Matthias Bartelmann

Cosmology – The Largest Possible Model?

1 Laws of Nature and the Foundations of Cosmology

Upon which foundation should one build a model for the Universe as a whole? The idea that such a model should exist seems bold itself. Can we really believe that we might be able to construct a physical model for a unique object that we cannot experiment with, that we are part of and of which we can only see a very small section? The goal of this article is to explain that this does indeed seem possible, that mathematical simplicity is used as a guiding principle in this construction, and that the resulting world model is remarkably consistent with a wealth of observations.

Let us begin with a detour through the foundation of laws of nature in physics. It is important to realise that laws of nature do not describe nature herself, but human concepts of nature. Otherwise it would not be possible to replace established laws by other, more general ones, as it has happened several times in the history of physics. Theories in physics are based on axioms chosen by physicists, and these axioms can be altered.

Newton's axioms underlie classical mechanics. They distinguish four entities; bodies, forces, space and time, and formulate how bodies move in time through space under the influence of forces. Field theory, initiated by Faraday, attaches forces to space and gives force fields their own dynamics. Special Relativity realises that space and time have no independent existence and thereby connects forces, space and time. General Relativity explains how the presence of bodies and energy affects the structure of space-time. Thus, in a general-relativistic field theory, the four initially separate entities of Newtonian physics are all linked together.

The dynamics of physical entities, i.e. their change in time, is described by differential equations. They themselves are not postulated, but derived from a more general concept overarching physics, namely that of an extremal principle. The best-known example is perhaps Fermat's principle, which states that light rays connect the source and the receiver in such a way that the light travel time is extremal along them. The extremal principle underlying essentially all of the established theories in physics is the so-called principle of least action, or Hamilton's principle. The action itself is an abstract quantity that can be constructed under very general rules. It must be independent of any observer's state of motion, and it is typically chosen to be invariant under certain symmetry operations. Both criteria are expressed by the mathematical concept of symmetry groups.

Symmetry groups and extremal principles currently form the deepest foundation of physical laws. Which symmetry operations a physical theory should obey exactly is largely the physicists' choice. Ultimately, however, no theory is acceptable that is in demonstrable conflict with experiments.

We know four fundamental interactions. The strong force keeps the fundamental building blocks of matter bound, the so-called hadrons. The weak force is responsible for certain conversions of particles into others, in particular through the so-called beta decay. Both act only on subatomic distances. Electromagnetism keeps atoms and molecules bound and is responsible for all interactions between charges and between matter and light. Gravity is by a large margin the weakest of the four interactions. Both electromagnetism and gravity have unlimited range and thus determine physical interactions in the macroscopic world. Since the sources of electromagnetism are positive and negative charges, the effects of one type of charge can be shielded or compensated by the other type. Effectively, therefore, the range of the electromagnetic interaction is typically also limited. Shielding is impossible with gravity, as it knows only one type of charge, i.e. the mass. Electromagnetism as well as the strong and the weak interactions are described by a unified quantum field theory called the Standard Model of Particle Physics. Gravity has so far withstood all attempts to cast it into the form of a quantum field theory as well.

If we now return to cosmology, we realise that any cosmological model must essentially be derived from a theory of gravity as the only long-range force that cannot be shielded. The most advanced theory of gravity is Albert Einstein's theory of General Relativity. We must thus begin with General Relativity in our construction of modern physical world models.

General Relativity can be seen as a prototypical example for a physical theory constructed as outlined above. It is based on the fundamental concept that the geometry of space-time, characterised by its metric, is a dynamical field determined by the presence of matter and energy. The dynamical equations of General Relativity, called Einstein's field equations, follow from the principle of least action, with an action that combines the geometry of space-time with the presence of matter in what seems to be the most straightforward and simple way.

Einstein's field equations form a set of ten independent, non-linear, partial differential equations that cannot be solved once and for all with a general scheme. Special solutions can be constructed once certain simplifying assumptions are being imposed. Typically, these come in the form of symmetry assumptions. We thus encounter symmetry considerations a second time at a more specific level. At a fundamental level, they were used to construct plausible physical theories themselves, for which General Relativity is one example. Now, symmetry assumptions are used a second time to identify classes of solutions of the dynamical equations of this theory. To construct a simple class of cosmological solutions of General Relativity, Friedman first assumed that they should be spatially isotropic and homogeneous. This means that arbitrary spatial rotations and translations should leave the observable universe unchanged for any observer. The solutions so obtained form a class characterised by certain parameters that describe the matter and energy content of the Universe and its expansion behaviour at one point in time. Once these parameters are known or set, the world model is fixed.

