Local and global definitions of time: Cosmology and quantum theory

William Nelson
Radboud University Nijmegen,
Institute for Mathematics, Astrophysics and Particle Physics.

April 3, 2013

Abstract

I will give a broad overview of what has become the standard paradigm in cosmology. I will describe the relational (à la Leibnitz) notion of time that is often used in cosmological calculations and discuss how the local nature of Einstein's equations allows us to translate this notion into statements about 'initial' data. Classically this relates our local definition of time to a quasi-local region of a particular spatial slice, however incorporating quantum theory comes at the expense of losing this (quasi-) locality entirely. This occurs due to the presence of two, apparently distinct, issues: (1) Seemingly classical issues to do with the infinite spatial volume of the universe and (2) Quantum field theory issues, which revolve around trying to apply renormalization in cosmology.

Following the cosmological principle - an extension of the Copernicus principle - that physics at every point in our universe should look the same, we are lead to the modern view of cosmology. This procedure is reasonably well understood for an exactly homogeneous universe, however the inclusions of small perturbations over this homogeneity leads to many interpretational/ conceptual difficulties. For example, in an (spatially) infinite universe perturbations can be arbitrarily close to homogeneous. To any observer, such a perturbation would appear to be a simple rescaling of the homogenous background and hence, physically, would not be considered an inhomogeneous perturbation at all. However, any attempt to choose the physically relevant scale at which perturbations should be considered homogeneous will break the cosmological principle i.e. it will make the resulting physics observer dependent. It amounts to 'putting the perturbations in a box' and a delicate practical issue is that the universe is not static, hence the scale of the box will be time dependent. Thus what appears physically homogeneous to an observer at one time will not appear so at another.

This issue is brought to the forefront by considering the canonical (space and time rather space-time) version of the theory. The phase space formulation of General Relativity, just as for any other theory, contains the shadow of the underlying quantum theory. This means that, although

the formulation is still classical, many of the subtleties that are present in the quantum theory are already apparent. In the cosmological context the infinite spatial volume renders almost all expressions formal or ill-defined. In order to proceed, we must restrict our attention to a cosmology that has some finite spatial extent, on which our relational notion of time is everywhere definable. In particular, this would constrain the permissible data outside our 'observable universe'.

This difficulty is an IR or large (spatial) scale issues in cosmology, however in addition there are UV or short (spatial) scale problems that need to be tackled. There are the usual problems of renormalization, which are further complicated by the fact that the universe is not static. In the cosmological setting this leads to new IR problems which again prevent one from taking the spatial extent of the universe to infinity. The physical relevance of this problem, the consequence for defining a time variable, and the distinction of homogeneous and inhomogeneous IR issues will be discussed.

1 Physical cosmology

The modern view of cosmology is a profoundly phenomenological one. One interprets observational data in terms of approximate solutions to Einstein's equations assuming the matter content can be thought of as a fluid. The observations themselves are often only proxies for the desired observable, themselves containing various levels of approximations or assumptions. In addition, cosmology represents an enormous extrapolation of scales - 15 orders of magnitude or more - in energy density, acceleration, length and time from the more familiar laboratory scale physics that has been tested. Given these caveats, one should not expect that cosmological observations perfectly agree with our currently established understanding of physics, and indeed this is not the case.

Observations have shown that on galactic scales, evolution does not follow from Einstein's theory for any know matter content. Whilst there is some debate as to whether this Dark Matter may, in fact, be due to our lack of understanding of gravitation on these scales (MOND, TeVeS etc.) or perhaps a consequence of our approximations, it is generally believed that it is a phenomenological description of some, as yet undiscovered, types of matter. Indeed it would be rather more fantastic to think that all the particles in existence have properties that can be probed in the laboratory with current technology, and that there is nothing left to discover.

On much larger scales, approaching the size of the observable universe, it has also been shown that the expansion of space does not follow our naive expectations. The phenomenological description of this effect is Dark Energy and its underlying explanation is rather more open to debate, with possibilities ranging from a new fundamental scale, modifications to general relativity, misuse of averaging in cosmology, undiscovered types of matter with radically different properties to those we currently observe etc. In addition to these possible classical resolutions, there is also much work on the possibility that Dark Energy is a consequence of the underlying quantum theory.

At this point one may start to be concerned. We have a phenomenological description of our universe that can match observations only if we postulate two new sources of matter/ energy, that together account for approximately 96% of the total current matter/ energy content of the universe. However, whilst we may know little of the underlying physics of these new sources, there are a myriad of observations that all agree on their overall cosmological effect. These observations paint a coherent picture of the evolution of the universe over the last 14 billion years or so, a picture that is continually being tested by new observations and predictions. Whilst we may not know what it is made of, we can at least see the whole elephant!

