Chapter 1

The physics of the early universe (an overview)

Silvio Bonometto
Department of Physics, University of Milan–Bicocca, Milan, Italy

1.1 The physics of the early universe: an overview

Modern cosmology has a precise birthdate, Hubble's discovery of Cepheids and ordinary stars in *Nebulae*. The nature of nebulae had been disputed for centuries. As early as 1755, in his *General History of Nature and Theory of the Sky*, Immanuel Kant suggested that nebulae could be galaxies. The main objection to this hypothesis has been supernovae. Today we know that, close to its peak, a supernova can exceed the luminosity of its host galaxy. But, while this remained unknown, single stars as luminous as whole nebulae were a severe objection to the claim that nebulae were made of as many as hundreds of billions stars. For instance, in 1893, the British astronomer Mary Clark reported the observation of two stellar bursts in a single nebula, one 25 years after the other. She wrote that: *The light of the nebula has been practically cancelled by the bursts, which… should have been of an order of magnitude so large, that even our imagination refuses in conceiving it.* Clark was not alone in having problems conceiving the energetics of supernovae.

After the recognition that most nebulae were galaxies, Hubble also claimed that they receded from one another, as fragments of a huge explosion. Such an expansive trend, currently named the *Hubble flow*, has been confirmed by the whole present data-set. Although there are no doubts that Hubble's intuition was great, the point is that his data-set did not show that much. At the distances where he pretended to see an expansive trend, the 'Hubble flow' is still dominated by peculiar motions of individual galaxies. Discovering the true nature of nebulae was, however, essential. It is the galactic scale which sets the boundary above which dynamical evolution is mostly due to pure gravity. Dissipative forces, of course, still play an essential role above such a scale. But even the huge x-ray emission from galaxy clusters, now the principal tool for their detection, bears limited dynamical effects.

Galaxies, therefore, are the inhabitants of a super-world whose rules are set by relativistic gravitation. Their average distances are gradually increasing, within the Hubble flow. The Friedmann equations tell us the ensuing rate of matter density decrease and how such a rate varies with density itself. No doubts, then, that the early universe must have been very dense. The cosmic clock, telling us how long ago density was above a given level, is set by the Hubble constant $H = 100h$ km s^{-1} Mpc^{-1}. Here h conveys our residual ignorance, but it is likely that $0.6 < h < 0.8$, while almost no one suggests that h lies outside the interval 0.5–0.9. (One can appreciate how far from reality Hubble was, considering that he had estimated that $h \simeq 5$.)

A realistic measure of h came shortly before the discovery of the cosmic background radiation (CBR). The Friedmann equations could then also determine how temperature varies with time and it was soon clear that, besides being dense, the early universe was hot. This defined the early environment and, until the 1980s, modern cosmologists essentially used known physics within the frame of such exceptional environments. In a sense, this extended Newton's claim that the same gravity laws hold on Earth and in the skies. On the basis of spectroscopical analysis it had already become clear that such a claim could be extended beyond gravity to the laws governing all physical phenomena, thereby leading cosmologists to extend these laws back in time, besides far in space.

1.1.1 The middle-age cosmology

This program, essentially based on the use of general relativity, led to great results. It was shown that, during its early stages, the universe had been homogeneous and isotropic, apart from tiny fluctuations, seeds of the present inhomogeneities. Cosmic times (t) can be associated with redshifts (z), which relate the *scale factor* $a(t)$ to the present scale factor a_0, through the relation

$$1 + z = a_0/a(t).$$

The redshift z also tells us the temperature of the background radiation, which is $T_0(1 + z)$ ($T_0 \simeq 2.73$ K is today's temperature).

On average, linearity held for $z > 30$–100. For $z > 1000$, the high-energy tail of the black body (BB) distribution contained enough photons, with an energy exceeding $B_{\rm H} = 13.6$ eV, to keep all baryonic matter ionized. Roughly above the same redshift, the radiation density exceeds the baryon density. This occurs above the so-called *equivalence* redshift $z_{\rm eq} = 2.5 \times 10^4 \Omega_{\rm b} h^2$. Here $\Omega_{\rm b}$ is the ratio between the present density of baryon matter and the present critical density $\rho_{\rm cr}$, setting the boundary between parabolic and hyperbolic models. It can be shown that $\rho_{\rm cr} = 3H_0^2/8\pi G$.

