Topology and the cosmic microwave background

Janna Levin[a,b,*]

[a]*DAMTP, CMS, Cambridge University, Wilberforce Rd, Cambridge CB3 0WA, UK*
[b]*Theoretical Physics Group, Imperial College, London, UK*

Abstract

Nature abhors an infinity. The limits of general relativity are often signaled by infinities: infinite curvature as in the center of a black hole, the infinite energy of the singular big bang. We might be inclined to add an infinite universe to the list of intolerable infinities. Many theories that move beyond general relativity naturally treat space as finite. In this review we discuss the mathematics of finite spaces and our aspirations to observe the finite extent of the universe in the cosmic background radiation. © 2002 Published by Elsevier Science B.V.

Contents

* Corresponding author. DAMTP, CMS, Cambridge University, Wilberforce Rd, Cambridge CB3 0WA, UK.
 E-mail address: j.levin@damtp.cam.ac.uk (J. Levin).

1. Introduction

Is the universe infinite? The universe had a birth and will have a death, whether from old age or a more catastrophic demise through a big crunch [1]. At least this is the standard picture of cosmology. Beyond the standard model there are a few theoretical scenarios of an eternal cosmos but all modern paradigms for cosmology are at the least dynamical. With the advent of general relativity, cosmology has accepted space as an evolving geometry. Although the universe is usually believed to begin and end a dynamical life, we continue to ascribe a property to the universe that we would never assign to any other natural object, and that is the property of being spatially infinite.

The universe is home to a plethora of structures all of which must be finite. It would be untenable to suggest that any physical structure be infinite, yet cosmologists often cavalierly assume space itself is infinite. Yet an infinite universe means an infinite number of stars and galaxies and people. By the

law of probability an infinite universe accommodates an infinite number of events each happening an infinite number of times [2,3]. Somewhere else in the cosmos, you are there. In fact, there are an infinite number of you littering space. An infinite number of us living the same lives and an infinite number living slightly different lives and all range of variants. To understand this, formally, we would have to properly define probabilities in an infinite volume system, which we do not yet know how to do. While this concept of infinity is manageable to some and may be resolveable, it reeks of something pathological to others. Is the universe a surreal object, unlike any of its inhabitants, or is it real and physical and limited and finite?

If the universe is finite, then light may wrap around the finite space decorating the sky with ghost images, sometimes called clone images.[1] In principle, we should be able to see copies of ourselves at different ages in different directions. A kaleidoscopic version of 'This is your life'. In practice, the universe does not appear to be so small. The universe, if finite, must at least be big. Big enough to house clusters of galaxies and perhaps even bigger than the entire light travel distance since the big bang. Therein lies our question. If the universe is finite, how small is it, and how do we measure its shape?

The question of the infinite extent of space is often answered dynamically by Einstein's theory. Assuming the weak energy condition is satisfied and the universe is not inflating, the standard model of cosmology presents three possibilities for the shape of space. If the amount of matter and energy is underdense then the universe will expand forever, if overdense then the universe will ultimately recollapse to a big crunch, and if critical then the universe will expand forever. To each of these dynamical possibilities corresponds a geometry of spacetime: the underdense cosmos is negatively curved and infinite, the overdense cosmos is positively curved and finite and the critical cosmos is flat and infinite.

Each of these three geometries can support topologies with finite volumes without altering the dynamics or the curvature. Multiconnected topologies are often overlooked in favor of the simply connected possibilities. This oversight is encouraged by a limitation of Einstein's theory which elevates gravity to a theory of geometry but does not provide a theory of topology. If unification of gravity with the other forces ever succeeds, topology will inevitably be integrated as a predictive feature of any cosmology. The first suggestion of compact extra dimensions began with Kaluza back in Einstein's day and have been dressed up in the modern guise of quantum gravity theories. The topology of additional compact dimensions is already understood to be a crucial feature in superstring theories and, in general, discretization of space into topologically finite bundles is a natural consequence of quantization [2]. No one is good enough at quantum gravity or string theory to make a prediction for the geometry and shape of the whole universe. Still the notion of topologically compact dimensions has become commonplace in fundamental physics even if it seems esoteric in astronomy. The connection with quantum gravity also suggests that even if we can never see the extent of the large three dimensions, we may be able to look for small extra dimensions to find clues about the geometry of the cosmos in direct cosmological observations.

In the absence of a predictive theory we can still explore the mathematical possibilities for a finite universe. This review covers topologies consistent with the standard cosmological geometries and our aspirations to observe the global extent of space in patterns in the oldest historical relic accessible,

---

[1] There are naturally exceptions to this rule. In the simply connected positively curved space light can only execute at most one journey across space in the history of the universe. Also if the universe inflates the expansion of the volume will outrace a lightbeam preventing multiple windings around space.

the cosmic microwave background (CMB). Methods to determine topology from the CMB range from the geometric such as the search for circles in the sky [3–5] and topological pattern formation [6–8] to the brute force statistical methods such as the method of images [9,10]. These methods will be explained in turn. This review is intended both for mathematicians interested in cosmology and cosmologists interested in the mathematics of topology. Not much will be taken for granted and we will take the time to explain even the standard cosmology.

Topology can be observed with the CMB or with the distribution of collapsed objects such as quasars or clusters of galaxies. There are reviews on galactic methods [11,12]. To keep the scope of this article manageable we will limit our discussion to searches for artifacts of topology in the CMB and defer a discussion of galactic methods to the reviews in Refs. [11–13]. (See also Refs. [14,15].) Also, much of the history on topology can be found in an earlier review [2] as well as [16]. Other accessible reviews can be found in Refs. [17,18]. Collections of papers for two international workshops can be found in [16,19]. The observational section in this review will focus on recent methods only, starting where Ref. [2] left off.

2. Topology and geometry

Cosmic topology aims to deduce the global shape of the universe by experimentally observing a pattern in the distribution of astronomical objects. To do so, it is helpful to first understand the possible topologies from a purely mathematical perspective. The isotropy of the cosmic microwave background radiation implies that the curvature of the observable universe is very nearly constant, and very nearly isotropic, so in the present article we consider only manifolds of constant curvature. The homogeneity of the observable universe does not, of course, exclude the possibility that the curvature of space varies enormously on a global scale beyond our view, but because that hypothesis is presently untestable we do not pursue it here. For the same reason, we are most interested in multiconnected spaces that are small enough to witness, although theories beyond general relativity, such as string theory, may help us to push beyond those limitations (see Section 7). There is a previous Physics Reports on cosmic topology which provides an excellent overview of the mathematical principles [2]. The reader is encouraged to consult that article as well as Refs. [20,21]. We will not repeat these detailed reviews but do survey the same topological methods for completeness and extend the discussion to include some additional modern methods.

The global shape of any space, including the ultimate outerspace, is characterized by a *geometry* and a *topology*. The term *geometry*, as used by mathematicians, describes the local curves while *topology* describes global features which are unaltered by smooth deformations. A continuous transformation is known as a *homeomorphism*, a smooth continuous and invertible map which deforms one manifold into another without cutting or tearing. General relativity is not invariant under homeomorphisms but is invariant under *diffeomorphisms*, which amount to a change of coordinates. It is this covariance of classical gravity which underlies the principles of general relativity. Covariance ensures that all observers experience the same laws of physics regardless of the worldline along which they travel and therefore regardless of the coordinate system within which they interpret the world. Relativity does not invoke such a principle with respect to topology. For this reason we do not have a fundamental principle to guide us when contemplating the topology of the universe. We might hope that such a principle or at least a prediction might precipitate from a complete theory
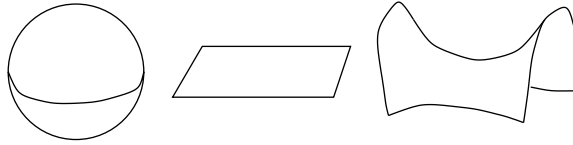
Fig. 1. A homogeneous isotropic space may be spherical (positive curvature), flat (zero curvature), or hyperbolic (negative curvature).

of gravity beyond classical. In the meantime, we consider any mathematically allowed topology, restricting ourselves to the constant curvature spaces as indicated by observational cosmology.

2.1. Overview of geometry

A cosmological principle asserts that the Earth should not be special in location or perspective. Cosmological observations on the largest scales indicate we live in a space with the symmetries of *homogeneity*, space looks the same in all directions, and *isotropy*, space looks the same in all directions. Homogeneous and isotropic manifolds have constant curvature. The *n*-dimensional geometries of constant curvature are the everywhere spherical $\mathbf{S}^n$, flat $\mathbf{E}^n$, and hyperbolic $\mathbf{H}^n$ geometries (Fig. 1). One natural consequence of curved geometries worth remembering is that the sum of the interior angles of a triangle are greater than π on $\mathbf{S}^n$, exactly π on $\mathbf{E}^n$, and less than π on $\mathbf{H}^n$.

In 2D, all Riemannian surfaces are homeomorphic to the three constant curvature geometries $\mathbf{S}^2$, $\mathbf{E}^2$, and $\mathbf{H}^2$. The infinite hyperbolic plane $\mathbf{H}^2$ cannot be drawn properly in 3D and so is difficult to visualize. It can be drawn schematically as a surface which everywhere has the curvature of a saddle. A nice way to treat $\mathbf{H}^2$ is as a pseudosphere embedded in $(2+1)$-Minkowski space. The pseudosphere has the equation $-z^2 + x^2 + y^2 = -1$. The full isometry group of the hyperboloid is $PSL(2, \mathbf{R}) \equiv SL(2, \mathbf{R})/Z_2$ with $SL(2, \mathbf{R})$, the special Lorentz group of real 2×2 matrices with unit determinant. The embedding in Minkowski space makes the symmetry group of $\mathbf{H}^2$ more obvious.

The sphere $\mathbf{S}^2$ can be embedded in Euclidean 3-space as a surface with radius $x^2 + y^2 + z^2 = 1$. The isometry group for the sphere is $O(3)$, all orthogonal 3×3 matrices with the absolute value of the determinant equal to one.

The plane $\mathbf{E}^2$ is the infinite surface (x, y) with the Minkowski time coordinate z fixed at unity (see Fig. 12). The Euclidean plane has the full Galilean group of isometries: translations, reflections, rotations and glide reflections. A glide reflection involves a translation with a reflection along the line parallel to the translation.

In 3D, all manifolds are not homeomorphic to the constant curvature manifolds. Instead, there are 8 homogeneous geometries some of which are anisotropic. These 8 geometries were classified by Thurston in the mathematical literature [22,23]. Cosmologists are more familiar with the equivalent classification of Bianchi into 8 homogeneous manifolds [24]. Out of respect for the cosmological principle, we will consider the fully homogeneous and isotropic spaces of constant curvature $\mathbf{S}^3, \mathbf{E}^3$, and $\mathbf{H}^3$. Similar to 2D, $\mathbf{H}^3$ can be embedded as a pseudosphere in $(3+1)$-Minkowski space, $\mathbf{S}^3$ as the sphere and $\mathbf{E}^3$ as the plane at fixed Minkowski time. The isometry group of $\mathbf{S}^3$ is $SO(4)$. The isometry group of $\mathbf{E}^3$ is $R^3 \times SO(3)$, which is the product of translations and the special orthogonal 3×3 matrices. The isometry group of $\mathbf{H}^3$ is $PSL(2, \mathbf{C}) \equiv SL(2, \mathbf{C})/Z_2$.
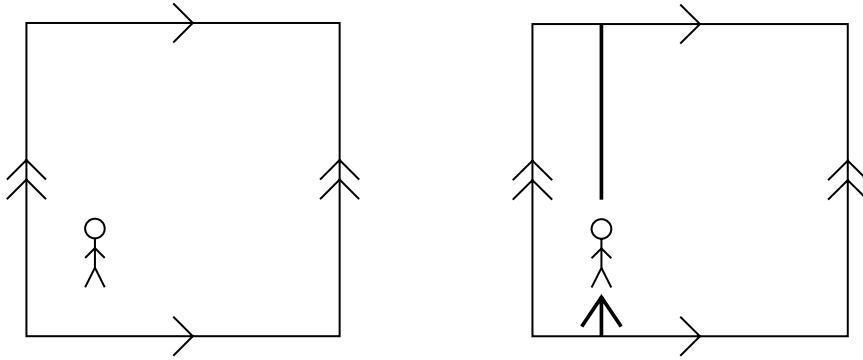
Fig. 2. The flat 2-dimensional torus is defined by abstractly gluing opposite edges of a square.

Fig. 3. An inhabitant of the 2-torus looks forward and sees herself from behind.
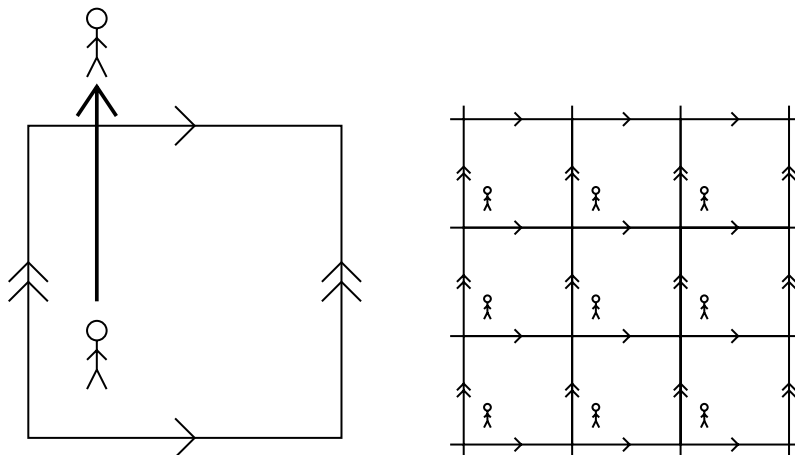
Fig. 4. The inhabitant of the 2-torus has the illusion of seeing a copy of herself.

Fig. 5. Indeed, the inhabitant of the 2-torus sees a lattice of images of herself.

2.2. Overview of topology

Topology is the pursuit of equivalence classes of spaces; that is, the classification of homeomorphic manifolds. To a topologist, all surfaces with one handle for instance are equivalent. Thus the doughnut and the coffee cup, in addition to providing a dubious breakfast, provide the quintessential example of equivalent topologies, despite their obviously different geometries (i.e. curves).

The doughnut and coffee cup are both manifestations of the 2-dimensional torus (Fig. 2), a pro-totypical multiply connected space. The torus has a finite area, yet has no boundary. When an inhabitant of the torus looks forward (Fig. 3), her line of sight wraps around the space and she sees herself from behind. She has the illusion of seeing another copy of herself directly in front of her (Fig. 4). Indeed, she also sees herself when she looks to the side, or along a 45° line, or along any line of rational slope. She thus has the illusion of seeing infinitely many copies of herself (Fig. 5).

Fig. 6. The torus of revolution has the same topology as the flat torus but different curvature.

Fig. 7. Left: The Klein bottle is similar to the torus, only the square's top and bottom edges are each glued to themselves, not to each other. The lattice of images in the Klein bottle is different from that in the torus because it contains glide reflections as well as translations.

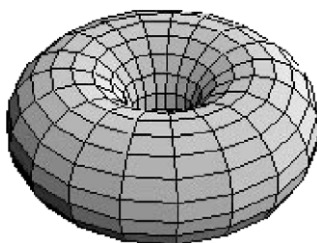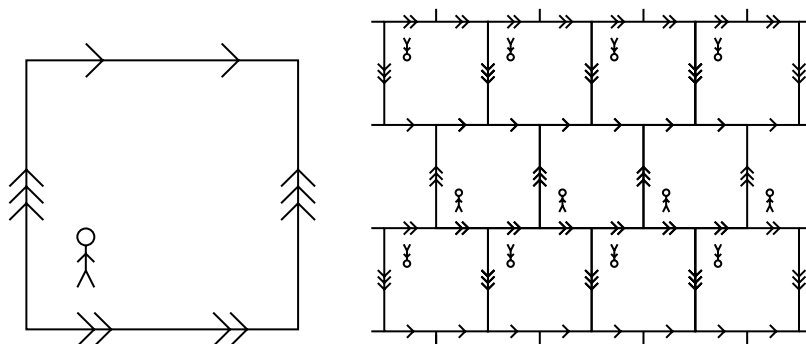In other words, even though her universe is finite, she has the illusion of living in an infinite universe containing an infinite lattice of repeating objects.

As described below, all multiconnected surfaces can be built by gluing the edges of a fundamental polygon. The torus is built by identifying opposite edges of a parallelogram. If we begin with a flat rectangle, we can bend this flat sheet into 3-dimensions to glue the top to the bottom and the left edge to the right. In doing so we have made a torus of revolution as in Fig. 6 which is neither flat nor constant curvature. The curvature of the torus clearly varies over the surface. This bending of the torus is an artifact of embedding the surface in 3-dimensions. A truly flat torus, better known as $T^2 = S^1 \times S^1$, is content to live in 2D with no such bending and projecting into 3D. The flat torus is akin to the video game with periodic boundary conditions where the left edge is identified with the right edge so that a flat explorer could stick their hand out the right edge only to have it appear poking out of the left edge. The edge in fact is as meaningless to the inhabitant of a flat torus as it is to an inhabitant of the torus of revolution. T^2 is truly flat and so has a different geometry from the torus of revolution although they are topologically equivalent, being characterized by the one handle. The genus g of the manifold is a topological invariant which counts the number of handles and holes. The torus has genus $g = 1$.

The 2-dimensional Klein bottle (Fig. 7) is similar to the torus, only now the top and bottom of the rectangle are glued not to each other, but each to itself with a shift of half a unit. The traditional way to make a Klein bottle is to start with the rectangle, glue top to bottom but glue the left to the
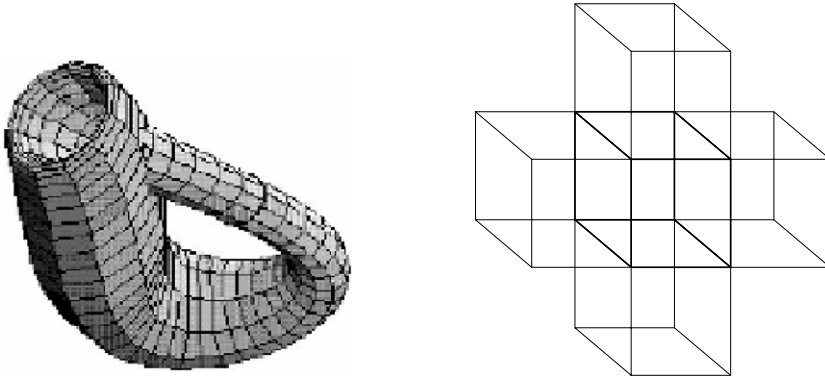
Fig. 8. A surface with the topology of the Klein bottle.

Fig. 9. 3-dimensional lattice of cubes.

right after a rotation through π. In the present article we use a different (but equivalent) construction (namely gluing the top and bottom each to itself with a shift of half a unit) because it will simplify the construction of flat 3-manifolds in Section 3.2. As in the torus an inhabitant has the illusion of seeing an infinite lattice of objects, only now the structure of the lattice is different: it contains glide reflections as well as translations. The truly flat Klein bottle is topologically equivalent to the bottle immersed in 3D and pictured in Fig. 8. The embedding again leads to a curved surface and in this case to a self-intersection of the surface. The flat Klein bottle, content to live in 2D, has no such curvature or self-intersection.

These examples generalize readily to 3-dimensions. A cube (Fig. 9) with opposite faces glued is a 3-dimensional torus, or 3-torus. Its inhabitant has the illusion of seeing an infinite 3-dimensional lattice of images of himself and of every other object in the space. Varying the gluings of the faces varies the structure of the lattice of images. Cosmic topology aims to observe these distinctive lattice-like patterns.

2.3. Tools of topology

In this preliminary discussion several important concepts have already begun to emerge. The first is the fundamental domain. A *fundamental domain* is a polygon or polyhedron from which a manifold may be constructed. For example, Fig. 2 shows a square fundamental domain for a 2-torus, and Fig. 9 shows a cube fundamental domain for a 3-torus. A *Dirichlet domain* is a special type of fundamental domain. To construct a Dirichlet domain, pick an arbitrary point in the manifold of interest to serve as a *basepoint*. Start inflating a balloon with center at the basepoint (Fig. 10) and let the balloon expand uniformly. Eventually, different parts of the balloon will bump into each other. When this happens, let them press flat against each other, forming a flat (totally geodesic) boundary wall. Eventually, the balloon will fill the whole manifold, at which point it will have the form of a polygon in two dimensions or polyhedron in 3-dimensions whose faces are the aforementioned boundary walls. This polyhedron is the Dirichlet domain. Gluing corresponding faces recovers the original manifold. Notice that the fundamental rectangular domain for the Klein bottle when glued
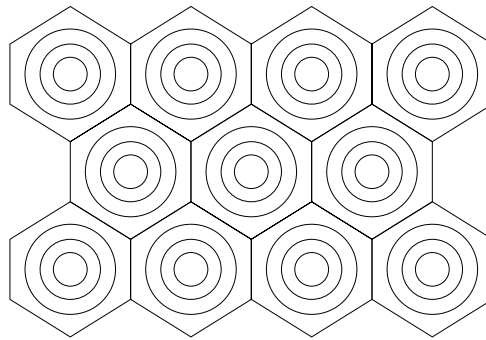
Fig. 10. Start inflating a balloon, with its center point fixed, and let it expand until it fills the entire space. The resulting polygon (in two dimensions) or polyhedron (in three dimensions) is called a Dirichlet domain. In some manifolds the shape of the Dirichlet domain depends on the choice of the center point; in other manifolds it does not.

after a shift of half a unit is not Dirichlet, but when glued after a rotation through π is Dirichlet. In some manifolds the Dirichlet domain depends on the choice of basepoint while in others it does not—more on this later.

By identifying the edges of the fundamental domain, a multiconnected space is constructed. Figs. 5 and 7 show how an inhabitant of a finite multiply connected space may have the illusion of living in an infinite simply connected space. This apparent space is called the *universal cover* and can be thought of as the simply connected manifold with the constant curvature geometry. The fundamental domain is cut from this cloth and glued together while preserving the geometry of the universal cover. So the flat torus preserves the geometry of the universal cover while the torus of revolution does not.

The full isometry group of the universal cover is usually denoted by *G*. An *isometry* is a one-to-one map which leaves the metric invariant and therefore preserves measures of lengths. The topological compactification of a space may preserve the geometry but will destroy some of the global symmetries of the universal cover. For instance, the infinite flat plane is symmetric under rotations but the flat torus is not.

An important structure is the group of *covering transformations*. The group of covering transformations is a property of both the universal cover and the topological space in question. It is the group of motions of the universal cover that take images of a given object to other images of the same object. For example, in the case of a torus the group of covering transformations is the obvious lattice of translations (a subset of the full symmetry group of the universal cover). In the case of the Klein bottle the group of covering transformations contains glide reflections as well as translations.

In 2-dimensions all topologies have been completely classified. We give that classification in Section 3.1. However, in 3-dimensions the classification of hyperbolic manifolds remains incomplete although huge advances were made last century. There is a collection of topological invariants which can be used to recognize the equivalence classes of spaces. We have already encountered the *genus* of the manifold which counts handles. A group structure can be used to capture handles and thereby detect the connectedness of the space. The most important of these is the loop group also known as the *fundamental group* or the first *homotopy group*. All loops which can be smoothly deformed into each other are called *homotopic*. A loop drawn on the surface of a sphere can be contracted to a point. In fact, all loops drawn on the surface of a sphere can be contracted to a point. Since
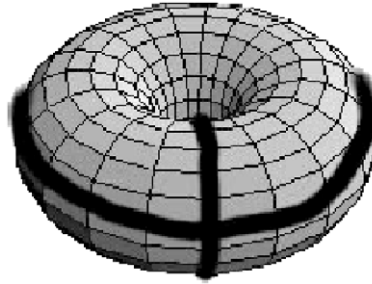
Fig. 11. Two homotopically distinct loops around the torus of revolution.

all loops are homotopic, the sphere is *simply connected* and in fact does not have any handles. If, however, we draw a loop around the torus of revolution as in Fig. 11, there is no way to smoothly contract that loop to a point. Similarly a loop drawn around the circumference of the toroid, also shown in Fig. 11, cannot be contracted to a point. The torus is therefore *multiply connected.* The group of loops is the fundamental group $\pi_1(\mathcal{M})$ and is a topological invariant. In more than 2D, the fundamental group is not sufficient to determine the connectedness of the manifold. Nonetheless, the fundamental group is a powerful tool. Poincaré has conjectured that if the fundamental group is trivial, then the n-manifold is topologically equivalent to S^n.

A closely related structure is the *holonomy group*. The holonomy group is essentially the same as the group of covering transformations. The only distinction to be made is that the elements of the holonomy group are isometries and are therefore geometric while the elements of the group of covering transformations are homeomorphisms and are therefore topological. The holonomy group can be thought of as the set of instructions for identifying the faces of the fundamental domain. The holonomies act discretely and without fixed point. The holonomy group is a subset of the full isometry group of the covering space. The full symmetries of the universal cover are broken by the identifications and it is customary to represent the compact manifold as a quotient space G/Γ where G is the isometry group of the universal cover and Γ is the holonomy group. The holonomy group is of extreme utility in the game of cosmic topology. The group provides the set of rules for distributing images throughout the manifold and thereby provides the mathematical means to determine the geometric distribution of astronomical images in the cosmos. The group is synonymous with the boundary conditions for the manifold and as we will see this is critical for building a prediction for the spectrum of the cosmic microwave background.

The holonomy group is so instrumental in cosmological applications that it is worth illustrating how the group can be used in practice. The group is expressed as a *presentation*, $\{\gamma_1 \ldots \gamma_n; r\}$, where $\gamma_i \in \Gamma$ are elements of the group. Implicitly, $r = 1$ following the semicolon and is a list of all of the relations among the group elements. There can be many different presentations for the holonomy group. There is often a simplest presentation with the fewest number of elements and relations. However, the simplest presentation is not always the most useful. The most natural presentation for cosmology has a group element for every identification rule and a correspondingly larger number of relations. The group elements are akin to letters in an alphabet and the relations define the grammar rules.

The holonomies can map the end-point of an orbit to the start of that orbit so that it becomes periodic in the compact space. In other words, a periodic orbit has $x_{\text{end}} = \gamma x_{\text{start}}$ where γ can be a composite word $\gamma = \prod_i^n \gamma_{k_i}$. Each word has a corresponding periodic orbit.
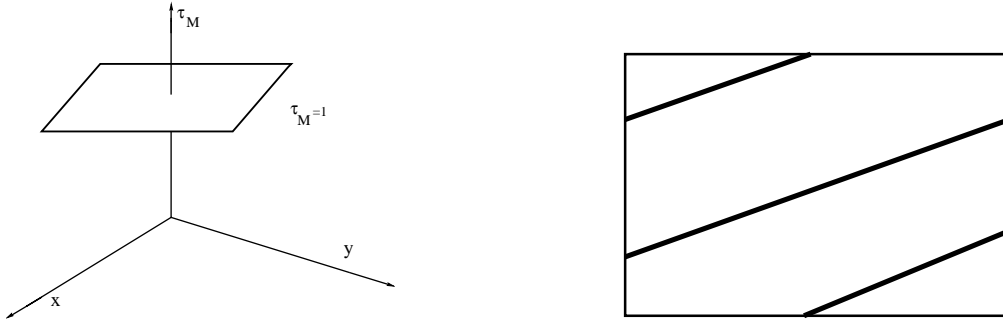
Fig. 12. The embedding of $\mathbf{E}^2$ into $(2+1)$-Minkowski space. The manifold appears as an infinite sheet fixed at $\tau_M = 1$.

Fig. 13. A periodic orbit on the torus which corresponds to $x_{\text{end}} = T_y T_x^2 x_{\text{start}}$.

As an explicit example we return to the canonical torus. We can make use of the embedding of $\mathbf{E}^2$ in a $(2+1)$-dimensional space. Specifically, the 2-dimensional coordinate is replaced with the $(2+1)$-dimensional coordinate

$$x^a = \begin{pmatrix} \tau_M \\ x \\ y \end{pmatrix} , \tag{1}$$

where τ_M is fixed at unity as in Fig. 12. In this coordinate system the holonomy group for the torus is the set of generators,

$$T_x = \begin{pmatrix} 1 & 0 & 0 \\ L_x & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad T_y = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ L_y & 0 & 1 \end{pmatrix} . \tag{2}$$

The boundary condition is then $x^a \to T_b^a x^b$ so that identification of the left edge to the right edge results in

$$\begin{pmatrix} \tau_M \\ x \\ y \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ L_x & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \tau_M \\ x \\ y \end{pmatrix} = \begin{pmatrix} \tau_M \\ x + L_x \\ y \end{pmatrix} . \tag{3}$$

The holonomy group for the torus in this presentation is $\{T_x, T_y; T_x T_y T_x^{-1} T_y^{-1}\}$. The relation $T_x T_y T_x^{-1} T_y^{-1} = 1$ is simply the commutativity of the generators and prunes the number of unique periodic orbits and therefore the number of clone or ghost images. Because the generators commute, the periodic orbits (homotopic loops) can be counted symbolically in terms of the number of windings taken around x and around y (m_x, m_y). The number of loops and therefore the number of clone images grows as a polynomial with length. For example, consider the periodic orbit of Fig. 13 which has $x_{\text{end}} = T_y T_x^2 x_{\text{start}}$ and so winding number ($m_x = 2, m_y = 1$).

The non-orientable Klein bottle first twists the z-faces through π before identification. The generators of Γ for the twisted space are $T_x, R(\pi)T_y$ with $R(\theta)$ the rotation matrix:

$$R(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix} . \tag{4}$$

All of the multiconnected, flat topologies can be built out of a combination of rotations and these translations.

In hyperbolic space, the generators cannot easily be written in functional form but can, nonetheless, be found with numerical entries in the matrix. The program SnapPea provides the generators explicitly in this embedded coordinate system for a census of known hyperbolic manifolds and orbifolds [25]. (Orbifolds can have singular points.) The number of periodic orbits and therefore of ghost images grows exponentially in a compact hyperbolic manifold. The exponential proliferation of images is a manifestation of the exponential growth in volume with radius and is related to the chaotic flows supported on compact hyperbolic spaces and discussed in Section 7.3.

2.4. Tessellations

The modest tools we have introduced go a long way in constructing cosmological models. To summarize we will characterize a manifold by

- The *fundamental domain*—the shape of space around a given basepoint.
- The *universal cover*—the simply connected manifold with the same constant curvature geometry.
- The *holonomy group*—face pairing isometries, i.e., the rules for identifying the faces of the fundamental domain.

Armed with these tools, we can begin to visualize life in a finite universe. An important visualization technique comes from a tessellation of the universal cover. Consider the torus again. The flow of an observer or of light can be followed through the compact flat space by conscientiously respecting the boundary conditions and drawing the motion within the fundamental polygon as in Fig. 14. The Dirichlet domain is the set of points

$$\{y \in \mathcal{M}\} , \tag{5}$$

such that

$$d(y,x) \leqslant d(y,\gamma x) \quad \forall \gamma \in \Gamma . \tag{6}$$

In T^2, light travels along straight lines as do all inertial observers. Another faithful representation of T^2 which does not distort the everywhere flat nature of the geometry is provided by a tiling picture. Beginning with the fundamental polygon, the periodic boundary condition which identifies the left edge of the rectangle to the right edge for instance can be represented by gluing an identical copy of the fundamental polygon left edge to right edge. This amounts to moving an entire copy of the fundamental domain with $T_x(FD)$. Similarly identical copies can be glued top to bottom $T_y(FD)$ ad infinitum until the entire infinite flat plane is tiled with identical copies of the rectangle. In general, the generators of the holonomy group can be thought of as defining an alphabet of $2n$ letters, for n generators and their n inverses. The alphabet will vary with different representations, as will the
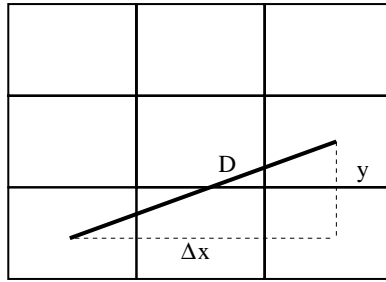
Fig. 14. The compact hypertorus can be represented as an identified parallelepiped. Alternatively, the compact topology can be represented by tiling flat space with identical copies of the fundamental parallelepiped. In the tiling picture above only the (x, y) directions are shown. A particular periodic orbit is drawn which corresponds to $x_{\text{end}} = T_y T_x^2 x_{\text{start}}$.

grammar (that is, the set of relations). Usually, though not always, the longer the word built out of the letters $\gamma_i \in \Gamma$, the more distant the corresponding image or tiling in the tessellation picture. Figs. 5, 7, and 10, demonstrate tessellations.

