



ELSEVIER

Nuclear Physics B 543 (1999) 275–320

NUCLEAR
PHYSICS B

Pre-Big-Bang bubbles from the gravitational instability of generic string vacua

A. Buonanno^a, T. Damour^{a,b}, G. Veneziano^{c,a}

^a *Institut des Hautes Etudes Scientifiques, 91440 Bures-sur-Yvette, France*

^b *DARC, CNRS-Observatoire de Paris, 92195 Meudon, France*

^c *Theory Division, CERN, CH-1211 Geneva 23, Switzerland*

Received 30 June 1998; accepted 13 November 1998

Abstract

We formulate the basic postulate of pre-Big-Bang cosmology as one of “asymptotic past triviality”, by which we mean that the initial state is a generic perturbative solution of the tree-level low-energy effective action. Such a past-trivial “string vacuum” is made of an arbitrary ensemble of incoming gravitational and dilatonic waves, and is generically prone to gravitational instability, leading to the possible formation of many black holes hiding singular space-like hypersurfaces. Each such singular space-like hypersurface of gravitational collapse becomes, in the string-frame metric, the usual Big-Bang $t = 0$ hypersurface, i.e. the place of birth of a baby Friedmann universe after a period of dilaton-driven inflation. Specializing to the spherically symmetric case, we review and reinterpret previous work on the subject, and propose a simple, scale-invariant criterion for collapse/inflation in terms of asymptotic data at past null infinity. Those data should determine whether, when, and where collapse/inflation occurs, and, when it does, fix its characteristics, including anisotropies on the Big-Bang hypersurface whose imprint could have survived till now. Using Bayesian probability concepts, we finally attempt to answer some fine-tuning objections recently moved to the pre-Big-Bang scenario. © 1999 Elsevier Science B.V.

PACS: 11.25.-w; 98.80.Cq

Keywords: Physics of the Early Universe; String theory

1. Introduction and general overview

Superstring theory (see Ref. [1] for a review) is the only presently known framework in which gravity can be consistently quantized, at least perturbatively. The well-known difficulties met in trying to quantize General Relativity (GR) – or its supersymmetric extensions – are avoided, in string theory, by the presence of a fundamental quantum of length [2] $\ell_s \sim (\alpha')^{1/2}$. Thus, at distances shorter than ℓ_s , string gravity is expected

to be drastically different from – and in particular to be much “softer” than – General Relativity.

However, as was noticed since the early days of string theory [3], a conspicuous difference between string and Einstein gravity persists even at low energies (large distances). Indeed, a striking prediction of string theory is that its “gravitational sector” is richer than that of GR: in particular, all versions of string theory predict the existence of a scalar partner of the spin-two graviton, i.e. of the metric tensor $g_{\mu\nu}$, the so-called dilaton, ϕ . This field plays a central rôle within string theory [1] since its present vacuum expectation value (VEV) ϕ_{now} fixes the string coupling constant $g^2 = e^{\phi_{\text{now}}}$, and, through it, the present values of gauge and gravitational couplings. In particular, it fixes the ratio of ℓ_s to the Planck length $\ell_P \sim G^{1/2}$ by $\ell_P \sim g\ell_s$. The relation is such that gauge and gravitational couplings unify at the string energy scale, i.e. at $E \sim M_s \sim gM_P \sim (\alpha')^{-1/2} \sim 3 \times 10^{17}$ GeV. It thus seems that the string way to quantizing gravity forces this new particle/field upon us.

We believe that the dilaton represents an interesting prediction (an opportunity rather than a nuisance) whose possible existence should be taken seriously, and whose observable consequences should be carefully studied. Of course, tests of GR [4] put severe constraints on what the dilaton can do today. The simplest way to recover GR at late times is to assume [5] that ϕ gets a mass from supersymmetry-breaking non-perturbative effects. Another possibility might be to use the string-loop modifications of the dilaton couplings for driving ϕ toward a special value where it decouples from matter [6]. These alternatives do not rule out the possibility that the dilaton may have had an important rôle in the previous history of the universe. Early cosmology stands out as a particularly interesting arena where to study both the dynamical effects of the dilaton and those associated with the existence of a fundamental length in string theory.

In a series of previous papers [7,8] a model of early string cosmology, in which the dilaton plays a key dynamical rôle, was introduced and developed: the so-called pre-Big-Bang (PBB) scenario. One of the key ideas of this scenario is to use the kinetic energy of the dilaton to drive a period of inflation of the universe. The motivation is that the presence of a (tree-level coupled) dilaton essentially destroys [9] the usual inflationary mechanism [10]: instead of driving an exponential inflationary expansion, a (nearly) constant vacuum energy drives the string coupling $g = e^{\phi/2}$ towards small values, thereby causing the universe to expand only as a small power of time. If one takes seriously the existence of the dilaton, the PBB idea of a dilaton-driven inflation offers itself as one of the very few natural ways of reconciling string theory and inflation. Actually, the existence of inflationary solution in string cosmology is a consequence of its (T -)duality symmetries [11].

This paper develops further the PBB scenario by presenting a very general class of possible initial states for string cosmology, and by describing their subsequent evolution, via gravitational instability, into a multi-universe comprising (hopefully) sub-universes looking like ours. This picture generalizes, and makes more concrete, recent work [12,13] about inhomogeneous versions of pre-Big-Bang cosmology.

Let us first recall that an inflation driven by the kinetic energy of ϕ forces both the coupling and the curvature to *grow* during inflation [7]. This implies that the initial state must be very perturbative in two respects: (i) it must have very small initial curvatures (derivatives) in string units and, (ii) it must exhibit a tiny initial coupling $g_i = e^{\phi_i/2} \ll 1$. As the string coupling g^2 measures the strength of quantum corrections (i.e. plays the rôle of \hbar in dividing the Lagrangian: $\mathcal{L} = g^{-2} \mathcal{L}'$), quantum (string-loop) corrections are initially negligible. Because of (i), α' corrections can also be neglected. In conclusion, dilaton-driven inflation must start from a regime in which the tree-level low-energy approximation to string theory is extremely accurate, something we may call an asymptotically trivial state.

In the present paper, we consider a very general class of such “past-trivial” states. Actually, perturbative string theory is well defined only when one considers such classical states as background, or “vacuum”, configurations. For the sake of simplicity the set of string vacua that we consider are already compactified to four dimensions and are truncated to the gravi-dilaton sector (antisymmetric tensor and moduli being set to zero). Within these limitations, the set of all perturbative string vacua coincides with the generic solutions of the tree-level low-energy effective action [14]

$$S_s = \frac{1}{\alpha'} \int d^4x \sqrt{G} e^{-\phi} [R(G) + G^{\mu\nu} \partial_\mu \phi \partial_\nu \phi], \quad (1.1)$$

where we have denoted by $G_{\mu\nu}$ the string-frame (σ -model) metric. The generic solution is parametrized by 6 functions of three variables. These functions can be thought of classically as describing the two helicity-2 modes of gravitational waves, plus the helicity-0 mode of dilatonic waves. (Each mode being described by two real functions corresponding, e.g., to the Cauchy data $(\phi, \dot{\phi})$ at some initial time.) The same counting of the degrees of freedom in selecting string vacua can be obtained by considering all the marginal operators (i.e. all conformal-invariance-preserving continuous deformations) of the conformal field theory defining the quantized string in trivial space-time. We therefore envisage, as initial state, the most general past-trivial classical solution of (1.1), i.e. an arbitrary ensemble of incoming gravitational and dilatonic waves.

Our aim will be to show how such a stochastic bath of classical incoming waves (devoid of any ordinary matter) can evolve into our rich, complex, expanding universe. The basic mechanism we consider for turning such a trivial, inhomogeneous and anisotropic, initial state into a Friedmann-like cosmological universe is gravitational instability (and quantum particle production as far as heating up the universe is concerned [8]). We find that, when the initial waves satisfy a certain (dimensionless) strength criterion, they collapse (when viewed in the Einstein conformal frame) under their own weight. When viewed in the (physically most appropriate) string conformal frame, each gravitational collapse leads to the local birth of a baby inflationary universe blistering off the initial vacuum. Thanks to the peculiar properties of dilaton-driven inflation (i.e. the peculiar properties of collapse with the equation of state $p = \epsilon$), each baby universe is found to contain a large homogeneous patch of expanding space which might constitute the beginning of a local Big Bang. We then expect each of these ballooning patches of

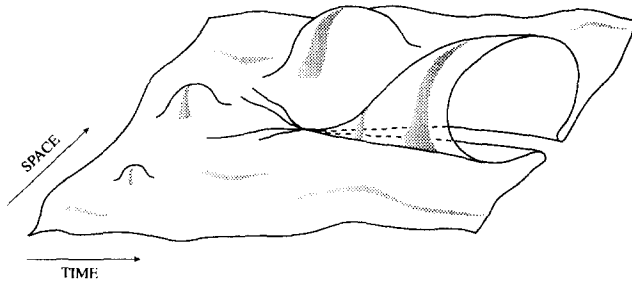


Fig. 1. Symbolic sketch of the birth of many pre-Big-Bang bubbles from the gravitational instability of a generic string vacuum made of a stochastic bath of classical incoming gravitational and dilatonic waves. Each local Einstein-frame collapse of sufficiently strong waves forms a cosmological-like space-like singularity hidden behind a black hole. The parts of those classical singularities where the string coupling grows sufficiently fast expand, when viewed in the string frame, and generate ballooning patches of space (here schematized as the stretching of one spatial dimension) which are expected to evolve into many separate quasi-closed Friedmann hot universes.

space to evolve into a quasi-closed Friedmann universe.¹ This picture is sketched in Figs. 1 and Fig. 2. In order to study in detail this scenario, we focus, in this paper, on the technically simplest case containing non-trivial incoming waves able to exhibit gravitational instability: spherically symmetric dilatonic waves. This simple toy model seems to contain many of the key physical features we wish to study.

Before entering the technicalities of this model, let us clarify a few general methodological issues. One of the goals of theoretical cosmology is to “explain” the surprisingly rich and special structure of our universe. However, the concept of “explanation” is necessarily dependent on one’s prejudices and taste. We wish only to show here how, modulo an “exit” assumption [8], one can “explain” the appearance of a hot, expanding homogeneous universe starting from a generic classical inhomogeneous vacuum of string theory. We do not claim that this scenario solves all the cosmological “problems” that one may wish to address (e.g., we leave untouched here the monopole and gravitini problems). We do not try either to “explain” the appearance of our universe as a quantum fluctuation out of “nothing”. We content ourselves by assuming, as is standard in perturbative string theory, the existence of a classical vacuum and showing how gravitational instability can then generate some interesting qualitatively new structures akin to those of our universe. In particular, we find it appealing to “understand” the striking existence of a preferred rest frame in our universe, starting from a stochastic bath of waves propagating with the velocity of light and thereby exhibiting no clear split between space and time. We shall also address in this paper the question of the naturalness of our scenario. Recent papers [18,19] have insisted on the presence of two large numbers among the parameters defining the PBB scenario. Our answer to this issue (which is, anyway, an issue of taste and not a scientifically well-posed problem) is twofold: on the one hand, we point out that the two large numbers in question are

¹ Our picture of baby universes created by gravitational collapse is reminiscent of earlier proposals [15–17], but differs from them by our crucial use of string-theory motivated ideas.

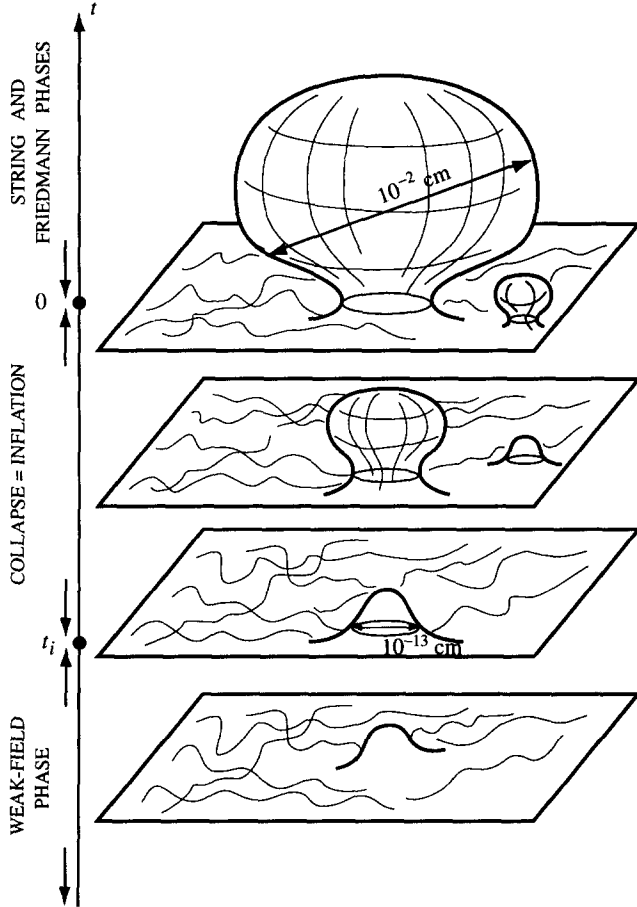


Fig. 2. A representation of pre-Big-Bang bubbles similar to that of Fig. 1, but in $2 + 1$ dimensions. The different horizontal planes represent different instants in the evolution from the asymptotic trivial past to the Friedmann phase. Two inflationary bubbles characterized by two different initial horizon sizes (both large in string units) are shown to lead to Universes of very different homogeneity scale at the time at which the Hubble radius reaches string-scale values ($\ell_s = \mathcal{O}(10^{-32}$ cm)).

classically undefined and therefore irrelevant when discussing the naturalness of a *classical* vacuum state; on the other hand, we show that the selection effects associated to asking any such “fine-tuning” question can render natural the presence of these very large numbers.

2. Asymptotic null data for past-trivial string vacua

Motivated by the pre-Big-Bang idea of an initial weak-coupling, low-curvature state we consider the tree-level (order g^{-2}) effective action for the gravitational sector of critical superstring theory, taken at lowest order in α' , Eq. (1.1). We set to zero the

antisymmetric tensor field $B_{\mu\nu}$, and work directly in four dimensions (i.e. we assume that the gravitational moduli describing the six “internal dimensions” are frozen). Though the physical interpretation of our work is best made in terms of the original string (or σ -model) metric $G_{\mu\nu}$ appearing in Eq. (1.1), it will be technically convenient to work with the conformally related Einstein metric

$$g_{\mu\nu} = e^{-(\phi - \phi_{\text{now}})} G_{\mu\nu}, \quad 16\pi G = \alpha' e^{\phi_{\text{now}}}. \quad (2.1)$$

In the following, we will set $G = 1$. In terms of the Einstein metric $g_{\mu\nu}$, the low-energy tree-level string effective action (1.1) reads²

$$S = \frac{1}{16\pi} \int d^4x \sqrt{g} \left[R - \frac{1}{2} \partial_\mu \phi \partial^\mu \phi \right]. \quad (2.2)$$

The corresponding classical field equations are

$$R_{\mu\nu} = \frac{1}{2} \partial_\mu \phi \partial_\nu \phi, \quad (2.3)$$

$$\square \phi \equiv \nabla^\mu \nabla_\mu \phi = 0, \quad (2.4)$$

with Eq. (2.4) actually following from Eq. (2.3) thanks to the Bianchi identity. As explained in the Introduction, we consider a generic solution of these classical field equations admitting an asymptotically trivial past. Such an asymptotic “incoming” classical state should allow a description in terms of a superposition of ingoing gravitational and dilatonic waves. The work of Bondi, Sachs, Penrose [20] and many others in classical gravitation theory indicates that this incoming state should exhibit a regular past null infinity \mathcal{I}^- , and should be parametrizable by some asymptotic null data (i.e. conformally renormalized data on \mathcal{I}^-). In plain terms, this means the existence of suitable asymptotically Minkowskian³ coordinates $(x^\mu) = (t, x, y, z)$ (which can then be used in the standard way to define polar coordinates r, θ, φ and the advanced time $v \equiv t + r \equiv t + (x^2 + y^2 + z^2)^{1/2}$) such that the following expansions hold when $r \rightarrow \infty$, at fixed v, θ and φ :

$$\phi(x^\lambda) = \phi_0 + \frac{f(v, \theta, \varphi)}{r} + o\left(\frac{1}{r}\right), \quad (2.5)$$

$$g_{\mu\nu}(x^\lambda) = \eta_{\mu\nu} + \frac{f_{\mu\nu}(v, \theta, \varphi)}{r} + o\left(\frac{1}{r}\right). \quad (2.6)$$

The null wave data on \mathcal{I}^- are: the asymptotic dilatonic wave form $f(v, \theta, \varphi)$, and the two polarization components $f_+(v, \theta, \varphi)$, $f_\times(v, \theta, \varphi)$, of the asymptotic gravitational wave form $f_{\mu\nu}(v, \theta, \varphi)$.⁴ (Introducing a local orthonormal frame $\mathbf{e}_{(1)}$, $\mathbf{e}_{(2)}$ on the sphere at infinity one usually defines $f_+ = \frac{1}{2}(f_{(1)(1)} - f_{(2)(2)})$, $f_\times = f_{(1)(2)}$.) The

² We will use the signature $(-, +, +, +)$ and the conventions $R^\mu{}_{\nu\rho\sigma} = \partial_\rho \Gamma^\mu_{\nu\sigma} - \dots$, $R_{\mu\nu} = R^\rho{}_{\mu\rho\nu}$.

³ These coordinates must be restricted by the condition that the incoming coordinate cones $v \equiv t + r = \text{const.}$ be asymptotically tangent to exact cones of $g_{\mu\nu}(x^\lambda)$.

⁴ The three functions f, f_+, f_\times of v, θ, φ are equivalent to six functions in \mathbb{R}^3 , i.e. six functions of r, θ, φ with $r \geq 0$, because the advanced time v ranges over the full line $(-\infty, +\infty)$.

other $o(1/r)$ pieces in the metric are gauge dependent, except for the “Bondi mass aspect” $m_\infty(v, \theta, \varphi)$ (defined below in a simple case) whose advanced-time derivative is related to the (direction-dependent) incoming energy fluxes:

$$\frac{\partial}{\partial v} m_\infty(v, \theta, \varphi) = \frac{1}{4} (\partial_v f)^2 + \frac{1}{4} (\partial_v f_+)^2 + \frac{1}{4} (\partial_v f_\times)^2 + \text{div.}, \quad (2.7)$$

where div. denotes an angular divergence, $(\sin \theta)^{-1} \partial_\theta (\sin \theta D^\theta) + \partial_\varphi D^\varphi$. Instead of the asymptotic wave forms f, f_+, f_\times (which have the dimension of length), one can work with the corresponding “news functions”

$$N(v, \theta, \varphi) \equiv \partial_v f(v, \theta, \varphi), \quad N_+ \equiv \partial_v f_+, \quad N_\times \equiv \partial_v f_\times, \quad (2.8)$$

which are dimensionless, and whose squares give directly the incoming energy flux appearing on the r.h.s. of Eq. (2.7).

