RESEARCH ARTICLE

George F. R. Ellis · **William R. Stoeger**

A note on infinities in eternal inflation

Abstract In some well-known scenarios of open-universe eternal inflation, developed by Vilenkin and co-workers, a large number of universes nucleate and thermalize within the eternally inflating mega-universe. According to the proposal, each universe nucleates at a point, and therefore the boundary of the nucleated universe is a space-like surface nearly coincident with the future light cone emanating from the point of nucleation, all points of which have the same proper-time. This leads the authors to conclude that at the proper-time $t = t_{\text{nuc}}$ at which any such nucleation occurs, an infinite open universe comes into existence. We point out that this is due entirely to the supposition of the nucleation occurring at a single point, which in light of quantum cosmology seems difficult to support. Even an infinitesimal space-like length at the moment of nucleation gives a rather different result—the boundary of the nucleating universe evolves in proper time and becomes infinite only in an infinite time. The alleged infinity is never attained at any finite time.

Keywords Inflation, Infinity

1 Introduction

It is commonly stated that infinities can occur in inflationary cosmology in two ways. Firstly, it is stated that quantum tunnelling can lead from a universe with finite
spatial sections to an imbedded expanding universe domain with infinite spatial sections. Secondly, it is claimed this will occur an infinite number of times: "most important of all is that once inflation starts it produces not just one universe, but an infinite number of universes" [6]. Thus what is proposed is the existence of

G. F. R. Ellis W. R. Stoeger University of Cape Town Rondebosch Cape Town 7701, South Africa george.ellis@uct.ac.za · W. R. Stoeger Vatican Observatory Research Group Steward Observatory, The University of Arizona Tucson, AZ 85721, USA wstoeger@as.arizona.edu

an infinite number of expanding universe domains like the one we see around us, but each with somewhat different characteristics, all imbedded in a surrounding inflationary universe; and each with infinite spatial sections.

The way this is stated implies, implicitly or explicitly, that the infinite spatial sections come into being instantaneously in each bubble; as soon as soon as a bubble nucleates in an inflating universe, negatively curved infinite spatial sections come into existence at that instant. This claim, which is repeated and provides the basis for further philosophical speculations in later work (e.g. [7; 9; 11]), is based essentially on identifying the future light cone of the point of nucleation—or more precisely, an exactly constant proper time surface very close to that light cone—as the effective boundary of the bubble universe. As such it is a surface of infinite extent that is everywhere at the same space–time distance from the origin, with a foliation of spacelike surfaces of constant negative curvature and infinite extent asymptotic to this light cone in its future. However, this conclusion is the result of assuming that the bubble universe originates at a single point, instead of as a region with some extension, no matter how small. When we substitute this alternative, the instantaneous emergence of the infinite bubble boundary is no longer apparent. We are left with a finite boundary which expands forever—but only becomes infinite in an infinite time, and hence is never in fact attained.

Thus in both cases, there is an ongoing process: at any finite time there never exists the claimed infinite number; it is what the situation tends to but never attains. The universe is evolving towards such a state, but never reaches it; for that is the essential nature of infinity.

2 Origination of infinite bubble universes: the null cone

In their development of open-universe inflation, Vilenkin and Winitzki [12] consider a mega-universe dominated by false vacuum inflating with an approximately de Sitter metric

$$ds^2 = -dt^2 + \exp(2H_0 t)(dr^2 + r^2 d\Omega^2), \tag{1}$$

where $H_0 = \sqrt{2\pi V_0/3}$, V_0 being the false vacuum potential. Within this inflating mega-verse, spherical bubbles of true vacuum nucleate to form daughter universes which quickly thermalize and evolve separately as Friedmann–Lemaître–Robertson–Walker (FLRW) universes. If we focus on one case of cosmic nucleation—at the moment of nucleation—at a particular point in time and in space, say at $t = 0$ and $r = r_0$—a bubble of true vacuum is formed, which looks like an open FLRW universe with

$$ds^s = d\tau^2 + a^2(\tau)(d\xi^2 + \sinh^2\xi d\Omega^2), \tag{2}$$

where for small values of τ, setting nucleation at $\tau = t = 0$, the bubble has yet to thermalize and is still inflating. So it is very close to de Sitter, with

$$a(\tau) = H_0^{-1}\sinh(H_0\tau). \tag{3}$$

Vilenkin and Winitzki [12] carefully relate the coordinates of the metrics, discuss the conditions for continued brief acceleration of the bubble universe and for its

thermalization. Similar analyses are given by other writers, for example Freivogel et al. [5], but in that case with tunnelling from the compact $k = +1$ form of the de Sitter universe.

