Causality and superluminal behavior in classical field theories: Applications to k-essence theories and modified-Newtonian-dynamics-like theories of gravity

Jean-Philippe Bruneton

To cite this version:
Jean-Philippe Bruneton. Causality and superluminal behavior in classical field theories: Applications to k-essence theories and modified-Newtonian-dynamics-like theories of gravity. Physical Review D, 2007, 75, 10.1103/PhysRevD.75.085013. hal-03646612

HAL Id: hal-03646612
https://hal.science/hal-03646612
Submitted on 29 Apr 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Causality and superluminal behavior in classical field theories: Applications to $k$-essence theories and modified-Newtonian-dynamics-like theories of gravity

Jean-Philippe Bruneton*

GReCO, Institut d’Astrophysique de Paris, UMR 7095-CNRS, Université Pierre et Marie Curie–Paris 6, 98 bis boulevard Arago F-75014, Paris, France

(Received 4 October 2006; published 19 April 2007)

Field theories with Lorentz (or diffeomorphism invariant) action can exhibit superluminal behavior through the breaking of local Lorentz invariance. Quantum induced superluminal velocities are well-known examples of this effect. The issue of the causal behavior of such propagation is somewhat controversial in the literature and we intend to clarify it. We provide a careful analysis of the meaning of causality in classical relativistic field theories and stress the role played by the Cauchy problem and the notion of chronology. We show that, in general, superluminal behavior threatens causality only if one assumes that a prior chronology in spacetime exists. In the case where superluminal propagation occurs, however, there are at least two nonconformally related metrics in spacetime and thus two available notions of chronology. These two chronologies are on equal footing, and it would thus be misleading to choose ab initio one of them to define causality. Rather, we provide a formulation of causality in which no prior chronology is assumed. We argue that this is the only way to deal with the issue of causality in the case where some degrees of freedom propagate faster than others. In that framework, then, it is shown that superluminal propagation is not necessarily noncausal, the final answer depending on the existence of an initial data formulation. This also depends on global properties of spacetime that we discuss in detail. As an illustration of these conceptual issues, we consider two field theories, namely, $k$-essence scalar fields and bimetric theories of gravity, and we derive the conditions imposed by causality. We discuss various applications such as the dark energy problem, modified-Newtonian-dynamics-like theories of gravity, and varying speed of light theories.

DOI: 10.1103/PhysRevD.75.085013

PACS numbers: 03.50.+z, 04.20.Gz, 95.30.Sf, 95.36.+x

I. INTRODUCTION

The question of the causal behavior of superluminal propagation has often been debated, in rather different contexts: tachyonic particles (e.g. [1–5]), field theories [6–13], and quantum induced superluminal propagations [3,14–17]. The issue is, however, still controversial and we intend to clarify it.

We will focus on classical field theories with Lorentz invariant action (or diffeomorphism invariant action whenever Einstein’s gravity is taken into account). Superluminal behavior may then arise when local Lorentz invariance is spontaneously broken by nontrivial backgrounds (or vacua), whereas some other sector of the theory is left unbroken [18]. In quantum electrodynamics, superluminal propagations induced by vacuum polarization provide well-known examples of this phenomenon [3,14–16].

We define superluminal behavior as going faster than gravitons (i.e., gravitational waves), and $c$ denotes the speed of gravitons in vacuum. This unconventional but convenient definition will not affect the argument that follows. In the standard theory of gravity, photons propagate along the gravitational metric so that, in this case, the definition agrees with the usual one.

Our general analysis will enable us to investigate the causal behavior of fields coupled to standard gravity. We will illustrate our arguments in three important cases: $k$-essence scalar fields, bimetric theories of gravity, and quantum induced superluminal propagations. We will make constant use of many important ideas and tools that were developed in order to investigate the causal behavior of general relativity, and we notably refer to [19,20].

In Sec. II, we analyze in detail the meaning of causality in classical field theories. We insist on the important role played by the Cauchy problem and by the notion of (local and global) chronology. As an illustration, we recall how causality is usually formulated in the theory of general relativity. We stress that superluminal propagation is automatically discarded if one assumes that causality should be defined with respect to the chronology induced by the gravitational metric. This postulate actually “sets the [gravitational] metric apart from the other fields on [the manifold] and gives it its distinctive geometrical character” [19]. This prior assumption cannot be supported by any mathematical reason but only by experiment. If superluminal propagation was found in the laboratory, it would thus not mean that causality, locality, or the entire framework of special relativity is lost, but simply that the gravitational field does not have any fundamental geochronological character. In that case, causality should simply not be expressed in terms of the chronology induced by the gravitational metric field. More generally, it is necessary to drop such prior assumptions to address the issue of the causal behavior of superluminal fields. In Sec. II C we thus

*Electronic address: bruneton@iap.fr
look for a minimal expression of causality in which no prior chronology is assumed. In that framework, superluminal propagation is generically allowed.

Before moving to superluminal but causal theories, we display two well-known examples of noncausal theories. This enables us to stress how noncausal behavior is related to constraints on initial data. We also comment on the causal “paradoxes” that arise in a noncausal theory. We finally briefly investigate what kind of field theory can lead to superluminal propagation.

We consider in Sec. III a \( k \)-essence scalar field \( \varphi \) which can propagate superluminally along an effective metric \( G^{\mu \nu}[\varphi] \). We find that the scalar field is causal, provided that the spacetime embedded with the effective metric \( G^{\mu \nu}_0 \) is globally hyperbolic. Here, \( G^{\mu \nu}_0 \) is the effective metric evaluated on a solution \( \varphi_0 \) of the field equation (the background). This condition puts some generic constraints on the free function defining the theory, and is generally satisfied on reasonable backgrounds. We emphasize that the claim that such scalar fields are not causal even on globally hyperbolic backgrounds [6,7,13,21] can only hold if a prior role is attributed to the gravitational (or flat) metric, an assumption that, we believe, was implicit in these works. The only threat for causality actually lies in global properties of backgrounds [7], and this can be related to the so-called chronology protection conjecture.

In Sec. IV, we investigate the question of causality in a bimetric theory of gravity, in which the matter sector is universally coupled to a metric \( \bar{g} \) that can differ from the Einstein-Hilbert metric \( g \). These two metrics generally define two distinct “causal” cones. Photons can travel faster than gravitons and vice versa. If causality is defined with respect to the chronology induced by the gravitational metric, then causality forces the matter metric \( \bar{g} \) to define a cone which coincides with or lies within the gravitational cone, and vice versa. These two choices of chronology have already been considered in the literature, and lead to opposite requirements on the theory. This illustrates the fact that there are no clear reasons why a metric, or a chronology, should be preferred over the other. We argue that the theory is actually causal if no prior chronology is assumed. The solution to the Cauchy problem, however, depends on the precise dynamics of the matter sector.

In Sec. V, we briefly show how our analysis applies to the case of quantum induced superluminal propagation. This enables us to conclude that superluminal velocity does not threaten causality, a point that is still controversial to date.

Let us emphasize that the question of causality in \( k \)-essence or bimetric theories is not just of academical interest, since these theories have drawn much attention recently. Quite generally, it is claimed that causality forces one of the cones to be wider than the other one (be it the gravitational or the matter one) [6–9,11,13], and this places some constraints on the theory. Our definition of causality, which is adapted to the multimetric case, will not, however, support these claims.

Such \( k \)-essence scalar fields were notably used as a dynamical dark energy fluid responsible for the late time acceleration of the Universe [22], as well as a fluid that can drive inflation [23]; see also [24–26]. It is also used as a new gravitational field in addition to the metric one in some relativistic theories of modified Newtonian dynamics (MOND), which are intended to account for the mass discrepancy at astrophysical scales without invoking dark matter [13,21,27]. Some of these models were notably discarded because of the presence of superluminal propagation. Bimetric theories are also an essential piece of some recent MOND-like theories [13,28] (see also [29]). Moreover, they represent the best motivated framework that reproduces a varying speed of light (VSL) scenario [30,31] which may address the problems of standard cosmology in a rather different way from inflation. Indeed, some VSL theories, in which the speed of light \( c \) is replaced by a changing velocity \( c(t) \) inside the equations of motion [32], however interesting phenomenologically, are not satisfying theoretically [31,33] (i.e., the resulting theory cannot be derived from an action).

In the last section, we consider some applications of our work to these field theories. We investigate the link between causality and stability in \( k \)-essence theories. We show that (\( k \)-essence) ghost stabilization might suffer from a serious problem. We also show that, if one reproduces the MOND phenomenology with the help of a \( k \)-essence scalar field, then a slight modification of Milgrom’s law is necessary for the theory to be causally well behaved. It leads to a nontrivial modification of the phenomenology in the very low acceleration regime. We also comment on a simpler theory that reproduces the MOND phenomenology, using only one scalar field. The initial value formulation of this theory still has to be checked and we leave it for future work. We briefly show how this framework can account for the Pioneer anomaly, and we finally make some comments about VSL theories.

