UCD-2003-10
September 2003
revised October 2003
gr-qc/0310002

Peaks in the Hartle-Hawking Wave Function from Sums over Topologies

M. Anderson*
Department of Mathematics, SUNY at Stony Brook
Stony Brook, NY 11794, USA

S. Carlip†
Department of Physics, University of California
Davis, CA 95616, USA

J. G. Ratcliffe‡
Department of Mathematics, Vanderbilt University
Nashville, TN 37240, USA

S. Surya§
Raman Research Institute, C.V. Raman Avenue, Sadashivanagar
Bangalore 560 080, INDIA

S. T. Tschantz¶
Department of Mathematics, Vanderbilt University
Nashville, TN 37240, USA

Abstract

Recent developments in "Einstein Dehn filling" allow the construction of infinitely many Einstein manifolds that have different topologies but are geometrically close to each other. Using these results, we show that for many spatial topologies, the Hartle-Hawking wave function for a spacetime with a negative cosmological constant develops sharp peaks at certain calculable geometries. The peaks we find are all centered on spatial metrics of constant negative curvature, suggesting a new mechanism for obtaining local homogeneity in quantum cosmology.

* *email: anderson@math.sunysb.edu*
† *email: carlip@dirac.ucdavis.edu*
‡ *email: John.G.Ratcliffe@Vanderbilt.edu*
§ *email: ssurya@rri.res.in*
¶ *email: Steven.T.Tschantz@Vanderbilt.edu*

1 Introduction

Quantum cosmology is a difficult subject, not least because we do not yet have a complete quantum theory of gravity. In the absence of such a theory, cosmologists must rely on plausible, but necessarily speculative, approaches to gravity in the very early Universe. One popular approach is Hawking's Euclidean path integral [1], which describes the wave function of the Universe in terms of a "Wick rotated" gravitational path integral over Riemannian (positive definite) metrics g on a spacetime manifold M, with an action

$$I_E[g,\phi;M] = -\frac{1}{16\pi G}\int_M d^nx\sqrt{g}(R[g]-2\Lambda) - \frac{1}{8\pi G}\int_{\partial M} d^{n-1}x\sqrt{h}K + I_{\mathrm{matter}}[\phi,g]. \quad (1.1)$$

Here $R[g]$ is the scalar curvature, Λ is the cosmological constant, h is the induced metric on ∂M, and K is the trace of the intrinsic curvature of ∂M, while ϕ represents a generic collection of matter fields.

A path integral ordinarily determines a transition amplitude between an initial and a final configuration, and to specify a unique wave function one must select appropriate initial conditions. The Hartle-Hawking "no boundary" proposal [2] is that there should be *no* initial geometry—the path integral should be evaluated for compact manifolds M with only a single, connected boundary component Σ. If we specify a metric h and a set of matter fields $\phi|_\Sigma$ on Σ, the path integral

$$\Psi[h,\phi|_\Sigma;M] = \int [dg][d\phi]\,\exp\{-I_E[g,\phi;M]\} \quad (1.2)$$

can be interpreted as a wave function, giving an amplitude for the universe with spatial topology Σ to have an "initial" spatial geometry h and matter configuration $\phi|_\Sigma$. There are plausible, although not conclusive, arguments that the spacetime M in the path integral should be orientable to allow the existence of ordinary quantum field theory [3,4]; we shall generally limit our attention to such manifolds in this paper. Apart from this restriction, however, there seems to be no natural way to select any one particular topology in the path integral. Following the general quantum mechanical prescription for dealing with such alternative "paths," the Hartle-Hawking proposal is therefore to sum over all manifolds M, subject to the condition that Σ be the sole boundary component:

$$\Psi[h,\phi|_\Sigma;\Sigma] = \sum_{M:\partial M=\Sigma} \Psi[h,\phi|_\Sigma;M]. \quad (1.3)$$

To obtain interesting physics, a further restriction on M must be imposed. The integration in (1.2) is over Riemannian metrics, and it is necessary to "analytically continue" to obtain the observed Lorentzian structure of spacetime. This will be possible if the Riemannian metrics in the path integral can be joined to Lorentzian metrics to the future of Σ. Gibbons and Hartle [5] have shown that a finite action continuation across Σ exists only if the extrinsic curvature K_{ij} of Σ vanishes, that is, if Σ is totally geodesic. One should thus limit the sum to "real tunneling geometries," manifolds M and metrics g such that $\partial M = \Sigma$ is totally geodesic.

Strictly speaking, this condition does not make sense in a quantum theory: the boundary metric h_{ij} and the extrinsic curvature K_{ij} are canonically conjugate variables, and cannot be specified simultaneously. Semiclassically, this problem manifests itself in the fact that the restriction $K_{ij} = 0$ often determines h_{ij} uniquely or almost uniquely. Real tunneling geometries may thus give saddle point contributions to the wave function $\Psi[h, \phi|_\Sigma; M]$ at only a few values of h; to determine Ψ at other values, one must consider complex metrics [6], with consequent ambiguities in the choice of integration contour. Fortunately, though, it has been shown in Ref. [7] that the restriction to real tunneling geometries still reproduces the locations of the extrema of the full wave function. In particular, in the semiclassical approximation considered below, it is not unreasonable to expect that the restriction to real Riemannian metrics with totally geodesic boundaries should correctly approximate the peaks of the wave function.

Let us now restrict ourselves to four-dimensional spacetimes with no matter and a negative cosmological constant. (The generalization to $n > 4$ is straightforward, but, as we shall explain below, the generalization to positive Λ is not.) The extrema of the path integral (1.2) are then Einstein metrics $\bar{g}$, with actions

$$\bar{I}_E(M) = -\frac{\Lambda}{8\pi G} Vol_{\bar{g}}(M) \tag{1.4}$$

where $Vol_{\bar{g}}(M)$ is the volume of M evaluated with respect to the metric $\bar{g}$. In the saddle point approximation, the Hartle-Hawking wave function is thus

$$\Psi[h; \Sigma] \approx \sum_{M:\partial M=\Sigma} \Delta_M \exp\left\{-\frac{|\Lambda|}{8\pi G} Vol_{\bar{g}}(M)\right\} \tag{1.5}$$

where the prefactors Δ_M are combinations of determinants coming from gauge-fixing and from small fluctuations around the extrema [8].