Before specializing this generic ingoing state, let us note [12], the existence of two important global symmetries of the classical field equations, (2.3) and (2.4). They are invariant both under global scale transformations, and under a constant shift of ϕ : $ds'^2 = a^2 ds^2$, $\phi' = \phi + b$. (They are also invariant under local coordinate transformations.) In terms of asymptotic data (whose definition requires a specific “flat” normalization of the coordinates at past infinity), the global symmetry transformations read

$$f(v, \theta, \varphi) \rightarrow f'(v', \theta, \varphi) = a f(v'/a, \theta, \varphi), \quad (2.9)$$

$$f_+(v, \theta, \varphi) \rightarrow f'_+(v', \theta, \varphi) = a f_+(v'/a, \theta, \varphi), \quad (2.10)$$

$$f_\times(v, \theta, \varphi) \rightarrow f'_\times(v', \theta, \varphi) = a f_\times(v'/a, \theta, \varphi) \quad (2.11)$$

with $r \rightarrow r' = ar$, $v \rightarrow v' = av$, and

$$\phi_0 \rightarrow \phi'_0 = \phi_0 + b. \quad (2.12)$$

Note that the three dimensionless news functions (2.8) are numerically invariant under the scaling transformations: $N'(v', \theta, \varphi) = N(v, \theta, \varphi)$, etc., with $v' = av$.

As the full time evolution of string vacua depends only on the null data, it is a priori clear that the amplitude of the dimensionless news functions must play a crucial rôle in distinguishing “weak” incoming fields – that finally disperse again as weak outgoing waves – from “strong” incoming fields that undergo a gravitational instability. To study this difficult problem in more detail we turn in the next section to a simpler model with fewer degrees of freedom.

Before doing so, let us comment on the “classicality” of an “in state” defined by some news functions (2.8). The classical fields (2.5), (2.6) deviate less and less from a trivial background $(\phi_0, \eta_{\mu\nu})$ as $r \rightarrow \infty$ (with fixed v, θ and φ). One might then worry that, sufficiently near \mathcal{I}^- , quantum fluctuations $\phi^q, h^q_{\mu\nu}$ might start to dominate over the classical deviations $\phi^c \equiv \phi(x) - \phi_0, h^c_{\mu\nu} \equiv g_{\mu\nu} - \eta_{\mu\nu}$. To ease this worry let us compute the mean number of quanta in the incoming waves (2.5), (2.6). Consider for simplicity the case of an incoming spherically symmetric dilatonic wave $\phi^c_{\text{in}} = f(v)/r + o(1/r)$. Such an incoming wave has evidently the same mean number of

quanta as the corresponding outgoing⁵ wave ϕ_{out}^c , with asymptotic behaviour $\phi_{\text{out}}^c = -f(u)/r + o(1/r)$, on \mathcal{I}^+ ($r \rightarrow \infty$, with $u \equiv t - r, \theta$ and φ fixed). As is well known [21], a classical outgoing wave can be viewed, quantum mechanically, as a coherent state in the incoming Fock space defined by the oscillator decomposition of the quantum field $\hat{\phi}$. Instead of parametrizing ϕ_{out}^c by the waveform $f(u)$, one can parametrize it by an effective source $J(x^\Lambda)$ such that $\phi_{\text{out}}^c(x) = \int d^4x' P_{\text{ret}}(x-x') J(x')$, where P_{ret} is the (action-normalized) retarded propagator of the field ϕ . (We use a normalization such that $S = -\frac{1}{2} \phi P^{-1} \phi + J \phi$.) The mean number of quanta in the coherent state associated to a classical source $J(x)$ has been computed in [21] for the electromagnetic case, and in [22] for generic massless fields. It reads (with $p x \equiv \mathbf{p} \cdot \mathbf{x} - \omega t$)

$$\bar{n} = \frac{\pi}{\hbar} \int \frac{d^4p}{(2\pi)^4} \delta(p^2) \hat{J}(-p) \mathcal{R} \hat{J}(p), \quad \hat{J}(p) \equiv \int d^4x e^{-ipx} J(x), \quad (2.13)$$

where \mathcal{R} denotes the residue of the propagator of the corresponding field, defined such that the field equation $P^{-1} \phi = J$ reads $\square \phi = -\mathcal{R} J$. (In the normalization of the present paper $\mathcal{R} = 16\pi G$.) Inserting $J(x) = -(4G)^{-1} f(t) \delta^{(3)}(\mathbf{x})$ into Eq. (2.13) yields

$$\bar{n} = \frac{1}{2G\hbar} \int_0^\omega \frac{d\omega}{2\pi} \omega |\hat{f}(\omega)|^2 = \frac{1}{2G\hbar} \int_0^\omega \frac{d\omega}{2\pi} \frac{1}{\omega} |\hat{N}(\omega)|^2, \quad (2.14)$$

where $\hat{N}(\omega) \equiv \int dt e^{i\omega t} N(t)$ is the one-dimensional Fourier transform of the news function $N(t) = f'(t)$. In order of magnitude, if we consider a pulse-like waveform $f(t)$, with characteristic duration $\Delta t \sim \ell_c$, and characteristic dimensionless news amplitude $N_c \sim f_c/\ell_c$, the corresponding mean number of quanta is

$$\bar{n} \sim \frac{N_c^2 \ell_c^2}{\ell_p^2} = \frac{N_c^2 \ell_c^2}{g^2 \ell_s^2}. \quad (2.15)$$

In the present paper, we shall consider incoming news functions with amplitude $N_c \sim 1$ and scale of variation $\ell_c \gg \ell_s$. Moreover, in order to ensure a sufficient amount of inflation, the initial value of the string coupling $g^2 = e^{\phi_0}$ has to be $\ll 1$. Therefore, Eq. (2.15) shows that such incoming states can be viewed as highly classical coherent states, made of an enormous number of elementary ϕ -quanta: $\bar{n} \gg 1$. This quantitative fact clearly shows that there is no need to worry about the effect of quantum fluctuations, even near \mathcal{I}^- where $\phi_{\text{in}}^c \sim f(v)/r \rightarrow 0$. In the case of the negative-curvature Friedmann-dilaton solution this fact has been recently confirmed by an explicit calculation of scalar and tensor quantum fluctuations [23].

⁵ In keeping with usual physical intuition, it is more convenient to work with an auxiliary problem where a classical source generates an outgoing wave which superposes onto the quantum fluctuations of an incoming vacuum state.

3. Equations for a spherically symmetric pre-Big-Bang

The simplest non-trivial model which can exhibit the transition between weak, stable data, and gravitationally unstable ones is the spherically symmetric Einstein-dilaton system. The null wave data for this system comprise only an angle-independent asymptotic dilatonic wave form $f(v)$, corresponding to the dimensionless news function

$$N(v) = \partial_v f(v). \tag{3.1}$$

The spherically symmetric Einstein-dilaton system has been thoroughly studied by many authors with the strongest analytic results appearing in papers by Christodoulou [24–26]. A convenient system of coordinates is the double null system, (u, v) , such that

$$\phi = \phi(u, v), \tag{3.2}$$

$$ds^2 = -\Omega^2(u, v) du dv + r^2(u, v) d\omega^2, \tag{3.3}$$

where $d\omega^2 = d\theta^2 + \sin^2 \theta d\varphi^2$. The field equations are conveniently re-expressed in terms of the three functions $\phi(u, v)$, $r(u, v)$ and $m(u, v)$, where the local mass function $m(u, v)$ is defined by⁶

$$1 - \frac{2m}{r} \equiv g^{\mu\nu} (\partial_\mu r) (\partial_\nu r) = -\frac{4}{\Omega^2} \left(\frac{\partial r}{\partial u} \right)_v \left(\frac{\partial r}{\partial v} \right)_u. \tag{3.4}$$

One gets the following set of evolution equations for m , r and ϕ :

$$2 \left(\frac{\partial r}{\partial u} \right)_v \left(\frac{\partial m}{\partial u} \right)_v = \left(1 - \frac{2m}{r} \right) \frac{r^2}{4} \left(\frac{\partial \phi}{\partial u} \right)_v^2, \tag{3.5}$$

$$2 \left(\frac{\partial r}{\partial v} \right)_u \left(\frac{\partial m}{\partial v} \right)_u = \left(1 - \frac{2m}{r} \right) \frac{r^2}{4} \left(\frac{\partial \phi}{\partial v} \right)_u^2, \tag{3.6}$$

$$r \frac{\partial^2 r}{\partial u \partial v} = \frac{2m}{r - 2m} \left(\frac{\partial r}{\partial u} \right)_v \left(\frac{\partial r}{\partial v} \right)_u, \tag{3.7}$$

$$r \frac{\partial^2 \phi}{\partial u \partial v} + \left(\frac{\partial r}{\partial u} \right)_v \left(\frac{\partial \phi}{\partial v} \right)_u + \left(\frac{\partial r}{\partial v} \right)_u \left(\frac{\partial \phi}{\partial u} \right)_v = 0. \tag{3.8}$$

This double-null evolution system is form-invariant under independent local reparametrizations of u and v : $u' = U(u)$, $v' = V(v)$. To freeze this gauge freedom it is convenient to require that $u = v$ on the central worldline of symmetry $r = 0$ (as long as it is a regular part of space-time), and that v and r asymptotically behave like ingoing Bondi coordinates on \mathcal{I}^- (see below).

The quantity

$$\mu(u, v) \equiv \frac{2m(u, v)}{r} \tag{3.9}$$

⁶ In order to avoid confusion we indicate, in standard notation, the variable kept fixed under differentiation.

plays a crucial rôle in the problem. Following Ref. [26], we shall sometimes call it the “mass ratio”. If μ stays everywhere below 1, the field configuration will not collapse but will finally disperse itself as outgoing waves at infinity. By contrast, if the mass ratio μ can reach anywhere the value 1, this signals the formation of an apparent horizon \mathcal{A} . The location of the apparent horizon is indeed defined by the equation

$$\mathcal{A} : \quad \mu(u, v) = 1. \tag{3.10}$$

The above statements are substantiated by some rigorous inequalities [26] stating that

$$\left(\frac{\partial r}{\partial u}\right)_v < 0, \quad \left(\frac{\partial m}{\partial v}\right)_u \geq 0, \tag{3.11}$$

$$\left(\frac{\partial r}{\partial v}\right)_u (1 - \mu) \geq 0, \quad \left(\frac{\partial m}{\partial u}\right)_v (1 - \mu) \leq 0. \tag{3.12}$$

Thus, in weak-field regions ($\mu < 1$), $(\partial_v r)_u > 0$, while, as $\mu > 1$, $(\partial_v r)_u < 0$, meaning that the outgoing radial null rays (“photons”) emitted by the sphere $r = \text{const.}$ become convergent, instead of having their usual behaviour.

It has been shown long ago that the presence of trapped surfaces implies the existence of some (possibly weak) type of geometric singularity [27,28]. In the case of the spherically symmetric Einstein-dilaton system, it has been possible to prove [25] that the presence of trapped surfaces (i.e. of an apparent horizon \mathcal{A} (3.10), the boundary of the trapped region) implies the existence of a future space-like singular boundary \mathcal{B} of space-time where the curvature blows up. Both \mathcal{A} and \mathcal{B} are “invisible” from future null infinity (\mathcal{I}^+), being hidden behind an event horizon \mathcal{H} (a null hypersurface). See Fig. 3.

One of our main purposes in this work is to give the conditions that the incoming dilatonic news function $N(v)$ must satisfy in order to create an apparent horizon, and thereby to lead to some localized gravitational collapse. Before addressing this problem let us complete the description of the toy model by re-expressing it in terms of the ingoing Bondi coordinates (v, r) . The metric can now be written in the form

$$ds^2 = - \left(1 - \frac{2m(v, r)}{r}\right) e^{2\beta(v, r)} dv^2 + 2e^{\beta(v, r)} dv dr + r^2 d\omega^2, \tag{3.13}$$

and the field equations (3.5)–(3.8) become

$$r \left(\frac{\partial \beta}{\partial r}\right)_v = \frac{r^2}{4} \left(\frac{\partial \phi}{\partial r}\right)_v^2, \tag{3.14}$$

$$2 \left(\frac{\partial m}{\partial r}\right)_v = \left(1 - \frac{2m}{r}\right) \frac{r^2}{4} \left(\frac{\partial \phi}{\partial r}\right)_v^2, \tag{3.15}$$

$$2 \left(\frac{\partial m}{\partial v}\right)_r = e^{-\beta} \frac{r^2}{2} \left[\left(\frac{\partial \phi}{\partial v}\right)_r^2 + e^{\beta} \left(1 - \frac{2m}{r}\right) \left(\frac{\partial \phi}{\partial r}\right)_v \left(\frac{\partial \phi}{\partial v}\right)_r \right], \tag{3.16}$$

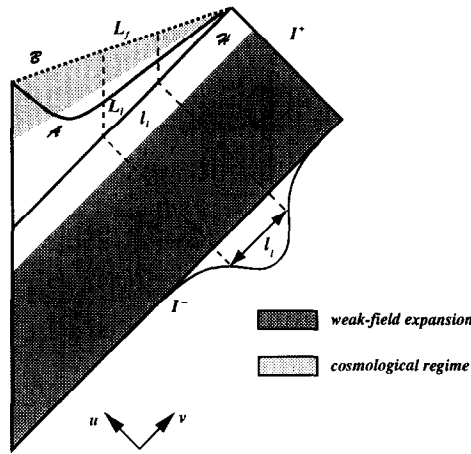


Fig. 3. Schematic representation of the space-time generated by the collapse of a spherically symmetric pulse of dilatonic waves. An incoming pulse of scalar news $N(v)$, which grows by $\mathcal{O}(1)$ on an advanced-time scale l_i , collapses to a space-like singularity \mathcal{B} after having formed an apparent horizon \mathcal{A} hidden behind the event horizon \mathcal{H} . In the vicinity of \mathcal{A} and \mathcal{H} , there is an abrupt transition between a weak-field region (dark shading) where the perturbation series (4.1)–(4.3) holds, and a strong-field one (light shading) where the cosmological-like expansion (5.34)–(5.36) holds.

$$2 \left[\frac{\partial^2 \phi}{\partial v \partial r} + \frac{1}{r} \left(\frac{\partial \phi}{\partial v} \right)_r \right] + e^\beta \left(1 - \frac{2m}{r} \right) \left(\frac{\partial^2 \phi}{\partial r^2} \right)_v + \frac{2e^\beta}{r} \left(1 - \frac{m}{r} \right) \left(\frac{\partial \phi}{\partial r} \right)_v = 0. \tag{3.17}$$

Eqs. (3.14), (3.15) imply

$$\partial_r \left(e^{\beta(v,r)} (r - 2m) \right) = e^{\beta(v,r)}. \tag{3.18}$$

In this coordinate system one can solve for the two functions $\beta(v, r)$ and $\mu(v, r)$ (or $m(v, r)$) by quadratures in terms of the third unknown function $\phi(v, r)$. Indeed, imposing that the coordinate system be asymptotically flat at \mathcal{I}^- (i.e. that $\beta(v, r) \rightarrow 0$ as $r \rightarrow \infty$), one first finds from Eq. (3.14)

$$\beta(v, r) = - \int_r^{+\infty} \frac{dr'}{r'} F(v, r'), \tag{3.19}$$

where

$$F(v, r) \equiv \frac{r^2}{4} \left(\frac{\partial \phi}{\partial r} \right)_v^2. \tag{3.20}$$

Eq. (3.15) can then be integrated to give

$$\mu(v, r) = \frac{[2m_\infty(v) - \int_r^{+\infty} dr' F(v, r') e^{\beta(v,r')}] }{r e^{\beta(v,r)}}, \tag{3.21}$$

where the “integration constant”

$$m_\infty(v) \equiv \lim_{r \rightarrow \infty} m(v, r) \tag{3.22}$$

denotes the incoming Bondi mass. By the Bondi energy-flux formula (2.7), which is also the limiting form of (3.16) for $r \rightarrow \infty$, $m_\infty(v)$ is given in terms of the news function by

$$m_\infty(v) = \frac{1}{4} \int_{-\infty}^v dv' N^2(v') = \frac{1}{4} \int_{-\infty}^v dv' \left(\frac{\partial f(v')}{\partial v'} \right)^2, \tag{3.23}$$

where we have inserted the asymptotic behaviour of ϕ on \mathcal{I}^- ,

$$\phi(v, r) = \phi_0 + \frac{f(v)}{r} + o\left(\frac{1}{r}\right), \tag{3.24}$$

and assumed that $N(v)$ decays sufficiently fast as $v \rightarrow -\infty$ (i.e. near past time-like infinity). From Eqs. (3.19) and (3.21) we find the leading asymptotic behaviour of the metric coefficients to be

$$1 - \frac{2m(v, r)}{r} = 1 - \frac{2m_\infty(v)}{r} + o\left(\frac{1}{r}\right), \tag{3.25}$$

$$e^{\beta(v, r)} = 1 + \mathcal{O}\left(\frac{1}{r^2}\right). \tag{3.26}$$

4. Data strength criterion for gravitational instability

The purpose of this section is to outline the condition that the initial data, i.e. the wave form $f(v)$ or the news function $N(v) \equiv f'(v)$, must satisfy in order to undergo, or not undergo, gravitational collapse. A lot of mathematical work has been done on this issue [24–26]. In particular, Ref. [24] and Ref. [26] gave two different “no collapse” criteria ensuring that the field configuration never collapses and finally disperses out as weak outgoing waves if some functional of the data is small enough. On the other hand, Ref. [25] gave a sufficient “collapse” criterion ensuring that the field configuration will collapse, if some functional of the data is large enough, thereby giving birth to a curvature singularity hidden behind a horizon. The problem with these nice and important mathematical results is threefold: (i) these criteria are sufficient but not necessary, so that they cannot answer our problem of finding (if possible) a sharp criterion distinguishing weak-field from strong-field data; (ii) the various measures of the “strength” of the data given in Refs. [24,25] and Ref. [26] are quite unrelated to each other and do not point clearly toward any sharp “strength criterion”; and (iii) these criteria are not expressed in terms of the asymptotic null data $N(v)$.

Our aim here is not to compete with Refs. [24–26] on the grounds of mathematical rigour, but to complement these results by a non-rigorous study which leads to the iterative computation of a sharp strength criterion, i.e. a functional $\mathcal{S}[f]$, such that data satisfying $\mathcal{S} < 1$ finally disperse out at infinity, while data satisfying $\mathcal{S} > 1$ are

gravitationally unstable and partly collapse to form a singularity. We will then compare our strength functional \mathcal{S} with the rigorous, but less explicit results of Ref. [25]. Note that the emphasis here is not on the question of whether the created singularity is hidden beyond a global event horizon when seen from outside, but rather on the quantitative criterion ensuring that past trivial data give birth to a space-like singularity which we shall later interpret as a budding pre-Big-Bang cosmological universe.

