

CHAOS, SYMMETRY AND STRING COSMOLOGY

THIBAUT DAMOUR

*Institut des Hautes Études Scientifiques,
35 route de Chartres, 91440 Bures-sur-Yvette, France*

We review the finding that the general solution of many physically relevant (bosonic) Einstein-matter systems in the vicinity of a cosmological, i.e. space-like, singularity exhibit a remarkable mixture of chaos and symmetry. On the one hand, the asymptotic behavior, as one approaches the singularity, of the general solution is found to be describable, at each (generic) spatial point, as a billiard motion in an auxiliary Lorentzian space. After a suitable “radial” projection, this evolution is described by a billiard motion on hyperbolic space (which is “chaotic” in many physically interesting cases). On the other hand, for certain Einstein-matter systems, notably for pure Einstein background (Ricci tensor = 0) in any space-time dimension D and for particular Einstein-matter systems arising in string theory, the billiard tables describing asymptotic cosmological behavior are found to be identical to the Weyl chambers of some Lorentzian Kač–Moody algebras. In the case of the bosonic sector of supergravity in 11-dimensional space-time the underlying Lorentzian algebra is that of the hyperbolic Kač–Moody group E_{10} , and there exists some evidence of a correspondence between the general solution of the Einstein-three-form system and a null geodesic in the infinite-dimensional coset space $E_{10}/K(E_{10})$, where $K(E_{10})$ is the maximal compact subgroup of E_{10} .

924 *Thibault Damour***Table of Contents**

1	Introduction and general overview	925
2	Models and Gauge Conditions	932
3	Asymptotic dynamics in the general case	934
3.1	Hamiltonian action	936
3.2	Appearance of sharp walls in the BKL limit	940
3.3	Constraints	944
3.4	Consistency of BKL behavior in spite of the increase of spatial gradients	945
4	Walls	947
4.1	Symmetry walls	947
4.2	Curvature (gravitational) walls	948
4.3	p -form walls	950
4.3.1	Electric walls	950
4.3.2	Magnetic walls	951
4.4	Subleading walls	953
5	Einstein (or cosmological) billiards	953
6	Kač–Moody theoretic formulation: the E_{10} case	956
	References	965

Ian Kogan was always interested in “confluences” between a priori different physical properties. The paradoxical “confluence” between Chaos and Symmetry appealed to him. In one of his last works (in collaboration with Dima Polyakov) [1], he suggested the presence of a chaotic behavior in the superstring dilaton β function in ghost-matter mixing backgrounds. I dedicate this review to Ian, in memory of the many enlightening physics discussions I was privileged to have with him.

1. Introduction and general overview

A remarkable connection between the asymptotic behavior of certain Einstein-matter systems near a cosmological singularity and billiard motions in the Weyl chambers of some corresponding Lorentzian Kač–Moody algebras was uncovered in a series of works [2–8]. This simultaneous appearance of *billiards* (with *chaotic* properties in important physical cases) and of an underlying *symmetry* structure (infinite-dimensional Lie algebra) is an interesting fact, which deserves to be studied in depth. Before explaining in detail the techniques that have been used to uncover this fact, we will start by reviewing previous related works, and by stating the main results of this billiard/symmetry connection.

The simplest example of this connection concerns the pure Einstein system, i.e. Ricci tensor = 0. Long ago, Belinskii, Khalatnikov and Lifshitz (BKL) gave a description [9–11] of the asymptotic behavior, near a space-like singularity, of the general solution of Ricci tensor = 0 in $(3 + 1)$ -dimensional space-time in terms of a continuous collection of second-order, nonlinear ordinary differential equations (with respect to the time variable). As argued by these authors, near the singularity the spatial points x^i , $i = 1, 2, 3$, essentially decouple, and this decoupling allows one to approximate a system of *partial* differential equations (PDE’s) in 4 variables (t, x^i) by a three-dimensional family, parametrized by $(x^i) \in \mathbb{R}^3$, of *ordinary* differential equations (ODE’s) with respect to the time variable t . The coefficients entering nonlinear terms of these ODE’s depend on the spatial point x^i but are the same, at each given x^i , as those that arise in some spatially homogeneous models. In the presently considered four-dimensional “vacuum” (i.e. Ricci tensor = 0) case, the spatially homogeneous models that capture the behavior of the general solution are of the Bianchi type IX or VIII (with homogeneity groups SU(2) or SL(2, \mathbb{R}), respectively). The asymptotic evolution of the metric was then found to be describable in terms of a chaotic [12, 13] sequence of generalized Kasner (i.e. power-law) solutions, exhibiting “oscillations” of the scale factors along independent spatial directions [9–11], or, equivalently, as a billiard

motion [14, 15] on the Lobatchevskii plane.

This picture, found by BKL in the four-dimensional pure Einstein case, was then generalized in various directions by adding more spatial dimensions, or by adding matter fields. Remarkably, it was found that there exists a *critical* space-time dimension for the asymptotic behavior of the general Ricci-flat solution [16]. When $D \leq 10$ the solution exhibits a BKL-type never-ending oscillatory behavior with strong chaotic properties, while when $D \geq 11$ the general solution ceases to exhibit chaotic features, but is instead asymptotically characterized by a monotonic Kasner-like solution [16, 17]. The addition of matter fields was also found to feature a similar subcritical/overcritical classification [2, 17, 18]. Let us consider, in some given space-time dimension D , the Einstein-dilaton- p -form system, i.e. let us add a massless scalar ϕ , and one or several p -form fields $A = A_{\mu_1 \dots \mu_p} dx^{\mu_1} \wedge \dots \wedge dx^{\mu_p} / p!$. It was found that the general solution, near a cosmological singularity, of this system is monotonic and Kasner-like if the dilaton couplings λ_p (defined below) of the p -forms belong to some (dimension dependent) open neighborhood of zero [17], while it is chaotic and BKL-like if the dilaton couplings belong to the (closed) complement of the latter neighborhood. In the absence of any p -forms, the Einstein-dilaton system is found to be asymptotically monotonic and Kasner-like [18, 19].

A convenient tool for describing the qualitative behavior of the general solution near a space-like singularity is to relate it to billiard motion in an auxiliary Lorentzian space, or, after projection on the unit hyperboloid, to the billiard motion on the Lobatchevskii space. This billiard picture naturally follows from the Hamiltonian approach to cosmological behavior (initiated long ago [20] in the context of the Bianchi IX models; see [21] for a recent review), and was first obtained in the four-dimensional case [14, 15] and then extended to higher space-time dimensions with p -forms and dilaton [4, 22–25]. Recent work has improved the derivation of the billiard picture by using the Iwasawa decomposition of the spatial metric and by highlighting the general mechanism by which all the “off-diagonal” degrees of freedom (i.e. all the variables except for the scale factors, the dilaton and their conjugate momenta) get “asymptotically frozen” [8].

The billiard dynamics describing the asymptotic cosmological behavior of D -dimensional Einstein gravity coupled to n dilatons ϕ_1, \dots, ϕ_n (having minimal kinetic terms $-\sum_j |d\phi_j|^2$), and to an arbitrary menu of p -forms (with kinetic terms $\sim -e^{\langle \lambda_p, \phi \rangle} |dA_p|^2$, where λ_p is a linear form) consists of a “billiard ball” moving along a future-directed null geodesic (i.e. a “light ray”) of a $(D-1+n)$ -dimensional Lorentzian space, except when it undergoes

specular reflections on a set of hyperplanar walls (or “mirrors”) passing through the origin of Lorentzian space. One can refer to this billiard as being an “Einstein billiard” (or a “cosmological billiard”).

As we will see below in detail the degrees of freedom of the billiard ball correspond, for $D - 1$ of them, to the logarithms, $-2\beta^a$, of the *scale factors*, i.e. the diagonal components of the $(D - 1) \times (D - 1)$ spatial metric in an Iwasawa decomposition, and, for the remaining ones, to n dilatons ϕ_j . The walls come from the asymptotic elimination (in the Hamiltonian) of the remaining “off-diagonal” degrees of freedom: off-diagonal components of the metric, p -form fields, and their conjugate momenta. It is often convenient to project the Lorentzian billiard motion (by a linear projection centered in the origin) onto the future half of the unit hyperboloid $G_{\mu\nu} \beta^\mu \beta^\nu = -1$, where $(\beta^\mu) \equiv (\beta^a, \phi_j)$ denote the position in, and $G_{\mu\nu}$ the metric of, $(D - 1 + n)$ -dimensional Lorentzian space. The signature of $G_{\mu\nu}$ is $- + + \dots +$. This projection leads to a billiard motion in a domain of hyperbolic space H_{D-2+n} , see Fig. 1. This domain (the “billiard table”^a) is the intersection of a finite number of half hyperbolic spaces, say the positive sides of the hyperbolic hyperplanar walls defined by intersecting the original Lorentzian hyperplanar walls with the unit hyperboloid $G_{\mu\nu} \beta^\mu \beta^\nu = -1$.

The distinction mentioned above between a chaotic BKL-like behavior and a monotonic Kasner-like behavior corresponds to the distinction between a finite-volume billiard table (in H_{D-2+n}) and an infinite-volume one. One could further distinguish: (i) the overcritical case (*compact* billiard table), (ii) the critical case (*finite-volume*, but *noncompact* billiard table) and (iii) the subcritical case (*infinite-volume* billiard table). Both cases (i) and (ii) lead to never-ending, chaotic oscillations of the scale factors (and of the dilatons), while the case (iii) leads, after a finite number of reflections off the walls (i.e. after a finite number of oscillations of the scale factors and of the dilatons), to an asymptotically free motion of β^μ (null geodesic), corresponding to an asymptotically monotonic, Kasner-like (power-law) behavior of the space-time metric near the singularity, see Fig. 1. In actual physical models, only the critical and subcritical cases are found.

Having introduced the concept of a Lorentzian (or, after projection, hyperbolic) billiard table associated with general classes of Einstein-matter systems, we can now describe the connection uncovered in [2–8] between certain specific Einstein-matter systems and Lorentzian Kač–Moody algebras.

^a We use the term “billiard table” to refer to the domain within which the billiard motion is confined. Note, however, that this “table” is, in general, of dimension higher than two, and is either considered in Lorentzian space or, after projection, in hyperbolic space.

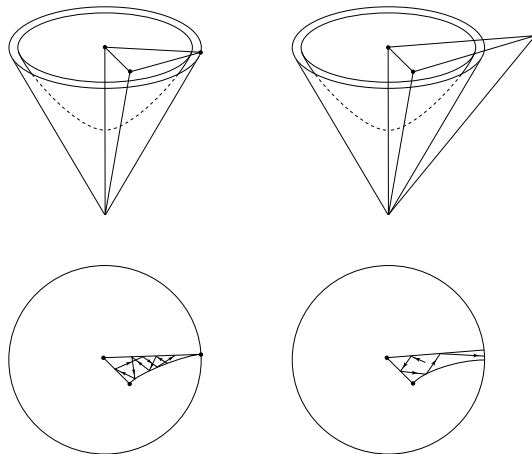


Figure 1. Sketch of billiard tables describing the asymptotic cosmological behavior of Einstein-matter systems. The upper panels are drawn in the Lorentzian space spanned by $(\beta^\mu) = (\beta^a, \phi)$. The billiard tables are represented as “wedges” in $[(2+1)\text{-dimensional}] \beta\text{-space}$, bounded by hyperplanar walls $w_A(\beta) = 0$ on which the (unrepresented) billiard ball undergoes specular reflections. The upper left panel is a (critical) “chaotic” billiard table (contained within the $\beta\text{-space}$ future light cone), while the upper right one is a (subcritical) “nonchaotic” one (extending beyond the light cone). The lower panels represent the corresponding billiard tables (and billiard motions) after projection onto hyperbolic space (H_2 in the case drawn here). The latter projection is defined in the text by central projection onto $\gamma\text{-space}$ (i.e. the unit hyperboloid $G_{\mu\nu} \gamma^\mu \gamma^\nu = -1$, see the upper panels), and is represented in the lower panels by its image in the Poincaré ball (disc). “Chaotic” billiard tables have finite volume in hyperbolic space, while “nonchaotic” ones have infinite volume.

In the leading asymptotic approximation to the behavior near the cosmological singularity, this connection is simply that the (Lorentzian) *billiard table* describing this behavior can be identified with the (Lorentzian) *Weyl chamber* of some corresponding (model-dependent) Lorentzian Kač–Moody algebra. We recall that the Weyl chamber of a Lie algebra with simple roots α_i , $i = 1, \dots, r$ (r being the rank of the algebra) is the domain of the Cartan subalgebra (parametrized by $\beta \in \mathbb{R}^r$) where $\langle \alpha_i, \beta \rangle \geq 0$ for all i 's. For this connection to be possible many conditions must be met. In particular, (i) the billiard table must be a Coxeter polyhedron, i.e. the dihedral angles between the adjacent walls must be integer submultiples of π (i.e., of the form π/k where k is an integer ≥ 2), and (ii) the billiard table must be a simplex, i.e. have exactly $D - 1 + n$ faces. It is remarkable that this seemingly very special case of a “Kač–Moody billiard” is found to occur in many physically interesting Einstein-matter systems. For instance, pure Einstein gravity in D -dimensional space-time corresponds to the Lorentzian Kač–Moody algebra AE_{D-1} [5]. The latter algebra is the canonical Lorentzian extension [26]

of the ordinary (finite-dimensional) Lie algebra A_{D-3} . Another interesting connection between qualitative PDE behavior and Kač–Moody theoretic concepts is that the transition between “critical” chaotic, BKL-like behavior in “low dimensions” ($D \leq 10$) and “subcritical” monotonic, Kasner-like behavior in “high dimensions” ($D \geq 11$) is found to be in strict correspondence with a transition between a *hyperbolic* Kač–Moody algebra (AE_{D-1} for $D \leq 10$) and a *nonhyperbolic* one (AE_{D-1} for $D \geq 11$). We recall here that V. Kač defines [26] a Kač–Moody algebra to be *hyperbolic* by the condition that any subdiagram obtained by removing a node from its Dynkin diagram be either of finite or affine type. Hyperbolic Kač–Moody algebras are necessarily Lorentzian (i.e. the symmetrizable Cartan matrix is of Lorentzian signature), but the reverse is not true in general.

Another connection between physically interesting Einstein–matter systems and Kač–Moody algebras concerns the low-energy bosonic effective actions arising in string and M theories. Bosonic string theory in any space-time dimension D is related to the Lorentzian Kač–Moody algebra DE_D [4, 6]. The latter algebra is the canonical Lorentzian extension of the finite-dimensional algebra D_{D-2} . The various superstring theories (in the critical dimension $D = 10$) and M -theory have been found [4] to be related either to E_{10} (when there are two supersymmetries in $D = 10$, i.e. for type IIA, type IIB and M -theory) or to BE_{10} (when there is only one supersymmetry in $D = 10$, i.e. for type I and II heterotic theories), see Fig. 2. A construction of the Einstein–matter systems related to the canonical Lorentzian extensions of *all* finite-dimensional Lie algebras $A_n, B_n, C_n, D_n, G_2, F_4, E_6, E_7$ and E_8 (in the above “billiard” sense) is presented in Ref. [6]. See also Ref. [27] for the identification of all hyperbolic Kač–Moody algebras whose Weyl chambers are the Einstein billiards.

