Response to ‘Local and global definitions of time: Cosmology and quantum theory’

Sarah Shandera
Pennsylvania State University

Abstract

I will discuss the notion of time and spatial finiteness from the perspective of observational cosmology. Our observed universe is well described at early times by small fluctuations in the spatial gravitational field, distributed homogeneously and isotropically on an otherwise smooth background spacetime. The dominant paradigm for an even earlier phase of the evolution of our universe typically generates a space similar to what we observe but with vastly larger spatial extent. Even in a classical theory, the fact that we observe only a finite volume of space and time means that there are statistical uncertainties in testing and constraining the theory. In the absence of evidence that our observable universe is all there is, the cosmological principle suggests that we should consider ourselves to be a typical region of a vastly larger space. I will consider the problems of time and largest spatial scale of the universe from the perspective of accepting an imperfect and incomplete theory on scales currently out of reach of our necessarily imperfect and incomplete observations. The question then is how to work in incremental steps to push for a broader range of understanding.

1 Introduction

In cosmology, as in the rest of physics, we believe that the scientific process works because we do not have to know everything in order to say something. More precisely, we believe we can describe the world we observe to nearly the accuracy at which we currently observe it, in terms of phenomena accessible at the scales we observe. The gap between observation and description, whose size fluctuates as new frontiers in energy, larger or smaller physical scales, or in complexity are opened and explored, is where new ideas evolve, but the gap is finite. For any given measurement there will be energy scales and time scales and spatial scales that we must be able to discard as parametrically unimportant compared to those we describe. The belief that this process is sensible in both the classical and the quantum domains is supported by the sum weight of evidence from our progress in physics so far: the idea itself is justified on the spatial, temporal, and energy scales we have successfully described. As with the rest of our laws of nature, it doesn’t have to continue to hold as we explore new phenomena, but it is hard to conceptualize what research would look like without this principle. Even within the range of scales we can describe well, we know from experience that any given description of a phenomena need not hold, or be the most calculable description, over a very large range of energy or length scales. There are often degrees of freedom whose dynamics need not be calculated precisely, but that can be replaced by an effective description in terms of fewer degrees of freedom.

In cosmology, we observe a finite region of space, a finite range of times, and we are able to make measurements only with a finite precision in both spatial and temporal resolution. So far none of these constraints are fundamental limitations: As our measurement techniques improve, we will be able to measure the properties of the universe to a better resolution. If we someday develop technology to detect the cosmic neutrino background we would have information from a slightly earlier time. As we wait, light from farther away will reach us (although, depending on the future expansion rate, we may not see an infinite volume even if we wait an infinite time). However, the rate at which we get information from larger spatial scales, thanks to the arrival of more photons, is insignificant compared to what many theories predict for the difference between the size of the region of space we currently see and the size of the region of space that we expect looks, statistically, more or less the same as what we see. For all practical purposes, our observable universe is spatially and temporally finite.

At the earliest time we can observe directly, and the earliest time about which we have indirect but compelling evidence for from thermodynamics, the universe is well described by a nearly homogeneous and isotropic distribution of a certain proportion of different types of matter and energy. The matter and energy determine how the universe evolves on the largest scales, and so the average total density can be used as a cosmological clock. On top of the smooth background at the earliest time are small inhomogeneities in the gravitational potential, which source small variations in the density of matter and radiation. These small perturbations evolve gravitationally according to well known rules and are responsible for both the inhomogeneities in the temperature of the cosmic microwave background and the inhomogeneities in the dark and visible matter, which grew into the galaxies and clusters of galaxies. In this paper, we will be concerned with how we make progress in uncovering a theory for an earlier phase in the history of the universe, which ideally provides an origin for both the homogeneous and inhomogeneous features we observe. For simplicity, we will consider the inhomogeneities in the density field as the representative observable quantity.

We only postulate theories that satisfactorily explain the homogeneous universe and compare them by asking how well they explain the inhomogeneities. This distinction is practical: we continue to gather increasingly detailed data about the inhomogeneities, but additional details about the homogeneous universe that would distinguish theories are so far out of observational reach. Our description of the inhomogeneities is statistical: models of the universe don't predict precisely which regions of space will contain objects of a certain size, but they do predict how many objects of a given size we are likely to see throughout the universe. We are satisfied with a theory that predicts the statistics of the inhomogeneities, with the understanding that the same theory could have resulted in an observable universe that is different in the particulars, but statistically compatible with ours. To study whether a given theory successfully matches observations, we create many realizations of the statistics predicted by the theory and calculate the probability that we observe the cosmology that we do, given the theory. We compare two theories by comparing the probability that each theory gives the sky we observe (possibly weighted by a prejudice for one theory or another based on something other than the cosmological data set).