We use, throughout the paper, the sign conventions of [34] and, in particular, the mostly + signature. The flat metric is denoted \( \eta \) or \( \eta_{\mu \nu} \) in a coordinate system. We denote by \( g \) the gravitational metric field (which obeys the Einstein equations) and \( \bar{g} \) the matter metric to which matter fields couple. Unless otherwise specified, indices are raised or lowered using the gravitational metric \( g \).

II. CAUSALITY AND SUPERLUMINAL BEHAVIOR

In the present paper, we are interested in the superluminal behavior of some matter fields in a general relativity-like context. The metric field \( g \) follows the dynamics induced by the Einstein-Hilbert action. By spacetime we will always mean a set \((\mathcal{M}, h)\), where \(\mathcal{M}\) is a four-dimensional differentiable manifold and \( h \) some nondegenerate Lorentzian metric on it. A superluminal signal, by
definition, propagates along spacelike curves of the gravitational metric $g$.

A. Causality, chronology, and the flow of time

By causality we usually mean the ability to find a cause to an effect. Since cause and effect are both described in terms of some physical variables, the usual principle of causality states that physical variables should be unambiguously determined at a given time from their values at a time before. Conversely, we should also be able to predict a future situation from a present one. The principle of causality thus requires that determinism holds “in both directions of time.” Causality demands the existence and uniqueness of solutions to the equations of motion given some initial data. In mathematical words, equations of motion must have a well-posed Cauchy problem, or initial value formulation.

Note that this definition shows that time ordering, or chronology, must exist between two spacetime points that are causally connected. Unlike Newton’s theory where a global chronology preexists the dynamics because of the prior topology assigned to spacetime, namely $\mathbb{R}^4 \times \mathbb{R}$, relativistic field theories (including Einstein’s gravity) do not involve a preexistent notion of time and chronology. On the contrary, any relativistic field $\psi$ defines its own chronology on $\mathcal{M}$ by means of the Lorentzian metric $h$ along which it propagates. Any Lorentzian metric indeed induces a local chronology in the tangent space through the usual special relativistic notions of absolute (i.e., Lorentz invariant) future and past.

Let us consider one particular field $\psi$ that propagates along a metric $h$. It is important to note that this metric may not induce a global chronology on the whole spacetime. Indeed, since spacetime may be curved and, moreover, may have nontrivial topology, the existence of a partial ordering in the tangent space (the local chronology of special relativity) does not imply that a partial ordering over the whole manifold exists. For instance, the local causal cones of $h$ can be distributed on the manifold in such a way that there exists a curve over $\mathcal{M}$ which is everywhere future-directed and timelike but closed. In that case, the field $\psi$ can propagate along a closed curve and an event could be both the cause and the effect of another event. Causality requires that this does not happen. The strongest way to prevent it is to require the global hyperbolicity of the spacetime $(\mathcal{M}, h)$.

Let us emphasize that causality also requires the notion of time flow. It is indeed worth recalling that, even if the above metric $h$ defines a global notion of future and past, we must also require that the field $\psi$ can only propagate in the future. If it could, on the contrary, propagate both in the future and in the past, one could always form a closed timelike curve with it. On the other hand, global hyperbolicity is enough to guarantee that no closed future-directed timelike curve exists.

When the manifold is embedded with a finite number of metrics $h_i$, the discussion of global properties is slightly more involved; see Sec. II C.

B. Causality in general relativity

Let us illustrate the previous discussion. In general relativity, causality is generally expressed by the following properties [19,20]:

(a) The gravitational metric $g$ must define a global chronology in spacetime.

(b) “The null cones of the matter equations coincide with or lie within the null cone of the spacetime metric $g$” [19].

(c) The whole set of equations of motion must admit a well-posed Cauchy problem.

Point (a) guarantees the existence of a global time ordering on $\mathcal{M}$ and point (c) is the formal expression of determinism. Note that this is a nontrivial mathematical property for differential systems. Satisfying (c) in general relativity critically depends on the precise dynamics of the matter sector.

Point (b) excludes faster-than-graviton propagation. Equivalently, it means that any initial data set on spacelike hypersurfaces with respect to the gravitational metric $g$ is allowed. Since the metric $g$ defines a global chronology in spacetime, so does any other metric $h_i$ associated with the propagation of some matter field $\psi_i$. Both conditions (a) and (b) therefore ensure that no closed future-directed and timelike curves exist for any metric.

Let us emphasize that, in the formulation (a)–(c) above, the gravitational metric field clearly plays a preferred role. Both conditions (a) and (b) are such that causality is actually defined with respect to the chronology induced by the metric $g$. This can be understood as a “postulate which sets the metric $g$ apart from the other fields on $\mathcal{M}$ and gives it its distinctive geometrical character” ([19], Sec. 3.2, emphasis added). The chronology defined by $g$ is thus assumed to be the preferred one on $\mathcal{M}$.

This postulate is, however, arguable because, as we pointed out, any relativistic field induces its own chronol-
ogy in spacetime. In particular, it means that, if a signal made up of waves of some field $\psi_i$ propagates between two spacetime points, then these points can be time ordered with the help of the metric $h_i$, and causality may be preserved even if the field propagates superluminally. The gravitational metric field $g$ is just one particular field on $\mathcal{M}$, and there are no clear reasons why it should be favored.

C. Causality and superluminal behavior

We have already stressed that superluminal behavior is automatically discarded by the above postulate. One may try to justify it by arguing that, since $g$ reduces locally to $\eta$ and therefore gives to the tangent space its special relativistic (Minkowskian) structure, going faster than gravitons would “undermine the entire framework of relativity theory” [20]. Let us, however, stress that any other metric $h_i$ can also be reduced locally to its fundamental form in the appropriate “inertial” coordinates. These coordinates transform under the action of the Lorentz group $SO(3, 1)$ with an invariant speed $c$; that can differ from $c$ if the cones of $h_i$ and $g$ do not coincide (due, for instance, to some spontaneous breaking of local Lorentz invariance; see Sec. II G). These coordinates are actually relevant if one uses rods and clocks made up of the field $\psi$ propagating along $h$. The fact that $g$ reduces locally to the flat spacetime metric can therefore not be used to claim that $g$ should be a preferred field on $\mathcal{M}$. We will come back to this important point in Sec. IV B.

We may therefore drop the above postulate and look for a minimal expression of causality with which no prior chronology is assumed. Let us consider a collection of fields $\psi_i$ that propagate along some metrics $h_i$ (gravity included). We are led to the picture of a finite set of causal cones at each point of spacetime. We do not want to prefer one metric with respect to the others, because there is no reason why, locally, some sets of coordinates, or some rods and clocks, should be preferred: coordinates are meaningless in general relativity.

The cones defined by the metrics $h_i$ may therefore be in any relative position with respect to one another. Moreover, these cones may even tip each other over depending on the location in spacetime. We also have to guarantee that no signal propagates along a closed curve. For this it is not sufficient to require the global hyperbolicity of each spacetime $(\mathcal{M}, h_i)$. Indeed, fields may interact together, and we may form a physical signal which propagates in spacetime along different metrics.

We shall instead define an extended notion of future-directed and timelike curves, by requiring that, at each point of the curve, the tangent vector is future-directed and timelike with respect to at least one of the metrics $h_i$. Such a construction has already been advocated in [35]. All the (extended) notions of future, past, domains of dependence, achronal sets, Cauchy surfaces, global hyperbolicity then follow. This corresponds to a “mixed” notion of chronology: a point $P$ is in the (extended) future of $Q$ if it is in the future of $Q$ for at least one of the metrics $h_i$; see Fig. 1. A spacetime interval is thus spacelike in the extended sense if it is spacelike with respect to all metrics $h_i$.

Causality then requires that a global (mixed) chronology in spacetime exists. In other words, spacetime must be globally hyperbolic in that extended sense. This straightforward generalization enables us to express the minimal formulation of causality by the following properties:

(i) The mixed chronology defined by the set of metrics $h_i$ must be a global chronology in spacetime.

(ii) The whole set of equations of motion must admit a well-posed Cauchy problem.