The argument for at infinite bubble universe originating at a point $t = \tau = 0$ with point-like radial position $r = r_0$ is straightforward. Since it has no spatial extension, from (1) $ds^2 = -dt^2$, and the proper interval s there can be identified with $t = 0$. So $s = t = 0$ at the origin. But the whole future light cone emanating from there is at the same proper distance from the origin, since it is given by $ds = 0$. And that surface extends to infinity. In a definite sense then, that whole surface comes into existence instantaneously at the instant $t = 0$. Vilenkin and Winitzki [12] show that the boundary of the bubble universe is in this case a constant proper-time spacelike surface infinitesimally close to that future light cone. Thus, an infinite bubble universe comes into existence at that moment. Spacelike surfaces of constant proper time occur within this boundary, each at a constant proper time from the origin; thus after say a finite time τ_1, no matter how small, infinite spacelike surfaces exist with every point produced through physical processes taking the identical time τ_1 since nucleation, no matter how far from the origin they are. An actual infinite spatial extent comes into existence in a finite time, originating from a universe with compact spatial sections—a truly remarkable claim, based on the nature of the hyperbolic geometry of special and general relativity theory.

3 Spatially extended finite alternative

Is this claim that an infinite bubble universe emerges instantaneously at $t = \tau = 0$ supportable? As emphasized above, this rests on the supposition that the surface of the bubble universe can be identified with the future light cone emanating from that point, which requires that it originate precisely at that point—without spatial extension. However, this supposition seems somewhat dubious from the point of view of quantum cosmology, taking into account the fact that such transitions really take place within finite spatial volumes—not at idealized points. Indeed according to Freivogel et al. [5], quantum tunnelling leads to a domain of finite radius coming into existence instantaneously. This is quite different than when it is supposed to occur at a point-like event.

So, what happens when we suppose, instead, that the bubble universe originates in a volume with some finite extension? Let's simply represent it as Δr^2. Then, from the point of view of the de Sitter mega-universe metric, we see that that originating surface is not at the same proper distance form the origin, since

$$\Delta s^2 = -\Delta t^2 + \exp(2H_0 t)\Delta r^2.$$

Clearly $t = \tau = 0$ cannot now be identified with the proper-interval of the surface of the originating bubble, and, as that surface evolves forward in time, expanding rapidly by inflation at least for a short while before thermalization, its boundary will not have the same proper distance from every part of that originating surface at $t = \tau = 0$. In fact, the proper time on that boundary measured from the various spatial points in the nucleated bubble, will diverge. There is then no support for considering it instantaneously infinite—it began finite, and continues to be finite

at every physically attainable time. It approaches infinity in an infinite time—but that is never actually reached or completed, in any real sense of that word.

Thus, the interpretation of an infinite $t = 0$ constant initial surface of a bubble universe coming into existence at that time within an eternally inflating mega-universe appears to be an artifact of treating its origination in a unwarranted idealized fashion—at a single space–time point. This is an example of how careful we need to be in drawing conclusions from idealizations which involve unextended points and the infinities which are connected with them.

4 Spacelike Surfaces and the emerging universe

The key underlying point is that the whole space–time does not come into being instantaneously: things take place physically, as events unroll along particle world lines (particles tunnel, scalar field roll down potential surfaces, particles collide, etc.). The outcome is determined as it happens (because quantum events are involved), determining both what happens in the space–time and (because space–time curvature is determined by the matter in the space–time) even the space–time structure itself. We do not in the real world have a block universe that instantaneously comes into being; rather we have an Evolving Block Universe that unrolls over time [2].

But which spacelike surfaces are the relevant ones for this process? Different choices of time are represented by different choices of spacelike surfaces of constant time, as depicted in Fig. 1 for the case of Minkowski (flat) space–time, which adequately illustrates the main points to be made here. Spacelike surfaces $S_0 = constant$ are at constant proper time from the origin of coordinates. In effect, the suggestion made by Vilenkin et al. is that we should take seriously these surfaces of constant distance from the origin rather than the surfaces of constant Minkowski time T.