This wealth of every more precise observations, and the remarkable success of the phenomenological approach, allows us to tackle the underlying conceptual challenges inherent in cosmology, with the real possibility of testing our conclusions in the near future. An excellent example of this was the question of why regions of space that have not been in causal contact look (statically) the same? How could different regions 'know' the properties of the others? One explanation is that there was an earlier period Dark Energy-like expansion - known as Inflation - which ensures that the entire observable universe was, in fact, in causal contact. Of course there are other possibilities, such as a universe with an infinite past (for example a bouncing or cyclic cosmology), however different solutions to this conceptual difficulty led to different predictions for other observables. Whilst nothing has yet been completely ruled out, the current observations are able to distinguish between the most basic versions of these (and other) possibilities.

2 Spatial extent and time

Of the many important conceptual challenges facing any approach to cosmology I will concentrate on two: the spatial 'size' of the universe and the notion of time. Here I will argue that it is very likely that these two issues are intimately related, with the relation coming from the underlying quantum gravity formulation of cosmology (what ever it may be).

2.1 Time

Here, in line with the phenomenological under pinnings of cosmology, I shall consider a very practical notion of 'time', taking it to be given entirely by relative change. I will consider time as a relational concept, which can be defined for any two sub-systems. If one of these sub-systems is considered a 'clock', the state of the others evolve with respect to the internal time given by the evolution of this clock. For example, consider a system containing as sub-systems a swinging pendulum, a revolving wheel and an observer capable of observing both. This observer may choose to use the pendulum as a clock and ask questions such as "How many revolutions of the wheel are there in one cycle of the pendulum?" Of course, the observer may equally decide to use the wheel as a clock and ask "How many cycles of the pendulum are there per revolution of the wheel?" The choice of which sub-system the observer chooses to use as a clock is irrelevant, only the relative change between them is important. Note that the inclusion of the observer here is important. If neither the pendulum nor the wheel were in motion, the observer would be unable to use either device as a clock, however the fact that the observer interacts with them (via observing), implies that there must be another sub-system that could be used for the observer to measure time. Namely the field propagating this interaction (i.e. photons traveling from the observer to the devices). Thus, should the clock be static, it may not be a useful means of measuring time, however the observer's act of seeing the clock means that time can still be defined, simply by using alternative degrees of freedom to register relative change. Consider the following example. A person observes a clock to read mid-day. A second, distinct, observation is subsequently made and again the clock reads mid-day. There are three conclusions that the observer might draw,

- $12N$ hours have elapsed between observations, for N a strictly positive integer,
- the clock is in fact broken and some indeterminate period of time has passed,
- no time has passed between observations.

The last case can be ruled out by the fact that the observer carries out *two distinct* observations. That is, other degrees of freedom - say the photon propagation from the observer to the clock - distinguish the two events. Similarly the first possibility does not present any difficult, since although this particular observation may produce in the same result - the clock reads mid-day - other observables would not (for example the charge in the battery). Finally the second possibility is the case in which the sub-system the observer has decided to use as a clock is static. Clearly this does not make for a useful method of measuring the passage of time, however, equally clearly, it does not mean that time is not definable. Other degrees of freedom that are responsible for the interaction of the clock and the observer could be used to define this relational notion of time. In practice this means that when we are choosing a particular degree of freedom of our system to be a clock, we want to ensure that its evolution is monotonic, at least for the period of our observation. If it is ever static, we must choose a different degree of freedom. In the event that *every* degree of freedom of the system is static, then, indeed, relative time would not be definable, however in such a situation the observer would also have to be static and evolution would have no meaning.

Typically we want to choose a clock whose evolution is largely unaffected by the state of the other sub-systems (we might prefer that our watches tick at a uniform rate despite swinging of our arms for example), however this needn't be the case. In cosmology for example, we may choose to use the expansion of the universe as a clock and ask "What is the state of some matter content when the expansion is such and such?" or "How does this state change when the expansion varies from x to y?". The evolution of the matter content is highly dependent on the expansion of the universe, however we can still describe the evolution of one parameter in terms of another. In cosmological notation this amounts to writing $\phi(a)$, where a is the scale factor and ϕ is some particular matter content. This practical, relational notion is very typical for cosmological observations, where one regularly refers to time via red-shift (closely related to expansion) and considers the relative change of energy density, temperature, galaxy number density etc. with respect to red-shift. In early universe cosmology it is also common to consider the evolution of matter fields (in particular the inflaton) as defining a clock, thus writting (for example) $a(\phi)$.

