Physics Letters B 749 (2015) 514–518

Contents lists available at ScienceDirect

Physics Letters B

www.elsevier.com/locate/physletb

A note on Boltzmann brains

Yasunori Nomura *

Berkeley Center for Theoretical Physics, Department of Physics, University of California, Berkeley, CA 94720, USA
Theoretical Physics Group, Lawrence Berkeley National Laboratory, CA 94720, USA
Kavli Institute for the Physics and Mathematics of the Universe (WPI), Todai Institutes for Advanced Study, University of Tokyo, Kashiwa 277-8583, Japan

ARTICLE INFO

Article history:
Received 11 August 2015
Accepted 12 August 2015
Available online 14 August 2015
Editor: M. Cvetič

ABSTRACT

Understanding the observed arrow of time is equivalent, under general assumptions, to explaining why Boltzmann brains do not overwhelm ordinary observers. It is usually thought that this provides a condition on the decay rate of every cosmologically accessible de Sitter vacuum, and that this condition is determined by the production rate of Boltzmann brains calculated using semiclassical theory built on each such vacuum. We argue, based on a recently developed picture of microscopic quantum gravitational degrees of freedom, that this thinking needs to be modified. In particular, depending on the structure of the fundamental theory, the decay rate of a de Sitter vacuum may not have to satisfy any condition except for the one imposed by the Poincaré recurrence. The framework discussed here also addresses the question of whether a Minkowski vacuum may produce Boltzmann brains.

1. Introduction

At first sight, the fact that we observe that time flows only in one direction may seem mysterious, given that the fundamental laws of physics are invariant under reversing the orientation of time.[1] Upon careful consideration, however, one notices that the problem is not the unidirectional nature *per se*. As discussed in Refs. [1,2], given any final state $|f\rangle$ whose coarse-grained entropy is lower than the initial state $|i\rangle$, the evolution history is overwhelmingly dominated by the *CPT* conjugate of the standard (entropy increasing) process $|\bar{f}\rangle \to |\bar{\imath}\rangle$. This implies that a physical observer, who is necessarily a part of the whole system, sees virtually always, i.e. with an overwhelmingly high probability, that time flows from the "past" (in which correlations of the observer with the rest of the system are smaller) to the "future" (in which the correlations are larger).

The problem of the arrow of time, therefore, is not to understand its unidirectional nature, but to explain why physical predictions are (probabilistically) dominated by what we observe in our universe, i.e. a flow from a very low coarse-grained entropy state to a slightly higher entropy state. In particular, it requires the understanding of the following facts:

- At least one set of states representing our observations, which are mutually related by time evolution spanning the observation time, are realized in the quantum state representing the whole universe/multiverse. (Here and below we adopt the Schrödinger picture.) This is the case despite the fact that these states have very low coarse-grained entropies.[2]
- The answer to a physical question, which may always be asked in the form of a conditional probability [6], must be determined by the class of low coarse-grained entropy states described above. In particular, the probability should not (always) be dominated by the states in which the unconditioned part of the system has the highest coarse-grained entropies.

* Correspondence to: Berkeley Center for Theoretical Physics, Department of Physics, University of California, Berkeley, CA 94720, USA.
E-mail address: ynomura@berkeley.edu.

[1] The operation discussed here is not what is called the time reversal T in quantum field theory, which we know is broken in nature. It corresponds to *CPT* in the standard language of quantum field theory.

[2] Because of the Hamiltonian constraint, the full universe/multiverse state is expected to be static, i.e. not to depend on any time parameter [3,4]. We may, however, talk about effective time evolution if we focus on branches of the whole universe/multiverse state, since they are not (necessarily) invariant under the action of the time evolution operator $e^{-iH\tau}$. This is the picture we adopt in this paper. Note that this time evolution still does not have to be the same as "physical time evolution" defined through correlations among physical subsystems, e.g., as in Ref. [5]. In the static-state picture, the statement here is phrased such that the state of the universe/multiverse contains components representing our observations despite the fact that they are not generic in the relevant Hilbert space.

http://dx.doi.org/10.1016/j.physletb.2015.08.029

These elements comprise (essentially) the well-known Boltzmann brain problem [7,8]. The problem of the arrow of time is thus *equivalent* to the Boltzmann brain problem [4] under (rather general) assumptions that went into the line of argument given above.

Any realistic cosmology must accommodate the two facts listed above. Is it trivial to do so? In a seminal paper [7], Dyson, Kleban, and Susskind pointed out that it is not. In particular, they considered a de Sitter vacuum representing our own universe and argued that if it lives long enough, thermal fluctuations in de Sitter space lead to Boltzmann brains observing chaotic worlds, who overwhelm ordinary, ordered observers like us. If true, this would give an upper bound on the lifetime of our universe which is much stronger than that needed to avoid the Poincaré recurrence (barring the possibility that the observed vacuum energy relaxes into a zero or negative value in the future). In this paper we argue that this consideration needs to be modified, based on the picture of the microscopic structure of quantum gravity advanced recently [9] to address the black hole information problem [10,11]. We discuss implications of this modification for our own universe and the eternally inflating multiverse. We also discuss implications of the framework for the possibility [12] of Boltzmann brain production in a Minkowski vacuum.

2. de Sitter space in quantum gravity

We first extend the discussion of Ref. [9], which mainly focused on a system with a black hole, to de Sitter space. In cosmology, de Sitter space appears as a meta-stable state in the middle of the evolution of the universe/multiverse, and in this sense it is similar to a spacetime with a dynamically formed black hole. Indeed, string theory suggests that there is no absolutely stable de Sitter vacuum in full quantum gravity; it must decay, at least, before the Poincaré recurrence time [13]. This implies that what we call de Sitter space cannot be an eigenstate of energy (at least in this context).