These copies are truly identical. If someone stands in the center of one rectangle he will see images of himself standing at the center of every other rectangle. If he moves to the left he will see every copy of himself move left. Because of the finite light travel time, it will take longer to see the more distant images moving. The light from these more distant images takes a longer path around the compact space before reaching the observer, the source of the very image. This creates the pattern of images already described so that the universe is an intricate hall of mirrors.

Notice that the tiling completely fills the plane without overlaps or gaps. We can also imagine tiling the plane with a perfect square or even a parallelogram without overlaps or gaps. A flat torus can therefore be made by identifying the edges of these other polygons. All have the same topology and the same curvature but some have different metrical features in the sense that the rectangle is truly longer than it is wide. An observer could line rulers up in space and unambiguously measure that their space was longer than it was wide while an observer in the perfect square would know that their space was equally long and wide. The torus is thus globally anisotropic even though the geometry of the metric is still locally isotropic. The torus is still globally homogeneous. The center could be moved arbitrarily and the fundamental polyhedron redrawn around that center to give an equally valid tiling. Most topologies globally break both homogeneity and isotropy.

The holonomies provide the list of rules for identifying the edges of the polygon. So the Klein bottle, while it can begin with the same fundamental polygon as the torus and has the same universal cover, it has a different set of rules for identifying the edges and therefore a different holonomy group. The tiling will be correspondingly different with identical copies glued left edge to right edge but top edge to bottom edge only after rotating the tile by π. So the tessellation creates the reflected and translated images specific to the Klein bottle (Fig. 7).

The flat plane can also be tiled by hexagons (Fig. 10) as many a bathroom floor demonstrates. The hexagon therefore also reproduces a flat torus. However, the general parallelogram and hexagon exhaust the possible convex polygons which can tessellate flat space. If we tried to fill the plane with octagons, we would find that they overlapped at a vertex and could not smoothly fill the plane, as in Fig. 15. If we glue the octagons together without confining them to the plane, they will create a curling surface that tries to have a hyperbolic structure. In fact, the 2D hyperbolic plane $\mathbf{H}^2$ can
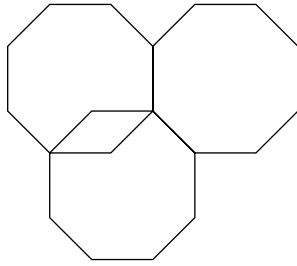
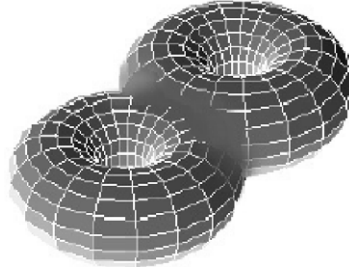Fig. 15. It is not possible to tile the plane with octagonal shaped tiles.

Fig. 16. The double doughnut.

be tiled by regular octagons. Recall that the angles of a triangle sum to less than π in hyperbolic geometry. Since the interior angles of the polygon narrow on the hyperbolic plane, they can be drawn just the right size relative to the curvature scale so that precisely the right number fit around a vertex to fill the negatively curved plane without gaps or overlaps. Note also that there is no scale in flat space and there is no intrinsic meaning to the area of the compact torus. It can be made arbitrarily large or small relative to any unit of measure. This is not the case when the surface is endowed with a curved metric. The area is then precisely determined by the topology. The area of the octagon is fixed relative to the curvature scale to properly tessellate the universal cover.

The topology created by the compact octagon is a double torus T_2 where the 2 refers to the number of handles or genus, g, of the manifold. Hyperbolic T_2 is topologically equivalent to the double doughnut of Fig. 16 which lives in R^3 although again T_2 has a different geometry, being flat, while the double doughnut is clearly curved.

A connected sum of tori can be constructed to make T_g, a flat 2D compact surface with g handles. In order to make a g handled object, a polygon with $4g$ edges is needed [23]. The polygons have to be drawn larger relative to the curvature scale the more edges they have in order that the interior angles be thin enough to fit together around a vertex and tile H^2. Consequently, a relationship emerges between the area of the surface and the topology. The area of the hyperbolic surface can be related to the Poincaré–Euler characteristic, $\chi = 2(1 - g)$ through the Gauss–Bonnet theorem

$$A = -2\pi\chi = 4\pi(g - 1) . \tag{7}$$

Although the area is a topological invariant in 2D, not all metrical quantities are topologically invariant. For instance, the asymmetric octagon and the regular octagon have the same topology, namely that of the double doughnut, and the same negative curvature, but they have a different spectrum of periodic orbits. So the lengths of the periodic orbits are not topologically invariant although the area is.

The tessellation of the universal cover is of particular importance when we consider 3-dimensional spaces. We simply do not have four dimensions available to us in which to bend the 3-dimensional volume and that visualization technique fails us. However, we can still visualize filling the 3-dimensional volume with 3D tiles which preserve the local geometry of the manifold. Therefore we rely heavily on the universal cover, the fundamental polyhedron and the holonomy group when investigating the topologies of these compact spaces.

2.5. Hyperbolic topologies and volumes

All 2-dimensional manifolds have been classified, as have flat and spherical 3-manifolds, but hyperbolic 3-manifolds are not yet fully understood. These continue to resist classification although enormous advances were made last century. An important realization was the rigid connection between metrical quantities and topological features in hyperbolic 3-manifolds. Metrical quantities describe lengths and would ordinarily fall under the domain of geometry. Yet in 3-dimensions the Mostow–Prasad [26] rigidity theorem ensures that once the topology is specified, then all metrical quantities are fixed for 3-D hyperbolic manifolds. This means that not only the volume but also the lengths of the shortest geodesics are immutable for a given topology. This is to be compared with 2D where a given topology can support an infinite number of metrics as already described.

Although the topology fixes the compact hyperbolic volume, the volume does not uniquely specify the topology. In other words, there can be different topologies with the same volume. Because of this powerful link between geometry and topology, compact hyperbolic manifolds can be ordered according to their volume. A remarkable result is that compact hyperbolic manifolds form a countably infinite number of countably infinite sequences ordered according to volume. A given sequence shows an accumulation of manifolds near a limiting volume set by a cusped manifold. A cusped manifold has finite volume but is not compact, having infinitely long cusped corners which taper infinitely thin. A cusp is topologically a 2-torus which is conformally shrunk to zero down the narrowing throat. The sequence of compact manifolds are built systematically through the process of Dehn surgery which is used in the next section to organize the spectrum of topologies. Dehn surgery is a formal process whereby a torus is drilled out of the cusped manifold and replaced with a solid torus. The surgery is taken along a periodic orbit of T^2 identified by the number of windings (m_x, m_y) the orbit takes around the torus before closure. The program SnapPea [25] names manifolds according to the number of windings so that the manifold m003$(3, -1)$ is a compact space constructed by Dehn filling the orbifold m003 along the periodic orbit with winding numbers $(3, -1)$.

The space m003$(3, -1)$ is also known as the Weeks space after the mathematician who discovered it and coincidentally has contributed to this article [27]. The Weeks space is of particular relevance to cosmology since it is the smallest known manifold with a volume $\mathscr{V} \sim 0.94$ in curvature units. The Weeks space displaced the Thurston space m003$(-2, 3)$ which has volume $\mathscr{V} \sim 0.98$. Although the Weeks space is the smallest known manifold it may or may not be the smallest possible. The current bound on the minimum volume for any compact hyperbolic manifold is about 0.3 in units of the curvature radius cubed. Contrast this with the flat hypertorus which can be made infinitesimally small. Again, the rigidity of compact hyperbolic spaces is at work.

Because of the rigidity of $3D$ hyperbolic manifolds, we can also characterize a topology by certain volume measures. Very useful quantities in cosmological investigations are the *in-radius* r_- and the *out-radius* r_+. The in-radius is the radius of the largest geodesic ball which can be inscribed within the Dirichlet domain. The out-radius is the radius of the smallest geodesic ball which can encompass the Dirichlet domain. The disparity between r_- and r_+ can be taken as a rough indication of the global anisotropy introduced by the topological identification. The *injectivity radius* is the radius of the largest ball, centered at the given point, whose interior embeds the manifold. The Dirichlet domain given by default in SnapPea is centered at a local maximum of the injectivity radius. In this way it often provides the most symmetric form for the Dirichlet domain. Moving the observer away from this basepoint will change the appearance of the global shape of space observationally.

A somewhat ambiguous but sometimes useful quantity is an estimate of the diameter of the manifold. Since the manifold is not isotropic this is not precisely defined. Nonetheless, one estimation is given by

$$d_{\mathcal{M}} = \sup_{x,y \in \mathcal{M}} d(x,y) \; ; \tag{8}$$

that is, it is the largest distance between any two points in the Dirichlet domain. Note that it is not a geodesic distance so no periodic orbit lies along the diameter. A more well-defined diameter is the maximum distance between two points in the manifold, where the distance is measured in the manifold itself, not in the Dirichlet domain:

$$diam = \sup_{x,y} \left(\inf_{\text{paths connecting } x \text{ to } y} (\text{length of path}) \right) \; . \tag{9}$$

The periodic geodesics are quite important and are closely related both to the eigenmodes in the space and the distribution of ghost or clone images. The periodic geodesics can be found with the holonomy group. The elements of the holonomy group can be expressed as matrices which transform a place on the manifold into its image as in Section 2.3. Eigenvectors for these matrices can be found in the usual way. The eigenvectors of the holonomy group span a plane in the $4D$ embedding which intersects the pseudosphere $\mathbf{H}^3$. The line of intersection occurs along the periodic orbit corresponding to that group element. A composite element can be thought of as a word formed from the fundamental alphabet of the presentation. The fact that the number of words grows exponentially with length means an exponential growth of long periodic orbits. The proliferation of periodic orbits, dense in phase space, are the canonical mark of chaotic dynamics. This leads us to remark that compact hyperbolic spaces are not just complicated mathematically but they are truly complex, supporting chaotic flows of geodesics. Since most cosmological investigations try to circumvent the chaotic nature of the motions, we mention chaos only occasionally as we go along. We reserve a final section for a brief discussion of this fascinating feature. A review of chaos on the pseudosphere can be found in Ref. [28].

3. Survey of compact manifolds

Having described the topological principles at work and some examples, we can compile a list of known manifolds, their geometries and topologies. The classification is clearly presented in Refs. [2,20,21]. We present a different mathematical approach for variety. The classification is organized by Weeks [29] in relation to four questions:

1. What constant curvature manifolds exist?
2. How flexible is each manifold? That is, to what extend can we change a manifold's shape without changing its topology or violating the constant curvature requirement?
3. What symmetries does each manifold have?
4. Does a Dirichlet domain for the manifold depend on the choice of basepoint?

It turns out that the answers to these questions depend strongly on both the curvature and the dimension. The remainder of Section 3 answers these questions in each of the six possible cases:
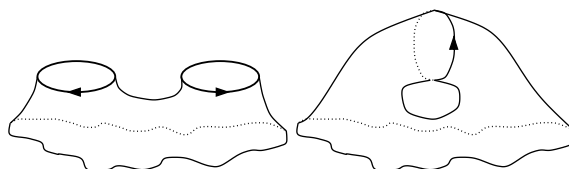
Fig. 17. To add a handle to a surface, remove two disks and glue the remaining boundary circles to each other.
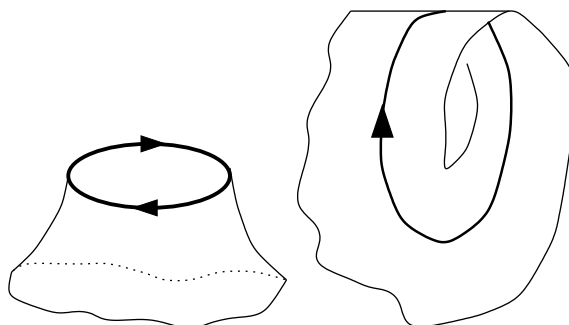
Fig. 18. To add a crosscap to a surface, remove one disk and glue the remaining boundary circle to itself by identifying antipodal points. The construction cannot be physically carried out in R^3 without self-intersections, but abstractly the resulting surface has a seam that looks locally like the centerline of a Möbius strip.

that is, in spherical, flat, or hyperbolic geometry in 2 dimensions (Section 3.1) or 3 dimensions (Section 3.2).

3.1. 2-Dimensional manifolds

3.1.1. Overview of 2-dimensional manifolds

Every closed surface may be obtained by starting with a sphere and adding handles (Fig. 17) or crosscaps (Fig. 18). Each handle is constructed by removing two disks from the sphere and gluing the resulting boundary circles to each other (Fig. 17), and each crosscap is constructed by removing one disk and gluing the resulting boundary circle to itself (Fig. 18). Conway found a particularly simple and elegant proof of this classification; for an illustrated exposition see Ref. [30].

The interaction between geometry and topology is a primary theme in our understanding of 2-manifolds. The promised interaction between geometry and topology is that each closed topological surface may be given a constant curvature geometry. Each closed 2-manifold admits a unique geometry. That is, a surface which admits a spherical geometry cannot also admit a flat geometry, and so on. For an elementary proof see Ref. [31]. Infinite 2-manifolds, by contrast, may admit more than one geometry. For example, the Euclidean plane and the hyperbolic plane are topologically the same 2-manifold. We organize the surfaces according to the total number of disks that get removed during the construction (Fig. 19).

In the case that no disks are removed, we have the sphere.

In the case that one disk is removed, we draw the sphere-minus-a-disk as a hemisphere, so that when the boundary circle gets glued to itself the spherical geometry continues smoothly across the
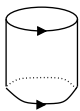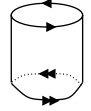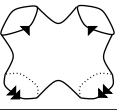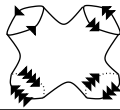
Fig. 19. Topological classification of closed 2-manifolds. Every closed 2-manifold may be obtained from a sphere by adding either handles (left column) or crosscaps (right column).

seam from one side to the other, giving the resulting sphere-with-a-crosscap a uniform spherical geometry.

In the case that two disks are removed we may have either a sphere-with-a-handle, better known as a torus, or a sphere-with-two-crosscaps, better known as a Klein bottle. We draw the sphere-minus-two-disks as a cylinder. The cylinder has an intrinsically flat geometry. You can construct one from a sheet of paper without stretching, and when you glue its top and bottom of the rectangle to each other (for the torus) or each to itself (for the Klein bottle) the flat geometry continues uniformly across the seams. Thus the torus and the Klein bottle each have a flat geometry.

In the case that three disks are removed we have a sphere-with-three-crosscaps, constructed from a three-way cylinder better known as a *pair of pants*. The pair of pants may in turn be constructed from two hyperbolic hexagons (Fig. 20), giving it an intrinsically hyperbolic geometry. When the pants' boundary circles (the cuffs) are glued to themselves to make the sphere-with-three-crosscaps, the hyperbolic geometry continues uniformly across the seams. Thus the sphere-with-three-crosscaps also gets a hyperbolic geometry.

All remaining surfaces—those for which four or more disks are removed in the construction—may be built from pairs of pants, so all get a hyperbolic geometry. In summary, we see that of the infinitely many closed 2-manifolds, two admit spherical geometry (the sphere and the sphere-with-crosscap), two admit flat geometry (the torus and the Klein bottle), and all the rest admit hyperbolic geometry.

The next three sections will examine the flat, spherical, and hyperbolic closed 2-manifolds in greater detail, in each case answering the four questions raised in Section 2.2.
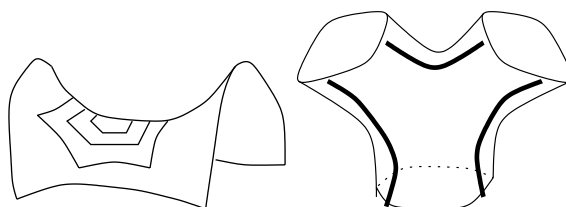
Fig. 20. A pair of pants can be made from two hexagons. For best results use hexagons cut from the hyperbolic plane, with all interior angles being right angles.
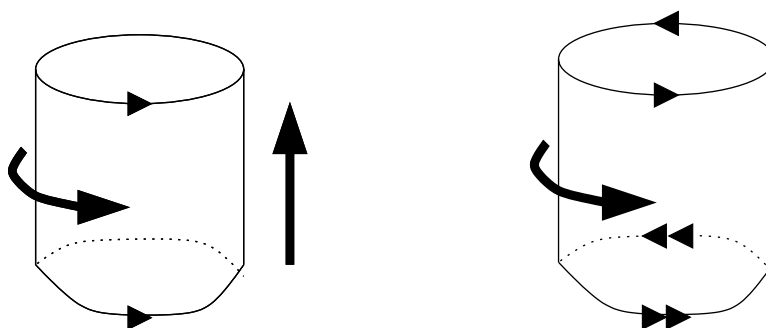
Fig. 21. The torus has a two-parameter continuous family of symmetries, consisting of rotations and vertical translations.

Fig. 22. The Klein bottle has continuous rotational symmetry, but any attempt at vertical translation is blocked by the crosscaps.

3.1.2. Flat 2-dimensional manifolds

Manifolds: *Torus and Klein bottle*.

Flexibility: The torus has three degrees of freedom in its construction: the height of the cylinder, the circumference of the cylinder, and the twist with which the top gets glued to the bottom. The Klein bottle has only two degrees of freedom in its construction: the height and circumference of the cylinder. Each end of the cylinder gets glued to itself, identifying diametrically opposite points, so there is no twist parameter.

Symmetry: Continuous symmetries. The torus has a two-parameter continuous family of symmetries, consisting of rotations and vertical translations (Fig. 21). This shows that the torus is *globally homogeneous* in the sense that any point may be taken to any other point by a global symmetry of the manifold. The Klein bottle has only a 1-parameter continuous family of symmetries, consisting of rotations (Fig. 22). It is globally inhomogeneous. When diametrically opposite points on a boundary circle are glued to make a crosscap, the seam is a closed geodesic that looks locally like the centerline of a Möbius strip. Points lying on this centerline are fundamentally different from other points in the manifold.

Discrete symmetries: As well as its continuous family symmetries, the torus has a discrete Z_2 symmetry given by a 180° rotation interchanging the cylinder's top and bottom (Fig. 23). The discrete and continuous symmetries may of course be composed. The best way to think of this is that the complete symmetry group of the torus consists of two disconnected components: symmetries that interchange
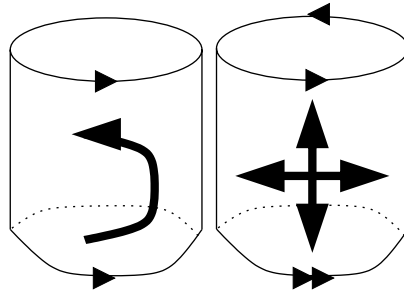
Fig. 23. The torus always has a discrete Z_2 symmetry given by a half turn, even if the cylinder's top and bottom circles are glued to each other with a twist. The Klein bottle has a discrete $Z_2 \times Z_2$ symmetry, which interchanges and/or inverts the crosscaps.
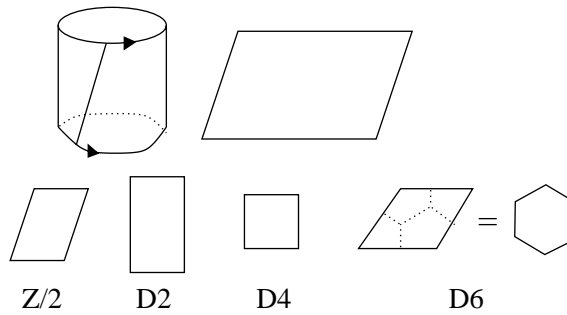
Fig. 24. Cut the cylinder along a line connecting a point on the bottom circle to the corrsponding point on the top circle, taking into the twist in the gluing. The resulting fundamental domain is a parallelogram. Every parallelogram admit a half turn, for a Z/2 symmetry group. If the parallelogram happens to be a rectangle, it also admits reflections, for a D2 symmetry group. If the parallelogram happens to be a square, it admits a quarter turn, for a D4 symmetry group. Finally, if the parallelogram happens to be a rhombus with angle $\pi/3$, the Dirichlet domain centered at a vertex is a regular hexagon, and the torus admits a one sixth turn, for a D6 symmetry group.

the boundary circles and symmetries that do not. In special cases the torus may have additional discrete symmetries (Fig. 24). The Klein bottle's discrete symmetry group is $Z_2 \times Z_2$, generated by vertical and horizontal reflections of the cylinder. Vertical reflection interchanges the two crosscaps. Horizontal reflection takes each crosscap to itself, but reverses the direction of its centerline.

Dirichlet domain: We have seen that the torus is globally homogeneous, so the shape of the Dirichlet domain does not depend on the choice of basepoint. On the other hand, the torus has three degrees of flexibility, and the Dirichlet domain's shape does depend on the shape of the torus. Typically, the Dirichlet domain is an irregular hexagon, but in special cases it may be a square or a regular hexagon. The Klein bottle is globally inhomogeneous, so even for a fixed geometry of the Klein bottle, the Dirichlet domain's shape depends on the choice of basepoint (Fig. 25).

3.1.3. Spherical 2-dimensional manifolds
Manifolds: Sphere and sphere-with-crosscap

Flexibility: The sphere and the sphere-with-crosscap have no flexibility whatsoever. Their geometry is completely determined by their topology. Here and throughout we make the assumption that
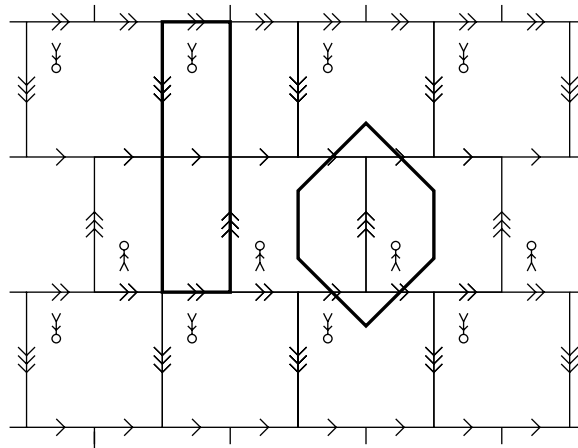
Fig. 25. In a Klein bottle, the shape of the Dirichlet domain depends on the choice of basepoint. For example, a Dirichlet domain centered at the triple arrowhead is a hexagon, while a Dirichlet domain centered at a single arrowhead is a rectangle.

spherical geometry refers to a sphere of radius one. Otherwise these two manifolds would have the flexibility of rescaling—they could be made larger or smaller.

Symmetry: The sphere has a symmetry taking any point to any other point, with any desired rotation. These symmetries form a 3-parameter continuous family. It also has a single discrete symmetry, given by reflection through the origin. The sphere-with-crosscap has a 3-parameter continuous family of symmetries taking any point to any other point with any desired rotation and either reflected or not reflected. It has no discrete symmetries. Indeed, the continuous family already takes a neighborhood of any point to a neighborhood of any other point in all possible ways, leaving no possibilities for discrete symmetries.

Dirichlet domain: The sphere and the sphere-with-crosscap both have somewhat degenerate Dirichlet domains (try the inflating-the-balloon experiment in each), but these manifolds are globally homogeneous so the degenerate Dirichlet domains do not depend on the choice of basepoint.

3.1.4. Hyperbolic 2-dimensional manifolds

Manifolds: There are infinitely many closed hyperbolic 2-manifolds, illustrated in Fig. 19.

Flexibility: All closed hyperbolic 2-manifolds are flexible. The shape of each manifold is parameterized by the circumferences of the cuffs of the pairs of pants used to construct it, and by the twists with which distinct cuffs are glued to each other. A cuff that is glued to itself to make a crosscap has no twist parameter.

Symmetry: Closed hyperbolic 2-manifolds never have continuous families of symmetries. The reason is that each seam (where cuffs are joined together) is the shortest loop in its neighborhood, and therefore cannot be slid away from itself. A seam cannot slide along itself either, because a pair of pants admits no continuous rotations. A generic closed hyperbolic 2-manifold also has no discrete symmetries, but nongeneric ones may have discrete symmetries if the cuff lengths and twists are chosen carefully.

Dirichlet domain: The Dirichlet domain of a closed hyperbolic 2-manifold always depends on the choice of basepoint. For carefully chosen basepoints it may exhibit some of the manifold's symmetries.

3.2. 3-Dimensional manifolds

3.2.1. Overview of 3-dimensional manifolds

Unlike 2-manifolds, not all 3-manifolds admit a constant curvature geometry. For a survey of the geometry of 3-manifolds, see Ref. [32]. Roughly speaking, any closed 3-manifold may be cut into pieces in a canonical way, and each piece admits a geometry that is homogeneous but not necessarily isotropic. For a very elementary exposition of homogeneous anisotropic geometries, see Ref. [31]. For a more detailed but still readable exposition see [33]. The observed isotropy of the microwave background radiation implies that at least the observable portion of the universe is approximately homogeneous and isotropic, so cosmologists restrict their attention to the three homogeneous isotropic geometries: spherical, flat and hyperbolic. However, as noted earlier, an approximately flat observable universe is also consistent with a very large universe of varying curvature.

3.2.2. Flat 3-dimensional manifolds

Manifolds: Fig. 26 illustrates the 10 closed flat 3-manifolds.

The first is the 3-torus introduced earlier—a cube with pairs of opposite faces glued.

The second and third manifolds are variations on the 3-torus in which the front face is glued to the back face with a half (respectively quarter) turn, while the other two pairs of faces are glued as in the 3-torus. Other possibilities for gluing the faces of a cube—for example, gluing all three pairs of opposite faces with half turns—fail to yield flat 3-manifolds because the cube's corners do not fit together properly in the resulting space. See Ref. [31] for a more complete explanation.

The fourth and fifth manifolds are similar in spirit to the first three, only now the cross section is a hexagonal torus instead of a square torus. That is, the cross section is still topologically a torus, but has a different shape (recall Fig. 24). The hexagonal cross section allows the front of the prism to be glued to the back with a one-sixth or one-third turn, producing two new flat 3-manifolds. The side faces are all glued to each other in pairs, straight across.

The sixth manifold, the famous Hantzsche–Wendt manifold, is the most interesting of all. Unlike the preceding five manifolds, which we defined by constructing fundamental domains, the most natural way to define the Hantzsche–Wendt manifold is to start in the universal cover and define its group of covering transformations. Specifically, its group of covering transformations is the group generated by screw motions about a set of orthogonal but nonintersecting axes (indicated by heavy lines in Fig. 24). Each screw motion consists of a half turn about an axis composed with a unit translation along that axis. Note that this group of covering transformations does *not* take a basic cube to all other cubes in the cubical tiling of 3-space. Rather, the images of a basic cube fill only half the cubes in the tiling, checkerboard style. Thus a complete fundamental domain would consist of two cubes, the basic cube and any one of its immediate neighbors; images of the basic cube would fill the black cubes in the 3-dimensional checkerboard, while images of its neighbor would fill the remaining white cubes. But we would really prefer a fundamental domain consisting of a single polyhedron. To get one we employ the balloon construction for a Dirichlet domain introduced in Section 3.1.2. Let the Dirichlet domain's basepoint be the center of a basic cube. As the balloon

Torus Bundles

3-torus   half turn   quarter turn

sixth turn   third turn

Hantzsche-Wendt manifold

or

Klein bottle bundles

plain   swap   invert   swap and invert
             cross caps   cross caps   cross caps
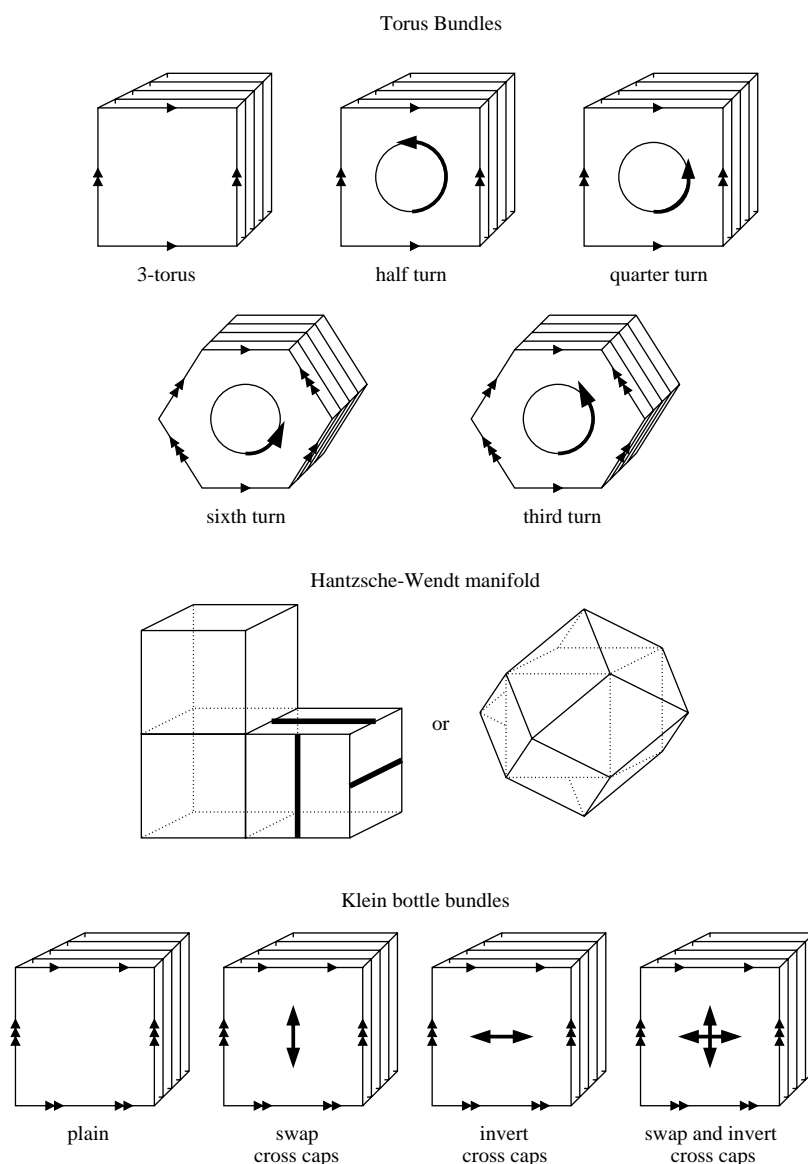
Fig. 26. The 10 closed flat 3-manifolds.

expands it fills that basic cell entirely, and also fills one-sixth of each of its six immediate neighbors. The resulting Dirichlet domain is a rhombic dodecahedron (Fig. 26). The face gluings are given by the original screw motions along the axes. Note that this construction of the Hantzsche–Wendt manifold corrects an error, appearing elsewhere in the cosmological literature, which takes the three screw axes to be coincident. This erroneous construction leads to a cube with each pair of opposite faces glued with a half turn. The cube's corners are therefore identified in four groups of two corners each instead of a single group of eight corners. The resulting space has four singular points and is thus an orbifold instead of a manifold.

The seventh through tenth manifolds are again defined by fundamental domains. They are similar to the first five manifolds, except now the cross section is a Klein bottle instead of a torus. That is, the fundamental domain's left and right faces are glued to each other as in a torus, but the top face (resp. bottom face) is glued to itself with a horizontal shift of half the width of the fundamental domain, taken modulo the full width, of course. This ensures that every cross section is a Klein bottle, which may be understood as a cylinder with one crosscap at the top and another at the bottom. The fundamental domain's front face, which is now a Klein bottle, may be glued to its back face in one of four ways: plain (seventh manifold), interchanging the crosscaps (eighth manifold), inverting the crosscaps (ninth manifold), or interchanging and inverting the crosscaps (tenth manifold). Unlike the first six manifolds, these four are all nonorientable.

Flexibility: All 10 closed flat 3-manifolds have at least two degrees of flexibility: the depth of each fundamental domain (Fig. 26) may be chosen independently of its width and height. For some of the manifolds, for example the one-sixth turn manifold, this is the only flexibility. Others have greater flexibility: for example, the cross section of the half turn manifold may be any parallelogram, and the fundamental domain for the 3-torus may be any parallelepiped.

