# Superluminal Behavior, Causality, and Instability*

Y. Aharonov, A. Komar, and L. Susskind
*Belfer Graduate School of Science, Yeshiva University, New York, New York 10033*

(Received 4 February 1969)

We distinguish two fundamentally different types of superluminal (group velocity$>c$) behavior. One type, called causal, is shown not to conflict with the usual condition that a signal originating in a region is felt only in the forward light cone of that region. This type of mode occurs as small oscillations about an unstable equilibrium configuration of high energy density. The second type of superluminosity, called noncausal, violates the causality conditions of conventional theory. It is shown that the noncausal superluminosity occurs only for field theories which are singular in the sense that there does not exist a unique one-to-one relationship between the momenta and the velocities. The field equations of such theories do not present a good Cauchy problem for arbitrary data on spacelike hyperplanes. They thus lead to ambiguous and/or singular solutions. An interesting feature of our first model is that it leads to tachyonlike behavior without one's having to introduce negative energies.

## I. TWO SIMPLE EXAMPLES OF SUPERLUMINAL BEHAVIOR

INTERESTING examples of systems admitting wave modes with $|v_g|>1$, where $v_g$ means group velocity and we take the speed of light to equal 1, have been given by Bludman and Ruderman.[1] Their arguments are based on the general relations between sound velocity, pressure, and energy, thus making it somewhat difficult to visualize the mechanisms leading to $v_g>1$, and their relation to the question of noncausal behavior. For purposes of elucidating these properties we have constructed two simple analyzable models.

Our models are unquantized field theories subject to the following requirements.

(i) The equations of motion are Lorentz-invariant and follow from a local Lagrangian.

(ii) The energy shall be positive definite.

### A. Causal Model

Consider a one-dimensional "lattice" of identical pendulums suspended from equally spaced fixed points along the $z$ axis, each free to swing in the $zy$ plane. Each pendulum bob is connected to its nearest neighbor by a spring. The entire system is in a gravitational field $g$ in the $-y$ direction. Thus, for $\theta_i-\theta_{i-1}\ll\pi$ the Lagrangian is

$$\mathcal{L}=\sum_i a\dot{\theta}_i^2-b(\theta_i-\theta_{i-1})^2-c(1-\cos\theta_i)\,, \qquad (1)$$

where $\theta_i$ is the angle of the $i$th pendulum with respect to the vertical, and $a$, $b$, and $c$ are related to the moments of inertia, spring constant, and $g$.

If we take the continuum limit of Eq. (1) in the usual manner, the Lagrange density becomes

$$\mathcal{L}(z)=\frac{1}{2}\left(\frac{d\theta}{dt}\right)^2-\frac{1}{2}\left(\frac{d\theta}{dx}\right)^2-(1-\cos\theta)\,, \qquad (2)$$

where we have chosen $a$, $b$, and $c$ so as to give unit coefficients in (2). We regard Eq. (2) as a relativistically invariant Lagrange density for a scalar field $\theta(x)$ containing, in addition to the usual free term $\Box\theta$, a nonlinear self-coupling $(1-\cos\theta)$.

The pendulum-spring interpretation is only used to help visualize the meaning of the solutions that we shall discuss.

The equation of motion resulting from this Lagrangian [Eq. (2)] is

$$d^2\theta/dt^2-d^2\theta/dx^2+\sin\theta=0\,. \qquad (3)$$

For $\theta$ small, i.e., all pendulums near the vertical stable equilibrium position, Eq. (3) is the usual Klein-Gordon equation with positive (mass)$^2$ equal to 1.

$$\Box\theta+\theta=0 \qquad (4)$$

For $\theta$ near $\pi$ (unstable equilibrium), if one defines

$$\phi=\pi-\theta\,,\quad \sin\theta=-\sin\phi\,, \qquad (5)$$

Eq. (3) becomes

$$d^2\phi/dt^2-d^2\phi/dx^2-\sin\phi=0\,, \qquad (6)$$

thus yielding the imaginary-mass Klein-Gordon equation for $\phi$ small.