The correspondence between the specific Einstein–three-form system (including a Chern–Simons term) describing the bosonic sector of 11-dimensional supergravity (also known as the “low-energy limit of M -theory”) and the hyperbolic Kač–Moody group E_{10} was studied in more detail in [7]. Reference [7] introduces a formal expansion of the field equations in terms of positive roots, i.e. combinations $\alpha = \sum_i n^i \alpha_i$ of simple roots of E_{10} , α_i , $i = 1, \dots, 10$, where the n^i ’s are integers ≥ 0 . It is then useful to *order* this expansion according to the *height* of the positive root $\alpha = \sum_i n^i \alpha_i$, defined as $\text{ht}(\alpha) = \sum_i n^i$. The correspondence discussed above between the *leading* asymptotic evolution near a cosmological singularity (described by a billiard) and Weyl chambers of Kač–Moody algebras involves only the terms in the field equation whose height is $\text{ht}(\alpha) \leq 1$. By contrast, the au-

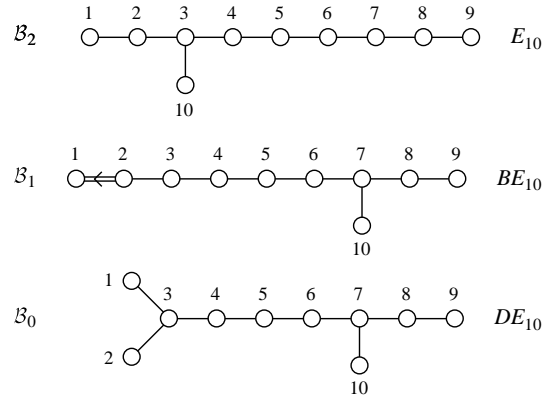


Figure 2. Coxeter–Dynkin diagrams encoding the geometry of the billiard tables describing the asymptotic cosmological behavior of three blocks of string theories: $\mathcal{B}_2 = \{M\text{-theory, type IIA and type IIB superstring theories}\}$, $\mathcal{B}_1 = \{\text{type I and the two heterotic superstring theories}\}$, and $\mathcal{B}_0 = \{\text{closed bosonic string theory in } D = 10\}$. Each node of the diagrams represents a dominant wall of the corresponding cosmological billiard. Each Coxeter diagram of a billiard table corresponds to the Dynkin diagram of a corresponding (hyperbolic) Kač–Moody algebra: E_{10} , BE_{10} and DE_{10} , respectively. In other words, the cosmological billiard tables can be identified with the Weyl chambers of the corresponding Lorentzian Kač–Moody algebras.

thors of Ref. [7] managed to show, by explicit calculation, that there exists a way to define, at each spatial point x , a correspondence between the field variables $g_{\mu\nu}(t, x)$, $A_{\mu\nu\lambda}(t, x)$ (and their gradients), and a (finite) subset of the parameters defining an element of the (infinite-dimensional) coset space $E_{10}/K(E_{10})$ where $K(E_{10})$ denotes the maximal compact subgroup of E_{10} , such that the (PDE) field equations of supergravity get mapped onto the (ODE) equations describing a null geodesic in $E_{10}/K(E_{10})$ up to terms of height 30. This tantalizing result suggests that the infinite-dimensional hyperbolic Kač–Moody group E_{10} may be a “hidden symmetry” of supergravity, in the sense of mapping solutions onto other solutions (the appearance of E_{10} as a possible symmetry group of supergravity was first hinted at by Julia long ago [28, 29]). Note that the conjecture here is that the *continuous* group $E_{10}(\mathbb{R})$ be a hidden symmetry group of *classical* supergravity. At the *quantum* level, i.e. for M theory, one expects only a discrete version of E_{10} , say $E_{10}(\mathbb{Z})$, to be a quantum symmetry. See [30] for recent work on extending the identification of [7] between roots of E_{10} and symmetries of supergravity/M-theory beyond height 30, and for references about previous suggestions of a possible role for E_{10} . For earlier appearances of the Weyl groups of the E series in the context of U -duality see [31–33]. A series of recent papers [34–38] has also explored the possible role of E_{11} (a nonhyper-

bolic extension of E_{10}) as a hidden symmetry of M theory.

It is also tempting to assume that the Kač–Moody groups underlying the other (special) Einstein-matter systems discussed above might be hidden (solution-generating) symmetries. For instance, in the case of pure Einstein gravity in $D = 4$ space-time, the conjecture is that AE_3 is such a symmetry of Einstein gravity. This case, and the correspondence between the field variables and the coset ones is further discussed in [8].

To end this introductory summary, it is important to add a significant mathematical caveat. Most of the results mentioned above have been obtained by “physicists’ methods”, i.e. nonrigorous arguments. (See, however, the mathematical results of [39] concerning the dynamics of the Bianchi type IX homogeneous model.) The only exception concerns the “nonchaotic”, monotonic Kasner-like behavior of the “subcritical” systems. The first mathematical proof that a general solution of the four-dimensional Einstein–dilaton system has a nonchaotic Kasner-like behavior was obtained by Andersson and Rendall [19]. By using and slightly extending the tools of this proof (Theorem 3 of [19], concerning certain Fuchsian systems), Ref. [17] gives mathematical proofs of the Kasner-like behavior of more general classes of Einstein–matter systems, notably the D -dimensional coupled Einstein–dilaton- p -form system when the dilaton couplings of the p -forms belong to the *subcritical* domain mentioned above. Reference [17] also gives a proof of the Kasner-like behavior of pure gravity (Ricci tensor = 0) when the space-time dimension $D \geq 11$.

An important mathematical challenge is to convert physicists’ arguments (summarized in the rest of this review) concerning the “chaotic billiard” structure of *critical* (and *overcritical*) systems into precise mathematical statements. The recent work [7, 8], which provides a simplified derivation of the billiard picture and a rather clean decomposition of the field variables into “ODE-described chaotic” and “asymptotically frozen” pieces, might furnish a useful starting point for formulating precise mathematical statements. In this respect, let us note two things: (i) the good news is that numerical simulations have fully confirmed the BKL chaotic-billiard picture in several models [40, 41], (ii) the bad news is that physicists’ arguments might oversimplify the picture by neglecting “nongeneric” spatial points where some of the leading walls disappear, because the spatially-dependent coefficients measuring the strength of these walls happen to vanish. A similar subtlety (vanishing at exceptional spatial points of wall coefficients) takes place in the subcritical, nonchaotic case (where a finite number of collisions on the billiard walls is expected to channel the billiard ball in the “good” Kasner-like

directions). See discussions of this phenomenon in [41, 42].

In the remainder of this review, we will outline various arguments leading to the conclusions summarized above. For a more complete derivation of the billiard results see [8].

2. Models and Gauge Conditions

We consider models of the general form

$$S[g_{MN}, \phi, A^{(p)}] = \int d^D x \sqrt{-g} \left[R(g) - \partial_M \phi \partial^M \phi - \frac{1}{2} \sum_p \frac{1}{(p+1)!} e^{\lambda_p \phi} F_{M_1 \dots M_{p+1}}^{(p)} F^{(p) M_1 \dots M_{p+1}} \right] + \dots, \quad (2.1)$$

where units are chosen such that $16\pi G_N = 1$ (where G_N is Newton's constant) and the space-time dimension $D \equiv d + 1$ is left unspecified. Besides the standard Einstein–Hilbert term the above Lagrangian contains a dilaton field ϕ and a number of p -form fields $A_{M_1 \dots M_p}^{(p)}$ (for $p \geq 0$). For simplicity, we consider the case where there is only one dilaton, i.e. $n = 1$ in the notation of Sec. 1. As a convenient common formulation we adopt the Einstein conformal frame and normalize the kinetic term of the dilaton ϕ with weight one with respect to the Ricci scalar. The Einstein metric g_{MN} has Lorentz signature $(- + \dots +)$ and is used to lower or raise the indices; its determinant is denoted by g . The p -form field strengths $F^{(p)} = dA^{(p)}$ are normalized as

$$F_{M_1 \dots M_{p+1}}^{(p)} = (p+1) \partial_{[M_1} A_{M_2 \dots M_{p+1}}^{(p)}] \equiv \partial_{M_1} A_{M_2 \dots M_{p+1}}^{(p)} \pm p \text{ permutations} . \quad (2.2)$$

The dots in the action (2.1) above indicate possible modifications of the field strength by additional Yang–Mills or Chapline–Manton-type couplings [43, 44], such as $F_C = dC^{(2)} - C^{(0)} dB^{(2)}$ for two 2-forms $C^{(2)}$ and $B^{(2)}$ and a 0-form $C^{(0)}$, as they occur in type IIB supergravity. Further modifications include Chern–Simons terms, as in the action for $D = 11$ supergravity [45]. The real parameter λ_p measures the strength of the coupling of $A^{(p)}$ to the dilaton. When $p = 0$, we assume that $\lambda_0 \neq 0$ so that there is only one dilaton. This is done mostly for notational convenience. If there were other dilatons among the 0-forms, these should be separated off from the p -forms because they play a distinct rôle. They would define additional space-like directions in the space of the (logarithmic) scale factors and would correspondingly increase the dimension of the relevant hyperbolic billiard.

The metric g_{MN} , the dilaton field(s) ϕ and the p -form fields $A_{M_1 \dots M_p}^{(p)}$ are

a priori arbitrary functions of both space and time, on which *no symmetry conditions are imposed*. Nevertheless it will turn out that the evolution equations near the singularity will be asymptotically the same as those of certain homogeneous cosmological models. It is important to keep in mind that this simplification does not follow from imposing extra-dimensional reduction conditions but emerges as a direct consequence of the general dynamics.

Our analysis applies both to past and future singularities, and in particular to Schwarzschild-type singularities inside black holes. To follow historical usage, we will assume for definiteness that the space-like singularity lies in the past, at finite distance in proper time. More specifically, we will adopt a space-time slicing such that the singularity “occurs” on a constant time slice ($t = 0$ in proper time). The slicing is built by use of pseudo-Gaussian coordinates defined by vanishing lapse $N^i = 0$, with metric

$$ds^2 = -(N(x^0, x^i)dx^0)^2 + g_{ij}(x^0, x^i)dx^i dx^j. \quad (2.3)$$

For simplicity, we will work in what follows with a *coordinate* spatial coframe dx^i , as indicated in Eq. (2.3). However, most of what we say would go through if we worked in a general *time-independent* spatial coframe $\omega^i(x^k) = \omega_j^i(x^k)dx^j$. The usefulness of using a noncoordinate spatial coframe will show up in the E_{10} case discussed below. In order to simplify various formulas later, we will find it useful to introduce a rescaled lapse function

$$\tilde{N} \equiv N/\sqrt{g} \quad (2.4)$$

where $g \equiv \det g_{ij}$. Note the distinction between space-time quantities g_{MN} , g and space ones g_{ij} , g . In the following, we work only with the spatial metric g_{ij} . We will see that in the Hamiltonian approach a useful gauge is that defined by requiring

$$\tilde{N} = \rho^2, \quad (2.5)$$

where ρ^2 is a quadratic combination of the logarithms of the scale factors and the dilaton(s), which we will define below in terms of the Iwasawa decomposition of g_{ij} . After fixing the time-zero hypersurface the only coordinate freedom left in the pseudo-Gaussian gauge (2.5) is that of making time-independent changes of spatial coordinates $x^i \rightarrow x'^i = f^i(x^j)$. Since the gauge condition (2.5) is not invariant under spatial coordinate transformations, such changes of coordinates have the unusual feature of also changing the slicing.

Throughout this paper, we will reserve the label t for the *proper time*,

$$dt = -Ndx^0 = -\tilde{N}\sqrt{g}dx^0, \quad (2.6)$$

whereas the time coordinate associated with the special gauge (2.5) will be denoted by T , viz.

$$dT = -\frac{dt}{\rho^2\sqrt{g}}. \quad (2.7)$$

Sometimes, it will also be useful to introduce an “intermediate” time coordinate τ that would correspond to the gauge condition $\tilde{N} = 1$. It is explicitly defined by

$$d\tau = -\frac{dt}{\sqrt{g}} = \rho^2 dT. \quad (2.8)$$

At the singularity the proper time t is assumed to remain finite and to decrease toward 0^+ . By contrast, the coordinates T and τ both increase toward $+\infty$, as ensured by the minus sign in (2.6). Irrespective of the choice of coordinates, the spatial volume density g is assumed to collapse to zero at each spatial point in this limit.

As for the p -form fields, we will assume, throughout this paper, a generalized temporal gauge, viz.

$$A_{0i_2\dots i_p}^{(p)} = 0, \quad (2.9)$$

where lower-case Latin letters i, j, \dots denote spatial indices from now on. This choice leaves the freedom to make time-independent gauge transformations, and therefore fixes the gauge only partially.

3. Asymptotic dynamics in the general case

The main ingredients of the derivation of the billiard picture are:

- use of the Hamiltonian formalism;
- Iwasawa decomposition of the metric; i.e. $g_{ij} \rightarrow (\beta^a, \mathcal{N}^a_i)$;
- decomposition of $\beta^\mu = (\beta^a, \phi)$ into radial (ρ) and angular (γ^μ) parts; and
- use of the pseudo-Gaussian gauge (2.5), i.e. of the time coordinate T as the evolution parameter.

More explicitly, with the conventions already described before, we assume that in some space-time patch, the metric is given by (2.3) (pseudo-Gaussian gauge), such that the local volume g collapses at each spatial point as $x^0 \rightarrow +\infty$, in such a way that the proper time t tends to 0^+ . We work in the Hamiltonian formalism, i.e. with the first order evolution equations in the phase space of the system. For instance, the gravitational degrees of freedom

are initially described by the metric g_{ij} and its conjugate momentum π^{ij} . We systematically use the Iwasawa decomposition of the spatial metric g_{ij} .^b Namely, we write

$$g_{ij} = \sum_{a=1}^d e^{-2\beta^a} \mathcal{N}^a{}_i \mathcal{N}^a{}_j \quad (3.1)$$

where \mathcal{N} is an upper triangular matrix with 1's on the diagonal, and where we recall that $d \equiv D - 1$ denotes the spatial dimension. We will also need the Iwasawa coframe $\{\theta^a\}$,

$$\theta^a = \mathcal{N}^a{}_i dx^i, \quad (3.2)$$

as well as the vectorial frame $\{e_a\}$ dual to the coframe $\{\theta^a\}$,

$$e_a = \mathcal{N}^i{}_a \frac{\partial}{\partial x^i} \quad (3.3)$$

where the matrix $\mathcal{N}^i{}_a$ is the inverse of $\mathcal{N}^a{}_i$, i.e., $\mathcal{N}^a{}_i \mathcal{N}^i{}_b = \delta^a_b$. It is again an upper triangular matrix with 1's on the diagonal.

