Dispersion Relations Alone Cannot Guarantee Causality

L. Gavassino,[1] M. Disconzi,[1] and J. Noronha[2]

[1]*Department of Mathematics, Vanderbilt University, 1326 Stevenson Center Lane, Nashville, Tennessee 37240, USA*
[2]*Illinois Center for Advanced Studies of the Universe & Department of Physics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801-3003, USA*

We show that linear superpositions of plane waves involving a single-valued, covariantly stable dispersion relation $\omega(k)$ always propagate outside the light cone unless $\omega(k) = a + bk$. This implies that there is no notion of causality for individual dispersion relations since no mathematical condition on the function $\omega(k)$ (such as the front velocity or the asymptotic group velocity conditions) can serve as a sufficient condition for subluminal propagation in dispersive media. Instead, causality can only emerge from a careful cancellation that occurs when one superimposes all the excitation branches of a physical model. This happens automatically in local theories of matter that are covariantly stable. Hence, we find that the need for nonhydrodynamic modes in relativistic fluid mechanics is analogous to the need for antiparticles in relativistic quantum mechanics.

Introduction.—The "practical" definition of relativistic causality is universally accepted: it is impossible to transmit information faster than the vacuum speed of light [1–3]. The question is how to translate this principle into mathematical constraints that our physical theories must obey. In some cases, this question has an unambiguous answer. In classical field theory, causality demands that the characteristics of the field equations lay inside or upon the light cone [4–8]. In quantum field theory, the commutator of spacelike-separated observables must vanish [9–11]. In other contexts, the mathematical nature of causality is less understood.

Consider a homogeneous system in thermodynamic equilibrium, and let $\omega(k)$ be the eigenfrequency of one of its (linear) excitations, which is a function of the wave number k [12]. Under which conditions is such dispersion relation compatible with causality? Most attempted answers revolve around imposing inequalities on the phase velocity $(\mathfrak{Re}\omega)/k$, or on the group velocity $d(\mathfrak{Re}\omega)/dk$ [13–15]. However, no fully consistent and universally reliable criterion has been found. The most widely accepted constraint is that the "front velocity" [16,17], or the "asymptotic group velocity" [18], both of which usually coincide by L'Hopital's rule, should not exceed the speed of light c ($= 1$, in our units), namely

$$v_f = \lim_{k \to \infty, k \in \mathbb{R}} \frac{\mathfrak{Re}\omega}{k} \overset{\text{L'H}}{=} \lim_{k \to \infty, k \in \mathbb{R}} \frac{d\mathfrak{Re}\omega}{dk} \in [-1, 1]. \quad (1)$$

Unfortunately, this condition is far from satisfactory as many famous acausal equations in physics fulfill (1) even if the theory of partial differential equations tells us that they propagate information at infinite speeds [4,8]. Three notable examples are the diffusion equation, the Euclidean wave equation, and the linearized Benjamin-Bona-Mahony (BBM) [19] equation, respectively:

$$(\partial_t - \partial_x^2)\varphi = 0 \Rightarrow \omega = -ik^2,$$
$$(\partial_t^2 + \partial_x^2)\varphi = 0 \Rightarrow \omega = \pm ik,$$
$$(\partial_t + \partial_x - \partial_t \partial_x^2)\varphi = 0 \Rightarrow \omega = \frac{k}{1+k^2}. \quad (2)$$

All these equations have $v_f = 0$. Furthermore, their phase and group velocities are (sub)luminal for all k. Nevertheless, these three models are strongly acausal. The BBM equation is particularly striking because one cannot attribute the causality violation to the imaginary part of ω, given that ω is real for real k. Yet, it is acausal as the lines $t = \text{const}$ are spacelike characteristics [20].

This Letter shows that the limitations of (1) are manifestations of a fundamental impossibility. Namely, unless $\omega(k) = a + bk$ for all k (with a, b constant), a single dispersion relation $\omega(k)$ cannot be causal. Rather, "causality" is a collective property of the system, which describes how all the excitation branches $\omega_n(k)$ *combine* when the full initial value problem is set up. Therefore, apart from $\omega(k) = a + bk$, it is impossible to formulate a sufficient condition for causality in the form of an

inequality that $\omega(k)$ should obey. This is why, given a causality criterion like (1), one can always find models that fulfill it and are acausal, such as (2).