The formulation of causality (a)–(c) is immediately obtained from these more general requirements whenever the matter light cones coincide with or lie within the gravitational one. We thus see that fields can propagate superluminally without threatening causality, provided that we do not refer to any preferred chronology in spacetime. That the Cauchy problem of some field $\psi$ be well posed depends on the precise equation of motion, and on whether initial data are set on surfaces that are spacelike with respect to the metric $h$ along which the field propagates. Moreover, if we are interested in the whole theory (gravity included), the Cauchy problem will be well posed if initial data are set on surfaces that are spacelike in the extended sense, that is, spacelike with respect to all metrics $h_i$.

D. Global properties

It must be stressed that, in general relativity, the fact that $g$ defines a global chronology cannot be proven, because local physics does not determine the topology of spacetime, which could, however, prevent the existence of a global chronology. This is the case, for instance, of non-

FIG. 1. The hatched part shows the extended future defined by two metrics (solid and dashed lines) in the case where one metric defines a wider cone than the other one (left) and vice versa (right).
time-orientable spacetimes. Moreover, there exist exact solutions to the Einstein equations that do not describe globally hyperbolic spacetimes, as explicitly shown by the Kerr solution, which possesses closed timelike curves, or Gödel’s universe. Einstein’s equations thus admit solutions that violate causality.

This difficulty is entirely subsumed into the so-called chronology protection conjecture [36], which asserts that the local laws of physics are such that they prevent the formation of closed future-directed and timelike curves in spacetime. 5 This has not been proven yet, so we usually assume from the beginning that spacetime is globally hyperbolic in the extended sense. Again, if we were able to prove the chronology protection conjecture, we would not have to make such a nontrivial assumption. We will discuss this point further in Sec. III D.

Let us stress that in our extended framework, and notably in the presence of superluminal matter fields, we will not be able to prove that condition (i) holds. We will have to assume that spacetime is globally hyperbolic in the extended sense. As long as the conjecture is not proven, such a restriction is not dynamical, but of epistemological nature [37].

In other words, condition (i) will have to be imposed by hand, just as in general relativity. As we already explained, we will not consider the issue of the existence of the flow of time, and thus, in the following, we will mostly be concerned with condition (ii). We provide in the next section two examples of theories that do not admit a well-posed Cauchy problem. It enables us to show the deep relationship between noncausality, closed curves in spacetime and constraints on initial data.

E. Noncausal theories: Closed curves and constraints on initial data

A manifestly noncausal theory is an elliptic Klein-Gordon scalar field \( \varphi \) with an equation of motion

\[
h^{\mu \nu} \partial_\mu \partial_\nu \varphi = 0,
\]

where the metric \( h \) has the signature \( \pm 4 \). This elliptical equation does not have a well-posed Cauchy problem. Such a Euclidean metric does not select the time coordinate as special compared to the spatial one, and initial data on three-surfaces cannot be propagated in four dimensions. On the contrary, a Lorentzian signature for \( h \) guarantees that the equation is hyperbolic and that, by virtue of well-known theorems [20], the Cauchy problem is well posed.

We now consider the more interesting case of tachyonic particles. By tachyons we mean particles or signals that can be sent at superluminal speeds relative to the emitter. 6 As correctly recognized in many papers in the literature, tachyons are not causal (see, e.g., [2,3]). Indeed, they can always be used to construct closed curves in spacetime along which a (tachyonic) signal propagates. Consider for instance an observer \( A \) sending, at time \( t_0 \) (event \( E_0 \)), a signal to observer \( B \) (event \( E_1 \)), who in turn sends a signal back to \( A \), at time \( t_1 \) (event \( E_1 \)). This last signal can be received by \( A \) at a time \( t_2 < t_0 \) (event \( E_2 \)), if signals can be sent from \( B \) at a superluminal speed and if \( A \) and \( B \) are in some relative motion at a speed \( v < c \); see Fig. 2. Information can therefore propagate along a closed curve in spacetime, in this case, curve \( (E_0, E_1, E_2, E_0) \).

The theory is not causal because an event along this curve will influence itself. Such an event is thus not freely specifiable, and we see that information propagation along a closed curve in spacetime always leads to constraints on initial data. On the contrary, if the Cauchy problem were well posed for such a theory, it would mean that there exist surfaces on which initial data can be specified freely (up to possible constraints arising from gauge invariance, if any) and unambiguously evolved in time.

FIG. 2. The closed curve followed by the tachyonic signal viewed in the rest frame of \( A \). The tachyon is sent by \( A \) at time \( t_0 \) (event \( E_0 \)) and received by \( B \) at time \( t_1 \) (event \( E_1 \)). Since \( B \) moves with respect to \( A \), the tachyonic signal he sends back to \( A \) is actually received before it was sent, at time \( t_2 < t_0 \) (event \( E_2 \)). The dashed line represents the Minkowski cone, and horizontal and vertical lines are the space and time axes in the frame of \( A \).

F. Closed curves and temporal paradoxes

Before going further, let us comment on the previous experimental device. Suppose (situation 1) that, when \( B \) receives a signal from \( A \), he sends a signal backward in time to \( A \). Then \( A \) receives a signal at time \( t_2 \) and is thereby forced to send a signal to \( B \) an instant later. It explicitly shows the constraints that appear in the case of propagation along closed curves.

A temporal (or causal) paradox can arise if we consider the following experimental device (situation 2): when \( B \) receives a signal from \( A \), he sends a signal back to \( A \), and in
turn $A$ sends a signal to $B$ only if he did not receive anything from $B$. One may wonder what happens in that case. If $A$ receives a signal from $B$, he will not send a signal to $B$ who in turn would not send a signal back to $A$, and this contradicts the hypothesis. The other possibility, when $A$ does not receive anything from $B$, is also inconsistent. This is, in fact, a version of the famous “grandfather paradox.”

Remarkably, there is no answer to the question “what happens?” It is, however, immediate to see that this question implicitly assumes that the knowledge of the initial data at time $t_2$, for instance ($A$ receives a signal from $B$ or not), is enough to determine what will happen “later” along the curve. This question thus assumes that causality (or determinism) holds, whereas it cannot even be defined, since there is no available notion of time ordering along a closed curve. On the contrary, we have seen that such a closed curve leads to constraints on initial data. The event at time $t_2$ is strongly correlated to itself so that the only relevant question reads “are either of the two initial conditions, $A$ receives—or does not receive—a signal from $B$, allowed?” Clearly, they are not.

Let us emphasize that situation 2 (the grandfather’s paradox) is often taken as an example of bad causal behavior induced by tachyonic particles (or, accordingly, induced by the use of time machines). This point is not, however, justifiable because the paradox only arises if we ask the wrong question. This experimental device simply cannot exist since there are no initial data consistent with it. This is not related to causality but only to logic: temporal paradoxes do not exist. On the other hand, the noncausal situation 1 can exist (constrained initial data exist that correspond to this experiment).

G. Superluminal behavior in field theories

Tachyonic particles propagate superluminally and violate causality. It does not mean, however, that any superluminal propagation is noncausal. Indeed, the closed curve in Fig. 2 can only be constructed if the speed of the tachyonic signal depends on the speed of the emitter. This is not the case in the metric description of superluminal propagation that we will mostly be concerned with.

Here we wish to investigate briefly what a field theory that involves superluminal propagation may look like. In a theory with Lorentz invariant action, the only possibility to open up the causal cones is to modify the propagational part of the dispersion relation. This can be achieved by adding higher-order derivatives in the action. Notice, however, that such theories may be unstable [38]. Superluminal behavior, for instance, arises from quantum corrections to electrodynamics in nontrivial vacua; see Sec. V.

On the other hand, we could still have second order, but nonlinear, equations of motion. We will consider two examples of such theories in the following sections. This nonlinearity, which is essential to obtain unusual dispersion relations, leads to a background-dependent Cauchy problem. Accordingly, Lorentz invariance of the action is spontaneously broken through nontrivial backgrounds and it allows superluminal propagation. Note also that the speed of small perturbations around the background only depends on the background and not on the motion of the emitter. Superluminal signals in field theories are therefore not tachyonic in the sense indicated in the previous section.

Note that we could also have considered non-Lorentz invariant Lagrangians like $L = -\phi^2 + u^2(\nabla \phi)^2$, where $u \neq c$. We will not consider such Lagrangians in the present paper. In practice, the theories we will consider will lead to similar field equations, but only through a spontaneous breaking of local Lorentz invariance by nontrivial backgrounds.