The relativistic theory of fluctuation growth, developed by Lifshitz, also showed that, in their linear stages, inhomogeneities would grow proportionally to $(1 + z)^{-1}$, if the content of the universe were assumed to be a single fluid. This moderate growth rate tells us that the actual inhomogeneities could not arise from purely statistical fluctuations. When the Lifshitz result was generalized to any kind of matter contents, it also became clear that fluctuations compatible with observed anisotropies in the CBR were too small to turn into galaxies, unless another material component existed, already fully decoupled from radiation at $z \simeq 1000$, besides baryons.

Various hypotheses were then put forward, on the nature of such *dark* matter, whose density, today, is $\Omega_c \rho_{cr}$. (The world is then characterized by an overall *matter* density parameter $\Omega_m = \Omega_c + \Omega_b$.) But, as far as cosmology is concerned, only the redshift z_d when the quanta of dark matter become non-relativistic matters. Let M_d be the mass scale entering the horizon at z_d and let us also recall that the mass scale entering the horizon at $z_{eq} = 2.5 \times 10^4 \Omega_m h^2$ is $\sim 10^{16} M_\odot$. Early fluctuations, over scales $< M_d$, are fully erased by free-streaming, at the horizon entry. If one wants to preserve a fluctuation spectrum extending to quite small scales, it is therefore important for z_d to be large.

As far as cosmology is concerned, the nature of dark matter can therefore be classified according to the minimal size of fluctuations able to survive. If fluctuations are preserved down to scales well below the galactic scale ($M_g \sim 10^8$–$10^{12} M_\odot$), we say that dark matter is *cold*. If dark matter particles are too fast, and become non-relativistic only at late times, so that $M_d > M_g$, we say that dark matter is *hot*. In principle, in the latter case galaxies could also form, because of the fragmentation of greater structures in their nonlinear collapse, which, in general, is not spherically symmetric. But such *top–down* scenarios were soon shown not to fit observational data. This is why cold dark matter (CDM) became a basic ingredient of all cosmological models.

This argument is quite independent from the assumption that Ω_m has to approach unity, in order for the geometry of spatial world sections to be flat. However, once we accept that CDM exists, the temptation to imagine that $\Omega_m = 1$ is great. There is another class of arguments which prevents Ω_b from approaching unity by itself alone. These are related to the early formation of light elements, like ^{2}H, ^{4}He, ^{7}Li. The study of big-bang nucleosynthesis (BBNS) has shown that, in order to obtain the observed abundances of light nuclides, we ought to have $\Omega_b h^2 \simeq 0.02$. BBNS occurred when the temperature of the universe was between 900 and 60 keV (ν decoupling and the opening of the deuterium bottleneck, respectively). At even larger temperatures, strongly interacting matter had to be in the quark–hadron plasma form. Going backwards in time we reach T_{ew}, when the weak and electromagnetic interactions separated. To go still further backwards, we need to speculate on physical theories, as experimental data are lacking. The physics of cosmology, therefore, starts from hydrodynamics and reaches advanced particle physics. In this book, a review of the physics of cosmology is provided in the contribution by John Peacock.

All these ages, starting from the quark–hadron transition, through the era when lepton pairs were abundant, then through BBNS, to arrive at the moment when matter became denser than radiation and finally to matter–radiation decoupling and fluctuation growth, are the so-called *middle ages* of the world. Their study, until the 1980s, was the main duty of cosmologists. Not all problems, of course, were solved then. Moreover, as fresh data flowed in, theoretical questions evolved. In his contribution Piero Rosati reviews the present status of observational cosmology, in relation to the most recent data.