The Iwasawa decomposition allows us to replace the $d(d+1)/2$ variables g_{ij} by the $d+d(d-1)/2$ variables $(\beta^a, \mathcal{N}^a{}_i)$. Note that $(\beta^a, \mathcal{N}^a{}_i)$ are *ultralocal* functions of g_{ij} , that is they depend, at each space-time point, only on the value of g_{ij} at that point, not on its derivatives. This would not have been the case if we had used a ‘‘Kasner frame’’ (i.e. a frame diagonalizing π_{ij} with respect to g_{ij}) instead of an Iwasawa one. The transformation $g \rightarrow (\beta, \mathcal{N})$ then defines a corresponding transformation of the conjugate momenta, as we will explain below. We then augment the definition of the β 's by adding the dilaton field, i.e. $\beta^\mu \equiv (\beta^a, \phi)$. The $(d+n)$ -dimensional space of the β^μ 's (where, as we said, we consider only one dilaton, i.e. $n = 1$) comes equipped with a canonical Lorentzian metric $G_{\mu\nu}$ defined by

$$d\sigma^2 \equiv G_{\mu\nu} d\beta^\mu d\beta^\nu \equiv \sum_{a=1}^d (d\beta^a)^2 - \left(\sum_{a=1}^d d\beta^a \right)^2 + d\phi^2. \quad (3.4)$$

This (flat) Lorentzian metric in the auxiliary β -space (of signature $- + \dots +$) plays an essential role in our study. Note that in models without a dilaton (such as pure gravity in space-time dimension $d+1$) one has a d -dimensional

^b If we were working with a noncoordinate spatial coframe $\omega^i = \omega^i_j(x^k) dx^j$, we would use the Iwasawa decomposition of the frame components of the spatial metric, $d\ell^2 = g_{ij} \omega^i \omega^j$.

936 *Thibault Damour*

Lorentzian space, with $(\beta^\mu) \equiv (\beta^a)$ and the metric

$$G_{\mu\nu} d\beta^\mu d\beta^\nu = \sum_a (d\beta^a)^2 - \left(\sum_a d\beta^a\right)^2.$$

We then decompose β^μ into hyperbolic polar coordinates (ρ, γ^μ) , i.e

$$\beta^\mu = \rho \gamma^\mu, \quad (3.5)$$

where γ^μ are coordinates on the future sheet of the unit hyperboloid^c which are constrained by

$$G_{\mu\nu} \gamma^\mu \gamma^\nu \equiv \gamma^\mu \gamma_\mu = -1, \quad (3.6)$$

and ρ is the time-like variable

$$\rho^2 \equiv -G_{\mu\nu} \beta^\mu \beta^\nu \equiv -\beta^\mu \beta_\mu > 0. \quad (3.7)$$

This decomposition naturally introduces the unit hyperboloid (“ γ -space”), see Fig. 1, which is a realization of the m -dimensional hyperbolic (Lobatchevskii) space H_m , with $m = d - 1 + n$ if there are $n \geq 0$ dilatons.

In terms of the “polar” coordinates ρ and γ^μ , the metric in the β -space becomes

$$d\sigma^2 = -d\rho^2 + \rho^2 d\Sigma^2 \quad (3.8)$$

where $d\Sigma^2$ is the metric on the γ -space H_m .

Note that ρ is also an ultralocal function of the configuration variables (g_{ij}, ϕ) . We assume that the hyperbolic coordinate ρ can be used everywhere in a given region of space near the singularity as a well-defined (real) quantity which tends to $+\infty$ as we approach the singularity. We then define the slicing of space-time by imposing the gauge condition (2.5). The time coordinate corresponding to this gauge is called T as above (see (2.7)). Our aim is to study the asymptotic behavior of all the dynamical variables $\beta(T)$, $\mathcal{N}(T)$, ... as $T \rightarrow +\infty$ (recall that this limit also corresponds to $t \rightarrow 0$, $\sqrt{g} \rightarrow 0$, $\rho \rightarrow +\infty$, with β^μ going to infinity inside the future light cone). Of course, we must also ascertain the self-consistency of this limit, which we will refer to as the “BKL limit.”

3.1. *Hamiltonian action*

To focus on the features relevant to the billiard picture, we assume here that there are no Chern–Simons terms or couplings of the exterior form gauge

^c Indeed, one finds that, near a space-like singularity β^μ tends to future time-like infinity.

fields through a modification of the curvatures $F^{(p)}$, which are thus taken to be Abelian, $F^{(p)} = dA^{(p)}$. See [8] for a proof that these interaction terms do not change the analysis. The Hamiltonian action in any pseudo-Gaussian gauge, and in the temporal gauge (2.9), reads

$$S \left[g_{ij}, \pi^{ij}, \phi, \pi_\phi, A_{j_1 \dots j_p}^{(p)}, \pi_{(p)}^{j_1 \dots j_p} \right] = \int dx^0 \int d^d x \left(\pi^{ij} g_{ij} + \pi_\phi \dot{\phi} + \frac{1}{p!} \sum_p \pi_{(p)}^{j_1 \dots j_p} \dot{A}_{j_1 \dots j_p}^{(p)} - H \right), \quad (3.9)$$

where the Hamiltonian density H is

$$H \equiv \tilde{N} \mathcal{H}, \quad (3.10)$$

$$\mathcal{H} = \mathcal{K} + \mathcal{M}, \quad (3.11)$$

$$\mathcal{K} = \pi^{ij} \pi_{ij} - \frac{1}{d-1} \pi_i^i \pi_j^j + \frac{1}{4} \pi_\phi^2 + \sum_p \frac{e^{-\lambda_p \phi}}{2p!} \pi_{(p)}^{j_1 \dots j_p} \pi_{(p) j_1 \dots j_p}, \quad (3.12)$$

$$\mathcal{M} = -gR + g g^{ij} \partial_i \phi \partial_j \phi + \sum_p \frac{e^{\lambda_p \phi}}{2(p+1)!} g F_{j_1 \dots j_{p+1}}^{(p)} F^{(p) j_1 \dots j_{p+1}}, \quad (3.13)$$

and R is the spatial curvature scalar. The dynamical equations of motion are obtained by varying the above action with respect to the spatial metric components, the dilaton, the spatial p -form components and their conjugate momenta. In addition, there are constraints on the dynamical variables,

$$\mathcal{H} \approx 0 \quad (\text{“Hamiltonian constraint”}), \quad (3.14)$$

$$\mathcal{H}_i \approx 0 \quad (\text{“momentum constraint”}), \quad (3.15)$$

$$\varphi_{(p)}^{j_1 \dots j_{p-1}} \approx 0 \quad (\text{“Gauss law” for each } p\text{-form}), \quad (3.16)$$

with

$$\mathcal{H}_i = -2\pi^j_{i|j} + \pi_\phi \partial_i \phi + \sum_p \frac{1}{p!} \pi_{(p)}^{j_1 \dots j_p} F_{i j_1 \dots j_p}^{(p)}, \quad (3.17)$$

$$\varphi_{(p)}^{j_1 \dots j_{p-1}} = \pi_{(p)}^{j_1 \dots j_{p-1} j_p} \Big|_{j_p}, \quad (3.18)$$

where the subscript $|j$ stands for spatially covariant derivative.

Let us now examine how the Hamiltonian action gets transformed when one performs, at each spatial point, the Iwasawa decomposition (3.1) of the

spatial metric. The kinetic terms of the metric and of the dilaton in the Lagrangian (2.1) are given by the quadratic form

$$\sum_{a=1}^d (d\beta^a)^2 - \left(\sum_{a=1}^d d\beta^a \right)^2 + d\phi^2 + \frac{1}{2} \sum_{a<b} e^{2(\beta^b - \beta^a)} (d\mathcal{N}^a_i \mathcal{N}^i_b)^2, \quad (3.19)$$

where we recall that \mathcal{N}^i_a denotes, as in (3.3), the inverse of the triangular matrix \mathcal{N}^a_i appearing in the Iwasawa decomposition (3.1) of the spatial metric g_{ij} . This change of variables corresponds to a point canonical transformation, which can be extended to the momenta in the standard way via

$$\pi^{ij} \dot{g}_{ij} \equiv \sum_a \pi_a \dot{\beta}^a + \sum_a P^i_a \dot{\mathcal{N}}^a_i. \quad (3.20)$$

Note that the momenta

$$P^i_a = \frac{\partial \mathcal{L}}{\partial \dot{\mathcal{N}}^a_i} = \sum_b e^{2(\beta^b - \beta^a)} \dot{\mathcal{N}}^a_j \mathcal{N}^j_b \mathcal{N}^i_b \quad (3.21)$$

conjugate to the nonconstant off-diagonal Iwasawa components \mathcal{N}^a_i are only defined for $a < i$; hence the second sum in (3.20) receives only contributions from $a < i$.

We next split the Hamiltonian density^d \mathcal{H} (3.10) in two parts, one denoted by \mathcal{H}_0 , which is the kinetic term for the local scale factors β^μ (including dilatons) already encountered in section 3, and a “potential density” (of weight 2) denoted by \mathcal{V} , which contains everything else. Our analysis below will show why it makes sense to group the kinetic terms of both the off-diagonal metric components and the p -forms with the usual potential terms, i.e. the term \mathcal{M} in (3.10). (Remembering that the off-diagonal components of the metric in one dimension higher become a one-form in the Kaluza-Klein reduction, it is not surprising that it might be useful to group together the off-diagonal components and the p -forms.) Thus, we write

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{V} \quad (3.22)$$

with the kinetic term of the β variables

$$\mathcal{H}_0 = \frac{1}{4} G^{\mu\nu} \pi_\mu \pi_\nu, \quad (3.23)$$

^d We use the term “Hamiltonian density” to denote both H and \mathcal{H} . Note that H is a usual spatial density (of weight 1, i.e. the same weight as \sqrt{g}), while $\mathcal{H} \equiv \sqrt{g}H/N$ is a density of weight 2 (like $g = (\sqrt{g})^2$). Note also that π^{ij} is of weight 1, while $\tilde{N} \equiv N/\sqrt{g}$ is of weight -1 .

where $G^{\mu\nu}$ denotes the inverse of the metric $G_{\mu\nu}$ of Eq. (3.4). In other words, the right hand side of Eq. (3.23) is defined by

$$G^{\mu\nu} \pi_\mu \pi_\nu \equiv \sum_{a=1}^d \pi_a^2 - \frac{1}{d-1} \left(\sum_{a=1}^d \pi_a \right)^2 + \pi_\phi^2, \quad (3.24)$$

where $\pi_\mu \equiv (\pi_a, \pi_\phi)$ are the momenta conjugate to β^a and ϕ , respectively, i.e.

$$\pi_\mu = 2\tilde{N}^{-1} G_{\mu\nu} \dot{\beta}^\nu = 2G_{\mu\nu} \frac{d\beta^\nu}{d\tau}. \quad (3.25)$$

The total (weight 2) potential density,

$$\mathcal{V} = \mathcal{V}_S + \mathcal{V}_G + \sum_p \mathcal{V}_p + \mathcal{V}_\phi, \quad (3.26)$$

is naturally split into a “centrifugal” part linked to the kinetic energy of the off-diagonal components (the index S referring to “symmetry,” as discussed below)

$$\mathcal{V}_S = \frac{1}{2} \sum_{a < b} e^{-2(\beta^b - \beta^a)} (P^j_b \mathcal{N}^a_j)^2, \quad (3.27)$$

a “gravitational” (or “curvature”) potential

$$\mathcal{V}_G = -gR, \quad (3.28)$$

and a term from the p -forms,

$$\mathcal{V}_{(p)} = \mathcal{V}_{(p)}^{el} + \mathcal{V}_{(p)}^{magn}, \quad (3.29)$$

which is a sum of an “electric” and a “magnetic” contribution

$$\mathcal{V}_{(p)}^{el} = \frac{e^{-\lambda_p \phi}}{2 p!} \pi_{(p)}^{j_1 \dots j_p} \pi_{(p) j_1 \dots j_p}, \quad (3.30)$$

$$\mathcal{V}_{(p)}^{magn} = \frac{e^{\lambda_p \phi}}{2 (p+1)!} g F_{j_1 \dots j_{p+1}}^{(p)} F^{(p) j_1 \dots j_{p+1}}. \quad (3.31)$$

Finally, there is a contribution to the potential linked to the spatial gradients of the dilaton,

$$\mathcal{V}_\phi = g g^{ij} \partial_i \phi \partial_j \phi. \quad (3.32)$$

We will analyze below in detail these contributions to the potential.

3.2. Appearance of sharp walls in the BKL limit

In the decomposition of the Hamiltonian given above, we split off the kinetic terms of the scale factors β^a and of the dilaton $\beta^{d+1} \equiv \phi$ from the other variables, and assigned the off-diagonal metric components and the p -form fields to various potentials, each of which is a complicated function of $\beta^\mu, \mathcal{N}^a{}_i, P^i{}_a, A_{j_1 \dots j_p}^{(p)}, \pi_{(p)}^{j_1 \dots j_p}$ and of some of their spatial gradients. The reason why this separation is useful is that, as we are going to show, in the BKL limit, and in the special Iwasawa decomposition which we have adopted, the asymptotic dynamics is governed by the scale factors β^μ , whereas all other variables “freeze.” Thus, in the asymptotic limit, we have schematically

$$\mathcal{V}\left(\beta^\mu, \mathcal{N}^a{}_i, P^i{}_a, A_{j_1 \dots j_p}^{(p)}, \pi_{(p)}^{j_1 \dots j_p}, \dots\right) \longrightarrow \mathcal{V}_\infty(\gamma^\mu), \quad (3.33)$$

where $\mathcal{V}_\infty(\gamma^\mu)$ stands for a sum of certain “sharp wall potentials” which depend only on the angular hyperbolic coordinates $\gamma^\mu \equiv \beta^\mu/\rho$. As a consequence, the asymptotic dynamics can be described as a “billiard” in the hyperbolic space of the γ^μ ’s, whose walls (or “cushions”) are determined by the energy of the fields that are asymptotically frozen.

This reduction of the complicated potential to a much simpler “effective potential” $\mathcal{V}_\infty(\gamma^\mu)$ follows essentially from the exponential dependence of \mathcal{V} on the diagonal Iwasawa variables β^μ , from its independence of the conjugate momenta of the β ’s, and from the fact that the radial magnitude ρ of the β ’s becomes infinitely large in the BKL limit.

