Philosophy of Science (Proceedings), vol. 70, December 2003, pp. 1086-1097

Understanding the Time-Asymmetry of Radiation

Abstract

I discuss the nature of the puzzle about the time-asymmetry of radiation and argue that its most common formulation is flawed. As a result, many proposed solutions fail to solve the real problem. I discuss a recent proposal of Mathias Frisch as an example of the tendency to address the wrong problem. I go on to suggest that the asymmetry of radiation, like the asymmetry of thermodynamics, results from the initial state of the universe.

Understanding the Time-Asymmetry of Radiation

1. The Problem

There is a puzzle about radiation. In our experience, waves display a clear time-asymmetry. Waves appear to spread outwards after their sources move; they do not converge on sources which then begin to move. Water waves diverge after a pebble is dropped in a pond; they do not travel inwards to a spot from which a pebble is then ejected. We see electromagnetic waves emerge after charges accelerate, not converge on charges which then begin to accelerate. Yet the equations governing wave phenomena are symmetric in time, allowing for both the kinds of waves we see and the time-reversal of these processes. Then where does the observed asymmetry of radiation come from?

We can state the asymmetry more formally. (I focus on electromagnetic radiation, though the discussion applies to wave phenomena generally.) The (inhomogeneous) equation for wave propagation is $(1/c^2 \cdot \partial^2/\partial t^2 - \nabla^2)\phi(\boldsymbol{r}, t) = 4\pi\rho(\boldsymbol{r}, t)$, where ϕ is the amplitude of the wave at $(\boldsymbol{r}, t)$, ρ is the source density, and c is the wave velocity. (The wave equation can be derived from Maxwell's equations.) One solution represents the field amplitude as determined by the source density at positions at times earlier than t. This is the retarded solution. The advanced solution characterizes the field amplitude in terms of the source density at later times. Advanced solutions describe waves converging from the past onto the charges with which they are associated; retarded solutions describe waves diverging from charges to the future. The asymmetry thus appears to consist in the fact that there is only retarded and no advanced radiation in our universe. The puzzle, as it is usually stated, is why this asymmetry holds when the laws permit both kinds of radiation. Ridderbos writes, "although mathematically the equation for wave propagation has two possible solutions...it is exclusively the [retarded] solution that seems to be observed in nature" (1997, 473). Likewise Frisch: "Even though both retarded and advanced solutions...are allowed by the Maxwell equations, the retarded solution...appears to represent the physical situation correctly" (2000, 384). And Price: "Maxwell's theory clearly permits both kinds of solution, but nature appears to choose only one. In nature it seems that radiation is always retarded rather than advanced" (1996, 50).

There is a difficulty with this characterization of the puzzle, however. We must be careful in spelling out precisely what it means to say that the radiation of our universe is retarded, and doing so is not as straightforward as it might seem.

In order for it to be the case that there is only retarded radiation, the fields we observe at every space-time point must be entirely attributed to retarded sources. As Sciama (1967) explains, for a given point, we must look at a spatial region containing it and consider whether, as the volume of that region tends to infinity, the retarded fields from all the charges within the region combine to yield the field at that point. If this is so for every point in the universe, then it seems we should conclude that there are only retarded fields. More formally: the field amplitude at a point can be expressed as the sum of a volume integral over a region containing the point, plus a surface integral over the surface of that volume: $\phi(\boldsymbol{r}, t) = \int_V \rho \, dV / r + \int_S \phi \, dS$. The volume integral represents the part of the field attributed to sources within *S*. The surface integral represents the radiation from sources outside *S* plus any free fields coming from infinity. We obtain the retarded solution if, for each point in the universe, the retarded surface integral (evaluated at the retarded time) vanishes as the volume *V* tends to infinity.

If we confine all the charges to a finite region, we can ensure that the contribution to the retarded surface integral from sources outside *V* will tend to zero as the volume goes to infinity. Then none of the field will be attributed to advanced sources. But this is insufficient to guarantee that we obtain the retarded solution. For part of the field at a given point might not be associated with sources: some of the field might consist of source-free radiation (solutions to the homogeneous wave equation, with the right-hand side equal to zero). If so, the retarded surface integral will not vanish. And, in fact, there *are* free fields in our universe: consider the fields that came into existence immediately following the big bang and which now make up the background radiation. This radiation is observed to be quite small—the initial fields have greatly cooled due to the universe's expansion—but it does exist.

