Subluminality of relativistic quantum tunneling

L. Gavassino [*] and M. M. Disconzi [†]
Department of Mathematics, Vanderbilt University, Nashville, Tennessee 37211, USA

(Received 23 August 2022; revised 5 November 2022; accepted 1 March 2023; published 13 March 2023)

We prove that the classical Dirac equation in the presence of an external (nondynamical) electromagnetic field is a relativistically causal theory. As a corollary, we show that it is impossible to use quantum tunneling to transmit particles or information faster than light. When an electron tunnels through a barrier, it is bound to remain within its future light cone. In conclusion, the relativistic quantum tunneling (if modeled using the Dirac equation) is an entirely subluminal process, and it is not instantaneous.

DOI: 10.1103/PhysRevA.107.032209

I. INTRODUCTION

There is some debate over whether the speed of tunneling could be faster than the speed of light [1–9]. Some authors claim that, yes, quantum tunneling allows for superluminal signaling [1]. Other authors argue that, no, there is no superluminal propagation of particles or signals going on [2]. More recently, some authors [8] have proposed an intermediate interpretation: when a particle tunnels through a barrier, it may indeed emerge on the other side faster than light, but the probability for this to happen is so low that in practice photons arrive first, preventing an actual superluminal signaling. What has made this subject so prone to interpretation is that, if one thinks just in terms of wave packets and dispersion relations, then it is hard to unambiguously define terms like signal or tunnelling time. On the other hand, the mathematical theory of partial differential equations provides us with all the tools that are needed for us to settle this matter once and for all. This is what we aim to do here.

Let us first clarify what is mean by superluminality in this context. Consider the following thought experiment. In the reference frame of Alice, there is a sequence of light bulbs at rest, at a distance of one meter from each other. Alice has synchronized them in such a way that they all turn on simultaneously at $t_A = 1$, according to Alice's clock (see Fig. 1, left panel). Now, let us move to Bob's frame, who travels with velocity $-v$ with respect to Alice. By relativity of simultaneity [10–12], the bulbs do not turn on all together, according to Bob. Instead, they turn on in sequence, and it looks as if there was a superluminal impulse traveling at speed $v^{-1} > 1$ (we set $c = 1$), which commands the bulbs to turn on one after the other. Clearly, this illusory impulse is just an artifact of synchronization. A similar phenomenon occurs whenever the phase velocity of a wave (or the group velocity of a wave packet) is larger than the speed of light while the underlying theory is causal [13]: different regions of the system are synchronized to generate what looks to be a superluminal wave, but no actual transfer of information or energy occurs [14].

The apparent superluminality discussed above is *not* what we are interested in. Also, we are not interested in issues related to the ontology of the wave function, which may render all quantum mechanics superluminal at the outset. Consider the following example. An electron is in a quantum superposition, with $1/2$ probability of being on Earth, and $1/2$ probability of being in the Andromeda Galaxy. If we make a measurement and we detect the electron, then automatically we know that it is not in the Andromeda Galaxy. However, was the electron already on Earth, say, one second before the measurement, or did our measurement itself localize the electron on Earth? Could it be that, maybe, the electron was in the Andromeda Galaxy one second before the measurement, and then it teleported itself on Earth at the instant of the measurement? The answer depends on the interpretation of quantum mechanics one is adopting, and it is intrinsically unobservable. For this reason, we will leave this kind of problem aside.

The question that we aim to answer here is more practical: taking into account the statistical nature of quantum mechanics, is it possible to *use* quantum tunneling as a means to effectively transfer particles or signals faster than light? To make this question more precise, we have identified three rigorous practical notions of superluminality, which will be assessed here, one by one:

(1) Can the support of the wave function propagate outside the light cone? If this could happen, it would be possible to make two consecutive measurements, where the electron is first detected inside some region $\mathcal{R}$, and then outside the causal future of $\mathcal{R}$. This would mean that we can actually *observe* an electron making a superluminal jump. It is well-known that, in the absence of a potential barrier, such a process is forbidden within relativistic quantum mechanics [15]. In Sec. III, point (i), we will prove that the same is true also in the presence of a barrier.

(2) Can Alice use quantum tunneling to send a message to Bob, assuming that Bob sits outside the future light cone of Alice? There is general consensus that this should not happen, as it would constitute a violation of the principle of causality.

[*]lorenzo.gavassino@vanderbilt.edu
[†]marcelo.disconzi@vanderbilt.edu

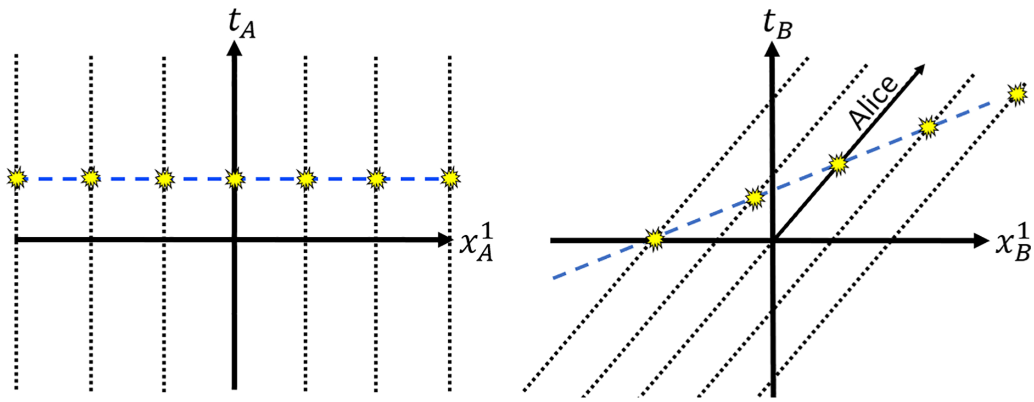

FIG. 1. Minkowski diagrams of the thought experiment outlined in the introduction. Left panel: Alice's viewpoint. The light bulbs (black dashed lines) are at rest, and they are all turned on simultaneously at $t_A = 1$ (blue dashed line). In Alice's coordinates, the space-time event at which the kth bulb is turned on is $(1, k)$, and it is marked with a yellow star. Right panel: Bob's viewpoint. The event at which the kth bulb is turned on is now $(\gamma + \gamma vk, \gamma v + \gamma k)$. These events are no longer simultaneous for different values of k, and it looks as if there was a signal traveling along the blue dashed line $x_B^1 = t_B/v - (\gamma v)^{-1}$.

However, at present there is no rigorous mathematical proof of this fundamental impossibility within tunneling models. We will provide such proof in Sec. III, point (ii).

(3) Suppose that the electron is on one side of the barrier, with probability $\mathcal{P}$. Can the probability of detecting the electron on the other side become larger than $1 - \mathcal{P}$ in less time than it would take for light to travel between the two edges of the barrier? If this were possible, we would be able to use the barrier to transfer probability (and, thus, particles) faster than light. In Sec. IV, we prove that this eventuality is indeed forbidden: probability flows subluminally between the edges of the barrier. As a corollary, quantum tunneling is not instantaneous.

Throughout the paper, we adopt the space-time signature $(-, +, +, +)$ and work in natural units: $c = \hbar = 1$. We use standard rectangular coordinates $\{x^\alpha\}_{\alpha=0}^3$ in Minkowski space $\mathbb{R}^{1+3}$, with $t := x^0$ denoting a time coordinate. Greek indices vary from 0 to 3, Latin indices from 1 to 3, and the sum convention is adopted.