To see the essence of this reduction, with a minimum of technical complications, let us consider a potential density (of weight 2) of the general form

$$\mathcal{V}(\beta, Q, P) = \sum_A c_A(Q, P) \exp(-2w_A(\beta)), \quad (3.34)$$

where (Q, P) denote the remaining phase space variables (that is, other than (β, π_β)). Here $w_A(\beta) = w_{A\mu}\beta^\mu$ are certain linear forms which depend only on the (extended) scale factors, and whose precise form will be derived in the following section. Similarly we will discuss below the explicit form of the prefactors c_A , which will be some complicated polynomial functions of the remaining fields, i.e. the off-diagonal components of the metric, the p -form fields and their respective conjugate momenta, and of some of their spatial gradients. The fact that the $w_A(\beta)$ depend *linearly* on the scale factors β^μ is an important property of the models under consideration. A second nontrivial fact is that, *for the leading contributions*, the prefactors are always nonnegative, i.e. $c_A^{\text{leading}} \geq 0$. Since the values of the fields for

which $c_A = 0$ constitute a set of measure zero, we will usually make the “genericity” assumption $c_A > 0$ for the leading terms^e in the potential \mathcal{V} . The third fact following from the detailed analysis of the walls that we will exploit is that all the leading walls are *time-like*, i.e. their normal vectors (in the Minkowski β -space) are space-like.

When parametrizing β^μ in terms of ρ and γ^μ , or equivalently

$$\lambda \equiv \ln \rho \equiv \frac{1}{2} \ln (-G_{\mu\nu} \beta^\mu \beta^\nu) \quad (3.35)$$

(with the conjugate momentum π_λ), and γ^μ , the part of the Hamiltonian describing the kinetic energy of the β 's, $H_0 = \tilde{N} \mathcal{H}_0$, takes the form

$$H_0 = \frac{\tilde{N}}{4\rho^2} [-\pi_\lambda^2 + \pi_\gamma^2]. \quad (3.36)$$

Choosing the gauge (2.5) to simplify the kinetic terms H_0 we end up with a Hamiltonian of the form

$$\begin{aligned} H(\lambda, \pi_\lambda, \gamma, \pi_\gamma, Q, P) &= \tilde{N} \mathcal{H} \\ &= \frac{1}{4} [-\pi_\lambda^2 + \pi_\gamma^2] + \rho^2 \sum_A c_A(Q, P) \exp(-\rho w_A(\gamma)), \end{aligned} \quad (3.37)$$

where π_γ^2 is the kinetic energy of a particle moving on H_m . In (3.37) and below we will regard λ as a primary dynamical variable (so that $\rho \equiv e^\lambda$).

The essential point now is that, in the BKL limit, $\lambda \rightarrow +\infty$ i.e. $\rho \rightarrow +\infty$, each term $\rho^2 \exp(-2\rho w_A(\gamma))$ becomes a *sharp wall potential*, i.e. a function of $w_A(\gamma)$ which is zero when $w_A(\gamma) > 0$, and $+\infty$ when $w_A(\gamma) < 0$. To formalize this behavior we define the sharp wall Θ -function^f as

$$\Theta(x) := \begin{cases} 0 & \text{if } x < 0, \\ +\infty & \text{if } x > 0. \end{cases} \quad (3.38)$$

A basic formal property of this Θ -function is its invariance under multiplication by a positive quantity. With the above assumption checked below that all the relevant prefactors $c_A(Q, P)$ are *positive* near each leading wall, we

^e As mentioned in Sec. 1, understanding the effects of the possible failure of this assumption is one of the subtle issues in establishing a rigorous proof of the BKL picture.

^f One should more properly write $\Theta_\infty(x)$, but since this is the only step function encountered in this article, we use the simpler notation $\Theta(x)$.

942 *Thibault Damour*

can formally write

$$\begin{aligned} \lim_{\rho \rightarrow \infty} \left[c_A(Q, P) \rho^2 \exp(-\rho w_A(\gamma)) \right] &= c_A(Q, P) \Theta(-2w_A(\gamma)) \\ &\equiv \Theta(-2w_A(\gamma)). \end{aligned} \quad (3.39)$$

Of course,

$$\Theta(-2w_A(\gamma)) = \Theta(-w_A(\gamma)),$$

but we will keep the extra factor of 2 to recall that the arguments of the exponentials, from which the Θ -functions originate, come with a well-defined normalization. Therefore, the limiting Hamiltonian density reads

$$H_\infty(\lambda, \pi_\lambda, \gamma, \pi_\gamma, Q, P) = \frac{1}{4} [-\pi_\lambda^2 + \pi_\gamma^2] + \sum_{A'} \Theta(-2w_{A'}(\gamma)), \quad (3.40)$$

where A' runs over the *dominant walls*. The set of dominant walls is defined as the *minimal set* of wall forms which suffice to define the billiard table, i.e. such that the restricted set of inequalities $\{w_{A'}(\gamma) \geq 0\}$ imply the full set $\{w_A(\gamma) \geq 0\}$. The concept of dominant wall will be illustrated below. Note that the concept of the “dominant” wall is a refinement of the distinction (which will also enter our discussion) between a leading wall and a subleading one.

The crucial point is that the limiting Hamiltonian (3.40) no longer depends on λ, Q and P . Therefore the Hamiltonian equations of motion for λ, Q and P tell us that the corresponding conjugate momenta, i.e. π_λ, P and Q , respectively, all become *constants of motion* in the limit $\lambda \rightarrow +\infty$. The total Hamiltonian density H_∞ is also a constant of motion (which must be set to zero). The variable λ evolves according to $d\lambda/dT = -\frac{1}{2}\pi_\lambda$. Hence, in this limit λ is a linear function of T . The only nontrivial dynamics resides in the evolution of (γ, π_γ) which reduces to the sum of a free (nonrelativistic) kinetic term $\pi_\gamma^2/4$ and a sum of sharp wall potentials, such that the resulting motion of the γ 's indeed constitutes a billiard, with geodesic motion on the unit hyperboloid H_m interrupted by reflections on the walls defined by $w_A(\gamma) = 0$. These walls are hyperplanes (in the sense of hyperbolic geometry) because they are geometrically given by the intersection of the unit hyperboloid $\beta^\mu \beta_\mu = -1$ with the usual Minkowskian hyperplanes $w_A(\beta) = 0$. A toy model which shows in more detail how the asymptotic constancy of the “off diagonal” phase-space variables (Q, P) arises is presented in Appendix A of [8].

The geodesic motion of a billiard in hyperbolic space has been studied in many works. It is known that this motion is chaotic or nonchaotic depending on whether the billiard has finite or infinite volume [46–49]. In the finite volume case, the generic evolution exhibits an infinite number of collisions with the walls with strong chaotic features (“oscillating behavior”). By contrast, if the billiard has infinite volume, the evolution is nonchaotic. For a generic evolution, the number of collisions with the walls is finite. The system generically settles after a finite time in a Kasner-like motion that lasts all the way to the singularity, see Fig. 1.

The above derivation relied on the use of hyperbolic polar coordinates (ρ, γ) . This is technically useful in that it represents the walls as being located at an asymptotically fixed position in hyperbolic space, namely $w_A(\gamma) = 0$. However, once one has derived the final result (3.40), one can re-express it in terms of the original variables β^μ , which run over a linear (Minkowski) space. Owing to the linearity of $w_A(\beta) = \rho w_A(\gamma)$ in this Minkowskian picture, the asymptotic motion takes place in a “polywedge,” bounded by the hyperplanes $w_A(\beta) = 0$. The billiard motion then consists of free motions of β^μ on straight light-like lines within this polywedge, which are interrupted by specular reflections off the walls. (See Eq. (5.3) below for the explicit effect of these reflections on the components of the velocity vector of the β -particle.) Indeed, when going back to β -space (i.e. before taking the BKL limit), the dynamics of the scale factors at each point of space is given by the Hamiltonian

$$H(\beta^\mu, \pi_\mu) = \tilde{N}\mathcal{H} = \tilde{N} \left[\frac{1}{4} G^{\mu\nu} \pi_\mu \pi_\nu + \sum_A c_A \exp(-2w_A(\beta)) \right]. \quad (3.41)$$

The β -space dynamics simplifies in the gauge $\tilde{N} = 1$, corresponding to the time coordinate τ . In the BKL limit, the Hamiltonian (3.41) takes the limiting form (in the gauge $\tilde{N} = 1$)

$$H_\infty(\beta^\mu, \pi_\mu) = \frac{1}{4} G^{\mu\nu} \pi_\mu \pi_\nu + \sum_{A'} \Theta(-2w_{A'}(\beta)), \quad (3.42)$$

where again the sum runs only over the dominant walls. When taking “equal time” slices of this polywedge (e.g., slices on which $\Sigma_i \beta^i$ is constant), it is clear that with increasing time (i.e. increasing $\Sigma_i \beta^i$, or increasing τ or ρ) the walls recede from the observer. The β -space picture is useful for simplifying the mathematical representation of the dynamics of the scale factors which takes place in a linear space. However, it is inconvenient both for proving that the exponential walls of (3.41) do reduce, in the large $\Sigma_i \beta^i$ limit, to sharp walls, and for dealing with the dynamics of the other phase-space variables

(Q, P) , which were suppressed in the coefficients c_A in Eq. (3.41) above. Let us only mention that in this picture, in order to prove the freezing of the phase-space variables (Q, P) , one must consider in detail accumulation of the “red shifts” of the energy-momentum π_μ of the β -particle when it undergoes reflections on the receding walls, and the effect on the evolution of (Q, P) of the resulting decrease of the magnitudes of the components of π_μ . By contrast, the γ -space picture that we used above allows for a more straightforward treatment of the effect of the limit $\rho \rightarrow +\infty$ on the wall sharpening and freezing of (Q, P) .

In summary, dynamics simplifies enormously in the asymptotic limit. It becomes ultralocal in that it reduces to a continuous superposition of evolution systems (depending only on a time parameter) for the scale factors and the dilatons, at each spatial point, with asymptotic freezing of the off-diagonal and p -form variables. This ultralocal description of the dynamics is valid only asymptotically. It would make no sense to speak of a billiard motion prior to this limit, because one cannot replace the exponentials by Θ -functions. Prior to this limit, the evolution system for the scale factors involves the coefficients in front of the exponential terms, and the evolution of these coefficients depends on various spatial gradients of the other degrees of freedom. However, one may contemplate setting up an expansion in which the sharp wall model is replaced by a model with exponential (“Toda-like”) potentials, and where the evolution of the quantities entering the coefficients of these “Toda walls” is treated as a next to leading effect. See [7] and below for the definition of the first steps of such an expansion scheme for maximal supergravity in eleven dimensions.

3.3. Constraints

We have just seen that in the BKL limit, the evolution equations become ordinary differential equations with respect to time. Although the spatial points are decoupled in the evolution equations, they are, however, still coupled via the constraints. These constraints just restrict the initial data and need only be imposed at one time, since they are preserved by the dynamical equations of motion. Indeed, one easily finds that *in the BKL limit*

$$\dot{\mathcal{H}} = 0, \quad (3.43)$$

since $[\mathcal{H}(x), \mathcal{H}(x')] = 0$ in the ultralocal limit. This corresponds to the fact that the collisions preserve the light-like character of the velocity vector. Furthermore, the gauge constraints (3.16) are also preserved in time since the

Hamiltonian constraint is gauge-invariant. *In the BKL limit*, the momentum constraint satisfies

$$\dot{\mathcal{H}}_k(x) = \partial_k \mathcal{H} \approx 0. \quad (3.44)$$

It is important that the restrictions on the initial data do not bring dangerous constraints on the coefficients of the walls in the sense that these may all take nonzero values. For instance, it is well known that it is consistent with the Gauss law to take nonvanishing electric and magnetic energy densities; thus the coefficients of the electric and magnetic walls are indeed generically nonvanishing. In fact, the constraints are essentially conditions on the spatial gradients of the variables entering the wall coefficients, not on these variables themselves. In some nongeneric contexts, however, the constraints could force some of the wall coefficients to be zero; the corresponding walls would thus be absent. For instance, for vacuum gravity in four dimensions, the momentum constraints for some Bianchi homogeneous models force some symmetry wall coefficients to vanish. But this is specific for the homogeneous case.

It is easy to see that the number of arbitrary physical functions involved in the solution of the asymptotic BKL equations of motion is the same as in the general solution of the complete Einstein-matter equations. Indeed, the number of constraints on the initial data and the residual gauge freedom are the same in both cases. Further discussion of the constraints in the BKL context may be found in [17, 19].

3.4. Consistency of BKL behavior in spite of the increase of spatial gradients

The essential assumption in the BKL analysis is the asymptotic dominance of time derivatives with respect to space derivatives near a space-like singularity. This assumption has been mathematically justified, in a rigorous manner, in the cases where the billiard is of infinite volume, i.e. in the (simple) cases where the asymptotic behavior is not chaotic, but is monotonically Kasner-like [17, 19].

On the other hand, one might a priori worry that this assumption is self-contradictory in those cases where the billiard is of finite volume, when the asymptotic behavior is chaotic, with an infinite number of oscillations. Indeed, it has been pointed out [50, 51] that the independence of the billiard evolution at each spatial point will have the effect of *infinitely increasing* the spatial gradients of various quantities, notably of the local values of the Kasner exponents $p_\mu(x)$. This increase of spatial gradients towards the

singularity has been described as a kind of turbulent behavior of the gravitational field, in which energy is pumped into shorter and shorter length scales [50, 51], and, if it were too violent, it would certainly work against the validity of the BKL assumption of asymptotic dominance of time derivatives. For instance, in our analysis of gravitational walls in the following section, we will encounter subleading walls, whose prefactors depend on spatial gradients of the logarithmic scale factors β .

To address the question of consistency of the BKL assumption we need to know how fast the spatial gradients of β , and of similar quantities entering the prefactors, grow near the singularity. We refer the reader to Ref. [8] for a discussion of this issue. The result is that the chaotic character of the billiard indeed implies an unlimited growth of the spatial gradients of β , but that this growth is only of *polynomial* order in ρ

$$\partial_i \beta = \mathcal{O}(1) \rho^2, \quad (3.45)$$

where the coefficient $\mathcal{O}(1)$ is a chaotically oscillating quantity. This polynomial growth of $\partial_i \beta$ (and of its second-order spatial derivatives) entails a polynomial growth of the prefactors of the sub-dominant walls. Because it is polynomial (in ρ), this growth is, however, negligible compared to the *exponential* (in ρ) behavior of the various potential terms. It does not jeopardize our reasoning based on keeping track of the various exponential behaviors. As we will see the potentially dangerously growing terms that we have controlled here appear only in subdominant walls. The reasoning of the appendix of [8] shows that the prefactors of the dominant walls are self-consistently predicted to evolve very little near the singularity.

