# Instantons and unitarity in quantum cosmology
# with fixed four-volume

Alan Daughton[*]

*Department of Physics, Syracuse University, Syracuse, New York 13244–1130, USA*
*and*
*Instituto de Ciencias Nucleares, UNAM, A. Postal 70-543, D.F. 04510, Mexico*[†]


Jorma Louko[‡]

*Department of Physics, Syracuse University, Syracuse, New York 13244–1130, USA*
*and*
*Department of Physics, University of Wisconsin–Milwaukee,*
*P.O. Box 413, Milwaukee, Wisconsin 53201, USA*
*and*
*Department of Physics, University of Maryland, College Park, Maryland 20742–4111, USA*
*and*
*Max-Planck-Institut für Gravitationsphysik, Schlaatzweg 1, D–14473 Potsdam, Germany*[§]


Rafael D. Sorkin[**]

*Department of Physics, Syracuse University, Syracuse, New York 13244–1130, USA*[††]
*and*
*Instituto de Ciencias Nucleares, UNAM, A. Postal 70-543, D.F. 04510, Mexico*
(Revised version, August 1998. Published in *Phys. Rev.* D **58** (1998) 084008.)

---

[*]Electronic address: daughton@nuclecu.unam.mx

[†]Present address.

[‡]Electronic address: louko@aei-potsdam.mpg.de

[§]Present address.

[**]Electronic address: sorkin@suhep.phy.syr.edu

[††]Permanent address.

# Abstract

We find a number of complex solutions of the source-free Einstein equations in the so-called unimodular version of general relativity, and we interpret them as saddle points yielding estimates of a gravitational path integral taken over a space of almost everywhere Lorentzian metrics on a spacetime manifold with topology of the "no-boundary" type. Within this interpretation, we address the compatibility of the no-boundary initial condition with the definability of the *quantum measure*, which reduces in this setting to the normalizability and unitary evolution of the no-boundary wave function $\psi$. We consider three spacetime topologies, $\mathbb{R}^4$, $\mathbb{RP}^4 \# \mathbb{R}^4$, and $\mathbb{R}^2 \times T^2$. (The corresponding truncated manifolds-with-boundary are respectively the closed 4-dimensional disk or ball, the closed 4-dimensional cross-cap, and the product of the two-torus with the closed two-dimensional disk.) The first two topologies we investigate within a Taub minisuperspace model with spatial topology $S^3$, and the third within a Bianchi type I minisuperspace model with spatial topology $T^3$. In each of the three cases there exists exactly one complex solution of the classical Einstein equations (or combination of solutions) that, to the accuracy of our saddle point estimate, yields a wave function compatible with normalizability and unitary evolution. The existence of such solutions tends to bear out the suggestion that the unimodular theory is less divergent than traditional Einstein gravity. In the Bianchi type I case, moreover, the distinguished complex solution is approximately real and Lorentzian at late times, and appears to describe an explosive expansion from zero size at $T = 0$. In this connection, we speculate that a fully normalizable $\psi$ can result only from the imposition of an explicit short distance cutoff. (In the Taub cases, in contrast, the only complex solution with nearly Lorentzian late-time behavior yields a wave function that is normalizable but evolves nonunitarily, with the total probability increasing exponentially in the unimodular "time" in a manner that suggests a continuous creation of new universes at zero volume.) The issue of the stability of these results upon the inclusion of more degrees of freedom is raised.

# I. INTRODUCTION

In formulating the gravitational functional integral on a compact manifold with boundary,

$$\int \mathcal{D}g \, \exp\left[iS(g)\right] \quad , \tag{1.1}$$

one may choose to limit the geometries $g$ that enter the sum by specifying a fixed value for the total 4-volume [1–3]. If $S(g)$ is the Einstein-Hilbert action, this restriction produces a theory whose classical limit is equivalent to Einstein's theory with a cosmological constant, the only difference being that the cosmological constant arises as a constant of integration and not as a prescribed parameter in the action [1–16]. This theory is often called unimodular gravity, owing to the fact that one can alternatively derive it by imposing the coordinate condition $\sqrt{-g} = 1$ in the action prior to variation.

In our view, the motivation for a unimodular modification of gravity is threefold [1,3]. First, it may help explain why the cosmological constant can be so small. Second, it is suggested by analogy with the structure of nonrelativistic quantum mechanics. Third, based on this analogy, it can be expected to improve the convergence of certain expressions that arise in the computation of the quantum measure of a set of histories (*i.e.*, of 4-geometries). This third motivation is the most relevant to the present paper.

In a histories framework for quantum theory (see Refs. [1,3,17–23] and references therein), the quantum measure $\mu$ plays a role analogous to that played by the classical probability-measure in a classically stochastic process such as diffusion. Within a histories framework, no wave function ever need be introduced, but it is often convenient to do so, because $\mu$ can often be computed as $||\psi||^2$ for a suitable $\psi$. In unimodular quantum gravity in particular, one can introduce a wave function on 3-geometries by summing over 4-geometries with fixed 4-volume [1], and the resulting $\psi$ will depend on the 4-volume $T$ of the spacetimes that enter into the sum. Now, if the relation between $\psi$ and the quantum measure $\mu$ in quantum gravity is like that in ordinary quantum mechanics, then, in order that $\mu$ be well-defined, it is necessary first of all that $||\psi||^2$ be finite, and secondly that it be independent of 4-volume, *i.e.*, that the "evolution" of $\psi$ with "time" $T = V$ be unitary. These are the principal questions that we explore in the present paper.

In the present paper, our considerations will be based on a Lagrangian formulation, both for its relative simplicity and because it is the most suitable formulation for dealing with the type of topology change that a "big bang" cosmology entails. Nevertheless, it may be of some interest to sketch here how the unimodular assumption manifests itself in Hamiltonian versions of gravity. To understand what happens to the constraints, it is useful to think in terms of the path integral: the condition of fixed spacetime volume removes one degree of freedom from the permissible deformations of the final hypersurface, and this in turn eliminates one of the infinity of Hamiltonian constraints that are present in the conventional formulation, or rather converts it into a Schrödinger equation expressing the dependence of $\psi$ on $T$. In one particular Hamiltonian scheme for unimodular gravity with closed spatial hypersurfaces, this works out in more detail as follows. The theory contains a pair of canonically conjugate fields that are not present in the canonical formulation of conventional Einstein gravity. One of the new fields specifies the value of the cosmological constant,

3

while the conjugate field carries the information about the spacetime volume bounded by the initial and final spacelike hypersurfaces. Dirac quantization of this Hamiltonian theory leads, in addition to a set of constraint equations, to a Schrödinger-type equation in which the "time" variable can be identified as the four-dimensional spacetime volume. One therefore would expect to adopt a Schrödinger-type Hilbert space in which the Hamiltonian would be a selfadjoint operator, and the wave function would evolve unitarily in "unimodular time." This "unfreezing" of the wave function raised hopes that the interpretational issues of quantum gravity, especially regarding time [24–26], might be more easily tractable within the unimodular theory than in the conventional theory. However, from a histories standpoint, no "problem of time" is evident, and the role of unimodularity would seem to be more technical than interpretational in nature. For some further discussion, see Refs. [15,25].

In this paper we explore the implications of unimodular gravity for quantum cosmology. Specifically we explore its implications for no-boundary initial conditions of the sort proposed by Hartle and Hawking [27–29], Linde [30–34], and Vilenkin [35–40], and more generally, for the framework for topology change set out in Refs. [3,41,42].

The condition we impose is that spacetime be a 4-manifold $M$ that is compact toward the past, with empty initial boundary.[1] When truncated toward the future in order to compute the quantum measure [1,19], $M$ will therefore acquire a future boundary $M_1$ that is *closed* in the technical sense of being compact without boundary.[2] A wave function obtained from a unimodular path integral over such a truncated manifold will have as arguments the 4-volume $T$ and the induced 3-geometry on the 3-manifold $M_1$.

Precisely what kinds of metrics are to be integrated over — or indeed, whether it is possible to define consistently a gravitational functional integral at all in a continuum theory — is a poorly understood issue [42,3,43–45] to which we shall return in section VII. For now, we mention only that the point of view we adopt in the following is that the path integral is originally over (almost everywhere) Lorentzian metrics, and any complex metrics one considers have meaning only insofar as they yield approximations to such a Lorentzian path integral. In the main part of the paper, we will simply assume that the path integral can analyzed in the crudest possible saddle-point approximation; and we will find the complex classical solutions for the unimodular theory, without attempting to control even the semiclassical prefactors. As the unimodular boundary condition requires the Lorentzian 4-volume to be real, the saddle-point geometries are in general *necessarily complex*, although saddle-points with Lorentzian signature will be seen to exist in certain special cases. We

---

[1]This condition can be made precise in the language of Morse theory: $M$ should admit a "height function" $h \geq 0$ with the property that $h^{-1}([0,r])$ is compact and $h^{-1}([0,r))$ is boundary-free for all real $r > 0$.

[2]The manifolds $M$ we consider in the present paper are all such that $M_1$ may be assumed to be a smooth 3-manifold. In cases where topological transitions are not limited to the "moment of birth" of the universe, the level surfaces of a Morse height function are not all manifolds. This suggests that, in computing the quantum measure for such spacetimes, one might want to consider wave functions defined on some sort of correspondingly generalized 3-geometries.

emphasize that this condition of real 4-volume excludes from our framework any geometry with purely Euclidean signature.

As explained above, the crucial consistency conditions for the quantum measure to be defined (and therefore for the path integral to lead to meaningful predictions) are that the path integral give a wave function that, with respect to a suitable measure, is square integrable and evolves unitarily. We investigate these features within two spatially homogeneous minisuperspace models: the Taub model (Bianchi type IX plus an additional U(1) symmetry) with $S^3$ spatial topology, and Bianchi type I with $T^3$ spatial topology and a certain additional discrete symmetry. As the (truncated) no-boundary 4-manifolds, we consider the closed 4-ball $\overline{B}^4$ and the closed 4-dimensional cross-cap $\mathbb{RP}^4 \# \overline{B}^4$ in the Taub model, and the closed disk times the two-torus in Bianchi type I. In all cases, finding the no-boundary saddle points reduces to solving a simple algebraic equation. Having found the saddle points, we first ask whether a saddle-point estimate to the path integral is compatible with a normalizable wave function, for any choice of saddle point(s). If yes, we then ask whether the corresponding wave function evolves unitarily, to the approximation in question. Also, we ask whether any of the saddle-point geometries are approximately Lorentzian at late times. Finally, we ask how the saddle point wave function behaves at $T = 0$, and in particular, whether this behavior seems compatible with the picture of a universe expanding from zero size that is implicit in the no-boundary topology.

In the Taub model, for both of our 4-manifolds, we find a unique saddle point that, to the accuracy of our estimate, is compatible with both normalizability and unitary evolution. This bears out the suggestion [1–3] that the unimodular theory is less divergent than traditional Einstein gravity, and tells us in each case what the approximate behavior of the wave function must be if the quantum measure is indeed well defined. Interestingly, this saddle point remains always in the quantum era, never making a spontaneous transition to classical behavior. In addition there is (for both 4-manifolds) a unique saddle point that is compatible with normalizability and does make a transition to classical behavior. However, the wave function corresponding to this saddle point turns out *not* to evolve unitarily: instead, probability is being injected into the configuration space at a rate that is exponentially increasing in the unimodular time. This injection appears to take place at a boundary of the configuration space, in a manner reminiscent of the tunneling boundary conditions advocated by Linde [30–34] and Vilenkin [35–40]. Physically, such an injection can perhaps be interpreted as a continuous creation of new "branch universes," all stemming from a single root.

In the Bianchi type I model, the unique saddle point that is compatible with normalizability turns out to be compatible also with unitary evolution. Further, the saddle-point geometries are, at late times, nearly Lorentzian, isotropically expanding universes. Thus, this saddle point exhibits many features normally regarded as desirable for quantum cosmology.