III. $k$-ESSENCE FIELD THEORY

A. Field equation and the Cauchy problem

The action for a $k$-essence scalar field theory reads

$$ S = -m^4 \int \sqrt{-g} d^4x \left[ F \left( \frac{X}{m^2} \right) + V(\varphi) \right], \quad (1) $$

where

$$ X = \frac{g^{\mu \nu} \partial_{\mu} \varphi \partial_{\nu} \varphi}{2}, \quad (2) $$

and $F$ and $V$ are some functions, $m$ is a mass scale, and $h = 1$. Hereafter, we also take $m = c = 1$. Note that the more general form $L = -F(X, \varphi)$ is often considered in the literature, but the following arguments would also apply to this case. The field equation reads

$$ G^{\mu \nu} \partial_{\mu} \varphi \partial_{\nu} \varphi - V'(\varphi) = 0, \quad (3) $$

where the effective metric $G^{\mu \nu}$ is given by

$$ G^{\mu \nu} = F'(X) g^{\mu \nu} + F''(X) \partial_{\mu} \varphi \partial_{\nu} \varphi, \quad (4) $$

and a prime denotes derivation with respect to $X$. Scalar waves propagate inside the “scalar cone” defined by this effective metric. This cone generally differs from the gravitational one unless $F''(X_0) = 0$; see Sec. III B.

This equation is a special case of quasilinear second order differential equations. If we first neglect the coupling to gravity which will be considered below, we may then invoke a theorem due to Leray that proves [20] that this equation has a well-posed Cauchy problem if spacetime

---

Footnote 7: Explicitly, one assumes that the tachyon propagates at some speed $u > c$ in the rest frame of $B$, and then deduces, using a Lorentz boost, that it travels backward in time in the frame of $A$; see [3].
(\mathcal{M}, G_0^{\mu \nu}) is globally hyperbolic, where \(G_0^{\mu \nu}\) is the effective metric evaluated on a solution \(\varphi_0\) to the field equation (3). Hereafter, we will refer to \(\varphi_0\) as the background. The initial data have to be specified on three-surfaces that are spacelike with respect to the background metric \(G_0^{\mu \nu}\), in connection with the discussion of Sec. II C.

A necessary but insufficient condition for global hyperbolicity is the Lorentzian signature of the background metric \(G_0^{\mu \nu}\), which ensures, in particular, that the field equation is hyperbolic. By diagonalization of the matrix \(G_0^{\mu \nu}\), we find that its signature is +2 everywhere over \(\mathcal{M}\) if and only if

\[
F'(X_0) > 0,
\]

\[
F'(X_0) + 2X_0F''(X_0) > 0,
\]

where \(X_0 = \eta^{\mu \nu} \partial_\mu \varphi_0 \partial_\nu \varphi_0 / 2\). In particular, if the function \(F\) is such that the above inequalities are satisfied for all \(X\), the metric \(G_0^{\mu \nu}\) will have a signature +2 on any background. Note that the sign of the above inequalities in Eq. (5) can be reversed. This would, however, correspond to a metric \(G_0^{\mu \nu}\) with signature -2, and the scalar waves would carry negative energy. The theory would therefore be unstable when coupled to other fields (notably gravity).

We exclude here this possibility, although it should be stressed that it comes from a stability argument, and not from some causal requirement.

In the presence of gravity the Cauchy problem has to be solved simultaneously for the gravitational variables and for the matter ones. It has been proved (see for instance [19]) that the Cauchy problem is well posed if matter fields satisfy “reasonable” equations of motion and if the stress-energy tensor of the matter fields only involves matter variables, the gravitational metric, and their first derivatives. This is the case for \(k\)-essence theories, as can easily be checked. Locally, the whole theory has a well-posed Cauchy problem.

B. Superluminal behavior

The analysis of the characteristic\(^8\) of the field equation (3) shows that the scalar field propagates superluminally if \(F''(X_0) > 0\), where we use \(F'(X_0) > 0\) by virtue of Eq. (5a).

As it will be useful later, let us remind the reader that any field which propagates along an effective metric of the form (up to some positive conformal factor)

\[
H^{\mu \nu} = g^{\mu \nu} + Bn^\mu n^\nu,
\]

where \(B\) and \(n^\mu\) are, respectively, some scalar and vector field, is superluminal if \(B > 0\). Note that the opposite sign for \(B\) is found in [6,7] because of the opposite choice of signature.

C. “Local causality” and the choice of initial data surfaces

The above theorem and our analysis of the meaning of causality in Sec. II enable us to conclude that \(k\)-essence theories are causal in flat spacetime whenever the background \((\mathcal{M}, G_0^{\mu \nu})\) is globally hyperbolic, even in the presence of superluminal scalar waves. Again, this conclusion holds only if we do not refer to any preferred chronology.

It was, however, claimed [6,7,13,21] that \(k\)-essence theories are not causal even if the background is globally hyperbolic. Here we emphasize that these claims can only be supported if it is assumed that the gravitational metric induces a preferred chronology on \(\mathcal{M}\), an assumption that, we believe, is implicitly made in these references. The main argument which rules out superluminal propagation is indeed that the Cauchy problem for the scalar field is not well posed for initial data that are set on surfaces which are spacelike with respect to the flat metric but timelike or null with respect to the background metric. In that case, initial data cannot be evolved because of caustics [7], and the Hamiltonian formalism is singular [6].

These difficulties are however not surprising. Indeed, this only shows that initial data surfaces for the scalar field must be spacelike with respect to the background metric, as stated by Leray’s theorem. If initial data are set on these surfaces, the theory is free of such a singular behavior. This claim therefore only arises from an unadapted choice of initial data surfaces, i.e., from the postulate that the gravitational metric defines a preferred chronology. This postulate actually prevents any superluminal propagation, as we explained in Sec. II B. In such a case, superluminal behavior is ruled out from the beginning in a rather ad hoc way, but not by some intrinsic (mathematical) argument.

D. Global properties

The only threat for causality lies, in fact, in the global properties of the background. It has been correctly recognized in [7] that the spacetime \((\mathcal{M}, G_0^{\mu \nu})\) may not always be globally hyperbolic. In particular, closed timelike curves (with respect to the effective metric) could exist.

Let us first stress that, in general, however, the spacetime \((\mathcal{M}, G_0^{\mu \nu})\) is globally hyperbolic. This is the case in trivial backgrounds \(\varphi_0 = \text{const}\) and in nontrivial but homogeneous (in a certain Lorentz frame) backgrounds \(\partial \varphi_0 = \text{const} \neq 0\), which may be relevant in a cosmological context. Note that these backgrounds have to be solutions to the equation of motion, Eq. (3). In the first case, \(\varphi_0\) thus has to be an extremum of the potential, whereas in the second one, \(\varphi_0\) is not constant in spacetime and the potential must be flat over some range or, more simply, vanish. In these

\(^{8}\)Notably, when they form a quasilinear, diagonal and second order hyperbolic system of equations.

\(^{9}\)See [39], Ch. 5, Appendix 1.
two important cases, the spacetime \((\mathcal{M}, G_0^{\mu \nu})\) is globally hyperbolic and, by virtue of the previous theorem, the Cauchy problem is well posed and the theory is causal.

Since the global hyperbolicity of spacetime \((\mathcal{M}, g)\) must be assumed in general relativity, as we explained in Sec. II D, we may also assume that spacetime is globally hyperbolic in the extended sense; see Sec. II C. This at least requires that inequalities in Eq. (5) hold, and automatically ensures that the whole theory of the \(k\)-essence scalar field and gravity is causal. The relevant Cauchy surfaces in that case are hypersurfaces that are spacelike with respect to all metrics (gravitational, scalar, and others matter metrics).

Note that this is a nontrivial (and maybe arguable) assumption. Let us, however, stress that it must also be assumed in the context of general relativity and standard matter alone so that the fact that nontrivial global properties may break causality does not appear to be rooted in superluminal propagation. This assumption may actually be linked to the chronology protection conjecture, as shown by the following example. Let us consider the highly nontrivial and nonglobally hyperbolic background invoked in [7]. It consists of two “bubbles” of nontrivial backgrounds \(\partial \varphi_0 = \text{const} \neq 0\) that move rapidly in opposite directions with a finite impact parameter. The space is otherwise empty (trivial background \(\varphi_0 = \text{const}\)). Small scalar perturbations thus travel superluminally inside the two bubbles and along null rays of the flat metric outside them. The fact that the two bubbles are in relative motion implies the existence of closed future-directed and timelike curves (timelike with respect to the background metric \(G_0^{\mu \nu}\)).

It should, however, be stressed that such a background is not relevant since it is not a solution to the equation of motion, Eq. (3). Indeed, the derivative of the field \(\varphi_0\) is not continuous at the boundary between the bubbles and empty space. Some Dirac-like source terms should be added to the equation of motion to accommodate such a background. On the contrary, a physical solution involving two such bubbles should exhibit a continuous transition from their interior to the empty space. Of course it may be the case that, even in this more realistic situation, the background is still not globally hyperbolic, but the contrary could also happen, and it would be an illustration of the chronology protection conjecture. We will not try to perform this analysis.

