SUPPLEMENT TO

NOT KNOT

David Epstein
Charlie Gunn

GEOMETRY CENTER

University of Minnesota

SUPPLEMENT TO

David Epstein

University of Warwick, England
and
Geometry Center
University of Minnesota, Minneapolis

Charlie Gunn

Geometry Center
University of Minnesota, Minneapolis

Jones and Bartlett Publishers
Boston

Editorial, Sales, and Customer Service Offices

Jones and Bartlett Publishers
20 Park Plaza
Boston, MA 02116

ISBN 0-86720-297-1 (pbk)

ISBN 0-86720-240-8 (video/pbk)

Credits for *Not Knot* (video)

Written by (in alphabetical order)
David Epstein
Charlie Gunn
Scott Kim
Silvio Levy
Stuart Levy
Delle Maxwell
Tobias Orloff
John Sullivan
William Thurston

Technical Director
Charlie Gunn

Artistic Director
Delle Maxwell

Modeling, Animation, and Rendering
Tobias Orloff
Delle Maxwell
Stuart Levy
Charlie Gunn

Beginning Titles
Scott Kim

Title Music
Tom Lonardo

Narrator
Chery Hays

Grateful acknowledgment of software used in movie:

Softimage	Softimage, Inc. Montreal, Canada
Renderman	Pixar, Inc. Richmond, CA
Mathematica	Wolfram Research Champaign, IL
Rayshade	Craig Kolb

Video recording by NCSA

Animation Recording and Segment Editing
Robert Patterson
National Center for Supercomputer Applications
University of Illinois at Urbana-Champaign

Post Production
KTCA
St. Paul, MN

Video Editor
Jerry Lasko

Audio
Joe Demko

Thanks to
Scott Hayes
Janelle Jurek
Jeff Weeks

Special thanks to
Albert Marden
Director
Geometry Center

The Geometry Supercomputer Project was funded by the National Science Foundation and the University of Minnesota.

Figure Acknowledgments

The pictures of knot diagrams with a rope-like texture on pages 6, 12, and 44 were created with Linktool, a program written by David Broman of Rice University. Linktool runs on NeXT computers, and is based in part on programs by Toby Orloff (link editing) and Charlie Gunn (link drawing). For readers with Internet access, Linktool is available via anonymous ftp from poincare.geom.umm.edu (128.101.25.31).

The pictures of the hyperbolic disk on pages 33 and 40 were produced by Mark Phillips of the Geometry Center, University of Minnesota, using a program written with Robert Miner while they were students at the University of Maryland.

The picture of the hyperbolic disk on page 35 was created by Steve Rumsby of the University of Warwick.

The illustrations in 68A were hand-drawn by George Francis of the University of Illinois and are reproduced here with his permission.

All the other illustrations, excluding movie frames, were created by David Epstein.

Printed in the United States of America

95 94 93 92 91　　10 9 8 7 6 5 4 3 2 1

Supplement to Not Knot

This supplement is designed to help you understand more completely the mathematical animation "Not Knot." It includes several different kinds of material.

- We have included a copy of the complete script. In the margin, next to the script, are key frames from the movie which will help you to match the video with the supplement.
- The script is broken into sections. After each section of script, there are questions and answers arising from that section of the script. We hope you'll find the answers to some of your questions here.
- For a broader introduction to some of the topics introduced in the movie, short expository articles on the history of knot theory, on mirrors, and on hyperbolic geometry are also included—we'll refer to these as *insets* in this supplement.
- We give some student activities, which involve working with easily available materials to build models illustrating the ideas of the movie.
- Finally, there are bibliographical references in the insets and in some of the answers which direct the interested viewer/reader to other sources of information related to the movie.

How to Use This Supplement

Our advice for using these materials:

- Watch the video once or twice. Write down any questions that occur to you. At this point it's very important to remember that all questions are valid.
- Then, watch the movie again using the supplement. That is, stop the video after each section of the script, and go over the questions and answers in the supplement. Compare them to your own. If your questions aren't answered here, you may find the answer in the insets. Or, the supplement may not have the answer to your question. You may be able to figure it out on your own by watching the video again, or you may have asked a deeper question. That's good too. We'd like to hear from you if you have a question which you don't think is adequately addressed by the supplement; if it's a good one and we use it in the next edition of the supplement, you'll get a free copy of the video. Send to:

  Not Knot
  Jones and Bartlett Publishers
  20 Park Plaza
  Boston, MA 02116-9792

- A square-root sign in the left-hand margin indicates a technical discussion directed at mathematicians and aspiring mathematicians. You may want to skip sections marked in this way.
- There are two books that we particularly wish to recommend as being suitable for those who have a general interest in the subject of the movie. These books discuss many of the topics which arise in the movie and the supplement.

  G.B. Francis, *A Topological Picturebook*, Springer 1987.
  J.R. Weeks, *The Shape of Space*, Marcel Dekker 1985.

- Keep in mind that even professional mathematicians, with expertise in this area, have to watch the video many times to grasp all the ideas. Many of the ideas in the movie are only found in college level math courses, and some would only be found in graduate courses. Watching the video is only a starting point for mastering its contents.
- If you remain puzzled, study the references and watch the video repeatedly. Talk to your friends, teachers, and colleagues, and cooperate in figuring out what is going on. You may discover new things, not known to those who made the movie. Write and tell us.

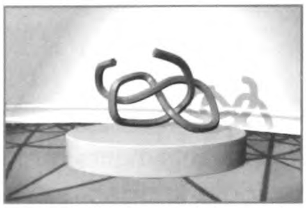
We thank Tamara Munzner and John Sullivan for extensive comments on early drafts.

David Epstein
Charlie Gunn
Geometry Center
1300 Second Street South
Minneapolis, MN 55454

Our work at the Geometry Center in Minneapolis is supported by the National Science Foundation, the Department of Energy, Minnesota Technology Inc., and the University of Minnesota. Epstein's work in the UK is supported by the University of Warwick and by the Science and Engineering Research Council, through its Mathematics Committee and its Computational Science Initiative.

1 (S) Welcome to the Exhibit of Geometry and Knots.

Knots are so much a part of our everyday world that we take them for granted. But mathematicians have found that studying even the simplest knots can lead into almost unimaginable spaces. In this movie we'll take you on a short guided tour of the world of knots, from a mathematician's perspective. Let's begin...



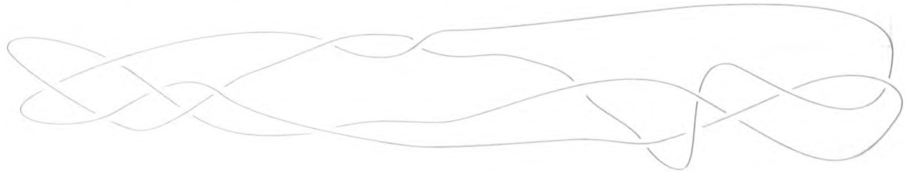
This figure eight knot can be untied, since its ends are free. From a mathematician's point of view this knot is no knot at all. But if we join the ends to form a loop, we can no longer undo the knot without cutting it. Now that's a knot.

2 Q **The film is pretty hard to understand. Can you give an overview, so that we can see what you are aiming at and where you are going?**

2 A Yes, but we will leave this to the last part of the supplement (see 74A), by which time we will have developed the ideas needed in the overview.

3 Q **What exactly is a knot?**

3 A By a "knot," we mean a length of cord with the two ends glued together. One way of constructing a knot is to take some nylon cord, tie a knot (in the everyday sense, not in the sense we are now trying to explain) and melt the two ends together. Or, tie a knot in an extension cord and then plug the two ends together. A special example of a knot is circular piece of string, lying flat on a plane. The name of this knot is "the trivial knot" or "the unknot." What do we mean by "undoing" a knot? We mean untying it, without cutting or breaking it, so that it can be laid out as a circle on a flat surface. That is to say,

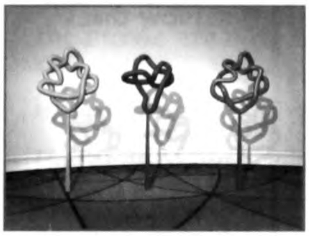


undoing a knot is the same thing as deforming it to the unknot. Given a knot, it is not always trivial to decide whether it is unknotted. What do you think about the knot shown above? (Answer in 73A.)

4 Q **How can we know that a knot is knotted, that it can't be undone?**

4 A To answer this question is one of the main objectives of the part of mathematics called knot theory. (See the inset on page 10.) There are many ways of proving that a knot is not trivial: the idea is to find some property which is unchanged by moving the knot around. Is the trefoil knot shown in 6A knot-

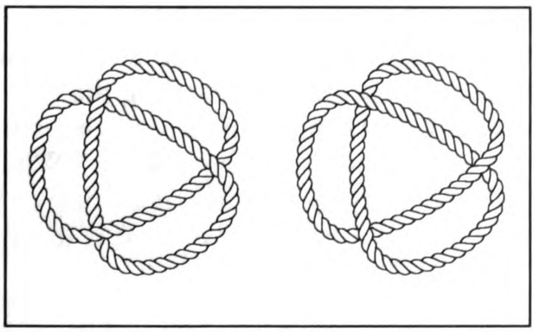
ted? One might be tempted to reply "It is obviously knotted." The same answer might be given for the knot shown in 3A. For the trefoil knot the unthinking answer is correct. It can be proved mathematically that the trefoil knot really is knotted. Why does the mathematician think it requires proof to know something that "can be seen at a glance"? Well, maybe one could stretch it out to be thousands of miles long, and then let millions of kittens play with it, tangling it up, and then invite millions of people to work together on undoing it; maybe, miraculously, it would come undone. Many of the most important phenomena in mathematics were once thought to be "clearly impossible." (One example is a space-filling curve.)



5 (S) These three knots look different. Are they? Two knots that look different may be just different arrangements of the same knot. Sometimes a knot that looks knotted really isn't.

6 Q **What does it mean for two knots to be the "same" as each other?**

6 A Most often, when mathematicians talk of two knots (or two links) being "the same," they mean that one knot can be physically moved to the other. In the



It is not possible to move the trefoil knot to its mirror image, but we still want to say they are the same knot.

movie, we talk of two knots being "the same" if one knot can be physically moved to the other knot *or to a mirror image of the other knot;* the terminology of the movie is also often used by mathematicians, but less often than the definition given in the first sentence of this paragraph. The reason for the choice made in the movie is that it fits more simply with the Gordon–Luecke Theorem stated in 14A. Mathematicians have always felt themselves free to make definitions, to make words mean what they want them to mean. Perhaps this is what Lewis Carroll was thinking of when he wrote:

> *"When I use a word," Humpty Dumpty said in a rather scornful tone, "it means just what I choose it to mean—neither more nor less....The question is—Which is to be master?—that's all."*

7 (S) Here are three intertwined loops. If you remove any one of them, the other two fall apart. Mathematicians call the union of several loops a "link." By rearranging this link we can see that the three loops are equivalent.

8 [Q] What does the word "equivalent" mean here and why is it significant?

8 [A] Take a look at the picture of the Borromean rings shown in 48S. One can rotate three-dimensional space rigidly, so that the red ring is moved to the blue, the blue to the green, and the green to the red. For those who like formulas, we take the plane of the blue ellipse to be given by the equation $x=0$, the plane of the green ellipse by $y=0$, and the plane of the red ellipse by $z=0$. One possible axis of rotation is the line $x=y=z$, which has the advantage of being symmetrically chosen with respect to the three coordinates. We rotate through one third of a revolution about this axis and this interchanges the colored ellipses.

The reason the equivalence is worth remarking on, is that it means we should expect the associated pictures (50S to the end of the movie) to be symmetrical with respect to interchanging the colors, and if we look carefully at the pictures, we see that they are.

9 (S) Here is another symmetric form of the same link. In this form it is called the Borromean rings.

10 [Q] Why are they called the Borromean rings?

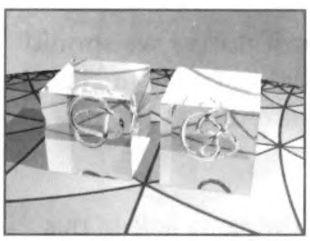
10 [A] The Borromean rings were named after the ancient Italian family of the Princes of Borromeo (Bor-oh-MAY-oh). They occur as heraldic symbols on the family coat of arms. Visitors to the Borromean castle on an island in the Lake Maggiore in Italy can see the rings cut into the stonework. There are different theories as to their significance. Some say they represent the indivisible Holy Trinity. Others say they represent the saying "United we stand, divided we fall," since if one ring is removed, the other two fall apart.

11 (S) When we look at a knot, we don't usually think about the space around it. But to a mathematician, the space around a knot is just as real as the knot itself. Sometimes it's easier to understand something by looking at what it's not.

The complement of a set is what's left when you take away the points of the set. So, what's not a knot is called its complement.

12 Q Can you explain the idea of a "complement" more precisely?

12 A Strictly speaking, the notion of a "complement" needs a universe U with respect to which complements are taken. The complement of X, a subset of U, is the set of all points that are in U but not in X. In our case, the universe is actually the three-sphere, and not three-dimensional euclidean space. The three-sphere can be obtained from euclidean space by adding a single "point at infinity." There is exactly one scene in the movie where this distinction is made visible. Can you say what it is? (The answer is given in 47A.)



13 (S) In 1988, Cameron Gordon and John Luecke proved that the complements of different knots can never be the same space. So studying the complements of knots helps us tell whether knots are the same or different.

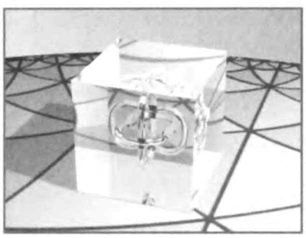
14 Q What does it mean for the complement of different knots or links to be the "same space"?

14 A The complement of a knot or link is a particular example of a space. In the movie a "space" is a topological space, or a metric space, or a smooth manifold with a riemannian metric. A space is not necessarily three-dimensional. Mathematicians also talk of surfaces as being spaces; and there are spaces of dimension four, five, *etc.* We say that two spaces are "the same" if there is a one-to-one continuous correspondence between them. That is to say, the correspondence sends nearby points to nearby points. Such a correspondence is called a *homeomorphism*. (Look at a topology textbook for a definition.) Here is a formal statement of the Gordon–Luecke result.

Theorem. *Let K' and K'' be two knots in the three-sphere S^3, and let S^3-K' be homeomorphic to S^3-K''. Then there is an isotopy of S^3 to itself, taking K' to K'' or to the mirror image of K''. Equivalently, there is a homeomorphism of S^3 with itself sending K' to K''.*

At first sight, the Gordon–Luecke result may appear obvious. In fact their result was a major goal of mathematicians for many years. Their article appears in Volume 2 (1989) of the Journal of the American Mathematical Society (pages 371-415).