$$\Box\phi-\phi=0\,. \qquad (7)$$

(We can note at this point that we have tachyonlike behavior,[2] although it is easy to check that the corresponding Hamiltonian is positive definite.) The dispersion relation for Eq. (7) is

$$\omega^2-k^2=-1 \qquad (8)$$

and the group velocity is

$$v_g=d\omega/dk=k/\omega>1\,. \qquad (9)$$

Thus the system is superluminal for modes oscillating about the point $\theta=\pi$. This point is the unstable vertical position for each pendulum. However, as long as $k$ is chosen so as to make $\omega$ real, small stable superluminal

* Supported in part by U. S. Air Force Office of Scientific Research Grant Nos. AF-68-1524 and AF-1282-67.

[1] S. A. Bludman and M. A. Ruderman, Phys. Rev. **170**, 1176 (1968).

[2] J. Dhar and E. C. G. Sudarshan, Phys. Rev. **174**, 1808 (1968).

oscillations about $\theta=\pi$ occur. The instability occurs when the perturbations involve $k^2<1$, thus making $\omega$ imaginary.

### B. Noncausal Model

Consider a scalar field $\phi(x)$ with a Lagrangian $\mathcal{L}=\frac{1}{2}[F(S)=m^2\phi^2]$, where $F$ is a nonlinear real function and $S=\eta^{\mu\nu}\phi_{,\mu}\phi_{,\nu}$. ($\eta_{\mu\nu}$ is the Minkowski metric.) The canonical momentum is

$$\Pi_\phi\equiv\Pi=(\partial F/\partial S)\dot{\phi}. \tag{10}$$

The Hamiltonian is

$$H=\Pi\dot{\phi}-\mathcal{L}=(\partial F/\partial K_S)\dot{\phi}^2-\tfrac{1}{2}F(S)+\tfrac{1}{2}m^2\phi^2. \tag{11}$$

We guarantee $H>0$ by choosing $F(S)<0$ and $\partial F/\partial S>0$ for all $S$.

From $\mathcal{L}$ we obtain the equations of motion

$$\eta^{\nu\mu}\left(\frac{\partial^2F}{\partial S^2}\frac{\partial S}{\partial x^\mu}\phi_{,\nu}+\frac{\partial F}{\partial S}\phi_{,\mu\nu}\right)-m^2\phi=0 \tag{12}$$

or

$$\left(\frac{2F''}{F'}\eta^{\sigma\nu}\eta^{\tau\mu}\phi_{,\mu}\phi_{,\nu}+\eta^{\sigma\tau}\right)\phi_{,\sigma\tau}=\frac{m^2\phi}{F'}\equiv C^{\sigma\tau}\phi_{,\sigma\tau}. \tag{13}$$

Evidently the characteristic lines for the second-order equation (13) are not the light cone but are the null geodesics of the "metric" $C_{\sigma\tau}$. If these characteristics lie outside the light cone then small perturbations of a solution originating from the action of a source at a point will propagate outside the light cone and will be superluminal. For example, choose as an initial condition $\phi=\text{const}=A$ and $\dot{\phi}=B=\text{const}$. Then $C_{\sigma\tau}$ is the matrix

$$\begin{bmatrix}1&&&\\&-1&&\\&&-1&\\&&&-1\end{bmatrix}+2B^2\frac{F''(B^2)}{F'(B^2)}\begin{bmatrix}1&&&\\&0&&\\&&0&\\&&&0\end{bmatrix}. \tag{14}$$

Clearly, if $F''/F'$ is negative the characteristics are outside the light cone.[3]

## II. CAUSALITY, SUPERLUMINOSITY, AND INSTABILITY

We shall now analyze the implications of our models for the questions of causality; i.e., whether information can be transmitted with $v>c$ through the superluminal modes.

Let us consider perturbing the equations of the model in Sec. I A by the addition of a point source at the origin. In the neighborhood of the unstable-equilibrium configuration the field equation becomes

$$\frac{d^2\phi}{dt^2}-\frac{d^2\phi}{dx^2}-\sin\phi=\delta(x)\delta(t). \tag{15}$$

To distinguish the signal from the background we seek a solution of Eq. (15) which vanishes for $t<0$. Such a solution exists and, on general principles, must vanish outside the forward light cone.[4] Thus, even though the theory is superluminal, it is causal in that small disturbances propagate with $v=c$.