The plan of the paper is as follows. In section II we introduce the unimodular Taub minisuperspace model and the unimodular positive curvature Friedmann model, which arises as the isotropic specialization of the Taub model. Sections III and IV discuss the Taub no-boundary saddle points when the 4-manifold is respectively the closed 4-ball and the closed cross-cap, and section V discusses the truncation of these no-boundary analyses to

the Friedmann model. The Bianchi type I model is analyzed in section VI. Our results are summarized and discussed in section VII.

We use throughout units such that $c = \hbar = 1$, but we keep Newton's constant $G$. A metric with signature $(-+++)$ is called Lorentzian, and a metric with signature $(++++)$ Riemannian.

## II. TAUB MINISUPERSPACE IN THE UNIMODULAR THEORY

In this section we describe the Taub minisuperspace model in the unimodular theory. Subsection II A presents the general case, with two independent scale factors. The truncation to the positive curvature Friedmann model is outlined in subsection II B. We assume throughout this section that the spacetime is everywhere Lorentzian and has topology $\mathbb{R} \times S^3$. The boundary conditions needed to express the no-boundary topologies we consider will be introduced in section III.

### A. The general Taub model

The Taub family of metrics can be written as [46,47]

$$ds^2 = \sigma^2 \left\{ -N^2 dt^2 + \tfrac{1}{4}a^2(\omega^1)^2 + \tfrac{1}{4}b^2\left[(\omega^2)^2 + (\omega^3)^2\right] \right\} \quad , \tag{2.1}$$

where $a$, $b$, and $N$ are functions of $t$, and the $\omega^i$ are the usual left-invariant one-forms on SU(2), satisfying

$$d\omega^i = -\frac{1}{2}\epsilon^i{}_{jk}\omega^j \wedge \omega^k \quad . \tag{2.2}$$

We use conventions in which the exterior derivative and the wedge product are $(d\omega)_{ab} = \partial_a\omega_b - \partial_b\omega_a$ and $(\omega \wedge \phi)_{ab} = \frac{1}{2}(\omega_a\phi_b - \omega_b\phi_a)$, and we have extracted in (2.1) the overall factor $\sigma^2 := 2G/3\pi$ for numerical convenience. As $\sigma$ has the dimension of length, we can take $a$, $b$, $N$, $t$, and $\omega^i$ to be dimensionless. In the special case $a = b$, the spatial sections are round 3-spheres with radius of curvature $\sigma a$.

The spacetime topology is $\mathbb{R} \times \text{SU}(2) \simeq \mathbb{R} \times S^3$, and the spacetime isometry group is that of the constant $t$ hypersurfaces, $\text{U}(2) \simeq \text{SU}(2)_L \times \text{U}(1)_R/\mathbb{Z}_2$. The SU(2) factor comes from the invariance of (2.1) under the left action of SU(2) on itself, and the further U(1) isometry (acting on the right) expresses the equality of the coefficients of $(\omega^2)^2$ and $(\omega^3)^2$.

Inserting the metric (2.1) into the gravitational action-integral,

$$S = \frac{1}{8\pi G}\left[\int(\tfrac{1}{2}R - \Lambda)dV + \oint \text{tr}K\right] \quad , \tag{2.3}$$

with (bare) cosmological constant $\Lambda$, yields the minisuperspace action-integral

$$S = \tfrac{1}{6}\int d\tau \left[-a\left(\frac{db}{d\tau}\right)^2 - 2b\frac{da}{d\tau}\frac{db}{d\tau} + 4a - a^3b^{-2} - 3\lambda ab^2\right] \quad , \tag{2.4}$$

where we have introduced the dimensionless proper time parameter $\tau$ by $d\tau := Ndt$ and written $\lambda := \frac{1}{3}\sigma^2\Lambda$. The true proper time is $\sigma\tau$.

Given the action-integral, it is easy to derive the classical equations of motion in both the unimodular and non-unimodular theories. In the non-unimodular theory, the classical equations of motion result from making arbitrary variations of $S$ that fix the metric on the boundary. Because the ansatz (2.1) expresses invariance under a compact symmetry group, it suffices to consider variations of the parameters $a$, $b$, and $\tau$. The condition of fixed boundary metric is then equivalent to fixing $a$ and $b$ (but not $\tau$) at the endpoints.[3]

The general solution to the variational equations can be written in the gauge $Na = 1$ as

$$b^2 = A^2 + t^2/A^2 \quad , \tag{2.5a}$$
$$a^2b^2 = A^2\left\{4\left(A^2 - t^2/A^2\right) + 3\lambda\left[t^4/(3A^4) + 2t^2 - A^4\right]\right\} + ABt \quad , \tag{2.5b}$$
$$N = 1/a \quad , \tag{2.5c}$$

where $A > 0$ and $B \in \mathbb{R}$ are the two constants of integration. This metric covers the spatially homogeneous region of the Taub-NUT-de Sitter solution [48]. In the notation of Ref. [48], the Taub-NUT "charge" $l$ and "mass-parameter" $m$ are given by $l = \frac{1}{2}\sigma A$ and $m = \frac{1}{4}\sigma B$.

For the purposes of the unimodular theory, it is convenient to write the action-integral in a slightly different form. First of all, we may set $\Lambda$ to zero in (2.3) without loss of generality, because it influences neither the classical solutions nor the quantum measure. (We obtain the classical equations of motion by varying $S$ subject to the condition that the total 4-volume is fixed.) Second, because of the special role played by the spacetime 4-volume, it is useful to adopt it as our time coordinate (see for example Ref. [1]). For numerical convenience, we may use the dimensionless 4-volume parameter $T$ defined by $dT = ab^2 d\tau$: the 4-volume bounded by the hypersurfaces $T = T_1$ and $T = T_2$, with $T_1 < T_2$, is then $2\pi^2\sigma^4(T_2 - T_1)$. Writing (2.4) in terms of $T$ (with $\lambda$ set to zero) we obtain the action-integral in the form

$$S = \tfrac{1}{6}\int dT\left(-a^2b^2b'^2 - 2ab^3a'b' + 4b^{-2} - a^2b^{-4}\right) \quad , \tag{2.6}$$

where the prime denotes derivative with respect to $T$. In the variation of (2.6), fixing the total spacetime volume is equivalent to fixing the difference between the initial and final values of $T$, and as the integrand does not involve $T$ explicitly, this is further equivalent to fixing individually the initial and final values of $T$. The general solution to the resulting variational equations is precisely (2.5), but $\lambda$ emerges now as an integration constant proportional to the "unimodular energy," and the general solution contains thus three constants of integration.

It is convenient to replace $a$ and $b$ by the coordinates (cf. [1,49])

---

[3]In an action that retains $t$ and $N$, the values of $t$ at the endpoints can be fixed, and the equation that results from variation with respect to $N$ is equivalent to the equation obtained by varying (2.4) with respect to $\tau$ at the endpoints.

$$u := a^2b \; , $$
$$v := b^3 \; , \tag{2.7}$$

which represent spatial volumes rather than lengths. In these coordinates, the action-integral (2.6) simplifies to

$$S = \int dT \left( -\tfrac{1}{18} u'v' - \tfrac{1}{6} uv^{-5/3} + \tfrac{2}{3} v^{-2/3} \right) \; . \tag{2.8}$$

The configuration space of the theory is therefore the future quadrant of a (1+1)-dimensional Minkowski space, and $(u, v)$ forms a pair of positive-valued null coordinates. (Notice that in this system, the "radial coordinate" $\sqrt{uv}$ represents the spatial volume.) The Hamiltonian operator corresponding to (2.8) is

$$\widehat{H} := 18 \frac{\partial^2}{\partial u \partial v} + \tfrac{1}{6} uv^{-5/3} - \tfrac{2}{3} v^{-2/3} \; , \tag{2.9}$$

where we have chosen to factor-order the kinetic part of the Hamiltonian as the D'Alembertian (or "Laplace-Beltrami operator") of the metric associated to the kinetic terms in the Lagrangian with $T$ used as time-parameter. Notice that a change of time-parameter would change this metric and consequently change what we would call the D'Alembertian. Conversely, the combination of 4-volume time with Laplace-Beltrami ordering with respect to that time provides a universal factor-ordering for all spatially homogeneous cosmological models [50]. The Hilbert space inner product that goes naturally with a metric is the $L^2$ integral with respect to the metric's volume element, and the Laplace-Beltrami operator is formally selfadjoint with respect to this inner product, as is easily seen in general. In the present case this inner product is

$$(\psi_1, \psi_2) := \int_{\substack{u>0 \\ v>0}} du\, dv\, \overline{\psi_1} \psi_2 \; . \tag{2.10}$$

As $\widehat{H}$ is real, it has selfadjoint extensions by von Neumann's theorem (Ref. [51], Theorem X.3). Choosing one such extension specifies a unitary evolution in the Hilbert space of square integrable functions $\psi$.

### B. Friedmann truncation

The Taub model can be truncated to its isotropic special case by setting $a = b$. The metric (2.1) becomes then the metric of the positive curvature Friedmann model,

$$ds^2 = \sigma^2 \left[ -N^2 dt^2 + a^2 d\Omega_3^2 \right] \; , \tag{2.11}$$

where $d\Omega_3^2$ is the metric of the unit 3-sphere.[4] The isometry group of the constant $t$ hypersurfaces is O(4). The action (2.6) becomes

---

[4]Our conventions in the Taub model were chosen so as to make the Friedmann metric (2.11) agree with the conventions of Refs. [28,29].

$$S = \tfrac{1}{2} \int dT \left(-a^4 a'^2 + a^{-2}\right) \quad . \tag{2.12}$$

The Hamiltonian operator corresponding to (2.12) is

$$\widehat{H} := \frac{1}{2} \left[ 9 \left( \frac{d^2}{dx^2} \right) - x^{-2/3} \right] \quad , \tag{2.13}$$

where $x := a^3$ and we have again chosen the Laplace-Beltrami factor ordering belonging to $T$ as time parameter. The matching inner product is

$$(\psi_1, \psi_2) := \int_0^\infty dx \, \overline{\psi_1} \psi_2 \quad . \tag{2.14}$$

and $\widehat{H}$ is formally selfadjoint with respect to it. The selfadjoint extensions of $\widehat{H}$ are specified by the boundary condition $\cos(\theta)\psi - \sin(\theta)d\psi/dx = 0$ at $x = 0$, where the parameter $\theta$ satisfies $0 \leq \theta < \pi$ (Ref. [51], Theorems X.8 and X.10). Choosing the value of $\theta$ specifies a unitary quantum evolution.

## III. TAUB NO-BOUNDARY SADDLE POINTS FOR $\overline{B}^4$ TOPOLOGY

In this section we discuss the Taub no-boundary saddle points and wave functions when the 4-manifold is the closed 4-dimensional ball $\overline{B}^4$ (often called disk $\mathbb{D}^4$). In subsection III A we find the saddle-point geometries, and in subsection III B we discuss the saddle-point estimates to the wave function.

As discussed in the Introduction, we introduce a wave function via the (formal) no-boundary path integral

$$\Psi_{\mathrm{NB}}(a, b; T) = \int \mathcal{D}g \, \exp\left(iS\right) \quad . \tag{3.1}$$

where the domain of integration is some appropriate class of Lorentzian, or almost Lorentzian, 4-geometries, possibly with a short-distance cutoff. In the histories entering the integral, the "final" values of the scale factors and the total elapsed volume parameter $T$ are required to coincide with the arguments of the wave function $\Psi_{\mathrm{NB}}$; in particular, $T$ in $\Psi_{\mathrm{NB}}$ is assumed always positive.

