How to see 3-manifolds

1998 Class. Quantum Grav. 15 2545

(http://iopscience.iop.org/0264-9381/15/9/004)

Class. Quantum Grav. **15** (1998) 2545–2571. Printed in the UK PII: S0264-9381(98)95879-8

How to see 3-manifolds

William P Thurston†

Mathematics Department, University of California at Davis, Davis, CA 95616, USA

Received 10 July 1998

Abstract. There have been great strides made over the past 20 years in the understanding of three-dimensional topology, by translating topology into geometry. Even though a lot remains to be done, we already have an excellent working understanding of 3-manifolds. Our spatial imagination, aided by computers, is a critical tool, for the human mind is surprisingly well equipped with a bit of training and suggestion, to 'see' the kinds of geometry that are needed for 3-manifold topology.

This paper is not about the theory but instead about the phenomenology of 3-manifolds, addressing the question 'What are 3-manifolds like?' rather than 'What facts can currently be proven about 3-manifolds?'

The best currently available experimental tool for exploring 3-manifolds is Jeff Weeks' program SnapPea. Experiments with SnapPea suggest that there may be an overall structure for the totality of 3-manifolds whose backbone is made of lattices contained in $PSL(2, \mathbb{Q})$.

PACS number: 0240

1. Introduction

Training the imagination. Our mental facilities for geometry and vision are remarkable. From an impersonal perspective, the act of walking through a crowd to meet up with someone on the other side is truly astounding. It is a far greater achievement than any merely intellectual achievement such as writing a PhD thesis in mathematics. Writing a PhD thesis is a great *intellectual* challenge, but the powerful intelligence needed to walk through a crowd is something that pooled human *intellect* has not come close to matching, despite years of effort and massive investment in related technologies.

Of course, we are no more able to program computers to 'do mathematics' than we are able to program them to walk down the street. Computers are powerful tools in mathematics, particularly for the symbolic aspects of mathematics (including numerical computations). It is ironic that the human use of symbols has often been touted as the unique human characteristic that makes us special. But our minds are complex organs, composed of many different cooperating modules. It is not the equations, symbols or logic that are hardest for computers, but the seemingly 'low-level' foundations of perception that prove hardest to match. 'Imagination', 'intuition' and 'instinct' are some of the words that are often used to allude to some of these perceptual foundations.

Geometric imagination is a powerful tool for three-dimensional geometry and topology—provided we teach it the foreign imagery it wants and needs to work with our intellect instead of rebelling. Our spatial–geometric instincts are rather strong-headed, and if we do not bring them along, they are bound to rebel.

† E-mail address: wpt@math.ucdavis.edu

Hyperbolic manifolds. Hyperbolic geometry, also called Lobachevskian geometry or non-Euclidean geometry, is the geometry of a complete Riemannian metric on $\mathbb{R}^3$ with constant curvature -1. (A metric is *complete* when you cannot get to the edge of space in a finite distance, that is, a metric ball of any finite radius is compact.) Hyperbolic space is symbolized as $\mathbb{H}^3$. This geometry is the crucial tool for seeing three-dimensional topology. We will visit $\mathbb{H}^3$ a little later and look around.

Most 3-manifolds are hyperbolic

A manifold is *hyperbolic* when it has been given a metric that locally is identical to $\mathbb{H}^3$. In other words, a small neighbourhood of any point matches hyperbolic space on the nose. This is the same as a metric of constant sectional curvature -1. But a more revealing description is that a hyperbolic 3-manifold comes from $\mathbb{H}^3$ modulo a discrete group of isometries, where all points equivalent under the group are identified to a single point of the manifold. One can think of the group 'rolling up' space into a compact bundle, as if it is a big floppy rug but much neater.

The geometry of our minds. Whether by design or accident, we have geometric modules in our minds that are remarkably well suited for use inside hyperbolic space and inside hyperbolic 3-manifolds. Three modules in particular—our sense of perspective, our sense of scaling and our sense of symmetry—connect directly to what we see when we visit a hyperbolic 3-manifold. Our minds must adjust their interpretations of these perceptual modules, but our brains are plastic and can readily make the needed adjustments.

Geometry gives us an understanding of 3-manifolds that we topologists of 25 and more years ago never imagined possible. This paper aims to present some of the basic geometric vision that lets us see, know and understand 3-manifolds. Predominantly, this means to see, know and understand hyperbolic 3-manifolds.

Disclaimer 1: other geometries. Not *all* 3-manifolds are hyperbolic. There are actually eight different flavours of three-dimensional geometry, describing eight different classes of 3-manifolds. Maybe the reason that for a long time nobody suspected that hyperbolic 3-manifolds are plentiful is that the simplest 3-manifolds are not hyperbolic, but typically have one of the seven non-hyperbolic flavours of geometry. It is easy to deduce that a manifold that has one of the seven non-hyperbolic flavours is topologically quite special. It is natural to assume by analogy that hyperbolic manifolds are special. They are indeed special, but they are plentiful.

We are lucky that most flavours are rare, because the geometry modules adapt much better to Euclidean space and hyperbolic space than to any of the other geometries†.

A 3-manifold is *geometric* if it has a Riemannian metric for which any two points have neighbourhoods that are identical. A metric satisfying this condition is a *locally homogeneous* metric. Any compact manifold with a locally homogeneous metric has other locally homogeneous metrics: you can always scale the metric by a constant factor. The total volume changes, so the new metric cannot be isometric with the old one. Some locally homogeneous metrics can be modulated in additional ways, while remaining locally

† Spherical geometry (S^3) is the runner-up, but S^3 when naively rendered is quite disorienting to visit because very distant objects look as big as very close objects. Our perception continues to rebel against this phenomenon long after our intellect has accepted it. Real-time special effects with fog and focus would undoubtedly remedy much of this difficulty.

homogeneous. A locally homogeneous metric is *isotropic* if there is a local isometry taking any given direction to any other direction, and *anisotropic* if there is some proper subset of directions distinguishable by the local geometry. In dimension 3, the tangent space to an anisotropic manifold has a line field constructible from the local geometry, and an orthogonal field of 2-planes. These scale factors in the line and the plane can be modulated independently. These variations are why we talk about 'flavours' of geometry rather than about the exact shapes of metrics.

Only three of the eight flavours are isotropic: spherical geometry, Euclidean geometry and hyperbolic geometry. It is a curious fact that all spherical 3-manifold (manifolds with metrics of constant positive curvature $+1$) also possess anisotropic locally homogeneous metrics. Euclidean 3-manifolds do not have anisotropic locally homogeneous metrics, but they all have a line field that is globally parallel to itself. In a certain sense, hyperbolic manifolds are the only ones that are genuinely anisotropic.

The anisotropic flavours of geometry, and 3-manifolds having those flavours, have the form either of a product or a fibre bundle, combining one of the two-dimensional geometries with one-dimensional geometry perhaps with one or another type of twisting.

Disclaimer 2: decomposable manifolds. Not *all* 3-manifolds are even geometric. There are two topological processes for joining geometric 3-manifolds, to form new, 'compound', 3-manifolds that do not have geometric structures.

The first compounding process is the *connected sum*, which means to remove balls from two 3-manifolds, and join their bounding spheres together—in other words, connecting them by a tunnel or wormhole whose cross section is the 2-sphere S^2.

The second process involves joining 3-manifolds with a boundary, to obtain a new manifold that may or may not have a boundary. A surface inside or on the boundary of a 3-manifold is *incompressible* if a curve on the surface cannot be the boundary of a disc in the 3-manifold unless it is already the boundary of a disc on the surface. For example, the surface of a doughnut in space is an example of a *compressible* torus, because for example a disc that cuts across the doughnut hole has a boundary that is a non-trivial curve on the torus. A *torus sum* of 3-manifolds is something that results from gluing together two incompressible component tori that are components of the boundary. This creates a second kind of tunnel or wormhole, whose cross section is a torus, T^2. Tunnels with cross sections that are multi-holed tori with two or more holes have no special status, because these surfaces do not interfere with geometric structures.

There is a complete topological theory of how to analyse 3-manifolds that have been combined by these two processes: topological criteria to recognize and undo these two ways of joining and to recover the original pieces. Undoing connected sums is called the *prime decomposition* of a 3-manifold. The theory of the prime decomposition was analysed by Kneser in the 1930s. A *prime* 3-manifold is a manifold which cannot be expressed as a connected sum except in a trivial way. Every compact 3-manifold is the connected sum of finitely many prime pieces, and the prime pieces are determined up to homomorphism by the 3-manifold.

The second decomposition process, undoing tunnels whose cross section is a torus, is called the *torus decomposition*, and was analysed in the 1970s, by Jaco, Shalen and Johannson. The theory is analogous to the prime decomposition, but we will skip giving the exact specifications, which are a little more involved.

Disclaimer 3: unproven. It has not been proved that *all* 3-manifolds are composed of geometric pieces. A number of years ago I proposed the *geometrization conjecture*, that the primitive parts produced by the prime decomposition and torus decomposition of any 3-manifold are always geometric.

To prove the geometrization conjecture in full generality is a great challenge in topology. One special case is the Poincaré conjecture, which has long been notorious as a tempting lure surrounded by hidden traps.

If you are a consumer rather than developer of 3-manifold theory, the status of the geometrization conjecture is probably fairly academic, because for most purposes you may as well assume it is true of the manifolds you need to use. The conjecture is overwhelmingly supported by the evidence we have seen to date. There is theoretical evidence, which has established the existence of the geometric decomposition for several broad classes of 3-manifolds. In addition, large numbers of particular examples have been tested. Mathematics is full of surprises, so we can never be certain of the geometrization conjecture until and unless it is proven. Nonetheless, the geometrization conjecture is a safe working hypothesis, and if you need to actually know for a particular case, you should probably just compute its structure anyway.

The words and the reality. The full description of the geometrization conjecture sounds complicated, since it involves two different decomposition processes and the eight geometries. Do not be fooled by this. The complexity of the verbal description is mismatched with the actual complexity of 3-manifolds. The prime decomposition and the torus decomposition may sound complex, but they are actually very orderly and straightforward processes. Similarly, the geometric manifolds having any of the seven non-hyperbolic flavours have been completely classified in an orderly and understandable way.

The true complexity of the structure of 3-manifolds—at least if the geometrization conjecture holds—is the structure of hyperbolic 3-manifolds.