One understands the depth of their result better if one notes that it is true for knots, but false for links. (The difference between a *knot* and a *link* is that a link has more than one component, while a knot has only one. Thus the trefoil knot shown in 6A is a knot, while the Borromean rings, with three components, form a link.) In 72A we give an example of two links for which the attempted extension of the Gordon–Luecke theorem to links is false.



15(S) What is left when you take away the points of a knot or link from three-dimensional space? What's left when you can't see the knot, not just because there is no matter there, but because space itself doesn't extend to where the knot used to be?

16 Q **Does the phrase "space itself doesn't extend to where the knot used to be" mean anything?**

16 A To answer this question, we need to digress a little, first explaining the difference between topology and geometry. In geometry, concepts such as "straight line" make sense (a straight line is the shortest path between two points) and you can measure lengths and angles. In contrast, in topology, which is sometimes called rubbersheet geometry, you are allowed to stretch the space (remember from 14A that a surface is also called a space) or compress it, but you mustn't tear it or glue together distinct points. When you stretch a rubber sheet, you change the distances between points on the sheet. You also change which paths on the sheet you call straight—for example, if you draw a straight line on a rubber sheet, and then stretch the sheet, it may no longer be straight. What we are planning to do is to station ourselves at a certain point in the complement of the link (or knot), and then stretch the complement, so as to steadily increase the distance between where we are and where the link (or knot) is. When this distance becomes infinite, the link (or knot) will be pushed to infinity, and will no longer be part of the space in which we are making observations. The stretching changes the geometry but not the topology of the link (or knot) complement. The space does not "extend to where the knot used to be" because, after the stretching, the knot is now infinitely far away; you never get there, no matter how long you travel.

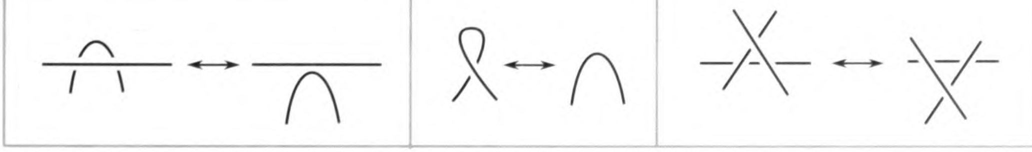
In technical terms, the metric we are aiming at has to be *complete* and have constant curvature. (The curvature is constant if any small piece of the space is isometric to a small piece of a geometry—see the discussion in the inset on hyperbolic space on page 34 for the meaning of a "geometry.") If you simply remove the knot, you have a space with constant curvature, but it is not complete—the edge of the universe is a finite distance away—and so we do not regard this as a satisfactory geometric structure.

Inset 1: Knot Theory

Knots have been a part of human culture for thousands of years, but the mathematical study of knots is relatively recent. Its origins can be traced to the 19th century. The impulse to study knots at that time came from physics, in particular from the vortex theory of matter. Lord Kelvin, a British physicist, conjectured that different physical elements corresponded to different knots tied in the vortices. Hence, it became important to make lists of all different possible knots. Such a list is known as a *knot catalog*, and the problem of creating a correct list without duplication is the problem of *knot classification*.

Knots are three-dimensional objects. One important simplification for the purpose of classifying knots is to work with 2-dimensional representations of knots, known as *knot diagrams*. A knot diagram is gotten by projecting the knot onto a two-dimensional plane. For example, if we shine a light on a knot held in front of a white screen, then the shadow which forms on the screen is such a projection. Wherever two strands of the knot cross in this projection is called a *crossing*, and is represented by leaving small gaps in the farther strand when we draw the diagram.

If you move parts of the knot relative to each other, or walk around the knot to look at it from a different point of view, then the knot diagram changes although the knot remains the same. So the same knot can yield different knot diagrams. Knot diagrams coming from the same knot can be transformed into one another by a set of three simple transformations known as *Reidemeister moves*, pictured below. They correspond to pulling a strand under another strand, straightening out a trivial loop, and moving a strand under a crossing, respectively. If two knot diagrams correspond to the same knot, then we can change one knot diagram to the other by a series of Reidemeister moves. However, given two knot diagrams representing the same knot, it can be very difficult to discover the sequence.



We can ask, "What is the minimum number of crossings in a knot diagram for a given knot?" This is called the *crossing number* of the knot and can be used to construct tables of knots, ordered by the crossing number. The figure on page 12 is the first page of a knot catalog, showing all prime knots with crossing number of 8 or less. The notation 7_3 refers to the third knot with 7 crossings. The early researchers created tables of knots and links up to 10 crossings; with new mathematical knowledge and computers, these tables have been extended to 13 crossings. This improvement may appear slow; however, the number of distinct knots with crossing number less than a given number *n* grows very rapidly with *n*, as the following table shows:

Crossing-number	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*
Number of knots	*1*	*1*	*2*	*3*	*7*	*21*	*49*	*165*	*552*	*2176*	*9988*

To give an explanation of the mathematical theory of knots which has arisen in the last 100 years is beyond the scope of this discussion. Knot theory is now a part of *topology*, an area of mathematics invented in the early 20th century, which is concerned with properties of objects which are not changed by stretching and deforming. Topological tools have been invaluable in distinguishing different knots. In the 1970's Bill Thurston made important discoveries about knot complements which are the subject matter of *Not Knot*. In the last decade, there has been a flurry of new activity in knot theory, which has led full circle. Physicists are again learning about knots because once more it is thought they may play a role in the fabric of nature.

Recommended Reading

Introductory

Clifford W. Ashley, *The Ashley Book of Knots*, Doubleday, 1944.

A wonderful collection of practical knot lore, chiefly in the form of hand-drawings of hundreds of different named knots from many areas of human activity, such as sailing and the decorative arts.

Vaughan F. R. Jones, "Knot Theory and Statistical Mechanics," *Scientific American*, November, 1990.

An introduction to the new developments of the 1980's, which have brought knot theory back into physics.

Lee Neuwirth, "Theory of Knots," *Scientific American*, June, 1979.

Intermediate

Louis H. Kauffman, *On Knots*, Princeton University Press, 1987.

Dale Rolfsen, *Knots and Links*, Publish or Perish, Berkeley, 1976.

Morton Thistlethwaite, "Knot Tabulations," In *Aspects of Topology: In Memory of Hugh Dowker*, Cambridge University Press, 1985.

An interesting account of the state of computation in knot theory. Includes interesting historical anecdotes.

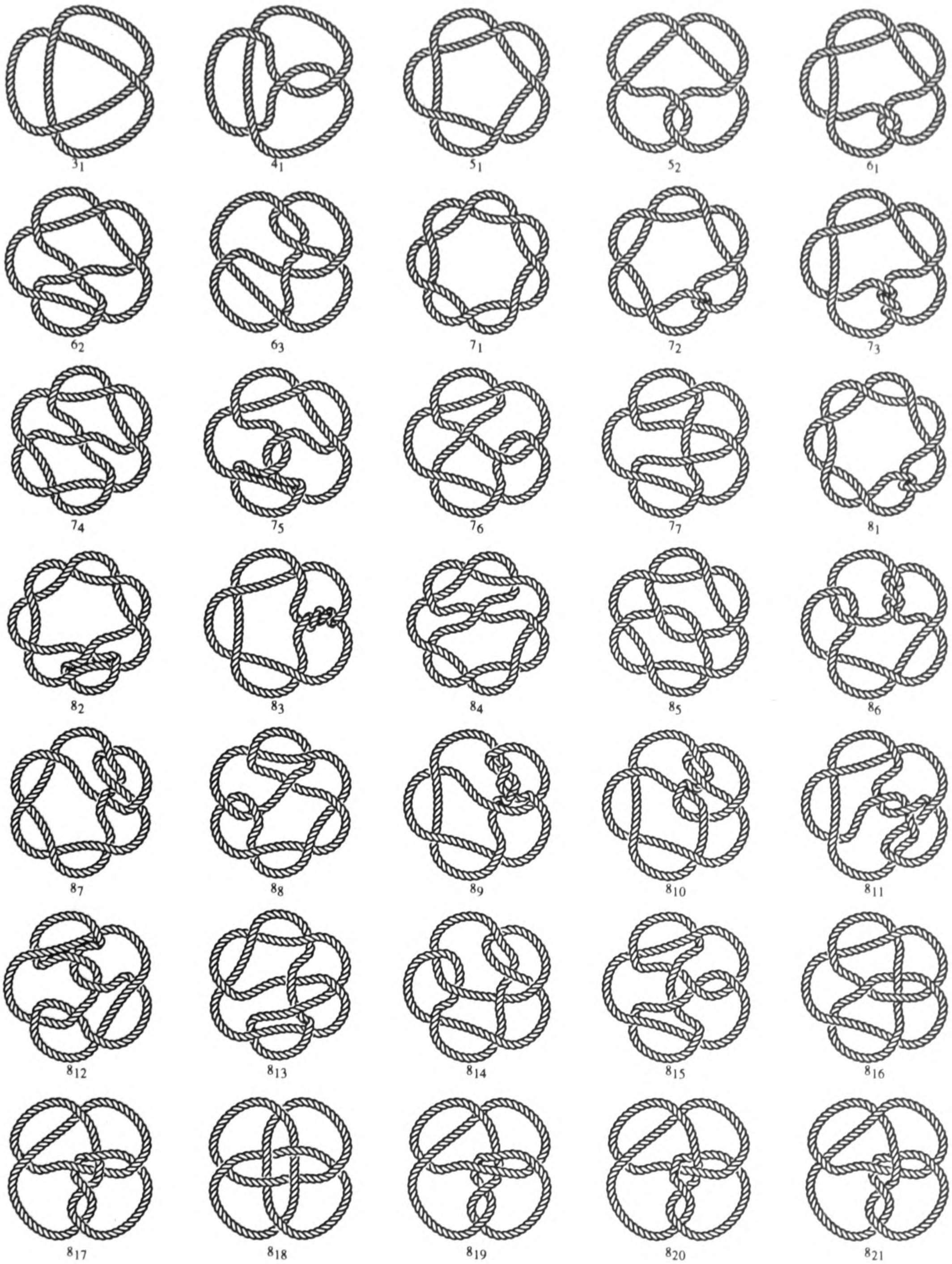
Advanced

Gerhard Burde and Heiner Zieschang, *Knots*, De Gruyter, 1985.

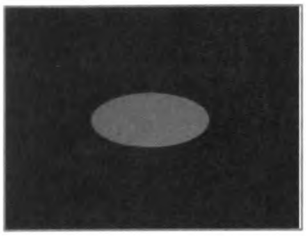
R.H. Crowell and R.H. Fox, *Introduction to Knot Theory*, Springer Verlag, 1977.

Louis H. Kauffman, *Knots and Physics*, World Scientific, 1991.

Knots to 8 Crossings



David Broman and Charlie Gunn



17 (S) We'll look first at a simpler picture: what is life like in a space with a single line missing, or in a plane with a single point removed?

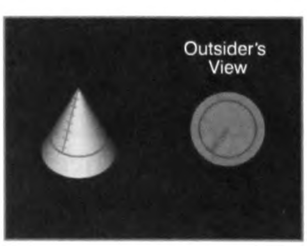
18 [Q] "What is life like?"; are you serious???

18 [A] Partly serious. We are referring to a single aspect of life, namely what the universe looks like to a creature living in it. The way the universe looks is closely connected with the distances between points, as a result of Fermat's principle, discussed in 28A. A more prosaic way of putting our question is to ask what geometry is obtained by stretching distances on the disk so that the distance from the observer to the center increases steadily, until it becomes infinite. The stretching will be done along radial directions only, that is, along rays going straight up the cone surface, directly to the cone point. The length of the blue circle remains fixed.

19 (S) We'll remove the point by pulling it upward, stretching the plane into a cone which grows sharper and sharper. Finally, the point disappears off to infinity and the cone becomes a cylinder. The point at the cone's tip is special: it's called a cone point.

20 [Q] Why pull the point upward?

20 [A] This is just one way of picturing the stretching of the original disk. The fact that the cone point is moved upward is not really relevant. It could just as well have been moved downward, for example. The important thing is that the length of the blue circle stays the same, but the material of which the cone is made stretches in such a way that the distance to the center point increases.



21 (S) As the cone steepens, the radius of a circle about the cone point increases though its circumference stays constant. Meanwhile, an outsider looking from above sees the picture unchanged.

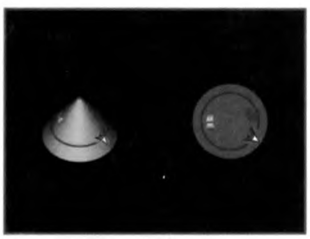
22 [Q] What's an outsider?

22 [A] An outsider is someone who views or thinks about the space in such a way that any object in the space exists in only one place. For example, in the movie frame shown in 21S, *both* views shown are outsider's views, although only the one on the right is labeled as such. The outsider's view of someone looking at a mirror is exactly that; the corresponding insider's view would consist of two universes, one on each side of the mirror, which are reflections of each other. The right-hand image of the movie frame in 21S is seen by someone stationed directly above the cone point, looking down. It is likely that the outsider's view will force light rays to look curved; for example, if light can travel from an object to itself, the ray will inevitably look curved to the outsider.

Very technical: the outsider sees the topological space underlying the relevant orbifold or cone manifold, with the singularities marked.

23 Q The narrator says that the outsider "sees the picture unchanged," but we clearly see movement. What is going on?

23 A What is moving are the marks on a ruler, showing how the scale changes. The actual view of the outsider is unchanged; but if we have a ruler lying on the cone surface (the left hand side of the movie frame in 21), then the markings, at a constant distance apart, appear to the outsider, stationed directly above the cone point, to move closer together as the cone point is moved upwards. If you hold a ruler in your hands, and start tilting it away from you until you are almost looking along the line of the ruler, you will observe the same phenomenon.

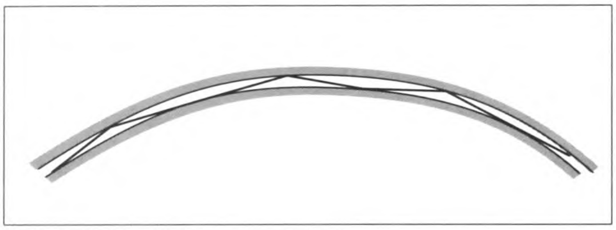


24 S We introduce an observer who lives inside the cone's surface, and a few objects for him to observe.

25 Q Don't you mean someone living *on* the surface? Or do you mean someone living under the cone, like living in a teepee?