Nevertheless, we also show that one can overcome these difficulties by appealing to specific structures present in many (but not all) physical theories, which guarantee that the dispersion branches combine "correctly" to ensure causality. In particular, if the operator governing the dynamics is *local* and the system is covariantly *stable* (in precise senses defined below), *all superluminal tails cancel out;* see Theorem 1 for a precise statement.

A key inequality.—Our analysis relies on the following inequality, which must hold for all dispersion branches describing disturbances around the equilibrium state of a stable system in relativity [22,23]:

$$\mathfrak{Im}\,\omega(k) \leq |\mathfrak{Im}\,k|. \tag{3}$$

This covariant bound can be derived from the study of retarded causal correlators of stable phases of matter, and it is textbook material [24], whose importance in constraining transport properties of matter was demonstrated in Ref. [22]. Independently from the principle of causality, (3) constitutes the physical requirement that a stable system should be simultaneously stable in every inertial frame of Ref. [25]. In fact, if (3) were violated, namely if there were some $k \in \mathbb{C}$ for which $\mathfrak{Im}\,\omega > |\mathfrak{Im}\,k|$, then a boost with velocity $v = \mathfrak{Im}\,k/\mathfrak{Im}\,\omega$ would lead us to a new reference frame where $\mathfrak{Im}\,\omega' > 0$ and $\mathfrak{Im}\,k' = 0$ [23]. This would imply that there is an observer who can detect a growing Fourier mode, signaling, an instability [26–28]. For this reason, we assume (3) holds as a basic stability property of the system.

Single dispersion branches are superluminal.—Fix some level of description of matter, which may be, e.g., quantum field theory, kinetic theory, or hydrodynamics. Using established techniques [29–32], one can compute all the (possibly infinite) dispersion branches predicted by such theory. Choose one of interest, $\omega(k)$. According to conventional wisdom [13–18], the relation $\omega(k)$ determines how the corresponding excitation "propagates", and there should be some causality criterion for $\omega(k)$, e.g., (1), which guarantees that the excitation propagates subluminally. Now we prove that this intuitive interpretation can be consistently maintained only in the trivial case $\omega(k) = a + bk$. In dispersive media, causality can never be argued from $\omega(k)$ alone.

First, let us make the above (incorrect) intuition about the causality of $\omega(k)$ more precise. Let $\varphi(x^\mu) \in \mathbb{C}$ be the linear perturbation to a local observable of interest. For example, $\varphi(x^\mu)$ may be the local energy density fluctuation. Then, consider a $1 + 1$ dimensional profile $\varphi(t, x)$ that is constructed by superimposing plane waves all belonging *solely* to the selected excitation branch $\omega(k)$, i.e.,
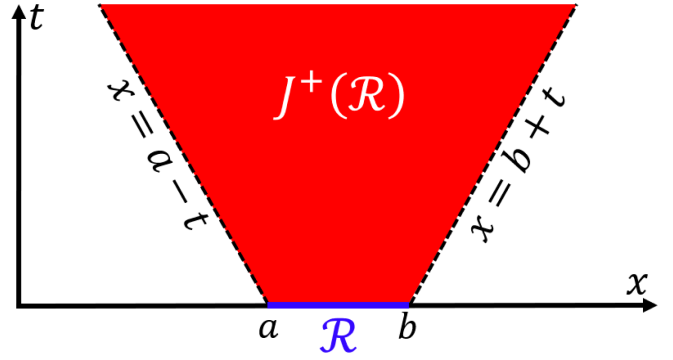

FIG. 1. The principle of causality. If a perturbation has initial support inside a region of space $\mathcal{R}$ (blue segment), then it cannot propagate outside the set $J^+(\mathcal{R})$ called the "causal future of $\mathcal{R}$" [6], or "future light cone of $\mathcal{R}$" [5] (red region).

$$\varphi(t, x) = \int_{-\infty}^{+\infty} \varphi(k) e^{i[kx - \omega(k)t]} \frac{dk}{2\pi}. \tag{4}$$

By setting $t = 0$, we find that $\varphi(k)$ is the Fourier transform of the initial data, $\varphi(0, x)$. The straightforward definition of "causal dispersion relation" is the following: If $\varphi(0, x)$ has support in a set $\mathcal{R}$, then the support of $\varphi(t, x)$ at later times should be contained inside the future light cone of $\mathcal{R}$, see Fig. 1. As a consequence, if $\varphi(0, x)$ has compact spatial support, one should find that $\varphi(t, x)$ has compact spatial support for each fixed $t > 0$. Now we will show that, in practice, this is never the case for φ given by (4). On the contrary, single-branched excitations of the form (4) always "travel" at infinite speed unless $\omega = a + bk$ (i.e., when the medium is not dispersive).