The geometry of a hyperbolic 3-manifold is a topological invariant, according to the Mostow rigidity theorem (extended to the non-compact lattices by Prasad). This fact allows one to extract a great deal of topological information; it makes it quite easy to tell whether or not two hyperbolic 3-manifolds are homeomorphic. The volume, in particular, gives an interesting measure that describes in a certain way the three-dimensional complexity of the manifold. Volumes seem to all be irrational. There are some rational relationships among volumes of different manifolds, but most pairs of manifolds have volumes that do not seem to be in any rational ratio. The set of all possible volumes is a countable, closed subset of $\mathbb{R}$. Each volume that occurs does so for only a finite set of examples. The accumulation points of volume are manifolds which have ‘cusps’—they are non-compact, with exponentially shrinking tubes going out to infinity whose cross sections are tori. The phenomenon of this convergence is part of the theory of continuous surgery on hyperbolic manifolds, which we will be experiencing.

Appearance, reality and imagination. It is a wonderful dream to see the topology of the universe some day. However, this paper is not about the topology of the physical universe, but about topology in our minds. We will imagine bending space, but this has the effect of bending the imagination. The purpose is not for science fiction diversions, but to develop true vision meeting the precise standards demanded by three-dimensional geometry and topology. Our instincts about appearance and reality are strong, so our minds need time and exposure to adjust to new possibilities.

The scale in our imagination can make a big difference in our thinking. An effective strategy is to think about 3-manifolds on the scales we might inhabit: perhaps the size of a house, the size of a stadium, or the size of a town. It is harder to attend as seriously to objects the size you might hold in your hand. It is interesting to sit back and imagine your surroundings—the streets and the land in the neighbourhood where you live—and then think of the same degree of imagination about a teacup. At the opposite pole, very huge objects—the size of the universe, or even of the earth—are so far removed from everyday experience that our imagination on those scales tends to be abstract and distant.

Other resources. The rather unique video 'Not Knot' [Geo81] journeys through territory that overlaps some of where we will go; it is worth seeking it out (A K Peters is the current publisher). Clips and pointers to this and other relevant resources, including some good downloadable software, can be found on the Geometry Center website at www.geom.umn.edu. My book *Three-dimensional Geometry and Topology* [Thu97] develops the broad sweep of three-dimensional topology from the geometric perspective. A 1982 review article [Thu82] summarized the geometric theory of 3-manifolds at that time. Theory has advanced since then, the broad picture has not changed, but the basic picture is close enough that it is still recommended as a summary of the geometric theory of 3-manifolds. See also the online version [Thu79] of lecture notes from seminars I gave in Princeton 1978–81. There is a large literature on the geometry of 3-manifolds that I will not even attempt to review; MathSciNet (http://www.ams.org/mathscinet/) is a much better window than anything I could write. I will only mention three of my own primary contributions [Thu86, Thu98a, Thu98b].

2. Geometry from the inside

Imagine walking in a barren desert when you see the space in front of you begin to shift. You are startled, and stop. You see a vertical, straight fracture where the left side does not quite match the right: the images overlap ever so slightly. At first you think your vision has gone bad, maybe you have become cross-eyed. However, when you turn your head and move from side to side, the fracture does not turn or move with your head and eyes. When you circle around at a wide distance, you see that the fracture is not fixed on the ground or on the distance scenery, but is localized on a line going straight up into the sky.

Paper models. When you get back to camp, you can make a model to help explain what you have seen. Cut a 350° sector of paper (i.e. a disc with a 10° angle removed) with edges joined to form a blunt cone. This is the cone with *cone angle* 350° or *curvature* 10°. You can trace geodesics on this cone either by stretching pieces of string, by flattening portions of it and drawing straight lines with a ruler or by using a folded strip of paper as a ruler that is flexible enough to fit the surface. If you draw geodesics that emanate from a point on the cone at an angle less than 10° and aimed to the left and right of the apex, they will cross again behind the apex. This behaviour is identical with the behaviour of geodesics in a certain distorted metric in the plane, as shown in figure 1.

In the same way, we can think of a distorted metric in space that is equivalent to the metric we would get from a sector, of say 359° around our singular line. This affects what you see, since light 'bends' to follow the geodesics in the metric; it is as if light near the axis is slowed down.

Return to the axis. Overcome by curiosity, you return to gaze at the singular axis. The singularity has progressed dramatically, and as you watch, the scenes along the two sides

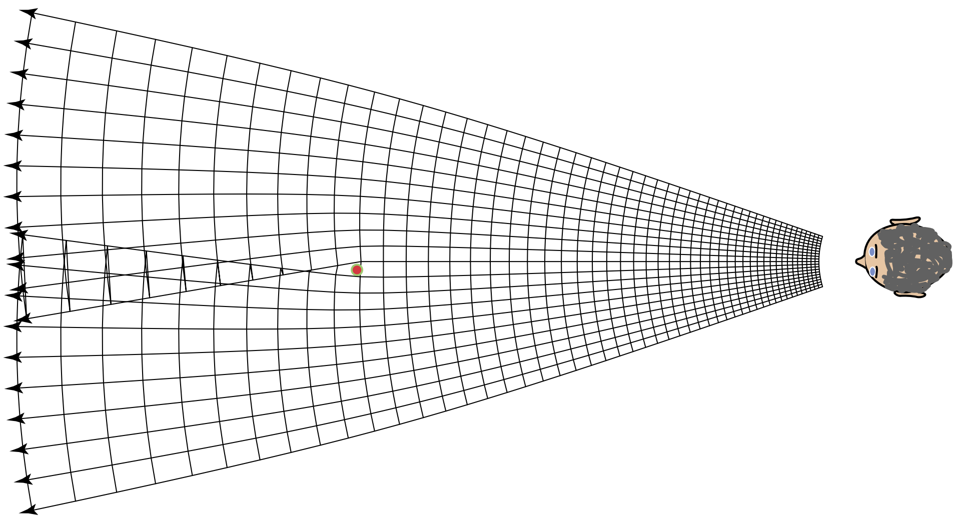


Figure 1. This figure shows paths of light near the apex of a cone. The metric of the cone is represented as a distortion of the metric in the plane. The long, nearly straight curves are geodesics, and the orthogonal trajectories can be thought of as wavefronts.

of the singular line scroll outward, squeezing into the surrounding panorama at your sides and behind you. You are startled to see arms and shoulders growing outward from the axis, then feet, a torso—and suddenly, there is your own head staring back! The image snaps into perfect alignment, and the motion stops. In a daze, you slowly turn around to take stock. Every single thing is repeated twice. For each boulder, there is an opposite boulder. Each mountain has an opposite twin mountain. There are even twin suns in the sky, shining from directions 180° apart, lighting up every shadow.

You realize that the singular line is a 180° cone axis (figure 2). Light 'bends' so much that you see another image of yourself directly behind the axis—lines of sight that go just to the right of the singular axis make a complete U-turn, heading almost straight back to you, so that you see the left side of your face. Similarly, your line of sight just to the left of the singular line, you see the right side of your face. With cone angle π, the two halves fit together seamlessly.

This is an example of an *orbifold*, which is a space locally modelled on $\mathbb{E}^n/\Gamma$, where Γ is a finite group of symmetries. Here $\mathbb{E}^n$ symbolizes Euclidean space equipped with its standard metric, and has a connotation somewhat different than $\mathbb{R}^n$. Orbifolds are generalizations of manifolds having the advantage that many phenomena are illustrated with much simpler examples of orbifolds than of manifolds.

Now the cone angle starts decreasing even further. The images to the left and right of the cone axis go out of sync once again (figure 3), but when the cone angle reaches $2\pi/3$, they are again coordinated and match each other seamlessly. You are in the orbifold $\mathbb{E}^3/C_3$, where C_3 is the cyclic group with three elements (that is, the group of integers modulo 3). As the cone angle continues to decrease, the images match when the cone angle is $2\pi/4, 2\pi/5, \ldots, 2\pi/n$. Figure 4 shows the case of cone angle $2\pi/7$.

The effect is reminiscent of two mirrors that meet at an edge, when the angle between them is varied. A single mirror can be thought of as a model for the orbifold $\mathbb{E}^3/C_2$, where

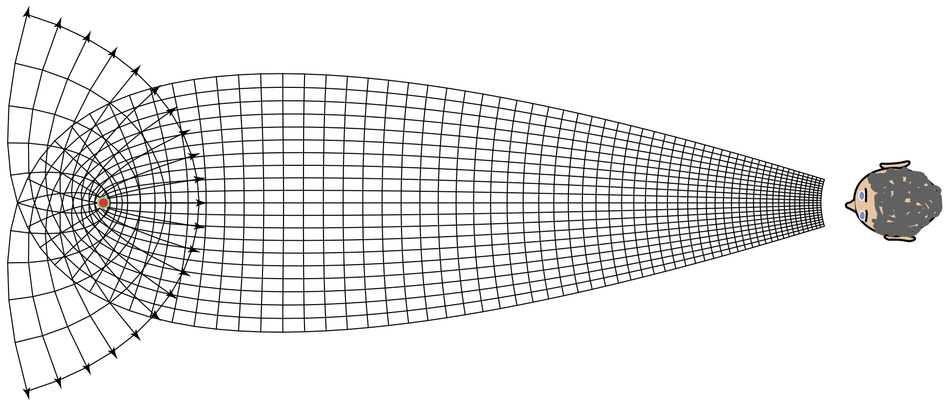


Figure 2. The behaviour of geodesics near the apex of a cone with cone angle π. The geodesics just to the left and right of the apex align perfectly. What you see is a perfect image of yourself. Unlike your reflection in a mirror, this image has the same orientation that you do.

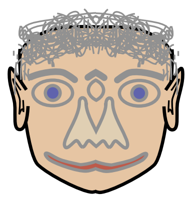


Figure 3. When there is a cone axis with angle $\pi - \epsilon$, you see an image of yourself on either side of the cone axis that is doubly covered in a narrow zone in the middle.

C_2 acts by reflection through the plane of the mirror. Two mirrors at an angle π/n give a model for the orbifold $\mathbb{E}^3/D_n$, where D_n is the dihedral group of $2n$ elements, generated by reflection in the two mirrors. Since $C_n \subset D_n$ is a subgroup of index 2, the optics of a cone angle of $2\pi/n$ can be visualized by ignoring every other image as seen in two mirrors that meet at an angle π/n (that is, ignore all images that have the opposite orientation from you).

Reality and appearance. Despite *Through the Looking Glass*, there is a significant distinction between the effects of a cone axis and the effects of mirrors. When you go with a companion to visit a cone axis with cone angle $\pi/2$, you see four images of your companion and three of yourself, plus your identity image. *The four images of your companion are all equally real*. You can walk in a straight line toward any of the images of your companion and shake hands. You can turn around and head toward any other image and shake hands again.

