Comment on: "Causality and the Arrow of Classical Time", by Fritz Rohrlich.

Carlo Rovelli

Dipartimento di Fisica dell'Università "La Sapienza", I-00185 Roma.
Centre de Physique Théorique de Luminy, Université de la Méditérranée, F-13288 Marseille.

January 16, 2004

Abstract

Rohrlich claims that "the problem of the arrow of time in classical dynamics has been solved". The solution he proposes is based on the equations governing the motion of extended particles. Rohrlich claims that these equations, which must take self-interaction into account, are are not invariant under time reversal. I dispute this claim, on several grounds.

In [1], Rohrlich proposes a solution to the "the problem of the arrow of time in classical dynamics". This is the problem generated by the conflict between the invariance properties under time reversal of the basic equations of classical dynamics, and the irreversibility of most natural phenomena. The solution Rohrlich proposes is based on the claim that fundamental classical equations are not trully time-reversal-invariant. First, he observes that classical dynamics is only relevant as a theory of extended particles, since point particles are trully quantum objects. Second, the motion of these extended particles is governed by equations that take into account the self-interaction of the particle: electromagnetic if the particle is charged, and gravitational in all cases. Rohrlich studies these equations, and observes that they are not time-reversal-invariant. He concludes that the fundamental equations of classical mechanics are not time-reversal-invariant. More precisely, he observes that these equations are nonlocal in time; they require initial data at earlier times. The time-reversed equations require initial data in the future, and therefore should be discarded on the basis of causality, namely the requirement that the effect cannot precede the cause. He concludes that lack of time reversal invariance at the fundamental level is a consequence of causality. To fully understand Rohrlich's point of view it is essential to notice that he begins by defining dynamics as "the study of the motion of a body under the influence of known forces".

On the basis of this argument Rohrlich claims in [1] that "the problem of the arrow of time in classical dynamics has been solved". I do not think that this claim is justified. I begin by summarizing Rohrlich's argument. Then I illustrate my objection with a simple example. Then I build up my counter-argument. At the end, I point out the difficulty in Rohrlich's argument.

Rohrlich's argument. The core of Rohrlich's observation is the following. Consider an extended charged particle of radius d at rest. Let's apply a force F_{ext} on the particle for a short finite time. What is the effect on the particle? Under the influence of the force F_{ext} the particle accelerates. The acceleration of the charged particle, generates electromagnetic radiation. The radiation, in turn, causes a force on the extended particle itself. This is called the self-force F_{self}. The self force is a time-delayed phenomenon: the force F_{self} that acts on the particle at time t depends on the acceleration of the particle at an *earlier* time. The reason is simple: the force F_{self} acting on a charge element de of the extended particle and caused by the acceleration of a different charge element de' is the effect of radiation emitted by de' and acting on de. Radiation takes a finite time to travel across the extended particle from de' to de. Hence the self-force on de is determined by the acceleration of de' at an earlier time. Since the time radiation takes to travel across the particle is $\tau = dc$, where c is the speed of light, it is clear that F_{self} acting on the extended particle at time t depends on the acceleration of the particle during the time interval $(t-\tau, t)$ that *precedes* the time t. Therefore the total force acting on the particle at time t will have the structure

$$F(t) = F_{ext}(t) + F_{self}^{earlier} \tag{1}$$

where the "delay" term $F_{self}^{earlier}$ depends on the motion of the particle at a time *earlier* than t. In fact, a Lorentz invariance equation with this structure has been recently derived. See [1] for details and full references. In the nonrelativistic limit in which the size d of the particle goes to zero, the last term converges to the well known self force term

$$F_{self}^{earlier} \quad \underset{d\to 0}{\longrightarrow} \quad \frac{2}{3}\frac{e^2}{c}\ddot{v}, \tag{2}$$

where e is the charge of the particle and $\ddot{v}$ the second derivative of its velocity. Now, Rohrlich claims that if we time reverse equation (1), replacing t with $-t$, we obtain the equation

$$F(t) = F_{ext}(t) + F_{self}^{later} \tag{3}$$

where F_{self}^{later} depends on the motion of the particle at a time *later* than t. That is, since the force on de at time t depends on the acceleration of de' at the earlier time $t-\tau$, then, argues Rohrlich, in the time reversed process the force on de at time t will depends on the acceleration of de' at the later time $t+\tau$. Rohrlich argues then that to solve this equation we need initial data in the future, and we can discard this possibility on the basis of causality: an effect (force on de)

cannot precede the cause (acceleration of de'). Therefore, argues Rohrlich, we can discard equation (3) on the basis of causality. Therefore at the fundamental level we have equation (1) but not its time reversed. Therefore the fundamental dynamical equations are not time reversible. This is the reason for which there is no contradiction between the observed irreversibility and classical dynamics. This is Rohrlich's argument.