II. A SIMPLE THEOREM

Our goal is to assess whether previous claims of superluminal physics (e.g., Refs. [7,8]) are mathematically rigorous. Since all such claims are derived within the framework of relativistic quantum mechanics, we will also stick to this approach (although the final word on the subject should come from quantum field theory [16]). In particular, following Dumont *et al.* [8], we will consider a single electron with quantum dynamics governed by the classical Dirac equation (we adopt the sign conventions of Weinberg [17]):

$$(\gamma^\mu \partial_\mu + ie\gamma^\mu A_\mu + m)\Psi = 0. \tag{1}$$

Here, γ^μ are Dirac's gamma matrices, $\Psi = \Psi(x)$, $x \in \mathbb{R}^{1+3}$ is a classical Dirac spinor (representing the electron [18]), e and m are the electron's charge and mass. The field $A_\mu = A_\mu(x)$, $x \in \mathbb{R}^{1+3}$, is the electromagnetic four-potential, and it is treated as a fixed, assigned, smooth function of the coordinates (it is not a dynamical degree of freedom). For tunneling models, one should take[1]

$$eA_\mu = (V, 0, 0, 0), \tag{2}$$

where $V(x)$ is the potential energy barrier. However, here we may also keep the potential A_μ completely general.

Our first task is to compute the characteristics of the system [19]. As a system of first-order partial differential equations, the Dirac equation (1) is naturally written in the standard matrix form (recall that Ψ is has four components):

$$\mathcal{M}^\mu \partial_\mu \Psi + \mathcal{N}\Psi = 0, \tag{3}$$

where $\mathcal{M}^\mu = \gamma^\mu$ and $\mathcal{N} = ie\gamma^\mu A_\mu + m$ are 4×4 complex matrices. Working in the Weyl basis, we can write $\mathcal{M}^\mu$ explicitly [20]:

$$\mathcal{M}^0 = \gamma^0 = -i\begin{bmatrix} 0_{2\times2} & \mathbb{I}_{2\times2} \\ \mathbb{I}_{2\times2} & 0_{2\times2} \end{bmatrix},$$
$$\mathcal{M}^j = \gamma^j = -i\begin{bmatrix} 0_{2\times2} & \sigma_j \\ -\sigma_j & 0_{2\times2} \end{bmatrix}, \tag{4}$$

where σ_j are the Pauli matrices:

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{5}$$

The characteristic surfaces are defined as the surfaces $\Phi =$ const of any scalar field Φ such that $\det[\mathcal{M}^\mu \xi_\mu] = 0$, with $\xi_\mu = \partial_\mu \Phi$. Using Eqs. (4), we obtain

$$\begin{aligned} 0 &= \det[\mathcal{M}^\mu \xi_\mu] \\ &= (-i)^4 \det \begin{bmatrix} 0 & 0 & \xi_0 + \xi_3 & \xi_1 - i\xi_2 \\ 0 & 0 & \xi_1 + i\xi_2 & \xi_0 - \xi_3 \\ \xi_0 - \xi_3 & -\xi_1 + i\xi_2 & 0 & 0 \\ -\xi_1 - i\xi_2 & \xi_0 + \xi_3 & 0 & 0 \end{bmatrix} \\ &= (\xi_\mu \xi^\mu)^2. \end{aligned} \tag{6}$$

[1] One should be careful about the sign: the potential energy of a particle with charge q in an electrostatic potential ϕ is $V = q\phi$. For the electron, $q = -e$. Furthermore, given that our metric signature is $(-, +, +, +)$, we have that $\phi = A^0 = -A_0$. Thus, $V = q\phi = eA_0$.

Hence, $\xi_\mu = \partial_\mu \Phi$ must be lightlike, meaning that the characteristic surfaces are null surfaces. This immediately allows us to derive the following.

Theorem 1 (Causality of the Dirac equation). Assume that A_μ is continuously differentiable and let Ψ be a continuously differentiable solution to (1). Let $\Sigma \subset \mathbb{R}^{1+3}$ be a Cauchy surface. Then, for any point x in the future of Σ, the value of Ψ at x, i.e., $\Psi(x)$, depends only on the values of Ψ in the region $J^-(x) \cap \Sigma$, and on the value of A_μ in the region $J^-(x) \cap J^+(\Sigma)$. Here, $J^-(x)$ is the causal past of x, and $J^+(\Sigma)$ is the causal future of Σ.

Remark 1. In practice, one usually takes Σ to be a surface where initial data for the system (1) is given (e.g., $\Sigma = \{t = 0\}$). In this case, the conclusion of the theorem can be rephrased in a more intuitive form as saying that the value of Ψ at x, i.e., $\Psi(x)$, depends only on the initial data in the region $J^-(x) \cap \Sigma$, and on the value of A_μ in the region $J^-(x) \cap J^+(\Sigma)$.

Proof. Fix x in the future of Σ and let Ψ_1 and Ψ_2 be two continuously differentiable solutions of the Dirac equation corresponding to two different choices of external potential. Then we have $\mathcal{M}^\mu \partial_\mu \Psi_1 + \mathcal{N}_1 \Psi_1 = 0$ and $\mathcal{M}^\mu \partial_\mu \Psi_2 + \mathcal{N}_2 \Psi_2 = 0$. Now assume that the external potential is the same on the space-time region $J^-(x) \cap J^+(\Sigma)$. Then, if we restrict our attention to such region, we have $\mathcal{N}_1 = \mathcal{N}_2$ and the field $\Psi_{\text{diff}} := \Psi_1 - \Psi_2$ is a solution of $\mathcal{M}^\mu \partial_\mu \Psi_{\text{diff}} + \mathcal{N}_1 \Psi_{\text{diff}} = 0$ on $J^-(x) \cap J^+(\Sigma)$. Finally, assume that Ψ_1 and Ψ_2 agree on $J^-(x) \cap \Sigma$. Then $\Psi_{\text{diff}} = 0$ on $J^-(x) \cap \Sigma$. At this point, we can just apply John's Global Holmgren Theorem (see Rauch [21], Sec. 1.8), considering that the characteristics of the Dirac equation are the same as those of the wave equation, and we find that $\Psi_{\text{diff}} = 0$ on $J^-(x) \cap J^+(\Sigma)$. This implies $\Psi_1(x) = \Psi_2(x)$. ■

We observe that the conclusion of Theorem 1 is coordinate independent, even if we employed standard coordinates in the computation of the characteristics. This follows from the invariance of the characteristics (see, e.g., Ref. [19]) and of $J^\pm$ [22], as well as from the standard theory of hyperbolic differential equations [23,24]. We also remark that Theorem 1 is not new. The Dirac equation is known to be a hyperbolic partial differential equation (see, e.g., Ref. [25]) and thus Theorem 1 follows from textbook theory (above we quoted Rauch [21] to provide the reader with a precise reference, but there are plenty of sources explaining the properties we used for the proof, e.g., Refs. [23–32]; see the Appendix of Ref. [33] for a summary). Nevertheless, we felt the need to state Theorem 1 and provide its proof because, as the literature review presented in the Introduction demonstrates, there seems to be some confusion in the literature regarding the causal properties of the Dirac equation. In particular, properties that follow from standard hyperbolic theory seem to be neglected in these discussions.

Theorem 1 coincides with the principle of relativistic causality that we meet in all textbooks of general relativity [22,34,35], and in the literature of relativistic hydrodynamics [12,36–42]. It is the mathematical condition that people have in mind when they say "no signal can exit the lightcone" [13] (see Fig. 2).