It is not quite the same if you try to shake hands with one of your own images, because whenever you move toward your image, your image moves away. If you stubbornly keep following anyway, you and your three images chase each other in an inward spiral until you eventually are close enough to the cone axis that you can reach out with your left hand

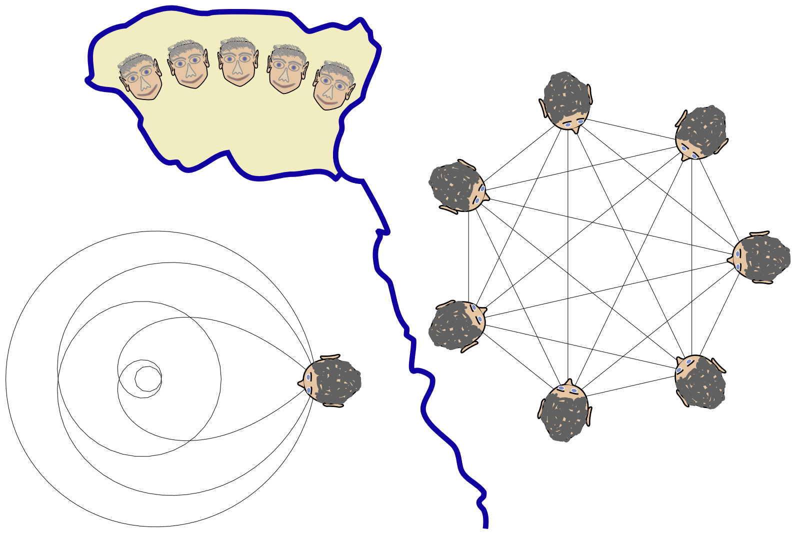


Figure 4. When a cone axis has cone angle $2\pi/7$, you see six images of yourself coming from geodesics that go ±1, ±2 or ±3 turns around the axis before reaching you. The appearance is the same as if you are part of a ring of septuplets. At the top is a snapshot of you near an n-fold axis, taken by a friend standing on a boulder behind you. Since the images are in perfect alignment, the cone axis is not visible.

and shake your right, forming a ring around the axis. You know that really there is only one of you, but it will take some time for your brain to readjust its model of how the world relates to seeing and touching.

The limit. With the right choice of normalization, the cone angle can keep shrinking all the way to 0 without crushing or impossibly distorting the space around you. The trick is that as the cone angle decreases, the cone axis recedes away from you further and further into the distance. To normalize, watch the cylinder through you around the cylinder axis. Initially, if the line of singularities is a distance r away from you, the cylinder has circumference $2\pi r$. For cone angle α, if the cone axis moves to a distance $2\pi r/\alpha$, the circumference of the cylinder in the cone manifold is still $2\pi r$; the intrinsic geometry of this cylinder stays the same, but the cylinder becomes less curved as $\alpha \to 0$. In the limit, the cone axis moves infinitely far away and vanishes and the two half-planes converge to parallel planes, and the cylinder you are on flattens out and become planar. The resulting space is now a manifold, $\mathbb{E}^3/\mathbb{Z}$, which can be thought of as a different metric for $\mathbb{R}^3 - \mathbb{R}$. What you see is that every object has an infinite repeating sequences of images lined up in a horizontal straight line, It is reminiscent of a barber shop with two opposite and parallel mirrors, although different because there is no obstacle to reaching an image of anything other than yourself.

Curvature can buffer from crushing. This choice of normalization protects you, but rather unfortunately, most of $\mathbb{E}^3$ is cataclysmically distorted. If you want to protect the rest of space from experiencing unbounded distortion, the only way is to give up local Euclidean geometry and allow space to become curved. With curvature, it is easy. To distort $\mathbb{E}^2$ in this way, just rest a cone on the plane with its point pointing up and round it off like a conical mountain. As the cone angle deforms to 0, the surface develops a long tube, growing like an asparagus sprout until it is asymptotically a cylinder.

These surfaces have a net negative curvature that exactly balances the curvature at the cone point; in the limit, the total curvature is -2π. In three dimensions, the product of the two-dimensional metrics with a line accomplishes a similar result. Whenever a cone angle along a singular axis decreases, the axis develops an increasing concentration of positive curvature, and it 'wants' to build a balancing cloud of negative curvature in planes transverse to the axis.

The process will work with great fluidity once we turn to the hyperbolic space, whose fabric is negatively curved.

A new cone axis appears. The space around us has transformed into $\mathbb{E}^3/\mathbb{Z}$, obtained by periodically identifying all points in $\mathbb{E}^3$ by a translation along a horizontal axis. The translation axis wraps around to form a closed geodesic loop C. Make sure you are standing some distance away from C. Now watch as space starts to develop a new cone singularity along C whose cone angle slowly decreases from 2π. The geometry of the changing metric can be represented concretely by slicing along a 'half-plane' radiating from C—that is, the surface has the local geometry of a half-plane, but it actually wraps around to form a half-infinite cylinder. After the singular metric is sliced along one of these planar cylinders, it matches the metric of an almost 360° wedge in the non-singular metric of $\mathbb{E}^3/\mathbb{Z}$ bounded by two planar cylinders. The singular metric is reconstructed by gluing the two cylinders back together.

The cone angle keeps on decreasing, while the geometry of space away from the singular axis maintains its locally Euclidean nature. When the cone angle reaches π, the cone manifold becomes an orbifold. At this instant, all visual images of any object are coordinated. You see two rows of images of yourself, rotated by 180° about the axis. Since the axis is horizontal, the 'other' row is upside down.

Carrying on, whenever the cone angle is $2\pi/n$, the images are again perfect: you see n rows of images that are symmetric by order n rotations about C, and repeating by translations along C. We make sure that the cone axis moves further and further away as $\alpha \to 0$, so that there is a limit for cone angle 0, normalized in a way that keeps the neighbouring rows of images so they are spaced a constant distance apart. The cylinder of points a constant distance from the axis is actually wrapped around to form a torus. In the limit, this torus flattens out and becomes planar. What one actually sees is a doubly periodic set of images of any object, repeating by an action of $\mathbb{Z}^2$ acting as a discrete group of translations of $\mathbb{E}^2$.

The Euclidean geometry of $\mathbb{E}^3$ is like a harbour on the edge of a small island in the middle of a vast hyperbolic ocean of hyperbolic manifolds. So far, we have taken a short stroll along the edge of the harbour. There are many fascinating tourist attractions here on the island—for instance, it is only a short trip from our hotel ($T^2 \times \mathbb{E}^1$) to the top of the mountain, S^3 that looms over us—but time is limited and we have far to go, so we will be wiser to work on preparations for going to sea.

3. Moving in two directions: continuous surgery

There are so many 3-manifolds that it is very easy to lose your bearings. It will help if we upgrade our navigational capabilities to be able to modulate the geometry at singular axes in the 'sideways' directions, in addition to increasing and decreasing cone angles. Without the new capability we would still be able to go wherever we like, just as even someone who is stubborn about choice of airline could still fly the 90 miles from San Francisco to Sacramento—by detouring thousands of miles through Chicago or Atlanta.

The new process is similar to what we have already encountered. Take a wedge of space between two half-planes, and glue one to the other, but shift one side up slightly with respect to the other. Optically, this creates a vertical displacement between images seen on the two sides of the axis. If the original cone angle was 2π, for example, and you shift the side to your right upward before identifying it with the left, then you see an image of yourself through the axis, with the portion of the image to your right shifted down (see figure 5).

Figure 5. This is the image seen when there an axis with cone angle π slightly shifted parallel to the vertical axis, with the side to your right joined to a slightly higher level on the right. Your lines of sight to the right of the axis whip around and come back at a higher level, so the topmost portion of your image is lower.

If we shift along an axis that is closed in a loop, as in $\mathbb{E}^3/2\mathbb{Z}$, the same effect occurs, but the detailed behaviour near the singular axis has a somewhat bizarre description. To effect the shift, cut along a half-infinite cylinder whose boundary is the axis and reglue with a vertical shift of a. Every point on the axis itself is identified with the point on the axis a distance a higher, so also with $2a$ higher, etc. If a is incommensurable with the length L of the axis, then these identifications are dense along the axis, so the axis collapses to a single point. But if we shift by a rational multiple $a = (p/q)L$ of the length of the axis, the axis now wraps around itself q times. Each point on the resulting axis has a neighbourhood formed from q wedges of the original axis, so the cone angle is multiplied by q. If the axis was non-singular before the shift, this results in a cone axis with a large cone angle $2\pi q$, bigger than 2π. On the other hand, if we start with a $2\pi/q$ (orbifold) axis and shift by $a = (p/q)L$, the new cone angle is 2π, and the axis becomes non-singular!

We can watch this process happening, starting with say with a sevenfold cone axis $\mathbb{E}^3/(\mathbb{Z} \times C_7)$. Let us orient the axis vertically. You are part of a ring of septuplets circled around the axis. Up above, you see another ring of images of yourself, and above them, yet another. Now the geometry begins to shear along the axis. Your ring now spirals slightly upward to your right, so the third septuplet to your right is slightly higher than the next one around, the third to your left. Since these images come from light paths that actually wrap several times around the axis before coming back to you, they shifted every time they hit the vertical half-cylinder which has been sliced and reglued. the vertical misalignment of images you see is a sevenfold amplification of the shift of the regluing. The shearing increases; when it reaches $L/7$, your images align again in an infinite spiral, very similar to a steep spiral staircase. The shearing continues. When the vertical shift reaches $2L/7$, the

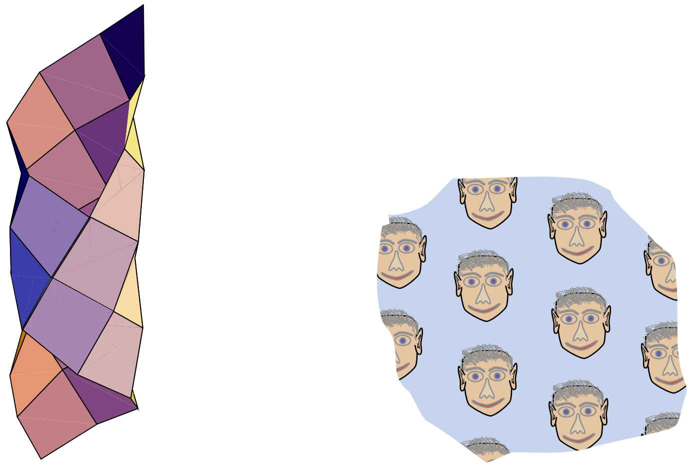


Figure 6. Left: $\mathbb{E}^3/\mathbb{Z}$ has undergone a compound shifting along a circular axis, first developing a $2\pi/5$ cone angle, then shearing vertically by $\frac{3}{5}$ the length of the axis. The central axis is not singular, but it is only $\frac{1}{5}$ as long as it used to be. You see multiple images of a single square whose sides are joined to form a torus. Before space shifted, this square extended full height and full circle around the original non-singular circular axis. Right: you have gone to investigate another compound axis close up. It is reassuring to belong to a whole flock of like-minded souls.

images go full circle after a vertical distance of $2L$, and you see two intertwined spirals, each twice as steep as the staircase before. The spirals are diametrically opposite. No two people are on the same height, but they alternate between the two spirals. When $a = 3L/7$, the images once again are all in alignment. You can make out three intertwined right-handed spirals, but this visual organization is not as automatic; you can see it instead as a parallelogram-grid of images wrapped around a cylinder. Perhaps you even see four intertwined left-handed spirals.