We conclude that the unlimited growth of some of the spatial gradients does not affect the consistency of the BKL analysis done here. This does not mean, however, that it will be easy to promote our analysis to a rigorous mathematical proof. The main obstacle to such a proof appears to be the existence of exceptional points, where a prefactor of a dominant wall happens to vanish, or points where a subdominant wall happens to be comparable to a dominant one. Though the set of such exceptional points is (generically) of measure zero, their density might increase near the singularity because of the increasing spatial gradients. This situation might be compared to the KAM (Kolmogorov–Arnold–Moser) one, where the “bad” tori have a small measure, but are densely interspersed among the “good” ones.

4. Walls

The decomposition (3.22) of the Hamiltonian gives rise to different types of walls, which we now discuss in turn. Specifically, we will derive explicit formulas for the linear forms $w_A(\beta)$ and the field dependence of the prefactors c_A entering the various potentials.

4.1. Symmetry walls

We start by analyzing the effects of the off-diagonal metric components which will give rise to the so-called ‘‘symmetry walls’’. As they originate from the gravitational action they are always present. The relevant contributions to the potential is the centrifugal potential (3.27). When comparing (3.27) to the general form (3.34) analyzed above, we see that firstly the summation index A must be interpreted as a double index (a, b) , with the restriction $a < b$, secondly that the corresponding prefactor is $c_{ab} = (P^j_a \mathcal{N}^b_j)^2$ is automatically nonnegative (in accordance with our genericity assumptions, we will assume $c_{ab} > 0$). The centrifugal wall forms read

$$w_{(ab)}^S(\beta) \equiv w_{(ab)\mu}^S \beta^\mu \equiv \beta^b - \beta^a, \quad (a < b). \quad (4.1)$$

We refer to these wall forms as the ‘‘symmetry walls’’ for the following reason. When applying the general collision law (5.3) derived below to the case of the collision on the wall (4.1) one easily finds that its effect on the components of the velocity vector v^μ is simply to permute the components v^a and v^b , while leaving unchanged the other components $\mu \neq a, b$.

The hyperplanes $w_{(ab)}^S(\beta) = 0$ (i.e. the symmetry walls) are time-like since

$$G^{\mu\nu} w_{(ab)\mu}^S w_{(ab)\nu}^S = +2. \quad (4.2)$$

This ensures that the symmetry walls intersect the hyperboloid

$$G_{\mu\nu} \beta^\mu \beta^\nu = -1, \quad \sum_a \beta^a \geq 0.$$

The symmetry billiard (in β -space) is defined to be the region of the Minkowski space determined by the inequalities

$$w_{(ab)}^S(\beta) \geq 0, \quad (4.3)$$

with $\sum_a \beta^a \geq 0$ (i.e. by the region of β -space where the Θ -functions vanish). Its projection on the hyperbolic space H_m is defined by the inequalities $w_{(ab)}^S(\gamma) \geq 0$.

The explicit expressions above for the symmetry wall forms also allow us to illustrate the notion of a “dominant wall” defined above. Indeed, the $d(d-1)/2$ inequalities (4.3) already follow from the *minimal* set of $d-1$ inequalities

$$\beta^2 - \beta^1 \geq 0, \quad \beta^3 - \beta^2 \geq 0, \quad \dots, \quad \beta^d - \beta^{d-1} \geq 0. \quad (4.4)$$

More precisely, each linear form which must be positive in (4.3) can be written as a linear combination, with positive (in fact, integer) coefficients of the linear forms entering the subset (4.4). For instance,

$$\beta^3 - \beta^1 = (\beta^3 - \beta^2) + (\beta^2 - \beta^1),$$

etc. As discussed in [8] this result can be reinterpreted by identifying the dominant linear forms entering (4.4) with the simple roots of $SL(n, \mathbb{R})$.

4.2. Curvature (gravitational) walls

Next we analyze the gravitational potential, which requires a computation of the curvature. To that end, one must explicitly express the spatial curvature in terms of the scale factors and the off-diagonal variables \mathcal{N}^a_i . Again, the calculation is most easily done in the Iwasawa frame (3.2). We use the shorthand notation $A_a \equiv e^{-\beta^a}$ for the (Iwasawa) scale factors. Let $C^a_{bc}(x)$ be the structure functions of the Iwasawa basis $\{\theta^a\}$, viz.

$$d\theta^a = -\frac{1}{2}C^a_{bc}\theta^b \wedge \theta^c, \quad (4.5)$$

where d is the spatial exterior differential. The structure functions obviously depend only on the off-diagonal components \mathcal{N}^a_i , but not on the scale factors. Using the Cartan formulas for the connection one-form ω^a_b ,

$$d\theta^a + \sum_b \omega^a_b \wedge \theta^b = 0, \quad (4.6)$$

$$d\gamma_{ab} = \omega_{ab} + \omega_{ba}, \quad (4.7)$$

where $\omega_{ab} \equiv \gamma_{ac}\omega^c_b$, and

$$\gamma_{ab} = \delta_{ab}A_a^2 \equiv \exp(-2\beta^a)\delta_{ab} \quad (4.8)$$

is the metric in the frame $\{\theta^a\}$, one finds

$$\begin{aligned} \omega^c{}_d = \sum_b \frac{1}{2} \left(C^b{}_{cd} \frac{A_b^2}{A_c^2} + C^d{}_{cb} \frac{A_d^2}{A_c^2} - C^c{}_{db} \frac{A_c^2}{A_d^2} \right) \theta^b \\ + \sum_b \frac{1}{2A_c^2} \left[\delta_{cd}(A_c^2)_{,b} + \delta_{cb}(A_c^2)_{,d} - \delta_{db}(A_d^2)_{,c} \right] \theta^b. \end{aligned} \quad (4.9)$$

In the square brackets above, the commas denote the frame derivatives $\partial_a \equiv \mathcal{N}^i{}_a \partial_i$.

The Riemann tensor $R^c{}_{def}$, the Ricci tensor R_{de} and the scalar curvature R are obtained through

$$\Omega^a{}_b = d\omega^a{}_b + \sum_c \omega^a{}_c \wedge \omega^c{}_b = \frac{1}{2} \sum_{e,f} R^a{}_{bef} \theta^e \wedge \theta^f, \quad (4.10)$$

where $\Omega^a{}_b$ is the curvature 2-form, and

$$R_{ab} = \sum_c R^c{}_{acb}, \quad R = \sum_a \frac{1}{A_a^2} R_{aa}. \quad (4.11)$$

Direct but somewhat cumbersome computations yield

$$R = -\frac{1}{4} \sum_{a,b,c} \frac{A_a^2}{A_b^2 A_c^2} (C^a{}_{bc})^2 + \sum_a \frac{1}{A_a^2} F_a(\partial^2 \beta, \partial \beta, \partial C, C), \quad (4.12)$$

where F_a is some complicated function of its arguments whose explicit form will not be needed here. The only property of F_a that will be of importance is that it is a polynomial of degree two in the derivatives $\partial \beta$ and of degree one in $\partial^2 \beta$. Thus, the exponential dependence on the β 's which determines the asymptotic behavior in the BKL limit, occurs only through the A_a^2 terms written explicitly in (4.12).

In (4.12) one obviously has $b \neq c$ because the structure functions $C^a{}_{bc}$ are antisymmetric in the pair $[bc]$. In addition to this restriction, we can assume, without loss of generality, that $a \neq b, c$ in the first sum on the right-hand side of (4.12). Indeed, the terms with either $a = b$ or $a = c$ can be absorbed into a redefinition of F_a . We can thus write the gravitational potential density (of weight 2) as

$$\mathcal{V}_G \equiv -gR = \frac{1}{4} \sum'_{a,b,c} e^{-2\alpha_{abc}(\beta)} (C^a{}_{bc})^2 - \sum_a e^{-2\mu_a(\beta)} F_a, \quad (4.13)$$

where the prime on \sum indicates that the sum is to be performed only over unequal indices, i.e. $a \neq b, b \neq c, c \neq a$, and where the linear forms $\alpha_{abc}(\beta)$

950 *Thibault Damour*

and $\mu_a(\beta)$ are given by

$$\alpha_{abc}(\beta) = 2\beta^a + \sum_{e \neq a, b, c} \beta^e, \quad (a \neq b, b \neq c, c \neq a) \quad (4.14)$$

and

$$\mu_a(\beta) = \sum_{c \neq a} \beta^c, \quad (4.15)$$

respectively. Note that α_{abc} is symmetric under the exchange of b with c , but that the index a plays a special role.

Comparing the result (4.13) to the general form (3.34) we see that there are, a priori, two types of gravitational walls: the α -type and the μ -type. It is shown in [8] that the α -walls dominate the μ -walls and, more precisely that $\mu_a(\beta) \geq 0$ within the entire future light cone of the β 's. (The proof uses the fact that each linear form $\mu_a(\beta)$ is lightlike.)

From these considerations we deduce the additional constraints

$$\alpha_{abc}(\beta) \geq 0, \quad (D > 3) \quad (4.16)$$

besides the symmetry inequalities (4.4). The hyperplanes $\alpha_{abc}(\beta) = 0$ are called the ‘‘curvature’’ or ‘‘gravitational’’ walls. Like the symmetry walls, they are time-like since

$$G^{\mu\nu}(\alpha_{abc})_{\mu}(\alpha_{abc})_{\nu} = +2. \quad (4.17)$$

The restriction $D > 3$ is due to the fact that in $D = 3$ space-time dimensions, the gravitational walls $\alpha_{abc}(\beta) = 0$ are absent, simply because one cannot find three distinct spatial indices. This is, of course, in agreement with expectations, because gravity in three space-time dimensions has no propagating degrees of freedom (gravitational waves).

4.3. *p*-form walls

While none of the wall forms considered so far involved the dilatons, the electric and magnetic ones do as we will now show. To make the notation less cumbersome we will omit the super-(or sub-)script (p) on the p -form fields in this subsection.

4.3.1. *Electric walls*

The electric potential density can be written as

$$\mathcal{V}_{(p)}^{el} = \frac{1}{2p!} \sum_{a_1, a_2, \dots, a_p} e^{-2e_{a_1 \dots a_p}(\beta)} (\mathcal{E}^{a_1 \dots a_p})^2, \quad (4.18)$$

where $\mathcal{E}^{a_1 \dots a_p}$ are the components of the electric field $\pi^{i_1 \dots i_p}$ in the basis $\{\theta^a\}$

$$\mathcal{E}^{a_1 \dots a_p} \equiv \mathcal{N}^{a_1}_{j_1} \mathcal{N}^{a_2}_{j_2} \dots \mathcal{N}^{a_p}_{j_p} \pi^{j_1 \dots j_p}, \quad (4.19)$$

(recall our summation conventions for spatial coordinate indices) and $e_{a_1 \dots a_p}(\beta)$ are the electric wall forms,

$$e_{a_1 \dots a_p}(\beta) = \beta^{a_1} + \dots + \beta^{a_p} + \frac{\lambda_p}{2} \phi. \quad (4.20)$$

Here all indices a_j 's are distinct because $\mathcal{E}^{a_1 \dots a_p}$ is completely antisymmetric. The variables $\mathcal{E}^{a_1 \dots a_p}$ do not depend on the β^μ . It is thus rather easy to take the BKL limit. The exponentials in (4.18) are multiplied by positive factors which generically are nonvanishing. Thus, in the BKL limit, $\mathcal{V}_{(p)}^{el}$ becomes

$$\mathcal{V}_{(p)}^{el} \simeq \sum_{a_1 < a_2 < \dots < a_p} \Theta[-2e_{a_1 \dots a_p}(\beta)]. \quad (4.21)$$

The transformation from the variables $(\mathcal{N}^a_i, P^i_a, A_{j_1 \dots j_p}, \pi^{j_1 \dots j_p})$ to the variables $(\mathcal{N}^a_i, \mathcal{P}^i_a, \mathcal{A}_{a_1 \dots a_p}, \mathcal{E}^{a_1 \dots a_p})$ is a point canonical transformation whose explicit form is obtained from

$$\sum_a P^i_a \mathcal{N}^{a_i} + \sum_p \frac{1}{p!} \pi^{j_1 \dots j_p} \dot{A}_{j_1 \dots j_p} = \sum_a \mathcal{P}^i_a \mathcal{N}^{a_i} + \sum_p \sum_{a_1, \dots, a_p} \frac{1}{p!} \mathcal{E}^{a_1 \dots a_p} \dot{\mathcal{A}}_{a_1 \dots a_p}. \quad (4.22)$$

The new momenta \mathcal{P}^i_a conjugate to \mathcal{N}^a_i differ from the old ones P^i_a by terms involving \mathcal{E} , \mathcal{N} and \mathcal{A} since the components $\mathcal{A}_{a_1 \dots a_p}$ of the p -forms in the basis $\{\theta^a\}$ depend on \mathcal{N} ,

$$\mathcal{A}_{a_1 \dots a_p} = \mathcal{N}^{j_1}_{a_1} \dots \mathcal{N}^{j_p}_{a_p} A_{j_1 \dots j_p}.$$

However, it is easy to see that these extra terms do not affect the symmetry walls in the BKL limit.

4.3.2. Magnetic walls

One can deal with the magnetic potential in a similar manner. Expressing it in the $\{\theta^a\}$ -frame, one obtains

$$\mathcal{V}_{(p)}^{magn} = \frac{1}{2(p+1)!} \sum_{a_1, a_2, \dots, a_{p+1}} e^{-2m_{a_1 \dots a_{p+1}}(\beta)} (\mathcal{F}_{a_1 \dots a_{p+1}})^2, \quad (4.23)$$

where $\mathcal{F}_{a_1 \dots a_{p+1}}$ are the components of the magnetic field $F_{m_1 \dots m_{p+1}}$ in the basis $\{\theta^a\}$,

$$\mathcal{F}_{a_1 \dots a_{p+1}} = \mathcal{N}^{j_1}_{a_1} \dots \mathcal{N}^{j_{p+1}}_{a_{p+1}} F_{j_1 \dots j_{p+1}}. \quad (4.24)$$

952 *Thibault Damour*

The $m_{a_1 \dots a_{p+1}}(\beta)$ are the magnetic linear forms

$$m_{a_1 \dots a_{p+1}}(\beta) = \sum_{b \notin \{a_1, a_2, \dots, a_{p+1}\}} \beta^b - \frac{\lambda_p}{2} \phi, \quad (4.25)$$

where again all a_j 's are distinct. One sometimes rewrites $m_{a_1 \dots a_{p+1}}(\beta)$ as $\tilde{m}_{a_{p+2} \dots a_d}$, where $\{a_{p+2}, a_{p+3}, \dots, a_d\}$ is the set complementary to $\{a_1, a_2, \dots, a_{p+1}\}$; e.g.,

$$\tilde{m}_{12 \dots d-p-1} = \beta^1 + \dots + \beta^{d-p-1} - \frac{\lambda_p}{2} \phi = m_{d-p \dots d}. \quad (4.26)$$

Of course, the components of the exterior derivative \mathcal{F} of \mathcal{A} in the nonholonomic frame $\{\theta^a\}$ involves the structure coefficients, i.e.

$$\mathcal{F}_{a_1 \dots a_{p+1}} = \partial_{[a_1} \mathcal{A}_{a_2 \dots a_{p+1}]} + C\mathcal{A} \text{ terms},$$

where $\partial_a \equiv \mathcal{N}^i_a \partial_i$ is the frame derivative.