How to choose between them? Well, in fact [2] events normally unroll along timelike or null world lines, rather than being based on spacelike surfaces, convenient as these are for setting up coordinate systems. Every point in each surface $S_0 = constant$ is at the same proper time from the origin of coordinates. So one can imagine a physical process that takes a proper time S_0 to develop after the initiation event at O; then the outcome will occur simultaneously at every point on the surface $S_0 = $ constant, no matter how far away in spatial terms. Thus one can claim that infinitely far away events (in spacelike terms) occur simultaneously with the nearest ones, in these surfaces. This is the rationale for claiming infinities of spatial events actually exist immediately after the tunnelling; for this argument holds no matter how small S_0 is.

It is in this context that it matters that the causal origin of the physical events after tunnelling is of finite extent, see Fig. 2 of [5] for this context. It is as a result of this finite size that infinities never exist at any finite time after the tunnelling event. We consider this first in the paradigmatic case of Minkowski space–time, and then in the context of general bifurcate Killing horizons, such as occur for example in de Sitter space–time.

Fig. 1 Constant time surfaces in Minkowski space–time. The surface S_0 of constant proper time from the origin of coordinates O lies inside the future light cone C of O, and is a spacelike surface of constant negative curvature. For events Q and R on that surface, there are alternative surfaces of constant time, such as the surface of constant Minkowski time T. The event R lies at a later Minkowski time than the event Q, even though both are at the same proper time from O

Fig. 2 Space–time distances in Minkowski space–time. The surface S_0 of constant proper time T_0 from the origin of coordinates O is not a surface of constant time from the event P. The proper time from O to R is the same as the proper time from O to Q; but the proper time from P to R is greater than the proper time from P to Q (n.b.: because this is a hyperbolic geometry, that statement is not obvious from the diagram)

4.1 Paradigmatic example: Minkowski space–time

Proper time from the origin O to a point Q on the surface at constant proper time from O in Minkowski space–time is by definition of that surface the same as the proper time from the origin O to any other point R on the surface (See Fig. 2), even though the first corresponds to a smaller Minkowski time T than second. However the proper time from P to R is greater than the proper time from P to Q. To show this in detail: let O be at $(T_O, X_O) = (0,0)$ and P be at $(T_P, X_P) = (0, \varepsilon)$ with $\varepsilon > 0$. The surface S_0 is given by

$$-T^2 + X^2 = -S_0^2. \tag{4}$$

Thus Q is at (T_Q, X_Q) where $X_Q = +\sqrt{T_Q^2 - S_0^2}$. The space–time distance from P to Q is given by

$$-\tau_{PQ}^2 = -(T_P - T_Q)^2 + (X_P - X_Q)^2 \tag{5}$$

Substituting for (T_P, X_P, X_Q), (5) becomes

$$\tau_{PQ}^2 = \left(S_0^2 - \varepsilon^2\right) + 2\varepsilon\sqrt{T_Q^2 - S_0^2}$$

Applying the same calculation to R and subtracting,

$$\tau_{PR}^2 - \tau_{PQ}^2 = 2\varepsilon \left(\sqrt{T_R^2 - S_0^2} - \sqrt{T_Q^2 - S_0^2} \right) \tag{6}$$

which diverges as $T_R \to \infty$ for fixed P and Q, provided $\varepsilon \neq 0$ (i.e. P is not at O); for large T_R,

$$\tau_{PR}^2 \simeq 2\varepsilon T_R. \tag{7}$$

Thus a process that takes a minimum proper time τ_1 to occur, but depending on all the points in an initial domain $U = (0, x)$, $0 < x < \varepsilon$, in the surface $T = 0$ will gradually progress up the points in the surface $S_0 = const$ and will take an arbitrarily long time to reach the furthest points in this surface, no matter how small ε is, provided it is non-zero. It will take an infinite time to affect all points in the surface in the true meaning of the word infinity: it will always be progressing to further and further points, and will never complete the process in any finite time. If however

ε $=$ 0 (i.e. P is at O), the originating events are of zero spacelike extent and we recover $\tau_{PR}^2 = \tau_{PQ}^2$ (both Q and R are at the same proper time from O); all events on the surface occur simultaneously in terms of proper time along the world lines from O. But this exactly pointlike origin of later physical events is unphysical; it cannot occur in practice.