Instead of cutting and regluing, we can think of this operation as deforming a metric on a fixed manifold, while always retaining its locally Euclidean character. Focus on the set of distance r from the cone axis; topologically this is a torus. Consider how we can modify the three-dimensional metric (singular at the core), while maintaining a constant two-dimensional intrinsic geometry on our chosen torus. The deformation in three dimensions will be determined by how the torus is bent; the bending of the torus always matches a cylinder. The bending is determined by specifying the two principal directions of curvature and the two principal curvatures. One of these principal directions has curvature 0 and points along the generating lines for the cylinder, so the geometry of bending is parametrized by a single tangent vector to the torus, up to a sign. (Note that the bending at one point determines the bending elsewhere along the torus: the principal directions are parallel in the intrinsic geometry of the torus, and the principal curvature is constant.)

If we go inward toward the singular core in the family of parallel tori on the inside of our given one, the metric on these tori changes by shrinking lengths in the curved direction at a steady rate while maintaining length in the perpendicular direction. If these lines of curvature close up, then ultimately the torus shrinks to a circle, which is a cone axis with cone angle equal to the integral of principal curvature around the closed line of curvature.

Otherwise, the torus shrinks to a point†.

Surgery coefficients. The convention is to describe surgery in terms of a choice of generators μ and λ (a *meridian* and *longitude*) for the fundamental group of our chosen torus. If we identify the torus with $\mathbb{R}^2/\mathbb{Z}^2$, we can identify μ and λ with lattice vectors in the plane. The bending is determined by a pair $\pm(m, l)$ of real numbers, determined up to a sign, such that the vector $m\mu + l\lambda$ points has total curvature 2π, i.e. if you go along that path you will have bent around full circle. The parameters (m, l) are the *continuous surgery coefficients* or *real surgery coefficients* for the singular geometry.

In the important special case that (m, l) is a pair of relatively prime integers, the tori shrink to a circle with cone angle 2π, so the geometry is in fact non-singular, and we have a manifold. This case is called (m, l) *Dehn surgery*, or sometimes m/l *rational surgery*. If (m, l) is a pair of integers with greatest common divisor n, the tori shrink to a cone axis with cone angle n, so we have an orbifold. In either of these two cases, what one actually sees is a regular array of images of any object, wrapped around in a cylindrical array: when the continuous surgery coefficients are integral, the paths of geodesics match up continuously on the two sides of the singular axis, so images align perfectly. In any other case, what one would actually see would be a linear dislocation or discontinuity along the axis—this is the same phenomenon as a branch cut in an analytic function. The connection is not superficial: continuous surgery is closely related to families of analytic functions of one complex variable that have singularities like z^a, where z is a complex variable and a is a complex power.

4. Hyperbolic space

Our vision has made quick progress in learning to see repetition and symmetry as topology. With only a little more training to recognize some common hyperbolic objects, and a little adaptation to seeing scaling as distance, our eyes will be ready to lead us into hyperbolic manifolds.

Hyperbolic perspective. Hyperbolic space has a succinct description as the projective geometry of the interior of the unit ball in $\mathbb{E}^3$. Hyperbolic lines map to Euclidean straight lines in the ball, and hyperbolic planes map to Euclidean planes; it is just that hyperbolic distances are longer than Euclidean distances, very much larger near the edge of the ball, which is actually at an infinite distance. The unit ball makes a good, compact 'cheat sheet' that is easy to carry in our heads for quick reference when we need it.

Isometries of the hyperbolic geometry are the same as projective automorphisms of the ball, i.e. homeomorphisms of the ball that take planes to planes. A hyperbolic plane is determined by its circle of intersection with the unit sphere. The angle between two planes can be seen as the angle between these two circles.

In a dimension one lower, this means that constructions using only a straight edge use only a straight edge in a plane where a single circle has been drawn translate into hyperbolic geometry†.

† It is interesting that this family of metrics constitutes a geodesic in the hyperbolic metric on the Teichmüller space for a torus. The 3-manifold metric has a cone axis if and only if the geodesic tends to a cusp in the modular orbifold $\mathbb{H}^2/PSL_2(\mathbb{Z})$.

† The tangent to the circle at a given point is an important ingredient for many hyperbolic constructions. It is a nice exercise to construct this tangent from a point on a circle using a straight edge alone. Of course if we could

In other words, the projective geometry of hyperbolic space is a fragment of the projective geometry of Euclidean space. Imagine yourself in the centre of a ball perhaps 100 m in radius. Objects have the same visual geometry whether interpreted as Euclidean or hyperbolic space: hyperbolic straight lines appear straight. Even the stereoscopic effect of parallax is correct: hyperbolic parallax makes everything look a bounded distance away to Euclidean-adjusted eyes.

A hyperbolic plane appears to you as a visually round disc where a Euclidean plane slices the ball. In hyperbolic perspective, figures on this plane are identical with the projective model of the hyperbolic plane.

With computer geometry tools such as Geometer's Sketch Pad, one can readily perform these constructions and imagine them to be hyperbolic perspective drawings. As we move control points, the projectively constructed figures move with them, and keys into our sense of perspective, so we can imagine looking at a hyperbolic plane while flying over it in hyperbolic space (figure 7).

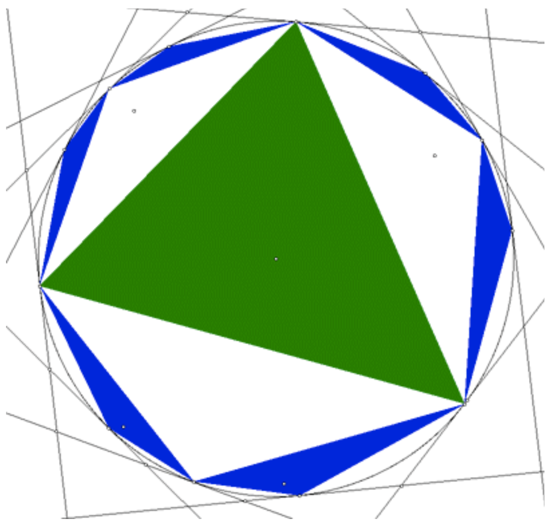


Figure 7. This is a screen shot from Geometer's Sketch Pad, showing the hyperbolic plane as seen in perspective from hyperbolic 3-space, showing the first stages of the construction of a tiling by congruent triangles. The construction starts with one circle and subsequently only requires a ruler, although it is more convenient to have automatic construction of tangents to the circle. The visual horizon of any plane in hyperbolic space is a circle like this. In hyperbolic perspective distant parts of the plane near the horizon circle are greatly foreshortened, since you have only a glancing view. In the program, moving corners of the centre triangle with a mouse helps trigger your sense of perspective and motion; the effect is stronger when the circle is large (or you put your eye closer to it) so that it more closely matches Euclidean perspective.

Hyperbolic space has lots of room. The big qualitative difference between $\mathbb{E}^3$ and $\mathbb{H}^3$ is that in $\mathbb{H}^3$ the area of a sphere of radius r grows incredibly faster: instead of scaling quadratically as in $\mathbb{E}^3$, the area grows exponentially fast. One consequence is that figures you see are closer than they appear, very much closer as their apparent size decreases.

not do it, we would dispense with the handicap and add it as a primitive operation.

This phenomenology is well captured in another model for $\mathbb{H}^3$, the upper half-space model, consisting of points in the half-space $z > 0$ in $\mathbb{R}^3$, where the hyperbolic arc length ds is given by the formula

$$\mathrm{d}s^2 = \big(1/z^2\big)\big(\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2\big).$$

What this says is that hyperbolic lengths scale in proportion to height above the xy-plane. Any similarlity of $\mathbb{E}^3$ that preserves the xy-plane also preserves this metric. In upper half-space, a geometric series of figures shrinking by factors of 2 toward the origin is a sequence of equally spaced hyperbolic figures.

All vertical lines are geodesics in upper half-space; the other geodesics are semicircles perpendicular to the xy-plane. The visual image of a figure in upper half-space, as seen by an observer high above the xy-plane, is very close to the perpendicular projection of the figure to the xy-plane. As an object moves downward, away from this observer, its visual image shrinks by similarities.

Thus for nearby objects in hyperbolic space, our Euclidean sense of perspective is a good guide to the geometry, while for distant objects, our sense of similarities can be mentally re-interpreted to give a good sense of the geometry. You can get a good feeling for this by using a hyperbolic-space viewer such as geomview (available free from http://www.geom.umn.edu), which enables you to put objects here and there, rotate them, translate them and fly around among them, all with the correct intrinsic hyperbolic geometry. To Euclidean-adjusted eyes, the distant scenery appears flattened. When we move toward distant objects, the appearance is very much like zooming a two-dimensional image. Large objects resemble Euclidean views that are in exaggerated perspective, taken with wide-angle lenses that are very close to their subject. However, this exaggerated perspective does not diminish when the object moves away; it simply records the exponential growth of the sphere of increasing radius, which corresponds to rapid shrinking of visual images with distance.

Complex coordinates. It is convenient to think of the xy-plane as the complex plane $\mathbb{C}^1$, because any orientation-preserving isometry of hyperbolic space extends to a meromorphic map, a fractional linear transformation $w \longmapsto (aw + b)/(cw + d)$ where $w = x + \mathrm{i}y$. Algebraically, the composition of fractional linear transformations is the same as matrix multiplication of their arrays of coefficients $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, thus giving an isomorphism with the group $PSL_2(\mathbb{C})$. Because of this, everything in three-dimensional hyperbolic geometry tends to have natural complex parameters. For example, a typical isometry of $\mathbb{H}^3$ to itself is a screw motion, translating a distance d along some axis and then rotating by an angle θ. The complex number $d + \mathrm{i}\theta$ is the *complex* translation length. Angles and length merge into a unified concept that yield holomorphic formulae for the geometry of any figure that varies freely in $\mathbb{H}^3$.

Banana cylinders. There are some specific surfaces that are important for understanding continuous surgery in the hyperbolic realm. Any line in $\mathbb{H}^3$ is the axis of a family of concentric cylinders of radius $\mathbb{R}$. In the upper half-space model, the cylinders about a vertical line appear as vertical cones. For the other (semi-circle) geodesics, these surfaces look something like bananas (figure 8), of varying thinness or fatness.

We can use the upper half-space to see how a banana-cylinder actually appears to a hyperbolic observer; simply think of the observer as stationed high above the xy-plane.