4.1. Perturbative analysis in the weak-field region

In view of the discussion of the previous section, the transition between “weak” and “strong” fields is sharply signaled by the appearance of an apparent horizon, i.e. of space-time points where $\mu \equiv 2m/r$ reaches the “critical” value 1. Therefore, starting from the weak-field region near \mathcal{I}^- , we can define the strength functional by computing $\mu(u, v)$ as a functional of the null incoming data and by studying if and when it can exceed the critical value. Specifically, we can set up a perturbative analysis of the Einstein-dilaton system in some weak-field domain connected to \mathcal{I}^- , and see for what data it suggests that μ will exceed somewhere the value 1. The expected domain of validity of such a weak-field expansion is sketched in Fig. 3.

The (u, v) system, Eqs. (3.2) and (3.3), lends itself easily to a perturbative treatment in the strength of ϕ . We think here of introducing a formal parameter, say λ , in the asymptotic data by $f(v) \rightarrow \lambda f(v)$ and of constructing the full solution $\phi(u, v)$, $r(u, v)$, $m(u, v)$, as a (formal) power series in λ (or by some better iteration method):

$$\bar{\phi}(u, v) \equiv \phi(u, v) - \phi_0 = \lambda \phi_1 + \lambda^3 \phi_3 + \lambda^5 \phi_5 + \dots, \tag{4.1}$$

$$r(u, v) = \frac{1}{2} (v - u) + \lambda^2 r_2 + \lambda^4 r_4 + \dots, \tag{4.2}$$

$$m(u, v) = \lambda^2 m_2 + \lambda^4 m_4 + \dots \tag{4.3}$$

Indeed, knowing ϕ to order λ^{2p-1} and r to order λ^{2p-2} we can use Eq. (3.5) or (3.6) to compute by quadrature m to order λ^{2p} . We can then rewrite Eqs. (3.7), (3.8) as

$$\partial_{uv} (r \bar{\phi}) = \omega r \bar{\phi}, \quad \partial_{uv} r = \omega r, \tag{4.4}$$

where

$$\omega \equiv \frac{\mu}{1 - \mu} \frac{\partial_u r \partial_v r}{r^2}, \tag{4.5}$$

and therefore the right-hand sides of Eqs. (4.4) are known to (relative) order λ^{2p} . Solving Eq. (4.4), then yields $\bar{\phi}$ and r to the next order in λ (λ^{2p+1} and λ^{2p} , respectively). To be more explicit, Eqs. (4.4) are of the form

$$\partial_{uv} \psi(u, v) = \sigma(u, v), \tag{4.6}$$

where the “source” $\sigma = \omega \psi$ is known at each order in λ from the lower-order expression of μ and ψ , and where the “field” $\psi(u, v)$ is restricted to vanish on the central worldline

of symmetry $v = u$ (where $r = 0$, and $\bar{\phi}$ is regular in the weak-field domain). The generic equation (4.6) can then be solved by introducing the retarded Green function

$$G(u, v; u', v') = \theta(u - u') \theta(v - v') - \theta(v - u') \theta(u - v'), \tag{4.7}$$

$$\partial_{uv} G(u, v; u', v') = \delta(u - u') \delta(v - v'). \tag{4.8}$$

This is the unique Green function in the domain $v \geq u, v' \geq u'$ which vanishes at $v = u$, and whose support in the source point (u', v') lies in the past of the field point (u, v) . The general solution of (4.6) can then be written as

$$\psi(u, v) = \psi_{\text{in}}(u, v) + \int_{v' \geq u'} \int du' dv' G(u, v; u', v') \sigma(u', v'), \tag{4.9}$$

where $\psi_{\text{in}}(u, v)$ is the free “incoming” field, satisfying $\partial_{uv} \psi_{\text{in}} = 0$ and vanishing when $v = u$, i.e.

$$\psi_{\text{in}}(u, v) = g_{\text{in}}(v) - g_{\text{in}}(u). \tag{4.10}$$

The incoming waveform $g_{\text{in}}(v)$ is uniquely fixed by the incoming data on \mathcal{I}^- . For instance, for $r(u, v)$ asymptotic flatness gives $r_{\text{in}}(u, v) = \frac{1}{2}(v - u)$ (we recall that in flat space $v = t + r$ and $u = t - r$), while for $\bar{\phi}(u, v) \equiv \phi(u, v) - \phi_0$ the asymptotic expansion (3.24) (in which $f(v)$ was supposed to vanish as $v \rightarrow -\infty$) yields $\bar{\phi}_{\text{in}}(u, v) = 2(f(v) - f(u))/(v - u)$.

In principle, this perturbative algorithm allows one to compute the mass ratio $\mu(u, v) = 2m/r$ to any order in λ :

$$\mu(u, v) = \lambda^2 \mu_2 + \lambda^4 \mu_4 + \dots \tag{4.11}$$

Evidently, the convergence of this series becomes very doubtful when μ can reach values of order unity, which is precisely when an apparent horizon is formed. One would need some resummation technique to better locate the apparent horizon condition $\mu(v, u) = 1$. However, we find interesting to have, in principle, a way of explicitly computing successive approximations to the possible location of an apparent horizon, starting only from the incoming wave data $f(v)$.

4.2. Strength criterion at quadratic order

Let us compute explicitly the lowest-order approximation to (4.11) (we henceforth set $\lambda = 1$ for simplicity). It is obtained by inserting the zeroth-order result $r_0(u, v) = \frac{1}{2}(v - u)$, and the first-order one for ϕ ,

$$\bar{\phi}_1(u, v) = \phi_1(u, v) - \phi_0 = \frac{f(v) - f(u)}{r_0} = \frac{2[f(v) - f(u)]}{v - u}, \tag{4.12}$$

into Eq. (3.5) or Eq. (3.6). This yields (with $f'(v) \equiv \partial f/\partial v$)

$$\left(\frac{\partial m_2}{\partial v}\right)_u = r_0^2 \left(\frac{1}{2} \frac{\partial \phi_1}{\partial v}\right)^2 = \frac{1}{4} \left[f'(v) - \frac{f(v) - f(u)}{v - u} \right]^2, \tag{4.13}$$

$$\left(\frac{\partial m_2}{\partial u}\right)_v = -r_0^2 \left(\frac{1}{2} \frac{\partial \phi_1}{\partial u}\right)^2 = -\frac{1}{4} \left[f'(u) - \frac{f(v) - f(u)}{v - u} \right]^2. \tag{4.14}$$

The compatibility of the two equations is easily checked, e.g., by using

$$\partial_v (r_0^2 (\partial_u \phi_1)^2) = r_0 \partial_u \phi_1 \partial_v \phi_1 = -\partial_u (r_0^2 (\partial_v \phi_1)^2), \tag{4.15}$$

which expresses the “conservation” of the ϕ -energy tensor in the (u, v) plane. Noting that $m(u, v)$ must vanish (by regularity of $r = 0$ in the weak-field domain) at the centre $v = u$ one gets, by quadrature, the following explicit result for m_2 :

$$m_2(u, v) = \tilde{m}(u, v) - \frac{1}{8} r_0 \overline{\phi_1^2}, \tag{4.16}$$

where

$$\tilde{m}(u, v) \equiv \frac{1}{4} \int_u^v f'^2(x) dx. \tag{4.17}$$

Note that Eq. (4.16) decomposes $m_2(u, v)$ in a “conserved” piece $\tilde{m}(u, v) = g(v) - g(u)$ (with $\partial_u \tilde{m} = 0$) and a ϕ -dependent one. This is an integrated form of the local “conservation” law (4.15) of the ϕ -energy tensor. Using (4.16) one gets a very simple result for the quadratic approximation to the mass ratio $\mu = 2m/r$, namely

$$\mu_2(u, v) = \frac{1}{v - u} \int_u^v dx f'^2(x) - \left[\frac{f(v) - f(u)}{v - u} \right]^2. \tag{4.18}$$

Let us introduce a simple notation for the *average* of any function $g(x)$ over the interval $[u, v]$:

$$\langle g \rangle_{[u,v]} \equiv \langle g(x) \rangle_{x \in [u,v]} \equiv \frac{1}{v - u} \int_u^v dx g(x). \tag{4.19}$$

Then the mass ratio, at quadratic order, can be simply written in terms of the scalar news function $N(v) \equiv f'(v)$:

$$\mu_2(u, v) = \langle N^2 \rangle_{[u,v]} - (\langle N \rangle_{[u,v]})^2, \tag{4.20}$$

$$= \langle (N(x) - \langle N \rangle_{[u,v]})^2 \rangle_{x \in [u,v]}, \tag{4.21}$$

$$\equiv \text{Var} (N(x))_{x \in [u,v]}, \tag{4.22}$$

where $\text{Var} (g)_{[u,v]}$ denotes the “variance” of the function $g(x)$ over the interval $[u, v]$, i.e. the average squared deviation from the mean.

As indicated above, having obtained μ to second order, say

$$\mu_2 = V(u, v) \equiv \text{Var} (N)_{[u,v]},$$

one can then proceed to compute r and ϕ to higher orders. Namely, from Eq. (4.9),

$$2r(u, v) = v - u - G * \left[\frac{V(u, v)}{v - u} \right] + \mathcal{O}(\lambda^4), \tag{4.23}$$

$$(r\phi)(u, v) = f(v) - f(u) - G * \left[V(u, v) \frac{f(v) - f(u)}{(v - u)^2} \right] + \mathcal{O}(\lambda^5), \tag{4.24}$$

where the star denotes a convolution. Then, one can get from these results m_4 , and thereby μ_4 , by simple quadratures. In the following we shall only use the quadratic approximation to μ , though we are aware that when the initial data are strong enough to become gravitationally unstable, higher-order contributions to μ are probably comparable to μ_2 .

Finally, at quadratic order, we can define the strength of some initial data simply as the supremum of the variance of the news over arbitrary intervals,

$$\mathcal{S}_2 = \sup_{\substack{u, v \\ u \leq v}} \mu_2 = \sup_{x \in [u, v]} \text{Var}(N(x)). \tag{4.25}$$

At this order $\mathcal{S}_2 < 1$ means that $\mu_2(u, v)$ stays always below one, i.e. that no trapped region is created, and that the field is expected to disperse, while $\mathcal{S}_2 > 1$ signals the formation, somewhere, of trapped spheres, and therefore (by the results of Ref. [25]) the formation of a singularity. To be more exact we should actually define, in analogy with (4.25), a quantity $\mathcal{S}_{2p} = \sup \sum_{k=1}^p \mu_{2k}$, at any finite order in the weak-field expansion, and state our collapse criterion as the inequality:

$$\mathcal{S}_{2p} > C_{2p}, \quad C_{2p} \xrightarrow{p \rightarrow \infty} 1. \tag{4.26}$$

In practice, even if we do not expect it to be quantitatively exact, we will use the criterion (4.26) with $p = 1$, taking $C_2 = 1$, hoping that it will be qualitatively correct in capturing the features of the news function which are generically important for producing a gravitational collapse. Let us note that the functional \mathcal{S}_2 is (like N) dimensionless (and therefore scale invariant) and that it is non-local. We note also that it is invariant under a constant shift of N , $N(v) \rightarrow N(v) + a$, which corresponds to adding a linear drift in $f : f(v) \rightarrow f(v) + av + b$. Such a shift is (formally) equivalent, in view of Eq. (4.12), to a constant shift of $\phi_0 \rightarrow \phi_0 + 2a$.

The non-locality of \mathcal{S}_2 is physically interesting because it indicates that it is not the instantaneous level of the energy flux $N^2(v)$ which really matters, but rather the possibility of having a flux which varies by $\sim 100\%$ over some interval of advanced time. This non-locality defines also some characteristic scales associated with the collapse (when $\mathcal{S}_2 > 1$). Indeed, if, by causality, we consider increasing values of v , and define $\mathcal{S}_2(v) = \sup_{u, u \leq v} \text{Var}(N)_{[u, v]}$, the first value of v for which $\mathcal{S}_2(v)$ exceeds 1 defines an advanced characteristic time when the collapse occurs, and the corresponding maximizing interval $[u, v]$ defines a characteristic time scale of (in)homogeneity. (Note that $\mu_2(u, v)$, Eq. (4.22), vanishes when $u = v$, and vanishes also generally when $u \rightarrow -\infty$ if $N(u)$ tends to a limit at $-\infty$.)

Let us also mention that our strength functional (4.25) is superficially similar to the one recently introduced by Christodoulou [26] to characterize sufficiently weak (i.e. non-collapsing) data. In the lowest-order approximation his criterion measures the strength of the data by $\mathcal{V} = TV(N(x))$, where TV denotes the total variation (i.e. essentially $\mathcal{V} = \int_{-\infty}^{+\infty} dv |N'(v)|$). Like our variance, this is a measure of the variation of $N(v)$ with, however, a crucial difference. If \mathcal{V} were a good criterion for measuring the strength of possibly strong data, we would conclude that a news function which oscillates with a very small amplitude for a very long time will be gravitationally unstable, while our criterion indicates that it will not, which seems physically more plausible. We hope that our strength functional (4.25) will suggest new gravitational stability theorems to mathematicians.

4.3. Comparison with a collapse criterion of Christodoulou

To check the reasonableness of our quadratic strength criterion (4.25), (4.26) we have compared it with the rigorous, but only sufficient, collapse criterion of Christodoulou [25], and we have applied it to two simple exact solutions. Ref. [25] gives the following sufficient criterion on the strength of characteristic data considered at some finite retarded time u :

$$\frac{2\Delta m}{\Delta r} \geq \left[\frac{r_1}{r_2} \log \left(\frac{r_1}{2\Delta r} \right) + \frac{6r_1}{r_2} - 1 \right], \tag{4.27}$$

where $r_1 \leq r_2, r_2 \leq 3r_1/2$, are two spheres, $\Delta r = r_2 - r_1$ is the width of the “annular” region between the two spheres, and $\Delta m = m_2 - m_1 \equiv m(u, r_2) - m(u, r_1)$ is the mass “contained” between the two spheres, i.e. more precisely the energy flux through the outgoing null cone $u = \text{const.}$, between r_1 and r_2 . Explicitly, from Eq. (3.15) in (u, r) coordinates, we have

$$2\Delta m = \int_{r_1}^{r_2} dr \frac{r^2}{4} \left(1 - \frac{2m}{r} \right) \left(\frac{\partial \phi}{\partial r} \right)_u^2. \tag{4.28}$$

We can approximately express this criterion in terms of the incoming null data $N(v) = f'(v)$ if we assume that the outgoing cone $u = \text{const.}$ is in the weak-field domain, so that we can replace ϕ on the r.h.s. of Eq. (4.28) by $\phi = r^{-1} [f(u + 2r) - f(u)]$. Then the l.h.s. of Eq. (4.27) becomes (at quadratic order)

$$\frac{2\Delta m}{\Delta r} = \frac{M(u, v_2) - M(u, v_1)}{v_2 - v_1} \tag{4.29}$$

where

$$M(u, v) \equiv (v - u) \text{Var}(N)_{\{u,v\}} \equiv \int_u^v dx (N(x) - \langle N \rangle_{\{u,v\}})^2. \tag{4.30}$$

Therefore, in this approximation, the criterion of Ref. [25] becomes (with $v_0 = u$)

$$\exists v_0, v_1, v_2 : \frac{M(v_0, v_2) - M(v_0, v_1)}{v_2 - v_1} \geq C(v_0, v_1, v_2), \tag{4.31}$$

with the constraints $2v_2 - 3v_1 \leq -v_0, v_1 \leq v_2$, and the definition

$$C(v_0, v_1, v_2) \equiv \frac{v_1 - v_0}{v_2 - v_0} \log \left(\frac{v_1 - v_0}{2(v_2 - v_1)} \right) + \frac{6(v_1 - v_0)}{v_2 - v_0} - 1. \tag{4.32}$$

If v_0 is at the boundary of its allowed domain, i.e. if $v_0 = 3v_1 - 2v_2$, this criterion is fully compatible with ours since $C(v_0, v_1, v_2) = 3$ implying through Eq. (4.30)

$$\text{Var}(N)_{[v_0, v_2]} \geq 1 + \frac{2}{3} \text{Var}(N)_{[v_0, v_1]} \geq 1. \tag{4.33}$$

In the opposite case (v_0 large and negative) we think that the criterion (4.31) is also compatible, for generic news functions, with the general form suggested above, i.e.

$$\exists v'_1, v'_2 : \text{Var}(N)_{[v'_1, v'_2]} = \langle N^2 \rangle_{[v'_1, v'_2]} - \langle N \rangle_{[v'_1, v'_2]}^2 \geq C_2, \tag{4.34}$$

with some positive constant C_2 of order unity. We first note that one should probably impose the physically reasonable condition that $N(v)$ decays⁷ sufficiently fast as $v \rightarrow -\infty$ (faster than $|v|^{-1/2}$) to ensure that the integrated incoming energy flux $\int_{-\infty}^{v_0} dv N^2(v)$ is finite. (This constraint freezes the freedom to shift $N(v)$ by a constant.)

When $-v_0$ is large, the question remains, however, to know how fast the function $N(v)$ decays when $v_1 - v$ (for $v_0 < v < v_1$) becomes much larger than $v_2 - v_1$. Let us first consider the physically generic case where $N(v)$ decays in a reasonably fast manner, say faster than a power $|v|^{-\kappa}$ with κ of order unity. Then choosing v_0 in the criterion (4.31) just large enough to allow one to neglect $N(v_0)$ with respect to $N(v)$, say $(v_1 - v_0) \sim (v_2 - v_1) \times 10^{1/\kappa}$, corresponding to $N(v_0) = 0.1 N(v_1)$, the l.h.s. of Eq. (4.31) becomes approximately

$$\frac{2\Delta m}{\Delta r} \simeq \langle N^2 \rangle_{[v_1, v_2]}, \tag{4.35}$$

while the function C on the r.h.s., which grows only logarithmically with v_0 , becomes

$$C(v_0, v_1, v_2) \simeq \log \left(\frac{v_1 - v_0}{2(v_2 - v_1)} \right) + 5. \tag{4.36}$$

Therefore, a function of the type $N_\kappa(v) \sim c [(v_2 - v)/(v_2 - v_1)]^{-\kappa}$ (when $v \leq v_1 < v_2$) with $c^2 \gtrsim 5 + (\log 10)/\kappa$ will fulfill the criterion (4.31). To see whether this is compatible with our criterion (4.34) we have studied (analytically and numerically) the variance, over arbitrary intervals $[v'_1, v'_2]$, of the function $N_\kappa(v)$. We found that the inequality (4.34) is satisfied for a constant $C_2^{(\kappa)}$ which is of order unity if κ stays of order unity. (For instance, in the extreme case $\kappa = 1/2$ corresponding to a logarithmically divergent incoming energy flux, we find $C_2^{(1/2)} \sim 0.2$.) We note, however, that if one

⁷ Actually, when comparing our criterion (based on asymptotic null data) to that of Ref. [26] (based on characteristic data taken at a finite retarded time $u = v_0$) we can consider, without loss of generality, that the news function $N(v)$ vanishes identically for $v \leq v_0$.

considers extremely slow decays of $N(v)$, i.e. very small exponents $\kappa \rightarrow 0$ (completed by a faster decay, or an exact vanishing, before some large negative cutoff v_c), the constant $C_2^{(\kappa)}$ needed on the r.h.s. of the variance criterion (4.34) tends to zero. This signals a limitation of applicability of our simple variance criterion based only on the quadratic order approximation $\mu = \mu_2 + \mu_4 + \dots \simeq \mu_2$. Indeed, one can check that, in such an extreme situation μ_2 is abnormally cancelled, while μ_4 will be of order unity. However, we believe that for generic, non-extremely slowly varying news functions, the simple “rule of thumb” (4.34) is a reasonable approximation to the (unknown) exact collapse criterion. To further check the reasonableness of our criterion (4.25) we turn our attention to some exact solutions.