It is apparent from (1.5) that contributions from "large" manifolds are exponentially suppressed, and it is natural to guess that the wave function is dominated by the smallest-volume Einstein manifold with the specified boundary data. In three spacetime dimensions, this was shown to be wrong in Refs. [9,10]: although the contributions of high-volume manifolds are exponentially suppressed, the number of such manifolds can be large enough to outweigh this suppression, and the sum over topologies can drastically change the behavior of the wave function.

In the more interesting case of four spacetime dimensions, a similar kind of topological dominance is known for one special case [11], and the partition function is also known to be dominated by complicated, high-volume manifolds [12, 13]. The purpose of this paper is to show that this behavior is much more general: for a wide class of manifolds, the sum over topologies produces sharp peaks in the Hartle-Hawking wave function that could not have been guessed by looking at any single contribution. Because of limits to our present understanding of the space of Einstein metrics, a complete, systematic understanding of this phenomenon is still lacking, but ultimately it may be possible to use this sort of analysis to make testable predictions about the geometry and topology of the Universe.

2 Einstein Dehn Filling

Our conclusions are based on an interesting new development in differential geometry, the discovery that one can use the method of Dehn filling on a cusped hyperbolic manifold to produce a large class of topologically distinct but geometrically close Einstein manifolds [14]. In this section, we shall briefly summarize this technique.

Let N be a complete, hyperbolic (that is, constant negative curvature) four-manifold of finite volume, with metric g_N. Such a manifold need not be compact: it may instead contain cusp ends, regions isometric to $\mathbb{R}^+ \times T^3$ with the AdS-like metric

$$ds^2 = \frac{3}{\Lambda}\left(dt^2 + e^{-2t} d\mathbf{x}\cdot d\mathbf{x}\right), \quad 0 < t < \infty. \tag{2.1}$$

More generally, the three-torus T^3 in such a cusp may be replaced with an arbitrary compact flat three-manifold, of which ten topological types exist, six of them orientable [16]. We shall postpone consideration of such nontoral cusps until later.

If one cuts a cusp E at constant t, the resulting manifold $N\backslash E$ acquires a new boundary with the topology of a three-torus. "Dehn filling" is the process of filling in this boundary by gluing on a solid torus $D^2 \times T^2$, whose boundary is also $\partial D^2 \times T^2 \simeq T^3$. This process is not unique: the boundary of the solid torus has a preferred simple closed curve, ∂D^2, which we may identify with any simple closed curve on the toral boundary of $N\backslash E$. If we specify a simple closed geodesic σ on this boundary,* however, and attach the solid torus by a diffeomorphism that identifies σ and ∂D^2, this is sufficient to determine the topology of the resulting manifold. The geodesic σ, in turn, is fixed by three integers $\mathbf{n} = (n_1, n_2, n_3)$: if one chooses three independent "circumferences" v_i of T^3 that intersect only at a single base point, then σ is homotopic to a linear combination $\mathbf{n}\cdot\mathbf{v}$, where $\mathbf{n}$ is a triple of relatively prime integers. Note that for $\mathbf{n}$ large, the length of the geodesic σ becomes large.

If N has only a single cusp, this process produces an infinite number of topologically distinct compact manifolds $N_{\mathbf{n}}$, labeled by triples of relatively prime integers $\mathbf{n}$. If N has more than one cusp, one can perform the same procedure on each, obtaining manifolds $N_{\{\mathbf{n}\}}$ labeled by a collection $\{\mathbf{n}\}$ of such triples. So far, the construction is purely topological. The main result of Ref. [14] is that for suitable $\{\mathbf{n}\}$, the filled manifold admits an Einstein metric $g_{\{\mathbf{n}\}}$, and that this metric is geometrically close to the original metric g_N.

Suppose N has k cusps $\{E_i\}$. For a fixed Dehn filling, the construction of the metric $g_{\{\mathbf{n}\}}$ starts with k copies of the toral AdS black hole metric [15] on the solid torus $D^2 \times T^2$, each appropriately "twisted" to match the desired filling. These metrics can be attached continuously to the hyperbolic metric on $N\backslash E_i$ at each cusp. The result is not quite smooth at the "seams," but it is shown in Ref. [14] that for suitable $\{\mathbf{n}\}$, it can be smoothed to give an Einstein metric.

More precisely, the topological Dehn filling is determined by k geodesics $\{\sigma_1, \ldots, \sigma_k\}$, or equivalently, as described above, k triples of relatively prime integers $\{\mathbf{n}_1, \ldots, \mathbf{n}_k\}$. Given such a filling, let R_{min} be the length of the shortest of the geodesics σ_i, in units $\Lambda = 3$, and

*Any simple closed curve on this T^3 boundary is homotopic to a geodesic.

let R_{max} be the length of the longest. Then the filled manifold $N_{\{\mathbf{n}\}}$ admits an Einstein metric provided that R_{min} is sufficiently large and

$$R_{max} \leq e^{c_0 {R_{min}}^3}, \tag{2.2}$$

where c_0 is a fixed constant. Note that all of the cusps must be filled—heuristically, an unfilled cusp is equivalent to a geodesic with $R_{max} = \infty$, and (2.2) has no solutions. On the other hand, by varying all of the fillings, one may easily obtain infinitely many manifolds satisfying (2.2), and thus infinitely many compact manifolds with Einstein metrics.