In order to better understand how this occurs, let us approximate Eq. (15) by

$$\ddot{\phi}-\phi''-\phi=\delta(x)\delta(t), \tag{16}$$

which in "momentum space" becomes

$$\ddot{\phi}(k)+(k^2-1)\phi(k)=\delta(t). \tag{17}$$

For $t>0$ the solution is

$$\phi(k)=(e^{i\omega_k t}-e^{-i\omega_k t})/i\omega_k, \tag{18}$$

with $\omega=(k^2-1)^{1/2}$.

Now for $k^2>1$, $\phi(k)$ has normal oscillatory motion. The excitation of such modes does not lead to an unstable motion. However, for $k^2<1$, Eq. (18) gives an exponential growth for $\phi(k)$. Since the local perturbation $\delta(x)\delta(t)$ has all Fourier modes $k$, the response will involve all modes, in particular, the exponentially growing ones. Hence any local disturbance will set off an instability. Of course, the field $\phi$ will not really grow indefinitely. As soon as $\phi$ is no longer small the nonlinearities of $\sin\phi$ become important and damp the increase of $\phi$. What is actually happening is that we have set off a motion in which the pendulums are falling over, domino fashion, into large oscillating swings, the fall spreading out from the source. Although the field contains the superluminal stable modes $k^2>1$ the Cauchy-Kowalewska theorem and the Holgrem theorem,[4] which assert that a solution is uniquely determined exclusively by the Cauchy data subtended by the backward characteristic cone, assure us that the sum of superluminal and unstable ($k^2<1$) modes cause the fall to spread with $v\leq c$.

Contrary to certain claims made in the literature, there exists a Green's function for the negative (mass)$^2$ Klein-Gordon equation which is strictly causal. However, it is exponentially divergent in any timelike direction. If exponentially growing modes are excluded ($k<1$), then no retarded Green's function exists and the field cannot couple locally to a source. In our nonlinear model there is no reason to exclude such exponentially divergent modes since the growth is soon damped. Furthermore, because of the nonlinearity, the unstable modes ($k^2<1$) and superluminal modes

[3] See Appendix A for a more general and detailed discussion of this question.

[4] See, e.g., *Partial Differential Equations*, Fritz-John, Chap. II. For the linearized case (tachyons), we explicitly demonstrate this property in Appendix B.

($k^2>1$) are coupled once we leave the small oscillation approximation.

On the other hand, if we turn to the noncausal model of Eq. (12) the effect of a source at a point will spread to fill the region between the characteristics defined by $C_{\mu\nu}$ which in general could be outside the light cone. Such a theory would then not be causal.

## III. NONCAUSALITY AND CANONICAL FORMALISM

We shall now show that the noncausal theory is necessarily singular and exhibits anomalies in the canonical formalism. Let us examine the relation between $\Pi_\phi$ and $\dot\phi$ given by Eq. (10):

$$\Pi_\phi = F'\dot\phi.$$

The condition that the time derivative be a single-valued function of $\Pi$ is necessary for an unambiguous transition from the Lagrangian formalism to the Hamiltonian formalism. Thus it is required that

$$\partial\Pi/\partial\dot\phi \neq 0 \qquad (19)$$

or

$$2F''\dot\phi^2 + F' \neq 0, \qquad (20)$$

which, for a general $F$, need not be satisfied. In fact, points at which $2F''\dot\phi^2+F'=0$ have a simple interpretation. From Eq. (13) we see that Eq. (20) is the condition that the time-time component of the metric $C_{\sigma\tau}$ does not vanish.

At points where $C_{00}$ does vanish, at least one of the characteristics is tangent to the $t=\text{const}$ lines and hence Cauchy data for the Euler-Lagrange equations cannot be propagated unambiguously off such $t=0$ surfaces.