A simple argument.—If $\varphi(t, x)$ vanishes outside the future light cone of a compact set $\mathcal{R}$, then also $\partial_t \varphi(t, x)$ must vanish there. Hence, to prove that $\varphi(t, x)$ exits the light cone, it suffices to show that, for some $t_0 > 0$, the spatial profiles of $\varphi(t_0, x)$ and $\partial_t \varphi(t_0, x)$ cannot both have compact support simultaneously. We assume that $\varphi(t, x)$ is smooth, but the argument can be generalized.

Fix $t_0 > 0$. From (4) and the uniqueness of the Fourier transform, we have that the spatial Fourier transform of $\varphi(t_0, x)$ is given by $\varphi(t_0, k) = \varphi(k) e^{-i\omega(k)t_0}$. Now, suppose that $\varphi(t_0, x)$ is compactly supported. Then, $\varphi(t_0, k)$ extends to an entire function of $k \in \mathbb{C}$ [33]. Under our assumptions, we can bring time derivatives under the integral to conclude that $\partial_t \varphi(t_0, x)$ has spatial Fourier transform $\dot{\varphi}(t_0, k) = -i\omega(k)\varphi(t_0, k)$. Corollary 1.1 of [22] tells us that, if $\omega(k)$ obeys (3), then it cannot be an entire function (unless $\omega = a + bk$). Therefore, $\dot{\varphi}(t_0, k)$ is the product of an entire function with a function that is not entire. Such a function can be entire only in the remote eventuality in which the discrete zeroes of $\varphi(k)$ happen to cancel the singularities of $\omega(k)$. This requires a perfect fine-tuning of the initial data, and it does not happen in general [34]. Furthermore, according to Theorem 2 of [22], if $\omega(k)$ satisfies (3), then

its singularities are never poles or essential singularities. Instead, they are expected to be branch points, which cannot be erased by multiplying $\omega(k)$ with an entire nonzero function. Thus, $\dot{\varphi}(t_0, k)$ is not an entire function and, therefore, $\partial_t \varphi(t_0, x)$ cannot have compact support, as desired.

Application 1: Hegerfeldt paradox.—The above argument is a generalization of the well-known result (due to Hegerfeldt [35]) that relativistic single-particle wave functions of the form

$$\varphi(t, x) = \int_{-\infty}^{+\infty} \varphi(k) e^{i(kx - \sqrt{m^2 + k^2}t)} \frac{dk}{2\pi} \qquad (5)$$

must propagate outside the light cone [9]. Indeed, it can be easily verified that the dispersion relation of the free particle, $\omega = \sqrt{m^2 + k^2}$, obeys (3), and this forces it to be nonanalytical, as testified by the square root. Hence, the support of (5) expands at infinite speed [36], even if the group velocity, $v_g(k) = k/\sqrt{m^2 + k^2}$, is subluminal.

Application 2: Necessity of nonhydrodynamic modes.—An immediate corollary of our analysis is that the retarded Green's function of any theory for diffusion having only one dispersion relation, $\omega(k) = -iDk^2 + \mathcal{O}(k^3)$, always exits the light cone. Thus, to build a subluminal Green's function, we need at least two dispersion relations (see [37], Sec. 7.4). This explains why an additional (usually gapped) mode is needed for causality [22].