In this paper we only consider $\Psi_{\mathrm{NB}}$ in the saddle-point estimate:

$$\Psi_{\mathrm{NB}} \approx \sum_k P_k \exp\left(iS^k\right) \quad , \tag{3.2}$$

where the $S^k(a, b; T)$ are the actions of some subset of the complex classical solutions on the no-boundary 4-manifold in question. The prefactors $P_k$ are assumed to be slowly varying compared with the exponential factors but otherwise are left unspecified. Which (if any) of the saddle point metrics actually contribute to (3.2), given a proper definition of the integral in (3.1), will be left a subject for future work (see the comments on this issue in Ref. [3]).

9

## A. Saddle-point geometries

To find the actions $S^k(a, b; T)$ that enter the saddle-point estimate (3.2), we need on $\overline{B}^4$ the Taub solutions with prescribed values of the boundary scale factors and total 4-volume. The boundary data for the solutions is Lorentzian, but the saddle-point geometries are allowed to be complex.

As a first step, we need the general complex Taub solution to the unimodular field equations. It is clear that one class of complex Taub solutions is obtained from the Lorentzian solution (2.5) by extending the parameters $A$, $B$, and $\lambda$ to complex values, with $A \neq 0$, and making $t$ a complex-valued function $t(s)$ of a new, real-valued time coordinate $s$. We may assume $dt/ds \neq 0$. It can be shown that the only complex Taub solutions not obtained in this way are obtained in the same way from the metrics given by

$$b^2 = \pm 2it \quad , \tag{3.3a}$$

$$a^2 = 2\lambda t^2 \pm 2it \pm iDt^{-1} \quad , \tag{3.3b}$$

$$N = 1/a \quad , \tag{3.3c}$$

where $D$ and $\lambda$ are complex parameters. The family (3.3) can be formally recovered from (2.5) by setting $B = \pm 8iA\left(1 - \lambda A^2\right) \pm 2iDA^{-3}$, adding $\pm iA^2$ to $t$, and taking the limit $A \to \infty$.

Now, the condition that the solution be defined on $\overline{B}^4$ means that $s$ must be interpretable as the radial coordinate of hyperspherical coordinates on $\overline{B}^4$. Without loss of generality we can take $s \in [0, 1]$, such that $s = 0$ occurs at the coordinate singularity at the origin of the hyperspherical coordinates, and $s = 1$ occurs on the boundary. It is then necessary that both $a$ and $b$ vanish at $s = 0$.

Let us first concentrate on the family (2.5). Writing $t(0) =: t_0$, the condition that $a$ and $b$ vanish at $s = 0$ implies

$$t_0 = i\eta A^2 \quad , \tag{3.4a}$$

$$B = 8i\eta A\left(1 - \lambda A^2\right) \quad , \tag{3.4b}$$

where $\eta$ is a parameter that may take the values $\pm 1$. When equations (3.4) hold, it is straightforward to verify that the metric is indeed extendible to $s = 0$ so that it defines a solution on $\overline{B}^4$. Thus the smoothness required of the no-boundary metric is automatic in this case, and does not further restrict the solution.

To find the action, we recall that when deriving (2.6) from (2.3) (with $\Lambda = 0$), we assumed the spacetime topology $\mathbb{R} \times S^3$, for which the boundary term in (2.3) consists of components both at the initial and final boundaries. In the action for the topology $\overline{B}^4$, (2.3) contains a boundary term only at a "final" boundary, and we therefore need to add to (2.6) at $s = 0$ the appropriate boundary term, $-\frac{1}{6}N^{-1}d(ab^2)/dt$. This boundary term however vanishes by virtue of (3.4). (As earlier with smoothness of the metric, assuming a no-boundary topology thus turns out to be indistinguishable, in this case, from simply assuming that the universe expands from zero diameter.) Evaluating (2.6) yields then

$$S = i\eta \left[\frac{\lambda A^4 \rho_*^2(\rho_* + 1)(\rho_* + 4)}{6(\rho_* + 2)} + \frac{2A^2 \rho_*(2\rho_* + 3)}{3(\rho_* + 2)}\right] \quad , \tag{3.5}$$

where we have parametrized the value of $t$ at $s = 1$ as $t_0(1 + \rho_*)$.

In terms of $A$ and $\rho_*$, the boundary scale factors and the total elapsed $T$ are given by

$$a^2 = -\frac{A^2 \rho_* [4 + \lambda A^2 \rho_*(\rho_* + 4)]}{\rho_* + 2} \quad , \tag{3.6a}$$

$$b^2 = -A^2 \rho_*(\rho_* + 2) \quad , \tag{3.6b}$$

$$T = -i\eta A^4 \rho_*^2 \left(\tfrac{1}{3}\rho_* + 1\right) \quad . \tag{3.6c}$$

Solving the algebraic equations (3.6) for $A^2$, $\rho_*$, and $\lambda$ yields

$$A^2 = \frac{b^2(\epsilon V - 1)^2}{\epsilon V - 2} \quad , \tag{3.7a}$$

$$\rho_* = -2 + \frac{2}{\epsilon V - 1} \quad , \tag{3.7b}$$

$$\lambda = \frac{4}{\epsilon V b^2} \left[ 1 - \frac{a^2}{b^2(\epsilon V - 1)^2} \right] \quad , \tag{3.7c}$$

where $\epsilon$ is a parameter that may take the values $\pm 1$,

$$V := \sqrt{1 + 12 i\eta T / b^4} \quad , \tag{3.8}$$

and we choose the branch of the square root in (3.8) so that the real part of $V$ is positive.

It is clear from (3.7) and (3.8) that all the four sign combinations for $\eta$ and $\epsilon$ yield complex geometries that satisfy our boundary conditions. In terms of the boundary data, the actions of these geometries read

$$S_\epsilon^\eta(a, b; T) = \frac{i\eta}{3} \left[ \frac{a^2}{2} \left( 1 - \frac{2}{\epsilon V - 1} \right) - b^2(\epsilon V - 1) \right] \quad . \tag{3.9}$$

Following the same strategy starting from the exceptional family (3.3) yields no solutions: regularity at $s = 0$ implies $t_0 = 0$ and $D = 0$, and positivity of the total elapsed $T$ is then incompatible with the positivity of $b^2$ at the boundary.

None of the solution geometries we have found in this section is Lorentzian, or indeed real with any signature. Rather all are genuinely complex. Although Euclidean signature metrics have played no role in their derivation, it may nevertheless be of interest to see what happens when these metrics are continued to purely imaginary values of the unimodular time $T$. (In particular, this may bear on the important question of whether one can reach our saddle point metric from a Lorentzian metric without leaving the kind of domain defined in Section 5 of Ref. [52].) When $T$ is analytically continued to imaginary values as $T = -i\eta T_R$ with $T_R > 0$, it can be verified that the geometries with $\epsilon = 1$ continue to Riemannian geometries that satisfy the unimodular boundary data on $\overline{B}^4$, with $T_R$ proportional to the total Riemannian 4-volume; conversely, it can be verified that these are the only Riemannian Taub solutions on $\overline{B}^4$ with the unimodular boundary data.[5] The continuation sends $i S_1^\eta$ to $-\eta I$, where

---

[5]The singularity in (3.7a) at $V = 2$ is a coordinate effect. A manifestly regular Riemannian form for the metric with $V = 2$ can be found as a Riemannian section of the exceptional family (3.3).

$I$ is the Riemannian unimodular action. Therefore, our complex solutions with $\epsilon = 1$ are related to solutions of the Riemannian theory by an analytic continuation in the unimodular time: the Wick rotation is in the usual direction for $\eta = 1$ and in the unusual direction for $\eta = -1$. When the same continuation is done to our solutions with $\epsilon = -1$, one obtains geometries on $\overline{B}^4$ that satisfy the Riemannian unimodular data, in particular with positive Riemannian volume, and $iS^\eta_{-1}$ again continues to real values. These geometries are however not Riemannian but complex.

## B. Saddle-point-estimate wave functions

We turn now to the saddle-point estimate (3.2) to the no-boundary wave function $\Psi_{\mathrm{NB}}$, with the classical actions $S^\eta_\epsilon$ of (3.9). We assume throughout the pre-exponential factors to be so slowly varying that the dominant behavior of each term in (3.2) arises from the exponentials. We assume also that the set of saddle points contributing to $\Psi_{\mathrm{NB}}$ (*i.e.*, the range of the index $k$ in (3.2)) does not depend on the values of the parameters $T$, $a$, and $b$.[6]

It can be verified that $S^\eta_\epsilon$ satisfies the time-dependent Hamilton-Jacobi equation with the Hamiltonian (2.9), as by construction it must. Consequently, the estimate (3.2) satisfies the Schrödinger equation

$$i\frac{\partial \psi}{\partial T} = \widehat{H}\psi \tag{3.10}$$

in the approximation to which we are working. We ask: which, if any, of the terms in (3.2) are compatible with $\Psi_{\mathrm{NB}}$ being a wave function in the Hilbert space $\mathsf{H}$ defined in subsection II A, evolving unitarily by some selfadjoint extension of the Hamiltonian (2.9)?

Consider first the case $\eta = -1$. As the first term in $S^{-1}_\epsilon$ has a negative imaginary part for either sign of $\epsilon$, the terms in (3.2) with $\eta = -1$ diverge exponentially as $a \to \infty$ with fixed $b$ (or, in the coordinates $(u, v)$ of subsection II A, as $u \to \infty$ with fixed $v$). These saddle points are therefore not compatible with $\Psi_{\mathrm{NB}}$ having finite $L^2$ norm.

Consider then the case $\eta = 1$. For fixed $T$, the imaginary part of $S^1_\epsilon$ is bounded below. Without further knowledge of the pre-exponential factor, one cannot therefore exclude the wave functions $\Psi_+ := P_+ \exp(iS^1_1)$ and $\Psi_- := P_- \exp(iS^1_{-1})$ from being normalizable. When $T/b^4 \ll 1$, we have

$$S^1_1 = -\frac{a^2 b^4}{18T} + \cdots \quad , \tag{3.11a}$$

$$S^1_{-1} = \frac{i(a^2 + 2b^2)}{3} + \cdots \quad . \tag{3.11b}$$

In this limit, the wave function $\Psi_+$ is therefore rapidly oscillating, while $\Psi_-$, on the other hand, is exponentially suppressed for large $a$ or $b$. One can also see that in this limit, the

---

[6]In section IV we will have to generalize this assumption slightly to cover the case where two saddle points degenerate for certain values of $a$, $b$ and $T$.

saddle point metric with $\epsilon = 1 = \eta$ is (except near $s = 0$) close to a Lorentzian Taub-NUT-de Sitter universe with $m = 0$. We note that a similar conclusion was reached within the conventional Einstein theory in Ref. [53].

To address the unitarity of the evolution, we note that for fixed $a$ and $b$, $S_\epsilon^1$ has the large $T$ expansion

$$S_\epsilon^1 = 2\epsilon e^{-i\pi/4}\sqrt{T/3} + \tfrac{1}{6}i(a^2 + 2b^2) + O\left(T^{-1/2}\right) \quad . \tag{3.12}$$

The contribution to the inner product $(\Psi_+, \Psi_+)$ from any compact region in the configuration space is therefore exponentially increasing in $T$. This means that $\Psi_+$ cannot evolve by any selfadjoint extension of the Hamiltonian (2.9): probability is being injected into the configuration space either from the finite boundaries or from infinity. As we do not have an estimate for the saddle-point prefactor, we shall not attempt to discuss exactly where the probability is entering the configuration space. However, if the saddle-point form is good for $T/b^4 \gg 1$, equation (3.12) shows that the probability is flowing in the direction of isotropic *expansion*, as measured by $\text{Imag}(\Psi^*\nabla\Psi)$; and for $T/b^4 \ll 1$, equation (3.11a) exhibits a similar direction of flow. This suggests that the flux is coming from somewhere on the finite boundary. The emerging picture is perhaps compatible with the tunneling proposals advocated by Linde [30–34] and Vilenkin [35–40]. Notice, however, that the latter proposals are couched in terms of the non-unimodular theory, where the relevant boundary is (for the Taub model) only one-dimensional. For us, it is two-dimensional (the extra dimension being parameterized by $T$), and there can be no simple correspondence between boundary conditions formulated in the two frameworks.