Ref. [14] does not construct the Einstein metric $g_{\{\mathbf{n}\}}$ explicitly, but proves that it is close to the original metric g_N, in several related senses. First observe that N embeds as a domain in $N_{\{\mathbf{n}\}}$; in fact, $N_{\{\mathbf{n}\}}$ with the core torus $\{0\}\times T^2$ removed from each filled cusp is diffeomorphic to N. As $\{\mathbf{n}\}$ grows, each core torus $\{0\} \times T^2$ becomes small in the metric $g_{\{\mathbf{n}\}}$, in that its diameter and area go to zero as $\{\mathbf{n}\} \to \infty$. In addition, the core tori recede progressively further down the cusps, so that the metrics $(N_{\{\mathbf{n}\}}, g_{\{\mathbf{n}\}})$ converge smoothly to the original metric g_N on any fixed compact set in N. One consequence of this behavior is that the volumes converge: there are constants $\delta_{\{\mathbf{n}\}} > 0$ such that

$$Vol_{g_{\{\mathbf{n}\}}}(N_{\{\mathbf{n}\}}) = Vol_{g_N}(N) - \delta_{\{\mathbf{n}\}} \quad \text{with } \delta_{\{\mathbf{n}\}} \to 0 \text{ as } \{\mathbf{n}\} \to \infty. \tag{2.3}$$

(We refer the reader to chapter 10 of Ref. [17] for a general discussion of norms used to define closeness in Riemannian geometry.)

As noted above, a cusp of a hyperbolic four-manifold need not be toral, but can have the topology $\mathbb{R}^+ \times F$, where F is any flat three-manifold. Any such F can be obtained as a quotient of $\mathbb{R}^3$ by a group of isometries. In Ref. [14] it is shown that the techniques described here for toral ends can be applied to four of the ten possible flat geometries of a cusp end, the geometries G_1, G_2, B_1, and B_2 of Ref. [16], or equivalently A, B, G, and H of Ref. [18]. Of these, only cusps of type A and B are orientable.

3 The Partition Function

One immediate consequence of this construction is that the semiclassical gravitational partition function Z—the path integral (1.1) for *closed* four-manifolds—diverges when $\Lambda < 0$. Indeed, given any hyperbolic "seed manifold" N with the appropriate cusp types, Einstein Dehn filling produces an infinite number of compact Einstein manifolds, all with volumes less than $Vol_{g_N}(N)$. Since the semiclassical contribution of each such manifold goes as

$$\exp\left\{-\frac{|\Lambda|}{8\pi G}\, Vol_{g_{\{\mathbf{n}\}}}(N_{\{\mathbf{n}\}})\right\}, \tag{3.1}$$

just as in (1.5), the sum over the $N_{\{\mathbf{n}\}}$ is infinite, at least at this order of approximation. The possibility remains, of course, that higher-order corrections will strongly suppress all but a few of the contributions to the sum. There seems to be no reason to expect that, however, particularly since all of the Einstein metrics $(N_{\{\mathbf{n}\}}, g_{\{\mathbf{n}\}})$ are strictly stable—that is, the second variation of the action has no zero or negative eigenmodes, which could otherwise lead to bad behavior or introduce phases in the prefactors Δ in (1.5).

The fact that the partition function diverges is not new. It was shown in Refs. [12, 13] that even the contributions of strictly hyperbolic manifolds lead to a divergent sum, since the number of manifolds with volume greater than V grows faster than exponentially with V [19]. That sum, however, is similar to a divergent sum occurring in string theory [20], and one might hope that a finite resummation similar to that coming from matrix models in string theory [21] might be possible. The divergence found here, on the other hand, has no obvious cure, although we shall discuss a few possible loopholes in section 6.

The physical meaning of the partition function, however, is not entirely clear. If one restricts oneself to manifolds of the form $S^1 \times \Sigma$, or more generally to circle bundles over a three-manifold, one can interpret the circumference of the S^1 factor as a local inverse temperature and Z as a thermal partition function. The manifolds we consider here are not of that form, however, and it is not obvious that a divergent partition function is a terrible thing. It is therefore important to look at quantities with clearer physical interpretations, such as the Hartle-Hawking wave function (1.3).

4 Totally Geodesic Boundaries

The discussion in Ref. [14] was restricted to manifolds with no boundaries apart from cusps. To understand the Hartle-Hawking wave function, we must instead consider manifolds with a totally geodesic boundary Σ. Fortunately, the generalization is fairly straightforward.

Let M be a finite-volume hyperbolic manifold with cusps that has a connected totally geodesic boundary Σ, with induced metric h_M on Σ. The double $N = 2M$ is formed by attaching two copies of M along Σ. This double can be characterized by the existence of a reflection isometry θ that permutes the two copies of M, with a fixed point set Σ, the "mirror" of the reflection. The existence of such an isometry with a two-sided, separating fixed point set ensures that $2M$ can be cut along this mirror to obtain the original manifold M with totally geodesic boundary Σ. If M is orientable, θ will be orientation-reversing.

The double $2M$ meets the conditions of section 2, and we can perform an identical Dehn filling on each pair of cusps related by the reflection symmetry. The smoothing procedure of Ref. [14] respects this symmetry, and thus produces a manifold $2M_{\{\mathbf{n}\}}$ that itself has a reflection symmetry. Cutting $2M_{\{\mathbf{n}\}}$ along the fixed point set of this symmetry, we obtain a manifold $M_{\{\mathbf{n}\}}$ with an Einstein metric $g_{\{\mathbf{n}\}}$ and a totally geodesic boundary $\Sigma_{\{\mathbf{n}\}}$. Since the Dehn filling affects the topology only at the cusp ends, the boundary $\Sigma_{\{\mathbf{n}\}}$ is diffeomorphic to Σ. Furthermore, the convergence conditions described in section 2 guarantee that the induced boundary metrics $h_{\{\mathbf{n}\}}$ converge to h_M, and that the volume $Vol(M_{\{\mathbf{n}\}})$ converges to $Vol(M)$ as in (2.3).