Explanation.—The superluminal behavior of (4) in causal matter seems absurd, but there is a simple explanation: excitations of the form (4) cannot be truly localized, unless $\omega = a + bk$. They may *seem* to have compact support, if $\varphi(0, x)$ is supported in $\mathcal{R}$, but, in principle, an observer can detect the excitation from outside $\mathcal{R}$ already at $t = 0$ by measuring some other observable. In fact, we recall that the dispersion relation $\omega(k)$ is derived from some underlying physical theory (e.g., quantum field theory, kinetic theory, or hydrodynamics), which may possess several other local observables besides φ. The fact that the initial profile $\varphi(0, x)$ has support inside $\mathcal{R}$ does not imply that all the measurable fields affected by the excitation are unperturbed outside $\mathcal{R}$. Instead, it may be the case that, due to this excitation, the perturbation to a second observable $\psi(x^\mu)$ of the theory has already unbounded support at $t = 0$. This is how, in principle, φ can propagate outside the light cone without necessarily violating the principle of causality in the full theory: there is no superluminal propagation of information if such information was already accessible through the measurement of $\psi(0, x)$ outside $\mathcal{R}$. Indeed, below we prove that if the perturbations to *all* the observables are initially supported inside a compact region $\mathcal{R}$ (i.e., the excitation is truly localized), and the dynamics is governed by a local operator, then $\varphi(t, x)$ cannot be expressed in the form (4), and it must always combine at least two dispersion branches, unless $\omega = a + bk$.

Compactly supported excitations.—We assume that the state of the system at a given time can be characterized, in the linear regime, by a collection of smooth perturbation fields $\Psi(x^\mu) \in \mathbb{C}^{\mathfrak{D}}$, which all vanish at equilibrium. We assume that $\mathfrak{D}$ is finite, although it can be as large as the number of particles in a material volume element. In most physical theories currently available (e.g., electrodynamics, elasticity theory, or hydrodynamics), the $1 + 1$ dimensional equation of motion of the system takes the form

$$\partial_t \Psi = \mathcal{L}(\partial_x)\Psi, \qquad (6)$$

where is $\mathcal{L}(\partial_x)$ a polynomial of finite degree in ∂_x, i.e., $\mathcal{L}(\partial_x) = A_0 + A_1 \partial_x + \cdots + A_M \partial_x^M$, where A_j are constant $\mathfrak{D} \times \mathfrak{D}$ matrices, and $M \in \mathbb{N}$ [38]. This is what we mean by a "local operator." In fact, operators involving an infinite series of derivatives can produce nonlocalities and causality violations. For example, if we set $\mathcal{L}(\partial_x) = e^{a\partial_x}$, Eq. (6) becomes $\partial_t \Psi(t, x) = \Psi(t, x + a)$, which is clearly a nonlocal theory. Indeed, the main reason why the BBM equation in (2) is acausal is that its dynamical operator, $\mathcal{L}(\partial_x) = (\partial_x^2 - 1)^{-1}\partial_x$, is nonlocal [39,40].

The general formal solution to (6) reads

$$\Psi(t, x) = \int_{-\infty}^{+\infty} e^{\mathcal{L}(ik)t} \Psi(k) e^{ikx} \frac{dk}{2\pi}, \qquad (7)$$

where $\Psi(k)$ is the Fourier transform of the initial data $\Psi(0, x)$. Now, the field $\varphi(t, x)$, being a linearized local observable, is a local linear functional of the degrees of freedom, namely $\varphi = \mathcal{V}(\partial_x)\Psi$, where $\mathcal{V}$ is also a polynomial of finite degree in ∂_x, i.e., $\mathcal{V}(\partial_x) = B_0 + B_1\partial_x + \cdots + B_N\partial_x^N$. Here, B_j are constant row vectors of length $\mathfrak{D}$, and $N \in \mathbb{N}$. Therefore, we have the following formula:

$$\varphi(t, x) = \int_{-\infty}^{+\infty} \mathcal{V}(ik) e^{\mathcal{L}(ik)t} \Psi(k) e^{ikx} \frac{dk}{2\pi}. \qquad (8)$$

Assume that the excitation is initially supported inside $\mathcal{R}$. Then, all the components of $\Psi(0, x)$ are compactly supported, and all the components of $\Psi(k)$ are entire functions. Furthermore, $\mathcal{L}(ik)$ and $\mathcal{V}(ik)$ are entire in $k \in \mathbb{C}$, being polynomials. Also, the matrix exponential is an analytic function of the components of the matrix in the exponent, so that $e^{\mathcal{L}(ik)t}$ is entire in k. Combining these results, we can conclude that the integrand of (8) is an entire function of k for all t. For this reason, it cannot coincide with (4), unless $\omega(k) = a + bk$ (i.e., in dispersion-free systems). This shows that, when we construct the state (4) in a dispersive medium, we implicitly allow some component of Ψ to have unbounded support already at $t = 0$, which is what we wanted to prove. In the Supplemental Material [21], we analyze the explicit example of the Klein-Gordon equation.