A very similar case was discussed in [3] in the context of the Casimir experiment; see Sec. V. Note that an analogous case has also been found in the context of general relativity by Gott [40], who showed that two straight infinite cosmic strings moving in opposite directions with a finite impact parameter lead to the formation of closed future-directed and timelike curves in spacetime. Again, it shows that difficulties with causality at a global level already exist in the context of general relativity and subluminal matter and should not be related to some intrinsic problems of superluminal propagation.

### IV. BIMETRIC THEORIES OF GRAVITY

#### A. Definition

By multimetric theories of gravity we mean theories of gravity in which some degrees of freedom in the matter sector are coupled to some matter metrics \(\tilde{g}\), that are distinct from the gravitational one \(g\). The fact that different matter fields are coupled to different metrics breaks the weak equivalence principle (WEP), which has been tested with great accuracy. As a special case, a bimetric theory of gravity is a theory where all the matter fields are coupled to the same metric \(g\). This ensures that the WEP is satisfied. The theory of general relativity just corresponds to the choice \(\tilde{g} = g\).

A typical example is a scalar-tensor theory of gravity in which \(\tilde{g} = \Omega^2 g\), at each point of \(\mathcal{M}\), and where \(\Omega\) is a smooth real-valued function over \(\mathcal{M}\). Dynamics of the scalar field \(\Omega\) arise from a standard kinetic term (but could also be of \(k\)-essence type). Because of the conformal relationship between \(g\) and \(\tilde{g}\), the two cones defined by these two metrics coincide, and there is no superluminal propagation.

In general, however, we could have a nonconformal relation between these two metrics. Consider, for instance, the so-called disformal relation [11]

\[
\tilde{g}_{\mu \nu} = A^2 (g_{\mu \nu} + BU_\mu U_\nu),
\]

where \(A\) and \(B\) are some functions of the scalar quantity \(U_\mu U^\mu\), and \(U\) is a vector field. When this matter metric is of Lorentzian signature, the matter light cone (defined by \(\tilde{g}\)) can be wider than the gravitational cone (defined by \(g\)), depending on the sign of \(B\). This disformal relation is an essential piece of some recent relativistic field theories of the MOND paradigm [13,28].

#### B. Superluminal behavior

Throughout this paper, we have defined superluminality as the propagation along spacelike curves of the gravitational metric. But we could also have defined it as going faster than light (photons). These definitions coincide in general relativity but not in bimetric theories since \(g \neq \tilde{g}\). In that framework, there are two opposite definitions of superluminal behavior. Photons can travel faster than gravitational waves and vice versa. It is thus not clear which of these two superluminal behavior we have to worry about.

This is closely related, again, to the choice of a preferred chronology. If one assumes that the gravitational metric field induces a preferred chronology with respect to which causality should be defined, then, as in Sec. II B, the matter light cone must coincide with or lie within the gravitational one everywhere in spacetime. On the contrary, it was claimed in [11,13] that, since rods and clocks are made of matter, the matter metric \(\tilde{g}\) should be somewhat favored, in the sense that, in order to preserve causality, no signal
should escape the matter light cone. In particular, this cone should be wider than the gravitational one. Very interesting is the fact that this postulate is the exact opposite of the previous one. It is perhaps the best way to show that there is no clear reason why one metric should be preferred to the other.

This confusion can be easily understood. Let us go further in the discussion of Sec. II C. Whenever $\tilde{g}$ is Lorentzian and disformally related to $g$, we have at hand two metrics which reduce locally to constant metrics $\tilde{g}_0$ and $g_0$. As a consequence, there exist two classes of inertial coordinates for which one of these two metrics, but not both, reduces to its fundamental form $\eta$. Inertial coordinates of each class transform under the action of the Lorentz group $SO(3,1)$ with a different invariant speed.\(^\text{10}\)

Let us emphasize that, in general relativity, superluminality is defined as the propagation on spacelike curves with respect to the gravitational metric, mainly because the latter is thought to be “the” spacetime metric, since it reduces to $\eta$ locally. We already stressed in Sec. II C that any Lorentzian metric can actually take this form locally, so that it could also be viewed as the spacetime metric. The framework of bimetric theories considerably enlightens this point. In a sense, indeed, we have two “natural” metrics and there is simply no way to decide which of these two metrics should be “the” spacetime metric. Accordingly, there are two natural chronologies and two locally invariant speeds. These two classes of inertial coordinates correspond to rods and clocks made up of matter or gravitons.

Bimetric theories thus clearly show the irrelevance of postulating a preferred chronology. Let us also stress that, without such a postulate, the notion of superluminal behavior itself becomes irrelevant. The only point, as far as bimetric theories are concerned, is actually that the ratio of the speed of gravitational waves to the speed of photons, for instance, varies in space and time. This is the reason why such a framework should be considered as the best motivated one that reproduces varying speed of light theories; see Sec. VI E.

C. The Cauchy problem and global properties

Following the discussion of Sec. II C, let us define causality by (i) and (ii). It is clear that the matter metric has to be Lorentzian for the Cauchy problem to be well posed. This condition reads $1 + BU_{\mu}U^{\mu} > 0$ and $A$ must be nonzero. We wrote the conformal factor as $A^2$ in order to ensure that matter fields carry positive energy. The matter stress-energy tensor only depends on the matter fields, the vector field, the gravitational metric, and their first derivatives. Depending on the precise form of the action of the vector and matter fields, the whole set of equations of motion (including gravity) may be a diagonal, second order, and quasilinear hyperbolic system. Note that, when $U$ is given by the gradient of a scalar field $U_{\mu} = \nabla_{\mu} \varphi$, the initial value formulation of the scalar equation is a complicated question; see Sec. VI C. On the other hand, when $U$ is a “true” vector field, the Cauchy problem is, in general, well posed.\(^\text{11}\)

We also have to consider the global structure of spacetime. As in the case of general relativity alone, we have to assume that spacetime is globally hyperbolic in the extended sense of Sec. II C. This assumption together with the above conditions on equations of motion of the matter fields and the vector field ensure that the whole theory is causal.

V. Quantum induced superluminal propagation

In quantum electrodynamics, it has been shown that taking into account finite temperature effects, or nontrivial electromagnetic or gravitational backgrounds, and also boundaries (e.g. the Casimir plates), may break local Lorentz invariance at the loop level through vacuum polarization, thus leading, in some cases, to faster-than-$c$ light propagation [3,14,15]; see also [16] and references therein.

Note that these results are derived within some range of approximation using the effective action formalism, up to the one or two loop level. Generically, the results only hold at frequencies less than the electron mass $m_e$. One therefore only derives the phase velocity of soft photons, whereas the relevant speed is the wave-front velocity, which is the phase velocity at infinite frequency and which corresponds to the analysis of the characteristics [15].

The computation of wave-front velocities is an arduous nonperturbative task that we will not be concerned with. An argument based on the standard Kramers-Kronig relation [16], however, gives some hint as to the value of the wave-front velocity. In the case of Casimir vacua, it has been shown that the wave-front velocity might be greater than $c$ in the direction orthogonal to the plates (breaking of Lorentz invariance by the boundaries). The wave-front velocity has to be equal to $c$ in the parallel direction because Lorentz invariance is left unbroken in that direction, at least if the Casimir plates are infinite (or if boundary effects in that direction are negligible).

Let us assume this result to be valid. Of course, “light does not travel faster than light,” but the point is that gravity does not see the plates in first approximation, so

\(^{10}\)In a sense, our analysis completes the interesting discussion about the different facets of $c$; see [33].

\(^{11}\)If the kinetic term of the vector field differs from the usual (Einstein-Maxwell) one, the vector field equation involves second derivatives of the metric field $g$ and is thereby not diagonal. It does not mean, however, that the Cauchy problem is not well posed, but rather that a careful analysis is in order. Generic vector field actions have been considered in [41] (and references therein).
that Lorentz invariance is not broken in the gravitational sector and gravitons still propagate at velocity $c$. The ratio of the speed of photons to the one of gravitons may thus be greater than 1.

Our discussion in Sec. II allows us to give an immediate answer to the question of causality in that case. It can be shown that photons propagate inside the plates along an effective metric of the form of Eq. (6), where $B$ is some positive constant in that case, and $n^\mu$ is the unit spacelike vector orthogonal to the plates (see [3] and references therein). Outside the plates, photons propagate along the flat metric $\eta$ (here we neglect curvature).