The only justification Friedman gave for the symmetry assumptions was mathematical simplicity. Do these assumptions and the world models constructed upon them correspond to reality in any way? As we saw, isotropy demands that the Universe, as seen by any observer, should exhibit the same physical properties in all directions. At first sight, this seems to be manifestly incorrect: The night sky does not at all appear independent of direction. However, if the properties observed in the Universe are averaged over sufficiently large scales, they do in fact approach isotropy. The most striking example for this statement is the cosmic microwave background (CMB), which will be introduced and discussed further below.

If we observe at least approximate isotropy, so should any other observer in the Universe. Since the Copernican revolution, we have grown used to the notion that our location in space and time is by no means unique or central. If, however, the Universe is isotropic about all of its points, as this concept suggests, then it must also be homogeneous.

With these considerations, we have come a long way already. We have seen that theories in physics are constructed upon very general concepts, expressed by symmetries and extremal principles, from which the dynamical differential equations follow. Of the modern theories of physics, only General Relativity is relevant for the construction of world models. In order to find appropriate solutions of its field equations, Friedman introduced the further symmetry assumptions of spatial homogeneity and isotropy, with the sole justification of mathematical simplicity. When combined with these symmetry assumptions, the field equations of General Relativity reduce to the class of Friedman world models. Two questions then arise: First, does our Universe exhibit at least the qualitative features of the Friedman models? Second, if so, is there a unique combination of the cosmological parameters appearing in these models that identifies a single Friedman model out of this class?

2 Empirical Evidence for the Standard Cosmology

Let us now investigate the class of Friedman solutions and the two questions raised above. These questions are asking whether the line of reasoning leading from the foundation of physical theories to the construction of physical world models finds its expression in nature. Two aspects of this procedure cannot be overemphasised: First, we have used symmetry assumptions and thus essentially mathematical concepts of regularity and simplicity as guiding principles. The Friedman models are a particularly simple class of solutions of Einstein's field equations. Why should they have any resemblance with the real world we find ourselves in?

The astounding result of decades of cosmological research, as shall be outlined now, is that it is indeed possible to draw a consistent, quantitative picture of our actual universe and its evolution within the class of Friedman models.

Second, while all other areas of physics can conduct experiments with their objects, cosmology cannot. It should be kept in mind that all statements concerning our universe as a whole are based on the comparatively tiny portion of it from which we can receive information. We thus vastly expand the realm of physical laws from our laboratories to the entire observable universe, and we extrapolate from the observable universe to the universe as a whole. The fact that the empirical evidence collected in cosmology does indeed seem to converge with the theoretical concepts underlying the class of Friedman models has a breath-taking aspect.

We shall now go through the most pronounced and relevant empirical pieces of evidence.

1. Friedman models turn out to be generically unstable. They must either contract or expand unless their parameters are very finely tuned. This means that any two points in space identified at a fixed time must either move towards or away from each other not because they move in space, but because space itself drags them along. In expanding models, every observer should see galaxies in his neighbourhood move away from himself with a velocity linearly increasing with distance. The most obvious question to begin with is thus whether our universe is in fact changing with time and whether the galaxies surrounding us do in fact move away from us in the linear fashion that the Friedman models predict.

This is a simple question in principle, but quite hard to address in practice. The problem is that our universe is not ideally homogeneous, as our existence demonstrates. The matter density is not constant, but fluctuates locally. Regions of higher matter density attract neighbouring galaxies and imprint a local motion on them that is superposed on any motion of cosmic origin. Since any cosmic velocity increases with distance in the Friedman models, the cosmic motion can be expected to dominate the local peculiar motion only beyond a certain distance. Thus, distant galaxies must be precisely observed and their distances measured, which is a demanding procedure. Slipher observed in the 1920s that galaxies typically move away from us, and Hubble found around 1930 that their velocities increase linearly with their distance, just as the Friedman models predict. If this is the correct interpretation of the mean motion of distant galaxies, we seem to be living in an expanding universe. The expansion rate, defined as the relative amount by which cosmological distances increase in time, is called the Hubble constant and is one of the fundamental parameters of any Friedman model.