Let us make an interesting remark. If we set $A_\mu = 0$, then we recover the free Dirac equation. It is well-known that, in this case, Ψ is also a solution of the free Klein-Gordon equation. Therefore, it is quite trivial to see that the free Dirac equation is a relativistically causal equation [43]. However, when $A_\mu \neq 0$ (e.g., inside a potential barrier), this becomes less intuitive. The key insight, here, is that the propagation of information is *entirely* determined by the characteristics of the system, which depend only on the principal part of the Dirac equation (the part with highest derivatives: $\gamma^\mu \partial_\mu \Psi$) and are completely unaffected by the presence of the external potential A_μ. In a nutshell, the presence of a potential barrier cannot increase (or shorten) the speed of information.

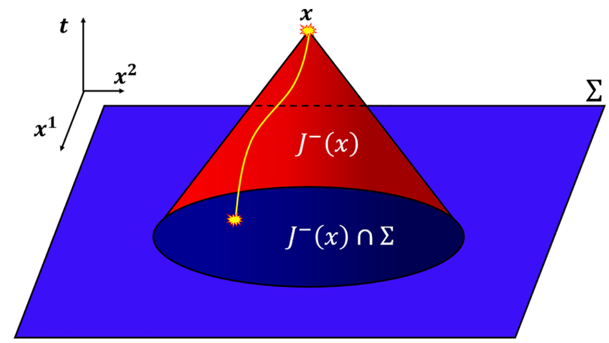

FIG. 2. Relativistic principle of causality. Let Σ (blue plane) be the initial-data hypersurface, e.g., the hyperplane $\{t = 0\}$. Pick an event x in the future of Σ. Such event can be influenced only by that portion of Σ that can be reached by a nonspacelike world line emitted from x (yellow line). In other words, the value of $\Psi(x)$ depends only on the initial data prescribed *inside* the past light cone of x (in red). Changes of initial data outside the past light cone of x cannot affect the value of $\Psi(x)$, when we solve the Dirac equation. Furthermore, we cannot change the value of $\Psi(x)$ even by altering the external potential A_μ outside the past light cone of x.

III. IMPLICATIONS

Let us apply Theorem 1 to the relativistic quantum tunneling, and let's see what we can argue on a purely mathematical basis.

(i) One direct implication of Theorem 1 is that, if $\Psi = 0$ in a region of space $\mathcal{R} \subset \Sigma$, then $\Psi = 0$ also on $\mathcal{D}^+(\mathcal{R})$, the Cauchy development of $\mathcal{R}$ [34]. This is a no-go theorem for superluminal motion: the electron cannot move faster than light because the *support* of the wave function cannot propagate outside the light cone. For example, consider the situation illustrated in Fig. 3, upper panel. A potential barrier extends over the region $0 \leqslant z \leqslant L$. At $t = 0$, the electron is on the left of such barrier with probability 1, so $\Psi = 0$ for $z > 0$.[2] Then, for arbitrary $t > 0$, Ψ must vanish in the region of space $z > t$. As a consequence, the probability for the electron to tunnel out of the barrier at a time $t < L$ is *exactly* zero.

(ii) Theorem 1 enables us also to answer the most important question: Can we use tunneling electrons to send signals faster than light? In the literature, the word "signal"

[2]Note that there is no obstruction to having wave functions that are of class C^∞, and yet they vanish for $z > 0$. The easiest way to construct them is to let Ψ decay like $\sim \exp(1/z)$, as $z \to 0^-$.

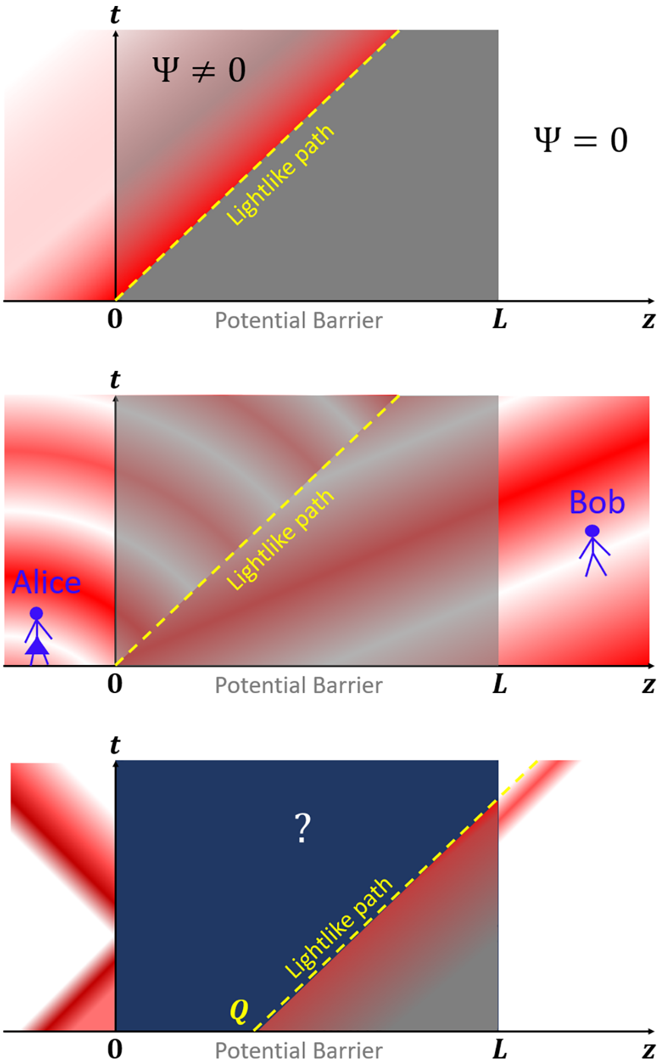

FIG. 3. Visual representations of our arguments (i)–(iii), respectively, which are direct consequences of Theorem 1. The shades of red represent the electron density $\Psi^\dagger\Psi$ (red large, white small). Upper panel: The field Ψ is bound to propagate within the lightcone. Hence, electrons cannot travel faster than light. Quantum tunneling through a barrier is no exception. Middle panel: Alice can generate a disturbance in Ψ by altering the value of A_μ at her location. However, such disturbance travels slower than light, and it cannot reach Bob, who is causally disconnected from Alice (no superluminal signaling). Lower panel: The only way for a tunneled wave packet to exit at a time $t < L$ is that $\Psi \neq 0$ on the right of Q already at $t = 0$.

often generates some debate. However, when we have a partial differential equation like (1), there is an unambiguous mathematical criterion to decide whether a superluminal signal can actually be sent or not. Consider the situation shown in Fig. 3, middle panel. Alice and Bob are on opposite sides of a potential barrier, and they are spacelike separated. The electron wave function Ψ fills the space between them. Can a decision of Alice affect a measurement of Bob? No! Alice can use a device to modify the value of the potential A_μ at her space-time location. This indeed generates a perturbation in Ψ. However, from Theorem 1, we know that changes in the value of A_μ outside the past light cone of Bob cannot affect the value of Ψ at Bob's space-time location. Thus, Alice has no way to influence Bob's measurements.[3]

(iii) When people say that the tunneling effect is superluminal, they typically have in mind the following scenario. A wave packet meets a potential barrier; most of it is reflected, but a small portion tunnels through, and it appears on the other side earlier than a hypothetical light beam emitted by the initial wave packet (see Fig. 3, lower panel). Recent findings already seem to question this picture [9], but let us say (for the sake of argument) that the idea is somehow correct. What does Theorem 1 have to say about that? Suppose that $\Psi(t = 0)$ were zero for $z > Q$ (Fig. 3, lower panel). Then, by Theorem 1, $\Psi(t)$ should vanish within the region $z > Q + t$, and there would be no tunneled wave packet. Therefore, the only way for us to observe a tunneled wave packet is to assume that Ψ was already nonzero on the right of Q at $t = 0$. In other words, to avoid a mathematical contradiction, we *must* assume that the incoming wave packet had a long tail, which extended largely inside the barrier, and that the tunneled wave packet is just the (subluminal) evolution of such long tail.