4.4. Exact solutions

A first exact, dynamical Einstein-dilaton solution is defined by a news function consisting of the simple step function

$$N(v) = p \theta(v) . \tag{4.37}$$

The corresponding solution has been independently derived by many authors [29]. Here, p is a real parameter and the value $|p| = 1$ defines the threshold for gravitational instability: no singularity occurs for $|p| < 1$, while $|p| \geq 1$ causes the birth of singularity. The metric for this solution takes the form (note that $\Omega = 1$)

$$ds^2 = -du dv + r^2 d\omega^2 , \tag{4.38}$$

with the solutions for ϕ , r and m given as follows:

$$\phi = 0 , \quad r = \frac{1}{2}(v - u) , \quad m = 0 , \quad \text{for } v < 0 , \tag{4.39}$$

while, for $v \geq 0$,

$$\phi(u, v) = -\log \left[\frac{(1 - p)v - u}{(1 + p)v - u} \right] , \tag{4.40}$$

$$\begin{aligned} r^2(u, v) &= \frac{1}{4} [(1 - p^2)v^2 - 2vu + u^2] \\ &= \frac{1}{4} [(1 - p)v - u] [(1 + p)v - u] , \end{aligned} \tag{4.41}$$

$$m(u, v) = -\frac{p^2 uv}{8r} , \tag{4.42}$$

$$\mu(u, v) = -\frac{p^2 uv}{[(1 - p^2)v^2 - 2vu + u^2]} . \tag{4.43}$$

Let us note that the perturbation-theory value of μ for $v > 0$, is

$$\mu_2(u, v) = \text{Var} [p \theta(x)]_{|u,v} = -\frac{p^2 uv}{(v - u)^2} \theta(-u) . \tag{4.44}$$

The maximum value of μ_2 , reached for $u = -v$, is $p^2/4$. This shows that, as expected, the quadratic order criterion (4.25) is only valid as an order of magnitude, but is

quantitatively modified by higher-order corrections. This example, and a related general theorem of Christodoulou [26], suggest that, if the exact criterion were of the type $\mathcal{S}_2 = C_2$, the constant C_2 should equal $\frac{1}{4}$. (Note that this value is also compatible with the constant $C_2^{(1/2)} \sim 0.2$ appropriate to $|v|^{-1/2}$ decay.)

A second exact solution [30,18] is a negative-curvature Friedmann-like homogeneous universe. It is defined by the following null data on \mathcal{I}^- (for $v < 0$)

$$f(v) = -\frac{\sqrt{3}}{v}, \quad N(v) = \frac{\sqrt{3}}{v^2}. \tag{4.45}$$

Here we view this solution as defined by incoming wave data in a flat Minkowski background. Actually the data are regular only in some advanced cone $v = T + R < 0$, and blow up when $v = 0$. Here, T and $R = \sqrt{X^2 + Y^2 + Z^2}$ are usual Minkowski-like coordinates in the asymptotic past. In terms of such coordinates the exact solution reads (when $T^2 - R^2 < 1$)

$$ds^2 = \left[1 - \frac{1}{(T^2 - R^2)^2} \right] [-dT^2 + dX^2 + dY^2 + dZ^2], \tag{4.46}$$

$$\phi = -\sqrt{3} \log \left(\frac{T^2 - R^2 - 1}{T^2 - R^2 + 1} \right). \tag{4.47}$$

Note that, for simplicity, we have set here $\phi_0 = 0$, and we have also set the length scale appearing in $f(v) = -\sqrt{3} \ell_0^2/v$ to $\ell_0 = 1$. We shall come back later to the cosmological significance of this solution. Let us only note here that, in terms of the null coordinates

$$u \equiv T - R, \quad v \equiv T + R, \tag{4.48}$$

the exact mass ratio reads

$$\mu(u, v) = \frac{uv(v-u)^2}{(u^2v^2-1)^2}, \tag{4.49}$$

while, starting from (4.45), one obtains

$$\mu_2(u, v) = \text{Var}(N)|_{u,v} = \frac{(v-u)^2}{u^3v^3}. \tag{4.50}$$

One finds that a strength criterion of the form $\sup_{u \leq v} \mu_2(u, v) = C_2$ is first satisfied (as one increases v from $-\infty$) when $v = v_{*2}$, with $u = u_{*2}$, where

$$v_{*2} = -\ell_0 \left(\frac{4}{27C_2} \right)^{1/4}, \quad u_{*2} = 3v_{*2}. \tag{4.51}$$

For C_2 of order 1, this is in qualitative agreement with the exact result that the apparent horizon is located at

$$r_{\mathcal{A}} = \frac{1}{v^3}, \quad r_{\mathcal{S}} = \frac{1}{2|v|^3} (1 - v^4)^{3/2}, \tag{4.52}$$

so that the apparent horizon and the singularity are first “seen” at $v_{* \text{ exact}} = -\ell_0$.

5. Transition from the weak-field to the cosmological regime

The following general picture emerges from the previous sections: Let us consider as “in state” a generic classical string vacuum, which can be described as a superposition of incoming wave packets of gravitational and dilatonic fields. This “in state” can be nicely parametrized by three asymptotic ingoing, dimensionless news functions $N(v, \theta, \varphi)$, $N_+(v, \theta, \varphi)$, $N_\times(v, \theta, \varphi)$. When all the news stay always significantly below 1, this “in state” will evolve into a similar trivial “out state” made of outgoing wave packets. On the other hand, when the news functions reach values of order 1, and more precisely when some global measure of the variation⁸ of the news functions, similar to the variance (4.25), exceeds some critical value of order unity, the “in state” will become gravitationally unstable during its evolution and will give birth to one or several black holes, i.e. one or several singularities hidden behind outgoing null surfaces (event horizons). Seen from the outside of these black holes, the “out” string vacuum will finally look, like the “in” one, as a superposition of outgoing waves. However, the story is very different if we look inside these black holes and shift back to the physically more appropriate string conformal frame. First, we note that the structure of black hole singularities in Einstein’s theory (with matter satisfying $p < \epsilon$) is a matter of debate. The work of Belinsky, Khalatnikov and Lifshitz [31] has suggested the generic appearance of an oscillating space-like singularity. However, the consistency of this picture is unclear as the infinitely many oscillations keep space being as curved (and “turbulent” [32,33]) as time. Happily, the basic gravitational sector of string theory is generically consistent with a much simpler picture. Indeed, it has been proven long ago by Belinsky and Khalatnikov [34] that adding a massless scalar field (which can be thought of as adding matter with $p = \epsilon$ as equation of state) drastically alters the BKL solution by ultimately quenching the oscillatory behaviour to end up with a much simpler, monotonic approach to a space-like singularity. When described in the string frame, the Einstein-frame collapse towards a space-like singularity will represent (if ϕ grows fast enough toward the singularity) a super-inflationary expansion of space. The picture is therefore that inside each black hole, the regions near the singularity where ϕ grows sufficiently fast will blister off the initial trivial vacuum as many separate pre-Big-Bangs. These inflating patches are surrounded by non-inflating, or deflating (i.e. without fastly increasing ϕ) patches, and therefore globally look like approximately closed Friedmann–Lemaître hot universes. This picture is sketched in Fig. 1 and Fig. 2. We expect such quasi-closed universes to recollapse in a finite, though very long, time (which is consistent with the fact that, seen from the outside, the black holes therein contained must evaporate in a finite time). To firm up this picture let us study in detail the appearance of the singularity in the simple toy model of Section 3.

⁸ The argument that the collapse criterion should be (at least) invariant under constant shifts of all the news functions (corresponding to classically trivial shifts of the background fields $g_{\mu\nu}^0$, ϕ_0) indicates that only the global changes of the news matter.

5.1. Negative-curvature Friedmann-dilaton solution

We have already seen that in the toy model there is an infinitely extended incoming region where the fields are weak and can be described as perturbations upon the incoming background values ($g_{\mu\nu}^0 = \eta_{\mu\nu}, \phi_0$). One expects that the perturbation algorithm described in Section 4.1 becomes unreliable when corrections become of order unity. In particular, one generically expects that the apparent horizon $\mu \equiv 2m/r = 1$ will roughly divide space-time in two domains: the perturbative incoming weak-field domain where $\mu(u, v) \ll 1$, and a strong-field domain where $\mu(u, v) \gg 1$. (Actually, we shall see below that the precise boundary of the domain of validity of perturbation theory may also depend on other quantities than just μ .) This separation in two domains is represented in Fig. 3.

For guidance let us study in detail this separation in two domains in the case of the $k = -1$ Friedmann-dilaton solution [30,18]. Let us first write explicitly the solution (4.46) and (4.47) in (u, v) coordinates

$$ds^2 = -\Omega^2(u, v) du dv + r^2 d\omega^2, \quad \phi(u, v) = -\sqrt{3} \log \left[\frac{uv - 1}{uv + 1} \right], \quad (5.1)$$

$$\Omega^2(u, v) = \left(1 - \frac{1}{u^2 v^2} \right), \quad r^2(u, v) = \frac{1}{4} \left(1 - \frac{1}{u^2 v^2} \right) (u - v)^2, \quad (5.2)$$

$$\mu(u, v) = \frac{uv(v - u)^2}{(u^2 v^2 - 1)^2}, \quad 1 - \frac{2m}{r} = \frac{(u^3 v - 1)(u v^3 - 1)}{(u^2 v^2 - 1)^2}. \quad (5.3)$$

The perturbative algorithm of Section 4.1 gives

$$\phi(u, v) \simeq \frac{2\sqrt{3}}{uv} \left(1 + \frac{1}{3} \frac{1}{u^2 v^2} \right), \quad (5.4)$$

$$r(u, v) \simeq \frac{1}{2} (v - u) \left(1 - \frac{1}{2} \frac{1}{u^2 v^2} \right), \quad (5.5)$$

which agree with the expansions of Eqs. (5.2) and (5.3). One sees that perturbation theory is numerically valid up to, say, $uv \leq \sqrt{2}$, at which point there is an abrupt transition towards the cosmological singularity located at $uv = 1$. Though the transition surface $uv \sim \sqrt{2}$ globally differs from the apparent horizon $\mu(u, v) = 1$, we note that our criterion points out to a specific event (u_{*2}, v_{*2}) on the apparent horizon \mathcal{A} (the point where \mathcal{A} is first “seen” from infinity) which lies, roughly, at the intersection of \mathcal{A} and of the transition surface. This confirms that our criterion is able, at least in order of magnitude, to correctly pinpoint when and where one should shift from the perturbative regime to a different, cosmological-type description. In the case of the solution Eqs. (5.1)–(5.3), one sees better the cosmological nature of the strong-field domain by introducing the following coordinates:

$$u = -e^{-\eta+\xi} = T - R, \quad v = -e^{-\eta-\xi} = T + R, \quad (5.6)$$

satisfying the useful relations

$$-T = e^{-\eta} \cosh \xi, \quad R = e^{-\eta} \sinh \xi, \tag{5.7}$$

$$\partial_u = e^{\eta-\xi} (\partial_\eta - \partial_\xi), \quad \partial_r = e^{\eta+\xi} (\partial_\eta + \partial_\xi), \tag{5.8}$$

$$e^{-2\eta} = uv, \quad e^{2\xi} = \frac{u}{v}. \tag{5.9}$$

In terms of the coordinates (η, ξ) the solution reads [30,18]

$$ds^2 = a^2(\eta) [-d\eta^2 + d\xi^2 + \sinh^2 \xi d\omega^2], \tag{5.10}$$

$$a^2(\eta) = 4 \cosh \eta \sinh(-\eta), \quad e^{-\phi} = \left(-\frac{\sinh \eta}{\cosh \eta} \right)^{\sqrt{3}}. \tag{5.11}$$

We note that the solution is regular in the domain $-\infty < \eta < 0, 0 \leq \xi < +\infty$ (which corresponds to the past of the hyperboloid $T^2 - R^2 = 1$ in Minkowski-like coordinates) and that a space-like singularity is reached at $\eta = 0$ (i.e. $uv = T^2 - R^2 = 1$). Let us also note the expressions

$$r^2(\xi, \eta) = a^2(\eta) \sinh^2 \xi = 4 \cosh \eta \sinh(-\eta) \sinh^2 \xi, \tag{5.12}$$

$$1 - \frac{2m}{r} = \sinh^2 \xi [\coth^2 \xi - \coth^2 2\eta], \tag{5.13}$$

telling us that the apparent horizon is located at $2\eta = -\xi$. Near the singularity, $\eta \rightarrow 0^-$, the dilaton blows up logarithmically while the metric coefficients have the following power-law behaviours:

$$\begin{aligned} \phi &\sim -\sqrt{3} \log(-\eta), & \Omega^2(\xi, \eta) \equiv a^2(\eta) &\sim -4\eta, \\ r^2(\xi, \eta) &\sim -4\eta \sinh^2 \xi. \end{aligned} \tag{5.14}$$

Before leaving this example we wish to emphasize some features of it which appear to follow from its homogeneity and could be misleading for the general case. The singular boundary terminates, in this case, on \mathcal{I}^- , instead of on \mathcal{I}^+ , as one generically expects. The apparent reason for this is the singularity in the flux at $v = 0$ which creates a future boundary on \mathcal{I}^- . We believe that, in this case, a better description of physics is obtained by going to new non-Minkowskian coordinates of Milne’s type (see Ref. [12]) which automatically incorporate the singularity on \mathcal{I}^- . However, if one restricts oneself to more regular initial data, having a finite integrated energy flux (generalized pulse-like data), the singular boundary should never come back to \mathcal{I}^- . (At most, it could end at space-like infinity i^0 , for non-integrable total energy flux).

5.2. Kasner-like behaviour of Einstein-dilaton singularities

Coming back to the generic case of an arbitrary Einstein-dilaton cosmological singularity, previous works [12] have shown that, in a suitable synchronous coordinate system (Gauss coordinates), the asymptotic behaviour of the string-frame metric near the singularity reads (in any space-time dimension D)

$$ds_S^2 \sim -dt_S^2 + \sum_{a=1}^{D-1} (-t_S)^{2\alpha_a(x)} (E_i^a(x) dx^i)^2, \tag{5.15}$$

$$\phi(x, t) = \phi(x, 0) + \sigma(x) \log(-t_S), \tag{5.16}$$

where $E_i^a(x)$ is some $(D - 1)$ -bein and where $(1 \leq a \leq D - 1)$

$$\sum_{a=1}^{D-1} \alpha_a^2 = 1, \quad \sigma = \left(\sum_{a=1}^{D-1} \alpha_a \right) - 1. \tag{5.17}$$

In the Einstein frame this asymptotic behaviour reads

$$ds_E^2 \sim -dt_E^2 + \sum_a (-t_E)^{2\lambda_a(x)} (e_i^a(x) dx^i)^2, \tag{5.18}$$

$$\phi(x, t) = \phi(x, 0) + \gamma(x) \log(-t_E), \tag{5.19}$$

where $e_i^a(x)$ is proportional to $E_i^a(x)$ and where

$$\sum_{a=1}^{D-1} \lambda_a = 1, \quad \gamma = \pm\sqrt{2} \sqrt{1 - \sum_{a=1}^{D-1} \lambda_a^2}. \tag{5.20}$$

The ‘‘Kasner’’ exponents, α_a in the string frame or λ_a in the Einstein frame, can vary continuously along the singularity. The string parametrization is the most global one as it shows that α_a runs freely over a unit sphere in \mathbb{R}^{D-1} while the exponent for e^ϕ is a linear function of the ‘‘vector’’ α_a . In the Einstein frame, the exponents λ_a are restricted by the linear equality $\sum_a \lambda_a = 1$ and the quadratic constraint $\sum_a \lambda_a^2 \leq 1$. A convenient geometric representation of these constraints, when $D = 4$, is to consider (in analogy with Mandelstam’s variables s, t, u) that the 3 λ_a ’s represent, in some Euclidean plane, the orthogonal distances of a point A from the three sides of an equilateral triangle (counted positively when A is *inside* the triangle). The quadratic constraint then means that A is restricted to stay inside the circle circumscribing the triangle. The sign ambiguity in Eq. (5.20) means that the parameter space is in fact a *two-sided disk*, namely the two faces of a coin circumscribed around the triangle. See Fig. 4.

The link between the two seemingly very different parameter spaces (the string-frame sphere parametrized by the α ’s and the Einstein-frame two-sided disk parametrized by the λ ’s) is geometrically very simple: one obtains the disk by a stereographic projection of the sphere, from the projection centre $(\alpha_a^c) = (1, 1, 1)$ (outside the sphere) onto the plane dual to this centre (i.e. the plane spanning the circle of tangency to the sphere of the straight lines issued from α^c). Algebraically (in any-space-time dimension D)

$$\lambda_a = \frac{\alpha_a(D - 2) - \sigma}{D - 2 - \sigma}, \quad \gamma = \frac{\sqrt{2(D - 2)} \sigma}{D - 2 - \sigma}, \tag{5.21}$$

where $\sigma \equiv (\sum_a \alpha_a) - 1$. Note that there is no sign ambiguity in the map $\alpha_a \rightarrow \lambda_a$. Coming back to the case $D = 4$, as α_a runs over a unit sphere in \mathbb{R}^3 , λ_a covers twice

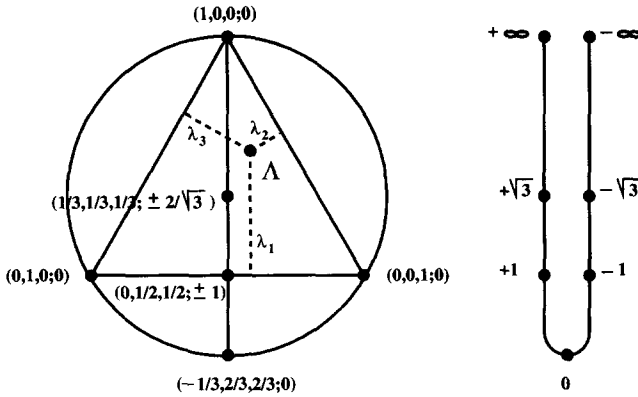


Fig. 4. Geometric representation of the Einstein-frame Kasner exponents $(\lambda_1, \lambda_2, \lambda_3; \gamma)$ of Einstein-dilaton cosmological singularities. The three λ 's, such that $\sum_a \lambda_a = 1$ are the orthogonal distances to the three sides of an equilateral triangle. The constraint $\gamma = \pm\sqrt{2} \sqrt{1 - \sum_a \lambda_a^2}$ restricts the representative Λ of λ_a to stay on a two-sided disk circumscribed around the triangle. The pure-Einstein cases ($\gamma = 0$) correspond to the circumscribing circle. In the spherically symmetric case ($\lambda_2 = \lambda_3$), Λ runs over a bissectrix of the triangle. The basic parameter α of Eqs. (5.24) and (5.33) runs over a full real line (shown folded on the right of the figure) and is mapped to the bissectrix via horizontal lines.