Causality criterion for stable matter.—The above analysis suggests that if the dynamics of the system is governed by a local operator $\mathcal{L}$, then all the dispersion branches will automatically combine in a way to cancel the infinite tails of the individual excitations (4). This intuition can be made rigorous through the following theorem, according to which, schematically,

$$\begin{pmatrix} \text{local} \\ \text{equations} \end{pmatrix} + \begin{pmatrix} \text{stability in} \\ \text{all frames} \end{pmatrix} \Rightarrow \begin{pmatrix} \text{relativistic} \\ \text{causality} \end{pmatrix}.$$

More precisely, note the following:

Theorem 1.—If $\mathcal{L}(ik)$ and $\mathcal{V}(ik)$ are polynomials of finite degree in ik, and the eigenvalues $\omega_n(k)$ of $i\mathcal{L}(ik)$ obey the stability requirement (3) for all $k \in \mathbb{C}$, then all smooth linear excitations propagate subluminally, in the sense that the support of $\varphi(t,x)$, as given by (8), is contained within the future light cone of the support of $\Psi(0,x)$.

Proof.—We will focus on the case where $\Psi(0,x)$ has support inside the interval $[-1,0]$. We will verify that $\varphi(t,x)$ vanishes for $x > t \geq 0$. More general cases can be recovered from here by invoking linearity, translation invariance, and closure of the solution space.

Consider the complex integral

$$\mathcal{I} = \int_\Gamma \mathcal{V}(ik)e^{\mathcal{L}(ik)t}\Psi(k)e^{ikx}\frac{dk}{2\pi}, \qquad (9)$$

where Γ is the closed loop in complex k space in Fig. 2, in the limit of large R. Since the integrand is entire, $\mathcal{I} = 0$. Let us now show that, if $x > t \geq 0$, then the contribution coming from the upper semicircle decays to zero as $R \to +\infty$, so that $0 = \mathcal{I} = \varphi(t,x)$; see Eq. (8). To this end, we first note that, according to the Jordan-Chevalley decomposition theorem, the matrix $\mathcal{L}(ik)$ can be expressed as

$$\mathcal{L}(ik) = -i\sum_n \omega_n(k)\mathcal{P}_n(k) + \mathcal{N}(k), \qquad (10)$$

where $\mathcal{P}_n$ are complementary eigenprojectors (so that $\mathcal{P}_m\mathcal{P}_n = \delta_{mn}\mathcal{P}_n$, $\sum_n \mathcal{P}_n = \mathbb{I}$), and $\mathcal{N}$ is a nilpotent matrix ($\mathcal{N}^a = 0$ for some $a \in \mathbb{N}$) which commutes with all $\mathcal{P}_n$. Thus, the integrand in (9) can be rewritten as

$$\sum_n \sum_{j=0}^{a-1} \mathcal{V}\frac{(\mathcal{N}t)^j}{j!}\mathcal{P}_n e^{i(kx-\omega_n t)}\Psi. \qquad (11)$$

The matrix elements of $\mathcal{N}^j$ and $\mathcal{P}_n$ grow at most like powers of $|k|$. This follows from [41] [Chap. 2, Eqs. (1.21) and (1.26)], applied to the matrix $(ik)^{-M}\mathcal{L}(ik)$ regarded as a polynomial in $(ik)^{-1} \to 0$, combined with the fact that $(ik)^{-M}\mathcal{L}(ik)$ and $\mathcal{L}(ik)$ have the same invariant subspaces. On the other hand, if $\mathfrak{Im}\,k \geq 0$, and $x > t \geq 0$, we have the following estimates:
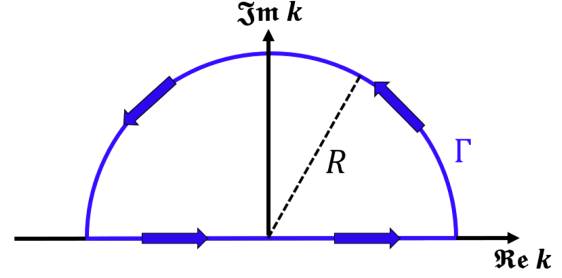