25 A Neither of these. The observer is living in the surface like an inkblot lives in blotting paper. Here is a physical model (in our three-dimensional space) of what it would look like to live in a two-dimensional surface, that is, with the surface being the observer's entire universe. Imagine a mirror, which is possibly bent into a curved surface. Spread a thin layer of transparent liquid over the mirror. On top of this thin layer spread a second layer of some perfectly reflecting material. This prevents light from leaving the surface. Then a tiny insect, living in the transparent layer, would see things in the way we are try-

ing to describe. Of course light sources are needed inside the transparent layer



A light ray (solid black line) reflected back and forth between two reflecting surfaces.

for things to be visible. (The physical model is only an approximation which becomes more and more accurate as the thickness of the transparent layer decreases. The reason it is a good approximation is that light will bounce back and forth between the two reflecting surfaces in the model at a glancing angle, as in the picture above, and will therefore approximate the correct path for a light ray within the surface. To make the model completely convincing, one should insist that the reflecting surfaces reflect glancing rays perfectly and absorb non-glancing rays completely.) When the movie talks of an "insider" while discussing cone surfaces, it is referring to this kind of (highly intelligent) insect observer. We should think of the insider as carrying a small ruler, or being able to measure distances by pacing them off.

26 (S) To the insider, light rays travel in straight lines.

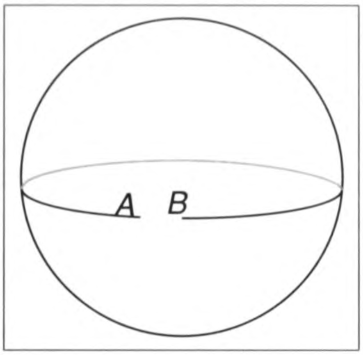
27 Q **What's an insider?**

27 A An insider is someone whose view is that of someone living inside the space, often called the "observer" in the movie and supplement. A light ray meeting the insider's eye is seen as a bright point and the insider thinks of light rays as straight. If light can travel from an object to itself, then, since the insider sees light rays as straight, it is *inevitable* that the insider's view of space will repeat in a symmetrical pattern. The insider sees the geometry of the space, but needs an effort of the imagination to see the topology. The outsider sees the topology, but needs an effort of the imagination to see the geometry. Both views are helpful, and both are necessary for full understanding.

In a space of constant curvature, any segment of any light ray path will look straight to the insider. If the curvature is not constant a segment of a light ray path not meeting the insider's eye may look curved to the insider. If the insider walks over to the light ray path, the path looks straighter and straighter as his or her eye gets nearer and nearer to the light ray path.

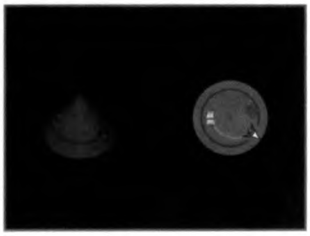
28 [Q] **How can the lines be straight in view of the fact that the surface of the cone is curved?**

28 [A] It is a principle of physics (Fermat's principle) that the path of a light ray in a homogeneous medium is along a shortest path, provided one only looks at rays travelling short distances.



Here is an example to show that light does not necessarily follow the shortest route over longer distances. If the earth were perfectly round, silvered to make a mirrored surface, and covered with a very thin transparent layer and a perfectly reflecting layer on top of that, then a light ray would travel around the equator. For any two points A and B nearby each other on the equator, a light ray could travel either the short route around the equator, or the long route, going all the way around the world.

A more familiar example of light not taking the shortest route is a reflection in a mirror. Hold a candle one foot in front of you, 20 feet from a mirror. The candle can be seen twice, once at a distance of one foot and once at a distance of 41 feet, so the 41 foot trajectory is certainly not a shortest route. How can we state the "shortest path" property so that it covers both the examples of the candle and the rays going round the earth's surface in two possible ways? The required statement is that there is a small interval around every point (except for a point on the surface of a mirror) on a light ray path, where light does take the shortest path. For a small interval around a point on a mirror, light again takes the shortest path, provided we insist that it has to strike the mirror at least once. Light rays from nearby objects are perceived by the insider as travelling straight towards him or her (depending on the sex of the insect). In fact, the insider has no other criterion for straightness, except the path of a light ray, or a shortest path, which is the same thing.

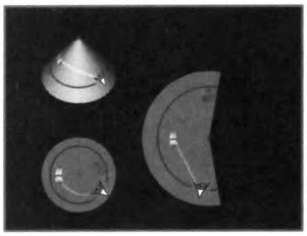


29 (S) But the outsider sees those lines as curving around the cone point.

30 [Q] **Can you explain further why the outsider sees the path of a light ray as curved?**

30 [A] The real question is, "How does light travel on the cone surface?" Once we know the answer to that, we can trace out the rays, and then look at them

from above. The outsider's view is like a map of one of the Earth's hemispheres in an atlas. Geographers and pilots are aware of the fact that a straight line drawn on the map, from London to San Francisco, is far from being the airplane's shortest route. In the same way, the shape of curves on the cone surface is distorted by the projection onto a flat surface, as in the outsider's view. It would be really surprising if the outsider were to see the path of every light ray as straight. The path of a light ray to the cone point *is* seen as straight by both the insider and the outsider.

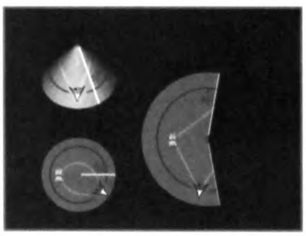


31 (S) To understand the insider's view, cut a cone made of paper and unroll it onto a plane. Since unrolling does not distort the paper, lines which are straight for the insider will be straight lines on the plane.

32 [Q] What is the point of cutting the cone and unrolling it?

32 [A] Once the insider's universe is laid out flat on the plane, as in 31S, it is easy to see how light travels. It travels in ordinary straight lines (by Fermat's principle—see 28A). The process of laying it out flat does not change the way light travels. Not all surfaces can be laid out flat like this—a cone surface can be cut and then laid flat, but the surface of a sphere cannot, no matter how small the cut pieces.

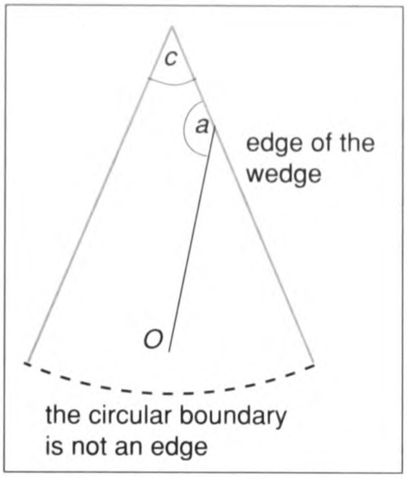
(√) Whether this can be done or not depends on the curvature of the surface. A constant curvature surface can be laid out (the technical word is *developed*) on a sphere or a hyperbolic plane if the curvature is not zero. In the same way, we have to be "escorted into hyperbolic space" in the words of the narrator (52S), in order to lay out the right-angled dodecahedra—it is not possible to flatten out this pattern of dodecahedra in ordinary three-dimensional space.



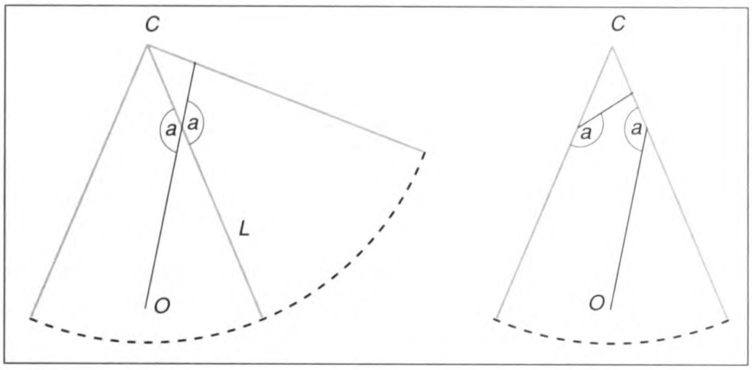
33 (S) This wedge is a building block for the cone and is called a "fundamental domain." The two edges resulting from the cut represent a single line back on the cone. Anything crossing one edge reappears at the other.

34 [Q] How do you use these wedges? And, while you're at it, can you explain what is meant by the "cone angle"?

34 [A] First of all let's get clear what we mean by an "edge" of the wedge. We do not mean the curved edge (dotted in the pictures of this supplement) which comes from the base of the cone—we mean the straight edges (grey in the pictures of this supplement and white in the movie) which result from cutting the cone surface. The curved edge is not really there, because the cone extends downwards infinitely far. For convenience of drawing, the movie shows the cone as a finite object, but the rest of it should be imagined.



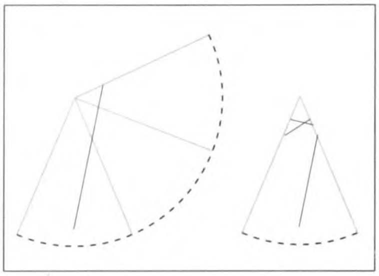
When we cut the cone open along a line L (the white line on the cone surface shown in 33) and lay it flat, we get a wedge like this, with two straight edges (grey), which are copies of L, and one circular edge (dotted), which is not actually supposed to be there. The grey edges meet at an angle c, called the ***cone angle.*** *The solid line represents the path of a light ray (yellow in the movie), starting from the eye and going to an edge of the wedge. To understand where the path goes next, wrap the wedge back over the cone. The two edges of the wedge become the single line L on the cone surface. The light ray meets L at an angle a. Next we consider the picture below.*



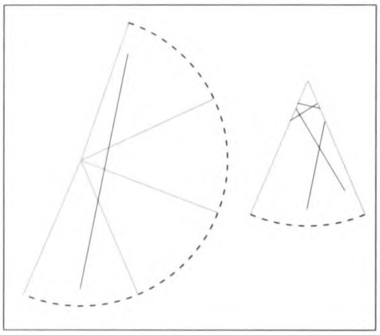
The single wedge on the right can be formed by taking the two wedges on the left, cutting them apart, making them transparent, and then stacking one on top of the other.

We should imagine the two stacked wedges above right as fused into each other to form one wedge. How do you use this to determine the path of a light ray on the cone surface? You wrap the fused wedge around the cone. The grey edges come together at the single grey edge *L* on the cone surface. The two pieces of light ray path on the fused wedge match perfectly when the wedge is wrapped onto the cone surface. Another way of getting exactly the same picture of a light ray path on the cone surface is to wrap the pair of wedges on the left of the picture above around the cone—it wraps neatly over the cone exactly twice.

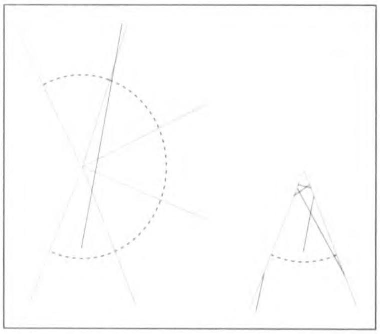
We continue the light ray until it hits the line L on the cone surface a third time. If we now cut the cone along L we get the wedge on the right in the picture below.



Each of the three segments of the light ray shown on the left gives rise to a corresponding segment on the right. Notice that the path of the light ray crosses itself once. (This picture and those above and below are not drawn to exactly the same scale.)



As the light ray has again reached the line L in the picture above, we need to add a fourth wedge to see how it travels. The light ray path crosses itself another two times in the fourth wedge, making three crossings in all.

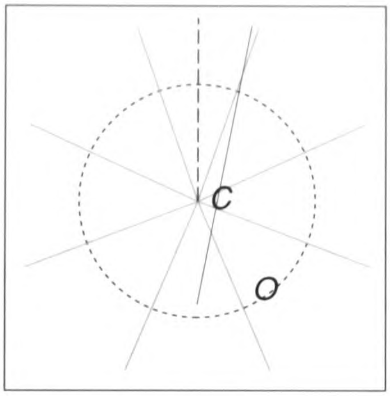


The light ray meets the line L a fourth time and then never again. To see this, we need to extend past the nonexistent circular boundary and add a fifth wedge. (We have reduced the scale, to make the various pictures have the same width. The dotted circular curve is at the same distance from the center in each of the pictures.)

35 Q **How can one see what the cone surface looks like from the insider's point of view?**

35 A The picture below in this section is not really the insider's view—it is a diagram from which the insider's view can be calculated. It enables one to see in an instant what the insider's view would be. Out of the 360 degrees of possible lines of sight from the observer O, there is exactly one line which is not covered by the discussion in 34A, namely the line going directly to the center C. This corresponds to a light ray going directly up the cone surface towards the cone point. What happens to it when it hits the cone point is

not described here—it depends on the precise microscopic optical properties at the cone point. We will just note that, except for special cone angles as stated in the movie script (37S), there is a discontinuity in the view at this one point, when the observer's line of sight changes from looking up the cone, just to the right of the cone point, to looking up the cone, just to the left of the cone point.

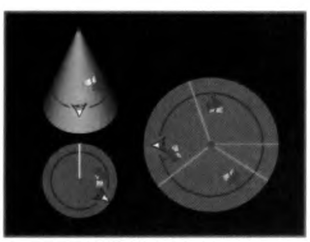


A diagram like this is very helpful for following the path of a light ray to the observer from any point on the cone. The line of sight from the observer O in every direction can be determined by drawing a straight line on this diagram starting at O, noting which wedges the light ray passes through, and then wrapping these wedges around the cone. For example, a light ray path going downwards from O could be understood using only one wedge. This would correspond to a ray on the cone surface, proceeding in a direction away from the cone point.

In the pictures in 34A and 35A, the wedge angle is 46°. The wedge at the top of the picture on this page, divided in two by the dashed vertical line through *C*, is smaller than the other wedges and is not a copy of the other wedges—its angle is 38°.

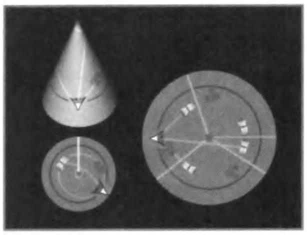
36 Q Why do the light rays in the movie travel from the eye to the object?

36 A Psychologically, one seems to send out searching rays from the eye, when one is looking for something. Ancient Greeks believed that light traveled from the observer to the object observed; but modern physics has shown that light actually travels from object to observer. In the movie we followed the Greeks, because it seemed the natural way to do the graphics.



37 S We can mend the cut by stretching, then gluing the edges. We get the outsider's view, where light rays look curved.

The position of the imaginary cut is arbitrary. These insiders, touring the cone in their car, don't notice as we move the cut to follow them. They look right through the cut, and see what seems to be another copy of the wedge. For special values of the cone angle, a whole number of wedges fit together neatly, but usually the remaining gap is filled with part of a wedge.