4.2 Bifurcate horizons

The more general context context worth considering is that of bifurcate Killing vectors, considered so clearly by Boyer [1], which includes the previous example and the de Sitter case considered in [12] as special cases.[1] The group of isometries acts as an isotropy group about the points in the bifurcate Killing horizon, which are fixed points of the group. As the group acts on the points in the space–time, the proper time $\tau(O, P_i)$ from the origin O to images P_i of a point in the surface of constant proper time from the origin stays invariant (Fig. 3). Similarly the proper time $\tau(Q_i, P_i)$ from images Q_i at constant spatial distance from the origin of a point Q_1, to images P_i of a point P_1 in the surface of constant time from O, stays invariant (Fig. 4). However the proper time $\tau(Q_1, Q_i)$ from Q_1 to Q_i increases linearly with the group parameter ξ along the timelike group orbits in the surfaces of constant distance from the origin. The proper time $\tau(P_1, Q_i)$ is greater than the sum of the times $\tau(P_1, P_2)$ and $\tau(P_2, Q_2)$ (as a timelike geodesic without conjugate points is longer than any other curve between its endpoints). Hence

$$\tau(P_1, Q_2) > \tau(P_1, P_2) + \tau(P_2, Q_2) = \alpha\xi + \tau(P_1, Q_1) \tag{8}$$

for some constant α; thus $\tau(P_1, Q_2) \to \infty$ as $\xi \to \infty$ and the proper time from P_1 to points on the surface $T0$ is unbounded in concordance with (7). The physical conclusions are the same as in the previous case. In particular this argument applies to the De Sitter Universe, where the static frame represents the transition to spatial surfaces of constant negative curvature.

---

[1] See for example Rindler [10], pages 49–54 and 306–309, respectively

Fig. 3 Group action about origin O for any bifurcate horizon. The symmetry group of a bifurcate horizon moves the point Q_1 in the surface S_0 (constant proper time $T0$ from O) to the point Q_2 in the same surface. The proper time OQ_1 is the same as the proper time OQ_2, because an isometry group leaves space–time distances invariant. The future null cone of O is mapped into itself by the group, and O is a fixed point of the group

Fig. 4 Group action on points near a bifurcate horizon. The symmetry group of a bifurcate horizon moves the point Q_1 in the surface S_0 (constant proper time $T0$ from O) to the point Q_2 in the same surface, and moves the point P_1 in the surface $D0$ (constant proper spatial distance $D0$ from O) to the point P_2. The proper time P_1Q_1 is the same as the proper time P_2Q_2, because an isometry group leaves space–time distances invariant, but is less than the proper time P_1Q_2

4.3 The null cone revisited

The same argument applies to the null cone, see Fig. 5 for the Minkowski space–time situation. The set of points represented by the null cone is instantaneous for the origin O, but more distant points on the null cone are at larger and larger proper distances from any point $P = (0, \varepsilon)$ in the surface $T = 0$, no matter how small ε, provided $\varepsilon \neq 0$. Essentially the same argument applies in both Figs. 2 and 5; hence events unroll along the null cone too, because they are based in what occurs in a finite rather than pointlike domain in the initial surface $T = 0$. In fact the same Eqs. (4)–(7) apply as before, but now with $S_0 = 0$. Thus we get

$$\tau_{PR}^2 = \tau_{PQ}^2 + 2\varepsilon\,(T_R - T_Q), \tag{9}$$

again leading to (7) for large T_R.

This argument will again extend to the case of any bifurcate Killing horizon as defined in [1]. Figure 6 is the relevant diagram, with the same argument as above in relation to Fig. 4 leading to the same result as before, but now as regards points on the light cone.

5 Conclusion

An infinite set of universes never exists at any finite time in chaotic inflation—it is a state that is never attained, but is what the physical situation tends to as time progresses. Similarly spatial infinities cannot occur at any finite time in any one of the universe domains resulting from quantum tunnelling; rather they may be the state that the models tend to as time goes to infinity. This is so even though the hyperbolic infinite space sections are at unit proper time from the origin, because what happens at later times is not determined only by what happens at a space–time point such as the origin.