FIG. 2. Path of integration for the proof of Theorem 1.

$$|e^{i(kx-\omega_n t)}| = e^{-x\mathfrak{Im}k + t\mathfrak{Im}\omega_n} \leq e^{-(x-t)\mathfrak{Im}k} \leq 1,$$

$$|\Psi(k)| = \left| \int_{-1}^0 e^{\tilde{x}\mathfrak{Im}k}\Psi(0,\tilde{x})e^{-i\tilde{x}\mathfrak{Re}k}d\tilde{x} \right| \leq L^1[\Psi(0,x)]. \qquad (12)$$

In the first line, we have invoked the inequality (3). In the second line, we have used the fact that $e^{\tilde{x}\mathfrak{Im}k} \leq 1$ inside the interval $[-1,0]$. Note that $\Psi(0,x)$, being continuous and compactly supported, has finite L^1 norm. From the estimates (12), we can conclude that (11) decays exponentially to zero when $\mathfrak{Im}k \to +\infty$. Furthermore, since $\Psi(\mathfrak{Re}k + i\mathfrak{Im}k)$, regarded as a function of $\mathfrak{Re}k$, is the Fourier transform of the Schwartz function $e^{x\mathfrak{Im}k}\Psi(0,x)$, it is itself a Schwartz function [33], meaning that (11) decays to zero faster than any power also when $\mathfrak{Re}k \to \infty$. It follows that, as $R^2 = (\mathfrak{Re}k)^2 + (\mathfrak{Im}k)^2 \to +\infty$, the integral over the semicircle converges to zero since the integrand decays faster than any power of R. ∎

Most derivations of (1) rely on the assumption that $\omega \approx v_f k$ for large $k \in \mathbb{C}$, so that (1) is a direct consequence of (3). However, (3) is a much more stringent condition, as it automatically rules out the acausal equations (2). Indeed, the apparent success of (1) in many situations can be traced back to (3) through the following theorem proven below:

Theorem 2.—If (6) is a hyperbolic first-order system, with $\mathcal{L}(\partial_x) = -\Xi - \mathcal{M}\partial_x$, then (3) implies (1), and the characteristic velocities coincide with the front velocities.

Proof.—The ratios ω_n/k are eigenvalues of the matrix

$$\frac{i\mathcal{L}(ik)}{k} = \mathcal{M} + (ik)^{-1}\Xi. \qquad (13)$$

If we regard the right-hand side as a polynomial in $(ik)^{-1}$, we can take the limit as $(ik)^{-1} \to 0$ and apply the continuity property of eigenvalues [41] to conclude that ω_n/k must converge to eigenvalues of $\mathcal{M}$ for large $k \in \mathbb{C}$. But the eigenvalues of $\mathcal{M}$ are the characteristic velocities of the system, and they are real (by hyperbolicity), so that

$$\lim_{k \to \infty, k \in \mathbb{C}} \frac{\omega_n}{k} = v_{\mathrm{ch},n} \in \mathbb{R}. \qquad (14)$$

Restricting the above limit to real k, and using the continuity of $\mathfrak{Re}$, we find that the characteristic velocities

coincide with the front velocities. Restricting the limit to imaginary k, we find that (3) implies causality.

$$v_{\mathrm{ch},n} = \mathfrak{Re}\left[\lim_{k\to\infty,k\in\mathbb{R}}\frac{\omega_n}{k}\right] = \lim_{k\to\infty,k\in\mathbb{R}}\frac{\mathfrak{Re}\,\omega_n}{k} = v_{f,n},$$

$$v_{\mathrm{ch},n} = \mathfrak{Re}\left[\lim_{k\to\infty,k\in i\mathbb{R}}\frac{\omega_n}{i\,\mathfrak{Im}\,k}\right] = \lim_{k\to\infty,k\in i\mathbb{R}}\frac{\mathfrak{Im}\,\omega_n}{\mathfrak{Im}\,k} \in [-1,1].$$

(15)

This completes our proof. ∎

While the above analysis was restricted to classical initial value problems, its broad implications may also be extrapolated to quantum systems. For example, some conformal field theories are known to be acausal [42]. Given that such theories are local, we can "apply" our Theorem 1 to conclude that such theories are not covariantly stable and violate the bound (3), in agreement with Sec. III.A of [25].