38 (S) Here's another way a cone surface differs from a plane. In a plane, only one straight line connects two points; but on a cone there may be several. There may even be straight paths from a point to itself.

We make our two-dimensional objects three-dimensional by stacking them up.

The stack of cone points forms a line, called a cone axis. The fundamental domain for this space is a solid wedge.

Our inside observer sees many copies of the car at once. Which is the real one? They are all equally real. But the light rays from the car reach the observer from several different directions.

Remember our question "What is life like in a space with a single line missing, or in a plane with a single point removed?"

As the cone angle decreases, the space we are looking at changes. Sometimes the wedge-shaped fundamental domains click together perfectly, like a jigsaw, and we get a symmetric picture. When *n* fundamental domains fit exactly around the cone axis, we say the axis has order-*n* symmetry.

As the order increases, the missing line recedes into the distance.

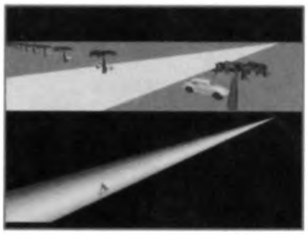
39 Q **What is "order-*n* symmetry," referred to above?**

39 A If this concept gives you trouble, it may help to read the inset on mirrors on page 22.

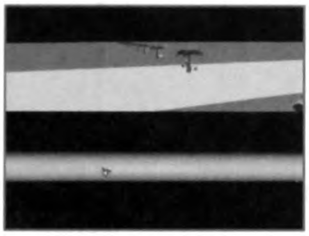
40 Q **Why does the missing line recede?**

40 A The missing line is the same as the cone axis. The distance of the insider to the cone axis is the same as the distance in the two-dimensional case from the insider to the cone point. As the cone becomes sharper, the cone angle decreases and the cone point moves further away. It is important that, as the surface changes, the distance along the path between each car and its succes-

sor is constant. This distance is the same as the length (that is, the circumference) of the blue circle on the cone.



41 (S) In the limit, the missing line disappears off to infinity. The inside observer sees an endless straight row of cars, which are all a single car seen by light travelling different paths.



Our building block or fundamental domain has become an infinite slab, and the cone is now a cylinder. We see how life would look in a space with a single line removed.

42 [Q] **Where does the infinite slab come from?**

42 [A] The fundamental domain in the case of the cylinder is an infinite strip, with parallel sides. This is formed by cutting the cylinder along an infinite vertical straight line. The strip has width equal to the circumference of the blue circle on the cone. In the infinite slab, the strip is laid out in a horizontal plane in three dimensions as in the frames of 41S. As its height varies, the horizontal strip moves through the infinite slab.

Inset 2: Mirror Activities

One difficulty in understanding cone axes is that they don't seem to occur naturally in the physical world. However, there are physical objects that illustrate many of the features of cone axes. In this inset we will describe how to construct such objects, and how to use them to understand cone axes better.

To perform these experiments, you'll need a pair of identical rectangular mirrors, 6"to 12" on a side. The bigger you can get them, the more effective the experiment will be. Glass companies are usually a good source for such mirrors. You'll also need a small amount of wide, sturdy tape (such as duct tape), a protractor, and some large sheets of blank paper.

The first step is to tape the two mirrors together along an edge, so that the resulting seam works like a hinge, and the two mirrors can be opened and shut like a pair of book covers. To do this, turn the mirrors face down on the table so they meet along the edges to be taped. Before applying the tape, separate the two mirrors by a distance approximately equal to twice their thickness. This will give the two mirrors room to move when they close. Then apply the tape. Two thicknesses are advisable to make it strong. Stand the resulting mirror book open on a table so that the mirrors face out towards the edge of the table. Using a little tape, tape one of the mirror backs to the table. This mirror will be fixed during our experiments; the other will be free to move. Finally, slip the large sheet of paper under the free mirror and align it along the fixed mirror.

To begin with, adjust the angle between the two mirrors to be slightly more than a right angle. One easy way to do this is to lay a book flat in the space between the two mirrors, with a corner of the book positioned at the junction of the two mirrors and an edge lying against the base of the fixed mirror. Push the free mirror snugly up against the book. This makes a right angle between the two mirrors. Then open the free mirror up slightly so that it no longer touches the book. Remove the book. Put your head at the level of the table top and look into the angle between the two mirrors. What do you see? You should see mirror, or reversed, images of your face to the left and right, and a partial copy of your face directly ahead. The latter image is not reversed. It is easier to see whether or not the image is reversed, if you apply makeup to one cheek only or wear only one earring.

Pay special attention to the line where the two mirrors meet. We'll call this line the *mirror axis*. With the two mirrors at an angle of slightly more than a right angle, you should see a partial copy of your own face crossing the mirror axis. Now slowly narrow the angle between the two mirrors by moving the free mirror and watch what happens along the mirror axis. As the angle between the two mirrors approaches a right angle, the missing part of your face should appear, as if streaming out of the mirror axis. Stop moving the mirrors when there is a complete copy of your face. The book should fit back snugly between the mirrors. Trace the edge of the free mirror on the white base paper and label this line with number 4. We arrive at the number 4 by counting the number of copies that we see and adding 1, for our "original" face.

As you move the mirrors to form an angle smaller than a right angle, you begin to see 2 additional copies of your face appearing near the mirror axis. Continue closing the mirrors until yet another copy of your face begins to appear around the mirror axis. Stop moving when the images on the left and right sides of the mirror axis join together exactly. Trace the line of the free mirror onto the base paper as before. Label this line with the number 6, the number of copies of your face which are visible in the mirrors plus one. Continue in this way until the two mirrors are too close to see between. Trace the line of the free mirror upon the base paper whenever an exact copy of your face appears around the mirror axis, and label this line with the number of copies of your face which are visible in the mirrors, plus one.

Between having 4 copies of your face and 6, you might have stopped at the position when there were 5 copies (4 images and the original you), and wondered about this. Your face may be too symmetrical to see what is going on, and it is important to introduce some asymmetry. We explained above how to use an earring or makeup, or you could put two different objects like a stone and a coin down on the table next to the mirrors, one to the right of your face and one to the left. You will then realize that the situation with 5 copies is not truly symmetrical.

Label	*Angle*
4	90°
6	60°
8	45°
10	36°
12	30°
14	26°

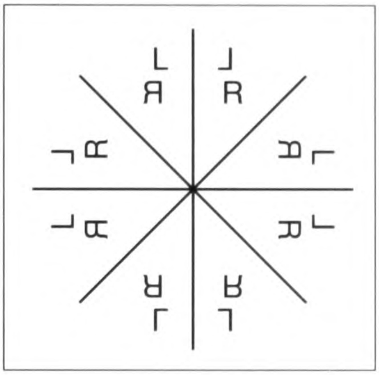
When you're done, remove the white paper and use the protractor to measure the angles formed between the labeled lines and the edge of the paper which was lined up against the fixed mirror. Arrange these measurements into a table with two columns: one, with the label of the number of the line, and two, the measure of the angle between that line and the fixed mirror line. Your table should look like the table in the right margin on page 23.

What is the pattern? We notice that in every case, the angle times the label number is equal to 360°. Another thing we notice is that the label number is always an even number.

In order to understand these observations, let's go back and think about how mirrors work. If we take a single mirror and look into it, we see a reflected image of the world in it. If it's a very good mirror and we don't see the edges, it appears to us as if there is a second universe on the other side of the mirror. The second universe can be constructed by imagining that the light rays which strike the mirror are not reflected but actually continue their path straight across the mirror. The image so constructed is called a *virtual* image. If we now introduce a second mirror at an angle to the first, the situation becomes more complicated. We have a virtual image associated to each of the mirrors, but each mirror also includes a virtual image of the other mirror, and hence includes any virtual image associated with that mirror, and so on.

Another important property of mirrors is that they reverse orientation. A left-handed person looking into a mirror sees a right-handed version of herself. But, if we have more than one mirror, then some of the images will be left-handed and others will be right-handed, depending on how many times the image has been reflected. If it's been reflected an even number of times, the orientation is not reversed; otherwise it is.

On the right we illustrate the experiments above. The letter R and a reversed letter L represent images of a right-handed person; the letter L and a reversed letter R represent a left-handed person. The space between the two mirrors is like the wedge-shaped region in 34A. In 34A we constructed the insider's view by adding copies of the wedge around the cone point; we extended light rays across the boundaries of the wedge. We can draw similar pictures for the pair of mirrors, extending the light rays to create virtual images. For example, when n=8, and the angle between the two mirrors is 45 degrees, we get a two-dimensional cross section through the mirror axis as in the figure on the right.



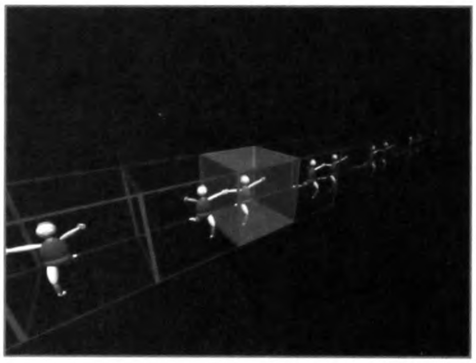
Notice that alternate copies, as we go around the mirror axis, are right-handed. This is because every time the image passes through a mirror, its orientation is reversed. Similar pictures can be drawn whenever the angle between the mirrors is 180° divided by some whole number.

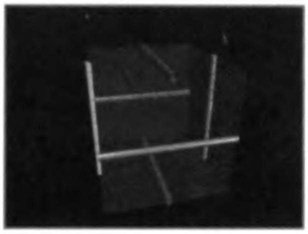
What is the relation of these mirror experiments with the discussion of cone points and cone axes? Suppose we construct mirrors which interchange vertically polarized light and horizontally polarized light as the light is reflected, and suppose we place a source of vertically polarized light between the mirrors. Next we put on spectacles that allow us to see only vertically polarized light. Then we would only be able to see half of the images. Notice how, in the picture above, we can think of the four right-handed images as being obtained from each other by rotation about the mirror axis, rather than by reflection. Polarizing light therefore turns the mirror axis in the picture into a cone axis of order four! In general, polarizing light turns a mirror axis of order $2n$ into a cone axis of order n.

Another Use of Mirrors

In the movie (50S) we visualize how life would be inside the space created by order-2 axes placed on the faces of a cube. Mirrors can also be helpful in understanding the sequence of images that arise here. We ask, what happens if we replace the cone axes on the faces by mirrored faces? That is, what do we see from inside a mirrored room, inhabited by the same figure as before? Each mirror considered separately, like each order-2 cone axis, generates one copy of the figure. The copy created by the cone axis is rotated by 180 degrees, while the copy generated by the mirror has reversed orientation.

To generate new pictures using mirrors instead of cone axes, we replace the red, green, and blue cylinders used to represent the cone axes with thin red, green, and blue frames around the walls themselves, to represent mirror frames. We imagine activating the red, green, and blue mirrors just as we activated the cone axes. The resulting three images are shown below. Notice that half of the figures stand on their left leg, while in the figure from the movie, all the figures stand on their right leg. What other changes do you notice?

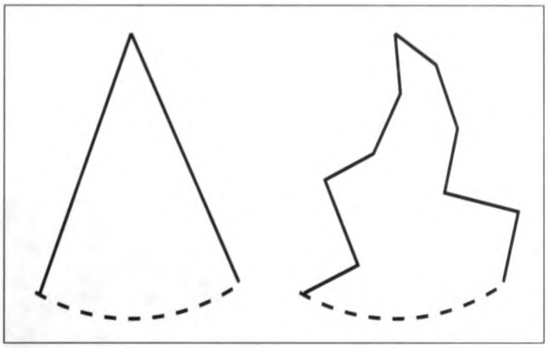




43 (S) Let's now turn to the question, "What is life like in a space from which the Borromean rings have been removed?" We place six cone axes with order-2 symmetry on the faces of a cube. This cube will be a fundamental domain for our experiment.

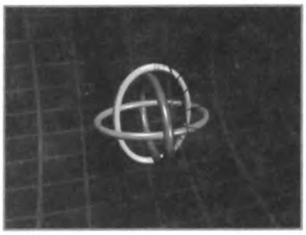
44 Q **Where does this cube come from?**

44 **A** In any given situation, there are usually many different possible fundamental domains. In general, choosing a nice fundamental domain takes some ingenuity, and a lot of work. The fact that the cube is a fundamental domain for the Borromean rings with order-2 symmetry was a mathematical discovery of Bill Thurston. Although it took insight and ingenuity to find the fundamental domain in this nice form, once it's found it's not so hard to check that it *is* a fundamental domain. That is what is done in 45A and 48A; the meaning of "fundamental domain" is exactly what is shown there, namely that when we glue up the faces of the fundamental domain in an appropriate way, we get the (outsider's view of the) space we are interested in.



Here are two different possible choices for the fundamental domain of a cone surface. On the left we have cut along a straight line on the cone surface through the cone point. On the right we have cut along a crooked line.

45 (S) Let's try and see what this space is from the outside. Remember that the walls of a fundamental domain determined by an axis of symmetry should be thought of as being glued together. We first glue the walls containing red axes, and then efface them as they are no longer necessary. Notice how the blue axes are joined together into an ellipse.



We now glue the walls containing green axes. This joins the red axes into an ellipse. We have folded along four of the six faces of the original cube, to form an ellipsoid. All that remains is to fold along the blue axis, which has also become an ellipse. To do this, we make the front and back hemisphere of the ellipsoid bulge up. The green axes too have joined, to form a green ellipse.

46 Q **This is hard to follow. How are the walls being glued?**

46 A Let's just look at the two red axes and the squares in which they lie. Even better, let's look at just one of these red axes and the square in which it lies. This face works like a hinge bending about its red axis. It is glued to *itself* and *not* to the opposite face. The two halves of the face bend just like a hinge until they exactly meet and match each other. The gluing takes place between the two plates of the hinge. (If this were done with superglue on a door hinge, you would never be able to open the door again.) Each face containing a red axis folds up and we get two separate folded-up hinges inside a solid cylinder. These two folded-up hinges are shown for an instant (use slow motion on your VCR) as almost transparent reddish sheets, and then disappear, leaving only the red axes. Notice that the red axes are now inside the solid cylinder, while the two green axes and the two blue axes (which have become a single blue ellipse) are on the surface of the cylinder. Each of the two faces containing a green axis is now a circular disk, with the green axis a diameter. Each of these two faces again behaves like a hinge, but this time the two hinges have circular rather than rectangular plates. Once again, each of these circular faces glues to itself, and not to the other face. Now we have an ellipsoid with a red ellipse inside it, a blue ellipse on the surface, and two green axes, each going through the inside of the ellipsoid. The final gluing of the boundary of the ellipsoid to itself is similar to the earlier cases, but this time there are not two separate faces, each glued to itself, but only one face (the entire surface of the ellipsoid) glued to itself. As indicated above, this sequence is easier to understand if you can play it using slow motion on your VCR.