The conclusion of our point (iii) is very similar to that of Büttiker and Washburn [2]: the tunneled wave packet is the causal evolution of the right tail of the incoming wave packet, which enters the barrier much earlier than the peak, so if we only focus on the peaks, we get the illusion of a superluminal motion. Dumont *et al.* [8] have criticized this interpretation, arguing that in quantum mechanics one should never say that one piece of the wave function originates from a corresponding piece in the past. Instead, the wave function should always be treated as whole. As a consequence, according to them, we cannot say that the tunneled wave packet originates from the right tail of the incoming wave packet.

We do not wish to enter philosophical debates over the ontology of the wave function. On the other hand, we would like to point out that, when we say that the tunneled wave packet originates from the right tail, we are just making two rigorous mathematical statements (which follow from Theorem 1). First, that if you change your initial data by removing the tail, i.e., by replacing $\Psi(t = 0)$ with $\Psi(t = 0)\Theta(Q - z)$, where Θ is the Heaviside step function[4], the tunneled wave packet disappears. Second, if you instead replace $\Psi(t = 0)$ with $\Psi(t = 0)\Theta(z - Q)$, leaving only the tail and cutting all the rest, the tunneled wave packet still remains, and it is completely unaffected. These facts may not establish an ontological relationship between the tunneled wave packet and tail of the incoming wave packet, but they tell us that the *existence* of the tunneled wave packet is a direct consequence of the

[3]Another thing that Alice may do is to make a measurement herself. However, here Quantum Field Theory comes to our aid, reassuring us that spacelike-separated observables always commute, meaning that their measurements cannot influence each other [16,44,45].

[4]Readers might object that, by introducing the Heaviside function, we are no longer dealing with continusouly differentiable functions, and thus Theorem 1 no longer applies. But since the Dirac equation is a linear equation, Theorem 1 remains true for distributional solutions (which will be the case for data involving the Heaviside function), see Ref. [24], Sec. 12.5. We assumed continuous differentiability only in order to avoid technicalities and keep the proof short.

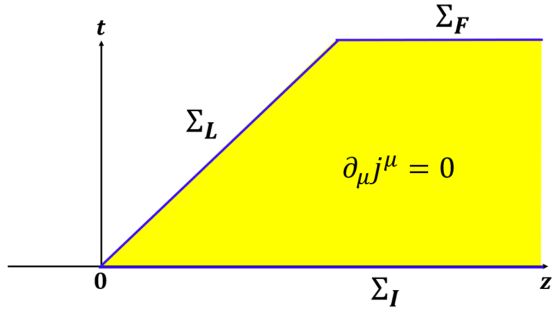

FIG. 4. Visualization of the Gauss-theorem argument discussed in Sec. IV. The yellow region represents the space-time volume where we integrate the divergence of the Dirac current (which vanishes). The hypersurfaces Σ_I and Σ_F have constant time, and thus they are spacelike. The hypersurface Σ_L is lightlike. Since the integral of $\Psi^\dagger\Psi$ across all space is normalized to 1, we know that Ψ decays to zero at space-like infinity, and thus we can extend the integration region up to $z = +\infty$. Note that we have located the down-left corner of the trapezium in the origin only for convenience. This argument still holds if we translate the trapezium anywhere else.

existence of such tail in its causal past. And this is enough to rule out any claim of superluminal behavior.

IV. SUBLUMINALITY IN AN INEQUALITY

There is a simple inequality which, we believe, will convince even the most ardent superluminalist that quantum tunneling is an entirely subluminal process. Let us consider the probability current $j^\mu = i\bar{\Psi}\gamma^\mu\Psi$, where $\bar{\Psi} = i\Psi^\dagger\gamma^0$. It can be easily shown that it has two properties [18,46–49]. First, it is conserved ($\partial_\mu j^\mu = 0$), also in the presence of an external potential A_μ. Second, it is nonspacelike, future-directed ($j^\mu j_\mu \leqslant 0$, $j^0 \geqslant 0$). In Appendix A, we verify explicitly that these properties hold also in the tunneling model of Dumont *et al.* [8]. Then, considering that Ψ must decay to zero at spacelike infinity, we can apply the Gauss theorem over the (infinitely long) trapezoidal region shown in Fig. 4, and we obtain (we adopt the orientation conventions of Misner *et al.* [50], Sec. 5)

$$-\int_{\Sigma_I} j^\mu d\Sigma_\mu + \int_{\Sigma_L} j^\mu d\Sigma_\mu + \int_{\Sigma_F} j^\mu d\Sigma_\mu = 0. \quad (7)$$

Since j^μ is nonspacelike future directed and $d\Sigma_\mu$, as a one-form, has a positive sense toward the future ("standard orientation" [50]), then the integral over Σ_L is non-negative, so[5]

$$\int_{\Sigma_F} j^\mu d\Sigma_\mu \leqslant \int_{\Sigma_I} j^\mu d\Sigma_\mu. \quad (8)$$

On the other hand, on both Σ_I and Σ_F the integrand is just $j^0 d^3x$. But $j^0(x) = \Psi^\dagger(x)\Psi(x)$ is the probability density of

[5]This inequality is analogous to Eq. (10.1.11) of Wald [34], which is used to prove well posedness and causality of the Klein-Gordon equation. There, instead of j^μ, Wald [34] uses an energy current, which is also conserved and future-directed nonspacelike. A similar theorem can also be found in Ref. [41].

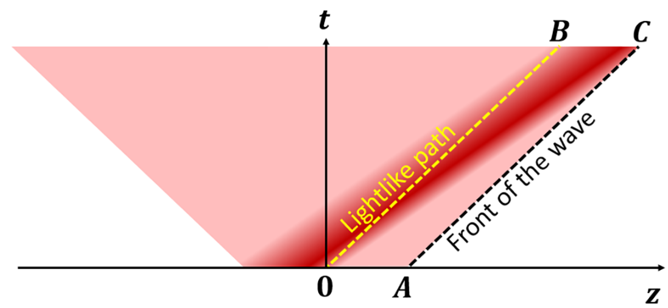

FIG. 5. A forbidden process. A wave function has support over the pink region, which expands at the speed of light. Most of the probability density j^0 is initially located to the left of the origin (darker region). Can there be a large transfer of probability (dark red beam) from the left to the right of the yellow lightlike path? According to our inequality (9), no. In fact, the probability stored in the segment BC can never exceed the probability stored in its nonzero causal past (the segment $0A$).

observing the electron at x. Therefore, the inequality (8) reduces the following constraint:

$$\mathcal{P}_t(z > t) \leqslant \mathcal{P}_0(z > 0). \quad (9)$$

In words: The probability of observing the electron on the right of $z = t$ at time t will *never* exceed the initial probability of observing that same electron on the right of $z = 0$ at time 0.

To better understand the meaning of this result, consider the Minkowski diagram in Fig. 5. The idea is the following. In Sec. III, implication (i), we proved that the front (i.e., the boundary) of the wave packet cannot travel faster than light. However, one could argue that, even if the front itself is luminal, perhaps the main body of the wave function can still drift superluminally *within* the support of the wave packet, transiting from the left front to the right front faster than light, as in Fig. 5 (dark red beam). Our inequality (9) forbids also this eventuality. In fact, a superluminal transfer of probability from the left to the right front would entail an increase of probability in a neighborhood of the right front (i.e., on the segment BC in Fig. 5), and this would constitute a violation of (9). Note that a similar result would hold also in the absence of fronts (e.g., for Gaussian wave packets): the probability on the right of B at time $t(B)$ cannot be larger than the probability on the right of 0 at time $t = 0$.