More care is needed if the double $2M$ has cusps that intersect the mirror hypersurface Σ. Let $\{E\}$ denote the cusps that intersect Σ, and let $\{\widehat{E}\}$ denote the cusps that do not. We can then look for a fixed Einstein Dehn filling of the $\{E\}$ that respects the mirror symmetry and proceed as above. If such a symmetric filling exists,[†] it will change the topology of Σ; for a fixed symmetric filling of the $\{E\}$ labeled by integers $\{\mathbf{m}\}$, we will obtain a new, compact boundary $\Sigma^{\{\mathbf{m}\}}$.

By now varying the fillings of the remaining cusps $\{\widehat{E}\}$, we can construct a large number of Einstein four-manifolds with induced boundary metrics that are close to a fixed $h_{M^{\{\mathbf{m}\}}}$, and with volumes close to a fixed $Vol(M^{\{\mathbf{m}\}})$. Because of the requirement (2.2), the number of such manifolds is no longer infinite: the fixed filling of the $\{E\}$ determines a minimum length R_{min}, and (2.2) then restricts the fillings of the remaining cusps $\{\widehat{E}\}$. The number of allowed fillings grows very rapidly with R_{min}, however; for a double $2M$ with $2k$ cusps that do not intersect the mirror hypersurface, the number of allowed fillings goes as $\exp\{3kc_0{R_{min}}^3\}$. Since R_{min} is typically very large, this number quickly becomes enormous.

The next question we must address is the existence of "seed manifolds" M of the type described here. Unfortunately, at least for now we cannot completely classify such manifolds. There are, however, two constructions that show that such manifolds exist. The first and most explicit of these is based on Ref. [18], in which 1171 complete hyperbolic four-manifolds of minimum volume are constructed by gluing together the sides of a regular ideal 24-cell. The manifolds of this class all contain totally geodesic hypersurfaces, and as described in [22], cutting along such a hypersurface will sometimes result in a complete hyperbolic manifold with a single totally geodesic boundary. For our construction to work, we must further check that the resulting manifold has the proper cusp types to allow the Einstein Dehn filling of Ref. [14].

By examining the manifolds of Ref. [18], we have found a large number of such manifolds. Most of these, however, have cusps of type B that intersect the mirror Σ. As we explain in Appendix B, such cusps do not admit Einstein Dehn fillings that respect the reflection symmetry, and hence cannot be filled to provide "seed manifolds" of the type we need. There are, however, five examples—described in detail in Appendix A—for which only cusps of type A intersect the mirror Σ, giving us precisely the structure we need.

Consider, for example, manifold N_{24} of Appendix A. This manifold is the orientable double cover of manifold 24 of [18], and has Euler characteristic two and a volume of $24\pi^2/\Lambda^2$. It has seven cusps, all of type A, three of which intersect the mirror. The mirror Σ in N_{24}, which will become the "spatial" boundary in the Hartle-Hawking wave function, is a hyperbolic three-manifold homeomorphic to a double cover of the complement of the four-component link 8^4_2. It has six cusps, and is therefore not compact. We can, however, make Σ compact by performing an Einstein Dehn filling of the three cusps of N_{24} that intersect Σ. As we discuss in Appendix B, it is possible to do this in a manner that respects the reflection symmetry of N_{24}, ensuring that the filled manifold is still a double. In fact, such a filling induces an ordinary (topological) Dehn filling of Σ. Conversely, any sufficiently large Dehn filling of Σ that has the symmetries described in Appendix B can be lifted to an Einstein Dehn filling of N_{24}. By choosing a large enough fixed filling of the cusps of Σ, one can make R_{min} arbitrarily large, and can thus find arbitrarily many fillings

[†]See Appendix B for a detailed description of some examples.

of the remaining cusps that satisfy (2.2).

The remaining manifolds described in Appendix A are different in detail, but qualitatively similar. Each has Euler characteristic two and a volume of $24\pi^2/\Lambda^2$. Each is a double, with a mirror Σ that has cusps; in fact, the topology of this mirror is identical in all five examples. In each case, the cusps of Σ can be filled by Einstein Dehn fillings that respect the reflection symmetry. By choosing different Dehn fillings of the cusps that intersect the mirror, as described in Appendix B, we can thus obtain an infinite collection of boundary topologies for which our construction applies.

Another set of potential "seed manifolds" are due to Long and Reid [23], who give algebraic descriptions of a variety of hyperbolic four-manifolds with totally geodesic boundaries. The manifolds discussed in this reference are compact, but the results may be extended to manifolds with cusps [24]. Further investigation of these examples would be of interest.

Long and Reid have also found a large set of three-manifolds for which our construction *cannot* work. In Ref. [25], they show that a three-manifold Σ can occur as a totally geodesic boundary of an orientable hyperbolic four-manifold only if the eta invariant $\eta(\Sigma)$ is an integer. As it is written, theorem 1.1 of that paper applies only to boundaries of compact orientable four-manifolds. But it follows from the proof of theorem 1.3 that for an orientable hyperbolic four-manifold M with totally geodesic boundary Σ and cusps with topologies $\mathbb{R}^+ \times F_i$,

$$\eta(\Sigma) + \sum_i \eta(F_i) \in \mathbb{Z}. \tag{4.1}$$

Long and Reid further show that $\eta(F_i)$ is an integer for cusps of type A and B, the two for which Einstein Dehn fillings are possible. Hence for our construction to apply, we must require that $\eta(\Sigma) \in \mathbb{Z}$. This is a strong restriction: of some $11,000$ small-volume hyperbolic three-manifolds in the SnapPea census [26], for example, only 41 have integral eta invariants, so only these 41 are even candidates for our construction. (It is not clear that this result generalizes in any easy way if one includes nonorientable four-manifolds, but as argued in section 1, there are good reasons to expect only orientable four-manifolds to contribute to the Hartle-Hawking wave function.)