In this connection notice that, precisely because the unimodular $\psi$ depends on a parameter time $T$, the early-time behavior of the universe shows up much more directly in $\psi$ than it would in a non-unimodular formulation, and one can ask whether $\psi$ is concentrated for $T = 0$ at universes of zero size, as the no-boundary picture would seem to require. For the wave function $\Psi_+$, equation (3.11a) does appear to describe a state for which the probability is concentrated near configurations of small 3-volume (in the sense that the rapid oscillations cause $\Psi_+$ to vanish uniformly in the distributional sense as $T \to 0$, in any region of compact support disjoint from the boundary $uv = 0$).[7] For the wave function $\Psi_-$, equation (3.11b) describes a universe that at birth is of Planckian dimensions in all directions, which for quantum gravitational purposes is probably as close as one would care to get to strictly zero size.

With $\epsilon = -1$, we see from (3.12) that the contribution to the inner product $(\Psi_-, \Psi_-)$ from any compact region in the configuration space is exponentially decreasing in $T$. One might see this as evidence for a different type of non-unitary evolution of $\Psi_-$, with probability now flowing out of the configuration space. Perhaps, however, one cannot exclude that $\Psi_-$ might still approximate a unitarily evolving wave function in some relevant sense: the probability would just be spreading out sufficiently fast to give an exponential suppression at late times

---

[7]More precisely, $\Psi_+$ seems to describe a violent explosion starting from zero volume, analogously to how a nonrelativistic particle initially at the origin at $t = 0$ is at any later moment uniformly distributed throughout space with a wave function $\propto \exp(ix^2/t)$.

in any fixed compact region of the configuration space. Resolving this question would seem to require, at a minimum, a better understanding of the prefactor in our saddle-point estimate.

In summary, the wave functions with $\eta = -1$ cannot represent states in the Hilbert space $\mathsf{H}$. The wave function with $\eta = \epsilon = 1$ may represent a state in $\mathsf{H}$, but it has an exponentially growing norm and thus cannot evolve unitarily. The wave function with $\eta = 1$ and $\epsilon = -1$ may represent a state in the Hilbert space, and, at the level of our semiclassical estimate, this state may evolve unitarily. It is therefore the only one that might be compatible, in a Lorentzian histories framework, with a universe consisting of a single macroscopic component with a single moment of birth.

## IV. TAUB NO-BOUNDARY SADDLE POINTS FOR CROSS-CAP TOPOLOGY

In this section we discuss the Taub saddle points and wave functions when the (truncated) 4-manifold is the closed 4-dimensional cross-cap, $\overline{M}_\otimes := \mathbb{RP}^4 \setminus B^4 \simeq \mathbb{RP}^4 \# \overline{B}^4$, where $B^4$ is the open 4-dimensional ball.[8] In subsection IV A we find the saddle-point geometries, and in subsection IV B we discuss the saddle-point estimates to the wave function.

### A. Saddle-point geometries

It is useful to view $\overline{M}_\otimes$ as a quotient space. To this end, let $\widetilde{M} := [-1, 1] \times S^3$. $\widetilde{M}$ is a compact orientable manifold with boundary $S^3 \cup S^3$. Consider the map $J : \widetilde{M} \to \widetilde{M}$; $(s, \mathbf{x}) \mapsto (-s, P\mathbf{x})$, where $P : \mathbf{x} \mapsto P\mathbf{x}$ is the antipodal map on $S^3$. $J$ is an involution with a free and properly discontinuous action, and the quotient space $\widetilde{M}/J$ is diffeomorphic to $\overline{M}_\otimes$. One can also visualize $\overline{M}_\otimes$ as built by closing off the upper half $s > 0$ of $\widetilde{M}$ by attaching $s = 0$ to an $\mathbb{RP}^3$. Clearly, $\overline{M}_\otimes$ is a nonorientable compact manifold with boundary $S^3$.

We need first the general complex Taub geometry on $\overline{M}_\otimes$. To begin, we recall from section III that the complex Taub geometries on $\widetilde{M}$ are obtained from (2.5) or (3.3) by letting the parameters take complex values and writing the coordinate $t$ as a complex-valued function $t(s)$, with $dt/ds \neq 0$. We take the boundaries of $\widetilde{M}$ to be at $s = \pm 1$. On these geometries, we realize $J$ as the map $(s, \mathbf{x}) \mapsto (-s, P\mathbf{x})$, where $P$ acts on $\mathrm{SU}(2) \simeq S^3$ as multiplication by $\mathrm{diag}(-1, -1)$ in the defining matrix representation. In order that $J$ be an isometry, it is necessary that $db/ds$ and $da/ds$ vanish at $s = 0$. This excludes (3.3), and shows that (2.5) is only possible with $t(0) = 0$ and $B = 0$. The condition that $J$ be an isometry implies, finally, $t(s) = -t(-s)$. Quotienting this geometry on $\widetilde{M}$ with respect to $J$ now gives the desired general complex geometry on $\overline{M}_\otimes$.

Next, we need to match the complex geometry on $\overline{M}_\otimes$ to the unimodular boundary data. From $t(1) = -t(-1) =: t_*$, we have

$$T = \int_0^{t_*} Nab^2 \, dt = t_* \left( A^2 + \frac{t_*^2}{3A^2} \right) \quad , \tag{4.1}$$

---

[8]This manifold was suggested to us by John Friedman.

and the action is evaluated from (2.6) to be

$$S = \frac{\lambda t_*}{6}\left(-5A^2 - \frac{t_*^2}{A^2} + \frac{8A^6}{A^4 + t_*^2}\right) + \frac{4t_*^3}{3(A^4 + t_*^2)} \quad . \tag{4.2}$$

The values of $\lambda$, $A^2$, and $t_*$ can now be found in terms of $T$ and the final scale factors from (2.5a), (2.5b), and (4.1). The general solution to this system of algebraic equations is

$$t_* = yb^2/z \quad , \tag{4.3a}$$

$$A^2 = b^2/z \quad , \tag{4.3b}$$

$$\lambda = \frac{z\left[(a^2/b^2)z^2 + 4z - 8\right]}{b^2\left(z^2 + 4z - 8\right)} \quad , \tag{4.3c}$$

where

$$z = 1 + y^2 \quad , \tag{4.4a}$$

and $y$ is any solution to

$$\frac{3T}{b^4} = \frac{y(3 + y^2)}{(1 + y^2)^2} \quad . \tag{4.4b}$$

Every root of (4.4b) does, with one exception, yield a saddle-point geometry on $\overline{M}_\otimes$ with our boundary conditions. The exception occurs when $z = 2\left(\sqrt{3} - 1\right)$, in which case the denominator of (4.3c) vanishes.

With (4.3) and (4.4), the action (4.2) can be written as

$$S = \frac{Tz}{2b^2(2 + z)}\left(4 - \frac{a^2 z}{b^2}\right) \quad . \tag{4.5}$$

As (4.5) depends on $y$ only through $z$, the distinct saddle-point values of the action are obtained by finding the distinct values of $z$ from (4.4). This is equivalent to solving for $z$ the quartic

$$\left(\frac{3T}{b^4}\right)^2 = \frac{(z - 1)(z + 2)^2}{z^4} \quad . \tag{4.6}$$

It is easily verified that the action (4.5), with $z$ given by any root of (4.6) that is a smooth function of $T/b^4$, satisfies the Hamilton-Jacobi equation.

The roots of (4.6) depend on the quantity $\alpha := 3T/b^4 > 0$. Let

$$\alpha_c := 2^{-5/2}3^{3/4}\left(\sqrt{3} + 1\right) \quad , \tag{4.7a}$$

$$z_c := 2\left(\sqrt{3} - 1\right) \quad . \tag{4.7b}$$

For $\alpha < \alpha_c$, there are two distinct real roots, denoted by $z_+$ and $z_-$, satisfying $1 < z_- < z_c < z_+$. For $\alpha = \alpha_c$, these two roots merge at $z_c$: this is the special case in which (4.3c)

15

becomes singular. For $\alpha > \alpha_c$, these two roots become a complex conjugate pair, denoted by $(z_1, z_2)$, where $z_1$ is in the first quadrant and $z_2 = \overline{z}_1$. The two remaining roots are always a complex conjugate pair, denoted by $(z_3, z_4)$, where $z_3$ is in the second quadrant and $z_4 = \overline{z}_3$. The only instance in which a root is not a smooth function of $\alpha$ occurs in the transition of the pair $(z_+, z_-)$ into the pair $(z_1, z_2)$ at $\alpha = \alpha_c$.

The roots $z_+$ and $z_-$ of (4.6) give Lorentzian saddle-point geometries on $\overline{M}_\otimes$. These geometries have the peculiarity that they are not time-orientable, and they therefore would fall into the framework of time-nonorientable cobordisms [54], if they were regarded as being in the domain of the Lorentzian path integral.[9] All the other saddle-point geometries are complex.

We denote the action (4.5) evaluated at the respective roots by $S^+$, $S^-$, $S^1$, $S^2$, $S^3$, and $S^4$. For later use, we note that the roots have for $\alpha \ll 1$ and $\alpha \gg 1$ the respective expansions

$$z_+ = \alpha^{-2} + 3 + O(\alpha^2) \tag{4.8a}$$

$$z_- = 1 + \tfrac{1}{9}\alpha^2 + O(\alpha^4) \tag{4.8b}$$

$$z_3 = \overline{z}_4 = -2 + \frac{4i\alpha}{\sqrt{3}}\left[1 - \frac{10i\alpha}{3\sqrt{3}} + O(\alpha^2)\right] \ , \tag{4.8c}$$

and

$$z_1 = \overline{z}_2 = e^{i\pi/4}\sqrt{2/\alpha}\left[1 + O(\alpha^{-1})\right] \ , \tag{4.9a}$$

$$z_3 = \overline{z}_4 = e^{3i\pi/4}\sqrt{2/\alpha}\left[1 + O(\alpha^{-1})\right] \ . \tag{4.9b}$$

### B. Saddle-point-estimate wave functions

We are now ready to consider the saddle-point estimate (3.2) to the no-boundary wave function. As in section III, we assume throughout that the pre-exponential factors are not important, and that the set of saddle points contributing to $\Psi$ does not change in ranges of $(T, a, b)$ where the corresponding saddle point geometries vary continuously. As before, we write $\Psi_+$, $\Psi_-$, $\Psi_1$, and so on, for the saddle point wave functions corresponding to $S^+$, $S^-$, $S^1$, etc. Notice, however, that while $\Psi_3$ and $\Psi_4$ are defined on all of configuration space, the pair $\Psi_\pm$ (respectively $\Psi_1$, $\Psi_2$) is defined only for $\alpha < \alpha_c$ (respectively $\alpha > \alpha_c$).

It can be shown that the second term in $S^1$ and $S^4$ has a negative imaginary part. From the limit $a \to \infty$ with fixed $b$ and $T$, it is therefore seen that $\Psi_1$ and $\Psi_4$ are not normalizable.

---

[9]For a discussion of quantum field theory on time-nonorientable spacetimes, see Ref. [55]. Because of their time-non-orientability, we suspect that these metrics are not valid saddle points for a path integral that is originally taken over almost everywhere Lorentzian metrics with a well-defined causal structure. See section VII for some further thoughts on how one might in principle recognize the valid saddle points.

The imaginary part of $S^3$, on the other hand, is positive, and the corresponding term $\Psi_3$ in (3.2) may represent a state in $\mathsf{H}$. For $\alpha \ll 1$ with $a \approx b$, we have

$$S^3 = \frac{i(a^2 + 2b^2)}{2\sqrt{3}} + \cdots \quad . \tag{4.10}$$

This again describes a Planck sized universe at $T = 0$. The contribution to the inner product $(\Psi_3, \Psi_3)$ from any compact region in the configuration space is exponentially decreasing in $T$ (see equation (4.12b) below). As in section III, we regard this as consistent with unitary evolution, although it also might signify a loss of probability through the finite boundary.