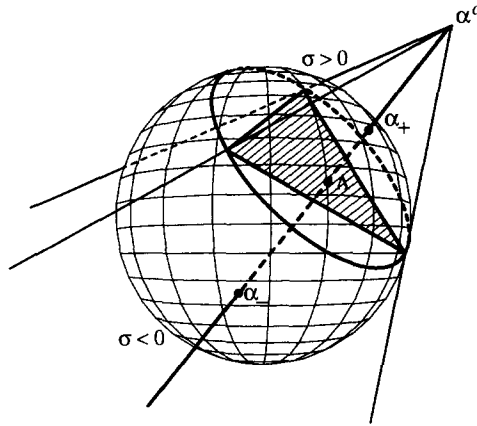


Fig. 5. Geometrical representation of the link between the two-sided disk of Fig. 4 on which the Einstein-frame Kasner exponents λ_a live, and the sphere, $\sum_a \alpha_a^2 = 1$, on which the string-frame Kasner exponents $(\alpha_1, \alpha_2, \alpha_3)$ live. The map $\alpha_a \rightarrow \lambda_a$ is a stereographic projection, from the centre $(\alpha_a^c) = (1, 1, 1)$. The upper side of the disk comes from the projection of the polar cap $\sigma \equiv (\sum_a \alpha_a) - 1 > 0$ located on the same side as α^c , while the lower side of the disk comes from projection of the antipolar cap $\sigma < 0$. The points α_+ on the polar cap, and α_- on the antipolar cap, are both projected on the same point Λ of the disk. The tangents to the sphere issued from α^c touch the sphere along the circle limiting the disk (on which $\sigma = 0$).

the disk circumscribed to an equilateral triangle. Each side of the disk is the image, via the stereographic projection, of a polar ($\sigma > 0$) or antipolar ($\sigma < 0$) “cap” of the sphere. This geometrical picture of the link between the α -sphere and the two-sided λ -disk is represented in Fig. 5. It is useful to keep in mind this geometrical picture because it shows clearly that one can continuously pass from one side of the disk to

the other, i.e. the sign of the Einstein-frame “Kasner” exponent of $e^\phi \sim (-t_E)^{\gamma(x)}$ (which is the same as the sign of its string-frame Kasner exponent σ , see Eq. (5.21)) can change as one moves along the singularity. Note also that the sign of the quantity $\sigma_0 \equiv \sum_a \alpha_a$ can change along the singularity. It is positive on the “upper hemisphere” in Fig. 5 (the half of the sphere directed toward α^c), and negative on the “lower hemisphere”. The sign of the quantity σ_0 is important because it determines whether (on average) space expands or not (in string units): if $\sigma_0 < 0$ (respectively > 0) the string-unit proper volume of a given coordinate spatial patch expands (respectively shrinks) toward the singularity. Because of the link $\sigma \equiv \sigma_0 - 1$ between σ_0 and the Kasner exponent of $g^2 = e^\phi \sim (-t_S)^\sigma$, a necessary condition for having an expanding space is to have a sufficiently fast growth of the string coupling: $\sigma < -1$. (In fact one needs a faster growth than that if one insists that all three dimensions of our space have expanded). Physically this means a very striking inhomogeneity near the singularity. Our mechanism generates at the singularity a “patchwork” of physically very different sub-universes (with different signs for various Kasner exponents). Only some of these patches (with sufficiently fast growing string coupling) can represent genuine pre-Big-Bangs. Finally, we note that Belinsky and Khalatnikov [34] have proven that, in the generic case, the Einstein-frame Kasner exponents $(\lambda_1, \lambda_2, \lambda_3)$ must all be positive for the ultimate stable asymptotic approach to the singularity. This corresponds to \mathcal{A} being inside the triangle of Fig. 4, i.e. in the hatched region of Fig. 5. However, this generic restriction does not apply in the special case of our toy model as we explain below.

5.3. Behaviour near the singularity in inhomogeneous, spherically symmetric solutions

Let us now restrict ourselves to our toy model and study its cosmological-like behaviour near the singularity. First, we note that spherical symmetry will impose that two of the metric Kasner exponents (those corresponding to the θ and φ directions) must be equal, say $\lambda_2 = \lambda_3$. Geometrically, this means that the Einstein-frame Kasner exponents run only over the intersection of the above two-sided disk with a bisectrix of the triangle. Second, we note that because of the vanishing, in spherical symmetry, of the crucial dreibein connection coefficient $e_1 \cdot (\nabla \times e_1)$, generically responsible for causing the expansion in the (radial) direction l to oscillate as $t \rightarrow 0^-$ [31], a negative value of λ_1 , corresponding to an expansion in the radial direction, is allowed as ultimate asymptotic behaviour. As a consequence, even the portion of the bisectrix outside the triangle is allowed as asymptotic state. We shall prove directly this fact below by constructing a consistent expansion near the singularity for λ_1 of any sign.

The mass-evolution equations (3.5), (3.6), together with Eqs. (3.11), (3.12) show that, starting from some regular event at the centre, $r = 0$, where $m = 0$, the mass $m(u, v)$ will grow if we follow a outgoing characteristic (i.e. a null geodesic) $u = \text{const}$. If this characteristic crosses the apparent horizon \mathcal{A} , Eqs. (3.11), (3.12) then show that $m(u, v)$ continues to grow within the trapped region \mathcal{T} ($2m/r > 1$) if one moves along *either* (future) outgoing or ingoing characteristics. This means that, within \mathcal{T} , $m(u, v)$

grows in all future directions. As the space-like singularity⁹ \mathcal{B} lies in the future of \mathcal{T} , we see that the mass function will necessarily grow toward \mathcal{B} , and therefore, either will tend to a finite limit on \mathcal{B} , or it will tend monotonically to $+\infty$ as $r \rightarrow 0$ with fixed v (or u). Motivated by the Kasner-like behaviours (5.18) and (5.19) we expect the following generic asymptotic behaviour of m on the singularity (here written in (v, r) coordinates)

$$m(v, r) \sim C_m(v) \left(\frac{r}{\ell_0}\right)^{-\alpha^2(v)} \quad \text{as } r \rightarrow 0, v \text{ fixed} \tag{5.22}$$

with some real v -dependent exponent $\alpha^2(v)$. (Here ℓ_0 denotes some convenient length scale.) It follows that $\mu = 2m/r$ always tends to $+\infty$ on \mathcal{B} (even when m tends to a finite limit). This suggests that we can consistently compute the structure of the fields near \mathcal{B} by using an “anti-weak-field” approximation scheme in which $\mu \gg 1$ (instead of the weak-field algorithm of Section 4.1 where we used $\mu \ll 1$). The expected domain of validity of such an “anti-weak-field”, or “cosmological-like”, expansion is sketched in Fig. 3.

To leading order in this large- μ limit, Eq. (3.15) yields

$$\frac{1}{4} \left(\frac{\partial\phi}{\partial \log r}\right)_v^2 = -\left(\frac{\partial \log m}{\partial \log r}\right)_v, \tag{5.23}$$

from which we obtain the following asymptotic behaviour of ϕ on \mathcal{B}

$$\phi(r, v) \sim 2\alpha(v) \log r + C_\phi(v). \tag{5.24}$$

We see now that the basic Kasner-like-exponent on \mathcal{B} is half the coefficient of $\log r$ in $\phi(r, v)$ as $r \rightarrow 0$, with fixed v . The Kasner-like exponent of $m(v, r)$ is the square of that basic exponent. Moreover, the function $C_m(v)$ appearing in the asymptotic behaviour of $m(v, r)$ on the singularity is not independent from the functions $\alpha(v)$ and $C_\phi(v)$. Indeed, to leading order, Eq. (3.16) reads

$$\frac{\partial \log m}{\partial v} = -\frac{1}{2} \frac{\partial\phi}{\partial \log r} \frac{\partial\phi}{\partial v}. \tag{5.25}$$

This relation is identically satisfied at order $r^0 \log r$, and yields at order r^0

$$\frac{\partial \log C_m}{\partial v} = -\alpha(v) \frac{\partial C_\phi}{\partial v}. \tag{5.26}$$

Using now Eq. (3.14), written as

$$\left(\frac{\partial\beta}{\partial \log r}\right)_v = \left(\frac{1}{2} \frac{\partial\phi}{\partial \log r}\right)_v^2, \tag{5.27}$$

⁹ We consider here only the “noncentral” singularity, i.e. the part of the singularity which can be reached by future outgoing characteristics issued from the regular centre $r = 0$. As shown by Ref. [26] there exists also, at the “intersection” between the regular centre and \mathcal{B} , a singularity reachable only via ingoing characteristics. Note that $r(u, v)$ tends to zero both at the centre and at the singularity (central or noncentral).

we further get the asymptotic behaviour of $\beta(v, r)$

$$\beta(v, r) \sim \alpha^2(v) \log r + C_\beta(v). \tag{5.28}$$

Finally, it is easy to see that the asymptotic behaviour of the metric on \mathcal{B} in (v, r) coordinates

$$ds^2 \sim \frac{2C_m(v)}{r} e^{2C_\beta(v)} \left(\frac{r}{\ell_0}\right)^{\alpha^2(v)} dv^2 + 2e^{C_\beta(v)} \left(\frac{r}{\ell_0}\right)^{\alpha^2(v)} dv dr + r^2 d\omega^2, \tag{5.29}$$

is indeed of the expected Kasner form (5.18) by introducing the cosmological time

$$(-t) \sim r \left(\frac{r}{\ell_0}\right)^{\frac{1}{2}(\alpha^2(v)+1)}. \tag{5.30}$$

The transformed metric reads, to dominant order as $r \rightarrow 0$, i.e. $t \rightarrow 0^-$,

$$ds^2 \sim -dt^2 + C_\rho(v) (-t)^{2\lambda_1(v)} \left(dv + \frac{r e^{-C_\beta} dr}{2C_m}\right)^2 + C_\omega(v) (-t)^{2\lambda_2(v)} [d\theta^2 + \sin^2 \theta d\varphi^2], \tag{5.31}$$

$$\phi \sim \gamma(v) \log(-t), \tag{5.32}$$

$$\lambda_1(v) = \frac{\alpha^2(v) - 1}{\alpha^2(v) + 3}, \quad \lambda_2(v) = \frac{2}{\alpha^2(v) + 3}, \quad \gamma(v) = \frac{4\alpha(v)}{\alpha^2(v) + 3}. \tag{5.33}$$

The basic parameter $\alpha(v)$ runs over the real line $-\infty < \alpha < +\infty$. One easily checks that the relations (5.20) are satisfied. Some cases having a particular significance are illustrated in Fig. 4: When $\alpha = 0$, one gets $(\lambda_1, \lambda_2, \lambda_3) = (-\frac{1}{3}, \frac{2}{3}, \frac{2}{3})$ and $\gamma = 0$ which describes a Schwarzschild-like singularity (when viewed with a cosmological eye). When $\alpha = \pm 1$, one gets $(\lambda_1, \lambda_2, \lambda_3) = (0, \frac{1}{2}, \frac{1}{2})$ and $\gamma = \pm 1$ which is locally similar to the scale-invariant solution [29] also described earlier. When $\alpha = \pm\sqrt{3}$, one gets $(\lambda_1, \lambda_2, \lambda_3) = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ and $\gamma = \pm 2/\sqrt{3}$ which describes an isotropic cosmological singularity of the type of the exact solution [30,18] we discussed above. As illustrated in Fig. 4, when α runs over \mathbb{R} its image, defined by Eq. (5.33), runs twice over the bisectrix of the Mandelstam triangle (intersected with the circumscribed disk). The only pure-Einstein case ($\gamma = 0$) reached at finite α is the Schwarzschild-like case $\alpha = 0$. In principle, another pure-Einstein behaviour occurs when $|\alpha| \rightarrow \infty$, leading to $(\lambda_1, \lambda_2, \lambda_3) = (1, 0, 0)$ i.e. $ds^2 \sim -dt^2 + t^2 d\rho^2 + d\omega^2 = -d\tau^2 + d\xi^2 + d\theta^2 + \sin^2 \theta d\varphi^2$. This would be an Einstein-cylinder-like universe. We believe, however, that $|\alpha(v)|$ will stay bounded ($|\alpha(v)| < \alpha_M < +\infty$) as v runs over \mathcal{B} and that the only Einstein-like case will be crossed when $\alpha(v)$ changes along \mathcal{B} .

We have confirmed our asymptotic analysis of the structure of the singularity (based on the ansätze (5.22), (5.24) and (5.28) in two ways. First, we have verified that it is consistent with the partial, but rigorous, results given in Ref. [26]. Namely, in (u, v) coordinates, our ansätze are compatible with the result that the function $r^2(u, v)$ is C^1

on \mathcal{B} if we posit that the conformal factor behaves like $\Omega^2 \sim r^{\alpha^2(v)-1}$ (while $m \sim r^{-\alpha^2(v)}$). A stronger check of our assumptions is obtained by going beyond the leading-order approximation, thus extending some results of [12]. In analogy to Section 4.1, where we set up a complete perturbation algorithm in the weak-field domain, we have shown that, starting from the leading-order terms, Eqs. (5.22), (5.24) and (5.28), containing three arbitrary “seed functions”, $\alpha(v)$, $C_\phi(v)$, $C_\beta(v)$ (from which $C_m(v)$ can be determined using Eq. (3.16)), it is possible to set-up an all-order iterative scheme generating a formal solution of the spherically symmetric Einstein-dilaton system in the strong-field, cosmological-like domain. Note that, $C_\beta(v)$ being a pure gauge function (it suffices to introduce $v' = \int e^{C_\beta(v)} dv$ to gauge $C_\beta(v)$ away), the formal solution thus generated contains two physically arbitrary functions of v , which is indeed the freedom that should be present in a generic solution.¹⁰ Details of this scheme are given in Appendix A. We only mention here the general structure of the expansions so generated near the singularity, i.e. as $r \rightarrow 0$, with fixed v : (see Eq. (A.34) in the appendix)

$$\phi(v, r) = \phi_0(v, r) + \sum_{n,m,l} a_{nml}^{(\phi)} r^{n(\alpha^2(v)+1)} r^{2m} P_l^{(\phi)}(\log r), \tag{5.34}$$

$$\beta(v, r) = \beta_0(v, r) + \sum_{n,m,l} a_{nml}^{(\beta)} r^{n(\alpha^2(v)+1)} r^{2m} P_l^{(\beta)}(\log r), \tag{5.35}$$

$$m(v, r) = m_0(v, r) + \sum_{n,m,l} a_{nml}^{(m)} r^{n(\alpha^2(v)+1)} r^{2m} P_l^{(m)}(\log r). \tag{5.36}$$

Here n, m, l are integers with $l \leq m$, $a_{nml}^{(\psi)}$ and the coefficients of the polynomials $P_l^{(\psi)}$ are functions of v , and (ϕ_0, β_0, m_0) are the leading-order terms given by Eqs. (5.22), (5.24) and (5.28). This expansion actually contains two intertwined series: an expansion in powers of $x = r^{1+\alpha^2}$ and a more complicated series in $r^{2m}(\log r)^l$. The second expansion is linked to the v -gradients of the seed functions $\alpha(v)$, $C_\phi(v)$, $C_\beta(v)$, while the first expansion is present even in the simple “homogeneous” case where α , C_ϕ and C_β do not depend on v . The fact, exhibited by Eqs. (5.34)–(5.36), that the expansion in the homogeneous case proceeds along powers of r^{α^2+1} confirms a recent conjecture by Burko [35].

5.4. Exact homogeneous cosmological solutions

Let us consider in more detail the “homogeneous” case where the seed functions α , C_ϕ and C_β have no spatial variation along the singularity. As it turns out, it is possible to resum exactly all the terms of the “homogeneous” expansion. Indeed, the analog of the Schwarzschild solution (t -independent, spherically symmetric solution) for the Einstein-dilaton system has been worked out analytically long ago by Just [36] (see also Ref. [37]), using the special gauge where $g_{00} g_{rr} = -1$. Contrary to the case of the

¹⁰ Here, contrary to what happened above, the variable v varies only on a half-line, $[v_*, +\infty]$ where v_* marks the birth of \mathcal{B} . As was said above two functions on a half-line (or one function on the full line) is the correct genericity in our toy model.

Schwarzschild solution, this solution cannot be continuously extended (through a regular horizon) down to a cosmological singularity at $r = 0$. However, it is easy to see that the following cosmological-like background (which is related to Just's original solution by formally extending the radial variable $\bar{r} = lx$ "below" the curvature singularity at $\bar{r} = a$) is still a solution of the Einstein-dilaton system,

$$ds^2 = l^2 \left[\left(\frac{1-x}{x} \right)^b dt^2 - \left(\frac{x}{1-x} \right)^b dx^2 + x^{1+b} (1-x)^{1-b} d\omega^2 \right], \quad (5.37)$$

$$\phi = \pm \sqrt{1-b^2} \log \left(\frac{1-x}{x} \right). \quad (5.38)$$

Here the (formerly radial) variable x varies between 0 and 1 and is "time-like", while the (formerly time) variable t is "space-like". The cosmological universe evolves from a Big Bang at $x = 0$ to a Big Crunch at $x = 1$. It has two arbitrary parameters, a scale parameter l , and the dimensionless b , $-1 \leq b \leq 1$. Near the singularity at $x = 0$, the link between b and the parametrization used above is

$$\alpha = \mp \sqrt{\frac{1-b}{1+b}}, \quad b = \frac{1-\alpha^2}{1+\alpha^2}. \quad (5.39)$$

Note that the behaviour near the other singularity at $x = 1$ is obtained by changing the sign of b and by changing α into $-1/\alpha$. Some cases of this homogeneous solution are of special significance: $b = 1$ is Schwarzschild (with $x = r/2m$), $b = 0$ belongs to the class of scale-invariant solutions [29] discussed above and $b = -1/2$ interpolates between a locally isotropic cosmological solution ($\alpha = \mp\sqrt{3}$) at $x = 0$ and an anisotropic one ($\alpha = \pm 1/\sqrt{3}$) at $x = 1$. Note that, in spite of the homogeneity and isotropy at $x = 0$, the latter special solution differs from the (everywhere) homogeneous-isotropic solution described earlier.

