Philosophica 63 (1999, 1) pp. 7-17

INFLATION. WHERE DID THAT COME FROM?

Gustaaf C. Cornelis

1. Introduction

It would have been perfect. Empirical data generate theoretical problems (the so-called cosmological paradoxes) and induce a crisis in scientific thinking, i.e. cosmology. A revolution occurs and a new paradigm, *in casu* big bang cosmology *annex* inflation theory (the idea that the universe grew exponentially in a split second) emerges. Did the search for dark matter generate the flatness problem, and, subsequently, did the flatness problem lead to inflation theory, did inflation theory solve the flatness paradox, and did it become the new paradigm? Unfortunately, this train of thought cannot be substantiated. So where did the 'ridiculous' idea of inflation come from?

2. The paradoxes

During the sixties and seventies the so-called cosmological paradoxes came forth, known as the horizon-problem, the flatness-problem and the smoothness-problem. The last one originated out of the first. In literature, these problems were treated as *genuine* cosmological paradoxes, although — as will be shown — only one can be proven to be a genuine paradox.

2.1. The horizon-problem

In 1969 Charles Misner presented the horizon-paradox (Misner, 1969). It states the question *why the universe seems to be more uniform than can be explained through thermodynamical processes*. Misner acknowledged the fact that the universe could only have a more or less equal distribution of matter and temperature *if* all parts could 'communicate' with each other at a certain time in the past — given, of course, that the universe did not originate out of a 'perfect' singularity. If the universe had an extent from the beginning, regions 'opposite' to each other could not ever connect in a young universe that expands initially with the velocity of light. To get a uniform universe in certain aspects, information regarding these had to be shared throughout the universe. Carriers have a limited speed, hence, there were no means at hand to make the young universe homogeneous.

2.2. The flatness-problem

The second riddle concerns the value of Ω, the density parameter of the universe. In 1922 Alexander Friedman proved that the mean density would determine the overall geometry. The universe could be open, closed or flat. The last solution needed unity for Ω.

Astronomers tried (and are still trying) to measure Ω. Beatrice Tinsley, J.R. Gott, D.N. Schramm and J.E. Gunn concluded that Ω was far too small to close the universe. If all light-producing material in the universe was counted, the density laid between one and ten percent of the ideal value.

If the universe is flat ($\Omega = 1$), it was always flat or it became flat. In general, it is accepted that the universe expands without gaining mass. Hence, the density decreases and the universe became (and still gets) slowly flatter all the time. But, if the universe is *now* almost flat, the value of Ω had to be very close to unity in the far past.

The following question was raised: why should the universe be flat in the past? The extreme conditions shortly after the 'big bang' should have made the universe strongly Riemannian. So, what made the universe (almost) flat? This is the so-called flatness-problem.

It is important to notice that in literature the flatness-problem is frequently mixed up with two other problems, *viz.* the question 'why is the universe now seemingly flat, while not enough matter can be found to account for its flatness?' and the so-called *age-problem*: the fact that the interdisciplinary time-scales are incompatible with each other (in the second case, the mix-up rose as a consequence of the fact that the flatness-problem is sometimes called the *age-paradox*).

2.3. The smoothness-problem

The smoothness-problem can be wrongly seen as the counterpart of the horizon-problem. They have both relevance to the homogeneity of the universe, but they do not 'neutralize' each other. The smoothness-problem tries to explain the local heterogeneity, given that the universe is globally homogeneous. In other words, we see galaxies and clusters of galaxies around us, so the universe is highly structured, but observations of the cosmic background radiation (the supposed remnant of an early era) show that the universe was fairly homogenous in the far past. The horizon-paradox has to do with the question how the universe could get so homogeneous after all, given that it did not start off being so and given that there was no possibility to get the observed early uniformity fast enough.

2.4. Why are they called 'paradoxes'?

The horizon problem concerns the conflict between on the one hand the observation of a uniform radiation-distribution and on the other hand the fact that the universe is so vast that, given that it did not arise out of a point-singularity, it could not get so homogeneous in time. As to the smoothness problem, there is the contradiction between the observation of a heterogeneous matter-distribution and the description of the earliest stages of the universe by the standard theory.

Once accepted that the universe has a density parameter equal to one, there is a conflict with standard theory: the big bang model can not account for this global flatness. The flatness-problem is clearly a genuine paradox, since there is no empirical reason to state that the universe globally is perfectly flat. Standard theory implies a value for Ω between 0.1 and 1.5 — roughly speaking (depending on the value of the cosmological constant and the Hubble constant, slightly different values are possible). The point is that standard theory does not fix the value of Ω. It does not follow that Ω is equal to one, nor that it was close to one in the past. The fact that standard theory does not offer any explanation to why Ω began close to one and therefore assumes that Ω's value was always close to one, is the essence of the flatness-paradox.