5 Peaks in the Hartle-Hawking Wave Function

We now turn to the implications of these results for the Hartle-Hawking wave function. Consider a universe described, at least in the low energy limit, by general relativity with a negative cosmological constant, as required, for example, by supergravity. Let Σ be a three-manifold of the sort discussed in the preceding section—that is, a manifold that can occur as a totally geodesic boundary of a hyperbolic manifold M with cusps of type A or B (or G or H if we permit nonorientable manifolds). Denote by h_M the metric induced on Σ by the hyperbolic metric g_M of M.

The manifold M will occur as a saddle point in the Hartle-Hawking path integral (1.2), giving a contribution, to lowest order, of

$$\exp\left\{-\frac{|\Lambda|}{8\pi G}\,Vol_{g_M}(M)\right\} \tag{5.1}$$

at the boundary value h_M. But it follows from the discussion above that there are a very large number of Einstein manifolds $M_{\{\mathbf{n}\}}$, also with boundary Σ, whose metrics have boundary values $h_{\{\mathbf{n}\}}$ that lie close h_M, in the sense of "closeness" discussed in section 2. By (2.3), the volumes of these filled manifolds are all less than $Vol_{g_M}(M)$, so at this order their contribution to the wave function is actually greater than that of M.

A large number of manifolds thus give nearly identical contributions to the sum (1.5) within a small neighborhood of the boundary metric h_M. Since the standard Hartle-Hawking prescription (1.3) tells us to simply add such contributions, the Hartle-Hawking wave function for Σ, at this order of approximation, is therefore sharply peaked around h_M. Note that such a peak occurs for any value of $\Lambda < 0$ and for any finite volume of the "seed manifold" M, and will appear for every spatial topology Σ that occurs as a boundary of an appropriate four-manifold.

If we could find a hyperbolic "seed manifold" whose double had no cusps on the mirror Σ, this peak would, in fact, be infinite. As discussed in the preceding section, however, the "seed manifolds" that we know how to construct explicitly all have boundaries Σ with cusps. To apply the construction of [14], we must fill these cusps, as discussed above and in Appendix B. If a fixed filling of these cusps respects the reflection symmetry, it gives a new "seed manifold" $M^{\{\mathbf{m}\}}$ with boundary $\Sigma^{\{\mathbf{m}\}}$ related to Σ by a topological Dehn filling. In the examples in Appendix A, such a partially filled four-manifold will still have a set of cusps that do not intersect Σ, and summing over the allowed fillings of these remaining cusps will give us a sharp peak in the Hartle-Hawking wave function on $\Sigma^{\{\mathbf{m}\}}$.

Note that while the $M^{\{\mathbf{m}\}}$ are not exactly hyperbolic, they are "nearly" hyperbolic in the sense of section 2. In particular, for large $\{\mathbf{m}\}$ the induced metric on $\Sigma^{\{\mathbf{m}\}}$ will be very close to the hyperbolic metric on Σ except in regions very far down the filled cusps. Moreover, the more nearly hyperbolic $M^{\{\mathbf{m}\}}$, the larger the geodesic length R_{min} in (2.2) will be, and the more Einstein Dehn fillings will be allowed. "Nearly hyperbolic" boundaries thus have very highly peaked wave functions.

Let us now point out an intriguing feature of this construction. The peaks that we have found are centered around metrics h_M (or $h_{M^{\{\mathbf{m}\}}}$) obtained from hyperbolic or very nearly hyperbolic "seed manifolds." Since the boundary Σ is always a hypersurface of vanishing extrinsic curvature, such an h_M is itself a hyperbolic or very nearly hyperbolic metric on Σ. This means that the peaks occur at or very near to locally homogeneous geometries of Σ. Of course, we do not know that these are the *only* peaks of the Hartle-Hawking wave function. Nevertheless, the appearance of a new, noninflationary mechanism for explaining spatial homogeneity of the early universe should be of some interest.

6 Conclusions and Caveats

The calculations we have described so far are, of course, only lowest order approximations to the full path integral (1.2)–(1.3). To the next order, we should take into account the prefactors $\Delta_{M_{\{\mathbf{n}\}}}$ of (1.5). The convergence of metrics is not, unfortunately, enough to guarantee the convergence of these determinants, which depend nonlocally on the metric throughout $M_{\{\mathbf{n}\}}$. In three spacetime dimensions, the prefactors combine to form a topological invariant, the Ray-Singer torsion [27], which does not converge under a similar Dehn filling procedure [9]. In that case, however, the prefactors densely fill a finite nonnegative interval, a behavior sufficient to show that the first-order infinite peak survives.

For four spacetime dimensions, no corresponding result is known, and it remains possible, although unlikely, that the prefactors become systematically small enough for large $\{\mathbf{n}\}$ to suppress the peaks we have found. As noted in section 3, the strict stability of our filled manifolds implies the absence of negative eigenvalues that could otherwise lead to nontrivial phases [1], eliminating one possible source of suppression. We can actually say a bit more about one-loop corrections. In an effective action formalism [28], such corrections lead to new terms of the form

$$\int_{M_{\{\mathbf{n}\}}} d^4x\sqrt{g}C_{abcd}C^{abcd} \tag{6.1}$$

in the effective action, where C_{abcd} is the Weyl tensor. But such terms converge to zero for large $\{\mathbf{n}\}$ [14], and thus do not affect our conclusions. Indeed, integrals of higher order polynomials in the Weyl tensor, which can appear at higher orders of perturbation theory, likewise converge to zero. Quantum corrections can also induce terms in the action (1.1) proportional to the Euler characteristic χ and the Hirzebruch signature τ [1]. The Euler characteristic is left invariant by Dehn filling, however, and the signature can be written in terms of integrals of the Weyl curvature that again go to zero for large fillings. Hence neither of these topological terms will affect our conclusions. While this does not guarantee the absence of higher-order nonlocal effects that could still suppress the peaks we have found, it makes our results considerably more robust.