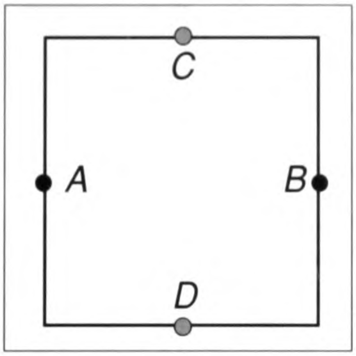
47 Q **What's the flash of light?**

47 A The flash of light which you see occurs as the surface passes through the point at infinity. It is at this moment that the movie acknowledges the essential point that the Borromean rings are being thought of as lying in the three-sphere, rather than in three-dimensional euclidean space.

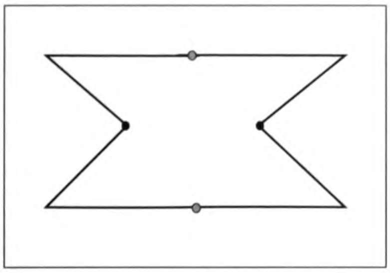
48 S Notice what we have now: the Borromean rings!

49 Q Are there analogues for this construction in other dimensions?

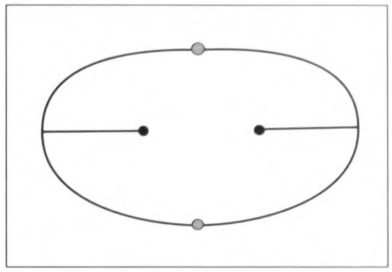
49 A Yes. Here is what happens in two dimensions.

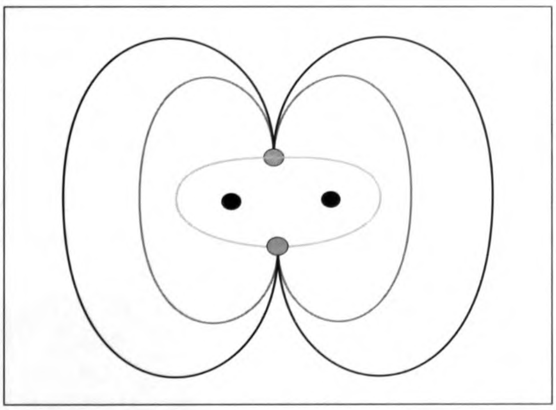


Start with a square, with the center points A and B of two parallel edges colored black and the center points C and D of the other two paralleled edges colored grey. Each of these four points is to be a point of order two.



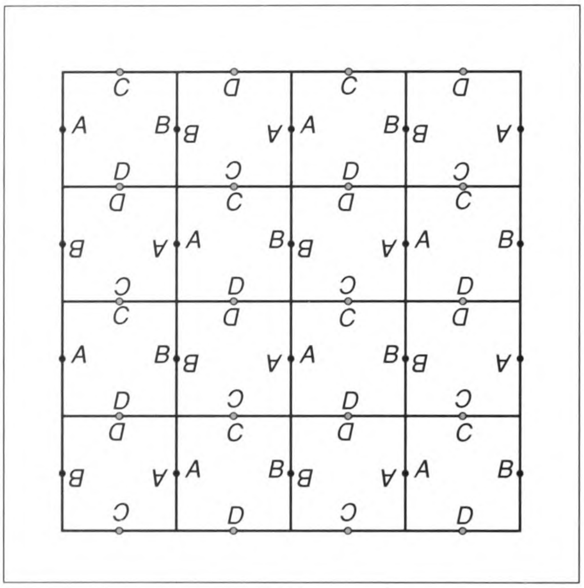
Fold up the edges containing the black dots A and B, until you get an ellipse, with the black dots in the middle of the ellipse.





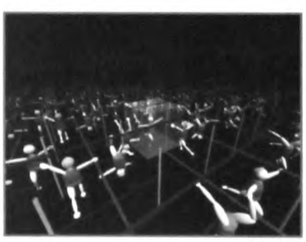
The ellipse above right is also shown at left (in the center of the picture in light grey). The top and bottom of the ellipse bulge out, trying to glue themselves together. Two positions are shown, the first in dark grey and the second black. The gluing cannot be completed in the plane, and we need the point at infinity, just as in the movie.

Alternatively, imagine the black boundary of the figure (the two ovals) to be a zipper, with the zip fastener at one of the grey points. Zipping it up gives us a pouch, and this is homeomorphic to a sphere. The sphere contains the four points, *A*, *B*, *C*, and *D*, analogous to the three-dimensional sphere containing the Borromean rings.



The picture on the left is the insider's view of the space formed by the gluing. Notice that alternate squares are upside down. They are not reflected, but are obtained by rotating the original square through 180°. This picture is analogous to the pictures in 50S, except that it is two-dimensional rather than three-dimensional.

50S Now we've seen that the outside view of our fundamental domain is the Borromean rings. What is the insider's view like? Remember, around each cone axis the insider sees two copies of every object. We first activate the red axes: the image from our fundamental domain gets replicated in the next-door cube by the front axis, and both of these images get replicated by the back axis, and so on, until we have copies of the cube extending all the way to infinity in both directions.

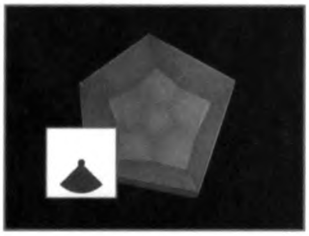


Next, turn on the green axes which face forward. This creates another infinite row of cubes. Turning on the other green axes reproduces these two rows to give four rows, and so on, until we have an infinite horizontal plane of cubes.

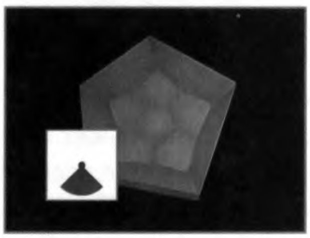
Finally, when the blue axes are turned on, they make the two-dimensional pattern be repeated in layers to fill up the whole space. This is what it looks like to live inside the space created by the order-2 axes on the sides of a cube.

51 Q **What happened to the cone axes in the final image above?**

51 A The red, green, and blue cone axes have been removed from the final image to make it easier to see the pattern formed by the little man. Strictly speaking, the cone axes can't be seen, since they are lines and don't have any width. However, they often appear in the movie as thick bright cylinders to emphasize their location.



52 (S) We can also work out what happens when the Borromean axes have higher-order symmetry. For instance, if we want them to be order-4 axes, we must build a fundamental domain with 90° angles along these axes.

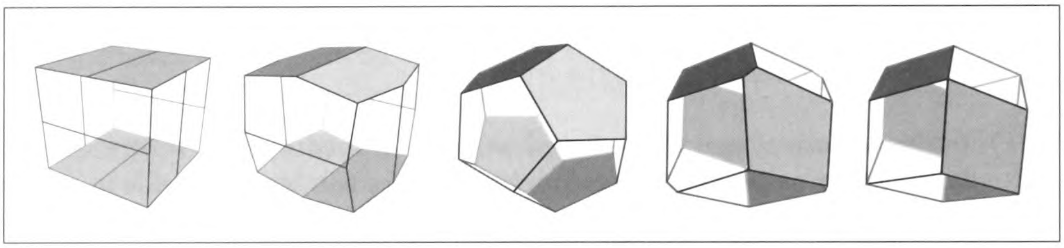


We must modify our cube so it has right angles along its six axes. Impossible, you say? You may not have noticed it, but we're escorting you into Lobachevskian, or hyperbolic, geometry, where this and many other things are possible. This dodecahedron, in true hyperbolic perspective, has 90° angles between every pair of adjacent faces.

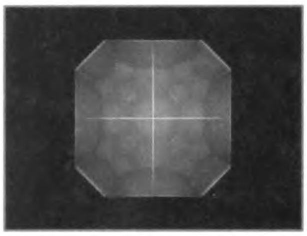
53 Q **How does the cube change into a dodecahedron?**

53 A In the movie, we change from an ordinary cube to a regular hyperbolic dodecahedron, and then to a rhombic dodecahedron. The viewer/reader may find it helpful to see the sequence of corresponding solids in ordinary space. The pictures below show a cube on the left, a regular pentagonal dodecahedron in the center and a rhombic dodecahedron on the right. The faces adja-

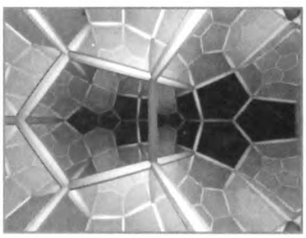
cent to the axes which are green in the movie are shaded. (The picture would have been similar for the blue or the red axes.)



Notice that six of the sides get shorter and shorter as we go from center rightwards in the picture above, until they disappear altogether, and we get four-sided faces, instead of five-sided faces.



54(S) When we look directly down on the red axis, we see that the faces meeting there make a right angle. We can glue three more copies of the dodecahedron around this axis. We can move our viewpoint inside this figure. Now we have four-fold symmetry around one axis. Before exploring further, we'll remove the walls and change the shape and color of the beams.



We can continue to add copies of the dodecahedron around each colored axis. First let's do this for some of the available green, blue, and red axes. Eventually the copies of the dodecahedron fill space without overlap. Just as we tiled ordinary space with cubes, we've tiled hyperbolic space with regular dodecahedra.

55 Q **How can we be sure this is a regular dodecahedron? The rear faces appear to be much smaller than the front faces.**

55 A Visual appearances in hyperbolic space follow different rules than in ordinary space. One of these rules affects the relative sizes of the parts of an object. In particular, this dodecahedron does have congruent faces, each one a pentagon with right-angled corners.

From any viewpoint in hyperbolic space, it is never possible to see more than three faces of an opaque right-angled dodecahedron at one time. In contrast, in ordinary space, it is possible to see half of the faces, that is, six faces, of an opaque regular dodecahedron simultaneously. (This is really due to the fact that any two faces meet at right angles, rather than to the fact that it is in hyperbolic space. In ordinary three-dimensional space as well, whenever we have a convex solid such that any two faces meet in a right angle or less, we can see at most three of its faces at a time.) A right-angled dodecahedron exists in hyperbolic space, but cannot exist in euclidean space.

Consult the inset on hyperbolic geometry on page 34 for a more detailed explanation of why the rear face of the dodecahedron always appears a constant fraction smaller than the front face.

56 Q **In the picture of space tessellated with dodecahedra, if we follow one edge or beam as far as possible in a straight path, we can see both ends in our field of view. How can that be?**

56 A The appearance of lines in hyperbolic and euclidean space is very different. If a straight line in hyperbolic space does not pass too near the observer's eye, the two ends of the line can both be seen at the same time, and so it appears to be a finite interval. In fact, the angle at the eye subtended by a line is always less than 180°, and decreases as the observer moves further away. (See the inset on hyperbolic geometry on page 34 for more details.) In contrast, in ordinary space, the two ends of an infinite straight line can only be seen if you turn your head. For example, look at the beams in 50S.

The same phenomenon is repeated with planes. A hyperbolic plane inside hyperbolic space appears to the observer as a disk of finite visual size. As the distance of the eye to the plane decreases, the size of the visual disk decreases. The edges of the disk bend away from the observer. If you look carefully at the tessellations of hyperbolic space in the movie, you will see many such planes. For example, any two infinite white beams that meet perpendicularly determine a plane. These two white beams give four points at infinity, coming from the ends of the beams. You should be able to see a pattern of distant beams forming a circle through these four points—this gives roughly the boundary circle of the visual disk corresponding to the plane in question. You should be able to make out a tessellation of the plane by pentagons or by four-sided figures (at the end of the movie). These pentagons and four-sided figures are faces of certain dodecahedral tiles.

If you watch the movie, you will see these circles changing into ellipses as they go off the side of the screen during the fly-through of hyperbolic space. This is not a phenomenon of hyperbolic space; it is a consequence of doing the computer graphics for a wide-angled camera lens. We used a 90° lens because we wanted you to see as much as possible of what was going on. In order to get the correct perspective, you need to sit right in front of the TV set, so that the angle subtended at your eye from the left and right edges of the TV screen is 90°.

See also the inset on hyperbolic geometry on page 34.

57 Q **Why have you made the beams thinner as they go to infinity?**

57 A They don't get thinner. They stay the same thickness, but their apparent thickness changes. In the euclidean plane, the set of points in a plane which are a

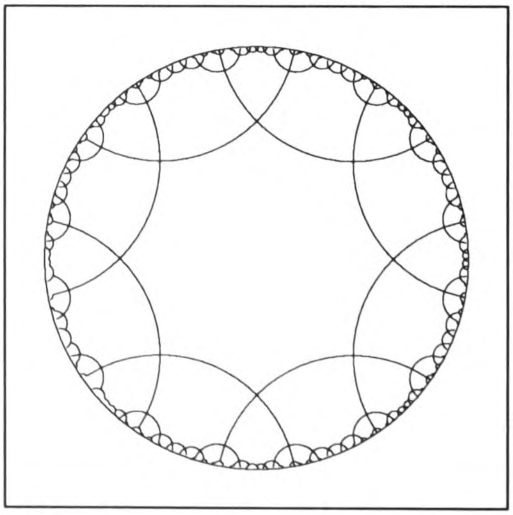
fixed distance to a line is another line parallel to the first (actually two lines, one on each side of the original one). This isn't true in hyperbolic geometry, where there are many lines parallel to a given line. (See the inset on page 34 for details.) In hyperbolic geometry the fixed distance construction gives us a pair of convex curves, one on each side of the original line. These curves are called *equidistant* curves—they are not straight lines.

The beams are created by sweeping an octagonal cross-section along the length of a line. In euclidean space this process yields a prism with eight faces, all of which are parallel to the central line. Sweeping an octagonal cross-section along a line in hyperbolic space creates a figure, which appears to shrink to a point at each end of the line. Slicing this figure with any plane containing the central line yields a pair of equidistant curves.

Later in the movie (63S), when hyperbolic space is rotated so that we look along the beams, we see that they are straight. Beams in hyperbolic space elsewhere in the movie, although straight, are seen as curving away from us! This remarkable phenomenon is of course impossible in everyday euclidean space.