2.5. Historiography

The horizon-problem *should have been taken seriously* as soon as the nature of the particle-horizon was understood by Rindler (Rindler, 1956). It should have been a acute problem when the high degree of isotropy of the background radiation was discovered (Partridge and Wilkinson, 1967). It did not, although Roger Penrose studied the problem in depth (Penrose, 1968) and Misner stressed the pertinence of it in view of the latest observations of the background radiation at that time (Misner, 1968). The majority of the astronomers and cosmologists alike became aware of the problem after a solution was given through inflation theory. We shall see that this solution was merely a spin-off.

The smoothness-problem arose after the horizon problem, after the mega-structures were discovered: where did the galaxies within the superstructures come from, given the fact that the universe was homogeneous at earlier times?

If the flatness problem is *now* a problem, it *should have been* a problem as early as the thirties, when a first estimate of the density was made. However, it seems that Dicke and Peebles were the first to draw real attention to it (Dicke and Peebles, 1979). The flatness problem was probably spelled out for the first time by Dicke (1968). Although cosmologists like Hawking and Carter reformulated the problem shortly after, only after the publication by Dicke and Peebles the problem was generally understood. Nonetheless, the problem made its impact after inflation theory was properly developed by Alan Guth (1981).

3. Inflation

3.1. The original inflation theory

The symmetry breaking process which happened shortly after the big bang could not have produced a unique asymmetry. Relatively small areas existed in the universe where the 'old' symmetry-laws still governed. They shared their characteristics with the theoretical particles Dirac searched for in 1931: the magnetic monopoles.

Henry Tye wanted to know whether the Higgs-field could produce magnetic monopoles when the unified force broke down to the strong and electroweak forces. He contacted Alan Guth, a physicist specialised in Higgs-fields. Guth proved — helping out his friend, but not interested in these matters at all — that magnetic monopoles did emerge out of the SU(5) model (a mathematical structure known as a group that describes operations on five objects. SU(5) is a *Grand Unification Theory*). They were quite heavy (10^{16} GeV), virtually undetectable (too slow) and therefore not interesting to Guth at all.

Guth met Steven Weinberg and discussed with him matters concerning symmetry-breaking and features of the Higgs-fields. Although several interesting questions were raised, the problems related to the monopoles could not succeed in fascinating Guth. When Robert Dicke mentioned the flatness problem to him, Guth said that it just did not relate to his work (Overbye, 1990). Nevertheless, Guth acknowledged the incompleteness of standard theory, as implied by the flatness problem: the big bang model could not account for the fact that the universe became flat.

3.2 The monopole-problem

At last (fall 1978), Tye could persuade Guth to compute the particle-output of SU(5). They found that there should be as much monopoles as there are protons. This was clearly not the case on an empirical base. Why had no one detected a monopole? Or, with other words, which cosmological phenomenon makes these particles virtually invisible?

This problem, having the same logical structure as the above mentioned problems, can indeed be considered to be a cosmological paradox as well: the monopole-problem.

Although SU(5) implied the paradox, this theory was not abandoned or at least revised — it was not a cosmological theory, anyway. Tye and Guth could not imagine to have made a miscalculation and after they found out that John Preskill came to the same conclusions in 1979, they knew that they were on the right track. Tye and Guth focused on the question whether there was anything that could be changed in the assumptions to make *grand unified theories* compatible with cosmology. Maybe the generating Higgs-field could do the trick? They focused on 'supercooling' (Weinberg's metaphor originally).

Guth and Tye supposed that the universe cooled down below 10^{27} K, without rigidifying the Higgs-field (i.e., without changing the symmetry).

The expansion of the universe evidently continued and brought the temperature further down. Any local asymmetry that was generated could spread quite far, although a unidirectional symmetry was completely out of the question. In any case, the abundance of monopoles was according to this new theory relatively small — small enough to make it almost impossible to find a monopole in this *region* of the universe.

Sidney Coleman had performed some independent research on the 'blocking' of symmetry-breaking processes. Coleman called the phase of decelerated cooling, the 'false vacuum'. From a classical point of view, the false vacuum appeared to be a stable phase. After a while, Coleman found out that it was possible that locally the conditions could generate through a quantum fluctuation a certain instability: a tunneling effect which made it possible to regain the 'frozen' energy of the Higgs-fields and to trigger off the symmetry-breaking process.

Tye and Guth got acquainted with Coleman's work and concluded that the phase of false vacuum combined with a tunneling effect could slow down the production of monopoles, just enough to disperse them through the universe. SU(5) was saved.