5.5. Discussion

Ideally, the two formal expansions we have constructed, the weak-field one, Eqs. (4.1)–(4.3), and the strong-field one, Eqs. (5.34)–(5.36), should match at some intermediate hypersurface, like the apparent horizon $2m/r = 1$ which looks like a natural borderline between the two near domains. If this matching were analytically doable, it would determine all the (so-far) arbitrary "seed functions" of the strong-field scheme in terms of the unique arbitrary function of the weak-field one, namely the asymptotic waveform $f(v)$ (or, equivalently, the news $N(v) = f'(v)$). However, it is clearly too naive to expect to perform this matching perturbatively: neither of the two expansions is expected to be convergent¹¹ (they are probably only asymptotic). Even if they converge on some domain they probably both break down before reaching a possible overlap region where they might be matched. At this stage, we can only state that, in principle,

¹¹ The possible presence of a finite number of BKL oscillations [34] near \mathcal{B} suggests that the formal strong-field expansion has very bad convergence properties.

all the seed functions of the strong-field scheme are some complicated non-linear and non-local functionals of $N(v)$. It would be particularly interesting to study the functional dependence of the Kasner exponent $\alpha(v)$ on $N(v)$. By causality (i.e. a domain of dependence argument) we know that $\alpha(v)$, at advanced time v , depends only on $N(v')$ on the interval $v' \leq v$. When starting from a generic $N(v)$ “of order unity”, we expect that the resulting $\alpha(v)$ will also be of order unity. A physically very important issue is the sign of $\alpha(v)$, i.e. the sign of $\gamma(v)$. Indeed $\alpha > 0$ means a decreasing ϕ , while $\alpha < 0$ means that ϕ grows near \mathcal{B} . Let us note that, to the lowest order of weak-field perturbation theory, the value of $\phi(v, r)$ at the (regular) centre ($r = 0$) is

$$\begin{aligned} \lim_{r \rightarrow 0} \phi_1(v, r) &= \lim_{r \rightarrow 0} \left[\phi_i + \frac{f(v) - f(v - 2r)}{r} \right] \\ &= \phi_i + 2f'(v) = \phi_i + 2N(v). \end{aligned} \quad (5.40)$$

Here ϕ_i denotes the background value at past infinity. From this result we expect that, if $N(v)$ is a simple Gaussian-like wave packet, $N(v) = A \exp[-(v - v_i)^2/\ell_i^2]$, with a large enough dimensionless amplitude for leading to collapse, i.e. $|A| \gtrsim 1$, the local values of ϕ near the collapsing central region will, at first, grow if $A > 0$, and decrease if $A < 0$. In this simple case, we therefore expect $\alpha(v)$ to have the opposite sign of A near the “central” region of \mathcal{B} , i.e. for $v \lesssim v_i$, where the gravitational instability sets in. But, further away on \mathcal{B} , i.e. for $v \gg v_i$, the sign of $\alpha(v)$ can change. We also expect that, as $v \rightarrow +\infty$, $\alpha(v)$ will tend to zero, corresponding to a Schwarzschild-like singularity (in the case of well-localized incoming $N(v)$ packets). This conjectured link between the sign of $\alpha(v)$ and the sign of $N(v)$ is confirmed by the exact isotropic solution equation (5.1) (with a positive growing $N(v) = \sqrt{3}/v^2$, and $\alpha = -\sqrt{3}$), as well as by the scale-invariant solution equation (4.40). Some numerical calculations [35] seem to confirm the general picture we propose (in particular the interesting possibility that the sign of $\alpha(v)$ changes several times on \mathcal{B} as v varies, before reaching a Schwarzschild-like asymptotic regime $|\alpha(v)| \rightarrow 0$ as $v \rightarrow +\infty$). We plan to study in more detail these issues in a future publication.

6. A Bayesian look at pre-Big-Bang’s “fine-tuning”

From its inception [38] it was pointed out that a successful pre-Big-Bang scenario must rely on a “reservoir” of inflationary e -folds during its perturbative phase. This is given by two small numbers, the initial curvature scale in string units, $H_i \ell_s$, and the initial string coupling constant $g_i^2 = e^{\phi_i} \lll 1$. (In this section, the index i is used for labelling “initial” quantities.) The need of large (or small) numbers has been recently discussed at length [18,19] and used to criticize the naturalness of the PBB scenario. In particular, it was pointed out [18,19] that, as soon as one goes beyond the simple spatially flat, homogeneous cosmology framework, the total duration of the perturbative dilaton-driven phase is finite so that the resolution of the homogeneity/flatness problems requires

$$g_i = e^{\phi_i/2} < 10^{-26}, \quad \frac{L_i}{\ell_s} \gtrsim \frac{M_s}{H_i} > 10^{19}. \quad (6.1)$$

Here, L_i denotes the spatial homogeneity scale of the PBB universe at the beginning of its inflationary phase, and $H_i \gtrsim L_i^{-1}$ its time-curvature scale (Hubble parameter).

In this Section we shall systematically work with string-frame quantities, even when we refer to results discussed in previous sections in the Einstein frame. For instance, the time-scale ℓ_i , characterizing the rate of variation of the news functions around the advanced time v_i and leading to gravitational instability, is now (locally) measured in string units. In any case, we are essentially working with dimensionless ratios which are unit-independent: $(\ell_i/\ell_s)_S = (\ell_i/\ell_s)_E = g_i(\ell_i/\ell_{\text{Planck}})_E$. We wish to emphasize here that, by combining our stochastic-like instability picture with a Bayesian approach to the a posteriori probability of being in a position of asking fine-tuning questions, the issue of the naturalness of the PBB scenario is drastically changed. By “Bayesian approach” we mean taking into account the selection effect that fine-tuning questions presuppose the existence of a scientific civilization. As emphasized long ago by Dicke [39] and Carter [40], civilization-related selection effects can completely change the significance of large numbers or of apparent coincidences. Linde and collaborators explored several aspects of the Dicke–Carter “anthropic principle” within the inflationary paradigm, and emphasized the necessity to weigh a posteriori probabilities by the physical volume of inflationary patches [41]. We shall here follow Vilenkin [42], and his “principle of mediocrity”, according to which the unnormalized a posteriori probability of a random scientific civilization to observe any values of the PBB parameters g_i , L_i and H_i is obtained by multiplying the corresponding a priori probability by the number of civilizations (over the whole of space and time) associated with the values g_i , L_i and H_i : $\mathcal{N}_{\text{civ}}(g_i, L_i, H_i)$.

6.1. A priori and a posteriori probability distributions for g_i and H_i

This approach can be applied to our case if we think of the initial past-trivial string vacuum as made of a more or less stochastic superposition of incoming waves (described by complicated news functions $N(v, \theta, \varphi)$, $N_+(v, \theta, \varphi)$, $N_-(v, \theta, \varphi)$ having many bumps, troughs and ramps). This stochastic bath of incoming waves will generate a rough sea of dilatonic and gravitational fields. If the input dimensionless wave forms can reach values of order unity, we expect that the local conditions for gravitational instability will be satisfied at several places in space and time. This will give birth to an ensemble of bubbling baby universes, with a more or less random distribution of initial parameters g_i and H_i , and with initial spatial homogeneity scales $L_i \gtrsim H_i^{-1}$. Indeed, the analysis of the previous sections has shown (in our toy model) that gravitational instability will set in on a spatio-temporal scale $\sim \ell_i$ when a rising wave of news function grows by $\mathcal{O}(1)$ on an advanced-time scale ℓ_i . From our variance criterion, Eq. (4.25), and the rough validity of weak-field perturbation theory nearly until the sharp transition to a cosmological-type behaviour, the work of the previous sections has shown that the initial, advanced-time scale ℓ_i is propagated via ingoing characteristics with little

deformation down to the strong-field domain (see Fig. 3), where it appears as a spatial homogeneity scale, i.e.

$$L_i \sim \ell_i. \tag{6.2}$$

From the leading-order result (5.40) we expect the local value of the Hubble parameter in the corresponding cosmological-like bubble to be

$$H_i \sim \sqrt{\dot{\phi}^2} \sim |N'(v_i)| \sim \frac{|N(v_i)|}{\ell_i}, \tag{6.3}$$

where v_i denotes the advanced time at which the instability sets in. Combining (6.2) and (6.3) we get

$$H_i L_i \sim |N(v_i)| \gtrsim 1. \tag{6.4}$$

These rough formulae show how, in principle, given the stochastic properties of the dimensionless news functions, one could deduce the distribution of H_i and L_i , naturally constrained by $H_i L_i \gtrsim 1$. The corresponding distribution of g_i is a priori independent from that of $\ell_i \sim L_i$, being linked to the presence of a slowly varying (“DC”) component N_0 in $N(v)$ (by contrast to the local variations of $N(v)$ leading to instability): $N(v) = N_0 + \text{fluctuations}$. The DC component N_0 corresponds to ramps in $f(v)$, $f(v) = N_0 v + f_0 + \text{fluctuations}$, i.e. to shifts of ϕ_i : $\phi_i \rightarrow \phi_i + 2N_0$. Finally, we can consider that the initial distribution of the news functions defines an a priori probability distribution for g_i , and H_i ,

$$dp_{\text{prior}} = w_i(g_i, L_i) dg_i dL_i, \tag{6.5}$$

with H_i generically of order L_i^{-1} (because $|N(v_i)| \sim 1$ is the threshold for instability).

An important remark must be made here. The two basic parameters we are talking about, say g_i and $L_i \sim H_i^{-1}$, precisely correspond to the two *global symmetries*, Eqs. (2.9)–(2.11) and Eq. (2.12), of the classical string vacua: a constant shift in ϕ , and a global coordinate rescaling. If we consider that the initial state of string theory is classical (rather than quantum), these symmetries mean that no particular values for ϕ_i or L_i are a priori preferred. One might even expect a “flat” distribution compatible with these global symmetries,

$$dp_{\text{prior}}^{\text{classical flat}} \propto \frac{dg_i}{g_i} \frac{dL_i}{L_i} \propto d\phi_i d \log L_i. \tag{6.6}$$

Evidently, the problem with such a flat distribution is that it is non-normalizable. Some cut-offs are needed to make sense of such a flat prior distribution but, while there are natural strong coupling/curvature cut-offs (the limits of validity of our approximation), it is not so easy to find natural small-curvature/coupling cut-offs. If we appeal to string theory for providing cut-offs for g_i and L_i , arguably the prior distribution (6.5) should be determined by the conjectured basic symmetries of string theory, i.e. by S - and T -duality. S -duality suggests that one should work with the complex quantity $z = \frac{\theta}{2\pi} + i \frac{4\pi}{g^2}$ and require modular invariance in $z = x + iy$. This selects the Poincaré

metric $ds^2 = (dx^2 + dy^2)/y^2$ as being special, and thereby defines a preferred measure in the fundamental domain (key-hole region) of the (x, y) plane: $d\mu = \sqrt{g} dx dy = dx dy/y^2 = dx |dy^{-1}|$. Integrating out the angular variable θ over $0 \leq \theta < 2\pi$ (which is the correct key-hole range when $g^2 < 4\pi$, i.e. $y > 1$) we end up with a preferred probability distribution for g : $dp(g) \propto dg^2 \propto g dg$, considered in the weak-coupling region $g \lesssim 1$. A similar argument based on T -duality selects as preferred probability distribution for the spatial length scale L : $dp(L) \propto d\widehat{L}^{-2} \propto \widehat{L}^{-3} d\widehat{L}$, where $\widehat{L} \equiv L/\ell_s$ is considered in the weak- σ -model-coupling region $\widehat{L} \gtrsim 1$. In conclusion, this type of argument would suggest the following normalizable, factorized prior distribution:

$$dp_{\text{prior}}^{S-T\text{-dual}} \propto (g_i dg_i) \widehat{L}_i^{-3} d\widehat{L}_i \propto g_i dg_i \widehat{H}_i d\widehat{H}_i, \quad (6.7)$$

with $0 < g_i \lesssim 1$, $0 < \widehat{H}_i \equiv H_i/M_s = \widehat{L}_i^{-1} \lesssim 1$.

For the sake of generality, we wish to leave open the nature of the prior distribution, and thus consider the general class of prior probability distributions:

$$dp_{\text{prior}} \propto g_i^a dg_i \widehat{H}_i^b d\widehat{H}_i, \quad (6.8)$$

with arbitrary powers a and b .

Any given initial distribution, say (6.8), will generate a corresponding ensemble of PBB inflationary bubbles. We assume, as usual, the existence of a successful “exit” mechanism by which dilaton-driven inflation, with growing g and H finally “exits” into a standard hot Big Bang, i.e. a radiation-dominated Friedmann universe. Current ideas about how this might happen [43] assume that the Hubble parameter reaches values of order of the string mass M_s , i.e. $\widehat{H}_s = H_s/M_s \sim 1$, before the string coupling constant reaches values of order unity. If the opposite were true ($g \sim 1$ occurring before $\widehat{H} \sim 1$), it is likely that quantum fluctuations would take over before H reaches the string scale, causing inhomogeneities to grow so large that the inflationary process “aborts” before a baby universe is born. In any event, we shall discard this possibility, assuming that the resulting cosmological universe would not evolve into anything able to harbour scientific civilizations.

After \widehat{H} reaches 1 at $g \ll 1$, the evolving universe is expected to be entering a so-called stringy De Sitter phase, i.e. a phase during which $\dot{\phi}$ and H are constant and given by some fixed-point values of the order of M_s : $|\dot{\phi}| = x M_s$, and $|H| = y M_s$ with $x \sim y \sim 1$. (See Ref. [43] for some suggestions on how to implement this mechanism.) Only the ratio between x and y enters our present phenomenological discussion. In keeping with a notation used in previous papers on PBB phenomenology, we introduce the (positive) parameter β by

$$|(\ln g) \dot{\phi}| = \frac{1}{2} |\dot{\phi}| = \beta H = \text{const}. \quad (6.9)$$

One finally assumes that this stringy De Sitter-like phase ends when $g = e^{\phi/2}$ reaches values of order unity, causing the amplified vacuum fluctuations to reach the critical density. After this moment, g should remain fixed at about its present value, either because of a non-perturbative ϕ -dependent potential, or possibly because of the attractor

mechanism of Ref. [6]. The fixing of g marks the end of the stringy modifications to Einstein’s theory, and the beginning of standard Friedmann-like cosmology. As we assume that g is attracted towards a unique fixed point, the resulting Friedmann universe will have no free parameter which could make it different from our universe, except for its initial homogeneity scale L_f , where the index f refers to the final state of the stringy phase. Therefore, the number of civilizations in the resulting Friedmann universe will be proportional to its total volume, at least for the large enough (in fact, old enough) universes that can harbour life

$$\mathcal{N}_{\text{civ}}(g_i, H_i) \propto \mathcal{V}_f(g_i, H_i) \theta[\mathcal{V}_f - \mathcal{V}_{\text{min}}] . \tag{6.10}$$

One assumes here, à la Dicke [39], that the time span during which civilizations can occur, being constrained by the lifetime of stars, is fixed. The minimum scale \mathcal{V}_{min} for a life-harboring universe is not known precisely, though it corresponds probably to universes whose total lifetime before recollapse is a few billion years. In conclusion, the a posteriori probability for a random scientific civilization to ask fine-tuning questions about the values of g_i and H_i is

$$dp_{\text{post}} \propto \mathcal{N}_{\text{civ}}(g_i, H_i) dp_{\text{prior}} \propto \mathcal{V}_f(g_i, H_i) \theta(\mathcal{V}_f - \mathcal{V}_{\text{min}}) dp_{\text{prior}} . \tag{6.11}$$

6.2. Computation of the a posteriori probability distribution in four dimensions

The computation of the final volume \mathcal{V}_f can be separated as

$$\mathcal{V}_f = L_f^3 = L_i^3 \frac{\mathcal{V}_1}{\mathcal{V}_i} z_s^3 . \tag{6.12}$$

Here, the index i refers to the (cosmological) time t_i of beginning of dilaton-driven inflation, the index 1 refers to the time t_1 of transition between dilaton-driven superinflation and a De Sitter-like phase, and $z_s = \exp(H(t_2 - t_1))$ denotes the total expansion during the De Sitter phase. The ending t_2 of the De Sitter-like phase is assumed to take place whence $g_2 \equiv g(t_2) \sim \mathcal{O}(1)$.

Using the leading power-law behaviour of the cosmological evolution during the dilaton-driven phase, Eq. (5.15), and imposing that this phase ends when $H_1 \sim (-t_1)^{-1} \sim M_s$, one gets

$$\frac{\mathcal{V}_1}{\mathcal{V}_i} \simeq \left(\frac{t_i}{t_s} \right)^{-\Sigma\alpha} \simeq (\hat{H}_i)^{\Sigma\alpha} , \tag{6.13}$$

where $\Sigma\alpha \equiv \alpha_1 + \alpha_2 + \alpha_3$ is the sum of the string-frame Kasner exponents. The α ’s are negative (expansion) and are constrained by Eq. (5.17). In the result (6.13) we have neglected a factor of order unity linked to the fact that the dilaton-driven inflation is generically anisotropic ($\alpha_1 \neq \alpha_2 \neq \alpha_3$) and therefore that the various Hubble expansion rates will reach the string scale at slightly different times. We assume here that a non-trivial basin of attraction from anisotropic expansion toward an isotropic stringy De Sitter-like phase exists as it is the case in examples [43].

During the subsequent stringy De Sitter phase, the growth of g is related, according to Eq. (6.9), to the spatial expansion by $g \propto a^\beta$ so that

$$z_s = \frac{a_2}{a_1} = \left(\frac{g_2}{g_1} \right)^{1/\beta} \sim g_1^{-1/\beta}, \tag{6.14}$$

where we used $g_2 \sim 1$. On the other hand, $g_1 = g(t_1)$ is expressible, through the use of Eq. (5.16), giving the evolution of g during the dilaton phase, as

$$e^{\phi_1 - \phi_i} = \left(\frac{g_1}{g_i} \right)^2 \simeq \left(\frac{t_1}{t_i} \right)^{2\alpha - 1} \sim \widehat{H}_i^{\Sigma\alpha - 1}. \tag{6.15}$$

Finally, we express, using $L_i \sim H_i^{-1}$, the volume at the beginning of standard cosmology in the form

$$\mathcal{V}_f(g_i, H_i) \sim \widehat{H}_i^{-A} [g_1(g_i, \widehat{H}_i)]^{-3/\beta} \theta(1 - g_1), \tag{6.16}$$

with

$$g_1(g_i, \widehat{H}_i) \sim g_i \widehat{H}_i^{-B}, \tag{6.17}$$

and where the exponents A and B are positive, and are given in terms of the α 's by

$$A = 3 - \Sigma\alpha, \quad B = \frac{1}{2} - \frac{1}{2} \Sigma\alpha. \tag{6.18}$$

If we assume, for illustration, an a priori distribution of values of g_i and \widehat{H}_i of the form (6.8), we get for the a posteriori probability distribution of g_i and \widehat{H}_i

$$dp_{\text{post}} \propto g_i^a \widehat{H}_i^{b-A} [g_i \widehat{H}_i^{-B}]^{-3/\beta} \theta(1 - g_1(g_i, \widehat{H}_i)) \theta(\mathcal{V}_f - \mathcal{V}_{\text{min}}) dg_i d\widehat{H}_i, \tag{6.19}$$

which can be written as

$$dp_{\text{post}} \propto g_i^{\bar{a}} \widehat{H}_i^{\bar{A}} \theta(1 - g_1) \theta(\mathcal{V}_f - \mathcal{V}_{\text{min}}) dg_1 d\widehat{H}_i. \tag{6.20}$$

In the last form we have expressed the measure dp_{post} in terms of the two independent variables g_1 and \widehat{H}_i (this is convenient because of the constraint $g_1 < 1$ carried by the step function). The exponents \bar{a} and \bar{A} appearing there are

$$\bar{a} \equiv a - \frac{3}{\beta}, \tag{6.21}$$

$$\bar{A} = A - (a + 1) B - b = 3 + \frac{1}{2} (a - 1) (\Sigma\alpha) - \frac{1}{2} (a + 1) - b. \tag{6.22}$$

The numerical values of the exponents \bar{a} and \bar{A} play a crucial rôle in determining the a posteriori plausibility of our universe having evolved from the seemingly “unnatural” values (6.1). Indeed, if $\bar{A} \geq 1$ the posterior probability distribution for \widehat{H}_i is peaked, in a non-integrable manner, at $\widehat{H}_i = 0$. Therefore, if $\bar{A} \geq 1$, it is natural to expect (within the present scenario) that the initial homogeneous patch, whose gravitational instability led to our universe, be extremely large compared to the string scale (and that the value of $g_i = \widehat{H}_i^B g_1$ with $B > 0$ and $g_1 < 1$ be correspondingly small). Most scientific civilizations are bound to evolve in such universes.