Genuine physical processes will originate from domains with non-zero spatial extent. And as soon as this domain has finite extent, arbitrarily near points to the origin in the Minkowski surfaces of constant time $T = const$ are infinitely distant in proper time from the infinitely distant spatial events in the surfaces of constant proper time from the origin $S = const$, and it will take an infinite time for their consequences to develop to the furthest reaches of these spacelike surfaces. The putative spatial infinities will not come into being at any finite time. The key point is that the physics does not take place instantaneously in the spatial sections at

Fig. 5 Space–time distances and the null cone of O in Minkowski space–time. Both Q and R lie on the future light cone of O, and hence are at zero proper distance from it, hence influences from O reach both Q and R simultaneously. The proper time PR however is greater than the proper time PQ, hence influences from P reach Q before they reach R

Fig. 6 Group action and the null cone of a bifurcate horizon. The symmetry group of a bifurcate horizon moves the point Q_1 in the future null cone of O to the point Q_2 in that null cone, and moves the point P_1 in the surface $D0$ (constant proper distance $D0$ from O) to the point P_2. The proper time $P_1 Q_1$ is the same as the proper time $P_2 Q_2$, because an isometry group leaves space–time distances invariant, but is less than the proper time $P_1 Q_2$

constant proper distance from the origin, because it cannot have an *exactly* point-like physical origin. Consequently, the potential infinity is never actually attained in physical terms. It is always in the future of what actually exists.

One of the main reasons that we study cosmology is because of our fascination with its philosophical aspects. The idea of infinite spatial sections has bizarre implications; if true, it plausibly implies that countless identical civilizations to ours are scattered in the infinite expanse of the cosmos, with semi-identical histories to ours replicated an infinite number of times out there [4]. Vilenkin et al. claim this is a necessary outcome of current inflationary theories [9; 11]. We claim that this is not the case, the catch lying in the idea that this is currently how things are: that it is the state at the present instant. As explained above, the real situation is that physical processes may be such that eventually an infinite number of galaxies, stars, planets, and civilizations will tend to come into existence; but that state is not achieved at any finite time through the supposed physical processes.

Any claims of actual existence of physical infinities in the real universe should be treated with great caution (cf. [3], Sec. 9.3.2), as emphasized by David Hilbert long ago ([8], p. 151):

> "*Our principal result is that the infinite is nowhere to be found in reality. It neither exists in nature nor provides a legitimate basis for rational thought…. The role that remains for the infinite to play is solely that of an idea … which transcends all experience and which completes the concrete as a totality…*"

Our results concur with this judgement.

References

1. R.H. Boyer (1969) Geodesic killing orbits and bifurcate killing horizons *Proc. R. Soc. (Lond.)* A311 245 – 252
2. Ellis, G.F.R.: Physics in the real universe: time and space–time. Gen. Relativ. Gravit. **38**, 1797–1824 (2006). arXiv:gr-qc/0605049
3. Ellis, G.F.R.: Issues in the philosophy of cosmology. In: Butterfield, J., Earman, J. (eds.) Handbook in Philosophy of Physics, pp. 1183–1285. Elsevier, Amsterdam (2006). arxiv:astro-ph/0602280
4. G.F.R. Ellis G.B. Brundrit (1979) Life in the infinite universe *Q. J. R. Asiat. Soc.* **20** 37 – 41
5. Freivogel, B., Kleban, M., Martinez, M.R., Susskind, L.: Observational consequences of a landscape. JHEP 0603:039 (2006). arXiv:hep-th/0505232
6. Guth, A.: Eternal inflation and its implications. J. Phys. A. **40**, 6811–6826 (2007). arXiv:hep-th/0702178
7. Garriga, J., Vilenkin, A.: Many worlds in one. Phys. Rev. **D64**, 043511 (2001). arxiv:gr-qc/0102010
8. Hilbert, D.: On the infinite. In: Benacerraf, P., Putnam, H. (eds.) Philosophy of Mathematics, pp. 134–151. Prentice-Hall, Englewood Cliff (1964)
9. Knobe, J., Ohm, K.D., Vilenkin, A.: Philosophical implications of inflationary cosmology. Br. J. Philos. Sci. **57**, 47–67 (2006). arxiv: physics/0302071

10. W. Rindler (2001) Relativity: Special, General, and Cosmological Oxford University Press Oxford
11. A. Vilenkin (2006) Many Worlds in One. The Search for Other Universes Hill and Wang New York
12. Vilenkin, A., Winitzki, S.: Probability distribution for Ω in open-universe inflation. Phys. Rev. **D55**, 548–559 (1997). arxiv:astro-ph/9605191