The only remaining possibility in (3.2) is a wave function that coincides with $\Psi_2 = P_2 \exp(iS^2)$ for $\alpha > \alpha_c$, and with some combination of $\Psi_+$ and $\Psi_-$ for $\alpha < \alpha_c$. We denote this wave function by $\Psi_0$. The curve $\alpha = \alpha_c$ in the configuration space is analogous to a turning point in a constant-energy WKB approximation, and a saddle-point estimate to $\Psi_0$ would not be expected to be good near this curve. However, beginning as $\Psi_2$ for $\alpha > \alpha_c$, $\Psi_0$ presumably resumes a saddle-point form for $\alpha < \alpha_c$, now as a linear combination of $\Psi_+$ and $\Psi_-$. Since the imaginary part of $S^2$ can be shown to be bounded below, and $S^\pm$ are purely real, $\Psi_0$ may thus be normalizable.

When $\alpha < \alpha_c$, the two terms in $\Psi_0$ have each an immediate semiclassical interpretation, as they each come from a Lorentzian universe (2.5) with $B = 0$. The two parameters in this family are $A$ and $\lambda$. The reason why *two* such solutions exist is that there are two choices for $A$ and $\lambda$, obtained from (4.3b) and (4.3c) with respectively $z = z_\pm$, that make a spacetime in this family pass through a prescribed point in the configuration space at a prescribed value of $T$. For $\alpha = 3T/b^4 \ll 1$ with $a \approx b$, we have

$$S^+ = -\frac{a^2 b^4}{18T} + \cdots \quad , \tag{4.11a}$$

$$S^- = \frac{T}{6b^2}\left(4 - \frac{a^2}{b^2}\right) + \cdots \quad . \tag{4.11b}$$

From (4.11) we also see that at $T = 0$, $\Psi_+$ behaves similarly to $\Psi_+$ in the previous section, as if the universe had exploded from zero size. On the other hand, $\Psi_-$ is quite different from its earlier namesake, looking like a zero momentum state spread out over all of configuration space.

Consider, finally, whether the norms of $\Psi_0$ and $\Psi_3$ can be independent of $T$. For fixed $a$ and $b$, (4.9) gives the large $T$ expansion

$$S^2 = e^{-i\pi/4}\sqrt{2T/3} + \frac{i(a^2 + 2b^2)}{6} + O(T^{-1/2}) \quad , \tag{4.12a}$$

$$S^3 = e^{3i\pi/4}\sqrt{2T/3} + \frac{i(a^2 + 2b^2)}{6} + O(T^{-1/2}) \quad . \tag{4.12b}$$

It follows that $\Psi_3$ dies out with time at any fixed $a$, $b$, whence its evolution can be unitary. On the other hand, the contribution to the inner product $(\Psi_0, \Psi_0)$ from any compact region in the configuration space is seen to be exponentially increasing in $T$, and $\Psi_0$ cannot evolve unitarily. As in section III, it is difficult to ascertain where the probability is entering the

17

configuration space, but the semiclassical discussion given above for $\alpha \ll 1$ suggests that the flux may be coming from somewhere on the finite boundary, as before.

In summary, the qualitative results with the 4-manifold $\overline{M}_{\otimes}$ are very similar to those obtained with the 4-manifold $\overline{B}^4$ in section III. Of the four saddle points, two lead to non-square-integrable wave functions, analogously to the case $\eta = -1$ earlier. The normalizable cases here are those of $\Psi_3$ and $\Psi_0$. The wave function $\Psi_3$ here is analogous to $\Psi_-$ there: it may be normalizable, it evolves consistently with unitarity, it is nowhere rapidly oscillating, and it describes a universe that has Planckian size at $T = 0$. Similarly, the wave function $\Psi_0$ here is analogous to $\Psi_+$ there: it may be normalizable, it has a WKB form expressing a classically evolving universe in a suitable limit, but its norm cannot be independent of $T$. Its behavior at $T = 0$ differs from that of the earlier $\Psi_+$, however, in a way related to the degeneracy of $z_1$ and $z_2$ at $\alpha = \alpha_c$.

## V. FRIEDMANN TRUNCATION OF THE TAUB SADDLE POINTS AND WAVE FUNCTIONS

The Taub saddle-point metrics found in sections III and IV clearly specialize to saddle-points of the Friedmann model by setting $a = b$, and the corresponding actions are the restrictions to $a = b$ of those found earlier.

For most of the saddle points, the discussion within the Friedmann model proceeds in parallel with that in the Taub model.[10] In particular, with the 4-manifold $\overline{B}^4$, the saddle point with $\eta = \epsilon = 1$ yields an exponentially growing probability flux, and this flux must now enter the configuration space at the boundary $a = 0$. With the 4-manifold $\overline{M}_{\otimes}$, a similar argument can be made for the wave function $\Psi_0$.

The only qualitative difference between the Taub analysis and the Friedmann analysis occurs for the saddle point with $\eta = -1$ and $\epsilon = 1$ on $\overline{B}^4$, and for the saddle point with $z_1$ on $\overline{M}_{\otimes}$. In the two Taub models, the corresponding wave functions (call them $\Psi_+^-$ and $\Psi_1^-$) were seen not to be normalizable. In the Friedmann situation, however, the imaginary part of the action turns out to be bounded below and the $a = b$ restrictions of these wave functions are in fact square integrable. (With $\overline{M}_{\otimes}$, $\Psi_1^-$ covers only part of the configuration space. However, when $\alpha < \alpha_c$, it becomes a linear combination of terms arising from $S^{\pm}$, so the full wave function $\Psi_0^-$ is also normalizable.) In the Friedmann restriction, therefore, these saddle points are both compatible with normalizability. Moreover, for fixed $a$, these wave functions are exponentially decreasing in $T$, and are even compatible with unitary evolution.

The drastic qualitative change in these two saddle-point contributions upon passing from the Friedmann model to the Taub model suggests that the "good" behavior of these saddle points in the Friedmann model should be seen as an artifact of the isotropic truncation. We shall return to this question in section VII.

---

[10] The results reported for the Friedmann model with the 4-manifold $\overline{B}^4$ in Ref. [56] only considered the saddle points with $\epsilon = 1$, inadvertently excluding the saddle points with $\epsilon = -1$.

# VI. BIANCHI TYPE I

In this section we discuss the unimodular no-boundary saddle points and wave functions in a Bianchi type I minisuperspace model. We take the spatial topology to be $T^3$, and the (truncated) no-boundary 4-manifold to be $\mathbb{D}^2 \times T^2$, where $\mathbb{D}^2$ is the closed disk. We set up the unimodular quantum theory in subsection VI A. The no-boundary saddle points and wave functions are analyzed in subsection VI B.

## A. Unimodular quantization of Bianchi type I

The local spatial homogeneity of Bianchi type I is compatible with ten distinct closed spatial topologies [57]. The number of minisuperspace degrees of freedom depends on the spatial topology [58–65], and the spatial topology also determines the group of large spatial diffeomorphisms that can be incorporated as gauge invariances of the model [59,60]. The topology also determines the possible ways of compactifying the manifold toward the past to obtain a manifold of no boundary type [59,60].

We shall here focus on a Bianchi type I model with an additional discrete symmetry group reminiscent of the additional U(1) symmetry that distinguishes the Taub class of metrics within Bianchi type IX. The results obtained for the conventional Einstein theory in Refs. [59,60] suggest that this specialized model should faithfully reflect the general Bianchi type I situation regarding the normalizability and unitary evolution of the wave function.

The metric of our Bianchi type I model reads

$$ds^2 = \rho^2 \left[ -N^2 dt^2 + a^2 dx^2 + b^2 \big( dy^2 + dz^2 \big) \right] \quad , \tag{6.1}$$

where $a$, $b$, and $N$ are functions of $t$, and the overall factor $\rho^2 := G/(2\pi^2)$ has been introduced for numerical convenience. In this subsection we take $a^2$, $b^2$, and $N^2$ to be positive, so that the metric is Lorentzian. The identifications made on the spatial coordinates are $(x, y, z) \sim (x + 2\pi, y, z) \sim (x, y + 2\pi, z) \sim (x, y, z + 2\pi)$, and the spatial topology is thus $T^3$. The metric (6.1) is obtained from the most general Bianchi type I metric with $T^3$ spatial topology by imposing the extra symmetry $\mathbb{Z}_2 \times D_8$ where the 8-element dihedral group $D_8$ is the symmetry group of the square. This is equivalent to demanding that the spatial metric have three *orthogonal* closed geodesics, and that two of these geodesics have equal length.

To derive the solutions of the conventional Einstein theory and the unimodular theory, we proceed as in section II. Inserting the metric (6.1) into the action-integral (2.3) with bare cosmological constant $\Lambda$, and introducing the proper time parameter $\tau$ by $d\tau = N dt$, we obtain the minisuperspace action

$$S = \tfrac{1}{2} \int d\tau \left[ -a \left( \frac{db}{d\tau} \right)^2 - 2b \frac{da}{d\tau} \frac{db}{d\tau} - \lambda a b^2 \right] \quad , \tag{6.2}$$

where $\lambda := \rho^2 \Lambda$. This action reproduces the full Bianchi type I Einstein equations with a cosmological constant under variations that fix the initial and final values of the scale factors but not those of $\tau$. For $\lambda \neq 0$, the general Lorentzian solution can be written in the gauge $Na = 1$ as

$$b = At \quad, \tag{6.3a}$$

$$a^2b = \tfrac{1}{3}\lambda At^3 + E \quad, \tag{6.3b}$$

$$N = 1/a \quad, \tag{6.3c}$$

where $A \neq 0$ and $E$ are constants. For $\lambda = 0$, the solutions not obtained from (6.3) with $\lambda = 0$ can be put in the form

$$b = B \quad, \tag{6.4a}$$

$$a^2b = Dt + E \quad, \tag{6.4b}$$

$$N = 1/a \quad, \tag{6.4c}$$

where $B \neq 0$, $D$, and $E$ are constants, $D$ and $E$ not both equal to zero.

In order to put the action-integral in a form convenient for the unimodular theory, we introduce the parameter time $T$ by $dT = ab^2 d\tau$. As before, we also simplify the action, without loss of generality in the unimodular theory, by setting the bare cosmological constant to zero. The integral (6.2) then takes the form

$$S = -\tfrac{1}{2} \int dT\, ab^2(ab'^2 + 2ba'b') \quad, \tag{6.5}$$

where the prime denotes derivative with respect to $T$. The 4-volume bounded by the hypersurfaces $T = T_1$ and $T = T_2$, with $T_1 < T_2$, is $8\pi^3\rho^4(T_2 - T_1)$, and fixing the 4-volume in the variation of (6.5) is therefore equivalent to fixing the initial and final values of $T$. The unimodular variational equations clearly reproduce the Einstein equations, with the cosmological constant now emerging as the integration constant proportional to the unimodular energy.

From (6.5), the unimodular Hamiltonian operator is

$$\widehat{H} := 6\frac{\partial^2}{\partial u \partial v} \quad, \tag{6.6}$$

where the coordinates $(u, v)$ are defined by (2.7), and we have adopted the Laplace-Beltrami factor ordering, as in section II. The matching inner product is again (2.10). As $\widehat{H}$ is symmetric and real, it has selfadjoint extensions by von Neumann's theorem.