Note that the duality-suggested values $a = 1$ and $b = 1$ of Eq. (6.8) lead to $\bar{A} = 1$ (independently of the α 's) for which such an a posteriori explanation works. From this point of view one can argue that, within string theory, it is “natural” to observe very small numbers such as in Eq. (6.1). The exponent \bar{a} also plays an important rôle. Indeed, if $\bar{a} > -1$, i.e. if

$$\beta > \frac{3}{a+1}, \tag{6.23}$$

the a posteriori probability distribution for g_1 is integrable over its entire possible range $(0, 1)$. This means that most scientific civilizations are expected to observe values $g_1 = \mathcal{O}(1)$, corresponding to a small number of e -folds during the stringy phase: $z_s \sim g_1^{-1/\beta} = \mathcal{O}(1)$. This is a phenomenologically interesting case, because it means that various interesting physical phenomena taking place during the dilaton-driven phase (such as the quantum amplification of various fields) might leave observable imprints on cosmologically relevant scales. (A very long string phase would essentially iron out all signals coming from the dilaton phase.) It is interesting to note that, in the string-duality-inspired case (6.7), the great divide between a short stringy phase ($\bar{a} > -1$), and a very long one ($\bar{a} \leq -1$, implying that g_1 is peaked in a non-integrable manner at $g_1 = 0$), lies at $\beta = 3/2$, which played already a special rôle in previous phenomenological studies [8].

Though the case $\bar{A} \geq 1$, $\bar{a} > -1$ appears as the conceptually most interesting case for the pre-Big-Bang scenario, other civilization-related selection effects might also render the case $\bar{A} < 1$ viable. In this case, the factors $\widehat{H}_i^{-\bar{A}} d\widehat{H}_i$ in Eq. (6.20) suggests that \widehat{H}_i should be of order unity (i.e. $L_i \sim H_i^{-1} \sim \ell_s$). However, in such a case, one must take into account the factor $\theta(\mathcal{V}_f - \mathcal{V}_{\min})$ (which could be neglected in the previous discussion). This factor means that scientific civilizations should a posteriori expect to find themselves in the smallest possible universe compatible with their appearance. This probably means that they should also expect to see inhomogeneities comparable to the Hubble scale, and to have appeared very late in the cosmological evolution, just before the universe recollapses. This seems to conflict with our observations. In conclusion, within this model, and limiting our discussion for simplicity to the class (6.8) with a and b integers, one finds the favorable situation $\bar{A} \geq 1$ realized (i.e. no a posteriori unnaturalness in observing very small g_i and \widehat{H}_i) when either $a = 0$ and $b \leq 2$, or $a = 1$ and $b \leq 1$. However, if $a \geq 2$, i.e. if the a priori distribution function of g_i vanishes faster than g_i^2 for $g_i \rightarrow 0$, the pre-Big-Bang scenario has to face a naturalness problem. (This is the case for quantum fluctuations of g_i , as discussed at the end of the next subsection.)

6.3. Computation of the a posteriori probability distribution in D dimensions

Assuming that the results discussed so far are qualitatively correct in higher dimensions, we have generalized the above considerations to the case in which the initial string vacuum is extended to d spatial dimensions (say $d = 9$), and where, through

gravitational instability, three of these dimensions are expanding and 6 are collapsing down to the string scale. This leads to introducing four special times: t_i (beginning of dilaton phase), $t_1 \sim -t_s$ (end of the dilaton-driven power-law evolution of space), t_2 (when the radius of curvature of the collapsing dimensions become $\mathcal{O}(\ell_s)$), and t_3 (when $g_3 = \mathcal{O}(1)$). In this scenario there are two stringy De Sitter-like phases: a first one during which the three “external” dimensions grow exponentially $a \propto \exp(H_1 t)$, while the $(d - 3)$ “internal” ones *shrink* exponentially $\bar{a} \propto \exp(-\bar{H}_1 t)$ (and while $\dot{\phi} = 2\beta_1 H_1 = 2\bar{\beta}_1 \bar{H}_1$), and a second phase during which the $(d - 3)$ string-scale-curved internal dimensions have frozen while the three external ones continue to grow exponentially $a \propto \exp(H_2 t)$ (and while $\dot{\phi} = 2\beta_2 H_2$). This scenario gives a result of the same form as above, namely

$$\mathcal{V}_f(g_i, H_i) \sim \widehat{H}_i^{-A} \left[g_2(g_i, \widehat{H}_i) \right]^{-3/\beta_2} \theta(1 - g_2), \tag{6.24}$$

with

$$g_2(g_i, \widehat{H}_i) = g_i \widehat{H}_i^{-B}, \tag{6.25}$$

and where

$$A = d - 3\alpha - (d - 3)\bar{\alpha} + (1 - \bar{\alpha})\bar{\beta}_1 \left(\frac{3}{\beta_1} - \frac{d - 3}{\bar{\beta}_1} \right), \tag{6.26}$$

$$B = \frac{1}{2} - \frac{3}{2}\alpha - \frac{1}{2}(d - 3)\bar{\alpha} + (1 - \bar{\alpha})\bar{\beta}_1. \tag{6.27}$$

Here one has assumed for simplicity that, during the dilaton phase, the three external dimensions grow with the same exponent.¹² $\alpha < 0$, i.e. proportionally to $(-t)^\alpha$, while the $(d - 3)$ internal ones decrease with the same exponent $\bar{\alpha} > 0$, proportionally to $(-t)^{\bar{\alpha}}$. Then, one can derive the analog of Eq. (6.20) with g_1 replaced by g_2 and with exponents

$$\bar{a} = a - \frac{3}{\beta_2}, \quad \bar{A} = A - (a + 1)B - b. \tag{6.28}$$

Once more, $\bar{A} \geq 1$ leads to no fine-tuning and $\bar{a} > -1$ to a short string phase. We shall not attempt a complete phenomenological discussion of the range of values of a and b compatible with $\bar{A} \geq 1$. (This would imply exploring the parameter space of allowed values for the Kasner exponents α and $\bar{\alpha}$.) Let us only note here that the introduction of more spatial dimensions generically helps because A (and \bar{A}) now contains the contribution $+d = +9$, say, instead of the previous $+3$. In this case, even distribution functions vanishing faster than g_i^2 (for $g_i \rightarrow 0$) can make very small values of g_i be a posteriori preferred. The power-law enhancement brought by the volume factor \mathcal{V}_f would be inadequate, however, for compensating an exponential suppression as $g_i \rightarrow 0$, say $\propto \exp(-c/g_i^2)$. In particular, the a priori distribution of g_i and H_i cannot be the one

¹² This string-frame Kasner exponent α should not be confused with the basic parameter $\alpha(v)$ introduced in Section 5.3.

expected (from the Lagrangian $\mathcal{L} \sim e^{-\phi}(\nabla\phi)^2 \sim (\nabla g^{-1})^2$) for quantum fluctuations of g on scales $L_i \sim H_i^{-1}$ around the trivial ground state. Indeed, this distribution vanishes for small g_i and small \hat{H}_i as $\exp(-c g_i^{-2} \hat{H}_i^{-2})$ with c of order unity. This confirms the standard idea of the pre-Big-Bang scenario, according to which the initial state should be a *classical* string vacuum, i.e. an arbitrary classical solution of the low-energy field equations.

7. Conclusion

We have proposed a new, concrete picture for the typical initial state of the Universe which can give rise to a cosmology of the so-called pre-Big-Bang type. We conceived this state as a very generic, past-trivial string vacuum, i.e. classically, as the most general solution of the low-energy tree-level string action endowed with a perturbative region in the asymptotic past. Such states can be parametrized (in three spatial dimensions) by three dimensionless “news functions”: a scalar news $N(v, \theta, \varphi)$ and two helicity-2 news $N_+(v, \theta, \varphi)$ and $N_-(v, \theta, \varphi)$. By studying the simple spherically symmetric tensor-scalar model we have argued that such a generic “in state” will be gravitationally unstable if (any of) the news functions vary (varies) by something of order unity on some scale $\Delta v \sim \ell$. More precisely, in the spherically symmetric model, we found, at the lowest order of perturbation theory, that the criterion for gravitational instability is given by the supremum over v_1 and v_2 of the *variance* of the function $N(x)$ over the interval $[v_1, v_2]$, see Eq. (4.25). When this supremum exceeds a threshold of order unity, the wave packet “sent” between the retarded times v_1 and v_2 will become gravitationally unstable when, later, it becomes focussed at the centre.¹³ We have verified that the infinite region between past infinity and such a centre can be described by weak-field perturbation theory. Then, near the centre, in a region of spatial extension $L_i \sim \Delta v \equiv v_2 - v_1$, one can abruptly shift from a weak-field description to a cosmological one. One can then describe the further evolution of the patch L_i by a complementary cosmological-like expansion of the type introduced by Belinsky, Khalatnikov and Lifshitz (BKL) and recently developed, for the PBB scenario, in [12]. We have proven to all orders of iteration (in our toy model) that the presence of the dilaton ensured the consistency of a (non-oscillatory) BKL expansion in log-corrected powers of the distance to the future (space-like) singularity. This result confirms the smoothing behaviour of a dilaton-driven superinflationary evolution. In the string frame, gravitationally unstable patches of vacuum, with homogeneity scale L_i , will expand (if $\dot{\phi}$ is sufficiently large and positive) into homogeneous patches on much larger scales. These blistering patches are surrounded by contracting regions, where $\dot{\phi}$ is not sufficiently large and positive or is even negative. Many such dilaton-driven pre-Big-Bang inflationary bubbles, looking like closed universes, can blister off a generic past-trivial string vacuum. By taking

¹³ In the absence of spherical symmetry, there may exist many “local centres” where gravitational instability sets in. They should each correspond to an annular section of \mathcal{I}^- with a $\mathcal{O}(1)$ fluctuation of the news looking roughly spherically symmetric when viewed from some centre and in some suitably boosted Lorentz frame.

into account the selection effects linked to the presence of scientific civilizations, we have shown that it can appear a posteriori probable to observe that our universe comes from a very large fluctuation $L_i \gg \ell_s$, with a very small g_i . In particular, we pointed out that string dualities suggest an a priori distribution function for g_i and L_i which is phenomenologically encouraging in predicting a short, intermediate De Sitter-like stringy phase, and the a posteriori naturalness of observing $L_i \gg \ell_s$ and $g_i \ll 1$.

Our scenario contains new features that might suggest new observable signatures of the pre-Big-Bang idea. In particular, we expect the dilaton-driven phase to be generically anisotropic. It will be interesting to study the possible fossils of this anisotropy. These could be of two kinds: (i) a macroscopic anisotropy of space-time which is neither washed out by the string phase (in particular if, as we argued, such a phase should be short) nor by thermalization during the radiation era, and (ii) an anisotropy in some of the amplified quantum fluctuations due to their exit during the anisotropic dilaton-dominated phase (for a recent computation see Ref. [44]). Some of these perturbations may isotropize after re-entry, if they are sufficiently coupled to the rest of matter, but others, like gravitons and axions, may be already decoupled at re-entry. In this case, if their spectra are sufficiently flat (e.g. for axions) they can contribute to CMB anisotropies in some distinctive way. We plan to come back to these important questions in some future work.

Our work suggests also new types of criteria for the gravitational instability of the Einstein-plus-scalar system. Concentrating on some non-cumulative measure of the variation of the news might lead to new theorems of the type of those of Refs. [25,26]. One of the most interesting well-posed technical problems suggested by our work is the study of the link between the incoming news function and the spatial dependence (along the singularity) of the Kasner exponents. We hope to study this problem in the near future.

Acknowledgements

We thank Valodia Belinsky, Demetrios Christodoulou, Nemanja Kaloper, Sergiu Klainerman, Andrei Linde and Slava Mukhanov for useful exchanges of ideas. Discussions with Amit Gosh, Federico Piazza, Giuseppe Pollifrone and Eliezer Rabinovici are also acknowledged.

Appendix A

In this appendix we will establish the structure of the perturbative expansion near the singularity in the spherically symmetric model analysed in the text. We found it more convenient to use the (v, r) coordinate system introduced in Section 3. The field equations read

$$2 \left[\frac{\partial^2 \phi}{\partial v \partial r} + \frac{1}{r} \left(\frac{\partial \phi}{\partial v} \right) \right] + e^\beta \left(1 - \frac{2m}{r} \right) \left(\frac{\partial^2 \phi}{\partial r^2} \right) + \frac{2e^\beta}{r} \left(1 - \frac{m}{r} \right) \left(\frac{\partial \phi}{\partial r} \right) = 0, \tag{A.1}$$

$$\left(\frac{\partial \beta}{\partial r} \right) = \frac{r}{4} \left(\frac{\partial \phi}{\partial r} \right)^2, \quad 2 \left(\frac{\partial m}{\partial r} \right) = \left(1 - \frac{2m}{r} \right) \frac{r^2}{4} \left(\frac{\partial \phi}{\partial r} \right)^2. \tag{A.2}$$

Let us expand the different fields (ϕ, β, m) around the leading-order solutions (5.22), (5.24) and (5.28)

$$\phi_0(v, r) = C_\phi(v) + 2\alpha(v) \log r, \tag{A.3}$$

$$\beta_0(v, r) = C_\beta(v) + \alpha^2(v) \log r, \tag{A.4}$$

$$m_0(v, r) = C_m(v) r^{-\alpha^2(v)}, \tag{A.5}$$

and define the corrections to it $\bar{\phi}, \bar{\beta}$ and \bar{m} by

$$\phi(v, r) = \phi_0(v, r) + \bar{\phi}(v, r), \tag{A.6}$$

$$\beta(v, r) = \beta_0(v, r) + \bar{\beta}(v, r), \tag{A.7}$$

$$m(v, r) = m_0(v, r) (1 + \bar{m}(v, r)). \tag{A.8}$$

Introducing the variable $e^\lambda = r$, and denoting differentiation with respect to v by a prime, Eqs. (A.1), (A.2) take the form

$$\frac{\partial^2 \bar{\phi}}{\partial \lambda^2} = S_\phi(v, \lambda) + L_\phi(v, \lambda) + NL_\phi(v, \lambda), \tag{A.9}$$

$$S_\phi = \frac{e^{-C_\beta}}{2C_m} \left[4\alpha' (1 + \lambda) e^{2\lambda} + 2C'_\phi e^{2\lambda} + 2\alpha e^{C_\beta} e^{(\alpha^2+1)\lambda} \right], \tag{A.10}$$

$$L_\phi = \frac{e^{-C_\beta}}{C_m} \left[e^{2\lambda} \frac{\partial \bar{\phi}'}{\partial \lambda} + e^{2\lambda} \bar{\phi}' + \alpha \bar{\beta} e^{C_\beta} e^{(\alpha^2+1)\lambda} + \frac{e^{C_\beta}}{2} e^{(\alpha^2+1)\lambda} \left(\frac{\partial^2 \bar{\phi}}{\partial \lambda^2} + \frac{\partial \bar{\phi}}{\partial \lambda} \right) \right], \tag{A.11}$$

$$NL_\phi = \frac{1}{2C_m} \left[e^{(\alpha^2+1)\lambda} (e^{\bar{\beta}} - 1) \left(\frac{\partial^2 \bar{\phi}}{\partial \lambda^2} + \frac{\partial \bar{\phi}}{\partial \lambda} \right) - 2C_m (e^{\bar{\beta}} - 1) (1 + \bar{m}) \frac{\partial^2 \bar{\phi}}{\partial \lambda^2} + 2\alpha e^{(\alpha^2+1)\lambda} (e^{\bar{\beta}} - 1 - \bar{\beta}) - 2C_m \bar{m} \frac{\partial^2 \bar{\phi}}{\partial \lambda^2} \right]; \tag{A.12}$$

$$\left(\frac{\partial \bar{\beta}}{\partial \lambda} \right) - \alpha \left(\frac{\partial \bar{\phi}}{\partial \lambda} \right) = S_\beta(v, \lambda) + L_\beta(v, \lambda) + NL_\beta(v, \lambda), \tag{A.13}$$

$$S_\beta = 0 = L_\beta, \quad NL_\beta = \frac{1}{4} \left(\frac{\partial \bar{\phi}}{\partial \lambda} \right)^2; \tag{A.14}$$

$$\left(\frac{\partial \bar{m}}{\partial \lambda} \right) + \alpha \left(\frac{\partial \bar{\phi}}{\partial \lambda} \right) = S_m(v, \lambda) + L_m(v, \lambda) + NL_m(v, \lambda), \tag{A.15}$$

$$S_m = \frac{\alpha^2 e^{(\alpha^2+1)\lambda}}{2C_m}, \quad L_m = \frac{\alpha e^{(\alpha^2+1)\lambda}}{2C_m} \left(\frac{\partial \bar{\phi}}{\partial \lambda} \right), \tag{A.16}$$

$$NL_m = \frac{e^{(\alpha^2+1)\lambda}}{8C_m} \left(\frac{\partial\bar{\phi}}{\partial\lambda}\right)^2 - \frac{1}{4} \left(\frac{\partial\bar{\phi}}{\partial\lambda}\right)^2 - \alpha\bar{m} \left(\frac{\partial\bar{\phi}}{\partial\lambda}\right) - \frac{\bar{m}}{4} \left(\frac{\partial\bar{\phi}}{\partial\lambda}\right)^2. \tag{A.17}$$

In the r.h.s.'s of Eqs. (A.9), (A.13) and (A.15) we have indicated by S the field-independent terms (“source terms”), by L the terms that are linear in the fields, but subleading with respect to the l.h.s. in the limit $r \rightarrow 0$ ($\lambda \rightarrow -\infty$), while by NL we refer to generic non-linear contributions.

By examining the source terms and the subleading linear terms, we see that they all contain either an explicit factor $e^{(\alpha^2+1)\lambda} \equiv r^{(\alpha^2+1)}$ or a factor $e^{2\lambda} \equiv r^2$. To keep track of the presence of these factors (which tend to zero at the singularity) let us introduce two formal bookkeeping parameters: σ_1 associated with each occurrence of $e^{(\alpha^2+1)\lambda}$ in S and L and, likewise, σ_2 associated with $e^{2\lambda}$. In the end we shall set $\sigma_1 = \sigma_2 = 1$, but the introduction of these parameters will be useful to delineate the structure of the expansion of the fields near $r = 0$.