More generally, if for any Lorentz-covariant field theory, a portion of its characteristic surface lies outside the light cone, there exists a Lorentz observer whose initial surface osculates the characteristic. It is evident that Cauchy data assigned on such an initial surface cannot propagate uniquely off it, for if such were the case it would not be possible to insert singularities, sources, or nonanalytic discontinuities on the characteristic surface.

We conclude that whenever truly noncausal behavior occurs, it is linked to a peculiarity in the equation of motion such that the solution at later times is not always determined by initial data. This in turn indicates that the Lagrangian of the theory is singular, or, equivalently, that the relation between canonical variables and Lagrangian variables must necessarily be singular.

## ACKNOWLEDGMENTS

One of us (Y.A.) would like to acknowledge useful discussions with S. Coleman. Two of us (Y.A.) and (L.S.) would like to acknowledge a useful discussion with M. Ruderman.

## APPENDIX A

If we consider for a moment the most general Lorentz-covariant Lagrangian for a real scalar field satisfying second-order field equations,

$$L = F(S,\phi),$$

the necessary and sufficient conditions for the resulting field equation

$$(2F'\eta^{\mu\nu} + 4F''\phi^{,\mu}\phi^{,\nu})\phi_{,\mu\nu} = \partial F/\partial\phi$$

(the prime denotes differentiation with respect to $s \equiv \eta^{\mu\nu}\phi_{,\mu}\phi_{,\nu}$) to have normal hyperbolic form requisite for a stable Cauchy problem can be stated Lorentz-covariantly as follows:

$$\text{(a) } F'>0, \quad \text{(b) } F''\geq 0, \quad \text{(c) } 2SF''/F'>-1.$$

If we let $\eta_\mu$ be the surface normal to the characteristic surface of this field equation, it must evidently satisfy

$$(\eta^{\mu\nu} + (2F''/F')\phi^\mu\phi^\nu)\eta_\mu\eta_\nu = 0$$

or, equivalently,

$$\eta^2 = -(2F''/F')(\phi^\mu\eta_\mu)^2.$$

Thus from inequalities (a) and (b) alone, we find that $n^2\leq 0$. That is, $\eta_\mu$ is either spacelike or null, and consequently the characteristic surface lies entirely within the usual Minkowski light cone, the theory being causal in the usual sense. These two inequalities ensure that we always have a Cauchy problem, or, equivalently, a nonsingular transition to the Hamiltonian formalism. Inequality (c), which greatly restricts the class of Lagrangians analytic in its arguments, assures the stability of the Cauchy problem—that is, small changes in the Cauchy data cannot produce large changes in the solution arbitrarily close to the initial surface. Inequality (c) also assures us that the energy is bounded from below and can therefore be effectively made positive definite. It can be easily seen that apart from the Klein-Gordon class, for which $F''=0$, all other analytic Lagrangians consistent with the above three inequalities cannot be finite polynomial functions of $s$. One of the simplest such analytic Lagrangians is

$$F(S) = \int_0^s \exp[\tfrac{1}{4}\arctan\tfrac{1}{2}s']ds'.$$

## APPENDIX B

We offer a proof here that the retarded Klein-Gordon Green's function differs from zero only inside the forward light cone, even when the sign of the mass term is negative.

Let $\Phi(x,t)$ be the solution for $\Box^2\Phi - m^2\Phi = \delta(x)\delta(t)$ which vanishes for $t<0$. Defining $f(x,t)=\Phi(x,t)e^{-mt}$, where for the retarded solution we require $m>0$, the equation that $f$ satisfies is

$$\Box^2+2m\partial f/\partial t=\delta(x)\delta(t).$$

Thus

$$f(x,t)\sim-\int\frac{e^{-i(k_0t-kx)}dk_0dk}{k_0^2-k^2+2imk}\quad\text{(modulo factors of }2\pi).$$