We might, in other words, wish to hold that the radiation of our universe is retarded if it comprises retarded radiation associated with charges plus source-free fields. But the wave equation is linear. Hence any linear superposition of a retarded and an advanced solution is a solution. Furthermore, any linear combination of an advanced or retarded solution (or a superposition of the two) with a solution to the homogeneous wave equation—a free field—is a solution. Therefore, if we can describe the total field at a point as a retarded-plus-free field, we can equivalently describe it as an *advanced*-plus-free field. We can represent any radiation that appears to be retarded with an advanced solution, so long as we include the right free-field term. So we can equally describe our world as containing only advanced radiation (plus free fields) as one in which there is only retarded radiation (plus different free fields). It may *look* as though all waves diverge from past sources; it may *seem* obvious that we should represent these waves by retarded solutions. But the linearity of the wave equation means that the advanced solutions can also represent this radiation accurately.

Some authors realize that both advanced and retarded solutions can describe actual radiative processes, but for the wrong reason. Price writes that "the advanced solutions...simply characterize absorption, and therefore do exist in nature" (1996, 60). Similarly, Callender (2002) says: "The advanced solutions describe the radiation sink's receiving waves, and this happens all the time." These claims misleadingly suggest that advanced solutions represent absorptions (and retarded solutions emissions). It is true that advanced solutions describe waves that travel inwards to converge on charges—to waves that appear, from our time sense, to be absorbed. But with the right free field, the retarded solution can represent radiation that seems to be absorbed, and the advanced solution radiation that appears to be emitted. *Any* radiation, regardless of whether it appears to us to be emitted or absorbed, can be represented by *either* solution, so long as we include the appropriate free-field term. For a free field can interfere with emissions and absorptions in such a way that a process in which radiation is "really" being emitted appears to us as one in which radiation is being absorbed, and vice versa. It is the linearity of the wave equation, not the nature of absorption, that allows for advanced solutions to correspond to actual phenomena.

Thus, it is not quite right to state the explanandum as the fact that our universe contains only retarded radiation even though the laws have advanced solutions. Since the fields we observe underdetermine their components, it is possible that we *do* see advanced radiation (combined with free fields). Empirical evidence underdetermines the "correct" description for the radiation of our universe. Whether we employ the retarded solution, the advanced solution, or a linear combination of the two depends on the free-field component. And for any choice of free field plus a given solution to the inhomogeneous wave equation, we can rewrite that description in terms of a different solution, with a different free field.

Nevertheless, it seems we do want to characterize the radiation of our universe as retarded. The linearity of the wave equation does not change the fact that we observe radiation only *after* charges accelerate. Our experience suggests that sources lie to the past of the waves they produce[1]; so that, *prima facie*, the retarded solution is the correct way of describing these waves. We now see, however, that the justification for employing the retarded solution cannot be based solely on this experience.

We can justify attributing retarded fields to sources by appealing to simplicity considerations. Consider using the advanced solution to characterize the light that fills a room after we turn on the switch. This solution corresponds to a wave collapsing onto the bulb just before the charges begin to accelerate, plus a free field that comes in from the beginning of the universe and just cancels out the advanced radiation so that the net field passing through the bulb is precisely what we observe—a field that appears to diverge from the bulb after the charges accelerate. Similarly for any process that seems to involve retarded radiation. Clearly, the free field that must be added to the advanced solution in order for it to describe the radiation we observe is extremely complicated and bizarre. The unnatural, conspiratorial character of this field is then what justifies our using the retarded solution to represent the radiation of our universe. For all that needs to be added to the retarded fields from accelerated charges is a free field that matches the tiny observed background radiation. Hence we have reason to believe that this is what the radiation of our universe is like. More generally, a universe should be characterized as containing retarded and not advanced radiation if, in describing its fields, the retarded solution requires a much more natural free-field component than the advanced solution.

[1] For convenience I sometimes say that waves are "due to" their "sources". It must be kept in mind (as the previous discussion shows) that it is misconceived to think of waves as having sources independent of background field: we can only attribute components to charges relative to a choice of free field.