### B. No-boundary saddle points and wave functions

The general complex solution with our Bianchi I symmetries is obtained from the Lorentzian solutions (6.3) and (6.4) by extending the parameters to complex values and making $t$ a complex-valued function $t(s)$ of a real-valued time coordinate $s$. We may assume $dt/ds \neq 0$. The condition that the solution be defined on $\mathbb{D}^2 \times T^2$ means that $s$ must be interpretable as the radial coordinate of polar coordinates on $\mathbb{D}^2$. Taking $s \in [0, 1]$ as in subsection III A, with $s = 0$ occurring at the coordinate singularity, it is then necessary that $a$ vanish at $s = 0$ but $b$ remain nonzero there. It is straightforward to show from (6.3) and (6.4) that the general complex solution with this property can be written as

20

$$b = B\left(1 + \lambda \tilde{A}t\right) \quad , \tag{6.7a}$$

$$a^2 b = t(B/\tilde{A})\left(1 + \lambda \tilde{A}t + \tfrac{1}{3}\lambda^2 \tilde{A}^2 t^2\right) \quad , \tag{6.7b}$$

$$N = 1/a \quad , \tag{6.7c}$$

where $B$ and $\tilde{A}$ are nonvanishing complex constants, and we have chosen $t(0) = 0$. The absence of a conical singularity at $s = 0$ requires $N^{-1}da/dt \to i\eta$ as $s \to 0$, where $\eta$ is a parameter that takes the values $\pm 1$: this implies $\tilde{A} = -\tfrac{1}{2}i\eta$. The metric then defines a solution on $\mathbb{D}^2 \times T^2$ in the sense we seek.

The total elapsed $T$ is

$$T = \int_0^{t_*} Nab^2\, dt = t_* B^2\left(1 - \tfrac{1}{2}i\eta\lambda t_* - \tfrac{1}{12}\lambda^2 t_*^2\right) \quad , \tag{6.8}$$

where we have written $t(1) =: t_*$. Solving for $B$, $\lambda$, and $t_*$ in terms of $T$ and the boundary values of the scale factors, we obtain

$$B = \frac{2i\eta T}{a^2 b} \quad , \tag{6.9a}$$

$$\lambda = \frac{8T}{3a^4 b^2}\left[\left(\frac{a^2 b^2}{2T}\right)^3 + i\eta\right] \quad , \tag{6.9b}$$

$$t_* = \frac{3a^4 b^2}{4T\left[\left(\dfrac{a^2 b^2}{2T}\right)^2 + \dfrac{i\eta a^2 b^2}{2T} - 1\right]} \quad . \tag{6.9c}$$

As discussed in subsection III A, the action contains the integral term (2.6) as well as a boundary term from $s = 0$, and for the metric (6.1) the boundary term reads[11] $-\tfrac{1}{2}N^{-1}d(ab^2)/dt$ [45,66–68]. The value of the integral term is $-\tfrac{1}{2}\lambda T$, and that of the boundary term is $-\tfrac{1}{2}i\eta B^2$. In terms of the boundary data, the action reads

$$S^\eta(a, b; T) = -\frac{a^2 b^4}{6T} + \frac{2i\eta T^2}{3a^4 b^2} \quad . \tag{6.10}$$

It is easily verified that this action satisfies the Hamilton-Jacobi equation.

---

[11] The boundary term contributes here because we are essentially in a 2-dimensional situation. Its presence marks a genuine difference between the point of view that the spacetime manifold is a cobordism with empty initial boundary, and the point of view that it has an initial boundary of zero spatial volume. In the Taub models we considered, this distinction was effectively moot because the analogous boundary term did not contribute. Similarly, the assumption that the saddle point metric must be smooth is also playing an important role here, in contrast to the Taub case. The boundary term is somewhat reminiscent of the pure imaginary topological contribution to the *Lorentzian* action-integral found in Ref. [52].

The solution geometries are genuinely complex, analogously to those found for the Taub model with $\mathbb{R}^4$ (untruncated) topology. As before, we have not tried to analyze directly which, if any, of the saddle points can be reached from an almost Lorentzian metric on the same manifold by an admissible deformation. We can, however, glean some indirect evidence on this by considering the Wick rotation to the *Riemannian* case. When $T$ is analytically continued to imaginary values as $T = -i\eta T_R$ with $T_R > 0$, the geometries continue to Riemannian geometries that satisfy the unimodular boundary data on $\mathbb{D}^2 \times T^2$, with $T_R$ proportional to the total Riemannian 4-volume; conversely, these are the only Riemannian solutions of our Bianchi type I model on $\mathbb{D}^2 \times T^2$ with the unimodular boundary data. The continuation sends $iS^\eta$ to $-\eta I$, where $I$ is the Riemannian unimodular action. The situation is thus as for $\epsilon = 1$ in section III: the complex spacetimes with the Lorentzian boundary conditions are related to solutions of the Riemannian theory by an analytic continuation in the unimodular time, with a Wick rotation in the usual direction for $\eta = 1$ and in the unusual direction for $\eta = -1$.

We can now turn to the saddle-point estimate (3.2). For fixed $T$, the wave function with $\eta = -1$ diverges exponentially as $a^2 b \to 0$, and consequently cannot be square integrable.[12] The wave function with $\eta = +1$, on the other hand, is compatible with normalizability. Moreover, this wave function decays exponentially as $T \to \infty$ at fixed $a$ and $b$.

As explained in section II, we officially regard such behavior as consistent with unitary evolution. In this case, moreover, it appears plausible that probability actually is flowing toward infinity, rather than escaping through the finite boundary. Indeed, if we limit ourselves to values of $a, b \geq 1$ (meaning that none of the dimensions of 3-space has sub-Planckian scale), then it is easy to see that the estimate $\Psi = O(1)e^{iS}$ implies that the probability for the 3-volume $V = ab^2$ to be less than $\sqrt{T}$ is exponentially small in $T^2/V^4$. Here it is helpful to rewrite $S^\eta$ in the form

$$iS^\eta(a, b; T) = -\frac{iuv}{6T} - \frac{2\eta T^2}{3u^2} \quad , \tag{6.11}$$

where $u = a^2 b$ and $v = b^3$ are the "light cone coordinates" introduced earlier, for which $V = \sqrt{uv}$. Thus, the universe inevitably expands as $T$ increases.

Moreover, if the universe expands enough so that $a^2 b^2$ becomes $\gg T$, then it enters a regime in which (for both $\eta = +1$ and $\eta = -1$) $\Psi$ oscillates rapidly, with the corresponding WKB trajectories forming (as (6.11) shows) a two-parameter family of classical Lorentzian solutions that are locally isometric to de Sitter, expanding exponentially in the cosmological time, with the ratio of the scale factors remaining constant. (One parameter of the family is the cosmological constant, and the other one is the ratio of the scale factors.) Indeed, the saddle point metric (6.7), (6.9) in this regime is itself (with our choice of "complex gauge" for it) very close to being Lorentzian at late times, and therfore close to some specific Lorentzian

---

[12]Unlike for the unnormalizable wave functions in the Taub cases, the divergence here is for small universes rather than large ones. In that sense, the argument for dismissing this saddle point is perhaps less compelling than before, because an ultraviolet cutoff on $a$ and $b$ would render the $\eta = -1$ wave function compatible with normalizability.

solution $\widehat{g}$ of the classical Einstein equations. This indicates that the major contribution to (3.1) for $a^2 b^2 \gg T$ comes from 4-geometries that at late times are close to $\widehat{g}$, and therefore are behaving essentially classically.

The behavior of our saddle point estimate for $T \to 0$ is also suggestive. In this limit, (6.11) shows, as before, that the distributional support of $\Psi$ shrinks down on $uv = 0$, describing the explosive birth of a universe at zero 3-volume $ab^2$ (although $a$ and $b$ need not vanish separately).

In summary, only the saddle point with $\eta = 1$ yields a square integrable wave function. This wave function is also compatible with unitary evolution at the level of our saddle-point estimate, and it can be construed as describing a universe that begins at 0 volume and ultimately enters a regime of classical isotropic expansion at late times.

In concluding this section, we mention that it would be straightforward to analyze also the Bianchi type I analog of the cross-cap manifold we considered in section IV, namely the 4-manifold that is the product of $T^2$ and the closed two-dimensional cross-cap. One could proceed as in section IV, quotienting $[-t_*, t_*] \times T^3$ by the map $J : (t, x, y, z) \mapsto (-t, x + \pi, y, z)$. The only saddle points are then flat, and the saddle-point action vanishes. This could be interpreted as a classical birth of a universe, if one were happy with the lack of time orientability of this metric (and the concomitant implication that the universe could die classically, in a time reversal of its birth).

## VII. SUMMARY AND DISCUSSION

In this paper we have discussed the no-boundary path integral within unimodular Einstein gravity in the Taub minisuperspace model with $S^3$ spatial topology and in a Bianchi type I minisuperspace model with $T^3$ spatial topology. The (future-truncated) 4-manifolds considered in the Taub model were the closed 4-dimensional ball and the closed 4-dimensional cross-cap, while in the Bianchi type I model we only considered the closed disk times $T^2$. In all three cases we found a saddle point (or combination of them) for which the resulting estimate to the wave function $\Psi$ is compatible with normalizability and unitary evolution. In the Bianchi type I model the estimate was rapidly oscillating for $a^2 b^2 \gg T$, and it corresponded there to a family of isotropically expanding Lorentzian universes. In the Taub model, on the other hand, the estimate did not appear to have such a WKB region with either choice for the 4-manifold.

In the Taub model, with either 4-manifold, we also found a saddle point for which the resulting estimate to the wave function is compatible with normalizability and corresponds at late times to a family of exponentially expanding Lorentzian universes. However, both these wave functions evolve nonunitarily, with probability (in the sense of $|\Psi|^2$) being injected into the configuration space at an exponentially increasing rate with respect to the unimodular time.

It should be emphasized that we did not attempt to define the path integral beyond the saddle-point approximation. In particular, we did not attempt to discuss how good our saddle-point estimate of $\Psi$ should be expected to be. It would be possible to make some estimates on the pre-exponential factor (assuming our choice of factor ordering in the Hamiltonian operator), but this would seem to contain little information beyond what

23

we already have. In particular, for the saddle points compatible with unitary evolution, the saddle-point estimate does not seem to contain enough information to single out a particular selfadjoint extension of the Hamiltonian.

When specializing the saddle points of the Taub model to those of the Friedmann model, we found, for each of the two 4-manifold topologies we considered, one saddle point for which an unnormalizable Taub wave function becomes a Friedmann wave function that is compatible with normalizability, and even compatible with unitary evolution. While these saddle points would thus have seemed highly appealing in the Friedmann model, the properties of interest disappear upon generalization to the Taub anisotropy. This should alert one to the need to understand whether our results in the Taub model and the Bianchi type I model would remain qualitatively unchanged upon the addition of still more degrees of freedom.

One avenue towards investigating this would be to include some inhomogeneous perturbations in the path integral as linearized "test" fields [44,45]. For example, if one adds to the Friedmann model a massless scalar test field, and takes for background the $\overline{B}^4$ saddle point metric with $\epsilon = 1$ and $\eta = -1$, then one does not obtain a normalizable saddle point wave function for the scalar field. In this case, therefore, the criterion of a normalizable scalar field perturbation wave function around the Friedmann saddle point metric agrees with the criterion of a normalizable Taub wave function.

The underlying question here is how one can actually recognize which saddle points, if any, yield a good approximation to the original path integral (3.1). In principle this reduces to the question whether a given saddle point metric $g$ can be reached by deforming the gravitational path integral from an originally Lorentzian "contour" to a complex contour passing through the saddle point in question. For such a deformation to be valid, the path integral would, at a minimum, have to be convergent for all intermediate values of the contour, and one might, in a preliminary formulation, reduce this to the question whether the complex metric $g$ can be reached by a curve of metrics $g(s)$ all of which admit a convergent path integral for a test scalar field (a type of perturbation that has much in common with perturbations of the metric itself). In Ref. [52] a criterion of this type was used to fix the sign of the imaginary part of a complex regulator that was there added to the metric.