Astronauts' tennis. Imagine two astronauts playing tennis in empty space. Let $X(t)$ an $Y(t)$ by the motions of the two astronauts, and $z(t)$ be the motion of the ball. Say the motions are one dimensional and the tennis ball simply bounces back and forth elastically between the two (a rather dull game, indeed). The dynamics of the system is simply the one of elastic collisions of three masses in one dimension. It is governed by simple time-reversible equations.

However, we can give a different description of this process. We can decide to ignore the tennis-ball variable $z(t)$. In fact, notice that the force on one player at time t is uniquely determined by the acceleration of the other player at a time $t-\tau$, where τ is the time of travel of the ball between the two relevant bounces. Let us call $F_{self}^{earlier}$ the force on one player due to the bounce of the tennis ball, and F_{ext} any other eventual force. The force on one player is therefore

$$F(t) = F_{ext}(t) + F_{self}^{earlier} \tag{4}$$

where the "delay" term $F_{self}^{earlier}$ is a function of the trajectory of the two players and the acceleration of the other player at an *earlier* time. Rohrlich's argument then applies immediately: the time reversal of equation (4) is the equation

$$F(t) = F_{ext}(t) + F_{self}^{later} \tag{5}$$

where the force on one player is due to the acceleration of the other player at a *later* time. This last equation must be discarded on the basis of causality, because to compute the force at time t we must know data in the future. Hence the process is governed by laws that are not time reversible.

This conclusion is manifestly false. Therefore there should be a mistake somewhere. The mistake is the following. We can always write an equation that connects a force at time t_2 with some events happened at an earlier time t_1. We can also argue that the event at time t_1 was the "cause" of the force acting at time t_2, if we like to think in term of "causes". But in the time reversed process, we cannot keep the same causal connections. If we want to think in terms of causes, causal connections must be reversed. If in the "forward" tennis game, say a bounce A happens first and a bounce B happens later. Then we can say that the bounce A is the "cause" of the later force at the bounce B. But in the time reversed process it is the bounce B that happens first. Therefore we cannot say anymore that A causes the force at B. This does not contradict the fact that there exists an equation connecting the force at B with the (later in the time reversed process) bounce at A.

To the very contrary, it is precisely because the process is time reversible that we can equally well write an equation that relates forces with events happened earlier, or later. In fact, nothing prevents us, in general, to write the force acting at a bounce as a function of the acceleration at the other player at the following bounce. That is, both equations (4) and (5) are correct in both the forward and the time reversed processes.

Back to extended particles. How does the example above relate to the extended particle case considered by Rohrlich? The situation is precisely the same in the two cases. In the case of the extended particle, the elements of the particle are like the two astronauts, and the field that propagates between them is like the tennis ball. The apparent lack of time reversibility, in fact, derives from the combination of various inputs.

First, it derives from the choice of describing the physics in a simplified form, dropping the field variables, as in the example the tennis ball variable was dropped. In fact, the full set of relevant variables are given by the center of motion $X(t)$ of the extended particle, as well as the electromagnetic field $F_{\mu\nu}(x,t)$. The set of variables $[X(t), F_{\mu\nu}(x,t)]$ satisfies a system of time reversible equations that contain no terms nonlocal in time. This system is given by the full Maxwell equations with a source term given by the charge *density* current formed by the charge distribution over the extended particle, the Lorentz force acting on each infinitesimal element of charge, plus some assumption on the structure of the particle itself. (If we assume, as Rohrlich does, that the extended particle is rigid, we are within some approximation that breaks relativistic invariance, since rigid bodies are not compatible with special relativity.) This set of equations forms the true basic dynamical description of the phenomenon. In fact, it is from this set of equations that the the delay equation can be derived. The fundamental set of equations describing the phenomenon is perfectly time reversible, and its evolution is fully determined by the data at a single instant of time. Now, we can separate the field $F_{\mu\nu}(x,t)$ in two components

