The universe cannot be a free lunch

G. Cavalleri (1) and S. Covino (2)
(1) Università Cattolica, Facoltà di Scienze - Brescia, Italy
(2) Università di Milano, Dipartimento di Fisica - Milano, Italy

(ricevuto l'8 Settembre 1995; approvato il 4 Marzo 1997)

Summary. — The total energy of the universe is definite-positive even for extremely relativistic matter. Hence its origin demands a violation of physical laws. The inflationary model which tries to justify an infinite age for the universe is criticized from both a theoretical and an observational point of view. Since all the other theories that tried to have a universe existing for ever in the past have been disproved, we may conclude that it is highly probable that the past life of the universe is finite.

PACS 98.80 – Cosmology.

1. – Introduction

The rather recent paper of Stenger's [1] is the clearest and simplest attempt to show that the universe is the ultimate free lunch. In fact, the author assumes that the universe could have begun from a state of zero energy and maximum entropy according to the inflationary model.

In the present paper we do not discuss the second law of thermodynamics because it is sufficient to show that the total energy of the universe is different from zero.

Moreover, we criticize the inflationary model which would imply the appearance of the universe from the quantum fluctuation of a vacuum with weird properties.

Consequently, we reaffirm the validity of common sense and of the relevant conclusions of nineteenth-century physics. The discovery of the cosmological expansion has reinforced Jeans’s idea that the past life of the universe if finite and has allowed us to quantify it: roughly 15 billion years.

2. – The total energy of the universe is positive-definite

One of the arguments used by the supporters of the “universe from nothing” is that its total energy be zero so that its appearance does not violate energy conservation. The energy of the vacuum would have been zero and such did the energy of the universe remain. Some authors were familiar with Cartan’s achievement that general relativity and Newton’s theory lead to the same result if applied to a uniform and isotropically
expanding universe. Consequently, they only considered the mutual bond energy per unit mass

$$E_{\text{mutual}}/m = -\frac{G}{r} \frac{4}{3} \pi q r^3 + \frac{v^2}{2},$$

which vanishes when the average mass density \( \overline{\rho} \) is equal to the critical value

$$\overline{\rho}_c = \frac{3H^2}{8\pi G},$$

where \( H \) is the Hubble constant and \( v = H r \) the cosmological recession speed.

These authors simply neglect the rest energy per unit rest mass \( E_{\text{rest}}/m_0 = c^2 \).

Stenger [1] is somewhat more refined but performs an analogous error. He considers an extremely relativistic matter but he confuses the cosmological recession speed \( v \) with the velocity \( v_p \) of the single relativistic particles.

Actually he takes the energy

$$E = mc^2 - \frac{8}{3} \pi m G \overline{\rho} r^2,$$

which with the critical mass density given by (2) yields

$$E = m(c^2 - H^2 r^2).$$

Now \( v = H r \) is the local average speed of recession for matter \( r \) apart from the observer and \( v \rightarrow 0 \) for \( r \rightarrow 0 \), even if the particles that produce \( \overline{\rho}_c \) have relativistic speed \( v_r \). Indeed \( v = H r \) is the vector average of the various \( v_p \), i.e.

$$v = \frac{1}{N} \sum_{s=1}^{N} m_s v_p,$$

Hence the various \( v_p \) may be relativistic although \( v \) may be very small. To conclude this first criticism, (4) tends to zero only at the horizon of the events and its integral over all space inside the horizon is positive.

This last result has been proved to be general in Einstein's gravitational theory. Actually, Schoen and Yau [2], in summarizing and improving their preceding work and that of other authors, have shown that the total energy is positive-definite in general relativity.

3. - Criticism to the theory of the inflationary universe

Any attempt, such as that of Stenger's [1], to have a universe appearing from vacuum fluctuations is based on the theory of the inflationary universe (with all its variants).

It is interesting to summarize all the hypotheses such an attempt needs.

1) The quantum vacuum must be present prior to the existence of the particles of the universe.

2) There must be tensile stresses on "nothing".
3) A repulsive cosmological constant term has to be present.

4) The quantum vacuum has to be always in rapid expansion in order to prevent the compenetration of the successive expanding universe bubbles.

Let us proceed to criticize all the four points.