Symmetry: All closed flat 3-manifolds except the Hantzsche–Wendt manifold have a continuous 1-parameter family of symmetries that pushes into the page in the fundamental domains in Fig. 26. These symmetries take each cross section to a cross section further along, with the cross sections that fall off the back of the stack reappearing cyclically at the front. Some of the manifolds have additional continuous families of symmetries as well. The 3-torus is the most symmetrical, with a 3-parameter family of symmetries taking any point to any other point. All 10 manifolds have discrete symmetries as well, which vary from case to case and may depend on the manifold's exact shape as well as its topology. For example, a cubical 3-torus has a 3-fold symmetry defined by a one-third turn about one of the cube's long diagonals, but a 3-torus made from an arbitrary parallelepiped lacks this symmetry.

Dirichlet domain: The 3-torus is exceptional in that its Dirichlet domain does not depend on the choice of basepoint. It is always the same, no matter what basepoint you choose, because there is a symmetry taking any point to any other point. For the nine remaining closed flat 3-manifolds, the Dirichlet domain always depends on the choice of basepoint.

Observational status: All 10 closed flat 3-manifolds are of course completely consistent with recent evidence of a flat observable universe. However, if the fundamental domain is larger than our horizon radius, the topology may, in practice, be unobservable.

3.2.3. Spherical 3-dimensional manifolds

Manifolds: This section requires only a minimal understanding of the 3-sphere, or hypersphere, defined as the set of points one unit from the origin in Euclidean 4-space. Readers wanting to learn more about the 3-sphere may refer Refs. [31] for an elementary exposition or Ref. [23] for a deeper understanding.

Spherical manifolds fall into two broad categories. The manifolds in the first category have very different properties from those in the second.

First category: The first category manifolds are most easily defined in terms of their groups of covering transformations, although it is not hard to work out their Dirichlet domains afterwards. Amazingly enough, their groups of covering transformations—which are symmetries of the 3-sphere—come directly from finite groups of symmetries of the ordinary 2-sphere. In other words, each finite group

of symmetries of the 2-sphere (that is, the cyclic, dihedral, tetrahedra, octahedral, and icosahedral groups) gives rise to the group of covering transformations of a spherical 3-manifold.

The bridge from the 2-sphere to the 3-sphere is provided by the quaternions. The quaternions are a 4-dimensional generalization of the familiar complex numbers. But while the complex numbers have a single imaginary quantity i satisfying $i^2 = -1$, the quaternions have three imaginary quantities $i, j,$ and k satisfying

$$i^2 = j^2 = k^2 = -1 , \quad ij = k = -ji , \quad jk = i = -kj , \quad ki = j = -ik . \tag{10}$$

The set of all quaternions $a + bi + cj + dk$ (for a, b, c, d real) defines 4-dimensional Euclidean space, and the set of all unit length quaternions, that is, all quaternions $a + bi + cj + dk$ satisfying $a^2 + b^2 + c^2 + d^2 = 1$, defines the 3-sphere.

Before proceeding, it is convenient to note that there is nothing special about the three quaternions $i, j,$ and k.

Lemma 3.2.1 (Quaternion change of basis). *Let c be a 3×3 orthogonal matrix, and define*

$$i' = c_{ii}i + c_{ij}j + c_{ik}k , \quad j' = c_{ji}i + c_{jj}j + c_{jk}k , \quad k' = c_{ki}i + c_{kj}j + c_{kk}k , \tag{11}$$

then i', j', and k' satisfy the usual quaternion relations

$$i'^2 = j'^2 = k'^2 = -1 , \quad i'j' = k' = -j'i' , \quad j'k' = i' = -k'j' , \quad k'i' = j' = -i'k' . \tag{12}$$

Lemma 3.2.1 says that an arbitrary purely imaginary quaternion $bi + cj + dk$ may, by change of basis, be written as $b'i'$. If the purely imaginary quaternion $bi + cj + dk$ has unit length, it may be written even more simply as i'. A not-necessarily imaginary quaternion $a + bi + cj + dk$ may be transformed to $a' + b'i$. If it has unit length it may be written as $\cos(\theta) + i'\sin(\theta)$ for some θ.

The unit length quaternions, which we continue to visualize as the 3-sphere, may act on themselves by conjugation or by left-multiplication.

Lemma 3.2.2 (Conjugation by quaternions). *Let q be an arbitrary unit-length quaternion. According to the preceding discussion, we may choose a basis $\{1, i', j', k'\}$ such that $q = \cos(\theta) + i'\sin(\theta)$ for some θ. It is trivial to compute how q acts by conjugation on the basis $\{1, i', j', k'\}$:*

$$(\cos(\theta) + i'\sin(\theta))1(\cos(\theta) - i'\sin(\theta)) = 1 ,$$

$$(\cos(\theta) + i'\sin(\theta))i'(\cos(\theta) - i'\sin(\theta)) = i' ,$$

$$(\cos(\theta) + i'\sin(\theta))j'(\cos(\theta) - i'\sin(\theta)) = j'\cos(2\theta) + k'\sin(2\theta) ,$$

$$(\cos(\theta) + i'\sin(\theta))k'(\cos(\theta) - i'\sin(\theta)) = -j'\sin(2\theta) + k'\cos(2\theta) . \tag{13}$$

We see that conjugation always fixes 1 (the north pole), so all the action is in the equatorial 2-sphere spanned by $\{i', j', k'\}$. The equatorial 2-sphere itself rotates about the i'-axis through an angle of 2θ.

Compare the preceding action by conjugation to the following action by left multiplication.

Lemma 3.2.3 (Left multiplication by quaternions). *Let q be an arbitrary unit length quaternion. According to the preceding discussion, we may choose a basis $\{1, i', j', k'\}$ such that $q = \cos(\theta) + i' \sin(\theta)$ for some t. It is trivial to compute how q acts by left multiplication on the basis $\{1, i', j', k'\}$.*

$$(\cos(\theta) + i' \sin(\theta))1 = 1\cos(\theta) + i' \sin(\theta) ,$$

$$(\cos(\theta) + i' \sin(\theta))i' = -1\sin(\theta) + i' \cos(\theta) ,$$

$$(\cos(\theta) + i' \sin(\theta))j' = j' \cos(\theta) + k' \sin(\theta) ,$$

$$(\cos(\theta) + i' \sin(\theta))k' = -j' \sin(\theta) + k' \cos(\theta) . \tag{14}$$

We see that left multiplication rotates 1 towards i' while simultaneously rotating j' towards k'. The result is a screw motion known as a Hopf flow. What is not obvious from the above computation is that the Hopf flow is completely homogeneous—it looks the same at all points of the 3-sphere.

All the tools are now in place to convert finite groups of symmetries of the 2-sphere to finite groups of symmetries of the 3-sphere. The algorithm is as follows:

Step 1: Choose a finite group of symmetries of the 2-sphere. (The only such groups are the cyclic groups, the dihedral groups, the tetrahedra group, the octahedral group, and the icosahedral group. You might also expect a cubical and a dodecahedral group, but they coincide with the octahedral and icosahedral groups because the octahedron is dual to the cube and the dodecahedron is dual to the icosahedron.)

Step 2: Write down the quaternions whose action by conjugation gives the chosen symmetries. Note that each symmetry corresponds to two quaternions q and $-q$, because q and $-q$ act identically when conjugating, i.e. $qrq^{-1} = (-q)r(-q)^{-1}$.

Step 3: Let the quaternions found in Step 3 act by multiplication. This action defines the group of covering transformations. Note that although a quaternion's action by conjugation always has a pair of fixed points on the 2-sphere, its action by left multiplication on the 3-sphere is always fixed point free.

The preceding algorithm is most easily understood in a concrete example.

Example 3.2.4. Consider the so-called *Klein four group*, the group of symmetries of the 2-sphere consisting of half turns about the i-, j-, and k-axis. (Remember that we are in the 3-space spanned by $\{i, j, k\}$, so the i-, j-, and k-axis play the role of the traditional x-, y-, and z-axis.) By Lemma 3.2.2 a half turn about the i-axis corresponds to conjugation by the quaternions $\pm i$. That is, when $\theta = \pi/2$, the quaternion $i = \cos(\pi/2) + i \sin(\pi/2)$ acts by conjugation as a rotation about the i-axis through an angle $2\theta = \pi$. Similarly, the half turn about the j-axis corresponds to the quaternions $\pm j$, and the half turn about the k-axis corresponds to $\pm k$. The trivial symmetry corresponds to conjugation by ± 1. Thus the complete set of quaternions is $\{\pm 1, \pm i, \pm j, \pm k\}$. If we now let these quaternions act by left multiplication on the 3-sphere, they give a group of covering transformations of order 8. (Exercise for the reader: compute the action of each of those eight quaternions on the basis $\{1, i, j, k\}$.) It is not hard to see that the fundamental domain is a cube. The action of the covering transformations shows you that each face of the cube is glued to the opposite face with a one-quarter turn. Eight copies of the cube tile the 3-sphere like the eight faces of a hypercube.

The following table shows the results of applying the algorithm to each finite group of symmetries of the 2-sphere:

Symmetry group of 2-sphere	Spherical 3-manifold
Cyclic Z/n	Lens space $L(2n, 1)$
	Fundamental domain is lens-shaped solid
Dihedral D_n	Prism manifold
	Fundamental domain is prism with $2n$-gon base
Tetrahedral (order 12)	Octahedral space
	Fundamental domain is octahedron
	24 of which tile the 3-sphere
	in the pattern of a regular 24-cell
Octahedral (order 24)	Snub cube space
	Fundamental domain is snub cube
	48 of which tile the 3-sphere
Dodecahedral (order 60)	Pöincaré dodecahedral space
	Fundamental domain is dodecahedron
	120 of which tile the 3-sphere
	in the pattern of a regular 120-cell

Second category: Second category manifolds do not arise from simple left multiplication by groups of quaternions. The simplest second category manifolds are generic *lens spaces*. The group of covering transformations of a lens space is cyclic, generated by a single screw motion. For the lens space $L(p, q)$, with p and q relatively prime integers satisfying $p > q > 0$, the screw motion translates a distance $2\pi/p$ along a given geodesic, while rotating through an angle $q(2\pi/p)$ about that same geodesic. The fundamental domain is a lens-shaped solid, p copies of which tile the 3-sphere in much the same way that p sectors tile the surface of an orange. In the special case that $q = 1$, the screw motion corresponds to left multiplication by the quaternion $\cos(2\pi/p) + i\sin(2\pi/p)$, so $L(p, 1)$ is a first category manifold. Otherwise $L(p, q)$ is a second category manifold. (Warning: $L(p, p-1)$ is the mirror image of $L(p, 1)$ and therefore has the properties of a first category manifold. It corresponds to right multiplication by $\cos(2\pi/p) + i\sin(2\pi/p)$ instead of left multiplication.)

The remaining second category manifolds are defined by simultaneous left and right multiplication by quaternions. That is, the holonomy group of each such manifold is a subgroup of a product $H_1 \times H_2 \times \pm 1$, where H_1 and H_2 are finite groups of unit-length quaternions. An element $(h_1, h_2, \pm 1)$ acts according to the rule $q \to \pm h_1 q h_2^{-1}$. For details, see Section 4.4 of [23] and Sections 3 and 4 of [34].

Flexibility: Spherical 3-manifolds are rigid; they have no flexibility. (As in 2-dimensions we make the assumption that spherical geometry refers to a sphere of radius one. Otherwise all spherical manifolds would have the flexibility of rescaling—they could be made larger or smaller.)

Symmetry: First category manifolds are always globally homogeneous, that is, each has a continuous 3-parameter family of symmetries taken from any point to any other point.

Proposition 3.2.5 (Global homogeneity of first category manifolds). *For any two points P and Q in a first category manifold M, there is a symmetry of M taking P to Q.*

Proof. Think of M as the quotient of the 3-sphere under the action of a finite group of unit-length quaternions $\{g_1 = 1, g_2, \ldots, g_n\}$. Point P has n images represented by the quaternions $\{g_0 x, \ldots, g_n x\}$, and point Q has n images represented by the quaternions $\{g_0 y, \ldots, g_n y\}$, for some x and y. Right multiplication by $(x^{-1} y)$ is a symmetry of the 3-sphere taking the images of P to the images of Q. More generally, right multiplication takes every set of equivalent quaternions to some other set of equivalent quaternions, so it defines not only a symmetry of the 3-sphere but also a symmetry of the original manifold M. □

Second category manifolds have smaller continuous families of symmetries. For example, the lens space $L(p, q)$ has a 2-parameter family of continuous symmetries defined by translations along, and rotations about, its central geodesic. Both first and second category manifolds may have discrete symmetries, which depend on the manifold.

Dirichlet domain: The Dirichlet domain of a first category manifold never depends on the choice of basepoint.

Proof. By Proposition 3.2.5 the manifold has a symmetry taking any point to any other point, so the manifold looks the same at all basepoints, and the Dirichlet domain construction must yield the same result. By contrast, the Dirichlet domain for a second category manifold does depend on the choice of basepoint.

Observational status: Recent evidence of an approximately flat observable universe implies that if the universe is spherical its curvature radius may be comparable to or larger than our horizon radius. So even though all spherical manifolds are viable candidates for the topology of the universe, the simplest ones, which have large fundamental domains, would be much larger than the horizon radius and could not be observed directly. A more complicated topology, with a larger group of covering transformations and a smaller fundamental domain, would be more amenable to direct observation.

3.2.4. Hyperbolic 3-dimensional manifolds

Manifolds: In 1924 Alexander Friedmann published a paper [35] complementing his earlier model of an expanding spherical universe with a model of an expanding hyperbolic universe. It is a tribute to Friedmann's foresight that he explicitly allowed for the possibility of a closed hyperbolic universe at a time when no closed hyperbolic 3-manifolds were known. Fortunately, examples were not long in coming. Löbell published the first in 1929 [36] and Seifert and Weber published a much simpler example soon thereafter [37]. Nevertheless, during the first half of the 20th century few closed hyperbolic 3-manifolds were known. The situation changed dramatically in the mid 1970s with the work of Thurston, who showed that most closed 3-manifolds admit a hyperbolic geometry. For an overview, see Ref. [32]. To study known low-volume closed hyperbolic 3-manifolds, see Ref. [25].

Flexibility: Unlike closed hyperbolic 2-manifolds, closed hyperbolic 3-manifolds are rigid; they have no flexibility. As usual we assume curvature radius one, to suppress the trivial flexibility of rescaling.

Symmetry: Closed hyperbolic 3-manifolds never have continuous families of symmetries. Truly generic closed hyperbolic 3-manifolds appear to have no discrete symmetries either, although the symmetry groups of low-volume manifolds are typically small cyclic or dihedral groups. Nevertheless, Kojima has shown that every finite group occurs as the symmetry group of a closed hyperbolic 3-manifold [38].

Dirichlet domain: The Dirichlet domain of a closed hyperbolic 3-manifold always depends on the choice of basepoint. For carefully chosen basepoints, the Dirichlet domain may display some of the manifold's symmetries. The Dirichlet domain may also display symmetries beyond those of the manifold itself; these so-called hidden symmetries are actually symmetries of finite-sheeted covering spaces.

4. Standard cosmology and the cosmic microwave background

In standard big bang theory, the universe is created in a hot energetic state, a giant primordial soup. The plasma is opaque to light as photons scatter off hot charged particles. As the universe expands, the soup cools and some 300,000 years into our history, the cosmos cools enough so that light no longer scatters efficiently and the universe becomes transparent to radiation. This transition, known as decoupling, marks the time when the primordial radiation can free stream throughout the universe unimpeded. Decoupling happens quickly but not instantaneously and the small spread in time will introduce some complication in our attempts to measure topology as discussed in later sections. This ancient radiation filling all of space has cooled today to a mere 2.73°K, in the microwave range from which it has acquired the name the cosmic microwave background (CMB). As the CMB carries information almost from the beginning of time, it is a relic of our deep past.

The light received at the Earth has traveled the same distance in all directions since last scattering and therefore defines a sphere. This sphere is known as the surface of last scatter (SLS) and is only defined relative to a given observer. An alien civilization in a neighboring cluster of galaxies, if coincidentally also conducting observations of the CMB today, would receive photons from its own SLS as shown in Fig. 27. This will be particularly relevant to the circles method of detecting topology described in Section 6.2.1.

The famed CMB carries much coveted information about the universe. It has been hoped that all local geometric quantities such as the local curvature, the expansion rate, the nature of the dark matter, and of the luminous matter, can be extracted from this one source. Clearly though the data will leave the full set of parameters underdetermined and a great deal of model dependence is unavoidable. What exactly the CMB can tell us is being put to the test with the recent high-resolution experiments such as TOCO [39], Boomerang and MAXIMA [40,41]. These experiments are consistent with an approximately flat observable universe. If the universe is hyperbolic, its curvature radius may be comparable to or larger than the horizon radius. So even though all hyperbolic manifolds are viable candidates for the topology of the universe, they may be too large to see, therein is the challenge. The CMB can also be used to decipher the global topology as well as fitting these many
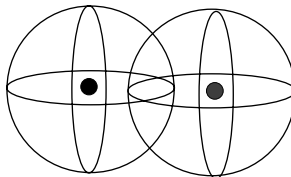
Fig. 27. Two different surfaces of last scattering, each centered on a different galaxy.

local parameters. Though statistical measures of topology are inherently model dependent, geometric methods are model independent. The pattern-based searches of circles in the sky and topological pattern formation require no assumptions about the primordial spectrum nor the nature of the dark matter. However, these searches will likely be plagued by difficulties of their own and it is unclear if these methods will succeed without a great deal of refitting.

The standard big bang theory also has a powerful successor, the standard inflationary theory. The standard inflationary theory suggests there was a period of accelerated expansion very early in the universe's history which dilutes all fluctuations, all matter, and all energy so that the cosmos suddenly goes from a tumultuous beginning to a nearly flat, empty, and biggish universe. Inflation is exited when it becomes energetically favorable for the energy driving the expansion to be released into a renewed soup of high-energy particles and light. This heating of the universe mimics the original primordial soup and the standard history of the universe resumes. There are proponents of inflation who would revise the big bang still further, essentially removing the moment of genesis altogether so that inflation is eternal [42,43]. The universe can be viewed as a ginger root of inflationary patches each with its own details. As inflation ends in any given patch, the energy is released to heat that region. As far as any given observer is concerned, the heating epoch on exit from inflation looks like a consequence of a big bang event. Since this is all we can ever observe with the CMB our discussion will apply to any of these early universe scripts.

While theorists still find it tricky to identify a specific working model of inflation and to devise a graceful exit to inflation, it is the favored story to tell and some of the reasons for this are good. Inflation elegantly resolves the mysterious homogeneity and isotropy of our observable universe. Without incredibly precise tuning of the initial conditions, the universe would tend to be extremely lumpy and would quickly cool if underdense or would quickly collapse if overdense. The observational fact that the universe is billions of years old and still just marginally curved, if not outright flat, seems special and so unsettling. A more rigorous and compelling argument is made by Guth in his original paper [44] and many subsequent articles [45]. Our homogeneous and isotropic space becomes remarkably unlikely, and yet here we are.

On the other hand, lumpiness seems natural since the many regions of the universe would not be in causal contact and therefore should have very different local properties such as temperature, density, etc. Yet when we look at the CMB we see that, in fact, the temperature appears identical to better than 1 part in 10^5 even for regions which would be separated far beyond causally connected distances. This would be like finding two ancient civilizations on opposite sides of the Earth with nearly identical languages. The civilizations must have been in causal contact. Likewise, two regions of the universe which seem to have equilibrated to precisely the same temperature must have been in causal contact. Inflation resolves this situation by taking a small causally connected region and stretching it so large that it exceeds the extent of the observable universe today. The CMB is essentially the same temperature throughout our patch of the cosmos precisely because it was in causal contact. The stretching also naturally renders the universe flat.

Another motivation for inflation is that, while it makes the universe on average homogeneous and isotropic, it also naturally generates fluctuations about this average. These minute perturbations are critical for initiating the collapse of matter into galaxies and clusters. These seemingly insignificant fluctuations become the catalysts for all the order and structure we observe in the universe. Because of the enormous stretching of scales involved, it manages to place fluctuations even on the largest possible scales and therefore gives a causal explanation for the seeds of structure formation. (Another

mechanism for generating the all important initial density fluctuations is via topological defect models although these have lately fallen into disrepute [46].)

To emphasize, the beauty is that inflation both explains the average homogeneity and isotropy as well as giving a prediction for the deviations from homogeneity, observed both in the CMB and in the form of galaxies. It seems a good question to ask, Is topology consistent with inflation? Precisely as inflation drives the curvature scale beyond observational reach, so too will it drive any topological scale beyond observational reach. The universe may still be finite, only we will never know it. Certainly, a model of inflation could be concocted which exited just at the critical moment so that the topology of the universe, and presumably the curvature too, is just within observation now, when we also happen to be here to look. However, such a contraption would be a perversion against the spirit of inflation as a restorer of naturalness. If an observable topology scale and inflation are to be conjoined, the reasons must be profound and topology must be an integral part of the inflationary mechanism. This is conceivable. For instance, the cosmic Casimir effect is the contribution to the vacuum energy due to any topological boundary. If this effect were to dilute as the topology scale expanded then one can envision inflation driven by this contribution to the vacuum and ending precisely as the topology scale became some natural big size and the vacuum energy became some correspondingly natural small size. To really wrap it all into one clean package one could even try to explain the seemingly unnatural $\Omega_\Lambda = 0.7$ this way [8] (see also Refs. [47,48]). Attempts to use compact topology to aid $\Omega < 1$ inflation have also been suggested [49]. A related mechanism has been put to use in a model with topologically compact extra dimensions [50]. Another recent argument to justify flatness and local isotropy with topology and without inflation can be found in Ref. [51]. Barrow and Kodama found that compact topology can severely restrict the anisotropies which an infinite universe will allow. Roughly speaking, the anisotropic modes will not fit in the compact space.

In any case, the inflationary predictions for the statistical distribution of fluctuations are often used in fixing the initial fluctuations and so will be relevant here. This assumption is a weakness of all statistical constraints on topology and we will discuss these issues in due course. The topic for this section is really, regardless of their origin, how do fluctuations in spacetime translate into temperature fluctuations in the CMB. And once this has been established, how does topology alter the standard predictions leaving an archeological imprint for us to uncover billions of years later.

4.1. Standard cosmological equations

This section provides a quick pedestrian review of the CMB and theories for the generation of its perturbations. There is nothing specific to topology in this section unless explicitly stated. Many detailed reviews on the CMB have been written and the reader is referred to these [52–55]. Observing topology in the CMB is discussed in Sections 4.4, 5 and 6.

The gravitational field in general relativity is determined by the local Einstein equation

$$G_{\mu\nu} = 8\pi G T_{\mu\nu} \ . \tag{15}$$

The tensor $G_{\mu\nu}$ describes the curvature and evolution of the metric $g_{\mu\nu}$ while $T_{\mu\nu}$ is the energy-momentum tensor and accounts for the matter fields. (For thorough discussions on general relativity see [56–58].) It is important to note here that the Einstein equations are local and therefore only determine the curvature of spacetime and do not fix the topology. Furthermore, the symmetries

of $g_{\mu\nu}$ are also local and are nearly always broken in the manifold by topological identifications. Today the universe appears to be homogeneous and isotropic on the largest scales. The CMB gives tremendous confirmation of these symmetries since the temperature appears identical in every direction to better than 1 part in 10^5. Homogeneity and isotropy imply that the Earth is not in a privileged position in the cosmos. The symmetries of homogeneity and isotropy severely restricts the class of possible solutions to the Friedman–Robertson–Walker (FRW) models defined by the gravitational metric

$$
\begin{aligned}
\mathrm{d}s^2 &= g_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu \\
&= -\mathrm{d}t^2 + a^2(\eta)\left[\frac{\mathrm{d}r^2}{1 - \kappa r^2} + r^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2)\right] ,
\end{aligned}
\tag{16}
$$

where $\kappa = 0, -1, 1$ corresponds to a flat, negatively curved (hyperbolic) and positively curved (elliptical) space, respectively (see Fig. 1). Negatively curved space is often termed "open" and positively curved space "closed" in reference to their simply connected geometries. Since we are interested in constructing compact, multiconnected spaces, we will avoid the "open", "closed" terminology.

The overall scale factor, $a(\eta)$, describing the expansion of the universe is determined by the energy of matter through the remaining Einstein equations. We can always operate in comoving units where the manifold is treated as a static constant curvature manifold and all of the dynamics is hidden in the conformal scale factor $a(\eta)$. An alternative expression for the metric is

$$
\mathrm{d}s^2 = a^2(\eta)[-\mathrm{d}\eta^2 + \mathrm{d}\sigma^2]
\tag{17}
$$

with the spatial part of the metric

$$
\mathrm{d}\sigma^2 = \mathrm{d}\chi^2 + f(\chi)(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2) ,
\tag{18}
$$

where we have used conformal time $\mathrm{d}\eta = \mathrm{d}t/a(t)$ and

$$
f(\chi) = \begin{cases}
\chi^2, & r = \chi & \text{flat} , \\
\sinh^2\chi, & r = \sinh\chi, & \text{hyperbolic} , \\
\sin^2\chi, & r = \sin\chi, & \text{spherical} .
\end{cases}
\tag{19}
$$

See also Appendix A.

The dynamical evolution of the space is given by the Einstein equation determining the scale factor

$$
H^2 + \kappa/a^2 = \frac{8\pi G}{3}\,\rho .
\tag{20}
$$

The different curvatures correspond to different values of the global energy density. Traditionally, $\Omega = \rho/\rho_c$ is defined with ρ_c the critical density required to render the universe flat, so that $\Omega > 1$ corresponds to $\kappa = 1$, $\Omega = 1$ corresponds to $\kappa = 0$, and $\Omega < 1$ corresponds to $\kappa = -1$. The curvature radius is $R_{\mathrm{curv}} = a/|\kappa|^{1/2}$. From Eq. (20), $8\pi G\rho_c/3 = H^2$ so that $H^2 a^2(\Omega - 1) = \kappa$ and the curvature radius can be expressed in terms of Ω as

$$
R_{\mathrm{curv}} = \frac{1}{H|\Omega - 1|^{1/2}} .
\tag{21}
$$

When working in comoving units we take the comoving curvature radius to be unity for curved space or ∞ for flat space.

Conservation of energy requires

$$\dot{\rho} + 3H(\rho + p) = 0 \tag{22}$$

for the perfect fluid energy-momentum tensor $T_\mu^\nu = (\rho, p, p, p)$. Decoupling occurs during the matter dominated error for which $p = 0$ and so $\rho \propto 1/a^3$ by Eq. (22). The solution for $a(\eta)$ during matter domination is

$$a(\eta) \propto \begin{cases} \cosh \eta - 1, & \kappa = -1 , \\ \eta^2/2, & \kappa = 0 , \\ 1 - \cos \eta, & \kappa = 1 \end{cases} \tag{23}$$

with the conformal Hubble expansion $\mathscr{H} = a'/a = aH$,

$$\mathscr{H} \propto \begin{cases} \sinh \eta/(\cosh \eta - 1), & \kappa = -1 , \\ 2/\eta, & \kappa = 0 , \\ \sin \eta/(1 - \cos \eta), & \kappa = 1 . \end{cases} \tag{24}$$

In comoving units, the universe is static. In physical units, the universe expands and the cosmic background radiation redshifts. The physical wavelength is shifted according to

$$\frac{\lambda(t)}{\lambda_{\text{initial}}} = \frac{a(t)}{a(t_{\text{initial}})} , \tag{25}$$

from which the cosmological redshift is defined as

$$1 + z = \frac{a_0}{a(t)} . \tag{26}$$

A subscript 0 will always be used to denote values today. The value of η during matter domination can be written in terms of Ω_0 and the redshift using $\eta = \int \mathrm{d}t/a = \int \mathrm{d}a/(\dot{a}a)$ and Eq. (20) [59]:

$$\eta = \begin{cases} \operatorname{arccosh}\left(1 + \dfrac{2(1 - \Omega_0)}{\Omega_0(1 + z)}\right), & \Omega_0 < 1 , \\[2ex] \dfrac{3}{(1 + z)^{1/2}}, & \Omega_0 = 1 , \\[2ex] \cos^{-1}\left(1 - \dfrac{2(\Omega_0 - 1)}{\Omega_0(1 + z)}\right), & \Omega_0 > 1 . \end{cases} \tag{27}$$

If curvature or topology are to be observable, the corresponding scales must be smaller than the diameter of the SLS. In negatively curved manifolds, the length of the shortest orbit does tend to be comparable to R_{curv} although the volume of the compact space tends to grow with topological complexity. By contrast, in positively curved manifolds, the volume of compact space tends to shrink with topological complexity as does the length of the shortest geodesic. Consequently, we might expect to see topology in a positively curved manifold regardless of the curvature radius [60]. In flat space, there is no natural scale since $R_{\text{curv}} = \infty$ and it would seem a random coincidence if the topology scale just happened to be observable. In curved space, one might expect the curvature and topology scales to be comparable. This may not make it any more natural to observe geometric effects, but it would mean that if curvature is observable, then topology may be also.

For the purpose of observational topology it is useful to list some relevant scales. The radius of the surface of last scatter is defined as the distance light travels between the time of decoupling and today,

$$D_\gamma = a(t) \int_{t_d}^{t_0} \frac{\mathrm{d}t}{a(t)} \ . \tag{28}$$

The distance depends on the curvature and evolution of the universe through Eq. (23). In comoving units the photon travels a distance

$$d_\gamma = \frac{D_\gamma}{a(t)} = \int_{t_d}^{t_0} \frac{\mathrm{d}t}{a(t)} = \int_{\eta_d}^{\eta_0} \mathrm{d}\eta = \Delta\eta \tag{29}$$

with $\Delta\eta \equiv \eta_0 - \eta_d$. The diameter of the SLS, $2\Delta\eta$, essentially defines the extent of the observable universe. A loose criterion that can be used to gauge if topology will influence the CMB is that the in-radius be less than $\Delta\eta$.

The volume enclosed by the SLS is the integral over $V_{\mathrm{SLS}} = \int \sqrt{-g}\, \mathrm{d}r\, \mathrm{d}\theta\, \mathrm{d}\phi$. For a radius of $\Delta\eta$,

$$\begin{aligned}
V_{\mathrm{SLS}} &= \pi(\sinh(2\Delta\eta - 2\Delta\eta)), \quad \Omega_0 < 1 \ , \\
&= \frac{4\pi}{3}\Delta\eta^3, \quad \Omega_0 = 1 \ , \\
&= \pi(2\Delta\eta - \sin(2\Delta\eta)), \quad \Omega_0 > 1 \ .
\end{aligned} \tag{30}$$

The number of clones of the fundamental domain that can be observed in a small compact cosmos can be estimated by the number of copies that can fit within the SLS: $V_{\mathrm{SLS}}/\mathscr{V}_{\mathscr{M}}$. We will often refer back to these scale comparisons in the coming discussions.

4.2. Fluctuations in the CMB

Regardless of the origin, there are small deviations from homogeneity and isotropy which are treated by perturbing the metric and matter tensors about the FRW solutions. The equations of motion for these perturbations to linear order are determined by $\delta G_{\mu\nu} = 8\pi G \delta T_{\mu\nu}$. While the equations are extremely complicated there is one quantity, the gauge invariant potential Φ, which is a coordinate invariant combination of components of $\delta g_{\mu\nu}$ and is of central importance for our considerations. It is so named since in the Newtonian limit it corresponds to the usual gravitational potential.

While there can be many causes of perturbations in the temperature of the CMB we restrict ourselves to the hot and cold spots caused by these primordial perturbations in the gravitational potential. The fluctuations Doppler shift the photons generating an anisotropy both at decoupling and as light traverses the time changing gravitational field. There are many carefully derived results for the temperature fluctuations in the literature based on relativistic kinetic theory and relativistic perturbation theory [61,54,55]. Since these topics comprise multiple review papers on their own, we will not go through the detailed derivations but will instead try to provide a cohesive, intuitive motivation for each of the relevant concepts.