2.2 The relation to spatial extent

To demonstrate the plausibility of a link between spatial extent and temporal evolution consider the following. General relativity provides us with local equations, the solution of which in any finite region can be calculated from data specified on the boundary of that region. Typically we consider foliating the manifold using 3-dimensional slices, labeled by a parameter t, in which case the required data can be specified at an 'initial' slice. The solution in any finite region on any other slice requires data to be specified only on a (different) finite region of this initial slice. For example, in order to predict the motion of an apple as it falls, one does not need to specify the data describing the sun at the time the apple is dropped (provided the apple hits the ground within 8 minutes - the light travel time from the sun to the Earth).

This heuristic explanation applies to any system that is 'well-posed' - of which general relativity is an example. It essentially says that solutions can be calculated by 'evolving' initial data in time and that if we are interested only in a finite temporal extent, then we need only specify a finite spatial extent of initial data (provided the speed of light is finite). The passage of time for some (finite) system can then be operationally defined by the expanding spatial region of initial data required to uniquely specify the evolution of that system. That is, given the state of an observable and a clock at some instant, to know how the observable will change in relation to the changing clock we need only specify the initial data required to calculate this evolution. The longer the evolution, the more initial data that is required.

Consider for example an observer and a clock on some slice Σ_t and suppose we provide initial data on a spatial slice Σ_i, prior to the slice containing the observer (i.e. some time before t). Both the observer and clock are a (presumably extremely complicated) solutions to the coupled general relativity and (classical) particle physics evolution equations, and they are given uniquely by the data contained within a certain (finite) region of the initial slice, $S_t \subset \Sigma_i$. Given this data, there is a unique solution describing the observer's evolution in time, as measured by the changing of the clock, from the initial slice to the slice labeled by t. Now we wish to understand what is required in order for this observer to evolve from the state of the clock given at t to the state at $t + \delta t$. The required information is simply the data specified in $S_{t+\delta t} \subset \Sigma$ and so we can consider the evolution in time of this particular observer, from t to $t + \delta t$, to be given by the additional data required in order to go from S_t to $S_{t+\delta t}$

i.e. $S_{t+\delta t} \cap (S_t)^c$, which is entirely defined on the initial spatial slice.

This transforming of the temporal extent of a solution to the spatial extent of the initial data is made possible by the existence of a fundamental velocity - the speed of light. This leaves aside issues such as singularities at which both spatial and temporal extent lose much of their meaning, however it does provide a useful definition of time, for all practical purposes, throughout most eras of cosmology. This entire discussion is a roundabout way of describing the very familiar concept of 'the observable universe' often used in cosmology - how much of the universe at some earlier time is in principle observable to us today? In the notation above, the current observable universe is just S_{today}, where the 'initial' data is specified on some slice given very early in the cosmological evolution (close to the big-bang singularity).

3 Infinite spatial extent

If we are interested only in evolution of a spatially finite system for finite times, only a finite region of the initial spatial slice is relevant, the rest is unimportant for this evolution and in particular the total spatial extent of the initial slice is not relevant. Of course, if we want to evolve a solution for infinite time, one must provide data on the entire spatial slice and in this case the total spatial extent of the universe is relevant. If our universe were exactly spatially homogeneous this question of spatial extent becomes redundant, even for evolution between infinite times. Since every point on the initial slice is equivalent, every point on any other slice must also be equivalent to each other (although there may of course be differences between slices). In this spatially homogeneous situation, all our solutions depend only on one function that relates any spatial region on different slices, the scale factor $a(t)$.

Before we move away from the classical theory it is worth noting that the above discussion also means that, for practical purposes, it does not matter whether the spatial extent of the foliating slices are infinite or not. Only a finite region of the 'initial' slice is relevant for the evolution of any particular observer (such as ourselves). Thus in classical cosmology there is essentially no difference between a universe that is spatially infinite and one that has some finite spatial extent, provided it is larger than the current size of the observable universe. Indeed observations can put limits on the minimum size of the spatial slice (and to some extent its topology), but these observational limits cannot extend beyond the observable spatial size of universe. There is of course a very important conceptual difference between a spatially infinite and a spatially finite universe, in particular when it comes to trying to define what constitutes a 'special' position or conditions. There are also some technical differences that effect the practical calculations typically used in cosmology such as Fourier series vs. Fourier transforms. However for cosmological observations the distinction is unnecessary.