It was correctly recognized in [3] that such a metric is stably causal, so that photons cannot propagate along closed curves, contrary to the claim made in [4] (see also the criticism of [5]). Actually, this spacetime is even globally hyperbolic and therefore perfectly causal, even if photons propagate faster than $c$ inside the plates.

Causality may, however, be lost if two Casimir vacua are moving rapidly towards each other [3]. In that case, spacetime may possess closed future-directed and timelike curves (with respect to the effective metric). This case is rather analogous to the case of two bubbles made of nontrivial background in $k$-essence theory; see Sec. III D. The authors of [3] then invoked the chronology protection conjecture, and, in particular, noted that the two Casimir vacua should be confined within plates that cannot be infinite, so that nontrivial boundary effects may prevent the formation of such closed curves. We essentially reached the same conclusion in Sec. III D. Similar arguments may be applied to other vacua. In particular, we note that vacuum polarization induces some effective metric along which photons propagate. This effective metric only differs from the flat one by terms of order $\alpha^2$, where $\alpha$ is the fine structure constant. These corrections are therefore small (at least in reasonable vacua), so that the effective metric is still Lorentzian and the spacetime is still globally hyperbolic.

VI. APPLICATIONS

We briefly comment on some applications of our results to recent interesting developments on possible modification of gravity.

A. $k$-essence theories and dark energy

$k$-essence scalar field theories have been suggested as promising candidates of the dark energy fluid [22] (see also the review [42] and relevant references therein). Such an effective action can also be motivated by the low energy regime of some string theories [23]. The hyperbolicity conditions, Eq. (5), on the function $F$ have been correctly derived in the literature, but in a different way (except in [43]).

Authors usually require the stability of scalar perturbations around some backgrounds. This is the case if the “sound speed”

$$c_s^2 = \frac{F'(X)}{F(X) + 2XF''(X)}$$

is real. This condition is thus equivalent to the Lorentzian character of the effective metric $G^{\mu\nu}$. This (in)stability is therefore directly related to the hyperbolic (resp. elliptical) character of the scalar field equation. Note that the signature can still be $+2$ or $-2$. Authors then demand that these small perturbations carry positive energy, which implies $F'(X) \geq 0$. Here we wish to stress that these results also arise from the analysis of causality, and are actually non-perturbative results. It can be shown, indeed, that the positivity of the whole Hamiltonian, and not only the one of perturbations, is guaranteed whenever the two conditions, Eq. (5), hold [44].

Unusual kinetic terms have also been considerably debated because phantom scalar fields can reproduce an equation of state $w < -1$ [22,42,45], which indeed reads

$$w = \frac{-F(X)}{F(X) - 2XF''(X)}.$$  (9)

Whenever the density $\rho = F(X) - 2XF'(X)$ is positive, $w < -1$ is equivalent to $F'(X) < 0$. Thus only phantom (ghost) matter can lead to superacceleration of the Universe. Since ghosts are usually associated with a fatal instability at the quantum level and notably in the UV regime, authors have suggested a stabilization mechanism of the ghost field at the UV scale. The function $F$ could be such that $F \sim -X/m^2 + O(X^2/m^4)$ (we reestablish the mass scale $m$). In that case the ghosts only appear at low energy $X \ll m^2$, but can be stabilized at higher energies by higher-order terms, and the time scale of the instability can be made arbitrarily high [46,47]. Such a mechanism may, however, suffer from serious diseases. Indeed note that the sign of $F'(X)$ must change for some values of $X_c$, and the sound speed squared may become negative. Accordingly, the effective metric may become Euclidean and the Cauchy problem will not be well posed anymore. The hyperbolicity may still be guaranteed if $F'(X) + 2XF''(X)$ also changes sign at the value $X_c$. Note that the function $F$ must then be fine-tuned: $X_c$ may be equal to zero, or $X_c$ must be a turning point as well as an inflexion point of $F$. Moreover, in that already fine-tuned case, the effective metric becomes totally degenerate (vanishes) at the point $X_c$. There is a caustic, and the theory is not well defined.

It has been shown in [48] that this point $X_c$ is not reached through cosmic evolution (or at a time $t = \infty$), and the theory may thus be free of singular behavior. However, the cosmological background is not the only relevant one. The theory must also apply at local (e.g. astrophysical) scales, and we expect the scalar field to have some inhomogeneities. Local physics then drives $X$ to positive values, and the above singular point $X_c$ will generically be crossed. Moreover, at a quantum level, it is not clear if we can still
make sense of summing momentum from zero to some cutoff, whereas the propagator is not defined for some value of the momentum.

In conclusion, $k$-essence field theories are quite relevant causal theories that can account for the dynamics of the dark energy fluid with an equation of state $w > -1$. On the other hand, $k$-essence phantom theories, even if stabilized, suffer from serious diseases and it is very unlikely that such a fluid could drive the superacceleration of the Universe $w < -1$. Note that it was recently found that inhomogeneities of the matter distribution in the Universe may be described by an effective scalar field which could be a phantom in certain cases [49]. Of course, such an effective field does not suffer from quantum instabilities.

B. $k$-essence theories and the MOND paradigm

$k$-essence theories are also used to account for the mass discrepancy in galaxies and clusters, without the need for dark matter. Milgrom [27] first pointed out that rotation curves of spiral galaxies exhibit a discrepancy at a universal acceleration scale $a_0 \sim 1.2 \times 10^{-10} \text{m.s}^{-2}$, and therefore that a MOND of the form

$$a \mu \left( \frac{a}{a_0} \right) = g,$$

where $g$ is the Newtonian gravitational field and $a$ the acceleration, could account for the observed discrepancy without dark matter. The $\mu$ function must behave asymptotically as $\mu(x) = 1$ if $x \gg 1$ and $\mu(x) = x$ if $x \ll 1$. The $\mu$ function is otherwise free. The choice $\mu(x) = x/\sqrt{1 + x^2}$ is standard and fits well with the data. Such a behavior in the low acceleration regime automatically leads to flat rotation curves far from the source, and also reproduces the well-established Tully-Fisher law $v^4 \propto L$, where $v$ is the plateau velocity and $L$ is the luminosity of the galaxy.

This successful phenomenology [50] can be reproduced with the help of relativistic aquadratic Lagrangians (RAQUAL) [21], i.e., a $k$-essence scalar field. Consider for instance a scalar field $\varphi$ coupled to matter via a conformal metric $\tilde{g} = \exp(-\alpha \varphi)g$. The scalar field equation then reads

$$\nabla_{\mu}(F(X)\nabla^{\mu} \varphi) = -4\pi G\alpha T,$$

where $\nabla$ is the covariant derivative corresponding to the metric $g$, $G$ is Newton’s constant appearing in the Einstein-Hilbert action, and $T$ is the trace of the stress-energy tensor of the matter fields (defined by variation and contraction with respect to $g$). If we consider a static distribution of matter $\rho$, this equation reduces to a static modified Poisson equation for the scalar gravitational potential $\varphi$:

$$\nabla(F((\nabla \varphi)^2)\nabla \varphi) = 4\pi G \alpha \rho.$$

It is then clear that, after restoring the appropriate constants, the function $F'(x^2)$ plays the role of Milgrom’s function $\mu(x)$. Note that $X > 0$ in the static case. If we require that $F(X)$ behaves as $X$ if $X \gg 1$, and as $2/3X^{3/2}$ if $X \ll 1$, we thus recover the MOND phenomenology.

Since the $\mu$ function is generally taken as a monotonic (increasing) function of its argument, the second derivative of $F$ is positive and the scalar waves propagate superluminally. The theory was thus thought to be acausal [13,21]. However, we argued in Secs. II and III that this conclusion is correct only if the gravitational metric defines a favored chronology, an assumption that may be dropped.

A critical point is, however, that the free function $F$ (related to $\mu$) must satisfy the conditions of Eq. (5). It is, however, immediate to note that the asymptotic conditions on $\mu$ imply that when $X$ goes to zero, $F'(X)$ and $F'(X) + 2XF''(X)$ also go to zero and the effective metric is completely degenerate.

This means that, in such a theory, in an astrophysical context, there must exist around each galaxy or cluster a singular surface along which the scalar degree of freedom does not propagate. The reason is that, near the source, $X$ must be positive, but negative far from it, due to the cosmological background. This theory therefore cannot lead to a consistent picture of local physics imbedded into a cosmological background. This major objection also applies to the recent relativistic model of MOND in [13].