$$F_{\mu\nu}(x,t) = F_{\mu\nu}^{ext}(x,t) + F_{\mu\nu}^{self}(x,t) \tag{6}$$

The first component is the external field, the second is the self-field generated by the acceleration of the particle itself. However, it is crucial to observe that this distinction is not time reversible. Going back to the tennis players, how can we distinguish the motion of the tennis ball "caused by the players" from the motion "independent from the players"? To do so, we could for instance assume that the tennis ball was at rest to start with, and therefore state that the subsequent motion is all "caused by the players". But this depends on a choice of a time direction, because if we insist that in the reversed process the motion is still caused by the players, we obtain the result that the earlier motion of the tennis ball is caused by a later bounce. This is precisely the difficulty in Rohrlich's argument: he assumes that the electromagnetic field can be consistently separated in the external part and the self part generated by the particle, and that the two components could be independently time reversed without affecting causality. This cannot be.

Finally, we can drop the explicit mention of the field variables $F_{\mu\nu}^{self}(x,t)$, and replace their effect on the motion of the extended particle by a self-force "delay" term that is non local in time. This is precisely what I did for the tennis ball. In doing so, we relate the force at some time with an acceleration of the particle at an earlier time. The relation is correct and remains true both in the forward and in the time-reversed process, but the interpretation as a causal relation is incorrect in the time inverse process. In fact, the time inverse process is governed by equation (3), but it also governed by equation (1), provided that the term $F_{ext}(t)$ is appropriately reinterpreted.

The dynamics of an extended particle is perfectly time reversible, and governed by time reversible equations.

A ship in still waters. Light is shed on the delay term by considering another physical phenomenon where the same term appears. Imagine that we want to write the equations governing the dynamics of a ship in the ocean. The ocean waves interact with the ship. Therefore we must consider a dynamical system where the variables are the ones describing the ship plus the waves' field. These form a system of equations which is local in time. Consider the particular case in which the ship is set in motion in still waters. If the ship is accelerated, it raises waves, which, in turn, affect the ship itself: a wave raised by the front of the ship generates a force on the ship side and bottom, as it travels by. If the ship is long L and the relevant waves travel at a speed v, then the effect will have a time scale of the order $T = L/v$. This is a sort of "self-force" on the ship, and is completely analogous to the electromagnetic self force. We can give a simplified effective description of this phenomenon without mentioning the variables describing the ocean waves, but simply adding a self-force "delay" term to the equations of motion of the ship. It is clear that this may be a useful effective description, bus has no bearing on the time reversal invariance of the basic equations.

Particles and fields. How is it possible that Rohrlich falls into this confusion? Let me examine his argument closer. The entities involved in the physical situation described by Rohrlich are two: the extended particle and the field. Initial conditions, equations of motion, and all considerations about time-reversal-invariance must include *both* actors. But Rohrlich defines dynamics as "the study of the motion of a body under the influence of known forces", and makes pretty clear that the "body" is just the extended particle. Where is the field gone?

In fact, what Rohrlich has in mind, is pre-relativistic dynamics, where fields do not have independent degrees of freedom, and are entirely determined by particles. In that context, dynamics can be correctly defined as Rohrlich does as "the study of the motion of a body under the influence of known forces". But in that context there is action-at-a-distance, and no time delay caused by field propagation. In the relativistic context, instead, dynamics is understood as something else: it is the the study of the reciprocal interactions of particles and fields. What Rohrlich describes is not a formulation of fundamental mechanics in which the extended size of particles is taken into account. It is just a hybrid approximation obtained trying to extract some corrections to the motion of extended particles alone, from a full fledged particle-field theory.

These corrections are derived under specific assumptions. The key assumption made, which is the one that breaks time-inversion invariance is to assume there is no incoming radiation. If there is no incoming radiation, then we have to restrict to retarded potentials and non-time-reversal-invariant equations follow. But this is just the effect of having chosen very peculiar initial conditions for the field.