A word on causality. The above discussion is very similar to the notion of causality, however it is not quite the same. Causality would usually be defined as follows. Consider a set of initial data on some slice and add to it a localized disturbance. Causality now describes the fact that, under evolution, the consequences of this disturbance propagates slower than (or equal to) the speed of light. It is certainly true that causality implies that any disturbance outside our observable universe cannot effect us, and hence that there is essentially no difference between an (spatially) infinite and a (spatially) finite universe. However the converse is not true. As we will see, a theory may be non-local and hence 'know' if the universe is infinite or not, whilst still obeying causality.

Finally, note in particular that our relational definition of time can be defined for any finite evolution, given by a finite initial data S_t, completely independently of whether the spatial slice is finite or infinite (provided it is sufficiently large to encompass S_t). Indeed it can be defined completely independently of whether regions on the initial slice, outside S_t have any definable notion of time at all. In this sense, relational time, for finite evolution, can be defined 'locally' (i.e. it is insensitive to the global properties of the spatial slice).

4 The consequences of quantum theory: Hamiltonian formulation

So much for classical cosmology, what changes when we try to incorporate quantum theory? Quantum theory is not a local theory and any attempt at quantization fails when we have a spatial slice that is infinitely large. Indeed, this difficulty can be seen even before going to the quantum theory, by considering the Hamiltonian or action formulation of the classical theory. Both these frameworks contain the 'shadow' of the underlying quantum theory and already at this (classical) level the problem of an infinite spatial extent arises. The Hamiltonian generates evolution of the system via Hamilton's equations which we may hope to use as a practical definition of time, since they describe how the observables evolve with respect to one and other, one of which we may choose to refer to as a clock. However the Hamiltonian is defined as an integral over the entire initial spatial slice. In a sense, the Hamiltonian is defined so that it can generate the evolution of all possible regions of the initial slice for all possible times. The only ways for this to be well-defined are that the initial slice has a finite spatial extent with specified boundary conditions or that the functions giving the Hamiltonian vanish sufficiently quickly 'at infinity' (both possibilities also require that the initial data be suitably regular everywhere). The latter is the usual condition used in mechanical systems, where the spatial extent of the system being consider is obviously finite, however in cosmology this would mean that our position in the universe is somehow 'special' which breaks cosmology's guiding principle and in any case, precludes the existence of a solution which is exactly homogeneous. Thus we are forced to consider only cosmologies in which the spatial extent is finite. Provided this finite spatial extent is sufficiently large, it will not yet be observable, however this does alter our description of time. The evolution of some finite region may explicitly depend only on the initial data in some finite region of the initial slice, however for it to be defined at all, initial data on the entire initial slice must be specified. The evolution of our falling apple may not depend on the state of the sun at the time the apple drops, *provided* the sun's contribution to the Hamiltonian is finite.

It is important to realize that although the Hamiltonian approach describes the classical evolution of the system, giving the same evolution described by Einstein's equations, it contains additional regularity conditions coming from the underlying quantum theory. These regularity conditions are non-local and we can no longer fully describe the evolution of a system independently of the evolution of other causally disconnected systems. It is not that the evolution of two causally disconnected systems are dependent on each other, however to be able to evolve one we must know that the evolution of the other exists and meets the regularity condition. Put another way, the existence of (this practical definition of) time for any system requires that the same definition exist for all systems. This additional requirement arises from the quantum theory, which then leaves its imprint in the classical theory via the Hamiltonian description. An exactly similar conclusion arises from the action formulation of the classical theory. In the language of quantum field theory, this phenomenon comes from the fact that the action contains additional 'off-shell' information, not present in the original classical equations of motion.

Thus we see that although the classical equations of motion allow one to define time via the evolution of a system, independently of causally disconnected regions, indeed independently of whether such regions exist at all or if they do, independently of whether time can be defined for systems inhabiting these regions, this is no longer the case once the quantum theory is considered. In the particular case of cosmology, if we wish to consider solutions that are (spatially) homogeneous, we are restricted to universes with a finite spatial extent and for inhomogeneous solutions the local existence of this practical notion of time i.e the evolution of our observable universe, implies that there exists a similar practical notion of time for all, causally disconnected regions on our spatial slice. That is, the existence of initial data sufficient to evolve into our universe implies that initial data exists everywhere on the initial spatial slice and most importantly that this slice is finite (or the data on it has compact support) and the initial data, when integrated over the entire slice, is also finite.

