The fundamental questions for tomorrow

Samir D Mathur

Department of Physics
The Ohio State University
Columbus, OH 43210, USA
mathur.16@osu.edu

1 Introduction

For many centuries the goal of science has been to move towards more and more fundamental laws that will explain all of nature. In this essay, I will argue that at the present juncture we need a sharp change in the nature of the questions we have been answering.

So far almost all of physics has been focused on the question: if we start with a given set of initial data, then how will this data evolve in the future? But we will at some point run out of questions of this kind; for example we might find the complete extension of the standard model, and unify it with gravity, perhaps using string theory. Will this be the end of physics?

I would say no; this will just set the stage for a more fundamental set of questions, some of which we can begin to ask and answer now. These more basic questions are of the following type. What determines the initial data for the evolution? What principle chooses the laws of dynamical evolution to be the one that we found?

It is true that we have very few tools to address such questions today. But there *are* some places that we can focus on to delve deeper into these directions. The goal of this essay is to identify such places. In particular we will see that the early Universe must have some kind of an instability, which we must try to understand in more detail. We will also note that there has been significant progress in understanding the quantum dynamics of black holes, and again the results here can be used to say something about the Universe and also about the nature of physical laws in general.

2 The local nature of today's laws

In the times of Aristotle, nature was described by a large set of rules. For example, it was argued that a rolling stone came to rest because its natural state was one of rest; similarly, birds fly up since they belong to the sky.

Today we would not consider such ideas as fundamental. Galileo noted that moving bodies tend to continue in their state of motion unless impeded. Newton developed a complete theory for moving bodies: the acceleration of a body was proportional to the force acting on the body.

Interestingly, the general pattern of Newton's law of motion has changed very little over the intervening centuries. Newton's law is a 'second order differential equation', which means that one has to specify two numbers for each particle – the initial position and the initial velocity – for the laws to then determine all future evolution of the particle. Special relativity did not change this: we still need two pieces on initial data to determine the evolution. The theory of fields – relevant to waves like radio waves – also starts with a second order equation; thus we need to know the initial fields and their rate of increase, and then the wave equation will yield the subsequent evolution. General relativity changed many fundamental beliefs about the nature of spacetime, but Einstein's equations of general relativity are still second order differential equations.

Quantum theory brought about a change which we may call a simplification: instead of two pieces of initial data, we need just one. Thus when describing a particle, we need to give the wavefunction at an initial time; the Schrodinger equation then determines all future evolution of the wavefunction. The initial wavefunction gave the probability amplitude for the particle to be at different positions, but by the principle of complementarity, we can use this wavefunction to also determine the probability amplitude for different possible *momenta* of the particle. Since the particle's position and momentum cannot be specified at the same time, the wavefunction can be cast in terms of one of these variables rather than both. This basic structure continues to quantum field theory and to string theory: in principle we just need to specify an initial 'wavefunctional', and then the evolution equations will determine the physics for all times in the future.

To summarize, it appears from what we have observed in experiments to date that nature is very *local*; if we know the state at a given time, then there are definite laws to tell us what happens at the next instant of time. This nice state of affairs did not have to be true. For example, if the laws of nature were based on nth order differential equations, then we would need n pieces on initial data before the further evolution could be determined. More generally, one could imagine a theory where one must know the state for an entire segment of time from t_1 to t_2 before the evolution for times $t > t_2$ could be determined. We do not know why nature has been kind and made physical evolution simple, but we do feel quite sure that all basic laws of physics will be such that initial data on one time slice will suffice to determine the future.

Today the idea of finding the 'theory of everything' is focused on finding laws of the type discussed in the above paragraphs. We ask ourselves if the standard model of particle physics is indeed described by the gauge groups incorporating the strong, weak and electromagnetic forces, or if string theory is indeed the correct description of quantum gravity. This search will, presumably, come to an end at some point, and we will indeed have the correct equations for describing evolution on the kind discussed above.

It is at this point that we will have to turn to the new fundamental questions mentioned in the introduction. We will have to ask what determines the initial state, and why the laws of evolution need just the initial state at one time.