2. The inverse of the Hubble constant sets the time scale for the cosmic expansion, which turns out to be on the order of 10 billion years. Is this time scale long enough to encompass the observable evolution of the universe, or are there any known objects whose age credibly exceeds the age of the universe? How old are the Earth, the Galaxy and the oldest objects we find in our observable universe?

The decay of suitably long-lived radioactive isotopes such as ^{235}U or ^{238}U provides the best constraints of the terrestric and the galactic ages. The Earth turns out to be 4.6 billion years old. The age of the Galaxy is less well constrained, but likely between 7 and 10 billion years. Older objects exist in the universe whose age we can determine. These are in particular certain end products of stellar evolution, the white dwarfs, and a certain class of co-eval stellar populations, the globular clusters. Upper limits on their age touch approximately 12 billion years. The fact that these age limits broadly agree with the time scale set by the inverse cosmic expansion rate is reassuring.

3. The observation that our universe is expanding today does not necessarily imply that it has been expanding during all of its past. Friedman models which are expanding today but were shrinking or stagnating for part of their history would also be possible. However, a few simple observations show that our universe cannot be of this type. The most intuitive of these is that objects exist whose spectra reveal that the universe was at least six or seven times smaller when their light was emitted than it is today. Thus, if our universe behaves like a Friedman model at all, its present expansion implies that it has always been expanding.

A monotonically expanding Universe keeps shrinking as we go back in time. Any two points then keep approaching each other until they come arbitrarily close after finite time. Any finite section of the Universe must have been very small at early times. Backward in time, the cosmic matter is compressed by the shrinking volume it is enclosed in. Thus, matter and all other ingredients of the Universe must have been hotter in the past than they are now. If it once was very small, the whole universe may have been as hot as the interior of stars is now. Then, nuclear fusion processes must have occurred throughout the Universe, leading to the formation of light elements such as deuterium, tritium or helium from hydrogen. In fact, interstellar gas contains about 25 % helium and 75 % hydrogen. This large amount of helium cannot have been fused by stars, but only if the entire Universe acted as a nuclear fusion reactor very early during its evolution.

It had been realised by Gamow and his collaborators already in the 1940s that the abundance of helium in the universe can be explained assuming that the Universe itself produced it in a hypothetical hot and dense, early phase. Effective fusion could have set in once the temperature of the Universe had dropped just below a billion degrees, and ended very quickly thereafter as the universe kept expanding and cooling. This happened when the Universe was between two and three minutes old.

4. Charged and sufficiently dense particles at temperatures so high produce energetic thermal radiation. Thus, if the universe was once indeed hot enough to fuse helium, the thermal radiation then produced must still be present, albeit cooled down considerably as the Universe expanded. In fact, it was possible already in the 1940s to predict from the observed helium abundance that the thermal leftover radiation should now have arrived at a temperature of a few degrees Kelvin. Thermal radiation with such a low temperature has characteristic wavelengths in the microwave regime. Thus, the existence of a so-called cosmic microwave background (CMB) could be predicted from the assumption of a hot beginning together with the observable amount of helium. An apparently isotropic, ambient radiation field with properties like the CMB was serendipitously discovered by Penzias and Wilson in 1965 while testing a telecommunication antenna. Immediately, Dicke and co-workers surmised that this radiation could indeed be thermal radiation left over from the very early Universe.

At that time, it was not possible to confirm that the radiation discovered was thermal radiation, as required by this interpretation. However, if one assumed that it was thermal, the temperature corresponding to the measured intensity was found to be approximately 3 Kelvin, in good agreement with the earlier prediction.

A satellite called the Cosmic Background Explorer, COBE in short, impressively demonstrated the thermal nature of the CMB. One of its three instruments measured the electromagnetic spectrum of the CMB and found it to be in perfect agreement with that of thermal radiation with a temperature of 2.7 K.