One more loophole remains. We have assumed that the sum (1.2) over topologies should include every manifold with a boundary Σ, and that all occur with equal weights. While this is a plausible assumption, it may not be correct. Sorkin et al. [29, 30] have suggested that considerations of causal structure may restrict the topologies occurring in the sum over histories, and that such a restriction may eliminate problems of infinite particle production [31] in topology-changing amplitudes. The restriction to "causally continuous" spacetimes requires, in particular, that the first Betti number b_1 of the spacetime relative to the boundary vanish. This condition would eliminate many of the examples considered here; we are currently investigating the question of whether an infinite number would remain. Observe also that the manifolds in our sum over topologies with large Dehn fillings differ significantly from each other only in extremely small regions "far down the cusps." If the correct quantum theory of gravity contains a cut-off in volume or diameter, say at the Planck scale, then perhaps these manifolds should be considered equivalent and not counted separately. Until we have a much better established quantum theory of gravity, these issues are likely to remain unresolved.

Assuming that the manifolds we have considered here really do appear in the sum over topologies, an analysis of this sort may eventually have predictive value, telling us what spatial geometries of the Universe are most probable. For now, though, several sizable obstacles remain.

First, of course, we do not have any particular reason to believe that the cosmological constant is negative. A negative Λ in the very early Universe is not excluded by observation—the apparently positive cosmological constant now may be a consequence of later dynamics—but neither is it supported. The analysis of this paper relies on particular properties of manifolds with negative curvature, and does not extend to those with positive curvature. In three spacetime dimensions, the case of positive Λ can be analyzed, and is qualitatively very different [10]. Ideally, one would like to do the same in four spacetime dimensions, but to the best of our knowledge the appropriate mathematical tools are not yet available.

Second, we do not yet have good control over which manifolds can appear in the constructions of this paper. In particular, while section 4 provides examples of three-manifolds Σ that can occur as boundaries in an Einstein Dehn filling, we have no reason to expect these examples to constitute a complete list. We clearly need a much better grasp of which three-manifolds can occur as totally geodesic boundaries of complete hyperbolic four-manifolds.

Third, while the method of Einstein Dehn fillings gives some peaks of the Hartle-Hawking wave function, it may not give all of the peaks. In general, the Hartle-Hawking prescription tells us to start with extrema of the action (1.1). But while such extrema are always Einstein manifolds, they need not be hyperbolic or "nearly" hyperbolic manifolds of the type we have discussed in this paper. For example, it is tempting to conclude from the results of Long and Reid described at the end of section 4 that a three-manifold Σ with $\eta(\Sigma) \notin \mathbb{Z}$ cannot have a peaked wave function. But this may not be true: while such a manifold cannot occur as a totally geodesic boundary of a hyperbolic spacetime, we do not know whether it can occur as a totally geodesic boundary of a nonhyperbolic Einstein spacetime, or, if so, how many four-manifold topologies can contribute. Ultimately, we need a much more thorough understanding of the space of Einstein metrics.

Despite these limitations, though, we believe that this work has some importance. We have demonstrated that the sum over topologies can have a clear and dramatic effect on the Hartle-Hawking wave function in the physically realistic case of four spacetime dimensions, and that it leads to at least some sharp peaks at spatially homogeneous geometries. And in the unlikely but still possible event that observations of the cosmic microwave background [32] or other of cosmic structure [33] demonstrate that our Universe has one of the spatial topologies described in section 4, we have a genuine prediction of a highly probable initial geometry.

Acknowledgments

We would like to thank Gary Gibbons and Alan Reid for valuable help. This work was supported in part by U.S. Department of Energy grant DE-FG03-91ER40674 and NSF grant DMS 0305865.

Appendix A Seed Manifolds

In this appendix we describe the "seed manifolds" of section 4. We start with the complete minimum-volume hyperbolic manifolds of Ref. [18], which were obtained by gluing together the sides of a single regular ideal 24-cell. As described in [22], each of these manifolds contains a collection of totally geodesic hypersurfaces, or "cross sections," inherited from the hyperplanes of symmetry of the 24-cell. If N is a nonorientable manifold chosen from this list with a cross section S, one can cut N along S to obtain a manifold M with boundary Σ. Ref. [22] gives the criteria for M to be a "seed manifold" of the type we need: S must be orientable, one-sided, and have an orientable complement. For the Einstein Dehn filling of Ref. [14] to apply, we must also demand that the cusps of the double $2M$ all be of type A or B. Moreover, as explained in Appendix B, any cusps of $2M$ that intersect the mirror Σ must be of type A.

In table 1 we give five examples that meet these criteria, together with the homology and cusp link types of their doubles. The column headed N gives the number from [18] of the nonorientable four-manifold N that we take as a starting point. The column headed by SP gives the side-pairing code for gluing the ideal 24-cell to obtain N, as explained in [18]. The pairings are taken so that the coordinate hyperplane cross section $x_4 = 0$ is a one-sided, nonseparating, totally geodesic hypersurface whose complement is orientable. Cutting open N along this cross section yields an orientable hyperbolic four-manifold M with a connected totally geodesic boundary Σ. The boundary Σ double covers the cross section of the starting four-manifold. The reflection in the $x_4 = 0$ cross section is a symmetry of N, and induces a symmetry of M that restricts to the covering transformation of Σ. This implies that the double of M along Σ is the orientable double cover of N.

N	SP	H_1	H_2	H_3	LT
24	9FC129	$\mathbb{Z}^7 \oplus \mathbb{Z}_2$	$\mathbb{Z}^{14}$	$\mathbb{Z}^6$	AAAAAAA
55	9FC12A	$\mathbb{Z}^5 \oplus \mathbb{Z}_2^3$	$\mathbb{Z}^{12}$	$\mathbb{Z}^6$	AAAAABB
237	9FC179	$\mathbb{Z}^3 \oplus \mathbb{Z}_2^5$	$\mathbb{Z}^{10}$	$\mathbb{Z}^6$	AAABBBB
1091	9FC124	$\mathbb{Z}^9$	$\mathbb{Z}^{18}$	$\mathbb{Z}^8$	AAAAAAAAA
1113	9FC127	$\mathbb{Z}^3 \oplus \mathbb{Z}_2^6$	$\mathbb{Z}^{12}$	$\mathbb{Z}^8$	AAABBBBBB

Table 1: Five double 24-cell manifolds.