3 The initial state

What determines the initial state of a system?

This question is most naturally posed in Cosmology. After all, quantum theory requires that the initial wavefunction be given on a complete slice through all of space. In practice we can ignore parts of this slice that are far away, but while we are asking questions at a fundamental level, we might as well be rigorous, and use the entire slice as the place where we must give our initial data. So let us ask: what determines the wavefunction of the Universe?

At first it may seem that we are asking a question that is outside the purview of physics; perhaps it is a question for the field of metaphysics. But such a view would be too narrow, and would be based on the assumption that finding the laws of evolution from the initial state is the only goal of physics. As we will now note, there have been several attempts to say something about the choice of initial state.

3.1 The no-boundary proposal

An intriguing proposal was made by Hartle and Hawking [1]. They argued that there *is* a preferred initial state for the whole Universe, one which starts with the idea that was *no* initial time slice at all. One first does an analytic continuation of time to imaginary values, so that it becomes a dimension on the same footing as the other directions of space. In this Euclidean space we can have a smooth hemispherical cap that can be used to 'cap-off' the Universe at early times. Quantum theory asks that we sum over all possible fluctuations of a system, and here one can sum over all the shapes that the cap can take. This determines a unique starting wavefunctional for the Universe, called the 'no-boundary wavefunctional'.

While this idea is pleasing, it cannot be complete in itself as an explanation for what we see around us. The natural length scale for quantum gravity is the planck length, which is miniscule. One might expect that an argument as abstract as the no-boundary proposal will generate a Universe that will have the size of planck length and last a time of order the planck time. Further, if the wavefunction of the Universe is unique, then should we say that everything that we see around us *had* to exactly how it is now? That is, if we see a certain pattern of stars in the sky, then is this pattern the unique one that had to evolve from the Hartle-Hawking wavefunction?

Such a conclusion looks strange, since the world around us exhibits structures that are are too rich and varied to be classified as 'unique'. Clearly, we need additional principles to explain the evolution of the Universe from a given simple initial wavefunction. We will now note two of these principles: the idea of *decoherence* and the physics of *inflation*.

3.2 Decoherence

The natural evolution of quantum mechanics leads to a proliferation of 'branches' of the wavefunction. Suppose a particle has an amplitude A_1 to be on the left of $x = 0$, and an amplitude A_2 to be on the right. Suppose further that a second particle passes by in the region $x > 0$. Then there is an amplitude A_1 that the particles do not interact (since they were not near each other) and an amplitude proportional to A_2 that they do (assuming the particles can interact when they are in the same region). Thus the overall wavefunction splits into two branches: one where no interaction has occurred, and one where they have scattered off each other. Clearly, as time progresses, the wavefunction of the Universe will split into an enormous number of branches. Does this lead to a very confused state for what we should see in the world today?

The answer is no, and the reason is decoherence. While the overall wavefunction is very complicated, the phase factor in different parts of this wavefunction has a very different rate of evolution. Most parts of the wavefunction become unobservable due to rapid phase cancellation. Configurations where the phases oscillate slowly define a decohered branch, and such a decohered branch is observable. But many decohered branches arise from the same overall wavefunction. Each of these branches gives rise to a configuration that can be close to classical; for example in a given branch the particles defining the earth may be here, and the particles defining the moon may be 200,000 miles away. In another branch all these particles may be close to each other, so that the moon is not a separate body from the earth. But these two different branches will not interfere with each other due to their mutually different phase coherences, and so we would see only one reality: in our case, that the moon is 200,000 miles away.

One importance of the idea of decoherence is that it allows classical physics to emerge from what must be a very quantum overall wavefunction. But there is a second output: we can explain how complicated patterns can emerge in the world without violating the uniqueness of the initial wavefunction. This initial wavefunction will likely have to be very symmetric, since we are trying to argue that it is unique. But this unique wavefunction can give rise to many branches. Each branch can break the symmetry in the sky, as long as the overall *set* of branches does not break any symmetry. Since we can observe only one of these branches by decoherence, we will see a non-unique looking pattern in the sky even if some theory has been able to predict a unique initial state for the Universe.