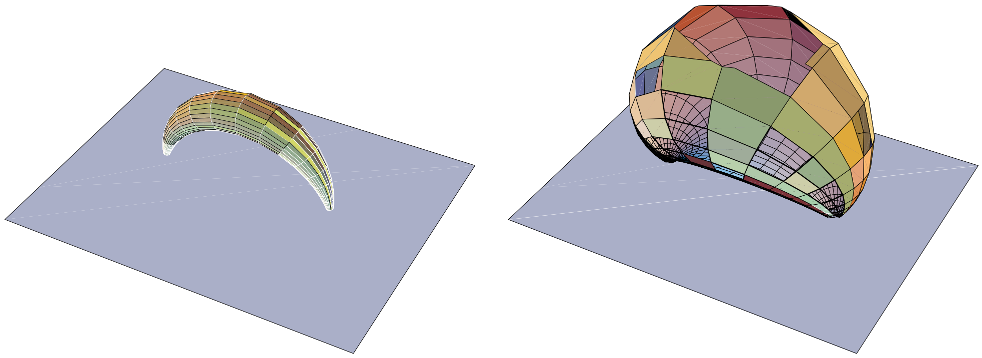


Figure 8. These bananas are actually hyperbolic cylinders concentric to a hyperbolic straight line, as seen in the upper half-space model. The banana is a union of arcs of the family of Euclidean circles that intersect the xy-plane at a fixed angle in a fixed pair of points.

A Euclidean cylinder of radius r curves in one direction only: its principal curvatures are 0 and $1/r$. A hyperbolic cylinder of radius r is shaped rather differently. In a cylindrical coordinate system (θ, t), where t measures arc length along the core axis and θ is the angle around it, the element of arc length ds on the hyperbolic cylinder satisfies

$$\mathrm{d}s^2 = (\sinh(r)\,\mathrm{d}\theta)^2 + (\cosh(r)\,\mathrm{d}t)^2.$$

As r grows large, the ratio of the scale factors $\sinh(r)$ and $\cosh(r)$ quickly converges to 1. The principal curvatures are the logarithmic derivatives of these scale factors: $\cosh(r)/\sinh(r)$ in the θ direction, and $\sinh(r)/\cosh(r)$ in the t direction. Before r is very large, both principal curvatures nearly equal 1, and the surface looks quite round even though it has the local intrinsic geometry of $\mathbb{E}^2$.

The limiting shape for a hyperbolic cylinder of radius $r \to \infty$, is a *horosphere*, with all principal curvatures equal to 1. The horosphere is also the limiting shape for spheres of radius tending to ∞, as well as the limiting shape for a surface at constant distance r from a plane. The reason these limiting shapes are all the same is that in $\mathbb{H}^3$, distant lines and distant planes all look like tiny dots, so the surfaces of constant radius about them are almost identical.

The intrinsic geometry of a horosphere is Euclidean, just like the intrinsic local geometry of a cylinder. In the upper half-space rendition of $\mathbb{H}^3$, most horospheres look like spheres tangent to $\mathbb{C}$, but there is a special case of horospheres tangent to ∞ that look like horizontal planes. In terms of hyperbolic geometry, these horizontal Euclidean planes are bent upward (that is, geodesics tangent to one of these horospheres tend downward from it).

No matter how far you are away from a horosphere in $\mathbb{H}^3$, most of it is hidden: the portion that is visible is a disc on the horosphere whose limiting radius as you go far away is 1, so the amount of area you can see from the 'outside' of a horosphere is always less than π. A similar effect is seen in the hyperbolic views of cylinders, as evident from the cut-away views in figure 9.

In $\mathbb{E}^3$, just to create one singular cone axis while maintaining locally Euclidean geometry requires a globally distributed distortion of space. In $\mathbb{H}^3$, cone axes can be arranged so that the distortion they force damps out exponentially with increasing distance from the axis. To construct a cone angle α along an axis, start with the map that adjusts angles and nothing

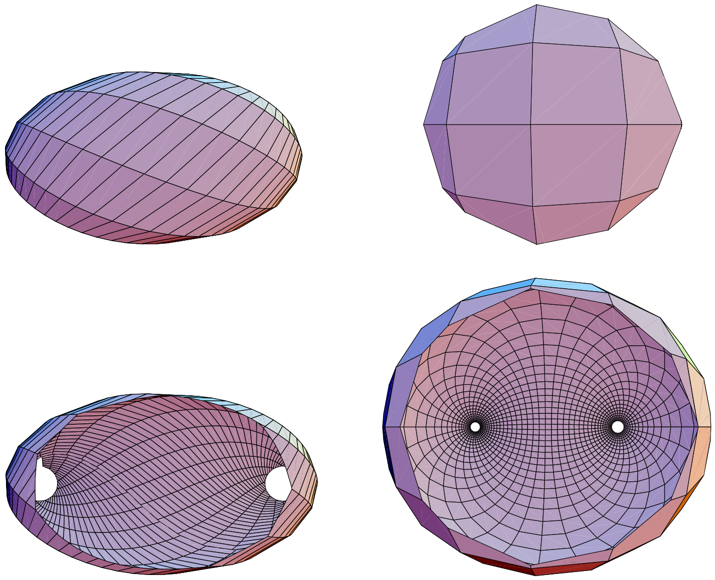


Figure 9. Two hyperbolic cylinders seen from within hyperbolic space. Each cylinder has been sliced into two pieces arranged one above the other so you can see inside. Left: a cylinder of radius 0.65, tiled with a skew grid of congruent parallelograms. Visualize the hyperbolic line that runs through the central axis of this banana. This line, like all hyperbolic geodesics, recedes from you quite quickly at its two ends. The turning away is forced by the explosive growth of the area of spheres of increasing radii about you. Staring at these pictures can help you sense the huge amount of 'stuff' at increasing distances in hyperbolic space. Right: a cylinder of radius 1.44, tiled with rectangles. In contrast to the analogous situation for Euclidean space, wider hyperbolic cylinders do not have the same visual shape as smaller cylinders, but appear rounder; this effect is exponentially rapid as a function of radius. When you are reasonably far away (as in these pictures) the diameter of a hyperbolic cylinder can be estimated as the log of ratio of the visual scaling factors from the front to the back, measured at the centre of the figure. Compare with figure 6, an analogous figure in $\mathbb{E}^3$.

else, given in cylindrical coordinates by the formula

$$(\theta, t, r) \longmapsto ((2\pi/\alpha)\theta, t, r).$$

Now adjust the value of r to get a new map that preserves the area element $\cosh(r)\sinh(r)\,\mathrm{d}\theta\,\mathrm{d}t$ on cylinders. As r increases, this converges to a motion of each cylinder by a constant distance. Since the ratio of scales in the two directions is very nearly 1, this is very nearly an isometry. Transporting the new metric by the map creates a cone axis by a distortion that decays exponentially away from the axis.

5. Embarking from T^3

The 3-torus T^3 has the simplest intrinsic description of any 3-manifold. It can be constructed by identifying each face of a cube directly to the opposite face. When you are in this geometric manifold, what you see is an array of multiple images, repeating in a regimented pattern in every direction.

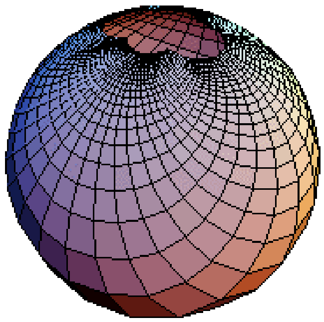


Figure 10. This is a horosphere in hyperbolic space, shown upside down (that is, in the lower half-space model). The tiles are congruent squares in the intrinsic geometry of the horosphere. Horospheres have the limiting shape of a sphere of large radius, and also the limiting shape of a cylinder of large radius.

This appearance is very special, and is not indicative of the qualitative appearance of a typical 3-manifold. Another way to understand the regimented quality of the 3-torus is to look at what happens when you move along a straight line within the 3-torus. It is unlikely that your journey is periodic, returning to exactly how it started, but your journey is almost periodic. With reasonably accurate initial data, you can predict where you will be for a long time to come. You can think of this like a ripple spreading in a pond. The set of possible places you could be after going a distance r has the local geometry of a sphere of radius r in $\mathbb{E}^3$, although in the torus wraps around and through itself. The sphere of radius r has area $4\pi r^2$: its area grows only quadratically as a function of r. Doubling r only quadruples the number of trajectories that separate within that time by at least some given small amount ϵ.

We will now watch as the metric of the torus deforms. Imagine yourself inside a cube (perhaps 20 m on a side) whose faces are identified to form a torus. A vertical line bisects the face of the cube in front of you; in the torus, it closes up to form a closed geodesic loop A, which also appears on the face of the cube behind you. A horizontal line bisects the faces to your left and right; in the torus, they are identified to make a closed geodesic circle. The floor and ceiling of the cube are bisected by a line running left to right, which identify to a closed geodesic loop C.

Now the metric of space starts to distort near loops A, B and C creating singularities along them with cone angles slowly decreasing from 2π. Geodesics in the new metric appear bent in the original Euclidean metric, bending toward the axes when they come near so that they cross just behind. What you see is a discontinuity of visual images along axis, with a double image for a narrow band (as in figures 1 and 3).

When we watched this happen earlier, we were in manifolds where the cone axes had single images. That is no longer the case: in the 3-torus, multiple images of axes A, B and C appeared, like three families of parallel lines in $\mathbb{E}^3$, translated in a pattern to interlace rather than intersect. As the metric deforms, the optical effect of image doubling occurs near *every* image of the axis. The effects are cumulative. We can understand the situation by thinking about how ripples (wavefronts) spread in the new metric, as a function of time t. Whenever the spreading ripple hits a singular axis, a thin wedge the space behind the axis is traversed doubly, because geodesics can arrive there after going to either side of the axis. This effect is compounded recursively. Every small area of the spreading ripple encounters singular axes and has portions that are doubled, at a small but regular rate. The compounding of growth of the area of the spreading ripple means that in the new metric, the area grows as an exponential function of time. In other words, the geodesic flow for the new metric now has positive topological entropy.

Another way to understand what is happening is to think about what happens if you trace a geodesic in the new metric but visualize it in the original flat metric that looks $\mathbb{E}^3$. Every time the trajectory comes near an image of one of the singular axes, it is deflects a little, in a way that a small change in the distance from an axis changes the angle, which causes a large change in position after some time. This means that successive deflections are poorly correlated with each other. After a large number of close encounters with axes, the trajectory is completely unpredictable: on a large scale in the image $\mathbb{E}^3$ image, the trajectory looks like a random walk, or a trajectory of Brownian motion.