Correlators in QFT.—Causality requires multiple dispersion relations also in QFT. Given a local observable operator $\hat{\varphi}(x^\mu)$, the correlator $G(x^\mu) = \langle[\hat{\varphi}(x^\mu),\hat{\varphi}(0)]\rangle$ has support inside the light cone [9]. But since the slices of the light cone at constant time are compact spheres, the spatial Fourier transform $G(t,\mathbf{k})$ must be entire in $\mathbf{k}$ for all t [33]. This is why introducing momentum cutoffs or "patching" correlators in momentum space leads to causality violations [43]: it breaks analyticity. Furthermore, if $G(t,\mathbf{k})$ can be expressed as a superposition of modes of the form $e^{-i\omega_n(\mathbf{k})t}$ (see e.g., [44]), then we know that all the nonanalyticities of the individual frequencies $\omega_n(\mathbf{k})$ must cancel out.

Final remarks.—Consider the following puzzle: All solutions of the relativistic Schrödinger equation $i\partial_t\varphi = \sqrt{m^2 - \partial_x^2}\,\varphi$ are also solutions of the Klein-Gordon equation $-\partial_t^2\varphi = (m^2 - \partial_x^2)\varphi$. Nevertheless, the former is notoriously acausal [9], while the latter is causal. This defies the intuition of causality as a statement about the propagation speed of φ. How can the same function $\varphi(t,x)$ be superluminal when viewed as a solution of one equation and subluminal when viewed as a solution of another equation?

Here, we solved this puzzle by showing that causality is not an intrinsic property of the fields themselves. Rather, it is a property of how we "attach information" to the fields by defining the physical state. The existence of faster-thanlight motion does not result in causality violation if the motion carries no new information about the state. Indeed, relativistic Schrödinger and Klein-Gordon differ by the way they define the physical state at a given time: $\{\varphi(x)\}$ in the former, and $\{\varphi(x),\partial_t\varphi(x)\}$ in the latter. The puzzle arises because compactly supported field states within relativistic Schrödinger (i.e., localized φ profiles) must have unbounded support within Klein-Gordon (i.e., cannot be localized in $\partial_t\varphi$); see Supplemental Material [21].

Starting from this intuition, we showed that nonhydrodynamic modes become necessary for relativistic viscous hydrodynamics for the same reason that antiparticles are

necessary for relativistic quantum mechanics: defining a notion of locality in dispersive systems requires at least two dispersion relations.

Note added.—Recently, other formulations of Theorems 1 and 2 were proposed [45,46].

[1] A. Einstein, Jahrb. Radioaktivität Elektron. **4**, 411 (1908).

[2] R. Tolman, *The Theory of the Relativity of Motion* (University of California Press, Berkeley, CA, 1917).

[3] L. Landau and E. Lifshitz, *Course of Theoretical Physics, Vol. 2: Classical Theory of Fields* (Butterworth Heinemann, Amsterdam, 1994).

[4] R. Courant and D. Hilbert, *Methods of Mathematical Physics, Vol. 2: Partial Differential Equations* (John Wiley and Sons, New York, 1989).

[5] Y. Aharonov, A. Komar, and L. Susskind, Phys. Rev. **182,** 1400 (1969).

[6] S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Space-Time*, Cambridge Monographs on Mathematical Physics (Cambridge University Press, Cambridge, England, 2011).

[7] R. M. Wald, *General Relativity* (Chicago University Press, Chicago, 1984).

[8] J. Rauch, *Partial Differential Equations*, Graduate Texts in Mathematics (Springer, New York, 1991).

[9] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory* (Addison-Wesley, Reading, MA, 1995).

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

[11] B. D. Keister and W. N. Polyzou, Phys. Rev. C **54,** 2023 (1996).

[12] For anisotropic systems, the dispersion relation is $\omega(k^x, k^y, k^z)$. In this case, we align the x axis along a direction of interest and define $\omega(k) := \omega(k,0,0)$. Then, our results apply to waves propagating in x.

[13] L. Brillouin, *Wave Propagation and Group Velocity* (Academic Press, New York, 1960).

[14] G. S. Denicol, T. Kodama, T. Koide, and P. Mota, J. Phys. G Nucl. Phys. **35**, 115102 (2008).