Another of COBE's instruments solved an acute problem that had accumulated since the discovery of the CMB. Since there are structures like galaxies, galaxy clusters and even larger objects in the Universe, the CMB is not supposed to be ideally homogeneous. Since the present cosmic structures should have originated from predecessors in the very early universe, those should have left their imprint on the CMB. It was estimated that temperature fluctuations with milli-Kelvin amplitudes around the mean temperature should be found. However, when detectors finally reached the required sensitivity, such fluctuations were not detected. Even at a level of one part in a thousand, the CMB was found to be perfectly isotropic.

Since isotropy is one of the primary symmetry assumptions underlying the Friedman models, the remarkable isotropy of the CMB was impressive evidence in their favour. The lack of temperature fluctuations at the level expected from the existing cosmic structures was highly disturbing at the same time. A solution was proposed by Peebles in the 1980s. If cosmic structures consisted not of ordinary matter as we know it, but of a form of matter that does not participate in the electromagnetic interaction, the present cosmic structures could be reconciled with considerably smaller temperature fluctuations in the CMB since then the imprint of the cosmic structures in formation on the CMB could be substantially lower. Fluctuations of one part in 100,000 would then be expected.

At that level, COBE finally found these fluctuations in 1992. This can be seen as a turning point for cosmology, and at the same time as a piece of evidence that cosmic structures are not dominated by the electromagnetically interacting forms of matter that we know, but by some dark matter of hitherto unknown composition.

5. The existence of dark matter was not surprising at that time. Rather, the important result was that dark matter cannot interact electromagnetically. Already in the 1930s, Zwicky had found that the member galaxies of the galaxy cluster in the constellation Coma moved so fast that much more matter was needed to keep them gravitationally bound than could be inferred from the amount of light emitted by the cluster and its galaxies. The amount of mass necessary for balancing the motion of the galaxies was approximately ten times higher than that necessary to produce the light observed. A similar observation was made later at the level of individual galaxies. Their stars also move considerably faster than they should if they moved under the influence of the gravity of their visible matter alone. Dark matter is thus seen on a hierarchy of levels in the Universe, but only the CMB requires it to be of a hitherto unknown form avoiding the electromagnetic interaction.
6. The physics of the CMB and its temperature fluctuations are simple and well understood. The CMB was set free when the universe had become cool

enough for hydrogen atoms to form from the cosmic plasma. It can easily be calculated that the temperature had to drop to approximately 3,000 K for this to happen. When this temperature was reached, the universe was just below 400,000 years old. The hydrogen plasma combined to form hydrogen gas within the relatively short time of about 40,000 years. Since then, the photons of the CMB could propagate almost freely throughout the universe.

Of course, it is not possible to predict the exact structure of the temperature fluctuations in the CMB since they depend on presumably random initial conditions whose exact realisation we cannot know in detail. However, predicting their statistical properties, in particular what the amplitudes of temperature fluctuations of a given size should be like, was possible as early as in 1970. This, however, depends on some of the most important cosmological parameters, such as the densities of ordinary and dark matter, the total matter and energy density in the universe, its expansion rate and the like. The statistical analysis of detailed and sensitive measurements of the CMB temperature fluctuations could thus reveal a good fraction of the cosmological parameters, once compared with theory.

It is an amazing fact on its own that precise measurements of CMB structures confirmed the theoretical predictions in detail and could in turn be used to accurately determine cosmological parameters. For this reason, measuring and interpreting CMB temperature fluctuations has developed into one of the main objectives of current cosmological research. After COBE, two further CMB satellites have been launched. The Wilkinson Microwave Anisotropy probe or WMAP has been observing between 2001 and 2010, while the Planck satellite began operations in 2009. The CMB data taken so far have greatly helped constraining the cosmological parameters with high precision. They have not revolutionised the cosmological model itself, but they were decisive for turning it into the cosmological standard model, whose parameters are now determined typically with relative uncertainties of 10 or less per cent.

7. So far, we have discussed only one piece of evidence probing the late universe, namely the cosmological expansion reflected by the systematic recession of the galaxies in our cosmic neighbourhood. In contrast, the fusion of helium and other light elements and the CMB both probe the early universe, albeit with a large separation in time. Helium fusion ended about three minutes after the beginning, while the CMB was released almost 400,000 years later.