1) The recent field of research called "stochastic electrodynamics" (SED) has explained all the linear (i.e. with forces proportional to their displacements from equilibrium positions) quantum phenomena by means of a stochastic electromagnetic (e.m.) field diffused all over the universe and having a power spectral density $\varphi$ given by

$$\varphi(\omega) = \hbar \omega^3 (2\pi^2 c^3)^{-1},$$

where $\omega$ is the angular frequency, $\hbar = h/2\pi$ and $c$ the speed of light [3, 4].

The power spectral density (6) is that of the zero-point field (ZPF) of quantum electrodynamics (QED). There are two differences between the ubiquitous, stochastic, e.m. field of SED and the ZPF of QED: i) It is sometimes considered as virtual in QED and always real in SED; ii) It has no upper cut-off in QED (the relevant infinite energy density is normalized to zero if the ZPF is considered as virtual) while it has an upper cut-off $\omega_c$ in SED in order to have a finite energy density.

The existence and value of $\omega_c$ as well as the explanation of nonlinear quantum phenomena are obtained if zitterbewegung (which is the solution of the Dirac equation for zero external field) is assumed as real, i.e. if the electron is assumed to move along a circumference at the speed of light $c$ with no special relativity (SR) [5]. An electron is considered at rest when the centre of the gyration orbit is at rest. An electron in motion means that the centre of gyration has a velocity $\mathbf{v}$ and the particle follows a helix. In common language, the electron consists of both the “true” particle moving with $c$ and the circumference (or the helix). SR is not present at the “subparticle” level but it arises because particle velocity is considered to be that of the gyration centre. By the above qualitative assumption the ZPF is generated by the gyration motion and the correct $\omega^3$-dependence appears [5]. Since the $\omega^3$ spectrum is Lorentz-invariant (it is the only one among all the possible $\varphi(\omega)$ spectra), SR is completed because the longitudinal contraction appears as well. Equating the radiated power to the absorbed power the gyration (or spin, or Compton) radius is determined and turns out to be given by [5]

$$R_s = \frac{3mcH(4\pi^2 27/4e^2N)^{-1}}{},$$

where $e$ and $m$ are the charge and mass of an electron, respectively, $N$ the average electron concentration in the universe, and $H$ the Hubble constant. Expressed in the usual terms it is

$$R_s = \frac{\hbar}{mc}.$$

The zitterbewegung implies a natural upper cut-off at the Compton frequency $\omega_c = c/R_s = mc^2/\hbar$. Actually, there is also the generation of the stochastic e.m. field by part of the quarks of the protons and neutrons with two other upper cut-offs so that the whole spectrum turns out to be [6]:

$$\varphi(\omega) = \frac{4}{3} \frac{N}{H} \frac{e^2}{c} \omega^3 \left[ \omega_s^{-2}\theta(\omega_s - \omega) + \omega_{\rho1}^{-2}\theta(\omega_{\rho1} - \omega) + \omega_{\rho2}^{-2}\theta(\omega_{\rho2} - \omega) \right],$$
where $\theta(x) = 1$ for $x > 0$ and $\theta(x) = 0$ for $x < 0$. However, the contribution of the quark motion at $\omega_4$ is almost negligible compared to the one due to electrons. The cut-offs are the maximum received frequencies due to nearby electrons and quarks. The lower part of $\varphi(\omega)$ is due to the radiation of distant electrons because of the Doppler-Fizeau effect. The $\omega^3$ spectrum is, in last analysis, due to the almost monochromatic frequencies of the electron and quark gyrations and to the expansion of the universe.

Once the $\omega^3$-dependence has been obtained SR appears. Indeed, the sizes of the atoms and their frequencies are obtained from the ZPF [4], and they are the same for two inertial observers $S$ and $S'$ in relative uniform motion, provided the coordinates of $S$ and $S'$ are connected via Lorentz transformations since the $\omega^3$-dependence is invariant together with its proportionality constant under Lorentz transforms.

The $\omega^3$-dependence together with the zitterbewegung leads to the Schrödinger equation [7], also including its coefficients [5]. It also leads to the explanation of the two-slit diffraction with electrons. It even predicts it if the slit through which the electrons pass is known [5,8].