[15] S. Koide and R. Morino, Phys. Rev. D **84**, 083009 (2011).

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

[17] E. Krotscheck and W. Kundt, Commun. Math. Phys. **60**, 171 (1978).

[18] S. Pu, T. Koide, and D. H. Rischke, Phys. Rev. D **81**, 114039 (2010).

[19] T. B. Benjamin, J. L. Bona, and J. J. Mahony, Phil. Trans. R. Soc. A **272**, 47 (1972).

[20] The superluminality of the BBM equation does not contradict [17]; see the Supplemental Material [21].

[21] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.132.162301 for we expand the discussion on the causality of the BBM equation and the mathematical relationship between the Relativistic Schrodinger and Klein-Gordon equations.

[22] M. P. Heller, A. Serantes, M. Spaliński, and B. Withers, arXiv:2212.07434.

[23] L. Gavassino, Phys. Lett. B **840**, 137854 (2023).

[24] C. Itzykson and J. B. Zuber, Quantum Field Theory, International Series in Pure and Applied Physics (McGraw-Hill, New York, 1980).

[25] L. Gavassino, Phys. Rev. X **12**, 041001 (2022).

[26] W. Hiscock and L. Lindblom, Phys. Rev. D **31**, 725 (1985).

[27] P. Kostädt and M. Liu, Phys. Rev. D **62**, 023003 (2000).

[28] L. Gavassino, M. Antonelli, and B. Haskell, Phys. Rev. D **102**, 043018 (2020).

[29] D. Birmingham, I. Sachs, and S. N. Solodukhin, Phys. Rev. Lett. **88**, 151301 (2002).

[30] P. K. Kovtun and A. O. Starinets, Phys. Rev. D **72**, 086009 (2005).

[31] P. Kovtun, J. High Energy Phys. 10 (2019) 034.

[32] G. Perna and E. Calzetta, Phys. Rev. D **104**, 096005 (2021).

[33] L. Hormander, The Analysis of Linear Partial Differential Operators I, Second Edition, Comprehensive Studies in Mathematics (Springer-Verlag, Berlin, 1990).

[34] Note that, if $\varphi(k)$ is the Fourier transform of a compactly supported function $\varphi(0, x)$, then $\varphi(k - a)$ is the Fourier transform of the function $\varphi(0, x)e^{iax}$, which is also compactly supported. Thus, if the zeroes of $\varphi(k)$ overlap the singularities of $\omega(k)$, we can always make a shift a in Fourier space, and construct a new solution for which this overlap no longer happens.

[35] G. C. Hegerfeldt, Phys. Rev. D **10**, 3320 (1974).

[36] B. Thaller, The Dirac Equation (Springer-Verlag, Berlin, 1992).

[37] P. M. Morse and H. Feshbach, Methods of Theoretical Physics, International Series in Pure and Applied Mathematics (McGraw-Hill Book Company, New York, 1953).

[38] Equations involving higher derivatives in time, e.g., $(\partial_t^2 - \partial_x^2)\varphi = 0$, can always be reduced to systems that are of first order in time through the introduction of more fields: $\partial_t \varphi = \Pi$, $\partial_t \Pi = \partial_x^2 \varphi$.

[39] T. Congy, G. A. El, M. A. Hoefer, and M. Shearer, arXiv:2012.14579.

[40] The equation $i\partial_t \varphi = \sqrt{m^2 - \partial_x^2}\varphi$ is also nonlocal for the same reason, and this is ultimately what gives rise to the Hegerfeldt paradox in relativistic quantum mechanics.

[41] T. Kato, Perturbation Theory for Linear Operators, Classics in Mathematics (Springer-Verlag, Berlin, 1995).

[42] M. Brigante, H. Liu, R. C. Myers, S. Shenker, and S. Yaida, Phys. Rev. Lett. **100**, 191601 (2008).

[43] P. A. Henning, E. Polyachenko, T. Schilling, and J. Bros, Phys. Rev. D **54**, 5239 (1996).

[44] P. Lowdon and O. Philipsen, EPJ Web Conf. **274**, 05013 (2022).

[45] D.-L. Wang and S. Pu, arXiv:2309.11708.

[46] R. E. Hoult and P. Kovtun, arXiv:2309.11703.