As we will now see, however, decoherence is not enough; we need in addition an an inherent instability in our system to split the wavefunction into a large number of branches.

3.3 Instability

The Universe presumably has its origins in a phase dominated by quantum gravity. Since the length scale of quantum gravity is planck length, 10^{-33} cm, how did we get a Universe that is 10^{28} cm across today? Further, the natural mass scale of quantum gravity is planck mass, 10^{-5} gm. How do we explain the vastly larger mass of the Universe today?

The key fact to explaining the latter discrepancy is the attractive nature of gravity. This attractive nature is manifested in the negative value of the gravitational potential. Thus we can take a large ball of dust with mass M and radius R. This mass corresponds to an intrinsic (i.e. non-gravitational) energy

$$E = mc^2$$

Using Newtons law of gravitation for the purposes of an estimate, we find the gravitational potential energy

$$V = -\frac{GM^2}{R}$$

By making R smaller, we can make the negative potential energy large enough so that the total energy

$$E_{total} \sim Mc^2 - \frac{GM^2}{R} \sim 0 \tag{1}$$

Thus we can start with no energy, and transition to a ball of large rest mass M without violating the conservation of energy. (The above argument can be made rigorous using general relativity.)

While this discussion shows that the vast Universe we see today can arise from an initial state with virtually no mass, it does not give a mechanism for such a transition to occur. In fact in the standard model of the big bang containing ordinary matter, the intrinsic energy of the matter decreases or at best stays constant. Thus in a radiation filled Universe the energy of each photon decreases as the Universe expands, while in a dust filled Cosmology the mass of the dust grains stays constant. In all such examples, if we wanted a Universe today with a large intrinsic energy of its matter, then we had to start with the same or more energy at the time of the big bang.

Remarkably, there is a kind of matter that *does* lead to an increase in the intrinsic energy – a scalar field trapped in a potential well that has a positive energy minimum. Such a configuration has effectively negative pressure. If matter has positive pressure then it does work as it expands, and its intrinsic energy decreases. But with negative pressure, the intrinsic energy increases as the Universe expands. If at some later time we convert the energy in the scalar field to the energy of ordinary matter, then we can end up with an intrinsic energy of this matter than it vastly in excess of planck energy.

While this idea of inflation [2] has some difficulties with issues of fine-tuning, the observation of baryon acoustic oscillations in the sky confirm that the essential features of inflation must be reproduced by some mechanism or other. One might therefore try to look for inflating configurations in string theory.

Surprisingly, it has not been easy to find inflation in string theory. Given that inflation is expected to be the natural mechanism for generating the energy of the Universe, one might expect that inflating configurations would be natural and easy to find in the theory.

But most configurations of strings behave like normal matter: strings behave as just like any other collection of particles. To get a configuration that inflates [3] one must prevent the compact extra dimensions from collapsing as the visible dimensions expand. But to get this stability of extra dimensions one must use nonperturbative effects in the theory, which are not easy to justify in a rigorous way with our present understanding of the theory. Further, it has recently been argued that the configurations which have been constructed are actually unstable, and quickly evolve to configurations that do not inflate further [4].

Given these difficulties, one might hope that there is an alternative and natural way to get something equivalent to inflation in string theory. We will now argue that an emerging theory – the fuzzball paradigm – might indeed have just such a mechanism.

4 The fuzzball paradigm

A fundamental problem in developing a quantum theory of gravity is the obstruction posed by the black hole information paradox. It has recently been proved, using inequalities from quantum information theory, that this paradox cannot be avoided as long as the black hole has a horizon to leading order in the semiclassical theory [6]. How then will string theory give a consistent theory of quantum gravity?

It turns out that in string theory we *cannot* make a horizon. Consider a shell of mass M that is collapsing to smaller values of its radius. As the radius of the shell approaches its horizon radius

$$R_h = \frac{2GM}{c^2} \tag{2}$$

the shell transitions to horizon sized stringy objects called fuzzballs [5]. The semiclassical expectation of horizon formation is violated due to the enormous number of fuzzballs: there is one fuzzball for each state of the black hole , so there are $Exp[S_{bek}]$ fuzzballs, where S_{bek} is the Bekenstein entropy. While the amplitude for the infalling shell to tunnel into any one fuzzball is small, this small number is cancelled by the large degeneracy factor to yield an overall probability for transition that is order unity [7].