Unfortunately, the issues raised in the previous two paragraphs are both clouded by the fact that each of our saddle points actually belongs to an entire family of saddle points (all with the same action $S$) whose members are related to each other via "complex diffeomorphisms", or in other words deformations of the complex path $t(s)$ that was used to parameterize the general complex solution of the Einstein equations in sections III and IV. Although $S$ itself does not depend on the choice of $t(s)$, the criteria of normalizable perturbation wave functions and convergent perturbation path integrals apparently do. We leave more systematic investigation of these questions to the future.

The results we have just summarized cannot be claimed to be realistic, of course, if only because they omit all other fields than gravity, and because they represent situations of artificially imposed symmetry. Nevertheless, the saddle points we have found, and the associated wave functions, contain enough interesting features to warrant some further comments of an interpretive nature.

In order to be convincing, an interpretation of our results would have to draw on a

more general interpretive framework for quantum gravity itself, in terms of which we could understand the significance of the saddle point metrics and wave functions we have been computing. From a histories point of view, a quantum wave function has no direct meaning at all. Rather, it is seen as an intermediary in the computation of the *quantum measure*, $\mu(C) = ||\psi_C||^2$, of the set $C$ of Lorentzian manifolds (or more general histories) whose path integral $\psi$ is.[13] It is in terms of this quantum measure $\mu(C)$ [not to be confused with the "measure-*factor*" $d\nu(g)$ that occurs in expressions like $\int d\nu(g)e^{iS(g)}$] that predictions must be made. In special situations $\mu$ reduces to a probability, and more generally it seems to represent a kind of propensity for the actual history to belong to $C$. In particular, one could probably interpret $|\psi(a, b; T)|^2$ in the present case as the probability density for the universe to find itself with the scale parameters $a$ and $b$ when the accumulated 4-volume reaches $T$. (For more details see Refs. [1,3,19,21,23].)

Now, in nonrelativistic quantum mechanics, $\psi_C$ depends parametrically on ordinary time $t$, and its squared norm $||\psi_C||^2$ must be independent of $t$ in order that $\mu(C)$ be defined consistently. This independence is guaranteed by unitarity. For gravity with $T = {}^4V$ playing the role of parameter time, an analogous unitarity would seem to be required in order that the quantum gravitational measure $\mu$ be well defined. In our minisuperspace truncation of general relativity, one can certainly impose a unitary evolution on $\psi$ if one neglects topology change, because the unimodular Hamiltonian operator is real in the Schrödinger representation, as pointed out earlier. However, it is not so clear how topology change affects the possibility of unitarity, nor is it clear what is the proper class of spacetimes over which the gravitational path integral should be taken in a cosmological setting. We believe that our results can shed some light on both these questions.

One natural idea, suggested by what we know of the big bang, is that the universe should expand from zero initial size. In a discrete setting this idea can perhaps be expressed by positing a single initial element or "origin", in a continuous setting it must translate into conditions on the topology and the metric of the spacetime manifold. Let us assume that the birth of a universe at zero size corresponds topologically to a cobordism with empty initial boundary (it is thus a special case of topology change). References [41,42,52] delineate a class of symmetric tensor fields that exist on any cobordism, that determine a well defined causal order, and that are globally smooth with Minkowskian signature almost everywhere. If we specialize them to the case of a manifold appropriate to a big bang cosmology — one without initial boundary and compact toward the past — then we arrive at a *Lorentzian* version of the so called no-boundary proposal for quantum cosmology [27–29,44], with a definite choice for the metrics to be integrated over.[14]

In a cosmology of this sort, the path integral is a sum over certain almost Lorentzian

---

[13]In addition to its technical role as "square root of the quantum measure", $\psi$ can serve as a summary of the past, useful for computing the measures $\mu(C)$ of sets $C$ defined in the future.

[14]Notice that the no-boundary condition, regarded in this manner, is a condition on the histories themselves; it need make no mention of any wave function. From a histories perspective, such a condition is more natural than any boundary condition couched in terms of the wave function $\Psi$.

metrics on manifolds without initial boundary. More fundamentally, one might expect the sum to be over an underlying discrete structure [3], a possibility that could manifest itself in the continuum in more than one way. For example, it might yield an amplitude for the universe to "bounce" (collapse and then re-expand) or to develop into a "bush-like", multi-component structure, emerging from a single "stem". In fact, some appeal to discreteness might be required just for consistency: in order that the (approximate) continuum wave function be truly square integrable (see below).

Now suppose for a moment that the only topology change that need be considered is the initial expansion we have just been considering, and that only a single macroscopic "component" of the universe comes into being thereby. If we further fix the 4-manifold topology, then we are left with a sum over almost Lorentzian 4-geometries on a given manifold, and if the quantum measure $\mu(C)$ is to be formed in the way suggested in [1], then for consistency, we need $||\psi_C||^2$ to be independent of $T$ (for $T$ sufficiently large), where in particular, $C$ can be the set of *all* 4-geometries with $V = T$. We also need, of course, that $||\psi||^2 < \infty$ in order that $\mu$ be defined at all. To make contact with the analysis in this paper, we need one further assumption, which is that the Lorentzian functional integral defining $\psi$ can be analytically continued to complex metrics and then approximated by deforming the "integration contour" to pass through a saddle point of the analytically continued integrand. This would justify (within a minisuperspace truncation) the kind of approximate wave function we have studied herein. What is more, an analysis of the conditions of validity for the contour deformation would tell us, in principle, which saddle points (if any) actually contribute to (3.2), and with what signs. In particular, it would tell us whether our conditions of a smooth complex metric on a smooth manifold without boundary correctly describe the saddle points of the analytically continued functional integral.

The two key questions then are whether $\psi$ is $L^2$ (so that $\mu$ can be defined) and if so, whether $(\psi, \psi)$ is independent of $T$ (so that the definition can be consistent). If the answer is yes, then the picture painted above is at least internally coherent. Lacking the deeper analysis that would tell us which saddle point(s) we must use, the best we can say is that, for each of the 4-manifold topologies we have studied in this paper, there is at least one saddle point consistent with these two key features at the level of approximation to which we are working. In fact, there is precisely one such saddle point for each manifold (actually a linear combination of saddle points in the Taub cross-cap case), so we get in effect a unique prediction of the no-boundary wave function in each case.

A further formal requirement, which would seem valid to the extent that post-natal topology changes can be ignored, or at least localized at the boundaries of configuration space, is that $\psi$ obey the unimodular Schrödinger equation (3.10).[15] We have seen that this demand is also met by all of our saddle points (or appropriate combinations of them in the Taub-cross-cap case), because, to the accuracy of our approximation, the Schrödinger equation reduces to the Hamilton-Jacobi equation, which all of our saddle point actions satisfied. Conversely, the requirement would not be met if we arbitrarily employed different

---

[15]See, however, the doubts raised in Ref. [69] about satisfaction of the Hamiltonian constraints in a path-integral formulation.

saddle points for different values of $T$, $a$, and $b$.

Another formal issue on which we might have hoped for guidance from our minisuperspace models is that of boundary conditions at the "edge" of configuration space. If unitarity is to obtain then the only freedom in the boundary conditions is that of a choice of selfadjoint extension for $\widehat{H}$. But that ignores the possibility of actual recollapse or, conversely, of a universe that remains "pre-geometric" for a long time $T$ and only then begins to expand. (For a causal set, the unimodular time $T$ would be identified with the total number of elements [3,70]. Even in a pre-geometric phase, therefore, $T$ retains its meaning.) A selfadjoint $\widehat{H}$ also allows for recollapse, of course, but it demands then an immediate "bounce", with no possibility of disappearance or of a (temporary or permanent) transition to a disordered, non-geometric phase. Unfortunately, it is unlikely that our saddle point estimates contain enough information to distinguish among these multiple possibilities.

Among our saddle points, there were ones exhibiting an exponential decay of $|\psi|^2$ with time in every compact region of configuration space. This occurred, in particular, for every one of our "unitary" saddle points. We chose to associate such decay with "an escape of probability toward $\infty$", but it might equally well signify an escape through the finite boundary — $i.e.$, a recollapse. Although these two alternatives do not seem to be conclusively resolvable at our level of approximation, the evidence points to a recollapse in the Taub cases and an unbounded expansion in the Bianchi type I case. If this is correct, then the evolution in the Taub cases is not actually unitary.

The wave functions $\Psi_-$ in the first Taub case and $\Psi_3$ in the second Taub case both look at $T = 0$ qualitatively like bound state wave functions localized near $uv = 0$, except that they die out exponentially as $T$ increases. The most direct interpretation of such behavior would seem to be a universe expanding to Planckian size and then rapidly shrinking to zero volume, at which point the saddle point approximation is helpless to tell us what happens next. (Beyond the possibilities mentioned earlier for what happens next, another is that probability escapes to inhomogeneous metrics — $i.e.$, the minisuperspace approximation breaks down.)[16]

In the Bianchi type I universe of section VI, the "unitary" wave function, called there $\Psi_+$, behaves very differently. For small $T$ (6.11) strongly resembles the action of a free nonrelativistic point particle (albeit with an indefinite mass tensor) released from $u = v = 0$ at $T = 0$, and we have noted in this connection that probability appears to flow toward infinity. This suggests that $\int dudv |\psi|^2$, which here diverges marginally, would continue to be infinite in a better approximation. One might attribute this divergence either to a failure of the no-boundary prescription, or to the idealization of spacetime as a continuum, for which the saddle point approximation is, disappointingly, not providing an effective cutoff. If this is correct, then to compensate, we might need to invoke discreteness explicitly, by

---

[16]If the universe really can die out entirely, then the rule given in [21] for forming the quantum measure $\mu(C)$ needs to be generalized in a way that allows paths to "exit" the configuration space at its boundary, without reappearing elsewhere. This does appear to be possible, but only if one requires both halves of the "Schwinger history" to exit at the same place and at the same value of $T$, which then serves as a premature "truncation time" for the exiting histories.

smearing $\psi$ out by hand from a delta-function to Planckian size at $T = 0$. This in turn would be expected to modify its behavior for $uv \gg T^2$.

We also found saddle points (and with many attractive features) that manifested an exponential *growth* of norm. For them, probability seems to be flowing *inward* from the finite boundary at an increasing rate. The nearest we can come to a picture of what this type of unitarity breakdown might mean would be a pre-geometric universe continually sending out branches that develop into continuum spacetimes. Such a "bush-like" universe could not really correspond to a one-configuration-space wave function at all, though, and it is not really clear whether any plausible interpretation can accommodate such saddle points.

A question often raised in connection with a quantum cosmological model is whether it predicts that the universe will, at late times, expand along an approximately classical trajectory. In effect, one is asking whether the universe, having once arrived at certain values of $T$, $a$, and $b$, can be expected to continue its expansion along some particular classical trajectory through these values. In the affirmative case, we may say that it makes a spontaneous transition to classical behavior, possibly after some initial era of non-classical expansion.[17] One can argue that such behavior is correlated with rapid, WKB-like oscillations of $\psi$ in the region in question. Another familiar criterion for classical evolution is the validity of a stationary phase approximation to the path integral, meaning that most of the contribution to the quantum measure comes from nearly classical histories, and this in turn can be related to the saddle point metric's being (up to "complex diffeomorphism") nearly Lorentzian at late times. To us, it seems unclear whether any of these features is really necessary or whether decoherence effects associated with gravitons or other matter could by themselves bring about a transition to a nearly classical universe like the one we inhabit. At any rate, it would seem helpful at least for the universe to attain a large size with high probability, and this happens (together with a spontaneous transition to classical expansion if $T \ll a^2 b^2$) in at least one of our normalizable and unitary cases, namely for the $\eta = +1$ saddle point in our Bianchi type I model.

As pointed out earlier, it is meaningful in the unimodular theory to ask what the wave function looks like at $T = 0$, the "moment of birth", and consistency would seem to require that the no-boundary prescription yield a universe which is of zero or Planckian size. This was the case for all of our "unitary" saddle points, so to that extent, the desired consistency seems to be present. (In contrast, the non-unitary but normalizable wave function $\Psi_0$ in the Taub cross-cap case seems to look at $T = 0$ like a combination of a delta-function with a "zero momentum" state spread out over all of configuration space. We recall that this $\Psi$ resulted for $\alpha < \alpha_c$ from a time-non-orientable, purely Lorentzian solution, which one suspects is not a valid saddle point at all.)