This reliance on simplicity considerations does render the distinction between a universe in which accelerated charges produce retarded radiation and one in which they produce advanced radiation somewhat imprecise. There may be cases for which we cannot determine whether the radiation is retarded or advanced, such as a universe with only source-free fields or a universe at thermal equilibrium. Nonetheless, these considerations suffice for determining which of the mathematically equivalent descriptions should characterize the radiation of our universe. Indeed, the linearity of the wave equation and the existence of free fields mean we *must* appeal to *something* above empirical considerations in order to single out one solution. And choosing the background radiation on simplicity grounds does the trick.

Note that physics textbooks generally miss this point. They try to rule out advanced solutions on the grounds that they violate basic causal or physical considerations. One book claims the advanced solutions "are not in accordance with elementary ideas of causality" (Schwinger et al. 1998, 346). Another asserts that the solution which "represents a diverging spherical wave" is called the retarded solution "because it exhibits the causal behavior associated with a wave disturbance" (Jackson 1962, 184-5). Another says there is "an *additional fact*—based on experience—that only the outgoing wave solution makes 'physical sense'" (Feynman et al. 1964, §20-14). Such claims beg the question since advanced solutions, with the appropriate free-field term, *do* describe the fields we observe; we cannot simply assume, on the basis of our radiative experience, that all waves propagate outwards to the future of their sources. And to discount the existence of advanced radiation on causal grounds invites the question of how to justify this causal asymmetry in a world of symmetric laws.

In any case, the asymmetry of our radiative experience is evident, and is puzzling in light of the symmetry of the laws governing that experience, regardless of the choice of background field. Yet since we do have reason for employing the retarded description, I will continue to put the puzzle as follows: Why do accelerated charges produce retarded and not advanced radiation? How can we account for this asymmetry with time-symmetric laws?

2. Frisch's Solution

Frisch has a different understanding of the puzzle. He states the explanandum as the fact that, "All accelerated charges…can be associated with fully retarded (but not with fully advanced) radiation fields" (2000, 387). In more detail:

> Even though both retarded and advanced solutions (and any linear combination of the two) are allowed by the Maxwell equations, the retarded solution is that solution to the inhomogeneous wave equation which for a source configuration in the absence of external fields appears to represent the physical situation correctly. The field associated with a single charge satisfies what is known as the 'Sommerfeld radiation condition': the free incoming field…is equal to zero. Of course, this field can alternatively be represented as the sum of an advanced and a source-free field…, but the two representations are not symmetric: the latter representation includes a source free-field, while the former does not (384).

Frisch maintains that this is the sense in which the radiation of our universe appears to be retarded: the advanced solution, but not the retarded one, requires a free-field component when representing a given field. The puzzle is to explain why this is so when the laws make no distinction between the two solutions.

Now we can see why this conception of the puzzle is confused. Frisch notes that the radiation of our universe can be characterized by any linear combination of retarded and advanced solutions. He realizes that we cannot state the asymmetry as the fact that accelerated charges produce retarded and not advanced radiation without some further justification for describing radiation this way. But he is wrong about what constitutes this justification, and this leads him to misunderstand the nature of the puzzle more generally.

Frisch thinks we use the retarded solution because it need not include any source-free term. Since our universe contains free fields, however, *some* such component must be added to the radiation attributed to charges in order to accurately describe the observed fields. This component is quite small when added to the retarded solution, but it is non-zero. Indeed, it is precisely the naturalness of this component that distinguishes the retarded from the advanced solution and justifies our using the former for our universe. *Contra* Frisch, the difference between the two solutions is not that one requires a source-free component and the other does not; the difference is that one requires a very large, unnatural source-free component and the other does not.

Once we acknowledge the existence of free fields, it becomes clear that accelerated charges *can* be associated with fully advanced fields as much as they can be associated with fully retarded ones. This is because, for any description of a field that attributes retarded radiation to accelerated charges, there is an equivalent description according to which they produce advanced radiation. *Both* descriptions, if they are to accurately represent the total field, will contend that some of it is source-free.

This misconception of the explanandum can be traced to Frisch's view on what it is for a field to be associated with a source. Frisch writes that, "The field component associated with a source is simply that component of the total field that would be absent, if the source were absent" (2000, 402). He suggests that we can empirically determine this association by removing a given charge and observing the resultant change in field.