Again, the BKL limit is quite simple and yields (assuming generic magnetic fields)

$$\mathcal{V}_{(p)}^{magn} \simeq \sum_{a_1 < \dots < a_{d-p-1}} \Theta[-2\tilde{m}_{a_1 \dots a_{d-p-1}}(\beta)]. \quad (4.27)$$

Just as the off-diagonal variables, the electric and magnetic fields freeze in the BKL limit since the Hamiltonian no longer depends on the p -form variables. The latter drop out because one can rescale the coefficient of any Θ -function to be one (when it is not zero), thereby absorbing the dependence on the p -form variables.

The scale factors are therefore constrained by the further ‘‘billiard’’ conditions

$$e_{a_1 \dots a_p}(\beta) \geq 0, \quad \tilde{m}_{a_1 \dots a_{d-p-1}}(\beta) \geq 0. \quad (4.28)$$

The hyperplanes $e_{a_1 \dots a_p}(\beta) = 0$ and $\tilde{m}_{a_1 \dots a_{d-p-1}}(\beta) = 0$ are called ‘‘electric’’ and ‘‘magnetic’’ walls, respectively. Both walls are time-like because their gradients are space-like, with the squared norm

$$\frac{p(d-p-1)}{d-1} + \left(\frac{\lambda_p}{2}\right)^2 > 0. \quad (4.29)$$

(For $D = 11$ supergravity, we have $d = 10$, $p = 3$ and $\lambda_p = 0$ and thus the norm is equal to $+2$.) This equality explicitly shows the invariance of the norms of the p -form walls under electric-magnetic duality.

4.4. Subleading walls

The fact that the leading walls originating from the centrifugal, gravitational and p - forms are all time-like, is an important ingredient of the overall BKL picture. We refer the reader to Ref. [8] for a discussion of the other contributions to the Hamiltonian (including the Chern–Simons and Yang–Mills couplings) and for a proof that they contribute only to subleading walls.

5. Einstein (or cosmological) billiards

Let us summarize the findings above. The dynamics in the vicinity of a space-like singularity is governed by the scale factors, while other variables (off-diagonal metric components, p -form fields) tend to become mere “spectators” which get asymptotically frozen. This simple result is most easily derived in terms of the hyperbolic polar coordinates (ρ, γ) , and in the gauge (2.5). In this picture, the essential dynamics is carried by the angular variables γ which move on a fixed billiard table, with cushions defined by the dominant walls $w_{A'}(\gamma)$. One can refer to this billiard as to an “Einstein billiard” (or a “cosmological billiard”).

It is often geometrically more illuminating to “unproject” the billiard motion in the full Minkowski space of the extended scale factors β^μ . In that picture, the asymptotic evolution of the scale factors at each spatial point reduces to a zigzag of null straight lines with respect to the metric $G_{\mu\nu}d\beta^\mu d\beta^\nu$. The straight segments of this motion are interrupted by collisions against the sharp walls

$$w_A(\beta) \equiv w_{A\mu}\beta^\mu = 0 \quad (5.1)$$

defined by symmetry, gravitational and p -form potentials, respectively. As we showed all these walls are *time-like*, i.e. they have space-like gradients,

$$G^{\mu\nu}w_{A\mu}w_{A\nu} > 0. \quad (5.2)$$

Indeed, the gradients of the symmetry and gravitational wall forms have squared norm equal to +2, independently of dimension d . By contrast, the norms of the electric and magnetic gradients, which are likewise positive, depend on the model. As we saw, there also exist subdominant walls which can be neglected as they are located “behind” the dominant walls.

In the β -space picture, the free motion before a collision is described by a null straight line. The effect of a collision on a particular wall $w_A(\beta)$ is easily obtained by solving (3.41), or (3.42), with only one term in the sum. This dynamics is exactly integrable. It suffices to decompose the motion of

the β -particle into two (linear) components: (i) the component parallel to the (time-like) wall hyperplane, and (ii) the orthogonal component. One easily finds that the parallel motion is left unperturbed by the presence of the wall, while the orthogonal motion suffers a (one-dimensional) reflection with a change of the sign of the outgoing orthogonal velocity with respect to the ingoing one. The net effect of the collision on a certain wall $w(\beta)$ then is to change the ingoing velocity vector $v^\mu = d\beta^\mu/d\tau$ entering the ingoing free motion into an outgoing velocity vector v'^μ given (in any linear frame) by the usual formula for a geometric reflection in the hyperplane $w(\beta) = 0$,

$$v'^\mu = v^\mu - 2 \frac{(w \cdot v) w^\mu}{(w \cdot w)}. \quad (5.3)$$

Here, all scalar products, and index raising, are done with the β -space metric $G_{\mu\nu}$. Note that the collision law (5.3) leaves invariant the (Minkowski) length of the vector v^μ . Because the dominant walls are time-like, the geometric reflections that the velocities undergo during a collision, are elements of the orthochronous Lorentz group. Each reflection preserves the norm and the time-orientation; hence, the velocity vector remains null and future-oriented.

From this perspective, we can also better understand the relevance of walls which are *not* time-like. *Light-like* walls (like some of the subleading gravitational walls) can never cause reflections because in order to hit them the billiard ball would have to move at superluminal speeds in violation of the Hamiltonian constraint. The effect of *space-like* walls (like the cosmological constant wall) is again different: they are either irrelevant (if they are “behind the motion”), or otherwise they reverse the time-orientation inducing a motion towards increasing spatial volume (“bounce”).

The *hyperbolic billiard* is obtained from the β -space picture by a radial projection onto the unit hyperboloid of the piecewise straight motion in the polywedge defined by the walls. The straight motion thereby becomes a geodesic motion on hyperbolic space. The “cushions” of the hyperbolic billiard table are the intersections of the hyperplanes (5.1) with the unit hyperboloid, such that the billiard motion is constrained to be in the region defined as the intersection of the half-spaces $w_A(\beta) \geq 0$ with the unit hyperboloid. As we already emphasized, not all walls are relevant since some of the inequalities $w_A(\beta) \geq 0$ are implied by others [4]. Only the dominant wall forms, in terms of which all the other wall forms can be expressed as linear combinations with nonnegative coefficients, are relevant for determining the billiard. Usually, these are the minimal symmetry walls and some of the p -form walls. The billiard region, as a subset of hyperbolic space, is in general

noncompact because the cushions meet at infinity (i.e. at a cusp); in terms of the original scale factor variables β , this means that the corresponding hyperplanes intersect on the light cone. It is important that, even when the billiard is noncompact, the hyperbolic region can have finite volume.

Given the action (2.1) with definite space-time dimension, the menu of fields and dilaton couplings, one can determine relevant wall forms and compute the billiards. For generic initial conditions, we have the following results as to which of the models (2.1) exhibit oscillatory behavior (finite volume billiard) or Kasner-like behavior (infinite volume billiard):

- Pure gravity billiards have finite volume for space-time dimension $D \leq 10$ and infinite volume for space-time dimension $D \geq 11$ [16]. This can be understood in terms of the underlying Kač–Moody algebra AE_d [5]: as shown there, the system is chaotic precisely if the underlying indefinite Kač–Moody algebra is hyperbolic.
- The billiard of gravity coupled to a dilaton always has infinite volume, hence exhibits Kasner-like behavior [17–19].
- If gravity is coupled to p -forms (with $p \neq 0$ and $p < D - 2$) *without a dilaton* the corresponding billiard has a finite volume [3]. The most prominent example in this class is $D = 11$ supergravity, whereas vacuum gravity in 11 dimensions is Kasner-like. The 3-form is crucial for closing the billiard. Similarly, the Einstein-Maxwell system in four (in fact any number of) dimensions has a finite-volume billiard (see [52–54] for a discussion of four-dimensional homogeneous models with the Maxwell fields exhibiting oscillatory behavior).
- The volume of the mixed Einstein-dilaton- p -form system depends on the dilaton couplings. For a given space-time dimension D and a given menu of p -forms there exists a subcritical open domain \mathcal{D} in the space of the dilaton couplings, i.e. an open neighborhood of the origin $\lambda_p = 0$, such that: (i) when the dilaton couplings λ_p belong to \mathcal{D} the general behavior is Kasner-like; but (ii) when the λ_p do not belong to \mathcal{D} the behavior is oscillatory [2, 17]. For all the superstring models, the dilaton couplings do not belong to the subcritical domain and the billiard has finite volume. Note, however, that the superstring dilaton couplings are precisely “critical”, *i.e.* on the borderline between the subcritical and the overcritical domain.

As a note of caution let us point out that some indicators of chaos must be used with care in general relativity, because of reparametrization invariance, and in particular redefinitions of the time coordinate; see [55, 56] for a

discussion of the original Bianchi IX model. In this respect, we refer to [8] for a discussion of the link between the various time coordinates used in the analysis above: t , τ and T .

The hyperbolic billiard description of the (3+1)-dimensional homogeneous Bianchi IX system was first worked out by Chitre [14] and Misner [15]. It was subsequently generalized to inhomogeneous metrics in [22, 24]. The extension to higher dimensions with perfect fluid sources was considered in [23], without symmetry walls. Exterior p -form sources were investigated in [25, 57] for special classes of metric and p -form configurations. The uniform approach (based on the Iwasawa decomposition) summarized above comes from [8].

6. Kač–Moody theoretic formulation: the E_{10} case

Although the billiard description holds for all systems governed by the action (2.1), the billiard in general has no notable regularity property. In particular, the dihedral angles between the faces, which can depend on the (continuous) dilaton couplings, need not be integer submultiples of π . In some instances, however, the billiard can be identified with the fundamental Weyl chamber of a symmetrizable Kač–Moody (or KM) algebra of indefinite type,^g with Lorentzian signature metric [4–6]. Such billiards are called “Kač–Moody billiards”. More specifically, in [4], superstring models were considered and the rank 10 KM algebras E_{10} and BE_{10} were shown to emerge,^h in line with earlier conjectures made in [28, 29]. This result was further extended to pure gravity in any number of space-time dimensions, for which the relevant KM algebra is AE_d , and it was understood that chaos (finite volume of the billiard) is equivalent to hyperbolicity of the underlying Kač–Moody algebra [5]. For pure gravity in $D = 4$ the relevant algebra is the hyperbolic algebra AE_3 first investigated in [58]. Further examples of emergence of Lorentzian Kač–Moody algebras, based on the models of [59, 60], are given in [6].

The main feature of the gravitational billiards that can be associated with the KM algebras is that there exists a group theoretical interpretation of the billiard motion: the asymptotic BKL dynamics is equivalent (in a sense to be made precise below), at each spatial point, to the asymptotic dynamics of a one-dimensional nonlinear σ -model based on a certain infinite-dimensional

^g Throughout this section, we will use abbreviations KM for Kač–Moody, and CSA for Cartan subalgebra.

^h Note that the Weyl groups of the E -family have been discussed in a similar vein in the context of U -duality [31–33].

coset space G/K , where the KM group G and its maximal compact subgroup K depend on the specific model. As we have seen, the walls that determine the billiards are the *dominant walls*. For the KM billiards, they correspond to the *simple roots* of the KM algebra. As we discuss below, some of the subdominant walls also have an algebraic interpretation in terms of higher-height positive roots. This enables one to go beyond the BKL limit and to see the beginnings of a possible identification of the dynamics of the scale factors *and* of all the remaining variables with that of a nonlinear σ -model defined on the cosets of the Kač–Moody group divided by its maximal compact subgroup [7, 8].

The KM theoretic reformulation not only enables us to give a unified group theoretical derivation of the different types of walls discussed in the preceding section, but also shows that *the β -space of logarithmic scale factors, in which the billiard motion takes place, can be identified with the Cartan subalgebra of the underlying indefinite Kač–Moody algebra*. The various types of walls can thus be understood directly as arising from the large field limit of the corresponding σ -models. It is the presence of gravity, which comes with a metric in scale-factor space of Lorentzian signature, which forces us to consider *infinite-dimensional* groups if we want to recover all the walls found in our previous analysis, and this is the main reason we need the theory of KM algebras. For finite-dimensional Lie algebras we obtain only a subset of the walls; one of the cushions of the associated billiard is missing, and one always ends up with a monotonic Kasner-type behavior in the limit $t \rightarrow 0^+$. The absence of chaotic oscillations for models based on finite-dimensional Lie groups is consistent with the classical integrability of these models. While they remain formally integrable for infinite-dimensional KM groups, one can understand the chaotic behavior as resulting from the projection of a motion in an infinite-dimensional space onto a finite-dimensional subspace.

For concreteness, we will only consider one specific example here: the relation between the cosmological evolution of $D = 11$ supergravity and a null geodesic on $E_{10}/K(E_{10})$ [7]. We refer the reader to Ref. [8] for a more general discussion of the link between the KM billiards and KM coset models (including a discussion of the AE_3 case relevant for pure Einstein gravity in $3 + 1$ dimensions). Note also that the relevance of nonlinear σ -models for uncovering the symmetries of M -theory has also been discussed from a different, space-time-covariant, point of view in [34–36], but there it is E_{11} rather than E_{10} that has been proposed as a fundamental symmetry.