The world we observe today is the result of fluctuation growth through linear and nonlinear stages. The initial *simplicity* of the model has been heavily polluted by nonlinear and dissipative physics. Tracing back the initial conditions from data requires both a theoretical and a numerical effort. In his contribution Anatoly Klypin presents such numerical techniques, the role of which is becoming more and more important. Using recent parallel computing programs, it is now possible to try to reproduce the events leading to the shaping of the universe.

The point, however, is that, once this self-consistent scenario became clear, cosmology was ready for another leap. Since the 1980s, it has become a new paradigm within which very high-energy physics could be tested.

1.1.2 Inflationary theories

The world we observe is extremely complex and inhomogeneous. The level of inhomogeneity gradually decreases when we go to greater scales (on this subject, see the contribution by Luigi Guzzo; another less shared point of view is exposed by Marco Montuori and Luciano Pietronero). But only the observations of CBR show a 'substance' close to homogeneity. In spite of this, the driving scheme of the cosmological quest had been that the present complexity came from an initial simplicity and much effort has been spent in developing a framework able to show that this is what truly occurred. When this desire for unity was fulfilled, cosmologists realized that it had taken them to a deadlock: the conditions from which the observed world had evidently arisen, which so nicely fulfilled their intimate expectations, were so exceptional as to require an exceptional explanation.

This is the starting point of the next chapter of cosmological research, which started in the 1980s and was made possible by the great achievements of previous cosmological research. The new quest took two alternative directions. The most satisfactory possibility occurred if, starting from generic metric conditions, their eventual evolution necessarily created the exceptional 'initial conditions' needed to give a start to the observed world. An alternative, weaker requirement, was that, starting from a generic metric, its eventual evolution necessarily created *somewhere* the exceptional 'initial conditions' needed to give a start to the observed world.

The basic paradigm for implementing one of such requirement is set by inflationary theories. The paradoxes such theories are called to justify can be listed as follows:

(i) Homogeneity and isotropy: apart from tiny fluctuations, whose distribution is itself isotropic, the conditions holding in the universe, at $z > 1000$, are substantially identical anywhere we can observe them. The domain our observations reach has a size $\sim ct_0$ (c, the speed of light; t_0, the present cosmic time). This is the size of the regions causally connected today. At $z \sim 10^3$, the domain causally connected was smaller, just because the cosmic time was $\sim 10^{4.5}$ times smaller than t_0. Let us take a sphere whose radius is $\sim ct_0$. Its surface includes ~ 1000 regions which were then causally disconnected one from another. In spite of that, temperature, fluctuation spectrum, baryon content, etc, were equal anywhere. What made them so?

(ii) Flatness: According to observations, the present matter density parameter $\Omega_{\rm m}$ cannot deviate from unity by more than a factor 10. (Recent observations on the CBR have reduced such a possible discrepancy further.) But, in order for $\Omega_{\rm m} \sim 0.1$ today, we need to *fine-tune* the initial conditions, at the Planck time, by 1:10^{60}. To avoid such tuning we can only assume that the spatial section of the metric is Euclidean. Then it remains as such forever.

(iii) Fluctuation spectrum: Let us assume that it reads:

$$P(k) = Ak^n.$$

Here $k = 2\pi/L$ and L are comoving length scales. This spectral shape, apparently depending on A and n only (spectral amplitude and spectral index, respectively), tries to minimize the scale dependence. But a fully scale-independent spectrum is obtained only if $n = 1$. It can then be shown that fluctuations on any scale have an identical amplitude when they enter the horizon. This fully scale-independent spectrum, first introduced by Harrison and Zel'dovich, approaches all features of the observed large-scale structure (LSS). How could such fluctuations arise and why did they have such a spectrum?

Apart from these basic requirements, there are a few other requests such as the absence of topological monsters that we shall not discuss here.

The scheme of inflationary theories amounts then to seeking a theory of fundamental interactions which eliminates these paradoxes. The essential ingredient in achieving such an aim is to prescribe a long period of cosmic expansion dominated by a false vacuum, rather than by any kind of *substance*. Early periods of vacuum dominance are indeed expected, within most elementary particle theories, and this sets the bridge between fundamental interaction theories and cosmological requirements.