We have touched an important argument that we need to accentuate further: It is possible to interpret these three types of observation in favour of the Friedman models. The recession of the galaxies agrees with the intrinsic instability of the Friedman models and exhibits the expected expansion behaviour. Moreover, it defines a time scale for the evolution of the universe which agrees reasonably well with the age determinations of old cosmic objects.

Its present expansion suggests that the universe originated in a hot and dense early state, which allowed the fusion of the large amounts of helium that are actually observed. This, in turn, gives rise to the prediction of left-over thermal radiation and thus of the CMB, whose temperature of a few degrees Kelvin is directly related to the amount of helium observed. The level of the temperature fluctuations in the CMB is a strong argument in favour of a form of dark matter that avoids the electromagnetic interaction. The statistics of the CMB temperature fluctuations depend on the details of the cosmic matter content in a precisely predictable way, enabling accurate constraints of cosmological parameters. The abundance of ordinary matter derived from the CMB agrees precisely with the abundance needed to understand the efficiency of the helium fusion.

This indicates that these pieces of evidence do not only individually support the *class* of Friedman models, but that they can be combined to jointly support a *single* Friedman model. This is an important step forward. Friedman models do not only allow the interpretation of snapshots of the universe taken at vastly different times, but they seem to single out one specific Friedman model that allows the consistent interpretation of all cosmological evidence discussed so far.

8. This picture can be extended by a few more colourful strokes. Further evidence is available that probes the Friedman models at epochs intermediate between the CMB and today.

Exciting and lively debated is the direct measurement of the cosmic expansion by means of a particular type of stellar explosion, the so-called supernovae of type Ia. We believe that such explosions arise when a white dwarf star is driven above its upper mass limit by matter overflowing from a companion star. Above this well-defined mass limit, the white-dwarf material is explosively ignited which disrupts the entire star. The amount of exploding material is thus known, approximately 1.4 solar masses, and therefore also the energy released, which sets the luminosity of the supernova. From the observed flux, we can then infer its distance. Its spectrum reveals when in cosmic history the supernova exploded. Type-Ia supernovae thus allow the reconstruction of the evolving of distances with the cosmic expansion, i.e. they directly probe the cosmic expansion history.

In doing so, they reveal an astonishing fact: When the universe was about half as old as it is now, its expansion began to accelerate. This is utterly counter-intuitive. We expect gravity to decelerate the cosmic expansion because of the usual gravitational attraction. Accelerated expansion is allowed, however, by General Relativity, provided there is a substance that Einstein introduced under the name of cosmological term or cosmological constant in order to stabilise the intrinsically unstable Friedman models.

We do not know what the cosmological constant could be. Attempts at explaining it in terms of a quantum field lead to the concept of dark energy, introduced for the sole purpose of interpreting the accelerated cosmic expansion indicated by type-Ia supernovae.

Strange as they may sound, these ideas receive substantial support from the CMB. Among the most solid conclusions from the statistical analysis of the CMB temperature fluctuations is the insight that the universe must be spatially flat. This is concluded directly from the size of the most pronounced warm and cool spots in the CMB. At fixed physical size, they appear larger if space is positively curved, and smaller if it is negatively curved. However, spatial flatness in the Friedman models is possible only if the matter or energy densities of all components of the cosmic fluid add up to a critical value of about one proton in five cubic metres. We thus know what the total matter and energy density in the universe is, but we know also what the densities of dark and ordinary matter are. Both together sum up to only about 30 % of the known total amount. If the difference is contributed by the cosmological constant or the dark energy, a model emerges which can precisely reproduce the expansion history probed by the type-Ia supernovae.

Another important class of observations probes the large-scale structure in the universe. Galaxies are not randomly distributed in space. Rather, they form galaxy clusters and extended filamentary structures, many millions of light-years long. Like the structures in the CMB, these structures in the distribution of cosmic objects carry most valuable statistical information. In particular, there is a characteristic length scale imprinted into the galaxy distribution which was set at a very specific epoch in cosmic history which is defined by the total matter density compared to the energy density of radiation in the universe. This implies that, if this characteristic scale in the galaxy distribution can be measured, the matter density can be inferred from it.

This approach requires galaxy surveys extending to distances that are substantially larger than the characteristic scale to be measured. Surveys of such size have become possible only in the recent past. They confirmed that the total matter density is about 30 % of the critical value, in agreement with the CMB data.