Let us see other implications of Eq. (9).

A. The tunneling probability comes from the causal past of the tunneled packet

Let us apply Eq. (9) to the setting discussed in point (iii) of Sec. III. It is immediate to see that we can transport our Gauss-theorem argument of Fig. 4 into Fig. 3 (lower panel), locating the lower-left corner of the trapezium at Q and overlapping Σ_L with the lightlike path shown in Fig. 3. Then, the inequality (9) becomes

$$\mathcal{P}(\text{tunneled packet}) \leqslant \mathcal{P}(\text{tail on the right of } Q). \quad (10)$$

Again, this is showing us that the emerging wave packet is just the subluminal evolution of the right tail of the incoming wave packet. But now we also know something more: the probability associated to the tunneled wave packet cannot exceed that of this initial tail. This means that the tail cannot be used as springboard or a means to push the electron through the barrier faster than light. No way. The probability stored in the tail is the maximum probability that the tunneled wave packet can carry.

There is a subtlety that we need to mention. In Fig. 3 (lower panel), the point Q falls inside the barrier (i.e., $Q > 0$). But this is true only if we set our clocks in such a way that at $t = 0$ the incoming wave packet is about to enter the barrier. In numerical experiments like the one performed by Dumont *et al.* [8], the wave packet is on the far left of the barrier at $t = 0$. In this case, Eq. (10) still holds, but point Q will also be on the far left of the barrier ($Q \ll 0$). In Appendix B, we calculate the position of Q for the numerical experiment of Dumont *et al.* [8], and we verify explicitly that their numerical analysis corroborates Eq. (10).

B. Luminal bound on the speed of tunneling

We are finally able to prove that quantum tunneling is not instantaneous, and that its speed is bounded above by the speed of light. Our proof works as follows.

Suppose that, at time $t = 0$, the electron is on the left of the barrier and it is about to tunnel through. The left edge of the barrier is located at $z = 0$ (as usual), while the right edge is at $z = L$. To keep our discussion completely general, we allow for a little portion of the electron's wave function to have already leaked inside the barrier. Hence, we just assume that the electron is on the left of the barrier with some probability $\mathcal{P} \leqslant 1$. Thus, we have the initial condition $\mathcal{P}_0(z \leqslant 0) = \mathcal{P}$, where the subscript 0 means that we are evaluating the probability at $t = 0$. Since the total probability is normalized to 1, we also know that $\mathcal{P}_0(z > 0) = 1 - \mathcal{P} =$ leaked probability. Plugging this initial condition into Eq. (9), we find that

$$1 - \mathcal{P} \geqslant \mathcal{P}_t(z > t). \tag{11}$$

Now, let us set $t < L$. Then, the probability $\mathcal{P}_t(z > t)$ cannot be smaller than the probability $\mathcal{P}_t(z > L)$:

$$\mathcal{P}_t(z > t) = \mathcal{P}_t(L > z > t) + \mathcal{P}_t(z > L) \geqslant \mathcal{P}_t(z > L). \tag{12}$$

Combining (11) and (12), we arrive at the inequality $1 - \mathcal{P} \geqslant \mathcal{P}_t(z > L)$ for $t < L$. On the other hand, $\mathcal{P}_t(z > L)$ is the probability of detecting the electron on the right of the barrier at time t. But this is just the tunneling probability at time t. In conclusion, we have that

$$\mathcal{P}_t(\text{tunneling}) \leqslant \mathcal{P}_0(\text{leaked}) \qquad \text{for } t < L. \tag{13}$$

In a nutshell, this is telling us that the best we can get in a time $t < L$ is that the part of the wave function that is already inside the barrier (at $t = 0$) will emerge on the right. But nothing more than this. If this leaked tail is negligible, then the tunneled wave packet cannot emerge in a time smaller than L/c (restoring nongeometric units).

V. SUBLUMINALITY AS AN ALGEBRAIC IDENTITY

All our analysis till this point has been carried out with explicit reference to the wave function Ψ. It is natural to wonder whether we can also express our results using Dirac's bra-ket notation. This is what we aim to do here. For clarity, we switch to $1+1$ dimensions and we adopt rectangular coordinates (t, z).

Since in the bra-ket notation one only deals with quantum states $|\Psi\rangle$, with no explicit reference to space-time events and locations, we need to first express Theorem 1 in a slightly different way. Our reasoning is the following. Since the Dirac equation is linear, we can always express a solution Ψ as the superposition of two other solutions, $\Psi = \Psi_L + \Psi_R$, provided that the initial data for Ψ_L and Ψ_R add up to the initial data of Ψ, namely, $\Psi(0, z) = \Psi_L(0, z) + \Psi_R(0, z)$. We choose for these two solutions the following initial data: $\Psi_L(0, z) = \Psi(0, z)\Theta(-z)$ and $\Psi_R(0, z) = \Psi(0, z)\Theta(z)$, which clearly add up to $\Psi(0, z)$. On the other hand, our theorem (which holds also for distributional solutions, see footnote 4) guarantees that, since $\Psi_L(0, z) = 0$ for $z > 0$, then $\Psi_L(t, z) = 0$ for $z > t$. This implies that

$$\Psi(t, z) = \Psi_R(t, z) \quad \text{for } z > t. \tag{14}$$

This clearly shows that the part of the wave function that is initially (at $t = 0$) in $z < 0$ cannot travel into the region $z > t$. Equivalently, the part of the wave function that enters the region $z > t$ is the causal evolution of $\Psi_R(0, z)$, namely, the portion of the initial wave function that was in the causal past of the region $z > t$.

Let us now switch to bra-ket notation. If we work in the Schrödinger picture, the initial wave functions $\Psi(0, z)$, $\Psi_L(0, z)$, $\Psi_R(0, z)$ correspond to three different quantum states, which may be represented by three corresponding state vectors: $|\Psi\rangle$, $|L\rangle$, and $|R\rangle$. The first state vector is normalized, $\langle\Psi|\Psi\rangle = 1$, while

$$\begin{aligned}
\langle L|L\rangle &= \int_{\mathbb{R}} \Psi^\dagger(0, z)\Psi(0, z)\Theta(-z)dz = \mathcal{P}_0(z < 0), \\
\langle R|R\rangle &= \int_{\mathbb{R}} \Psi^\dagger(0, z)\Psi(0, z)\Theta(z)dz = \mathcal{P}_0(z > 0), \\
\langle L|R\rangle &= \int_{\mathbb{R}} \Psi^\dagger(0, z)\Psi(0, z)\Theta(z)\Theta(-z)dz = 0.
\end{aligned} \tag{15}$$

Clearly, the initial condition $\Psi(0, z) = \Psi_L(0, z) + \Psi_R(0, z)$ is equivalent to $|\Psi\rangle = |L\rangle + |R\rangle$. Then, the condition $\Psi = \Psi_L + \Psi_R$ just expresses the fact that unitary time evolution (in the Schrödinger picture) is linear,

$$e^{-i\hat{H}t}|\Psi\rangle = e^{-i\hat{H}t}|L\rangle + e^{-i\hat{H}t}|R\rangle, \tag{16}$$

where $\hat{H}$ is the Hamiltonian which generates the dynamics of (1). Finally, the fact that $\Psi_L(t, z) = 0$ for $z > t$ translates into the condition $\hat{\mathcal{P}}(z > t)e^{-i\hat{H}t}|L\rangle = 0$, where $\hat{\mathcal{P}}(z > t)$ is the orthogonal projector[6] onto the region $z > t$. Therefore, if we apply $\hat{\mathcal{P}}(z > t)$ on both sides of (16) we recover Eq. (14), namely,

$$\hat{\mathcal{P}}(z > t)e^{-i\hat{H}t}|\Psi\rangle = \hat{\mathcal{P}}(z > t)e^{-i\hat{H}t}|R\rangle. \tag{17}$$

[6]The orthogonal projector $\hat{\mathcal{P}}(z > Q)$ maps an arbitrary wave function Ψ into $\Theta(z - Q)\Psi$, while the orthogonal projector $\hat{\mathcal{P}}(z < Q)$ maps an arbitrary wave function Ψ into $\Theta(Q - z)\Psi$.