So far this discussion of exponential growth and unpredictability applies equally well to many different geometries for the torus with singular cone axes as described. However, the geometry of our 3-torus happens to be following a very special path that is far superior to any other. As we look off in the distance, images of round objects appear round, except where they are sliced by images of the singular axes. This feature of our geometry is very special and rare; a generic, variably curved geometry would show local astigmatic effects, oriented this way and that, and growing more pronounced the further we look. The shapes of distant objects would be unrecognizable.

Our torus accomplishes this by using hyperbolic metrics, of constant negative curvature. When the cone angles at the three axes decrease, they are 'trying' to create surrounding clouds of negative curvature in their transverse directions. This negative curvature has been homogenized and spread evenly around everywhere! Metrics of this quality can be constructed from polyhedra in hyperbolic space that are combinatorially described as cubes with six faces each divided into half, making 12 faces in all. Each of these seemingly rectangular faces is actually a pentagon, when you take into account the extra vertex in one of its sides. The pattern of subdivision is the same as for a regular dodecahedron with regular pentagonal faces. It is not hard to construct a hyperbolic dodecahedron where the six edges that identify to the three loops A, B and C take a value $0 < \alpha/2 < \pi$, and all other angles are $\pi/2$. The shapes of these hyperbolic polyhedra are determined by their angles; when the faces are glued together, it gives a hyperbolic cone-metric for the torus. When a is very small, the polyhedron is very small, and its shape is nearly cubical. When rescaled to constant diameter (longest edge length 20 m), these tori converge to our cubical torus. Figure 11 show some examples of dodecahedra of this form.

The cone angles along axes A, B and C keep decreasing, until they eventually attain the value π. At this moment, the metric can be constructed from a regular, right-angled hyperbolic dodecahedron with all angles equal to $\pi/2$. Suddenly, all the double images align with each other so that every image is perfect. What we see is a pattern repeating in hyperbolic space. Our right-angled dodecahedron is a fundamental domain, or basic tile; we can fill up $\mathbb{H}^3$ with non-overlapping dodecahedra, matching up on their faces. We see one image of ourself in each of these dodecahedral cells. The sizes of the images decreases exponentially fast with distance, since there are exponentially many at a given distance. Since the long-term behaviour of geodesics is random, distant images face every which way, no matter in what direction we look.

The cone angles start to decrease again. As they decrease, the three axes A, B and C also decrease in length. When the cone angles reach $\pi/3$, all images are in alignment; then they go out again. They come back in alignment for an infinite sequence of values π/n, the

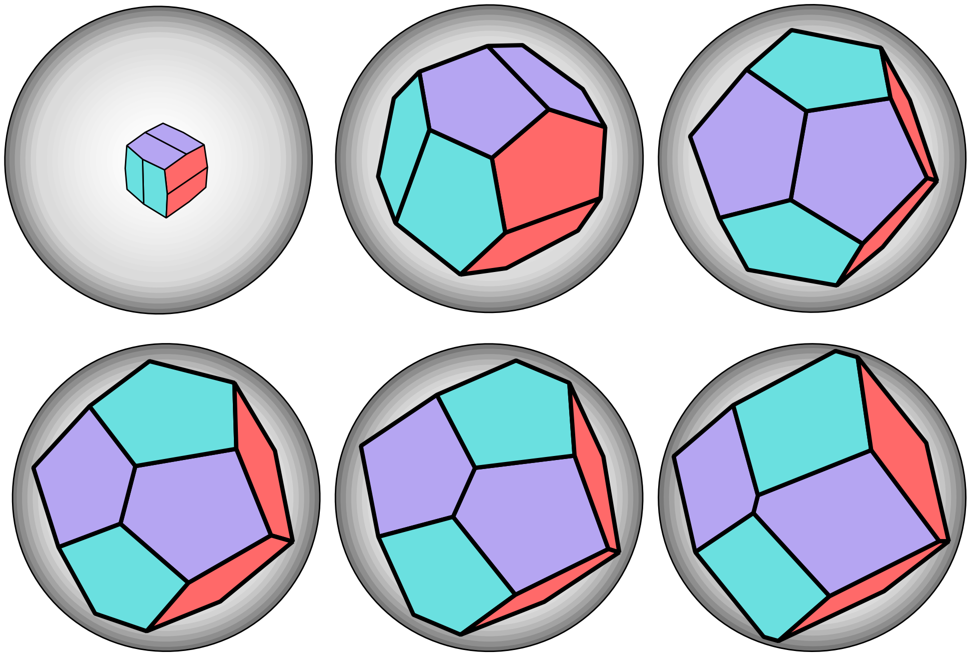


Figure 11. Six hyperbolic dodecahedra with varying angles, seen in the projective model of $\mathbb{H}^3$. Each dodecahedron has a set of six edges which decrease, in successive frames, through values $\pi - \epsilon$, $2\pi/3$, $\pi/4$, $\pi/5$, $\pi/6$ and $\pi/12$, while the other 24 edges remain at $\pi/2$. When opposite pairs of like-coloured faces are glued, one obtains T^3 with evolving cone angles at three edges. Notice that for the last five dodecahedra an observer in $\mathbb{H}^3$ would never be able to see more than three faces at a time without going inside the dodecahedron. A hyperbolic perspective drawing facing the centre of the ball is the same as the Euclidean perspective drawing from the same point, but for the correct Euclidean interpretation of the drawing you need to stand close to it, while for the hyperbolic interpretation you stand far away. The first figure is just a sketch, while the others are traced from output of SnapPea's Dirichlet domain module.

axes growing shorter and shorter every time, and receding off into the distance from where we are standing. There is a limit, when the cone angles go to 0. The limiting manifold is homeomorphic to $T^3 \setminus A \cup B \cup C$. The limiting metric is *complete*, meaning that the holes where A, B and C have vanished are infinitely deep, and it is not possible to reach the edge of the universe in a finite distance. These holes are surrounded by a family of concentric tori, bent into the shape of horospheres. The intrinsic metric of each torus is flat. If you proceed inward down the hole, the concentric tori all have the same shape, but they shrink exponentially fast. It follows that the total volume for this metric is finite. If the metric is scaled so that its curvature is -1 (that is, the hyperbolic radian is used as a unit of distance), the metric is uniquely determined by the topology; in particular, its volume is a topological invariant. Our current manifold has volume $7.327\,724\,753\ldots$.

The program SnapPea, when given the topological description of a manifold such as this, can almost instantaneously compute its complete hyperbolic metric (which is unique), computes numerical invariants such as its volume and displays various pictures. The most convenient way to describe the topology is usually by specifying a Dehn surgery to produce it from a link in S^3. In this case, $T^3 \setminus A \cup B \cup C$ is homeomorphic with $S^3 \setminus B$, where B

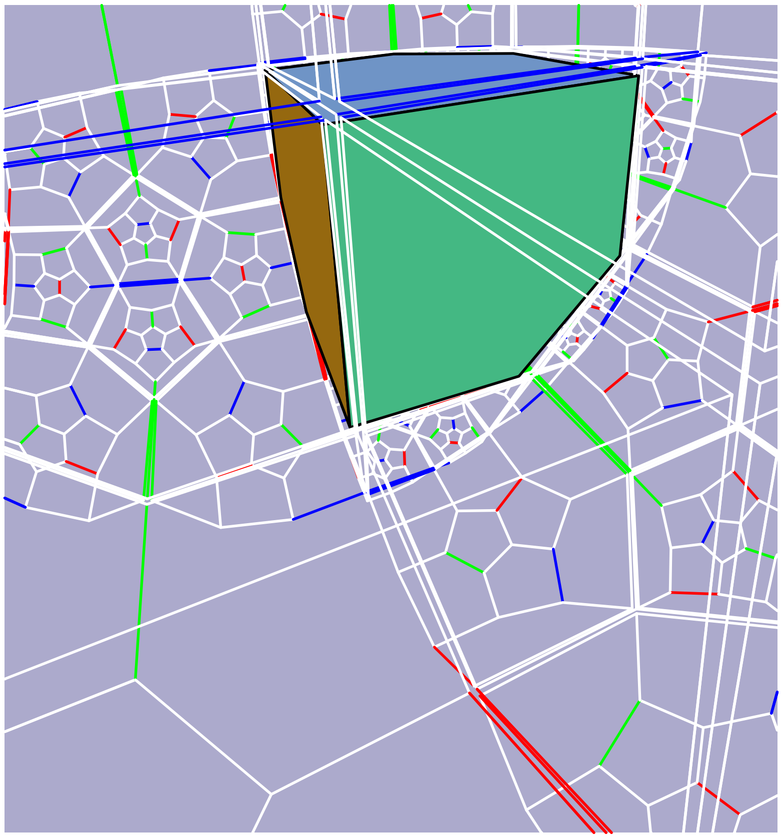


Figure 12. Hyperbolic space tiled by right-angled dodecahedra. This is a true perspective drawing inside T^3 when it has a hyperbolic metric with three singular axes of cone angle π. Compare with figure 11, frame 3. The fundamental domain (the dodecahedron) has many images, which appear in distorted perspective to Euclidean-adjusted eyes. The prominent circle is the horizon circle for one of the planar sheets of faces. Can you pick out any 'bananas', or rough hyperbolic cylinders, made from strings of dodecahedra? This figure is a postscript snapshot from geomview, a program available for common unix platforms from http://www.geom.umn.edu. Click on the Not Knot Flythrough module (written by Charlie Gunn) and you will be conducted on a flying tour through this scene (minus the solid dodecahedron). From the File menu open data/geom/dodec; with patience you can translate and rotate dodec in one of the bays. Fancier versions of scenes like this occur in the video NotKnot.

is a famous link of three components called the Borromean rings. These three interlocked circles have the property that when any one is removed, the other two are unlinked. T^3 is obtained from S^3 by doing $(0, 1)$ surgery on the Borromean rings.

Perhaps the most informative picture in SnapPea is the cusp view. A *cusp* is the jargon for a deep 'hole' in a hyperbolic manifold that is enclosed by a family of concentric horospherical-shaped tori. SnapPea has sliders to choose one horospherical torus concentric to each cusp, identifying which torus by the volume of the portion of the cusp it encloses. For each cusp, a display can be produced of the view from deep within that cusp of the images of all other horospherical tori, except for the one you are inside (otherwise it would block the view of everything). The cusp views can have quite distinctive qualitative features. By studying and comparing them, you can 'triangulate', and see how the geometry of the entire manifold fits together.