An experiment is in progress to discriminate between this new interpretation and that of quantum mechanics (QM) which predicts no diffraction at all if the electron beam is much smaller than the size of a slit.

A single elementary particle in equilibrium (i.e. with $P_{\text{radiated}} = P_{\text{absorbed}}$) in the first region of frequency $0 < \omega < \omega_s$ characterized by (9) is called electron, in the second region $\omega_s < \omega < \omega_{p1}$ is called $\mu$ meson, and in the third region $\omega_{p1} < \omega < \omega_{p2}$ is called $\pi$ particle. The three families of matter are thus explained. The presence of discontinuities of $\varphi(\omega)$ as given by (9) does not affect usual phenomena since any oscillator of proper frequency $\omega_0$ probes a small interval of $\omega$ around $\omega_0$. The first discontinuity (at $\omega = \omega_s$) could be detected by accelerating an ion with atomic number $Z > 10$ in the synchrotron of the F. ermilab, and testing the proportionality between the communicated and the acquired energy up to a threshold depending on $Z$. Saturation should occur after this energy acquisition since the energy further acquired should be absorbed by the ZPF, i.e. a friction arises in vacuo after the $Z$-dependent threshold [9]. The detection of this predicted effect demands an ion injector and one of the most powerful synchrotrons.

To conclude, this new field of research, in which a gyration at the speed of light of the elementary particles causes both SR and QED, implies some experiments which can discriminate this new point of view from SR and QED by putting limits to their validity. Moreover, $\hbar$ is derived from cosmology [5]. Using the new relation in a reversed way allows an improvement for the accuracy of the Hubble constant and of the average baryonic density of the universe. Well, this new interpretation of SR and QED is incompatible with the theory of the inflationary universe. Actually, the quantum vacuum is, in the new theory, nothing but the electromagnetic radiation generated by the gyration (or spin) rotation of all the particles of the universe. Consequently, the quantum vacuum could not exist before the creation of the elementary particles.

2) There must be tensile stresses in vacuo, i.e. traction forces between “nothing”. This assumption is much harder to be conceived than the direct postulate of spontaneous self-creation from nothing.

3) A repulsive cosmological term (at variance with point 2) has to be present. Consequently, the vacuum should consist of two fluids, one with tensile stresses (due to assumption 2) and the other with compressive stresses. Apart from this absurdity, let us discuss the plausibility of a positive cosmological term $\Lambda$. 


A positive $\Lambda$ would be the consequence of a negative energy density. But, as discussed in sect. 2, the total energy of the universe is positive-definite. This implies a positive-definite energy density if the universe is homogeneous. On the other hand, an ad hoc introduction of $\Lambda$ as done and then refused by Einstein is unphysical. This can clearly be seen in the alternative approach to gravity started by Pauli and Thirring, developed by Sexl, Belifante and Deser and brought to completion by Cavalleri and Spinelli [10]. In this alternative approach gravitation is treated, as any other field of physics, in the flat pseudo-Euclidean space-time. The gravitational field shortens the lengths of material bodies and slows down the real clocks so that the geometry measured by these instruments affected by gravity turns out to be curved. In precise terms, if the rest masses are assumed to be unaffected by gravity the geometry becomes Riemaniann and Einstein’s theory is obtained. On the contrary, if one lets the rest masses depend on gravity a Finsler space results. However, even in this case we may practically recover the results of general relativity if we assume that all the elementary particles have a zitterbewegung such as that of an electron in Dirac’s theory [10, 11]. Well, in this field-theoretic approach to gravity, a cosmological constant can only arise because of diffused matter whose energy is always positive so that $\Lambda$ can only be negative.

4) According to the inflationary model, the universe enucleated from the quantum vacuum because of instability. Consequently, infinite universe bubbles should have arisen in the past if the quantum vacuum is assumed to have existed for ever. With a quantum vacuum in relative rest all the expanding universe bubbles would have superposed and the matter density would be infinite. To avoid this Olber paradox for matter, the quantum vacuum has to be assumed in rapid expansion (in order to avoid the superposition of the various expanding bubbles). In spite of its expansion, the quantum vacuum has to keep its properties unaltered without any attenuation due to its increasing rarefaction. A kind of continuous creation of the quantum vacuum is demanded and this requirement is no less difficult to be accepted than the creation of the universe ex nihilo.