In this book, inflationary theories and their framework are discussed in detail by Andrei Linde and George Ellis, and therefore we refrain from treating them further in this introduction. Let us rather outline what is the overall resulting scheme. One assumes that, around the Planck time, the universe emerges from quantum gravity in a *chaotic* status. Hence, anisotropies, inhomogeneities, discontinuities, etc, were dominant then.

However, such a variety of initial conditions has nothing to do with the present observed variety. The universe is indeed anisotropic, inhomogeneous, discontinuous, etc, today; and more and more so, as we go to smaller and smaller scales. But such *secondary* chaos has nothing to do with the *primeval* chaos. It is a kind of *moderate* chaos that we have reached after passing through intermediate highly symmetric conditions. The sequence *complex* $\to$ *simple* $\to$ *complex* had to run, so that today's world could arise.

1.1.3 Links between cosmology and particle physics

There are, therefore, at least two fields where the connections between particle physics and cosmology have grown strong. As we have just outlined, explaining why and how an inflationary era arose and runs is certainly a duty that cosmologists and particle physicists have to fulfill together.

In a sense, however, this is a more speculative domain, compared with the one opened by the need for a dark component. The first idea on the nature of dark matter was that neutrinos had mass. A neutrino background, similar to the CBR, must exist, if the universe ever had a temperature above $\sim$1 MeV. Such a background would be made by $\sim$100 neutrinos/cm^3, for each neutrino flavour. It is then sufficient to assume that neutrinos have a mass $\sim$10–100 eV, to reach $\Omega_{\rm m} \sim 1$.

Such an appealing picture, which needs no hypothetical new quanta, but refers to surely existing particles only, was, however, shown not to hold. Neutrinos could be *hot* dark matter, as they become non-relativistic around $z_{\rm eq}$. As we have already stated, the *top–down* scenario, where structures on galactic scales form thanks to greater structure fragmentation, is widely contradicted by observations.

This does not mean that massive neutrinos may not have a role in shaping the present condition of the universe. Models with a mix of cold and hot dark matter were considered quite appealing until a couple of years ago. Their importance, today, has somehow faded, owing to recent data on dark energy. Recent data on the neutrino mass spectrum are reviewed by Gianluigi Fogli in his contribution.

Alternative ideas on the nature of dark matter then came from supersymmetries. The lightest neutral supersymmetric partner of existing bosons is likely to be stable. In current literature this particle is often called the *neutralino*. There are quite a few parameters, concerning supersymmetries, which are not deducible from known data and, after all, supersymmetries themselves have not yet been shown to be viable. However, well within observationally acceptable values, it is possible for neutralinos to have mass and abundance such as to yield $\Omega_{\rm m} \sim 1$.

In their contribution Antonio Masiero and Silvia Pascoli focus on the interface between particle physics and cosmology, discussing in detail the nature of CDM. Andrea Giuliani's paper deals with current work aiming at detecting dark matter quanta in laboratories and the contribution by Rita Bernabei *et al* relates possible evidence for the detection of neutralinos. Various hypotheses were considered, about dark matter setting. Its distribution may differ from visible matter, on various scales. By definition, its main interaction, in the present epoch, occurs via gravity and gravitational lensing is the basic way to trace its presence. In his contribution Philippe Jetzer reviews the basic pattern to detect dark matter, over different scales, using the relativistic bending of light rays.

1.1.4 Basic questions and tentative answers

There can be little doubt that the last century has witnessed a change of the context within which the very word 'cosmology' is used. Man has always asked basic questions, concerning the origin of the world and the nature of things. The only answers to such questions, for ages, came from metaphysics or religious beliefs. During the last century, instead, a large number of such questions could be put into a scientific form and quite a significant number could be answered.