Introducing the notation $\psi = (\bar{\phi}, \bar{\beta}, \bar{m})$, considered as a column matrix, and denoting D_λ the following derivative operator with respect to λ :

$$D_\lambda = \begin{pmatrix} \partial_\lambda^2 & 0 & 0 \\ -\alpha(v)\partial_\lambda & \partial_\lambda & 0 \\ +\alpha(v)\partial_\lambda & 0 & \partial_\lambda \end{pmatrix}, \tag{A.18}$$

we can rewrite Eqs. (A.9), (A.13) and (A.15) in the compact matrix form

$$D_\lambda\psi = \sigma_1 S_\psi^{(1)} + \sigma_2 S_\psi^{(2)} + \sigma_1 L_\psi^{(1)} + \sigma_2 L_\psi^{(2)} + NL_\psi. \tag{A.19}$$

The non-linear terms in (A.19) depend linearly on σ_1 and σ_2 : $NL_\psi(\psi, \partial_\lambda\psi, \partial_\lambda^2\psi, \sigma_1, \sigma_2) = NL_\psi^{(0)} + \sigma_1 NL_\psi^{(1)} + \sigma_2 NL_\psi^{(2)}$. The differential matrix system (A.19) can be rewritten as an integro-differential system by introducing the “retarded” Green function $G(\lambda, \lambda')$, i.e. the unique inverse of D_λ which vanishes when $\lambda' > \lambda$,

$$\psi = G * \left(\sigma_1 S_\psi^{(1)} + \sigma_2 S_\psi^{(2)} + \sigma_1 L_\psi^{(1)} + \sigma_2 L_\psi^{(2)} + NL_\psi \right), \tag{A.20}$$

where we indicate by $*$ the operation of integrating over λ' . This choice of boundary conditions is imposed by consistency: We defined the background solution (A.3)–(A.5) as giving the asymptotic behaviour of (ϕ, β, m) when $r \rightarrow 0$, i.e. $\lambda \rightarrow -\infty$. We must therefore require that the deviation field $\psi(\lambda)$ vanish when $\lambda \rightarrow -\infty$.

The system (A.20) can be formally solved by successive iterations, leading to a solution ψ in the form of a double power series in the bookkeeping parameters σ_1 and σ_2 ,

$$\psi = \sum_{n,m \geq 1} \sigma_1^n \sigma_2^m \psi_{nm}. \tag{A.21}$$

For instance, the first-order (in σ_1 and σ_2) corrections $\psi_1 = \sigma_1 \psi_{10} + \sigma_2 \psi_{01}$, i.e. the next-to-leading-order terms in Eqs. (A.6)–(A.8), are obtained by neglecting the L and NL terms. They satisfy the equations

$$\frac{\partial^2 \phi_1}{\partial \lambda^2} = S_\phi(v, \lambda), \tag{A.22}$$

$$\left(\frac{\partial \beta_1}{\partial \lambda}\right) - \alpha(v) \left(\frac{\partial \phi_1}{\partial \lambda}\right) = S_\beta(v, \lambda), \tag{A.23}$$

$$\left(\frac{\partial m_1}{\partial \lambda}\right) + \alpha(v) \left(\frac{\partial \phi_1}{\partial \lambda}\right) = S_m(v, \lambda). \tag{A.24}$$

The unique solution of this system which vanishes when $\lambda \rightarrow -\infty$ reads

$$\phi_1(v, \lambda) = \int_{-\infty}^{\lambda} d\lambda' \int_{-\infty}^{\lambda'} d\lambda'' S_\phi(v, \lambda'') = \int_{-\infty}^{\lambda} d\lambda' (\lambda - \lambda') S_\phi(v, \lambda'), \tag{A.25}$$

$$\beta_1(v, \lambda) = \alpha(v) \phi_1(v, \lambda) + \int_{-\infty}^{\lambda} d\lambda' S_\beta(v, \lambda'), \tag{A.26}$$

$$m_1(v, \lambda) = -\alpha(v) \phi_1(v, \lambda) + \int_{-\infty}^{\lambda} d\lambda' S_m(v, \lambda'). \tag{A.27}$$

The form of this solution defines the action $G * S \equiv \int d\lambda' G(\lambda, \lambda') S(\lambda')$ of the retarded matrix Green function $G(\lambda, \lambda')$ on any source term $S(\lambda') = (S_\phi(\lambda'), S_\beta(\lambda'), S_m(\lambda'))$. As a check on the consistency of the iteration method, note that, if one had not imposed any boundary conditions, the generic solution ψ_1 would have contained an arbitrary homogeneous solution ψ_{hom} (“zero-mode”). However, it is easy to check that the addition of such a generic ψ_{hom} would be absorbable in suitable redefinitions of the “seed” functions $\alpha(v), C_\phi(v), C_\beta(v), C_m(v)$ entering the background solution equations (A.3)–(A.5). Finally, the explicit form of the next-to-leading order solution (with $\sigma_1 = \sigma_2 = 1$) is

$$\phi_1(v, \lambda) = \frac{\alpha}{(\alpha^2 + 1)^2 C_m} e^{(\alpha^2 + 1)\lambda} + \frac{C'_\phi e^{-C_\beta}}{4C_m} e^{2\lambda} + \frac{\alpha' e^{-C_\beta}}{2C_m} \lambda e^{2\lambda}, \tag{A.28}$$

$$\beta_1(v, \lambda) = \alpha \phi_1(v, \lambda), \tag{A.29}$$

$$m_1(v, \lambda) = -\alpha \phi_1(v, \lambda) + \frac{\alpha^2}{2(\alpha^2 + 1)C_m} e^{(\alpha^2 + 1)\lambda}. \tag{A.30}$$

Note that the coefficients of the terms proportional to $e^{2\lambda}$ and $\lambda e^{2\lambda}$ contain derivatives with respect to v , hence they determine the inhomogeneous expansion near the singularity, while the terms proportional to $e^{(\alpha^2 + 1)\lambda}$ are present also in the purely homogeneous case.

Inserting the double power series expansion (A.21) into (A.19) or (A.20) yields, for each contribution ψ_{nm} ,

$$D_\lambda \psi_{nm} = s_{nm}, \quad \psi_{nm} = G * s_{nm}, \tag{A.31}$$

where the right-hand-side s_{nm} plays, at each order, the rôle of an *effective source term* which is (in principle) known, being some algebraico-differential expression in the lower-order terms $\psi_{n'm'}$, with $n' + m' < n + m$. Let us make the inductive assumption that s_{nm} is a sum of terms of the form

$$s_{nml} = e^{n(\alpha^2+1)\lambda} e^{2m\lambda} Q_l(\lambda), \tag{A.32}$$

with $l \leq m$, where $Q_l(\lambda)$ is a polynomial in λ of order l . (Here, n, m, l are all integers.) It is straightforward to check, from the explicit action (A.25)–(A.27) of the Green function on source terms, that the solution ψ_{nml} corresponding to any term s_{nml} , Eq. (A.32), in the r.h.s. of Eq. (A.31) is again of the form

$$\psi_{nml} = G * s_{nml} = e^{n(\alpha^2+1)\lambda} e^{2m\lambda} \tilde{Q}_l(\lambda), \tag{A.33}$$

where $\tilde{Q}_l(\lambda)$ is another polynomial in λ , of the same order l . Examining the explicit form of the L and NL terms on the r.h.s. of Eqs. (A.9), (A.13) and (A.15), one can then prove that the induction hypothesis consistently propagates at the next order in $n + m$ (still with the constraint $l \leq m$).

In conclusion, coming back to the variable r , we find that the inhomogeneous expansion near the singularity contains two intertwined series: an expansion in powers of $r^{(\alpha^2+1)}$ and one in powers of r^2 mixed with powers of $\log r$. The solution to all orders will be of the form

$$\psi = \sum_{n, m, l \leq m} a_{nml} r^{n(\alpha^2+1)} r^{2m} P_l(\log r), \tag{A.34}$$

where a_{nml} , and the coefficients of the polynomial P_l , are functions of v that in principle can be iteratively evaluated. As shown in Section 5.4, the full subseries $\sum_n a_{n00} r^{n(\alpha^2+1)}$ can be exactly summed in terms of the variable x ($x \rightarrow 0^+$) such that $r^2 = l^2 x^{1+b} (1-x)^{1-b}$ with $b = (1 - \alpha^2)/(1 + \alpha^2)$ (see Eqs. (5.37), (5.38)). On the other hand, the terms with $m \neq 0$ are more complicated, as they contain more and more v -derivatives of $\alpha(v)$, $C_\phi(v)$ and $C_\beta(v)$ as m increases. We note, in passing, that this monotonous increase of v -derivatives shows that the series in m (and l) (inhomogeneous expansion) cannot be convergent for generic (non-analytic) seed functions $\alpha(v)$, $C_\phi(v)$ and $C_\beta(v)$.

To complete the proof that the algorithm defining ψ solves all the initial field equations (3.14)–(3.17), we finally need to check that the $\bar{m}(v, r)$ constructed above satisfies the constraint (3.16) involving $\partial m / \partial v$. Let us denote the quantity which must vanish as

$$\mathcal{E} \equiv \frac{1}{m} \left\{ 2 \left(\frac{\partial m}{\partial v} \right)_r - e^{-\beta} \frac{r^2}{2} \left[\left(\frac{\partial \phi}{\partial v} \right)_r^2 + e^\beta \left(1 - \frac{2m}{r} \right) \left(\frac{\partial \phi}{\partial r} \right)_v \left(\frac{\partial \phi}{\partial v} \right)_r \right] \right\}. \tag{A.35}$$

It is easily checked that, when the other field equations (3.14), (3.15) and (3.17) are satisfied, \mathcal{E} fulfills the following identity (which could also be derived from the Bianchi identities):

$$\frac{\partial \mathcal{E}}{\partial r} + A(v, r) \mathcal{E} \equiv 0, \quad A(v, r) = \frac{r^2}{8m} \left(\frac{\partial \phi}{\partial r} \right)^2. \quad (\text{A.36})$$

Noting that the coefficient $A(v, r)$ is bounded near $r = 0$ ($A(v, r) \sim r^{\alpha^2(v)}$), and that the relation (5.26) ensures the vanishing of \mathcal{E} as $r \rightarrow 0$, we conclude from the identity (A.36) (viewed as an ODE in r with data given at $r = 0$) that \mathcal{E} vanishes everywhere.

References

- [1] M.B. Green, J.H. Schwarz and E. Witten, *Superstring Theory* (Cambridge University Press, Cambridge, 1987).
- [2] G. Veneziano, *Quantum Strings and the Constants of Nature*, in *The Challenging Questions*, Erice, 1989, ed. A. Zichichi (Plenum, New York, 1990).
- [3] J. Scherk and J.H. Schwarz, *Nucl. Phys. B* 81 (1974) 118.
- [4] T. Damour, *Gravitation and Experiment*, in *Critical Problems in Physics*, ed. V.L. Fitch et al. (Princeton University Press, Princeton, 1997) pp. 147–166 [gr-qc/9711061];
T. Damour and G. Esposito-Farèse, *Phys. Rev. D* 58 (1998) 042001.
- [5] T.R. Taylor and G. Veneziano, *Phys. Lett. B* 213 (1988) 459.
- [6] T. Damour and A.M. Polyakov, *Nucl. Phys. B* 423 (1994) 532; *Gen. Rel. Grav.* 26 (1994) 1171;
T. Damour and A. Vilenkin, *Phys. Rev. D* 53 (1996) 2981.
- [7] G. Veneziano, *Phys. Lett. B* 265 (1991) 287;
M. Gasperini and G. Veneziano, *Astropart. Phys.* 1 (1993) 317; *Mod. Phys. Lett. A* 8 (1993) 3701;
Phys. Rev. D 50 (1994) 2519.
- [8] An updated collection of papers on the pre-Big-Bang scenario is available at the following address:
<http://www.to.infn.it/~gasperin/>.
- [9] B.A. Campbell, A.D. Linde and K.A. Olive, *Nucl. Phys. B* 355 (1991) 146;
R. Brustein and P.J. Steinhardt, *Phys. Lett. B* 302 (1993) 196.
- [10] For reviews of inflation, see A. Linde, *Particle Physics and Inflationary Cosmology* (Harwood, 1990);
E.W. Kolb and M.S. Turner, *The Early Universe* (Addison-Wesley, Redwood City, CA, 1990).
- [11] G. Veneziano, *Phys. Lett. B* 265 (1991) 287;
A.A. Tseytlin, *Mod. Phys. Lett. A* 6 (1991) 1721;
K.A. Meissner and G. Veneziano, *Phys. Lett. B* 267 (1991) 33; *Mod. Phys. Lett. A* 6 (1991) 3397;
A. Sen, *Phys. Lett. B* 271 (1991) 295;
S.F. Hassan and A. Sen, *Nucl. Phys. B* 375 (1992) 103;
A.A. Tseytlin and C. Vafa, *Nucl. Phys. B* 372 (1992) 443;
M. Gasperini and G. Veneziano, *Phys. Lett. B* 277 (1992) 256.
- [12] G. Veneziano, *Phys. Lett. B* 406 (1997) 297;
A. Buonanno, K.A. Meissner, C. Ungarelli and G. Veneziano, *Phys. Rev. D* 57 (1998) 2543;
J. Maharana, E. Onofri and G. Veneziano, *JHEP* 4 (1998) 4.
- [13] A. Feinstein, R. Lazkoz and M.A. Vazquez-Mozo, *Phys. Rev. D* 56 (1997) 5166;
J.D. Barrow and K.E. Kunze, *Phys. Rev. D* 56 (1997) 741.
- [14] C. Lovelace, *Phys. Lett. B* 135 (1984) 75;
E.S. Fradkin and A.A. Tseytlin, *Nucl. Phys. B* 261 (1985) 1;
C.G. Callan, E.J. Martinec, M.J. Perry and D. Friedan, *Nucl. Phys. B* 262 (1985) 593;
A. Sen, *Phys. Rev. Lett.* 55 (1985) 1846.
- [15] V.P. Frolov, M.A. Markov and V.K. Mukhanov, *Phys. Lett. B* 216 (1989) 272; *Phys. Rev. D* 41 (1990) 383.
- [16] V.K. Mukhanov and R. Brandenberger, *Phys. Rev. Lett.* 68 (1992) 1969;
R. Brandenberger, V.K. Mukhanov and A. Sornborger, *Phys. Rev. D* 48 (1993) 1629.
- [17] L. Smolin, *Class. Quant. Grav.* 9 (1992) 173.
- [18] M.S. Turner and E.J. Weinberg, *Phys. Rev. D* 56 (1997) 4604.

- [19] N. Kaloper, A. Linde and R. Bousso, Pre-Big-Bang requires the Universe to be exponentially large from the very beginning, [hep-th/9801073].
- [20] H. Bondi, M.G.J. van der Burg and A.W.K. Metzner, Proc. Roy. Soc. London A 269 (1962) 21; R.K. Sachs, Proc. Roy. Soc. A 270 (1962) 103; R. Penrose, Proc. Roy. Soc. London A 284 (1965) 159.
- [21] C. Itzykson and J.B. Zuber, Quantum Field Theory (McGraw-Hill, New York, 1980) Chapter 4.
- [22] A. Buonanno and T. Damour, Phys. Lett. B 432 (1998) 51.
- [23] A. Ghosh, G. Pollifrone and G. Veneziano, Phys. Lett. B 440 (1998) 51.
- [24] D. Christodoulou, Comm. Math. Phys. 105 (1986) 337.
- [25] D. Christodoulou, Comm. Pure Appl. Math. 54 (1991) 339.
- [26] D. Christodoulou, Comm. Pure Appl. Math. 56 (1993) 1131.
- [27] R. Penrose, Phys. Rev. Lett. 14 (1965) 57.
- [28] S.W. Hawking and R. Penrose, Proc. Roy. Soc. London A 314 (1970) 529.
- [29] P.S. Wesson, J. Math. Phys. 19 (1978) 2283; M.D. Roberts, Gen. Rel. Grav. 21 (1989) 907; R.A. Sussman, J. Math. Phys. 32 (1990) 223; D. Christodoulou, Comm. Pure Appl. Math. 56 (1993) 1131; Y. Oshiro, K. Nakamura and A. Tomimatsu, Prog. Theor. Phys. 91 (1994) 1265.
- [30] E.J. Copeland, A. Lahiri and D. Wands, Phys. Rev. D 50 (1994) 4868.
- [31] V.A. Belinskii and I.M. Khalatnikov, Sov. Phys. JETP 30 (1970) 1174; I.M. Khalatnikov and E.M. Lifshitz, Phys. Rev. Lett. 24 (1970) 76; V.A. Belinskii, E.M. Lifshitz and I.M. Khalatnikov, Sov. Phys. Uspekhi 13 (1971) 745.
- [32] A.A. Kirillov and A.A. Kochnev, Pis'ma Zh. Eksp. Teor. Fiz. 46 (1987) 345 (JETP Lett. 46 (1987) 435).
- [33] V.A. Belinskii, Pis'ma Zh. Eksp. Teor. Fiz. 56 (1992) 437 (JETP Lett. 56 (1992) 421).
- [34] V.A. Belinskii and I.M. Khalatnikov, Sov. Phys. JETP 36 (1973) 591.
- [35] L.M. Burko, Phys. Rev. D 58 (1998) 084013.
- [36] K. Just, Z. Nat. 14 (1959) 751.
- [37] R. Coquereaux and G. Esposito-Farèse, Ann. Inst. H. Poincaré 52 (1990) 113; T. Damour and G. Esposito-Farèse, Class. Quant. Grav. 9 (1992) 2093.
- [38] G. Veneziano, Phys. Lett. B 265 (1991) 287.
- [39] R.H. Dicke, Nature 192 (1961) 440.
- [40] B. Carter, in I.A.U. Symposium 63: Confrontation of Cosmological Theories with Observational Data, ed. M.S. Longair (Reidel, Dordrecht, 1974) p. 291.
- [41] A. Linde, Ref. [1] in Ref. [10] (particularly Chapter 10), and references therein; A.D. Linde and A. Mezhlumian, Phys. Lett. B 307 (1993) 25; A.D. Linde, D.A. Linde and A. Mezhlumian, Phys. Rev. D 49 (1994) 1783; J. García-Bellido, A. Linde and D. Linde, Phys. Rev. D 50 (1994) 730 (see especially the appendix); J. García-Bellido and A. Linde, Phys. Rev. D 51 (1995) 429.
- [42] A. Vilenkin, Phys. Rev. Lett. 74 (1995) 846.
- [43] M. Gasperini, M. Maggiore and G. Veneziano, Nucl. Phys. B 494 (1997) 315; R. Brustein and R. Madden, Phys. Lett. B 410 (1997) 110; Phys. Rev. D 57 (1998) 712.
- [44] M. Giovannini, Blue spectra of Kalb–Ramond axions and fully anisotropic string cosmologies, [hep-th/9809185].