Yet another probe of cosmic structures dominated by dark matter is provided by the so-called gravitational lensing effect. General Relativity implies that concentrations of mass or energy deflect light in a way comparable to convex optical lenses. This gives rise to a multitude of interesting effects of different magnitude. In our context, the most important one is that any light ray propagating from a distant source to us must be deflected multiple times by the intervening large-scale structures, irrespective of what kind of matter they are composed of. This deflection gives rise to faint distortions of background galaxies which are indeed measurable, albeit with a formidable effort.

This cosmological weak lensing effect cannot distinguish between diluted matter that is clumped to a large degree and dense matter that is less clumpy because it is only sensitive to the absolute amount of inhomogeneity in the matter distribution. However, the results are well in agreement with a Friedman model in which the matter density reaches approximately 30 % of the total, critical density, while the rest is contributed by the cosmological constant or dark energy.

3 Consequences and Perspectives

What does it all mean? We have collected a substantial body of evidence in favour of the Friedman class of cosmological models. It is worth recalling what they are based upon: The only two ingredients were General Relativity, combined with spatial isotropy and homogeneity. Going one level deeper, General Relativity itself is built upon the concepts that the geometry of space-time adapts to the presence of matter and energy and that the experimentally well-established theory of Special Relativity remains locally valid. The dynamical equations governing the way how geometry reacts to the presence of matter and energy again follow from underlying symmetry and extremal principles extending far beyond General Relativity itself. Interestingly, it can be mathematically proven that General Relativity is unique in a quite general sense. Under broadly acceptable assumptions, Einstein's field equations are even the only dynamical equations possible.

The class of Friedman world models thus seems to stand on a rock-solid theoretical foundation, supported by a large body of empirical evidence. However, ways out are possible along the paths sketched in the beginning.

Either, one remains within General Relativity, then at least one of the two symmetry assumptions must be abandoned that the Friedman models are built upon. Any deviations from symmetry must, however, obey the tight limits on isotropy set in particular by the temperature of the CMB and its fluctuations. More vulnerable is the assumption of homogeneity, which is much harder testable, if at all. If we decided to give it up, we would have to accept being located not at a random, but at a fairly special place in the universe. While this is not at all impossible, it is quite unlikely that a sufficiently special place exists from where the universe looks as peculiar as it does, in particular in view of its accelerated expansion.

Alternatively, we could give up or modify General Relativity. The most gentle way of doing so consists in adding terms to the action that are still in agreement with the general underlying symmetries. The principle of least action then provides a standardised way of deriving modified field equations to replace Einstein's equations. Based upon them, Friedman's symmetry assumptions could be re-established to arrive at modified or generalised Friedman models. Alternatively, at least isotropy could be questioned in addition. However, General Relativity has so far survived all experimental tests it was subjected to. Admittedly, the most stringent tests all concerned local, weak gravity in the Solar System, but nonetheless these must also be met by alternative theories.

More radical approaches are also possible and are being pursued. One consists in extending General Relativity to more than four space-time dimensions. This was already suggested by Klein and Kaluza in the 1920s in attempts to unify electromagnetism and gravity and to explain quantum aspects of matter. The additional, fifth, dimension then introduced had to be considered as compactified, or rolled up, in order to be macroscopically hidden. This concept has been revived in current theories. Another approach aims at a quantum theory of gravity, which still seems well beyond the horizon.

Perhaps the most conservative point of view accepts that the foundations of the Friedman models are hard to shatter. Then, accepting the Friedman models and testing them against the empirical evidence leads to the single, standard model of cosmology that provides a consistent framework for virtually all cosmological observations. It comes, however, at the considerable price that dark matter and dark energy must then be accepted. We have some promising and testable ideas regarding the nature of the dark matter. Most likely, it is composed of weakly interacting, massive elementary particles. No suitable particle has yet been discovered, but it seems plausible that if dark-matter particles exist, either indirect evidence for them will be found at the Large Hadron Collider, or direct evidence in dedicated recoil experiments.

Dark energy, in contrast, remains essentially mysterious to us. It could be Einstein's cosmological constant, but then its theoretical foundation seems unsatisfactory. It could be some quantum field taking part in the cosmological evolution in some way, but so far there is no empirical evidence whatsoever that the dark energy might depend on time. It is well possible that, if General Relativity persists, we have to accept the cosmological constant in just the same way as we have to accept other constants of nature, such as the fine-structure constant or the elementary charge.