So far, we have examined the assumption of the theory of the inflationary universe, which has the only merit of explaining in a qualitative way the flatness and the horizon problems without any fine tuning (other explanations of these two problems are, however, possible [6,12]). Let us now discuss the wrong predictions of the theory, taking into account that a single disagreement with well-proved experiments (or observations) is sufficient to rule out a theory (see, in particular, Popper’s confutation method).

1) The huge cosmological constant $\Lambda$ necessary to have the extremely short inflationary period can be damped during the energy release in the form of present matter. However, when the false vacuum is re-established $\Lambda$ should acquire its initial value. Many attempts to have a fine-tuning compensation with the other fields have failed. The ratio between the $\Lambda_p$ predicted, and the maximum $\Lambda_M$ value allowed by observations is $\Lambda_p/\Lambda_M = 10^{105}$. Never in the history of physics an error of this magnitude has been heard of. A single wrong order of magnitude is sufficient to consider a theory deprived of physical meaning.
2) Quantum fluctuations had to be present in the inflation period and this demand gave hope to account for the initial inhomogeneities necessary to explain the observed structure of the universe. However, the predicted inhomogeneities in the background microwave radiation at 2.73 K produced by all realistic inflationary models are too large [13]. Smoot [14], actually, reported that cosmic background radiation anisotropies on all angular scales \( \lesssim 7^\circ \) to \( \delta T / T \lesssim 10^{-5} \) are consistent with a universe described by Robertson-Walker metric, inflationary models, gravitational instability and show no positive evidence of anisotropic expansion, rotation, or defects (strings, walls, texture). However, COBE DMR data does not necessarily confirm it. As shown by Vittorio [15], anisotropy data can rule out Cold Dark Matter models with \( \Omega \approx 1 \) and spectral index \( n \approx 1 \). It seems they do not provide a viable scenario for the large-scale structure of the universe.

3) Although the inflation may lead to any value of \( \Omega = \rho / \rho_{cr} \) [16], in order to solve the above problems the inflation has to be complete, thus leading to \( \Omega = 1 \), otherwise a new kind of “fine tuning” is still required. Otherwise, one can introduce a second scalar field (as done by Linde [17]) and obtain any \( \Omega \) value without a fine tuning. But the fine tuning is simply masked by the second arbitrary scalar field that appears as a new phenomenological parameter added ad hoc and without any physical basis. Now by \( \Omega = 1 \) there is no possibility of agreement between the age of the universe obtained by the Hubble constant, the value obtained from the star evolution, and that derivable from the decay of the radioactive elements. Let us discuss these statements in detail. The value of the Hubble constant is still somewhat uncertain. According to various measurements it is in the range \( \sim 40-90 \) km s\(^{-1}\) Mpc\(^{-1}\) [18] and some of the most recent results yield a value \( 73 \pm 10 \) km s\(^{-1}\) Mpc\(^{-1}\) [19]. This value was obtained by the analyses of a set of galaxies observed by the Hubble Space Telescope and is based on a variety of methods, including a Cepheid calibration of the Tully-Fisher relation, type-Ia supernovae, a calibration of distant cluster tied to others closer to us, and direct Cepheid distances out to \( \sim 20 \) Mpc. If we take the average value \( 73 \) km s\(^{-1}\) Mpc\(^{-1}\), the corresponding Hubble time is \( T_H = H^{-1} = 13.7 \) Gyr. Since complete inflation demands \( \Omega = \rho / \rho_{cr} = 1 \), the universe age should be \( T_u = 2T_H / 3 \approx 9.1 \) Gyr, much shorter than the age of the oldest stars. Even if we take the smallest \( H \) value (\( H \sim 73 \)) the corresponding Hubble time is \( T_H \approx 15.8 \) Gyr and, according to full inflation, \( T_u = T_H / 1.5 \approx 10.5 \) Gyr, would be still less than the age of the oldest stars. Taking into account that the relative abundance of helium is \( 0.232 \pm 0.003 \pm 0.0005 \) [20], the corresponding average baryonic density turns out to be \( \rho = (0.015-0.045) \rho_c \) (\( H = 73 \)) [21]. This value is inside the uncertainty interval obtainable from the relative abundances of \( ^3\)D, \( ^3\)He, \( ^7\)Li, \( ^4\)He and is in agreement with the minimum value derivable from the free fall of the Local Group toward the centre of the Virgo cluster \( 0.042 \rho_c < \rho < 0.085 \rho_c \) [22]. Since the attraction of the Virgo cluster is due to all kinds of masses, the hypothetical nonbaryonic dark matter could at maximum be equal to the baryon mass. The corresponding value of the deceleration parameter \( q = \Omega / 2 = \rho / 2 \rho_c \approx 0.02 \) is so small that the universe age \( T_u \) should be slightly less than the Hubble time \( T_H = H^{-1} \). Consequently, \( T_u \) can be somewhat larger than the age of the oldest stars and of the radioactive elements. If we take the average value \( H \approx 73 \) which gives \( T_H \approx 14 \) Gyr this agreement is impossible if \( \Omega = 1 \) because \( T_u = 2T_H / 3 \approx 9 \) Gyr.