Equation (17) expresses the subluminality of the Dirac equation, both in the presence and in the absence of a potential barrier. In fact, it tells us that, if we perform a measurement inside the region $z > t$, there are no contributions coming from $|L\rangle$. In other words, all probabilities computed inside the region $z > t$ depend only on state $|R\rangle$, which describes that part of the wave function which was in the causal past of such region. Again, this shows that relativistic quantum tunneling is an entirely subluminal process. This conclusion is also corroborated by the observations below:

(1) If we take the norm of Eq. (17), we obtain

$$\langle\Psi|\, e^{i\hat{H}t}\hat{\mathcal{P}}(z>t)e^{-i\hat{H}t}\,|\Psi\rangle = \langle R|\, e^{i\hat{H}t}\hat{\mathcal{P}}(z>t)e^{-i\hat{H}t}\,|R\rangle \leqslant \langle R|\, e^{i\hat{H}t}e^{-i\hat{H}t}\,|R\rangle = \langle R|R\rangle\,. \tag{18}$$

On the other hand, the quantity $\langle\Psi|\, e^{i\hat{H}t}\hat{\mathcal{P}}(z>t)e^{-i\hat{H}t}\,|\Psi\rangle$ is just the probability of observing the electron in the region $z > t$ at time t, namely, $\mathcal{P}_t(z > t)$. Hence, recalling the second equation of (15), we recover the inequality (9). The interpretation is simple: the probability stored in $z > t$ at time t comes only from $|R\rangle$, and hence cannot exceed $\langle R|R\rangle = \mathcal{P}_0(z > 0)$.

(2) It is evident that $|L\rangle = \hat{\mathcal{P}}(z < 0)\,|\Psi\rangle$. Hence, the condition $\hat{\mathcal{P}}(z>t)e^{-i\hat{H}t}\,|L\rangle = 0$, which is valid for any $|\Psi\rangle$, can be expressed as an operatorial identity:

$$\hat{\mathcal{P}}(z>t)e^{-i\hat{H}t}\hat{\mathcal{P}}(z<0) = 0. \tag{19}$$

In words: No contribution coming from $z < 0$ can reach the region $z > t$ in a time t, independently from the presence of a potential barrier.

Let us make one final remark. In the analysis above, we have expressed the initial state $|\Psi\rangle$ as the quantum superposition of two other states, $|L\rangle$ and $|R\rangle$. Dumont *et al.* [8] have criticized this kind of approach. Here we report their reasoning [8]: "This argument appears to imply that we are actually able to track individual components of the initial wave packet ... and identify parts of the final wavepacket as having come from the front or back of the initial distribution.... 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."

In response to this, we would like to remark that the principle of superposition is a defining feature of quantum mechanics, and the ability to track individual parts of a wave packet follows directly from the linearity property of the unitary evolution, see Eq. (16). Equation (17), then, only expresses the fact that the outcome of a measurement performed inside the space-time region $z > t$ depends only on $|R\rangle$, which is the projection of $|\Psi\rangle$ into the causal past of such region. This is an observable physical fact. No reference to interpretations of quantum mechanics is needed.

VI. CONCLUSIONS

General relativity [22,34,35] and the theory of partial differential equations [19,21] provide us with all the machinery necessary to assess the speed of a physical process, and the propagation of information within a given theory. Here, we have applied these techniques to relativistic quantum tunneling, modeled using the classical Dirac equation in a background electromagnetic field. This has allowed us to establish rigorously three mathematical facts.

(1) It is impossible to use quantum tunneling to send information faster than light [see Sec. III, point (ii)]. In fact, if an observer (Alice) perturbs the wave function Ψ at a point, the perturbation is bound to travel inside the light cone. This implies that, if another observer (Bob) sits somewhere outside the light cone, he cannot know that Alice has acted on the electron because the perturbation cannot reach him. Bob has no way to tell whether Alice perturbed the wave function or not. We would like to stress that this result is completely independent from any bound on the speed of tunneling. In fact, this theorem just tells us that, if quantum tunneling were superluminal, Alice would have absolutely no influence on that part of the wave function that exits her future light cone. She would not be able to control it in any way. It would be impossible for her to manipulate its shape or even to stop it. The very fact that a superluminal wave packet reaches Bob would be independent from Alice's decisions. So, even if quantum tunneling were superluminal, it would be impossible to use it to send *information*, because Bob would have no way to infer what Alice has done.

(2) Relativistic quantum tunneling is not instantaneous. Instead, its speed is bounded by the speed of light. In particular, if at a given time the electron is on the left of the barrier with probability $\mathcal{P}$, then we need to wait at least a time L/c (length of barrier/speed of light) before the probability of having the electron on the right of the barrier can become larger than $1 - \mathcal{P}$ [see Sec. IV B]. This guarantees that the probability flows subluminally between the two edges of the barrier. As a particular case, if the electron is on the left of the barrier with probability 1, we need to wait a time L/c before it can emerge on the right with nonzero probability [see Sec. III, point (i)].

(3) If a photon is traveling toward the right, the probability of observing an electron on the right of such photon can only decrease in time (or stay constant). In other words, photons always overtake electrons. The reversal cannot happen, even during quantum tunneling. This is shown in Fig. 5.

Hence, we believe that this paper has finally settled a 20-year-old debate. The Dirac equation is a perfectly causal field equation, also when we turn on an extremely high potential barrier. That is because the electromagnetic four-potential $A_\mu(x)$ does not enter the equation of the characteristics. Indeed, recent numerical tests [9] confirm our main message: tunneling electrons cannot be faster than photons in vacuum. Interestingly, if we assume that a photon and an electron start with the same initial distribution, then our inequality (9) can be equivalently rewritten in the form[7]

$$\mathcal{P}_t(\text{arrival electron}) \leqslant \mathcal{P}_t(\text{arrival photon}), \tag{20}$$

which is precisely what has been observed in all tests performed by Dumont and Rivlin [9].

[7]To show this, just consider one electron and one photon with same initial probability distributions, and use the fact that, for photons, the inequality (9) is saturated.

We hope that our work will also foster the interdisciplinary communication between the quantum physics community and the mathematical relativity community.

ACKNOWLEDGMENTS

M.M.D. is partially supported by a Sloan Research Fellowship provided by the Alfred P. Sloan foundation, NSF Grant No. DMS-2107701, and a Vanderbilt's Seeding Success Grant. L.G. is partially supported by a Vanderbilt's Seeding Success Grant. The authors would like to thank the anonymous referees for several comments that led to a significant improvement of the paper over its first version.