To derive the famous Sachs–Wolfe effect [62], we followed the pedagogical discussion of White and Hu [55] (see also Ref. [63]). Photons gain energy and are therefore blueshifted as they fall into a potential well and lose energy and are therefore redshifted as they climb out. The Doppler shifts will cancel in a static homogeneous and isotropic space. However, at the time that the photons last

scatter, some photons will be in potential wells and some will be in potential peaks. Therefore, this snapshot of the metric fluctuations becomes frozen into the CMB. It is this snapshot or fossil record we observe. Energy conservation gives a Doppler shift of

$$\left.\frac{\delta T}{T}\right|_f - \left.\frac{\delta T}{T}\right|_i = \Phi_f - \Phi_i \ . \tag{31}$$

The local gravitational potential Φ_f makes an isotropic contribution to $\delta T/T$ and will be left off hereafter. By adiabaticity, $aT = \text{constant}$, it follows that

$$\left.\frac{\delta T}{T}\right|_i = -\frac{\delta a}{a} \ . \tag{32}$$

Assuming flat space for simplicity, $a \propto t^{2/3}$ during matter domination. The shift in a is then

$$\frac{\delta a}{a} = \frac{2}{3}\frac{\delta t}{t} \ . \tag{33}$$

The time shift δt can be understood as a shift relative to the cosmic time due to the perturbed gravitational potential. In a gravitational potential clocks appear to run slowly by

$$\frac{\delta t}{t} \sim \Phi \tag{34}$$

from which it follows that

$$\left.\frac{\delta T}{T}\right|_f = \frac{\Phi_i}{3} \ . \tag{35}$$

This term is the surface Sachs–Wolfe effect from the time of decoupling. There is an additional contribution to the temperature fluctuation known as the integrated Sachs–Wolfe (ISW) effect which is accumulated as the photon traverses space through a decaying potential. Including this latter contribution the full Sachs–Wolfe effect is

$$\frac{\delta T(\hat{n})}{T} = \frac{1}{3}\,\Phi(\eta_o\hat{n}) + 2\int_{\eta_{\mathrm{SLS}}}^{\eta_o} \mathrm{d}\eta\,\Phi'(\eta,\vec{x}) \ , \tag{36}$$

where a prime denotes differentiation with respect to conformal time [62]. In flat space with ordinary energy, the only contribution to $\delta T/T$ is from the surface Sachs–Wolfe effect (the first term in Eq. (36)). The ISW (the second term in Eq. (36)) makes a significant contribution to large-scale fluctuations if space is curved and/or when a cosmological constant dominates the energy density. In the absence of a significant cosmological constant, the ISW is important as perturbations decay along the line of sight during the curvature dominated epoch at

$$1 + z \sim (1 - \Omega_0)/\Omega_0 \tag{37}$$

for underdense cosmologies. This will prove to be important to the issue of topology since the ISW may camouflage the conspicuous marks of topology imprinted in the surface Sachs–Wolfe effect.

Therefore the temperature fluctuations observed today can be predicted from the primordial fluctuations in the metric. The shape of $\Phi(\eta,\vec{x})$ will depend on the geometry of space as well as some

initial spectrum. The perturbed Einstein equations give the equation of motion

$$\Phi'' + 3\mathscr{H}(1 + c_s^2)\Phi' - c_s^2\nabla^2\Phi + (2\mathscr{H}' + (1 + 3c_s^2)(\mathscr{H}^2 - \kappa))\Phi = 0 , \tag{38}$$

where $\mathscr{H} = a'/a$ and c_s is the speed of sound in the cosmological fluid and can be written as $c_s = \rho_r/(3(\rho_r + (3/4)\rho_m))$ with subscript r denoting radiation and m denoting matter. Notice that in a matter dominated era $c_s = 0$. The potential is separable and can be written $\Phi = F(\eta)\Psi(\vec{x})$ with $F(\eta)$ the solution to Eq. (38). The $\Psi(\vec{x})$ can always be expanded in terms of the eigenmodes to the Laplacian; that is,

$$\Phi(\eta, \mathbf{x}) = F(\eta) \int \mathrm{d}^3\vec{k}\,\hat{\Phi}_{\vec{k}}\psi_{\vec{k}}(\vec{x}) \tag{39}$$

with the eigenmodes ψ_k satisfying

$$\nabla^2\psi_{\vec{k}} = -k^2\psi_{\vec{k}} . \tag{40}$$

The Laplacian $\nabla^2 = g_{\mu\nu}D^\mu D^\nu$ depends on the curvature through the covariant derivatives D_μ. For now we assume all the eigenvalues $\vec{k}$ are part of a continuum although this will not be the case when the space is topologically identified. The $\hat{\Phi}_k$ are initial amplitudes given by the statistical profile of the initial fluctuation spectrum. All assumptions and/or predictions for the initial perturbations are contained in the $\hat{\Phi}_{\vec{k}}$.

During inflation quantum fluctuations in the field are amplified by the accelerated expansion. These fluctuations about the ground state of the field theory are known to be Gaussian distributed. The amplitude of the fluctuations are related to the specific inflationary model but for the most part are independent of the scale k. So inflation predicts a Gaussian distribution of fluctuations nearly independent of scale. Specifically, this means the $\hat{\Phi}_{\vec{k}}$ are drawn from a random Gaussian ensemble consistent with

$$\langle\hat{\Phi}_{\vec{k}}^*\hat{\Phi}_{\vec{k}}\rangle \propto \mathscr{P}_\Phi(k)\delta^3(\vec{k} - \vec{k}') . \tag{41}$$

The angular bracket $\langle\rangle$ denotes an ensemble average and $\mathscr{P}_\Phi$ is the predicted power spectrum. De Sitter inflation delivers a flat, Harrison–Zeldovich spectrum which corresponds to $\mathscr{P}_\Phi = $ constant. Other forms for $\mathscr{P}_\Phi \propto k^{n-1}$ have been derived from specific inflationary models with the spectral index $n \neq 1$ and give tilted as opposed to flat spectra. If an inflationary prior is used in conjunction with topology, then it is important to bear in mind that any statistical analysis will constrain topology *with this prior*. The scope of any such bound is limited. It is this weakness in assumptions that has lead some to pursue pattern-driven approaches where no prior assumptions are needed (see Sections 5.2 and 6.2).

From the preceding we can fully predict Φ and therefore $\delta T/T$ given the curvature of spacetime and the initial spectrum. Armed with this prediction, fluctuations on a given manifold can be compared to the data. For data comparison it is customary to consider the correlation function. Although $\delta T(\mathbf{x})/T$ permeates space, we only observe fluctuations from our SLS at location $\mathbf{x} = \Delta\eta\hat{n}$ with $\hat{n}$ a unit directional vector. Since the fluctuations are taken to be Gaussian, the correlation function contains all of the information about the temperature fluctuations. The ensemble average correlation

function between any two points on the sky is

$$C(\hat{n}, \hat{n}') = \left\langle \frac{\delta T}{T}(\hat{n}) \frac{\delta T}{T}(\hat{n}') \right\rangle . \tag{42}$$

The theoretically predicted $C(\hat{n}, \hat{n}')$ can then be statistically compared to the data to estimate the likelihood that a given model is responsible for the world we live in.

Because fluctuations are observed on a sphere it is customary to decompose the data into spherical harmonics,

$$\frac{\delta T}{T}(\hat{n}) = \sum_{\ell m} a_{\ell m} Y_{\ell m}(\hat{n}) , \tag{43}$$

where the $Y_{\ell m}$'s are the usual spherical harmonics. The spherical harmonics form a complete set of states on the sphere and are orthogonal so that

$$\int_0^{2\pi} d\phi \int_0^\pi \sin\theta \, d\theta Y_{\ell' m'}^*(\theta, \phi) Y_{\ell m}(\theta, \phi) = \delta_{\ell' \ell} \delta_{m' m} \tag{44}$$

and unit normalized. Using the orthogonality of the $Y_{\ell m}$, invert (43) to find

$$a_{\ell m} = \int d\Omega Y_{\ell - m}(\hat{n}) \frac{\delta T}{T}(\hat{n}) . \tag{45}$$

If the probability distribution is Gaussian, then the Fourier correlation function

$$C_\ell = \langle |a_{\ell m}|^2 \rangle \tag{46}$$

contains complete information about the fluctuations. The $\langle \rangle$ again denotes an ensemble average. Either non-Gaussianity or the breaking of homogeneity and isotropy will mean that the C_ℓ's do not provide complete information. Topological identifications nearly always break homogeneity and isotropy and so the C_ℓ's are not a complete description of the sky in a compact space as we will emphasize in Section 6.

Assuming for now a simply connected topology, homogeneity and isotropy of the metric allows one to perform an angular average which is in effect an average over m's without loss of information. Such an average gives an estimator for the ensemble average angular power spectrum C_ℓ,

$$C_\ell = \sum_{m=-\ell}^{m=\ell} |a_{\ell m}|^2 / (2l + 1) . \tag{47}$$

Using Eq. (45) gives

$$C_\ell = \frac{1}{(2l+1)} \sum_m \left[\int d\Omega' Y_{\ell - m}(\hat{n}') \int d\Omega Y_{\ell m}(\hat{n}) \left\langle \frac{\delta T}{T}(\hat{n}) \frac{\delta T}{T}(\hat{n}') \right\rangle \right] . \tag{48}$$

The parameter ℓ can be associated with the angular size of a given fluctuation, with $\ell \sim \pi/\theta$. Large angle fluctuations are associated with low ℓ and small angle fluctuations are associated with high ℓ. The lower the value of ℓ, the fewer contributions there are to sum (47). As a result, the estimator of the true ensemble average is poorer for low ℓ than it is for high ℓ where there are many contributions to the average. This is known as cosmic variance and can be included as an error bar. For a homogeneous, isotropic space cosmic variance can be estimated as

$$C_\ell \sqrt{2/(2\ell + 1)} . \tag{49}$$

We only have one universe available to measure and we cannot be sure that our universe is not just a randomly large deviation from average. It is difficult to interpret the significance of low ℓ observations such as the lack of power in the quadrupole as observed by COBE–DMR. The ambiguity of cosmic variance is worsened with topological identifications as emphasized in [10] and explained in Section 4.4.

These are all the tools needed to determine $\delta T/T$ for a specific theoretical model, the correlation function $C(\hat{n}, \hat{n}')$, and the angular average power spectrum C_ℓ. These predictions can then be compared with data. When we come to the influence of topology we will see the increasing importance of using full sky maps of $\delta T/T$ and the full correlation function $C(\hat{n}, \hat{n}')$.

We have described the scalar modes above. There are also tensor modes which are gravitational waves and vector modes which are rotational perturbations and do not grow with time. The scalar modes are the only ones which couple to the energy density and pressure and we restrict ourselves to scalar modes hereafter. Topology may, in fact, influence the gravitational wave background but to date this remains unexplored territory.

The primordial fluctuations catalyze the formation of galaxies and galaxy clusters. So the minute quantum fluctuations are amplified into the gigantic structures we see today. The primordial spectrum can be tested by measuring the fluctuations in the CMB but there are other manifestations. For instance, the eventual nonlinear growth of perturbations can be simulated numerically to test theoretical predictions for structure formation against astronomical observations. It would be interesting to know if topological features could sculpt the distribution of structure as well [64].

In the following subsections we find the eigenmode decomposition in simply connected spaces. Perpetuating a cruel prejudice against $\mathbf{S}^3$ we consider only $\mathbf{H}^3$ and $\mathbf{E}^3$. However, the authors of Ref. [34] catalog all of the $\mathbf{S}^3$ topologies and discuss detection strategies based on the crystallographic methods. The motivation for neglecting $\mathbf{S}^3$ topologies in connection with the CMB is observational as there has been no outstanding evidence which supports an $\Omega > 1$ universe [65]. The current debate is whether $\Omega = 1$ or whether $\Omega < 1$. The future satellites which aim to refine measurements of the CMB intend to resolve this debate. The overriding theoretical prejudice is for $\Omega = 1$ consistent with inflation. More recently the supernova data has drummed up enthusiasm for $\Omega_\Lambda = 0.7$ and $\Omega_{\mathrm{m}} = 0.3$ so that the total $\Omega = 1$ [66]. The recent small angle experiments Boomerang and MAXIMA corroborate these values if certain prior assumptions constrain the statistical analysis [40,41,67]. More quantitatively, the current numbers are $\Omega = 1.11 \pm 0.07$, $\Omega_{\mathrm{b}} h_0^2 = 0.032^{+0.005}_{\pm -0.004}$, and a spectral index of $n = 1.01^{+0.09}_{-0.07}$ [41] where the subscript b denotes the baryonic contribution. Oddly enough these latest numbers put $\Omega > 1$ but again the $\Omega \leqslant 1$ cases are within experimental error. Still, the data are clearly underdetermined and it is an unresolved issue how many different models can match the data.

In Ref. [68] an argument is made for a constraint on Ω_0 with some model independence. They argue that a constraint is imposed if one requires that a local maximum detected in the correlation function of large-scale structure ($\sim 100 - 200 h^{-1}$ Mpc) occurs at the same comoving positions at different redshifts. The argument is independent of both CMB and supernova Ia data, and favors a hyperbolic universe. Their result is a cosmology with $\Omega_0 = 0.9 \pm 0.15$ (95% confidence) (with Ω_Λ in the vicinity of ~ 0.65).

Instead of entering further into the parameter estimation debate, we will review the recent investigations in cosmic topology which have focused on the flat and the hyperbolic manifolds. For an interesting catalog of topologies and for more general Bianchi classes see Ref. [51].

4.3. Observing the CMB

Many billions of years after last scattering, the COsmic Background Explorer (COBE) satellite is built to refine the earlier detections by Penzias and Wilson of the microwave background radiation. COBE was launched to measure the spectrum of the radiation and search for fluctuations. The all sky map generated by the COBE satellite confirmed that the average temperature was homogeneous and isotropic at 2.728 K and that the spectrum was extremely thermal, a result that drew spontaneous applause when presented in 1992. The data also confirmed the theoretical expectation that there are in fact minute fluctuations as predicted by inflation at the 10^{-5} level that appear to be scale invariant (Fig. 28). COBE measures fluctuations on large scales, $\ell \leqslant 30$. Low ℓs correspond to fluctuations on scales far outside the horizon at the time of decoupling. So COBE which measures $\ell \lesssim 30$ is a probe of the largest geometric features and therefore is quite important as a probe of topology as well. Causal microphysics becomes important for $\ell \gtrsim 100$. At these large ℓ our calculation of $\delta T/T$ in Eq. (36) is insufficient as it neglects microphysical effects and a more detailed analysis is required. Of particular importance are the high ℓ Doppler peaks. Before recombination, sound waves in the baryon–photon fluid induce additional Doppler shifts which produce a peak in power on a scale related to the size of the horizon at the time of decoupling. The height of the peaks depends on the specific model parameters. In a standard inflationary model, the location of the first peak depends primarily on curvature, $\ell_{\text{peak}} \simeq 220\Omega_0^{-1/2}$, with some small H_0 dependence. COBE–DMR does not measure high ℓ fluctuations but two important all-sky satellites will, microwave anisotropy probe (MAP) and *Planck surveyor*. The recent balloon-borne experiments locate $\ell \sim 200$ and therefore contribute more evidence in favor of $\Omega_0 \sim 1$.

The task of the differential microwave radiometer (DMR) experiment on the COBE satellite was to measure the large angle anisotropy of the entire sky [69,70]. The temperature COBE–DMR measures in a given pixel can be written as

$$\left(\frac{\delta T}{T}\right)_i = \sum_{\ell m} a_{\ell m} B_\ell Y_{\ell m}(\vec{x}_i) + n_i \; , \tag{50}$$

where B_ℓ is the experimental beam pattern and the noise in each pixel is n_i. The DMR horns are characterized by an imperfect Gaussian beam pattern. The noise is assumed to be uncorrelated and Gaussian with mean $\langle n_i \rangle = 0$ and variance $\langle n_i n_j \rangle = \sigma \delta_{ij}$.

The experiment consisted of three pairs of antennae. Each pair measures the temperature difference in two directions separated by $60°$. Each antenna has a $7°$ beam and the data is smoothed on $10°$. A full sky scan was performed several times over 4 years. The COBE–DMR data is provided in the form of 6 maps in 3 frequencies, 31, 53 and 90 GHz [69,70]. Each map has $N_p = 6144$ of size $(2.6°)^2$. Compressing to resolution 5 pixels, $N_p = 1536$ pixels of size $(5.2°)^2$, loses no information and is sometimes implemented in data analysis.

The observed C_ℓs have been determined from the COBE–DMR data by Gorski [71], by Tegmark [72], and by Bond, et al. [73]. The monopole ($\ell = 0$) is just the average temperature itself and can be discarded from the data. The largest scale anisotropy is the dipole ($\ell = 1$) generated by our solar system's peculiar velocity relative to the CMB. Since this is not cosmological in origin, the entire dipole is discarded from the data. The first relevant cosmological observations begin with $\ell = 2$, the quadrupole, and it is curious to note that the measured quadrupole is low. This may just be cosmic variance in action but compact manifolds do happen to predict low power on large scales. This is
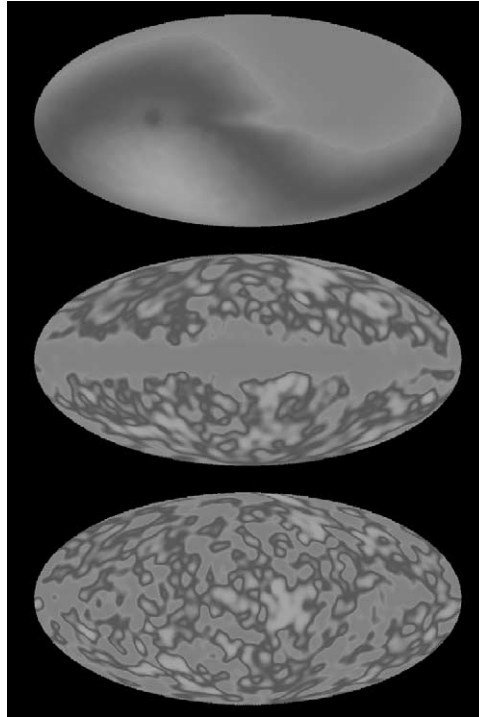
Fig. 28. The three panels are, from top to bottom, the combined data before the dipole is subtracted, the data after dipole subtraction but before the galactic cut, the map minus both dipole and galactic emission. Taken from the COBE webpage at http://space.gsfc.nasa.gov/astro/cobe/.

discussed at length in Section 5. The six maps can be compressed into one weighted-sum map with the monopole and dipole subtracted. A galactic cut of a region $\pm 20°$ around the Galactic plane is also needed to eliminate Galactic emission. This finally messaged data is ready for comparison to theoretical predictions.

The signal-to-noise of COBE–DMR is $\xi = 2$. By comparison, the projected sensitivity of MAP is $\xi = 15$ with a resolution of $0.5°$, a tenfold improvement in signal-to-noise and 30 in resolution. Planck has even higher resolution but will launch much later.

4.4. Topology and the CMB

Several aspects of the fluctuation spectrum are altered when the manifold is multiconnected. The most conspicuous and consistent signatures of topological identifications are as follows: (1) Multiconnectedness destroys global isotropy for all but the projective space and destroys global homogeneity for all but the projective space and the hypertorus. (2) The spectrum of fluctuations is discrete reflecting the natural harmonics of the finite space. (3) Since the fluctuations must fit within the finite space, there is a cutoff in perturbations with wavelengths which exceeds the topological scale in a given direction. (4) Geometric patterns are encrypted in the spatial correlations. The patterns reflect the multiple lines of sight to the topologically lensed hot and cold spots.

The proposed methods of scanning the CMB for evidence of topology can be split into direct statistical methods and geometric methods. The direct methods begin with a theoretical model, compute the temperature fluctuations as was done in Sections 5.1.1 and 6.1.1 for the simply connected spaces, and determine the likelihood of the model parameters against the data. There are shortcomings with this brute force approach. All conclusions are model dependent and there are an infinite number of models. The model is not just the manifold but also the orientation and the location of the Earth and the initial primordial spectrum. While some conclusions can still be drawn there is an appeal to a model independent attempt at observing topology. That is where the geometric methods come in. The geometric methods treat topological lensing much like gravitational lensing. One just observes the lensed images and their distribution to reconstruct the geometry of the intervening 'lens', in this case the shape of space. The most amusing example is the correlated circles of [3–5] Sections 6.2.1 although other patterns can emerge as well [6–8] as described in Section 6.2.2.

Before preceding to review the known approaches to observing topology it is worth expanding on point (1) above. The global anisotropy and inhomogeneity means that the correlation function depends on the location of the observer and the orientation of the manifold. Additionally, $C(\hat{n}, \hat{n}')$ depends on both $\hat{n}$ and $\hat{n}'$ and not just the angle between them. As a result, the angular average performed in the definition of the multipole moments C_ℓ discards important topological information. The information lost in the angular average can be quantified by an enhanced cosmic variance $\langle C_\ell^2 \rangle$ [10]. The correlation function can be decomposed into isotropic and anisotropic pieces,

$$C(\hat{n}, \hat{n}') = C^{\mathrm{I}}(\hat{n}, \hat{n}') + C^{\mathrm{A}}(\hat{n}, \hat{n}') \ . \tag{51}$$

By isotropy

$$C^{\mathrm{I}}(\hat{n}, \hat{n}') = \sum_\ell \frac{\ell + 1/2}{\ell(\ell+1)} C_\ell P_\ell(\hat{n} \cdot \hat{n}') \tag{52}$$

and by anisotropy

$$C^{\mathrm{A}}(\hat{n}, \hat{n}') = \int \mathrm{d}\Omega_{\hat{n}} \int \mathrm{d}\Omega_{\hat{n}'} C^{\mathrm{A}}(\hat{n}, \hat{n}') P_\ell(\hat{n} \cdot \hat{n}') = 0 \ , \tag{53}$$

so that the anisotropic piece is orthogonal to the Legendre polynomials. The expectation value of the estimator for the C_ℓ depends solely on the isotropic piece by construction

$$\langle \tilde{C}_\ell \rangle = \frac{\ell(\ell+1)}{8\pi^2} \int \mathrm{d}\Omega_{\hat{n}} \int \mathrm{d}\Omega_{\hat{n}'} C(\hat{n}, \hat{n}') P_\ell(\hat{n} \cdot \hat{n}') \ , \tag{54}$$

but the variance

$$\mathrm{var}(\tilde{C}_\ell) = \langle \tilde{C}_\ell^2 \rangle - \langle \tilde{C}_\ell \rangle^2 \tag{55}$$

contains anisotropic pieces. This can be interpreted as very large error bars due to cosmic variance. As a result, conclusions based on the C_ℓ's alone are weak [9,10]. Many of the statistical analyses described below do only examine the C_ℓs and so can only rule out a topology but never really confirm topology.

However, others have argued that this increased error is not very large for compact hyperbolic spaces. An argument by Inoue, for instance, expresses the variance as the sum of a geometric variance and the usual cosmic variance. The geometric variance contains the additional variance due to topology and is intrinsically small [74]. The smallness of the geometric variance can be traced

to the randomness of the eigenmodes on compact hyperbolic models (explored further in Section 6.1.3) which generates what Inoue refers to as geometric Gaussianity [75].

5. Observing flat topologies in the CMB

Since the flat spaces are most easily constrained using the COBE–DMR data we begin with them and discuss hyperbolic manifolds in all their glory separately. While small universes can be ruled out, it is rather fascinating to note that large cases are marginally consistent with the data. After all, the observed quadrupole is low and so the infrared truncation in power seen in flat spaces and described below can actually create a better fit to the data. However, such a marginal flat space, just coinciding with the size of the observable universe is unaesthetic at best. In fact, there is no natural scale for the size of a flat universe and they are not the favored small universe candidates for this reason. Still, they provide an important and accessible testing ground for methods of observation and we discuss them next.

5.1. Direct methods in flat space

5.1.1. Simply connected flat space

In flat, simply connected space, the spectrum of fluctuations in an inflationary universe is well known. We need to find the elements of Eq. (39) with $F(\eta)$ the solution to Eq. (38). In flat space $\mathscr{H} = 2/\eta$ and the solution to Eq. (38) during matter domination when $c_s = 0$ decays as $\Phi' \propto \eta^{-6}$. (During radiation domination, $c_s \neq 0$; however, it can easily be shown that for long-wavelength modes, $F(\eta) \sim$ const. See for instance Ref. [52].) Consequently, in flat space $F(\eta)$ is effectively constant and there is no ISW effect during radiation or matter domination. (There is, however, an ISW effect if the universe is dominated by a cosmological constant. This may turn out to be important in salvaging finite flat models if the universe is accelerating as the recent supernovae observations indicate [66].) The Laplacian in flat space is simply

$$(\partial_x^2 + \partial_y^2 + \partial_z^2)\psi_{\vec{k}} = -k^2\psi_{\vec{k}} \tag{56}$$

with solutions

$$\psi_{\vec{k}} = \exp(i\vec{k} \cdot \mathbf{x}) \ . \tag{57}$$

The potential of Eq. (39) can then be expanded in terms of these as

$$\Phi(\vec{\mathbf{x}}) = \int_{-\infty}^{\infty} \mathrm{d}^3k \, \hat{\Phi}_{\vec{k}} \exp(i\vec{k} \cdot \mathbf{x}) \ . \tag{58}$$

All of the assumptions about the statistics and shape of the spectrum are contained in the $\hat{\Phi}_{\vec{k}}$. All other quantities are determined by the geometry of the space. In accordance with the inflationary prediction the $\hat{\Phi}_{\vec{k}}$ are assigned an independent Gaussian probability distribution consistent with the flat space normalization [52],

$$\langle \hat{\Phi}_{\vec{k}} \hat{\Phi}_{\vec{k}'}^* \rangle = \frac{2\pi^2}{k^3} \mathscr{P}_\Phi(k) \delta^3(\vec{k} - \vec{k}') \ . \tag{59}$$

From Eq. (36) the Sachs–Wolfe effect is simply

$$\frac{\delta T(\hat{n})}{T} = \frac{\Phi(\hat{n})}{3} \ . \tag{60}$$

The correlation function between any two points on the SLS is then, from Eq. (42),

$$C(\hat{n}, \hat{n}') \propto \int \frac{\mathrm{d}^3 k}{k^3} \, \mathscr{P}_\Phi \exp(\mathrm{i}\Delta\eta\vec{k} \cdot (\hat{n} - \hat{n}')) \tag{61}$$

up to a normalization. For a homogeneous and isotropic space, the correlation function depends only on the angular separation between $\hat{n}$ and $\hat{n}'$. Consequently, all the information on the theoretical sky is in the angular average. Using the orthogonality relations of the Legendre polynomials in Eq. (48) gives

$$\begin{aligned}
C_\ell &= \frac{1}{4\pi} \int \mathrm{d}\Omega \int \mathrm{d}\Omega' C(\hat{n}, \hat{n}') P_\ell(\mu) \\
&= \frac{4\pi}{25} \int_0^\infty \frac{\mathrm{d}k}{k} \, j_\ell^2(\Delta\eta k) \mathscr{P}_\Phi \ .
\end{aligned} \tag{62}$$

If we assume the initial fluctuation is a powerlaw $\mathscr{P}_\Phi \propto k^{n-1}$, this can be integrated to give [76].

$$C_\ell \propto \frac{\Gamma(\ell + (n-1)/2)}{\Gamma(\ell + (5-n)/2)} \frac{\Gamma((9-n)/2)}{\Gamma((3+n)/2)} \ . \tag{63}$$

As described in Section 4.3, the COBE–DMR data has been analyzed over the years by many groups to determine the C_ℓ's observed. Statistical likelihood analyses are then performed to compare the theoretical model with the COBE–DMR data. The infinite, flat model is generally found to be consistent with the data for a nearly flat power spectrum ($n \sim 1$). The question addressed in the next section, is how well finite flat models match the data. To flash forward, the answer is essentially that finite flat models match the data well if they are comparable in size to half the observable universe although some studies have put even stronger limits as described below. All limits are subject to caveats based on the assumptions made about the perturbation spectrum—none of which have yet been tested observationally.

5.1.2. Compact, flat spaces

We have already reviewed the standard decomposition of the temperature fluctuations in a simply connected space. When the manifold is compact and the SLS exceeds the dimensions of the space, then the global topology is reflected in the CMB sky. A classification of all of the flat orientable manifolds can be found in Ref. [20] with a correction to the construction of the Hantzsche–Wendt flat manifold in Ref. [77]. The exact eigenmodes can be found for all of the 6 compact orientable flat spaces [78,79]. This allows a direct statistical comparison of the theoretical predictions for compact, orientable spaces with the COBE–DMR data. With any statistical comparison of a model with the data, the conclusions are model dependent. The reader should bear in mind that the bounds quoted assume equal-sided spaces, $\Omega_\Lambda = 0$, and a flat Gaussian distributed power spectrum.

The earliest bounds on the hypertorus constrained the topology scale to be $\gtrsim 0.8\Delta\eta$ [80] which is still less than the diameter of the SLS. There could still be as many as eight copies of our universe within the observable horizon. The analysis was later extended to the other compact, orientable flat spaces where similar bounds were obtained [78,79]. Stronger bounds were put on the hypertorus by

comparing the full covariance matrix against the 2-year COBE–DMR data [81]. The length of a side was set to be $\gtrsim 1.2\Delta\eta$.

The most conspicuous feature in the fluctuation spectrum is a suppression of power on large scales since large fluctuations cannot fit into the finite box. Such an infrared cutoff can be deduced from Cheeger's inequality [82]

$$k_{\min} \geqslant \frac{h_C}{2}, \quad h_C = \inf_X \frac{A(S)}{\min(V(M_1), V(M_2))} \tag{64}$$

with the infimum taken over all possible surfaces S that divide the space into two subspaces $\mathcal{M}_1$ and $\mathcal{M}_2$ where $\mathcal{M} = \mathcal{M}_1 \cap \mathcal{M}_2$. S is the boundary of the two subspaces, $S = \partial\mathcal{M}_1 = \partial\mathcal{M}_2$. The isoperimetric constant h_C depends on the geometry more than the topology. Intuitively speaking, in very long thin manifolds h_C can be quite small leading to a lowered spectrum of eigenvalues. However, this is highly correlated with thin bottleneck structures and is not a common feature of spaces with a more regular shape [82,83]. From Cheeger's inequality it follows that all flat hypertori have $k_{\min} \geqslant 2/L$ with L the longest side of the torus as argued in [9].

Although the cutoff would seem a good indicator of topology, it just so happens that the longest wavelength fluctuation observed, namely the quadrupole, is in fact low. Some might even take this as evidence for topology [84]. Cosmic variance is also large on large scales. Consequently, a fundamental domain which is the size of the observable universe is actually consistent with the observed COBE–DMR C_ℓ's [78]. However, since the fundamental domain has a particular orientation on the sky, the correlation is not simply a function of the angular separation between $\hat{n}$ and $\hat{n}'$ as it is in the infinite case. Therefore conclusions based on the C_ℓs alone are weaker than a statistical analysis based on the full $C(\hat{n}, \hat{n}')$. The strongest bounds on the hypertorus were placed by comparing the full correlation function $C(\hat{n}, \hat{n}')$ to the data. It was found that $L \gtrsim 1.3\Delta\eta$ [81,9,10], which is still less than the diameter of the observable universe, $\sim 2\Delta\eta$. Asymmetric spaces were constrained by de Oliveira–Costa et al. as described in Section 5.2.

Despite the bounds placed on the equal-sided compact, flat spaces, they still provide an excellent testing ground for geometric measures of topology. They may also be saved by $\Omega_A \neq 0$ models. We compile a list of all the eigenmodes, eigenvalues and relations for the compact flat spaces for completeness. These solutions are taken from Ref. [78,79]. The spectra were also found in Ref. [80] for the other compact spaces; however, the critical relations were overlooked leading to confusion about the cutoff in the long wavelength modes. It was mistakenly concluded because of the missing relations that in the twisted spaces longer wavelengths could fit in the fundamental domain since a wave needs to wrap more than once before coming back to a fully periodic identification [5]. However, it was shown in Ref. [78,79], that these long modes are forbidden by the relations among the $\hat{\Phi}_{\vec{k}}$, and all of the compact topologies have roughly the same long wavelength cutoff.