Given the existence of free fields and the linearity of the wave equation, however, the association of field components to sources is not so clear-cut. After all, there are many mathematically equivalent representations of any given field. Each representation accurately describes the observed field even though each says that different components combine to produce it. Specifically, each includes a different free-field component and as a result attributes different components to accelerated charges. Which component is associated with a charge therefore depends on the representation of the total field. And this representation, in turn, depends on the choice of background field. Hence which component is associated with a charge also depends on this choice.

The problem is that the choice of free field is underdetermined by what we observe, so that the association of components to charges is similarly underdetermined by the empirical evidence. We can only determine the field component associated with a charge *according to a given description of the total field*, and, so, *relative to a choice of background field*. We can go on to justify one solution over the others, thereby warranting the claim that accelerated charges produce a certain kind of field. In our universe, choosing the simplest free field allows us to say that accelerated charges produce retarded radiation. But we must remember that this association of components to charges is relative to a choice of background field; and we can make this choice so that charges produce advanced fields while saving the empirical phenomena. Frisch cannot simply assume, independent of simplicity considerations, that charges cannot be associated with advanced radiation.

There is a sense in which Frisch's statement that the retarded solution "for a source configuration in the absence of external fields appears to represent the physical situation correctly" (2000, 384) can be understood as true. By "external fields," Frisch might mean the free-field component of the otherwise fully retarded solution. And it is true that without this part of the total field, the retarded components from accelerated charges sum to the remaining field. But in this sense it is also true that in the absence of (different) external fields, the *advanced* solution will accurately represent the situation. Recall that both the retarded and the advanced solution characterize the observed field, though they differ on what free field exists above the radiation from charges. In order to select one of the representations, we must choose a background field. Once we make this choice in favor of, say, the retarded description, then clearly in the absence of *this* field, the retarded solution will represent the physical situation. For this is the very field whose choice justifies the representation according to which accelerated charges produce retarded fields. Similarly, absent the free field of the advanced description, the advanced fields from accelerated charges combine to yield the resultant field; for *this* is the field whose choice allows us to say that accelerated charges produce *advanced* radiation. Thus, we can interpret Frisch's statement generously, but in this sense it is obviously true and fails to draw a distinction between retarded and advanced solutions. Otherwise, it is simply false, contending that the fully retarded solution—with no free-field component—accurately represents any field of our experience.

This misconception of the puzzle leads Frisch to misunderstand what his own solution amounts to. Frisch thinks we want to explain why advanced solutions, with no free-field term, fail to characterize actual fields. For him, the puzzle is generated by the existence of solutions to the laws that do not correspond to actual phenomena. His answer is that there is a further law—the retardation condition—that deems the advanced solutions impossible. So that when we take all the laws into account, all the solutions *do* correspond to actual processes. He concludes that his account is not so much a solution to our puzzle as a means of denying there was any real puzzle to begin with. In his words: "If the radiation condition were a law just like the Maxwell equations, then the fact that radiative phenomena satisfy the condition would be no more in need of an explanation than the fact that they satisfy the Maxwell equations" (2000, 406).

Contra Frisch, the puzzle cannot be to explain why advanced solutions fail to correspond to what we observe, since they *do* (or *can*) characterize this radiation. More generally, the problem is not simply to explain why some physically allowable processes fail to be actual. The puzzle is to explain why accelerated charges produce retarded and not advanced radiation: we want an account of the asymmetry of our radiative experience given that the laws are temporally symmetric. This observed asymmetry remains unaccounted for despite our views on the relation between fundamental laws and physical possibility.

There is, then, a genuine puzzle here. And Frisch has proposed an answer to this puzzle, namely: it is a law that accelerated charges produce retarded radiation.

But this solution cannot succeed. Given the existence of free fields, Frisch's retardation condition is not true of our universe: not all electromagnetic fields satisfy this constraint. The vagueness behind the statement that charges produce retarded radiation further suggests that it is wrongheaded to posit an exact law of retarded radiation. Finally, the retardation condition, as a non-statistical generalization, should be derivable from initial conditions plus the deterministic Maxwell equations. This throws doubt on Frisch's claim to be positing an additional fundamental law on a par with Maxwell's equations. What is more, this would mean that there *is* an explanation for the retardation condition's holding, since it follows from initial conditions plus the fundamental dynamical laws.