APPENDIX A: PROBABILITY CURRENT FOR TUNNELING MODELS

The tunneling model of Dumont *et al.* [8] is (1+1)-dimensional, and it evolves only two components of Ψ in the Dirac basis (as the other two components are fully decoupled). We call such components f and h. The Dirac equation then reads

$$\begin{aligned} i\partial_t f &= i\partial_z h + (V+m)f \\ i\partial_t h &= i\partial_z f + (V-m)h, \end{aligned} \tag{A1}$$

where $V = V(x)$ is the potential barrier. Taking the complex conjugate, we get

$$\begin{aligned} i\partial_t f^* &= i\partial_z h^* - (V+m)f^* \\ i\partial_t h^* &= i\partial_z f^* - (V-m)h^*. \end{aligned} \tag{A2}$$

Thus, it is immediate to verify that

$$\begin{aligned} \partial_t(f^*f) &= f^*\partial_z h + f\partial_z h^* \\ \partial_t(h^*h) &= h^*\partial_z f + h\partial_z f^*. \end{aligned} \tag{A3}$$

As we can see, all the terms with V cancel out. Taking the sum of these two equations, and bringing every term on the left-hand side, we obtain an equation of the form $\partial_\mu j^\mu = 0$, with

$$j^\mu = \begin{pmatrix} f^*f + h^*h \\ -f^*h - h^*f \end{pmatrix}. \tag{A4}$$

As we can see, $j^0 = f^*f + h^*h$ is non-negative definite, and it has the usual form of a probability density. To prove that j^μ is nonspacelike future-directed, we only need to show that $j^0 \geqslant |j^z|$, namely, $f^*f + h^*h \geqslant |f^*h + h^*f|$. But this follows immediately from the chain of identities below:

$$\begin{aligned} 0 \leqslant |f \pm h|^2 &= (f^* \pm h^*)(f \pm h) \\ &= f^*f + h^*h \pm (f^*h + h^*f). \end{aligned} \tag{A5}$$

1. Application: Superluminal interference fringes do not transport probability

Some solutions of the Dirac equation can exhibit interference fringes whose phase velocity is larger than the speed of light [51]. We know from Theorem 1 that such fringes cannot be used to transport information faster than light. Now we are also in a position to show that they cannot even be used to transport the electron itself (namely, the probability density) faster than light. The proof is very simple. We have shown that the probability satisfies a continuity equation of the form $\partial_t j^0 + \partial_z j^z = 0$ (in 1+1 dimensions, for simplicity). If we define the velocity $v := j^z/j^0$ (probability flux/probability density), the continuity equation acquires the usual form from hydrodynamics textbooks [52]: $\partial_t j^0 + \partial_z(j^0 v) = 0$. This justifies the interpretation of v as the velocity of probability, since it quantifies how fast a given probability density is crossing the boundary of a certain region of space:

$$\frac{d}{dt}\mathcal{P}_t(z<a) = \frac{d}{dt}\int_{-\infty}^{a} j^0(t,z)dz = -j^0(t,a)v(t,a). \tag{A6}$$

On the other hand, we have shown that $|j^z| \leqslant j^0$, which implies $|v| \leqslant 1$. Hence, the speed of probability can never exceed the speed of light. This completes our proof.

It is instructive to analyze a concrete example. To simplify the calculations, let us set $m = V = 0$ in Eq. (A1). Then, it is immediate to see that

$$\begin{aligned} f(t,z) &= a(z+t) + b(z-t) \\ h(t,z) &= a(z+t) - b(z-t) \end{aligned} \tag{A7}$$

is a solution of the massless Dirac equation, for any couple of complex functions a and b. Equation (A7) describes a quantum superposition of a left-travelling state a and a right-travelling state b. It is immediate to verify that

$$\begin{aligned} f^*f &= a^*a + a^*b + b^*a + b^*b, \\ h^*h &= a^*a - a^*b - b^*a + b^*b, \\ f^*h &= a^*a - a^*b + b^*a - b^*b, \\ h^*f &= a^*a + a^*b - b^*a - b^*b, \end{aligned} \tag{A8}$$

where it is understood that all functions a are evaluated at $z+t$, and all functions b are evaluated at $z-t$. If we plug the above formulas into (A4), we finally obtain a formula for the velocity of probability:

$$v = \frac{b^*b - a^*a}{b^*b + a^*a} \in [-1, 1]. \tag{A9}$$

Now, let us set $a(z) = \exp(ikz)$ and $b(z) = \exp(ipz)$, with $p > k > 0$. Then, the first line of (A8) becomes

$$f^*f = 4\cos^2\left[\frac{p-k}{2}\left(z - \frac{p+k}{p-k}t\right)\right]. \tag{A10}$$

As we can see, the wave function exhibits an interference fringe, which drifts with phase velocity $(p+k)/(p-k) > 1$. On the other hand, since $a^*a = b^*b = 1$, the probability has vanishing net velocity, i.e., $v = 0$. This result is completely analogous to what we see in hydrodynamics: a sound wave travels at the speed of sound, but this does not mean that matter itself is transported at the speed of sound. Indeed, during the passage of a sound wave, the matter elements oscillate, but on average they do not move. It is the same here. The phase velocity of an interference fringe can exceed the speed of light, but such fringe does not transport probability. The particle is, on average, at rest.

APPENDIX B: A NUMERICAL EXPERIMENT

Here we show that the numerical experiment of Dumont *et al.* [8] corroborates our Eq. (10). We choose our units of space, time, and energy in such a way that $c = \hbar = \lambda_{rC} = 1$,

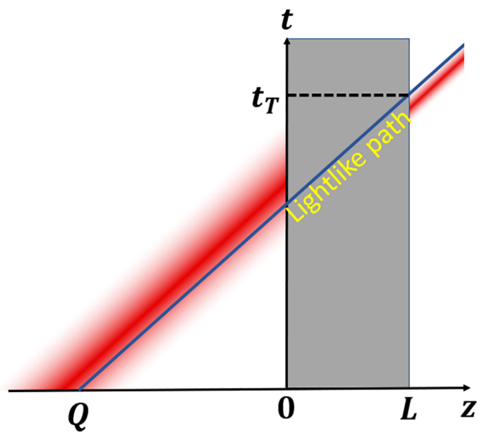

FIG. 6. Location of point Q in the numerical experiment of Dumont *et al.* [8]. The mathematical procedure for computing it is the following. First step: Identify the position of the tunneled wave packet in the Minkowski diagram (red "beam" in the upper right corner). Second step: Draw a right-travelling lightlike path (in blue) located immediately on the left of the tunelled wavepacket. Third step: find the location Q where this lightlike path intersects the line $t = 0$. The region $z > Q$ defines the "tail" of the incoming wave function in the causal past of the tunneled wavepacket. Equation (10) states that the tunneling probability cannot exceed the probability associated to such tail (because the tunneled wave packet is the causal evolution of such tail). In other words, the probability flows *subluminally* from the tail to the tunneled wave packet. We note that, since the incoming wave packet is very far from the barrier at $t = 0$, the point Q that marks the beginning of the tail lies outside the barrier ($Q < 0$).

where λ_{rC} is the reduced Compton wavelength of the electron. Furthermore, we set the origin of the coordinate z to coincide with the left edge of the barrier.

The incoming wave packet is initially centered around $z_0 = -120$. The barrier starts at $z = 0$ and ends at $z = L = 15$ (we consider the most superluminal case). The tunneled wave packet emerges at a time $t_T \approx 110$ (see Fig. 1.c of Ref. [8]). Therefore, point Q is just (see Fig. 6)

$$Q = L - t_T = -95. \tag{B1}$$

This is telling us that the tail of the wave function that lies inside the causal past of the tunneled wave packet covers the region $z > -95$. If our Eq. (10) is correct, the probability associated to such a tail should not be smaller than the probability associated to the tunneled wave packet. Now, the initial probability distribution is essentially a Gaussian,

$$j^0(0, z) \approx \frac{1}{\sqrt{\pi}\,\Delta z^2} \exp\left[-\frac{(z - z_0)^2}{\Delta z^2} \right], \tag{B2}$$

with width $\Delta z = 15$. Integrating it for $z > Q$, we get

$$\mathcal{P}_0(z > -95) \approx 0.009. \tag{B3}$$

This is a small tail, as it encompasses only $\sim 1\%$ of the incoming wave packet, but it is much larger than $\mathcal{P}(\text{tunneled packet}) \sim 10^{-15}$ (see Fig. 1.c of Dumont *et al.* [8]). This corroborates Eq. (10) and confirms that no superluminal flow of probability has taken place. The probability associated to the tunneled wave packet never exceeds the probability that was already contained in its causal past.