58 Q **Can the analogy you were presenting in 49A be extended to the right-angled dodecahedron?**

58 A Yes. The square in 49A with each side broken into two can be considered as an octagon in which four of the angles are 90° and four of the angles are 180°. This is analogous to the cube in 53A. If we make all of the angles of the octagon into right angles, we obtain a two-dimensional analogue to the right-angled dodecahedron in the movie (the lower of the two movie frames in 52S). The associated tiling on the hyperbolic plane is shown in the picture below. Pictures developing this sequence of ideas further are in 64A.



The straight lines of the hyperbolic plane are shown here as circular arcs, meeting the boundary perpendicularly. This is not an accurate representation of the hyperbolic plane, owing to the distorting effect of a map on flat paper (see the inset on hyperbolic geometry on page 34). There are an infinite number of octagons, each having the same size. (Check that each region does indeed have eight sides.)

Inset 3: Hyperbolic Space

What properties should one demand of space in a theory of the universe? We might demand that the space be *homogeneous*, that is, all measurements are independent of where the observer is in the universe. We can also require that the space be *isotropic*, that is, the measurements are also independent of the direction in which she is facing. For example, we want the results to be the same if the observer and all the equipment are turned upside down, or turned through a right angle, so that the observer is lying down with respect to her previous position. We also want the mirror image observer and equipment to record the same measurements as before, when the same experiments are done. Geometry is the name of the branch of mathematics in which such universes are studied, and each such universe is also called a *geometry*.

In dimension one there are two possible geometries, namely the real line and the circle. A one-dimensional observer can only face in two directions, forward and back. In dimension two there are two familiar universes, namely the plane and the sphere. The relevant geometries are called *euclidean* and *spherical*. Spherical geometry is extremely important, because it is used for navigation of the earth's surface; it also comes into many aspects of engineering. But only euclidean geometry is normally taught in high schools. A comparison of some of the properties of euclidean and spherical geometry can be seen in the table on page 37. In spherical geometry, the role of straight lines is played by great circles. (The equator and any circle on the sphere of the same length as the equator is a great circle, but the other lines of latitude are not.)

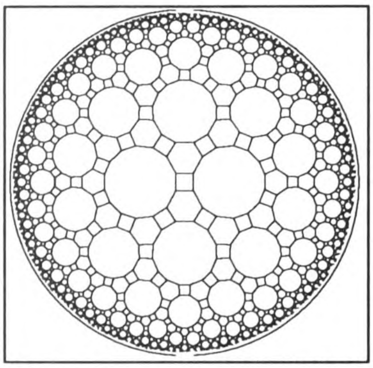
At the beginning of the 19th century, Karl Friedrich Gauss in Germany, Janos Bolyai in Hungary, and Nicolai Lobachevsky in Russia independently discovered a third geometry, namely *hyperbolic* geometry. Their discovery solved a problem which had been troubling mathematicians for more than 2000 years. One of the great achievements of ancient Greek mathematics was to give *axioms* for euclidean geometry. This is a short list of properties from which all other properties can be deduced. One of these axioms was the **Parallel Postulate.** This states that, given a line L and a point P not on L, there is one and only one line through P that does not meet L (such a line is called *parallel* to L). In hyperbolic geometry, there are an infinite number of lines through P which do not meet L. Since hyperbolic geometry does satisfy all the other axioms of euclidean geometry, this proves that the parallel postulate is not a consequence of the other axioms. Gauss apparently was the first to make the discovery but withheld publication. Of the three discoverers, Lobachevsky made the most systematic developments of the new science. In honor of this achievement, hyperbolic geometry is also known as Lobachevskian geometry.

In each dimension greater than one there are exactly three geometries satisfying the properties of homogeneity and isotropy, namely, euclidean, spherical and hyperbolic.

To be strictly accurate, there is also elliptic geometry, a close relative of spherical geometry. The main difference is that in elliptic geometry any two lines intersect in exactly one point instead of two, as in spherical geometry. One might also be tempted to think that there are other geometries based on quotients of spherical, euclidean, or hyperbolic space by a group. But in any other case, the observer would see copies of herself in particular directions, violating the condition of isotropy.

Although hyperbolic geometry is extremely important, it is less well-known than either of the other two; whereas the whole of the plane and the whole of the sphere can be embedded in ordinary three-dimensional space, it is possible to embed only part of the hyperbolic plane without distorting distances.

Although we cannot embed it, we can make maps of the hyperbolic plane. In describing the surface of the earth or the moon, we work with many different types of maps or *projections* as they are called. Which projection is chosen depends on what properties are under consideration. For example, one can use projections where area is correctly represented, or projections where the angles between curves is correctly represented, but one cannot do both at the same time. It is not possible to correctly represent distances on the earth's surface using a map on a flat piece of paper. In the same way, our maps of the hyperbolic plane will inevitably be inaccurate in some respect or other. There are several different maps of the hyperbolic plane that are popular among mathematicians. Mathematicians often call such maps *models* because the word "map" is usually used by mathematicians for other purposes. Here is a picture of one of them.



The picture on the left is a map of the hyperbolic plane, divided up into squares, hexagons and 14-gons. Each hexagon is the same size in the hyperbolic plane, but they appear to be different sizes because of distortions in the map. This map is called the **conformal,** *or* **Poincaré,** *model of the hyperbolic plane. It represents angles correctly but not distances. Straight lines are represented by circular arcs that intersect the boundary circle at right angles. The distance of the boundary circle to any point inside the figure is infinite, as can be seen by counting the number of regions which must be crossed to reach the boundary.*

A map of the surface of the earth cannot show *all* points on the surface—at least one has to be omitted, no matter how badly we are prepared to distort the scale. This is because the sphere cannot be embedded in the ordinary plane. We are in a better situation with the hyperbolic plane—the whole of it can be embedded in the ordinary plane, although, as already stated, this requires increasingly more distortion as we go out to infinity in the hyperbolic plane. In the picture above, the boundary of the circle represents points which are infinitely far away: they are not actually in the hyperbolic plane at all.

The other favorite model of hyperbolic space is known as the *projective*, or *Klein*, model. In contrast to the conformal model, it represents straight lines as pieces of euclidean straight lines, while angle measurements are not correct. The projective model is a better choice for visualizing the insider's view of hyperbolic space and is used throughout *Not Knot*.

With this all-too-brief introduction to hyperbolic geometry, let's find examples of hyperbolic geometry in the movie. We'll use the table on page 37 to identify distinguishing features.

Our journey into hyperbolic space begins as we increase the order of symmetry around the cone axes on the faces of the cube in 52S. We transform the cube into a right-angled dodecahedron (52S and 53A). Can we use the table to see that this dodecahedron is hyperbolic? Each of its faces is a regular pentagon that has right angles at each corner. In euclidean space, a regular pentagon has 108° at each vertex. By connecting the central point of the pentagons to the vertices with line segments, we can decompose each pentagon into 5 isosceles triangles. Each triangle has central angle of $360°/5 = 72°$. In the euclidean case, the base angles are each $108°/2 = 54°$, while in the right-angled pentagon the base angles are $90°/2 = 45°$. Hence the triangles in the latter case have angle sum equal to 162°, which is less than 180°. According to the table on page 37, if a triangle has angle sum less than 180° it is a hyperbolic triangle. So these triangles, and the right-angled dodecahedron that they form, can only exist in hyperbolic space.

The tessellation of hyperbolic space by these dodecahedra shows that there are many small copies of the dodecahedra at a relatively close distance to the observer. This expresses one of the fundamental qualitative properties of hyperbolic space: there is more space there. If we look on page 37 at the expression for the circumference of a circle of radius r in hyperbolic space, we see that it is proportional to $\sinh(r)$, which grows very quickly with r. It's instructive to work out an example. Suppose that the unit of length in three-dimensional hyperbolic space is the same as one meter in ordinary space. We construct a hemispherical swimming pool in hyperbolic space 50 meters across. Then the volume of water it contains is more than three times the volume of our planet!

A side effect of the above property is evident when we begin to move around in hyperbolic space. When one retreats a distance r from an object, its apparent size drops off much more rapidly than in euclidean space, because it must share the visual field with so many more other objects than it would at the same distance in euclidean space. In order to halve the apparent size of an object in ordinary space, you need to stand roughly twice as far from it. In contrast, in hyperbolic space, if an object is moved .7 units of length further away, no matter how far away it is to begin with, its apparent size is more than halved. The consequence is that when we look at the right-angled dodecahedron in 52S, the face furthest from us looks *much* smaller than the face nearest us. A face that is two dodecahedra away looks much smaller than that, and the faces rapidly become too small to see as they get further away. This is the geometric explanation behind 55A. This is also equivalent to the property described as *divergence of lines* in the table on page 37.

Another related signature of hyperbolic geometry present in the movie is that it is possible to see straight lines in their entirety, without having to turn your head to look in the opposite direction to see the other end. This is discussed in 56A.

The edges of the dodecahedra in hyperbolic space are covered by beams. See 57A for a discussion of how the shape of these beams expresses the hyperbolic geometry of the surrounding space.

A more subtle feature of hyperbolic space is the existence of *horospheres*, or embedded copies of the euclidean plane. This is discussed more fully in 65A.

This touches on some of the many ways that hyperbolic geometry appears in *Not Knot*. If you want to understand hyperbolic geometry more deeply, the references below can serve as a starting point for further exploration.

Recommended Reading

Introductory

Jeffrey Weeks, *The Shape Of Space*, Marcel Dekker, 1985.

This is perhaps the best introduction to the ideas presented in this inset.

Intermediate

H.M.S. Coxeter, *Non-Euclidean Geometry*, University of Toronto Press, 1965.

Marvin Jay Greenberg, *Euclidean and Non-Euclidean Geometries*, W. H. Freeman, 1980

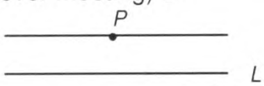
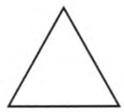
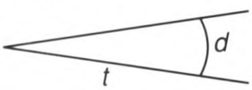
Advanced

Alan Beardon, *The Geometry of Discrete Groups*, Springer-Verlag, 1983.

William Thurston, *The Geometry and Topology of Three-Manifolds*, Volume I, Princeton University Press, to appear.

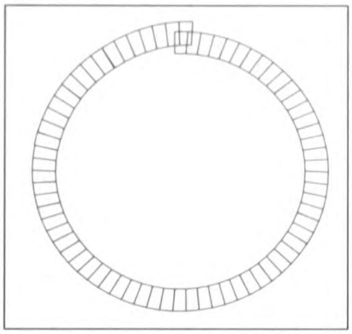
This is the primary source for the mathematical exposition of the ideas in this inset and in "Not Knot" in general. It is very readable and includes many illustrations.

Properties of 2D Geometries

Feature	Sphere	Euclidean plane	Hyperbolic plane
Number of lines through point P parallel to (that is, never meeting) L 	None	1	∞
S = sum of angles of a triangle	$> \pi$  Area = $(S - \pi)$	$= \pi$  Area not determined by angles. Triangles with the same angles and different areas are called "similar."	$< \pi$  Area = $(\pi - S)$
Circumference of circle of radius r	$2\pi \sin(r)$	$2\pi r$	$2\pi \sinh(r)$
Divergence of lines 	Bounded $d < \pi$	Linear $d = kt$	Exponential $d \sim ke^t$
Area of entire surface	4π	Infinite	Infinite
Area of circle radius r	$2\pi (1 - \cos(r))$	πr^2	$2\pi (\cosh(r) - 1)$
Gaussian curvature	+1	0	-1
A portion of a surface in euclidean three-space with the same gaussian curvature	"round" 	"flat" 	"saddle-shaped" 
Everyday objects with this curvature	Cabbage leaf, surface of a ball	Paper, surface of cylinder, surface of cone	Kale leaf, cooling tower at power station

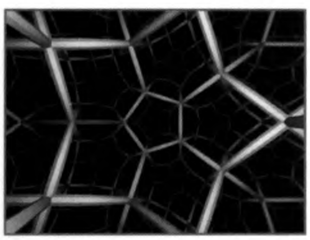
59 Q **How do we know that these dodecahedra tile the space?**

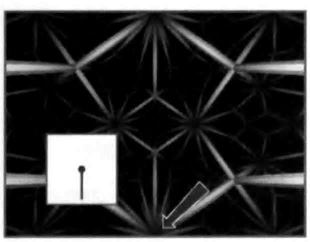
59 A They certainly look as though they tile, but that is not enough for the mathematician. Maybe if we keep on laying down these dodecahedra, they will curl around and bump into ones already there.



Here's the kind of situation we have to avoid. The tiles start out looking as though there is no problem, but after laying down a very large number, they start to overlap incorrectly.

To show that overlapping does not occur, we apply a theorem due to Henri Poincaré, which tells us that if we can fit the dodecahedra together perfectly around each edge, then they tessellate. This condition is satisfied because each edge has order-n symmetry, where $n=4$ for the white edges and $n=4$ for the colored edges in the movie frame shown in 60. In later frames, the symmetry is still order 4 for the white beams, but the order for the colored beams is higher. A proof of Poincaré's theorem is given by B. Maskit in "On Poincaré's Theorem for Fundamental Polygons," *Advances in Mathematics*, (1971) 7 pp. 219–230.





60 (S) Let's fly around a little in hyperbolic space to get a better feel for it. Notice how quickly apparent size changes as we move. This is one of the biggest qualitative differences between our everyday space and hyperbolic space. This is what it looks like to live inside the space created by order-4 axes along the edges of the dodecahedron.

By adjusting the angles at the colored axes, we can derive similar pictures for order-5 symmetry... order-6 symmetry ... and so on for all higher orders. Notice that as the order of the symmetry increases, the colored axes of the dodecahedron grow very short and move far away. Finally, in the limit, the red, green, and blue axes have receded to infinity!

The resulting figure is called a rhombic dodecahedron.

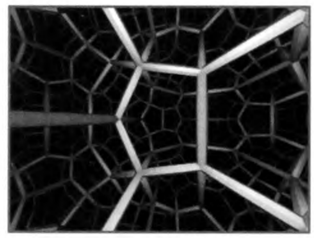
61 Q **What is a rhombic dodecahedron?**

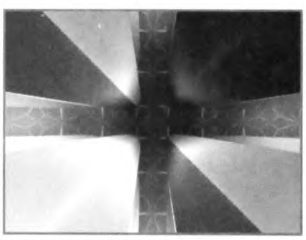
61 A The best explanation is provided by making a model—you can use the cutout diagrams provided in the inset on Activities (page 47). It is a solid with 12

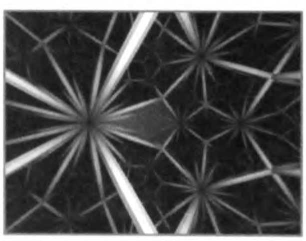
faces, 24 edges and 14 vertices. In euclidean space, each face is a rhombus (a parallelogram with all sides having equal length). There are two types of vertex—there are 6 of the first type, each having 4 edges coming into it, and there are 8 of the second type, each have 3 edges coming into it. A picture is shown on the right in 53A. In the movie, the rhombic dodecahedron discussed is a hyperbolic version of this, in which the vertices with 4 incoming edges are at infinity, and the edges at a vertex with 3 incoming edges are mutually perpendicular, like the 3 edges coming into a vertex of the ordinary cube. Recall from 55A our discussion of how the back face of the regular hyperbolic right-angled dodecahedron appears much smaller than the front face. The same phenomenon occurs here: once again the back face appears much smaller than the front face, though they are of the same size.