In each of these examples, the coordinate hyperplane cross section is the hyperbolic three-manifold M^3_{10} described in [18] with side-pairing code 174 for the gluing of the sides of an ideal rhombic dodecahedron. The ideal rhombic dodecahedron is the coordinate hyperplane cross section of the ideal 24-cell fundamental domain. The pairings of sides of the ideal 24-cell for our starting four-manifold N that give rise to the side-pairing of the rhombic dodecahedron cross section all reflect in the direction perpendicular to the cross section, thus joining the top and bottom sides of the cross section in the 24-cell. This implies that the boundary Σ of M can be obtained by gluing two copies of the ideal rhombic dodecahedron fundamental domain of M^3_{10} so that each side glues to a side of the other dodecahedron by the gluing that applies for M^3_{10}. See Figure 5 in [18] for an illustration of the ideal rhombic dodecahedron fundamental domain of M^3_{10}.

The mirror Σ has six cusps, each of whose link is a torus. The first and second homology groups of Σ are $H_1(\Sigma) = \mathbb{Z}^6$ and $H_2(\Sigma) = \mathbb{Z}^5$. In units such that $\Lambda = -2$, the volume of Σ is $16L(2) = 14.6554494\ldots$, where $L(s)$ is the Dirichlet L-function

$$L(s) = 1 - \frac{1}{3^s} + \frac{1}{5^s} - \frac{1}{7^s} + \cdots . \tag{A.1}$$

Doubling the four-manifold M by reflecting across the mirror Σ gives the orientable double cover of our starting nonorientable four-manifold N. For a more explicit description, we can read off a side-pairing of two ideal 24-cells from the side-pairing code of N: we use the original side-pairing, but glue a given side to the appropriate side of the same 24-cell if the side-pairing code is at most 7, and to the corresponding side of the other 24-cell if the side-pairing code is 8 or more in hexadecimal notation.

The columns of Table 1 headed H_i list the ith homology groups of the manifolds. The column headed LT lists the link types of the cusps of the manifolds. Here A and B represent the first two closed orientable flat three-manifolds in the order given by Wolf [16]: A is the three-torus and B is the half-twisted three-torus. See Figure 1 of [22] for an illustration of a fundamental domain for the flat three-manifold B.

The manifolds in the table have been selected so that only link type A appears for the three cusps intersecting the mirror (the first three listed), and only A and B link types appear for the remaining cusps off the mirror. Each cusp that intersects the mirror contains two cusps of the mirror. Each link of a cusp that intersects the mirror is itself a double of a flat three-manifold with a totally geodesic boundary. By Theorem 1 of [22], the link of a cusp that intersects the mirror is the double of a right cylinder $T^2 \times [0, 1]$ whose boundary components are the links of the cusps of the mirror that are contained in this cusp.

The five hyperbolic four-manifolds in Table 1 have Euler characteristic two, and so have volume $8\pi^2/3$ by the Gauss-Bonnet theorem, in units $\Lambda = -3$. These manifolds have the smallest possible volume for a hyperbolic four-manifold that is a double of a hyperbolic four-manifold with a totally geodesic boundary.

Appendix B Filling Cusps on the Mirror

The "seed manifolds" of Appendix A all have cusps that intersect the mirror hypersurface, which therefore itself has cusps. The construction of Ref. [14], on the other hand, requires that all cusps, including those that intersect the mirror, be filled. Here we describe the conditions for such a filling to respect the mirror symmetry.

Start with a "seed manifold" N with a mirror Σ fixed by a reflection isometry θ. Let $E \simeq \mathbb{R}^+ \times F$ be a cusp of N that intersects Σ. Such a cusp must intersect Σ symmetrically, so θ is also a reflection isometry of F. The means that F is itself the double of a manifold with boundary, or a "flat gravitation instanton" in the sense of [22]. Such manifolds have been classified in [22]: if F is type A it must be a flat three-torus obtained by doubling a flat cylinder $T^2 \times [0,1]$, while if it is type B it must be a flat closed three-manifold obtained by doubling a twisted I-bundle over a flat Klein bottle.

Suppose first that F is of type A. We can write $F \simeq T^2 \times S^1$, where the last S^1 is to be understood as the double of $[0,1]$. The fixed points of the action of θ on this S^1 are the endpoints $\{0\}$ and $\{1\}$ of the interval, so F intersects Σ twice, at $T^2 \times \{0\}$ and $T^2 \times \{1\}$. Hence a cusp of type A in N that intersects Σ does so at two cusps of Σ.

Now, recall that a Dehn filling of E requires that we identify the boundary of a solid torus with F. Let us think of the solid torus as $(D^2 \times S^1) \times S^1$. We can then fill the cusp E in a way that respects the reflection symmetry by gluing $\partial D^2 \times S^1$ to the T^2 factor in F, with ∂D^2 identified with some geodesic in T^2, and gluing the second S^1 factor in the solid torus to the S^1 in F by the identity. Such a filling is determined by two relatively prime integers (m_1, m_2) that label the geodesic in T^2. Note that this filling of a cusp of the four-manifold N is automatically also a Dehn filling of the two cusps $\mathbb{R}^+ \times T^2 \times \{0\}$ and $\mathbb{R}^+ \times T^2 \times \{1\}$ of the mirror Σ, with surgery invariants (m_1, m_2) and $(m_1, -m_2)$ that are identical up to a sign change coming from a difference of orientation. Conversely, any such symmetric filling of these two cusps of Σ determines a filling of E.