Let $u=k_0+k$, $\theta=k_0-k$, $x_1=t+k$, and $x_2=t-x$. Then

$$f=-\int\frac{e^{i-(ux_2+\theta x_1)}}{u\theta+2im(u+\theta)}dud\theta.$$

The future light cone is characterized by $x_1>0$ and $x_2>0$. Thus we would like to show that we can define $f$ so that it vanishes when either $x_1<0$ or $x_2<0$. Let us choose $x_1<0$. If we perform the $\theta$ integral first by closing the contour of integration on the upper half of the complex $\theta$ plane, we see that $f$ vanishes, for there is no pole for $u$ real and $\text{Im}\theta>0$. To prove this, let $\theta=\theta_1+i\theta_2$, where $\theta_2>0$. The condition for a pole, $u\theta+2im(u+\theta)=0$, becomes, upon separation into real and imaginary parts, $u\theta_1-2m\theta_2=0$ and $u\theta_2+2m(u+\theta_1)=0$. Eliminating $\theta_1$ from these last two equations, we obtain $u^2(\theta_2+2m)+4m^2\theta_2=0$, which, in view of our assumptions $m>0$, $\theta_2>0$, cannot be satisfied. Thus all the poles are in the lower half of the complex $\theta$ plane. Similarly $f=0$ for $x_2=t-x<0$ since the integral is invariant under interchanging $x_2$ and $u$ with $x_1$ and $\theta$. Thus we have demonstrated the existence of a Green's function for our equation with support only in the future light cone.

---

PHYSICAL REVIEW VOLUME 182, NUMBER 5 25 JUNE 1969

# Response to the Variation of an External Gravitational Field and Its Implications for Equal-Time Commutators

J. C. Katzin and William B. Rolnick
*Wayne State University, Detroit, Michigan 48202*
(Received 25 November 1968; revised manuscript received 23 January 1969)

Conjectures on the response of currents and the stress-energy tensor to variations of external gravitational fields have led to important statements about equal-time commutators (ETC's). We consider the case where only $g_{00}$ is varied and prove that those conjectures are *required* by the general covariance of tensor fields. Although currents are known to be limits of nonlocal expressions, these limits may be taken before we vary $g_{00}$, and our theorem is applicable. However, physical properties of the vacuum preclude $T^{ij}$ from behaving this way. To resolve this inconsistency, we conjecture that $T^{\mu\nu}$ is the limit of a nonlocal expression. (A reformulation of Lagrangian field theory may be necessary.) We also discuss possible difficulties arising in the derivation of the ETC's using Schwinger's variational method, and show that our theorem ensures that these difficulties are not present for ETC's of $T^{00}$ with tensor fields.

## I. INTRODUCTION

An elegant method for obtaining equal-time commutators (ETC's) results from Schwinger's variational method.[1] The extra terms arising in ETC's of currents (Schwinger terms) are relevant to any attempts to formulate a theory of currents. In order to prove various properties of Schwinger terms, Gross and Jackiw[2] have assumed a form for the commutators of currents with the energy density. They have indicated the behavior of the currents, as the metric tensor is varied, which would lead to this form (employing Schwinger's variational method). Conjectures have also been made about the behavior of the stress-energy tensor as the metric tensor is varied.[1]

In this paper we investigate the behavior of local tensor operators as $g_{00}$ is varied. We will prove that the behavior of any local tensor field (which is to be a tensor under general coordinate transformations), as we vary $g_{00}$, is completely determined by the number of indices which are zero. (The sense in which we use the word "local" may be found in Sec. II.) We argue that currents other than $T^{\mu\nu}$ may be considered local. Schwinger's variational method, in the presence of an external gravitational field, is discussed in light of the above. We show that the usual derivation of ETC's involving $T^{00}$ depend upon the absence of $\partial_0 g_{00}$ from the Lagrangian and the field equations. We prove that $\partial_0 g_{00}$, indeed, does not appear in the Lagrangian or the field equations, thus validating the ETC's of $T^{00}$ with local operators. Our results imply that the proof in Ref. 2 is valid. The validity of the derivations of those ETC's of special relativity which involve $T^{00}$ (but not $T^{ij}$) follows as well.

---

[1] J. Schwinger, Phys. Rev. **130**, 406 (1963).

[2] D. J. Gross and R. Jackiw, Phys. Rev. **163**, 1688 (1967). For currents associated with gauge fields, Jackiw has verified the ETC of $J^0$ with $T^{00}$ (with a particular form of $T^{00}$): R. Jackiw, *ibid.* **175**, 2058 (1968).