 1202
[14] Ramos R, Spierings D, Racicot I and Steinberg A M 2020 *Nature* **583** 529–32
[15] Pollak E and Miller W H 1984 *Phys. Rev. Lett.* **53** 115
[16] Pollak E 1985 *J. Chem. Phys.* **83** 1111
[17] Landauer R and Martin T 1994 *Rev. Mod. Phys.* **66** 217
[18] Wigner E P 1955 *Phys. Rev.* **98** 145
[19] Büttiker M 1983 *Phys. Rev.* B **27** 6178
[20] Hauge E H and Støvneng J A 1989 *Rev. Mod. Phys.* **61** 917
[21] Martin T 1996 *Int. J. Mod. Phys.* B **10** 3747–63
[22] Sokolovski D and Baskin L M 1987 *Phys. Rev.* A **36** 4604
[23] Leavens C R and Aers G C 1987 *Solid State Commun.* **63** 1101–5
[24] Jauho A P and Jonson M 1989 *J. Phys.: Condens. Matter.* **1** 9027
[25] Fertig H A 1990 *Phys. Rev. Lett.* **65** 2321
[26] Gasparian V, Ortuño M, Ruiz J, Cuevas E and Pollak M 1995 *Phys. Rev.* B **51** 6743
[27] Balcou P and Dutriaux L 1997 *Phys. Rev. Lett.* **78** 851
[28] MacColl L A 1932 *Phys. Rev.* **40** 621
[29] Hartman T E 1962 *J. Appl. Phys.* **33** 3427
[30] Dumont R S and Marchioro II T L II 1993 *Phys. Rev.* A **47** 85
[31] Maji K, Mondal C K and Bhattacharyya S P 2007 *Int. Rev. Phys. Chem.* **26** 647
[32] Petersen J and Pollak E 2017 *J. Phys. Chem. Lett.* **8** 4017

[33] Petersen J and Pollak E 2018 *J. Phys. Chem.* A **122** 3563
[34] De Leo S and Rotelli P P 2007 *Eur. Phys. J.* C **51** 241
[35] De Leo S 2013 *J. Phys. A: Math. Theor.* **46** 155306
[36] del Barco O and Gasparian V 2011 *J. Phys. A: Math. Theor.* **44** 015303
[37] Thaller B 1992 *The Dirac equation* (*Texts and Monographs in Physics*) (Berlin: Springer)
[38] Low F E and Mende P F 1991 *Ann. Phys., NY* **210** 380
[39] Landauer R and Martin T 1992 *Solid State Commun.* **84** 115
[40] Aharonov Y, Oppenheim J, Popescu S, Reznik B and Unruh W G 1998 *Phys. Rev.* A **57** 4130
[41] Lozovik Y E and Filinov A V 1999 *J. Exp. Theor. Phys.* **88** 1872
[42] Mayato R S, Alonso D and Egusquiza I L 2008 Quantum clocks and stopwatches *Time in Quantum Mechanics, Volume 1* (Lecture Notes in Physics) vol 734 (Berlin: Springer) p 235
[43] Chiao R 2011 arXiv:1111.2402
[44] Ghirardi G 2013 Entanglement, nonlocality, superluminal signaling and cloning *Advances in Quantum Mechanics* ed P Bracken (Rijeka: IntechOpen) ch 24
[45] Diener G 1996 *Phys. Lett.* A **223** 327–31
[46] Winful H G 2006 *Phys. Rep.* **436** 1–69
[47] Nimtz G 2003 *Prog. Quantum Electron.* **27** 417–50
[48] Nimtz G and Stahlhofen A A 2007 arXiv:0708.0681
[49] Nimtz G and Haibel A 2002 *Ann. Phys.* **11** 163–71
[50] Chiao R Y and Steinberg A M 1997 *VI: Tunneling Times and Superluminality* (*Progress in Optics*) vol 37 ed E Wolf (Amsterdam: Elsevier) pp 345–405
[51] Aichmann H and Nimtz G 2014 *Found. Phys.* **44** 678–88
[52] Nimtz G and Aichmann H 2017 *Z. Naturforsch.* A **72** 881–4
[53] Büttiker M and Washburn S 2003 *Nature* **422** 271–2
[54] Sala R, Brouard S and Muga J G 1995 *J. Phys. A: Math. Gen.* **28** 6233
[55] Anderson M 2007 *New Sci.* **195** 10
[56] Chiao R Y, Kwiat P G and Steinberg A M 1993 *Sci. Am.* **269** 52–60
[57] Steinberg A M, Kwiat P G and Chiao R Y 1993 *Phys. Rev. Lett.* **71** 708
[58] van Dijk W, Sprung D W L and Castonguay-Page Y 2020 *Phys. Scr.* **95** 065223