As a result of the global topology, all of these spaces are anisotropic and all except for the hypertorus are inhomogeneous. Topology can be implemented by imposing the boundary conditions

$$\Phi_{\vec{k}}(\vec{x}) = \Phi_{\vec{k}}(g\vec{x}) \quad \forall g \in \Gamma . \tag{65}$$

Compact topology always restricts the eigenvalues to a discrete spectrum:

$$\Phi(\vec{x}) = \sum_{\vec{k}} \hat{\Phi}_{\vec{k}} e^{i\vec{k}\cdot\vec{x}} . \tag{66}$$

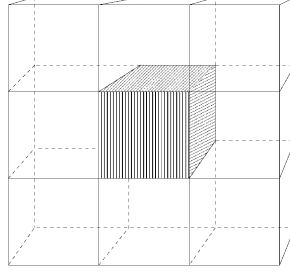In addition, relations are imposed on the $\hat{\Phi}_{\vec{k}}$.

Fig. 29. Tiling flat space with parallelepipeds.

Demonstrating the discretization explicitly for the hypertorus, the identifications of the cube (Fig. 29) are expressed in the three boundary conditions $\Phi(x, y, z) = \Phi(x + L_x, y, z) = \Phi(x, y + L_y, z) = \Phi(x, y, z + L_z)$. Imposing the first boundary condition on $\Phi(\vec{x})$ of Eq. (66) gives $e^{-ik_x x} = e^{-ik_x(x+L_x)}$ which requires $k_x = (2\pi/l_x)n_x$ with n_x an integer. The other two boundary conditions provide the discrete spectrum

$$k_x = \frac{2\pi}{L_x}n_x, \quad k_y = \frac{2\pi}{L_y}n_y, \quad k_z = \frac{2\pi}{L_z}n_z \tag{67}$$

with the n_i running over all integers [80,85,86]. It is clear that there is a minimum eigenvalue and hence a maximum wavelength which can fit inside the fundamental domain defined by the parallelepiped [80]:

$$k_{\min} = 2\pi \min\left(\frac{1}{L_i}\right), \quad \lambda_{\max} = \max(L_i) .$$

Although global anisotropy is broken by topology, we can still form the C_ℓ's to make a comparison, if incomplete, with the COBE–DMR C_ℓ's. Expanding the exponential and Legendre polynomials in terms of spherical harmonics, the C_ℓ equation (48) becomes

$$C_\ell \propto \sum_{\vec{k}} \sum_{\vec{k}'} \frac{\mathscr{P}_{\vec{k}}}{k^3} j_\ell^2(\Delta\eta k) . \tag{68}$$

Three other spacetimes are constructed from a parallelepiped. The first twisted parallelepiped has opposite faces identified with one pair of faces twisted by π before gluing. The eigenmodes are $\vec{k} = 2\pi(j/L_x, w/L_y, n/2L_z)$, with the additional relation $\hat{\Phi}_{jwn} = \hat{\Phi}_{-j-wn}e^{i\pi n}$. The minimum mode consistent with the relations is $k_{\min} = 2\pi/L$ where L is the minimum of $L_{x,y,z}$. The C_ℓ's become

$$C_\ell \propto \sum_{jwn} \frac{\mathscr{P}(k)}{k^3} \frac{j_\ell(\Delta\eta k)^2}{2\ell+1} \sum_{m=-\ell}^{\ell} |Y_{\ell,m}(\hat{\mathbf{k}})|^2 (1 + e^{i\pi(n+m)}) . \tag{69}$$

A parallelepiped with one face rotated by $\pi/2$ has discrete eigenmodes $\vec{k} = 2\pi(j/L_x, w/L_y, n/4L_z)$, with the additional relations $\hat{\Phi}_{jwn} = \hat{\Phi}_{w-jn}e^{in\pi/2} = \hat{\Phi}_{-w-jn}e^{in\pi} = \hat{\Phi}_{-wjn}e^{i3n\pi/2}$. Again, the minimum
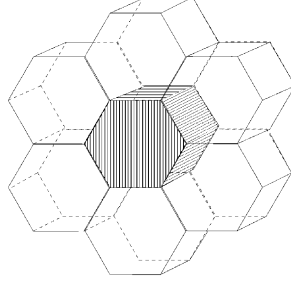
Fig. 30. Tiling flat space with hexagonal prisms.

mode is $k_{\min} = 2\pi/L$. The C_ℓ's are given by

$$C_\ell \propto \sum_{jwn} \frac{\mathscr{P}(k)}{k^3} \frac{j_\ell(\Delta\eta k)^2}{2\ell+1} \sum_{m=-\ell}^{\ell} |Y_{\ell,m}(\hat{\mathbf{k}})|^2 (1 + \mathrm{e}^{\mathrm{i}(n+m)\pi/2} + \mathrm{e}^{\mathrm{i}(n+m)\pi} + \mathrm{e}^{\mathrm{i}3(n+m)\pi/2}) \ . \tag{70}$$

The last parallelepiped has a fundamental domain of volume $2hbc$ and is thus a double parallelepiped. The identification rules involve three rotations through π as shown in [20]. They are so that $(x, y, z) \rightarrow (x + L_x, -y, -z)$, $(x, y, z) \rightarrow (-x, y + L_y, -(z + L_z))$, and $(x, y, z) \rightarrow (-(x + L_x), -(y + L_y), z + L_z)$. The resultant discrete spectrum is $\vec{k} = \pi(j/L_x, w/L_y, n/4L_z)$ with relations $\hat{\Phi}_{jwn} = \hat{\Phi}_{j-w-n} \, \mathrm{e}^{\mathrm{i}\pi j} = \hat{\Phi}_{-jw-n} \, \mathrm{e}^{\mathrm{i}\pi(w+n)} = \hat{\Phi}_{-j-wn} \, \mathrm{e}^{\mathrm{i}\pi(j+w+n)}$. The minimum mode is $k_{\min} = \sqrt{2}\pi/L$. Since the volume of the space is $2L^3$, the angular averaged long wavelength is $\sqrt{2}/2^{1/3} \simeq 1.1$ times the length of the fundamental domain. The angular power spectrum is

$$C_\ell \propto \sum_{jwn} \frac{\mathscr{P}(k)}{k^3} \frac{j_\ell(\Delta\eta k)^2}{2\ell+1} \sum_{m=-\ell}^{\ell} Y_{\ell,m}^*(\hat{\mathbf{k}})$$
$$\times (Y_{\ell,m}(\hat{\mathbf{k}})(1 + \mathrm{e}^{\mathrm{i}(m+j+w+n)\pi}) + Y_{\ell,-m}(\hat{\mathbf{k}})\mathrm{e}^{\mathrm{i}\ell\pi}(\mathrm{e}^{\mathrm{i}(m+j)\pi} + \mathrm{e}^{\mathrm{i}(w+n)\pi})) \ . \tag{71}$$

The last two possible compact flat spaces are based on a hexagonal tiling (Fig. 30). With the opposite sides of the hexagon identified and the prism faces rotated relative to each other by $2\pi/3$, the potential can be written as

$$\Phi = \sum_{n_2 n_3 n_z} \hat{\Phi}_{n_2 n_3 n_z} \mathrm{e}^{\mathrm{i}k_z z} \exp\left[\mathrm{i}\frac{2\pi}{L_x}\left[n_2\left(x - \frac{1}{\sqrt{3}}y\right) + n_3\left(x + \frac{1}{\sqrt{3}}y\right)\right]\right] \tag{72}$$

with the eigenmodes $\vec{k} = 2\pi((n_2 + n_3)/L_x, (-n_2 + n_3)/L_y, n_z/3L_z)$ and with $L_x = L_y$. The relations on this space are $\hat{\Phi}_{n_2,n_3,n_z} = \hat{\Phi}_{n_3,-(n_2+n_3),n_z} \mathrm{e}^{\mathrm{i}2\pi n_z/3} = \hat{\Phi}_{-(n_2+n_3),n_2,n_z} \mathrm{e}^{\mathrm{i}4\pi n_z/3}$ and lead to

$$C_\ell \propto \sum_{n_2 n_3 n_z} \frac{\mathscr{P}(k)}{k^3} \frac{j_\ell(\Delta\eta k)^2}{2\ell+1} \sum_{m=-\ell}^{\ell} |Y_{\ell,m}(\hat{\mathbf{k}})|^2 (1 + \mathrm{e}^{\mathrm{i}2(n_z+m)\pi/3} + \mathrm{e}^{\mathrm{i}4(n_z+m)\pi/3}) \ . \tag{73}$$

Lastly, the prism faces are glued after rotation by $\pi/3$. The potential can still be written as (72) with eigenmodes $\vec{k} = 2\pi((n_2 + n_3)/L_x, (-n_2 + n_3)/L_y, n_z/6L_z)$ and a set of relations $\hat{\Phi}_{n_2,n_3,n_z} = \hat{\Phi}_{(n_2+n_3),-n_2,n_z} \mathrm{e}^{\mathrm{i}\pi n_z/3} = \hat{\Phi}_{n_3,-(n_2+n_3),n_z} \mathrm{e}^{\mathrm{i}2\pi n_z/3} = \hat{\Phi}_{-n_2,-n_3,n_z} \mathrm{e}^{\mathrm{i}\pi n_z} = \hat{\Phi}_{-(n_2+n_3),n_2,n_z} \mathrm{e}^{\mathrm{i}4\pi n_z/3} = \hat{\Phi}_{-n_3,(n_2+n_3),n_z} \mathrm{e}^{\mathrm{i}5\pi n_z/3}$.

The C_ℓ's are given by

$$C_\ell \propto \sum_{n_2 n_3 n_z} \frac{\mathscr{P}(k)}{k^3} \frac{j_\ell(\Delta\eta k)^2}{2\ell+1} \sum_{m=-\ell}^{\ell} |Y_{\ell,m}(\hat{\mathbf{k}})|^2$$
$$\times (1 + e^{i(n_z+m)\pi/3} + e^{i2(n_z+m)\pi/3} + e^{i(n_z+m)\pi} + e^{i4(n_z+m)\pi/3} + e^{i5(n_z+m)\pi/3}) \ . \tag{74}$$

The volume of both of these topologies is $L_x L_y L_z \sqrt{3}/2$ (again $L_x = L_y$).

In summary, both the π- and $\pi/2$-twisted tori have C_ℓ's that are almost identical to that of the torus. The harmonics do show distinctions between the parallelepiped topologies, but unfortunately the variations fall well within cosmic variances. In addition to the damping at low ℓ, harmonics of the discrete spectrum create both dips and enhancements extending up to high values of ℓ. The dips are a natural consequence of the discretization and the resultant absence of certain modes in the spectrum. The enhancements are due to the distribution of multiple images dictated by the geometry of the fundamental domain.

As a result of the low observed quadrupole, the predicted cutoff alone is not enough to rule out compact, flat models. The likelihood of the C_ℓ in the full COBE–DMR range was compared to the relative likelihood of a flat, infinite cosmology in Refs. [78–80]. Compact, flat spaces are tens of times less likely than their infinite counterparts if the topology scale is about half the diameter of the SLS. Since compact topologies do not give isotropic temperature fluctuations, this lends ambiguity to any likelihood analysis. The conclusions drawn were quite conservative, ruling out $L \gtrsim 0.8\Delta\eta$; that is, the length of a side must be $\gtrsim 0.4$ the diameter of the observable universe. Again, it should be emphasized that these conclusions are contingent on the correctness of the assumptions about the initial perturbation spectrum.

While the angular power spectrum is sufficient to constrain symmetric, flat topology it is in general a poor discriminant. The average over the sky fails to recognize the strong inhomogeneity and anisotropy manifest in these cosmologies. Direct attacks on anisotropic spaces inspires more geometric approaches as discussed next.

5.2. Geometric methods in flat space

One of the first geometric, or pattern-driven searches for topology was initiated in Ref. [87]. For an anisotropic hypertorus, a symmetry plane or a symmetry axis can be identified in the CMB [47,88,89]. For these asymmetric spaces, a search for patterns was emphasized in Ref. [87]. de Oliveira Costa, Smoot, and Starobinsky develop a statistic that is independent of cell orientations but is sensitive to the plane and axis symmetries of the rectangular spaces. For anisotropic models there is a stronger dependence on the orientation of the manifold relative to the Galaxy cut. The statistic $S(\hat{n}_i)$ searches for reflection symmetries in a plane perpendicular to $\hat{n}_i$ and is defined by

$$S(\hat{n}_i) = \frac{1}{N_{\text{pix}}} \sum_{j=1}^{N_{\text{pix}}} \frac{[\delta T(\hat{n}_j)/T - \delta T(\hat{n}_{ij})/T]^2}{\sigma^2(\hat{n}_j) + \sigma^2(\hat{n}_{ij})} \ . \tag{75}$$

N_{pix} is the number of pixels after the Galaxy cut and $\sigma(\hat{n})$ is the r.m.s. error associated with the pixels in the direction $\hat{n}$. The object $\hat{n}_{ij}$ is the reflection of $\hat{n}_j$ in the plane with normal $\hat{n}_i$,

$$\hat{n}_{ij} = \hat{n}_j - 2(\hat{n}_i \cdot \hat{n}_j)\hat{n}_i \ . \tag{76}$$

Lower values of $S(\hat{n}_i)$ corresponds to a higher degree of symmetry. The temperature fluctuation is computed using the eigenmode expansion and a flat spectrum. The resultant $\delta T/T$ depends on six parameters: the three spatial orientations fixing the domain orientation and the three topology scales.

The smaller a given direction the more extreme is the asymmetry. In T^1 for instance, if the z direction is extremely small, then the size of fluctuations will be notably smaller in this direction leading to a map where there is essentially less and less structure at COBE–DMR resolution in this one direction and the structure will be drawn out on the (x, y) plane. Similarly for T^2 there will essentially only be structure along the large direction and so the pattern will appear to be rings of fluctuations along the small directions.

Using this statistic, T^1 and T^2 models were constrained to be greater than half the radius of the SLS in their small dimensions. Subsequently, the circle method was used to study specific asymmetric models [90] where again pessimistic conclusions were reached as discussed in Section 6.2.1.

6. Observing hyperbolic topologies in the CMB

The eigenmode decomposition is well known on simply connected $\mathbf{H}^3$ and is given in Section 4.2. However, when hyperbolic space is fully compact the decomposition becomes intractable. This is a stronger statement than simply saying the eigenmodes and eigenvalues are *difficult* to find. It is actually formally impossible to write down the eigenmodes analytically. The boundary conditions are so intricate they resist decomposition [28].

The absence of an analytic solution to the eigenmode spectrum can be directly related to the incipient chaos on compact manifolds. On simply connected $\mathbf{H}^3$ geodesics show the first critical ingredient for the onset of chaos; that is, extreme sensitivity to initial conditions. The geodesic deviation equation shows that two nearby trajectories diverge away from each other exponentially quickly:

$$\frac{D^2 \zeta^\mu}{ds^2} = -R^\mu_{\alpha\beta\gamma} u^\alpha \zeta^\beta u^\gamma \ , \tag{77}$$

where $u^\mu = dx^\mu/ds$ and ζ^μ are the separation of neighboring geodesics. Normal to the geodesic flow this becomes [91]

$$\frac{d^2 \|\zeta^\mu_N \zeta_{N\mu}\|}{ds^2} = -2\kappa \|\zeta^\mu_N \zeta_{N\mu}\| \ , \tag{78}$$

where κ is the curvature. If $\kappa < 0$, these geodesics diverge exponentially and a coordinate invariant Lyupanov exponent can be interpreted as $\lambda = \sqrt{|\kappa|}$. Still, there is no chaotic motion on infinite $\mathbf{H}^3$ since there is no mixing and no folding of trajectories. In other words, there is no loss of predictability as a result of the exponential deviation. By contrast, when the space is made topologically compact, the mixing and folding of trajectories is assured and the flows are well known to be fully chaotic. An entropic measure of the chaotic flow can be related to the volume of the manifold through the Kolmogorov entropy $S_K \propto \mathcal{V}^{-1/3}$ where $\mathcal{V}$ is the volume of the spacetime [92]. Note that if the space is infinite this entropy vanishes.

Chaotic flows on compact 2-dimensional manifolds in particular have been studied at great length [28,93,95,96]. The cosmological implications of chaos have only been touched upon [49,64] and remain a largely unexplored terrain. By and large, people have tried to obviate the chaotic flows entirely when analyzing the CMB. Methods include brute force numerical determination of the modes,

the method of images construction of CMB maps, and geometric methods. We will discuss a catalog of such studies.

6.1. Direct methods in hyperbolic space

6.1.1. Simply connected hyperbolic space

Since all of the numerical methods will rely on the expansion of the eigenmodes on the universal cover, we summarize those results here. From the standard expansion (39), we need to determine each of the factors in the decomposition. In negatively curved space perturbations are time dependent according to Eq. (38) with solution

$$F(\eta) = \frac{5(\sinh^2 \eta - 3\eta \sinh \eta + 4\cosh \eta + 4)}{(\cosh \eta - 1)^3} \ . \tag{79}$$

In a negatively curved space the Helmholtz equation becomes

$$\frac{1}{\sinh^2 r}\left[\partial_r(\sinh^2 r \, \partial_r) + \frac{1}{\sin^2 \theta}\partial_\theta(\sin\theta\partial_\theta) + \frac{1}{\sin^2 \theta}\partial_\phi^2\right]\psi_{\vec{k}} = -k^2\psi_{\vec{k}} \ . \tag{80}$$

(For other coordinate systems in which to express the Laplacian see Appendix A.) As emphasized in Ref. [97], a complete orthonormal basis for the simply connected space is formed by modes with real $k^2 > 1$ and so the eigenvalue range is $k^2 = [1, \infty]$. These modes vary on a scale below the curvature radius and are therefore subcurvature modes. Standard theories such as inflation will not seed fluctuations on scales $k < 1$ although some other unforeseen mechanism may. It is customary to introduce another parameter

$$q^2 = k^2 - 1 \tag{81}$$

with the range $q^2 = [0, \infty]$. The eigenmodes on $\mathbf{H}^3$ can be expressed as

$$\psi_{q\ell m} = X_{q\ell}(r)Y_{\ell m}(\hat{n}) \ , \tag{82}$$

where the $Y_{\ell m}$'s are the usual spherical harmonics and the $X_{q\ell}$ are the radial functions

$$X_{q\ell} = \frac{\Gamma(\ell + 1 + iq)}{\Gamma(iq)}\sqrt{\frac{1}{\sinh r}}P_{iq-1/2}^{-\ell-1/2}(\cosh r) \ , \tag{83}$$

where Γ denotes the Gamma function and P denotes the Legendre functions [97].

The $\hat{\Phi}_q$ are drawn from a Gaussian probability distribution with proper normalization for hyperbolic space

$$\langle\hat{\Phi}_{q\ell m}\hat{\Phi}_{q'\ell'm'}^*\rangle = \frac{2\pi^2}{q(q^2+1)}\mathcal{P}_\Phi(q)\delta(q-q')\delta_{\ell\ell'}\delta_{mm'} \ . \tag{84}$$

We can summarize the Sachs–Wolfe effect as

$$\frac{\delta T(\hat{n})}{T} = \int \mathrm{d}^3\vec{k}\,\hat{\Phi}_{\vec{k}}(\hat{n})L_{\vec{k}} \tag{85}$$

with both the surface Sachs–Wolfe and the integrated Sachs–Wolfe accounted for in

$$L_{\vec{k}} = \left[\frac{1}{3}F(\eta_{sls}) + 2\int_{\eta_{sls}}^{\eta_0}\mathrm{d}\eta F'(\eta)\right]\psi_{\vec{k}}(\vec{x}) \ . \tag{86}$$
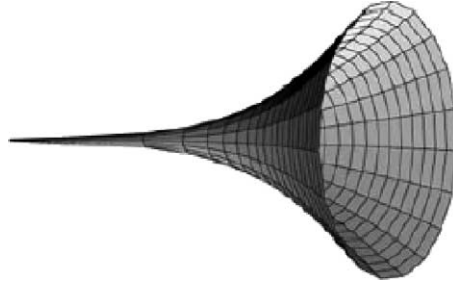
Fig. 31. Embedding of the cusp topology with one of the dimensions suppressed.

The Fourier multipoles are

$$C_\ell = 2\pi^2 \int_1^\infty \frac{\mathrm{d}k}{k(k^2-1)} \mathscr{P}_\Phi(k) \left[\frac{1}{5} X_{k\ell}(\Delta\eta) + \frac{6}{5} \int_{\eta_d}^{\eta_0} \mathrm{d}r\, X_{k\ell}(r) F'(\Delta\eta - r) \right]^2 . \tag{87}$$

The COBE–DMR data alone tends to favor $\Omega \sim 0.3$–0.4 (see for instance [98]) although both very low Ω and an $\Omega \sim 1$ are still compatible. For a long time, the COBE results along with other astronomical observations [99] put low Ω cosmologies in the limelight. However, more recent high ℓ observations [40,41] in conjunction with the supernovae data [66] have pushed public opinion back towards $\Omega = 1$ cosmologies. The recent high ℓ observations only examine small patches of the sky. A full sky map is required to advance topology observations and so we reserve more commentary until the launch and hopeful success of the future MAP and *Planck* missions.

6.1.2. Cusps

There are multiconnected hyperbolic spaces which are not completely compact (see Ref. [2] for some examples such as those of Refs. [100,101]). For some of these the motion is not chaotic and it is possible to find the eigenmodes. One topology of particular interest is the toroidal horn [102,103]. The horn is interesting since many manifolds have cusped corners (Fig. 31) as stressed by [5]. Despite their frequency in a set of generic manifolds, it was argued in Ref. [5] that according to the thick–thin decomposition, it would very improbable for us to live in the thin part of the manifold. This argument is flawed. Since the cusp narrows exponentially, an observer can live in a fat part of the manifold and still see photons coming from a constricted part of the cusp. Olson and Starkman have developed an approximation scheme to study cusps on full compact manifolds and show the generic patterns do in fact persist [104].

The cusp is most easily understood in the upper half-space representation of $\mathbf{H}^3$ (see Appendix A)

$$\mathrm{d}s^2 = -\mathrm{d}\eta^2 + \mathrm{d}z^2 + \mathrm{e}^{-2z}(\mathrm{d}x^2 + \mathrm{d}y^2) \tag{88}$$

with the coordinate transformation $\mathrm{e}^{-z} = \cosh r - \sinh r \cos\theta, \mathrm{e}^{-z} x = \sin\theta \cos\phi \sinh r, \mathrm{e}^{-z} y = \sin\theta \sin\phi \sinh r$. The (x,y) subspace looks like a conformally stretched flat space. The topological identifications $x \to x + L_x$ and $y \to y + L_y$ make a 2-torus which is squeezed in area by the exponential factor e^{-2z} into a tightening cusp. Since the subspace is conformally flat, there are no tangled geodesics and it is possible to decompose the eigenvalues analytically.
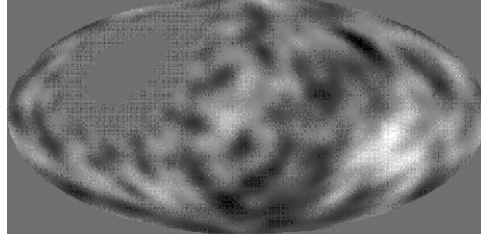
Fig. 32. Flat spot in a map of the sky in the cusp topology with scales $L_x = L_y = 1$ and $\Omega_0 = 0.3$.

In the coordinate system (88), the temperature fluctuations can be decomposed as

$$\frac{\delta T}{T}(\hat{n}_y) = \int_0^\infty \mathrm{d}q \sum_{n_x n_y} \hat{\Phi}_{q n_x n_y}(\hat{n}) N_{q n_x n_y} L_{q n_x n_y} \tag{89}$$

with the normalization

$$N_{q n_x n_y} = \left(\frac{2q \sinh(\pi q)}{\pi^2} \frac{(2 - \delta_{n_x 0})(2 - \delta_{n_y 0})}{L_x L_y} \right)^{1/2} \tag{90}$$

and

$$L_{q n_x n_y} = \left[\frac{1}{3} + 2 \int_{\eta_i}^{\eta_0} \mathrm{d}\eta F'(\eta) \right] \mathrm{e}^z K_{iq}(Q \mathrm{e}^z) \times \begin{pmatrix} \sin\left(\frac{2\pi n_x}{L_x} x\right) \\ \cos\left(\frac{2\pi n_x}{L_x} x\right) \end{pmatrix} \begin{pmatrix} \sin\left(\frac{2\pi n_y}{L_y} y\right) \\ \cos\left(\frac{2\pi n_y}{L_y} y\right) \end{pmatrix}, \tag{91}$$

where K_{iq} is a modified Bessel function with imaginary index. The entire function is included here in the integration over η. The argument of the Bessel function is $Q^2 = 4\pi^2(n_x^2/L_x^2 + n_y^2/L_y^2)$. Notice that q is still a continuous index since the z direction is infinite. There is no cutoff in the range of q but there is a cutoff in the range of (n_x, n_y). This cutoff results in a severe suppression of power as an observer lives nearer the cusp. The suppression appears as a flat spot in simulations of the COBE–DMR sky with concentric rings of larger and larger structures as the cusp widens [103] (Fig. 32).

A standard likelihood analysis of the C_ℓ's places the size of the torus at the location of the Earth to be $\gtrsim \Delta\eta$. This is not a particularly narrow part of the cusp relative to the SLS but, nonetheless, an observer can still see exponentially deep down the throat putting large quiet regions in the CMB maps [103].

The generic property of flat spots was tested in Ref. [104]. They showed that from a typical location in a cusped manifold m003, flat spots of angular size ~5° would be visible for $\Omega_0 = 0.3$. They conjecture that observable flat spots on this scale are typical of cusped manifolds. To handle the fully compact case Olson and Starkman found a means to approximate the modes in the cusp without tackling the full eigenmode solutions. They consider a horosphere, a sphere within the Poincaré representation of $\mathbf{H}^3$, which is tangent to the cusp and passes through the point on the SLS. On the horosphere, the transformation group simplifies and they are able to isolate modes.

The specific space they consider is a cusped manifold, namely m003, constructed from two tetrahedra with total volume $\mathcal{V} \approx 2.0299$. The restriction to the horosphere results in a hexagonal tiling of $\mathbf{E}^3$ and the eigenmodes are reminiscent of those of Eq. (72). Consequently, the minimum eigenmode

from the hexagonal tiling provides an estimate for the extent of the flat spot. Situating the observer at a symmetric point in the manifold they find that flat spots in the CMB from cusped regions subtend an average angle of roughly $5°$ for $\Omega_0 = 0.3$. Larger values of Ω_0 will naturally diminish the angular scale. They argue that in general cusped manifolds will show flat spots as characteristic features in the CMB maps.

6.1.3. Numerical eigenvalue solutions

The low-lying eigenmodes can always be obtained by brute force numerically. Three groups have developed numerical methods to isolate a list of eigenmodes and eigenvalues for a small collection of manifolds [105–109]. A comparison of the simulated CMB sky with the COBE–DMR data shows the specific compact hyperbolic manifolds studied were consistent with the data. However, some of the statistical analyses compare only the C_ℓ's and not the full correlation functions. Using the method of images to simulate the CMB, a full statistical analysis is performed by Bond, Pogosyan, and Souradeep as described in Section 6.1.4 where more negative conclusions are obtained; namely that the spaces studied were inconsistent with COBE except for maybe one orientation of the manifold. It would be interesting if numerical eigenmode constructions would compare a full statistical analysis to that performed using the method of images to check for consistency.

Eigenmodes of the Thurston space: Inoue was the first to find precise eigenmodes on a compact hyperbolic 3-space for the purposes of simulating a cosmological model [105]. (Aurich and Marklof were the first to compute the eigenmodes on an orbifold [110].) His method was based on the direct boundary element method initially developed by Aurich and Steiner for the study of 2-dimensional compact hyperbolic spaces. The numerical method is based on solving the Helmholtz equation

$$(\nabla^2 + k^2)\psi_{\vec{k}}(\vec{\mathbf{x}}) = 0 \tag{92}$$

on a compact manifold, i.e. with periodic boundary conditions on the universal covering space. Various methods for doing so include the finite element method and the finite difference method. He uses the more precise but numerically more time-consuming method called the direct boundary element method (DBEM) also used by Aurich and Steiner to study quantum chaos on a 2D compact space [93]. He isolated the first 36 eigenmodes of the manifold m003($-2, 3$) from the SnapPea census, otherwise known as the Thurston space (Fig. 33) in homage to the Fields Medalist W. Thurston. There are two particularly interesting results. Firstly, he finds a cutoff in the low k eigenvalues and therefore a maximum wavelength at a value near the average diameter of the space. Secondly, he discovered that the expansion coefficients of the eigenmodes (ξ_{qlm} below) are well described as pseudo-random in behavior [105]. Despite the long-wavelength cutoff, the integrated Sachs–Wolfe effect is able to compensate for low Ω_0, as expected. Consequently, assuming an Harrison–Zeldovich power spectrum, he finds that the C_ℓ's are consistent with COBE–DMR for $0.1 \leqslant \Omega_0 \leqslant 0.6$.

As in flat space, the eigenvalues are a discrete subset of the continuous eienvalues. Since, the eigenmodes of a compact manifold are a subset of the eigenmodes on the universal cover (Section 6.1.1), they can be expanded as

$$\psi_q^{\mathscr{M}} = \sum_{\ell m} \xi_{q\ell m} X_{q\ell}(r) Y_{\ell m}(\hat{n}) \ . \tag{93}$$

The $\xi_{q\ell m}$ can be thought of us as containing the analogue of the relations found explicitly in flat space in Section 5.1.2. The challenge of numerically isolating the modes can be reduced to the still
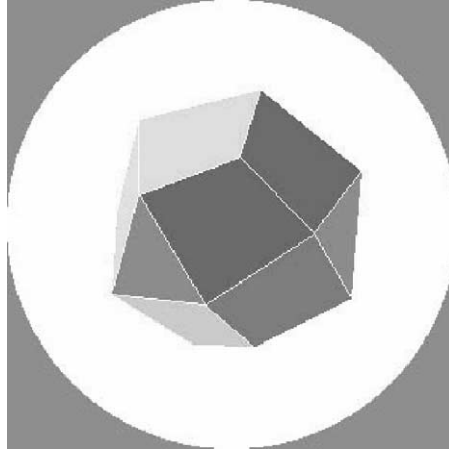
Fig. 33. Dirichlet domain for the Thurston space obtained from SnapPea.

difficult task of finding the $\xi_{q\ell m}$. Inoue explicitly finds that the $\xi_{q\ell m}$ behaves as random Gaussian numbers for ($\ell < 19$ and $q > 9.94$). He extrapolates that this continues to be true, even more so, for all higher modes and introduces this as an approximation to determine modes above $k = 10$. A distinction should be made between the random Gaussian initial conditions contained in the $\hat{\Phi}_{\bar{q}}$ and the random Gaussian behavior of the $\xi_{\bar{q}}$. The former is an assumption which may or may not be true depending on the history of the universe while the latter is a property of the manifold.

The DBEM was first used to find the pseudo-Gaussian random numbers $\xi_{q\ell m}$ for $k < 10$ [105] and the analysis was then extended to include the first 36 eigenmodes with $k \leqslant 13$ [106]. (As discussed below, it is claimed in Ref. [108] that 2 modes were originally overlooked in this method.) A numerical cutoff in the low k spectrum was found at the value $k_1 = 5.41$. The corresponding maximum wavelength $\lambda_{\max} = 2\pi/k_1$ can be compared to an approximation for the diameter of $\mathscr{M}$ which he computes as the average of the in-radius and the out-radius. The manifold has in-radius $r_{\text{in}} = 0.535$, out-radius $r_{\text{out}} = 0.7485$ (and volume $\mathscr{V} = 0.98139$). This yields $k_{\min} = 4\pi/(r_+ + r_-) = 4.9$, an agreement to within 10% of his k_1. There are no supercurvature modes in the Thurston space.