3. The Radiative Asymmetry and the Initial Condition of the Universe

I think we can do better. I think we can account for the radiative asymmetry with the laws we already have rather than positing the retardation condition as an additional law. I want to suggest that this asymmetry results from the initial condition of the universe.

First, let us review how the nature of the initial state helps solve the puzzle about the asymmetry of thermodynamics. This problem is to explain why entropy does not decrease towards the future when the laws governing the particles of thermodynamic systems are symmetric in time.

The traditional solution relies on a notion of entropy defined in terms of volumes of regions of phase space. One point in a 6N-dimensional phase space represents the microcondition of a (classical) N-particle system. Its macrocondition corresponds to a region of phase space. Each point in this region picks out a microcondition compatible with the system's being in that macrocondition. Boltzmann found that the entropy of a thermodynamic system is a function of the number of arrangements of its particles compatible with its macrocondition. Hence regions corresponding to higher-entropy macrostates take up much larger volumes of phase space than regions corresponding to lower-entropy macrostates. The equilibrium condition takes up by far the largest such volume.

The partitioning of phase space into distinct macroconditions is therefore extremely uneven. This allows us to understand entropy increase as the progression towards more and more probable macrostates. If we count each microcondition that a system might evolve into as equally probable, then it is overwhelmingly likely that the system will evolve into a microcondition compatible with the macrocondition occupying the largest region of phase space. And this is the highest-entropy macrocondition.

This is not enough to account for the *asymmetry* of thermodynamics, however. Since the dynamical laws are time reversal invariant, the uniform distribution plus the laws yield overwhelmingly high probability of entropy increase in *both* temporal directions.

We can solve this difficulty by appealing to an asymmetry in the boundary conditions. If we assume that entropy was lower in the past, and take the uniform distribution over microstates at that time, then the dynamical laws predict overwhelmingly likely entropy increase to the future of that time. In order to obtain an empirically adequate account of thermodynamics for the entire history of our universe, then, we must assume that the universe began in an extremely low-entropy state, what Albert (2000) calls the Past Hypothesis.

Perhaps the Past Hypothesis can similarly account for the asymmetry of radiation. Consider the state of the universe immediately following the big bang. The matter and fields were evenly distributed at uniform density and temperature. There was a lot of gravitational potential energy: energy could be converted into heat (as matter came together under gravity); once created, this heat is unlikely to be converted back into gravitational energy. So this was a state of extremely low gravitational entropy[2]; a clumped up state would have high gravitational entropy. However, since everything was in one uniformly hot "soup", the universe was at thermal equilibrium. It is the extremely low entropy due to gravity that gives the early universe the low entropy of the Past Hypothesis.

From this state, the material particles began to clump up under gravity, forming hot masses like stars. This clumping constitutes a progression towards gravitationally higher-entropy macrostates and can be understood similarly to the progression towards thermodynamically higher-entropy states—as overwhelmingly likely given the initial low-entropy state. This clumping results in a decrease in thermal entropy, however, since as the universe expands, it moves to a state of hot clumps of matter in a colder surrounding space.

[2] 'Gravitational entropy' is not meant to indicate something different from Boltzmannian entropy, only that it results from gravitational forces. Higher gravitational entropy states should still correspond to larger volumes of phase space (think of it in momentum space).

Thus, as material particles clump up due to the initial low gravitational entropy, the universe moves farther and farther away from thermal equilibrium.