5 Consequences of quantum theory: renormalization

We have seen that there is a surprising relationship between the total spatial extent of our universe and the existence of a practical definition of time, that is now required to be globally defined. This is apparent already when the classical theory is written in terms of the Hamiltonian approach, however it arises from the quantum theory, and the requirement that evolution be unitairily implementable. The conclusion we have reached is that the total spatial extent must be finite, even if we wish to focus only on finite evolution of finite spatial regions. Previously the possibility of an infinite spatial extent was not relevant for finite evolution, however now a divergence appears. Perhaps we should not be surprised that, in the transition from classical to quantum theory, previously finite expressions become divergent. This occurs also in the subtly of defining products of operators at coincident (spatial) points, which also lead to divergences in classically finite expressions. Since this is an issue to do the small separation scales they are usually referred to as UV divergences. These need to be regulated in a suitable manner, consistent with the basic axioms of quantum theory, through a process known as renormalization. The second source of divergence that we have discussed, due to an infinite spatial extent is usually referred to as IR divergences. Perhaps this too can be 'renormalized' in some suitable manner? Or perhaps the two divergences are related in some as yet unexpected way? The physical significant of our inability to define observables in a spatially infinte universe may well point to a failing of the renormalization scheme. It may be that the correct approach to renormalization, presumably arising from some full theory of quantum gravity, will show us how to regulate not only the usual UV divergences, but also the IR divergences that appear in cosmology. If this were indeed the case, it may allow the local notion of time that can be defined for classical systems to be harmonized with the global requirement that currently result from quantum theory.

In the following I will describe how once again, the spatial extent of the universe can enter the game when (one particular version of) the usual renormalization procedure is carried out in cosmology. Again this leads us to requiring that the total spatial extent of the universe be finite if physical quantities such as (expectation values of) energy density are to be well defined. At first sight this will suggest that we should alter our definition of renormalization, and indeed it may well be possible to do so in such a way as to ensure that all (finite) physical observables and their (finite) evolution are well defined, even in spatially infinite cosmologies. However, the point of the entire procedure of renormalization is that there is a universal method for removing UV divergences. This universality can be traced to the basic fact that all Riemannian manifolds are locally approximated by Euclidean space and hence that, in some loose sense, all UV divergences are the same. This universality is *not* present in considerations of IR divergences and without something similar it would be difficult to justify any particular approach.

5.1 Adiabatic renormalization

If one considers an almost homogeneous cosmology, then it is very natural to decompose the system into a homogeneous part (the spatial average) and the inhomogeneous part, which are considered small perturbations. This matches extremely well with all observations of the early universe and forms the basis of our modern understanding of the early universe. The next step is to consider these small perturbations to be a quantum field, now evolving on the homogeneous background, whose evolution provides us with a 'time' parameter with which to evolve the perturbations. This step is a particularly thorny one and it leads to many difficult conceptual issues that collectively go under the name of the 'quantum to classical transition', which will not be discussed here. Once this step has been taken, our cosmological system contains the classical evolution of a homogeneous universe and a quantum field (the perturbations) evolving on the dynamical background. As with any quantum field theory, the definition of many important quantities require renormalization and in particular the definition of the energy density of the quantum field is formally divergent without this procedure. Fortunately there is a well defined method of performing this renormalization, even in for a system with a dynamical background, and one can calculate the (expectation value of the) energy density in the perturbations. For self-consistency of the basic approximations one needs to ensure that this renormalized energy density is indeed a perturbation to the homogeneous background.

One explicit method for performing this renormalization of products of operators, known as 'adiabatic renormalization' is the following. Perform a Fourier transformation on the perturbation field and calculate the formal expression for the desired operator (in this case the energy density). This formal expression contains an integral (or sum) over the values of the Fourier modes, k, which diverges. However, the integrand (or each term in the sum) is finite and well defined for every k, it is only the combination of all these finite contributions that diverges. Adiabatic renormalization then provides an explicit factor to subtract from the integrand (or from each term in the sum) *for all* k, that ensures that the resulting integral is well defined. In this scheme, the large values of k correspond to small spatial separations, and it is in this limit ($k \to \infty$) that the subtraction factor tends to a universal form. Indeed, the entire procedure for calculating this subtraction factor is based on the idea that, for large enough k (i.e. small enough scales), the spatial slice must approximate a flat slice and hence, on these scales, the quantum field should approximate the results of the well defined, flat space, quantum theory.