At first it was not clear whether the supercooling had some side-effects. So Guth and Tye did not mention anything about possible implications for standard cosmology in their paper of 1980. However, they proved that a supercooling of the symmetry-breaking process was in fact an exothermic process (the vacuum would not generate an implosion). The accumulated energy would 'inflate' the young universe to enormous proportions in a flick of time. Every 10^{-34} s, the universe would double its volume. Obviously, the cosmological implications were devastating. However, the theory was not inconsistent with standard big bang cosmology; it could be implemented in the overall structure.

What about the initial problem? During the supercooling stage, only a few monopoles would be generated and during the inflation phase, they would be scattered throughout the universe, disappearing behind the horizon (they can always show up later). The monopole problem was solved.

In a way, the flatness problem was resolved too: the inflation theory was compatible with Ω being equal to one during the first moments after the inflation phase. Guth realised only after developing the theory that inflation "would solve the flatness problem" (Lightman and Brawer, 1990:472). It is interesting to see that only a few cosmologists know that inflation theory does not *necessarily* imply an ideal value for Ω. In fact, Ω could have taken any value, *a fortiori* a value below one, compatible with the observations (e.g., Ω=0.1) (Madsen and Ellis, 1984; Ellis, 1989:400):

> As in the past in the history of cosmology, some of the widely cherished beliefs are based on dogma rather than physics or observational evidence. An example is *the widely made claim that an inflationary universe necessarily leads to a present density parameter Ω that is very close to unity. This is not true: while inflation makes it much more likely that Ω will be close to unity than if there is no inflation, this does not imply that Ω=1 today*; on the contrary, there are inflationary universe models leading to any desired present-day value for Ω. The inflationary idea is important because it can solve the horizon problem, but it does not necessarily imply that most of the matter in the universe is 'dark matter' [or 'missing matter']. Thus a great deal of current astronomical and theoretical activity aimed at searching for such matter (which is quite worthwhile on other grounds) is not necessitated by the inflationary-universe idea, contrary to many claims in the published literature.

So, inflation theory does not make Ω equal to one necessarily.

Finally, the horizon problem was solved: inflation made the universe expand beyond speed of light. The observable universe developed out of a small part of the embryonic (or pre-inflationary) universe, small enough to get homogeneous in time.

3.3. New inflation theory

Independently, the Russian cosmologist Andrei Linde came up with inflation theory too. In contrast to Guth, Linde had always been working on cosmological matters. In 1978, Linde tried to solve the ratio of photons to baryons (Linde, 1979). When he learned about the flatness problem (after he initiated his project), he thought for a while that it concerned the same question.

The solution he found was problematic. Apparently, inflation generated 10^{80} bubbles which had to melt together in some way or another to make this theory acceptable. Guth came to the same conclusion only a few months earlier: inflation, indeed, was in big trouble. Only after getting acquainted with Guth's work, Linde continued his own research: the fact that a colleague had stumbled upon the same complex questions, made the research worth the effort. Paul Steinhardt and Andreas Albrecht teamed up and continued where Guth and Linde got off the rails (Albrecht and Steinhardt, 1982). The independent research converged in the so-called new inflation theory. It is interesting to see how two different research programs came about, leading to the same conclusions (one in America — Guth, one in Russia — Andrei Linde).

3.4. Chaotic inflation theory

Inspired by the eminent cosmologist Markov and by Indian mythology, Andrei Linde went beyond new inflation and devised the so-called *chaotic inflation model*. He acknowledged the fact (Lightman and Brawer, 1990:490) that he was exceedingly influenced by Indian thought, the idea of a universe that lives and then dies, expanding and contracting.

In this model, the chaotic conditions in the early universe generate large fluctuations in the values of certain parameters and a complicated network of mini-universes would be produced. This process can continue forever, leading to what Linde calls the 'eternally existing self-producing chaotic inflationary universe'. Inside the universe, several regions undergo inflation during a short period, creating spaces that share features with the region we live in.

4. Conclusions

4.1. Initial role of the monopole-problem

The monopole-problem differs form the others in two ways: it has an interdisciplinary nature and served as the direct cause to inflation theory. It is clear that at least one theoretical problem served as a starting point for cosmological research.

4.2. Confirming roles of the flatness-problem and the horizon-problem

The flatness problem (that is, Linde's own faulty interpretation of it) was only stimulating to Linde's research: in reality he tackled the photon/baryon problem first. Guth tried to solve the monopole problem. After developing the inflation theory, he saw that the flatness problem and the horizon problem disappeared. The solution to the two cosmological paradoxes was a confirmation for the theory. The theoretical problem of the bubbles originated the search for a new inflation theory.