We have noted that the transition to fuzzballs happens when the shell reaches close to its horizon radius (2). But note that this is also the length scale that we find from the requirement that we get intrinsic energy out of gravity in Cosmology; from equation (1) we find

$$R \sim \frac{GM}{c^2} \tag{3}$$

which is the same scale as (2). Thus the entropy of fuzzballs should be an important ingredient is the dynamics of the early Universe! Let us see what effect we can expect from fuzzball dynamics in Cosmology.

In the black hole problem, one finds that one must study dynamics on the full space H_F which describes all the possible configurations of fuzzballs. The wavefunction of the infalling shell has a small 'under the barrier' leakage into the space of fuzzball configurations

H_F. This leakage becomes larger and larger as the shell approaches its horizon radius R_h, and when the shell reaches radius R_h then the wavefunction spreads over the space H_F and the infalling shell ceases to exist as a semiclassical configuration.

The aspect of this dynamics that is relevant for Cosmology is the following. The leakage of the wavefunction describing the actual matter (the shell) into the configuration space of fuzzballs H_F causes a reduction in the energy of the wavefunction; this reduction of energy is just the lowering of energy that comes from a reduced kinetic term when the wavefunction is allowed to spread. This lowering of energy can be directly identified with the gravitational redshift, and the diverging of this redshift at the horizon signals the fact that the wavefunction has spread completely over H_F. If we limit the size of the Universe to a ball of radius R_{univ}, then we cannot have fuzzballs of radius greater than R_{univ}. Thus for smaller R_{univ} there are less directions in H_F for the wavefunctions to spread in and thus a smaller lowering of the energy. This effect gives rise to a pressure forcing R_{univ} to larger values. This force is not part of semiclassical dynamics, just as the transition to fuzzballs is not seen in semiclassical black hole collapse. Thus we have a new force arising directly from the black hole states present in string theory, and this force can drive an expansion that may reproduce the observations of inflation.[1]

5 Iterative theories

Finally we come to the question: what should determine the dynamics of the theory; i.e., the rules which tell us how the initial state evolves with time? Is there any principle which selects some rule over all the others that are possible? Here we will find that fuzzball dynamics might incorporate a new principle: the theory should be *iterative* in its quantum structure.

To explain what this means, we first look at AdS/CFT duality as a toy example of the principle. Usually one thinks of this duality as a map between a gravity theory – string theory in AdS – and a field theory – Yang Mills theory. But we will take a different perspective on this duality to illustrate the point we wish to make.

We start with a collection of N D3 branes, which are excitations of our gravity theory – string theory. For large N, we have a large virtual cloud of open strings between these D3 branes. An infalling graviton hits these open strings and converts its energy to collective vibrations of this cloud. These collective vibrations live on the vast space of open string configurations between N^2 D-branes. But the low energy dynamics of these vibrations has a simple description: *it can be mapped to string theory on gently curved AdS spacetime.*

Thus we can study the same system in two ways: (i) We can study the dynamics of the complicated set of D-branes and the cloud of open strings between them, or (ii) replace all

[1] This force from virtual fuzzballs is different from the entropic force conjectured in [8]. The entropic force arises for non-Hamiltonian systems, while here we have just the normal pressure from the spreading of quantum wavefunction over the space of states H_F.

these things by AdS space on which we have just a few moving objects representing the excitations of the cloud.

Note that this effective AdS spacetime emerging from the collective vibrations allows all the excitations of the original string theory: gravitons, strings branes etc. Thus the dynamics of a complicated set of degrees of freedom of string theory – the cloud of open strings – has been mapped back to the dynamics of string theory around the local vacuum. We will call this behavior of the theory 'iterative'. A general theory will not have such a behavior: there will always be a cloud of virtual excitations around any massive object, but there need not be any simple rule governing the dynamics of this cloud.