62 Q **I kept expecting to see order-3 axes, but it looks like you skipped this case entirely. Why?**

62 A You're correct. The movie moves directly from the order-2 case to the order-4 case. The order-3 case has slightly less interesting symmetry than the order-4 case, whose fundamental domain is a regular polyhedron, to be exact, a regular right-angled dodecahedron. However the order-3 case certainly exists; the picture at the right shows a tessellation of hyperbolic space given by it.







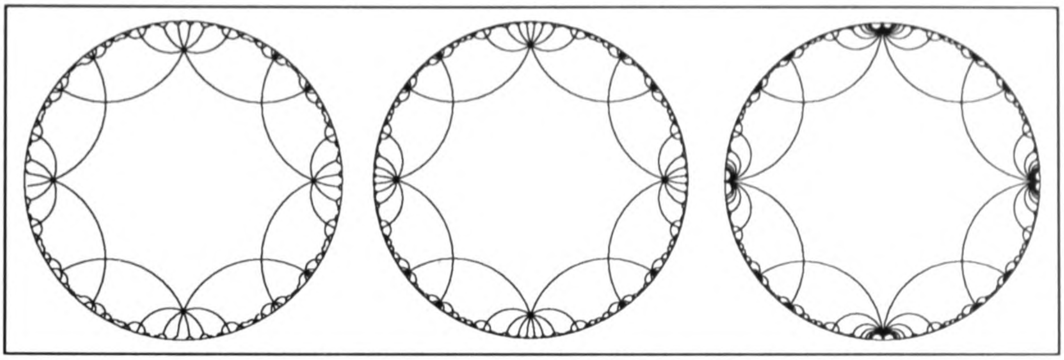
63 S The six colored axes have been transformed into 6 vertices at infinity. To better understand the geometry of this shape, we add transparent walls to one copy of the figure and rotate hyperbolic space around its center. As the vertices at infinity pass behind our eye, we see interesting patterns.

From inside the rhombic dodecahedron the link has become infinitely far away, so that light can never reach it. This is the picture of the complement of the Borromean rings: we have succeeded in bending light so that it continues forever without hitting the link.

64 Q **How does the analogy (begun in 49A and 58A) with two dimensions continue?**

64 A Just as in three dimensions, we can push points off to infinity by increasing the order of symmetry. In each of the following three pictures, alternate

angles in the central octagon (and in every octagon) are right angles. In the



picture on the left, the other four angles are 45°. In the central picture, the other four angles are 36° and in the picture on the right, the other four angles are 0°. In the picture on the right, the four points have been pushed to infinity.

65 Q Can you say a little more about the "interesting patterns" referred to in the first paragraph of 63?

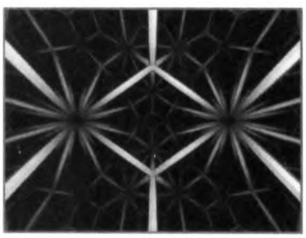
65 A In order to follow this discussion, you really need to pause the movie at the first frame of 63S. (If you can find a high-quality VCR, you will be able to see much more in this frame.) The frame shows a horizontal and a vertical strip of little squares, surrounded by 4 large white rectangles occupying the 4 corners of the frame. (The white rectangles of the movie are not adequately reproduced in black and white. The image in the supplement shows the four rectangles divided into regions by diagonal lines, which are the edges of the beams. Some faces of the beams are in shadow and show up as dark regions in the black and white image. It is much easier to understand the corresponding frame in the movie itself, using the pause button on the VCR.)

The eye of the observer at this moment is located between 4 beams, each connected to the same vertex at infinity of the rhombic dodecahedron, and this vertex is behind the observer's head. These 4 beams occupy the 4 corners of the frame. Each square in the strips between these 4 beams represents another rhombic dodecahedron sharing this same vertex at infinity. If we could see through the 4 beams in the corners of the frame, we would see a complete two-dimensional array of these squares.

In hyperbolic space it is not possible to have a square with all angles right angles. (Consult the inset on hyperbolic geometry on page 34 for details.) What we are seeing here is a euclidean pattern of squares on a surface inside hyperbolic space known as a *horosphere*. A horosphere is a copy of euclidean space that has been embedded inside hyperbolic space with the correct euclidean distances between points. It is a convex surface, which is the limiting case of the surface of a ball in hyperbolic space, where the surface stays a constant distance from the observer, while the radius of the ball gets bigger and its center goes to infinity. (The corresponding limit in euclidean space is a euclidean plane.) The limit of the centers is a point at infinity called the *center of the horosphere*. In our case the center of the horosphere is a vertex at infinity of the rhombic dodecahedron, and this vertex is directly behind the observer's eye. The view is a very, very close approximation to a euclidean view of a euclidean plane, though in theory it is not 100% right. The approximation becomes better and better as the camera moves towards the vertex at infinity. (But the camera can't go too much further out, because the beams occupy a greater and greater proportion of the space, until there is no gap at all between them, and after that they overlap, filling out completely the part of hyperbolic space behind the observer.)

66 Q **What, if anything, does the process of changing a cone into a cylinder have to do with this sequence of spaces talked about in the second paragraph of 60S?**

66 A The process of passing through this sequence of dodecahedra with the cone angles decreasing in the sequence 90° (order-4 symmetry), 72° (order-5 symmetry), 60° (order-6 symmetry) and so on, is very much the same as the process which leads from a cone to a cylinder, through the same angles. There were in fact three scenarios, with steadily increasing complication. The first was the cone surface, changing to a cylinder. We changed the cone angle continuously to zero, but we could have jumped through various cone angles, 90°, 72°, 60° and so on, just as in the case of the dodecahedra, omitting the intermediate angles. The second scenario was the three-dimensional version with a single cone axis; again the change of cone angle was continuous, but could have been in jumps. The third was the changing dodecahedra. In theory, it would have been possible to make the changes here also continuous instead of in jumps, but the pictures would have been much harder to make.

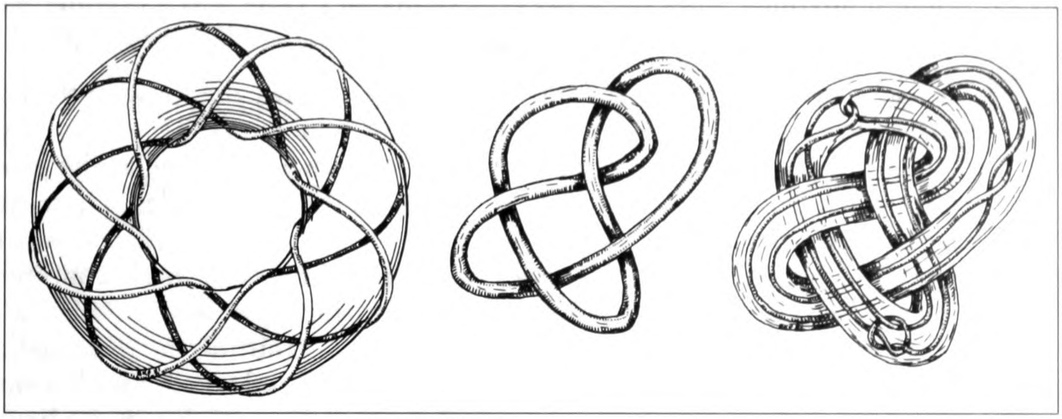


67 S This geometry is called a hyperbolic structure for the link complement. According to theorems of Mostow, Marden, and Prasad, there is at most one hyperbolic structure for any link complement. According to a theorem of Bill Thurston, all knots and links, with some simple exceptions, have complements that admit hyperbolic structures.

68 Q **What are the exceptions?**

68 A The exceptions are as follows. First we have those links which consist of more than one piece; that is, if you have two pockets, it's possible to put part of the link in one pocket and part in the other, without any of it hanging outside the two pockets. Secondly we have *torus knots*; these are knots which can be drawn on the surface of an ordinary, unknotted solid torus, as on the left in the illustrations below. The third (and only other) type of exception is a *companion knot* (or link). To give the idea of what a companion knot is, we give

one example. In the center illustration below you see a knot, which we call M (Medium size cross-section). On the right we have a thick knot L (Large cross-section). L is a thickened-up copy of M and they are the same knot in the sense explained in 6A. Inside L you see another knot S (Small cross-section). S is called a *companion knot* to M. Notice that we could also move M rigidly sideways to lie inside L; this does not make M into a companion knot of itself or of L. Once M is moved inside L, every cross-section of L meets M in a single circle. It is not possible to move the companion knot S inside L so that every cross-section of L meets S in a single circle. We also have to insist that every cross-section of the knot playing the role of L meets the knot playing the role of S. Finally we have to insist that the knot playing the role of M really is knotted.



A *companion link* is defined as a link S which is associated to a simpler knot or link M, by placing one or more components of S inside a component of M, in a similar manner to that already described for knots.

69 Q Can you give references and precise statements for the theorems of Mostow, Marden, and Prasad? And what about Thurston's theorem?

69 A The theorems of Mostow, Marden, and Prasad state that a homotopy equivalence between two complete hyperbolic manifolds of finite volume is homotopic to an isometry, provided the manifolds have dimension at least three. Mostow proved the result in all dimensions greater than three, for compact manifolds only. Prasad extended this to manifolds of finite volume. Marden proved the same result independently for hyperbolic manifolds of dimension three. Later Mostow proved much more general results applying to all Lie groups. References are as follows:

G.D. Mostow, *Quasi-conformal mappings in n-space and the rigidity of hyperbolic spaceforms*, Publications Mathématiques IHES, 1968, Vol. 34, pp.53-104.

G.D. Mostow, *Strong rigidity of locally symmetric spaces*, Princeton University Press, Princeton, 1973.

G. Prasad, *Q-rank-one lattices*, Inventiones Mathematicae, 1973, Vol. 21, pp. 255-286.

A. Marden, *The geometry of finitely generated Kleinian groups,* Annals of Math., 1974, Vol. 99, pp.383-462.

Thurston's theorem states that an irreducible and boundary-incompressible compact Haken manifold, such that each incompressible torus is homotopic to a boundary component, can be given a geometric structure. Since complements of knots and non-splittable links are Haken, we can easily deduce that a link or knot complement has a hyperbolic structure, unless it is splittable or is the unknot, or a torus knot, or a companion knot, or link. (In a number of specific examples, including the Borromean rings, Bob Riley first noticed the hyperbolic structure.)

Thurston described his results and discussed them without proofs in:

W. P. Thurston, *Three-dimensional manifolds, Kleinian groups and hyperbolic geometry,* Bull. AMS, Vol. 6, 1982, pp. 357-381.

His results are supposed to appear in a series of papers. However only the first of these has appeared so far, though some of the others are available in preprint form. The first in the series is:

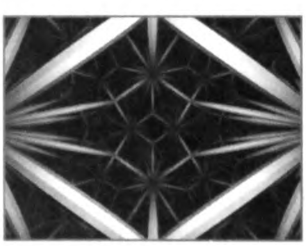
W. P. Thurston, *Hyperbolic structures on 3-manifolds I: Deformations of acylindrical manifolds*, Annals of Math., Vol. 124, 1986, pp. 203-246.

A fairly complete proof of Thurston's theorem can be put together from the following sources:

J. Morgan, *On Thurston's uniformization theorem for three dimensional manifolds*, pp. 37-125, in *The Smith Conjecture*, Academic Press, 1984.

C. McMullen, *Amenability, Poincaré series and quasiconformal maps*, Invent. Math., 1989, Vol. 97, pp. 95-127.

C. McMullen, *Iteration on Teichmüller space*, Invent. Math., 1990, Vol. 99, pp.425-454.



70 (S) Therefore, knots and links are completely determined by pictures such as this one. This conclude our guided tour. We have had a glimpse of how mathematicians understand knots and links, and the spaces around them.

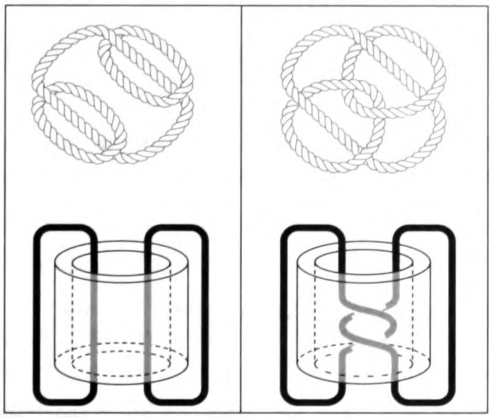
71 Q What exactly does the phrase "pictures such as this one" mean?

71 A The phrase "pictures such as this one" needs a bit of interpretation if it is to remain strictly correct. It is correct without qualification for knots by the theorem of Gordon and Luecke, but the statement needs some qualification for a link. In the case of the complement of a link, the structure is once again hyperbolic (with the exceptions described in 68A). But the hyperbolic structure, although it determines the complement as a topological space, does not necessarily determine the link. A single link complement may correspond to an infinite number of different links (see 72A for a more extended discussion of this), and so a single picture, like the picture of the Borromean link complement in the movie, may correspond to an infinite number of different links. To specify the link completely, a small amount of extra information is needed; namely we need to specify, for each component of the link, the direction of a short circle going round (that is, linking) the component. Such information could have been added visually to the pictures presented in the movie.

Miscellaneous Questions

72 Q **If two links have the same complement, are they the same link?**

72 A In 14A, we promised an example of two links which are different, but whose complements are the same. Such an example shows that the Gordon–Luecke Theorem, stated in 14A, does not apply to links, and helps us to realize what a deep result they have proved.



Each of the two links shown here made of rope has three components. The two links are not the same as each other, and not the same as the Borromean rings. The way we know this is by considering what happens when we delete one component (each link gives us three possible deletions).

Below each rope picture, we show another picture in which the largest rope link has been replaced by a transparent tube.