As an example, it is now clear that the universe is evolutionary. At the beginning of modern cosmology, models claiming a steady state (SS) of the universe had been put forward. They have been completely falsified, although it is now clear that the stationary expansion regime, introduced by SS models, is not so different from the inflationary expansion regime, needed to make big-bang models self-consistent. Furthermore, if recent measures of the deceleration parameter are confirmed, we seem to be living today in a phase of accelerated expansion, quite similar to inflation. It ought to be emphasized that the strength of the data, supporting this kind of expansion, is currently balanced by the theoretical prejudices of wise researchers. In fact, an accelerated expansion requires a desperate fine-tuning of the vacuum energy, which seems to spoil all the beauty of the inflationary paradigm.

Since Hubble's hazardous conclusion that the universe was expanding, the century which has just closed has seen a number of results, initially supported more by their elegance than by data. The Galilean scheme of experimental science is not being forgotten, but one must always remember that such a scheme is far from requiring pure experimental activity. The basic pattern to physical knowledge is set by the intricate network of observations, experiments and predictions that the researcher has to base on data, but goes well beyond them. With the growing complication of current research, the theoretical phase of scientific thought is acquiring greater and greater weight. During such a stage, the lead is taken by the same criteria which drove mathematical research to its extraordinary achievements.

Besides Hubble's findings, within the cosmological context, we may quote Peebles' discovery of the correlation length r_0, based on angular data, which have recently been shown to allow quite different interpretations. Outside cosmology, the main example is given by gauge theories, which are now the basic ingredient of the standard model of fundamental interactions, and were deepened, from 1954 to the early 1970s, only because they were *too beautiful not to be true*. At least two other fields of research in fundamental physics are now driven by similar criteria—supersymmetries and string theories (see the paper by Renata Kallosh).

While supersymmetries can soon be confirmed, either by the discovery of neutralinos by passive detectors or at CERN's new accelerator, string theories might only find confirmation if signals arriving from the Planck era can be observed. This might be possible if future analyses of CBR anisotropies and polarization show the presence of tensor modes. In this book a review of current procedures for CBR analysis is provided by Arthur Kosowsky.

Also within the cosmological domain, leading criteria linked to aesthetical categories are now being pursued. However, in this field, the concept of beauty is often directly connected with ideological prejudices. Questions such as '*can the universe tunnel from nothing*' have been asked and replied within precise physical contexts. It is, however, clear that the ideological charge of such research is dominant. Moreover, when theoretical results, in this field, are quoted by the media, the distinction between valid speculations and scientific acquisitions often fully fades.

But the main question, for physicists, is different. For at least two centuries, basic mathematics has developed without making reference to experimental reality. The criterion driving mathematicians to new acquisitions was the *mathematical beauty*. Only a tiny part of such mathematical developments then found a role in physics. Tensor calculus was developed well before Einstein found a role for it in special and general relativity. Hilbert spaces found a role in quantum mechanics. Lie groups found a role in gauge theories. But there are plenty of other chapters of beautiful advanced mathematics which are, as yet, unexplored by physicists and may remain so forever.

There is, however, no question about that. Mathematics is an *intellectual* construction and its advancement is based on *intellectual* criteria. The problem arises when physicists begin to use similar criteria to put order in the physical world. Let us emphasize that this is not new in the history of research. The Pythagorean school, in ancient Greece, centered its teaching on mathematical beauty. They also found important physical results, e.g. in acoustics, starting from their criterion that the world should be a reflection of mathematical purity. In the ancient world, the views of Pythagoreans were then taken up by the whole Platonic school, in opposition to the Aristoteleans who thought that the world was ugly and complicated, so that attempting a quantitative description was in vain.

Even though we now believe that the final word has to be provided by the experimental data, there is no doubt that theoretical developments, often long and articulate, are grounded on mathematical beauty. This is true for any field of physics, of course, but the impact of such criteria in the quest for the origin is intellectually disturbing. What seems implicit in all this is that the human mind, for some obscure reason, although in a confused form, owns in itself the basic categories enabling it to distinguish the truth and to assert what is adherent to physical reality.

It is not our intention to take a stand on such points. However, we believe that they should be very present in the mind of all readers, when considering recent developments in basic physics and modern cosmology.