However, in a cosmology with an infinite spatial extent a new difficulty arises. This adiabatic renormalization scheme is designed to ensure that the large k (small scale) limit of the quantum field theory is well defined, however it does not guarantee that the explicit subtraction term is finite away from this limit. It may happen that the factor that needs to be subtracted from integrand at some low value of k (i.e. on some large scales) itself diverges, in which case the renormalization scheme may successfully remove the UV divergences of the theory, only to introduce new types of IR divergences. But should this be a concern? We have already encountered IR divergences in cosmology and have thus had to restrict our attention to cosmologies with finite spatial extent, is this more of the same? Unfortunately not. One can remove these new forms of IR divergences by considering a spatial slice that has a finite spatial extent, at the technical price of replacing the continuous Fourier transform with a discrete Fourier series, however this spatial extent cannot be arbitrarily large. If the renormalization procedure is to be well defined, not only must the spatial slice be finite, but it must have a spatial extent *smaller* than some specific value (given by the background dynamics).

But should this new IR divergence be considered physical? It arose from our attempt to remove UV divergences in some universal manner, but this universality applies only to the UV limit of the theory. If we choose to alter the subtraction terms for low k scales, whilst ensuring that the large

k limit remains intact, we will not have spoiled the universality property of the renormalization scheme. Indeed it is possible to perform an alteration in such a way that the IR divergences are removed, thus ensuring that the expectation value of all (relevant) renormalized operators are well defined. The difficultly is that there are many ways of doing so and unlike the removal of the UV divergences, there is no universality argument that can be appealed to choose one (class). One needs to use additional physical or conceptual inputs to choose the particular, 'natural', scheme.

6 Summary

As it has been presented here there appears to be two sources of IR divergences in cosmology, one coming from the homogeneous sector and an infinite spatial extend, the other from the renormalization scheme of the inhomogeneous perturbations, however the distinction is perhaps one only of mathematical convenience. Without splitting the cosmological model into an homogeneous background and inhomogeneous perturbations, it is difficult to formulate the quantum system, however physically it is perhaps not illuminating. In particular when one is forced to consider spatial regions larger than the currently observable universe. What an observer considers 'homogeneous' is tied to their observable universe. After all, we cannot say whether the universe is homogeneous or not on scales larger than our currently observable universe. As we saw in Section (4) quantum theory would seem to require us to restrict to a finite spatial slice if we wish to consider an homogeneous universe, whilst in Section (5) we saw that if we consider small perturbations over this background, then either one of the following must hold.

- Not only must the spatial extent be finite, it must be smaller than some specific scale or,
- Renormalization must be altered according to some, as yet unknown, new principles.

One possibility is that we may try to avoid this artificial decomposition into homogeneous and inhomogeneous sectors, defining renormalization in such a way as to ensure that (for example) the energy density in the zero-mode (the purely homogeneous sector) vanishes in some suitable manner. This may cure both forms of IR divergences (whilst not effecting the removal of the UV divergences of the theory) and *if* the resulting cosmological predictions were consistent with observations it would be a viable model. However, such an approach would impose our prejudice on the form of the universe on large (unobservable) scales. In such a case we would have succeeded in rendering quantum theory in an infinite universe well defined, however we would still have the fundamental discrepancy between the purely classical, local notion of time or evolution and the quantum requirement that this notion be applicable over the entire spatial slice. There is an essential non-locality built into quantum theory, which manifests itself at the classical level not by any change in causality, but rather by a requirement that our notion of time be globally definable.

These IR issues arise from the implementation of quantum theory to cosmology. In all other quantum systems, we do not need to consider the consequences of an infinite spatial extent, however in cosmology we must face this head on. There two possible resolutions to the discrepancy between classical cosmology, in which infinite and finite spatial extents are indistinguishable and the application of quantum theory, in which a difference appears:

- The classical locality is incorrect. We do need to constrain data outside our observable universe to be able to define evolution and hence relational time everywhere on a spatial slice, even for finite evolution of a finite region. In additional to the regularity constraints of the data, this requires either that the universe have a finite spatial extent or that the data defining the basics variables have compact support in a universe with infinite spatial extent (in particular it cannot be homogeneous).
- Quantum theory needs to be reformulated in a more 'local' way, so as to ensure that, when applied to cosmology, our observable universe is not dependent on unobservable data. However any reformulation must agree with the wealth of experimental data that quantum theory has provided on laboratory scales, whilst simultaneously allowing for independent definitions of evolution for causally disconnected regions *and* the compatibility between these evolutions if these regions overlap.