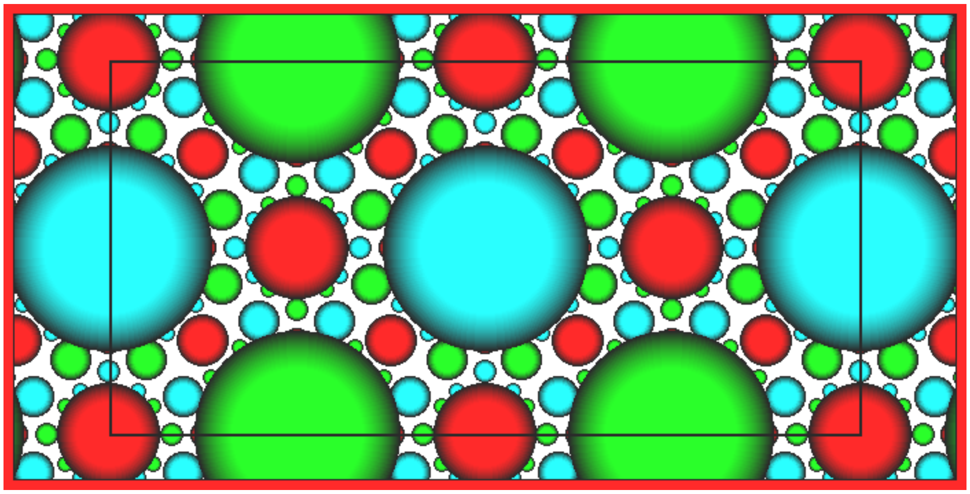


Figure 13. Axes A, B and C in T^3 have gone off to infinity as their cone angles went to 0. This is the view from deep inside cusp A, looking out, showing multiple images of three horospherical tori that enclose axes A, B and C. We are inside the region enclosed by the A torus, whose fundamental domain is indicated by the rectangle. Luckily, it is made of one-way glass so we see out to the rest of space. This view is similar to the panorama wrapping around the vertical pillar A in the front of the cube from which the original Euclidean geometry of T^3 was formed. Large images of horizontal cylinders B on the left and right of the room appear at mid-height. Large images of cylinders C on the floor and ceiling appear at the top and bottom of the rectangle. In between these large images we can make out images of our own cusp A, at the front and back of the room. The view from each cusp is the same. Imagine how it looks to an observer stationed in another cusp, and match the visual geometry as she would see it to the geometry you see. The sky above each horosphere has a closest layer of horospheres, arranged in a diamond pattern just like what you see.

Continuous hyperbolic surgery works in quite a nice way. Hyperbolic manifolds with singular axes are controlled by holomorphic parameters, one complex parameter for each cusp that enables its singular behaviour to change in two real dimensions, in the cone angle direction and in the shear direction. SnapPea implements continuous surgery in a way that enables one to navigate quite readily among 3-manifolds, and to identify where you are. For a hyperbolic manifold, a view such as in figure 13 is determined by the geometry, plus the user-controlled parameters of how big the cusp neighbourhoods are. The topology determines the geometry, by the rigidity theorem of Mostow and Prasad. Therefore, it is mathematically a straightforward computation to decide whether or not two hyperbolic manifolds are homeomorphic; SnapPea's Isometry module answers a particular instance of this question in short order. A human looking at the cusp-view pictures can also usually tell at a glance whether or not two 3-manifolds are the same or different, although there are some tricky special cases when manifolds are hard to compare accurately by eye, involving 3-manifolds are built out of identical polyhedra that are glued together in different ways.

To make sense of the view inside a hyperbolic 3-manifold, it helps to keep the imagery of hyperbolic cylinders (bananas) in the front of the mind.

Starting with a complete hyperbolic manifold of finite volume, all but a bounded (compact) set of surgery coefficients yield a singular hyperbolic manifold, and in particular, all but a finite number of Dehn fillings for each cusp give hyperbolic manifolds. Often there are no exceptions; at most there are a handful. The volume of a manifold obtained by Dehn filling is always less than the volume of the original. The set of all hyperbolic 3-manifolds with volume less than a constant can all be obtained by Dehn filling from a finite set of manifolds with cusps.

The formula for the volume of a hyperbolic polyhedron involves the dilogarithm function, which is a non-elementary but readily computed function. Much simpler than the volume is the formula for the derivative of volume, as the polyhedron varies while maintaining a combinatorial type. The formula, discovered by Schlafly, says that the derivative of volume of a polyhedron is the sum over all its edges of the edge length times the derivative of its exterior angle, i.e. as the edges get sharper, the volume gets bigger. The same formula applies when polyhedra are glued together to form a hyperbolic cone-manifold: the derivative of its hyperbolic volume is the sum, over all cone axes, of the length of the derivative of the curvature concentrated at the axis (curvature is 2π minus its cone angle). It is as if there is a secret tunnel to another world along each cone axis; volume flows in or out of a segment of a cone axis in direct proportion to the change in its cone angle.

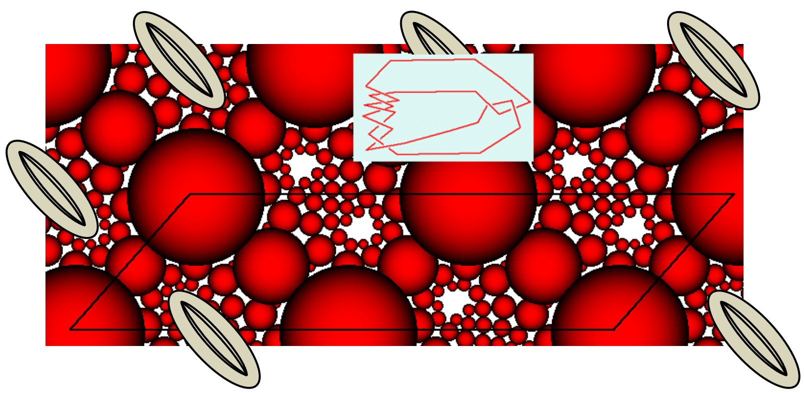


Figure 14. The cusp view of a knot (inset) from SnapPea, with bananas. The cut-away form of hyperbolic cylinders is plainly visible in the hollows where bananas were omitted: cf figure 9. Each banana has many images of horospheres spiralling around it. All horospheres wrap up to a single torus in the manifold, so the picture strongly suggests that the axis of the bananas wrap into a fairly short geodesic. The knot diagram suggests a short geodesic looping around the twisty area. spiralling fairly tightly (see figure 15).

The knot in the inset of figure 14 has two strands that twist three times around each other. The twisting is reflected by the geometry of the complete hyperbolic structure for its complement. If you have memorized the appearance of a cylinder in $\mathbb{H}^3$, you can plainly make out the shape of a fairly fat cylinder in the cusp view of the knot. This strongly suggests that there is a short geodesic. In fact, a loop encircling the two twisted strands is homotopic to a closed geodesic. This can be calculated in SnapPea by drawing the

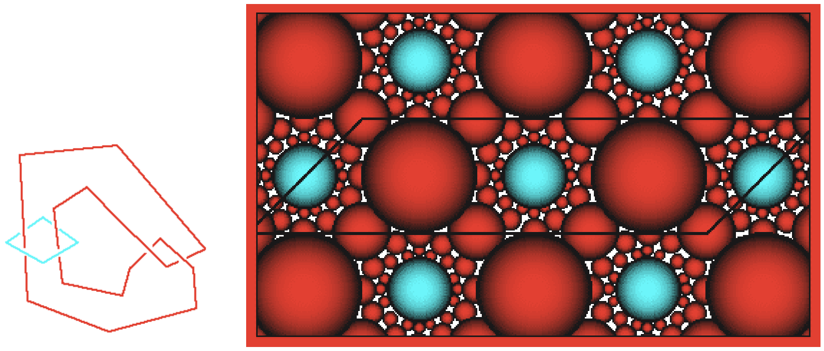


Figure 15. Left: the Whitehead link consists of two circles, which separately are unknotted. The link in figure 14 is obtained by $(1, 3)$ Dehn filling of the shorter loop. This is the same as slicing along a horizontal disc spanning this loop, twisting three times and gluing back. Right: the cusp view of the Whitehead link, as seen from the cusp for the 'long' loop. The bananas have turned into horospheres. SnapPea's sliders for cusp volumes have been adjusted to maximize the 'big' cusp and keep the other one small, for better comparison with figure 14. (It is not obvious from these pictures, but the two strands are actually symmetric with each other.)

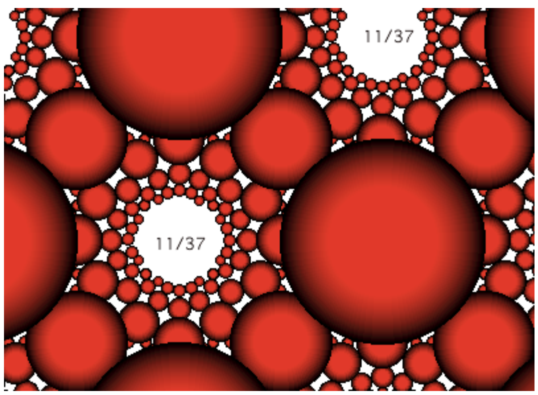


Figure 16. This is the cusp view of $(11, 37)$ surgery on the small component of the Whitehead link. There is a closed geodesic that is quite short (length is $0.000\,989\,647\ldots$) and far away, so that the cylindrical neighbourhood is not easy to distinguish from a horosphere. This view bears a striking resemblance to comparable cusp views for the complement of the Borromean rings. A picture in figure 13 would be obtained by adjusting the objects that are drawn: shrinking the single horosphere in the present picture, and drawing a hyperbolic cylinder of appropriate diameter (a very bloated banana) around the short geodesic that is the core of the $(11, 37)$ filling. The nearly congruent geometry reflects the fact that the complement of the Borromean rings is a twofold cover of the complement of the Whitehead link.

loop in question, forming the complement of the resulting link, filling the new component using $(1, 0)$ surgery (which means filling that component back the way it was). There is a core geodesic view in SnapPea which now will tell you the length of the filled-in core: it indeed reports a short hyperbolic length $0.1535\ldots$ with a twist of $2.2071\ldots$ rad. In comparison, SnapPea's length spectrum module reports a second shortest length of

1.138 084 551 958 0791 for a closed geodesic—still a reasonably short hyperbolic length.

A famous theorem of Gordon and Luecke says that knots are equivalent if and only if there is a homomorphism between their complements. However, the situation is quite different for links with more than one component. Any time there is an unknotted component to a link, you can find a disc that has that unknot as its boundary, intersecting other strands in some collection of points. If you slice along the disc, twist one side of the cut n full times, then glue the two cuts back together, you obtain a new link, usually different from the old one since you have added some extra twists. But the complement of the new link is homeomorphic to the old one, since apart from discontinuous behaviour at the boundary, the new gluing map is the same as the old. Thus, when a new unknotted component is added to our link, we can untwist it to make a simpler picture, known as the Whitehead link (figure 15). The $(11, 37)$ Dehn surgery on the second component of the Whitehead link is shown in figure 16. The length of the core geodesic in that case is quite short, only 0.000 989 647, and the geometry is almost indistinguishable from that of the Whitehead link except near the central part of the core.