A key point is that the inhomogeneities as we observe them today, and that we use to describe the universe at a much earlier time, are classical and are necessarily defined with respect to our observed universe. That is, we define the perturbations as deviations from respect to the mean density, which in turn can be defined by smoothing out, or averaging, the total density field on the scale of our current observed universe.

Although we can describe the observed history of our universe by postulating the matter and energy content, together with the statistics of the inhomogeneities, at some earlier time, this is aesthetically unsatisfactory because of an important observed feature: the inhomogeneities look statistically the same on scales that were never in causal contact if we just run the cosmology we see back in time. If we just choose a time to call the initial time, and postulate the statistics as 'initial conditions', it looks extremely ad hoc. Instead, we prefer to take the similarity of the inhomogeneities as evidence for an earlier era, dominated by different energy sources than those in the era we observe directly, during which the perturbations that we now see on large scales were in causal contact. The simplest idea for the primordial era is called inflation, a phase of accelerated expansion. Furthermore, the simplest realizations of this idea are quite appealing because they simultaneously generate a nearly homogeneous background together with small inhomogeneities. The inhomogeneities are understood to originate in quantum fluctuations that were unavoidably present during inflation and stretched to very large scales where they effectively become classical. Although formal details of this scenario remain uncertain in some respects, the qualitative picture has been enormously successful in providing a route to explain both qualitative and quantitative features consistent with observations. Inflation may not be the correct or complete description of the earlier phase of the evolution of our universe, and it is not the only idea available. The ideas discussed below may be relevant for discussing to what extent we could ever 'prove' inflation correct or falsify it, but for simplicity in what follows I will largely assume that inflation is the paradigm we want to test.

2 The primordial era, space and time

The primordial era need not last any longer than what is strictly required to put the largest scales we see today in causal contact. However, in most scenarios the inflationary phase lasts much longer. In that case the entire universe we currently observe is a small region of an exponentially larger space that is well described by a homogeneous background and small fluctuations. Inflation itself may have been preceded by an even earlier phase with a completely different evolution. As long as the pre-inflationary era was followed by a long inflationary phase, we may well have no observational evidence of anything earlier than inflation now or for millennia to come. This uncertainty means that, from an observational point of view, we need only require that a theory gives a consistent description on the scales relevant for observations. Anything more is untestable. While we cannot resist trying to extend our theories to higher energies and earlier times through pure logical consistency, there is no reason to expect that there is only one such theory compatible with current observations. The search for a 'theory of everything', and the existence of many, many solutions in field theory and string theory have shown it may not be that all things about our observed universe are deterministically explainable as the only possible universe the laws of physics allow. A century after the discovery of quantum mechanics, we still don't know how to make sense of the possibility that our best theory may ultimately leave us with more than one possible universe, and no way to decide between them.

In practice, our description of the inflationary universe makes sense only if we can implement some smoothing procedure so that when we course-grain the description on sufficiently large scales it appears homogeneous. Then there is a good notion of a clock in that region. Although this smoothing procedure may be difficult to implement in the action, during the dynamical phase, we expect to be able to apply it to a region that is well described in terms of a homogeneous background with small fluctuations on some time slice after inflation.

In the current understanding of the theory, there are two ways in which we think the spacetime may in fact be infinite. There may be perturbatively connected regions that continue to inflate even though in some regions of spacetime inflation has ended. Over scales that connect our observed universe to the continually inflating patches, the description of a common homogeneous background with small fluctuations breaks down. In this case, there is not a global notion of time. There may also be non-perturbatively connected regions in which there may again be an entirely different notion of time, or even none at all (if the region is locally classically static). In both cases, we have techniques for computing the fluctuations expected in a local region where inflation ends, corresponding to our observable universe, that do not suffer for infrared divergences. But, the infinite spaces still raise very difficult and unresolved questions if we want to ask why we find ourselves in region of the universe in which inflation ended. However, we will see that the map between observation and theory can be fuzzy long before reaching that point, which is so far intractable.