It was a mistake to think that standard theory combined with inflation theory implied a value for Ω equal to one. Two possible explanations can be given. Guth mentioned himself (Guth and Steinhardt, 1984) that inflation theory solves the flatness problem. For sure, standard theory combined with the inflation concept was compatible with a value of one for Ω. But this was also the case without inflation theory altogether.

There were still other research programmes which *really* did tackle the flatness problem. Several scientists tried to account for the missing matter (needed to close the universe, 'making' Ω equal to one). Standard theory was enlarged by all sorts of weird theories in the realm of particle physics. A zoo of so-called exotic particles originated out of the blue. Others tried to match up standard theory with the ideal value for Ω through varying over the fundamental constants, including the Hubble-constant and the cosmological constant. The Hubble-constant was *highly* subjected to observations, but the other parameters were not. It is obvious that this methodology only shifted the problem. Not willing to choose a value for Ω, other parameters were fixed in a rather arbitrary way.

The flatness-problem has a numerological ring to it. Why should anyone bother to do some research just to prove that a certain parameter has an ideal value? Why should it be precisely equal to one and not, say, 1.23456...? It is very intriguing to see that originally the value had to be one (on aesthetical grounds), while after doing research some astronomers seemed to settle for values differing from one. Clearly, any divergence from an ideal value weakens (and in my opinion even annihilates) previously made aesthetical considerations. To put it another way, the reason to search for the value of the density-parameter *nowadays* does not supply the primary criterion any longer to evaluate the empirical found values.

4.3. Synthesis

The inflation concept was merely a side effect of the research results (found by Alan Guth) concerning the monopole problem (which is a paradox to particle physics). The pertinence of inflation to cosmology and the other paradoxes came later. Reflecting on the work of George Ellis, we conclude that the relevance of inflation, as far as the paradoxes are concerned, is of minor importance. The inflation concept does not necessarily imply the boundary conditions for the universe that make the paradoxes work.

Vrije Universiteit Brussel

REFERENCES

Albrecht, A. and P.J. Steinhardt (1982). 'Cosmology for Grand Unified Theories with Radiatively Induced Symmetry Breaking.' *Physical Review Letters* **48**: 1220.

DeWitt, C. and J.A. Wheeler (1968). *Battelle Rencontres: 1967 Lectures in Mathematics and Physics*. New York: Benjamin.

Dicke, R.H.(1968). 'Gravitation and the Universe.' In: *The Jayne Lectures for 1969. American Philosophical Society.*

Dicke, R. H. and P.J.E. Peebles (1979), 'The Big Bang Cosmology—Enigmas and Nostrums.' In: Hawking and Israel (1979).

Dirac, P. A.M. (1931). 'Quantized Singularities in the Electromagnetic Field.' *Proceedings of the Royal Society* **A133**: 60.

Ellis, G. F.R. (1989), 'Cosmology from 1917 to 1960.' In: Howard and Stachel (1989).

Guth, A.H. (1981). 'Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems.' *Physical Review D* **23**: 347-356.

Guth, A.H. and S.-H.H. Tye (1980). 'Phase Transitions and Monopole Production in the Very Early Universe.' *Physical Review Letters* **44**: 631.

Hawking, S.W. and W. Israel (1979). *General Relativity: An Einstein Centenary Survey*. Cambridge: Cambridge University Press.

Howard, D. and J. Stachel (1989). *Einstein and the History of General Relativity*. Boston: Birkhäuser.

Lightman, A. and R. Brawer (1990). *Origins. The Lives and Worlds of Modern Cosmologists*. Cambridge: Harvard University Press.

Linde, A.D. (1979). 'Phase Transitions in Gauge Theories and Cosmology.'

Reports on the Progress of Physics **42**: 389.

Madsen, M.S. and G.F.R. Ellis (1984). 'The Evolution of Omega in Inflationary Universes.' *Monthly Notices of the Royal Astronomical Society* **234**: 67-77.

Misner, C.W. (1969). 'The Mixmaster Universe.' *Physical Review Letters* **22**: 1071.

Overbye, D. (1990). *Lonely Hearts of the Cosmos.* New York: Harper & Row.

Partridge R.B. and D.T. Wilkinson (1967). 'Isotropy and Homogeneity of the Universe from Measurements of the Cosmic Background Radiation.' *Physical Review Letters* **18**: 557-559.

Penrose, R. (1968). 'Structure of Space-Time.' In: DeWitt and Wheeler (1968).

Preskill, J.P. (1979). 'Cosmological Production of Superheavy Magnetic Monopoles.' *Physical Review Letters* **43**: 1365.

Rindler, W. (1956). 'Visual Horizons in World Models.' *Monthly Notices of the Royal Astronomical Society* **116**: 662-677.