Now suppose instead that F is of type B. In Ref. [14], an Einstein Dehn filling of such a cusp was defined by first going to a covering space, a flat manifold of type A, and then considering fillings that respect the covering projection. For a cusp that intersects Σ, however, we must also demand that the filling respect the reflection symmetry. These two requirements are nearly incompatible; it can be shown that only two fillings are allowed, and there is no reason to expect these to admit Einstein metrics. We must therefore discard manifolds for which cusps of type B intersect the mirror.

Finally, we return to the "seed manifold" N, and now assume that only cusps of type A intersect Σ. We divide the cusps of N into two sets: n cusps $\{E\}$, all of type A, that intersect Σ and $\hat{n}$ cusps $\{\widehat{E}\}$, of type A or B, that do not intersect Σ. Pick a *fixed* Einstein Dehn filling of the $\{E\}$ that respects the reflection symmetry of N. Such a filling is determined by n pairs of relatively prime integers $\{\mathbf{m}\}$, and automatically fills the $2n$ cusps of Σ.

The resulting manifold $N^{\{\mathbf{m}\}}$ has a reflection isometry, with a mirror $\Sigma^{\{\mathbf{m}\}}$ that is now compact. Furthermore, $N^{\{\mathbf{m}\}}$ still has $\hat{n}$ cusps that do not intersect $\Sigma^{\{\mathbf{m}\}}$. The allowed fillings of these remaining cusps are restricted by (2.2), where for most fillings R_{min} is determined by the $\{\mathbf{m}\}$. For a "typical" example, $\{\mathbf{m}\}$ is large, $R_{min} \gg 1$, and the number of such fillings increases exponentially with ${R_{min}}^3$. Thus, as in section 5, summing over Einstein Dehn fillings of these remaining cusps will lead to a sharp peak in the Hartle-Hawking wave function $\Psi[h; \Sigma^{\{\mathbf{m}\}}]$.

This peak is centered on the metric $h^{\{\mathbf{m}\}}$ induced from the metric of the partially filled manifold $N^{\{\mathbf{m}\}}$, which is not quite hyperbolic. But for "most" manifolds obtained in this manner—that is, for large $\{\mathbf{m}\}$—$h^{\{\mathbf{m}\}}$ is arbitrarily close to a hyperbolic metric on compact sets, that is, not too far down the cusps. In particular, the regions with large deviations from homogeneity become arbitrarily small: for any fixed ϵ, the volume of the region of Σ in which $h^{\{\mathbf{m}\}}$ differs from a hyperbolic metric by more than ϵ goes to zero as $\{\mathbf{m}\}$ goes to infinity.

References

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

[2] J. B. Hartle and S. W. Hawking, Phys. Rev. D28 (1983) 2960.

[3] G. W. Gibbons, Nucl. Phys. B271 (1986) 497.

[4] J. L. Friedman and A. Higuchi, Phys. Rev. D52 (1995) 5687.

[5] G. W. Gibbons and J. B. Hartle, Phys. Rev. D42 (1990) 2458.

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

[7] S. Carlip, Class. Quant. Grav. 10 (1993) 1057.

[8] P. O. Mazur and E. Mottola, Nucl. Phys. B341 (1990) 187.

[9] S. Carlip, Phys. Rev. D46 (1992) 4387.

[10] S. Carlip, Class. Quant. Grav. 10 (1993) 207.

[11] J. G. Ratcliffe and S. T. Tschantz , Class. Quant. Grav. 17 (2000) 2999.

[12] S. Carlip, Phys. Rev. Lett. 79 (1997) 4071.

[13] S. Carlip, Class. Quant. Grav. 15 (1998) 2629.

[14] M. T. Anderson, math.dg/0303260.

[15] D. R. Brill, J. Louko, and P. Peldan, Phys. Rev. D56 (1997) 3600.

[16] J. A. Wolf, *Spaces of Constant Curvature* (Publish or Perish, Houston, 1974).

[17] P. Petersen, *Riemannian Geometry* (Springer, 1998).

[18] J. G. Ratcliffe and S. T. Tschantz , Experiment. Math. 9 (2000) 101.

[19] M. Burger, T. Gelander, A. Lubotzky, and S. Mozes, Geom. Funct. Anal. 12 (2002) 1161.

[20] D. J. Gross and V. Periwal, Phys. Rev. Lett. 60 (1988) 2105.

[21] See, for example, the articles by D. J. Gross, V. Kazakov, and M. R. Douglas in *Strings 90*, edited by R. Arnowitt et al. (World Scientific, 1991).

[22] J. G. Ratcliffe and S. T. Tschantz , Class. Quant. Grav. 15 (1998) 2613.

[23] D. D. Long and A. W. Reid, Math. Res. Lett. 8 (2001) 443.

[24] A. W. Reid, personal communication.

[25] D. D. Long and A. W. Reid, Geom. Topol. 4 (2000) 171.

[26] D. Coulson, O. A. Goodman, C. D. Hodgson, and W. D. Neumann, Experiment. Math. 9 (2000) 127.

[27] D. Bar-Natan and E. Witten, Commun. Math. Phys. 141 (1991) 423.

[28] B. S. DeWitt, in *Relativity, Groups, and Topology II*, edited by B. S. DeWitt and R. Stora (North-Holland, 1984).

[29] F. Dowker and S. Surya, Phys. Rev. D58 (1998) 124019.

[30] A. Borde, H. F. Dowker, R. S. Garcia, R. D. Sorkin, and S. Surya, Class. Quant. Grav. 16 (1999) 3457.

[31] A. Anderson and B. S. DeWitt, Found. Phys. 16 (1986) 91.

[32] N. Cornish, D. Spergel, and G. Starkman, Class. Quant. Grav. 15 (1998) 2657.

[33] B. F. Roukema, Bull. Astron. Soc. India 28 (2000) 483.