[1] G. Nimtz and A. Haibel, Ann. Phys. (Leipzig) **11**, 163 (2002).
[2] M. Büttiker and S. Washburn, Nature (London) **422**, 271 (2003).
[3] H. G. Winful, Phys. Rev. Lett. **90**, 023901 (2003).
[4] M. Büttiker and S. Washburn, Nature (London) **424**, 638 (2003).
[5] H. G. Winful, Phys. Rep. **436**, 1 (2006).
[6] O. del Barco and V. Gasparian, J. Phys. A: Math. Theor. **44**, 015303 (2011).
[7] S. De Leo, J. Phys. A: Math. Theor. **46**, 155306 (2013).
[8] R. S. Dumont, T. Rivlin, and E. Pollak, New J. Phys. **22**, 093060 (2020).
[9] R. S. Dumont and T. Rivlin, arXiv:2205.15375.
[10] E. Gourgoulhon, *Special Relativity in General Frames: From Particles to Astrophysics*, 1st ed., Graduate Texts in Physics (Springer-Verlag, Berlin, 2013).
[11] L. Gavassino and F. Giacosa, Phys. Rev. A **106**, 042215 (2022).
[12] L. Gavassino, Phys. Rev. X **12**, 041001 (2022).
[13] Y. Aharonov, A. Komar, and L. Susskind, Phys. Rev. **182**, 1400 (1969).
[14] R. L. Tolman, *The Theory of the Relativity of Motion* (University of California Press, Berkeley, 1917).
[15] B. Thaller, *The Dirac Equation* (Springer-Werlag, Heidelberg, 1992).
[16] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory* (Addison-Wesley, Reading, 1995).
[17] S. Weinberg, *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity* (John Wiley and Sons, New York, 1972).
[18] L. D. Landau and E. M. Lifshitz, *Relativistic Quantum Theory*, Course of Theoretical Physics, Vol. 4 (Pergamon Press, New York, 1959).
[19] R. Courant and D. Hilbert, *Partial Differential Equations*, Methods of Mathematical Physics, Vol. 2 (John Wiley and Sons, New York, 1989).
[20] S. Weinberg, *The Quantum Theory of Fields* (Cambridge University Press, Cambridge, 1995), Vol. 1.
[21] J. Rauch, *Partial Differential Equations*, Graduate Texts in Mathematics (Springer, New York, 1991).
[22] S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Space-Time*, Cambridge Monographs on Mathematical Physics (Cambridge University Press, Cambridge, 2011).
[23] J. Leray, *Hyperbolic Differential Equations* (Institute for Advanced Study, Princeton, N.J., 1953), p. 240.
[24] L. Hörmander, *The Analysis of Linear Partial Differential Operators. III*, Classics in Mathematics (Springer, Berlin, 2007), pp. viii and 525, pseudo-differential operators, Reprint of the 1994 edition.
[25] Y. Choquet-Bruhat, *General Relativity and the Einstein equations*, Oxford Mathematical Monographs (Oxford University Press, Oxford, 2009), pp. xxvi and 785.

[26] A. Lichnerowicz, *Relativistic Hydrodynamics and Magnetohydrodynamics: Lectures on the Existence of Solutions* (W. A. Benjamin, New York, 1967).

[27] F. John, *Partial Differential Equations*, 4th ed., Applied Mathematical Sciences, Vol. 1 (Springer-Verlag, New York, 1982), pp. x and 249.

[28] K. Kajitani and T. Nishitani, *The Hyperbolic Cauchy problem*, Lecture Notes in Mathematics, Vol. 1505 (Springer-Verlag, Berlin, 1991), pp. viii and 168.

[29] S. Mizohata, *On the Cauchy Problem*, Notes and Reports in Mathematics in Science and Engineering, Vol. 3 (Academic Press, Inc., Orlando, FL; Science Press Beijing, Beijing, 1985), pp. vi and 177.

[30] F. Trèves, *Basic Linear Partial Differential Equations*, Pure and Applied Mathematics, Vol. 62 (Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London, 1975) pp. xvii and 470.

[31] S. L. Sobolev, *Some Applications of Functional Analysis in Mathematical Physics*, Translations of Mathematical Monographs, Vol. 90 (American Mathematical Society, Providence, RI, 1991), pp. viii and 286, translated from the third Russian edition by Harold H. McFaden, With comments by V. P. Palamodov.

[32] M. E. Taylor, *Partial differential equations. I*, Applied Mathematical Sciences, Vol. 115 (Springer-Verlag, New York, 1996), pp. xxiv and 563, basic theory.

[33] M. M. Disconzi, Commun. Pure Appl. Anal. **18**, 1567 (2019).

[34] R. M. Wald, *General relativity* (Chicago University Press, Chicago, IL, 1984).

[35] S. M. Carroll, *Spacetime and Geometry: An Introduction to General Relativity* (Pearson, Harlow, 2014).

[36] F. S. Bemfica, M. M. Disconzi, and J. Noronha, Phys. Rev. D **98**, 104064 (2018).

[37] F. S. Bemfica, M. M. Disconzi, and J. Noronha, Phys. Rev. D **100**, 104020 (2019).

[38] F. S. Bemfica, M. M. Disconzi, and J. Noronha, Phys. Rev. Lett. **122**, 221602 (2019).

[39] F. S. Bemfica, M. M. Disconzi, V. Hoang, J. Noronha, and M. Radosz, Phys. Rev. Lett. **126**, 222301 (2021).

[40] F. S. Bemfica, M. M. Disconzi, and J. Noronha, Phys. Rev. X **12**, 021044 (2022).

[41] L. Gavassino, M. Antonelli, and B. Haskell, Phys. Rev. Lett. **128**, 010606 (2022).

[42] L. Gavassino, Class. Quantum Grav. **39** 185008 (2022).

[43] R. Fox, C. G. Kuper, and S. G. Lipson, Proc. R. Soc. London A **316**, 515 (1970).

[44] P. H. Eberhard and R. R. Ross, Found. Phys. Lett. **2**, 127 (1989).

[45] S. Coleman, *Lectures of Sidney Coleman on Quantum Field Theory*, edited by B. G.-g. Chen, D. Derbes, D. Griffiths, B. Hill, R. Sohn, and Y.-S. Ting (WSP, Hackensack, 2018).

[46] J. J. Sakurai, *Advanced Quantum Mechanics* (Addison-Wesley Publishing Company, Menlo Park, 1967).

[47] P. A. M. Dirac, *The Principles of Quantum Mechanics* (Oxford University Press, London, 1947).

[48] F. Finster, J. Smoller, and S.-T. Yau, arXiv:gr-qc/9910030.

[49] F. Finster, J. Smoller, and S.-T. Yau, arXiv:gr-qc/0211043.

[50] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (W.H. Freeman and Co., San Francisco, 1970).

[51] M. V. Berry, Eur. J. Phys. **33**, 279 (2012).

[52] L. Landau and E. Lifshitz, *Fluid Mechanics* (Pergamon Press, Oxford, 1987), Vol. 6.