Let us examine separately the two age constraints, the one due to the oldest stars (contained in the globular clusters) (i) and the one due to nuclear chronometers (ii).
i) Galactic globular clusters are the oldest stellar systems with rather reliably
determined ages. The two methods for absolute cluster dating that are currently most
commonly used involve the direct main-sequence (MS) fitting of observed colour-
magnitude diagrams (CMD) to theoretical isochrones, and the use of a calibration
of the magnitude difference between horizontal-branch (HB) stars and the MS turn-off,
at the colour of the turn-off. The MS phase is the longest of stellar evolution,
characterized by stars still burning hydrogen in their cores. The HB stars, instead, are
in the core helium-burning stage. For old stellar population the HB luminosity is
weakly dependent on age while the luminosity of the brightest stars on the MS
progressively reduces as the population ages. The turn-off is a point on a CMD
populated by stars that are exhausting their core hydrogen content and are leaving the
MS toward their following evolutionary stages. Vandenberg [23] found 15–18 Gyr and
Janes and Demarque [24] found 14–19 Gyr from isochrone fitting, while Iben and
Renzini [25] found 16 ± 3.5 Gyr and Vandenberg [26] found 15–20 Gyr from the
magnitude difference. The situation is complicated by the possible existence of large
variations in the chemical composition of elements heavier than helium (“metals”)
among globular clusters. Lee et al. [27] suggested that the metal-rich clusters may be
younger than the metal-poor clusters by several billion years. Sandage and
Cacciari [28] found a mean age for their cluster sample of 17 Gyr, for [O/Fe] = 0, and a
lower mean age of 14 Gyr, for the choice [O/Fe] = + 0.6. [O/Fe] is a short form for the
decimal logarithm of the ratio of the abundances of oxygen and iron nuclei with respect
to the analogous solar abundance. Similarly, from main-sequence fitting they obtained
19 Gyr, for no oxygen enhancement, and 15.5 Gyr for the oxygen enhanced.
In any case, a spread of ~3–4 Gyr seems appropriate, with a mean age of
≈14–17 Gyr [29].
In spite of common belief, the ages of the globular clusters estimated by theoretical
models have at most an uncertainty of ~15% [30]. The reported error bars are
substantially due to observational uncertainties regarding the chemical composition
and/or the distance moduli.

ii) Nucleocosmochronology is the use of the relative abundances of radioactive
nuclear species and their radiogenic decay daughters to establish the finite age of the
elements and time scale for their formation [31]. Nuclear chronometers also give the
age of the Solar System, 4.6 Gyr, with great precision [32]. Some chronometric pairs
(²³²Th-²³⁸U, ²³⁵U-²³⁸U and ¹⁸⁷Re-¹⁸⁷Os) lead to a mean age prior to the condensation of
the Solar System. However, due to both observational uncertainties in their relative
abundances, and to the necessity of relying on highly speculative and poorly
constrained models for the Galactic history of nucleosynthesis, the indicated age of the
universe is quite uncertain. Cowan et al. [33] found the Galactic age, t_G, in the range
12.4–14.7 Gyr with a simple model for the initial Galactic enrichment, while Thielemann
et al. [34], with different choices on physical parameters regulating the productions of
the chronometric elements involved, proposed an age of 20.8±2 Gyr. Clayton [35], with a
more complex treatment of the initial enrichment, determined t_G = 20 Gyr in
agreement with the range deduced by Lawler et al. [36] of 15–20 Gyr. Taking into
account the various uncertainties still involved in many aspects of nucleocosmo-
chronology, Cowan et al. [31] quoted that the nucleocosmochronology estimated
Galactic age can only be constrained to be in the range 10–20 Gyr.
There is a good agreement between the ages of the oldest stars i) (14–17 Gyr) and
the age of formation of the radioactive nuclear species ii) (10–20 Gyr and, preferably,
15–20 Gyr). Since the universe age $T_U$ has to be somewhat larger, it must be $T_U > 14$ Gyr.