The action defining the bosonic part of $D = 11$ supergravity reads

$$S = \int d^{11}x \left[\sqrt{-G} R(G) - \frac{\sqrt{-G}}{48} \mathcal{F}_{\alpha\beta\gamma\delta} \mathcal{F}^{\alpha\beta\gamma\delta} + \frac{1}{(12)^4} \varepsilon^{\alpha_1 \dots \alpha_{11}} \mathcal{F}_{\alpha_1 \dots \alpha_4} \mathcal{F}_{\alpha_5 \dots \alpha_8} \mathcal{A}_{\alpha_9 \alpha_{10} \alpha_{11}} \right], \quad (6.1)$$

where the space-time indices $\alpha, \beta, \dots = 0, 1, \dots, 10$, where $\varepsilon^{01\dots 10} = +1$, and where the four-form \mathcal{F} is the exterior derivative of \mathcal{A} , $\mathcal{F} = d\mathcal{A}$. Note the presence of the Chern–Simons term $\mathcal{F} \wedge \mathcal{F} \wedge \mathcal{A}$ in the action (6.1). Introducing a zero-shift slicing ($N^i = 0$) of the eleven-dimensional space-time, and a *time-independent* spatial zehnbain $\theta^a(x) \equiv E^a_i(x) dx^i$, the metric and four-form $\mathcal{F} = d\mathcal{A}$ become

$$ds^2 = G_{\alpha\beta} dx^\alpha dx^\beta = -N^2(dx^0)^2 + G_{ab}\theta^a\theta^b \quad (6.2)$$

$$\mathcal{F} = \frac{1}{3!} \mathcal{F}_{0abc} dx^0 \wedge \theta^a \wedge \theta^b \wedge \theta^c + \frac{1}{4!} \mathcal{F}_{abcd} \theta^a \wedge \theta^b \wedge \theta^c \wedge \theta^d.$$

We choose the time coordinate x^0 so that the lapse $N = \sqrt{G}$, with $G := \det G_{ab}$ (note that x^0 is not the proper timeⁱ $T = \int N dx^0$; rather, $x^0 \rightarrow \infty$ as $T \rightarrow 0$). In this frame the complete evolution equations of $D = 11$ supergravity read

$$\begin{aligned} \partial_0(G^{ac} \partial_0 G_{cb}) &= \frac{1}{6} G \mathcal{F}^{\alpha\beta\gamma\delta} \mathcal{F}_{\beta\gamma\delta} - \frac{1}{72} G \mathcal{F}^{\alpha\beta\gamma\delta} \mathcal{F}_{\alpha\beta\gamma\delta} \delta_b^a - 2GR^a_b(\Gamma, C), \\ \partial_0(G\mathcal{F}^{0abc}) &= \frac{1}{144} \varepsilon^{abca_1 a_2 a_3 b_1 b_2 b_3 b_4} \mathcal{F}_{0a_1 a_2 a_3} \mathcal{F}_{b_1 b_2 b_3 b_4} \\ &\quad + \frac{3}{2} G \mathcal{F}^{de[ab} C^c]_{de} - G C^e_{de} \mathcal{F}^{dabc} - \partial_d(G\mathcal{F}^{dabc}), \\ \partial_0 \mathcal{F}_{abcd} &= 6\mathcal{F}_{0e[ab} C^e_{cd]} + 4\partial_{[a} \mathcal{F}_{0bcd]}, \end{aligned} \quad (6.3)$$

where $a, b \in \{1, \dots, 10\}$ and $\alpha, \beta \in \{0, 1, \dots, 10\}$, and $R_{ab}(\Gamma, C)$ denotes the spatial Ricci tensor; the (frame) connection components are given by $2G_{ad}\Gamma^d_{bc} = C_{abc} + C_{bca} - C_{cab} + \partial_b G_{ca} + \partial_c G_{ab} - \partial_a G_{bc}$ with $C^a_{bc} \equiv G^{ad} C_{dbc}$ being the structure coefficients of the zehnbain $d\theta^a = \frac{1}{2} C^a_{bc} \theta^b \wedge \theta^c$. (Note the change in sign convention here compared to above.) The frame derivative is $\partial_a \equiv E^i_a(x) \partial_i$ (with $E^a_i E^i_b = \delta^a_b$). To determine the solution at any *given* spatial point x requires knowledge of an infinite tower of spatial gradients; one should thus augment (6.3) by evolution equations for $\partial_a G_{bc}, \partial_a \mathcal{F}_{0bcd}, \partial_a \mathcal{F}_{bcde}$, etc., which in turn would involve higher and higher spatial gradients.

ⁱIn this section we denote the proper time by T to keep the variable t for denoting the parameter of the one-dimensional σ -model introduced below.

The geodesic Lagrangian on $E_{10}/K(E_{10})$ is defined by generalizing the standard Lagrangian on a finite-dimensional coset space G/K , where K is a maximal compact subgroup of the Lie group G . All the elements entering the construction of \mathcal{L} have natural generalizations to the case where G is the group obtained by exponentiation of a hyperbolic KM algebra. We refer the reader to Ref. [26] for basic definitions and results of the theory of the KM algebras, and here only recall that a KM algebra $\mathfrak{g} \equiv \mathfrak{g}(A)$ is generally defined by means of a Cartan matrix A and a set of generators $\{e_i, f_i, h_i\}$ and relations (Chevalley–Serre presentation), where $i, j = 1, \dots, r \equiv \text{rank } \mathfrak{g}(A)$. The elements $\{h_i\}$ span the Cartan subalgebra (CSA) \mathfrak{h} , while the e_i and f_i generate an infinite tower of raising and lowering operators, respectively. The “maximal compact” subalgebra \mathfrak{k} is defined as the subalgebra of $\mathfrak{g}(A)$ left invariant under the Chevalley involution $\omega(h_i) = -h_i$, $\omega(e_i) = -f_i$, $\omega(f_i) = -e_i$. In other words, \mathfrak{k} is spanned by the “antisymmetric” elements $E_{\alpha,s} - E_{\alpha,s}^T$, where $E_{\alpha,s}^T \equiv -\omega(E_{\alpha,s})$ is the “transpose” of some multiple commutator $E_{\alpha,s}$ of the e_i ’s associated with the root α (i.e. $[h, E_{\alpha,s}] = \alpha(h)E_{\alpha,s}$ for $h \in \mathfrak{h}$). Here $s = 1, \dots, \text{mult}(\alpha)$ labels the different elements of $\mathfrak{g}(A)$ having the same root α . The (integer-valued) Cartan matrix of E_{10} is encoded in its Dynkin diagram. See the top diagram in Fig. 2.

The σ -model is formulated in terms of a one-parameter dependent group element $\mathcal{V} = \mathcal{V}(t) \in E_{10}$ and its Lie algebra valued derivative,

$$v(t) := \frac{d\mathcal{V}}{dt} \mathcal{V}^{-1}(t) \in \mathfrak{e}_{10} \equiv \text{Lie } E_{10}. \quad (6.4)$$

The action is $\int dt \mathcal{L}$ with

$$\mathcal{L} := n(t)^{-1} \langle v_{\text{sym}}(t) | v_{\text{sym}}(t) \rangle, \quad (6.5)$$

with a “lapse” function $n(t)$ (not to be confused with N), whose variation gives rise to the Hamiltonian constraint ensuring that the trajectory is a null geodesic. The “symmetric” projection

$$v_{\text{sym}} := \frac{1}{2}(v + v^T)$$

eliminates the component of v corresponding to a displacement “along \mathfrak{k} ”, thereby defining an evolution on the coset space $E_{10}/K(E_{10})$. $\langle \cdot | \cdot \rangle$ is the standard invariant bilinear form on the KM algebra [26].

Because no closed form construction exists for the raising operators $E_{\alpha,s}$, nor their invariant scalar products $\langle E_{\alpha,s} | E_{\beta,t} \rangle = N_{s,t}^{\alpha} \delta_{\alpha+\beta}^0$, a recursive approach based on the decomposition of E_{10} into irreducible representations

of its $SL(10, \mathbb{R})$ subgroup was devised in [7]. (See also Ref. [61] for a decomposition of E_{10} based on its D_9 subgroup.) Let $\alpha_1, \dots, \alpha_9$ be the nine simple roots of $A_9 \equiv sl(10)$ corresponding to the horizontal line in the E_{10} Dynkin diagram, and α_0 the “exceptional” root connected to α_3 . (This root is labeled 10 in Fig. 2.) Its dual CSA element h_0 enlarges A_9 to the Lie algebra of $GL(10)$. Any positive root of E_{10} can be written as

$$\alpha = \ell\alpha_0 + \sum_{j=1}^9 m^j \alpha_j \quad (\ell, m^j \geq 0). \quad (6.6)$$

We call $\ell \equiv \ell(\alpha)$ the “level” of the root α . This definition differs from the usual one, where the (affine) level is identified with m^9 and thus counts the number of appearances of the overextended root α_9 in α [58, 62]. Hence, our decomposition corresponds to a slicing (or “grading”) of the forward light cone in the root lattice by space-like hyperplanes, with only finitely many roots in each slice, as opposed to the light-like slicing for the E_9 representations (involving not only infinitely many roots but also infinitely many affine representations for $m^9 \geq 2$ [58, 62]).

The adjoint action of the A_9 subalgebra leaves the level $\ell(\alpha)$ invariant. The set of generators corresponding to a given level ℓ can therefore be decomposed into a (finite) number of irreducible representations of A_9 . The multiplicity of α as a root of E_{10} is thus equal to the sum of its multiplicities as a weight occurring in the $SL(10, \mathbb{R})$ representations. Each irreducible representation of A_9 can be characterized by its highest weight Λ , or equivalently by its Dynkin labels (p_1, \dots, p_9) where $p_k(\Lambda) := (\alpha_k, \Lambda) \geq 0$ is the number of columns with k boxes in the associated Young tableau. For instance, the Dynkin labels (001000000) correspond to a Young tableau consisting of one column with three boxes, i.e. the antisymmetric tensor with three indices. The Dynkin labels are related to the 9-tuple of integers (m^1, \dots, m^9) appearing in (6.6) (for the highest weight $\Lambda \equiv -\alpha$) by

$$S^{i3}\ell - \sum_{j=1}^9 S^{ij}p_j = m^i \geq 0, \quad (6.7)$$

where S^{ij} is the inverse Cartan matrix of A_9 . This relation strongly constrains the representations that can appear at level ℓ , because the entries of S^{ij} are all positive, and the 9-tuples (p_1, \dots, p_9) and (m_1, \dots, m_9) must both consist of nonnegative integers. In addition to satisfying the Diophantine equations (6.7), the highest weights must be roots of E_{10} , which implies

the inequality

$$\Lambda^2 = \alpha^2 = \sum_{i,j=1}^9 p_i S^{ij} p_j - \frac{1}{10} \ell^2 \leq 2. \quad (6.8)$$

All representations occurring at level $\ell + 1$ are contained in the product of the level- ℓ representations with the $\ell = 1$ representation. Imposing the Diophantine inequalities (6.7), (6.8) allows one to discard many representations appearing in this product. The problem of finding a completely explicit and manageable representation of E_{10} in terms of an infinite tower of A_9 representations is thereby reduced to the problem of determining the outer multiplicities of the surviving A_9 representations, namely the number of times each representation appears at a given level ℓ . The Dynkin labels (all appearing with outer multiplicity one) for the first six levels of E_{10} are

$$\begin{aligned} \ell = 1 & : (001000000); \\ \ell = 2 & : (000001000); \\ \ell = 3 & : (100000010); \\ \ell = 4 & : (001000001), (200000000); \\ \ell = 5 & : (000001001), (100100000); \\ \ell = 6 & : (100000011), (010001000), (100000100), (000000010). \end{aligned} \quad (6.9)$$

The level $\ell \leq 4$ representations can be easily determined by comparison with the decomposition of E_8 under its A_7 subalgebra and use of the Jacobi identity, which eliminates the representations (000000001) at level three and (010000000) at level four. By use of a computer and the E_{10} root multiplicities listed in [62, 63], the calculation can be carried much further [64].

From (6.9) we can now directly read off the $GL(10)$ tensors making up the low level elements of E_{10} . At level zero, we have the $GL(10)$ generators K^a_b obeying $[K^a_b, K^c_d] = K^a_d \delta_b^c - K^c_b \delta_d^a$. The \mathfrak{e}_{10} elements at levels $\ell = 1, 2, 3$ are the $GL(10)$ tensors $E^{a_1 a_2 a_3}$, $E^{a_1 \dots a_6}$ and $E^{a_0 | a_1 \dots a_8}$ with the symmetries implied by the Dynkin labels. The σ -model associates to these generators a corresponding tower of “fields” (depending only on the “time” t): a zehnbein $h^a_b(t)$ at level zero, a three form $A_{abc}(t)$ at level one, a six-form $A_{a_1 \dots a_6}(t)$ at level two, a Young-tensor $A_{a_0 | a_1 \dots a_8}(t)$ at level 3, etc. Writing the generic E_{10} group element in Borel (triangular) gauge as

$$\mathcal{V}(t) = \exp X_h(t) \cdot \exp X_A(t)$$

962 *Thibault Damour*

with $X_h(t) = h^a{}_b K^b{}_a$ and

$$X_A(t) = \frac{1}{3!} A_{abc} E^{abc} + \frac{1}{6!} A_{a_1 \dots a_6} E^{a_1 \dots a_6} + \frac{1}{9!} A_{a_0 | a_1 \dots a_8} E^{a_0 | a_1 \dots a_8} + \dots,$$

and using the E_{10} commutation relations in $GL(10)$ form together with the bilinear form for E_{10} , we find up to third order in level

$$\begin{aligned} n\mathcal{L} = & \frac{1}{4}(g^{ac}g^{bd} - g^{ab}g^{cd})\dot{g}_{ab}\dot{g}_{cd} + \frac{1}{2}\frac{1}{3!}DA_{a_1 a_2 a_3}DA^{a_1 a_2 a_3} \\ & + \frac{1}{2}\frac{1}{6!}DA_{a_1 \dots a_6}DA^{a_1 \dots a_6} + \frac{1}{2}\frac{1}{9!}DA_{a_0 | a_1 \dots a_8}DA^{a_0 | a_1 \dots a_8}, \end{aligned} \quad (6.10)$$

where $g^{ab} = e^a{}_c e^b{}_c$ with

$$e^a{}_b \equiv (\exp h)^a{}_b,$$

and all “contravariant indices” have been raised by g^{ab} . The “covariant” time derivatives are defined by (with $\partial A \equiv \dot{A}$)

$$\begin{aligned} DA_{a_1 a_2 a_3} & := \partial A_{a_1 a_2 a_3}, \\ DA_{a_1 \dots a_6} & := \partial A_{a_1 \dots a_6} + 10A_{[a_1 a_2 a_3} \partial A_{a_4 a_5 a_6]}, \\ DA_{a_1 | a_2 \dots a_9} & := \partial A_{a_1 | a_2 \dots a_9} + 42A_{\langle a_1 a_2 a_3} \partial A_{a_4 \dots a_9 \rangle} \\ & \quad - 42\partial A_{\langle a_1 a_2 a_3} A_{a_4 \dots a_9 \rangle} + 280A_{\langle a_1 a_2 a_3} A_{a_4 a_5 a_6} \partial A_{a_7 a_8 a_9 \rangle}. \end{aligned} \quad (6.11)$$

Here antisymmetrization $[\dots]$, and projection on the $\ell = 3$ representation $\langle \dots \rangle$, are normalized with strength one (e.g. $[[\dots]] = [\dots]$). Modulo field redefinitions, all numerical coefficients in (6.10) and (6.11) are uniquely fixed by the structure of E_{10} .