The pseudo-Gaussian behavior can be interpreted in a quantum mechanical context [105]. Let the Laplace–Beltrami operator be the Hamiltonian of a quantum system, then the eigenmode is a wavefunction eigenstate. Since the underlying classical dynamics is chaotic, it has been conjectured that the quantum mechanical system will be governed by the predictions of random matrix theory (RMT) [111,96]. One such prediction is that the square expansion coefficients are Gaussian distributed if expanded with respect to a generic bases. In particular, the coefficients should be given by the Gaussian orthogonal ensemble (GOE)

$$P(x) = \frac{1}{\sqrt{2\pi x}} e^{-x/2} \ . \tag{94}$$

To check this prediction for modes (93), Inoue first recasts the eigenmodes in terms of real independent coefficients $a_{q\ell m}$ and real functions $R_{q\ell m}$ as

$$\psi_q^{\mathscr{M}} = \sum_{\ell m} a_{q\ell m} R_{q\ell m} \ , \tag{95}$$

where the $a_{q\ell m}$ can be expressed in terms of the $\xi_{q\ell m}$ and the $R_{q\ell m}$ can be expressed in terms of the $X_{q\ell} Y_{\ell m}$. Following [93], he examines the statistical behavior of

$$x = \frac{|a_{q\ell m} - \bar{a}_q|^2}{\sigma_q^2} \; , \tag{96}$$

where $\bar{a}_q$ is the average and σ_q^2 the variance. The probability is singular at $x=0$ and so it is customary to compare the numerics to the cumulative distributions

$$I(x) = \int_0^x P(x)\,\mathrm{d}x = \mathrm{erf}(\sqrt{x/2}) \; . \tag{97}$$

To test the RMT prediction, the numerically determined cumulative distributions are compared for a goodness of fit. Very good agreement with the GOE prediction is found confirming Gaussian behavior for the low-lying states. The Gaussian behavior is expected for the highly excited states as a consequence of the classical chaos but Gaussian behavior for the low-lying states is less obvious. Yet this is in fact what his observations confirm. He also tests the randomness of the $a_{q\ell m}$ and finds that they do behave as random variables. The randomness is not a property of the eigenmodes but is due rather to the almost random distribution of images in the universal cover. The number of copies of the fundamental domain inside a sphere with radius η_0 is

$$n_1 = \frac{\pi(\sinh(2\eta_0) - 2\eta_0)}{\mathscr{V}} \tag{98}$$

giving $n_1 \sim 29$ for $\eta_0 = 1.6$.

The number of eigenmodes increases as k^3 and the number of boundary elements increases as k^2. The task of computing high k modes becomes unmanageable with this method. For higher k, Ref. [105] suggests taking the clue from the behavior of low k modes and approximating the highly excited states as random Gaussian numbers with a variance proportional to q^{-2}. Using Weyl's asymptotic formula,

$$N[q] = \frac{\mathscr{V} q^3}{6\pi^3}, \quad q \equiv \sqrt{k^2 - 1}, \quad q \gg 1 \tag{99}$$

with N being an integer. Since for a compact hyperbolic space the volume is fixed, Weyl's formula effectively relates the number of states at a given k with a topological feature of the space. It also allows an estimate of the k_j. The spacing between discrete eigenvalues decreases as the inverse of k for large k and so approaches a continuous spectrum in the large limit, as expected. Weyl's formula is well obeyed and 30 of the 36 modes show random Gaussian behavior. Interestingly, six degenerate states are found which correspond to a nearly symmetric mode reflecting the global symmetry of the fundamental domain. A linear combination of the degenerate modes shows Gaussian behavior again. This is typical of classically chaotic systems where the classical chaos leads to Gaussian behavior in the quantized eigenmodes although occasionally the global symmetry of the space can surface in the eigenmodes as nonGaussian behavior. (The connection with quantum chaos led Inoue to use the Selberg Trace Formula to compute eigenvalues for a large number of compact hyperbolic manifolds [94]. This method proved to be quicker and easier to implement numerically.)

The C_ℓ's are calculated by sewing together the numerically obtained eigenmodes for the eigenvalues in the range $5.4 \leqslant k \leqslant 13$ with the above approximation for the expansion coefficients for $13 \leqslant k < 20$. With the assumption that $\mathscr{P}_\Phi = $constant, the resultant C_ℓ is compared with the COBE–DMR

data to conclude that the spectrum is consistent with COBE–DMR for $0.2 \leqslant \Omega_0 \leqslant 0.6$. The cutoff in the spectrum due to the minimum mode $k = 5.4$ is buried under the contributions to low ℓ power from the integrated Sachs–Wolfe effect. The ISW becomes important during the curvature dominated epoch at about $1 + z \sim (1 - \Omega_0)/\Omega_0$. They find the contribution from $k \leqslant 13$ modes to C_ℓ for $2 \leqslant \ell \leqslant 20$ is 7 % for $\Omega = 0.2$ and 10% for $\Omega_0 = 0.4$. The rest is due to the ISW. For this reason, the spectra appear to have a gradual peak near the long-wavelength cutoff or are nearly flat down to low k and are less severely constrained for $\Omega \geqslant 0.1$ than the flat models of Section 5 [105,106]. In fact, they find very good agreement with COBE–DMR for $\Omega_0 = 0.6$, better than other FRW models. In particular, an alignment of the peak in the COBE–DMR data at $\ell \sim 4$ with the peak due to the cutoff accounts for the better fit.

The conclusion cannot be stated as proving that the Thurston space is consistent with the observations, only that it does not appear inconsistent. A comparison of the full correlation function to the data may not survive consistency and in fact does not according the method of images analysis of [9,10].

Eigenmodes of an orbifold: Aurich was able to compute a huge number of eigenmodes, the first 749, in an orbifold with negative curvature. Orbifolds can have compact volume but possess points which are not locally R^3. They can also have rotation group elements among their isometries, unlike compact manifolds which have no groups with fixed points in their isometries. In the absence of a predictive theory for the topology of the cosmos, there is no convincing reason to omit them from the catalog. Interestingly, they can have even smaller volumes than the minimum bound on compact hyperbolic spaces.

The fundamental cell for the orbifold is a pentahedron which is symmetric along a plane dividing the fundamental domain into two identical tetrahedra. This allows a desymmetrizing of the pentahedron useful for deducing the eigenmodes following early work [110]. There are 9 compact hyperbolic tetrahedra and the one built into the pentahedral orbifold is known as T_8. The volume of T_8 is $\mathscr{V}_{\text{tetrahedron}} \simeq 0.3586524$ which is smaller than the Weeks space. The smallest compact hyperbolic tetrahedron T_3 has volume $\mathscr{V} = 0.03588506$, 10 times smaller than T_8 and smaller even than the existing bound on compact hyperbolic manifolds ($V_{\mathscr{M}} \sim 0.3$). Using the boundary element method 749 eigenmodes on T_8 are obtained.

Both radiation and matter are included in his analysis which evolves the metric perturbations according to the usual adiabatic, linear perturbation theory and solves $F(\eta)$ numerically. In his analysis, the eigenmodes are expanded as

$$\Phi(\eta, \vec{\mathbf{x}}) = \sum F_q(\eta) \psi_q(\vec{\mathbf{x}}) , \tag{100}$$

where the ψ_q are the eigenmodes and all of the time dependence is in $F_{\vec{q}}(\eta)$. Notice that in comparison to the standard expansion (16), he does not include the usual randomly seeded fluctuation amplitude $\hat{\Phi}_k$. Instead, all of the initial conditions are absorbed into $F(\eta)$ with

$$F_q(\eta_i) = \alpha/q^{3/2}, \quad F_q'(\eta_i) = 0 . \tag{101}$$

There is however some element of randomness in the sign of the fluctuation. The eigenfunctions have a freedom in the phase. Since the phase is always chosen to be real, this freedom results in a $\pm$ ambiguity in the eigenfunction which is chosen randomly. The constant α is normalized against the COBE–DMR data. This is consistent with the Harrison–Zeldovich flat spectrum in terms of normalization but lacks any of the randomness that real fluctuations would have. Because the

randomness of the temperature fluctuations about the mean is neglected, it would not be quite right to interpret the simulated maps of Ref. [109] in terms of an actual universe. The advantage however is that the maps show a random nature which can only be due to the random character of the modes themselves. This is important for the reasons discussed at length in the previous Section 6.1.

Since the C_ℓ produces a global averaging, an ensemble average is generated by the expansion (100)–(101) even if the standard random initial fluctuations were ignored. Although the simulations depend on the location of the observer and the orientation of the manifold, this dependency is not examined and the observer is fixed, offset from the origin. The C_ℓ's are compared to the $\ell < 30$ COBE–DMR data, the Saskatoon data around $\ell \sim 100$ [112] and the QMAP data [113] above $\ell \sim 80$. All modes are computed up to $k_{max} = 55$. Aurich varies Ω_0 between 0.2 and 0.6 and finds reasonable agreement for $\Omega_0 \simeq 0.3$–0.4. For smaller Ω_0 he finds that the numerical C_ℓ increases too quickly with ℓ. We suggest that this is evidence of a long wavelength cutoff in the spectrum. Although for a lower Ω_0, the ISW is important in terms of building up the low ℓ power, it is also true that the suppression is more severe. To put it another way, if the constant α were normalized to COBE–DMR at high ℓ, as it should be in order to minimize the effects of cosmic variance, then the steep slope would be seen as a suppression of large angle power due to the finite extent of the space.

The same orbifold was studied again in Ref. [114] where both a cosmological constant and a smooth dark energy component were added to the radiation and matter energy density. The total energy was taken to be nearly but not quite flat to address the recent results from the small-angle CMB experiments [40,41,67]. They still found a suppression in power for $\ell \lesssim 10$, as illustrated in Fig. 34.

Although the results of Ref. [109] to be at variance with the results of [9,10] we emphasize again the limitations of comparing only the C_ℓ's. This is expanded upon in Section 6.1.4.

Fast method for isolating eigenmodes: A fast numerical method for obtaining the eigenvalues on compact hyperbolic manifolds was developed by Cornish and Spergel [108]. In comparison to the boundary element method developed by Aurich and Steiner [93], it is technically inferior but far quicker and easier to implement numerically. The DBEM has only been applied to a few cosmological spaces, the tetrahedral orbifold of [109], the Thurston manifold [105], and the Weeks manifold to
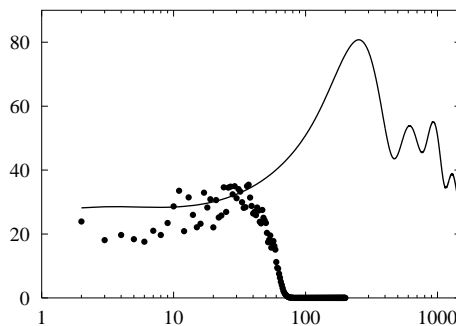

Fig. 34. This figure was supplied courtesy of R. Aurich. The figure shows T_ℓ measured in micro-Kelvin versus ℓ as computed by Aurich and Steiner for an $\Omega_0 = 0.9$ cosmology with $\Omega_\Lambda = 0.6$. The angular power spectrum C_ℓ for the orbifold is shown (dots) in comparison to an infinite topology as obtained by CMBFAST (smooth line). The suppression for $\ell > 30$ is only due to the truncation of the sum over the eigenmodes since only the first 749 eigenmodes have been taken into account.
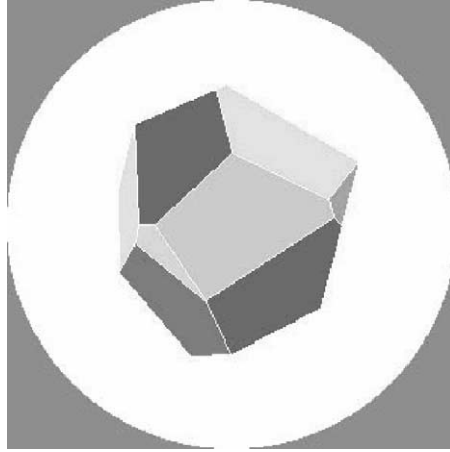
Fig. 35. Dirichlet domain for the Weeks space.

name a few [94], while the method of Ref. [108] allows them to obtain the lowest eigenvalues and eigenmodes for 12 manifolds, requiring only the generators as input.

The method is to simply solve

$$\Psi_k(x) = \Psi_k(gx) \tag{102}$$

using a singular value decomposition. The eigenmodes are again expanded in terms of the eigenmodes on the universal cover $\mathbf{H}^3$ as in Eq. (93). Random points are selected within the fundamental domain and all of the images out to some distance are located in the covering space. For each point p_j there are n_j such images located with the generators g_α. The $n_j(n_j + 1)/2$ boundary conditions are

$$\Psi_k(g_\alpha p_j) = \Psi_k(g_\beta p_j), \quad \alpha \neq \beta . \tag{103}$$

Using the expansion of the eigenmodes in terms of the eigenmodes of the universal cover, Section 6.1.1, they write (103) as a collection of difference equations which they then solve using a standard singular value decomposition method for handling over constrained systems of equations.

The example they studied in detail is the Weeks space m003(3, −1) (Fig. 35). They also looked at the Thurston space and found agreement with the DBEM method applied by Inoue except they found 2 new modes missed previously but later confirmed [75]. Again, for higher k, the number of modes obeyed Weyl's formula well. Also, the modes were well described by random matrix theory which predicts the expansion coefficients obey a Gaussian orthogonal ensemble.

The Weeks space has a large symmetry group which ensures a large degeneracy of eigenmodes. The symmetry group is the dihedral group of order 6. They found the first 74 eigenmodes, many of which are degenerate. The higher the mode, the more degeneracies in agreement with Weyl's asymptotic formula. A view of the lowest eigenmode in the Weeks space can be found in Fig. 36.

Implementing their eigenmodes in a simulation of a cosmological model, they compare the numerically generated C_ℓ's to the COBE–DMR data [108]. Their numerically modeled cosmologies are based on the first 100+ modes of the following small spaces: The Weeks space m003(3, −1) with
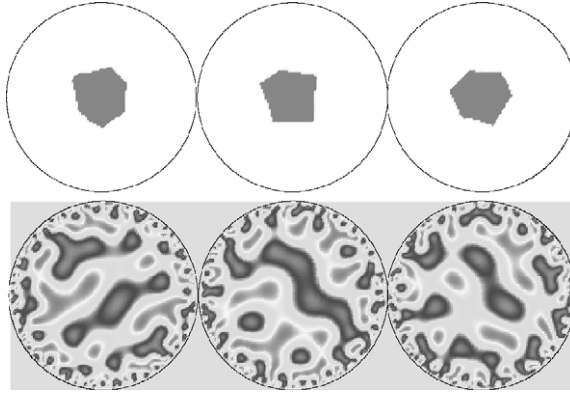
Fig. 36. This figure was supplied courtesy of N. Cornish. The lowest eigenmode in the Weeks space is shown. The space is drawn in the Poincaré representation. The three panels represent different slices through the Poincaré ball. The upper figure shows the slice through the fundamental domain while the lower figure shows the slice through the eigenmode. Specifically, the leftmost panel is the $x = 0$ slice, the middle is the $y = 0$ slice, and the rightmost panel is the $z = 0$ slice.
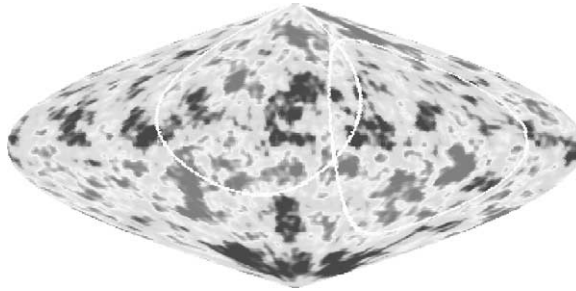

Fig. 37. This figure was supplied courtesy of N. Cornish. A numerical realization of the CMB in the Weeks topology with $\Omega_0 = 0.3$. One pair of matched circles is indicated by the white lines in this projection.

$\mathcal{V} = 0.9427$, the Thurston space m003$(-2, 3)$ with $\mathcal{V} = 0.9814$, s718$(1,1)$ with $\mathcal{V} = 2.2726$ and v3509$(4,3)$ with $\mathcal{V} = 6.2392$. One realization of the Weeks space is shown in Fig. 37. Generically, a kind of statistical isotropy prevails even though the spaces are globally anisotropic. Since the expansion coefficients are pseudo-random, Gaussian distributed numbers which are statistically independent of ℓ and m, they generate nearly isotropic eigenmodes and lead to a kind of isotropy across the microwave sky. The authors suggest that inflation may not be need to explain why the universe is nearly isotropic. However, inflation is still needed to explain why local values are so marginally near flat.

A standard likelihood analysis based on 100 realizations of each topology was performed. A mode cutoff in the spectrum is again evident and gets worse as Ω_0 is lowered. For all of the manifolds studied, they find that

$$\lambda_1^q = (1.3 \to 1.6)D \,, \tag{104}$$

where $\lambda_1^q = 2\pi/q_1$ (although it is customary to define the wavelength corresponding to a given mode as $\lambda^k = 2\pi/k$) which obeys the bound

$$\frac{4\tilde{D}}{D^2(\sinh \tilde{D} + \tilde{D})^2} \leqslant k_1^2 \leqslant 1 + \left(\frac{2\pi}{D}\right)^2 \tag{105}$$

with $\tilde{D}$ the square root of the smallest integer $\geqslant D^2$. They do caution that they cannot find modes with $q = [0, 1/4]$ using their method, and so perhaps supercurvature modes lurk. However, the ISW becomes increasingly more important as Ω_0 is lowered. There is an optimal value where the two effects compete to create a spectrum with a slight tilt at low ℓ and give a better match to the COBE–DMR data. For the Weeks space this occurs at $\Omega_0 = 0.3$.

Additionally, once the spectrum is normalized to COBE–DMR they find that the size of fluctuations on $8h^{-1}$ Mpc is naturally increased above the value of $\sigma_8 = 0.6$ for a simply connected universe with $\Omega = 0.3$ to a value of $\sigma_8 = 0.75$ for the Weeks manifold with $\Omega = 0.3$. The present day cluster abundance seems to imply $\sigma_8 = 0.9 \pm 0.1$ for $\Omega_0 = 0.3$ [115] and so higher values are desirable.

They find the C_ℓ's have a better fit to the COBE–DMR data with a relative likelihood of ~ 20 and a better fit to the large-scale structure date (σ_8 increases by $\sim 25\%$). However, as with all of the other statistical analyses they have applied a weak test by analyzing the angular averaged C_ℓ's and not the full correlation function. Their conclusions can be interpreted as finding the manifolds are not ruled out by the C_ℓ alone.

6.1.4. The method of images

The fluctuations can be simulated using the method of images. The correlation functions can be calculated without the explicit eigenmodes and eigenvalues. The procedure is to sum the correlation function on the universal cover over all images out to some large radius. The more distant images are handled in a continuous approximation. The eigenspectrum can be calculated but the correlation functions are obtained to better accuracy for a given order in the sum. Bond et al. implemented a detailed method of images [9,10,116–118]. They emphasize that the full correlation function $C(\hat{n}, \hat{n}')$ must be compared to the data for a meaningful statistical analysis. The philosophy is to perform a statistical search for patterns. They find two principle effects (1) anisotropic patterns and (2) a long wavelength cutoff in the power spectrum. The patterns appear as spikes of positive correlation when a point on the SLS and one of its images is correlated. The larger the SLS relative to the out-radius, the more statistically significant will the patterns be. If the SLS is smaller than the in-radius then naturally the correlation function is very close to that for the simply connected space.

The angular correlation function $C(\hat{n}, \hat{n}')$ is computed from the spatial two-point correlation function $\xi_\Phi \equiv \langle \Phi(\vec{x}, \tau_{LS}) \Phi(\vec{x}'', \tau_{LS}) \rangle$ which can be expressed as

$$\xi_\Phi = \sum_i \mathscr{P}_\Phi(k_i) \sum_{j=1}^{m_i} \Psi_{ij}(\mathbf{x}) \Psi_{ij}^*(\mathbf{x}') \tag{106}$$

with the sum over a discrete ordered set with multiplicities m_i. Notice that the Ψ_{ij} obey

$$(\nabla^2 + k_i^2)\Psi_{ij} = 0 \tag{107}$$

as always, however with the slightly different two-index notation. The j index accounts for any degeneracies. The spatial correlation function ξ^c on the compact manifold $\mathcal{M}^c$, can be expressed in terms of the spatial correlation function ξ^u on the universal cover. A derivation of the relation between ξ^c and ξ^u exploits orthonormality and completeness with the following equations:

$$\int_{\mathcal{M}} d\vec{\mathbf{x}}' \xi^c_\Phi(\vec{\mathbf{x}}, \vec{\mathbf{x}}') \Psi^c_{ij}(\vec{\mathbf{x}}') = \mathcal{P}_\Phi(k_i) \Psi^c_{ij}(\vec{\mathbf{x}}') \,, \tag{108}$$

$$\int_{\mathcal{M}^u} d\vec{\mathbf{x}}' \xi^u_\Phi(\vec{\mathbf{x}}, \vec{\mathbf{x}}') \Psi^u_j(k, \vec{\mathbf{x}}') = \mathcal{P}_\Phi(k_i) \Psi^u_j(k, \vec{\mathbf{x}}') \,. \tag{109}$$

Since eigenfunction on the compact space must also be eigenfunctions on the universal cover (although the converse is not true) it follows that

$$\int_{\mathcal{M}^u} d\vec{\mathbf{x}}' \xi^u_\Phi(\vec{\mathbf{x}}, \vec{\mathbf{x}}') \Psi^c_j(k, \vec{\mathbf{x}}') = \mathcal{P}_\Phi(k_i) \Psi^c_j(k, \vec{\mathbf{x}}')$$
$$= \int_{\mathcal{M}} d\vec{\mathbf{x}}' \xi^c_\Phi(\vec{\mathbf{x}}, \vec{\mathbf{x}}') \Psi^c_{ij}(\vec{\mathbf{x}}') \,. \tag{110}$$

Combining (110) with (109) gives

$$\int_{\mathcal{M}} d\vec{\mathbf{x}}' \xi^c_\Phi(\vec{\mathbf{x}}, \vec{\mathbf{x}}') \Psi^c_j(k, \vec{\mathbf{x}}') = \int_{\mathcal{M}^u} d\vec{\mathbf{x}}' \xi^u_\Phi(\vec{\mathbf{x}}, \vec{\mathbf{x}}') \Psi^c_{ij}(\vec{\mathbf{x}}') \,. \tag{111}$$

Since $\mathcal{M}$ tessellates $\mathcal{M}^u$ we can re-express the left-hand side of (111) as

$$\sum_{g \in \Gamma} \int_{\mathcal{M}} d\vec{\mathbf{x}}' \xi^u_\Phi(\vec{\mathbf{x}}, g\vec{\mathbf{x}}') \Psi^c_{ij}(\vec{\mathbf{x}}') = \int_{\mathcal{M}} d\vec{\mathbf{x}}' \left[\sum_{g \in \Gamma}^{\sim} \xi^u_\Phi(\vec{\mathbf{x}}, g\vec{\mathbf{x}}') \right] \Psi^c_{ij}(\vec{\mathbf{x}}') \,, \tag{112}$$

where in the last step a regularization is needed and denoted by the tilde above the summation. Identifying integrands gives

$$\xi^c_\Phi(\vec{\mathbf{x}}, \vec{\mathbf{x}}') = \sum_{g \in \Gamma}^{\sim} \xi^u_\Phi(\vec{\mathbf{x}}, g\vec{\mathbf{x}}') \tag{113}$$

and the ξ^c_Φ can be calculated as the sum over images in the universal cover with only a knowledge of the group elements. The spatial correlation function on the universal cover is known to be

$$\xi^u_\Phi(\vec{\mathbf{x}}, \vec{\mathbf{x}}') \equiv \xi^u_\Phi(r) = \int_0^\infty \frac{dq\, q}{(q^2 + 1)} \frac{\sin(qr)}{q \sinh r} \mathcal{P}_\Phi(q) \tag{114}$$

and $\mathcal{P}_\Phi$ is taken to be a Harrison–Zeldovich spectrum. This is the core of the method of images.

The regularizer requires some additional effort. The correlation function on the universal cover does not have compact support:

$$\int_{\mathcal{M}^u} d\vec{\mathbf{x}}' \xi^u_\Phi(\vec{\mathbf{x}}, \vec{\mathbf{x}}') = \infty \,. \tag{115}$$
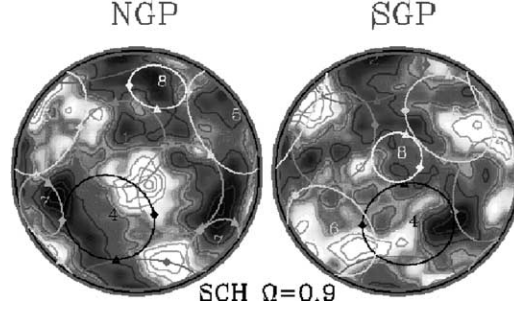
Fig. 38. This figure was supplied courtesy of T. Souradeep. The full CMB sky is represented by the two hemispherical caps—one in the direction of the South Galactic Pole (SGP) and the other in the direction of the North Galactic Pole (NGP). The label SHC refers to a small compact hyperbolic model, namely m004$(-5, 1)$. Along with the fluctuations, this figure also shows correlated circle pairs explained in more detail in the following section.

The need for regularization is not a result of the compact hyperbolic space having a large number of periodic orbits as incorrectly claimed in [3] nor is it due to the chaotic nature of the trajectories but rather is a result of the correlation function not having compact support. Even flat spaces require regularization. The regularized spatial correlation function can be written as

$$\tilde{\xi}^{\mathrm{u}}_{\Phi}(\vec{\mathbf{x}}, \vec{\mathbf{x}}') \equiv \xi^{\mathrm{u}}_{\Phi}(\vec{\mathbf{x}}, \vec{\mathbf{x}}') - \frac{1}{V_{\mathcal{M}}} \int_{g\mathcal{M}} \mathrm{d}\vec{\mathbf{x}}'' \, \xi^{\mathrm{u}}_{\Phi}(\vec{\mathbf{x}}', \vec{\mathbf{x}}'') \tag{116}$$

for g such that $\vec{\mathbf{x}}'$ lies in $g\mathcal{M}$. The regularization ensures

$$\int_{\mathcal{M}^{\mathrm{u}}} \mathrm{d}\vec{\mathbf{x}}' \, \tilde{\xi}^{\mathrm{u}}_{\Phi}(\vec{\mathbf{x}}, \vec{\mathbf{x}}') = 0 \ . \tag{117}$$

Finally,

$$\xi^{\mathrm{c}}_{\Phi}(\vec{\mathbf{x}}, \vec{\mathbf{x}}') = \sum_{g \in \Gamma} \tilde{\xi}^{\mathrm{u}}_{\Phi}(\vec{\mathbf{x}}, \vec{\mathbf{x}}') = \sum_{g \in \Gamma} \xi^{\mathrm{u}}_{\Phi}(\vec{\mathbf{x}}, \vec{\mathbf{x}}') - \frac{1}{V_{\mathcal{M}}} \int_{\mathcal{M}^{\mathrm{u}}} \mathrm{d}\vec{\mathbf{x}}'' \, \xi^{\mathrm{u}}_{\Phi}(\vec{\mathbf{x}}', \vec{\mathbf{x}}'') \ . \tag{118}$$

The regularization prescription is not unique. The actual limiting procedure utilized in Ref. [9] sums images up to a radius r_* and then regularizes by subtracting the integral of $\xi^{\mathrm{u}}_{\Phi}(r)$ over a spherical ball of radius r_*:

$$\xi^{\mathrm{c}}_{\Phi}(\vec{\mathbf{x}}, \vec{\mathbf{x}}') = \lim_{r_* \to \infty} \left[\sum_{r_j < r_*} \xi^{\mathrm{u}}_{\Phi}(r_j) - \frac{4\pi}{V_{\mathcal{M}}} \int_0^{r_*} \mathrm{d}r \, \sinh^2 r \, \xi^{\mathrm{u}}_{\Phi}(r) \right] \tag{119}$$

with $r_j = d(\vec{\mathbf{x}}, g_j \vec{\mathbf{x}}') \leqslant r_{j+1}$. The value of r_* is numerically pushed out to 4 or 5 times the out-radius in order to get a convergent result. This procedure is simpler and does well for large r_*. It also does not require a detailed dependence on the complicated shape of the fundamental domain.

An example of the maps they generate using the method of images is shown in Fig. 38.

They are also able to estimate the density of states and therefore obtain a rough estimate of the power spectrum using the method of images. This is in addition to their primary result of having

computed the correlation function. When they do estimate the density of states they always find an infrared cutoff as expected for long-wavelength modes.

This is consistent with the expectations. For a 3-dimensional compact hyperbolic space there is a bound on Cheeger's isoperimetric constant of [119,120]

$$k_{\min} \geqslant h_C/2 \geqslant \frac{1}{d_{\mathcal{M}}} \left[2 \int_0^{1/2} dt \cosh^2(t) \right]^{-1} = 0.92/d_{\mathcal{M}} \tag{120}$$

and so there are no supercurvature modes for $d_{\mathcal{M}} < 0.92$. The suppression of power is covered in part by the ISW and so is less prominent than in flat models.

Given their numerical calculation, they do a Bayseian probability analysis comparing a few compact hyperbolic cosmologies to the COBE–DMR data. Their analysis represents by far the most complete of the statistical tests that has been performed and, in principle, is the most complete test that can be performed. They find the compact hyperbolic models they study are inconsistent with the COBE–DMR data for most orientations although for a small set of special orientations they do find a better fit than the standard infinite models. Their results are qualitatively independent of the matter content. The analysis magnifies the inadequacy of using the C_ℓ alone. In particular, they point to the large cosmic variance that results from the break down associated with isotropy.

With the assumption of Gaussianity both in the noise and in the raw $\delta T/T$, the probabilistic comparison used is

$$\mathscr{P}(\delta|C_T) = \frac{1}{(2\pi)^{N_p/2}\|C_T\|^{1/2}} e^{(1/2)\delta^\dagger C_T^{-1}\delta} , \tag{121}$$

where $C_T = C(\hat{n}, \hat{n}')$ and δ is the data. To numerically obtain the full correlation function requires an evaluation of the spatial correlation function at $N_p(N_p + 1)N_L^2/2$ pairs of points where N_p is the number of pixels and N_L is is the number of points along the line of sight used to integrate the ISW. They find an $N_L \sim 10$ is satisfactory. They estimate the likelihood function

$$\mathscr{L}(C_T) \equiv \mathscr{P}(\bar{\delta}|C_T) = \int d\delta \mathscr{P}(\bar{\delta}|\delta)\mathscr{P}(\delta|C_T) \tag{122}$$

with $\bar{\delta}$ the map which maximizes the conditional probability and the integration is performed over all realizations of the simulated sky δ. The resultant likelihood function they use is

$$\mathscr{L}(C_T) = \frac{1}{(2\pi)^{N_p/2}\|C_N + C_T\|^{1/2}} e^{(1/2)\bar{\delta}^\dagger(C_N+C_T)^{-1}\bar{\delta}} \tag{123}$$

with C_N the noise covariance matrix. What is actually obtained is model-dependent relative likelihoods. The model-dependent parameters are $\mathcal{M}$, the orientation, the location of the observer, Ω_m, Ω_0 and Ω_Λ, and the initial assumptions regarding the spectral shape and character. The manifolds studied were m004(−5, 1), which is relatively small, and v3543(2,3), which is comparatively large. (Always $\Omega_\Lambda = 0$. Adding Ω_Λ relaxes the constraints.) They varied $\Omega_m = \Omega_0$ over three values arranged so that the SLS was comparable to the out-radius and varied over 24 different orientations. They do leave the observer at the maximum of the injectivity radius. This location gives the most symmetric

perceived shape to the Dirichlet domain which one might expect to lead to the most conservative bounds. It seems fair to assume that moving the observer to a thinner region in the manifold for instance will only amplify asymmetries and the constraint on the correlation function should only be more severe. However, others have argued that locating the Earth away from the local maximum of the injectivity radius can increase the likelihood fit of a compact hyperbolic model to the data. For instance, to check the effect of the inhomogeneity, Inoue moved the observing point for the same model m004($-5, 1$)=Thurston manifold and performed the Bayesian analysis as Bond et al. did. He found that there are a few choices of the position and orientation for which the likelihood is much larger than that of the infinite counterpart [74] (as did [10]). The best-fit positions are scattered in the manifold and far from the local maximum of the injectivity radius.

In any case, the results of [9,10] are presented as the relative likelihood of a given model in comparison with a simply connected hyperbolic CDM model with the same $\Omega_m = \Omega_0$. They uniformly find that the compact hyperbolic models have very small relative likelihoods as compared with the simply connected models. The rare exception occurred near $\Delta\eta \approx r_+$ with a particular orientation. The statistical significance is unclear since a fortuitous alignment of the measured fluctuations with simulated topological images could enhance the likelihood when taken over all realizations which reinforce this correlation. The authors defer conclusions to future tests such as the circle method. By contrast, Inoue argues the statistical significance *is* clear. If the C_ℓ's fit the data poorly, there will be no orientation of the manifold that leads to an alignment of measured fluctuations with topological images. Put another way, it may be unfair to suggest that a good fit to the data is a fortuitous alignment instead of acknowledging this good fit to be a good fit.