A trivial modification of the asymptotic form of the function $\mu$, however, cures the problem. Let us consider that $\mu(x) \sim x^\varepsilon$ if $x \ll 1$, or equivalently that $F'(X) \sim \sqrt{X + \varepsilon}$ if $X \ll 1$. This ensures that the theory is well behaved at the transition between local and cosmological physics. This slight modification induces a significant change in the phenomenology because there is a return to Newtonian behavior very far from the source, with a renormalized value of the gravitational constant. Such a theory thus predicts that rotation curves are only approximatively flat on a finite range of $r$, and current data require $\varepsilon$ to be at most of order 1/100. More details on this can be found in [51]. Interestingly, in that kind of theory, MOND only appears as an intermediate regime between two Newtonian ones only differing by the value of the gravitational constant, with the transition, driven by the scalar field, occurring at Milgrom’s acceleration scale.

C. Bimetric theories and MOND

MOND-like theories of gravity must also predict enhanced light deflection in order to be consistent with the data. In the previous theory, however, the conformal coupling of the scalar field to the matter metric implies that light is not coupled to the scalar field because of the conformal invariance of electromagnetism in four dimensions.

This has led various authors to consider more general bimetric theories of gravity in which matter is coupled to a...
disformal metric of the type Eq. (7) [11,12]. In these models, the vector field $U$ in Eq. (7) was assumed to be the gradient of the $k$-essence scalar field, and $A$ and $B$ some functions of $\varphi$ and $X$. It was, however, proven that when the matter light cone is wider than the gravitational one (see Sec. IV) there is actually less light deflection than in general relativity.

However, we have seen that causality does not require the matter light cone to be wider than the gravitational one. On the contrary, it could be inside the gravitational one and the light deflection would be enhanced (compared to general relativity). Such a framework is therefore relevant for relativistic theories of the MOND paradigm, as we stress in [44]. The Cauchy problem is very likely to be well posed, since the (Einstein) equation of the metric field is still diagonalized, hyperbolic, and of the second order. The scalar field equation in vacuum has a well-posed Cauchy problem inside the matter is thus a quite involved problem. The scalar field equation inside the matter is, however, complicated and is not diagonalized. The Cauchy problem inside the matter is thus a quite involved question, but we expect that some generic requirements on the free functions $A$ and $B$ and on energy conditions of the matter sector will guarantee its well-posedness. This is left for further investigation (see also [44]).

**D. Bimetric theory of gravity and the Pioneer anomaly**

Let us briefly consider another application of bimetric theories of gravity. Conventional physics has not succeeded, so far, in explaining the anomalous motion of the two Pioneer spacecrafts [52] that experience a small anomalous acceleration (roughly directed towards the sun).

Unconventional theories of gravity could however explain it. The two Pioneer spacecrafts are moving along geodesics of the bare metric $g$, and we shall look for a slight modification of the Schwarzschild solution to explain the Pioneer anomaly. It is worth noting that this anomaly was detected soon after Jupiter’s flyby [52] so we have to check that the suggested modification is compatible with the motion of outer planets. Outer planets are moving along quasicircular orbits and are thereby mostly sensitive to the time-time component of the matter metric $g_{00}$. Very sensitive tests of Kepler’s third law [53] then place stringent bounds on the deviation from the leading order of the time-time component of the Schwarzschild metric. These bounds are actually too small to account for the Pioneer anomaly.

This fact leads to the wrong statement in [54] that the Pioneer anomaly cannot be of gravitational origin. This is indeed not justifiable since the two spacecrafts evolve on hyperbolic trajectories and are thereby also sensitive to the radial-radial component of the matter metric, which is not well constrained in the outer solar system. Let us stress that, if spherical symmetry is assumed, a slight modification of $g_{rr}$ compared to the Schwarzschild solution is the only way to give the Pioneer anomaly a gravitational origin in a metric theory of gravity. This was realized in [55].

Here we provide, with the help of bimetric theories of gravity, a framework that realizes this modification of the radial-radial component. The action of the theory is given by the Einstein-Hilbert action, a canonical action for the scalar field, and the matter fields are coupled to the matter metric

$$\tilde{g}_{\mu \nu} = A^2(\varphi)(g_{\mu \nu} + B(X)\nabla_\mu \varphi \nabla_\nu \varphi),$$

where $X$ is still defined by Eq. (2). As we have already stressed, nontrivial properties of $A$, $B$ and energy conditions of the matter sector may be required to ensure the hyperbolicity of the scalar equation inside matter. In addition, the matter metric has to be Lorentzian and this reads $1 + 2XB(X) > 0$. In a static and spherically symmetric situation, we have $\tilde{g}_{00} = A^2g_{00}$ and $\tilde{g}_{rr} = A(g_{rr} + B(X)\nabla_r \varphi \nabla_r \varphi)$. The bare metric $g$ is given by the solution of Einstein’s equations and it can be shown that it coincides with the Schwarzschild solution, to leading order. This theory is thus a realization of the above phenomenology. Such models predict “Pioneer-like” anomalies in the precession of perihelion that could, for instance, be found in precise measures of the orbit of Mars. More details on the Pioneer anomaly and disformal theories can be found in [44].

**E. Varying speed of light theories**

The first varying speed of light theories were constructed by replacing $c$ by $c(t)$ in the equations of motion of general relativity, where $t$ could be the cosmic time [32]. This, however, leads to equations of motion that do not conserve stress energy anymore. In other words, this theory cannot be obtained from a variational principle [31,33].

Some authors thus made use of the disformal matter metric, Eq. (13), to reproduce VSL within a consistent framework [30,31,56]. Note that, in all of these models, the free functions were taken as constants $A = 1$ and $B = -1/m^2$. Let us stress that if one insists on the weak equivalence principle there cannot be any coupling between $\varphi$ and the standard model matter other than through the matter metric. Then, by varying the action with respect to $\varphi$, one finds that the scalar field is not created by the matter sources ($\nabla_\mu \varphi = 0$ is always a solution). The scalar field can only be generated by some (nonconstant) function $A(\varphi)$, like in scalar-tensor theories.

Note also that the choice of $B = -1/m^2$ in these works does not guarantee the Lorentzian character of the matter metric, which can actually be Euclidean in a cosmological background such that $(\partial_0 \varphi)^2 > m^2$. One could, however, circumvent this problem by arguing that the theory is only an effective model valid for $(\partial \varphi)^2 \ll m^2$, with $m$ of the order of the grand unification scale.

As we have already stressed, the initial value formulation of the scalar field equation in such models is an involved question. This has not been pointed out previ-
ously, although the existence of an initial value formulation is crucial in order to give any sense to such VSL models.

Let us finally remark that, since the fine structure constant \( \alpha \) is proportional to the inverse of the speed of light, it has been argued that VSL theories could also account for the variation of \( \alpha \) with cosmic time, if such a variation exists [57,58]. However, in this bimetric framework of VSL theories, if one analyzes atomic rays at some redshift by usual techniques, one is observing electromagnetic phenomena using matter rods and clocks, so that no variations of \( \alpha \) are actually observable. The fact that the ratio of the speed of light to the speed of gravitational waves varies in space and time does not lead to a variation of \( \alpha \) (if it is measured in the usual way), but simply to a redefinition of the redshift \( z \) of distant objects.

**VII. CONCLUSIONS**

We have provided a careful analysis of the meaning of causality in classical field theories. This has led us to the conclusion that superluminal behavior is found to be non-causal only if one refers to a prior chronology in spacetime. This postulate actually states that, locally, some sets of inertial coordinates must be the preferred ones, and this appears to be in great conflict with the spirit of general relativity and, more precisely, with general covariance. On the contrary, we derived in Sec. II C, by means of the conditions (i) and (ii), a formulation of causality in which coordinates are still physically meaningless.

Note that, while referring to a preferred chronology may seem natural in general relativity since all fields (both gravity and matter) propagate along the gravitational metric \( g \), it becomes somewhat unnatural whenever spontaneous breaking of Lorentz invariance occurs, be it driven by quantum polarization (and, in general, the solution of field equations), nonlinearities of some new fields, etc. The resulting spacetime is generically endowed with a finite set of Lorentzian metrics \( h \), which may not be conformally related to each other. In that case, rods and clocks made up of different fields lead to different systems of coordinates that do not transform under the same Lorentz group. There is not only one, but at least two invariant speeds in that case. Note that it would be misleading to think that there is a preferred Lorentz invariance in the theory because of the symmetry of the action. Actually, in a general relativity-like context, the action is not Lorentz invariant but diffeomorphism invariant. This notably means that all local systems of coordinates are equivalent. On the contrary, the existence of a preferred chronology means that (as far as causality is concerned) there exists some preferred class of inertial coordinates, or equivalently some preferred rods and clocks. This seems to be physically unacceptable.