In order to get back to thermal equilibrium, accelerating charges in the hot masses would have to radiate energy into the surrounding space. The temporally symmetric laws say that both advanced and retarded radiation could be emitted. However, given the universe's thermal disequilibrium, the charges are overwhelmingly likely to radiate towards the future, as part of the overwhelmingly likely progression towards equilibrium in that temporal direction. They are overwhelmingly unlikely to radiate towards the past because the universe was at thermal equilibrium in that direction. Note that on this view the retarded nature of radiation is statistical: advanced radiation is not prohibited but given extremely low probability.[3]

This explanation should carry over to the quantum realm. Indeed, we must turn to quantum theory. For there is no adequate classical account of an equilibrium state between matter and fields. Hence there is no classical description of the initial condition of the universe. Quantum electrodynamics, however, treats matter and field particles in the same way and can describe an equilibrium between them. This theory, as explained in Feynman (1985), gives the probability amplitudes for photons to be at given space-time locations; for electrons to be at given locations; and for the emission or absorption of a photon by an electron. An accelerated charge has a probability amplitude of emitting a real photon, which we would call the emission of retarded radiation. It also has a probability amplitude of absorbing a photon, or emitting

---

[3] The analogy with thermodynamics underscores a difficulty with Frisch's position. Callender (2002) notes two viable approaches to the thermodynamic asymmetry: (1) posit an asymmetry in boundary conditions; (2) posit additional, non-time-reversal invariant laws restricting physically possible worlds to entropy-increasing ones. For the radiative asymmetry, I advocate (1). Frisch considers his solution an example of (2). Since the retardation condition should be derivable from initial conditions plus Maxwell's equations, this renders his solution more along the lines of (1).

advanced radiation. The probability for either kind of radiation depends on the probability amplitudes for the number of photons around and on whether the charge is in an excited state.

In the quantum picture, the early universe comprised photons in thermal equilibrium with material particles: there were, on average, just as many photons emitted as absorbed. As material particles clumped up and the universe expanded, the universe moved to a state containing charges in excited states (in the hot masses) and photons less densely distributed throughout space. These factors result in a high probability of emission of photons towards the future by accelerating charges, and a low probability of emission of photons to the past. The probability for retarded radiation will remain high until thermal equilibrium is reestablished. In quantum electrodynamics, then, just as in classical electromagnetism, which kind of radiation is likely to occur depends on the initial state. The observed asymmetry is then explained as overwhelmingly likely given the extremely low entropy of that state.

On this story, the Past Hypothesis is the source of the retarded nature of radiation in our universe. This is therefore a simpler, more unified account than Frisch's. First, it accounts for the radiative asymmetry with the laws we already have. At least, this is so if we regard the Past Hypothesis as a law, and there seem many reasons to do so, e.g., it is a simple, informative generalization that supports counterfactuals and yields successful predictions. Second, this account provides a similar explanation of the field and particle arrows of time. This unification is particularly appropriate when we turn to quantum theory and its comparable treatment of matter and field particles. Both the asymmetry of radiation and the asymmetry of thermodynamics can therefore be explained within a universe of time-symmetric laws given the Past Hypothesis.[4]

References

Albert, David. Z. (2000), *Time and Chance*. Cambridge, MA: Harvard University Press.

Callender, Craig (2002), "Thermodynamic Asymmetry in Time", *The Stanford Encyclopedia of Philosophy* (spring 2002 edition), Edward N. Zalta (ed.), URL = http://plato.stanford.edu/archives/win2001/entries/time-thermo/.

Feynman, Richard P. (1985), *QED: The Strange Theory of Light and Matter*. New Jersey: Princeton University Press.

Feynman, Richard P., Robert B. Leighton and Matthew Sands (1964), *The Feynman Lectures on Physics*, vol. II. Reading, MA: Addison-Wesley.

Frisch, Mathias (2000), "(Dis-)Solving the Puzzle of the Arrow of Radiation", *British Journal for the Philosophy of Science* 51: 381-410.

Jackson, John D. (1962), *Classical Electrodynamics*. New York: John Wiley & Sons.

Price, Huw (1996), *Time's Arrow and Archimedes' Point: New Directions for the Physics of Time*. New York: Oxford University Press.

Ridderbos, T. M. (1997), "The Wheeler-Feynman Absorber Theory: A Reinterpretation?", *Foundations of Physics Letters* 10 (5): 473-486.

Schwinger, Julian, *et al*. (1998), *Classical Electrodynamics*. Reading, MA: Perseus Books.

Sciama, D. W. (1967), "Retarded Potentials and the Expansion of the Universe", in T. Gold (ed.), *The Nature of Time*. Ithaca: Cornell University Press, 55-67.

[4] I thank David Albert, Frank Arntzenius, Mathias Frisch, Tim Maudlin, Bradley Monton, and Huw Price for helpful comments.