6. The ocean of hyperbolic manifolds

The complement of the Borromean rings is a twofold cover of the complement of the Whitehead link. This kind of relationship is not an isolated phenomenon. Further exploration of 3-manifolds via continuous surgery reveals quite a lot of convergence, if one aims toward bright beacons. it is as if there are just a few major shipping channels marked by powerful beacons, but with large numbers of interesting destinations off to the sides.

In SnapPea there is a module for *drilling* out geodesics loops that are identified geometrically. Given a hyperbolic manifold, the program will present a list of all loops up to a specified combinatorial complexity (as determined by an internal representation of the manifold), sorted in order of hyperbolic length. SnapPea can remove any of these geodesics from the manifold, in essence deforming its cone angle to 0 so as to obtain a complete hyperbolic structure for its complement.

Hyperbolic volume increases, by an amount that is typically in rough proportion to the length of the geodesic you drill.

Here is one phenomenon that seems to always occur. This is a striking effect that others (among them Jeff Weeks) have also noticed, but I do not know if it has found its way into print:

Conjecture 6.1. When the shortest simple closed geodesics are repeatedly removed from any complete hyperbolic 3-manifold of finite volume, eventually one obtains a manifold whose shortest closed geodesic has length $2.122\,55\ldots$ with angle of twist $\pm 1.809\,11\ldots$, and has self-replicating behaviour when removed.

This length is one that occurs in the complement of the Borromean rings; it is the length of a loop that links around two of the rings where they cross, and comes from an element of $SL_2(\mathbb{C})$ whose trace is $2 + 2i$. This length also occurs in the complement of the Whitehead link. If you cut the Whitehead link open along a twice-punctured horizontal disc with boundary the ‘small’ loop in figure 15, you obtain a fragment of a hyperbolic manifold bounded by two twice-punctured discs. This particular fragment has identical geometry in any manifold where it occurs. This module is a *clasp*, and is fairly common in 3-manifold topology. If a short curve encircling the two arcs of the clasp is removed, the result has the

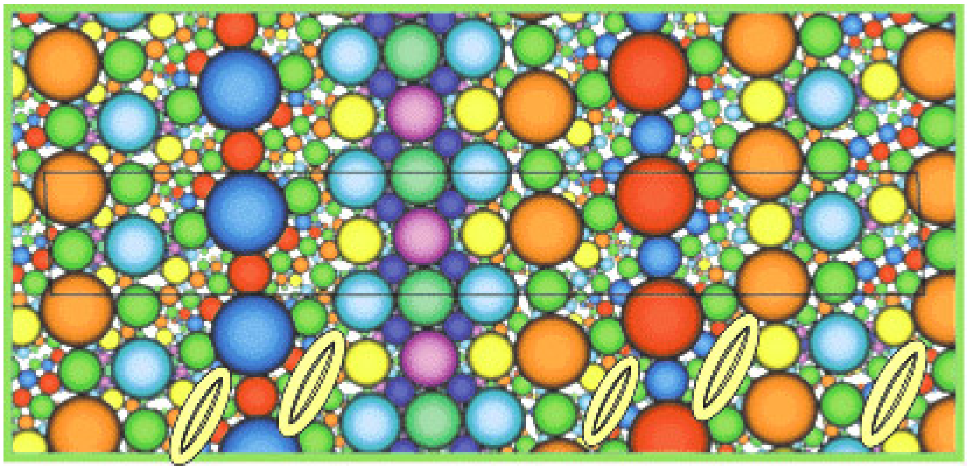
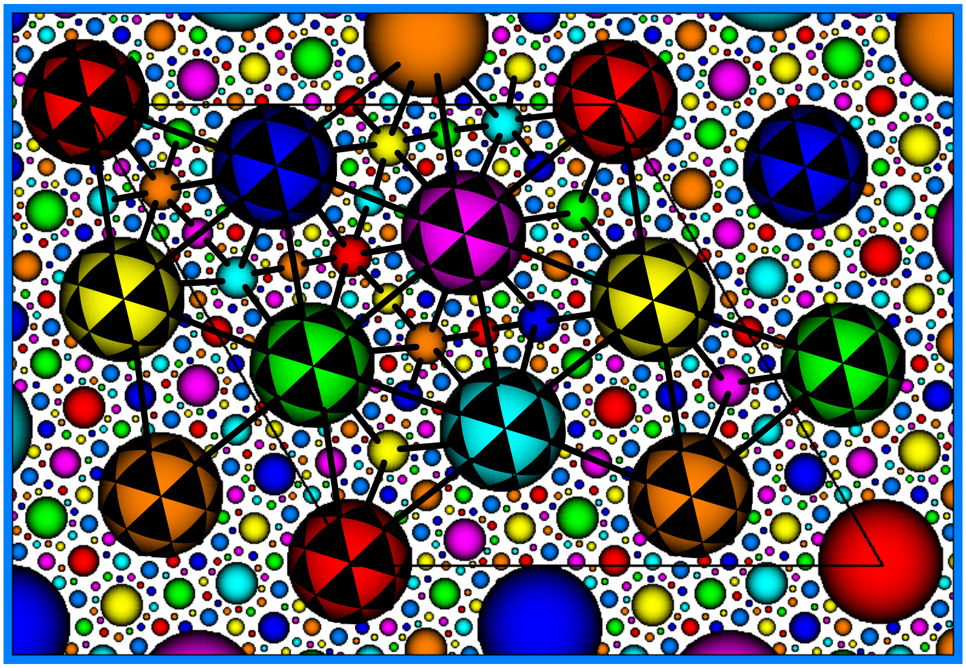


Figure 17. One of the cusp views of a nine-cusped hyperbolic manifold of volume 44.0227. The horosphere that walls off this cusp encloses a volume of 7.5; the next largest cusp volume is 4.2. Features that look like hyperbolic cylinders are scattered around the image, as indicated by the bananas sketched in at the bottom. Notice how the bananas together with the horoballs seem to fit in a semi-regular pattern, as for the cusp view for the Borromean rings but with dislocations where the grid turns 45°. Patterns like this are common.

same topology as two clasps strung together, joined on twice-punctured discs. Like Hydra, it only multiplies when it is zapped.

There are other patterns that occur under drilling. The Borromean rings is an instance of a phenomenon called a *universal link*: every possible 3-manifold can be obtained by Dehn filling of some covering space of finite degree of the Borromean rings. To a surprising degree, one can learn about covering spaces and certain relatives of covering spaces by drilling geodesics based only on the patterns of their lengths. For the simplest manifolds repeated drilling of the shortest geodesic soon arrives at the complement of an n-link chain, in a traditional circular chain arrangement. The topology of such a chain depends on the number of links and the amount of twist. The least-twisted chain of five links is particularly important for the simpler hyperbolic manifolds. Its complement has volume 10.149 . . . . Its cusp view looks like figure 18, except that its volume is $\frac{10}{28}$ the size and the colouring repeats on a shorter scale. This manifold has an amazing symmetry group of order 120 that performs all possible permutations of the cusps. Very many of the simpler hyperbolic manifolds arrive at this topology/geometry after recursively drilling out shortest geodesics until there are five cusps (see the discussion of the census of the simplest 3-manifolds whose data are available within SnapPea [HW89]). When one more shortest geodesic is drilled out, the manifold recrystallizes into a 6-cups manifold with the same cusp pattern as the Borromean rings.

More regularities occur for larger examples. Repeatedly drilling out of relatively short geodesics (but *not* always the very shortest, since the shortest geodesics eventually start to be confined to certain regions) always seems to start ‘crystallizing’ a 3-manifold, so that its cusp neighbourhoods can be adjusted in size to pack together in regular crystalline patterns, most commonly in a square-looking pattern as for the Borromean rings. There are several other patterns that occur frequently, but less frequently, notably the equilateral pattern of



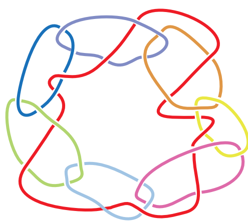


Figure 18. Eight linked full tori (right) when inflated fill S^3 so that each meets each other along a single (curved) hexagon. The complement of the link is a hyperbolic manifold (above) of volume 28.418 . . . , having extraordinary 336-fold symmetry. The seven nearest cusps describe a seven-colour map of the torus surrounding our own cusp. As seen from a neighbouring cusp, we form part of a triangular grid along with a hexagon of common neighbours, as indicated by the painted triangles. Beams join pairs of nearest neighbours. The closest images of ourselves are in the fourth tier, quite distant for a manifold of this volume.

figure 18. There are often dislocations, where two different patterns meet or where the crystalline grids are at different angles.

The patterns, constructed and viewed with SnapPea, are striking. The phenomena are similar even for large manifolds of volume 100 or 200. It is often hard to formally capture and describe patterns that we see plainly with our eyes, but these observations suggest that there might be a comprehensible systematic organization for hyperbolic crystallography. Possibly every compact hyperbolic 3-manifold can be thought of as a 'crystallized' arrangement of solid tori surrounding relatively short geodesics, fitting with each other in a small repertoire of local arrangements similar to figures in this paper. Any universal pattern of this nature would need to match certain manifolds in more than one way, but with luck the ambiguity would be manageable.

Acknowledgment

Partially supported by NSF grant DMS-9704135.

References

[Geo81] The Geometry Center (University of Minnesota) 1981 *Not Knot* (Boston: Jones and Bartlett) video tape (16 min)

[HW89] Hildebrand M and Weeks J 1989 A computer generated census of cusped hyperbolic 3-manifolds *Computers and Mathematics (Cambridge, MA)* (New York: Springer) pp 53–9

[Thu79] Thurston W P 1979 *Geometry and Topology of Three-manifolds (Princeton Lecture Notes)* http://www.msri.org/publications/books/gt3m

[Thu82] Thurston W P 1982 Three-dimensional manifolds, Kleinian groups and hyperbolic geometry *Bull. Am. Math. Soc.* 6

[Thu86] Thurston W P 1986 Hyperbolic structures on 3-manifolds, I: Deformation of acylindrical manifolds *Ann. Math.* **124** 203–46

[Thu97] Thurston W P 1997 *Three-dimensional Geometry and Topology* vol 1 *(Princeton Mathematical Series 35)* ed S Levy (Princeton, NJ: Princeton University Press)

[Thu98a] Thurston W P 1998 Hyperbolic structures on 3-manifolds, III: deformations of 3-manifolds with incompressible boundary *Preprint*

[Thu98b] Thurston W P 1998 Hyperbolic structures on 3-manifolds, II: surface groups and 3-manifolds which fiber over the circle *Preprint* (orig. *Preprint* 1981, revised 1986)