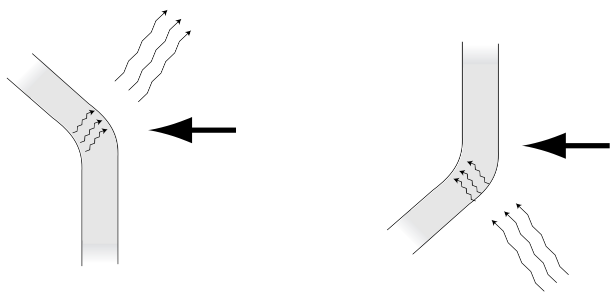

Figure 1: Pictorial illustration of the physics of self-interaction and its time reversal. Time runs upward in these two spacetime diagrams. The world-history of an extended particle is represented by the grey strip. The external force on the particle is represented by the thick arrow. Radiation is represented by weaving arrows. In the left panel, there is no incoming radiation. The particle, initially at rest, is accelerated by the force. The emitted radiation exerts a force back on the particle itself, as well as escaping as outgoing radiation. The right panel represents the time reversed process. The particle is initially in motion and there is incoming radiation. The incoming radiation and the force bring the particle at rest. Both sequences of events are perfectly compatible with causality.

In fact, the key detail that Rohrlich does not mention in his paper is that the field generated by the accelerating particle doesn't have the sole effect of acting back on the particle itself. It also escapes towards the future, as full fledged outgoing radiation. The physics described by the self-interaction equations is therefore a physics in which accelerated particles *emit radiation.* The assumption made in deriving these equations is that we allow outgoing radiation but we assume no incoming radiation. It is this *assumption* that breaks time reversal invariance. Figure 1 illustrates pictorially the physics of the process as well as its time-reversal, which is perfectly causal.

By defining dynamics as "the study of the motion of a body under the influence of known forces", and then using partial field theoretical results, Rohrlich is mixing two different theoretical frameworks. One is the dynamics of particles under the influence of known forces. This is a theory with a finite number of degrees of freedom. The second is a theory of particles and fields, which is a theory with many (in fact, an infinite number of) degrees of freedom.

Now, irreversible processes appear in theories with many microscopical degrees of freedom when we neglect these degrees of freedom and concentrate on a few macroscopic variables. This is exactly what Rohrlich does. He concentrates on the particle motion, and disregards the fields degrees of freedom.

It is well known that phenomenological equations that neglect effects on systems with many degrees of freedom are non-time-reversal-invariant. Friction forces are the prototypical case. Here we are exactly in the same situation: the irreversibility is caused by a sort of "friction" with the field degrees of freedom.

More precisely, Rohrlich puts himself in a context which is non-time-reversal invariant *by definition*. This context is the motion of particles with arbitrary initial conditions, interacting with fields with specific and non-time-reversal-invariant initial conditions (arbitrary outgoing radiation but no incoming radiation). This is like saying that the dynamics the waves raised by a stone falling into still water is governed by non-time-reversal-invariant equations. True, but trivial, because time reversal invariance is broken by the assumption of the initial stillness of the water.

In summary, if we are allowed to choose peculiar initial conditions for the field (no incoming radiation), then we have retarded potential and non-time-reversal-invariant equations. In this case, Rohrlich's reference to the peculiar self-force equations is redundant: we already knew that non time reversal invariant assumptions on the field leads to the breaking of time-reversal invariance. If, on the other hand, we allow arbitrary initial configurations for particle and field, then we must include also advanced potentials, and the time-reversed self-force equations. In this case, there is no violation of time-reversal-invariance.

In conclusion, the reference to the self-force equations does not add anything to what was already well understood, or not understood, about time-reversal invariance.

In addition, one could also observe that self-interaction forces are small relativistic effects that play little or no role in the common irreversible processes that surround us (such as lake's wave moved by a stone). These irreversible processes can be modeled by dynamical systems where no self-interaction appear, and their full fledged macroscopical irreversibility manifests itself nevertheless. I think it is save to say that the self-interaction of extended classical particles does not teach us anything new about the arrow of time.

References

[1] Fritz Rohrlich, *Causality and the Arrow of Classical Time*, Studies in History and Philosophy of Modern Physics, 31b (2000) 1.