Bond et al. draw the general conclusion that $d_{\mathcal{M}}/2 > 0.7\Delta\eta$. The compact models they tested are excluded at the 3σ level with the exception of those with very special orientations. By contrast, if they were to only analyze the statistical likelihood of the C_ℓ they would mistakenly conclude that the compact models were preferred at the 1σ level over the simply connected cosmology. They emphasize that the error bars on the C_ℓ's are huge because of the exaggerated cosmic variance in the topologically connected models. It may also be worth noting here that there is some argument over the interpretation of these error bars (see Ref. [74]).

Due to the global breaking of homogeneity, they find the variance is spatially dependent. This is another reason to be wary of conclusions based on C_ℓ alone. They found characteristic loud spots, that is regions in the sky with larger variance than others. Loud regions correspond to intersections of the SLS through smaller regions of the fundamental domain. The regions are geodesicaly small, that is to say, there are shorter geodesics relative to other regions in the volume. Unless the region is very very small, the ISW can obscure this particular feature. These loud and quiet regions are familiar from the cusp topology studied in Ref. [103]. There, even the ISW cannot compensate for the deepest regions of the cusp.

Another attitude to take would be to argue that if we do live in a compact hyperbolic manifold, some orientation is necessarily going to be much more likely, namely the right one. The special status of the "correct" orientation for our manifold, assuming it is compact hyperbolic, was taken seriously by one of the pioneers in cosmic topology, Fagundes [121,122]. He reasoned that the particularly significant hot and cold spots in the COBE maps found by Cayón and Smoot [123] may shed light on the orientation of the manifold. These loud spots may be patches of high and low regions in the physical density, as opposed to being just statistical fluctuations. Since the density fluctuations eventually evolve into large-scale structure, then the hot and cold spots should correspond
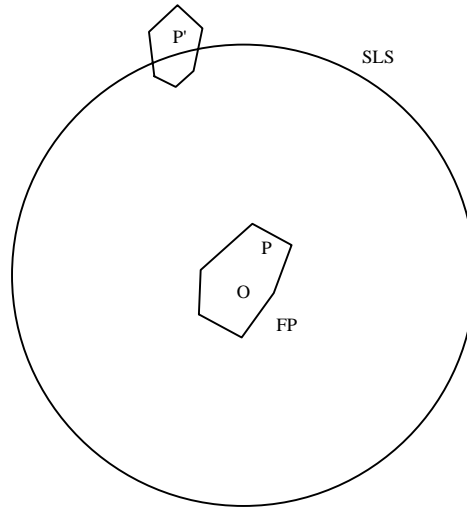
Fig. 39. This figure was supplied courtesy of H. Fagundes. It shows the fundamental polyhedron as well as a copy of the fundamental polyhedron that intersects the spherical surface of last scattering. The point of intersection identifies the original location of a CMB spot in the universe.

to physical superclusters and voids, respectively. Fagundes used this idea to try to match sources in catalogs of galaxy superclusters and voids and thereby fix the orientation and location of the earth in the manifold. His algorithm begins with a point P' centered on one of the spots isolated in Ref. [123] as illustrated in the schematic Fig. 39. The point $P' \in SLS$. He then maps ghost images of $P' = \gamma P$ for $\gamma \in \Gamma$ so the γ are composite words built from the face-pairing matrices. He tries to align the image points with voids and superclusters for different orientations of the space and different basepoints, i.e. different locations of the earth. He performed these scans for the 10 smallest compact hyperbolic manifolds known [124]. No firm conclusions can be drawn without very deep surveys of superclusters. Fagundes' idea remains a plausible avenue to pursue if such data ever become available.

6.2. Geometric methods in hyperbolic space

Given the likelihood analysis of the previous section, should we generically conclude that all compact hyperbolic spaces must be large relative to the observable universe? The answer is really "no" since there is so much model dependence in any of the direct methods of the previous section. We cannot be confident about our assumptions for the initial fluctuation spectrum, the choice of manifold, the orientation, location of observer, local parameter values or even the statistics themselves. (For interesting discussions on related issues please see Ref. [125].) Remembering that there are an infinite number of manifolds to consider and the ambiguities in the model parameters, statistical conclusions are limited to the specific. These attributes beg for a template-independent method of searching for topology. Much like gravitational lensing, we might hope to simply look at the sky and see evidence of topological lensing without prior assumptions about the shape of the "lens"
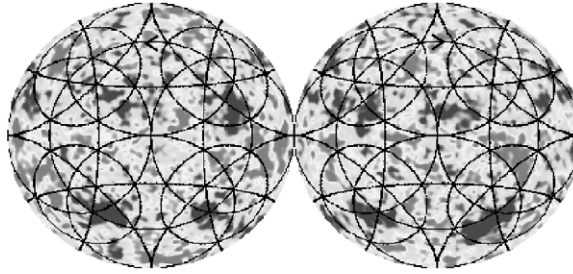
Fig. 40. This figure was supplied courtesy of N. Cornish and shows the northern and southern hemispheres of the microwave sky in flat hypertorus. There are 13 matched circle pairs indicated by black lines.

(in this case, the shape of space). There is hope for such model-independent observations as exemplified in the circles in the sky described below. Generically, we distinguish these geometric methods which search for patterns from statistical methods which rely on a specific model.

6.2.1. Circles in the sky

Possibly the nicest geometrical observation made thus far has been the prediction of circles in the sky of Cornish et al. [3–5]. It has quickly become a popular topic in conversations on the topology of the universe. The circles are most easily seen in the tiling representation of a compact space. Each copy of the Earth will come complete with its own surface of last scatter. If the two images of the Earth are near enough, these identical copies of the SLS will intersect. The intersection of two spheres occurs along a circle. Since the observers at the center of the intersecting spheres are actually just copies of one observer, the circle of intersection must always come in pairs as illustrated in Fig. 27. To emphasize, we will not look up in the sky and see intrinsic circles in the microwave; that is, the circle pairs do not have identical temperatures along a circle but rather the temperature varies identically when taken along correlated circles [3]. An illustration of the location of circle pairs in a finite flat torus is shown in Fig. 40 from Ref. [126].

The most important aspect of the circles approach is that it applies to all multiconnected topologies and does not require a template nor any a priori assumptions about a model. All compact spaces will have circles if any part of the geometry is smaller than the SLS. The radius, number, and distribution of the circles will vary from space to space and hence the topology can be reconstructed from the circle pairs. A distribution of clone images in the Thurston space can be seen in Fig. 41. Since the shape of the Dirichlet domain is not a topological invariant but rather depends on the location and orientation of the observer, even the Earth's location in the universe can be deduced.

A statistical scan of the sky must be performed to draw the correlated circles out of the maps. The circle pairs will be completely hidden. To pull them out, consider two rings each with angular radius α and with relative phase ϕ_* centered on arbitrary points $\vec{x}$ and $\vec{y}$. To test whether these arbitrary rings are in fact correlated circles of intersection, the comparison statistic

$$S(\phi_*) = \frac{\langle 2T_1(\pm\phi)T_2(\phi + \phi_*)\rangle}{\langle T_1(\phi)^2 + T_2(\phi + \phi_*)^2\rangle} \tag{124}$$
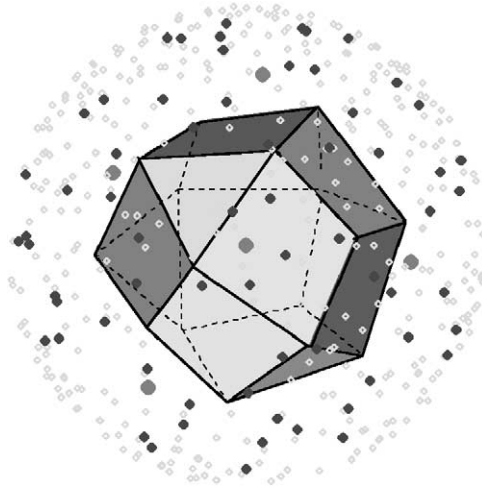
Fig. 41. This figure was supplied courtesy of N. Cornish and illustrates just how many circles are to be expected. A fundamental domain for the Thurston space is shown. The distribution of points mark our clones out to a distance of 3 times the curvature radius. The large points are within one curvature radius, the medium-sized points are within two curvature radii and the small points are within three curvature radii. Each of these clones will have a clone surface of last scatter. The clones which generate circle pairs will depend on the size of surface of last scatter compared to the clone distance and therefore depends on the value of Ω_0.

has been proposed where $\langle \rangle = \int_0^{2\pi} \mathrm{d}\phi$ with S range $[-1, 1]$ [3]. Perfectly matched circles have $S = 1$ while an ensemble of uncorrelated circles will have a mean value of $S = 0$. (Roukema includes the small-scale Doppler effect in a slight alteration of the statistic [90].) Orientable topologies will have clockwise–anticlockwise correlations while nonorientable topologies will have a mixture of clockwise–clockwise and clockwise–anticlockwise correlations. In flat space, matched circle pairs have angular radius

$$\alpha = \arccos\left(\frac{X}{2\Delta\eta}\right) \tag{125}$$

with X the distance between the Earth and its image. In hyperbolic geometry

$$\alpha = \arccos\left(\frac{\cosh X - 1}{\sinh X \tanh \Delta\eta}\right) \ . \tag{126}$$

There are no circle pairs if the image is too far for the spheres to intersect and the expressions are invalid for $X > \Delta\eta$. The number of images grows exponentially with X in hyperbolic space so that most circles have small radii, although the statistic works best for large circles. It is not clear which effect dominates.

Noise will degrade the circle statistic so that $S \neq 1$. The experimental noise in each pixel can be approximated as random Gaussian noise

$$P(n) = \frac{1}{\sigma_n\sqrt{2\pi}}\mathrm{e}^{-n^2/2\sigma_n^2} \tag{127}$$

with variance $\sigma_{\rm n}$. The true temperature fluctuations are also taken to have a Gaussian distribution with variance σ_s. The probability for the matched circle is then

$$P^m(S)\,{\rm d}S = \frac{\Gamma(N)2^{-N+1}}{\Gamma(N/2)^2}\frac{(1+2\xi^2)^{N/2}}{(1+(1-S)\xi^2)^N}(1-S^2)^{N/2-1}\,{\rm d}S \tag{128}$$

with $\Gamma(N)$ the gamma function and $\xi = \sigma_s/\sigma_{\rm n}$ is the signal to noise of the detector while N is the number of pixels. For large N, the distribution has a maximum

$$S_{\rm max}^m = \frac{\xi^2}{1+\xi^2} + \mathcal{O}(N^{-1})\ . \tag{129}$$

The higher the resolution and the signal to noise, the better this statistic will fare. If the experimental angular resolution is $\delta\theta$, then there are $N \simeq 2\pi\sin\alpha/\delta\theta$ data points around each circle of angular radius α. For COBE–DMR, the signal-to-noise ratio is $\xi = 2$, $\delta\theta = 10°$ and $N{\sim}36\sin\alpha$ pixels while for MAP, $\xi{\simeq}15$, $\delta\theta = 0.2°$ and $N{\simeq}1800\sin\alpha$ in its highest frequency channel. Using MAP parameters, only circles with $\alpha > 4°$ are detectable.

To make a relative comparison with a probability distribution for unmatched circles they advocate

$$P^{\rm u}(S)\,{\rm d}S = \frac{\Gamma(N)2^{-N+1}}{\Gamma(N/2)^2}(1-S^2)^{N/2-1}\,{\rm d}S\ , \tag{130}$$

which is centered at $S = 0$ with a FWHM$\simeq (8\ln 2/n)^{1/2}$. The unmatched probability distribution $P^{\rm u}$ was derived by assuming the temperature at each point is an independent Gaussian random variable. There are some weaknesses in this assumption since there are known correlations in the CMB.

Cornish and Spergel applied the circles test to maps generated with their numerically isolated eigenmodes of Section 6.1. For a realization of the Weeks space they found the uncorrelated ISW clouded the statistic resulting in a poor match for circle pairs. To combat this pollution they remove all power in modes below $\ell = 21$ and find a substantial improvement in the match with values of $S \approx 0.9$. Their expectations for the future satellite missions are high.

The circles method was first applied to simulated maps by Bond, Pogosyan and Souradeep using the method of images of Section 6.1.4. From their correlation function they are able to make full sky maps. Doing so they search for circles using the cross-correlation coefficient between fluctuations along two circles C_1 and $C_2 = gC_1$,

$$\rho_{12} \equiv \frac{\langle\delta T(\vec{\mathbf{x}})\delta T(g\vec{\mathbf{x}})\rangle}{\left[\langle\delta T(\vec{\mathbf{x}})^2\rangle\langle\delta T(g\vec{\mathbf{x}})^2\rangle\right]^{1/2}} \tag{131}$$

with $\vec{\mathbf{x}} \in C_1$ and $g\vec{\mathbf{x}} \in C_2$. A statistical average is implied by the angular brackets. For one given realization an integration over $\vec{\mathbf{x}}$ along the circles is performed instead. They find ρ_{12} is in the range 0.6–0.95 for $\Omega_0 = 0.9$ and m004$(-5, 1)$ and 0.2–0.6 for $\Omega_0 = 0.6$ for v3543(2,3). Notice that Ω_0 is selected for each manifold so that $\mathcal{V}$ is comparable to the volume of the SLS. The matches they find are good even at COBE–DMR resolution. The correlations along circles gets worse as the ISW contribution is enhanced at low Ω_0.

In addition to the circular intersections of the boundary of the SLS, there are intersections of the volume. Let S represent the collection of points contained within the SLS. Then the intersection of

$gS \cup S$ defines a lens-shaped region in the volume and must be identical to the intersection of $S \cup g^{-1}S$. As long as $r_- < \Delta\eta$, there will be lens-shaped intersections of the volumes and circular intersections of the boundaries of the copies of the SLS. Since the anisotropies are generated along the entire line of sight, Bond, Pogosyan and Souradeep consider the correlations in the full lens-shaped volume. Based on the COBE–DMR constraints obtained via the method of images as described in Section 6.1.4 they take the topology scale comparable to the SLS. They argue that correlations must then be very near the faces of the Dirichlet domain and so the correlations provide a quick sketch of the shape of the universe [10].

The volume intersections are relevant since, contrary to intuition, they found the ISW is not entirely uncorrelated. While there are an infinite number of lines of sight which share at least one common point, there are also pairs of lines which have segments in common and so an enhanced correlation. Every pair of lines of sight which are directed toward the center of matched circles, necessarily contain segments of identical points and leads to correlated patterns. These substantial anticorrelated features tend to lie at the centers of matched circles. The anticorrelation comes from the interference term between the surface Sachs–Wolfe effect and the ISW. As described in Section 6.1.4, their likelihood comparison of a handful of spaces to the existing data led to pessimistic conclusions.

Another application of the circles method using COBE–DMR sought to identify an asymmetric flat 3-space [90]. COBE–DMR is not ideal for detecting circles in hyperbolic space since the ISW accounts for most of the power in the COBE–DMR range of detection. However, for flat spaces with no cosmological constant the sky is determined by the surface Sachs–Wolfe effect alone without the obscuring effects of the ISW. Despite the low resolution, COBE–DMR might still see circles if we live in a compact flat space.

Roukema considered a specific asymmetric torus which was put forth as a candidate model to match cluster observations [127,128]. Although he used the recent circles prediction to test his hypothesis, the symmetries of the space and the geometric approach are reminiscent of the earlier work of [87] described in Section 5.2. Hot X-ray bright gas in large galaxy clusters were used to search for topological images [127]. Two clusters at redshifts $z \sim 0.4$ very nearly form a right angle with the Coma cluster with very nearly equal arms. On the basis of this geometric relation, they take a toroidal geometry for the universe to explain these clusters as topological images of the Coma cluster. The distance from the Coma cluster to CL 09104+4109 and the distance from Coma to the cluster RX J1347.5-1145 are both $\approx 960h^{-1}$ Mpc for zero cosmological constant. The third dimension is taken to be larger than the diameter of the SLS and hence topological effects from this direction are essentially unobservable. The size of the small dimension was taken to be roughly $\Delta\eta/13.2$. Using the circles statistic, the candidate was ruled out at the 94% confidence level, provided that the cosmological constant is zero. Specifically, the statistic Roukema used was

$$d \equiv \left\langle \frac{(\delta T/T)_i - (\delta T/T)_j}{\{[\Delta(\delta T/T)]_i^2 + [\Delta(\delta T/T)]_j^2\}^{1/2}} \right\rangle \tag{132}$$

where $[\Delta(\delta T/T)]_i$ are the observational error estimates on the temperature fluctuations $(\delta T/T)_i$ as estimated by the COBE team. This statistic directly tests the consistency of temperature values within observational error bars [90]. If the ISW effect is treated as noise as in Ref. [90], the application of
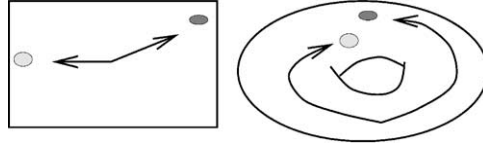
Fig. 42. Two temperature spots which appear to originate far from each other may actually originate quite near each other on a finite space.

the circles principle using the statistic (132) can still enable rejection of a specific partly compact model or show that a model one-tenth of the horizon diameter is consistent with the COBE data.

In general, there are obstacles to implementing the circle method in practice. For one, the much emphasized ISW effect is not correlated in this way and so can obscure the circle pairs. Other problematic effects include the velocity and thickness of the SLS. In Ref. [90], it was noted that many circles are partially lost along with the 20° galactic cut (the size of the galactic cut may be significantly less for MAP and *Planck*) and that both detector noise and foreground contamination posed difficulties for the circle detections. Realistically, it may not be possible to observe circles even if they are there and this will be the challenge faced in realistic analyses of the future satellite data.

6.2.2. Pattern formation

Developmental biology and condensed matter physics have long exploited pattern formation induced by periodic boundary conditions. Since compact topologies can be understood as a set of intricate boundary conditions, pattern formation has a cosmological analogue. The emergence of patterns in the sky are not so clear since many modes are competing for attention. The superposition of many otherwise distinct geometric patterns can lead to something apparently random. The task of geometric methods is to separate the patterns out of the sky. The statistic of Eq. (124) manages to draw out circles. Other patterns in addition to the circles can emerge as described in this section.

The patterns are best siphoned off a CMB map by scanning for correlations [7]. As an example consider a map of the antipodal correlation

$$A(\hat{n}) = C_{\mathscr{M}}(\hat{n}, \hat{n}') ,\tag{133}$$

which measures the correlation of fluctuations received from opposite points on the SLS [7]. In a simply connected space opposite sides of the SLS should have no communication between them and a map of antipody would generate nothing more than a monopole. It is possible that accidental correlations appear at random but no geometric structure would emerge. By contrast, if the universe is topologically connected, then points on the manifold which seem to be far apart in the tiling picture may actually be quite close together in the fundamental domain. Therefore, opposite points on the SLS may be strongly correlated, may in fact be the same point. This is another example of ghost images but in an antipodal map collections of ghosts are caught and a picture of the symmetries of the space emerges [7] (Fig. 42).

The size of a spot can be estimated at the Silk damping scale below which fluctuations have smoothed some. The angular size of these spots is too small for COBE–DMR to have detected but will be visible to the high-resolution MAP and *Planck Surveyor*.
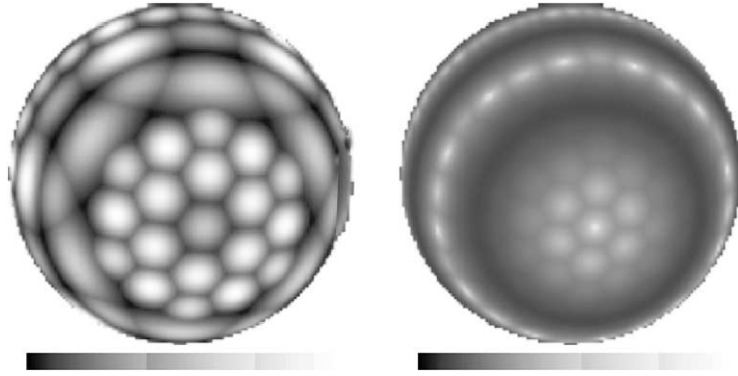
Fig. 43. Left: Orthographic projection of $A(\hat{n})$ for a hexagonal prism with $L=0.6\Delta\eta$. Right: $A(\hat{n})$ for $\pi/3$-twisted hexagonal prism. The length of the prism direction is $0.24\Delta\eta$ while $L=1$. There are circles.

Although the search for pattern formation in correlated maps is model independent, a zoo illustrating the variety of structures compact manifolds produce can be built with some simple approximations. The correlation between two points on a compact manifold can be estimated as the correlation they would have on the universal cover given their minimum separation:

$$C_{\mathcal{M}}(\hat{n},\hat{n}') \approx C^{\mathrm{u}}[d_{\min}(\vec{x}(\hat{n}),\vec{x}'(\hat{n}'))] \, , \tag{134}$$

where C^{u} is the correlation function on the universal cover and $d_{\min}$ is the minimum distance between the two points in the topological space. The estimate is effectively the lowest order term in the method of images. It is inadequate for use in a likelihood analysis but is sufficient for predicting the types of patterns which emerge from topological lensing. The images of a given point out to order m are found with the generators of the identifications as

$$\vec{y}_{k_m,\dots,k_i} = \prod_{i}^{m} g_{k_i}\vec{x}'(\hat{n}') \, . \tag{135}$$

The image point which lands closest to $\vec{x}(\hat{n})$ determines $d_{\min}$.

As an example, we show the antipodal map for a $2\pi/3$-twisted hexagonal prism in Fig. 43. For the antipodal map we prefer the orthographic projection which shows the genuine shape of the surface of last scattering instead of the Aitoff projection customary in $\delta T(\hat{n})/T$ maps. Notice the clear hexagonal face drawn out by the correlated map. Since antipody is symmetric under π, the back is a copy of the front. For the $\pi/3$-twisted hexagon, if the space is small enough, pairs of circles appear in $A(\hat{n})$, as shown in the right-most panel of Fig. 43. These are the circles in the sky of Section 6.2.1.

A compact hyperbolic space shows distinct patterns. The compact icosahedron known as the Best space after the mathematician who identified the manifold provides the best testing ground (Fig. 44) [129]. The map of $A(\hat{n})$ is shown in the left of Fig. 45. Antipody outlines pairs of identified triangular faces and also locates circles. Clearly, a symmetry group for the Best space is located in this map. Another correlation function is also shown which compares one point on the SLS to the rest of the sphere. The point selected is a copy of the origin and so reflects the most
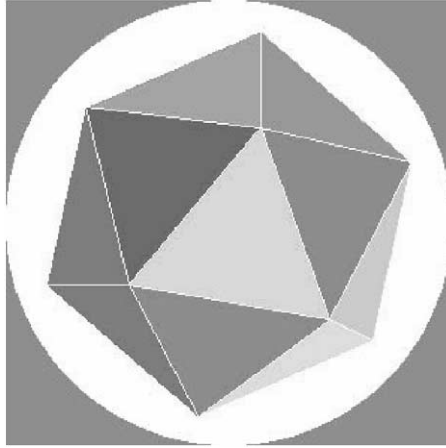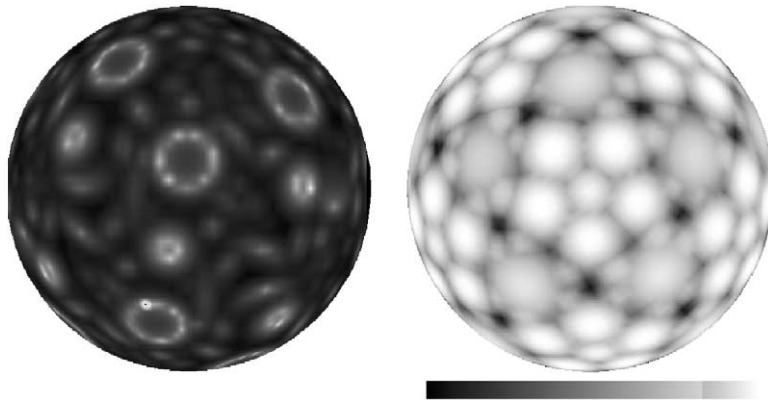
Fig. 44. Dirichlet domain for the Best space from SnapPea.

Fig. 45. Left: $A(\hat{n})$ for the Best space with $\Omega_0 = 0.3$. Right: The point-to-sphere correlation.

symmetric observation of the fundamental domain. Another example is given by the antipodal map for the Thurston space in Fig. 46. Some of the correlated features such as the arcs in Fig. 46 may be secondary correlations and it is not clear they will ever be bright enough to be observed.

In fairness, it is difficult to know if any of these patterns will really be measurable in a realistic experiment with physical complications such as the thickness of the surface of last scatter, additional Doppler effects, noise, fictitious correlations, etc. To read the correlations from the future data, real-space statistics will need to be developed which handle smoothings, subtractions of low-order multipoles, noise and fictitious correlations. While some statistics have been promoted, a realistic approach will likely develop only when the data are actually available.
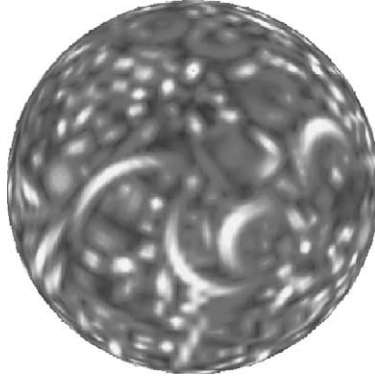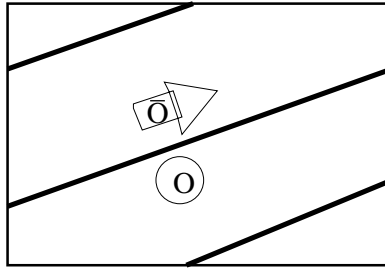
Fig. 46. Antipody in the Thurston manifold.

Fig. 47. An observer O at rest with respect to the compact spacelike hypersurface Σ and an observer $\bar{O}$ moving at constant velocity with respect to Σ on a periodic orbit of the torus. The orbit corresponds to $x_{\text{end}} = T_y T_x^2 x_{\text{start}}$.

7. Beyond standard cosmology

7.1. The twin paradox and compact time

So far we have ignored the issue of time. We have explicitly considered spacelike hypersurfaces Σ of constant curvature foliated by a natural conformal time. Compactification of these surfaces leads to the pictures we have described without compactifying time. However, an observer moving on an inertial worldline which is not at rest with respect to the cosmic expansion will perceive an identification which mixes spacetime coordinates as dictated by the Lorentz transformations Λ so that $\bar{x} = \Lambda x$ with $x = (\eta, \vec{x})$ comoving coordinates. As a result, a compactification of Σ in the comoving coordinates of the form $(\eta, \vec{x}) \rightarrow (\eta, \vec{x} + \vec{L})$ for instance will result in identification of a time shift as measured by the noncomoving observer of $(1 - \beta^2)^{-1/2}(\eta - \vec{\beta} \cdot \vec{x}) \rightarrow (1 - \beta^2)^{-1/2}(\eta - \vec{\beta} \cdot (\vec{x} + \vec{L}))$ where $\vec{\beta}$ is the velocity relative to Σ.

As an explicit demonstration, consider the twin paradox [130–132]. Let O be at rest in a comoving frame where a flat spacelike hypersurface is compactified into a torus. Let $\bar{O}$ be an inertial observer on a periodic orbit with respect to O as in Fig. 47. The periodic orbit will obey the boundary condition $\bar{x} = \Lambda x \rightarrow \Lambda \gamma x$ where γ is the corresponding word. From Section 2.3 we have $\gamma = T_y T_x^2$

(to include the physical time in the embedded coordinates and in T_x, T_y, see Ref. [130]). According to $\bar{O}$, both space and time points have been identified. As a result it becomes impossible for $\bar{O}$ to synchronize her clocks [133]. The lack of synchronicity will be given by the time component of $\Lambda(1 - \gamma)x$ [130].

Since both O and $\bar{O}$ are inertial, by the principles of relativity, each should believe the other's clocks run slower and therefore each could expect the other to be younger at their reunion leading to a paradox. However, the topological identification breaks the general invariance and selects a preferred frame, namely the frame in which the topological identification is purely on Σ. In that frame clocks can be synchronized and the volume of space looks smallest. The observer on the periodic orbit, though inertial, will discover that their clocks cannot be synchronized and this additional shift leads them to compute an older age for their twin. Both agree the twin at rest with respect to Σ is older.

This simple thought experiment emphasizes that compact topology selects a preferred frame, namely the frame in which the universe looks smallest and in which observers can synchronize their clocks [133] (see also Refs. [134]). To generalize to curved space, Λ can be replaced by an appropriate diffeomorphism and the spacetime topology generalizes to $\mathcal{M}_c = R \otimes G/\Gamma$.

Finite spaces reverse some of Copernicus' philosophical advances by selecting a preferred location at the center of the space, a preferred observer at rest with respect to the compactification and a preferred time. While Copernicus may have removed us from the center of the universe, topology puts some observers back there.

This also raises the question of compactifying time outright. Probably, time could be compactified so that there were closed timelike curves that were not causality violating. Very restrictive possibilities would result since only events which could repeat ad infinitum would obey the boundary conditions. It would be hard to envision a compact time model being consistent with the laws of thermodynamics, except perhaps on a cosmological timescale. The compact timescale would have to be much shorter or much larger than biological timescales or no children could sensibly be born since they would somehow have to grow young again. A universe could go through a big bang and eventual big crunch only to repeat the history of the cosmos with another big bang. The same galaxies, stars and the same people would be born, live and die. Fated to repeat their paths ad infinitum. Quantum gravity may reset the initial conditions at each big bang allowing new galaxies, new stars and new organisms to form, sparing us the relentless cosmic boredom. The fanciful possibility of a compact time magnifies the already strange and distinct nature of time.

7.2. Extra dimensions

It has been suggested that topology is a discrete feature of space [2]. As such, it may be better integrated in a quantized theory of gravity while there is no prediction for topology in classical relativity. The earliest attempts at creating a finite universe grew out of semiclassical quantum cosmology [135–138]. It may seem intuitive that a smaller universe should be easier to create from nothing than a larger universe. Therefore, the probability for creating a small, finite cosmos may be relatively high, if only we could compute the wave function. However, technical and conceptual difficulties dog the semiclassical approach such as defining a measure on the space of states, normalizing the wavefunction, etc. A convincing statement about the topology of the universe will very likely require a fully quantized theory of gravity.

Interestingly, additional dimensions have featured prominently in attempts to quantize gravity. These extra dimensions are always topologically compact and very small. While no quantum gravity theory is yet able to predict the topology of space, the possibility that compact internal dimensions will have topologically compact external (that is, large) dimensions is certainly alluring. Ultimately, a fundamental theory should predict the global topology of the entire manifold whether it be $3D$ or $11D$. In the meantime, the hierarchy between small and large dimensions remains mysterious although some recent suggestions have created a bit of a stir [139].

Kaluza–Klein theories introduced extra compact dimensions in an attempt to unify fundamental theories of physics [140]. Upon compactification, the radii of the small dimensions behave as scalar and tensor field theories. More modern string theories naturally invoke extra dimensions in a manner reminiscent of these early Kaluza–Klein models. Recent fervor in string phenomenology has involved compact extra dimensions of moderate to large size in an attempt to explain the hierarchy problem in standard particle physics. The hierarchy problem questions the disparity in scales from the Planck mass of 10^{19} GeV to the mass of the electron at a few electronvolts. A unification of fundamental theories has to naturally justify this span over 25 decades of energy scales. As in Ref. [141] we consider a spacetime with $4 + \mathcal{N}$ dimensions of the form $\mathcal{M} = R \otimes M^3 \otimes \mathcal{M}^{\mathcal{N}}$ where R represents time, M^3 is a constant curvature Friedman–Robertson–Walker metric and $\mathcal{M}^{\mathcal{N}} = G^{\mathcal{N}}/\Gamma$ is an $\mathcal{N}$-dimensional compact internal space. The $(4 + \mathcal{N})$-dimensional gravitational action becomes

$$A = \int \mathrm{d}^{(4+\mathcal{N})}x \sqrt{-g^{(4+\mathcal{N})}} \mathcal{R}^{(4+\mathcal{N})} M_*^2 \ . \tag{136}$$

In the simplest Kaluza–Klein picture, integration over the compact extra dimensions leads to the $(3 + 1)$-dimensional action, we experience

$$A \sim \int \mathrm{d}^4 x \sqrt{-g^{(4)}} \mathcal{R}^{(4)} M_{\mathrm{pl}}^2 + \cdots \tag{137}$$

plus additional dynamical terms where $M_{\mathrm{pl}}^2 = M_*^{2+\mathcal{N}} R^{\mathcal{N}}$ and R is the radius of the internal dimensions. The hierarchy between M_{pl} and M_* is large if RM_* is large [141]. The disparity in energy scales then becomes a dynamical and geometric question.