Consider a semiclassical de Sitter space with Hubble radius α. (We focus on 4-dimensional spacetime for simplicity, but the extension to other dimensions is straightforward.) Following the complementarity hypothesis [14], and in particular its implementation in Refs. [2,6], we adopt a "local description," in which quantum states represent physical configurations on equal-time hypersurfaces foliating the causal patch associated with a freely falling frame. We assume that the timescale for the evolution of microstates representing the de Sitter space is of order $\Delta t \approx \alpha$, where t is the proper time measured at the spatial origin, p_0, of the reference frame. The uncertainty principle then implies that a state representing this space must involve a superposition of energy eigenstates with a spread of order $\Delta E \approx 1/\alpha$. Associating this energy with the vacuum energy density ρ_Λ integrated over the Hubble volume, $E \approx O(\rho_\Lambda \alpha^3) \approx O(\alpha/l_{\rm P}^2)$, this spread is translated into $\Delta\alpha \approx O(l_{\rm P}^2/\alpha)$, where $l_{\rm P}$ is the Planck length.

How many different independent ways are there to superpose the energy eigenstates to arrive at the semiclassical de Sitter space described above? As in the black hole case, we assume that the Gibbons–Hawking entropy [15]

$$S_{\rm GH} = \frac{\mathcal{A}}{4l_{\rm P}^2} = \frac{\pi\alpha^2}{l_{\rm P}^2}, \tag{1}$$

gives the logarithm of this number (at the leading order in expansion in inverse powers of $\mathcal{A}/l_{\rm P}^2$), where $\mathcal{A} = 4\pi\alpha^2$ is the area of the de Sitter horizon [16]. In particular, there are exponentially many independent de Sitter *vacuum* states—the states that do not have a field or string theoretic excitation in the semiclassical background—which all represent the same de Sitter vacuum at the semiclassical level.

The analysis of physics in this de Sitter vacuum is parallel to that on a black hole background in Ref. [9]. Denoting the index representing the exponentially many de Sitter vacuum states by

$$k = 1, \dots, e^{S_0}, \tag{2}$$

where $|S_0 - S_{\rm GH}| \approx O(\mathcal{A}^q/l_{\rm P}^{2q};\, q < 1)$, states at late times on this vacuum can be expanded in terms of the microstates of the form

$$|\Psi_{\bar{a}a;k}(\alpha)\rangle. \tag{3}$$

Here, $\bar{a}$ and a label excitations of the stretched horizon, located at $r = \alpha - O(l_{\rm P}^2/\alpha) \equiv r_{\rm s}$, and the interior region, $r < r_{\rm s}$, respectively, where r is the static radial coordinate with $r = 0$ taken at p_0. Note that excitations here are defined as fluctuations with respect to a fixed background, so their energies as well as entropies can be either positive or negative, although their signs must be the same. The contribution of the excitations to the entropy is subdominant in the $l_{\rm P}^2/\mathcal{A}$ expansion, so that the total entropy of this de Sitter system (not necessarily of the vacuum states) is still given by $S = \mathcal{A}/4l_{\rm P}^2$ at the leading order.

The indices for the excitations, $\bar{a}$ and a, and the vacuum, k, do not fully "decouple". In particular, operators in the semiclassical theory representing modes whose energies defined at $r = 0$ are

$$\omega \lesssim T_{\rm GH}, \tag{4}$$

act nontrivially on both a and k indices, where $T_{\rm GH} = 1/2\pi\alpha$ is the Gibbons–Hawking temperature. This allows us to understand the thermal nature of the semiclassical de Sitter space in the following manner. The fact that all the independent microstates with different k lead to the same geometry (within the quantum mechanical uncertainty) suggests that the semiclassical picture is obtained after coarse-graining the degrees of freedom represented by this index, which we call the *vacuum degrees of freedom*. In this picture, the de Sitter vacuum in the semiclassical description is represented by the density matrix

$$\rho_0(\alpha) = \frac{1}{e^{S_0}} \sum_{k=1}^{e^{S_0}} |\Psi_{\bar{a}=a=0;k}(\alpha)\rangle\langle\Psi_{\bar{a}=a=0;k}(\alpha)|. \tag{5}$$

To obtain the response of this state to the operators in the semiclassical theory, we may trace out the subsystem $\bar{C}$ on which they do not act. Consistently with our identification of the origin of the Gibbons–Hawking entropy, we identify the resulting reduced density matrix as the thermal density matrix

$$\tilde{\rho}_0(\alpha) = {\rm Tr}_{\bar{C}}\, \rho_0(\alpha) \approx \frac{1}{Z} e^{-\frac{H_{\rm sc}(\alpha)}{T_{\rm GH}}}, \tag{6}$$

where $Z = {\rm Tr}\, e^{-H_{\rm sc}(\alpha)/T_{\rm GH}}$, and $H_{\rm sc}(\alpha)$ is the Hamiltonian of the semiclassical theory.

Another manifestation of the non-decoupling nature of the a and k indices is that for states having a negative energy excitation, the range over which k runs is smaller than that in Eq. (2)—this is the meaning that a negative energy excitation carries a negative entropy. As discussed in the next section, this fact is important in ensuring unitarity in the process in which a physical detector held at constant r is excited due to interactions with the de Sitter spacetime.

3. Vacuum degrees of freedom

The expression in Eq. (6) implies that the spatial distribution of the information in k follows the thermal entropy calculated using the local temperature