4. - Conclusions

Stenger [1] has re-emphasized that the universe has zero total rest energy and could therefore have emerged from vacuum because of a quantum fluctuation. That is why he repeated the notorious sentence: "the universe is the ultimate free lunch". We have shown in sect. 2 that the total energy of the universe is positive-definite for both non-relativistic and relativistic particles. A violation of physical laws is therefore demanded for its appearing into existence.

In sect. 3 we have criticized both the assumptions and the predictions of the inflationary theory which gives the hypothetical mechanism according to which the universe appears from "nothing". The assumptions are criticized on the conceptual and qualitative level while the predictions, on the quantitative level.

The first hypothesis (a quantum vacuum pre-existing the real particles) is shown to be an exchange of the cause for the effect. Actually, stochastic electrodynamics (SED) implemented by the spin (or gyration) motion explains all the quantum phenomena, and even new effects, because of the random electromagnetic field (or zero-point field: ZPF) radiated by the gyration motion of all the particles of the universe. The ZPF is nothing but the quantum vacuum which therefore appeared after the existence of the particles.

The second hypothesis (tensile stresses on nothing) is much harder to accept and perhaps more difficult to conceive than to have a creation ex nihilo.

Similar difficulties hold for the third hypothesis (the existence of a positive, or repulsive, cosmological term) which is unphysical, as shown by the flat space-time approach to gravitation [10], and the fourth hypothesis (a quantum vacuum in rapid expansion but always recreating its properties). Actually, a continuous creation of the negative energy and of the stresses in vacuo (hence on nothing) is demanded otherwise its rapid expansion would imply a weakening of its properties.

There are only two qualitative right predictions of the inflationary theory: it solves the horizon and the flatness problems. The latter is not, however, solved for its order of magnitude but a "fine tuning" is required if, as turns out in a way progressively more convincing, $\Omega = \rho/\rho_c \simeq 0.02$. Actually, the $\Omega = 1$ prediction of the theory has to be considered as a wrong quantitative prediction since the free fall of the Local Group toward Virgo cluster depends on all kinds of matter and the corresponding $\Omega$ value is $0.042 < \Omega < 0.085$. Moreover, only by a deceleration parameter $q = \Omega/2 \simeq 0.015$ may we have agreement between the age of the universe obtained from the Hubble constant (whose mean recent value is $\approx 73$ km s$^{-1}$ Mpc$^{-1}$ corresponding to $T_H \approx 14$ Gyr) and the one obtainable from the stellar evolution applied to the globular clusters and the one derivable from the decay of the radioactive elements ($T_U > 14$ Gyr). With $\Omega = 1$ it would be $T_U = 2T_H/3 \approx 9$ Gyr, hence shorter than the other two ages which demand $T_U > 14$ Gyr. The second wrong prediction regards the intensity and the angular spread of the density fluctuations. But it is the first wrong prediction considered in sect. 3 that is astonishing: a cosmological constant $\Lambda$ wrong by 105 orders of magnitudes!

To conclude, all the attempts to have a universe existing from eternity have failed. The inflationary theory is the last and the most absurd of these attempts. It is
therefore highly probable that the universe has a finite past life. Moreover, as shown in sect. 2, its total energy is positive-definite and its coming into existence demands a violation of energy conservation.

***

We thank Prof. S. Bonometto for useful discussions during the development of this work. We also declare, according to the Italian law, that the realization of this paper is equally shared by the authors.

REFERENCES