The Lagrangian (6.5) is invariant under a nonlinear realization of E_{10} such that $\mathcal{V}(t) \rightarrow k_g(t)\mathcal{V}(t)g$ with $g \in E_{10}$; the compensating “rotation” $k_g(t)$ being, in general, required to restore the “triangular gauge”. When g belongs to the nilpotent subgroup generated by the E^{abc} , etc., this symmetry reduces to the rather obvious “shift” symmetries of (6.10) and no compensating rotation is needed. The latter are, however, required for the transformations generated by $F_{abc} = (E^{abc})^T$, etc. The associated infinite number of conserved (Noether) charges are formally given by $J = \mathcal{M}^{-1}\partial\mathcal{M}$, where $\mathcal{M} \equiv \mathcal{V}^T\mathcal{V}$. This can be formally solved in closed form as

$$\mathcal{M}(t) = \mathcal{M}(0) \cdot \exp(tJ). \quad (6.12)$$

The compatibility between (6.12) (indicative of the integrability of (6.10)) and the chaotic behavior of $g_{ab}(t)$ near a space-like singularity is discussed in [8].

The main result of concern here is the following: there exists a *map* between geometrical quantities constructed at a given spatial point x from the supergravity fields $G_{\mu\nu}(x^0, x)$ and $\mathcal{A}_{\mu\nu\rho}(x^0, x)$ and the one-parameter-dependent quantities $g_{ab}(t), A_{abc}(t), \dots$ entering the coset Lagrangian (6.10), under which the supergravity equations of motion (6.3) become *equivalent, up to 30th order in height*, to the Euler-Lagrange equations of (6.10). In the gauge (6.2) this map is defined by $t = x^0 \equiv \int dT/\sqrt{G}$ and

$$\begin{aligned} g_{ab}(t) &= G_{ab}(t, x), \\ DA_{a_1 a_2 a_3}(t) &= \mathcal{F}_{0a_1 a_2 a_3}(t, x), \\ DA^{a_1 \dots a_6}(t) &= -\frac{1}{4!} \varepsilon^{a_1 \dots a_6 b_1 b_2 b_3 b_4} \mathcal{F}_{b_1 b_2 b_3 b_4}(t, x), \\ DA^{b|a_1 \dots a_8}(t) &= \frac{3}{2} \varepsilon^{a_1 \dots a_8 b_1 b_2} (C^b{}_{b_1 b_2}(x) + \frac{2}{9} \delta^b{}_{[b_1} C^c{}_{b_2]c}(x)). \end{aligned} \quad (6.13)$$

The expansion in height $\text{ht}(\alpha) \equiv \ell + \sum m^j$, which controls the iterative validity of this equivalence, is as follows: the Hamiltonian constraint of the coset model (6.10) contains an infinite series of exponential coefficients $\exp(-2\alpha(\beta))$, where α runs over all positive roots of E_{10} , and where $\beta^a \equiv -h^a{}_a$ parametrize the CSA of E_{10} . The *billiard picture* discussed above is equivalent to saying that, near a space-like singularity ($t \rightarrow \infty$), the dynamics of the supergravity fields and of truncated versions of the E_{10} coset fields is asymptotically dominated by the (hyperbolic) Toda model defined by keeping only the exponentials involving the *simple roots* of E_{10} . Higher roots introduce smaller and smaller corrections as t increases. The “*height expansion*” of the equations of motion is then technically defined as a formal BKL-like expansion that corresponds to such an expansion in decreasing exponentials of the Hamiltonian constraint. On the supergravity side, this expansion amounts to an expansion in gradients of the fields in appropriate frames. Level one corresponds to the simplest one-dimensional reduction of (6.3), obtained by assuming that both $G_{\mu\nu}$ and $\mathcal{A}_{\lambda\mu\nu}$ depend only on time [3]; levels 2 and 3 correspond to configurations of $G_{\mu\nu}$ and $\mathcal{A}_{\lambda\mu\nu}$ with a more general, but still very restricted x -dependence, so that e.g. the frame derivatives of the electromagnetic field in (6.3) drop out [65]. In [7] it was checked that, when neglecting terms corresponding to $\text{ht}(\alpha) \geq 30$, the map (6.13) provides a *perfect match* between the supergravity evolution equations (6.3) and the E_{10} coset ones, as well as between the associated Hamiltonian constraints. (In fact, the matching extends to *all real roots* of level ≤ 3 .)

It is natural to view the map (6.13) as embedded in a hierarchical se-

quence of maps involving more and more spatial gradients of the basic supergravity fields. The height expansion would then be a way of revealing step by step a hidden hyperbolic symmetry, implying the existence of a huge nonlocal symmetry of Einstein's theory and its generalizations. Although the validity of this conjecture remains to be established, one can at least show that there is "enough room" in E_{10} for all the spatial gradients. Namely, the search for affine roots (with $m^9 = 0$) in (6.7) and (6.8) reveals three infinite sets of admissible A_9 Dynkin labels $(00100000n)$, $(00000100n)$ and $(10000001n)$ with highest weights obeying $\Lambda^2 = 2$, at levels $\ell = 3n+1, 3n+2$ and $3n+3$, respectively. These correspond to three infinite towers of ε elements

$$E_{a_1 \dots a_n}{}^{b_1 b_2 b_3}, E_{a_1 \dots a_n}{}^{b_1 \dots b_6}, E_{a_1 \dots a_n}{}^{b_0 | b_1 \dots b_8} \quad (6.14)$$

which are symmetric in the lower indices and all appear with outer multiplicity one (together with three transposed towers). Restricting the indices to $a_i = 1$ and $b_i \in \{2, \dots, 10\}$ and using the decomposition $\mathbf{248} \rightarrow \mathbf{80} + \mathbf{84} + \overline{\mathbf{84}}$ of E_8 under its $SL(9)$ subgroup one easily recovers the affine subalgebra $E_9 \subset E_{10}$. The appearance of higher order dual potentials (*à la* Geroch) in the E_9 -based linear system for $D = 2$ supergravity [66] indeed suggests that we associate the E_{10} Lie algebra elements (6.14) to the higher order spatial gradients $\partial^{a_1} \dots \partial^{a_n} A_{b_1 b_2 b_3}$, $\partial^{a_1} \dots \partial^{a_n} A_{b_1 \dots b_6}$ and $\partial^{a_1} \dots \partial^{a_n} A_{b_0 | b_1 \dots b_8}$ or to some of their nonlocal equivalents. Finally, we refer to [8] for a more general discussion of the height expansion of Kač–Moody σ -models, and we note that the approach of [7, 8] can be extended to other physically relevant KM algebras, such as BE_{10} [4, 60] and AE_n [5].

References

1. I. I. Kogan and D. Polyakov, Phys. Atom. Nucl. **66**, 2062 (2003) [Yad. Fiz. **66**, 2111 (2003); hep-th/0212137].
2. T. Damour and M. Henneaux, Phys. Rev. Lett. **85**, 920 (2000) [hep-th/0003139]; see also a short version in Gen. Rel. Grav. **32**, 2339 (2000).
3. T. Damour and M. Henneaux, Phys. Lett. B **488**, 108 (2000) [hep-th/0006171].
4. T. Damour and M. Henneaux, Phys. Rev. Lett. **86**, 4749 (2001) [hep-th/0012172].
5. T. Damour, M. Henneaux, B. Julia and H. Nicolai, Phys. Lett. B **509**, 323 (2001) [hep-th/0103094].
6. T. Damour, S. de Buyl, M. Henneaux and C. Schomblond, JHEP **0208**, 030 (2002) [hep-th/0206125].
7. T. Damour, M. Henneaux and H. Nicolai, Phys. Rev. Lett. **89**, 221601 (2002) [hep-th/0207267].
8. T. Damour, M. Henneaux and H. Nicolai, Class. Quant. Grav. **20**, R145 (2003) [hep-th/0212256].
9. V.A. Belinskii, I.M. Khalatnikov and E.M. Lifshitz, Adv. Phys. **19**, 525 (1970).
10. V.A. Belinskii, I.M. Khalatnikov and E.M. Lifshitz, Sov. Phys. JETP **35**, 838 (1972).
11. V.A. Belinskii, I.M. Khalatnikov and E.M. Lifshitz, Adv. Phys. **31**, 639 (1982).
12. E. M. Lifshitz, I. M. Lifshitz and I. M. Khalatnikov, Sov. Phys. JETP **32**, 173 (1971).
13. D. F. Chernoff and J. D. Barrow, Phys. Rev. Lett. **50**, 134 (1983).
14. D. M. Chitre, Ph. D. Thesis, University of Maryland, 1972.
15. C.W. Misner, in: D. Hobill et al. (Eds), *Deterministic chaos in general relativity*, (Plenum, 1994), p. 317 [gr-qc/9405068].
16. J. Demaret, M. Henneaux and P. Spindel, Phys. Lett. B **164**, 27 (1985).
17. T. Damour, M. Henneaux, A. D. Rendall and M. Weaver, Annales Henri Poincaré **3**, 1049 (2002) [gr-qc/0202069].
18. V.A. Belinskii and I.M. Khalatnikov, Sov. Phys. JETP **36**, 591 (1973).
19. L. Andersson and A.D. Rendall, Commun. Math. Phys. **218**, 479 (2001) [gr-qc/0001047].
20. C.W. Misner, Phys. Rev. Lett. **22**, 1071 (1969).
21. R. T. Jantzen, *Spatially homogeneous dynamics: A unified picture*, gr-qc/0102035.
22. A. A. Kirillov, Sov. Phys. JETP **76**, 355 (1993).
23. A. A. Kirillov and V. N. Melnikov, Phys. Rev. D **52**, 723 (1995) [gr-qc/9408004].
24. V. D. Ivashchuk, A. A. Kirillov and V. N. Melnikov, JETP Lett. **60**, 235 (1994) [Pisma Zh. Eksp. Teor. Fiz. **60**, 225 (1994)].
25. V.D. Ivashchuk and V.N. Melnikov, Class. Quantum Grav. **12**, 809 (1995).
26. V.G. Kač, *Infinite Dimensional Lie Algebras*, Third Edition, (Cambridge University Press, 1990).
27. S. de Buyl and C. Schomblond, *Hyperbolic Kač Moody algebras and Einstein billiards*, hep-th/0403285.
28. B. Julia, Report LPTENS 80/16, Invited Paper Presented at the Nuffield Gravity Workshop, Cambridge, England, June 22 – July 12, 1980.
29. B. Julia, in *Lectures in Applied Mathematics*, AMS-SIAM, vol. 21 (1985), p. 355.
30. J. Brown, O. J. Ganor and C. Helfgott, *M-theory and E_{10} : Billiards, Branes, and Imaginary Roots*, hep-th/0401053.
31. H. Lu, C. N. Pope and K. S. Stelle, Nucl. Phys. B **476**, 89 (1996) [hep-th/9602140].
32. N. A. Obers, B. Pioline and E. Rabinovici, Nucl. Phys. B **525**, 163 (1998) [hep-th/9712084].
33. T. Banks, W. Fischler and L. Motl, JHEP **9901**, 019 (1999) [hep-th/9811194].

34. P. C. West, *Class. Quant. Grav.* **18**, 4443 (2001) [hep-th/0104081].
35. I. Schnakenburg and P. C. West, *Phys. Lett. B* **517**, 421 (2001) [hep-th/0107181].
36. I. Schnakenburg and P. C. West, *Phys. Lett. B* **540**, 137 (2002) [hep-th/0204207].
37. F. Englert, L. Houart, A. Taormina and P. West, *JHEP* **0309**, 020 (2003) [hep-th/0304206].
38. F. Englert and L. Houart, *JHEP* **0405**, 059 (2004) [hep-th/0405082].
39. H. Ringström, *Ann. H. Poincaré* **2**, 405 (2001) [gr-qc/0006035].
40. B. K. Berger, D. Garfinkle, J. Isenberg, V. Moncrief and M. Weaver, *Mod. Phys. Lett. A* **13**, 1565 (1998) [gr-qc/9805063].
41. B. K. Berger, *Living Rev. Relativity* 2002-1 [gr-qc/0201056].
42. A. D. Rendall and M. Weaver, *Class. Quant. Grav.* **18**, 2959 (2001). [gr-qc/0103102].
43. E. Bergshoeff, M. de Roo, B. de Wit and P. van Nieuwenhuizen, *Nucl. Phys. B* **195**, 97 (1982).
44. G. F. Chapline and N. S. Manton, *Phys. Lett. B* **120**, 105 (1983).
45. E. Cremmer, B. Julia and J. Scherk, *Phys. Lett. B* **76**, 409 (1978).
46. G. A. Margulis, *Funct. Anal. Appl.* **4**, 335 (1969).
47. R. E. Howe and C. C. Moore, *J. Functional Analysis* **32**, 72 (1979).
48. R. Zimmer, *Ergodic Theory and Semisimple Groups*, (Birkhauser, Boston, 1984).
49. A. Eskin and C. McMullen, *Duke Math. J.* **71**, 181 (1993).
50. A. A. Kirillov and A. A. Kochnev, *JETP Letters* **46**, 436 (1987).
51. V. A. Belinskii, *JETP Letters* **56**, 422 (1992).
52. R.T. Jantzen, *Phys. Rev. D* **33**, 2121 (1986).
53. V.G. LeBlanc, *Class. Quantum Grav.* **14**, 2281 (1997); **15**, 1607-1626 (1998).
54. M. Weaver, *Class. Quantum Grav.* **17**, 421 (2000) [gr-qc/9909043].
55. N. J. Cornish and J. J. Levin, *Phys. Rev. Lett.* **78**, 998 (1997) [gr-qc/9605029]; *Phys. Rev. D* **55**, 7489 (1997) [gr-qc/9612066].
56. G. Imponente and G. Montani, *Phys. Rev. D* **63**, 103501 (2001) [astro-ph/0102067].
57. V. D. Ivashchuk and V. N. Melnikov, *J. Math. Phys.* **41**, 6341 (2000) [hep-th/9904077].
58. A.J. Feingold and I.B. Frenkel, *Math. Ann.* **263**, 87 (1983).
59. P. Breitenlohner, D. Maison and G. W. Gibbons, *Commun. Math. Phys.* **120**, 295 (1988).
60. E. Cremmer, B. Julia, H. Lu and C. N. Pope, *Higher-dimensional origin of $D = 3$ coset symmetries*, hep-th/9909099.
61. A. Kleinschmidt and H. Nicolai, *$E(10)$ and $SO(9,9)$ invariant supergravity*, hep-th/0407101.
62. V.G. Kač, R.V. Moody and M. Wakimoto, in: *Differential Geometric Methods in Theoretical Physics*, NATO Advanced Research Workshop, Eds. K. Bleuler and M. Werner, (Kluwer, 1988).
63. O. Bärwald, R.W. Gebert and H. Nicolai, *Nucl. Phys. B* **510**, 721 (1998).
64. T. Fischbacher and H. Nicolai, *Low level representations for E_{10} and E_{11}* , hep-th/0301017.
65. J. Demaret, J.L. Hanquin, M. Henneaux and P. Spindel, *Nucl. Phys. B* **252**, 538 (1985).
66. P. Breitenlohner and D. Maison, *Ann. Inst. Poincaré* **46** 215 (1987); H. Nicolai, *Phys. Lett.* **194**, 402 (1987).