3 Mode coupling, particle interactions, and statistical uncertainties

Even in a purely classical picture, the finiteness of our universe together with the statistical nature of the inhomogeneities may make it difficult or impossible to find a one-to-one map between observations and any theory for a space larger than what we observe. This is true if inhomogeneities on different wavelengths (different Fourier modes) are coupled. In the theory of inflation, or any competing ideas, mode coupling is inevitable when there are interactions between the fields relevant for the early universe. The fact that we observe fluctuations at all implies that at least there was a primordial matter field coupled to gravity, so some level of mode-coupling is inevitable. Many scenarios have more than one light particle relevant during or after inflation, which can lead to much higher levels of coupling. In fact, constraints from measurements of the temperature fluctuations in the cosmic microwave background limit the amount of mode-coupling to a level that is quite small compared to typical interaction strengths found in the particle physics we probe in colliders.

Interestingly, this observational limit may or may not mean that there is no mode-coupling in the universe. It may just mean that the region of the universe that has properties close to what we observe is exponentially larger than what we observe. This would be a fitting extension of the history of cosmology so far, where we have discovered that the observable universe is exponentially larger than the typical scale of human activity, some 10^{13} times larger than the size of our solar system, and about 10^5 times the size of our galaxy. The cosmological principle is usually taken to mean that our place in the observed universe is not special. One might extend that to say that our place in the entire universe is not special. The extension is relevant for theories where the primordial fluctuations on scales within our observable universe are coupled to long wavelength scales that are outside our observable universe and by definition un-measurable.

To make this idea more concrete, we assume that there is a finite region, exponentially larger than what we observe, that is well described with a homogeneous background that can be used to define a notion of time. Over this entire region we have a statistical prediction for the properties of small fluctuations. One might define the boundary of this large region purely as the largest scale on which fluctuations are small. In a particular inflationary scenario, one might define the size of the region by the length of time inflation lasted, since the description of the physics on larger scales is certainly different. However, it is important to note that this description, and the conclusions that follow below, are independent of the dynamics for generating the inhomogeneities. In addition, it applies to classical inhomogeneities. In that sense, the results apply to any theory that generates inhomogeneities consistent with observations so far.

If short wavelength fluctuations are coupled to long wavelength fluctuations, the amplitude and coupling of modes within a small volume may depend on the amplitude of long wavelength modes. However, any long wavelength modes that are constant over the small volume are not independently observable: they only contribute to the local background and shift the values of properties of the statistics in the small volume. If this shift changes parameters that we consider crucial evidence for or against inflation, it throws into doubt our ability to test theories of the primordial universe in a meaningful way. We also do not know the extent of the universe beyond what we observe that looks statistically similar to what we observe. Since the amplitude and number of long-wavelength modes that our local physics may be coupled to are all unknown, the total uncertainty in the relation between our observed universe and the statistics averaged over a much larger volume may be large. These effects have not yet been fully explored, but they must be understood before we can claim that our interpretation of the beautifully precise cosmological measurements is water tight. As before, the best we can do is to compare competing theories by computing the probability that those theories produce the observations we see.

4 Conclusions

Cosmology is a field of physics where questions of scale and finiteness in space, time, and energy take on a different flavor from many other fields of physics. Much of the observational power to distinguish between different theories for the primordial universe comes from observations of inhomogeneities, which we describe statistically. However, the number and scale of the observations we make are finite, and we can never access additional information on large scales. Furthermore, our observable universe may have properties that are coupled to long-wavelength information that is forever unobservable to us. And, we have no way of 'repeating the experiment' to average over all possible long-wavelength physics. The consequence of this is that, in the absence of evidence that the entire universe is finite on scales close to the largest scales we observe, we must apply the cosmological principle and assume our observed universe is a typical subset of whatever space is predicted by a theory of the primordial era. When modes of different wavelengths are coupled, the statistics we observe are likely to deviate from the statistics averaged over a much larger volume in a way that is calculable given a theory. The effects of mode coupling may be relevant for interpretations of key cosmological parameters, like the amplitude of fluctuations as a function of scale. On the other hand, the cosmological principle together with some types of mode coupling may lead us to expect that it is 'statistically natural' to see only very weak mode coupling in small subvolumes.

The statements above all apply to a purely classical description of the matter and energy distributions in the universe. They leave aside additional questions about the evolution of quantum fields in some pre-inflationary era that may give rise to the classical fluctuations. However, there are large uncertainties in the matter content relevant during the primordial era, as well as in the thermodynamics involved in the transition from the primordial era to the presently observable era. Given that there is very little hope of observationally constraining those aspects in the near future, it is worth a more careful study of what aspects of *any* primordial theories are truly robust when only a finite range of scales and number of modes can be measured. These issues are, pragmatically speaking, important to address before asking the more difficult questions about quantum mechanical quantities defined over unobservably large or infinite volumes.