AdS/CFT is a special case of the iterative principle; we conjecture that the general case of iterative behavior is expected to occur for black holes. We have noted that the wavefunction of a collapsing shell spreads over the vast space of fuzzballs H_F as the radius of the shell approaches R_h, the horizon radius for the shell. But what can we say about the evolution of the wavefunction on this space H_F?

The conjecture of 'fuzzball complementarity' says the following: *The evolution of sufficiently smooth wavefunctions on H_F can be mapped onto the gravitational dynamics of ordinary 3+1 dimensional spacetime.* Note that H_F is not 3-dimensional; it is a space with a very large dimension. But according to our conjecture, the low energy collective excitations on this space have the same approximate spectrum as the spectrum of gravitational excitations around the *vacuum*.

The conjecture of fuzzball complementarity allows us to recover, in an approximate way, the classical dynamics of the black hole interior. We have already seen that the infalling shell transitions to fuzzballs with no horizon; this resolves the information paradox while removes the classically expected black hole interior. But the wavefunction continues to evolve on the space H_F, and this evolution can be approximately mapped to a continued infall into the classical black hole. This is the sense in which the black hole has an effective 'interior'. Such an interior can be recovered only if the dynamics of fuzzballs is iterative; i.e., the collective excitations on the space of fuzzballs H_F behave just like the low energy excitations of string theory in locally empty spacetime.

In short, an iterative theory is one where the collective modes of the cloud of virtual excitations are described by the original Lagrangian of the theory. It appears that string theory is iterative, and it may be that requiring iterative behavior uniquely fixes the dynamics that we are allowed.

6 Summary

The last few centuries have seen a dramatic progress in our understanding of how a system evolves once we are given its initial state. But we are now at the threshold where we will have to confront much more fundamental questions. What, if anything determines the initial state? Does this initial state evolve from a principle like the no-boundary proposal of quantum Cosmology, and then explode by a process of inflation? If so, can this inflation be caused by the novel effects recently discovered in black hole physics? Can the apparent randomness of the sky be explained by decoherence? Could it be that the requirement of iterative physics determines the ultimate laws of evolution?

These questions may have seemed metaphysical at one point, but I believe they are the fundamental questions for tomorrow, and we are at the point where we can begin to make some progress towards them.

References

[1] J. B. Hartle and S. W. Hawking, Phys. Rev. D **28**, 2960 (1983). doi:10.1103/PhysRevD.28.2960

[2] A. H. Guth, Phys. Rev. D **23**, 347 (1981). doi:10.1103/PhysRevD.23.347

[3] S. Kachru, R. Kallosh, A. D. Linde and S. P. Trivedi, Phys. Rev. D **68**, 046005 (2003) doi:10.1103/PhysRevD.68.046005 [hep-th/0301240].

[4] I. Bena and G. Pasini, JHEP **1604**, 181 (2016) doi:10.1007/JHEP04(2016)181 [arXiv:1511.01895 [hep-th]].

[5] S. D. Mathur, Fortsch. Phys. **53**, 793 (2005) doi:10.1002/prop.200410203 [hep-th/0502050];

[6] S. D. Mathur, Class. Quant. Grav. **26**, 224001 (2009) [arXiv:0909.1038 [hep-th]].

[7] S. D. Mathur, arXiv:0805.3716 [hep-th]; S. D. Mathur, Int. J. Mod. Phys. D **18**, 2215 (2009) [arXiv:0905.4483 [hep-th]]; P. Kraus and S. D. Mathur, Int. J. Mod. Phys. D **24**, no. 12, 1543003 (2015) doi:10.1142/S0218271815430038 [arXiv:1505.05078 [hep-th]]; I. Bena, D. R. Mayerson, A. Puhm and B. Vercnocke, JHEP **1607**, 031 (2016) doi:10.1007/JHEP07(2016)031 [arXiv:1512.05376 [hep-th]].

[8] E. P. Verlinde, JHEP **1104**, 029 (2011) doi:10.1007/JHEP04(2011)029 [arXiv:1001.0785 [hep-th]].