In concluding, it seems fitting to remark on the rather lifelike nature of some of our models. In contrast to non-unimodular versions of quantum cosmology, where the wave function is typically non-normalizable and otherwise very hard to interpret, we have found here many saddle points yielding $\psi$'s which are either $L^2$ or marginally so, and which evolve

---

[17]By saying that a history (Lorentzian geometry) evolves "non-classically" we merely mean that it fails to satisfy the classical Einstein equations.

28

consistently with unitarity at our degree of approximation. This seems encouraging for the account of topology change sketched above.

Among all of our saddle points, there is precisely one that is consistent with normalizability and unitarity and that spontaneously makes a transition to classical expansion. Interestingly, it belongs to spatial topology $T^3$ and not to $S^3$. In this way, we might imagine predicting something about the large scale topology of the universe, if it turned out that this distinction between $T^3$ and $S^3$ persisted in more realistic models.

## ACKNOWLEDGMENTS

# REFERENCES

[1] R. D. Sorkin, "On the Role of Time in the Sum-over-histories Framework for Gravity", paper presented to the conference on The History of Modern Gauge Theories, held Logan, Utah, July 1987, published in Int. J. Theor. Phys. **33**, 523 (1994).

[2] R. D. Sorkin, "A Modified Sum-Over-Histories for Gravity", reported in *Proceedings of the International Conference on Gravitation and Cosmology, Goa, India, 14-19 December, 1987*, edited by B. R. Iyer, A. Kembhavi, J. V. Narlikar, and C. V. Vishveshwara, see pages 184-186 in the article by D. Brill and L. Smolin: "Workshop on quantum gravity and new directions", pp. 183-191 (Cambridge University Press, Cambridge, England, 1988).

[3] R. D. Sorkin, Int. J. Theor. Phys. **36**, 2759 (1997). (gr-qc/9706002)

[4] A. Einstein, Sitzungsber. K. Preuss. Akad. Wiss. 349 (1919); English translation in *The Principle of Relativity* (Methuen, 1923; reprinted by Dover, New York, 1952), p. 189.

[5] J. J. van der Bij, H. van Dam, and Y. J. Ng, Physica (Utrecht) **116A**, 307 (1982).

[6] F. Wilczek, Phys. Rep. **104**, 111 (1984).

[7] A. Zee, in *Gauge Interactions, Proceedings of the Twentieth Orbis Scientiae*, Miami, Florida, 1983, edited by B. Kursunoglu, S. L. Mintz, and A. Perlmutter (Plenum, New York, 1984).

[8] W. Buchmüller and N. Dragon, Phys. Lett. B **207**, 292 (1988).

[9] S. Weinberg, Rev. Mod. Phys. **61**, 1 (1989).

[10] W. G. Unruh, Phys. Rev. D **40**, 1048 (1989); in *Quantum Gravity*, Proceedings of the Fourth Seminar on Quantum Gravity, Moscow, USSR, 1987, edited by M. A. Markov, V. A. Berezin, and V. P. Frolov (World Scientific, Singapore, 1988).

[11] W. G. Unruh and R. M. Wald, Phys. Rev. D **40**, 2598 (1989).

[12] M. Henneaux and C. Teitelboim, Phys. Lett. B **222**, 195 (1989).

[13] J. D. Brown and J. W. York, Phys. Rev. D **40**, 3312 (1989).

[14] L. Bombelli, in *Gravitation — A Banff Summer Institute*, edited by R. Mann and P. Wesson (World Scientific, Singapore, 1991).

[15] K. V. Kuchař, Phys. Rev. D **43**, 3332 (1991).

[16] M. A. Castagnino and F. Lombardo, Phys. Rev. D **48**, 1722 (1993).

[17] R. P. Feynman, Rev. Mod. Phys. **20**, 367 (1948).

[18] R. Griffiths, J. Stat. Phys. **36**, 219 (1984).

[19] R. D. Sorkin, in *Conceptual Problems of Quantum Gravity* (Einstein Studies II), edited by A. Ashtekar and J. Stachel (Birkhauser, Boston, 1991).

[20] C. J. Isham, J. Math. Phys. **35**, 2157 (1994). (gr-qc/9308006)

[21] R. D. Sorkin, Mod. Phys. Lett. A **9**, 3119 (1994). (gr-qc/9401003)

[22] J. B. Hartle, in *Gravitation and Quantizations, Les Houches, session LVII, 1992*, edited by B. Julia and J. Zinn-Justin (Elsevier Science B.V., Amsterdam, 1995). (gr-qc/9304006)

[23] R. D. Sorkin, in *Proceedings of the Fourth Drexel Symposium on Quantum Nonintegrability: Quantum Classical Correspondence*, held Philadelphia, September 8-11, 1994, edited by D. H. Feng and B-L. Hu (International Press, Cambridge, Massachusetts, 1997), pp. 229–251. (gr-qc/9507057)

[24] *Conceptual Problems of Quantum Gravity* (Einstein Studies II), edited by A. Ashtekar and J. Stachel (Birkhauser, Boston, 1991).

[25] K. V. Kuchař, in *Proceedings of the 4th Canadian Conference on General Relativity and Relativistic Astrophysics*, edited by G. Kunstatter, D. E. Vincent and J. G. Williams (World Scientific, Singapore, 1992).

[26] C. J. Isham, in *Integrable Systems, Quantum Groups and Quantum Field Theory*, edited by L. A. Ibart and M. A. Rodrigues (Kluwer, Dordrecht, 1992). (gr-qc/9210011)

[27] S. W. Hawking, in *Astrophysical Cosmology: Proceedings of the Study Week on Cosmology and Fundamental Physics,* edited by H. A. Brück, G. V. Coyne and M. S. Longair (Pontificiae Academiae Scientiarum Scripta Varia, Vatican City, 1982).

[28] J. B. Hartle and S. W. Hawking, Phys. Rev. D **28**, 2960 (1983).

[29] S. W. Hawking, Nucl. Phys. **B239**, 257 (1984).

[30] A. Linde, Zh. Eksp. Teor. Fiz. **87**, 369 (1984) [Sov. Phys. JETP **60**, 211 (1984)].

[31] A. Linde, Lett. Nuovo Cimento **39**, 401 (1984).

[32] A. Linde, Rep. Prog. Phys. **47**, 925 (1984).

[33] A. Linde and A. Mezhlumian, Phys. Lett. B **307**, 25 (1993). (gr-qc/9304015)

[34] A. Linde, D. Linde, and A. Mezhlumian, Phys. Rev. D **49**, 1783 (1994). (gr-qc/9306035)

[35] A. Vilenkin, Phys. Rev. D **30**, 509 (1984).

[36] A. Vilenkin, Phys. Rev. D **33**, 3560 (1986).

[37] A. Vilenkin, Phys. Rev. D **37**, 888 (1988).

[38] A. Vilenkin, Phys. Rev. D **50**, 2581 (1994). (gr-qc/9403010)

[39] A. Vilenkin, in *String Gravity and Physics at the Planck Energy Scale* (NATO Advanced Study Institute, series C: Mathematical and physical sciences, Vol. 476), edited by N. Sanchez and A. Zichichi (Kluwer Academic, Dordrecht, Netherlands, 1996). (gr-qc/9507018)

[40] A. Vilenkin, "The wave function discord", gr-qc/9804051.

[41] R. D. Sorkin, in *Proceedings of the Third Canadian Conference on General Relativity and Relativistic Astrophysics* (held Victoria, Canada, May 1989), edited by A. Coley, F. Cooperstock, and B. Tupper (World Scientific, Singapore 1990), pp. 137-163.

[42] A. Borde and R. D. Sorkin, "Causal Cobordism: Topology Change Without Causal Anomalies" (in preparation).

[43] S. W. Hawking, in *General Relativity: An Einstein Centenary Survey,* edited by S. W. Hawking and W. Israel (Cambridge University Press, Cambridge, England, 1979).

[44] J. J. Halliwell and J. B. Hartle, Phys. Rev. D **41**, 1815 (1990).

[45] J. J. Halliwell and J. Louko, Phys. Rev. D **42**, 3997 (1990).

[46] M. P. Ryan and L. C. Shepley, *Homogeneous relativistic cosmologies* (Princeton Univerity Press, Princeton, New Jersey, 1975).

[47] R. T. Jantzen, Commun. Math. Phys. **64**, 211 (1979).

[48] D. Kramer, H. Stephani, E. Herlt, and M. MacCallum, *Exact Solutions of Einstein's Field Equations,* edited by E. Schmutzer (Cambridge University Press, Cambridge, England, 1980), Section 11.3.1.

[49] R. Laflamme, in *Quantum Gravity and Cosmology: Proceedings of the XXII GIFT International Seminar on Theoretical Physics,* edited by J. Pérez-Mercader, J. Solà and E. Verdaguer (World Scientific, Singapore, 1992).

[50] M. P. Ryan and R. D. Sorkin, "On Factor Ordering and the use of Spacetime Volume

as a Time Parameter" (in preparation)

[51] M. Reed and B. Simon, *Methods of Modern Mathematical Physics* (Academic, New York, 1975), Vol. II.

[52] J. Louko and R. D. Sorkin, Class. Quantum Grav. **14**, 179 (1997). (gr-qc/9511023)

[53] L. G. Jensen, J. Louko, and P. J. Ruback, Nucl. Phys. **B351**, 662 (1991).

[54] R. D. Sorkin, Int. J. Theor. Phys. **25**, 877 (1986).

[55] J. L. Friedman and A. Higuchi, Phys. Rev. D **52**, 5687 (1995). (gr-qc/9505035)

[56] A. Daughton, J. Louko, and R. D. Sorkin, in *Proceedings of the 5th Canadian Conference on General Relativity and Relativistic Astrophysics*, edited by R. B. Mann and R. G. McLenaghan (World Scientific, Singapore, 1994). (gr-qc/9305016)

[57] J. A. Wolf, *Spaces of Constant Curvature* (McGraw-Hill, New York, 1967).

[58] G. F. R. Ellis, Gen. Relativ. Gravit. **2**, 7 (1971).

[59] J. Louko and P. J. Ruback, Class. Quantum Grav. **8**, 91 (1991).

[60] D. Giulini and J. Louko, Phys. Rev. D **46**, 4355 (1992). (hep-th/9203007)

[61] T. Koike, M. Tanimoto, and A. Hosoya, J. Math. Phys. **35**, 4855 (1994). (gr-qc/9405052)

[62] H. Waelbroeck, Commun. Math. Phys. **170**, 63 (1995). (gr-qc/9311033)

[63] M. Tanimoto, T. Koike, and A. Hosoya, J. Math. Phys. **38**, 350 (1997). (gr-qc/9604056)

[64] M. Tanimoto, T. Koike, and A. Hosoya, J. Math. Phys. **38**, 6560 (1997). (gr-qc/9705052)

[65] H. Kodama, Prog. Theor. Phys. **99**, 173 (1998). (gr-qc/9705066)

[66] R. Laflamme, in *Origin and Early History of the Universe: Proceedings of the 26th Liège International Astrophysical Colloquium (1986)*, edited by J. Demaret (Université de Liège, Institut d'Astrophysique, 1987); Ph.D. thesis (University of Cambridge, 1988).

[67] J. Louko, Ann. Phys. (N.Y.) **181**, 318 (1988); Class. Quantum Grav. **8**, 1947 (1991).

[68] J. Louko, Phys. Lett. B **202**, 201 (1988).

[69] C. Teitelboim, Phys. Rev. Lett. **50**, 705 (1983).

[70] L. Bombelli, J. Lee, D. Meyer, and R. D. Sorkin, Phys. Rev. Lett. **59**, 521 (1987).