The modern ideas inspired by string theory involve the localization of matter to a 3-brane (a 3D extended object) nested in the full $(4 + \mathcal{N})$-dimensional space. The scale M_* is expected to be $\sim$TeV. Since ordinary matter is confined to a 3-brane, we would be unaware of these extra dimensions regardless of their size unless we try very hard to look for them. Some laboratory experiments are underway to probe any additional dimensions which may be lurking there.

The topology of the compact extra dimensions is not well understood. There are only a few requirements the internal dimensions must satisfy in order to break supersymmetry at a physically sensible scale. The internal spaces fall loosely under the broad category of Calabi–Yau manifolds. Nearly, all topologies investigated so far are modeled on flat geometries. An outgrowth of cosmic topology has been the suggestion that these internal dimensions be compact and negatively curved [141]. The hyperbolic internal spaces have some advantages over flat space including a less demanding tuning of geometric parameters and a suppression of astrophysically harmful graviton modes from the extra dimensions [141]. Another advantage of the compact hyperbolic extra dimensions over compact

flat extra dimensions may be that chaotic mixing on finite hyperbolic manifolds could explain the smoothness and flatness of the large dimensions as argued in Ref. [50]. We briefly discuss chaos in the next section.

The profound connection between small dimensions and large have only begun to be forged. As many times in the past, cosmology provides a unique terrain on which to test fundamental theories [142].

7.3. Chaos

We have touched upon the chaotic motions of particles on a compact hyperbolic space. Chaos refers to the thorough mixing of orbits which show an extreme sensitivity to initial conditions. Geodesics deviate on a surface of negative curvature as mentioned in Section 6 so that they are extremely sensitive to initial conditions. The motion becomes fully ergodic upon compactification of the surface. Formally, chaos in these Hamiltonian systems means the geodesics equations are nonintegrable; there is no smooth analytic function which can interpolate between orbits with different initial conditions.

Despite the resistance of chaotic systems to conventional integration methods a great deal about the structure of phase space can be determined. Much like thermodynamics, chaotic systems can be understood in terms of a set of states dense in the phase space. For chaos this set is provided by the collection of unstable periodic orbits which grow exponentially with length. All aperiodic orbits can be understood in terms of this special subset. The periodic orbits pack themselves into a fractal set in order to fit within the finite phase space. The fractional dimension maintains the set at a finite volume but allows for an infinite area. In this way fractals try to maximize the information content, so to speak, while minimizing the volume. The explicit fractal structure on a compact octagon was isolated in Ref. [130].

Pursuing the analogy with thermodynamics, entropy can be defined as

$$h(\mu) = \lim_{k \to \infty} \frac{1}{k} \sum_{i}^{N(k)} \mu \ln(1/\mu(k)) \ . \tag{138}$$

The calculation goes like this. Draw a fundamental domain. Now use the generators $g_1 \ldots g_n$ to tile $\mathbf{H}^3$ with copies. The measure at order k is then defined as the fraction of the total volume where each tile has the same volume:

$$\mu(k) = \frac{1}{N(k)} \tag{139}$$

and $N(k)$ is the number of unique tiles. The sum in Eq. (138) reduces to

$$h(\mu) = \lim_{k \to \infty} \frac{1}{k} \ln(N(k)) = H_{\mathrm{T}} \ . \tag{140}$$

This is equivalently the topological entropy, a symbolic entropy which counts the number of accessible states; that is, closed loops or equivalently tiles in the tessellation.

Recall that periodic orbits can then be counted symbolically with the homotopy group. We do not distinguish between loops of varying lengths if they are homotopic. The spectrum of periodic orbits so defined is therefore a topological feature. The number of unique words that can be built of length k out of an n-letter alphabet with r relations is equivalent to the number of neighbors at order k in the tiling. Not all of these neighbors will be unique. The repeats are accounted for by the relations and we know that the number of neighbors at order k is bounded by

$$(2n)k \geqslant N(k) \leqslant (2n)^k , \tag{141}$$

the lower bound being simple spaces such as the flat nonchaotic torus and the upper bound the unpruned maximally chaotic case. The topological entropy of the torus is zero while the entropy of an unpruned hyperbolic space is $H_T = \ln(2n)$. The compact octagon was shown to have a topological entropy of $H_T = \ln 7$ [130].

Another kind of entropy, the metric entropy introduced in Section 6, describes how quickly mixing takes place in a chaotic system and can be expressed as the sum of the positive Lyapunov exponents: $h = \sum \lambda$.

This is just a taste of a rich area in dynamical systems theory and the reader is referred to the many excellent texts on the subject [96,95].

Although cosmologists gingerly avoid this complex feature of the dynamics, chaos does have profound and unavoidable consequences. One of the original motivations for a compact topology exploited the chaotic motions. The chaotic mixing could lead to a dilution of initial anisotropies leading to the symmetric universe we observe today. These initial attempts failed since the space could not be made small enough to allow for sufficient mixing and still be the large cosmos we observe today [21]. Variants on this idea fuse the chaotic mixing with an $\Omega < 1$ inflation model where an early episode of chaotic mixing provides the moderate initial conditions needed to permit a subsequent inflationary phase to succeed [49].

We have already encountered the nonintegrability in trying to find eigenmodes on compact hyperbolic cosmologies. For the initial spectrum of fluctuations, researchers continue to use the assumption of a flat Gaussian spectrum as motivated by inflation. Still, since the consistency of inflation and an observably small topology is shaky, a more consistent approach would be to examine a distribution of initial fluctuations on a compact space in the absence of inflation. The numerical results of Section 6.1 do suggest that a Gaussian flat spectrum is in fact natural on compact hyperbolic spaces, even in the absence of inflation. While this work is very suggestive, a primordial quantum system has never been very thoroughly thought through. Because of the importance of this issue we take a moment to discuss quantum chaos. The quantum system is relevant to our discussion regardless since the expansion of fluctuations will be analogous to the expansion of the semiclassical wave function. The stationary Schrödinger equation is the usual Helmholtz Equation (92).

The quantization of the chaotic system is still not fully understood but interesting features have been conjectured and confirmed. Reminiscent of the Feynman path integral approach, the wavefunction can be thought of as the sum of classical trajectories. Since the classical trajectories chaotically mix in phase space, it has been conjectured that the wave function would be a random Gaussian function of the eigenvalue q with a spectrum given by the Wigner distribution function. Wigner's function is an attempt to generalize Boltzmann's formula for the classical distribution of a statistical system to a quantum system [143]. It provides a description of general attributes of wave functions

[95] and shows how the function tends to distribute itself over classical regions of phase space. Note the similarity between this suggestion and the method of images as well as the numerical results of Refs. [105,108].

The nature of the discrete spectrum of energy levels can also be related to the classical chaos. As already argued in the numerical work on the eigenmodes of a compact space, the number of eigenstates can be related to the volume of the space through Weyl's asymptotic formula. Additional features, such as the spacing between energy levels, can be directly associated with the underlying chaos.

Despite this conjecture of randomness, distinct remnants of the underlying chaotic dynamics have been found in the form of scars [144]. Scars are regions of enhanced intensity in high k states along periodic orbits. It is hard to imagine the very high k modes having cosmological significance since it is the low k modes which probe the largest cosmological distances. Scaring is expected to be less prominent in low k modes since the width of the enhancement will be correspondingly more diffuse. Although diffuse, scars on large scales could provide the small catalyst needed to order structure on large scale. Scarring on a 2D double doughnut was investigated in Ref. [93]. The underlying fractal structure of periodic orbits on the double doughnut was studied in detail in Ref. [130]. There it was suggested that the filamentary structure we observe in the distribution of galaxies may be a consequence of this slight enhancement of the seeds of structure formation.

For other important chaos articles see Refs. [145–149].

8. Summary

The creation of the universe is still not well understood. Clarity on the earliest moments will likely come only with a fully functional quantum theory of gravity. In the meantime, we know that space is curved and evolving and most also possess *some* topology. If the curvature of the universe falls within the observable horizon, then topology may also. The CMB provides the deepest probe of the universe on the largest scales and we have reviewed the many ideas on how to extract the topology of space from maps of the microwave sky. The methods fall into two primary categories: direct statistical methods and generic geometric methods. The salient features in CMB maps which reveal topology are (1) a discretization of the sizes of hot and cold spots, (2) a cutoff in the spectrum for wavelengths too big to fit within the finite space and (3) an anisotropic and inhomogeneous distribution of correlations corresponding to repeated ghost images of the same spots. Data from the future satellite missions MAP and *Planck Surveyor* are needed to determine if the CMB does in fact encode such features. When the new high-resolution maps are in hand, we will have to face the potentially prohibitive difficulty of foreground contaminations, the thickness of the surface last scattering, and fortuitous correlations. If we are lucky enough to surmount these observational trials, we may be able to see the entire shape of space.

Still, as many fear, the topology scale may naturally be far beyond the observable universe. If this is the case, we can turn to physics on the smallest scales to learn something about what we will never see on the largest scales. If an ultimate theory of gravity beyond Einstein's is able to predict the geometry and topology of small extra dimensions, there is every reason to hope we will learn, if only indirectly, the geometry and topology of the large dimensions.

Acknowledgements

Appendix A. Representations of hyperbolic space

Because of the recent emphasis on hyperbolic models, we include a section here on other useful coordinate models for manipulating $\mathbf{H}^3$. We have already seen two different coordinate systems for the metric equations (16) and (18). Another important coordinate system comes from embedding $\mathbf{H}^3$ in the $(3+1)$-Minkowski space as discusssed in Section 2.1. The Minkowski metric is

$$d\sigma^2 = -dx_0^2 + dx_1^2 + dx_2^2 + dx_3^2 \ . \tag{A.1}$$

The $3D$-hypersurface is constrained to have pseudo-radius

$$-x_0^2 + x_1^2 + x_2^2 + x_3^2 = 1 \tag{A.2}$$

and the curvature radius has been taken to 1. In Minkowski coordinates the generators $\gamma \in \Gamma$ take the convenient form of $O(4)$, orthogonal 4×4 matrices. These coordinates can be related to those of Eq. (18) with the transformation

$$x^\mu = \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} \cosh \chi \\ \sinh \chi \sin \theta \cos \phi \\ \sinh \chi \sin \theta \sin \phi \\ \sinh \chi \cos \theta \end{pmatrix} \ . \tag{A.3}$$

There are several other useful representations of the hyperbolic plane including the Poincaré ball with

$$\vec{x} = \tanh(r/2)\hat{n} \tag{A.4}$$

with $\hat{n}$ the usual unit vector in spherical coordinates. The Poincaré ball model maps $\mathbf{H}^3$ to the open ball $\{\vec{x} \in \mathbf{E}^3 | \vec{x} \cdot \vec{x} < 1\}$ with

$$d\sigma^2 = \frac{4d\vec{x} \cdot d\vec{x}}{(1 - \vec{x} \cdot \vec{x})} \ . \tag{A.5}$$

It is useful to know that the geodesic distance is

$$d(\vec{x}, \vec{x}') = \operatorname{arccosh} \left[1 + \frac{2|\vec{x} - \vec{x}'|}{(1 - |\vec{x}|^2)(1 - |\vec{x}'|^2)} \right] \ . \tag{A.6}$$

All geodesics intersect the boundary orthogonally and are therefore semicircles or straight lines which are the diameters of the ball.

There is also the upper half-space representation $\{\vec{\mathbf{x}} \in \mathbf{E}^3 | x_3 > 0\}$,

$$
\begin{aligned}
\mathrm{d}\sigma^2 &= \frac{(\mathrm{d} + x^2 + \mathrm{d}\,y^2 + \mathrm{d}\,z^2)}{z^2} \\
&\equiv \mathrm{d}\rho^2 + \mathrm{e}^{-2\rho}(\mathrm{d}x^2 + \mathrm{d}y^2)
\end{aligned}
\tag{A.7}
$$

with $\mathrm{e}^{\rho} = z$. This coordinate system is particularly useful for cusped manifolds which are discussed in Section 6.1.2.

Finally, there is the 3-dimensional Klein model with coordinates

$$
\vec{x} = \tanh \chi \hat{n} \,,
\tag{A.8}
$$

which is often used in constructing the Dirichlet domain in 3-dimensions and in mapping the periodic geodesics. Geodesics are mapped into straight lines in this model.

References

[1] F.C. Adams, G. Laughlin, Rev. Mod. Phys. 69 (1997) 337;
    L. Krauss, G.D. Starkman, Astrophys. J. 521 (2000) 22.
[2] M. Lachieze-Rey, J.-P. Luminet, Phys. Rep. 254 (1995) 136.
[3] N.J. Cornish, D.N. Spergel, G.D. Starkman, Class. Quant. Grav. 15 (1998) 2657.
[4] N.J. Cornish, D.N. Spergel, G.D. Starkman, gr-qc/9602039.
[5] N.J. Cornish, D.N. Spergel, G.D. Starkman, Phys. Rev. D 57 (1998) 5982.
[6] J. Levin, E. Scannapieco, J. Silk, Class. Quant. Grav. 15 (1998) 2689.
[7] J. Levin, E. Scannapieco, G. de Gasperis, J. Silk, J.D. Barrow, Phys. Rev. D 58 (1998) 123006.
[8] I. Heard, J. Levin, Proceedings for Cosmological Topology in Paris (CTP98), astro-ph/9907166.
[9] J.R. Bond, D. Pogosyan, T. Souradeep, Phys. Rev. D 62 (2000) article 043005.
[10] J.R. Bond, D. Pogosyan, T. Souradeep, Phys. Rev. D 62 (2000) article 042006.
[11] J.-P. Uzan, R. Lehoucq, J.-P. Luminet, in: E. Aubourg, T. Montmerle, J. Paul, P. Peter, (Eds.), Proceedings of the XIXth Texas meeting, Paris 14–18 December 1998, article-no: 04.25.
[12] J.-P. Luminet and B.F. Roukema, Proceedings of Cosmology School held at Cargese, Corsica, August 1998 astro-ph/9901364.
[13] B.F. Roukema, Bull. Astr. Soc. India, 28 (2000) 483 (astro-ph/0010185);
    B.F. Roukema, in: Ruffini et al. (Ed.), Marcel Grossmann IX Conference, 2001 (astro-ph/0010189).
[14] L.Z. Fang, H. Sato, Comm. Theor. Phys. (China) 2 (1983) 1055;
    H.V. Fagundes, Astrophys. J. 291 (1985) 450;
    H.V. Fagundes, Astrophys. J. 338 (1989) 618;
    H.V. Fagundes, U.F. Wichoski, Astrophys. J. Lett. 322 (1987) L5.
[15] H.V. Fagundes, Phys. Rev. Lett. 70 (1993) 1579;
    R. Lehoucq, M. Lachiese-Rey, J.-P. Luminet, Astron. Astrophys. 313 (1996) 339;
    B.F. Roukema, A.C. Edge, Mon. Not. Roy. Astron. Soc. 292 (1997) 105;
    B.F. Roukema, V. Blanloeil, Class. Quantum Grav. 15 (1998) 2645;
    B.F. Roukema, MNRAS 283 (1996) 1147;
    B.F. Roukema, S. Bajtlik, MNRAS 308 (1999) 309.
[16] G. Starkman (Ed.), The entire issue of Class. Quant. Grav. 15 (1998) 2589.
[17] W.P. Thurston, J.R. Weeks, Sci. Am. 251(1) (1984) 108 (July issue).
[18] J.P. Luminet, G.D. Starkman, J.R. Weeks, Sci. Am. 280(4) (1999) 90 (April issue).
[19] V. Blanlœil, B.F. Roukema (Eds.), Proceedings of the Cosmological Topology in Paris 1998 meeting, astro-ph/0010170, 2000.

[20] J.A. Wolf, Space of Constant Curvature, 5th Edition, Publish or Perish, Berkeley CA, 1994.
[21] G.F. Ellis, Quart. J. Rev. Astron. Soc. 16 (1975) 245;
     G.F.R. Ellis, Gen. Rel. Grav. 2 (1971) 7.
[22] W.P. Thurston, Bull. Am. Math. Soc. 6 (1982) 357.
[23] W.P. Thurston, in: Silvio Levy (Ed.), Three-Dimensional Geometry and Topology, Princeton University Press, Princeton, NJ, 1997.
[24] L. Bianchi, Mem. Soc. It. Della. Sc. (Dei. XL) 11 (1897) 267.
[25] J.R. Weeks, SnapPea: A computer program for creating and studying hyperbolic 3-manifolds, Univ. of Minnesota Geometry Center (freely available at http://www.northnet.org/weeks).
[26] G.D. Mostow, Ann. Math. Studies, Vol. 78, Princeton University Press, Princeton, NJ, 1973;
     G. Prasad, Invent. Math. 21 (1973) 255.
[27] J.R. Weeks, Ph.D. Thesis, Princeton University, Princeton, NJ, 1985.
[28] N.L. Balazs, A. Voros, Phys. Rep. 143 (1986) 109.
[29] J. Weeks, The mathematical section was contributed largely by Jeff Weeks who unfortunately was unable to coauthor this paper, private communication.
[30] G. Francis, J. Weeks, Amer. Math. Monthly 106 (1999) 393–399.
[31] J. Weeks, The Shape of Space, Marcel Dekker, New York, 1985.
[32] W.P. Thurston, Bull. Am. Math. Soc. 6 (1982) 357.
[33] P. Scott, London Math. Soc. 15 (1983) 401.
[34] E. Gausmann, R. Lehoucq, J.-P. Luminet, J.-P. Uzan, J. Weeks, Class. Quant. Grav. 18 (2001) 5155.
[35] A. Friedmann, Zeitschrift für Physik 21 (1924) 326.
[36] F. Löbell, Ber. d. Sächs. Akad. d. Wiss. 83 (1931) 167.
[37] C. Weber, H. Seifert, Math. Zeit. 37 (1933) 237.
[38] S. Kojima, Isometry Transformations of Hyperbolic 3-Manifolds, Topol. Appl. 29 (1988) 297.
[39] Miller, et al. (TOCO experiment), Astrophys. J. 524 (1999) L1.
[40] P. de Bernardis, et al., Proc. of the CAPP2000 Conference, Verbier, 17–28 July 2000, astro-ph/-0011469;
     J.R. Bond, et al., Astrophys. J. 545 (2000) L5;
     S. Hanany, et al., Phys. Rev. Lett. 86 (2001) 3475.
[41] A.H. Jaffe, et al. astro-ph/0007333.
[42] A. Vilenkin, Phys. Rev. D 27 (1983) 2848.
[43] A.D. Linde, Mod. Phys. Lett. A 1 (1986) 81;
     A.D. Linde, Phys. Lett. 175B (1986) 395.
[44] A.H. Guth, Phys. Rev. D 23 (1981) 347.
[45] A.H. Guth, Phys. Rep. 333 (2000) 555.
[46] C. Contaldi, astro-ph/0005115.
[47] I.Y. Sokolov, JETP Lett. 57 (1993) 617.
[48] D. Muller, H.V. Fagundes, R. Opher, Phys. Rev. D 63 (2001) 123508.
[49] N.J. Cornish, D.N. Spergel, G. Starkman, Phys. Rev. Lett. 77 (1996) 215.
[50] G.D. Starkman, D. Stojkovic, M. Trodden, Phys. Rev. Lett. 87 (2001) 231303.
[51] J.D. Barrow, H. Kodama, Class. Quant. Grav. 18 (2001) 1753;
     J.D. Barrow, H. Kodama, Int. J. Mod. Phys. D 10 (2001) 785.
[52] V.F. Mukhanov, H.A. Feldman, R.H. Brandenberger, Phys. Rep. 215 (1992) 203.
[53] W. Hu, in: E. Martinez-Gonzalez, J.L. Sanz (Eds.), The Universe at High-z, Large Scale Structure and the Cosmic Microwave Background, Springer, Berlin, 1995.
[54] A.R. Liddle, D.H. Lyth, Phys. Rep. 231 (1993) 1.
[55] M. White, W. Hu, Astron. Astrophys. 321 (1997) 8.
[56] R.M. Wald, General Relativity, The University of Chicago Press, Chicago, 1984.
[57] C.W. Misner, K.S. Thorne, J.A. Wheeler, Gravitation, Freeman, New York, 1970.
[58] S. Weinberg, Gravitation and Cosmology, Wiley, New York, 1972.
[59] E.W. Kolb, M.S. Turner, The Early Universe, Addison-Wesley, Reading, MA, 1994.
[60] I am grateful to Glenn Starkman for emphasizing this point.
[61] P.J.E. Peebles, The Large-Scale Structure of the Universe, Princeton University Press, Princeton, NJ, 1980.

[62] R.K. Sachs, A.M. Wolfe, Astrophys. J. 147 (1967) 73.

[63] J. Hwang, T. Padmanabhan, O. Lahav, H. Noh, Phys. Rev. D 65 (2002) 043005.

[64] J. Levin, J.D. Barrow, Class. Quantum Grav. 17 (2000) L61.

[65] A. Dekel, D. Burstein, S.D.M. White, in: N. Turok, N. Bahcall, X. Fan (Eds.), Critical Dialogues in Cosmology, World Scientific, Singapore, 1996 (astro-ph/9804082).

[66] S. Perlmutter, et al., Astrophys. J. 517 (1999) 565.

[67] A. Balbi, et. al., Astrophys. J. 545 (2000) L1.

[68] B.F. Roukema, G.A. Mamon, S. Bajtlik, Astron. & Astrophys. 382 (2002) 397.

[69] C.L. Bennett, et al., Astrophys. J. 391 (1992) 466.

[70] G.F. Smoot, et al., Astrophys. J. 360 (1990) 685;
G.F. Smoot, et al., Astrophys. J. 396 (1992) L1.

[71] K. Gorski, in: F.R. Bouchet et. al. (Eds.) Proceedings of the Moriond XVI, Gif-Sur-Yvette: Editions Frontièrs, 1997.

[72] M. Tegmark, Phys. Rev. D 55 (1997) 5895.

[73] J.R. Bond, A. Jaffe, L. Knox, preprint CfPA-97-TH-11;
J.R. Bond, A. Jaffe, L. Knox, in preparation;
J.R. Bond, A. Jaffe, in: F.R. Bouchet et. al. (Eds.), Proceedings of the Moriond XVI, Gif-Sur-Yvette: Editions Frontièrs, 1997.

[74] K.T. Inoue, Prog. Theoret. Phys. 106 (2001) 39.

[75] K.T. Inoue, Phys. Rev. D 103001 (2000) 1–15.

[76] J.R. Bond, G.P. Efstathiou, Mon. Not. R. Astron. Soc. 226 (1987) 655.

[77] G.I. Gomero, M.J. Reboucas, A.F.F. Teixeira, Phys. Lett. A 275 (2000) 355;
G.I. Gomero, M.J. Reboucas, A.F.F. Teixeira, Class. Quant. Grav. 18 (2001) 1885;
H. Fagundes, E. Gaussman, Phys. Lett. A 238 (1998) 235.

[78] J. Levin, E. Scannapieco, J. Silk, Phys. Rev. D 58 (1998) article 103516.

[79] E. Scannapieco, J. Levin, J. Silk, Mon. Not. R. Astron. Soc. 303 (1999) 797.

[80] D. Stevens, D. Scott, J. Silk, Phys. Rev. Lett. 71 (1993) 20.

[81] A. de Oliveira-Costa, G.F. Smoot, Astrophys. J. 448 (1995) 447.

[82] J. Cheeger, Problems in Analysis, (A Symposium in honor of S. Bochner), Princeton University Press, Princeton, NJ, 1970.

[83] P. Buser, Proc. Symp. Pure Math. 36 (1980) 29.

[84] G.D. Starkman, Quant. Grav. 15 (1998) 2529.

[85] Ya B. Zel'dovich, Comm. Astrophys. Space Sci. 5 (1973) 169.

[86] L.Z. Fang, M. Houjun, Mod. Phys. Lett. A 2 (1987) 229.

[87] A. de Oliveira Costa, G.F. Smoot, A.A. Starobinsky, Astrophys. J. 468 (1996) 457.

[88] A.A. Starobinsky, JETP Lett. 57 (1993) 622.

[89] L.Z. Fang, Mod. Phys. Lett. A 8 (1993) 2615.

[90] B.F. Roukema, Mon. Not. R. Astron. Soc. 312 (2000) 712;
B.F. Roukema, Class. Quant. Grav 17 (2000) 3951.

[91] G.F.R. Ellis, R. Tavakol, in: D. Hobbill, A. Burd, A. Coley (Eds.), Deterministic Chaos in General Relativity, Plenum Press, New York, 1994.

[92] Y.G. Sinai, Sov. Math. Dok. 1 (1960) 335.

[93] R. Aurich, F. Steiner, Physica D 39 (1989) 169;
R. Aurich, F. Steiner, Physica D 64 (1993) 185.

[94] K.T. Inoue, Classical and Quantum Gravity 18 (2001) 629.

[95] M.C. Gutzwiller, Chaos in Classical and Quantum Mechanics, Springer, New York, 1990;
M.C. Gutzwiller, J. Math. Phys. 12 (1971) 343.

[96] E. Ott, Chaos in Dynamical Systems, Cambridge University Press, Cambridge, 1993.

[97] D.H. Lyth, A. Woszczyna, Phys. Rev. D 52 (1995) 3338.

[98] K.M. Gorski, B. Ratra, N. Sugiyama, A.J. Banday, Astrophys. J. 444 (1995) L65.

[99] D.N. Spergel, Class. Quant. Grav. 15 (1998) 2589.

[100] F. Löbell, Ber, Sächs. Akad. Wiss. Leipzig 83 (1931).

[101] H. Seifert, C. Weber, Math. Z. 37 (1933) 237.

[102] D.D. Sokolov, A.A. Starobinskii, Sov. Astron. 19 (1976) 629.

[103] J. Levin, J.D. Barrow, E.F. Bunn, J. Silk, Phys. Rev. Lett. 79 (1997) 974.

[104] D. Olson, G.D. Starkman, Class. Quant. Grav. 17 (2001) 3093.

[105] K.T. Inoue, Class. Quantum Grav. 16 (1999) 3071.

[106] K.T. Inoue, K. Tomita, N. Sugiyama, Mon. Not. R. Astron. Soc. 314 (2000) L21.

[107] K.T. Inoue, Ph.D. Thesis, astro-ph 0103158.

[108] N.J. Cornish, D.N. Spergel, math.DG/9906017;
       N.J. Cornish, D.N. Spergel, Phys. Rev. D 62 (2000) 087304.

[109] R. Aurich, Astrophys. J. 524 (1999) 497.

[110] R. Aurich, J. Marklof, Physica D 92 (1996) 101.

[111] M.V. Berry, J. Phys. A 12 (1977) 2083;
       M.V. Berry, in: G. Iooss, R. Hellman, R. Stora (Eds.), Chaotic Behavior of Deterministic Systems, Les Houches Proc. 36, North-Holland, New York, 1981.

[112] C.B. Netterfield, M.J. Devlin, N. Jarosik, L.A. Page, E.J. Wollack, Astrophys. J. 474 (1997) 47.

[113] A. de Oliveira-Costa, M.J. Devlin, T. Herbit, A.D. Miller, C.B. Netterfield, L.A. Page, M. Tegmark, Astrophys. J. Lett. 509 (1998) L77.

[114] R. Aurich, F. Steiner, Mon. Not. R. Astron. Soc. 323 (2001) 1016.

[115] U.-L. Pen, Astrophys. J. 498 (1998) 60.

[116] J.R. Bond, D. Pogosyan, T. Souradeep, in: A. Olinto, J. Frieman, D.N. Schramm (Eds.), Proceedings of the XVIIIth Texas Symposium on Relativistic Astrophysics, World Scientific, Singapore, 1997.

[117] J.R. Bond, D. Pogosyan, T. Souradeep, Proceedings of the XXXIIIrd Recontre de Moriond, Fundamental Parameters in Cosmology, January 17–24, 1998, Les Arc, France.

[118] J.R. Bond, D. Pogosyan, T. Souradeep, Class. Quant. Grav. 15 (1998) 2671.

[119] I. Chavel, Eigenvalues in Riemannian Geometry, Academic Press, New York, 1984.

[120] P.H. Berard, Spectral Geometry: Direct and Inverse Problems, Lecture Notes in Mathematics, Vol. 1207, Springer, Berlin, 1980.

[121] H.V. Fagundes, Astrophys. J. 470 (1996) 43.

[122] H.V. Fagundes, IX Marcel Grossman Meeting, Rome, 2–8 July, 2000, astro-ph/0007443.

[123] L. Cayón, G. Smoot, Astrophys. J. 452 (1995) 494.

[124] C.D. Hodgson, J.R. Weeks, Exp. Math. 3 (1994) 261.

[125] G.I. Gomero, M.J. Reboucas, R. Tavakol, Class. Quant. Grav. 18 (2001) 4461;
       G.I. Gomero, M.J. Reboucas, R. Tavakol, Class. Quant. Grav. 18 (2001) L145.

[126] N.J. Cornish, J. Weeks, Notices of the American Mathematical Society, astro-ph/9807311.

[127] B.F. Roukema, A.C. Edge, Mon. Not. R. Astron. Soc. 292 (1997) 105.

[128] B.F. Roukema, V. Blan loeil, Class. Quant. Grav. 15 (1998) 2645.

[129] L.A. Best, Can. J. Math. 23 (3) (1971) 451.

[130] J.D. Barrow, J. Levin, Phys. Rev. A 63 (2001) 044104.

[131] J. Weeks, unpublished.

[132] J.P. Uzan, J.P. Luminet, R. Lehoucq, P. Peter, physics/0006039.

[133] P.C. Peters, Am. J. Phys. 51 (1983) 791;
       P.C. Peters, Am. J. Phys. 54 (1986) 334.

[134] J.R. Lucas, P.E. Hodgson, Spacetime and Electromagnetism, Oxford University Press, Oxford, 1990, pp. 76–83.

[135] S. Carlip, Class. Quant. Grav. 15 (1998) 2629, and references therein.

[136] G.W. Gibbons, Nucl. Phys. B 472 (1996) 683;
       G.W. Gibbons, Class. Quantum Grav. 15 (1998) 2605.

[137] H.V. Fagundes, S.S. e Costa, Gen. Rel. Grav. 31 (1999) 863.

[138] S.S. e Costa, H.V. Fagundes, Gen. Rel. Grav. 33 (2001) 1489.

[139] L. Randall, R. Sundrum, Phys. Rev. Lett. 83 (1999) 3370.

[140] Th. Kaluza, Situngsber. Preuss. Akad. Wiss. Phys. Math. Kl. 966 (1921);
       O. Klein, Z. Phys. 37 (1929) 895.

[141] N. Kaloper, J. March-Russell, G.D. Starkman, M. Trodden, Phys. Rev. Lett. 85 (2001) 928;
      G.D. Starkman, D. Stojkovic, M. Trodden, Phys. Rev. D 63 (2001) 103 511.
[142] G.T. Horowitz, D. Marolf, J. High Energy Phys. 9807 (1998) 014.
[143] E.P. Wigner, Phys. Rev. 40 (1932) 749.
[144] E.J. Heller, Phys. Rev. Lett. 53 (1984) 1515.
[145] G.F.R. Ellis, R. Tavakol, Class. Quantum Grav. 11 (1994) 675.
[146] G.F.R. Ellis, G. Schreiber, Phys. Lett. A 115 (1986) 97.
[147] D. Hobill et al. (Eds.), Deterministic Chaos in General Relativity, Plenum Press, New York, 1994.
[148] J. Richard Gott, III Mon. Not. R. Astron. Soc. 192 (1980) 153.
[149] C.M. Lockhart, B. Misra, I. Prigogine, Phys. Rev. D 25 (1982) 921.