Unveiling the nature of the dark matter and the dark energy are at the heart of current cosmological research. In the context of cosmological model building, this is perhaps an irrelevant detail. What is important, however, is a remarkable reversal in the order of arguments usually leading to the construction or the dismissal of a model. The foundations of the Friedman class of cosmological models appear so solid that it seems more appropriate to accept the seemingly exotic consequences of dark matter and dark energy than to abandon the model. The situation reminds of a letter Einstein wrote to Sommerfeld after he had completed General Relativity. In this letter, Einstein remarked that he would lose no words in defending the theory because Sommerfeld would be convinced of it at a glance. In the cosmological standard model, the simplicity and the high degree of symmetry of the primary assumptions seem more appealing than the apparently preposterous consequences might be repelling.

4 Dark matter, dark energy, and the future of the Universe

Up to this point, the cosmological standard model may appear impressive by its simplicity and its consistency throughout almost all of cosmic history. We have, however, swept one of its major problems under the rug, which has to do with an apparent violation of causality.

As we have seen before, the CMB was released from the cosmic plasma when the universe was approximately 400,000 years old. During this time, light can obviously travel by no more than 400,000 light years. When compared to the full CMB sky, this distance corresponds to a very small angular scale. It spans an angle approximately as large as twice the full moon. Since no information can propagate faster than the speed of light, two points on the CMB separated by more than twice the so-called horizon radius of 400,000 light years could never communicate prior to the release of the CMB. How was it possible then that any two points on the CMB separated by more than a few angular diametres of the full moon could ever have arranged to attain the same temperature? How could the temperature information at one point on the CMB sky ever have propagated far enough to adjust the temperature to the same value everywhere?

One might object that this is of course necessarily so – in a model universe that has been set up to be ideally isotropic. By construction, the temperature must then be the same everywhere on the observer's sky, thus the observation of a CMB sky with constant temperature just appears as a consequence of the far-reaching symmetry assumptions that we started out with. This is not a way out, however, because coherent structures exist in the CMB that are larger than the horizon radius. Thus, even if we would be willing to accept isotropy of the CMB temperature without asking further how it could have been established in absence of causal mechanisms, the existence of coherent structures larger than the horizon radius implies that the processes creating these structures must have acted in a causally connected way even though the structures extend well beyond the scale of causal connection. This is an unbearable imposition.

The only feasible way out seems to be postulating a very early epoch in the cosmic evolution in which the universe expanded in a very strongly accelerated fashion. This epoch is called cosmological inflation. Among the primary purposes of its introduction was the causality problem just sketched. It solves this problem by assuming that tiny, causally connected regions in the primordial universe were stretched to cosmological size by the inflationary expansion, turning a potentially small section of a region that was previously in causal contact into our observable universe.

It is unclear what this epoch of cosmological inflation could have been driven by. A suitable quantum field called the inflaton is postulated for this purpose. This may seem helplessly unsatisfactory, but it has observable consequences. One of them is that any quantum field must undergo fluctuations because of Heisenberg's uncertainty principle. During the inflationary epoch, these quantum fluctuations would have been stretched to macroscopic and even cosmological scales such that they could later form the seeds for the rich variety of cosmic structures we see today. Even though it may appear ludicrous, this hypothesis allows a calculation of the expected statistical properties of cosmic structures produced that way. Measurements of the temperature fluctuations in the CMB confirm this expectation precisely.

This gives rise to the truly breath-taking notion that cosmic structures may have originated in quantum fluctuations of a primordial inflaton field that drove the early phase of accelerated expansion. If this seems incredible, we must recall that we have now entered another phase of accelerated cosmic expansion, as demonstrated directly by the type-Ia supernovae and indirectly by the CMB.

References

Singh, Simon (2005). *Big Bang. The Origin of the Universe*. New York: Harper Perennial.

Prof. Dr. Matthias Bartelmann
Center for Astronomy of the University of Heidelberg
Institute of Theoretical Astrophysics
Philosophenweg 12
69120 Heidelberg
Germany
bartelmann@uni-heidelberg.de