To see that each rope link is the same as the link shown beneath it, imagine the transparent tube to be made of some putty-like material, so that it can be molded into a ring. Now tilt this ring so that the part nearest you is tilted upwards and the part away from you is tilted downwards. Each picture below then becomes the same as the picture above it. To show that the complement of the rope link on the left is the same as (homeomorphic to) the complement of the rope link on the right, we need only show the same thing for the pictures below. The homeomorphism we choose is the identity map outside the (convex hull of the) transparent tube. Inside the (convex hull of the) transparent tube, the map is defined by a twist that increases steadily as we proceed down the tube, starting with a twist of 0°, and ending with a twist of 360°.

73 Q **Is the knot shown in 3A knotted?**

73 A No, it is unknotted, that is, it can be moved to the position of a circle lying on a flat surface. The easiest way to see this is to knot some nylon cord in exactly this form, melt the two ends together, and then untie it. An alternative useful method is to use an extension cord, tying it in the correct form, and then plugging the two ends together. Be careful to create the correct crossings.

74 Q **In 2A you promised an overview. How about it?**

74 A The aim of the movie is to show how to introduce a hyperbolic structure on the complement of the Borromean rings. Thurston's theorem states that, with the exceptions listed in 68A, every knot complement and every link complement has a (complete) hyperbolic structure. This shows that there are many, infinitely many, examples of knots and links that could have been used instead of the Borromean rings. But the Borromean rings have particular symmetry and the pictures are particularly pleasing. In addition, we thought it would be nice to use a link associated to several important solids—the cube, the regular dodecahedron, and the rhombic dodecahedron. Since we used a link complement, rather than a knot complement, several people have suggested that the title should have been "Not Link" rather than "Not Knot." This is true, and you can call it "Not Link" if you prefer, and especially if you don't like puns.

The hyperbolic structure on the link complement is approached by a sequence of hyperbolic structures, where each of the Borromean rings is a cone axis of a certain order. These give rise to the sequence of the movie in 60S. In order to understand this process more clearly, we precede the discussion of the Borromean rings, where there are three cone axes, with the simpler situation where there is only a single cone axis. But this can be understood most easily in terms of a cross-section through the axis.

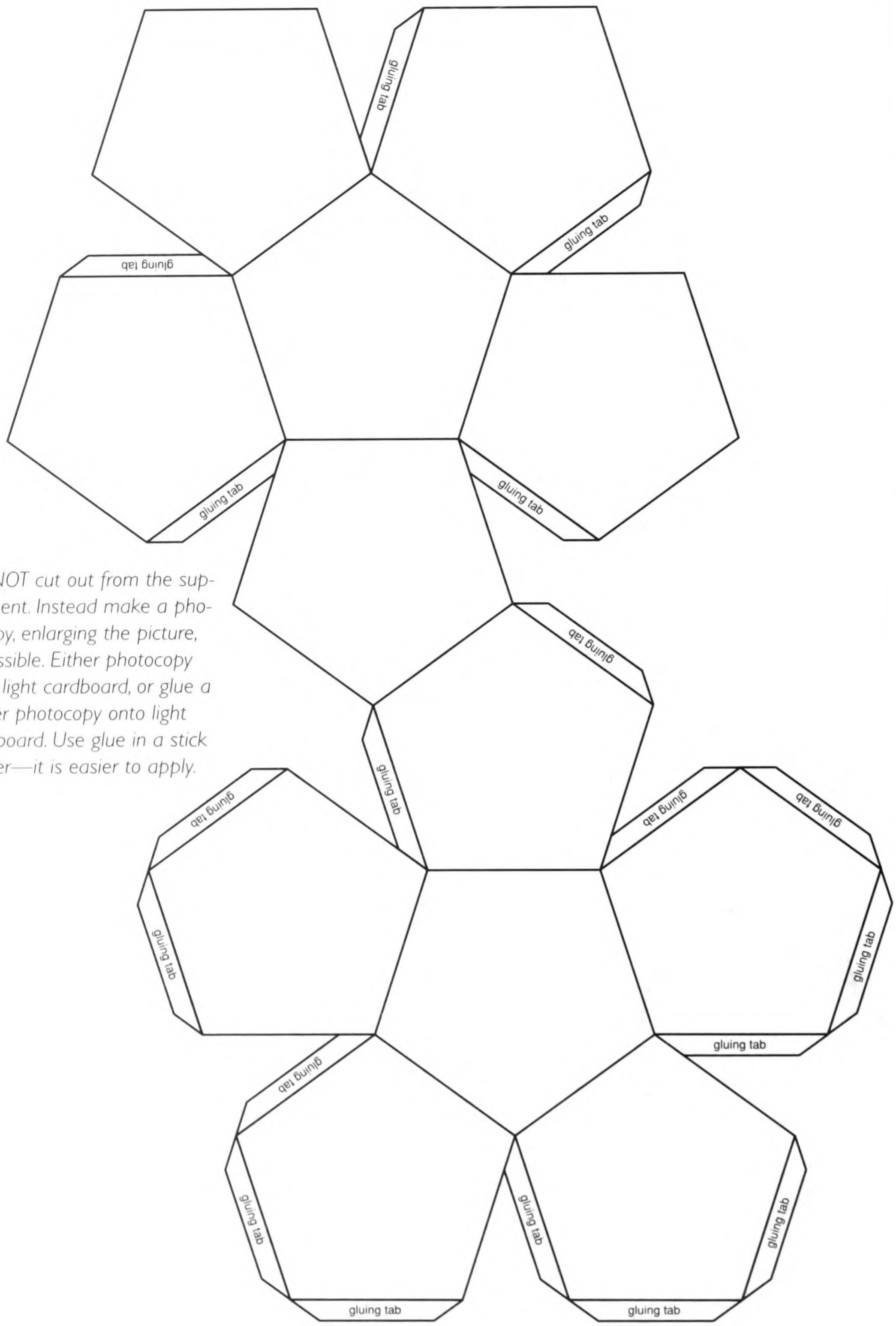
In this way, we see that the first thing to understand is a disk with a cone point. So the general structure of the movie is: First we introduce knots and links and explain what they are. Then we talk about the aim of putting a structure on the space, so that the knot or link becomes infinitely far away. Then we turn to a point in a disk, and show how that can be pushed infinitely far away. Next, we push a single axis in three-space infinitely far away. Finally we do the same thing for the Borromean rings, pushing all three rings to infinity at the same time.

The phenomena which occur in the simpler situations reproduce themselves in the more complicated situations, and so the simpler situations are discussed first.

Inset 4: Activities

- Make the cutouts on the next two pages.
- Figure out how to color the faces of the pentagonal dodecahedron so that corresponding faces on the rhombic dodecahedron have the same colors.
- Make cones of many different cone angles using light cardboard or paper. Follow paths of light rays using the technique shown in 34A. Find out how many times the light rays cross each other. Glue together the two straight edges of the wedge and see how the light ray crosses the edge of the cut. It should cross the cut without bending, and the light rays on each side of the cut should match correctly.
- Draw an observer's eye and a car on one wedge, and then figure out how many times the observer would see the car.
- Make wedges of the same size which fit together so that a disk is formed from a whole number of wedges. Draw light rays on these. Convince yourself that when *n* wedges fit together perfectly, you get exactly *n* images.
- Fix a length (say 2 inches). Make a sequence of wedges of different angles. On each wedge, construct a blue arc of circle, centered at the point of the wedge, so that the length of the arc of blue circle on the wedge is 2 inches. (You may want to work out a formula which gives the radius of the required circular arc in terms of the wedge angle.) See how the point of the wedge gets further away as the angle is decreased.
- Extend the exercise of 34A to the cylinder. This is formed from an infinite strip, rather than a wedge. Show how to construct light rays on a cylinder by drawing straight lines on a strip of paper. Convince yourself that a light ray on the cylinder travels in a spiral on the cylinder. Convince yourself that if you place an observer's eye and a car on the cylinder, then there are an infinite number of different light rays connecting them. (If the cylinder is vertical, it is best if the eye and the car are at different heights.)
- Make hyperbolic paper from a large number of equilateral triangles of the same size, by gluing them together seven at a vertex, instead of the euclidean number, which is six at a vertex. Use strong paper, not card, so that it can bend. If the triangles are reasonably large, you will find it easier to make the paper. This hyperbolic paper is an approximation to a piece of the hyperbolic plane.
- Carry out the mirror activities described in the inset on page 23.
- Get an electric extension cord. Make a knot and plug the two ends together. Then try to unknot it. Try this with the knot shown in 3A. Show that you can move a figure eight knot (shown in the center of 68A) to its mirror image. Try to do the same for the trefoil (shown in 6A)—it can't be done.
- Take a square piece of cloth as in 49A, and sew on two black buttons to represent the points *A* and *B* and two grey buttons to represent the points *C* and *D*. Sew up the sides of the square, using the description in 49A as far as you can. Notice that you get a surface homeomorphic to a sphere.
- Make lots of copies of the square shown in 49A and place them together to get the pattern shown in the last picture in 49A.

Cutout model for the regular dodecahedron

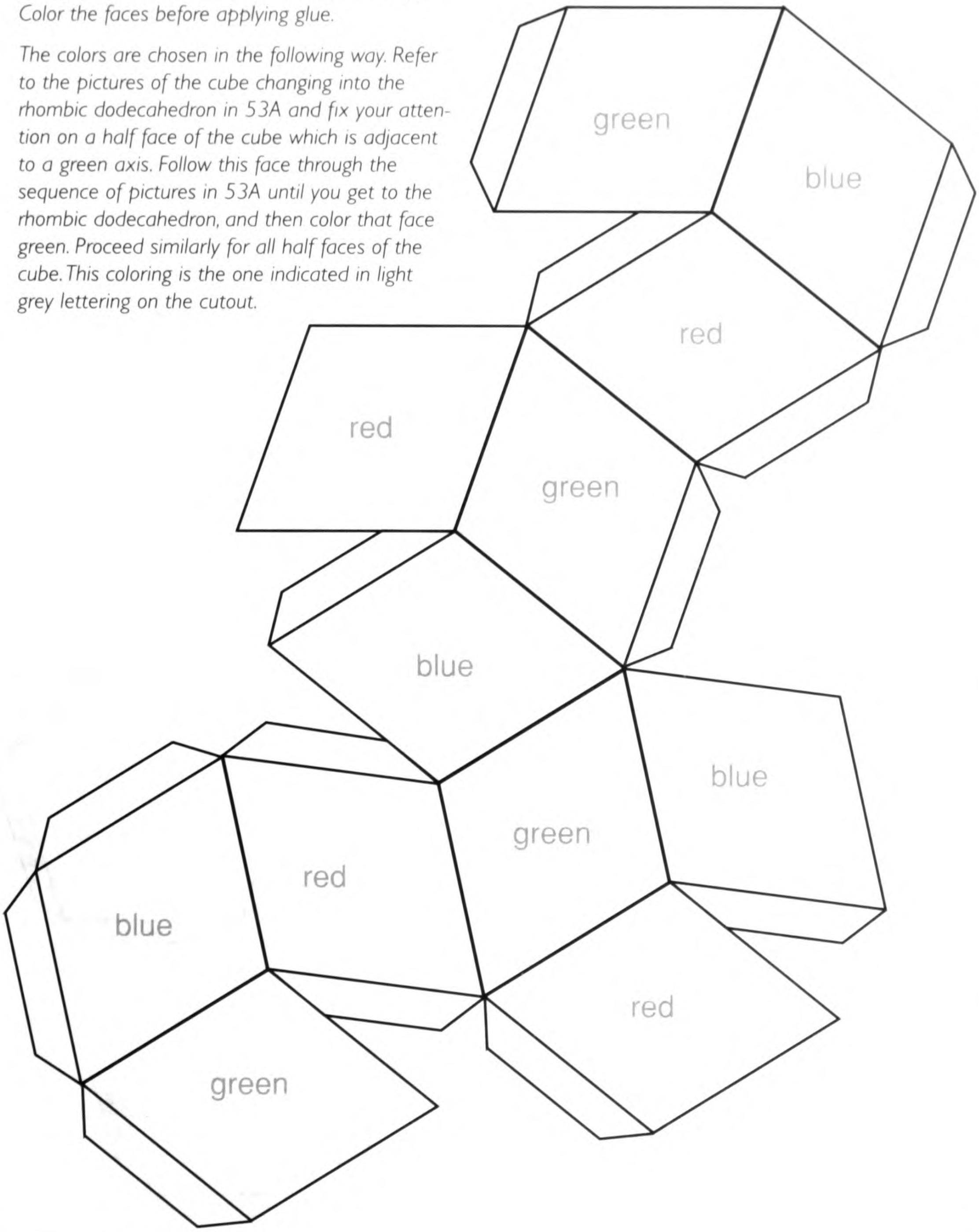


Do NOT cut out from the supplement. Instead make a photocopy, enlarging the picture, if possible. Either photocopy onto light cardboard, or glue a paper photocopy onto light cardboard. Use glue in a stick holder—it is easier to apply.

Cutout model for the rhombic dodecahedron

Do NOT cut out from the supplement. Instead make a photocopy, enlarging the picture, if possible. Either photocopy onto light cardboard, or glue a paper photocopy onto light cardboard. Use glue in a stick holder—it is easier to apply. Color the faces before applying glue.

The colors are chosen in the following way. Refer to the pictures of the cube changing into the rhombic dodecahedron in 53A and fix your attention on a half face of the cube which is adjacent to a green axis. Follow this face through the sequence of pictures in 53A until you get to the rhombic dodecahedron, and then color that face green. Proceed similarly for all half faces of the cube. This coloring is the one indicated in light grey lettering on the cutout.



NOT KNOT is a guided tour into computer animated hyperbolic space. It proceeds from the world of knots to its complementary space, what's not a knot. Profound theorems of recent mathematics show that most knot complements carry the structure of hyperbolic geometry, a geometry in which the sum of the three angles of a triangle is always less than 180 degrees and in which there is so much room that, with unit of length one meter; a hyperbolic hemispherical swimming pool 25 meters in diameter contains 23 times the (ordinary) volume of the earth. The video shows the geometry of the knot complement, the space around the knot, changing into hyperbolic space, and then you see what it is like to "fly through" the hyperbolic space. This video goes beyond the Euclidean geometry of your school days and you see a curved space sometimes studied in modern cosmology. Watch it repeatedly: each time you will think of more questions and discover answers to some of your previous questions. You may find the answers to your questions in this supplement. It leads you through the video script and provides explanations for many puzzling questions that have been asked by students and mature mathematicians alike. You will be rewarded with a better understanding of how mathematicians see and explore space.

DISTRIBUTED BY
Jones and Bartlett Publishers, Inc.
20 Park Plaza, Suite 1435, Boston, MA 02116
617-482-3900 or 1-800-832-0034

ISBN(video)0-86720-240-8
ISBN(pbk)0-86720-297-1