As a consequence, one cannot refer to causality in order to assert that nothing can travel faster than one of these speeds. Accordingly, causality does not require one of these metrics to define a cone wider than the others everywhere in spacetime. Bimetric theories of gravity greatly enlighten this point, since both the gravitational and the matter metric could be used to define a natural chronology in spacetime. Depending on the choice made, one then finds that nothing can travel faster than gravity or light. These two opposite requirements can be found in the literature.

On the contrary, our definition of causality in which no prior chronology is assumed (by means of our mixed chronology; see Sec. II C), enables a causal theory to include "superluminal" propagation. Actually the very notion of superluminal behavior is no more meaningful in that framework, and the only point is that some degrees of freedom can propagate faster than other ones. Moreover, the causal cones may even tip over each other depending on the location in spacetime. As an application, in \( k \)-essence theories, causality does not require the sign of \( F''(X) \) to be fixed (and notably negative), and in bimetric theory causality does not require the gravitational light cone to be wider than the matter one (or conversely). Causality first requires that the equations of motion have a well-posed Cauchy problem. This strongly depends on the precise form of the dynamics, and throughout this paper we used classical results [20] on that subject. We also discussed in detail the fact that global properties of spacetimes may break causality. We pointed out that this already occurs in standard general relativity so that it cannot be related to some intrinsic disease of superluminal propagation. We discuss three very similar cases: Gott’s cosmic strings [40], the two bubbles of nontrivial vacua in \( k \)-essence theories [7], and the two Casimir experiments of [3]. What became clear was that such nontrivial global properties may be suppressed by subtle boundary effects. This would be an illustration of the chronology protection conjecture.

**ACKNOWLEDGMENTS**

The author is indebted to G. Esposito-Farèse for many useful discussions and comments, and thanks C. Deffayet, R. Durrer, E. Flanagan, J. Larena, J. Moffat, and J.-P. Uzan for useful discussions, and the anonymous referees for helpful remarks.
[1] R. C. Tolman, *The Theory of Relativity of Motion* (University of California Press, Berkeley, 1917).
[2] G. A. Benford, D. L. Book, and W. A. Newcomb, Phys. Rev. D 2, 263 (1970).
[3] S. Liberati, S. Sonnino, and M. Visser, Ann. Phys. (N.Y.) 298, 167 (2002).
[4] A. D. Dolgov and I. D. Novikov, Phys. Lett. B 442, 82 (1998).
[5] M. Yu Konstantinov, Russian Physics Journal 45, 23 (2002); gr-qc/9810019.
[6] Y. Aharonov, A. Komar, and L. Susskind, Phys. Rev. 182, 1400 (1969).
[7] A. Adams, N. Arkani-Hamed, S. Dubovsky, A. Nicolis, and R. Rattazzi, J. High Energy Phys. 10 (2006) 014.
[8] E. A. Lim, Phys. Rev. D 71, 063504 (2005).
[9] C. Bonvin, C. Caprini, and R. Durrer, Phys. Rev. Lett. 97, 081303 (2006).
[10] J. W. Elliott, G. D. Moore, and H. Stoica, J. High Energy Phys. 08 (2005) 066.
[11] J. D. Bekenstein, Phys. Rev. D 48, 3641 (1993).
[12] J. D. Bekenstein, Astrophys. J. 429, 480 (1994).
[13] J. D. Bekenstein, Phys. Rev. D 70, 083509 (2004).
[14] I. T. Drummond and S. J. Hathrell, Phys. Rev. D 22, 343 (1980); K. Scharmhorst, Phys. Lett. B 236, 354 (1990); J. I. Lattore, P. Pascual, and R. Tarrach, Nucl. Phys. B437, 60 (1995); W. Dittrich and H. Gies, Phys. Rev. D 58, 025004 (1998).
[15] G. M. Shore, gr-qc/0302116.
[16] K. Scharmhorst, Ann. Phys. (Berlin) 7, 700 (1998).
[17] G. M. Shore, hep-th/0701185.
[18] J. W. Moffat, Found. Phys. 23, 411 (1993).
[19] S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Spacetime* (Cambridge University Press, Cambridge, England, 1973).
[20] R. M. Wald, *General Relativity* (The University of Chicago Press, Chicago, 1984).
[21] J. D. Bekenstein and M. Milgrom, Astrophys. J. 286, 7 (1984).
[22] T. Chiba, T. Okabe, and M. Yamaguchi, Phys. Rev. D 62, 023511 (2000); C. Armendáriz-Picón, V. Mukhanov, and P. J. Steinhardt, Phys. Rev. D 63, 103510 (2001).
[23] C. Armendáriz-Picón, T. Damour, and V. Mukhanov, Phys. Lett. B 458, 209 (1999).
[24] A. D. Rendall, Classical Quantum Gravity 23, 1557 (2006).
[25] A. Vikman, Phys. Rev. D 71, 023515 (2005).
[26] E. Babichev, V. F. Mukhanov, and A. Vikman, J. High Energy Phys. 09 (2006) 061.
[27] M. Milgrom, Astrophys. J. 270, 365 (1983).
[28] R. H. Sanders, Astrophys. J. 480, 492 (1997).
[29] I. T. Drummond, gr-qc/9908058; Phys. Rev. D 63, 043503 (2001).
[30] M. A. Clayton and J. W. Moffat, Phys. Lett. B 460, 263 (1999); 477, 269 (2000).
[31] B. A. Bassett, S. Liberati, C. Molina-París, and M. Visser, Phys. Rev. D 62, 103518 (2000).
[32] J. D. Barrow and J. Magueijo, Phys. Lett. B 443, 104 (1998); J. D. Barrow, Phys. Rev. D 59, 043515 (1999); A. Albrecht and J. Magueijo, Phys. Rev. D 59, 043516 (1999); J. D. Barrow and J. Magueijo, Classical Quantum Gravity 16, 1435 (1999).
[33] G. F. R. Ellis and J.-P. Uzan, Am. J. Phys. 73, 240 (2005).
[34] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973).
[35] M. A. Clayton and J. W. Moffat, Int. J. Mod. Phys. D 11, 187 (2002).
[36] S. W. Hawking, Phys. Rev. D 46, 603 (1992).
[37] E. Bois and E. Trelut, astro-ph/0304520.
[38] J. Z. Simon, Phys. Rev. D 41, 3720 (1990); R. P. Woodard, astro-ph/0601672.
[39] R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Interscience, New York, 1962), Vol. II.
[40] J. R. Gott, Phys. Rev. Lett. 66, 1126 (1991).
[41] T. Jacobson and D. Mattingly, Phys. Rev. D 64, 024028 (2001).
[42] E. J. Copeland, M. Sami, and S. Tsujikawa, Int. J. Mod. Phys. D 15, 1753 (2006).
[43] C. Armendáriz-Picón and E. A. Lim, J. Cosmol. Astropart. Phys. 08 (2005) 007.
[44] J.-P. Bruneton and G. Esposito-Farese (unpublished).
[45] R. R. Caldwell, Phys. Lett. B 545, 23 (2002).
[46] S. M. Carroll, M. Hoffman, and M. Trodden, Phys. Rev. D 68, 023509 (2003).
[47] J. M. Cline, S. Jeon, and G. D. Moore, Phys. Rev. D 70, 043543 (2004).
[48] N. Arkani-Hamed, H-C. Cheng, M. A. Luty, and S. Mukohyama, J. High Energy Phys. 05 (2004) 074.
[49] T. Buchert, J. Larena, and J-M. Alimi, Classical Quantum Gravity 23, 6379 (2006).
[50] R. H. Sanders and S.S. McGaugh, Annu. Rev. Astron. Astrophys. 40, 263 (2002).
[51] B. Famaey, J.-P. Bruneton, and H. Zhao, astro-ph/0702275.
[52] J. D. Anderson, P.A. Laing, E.L. Lau, A.S. Liu, M.M. Nieto, and S.G. Turyshhev, Phys. Rev. D 65, 082004 (2002); V.T. Toth and S.G. Turyshhev, gr-qc/0603016.
[53] C. Talmadge, J.-P. Berthias, R.W. Hellings, and E.M. Standish, Phys. Rev. Lett. 61, 1159 (1988).
[54] K. Tangen, gr-qc/0602089.
[55] M.-T. Jaekel and S. Reynaud, Mod. Phys. Lett. A 20, 1047 (2005); Classical Quantum Gravity 23, 777 (2006).
[56] N. Kaloper, Phys. Lett. B 583, 1 (2004).
[57] J. K. Webb, V. V. Flambaum, C.W. Churchill, M.J. Drinkwater, and J.D. Barrow, Phys. Rev. Lett. 82, 884 (1999).
[58] A. V. Ivanchik, E. Rodriguez, P. Petitjean, and D.A. Varshalovich, Astronomy Letters 28, 423 (2002).