You Can't Hear the Shape of a Drum

Spectral geometry yields information about geometric shapes from the vibration frequencies they produce

Carolyn Gordon and David Webb

Around the middle of the last century, the French philosopher Auguste Comte speculated that knowledge of the chemical composition of stars would be forever beyond the reach of science. Scarcely a decade later, however, the new science of spectroscopy was born—and with it the ability to determine the chemical makeup of stellar atmospheres. This development precipitated a revolution in physics. The second most abundant element in the universe, helium, was first discovered spectroscopically in the solar atmosphere, and eventually an enormous database accumulated, permitting the identification of molecules in interstellar space by extraordinarily indirect means: observation of the natural vibration frequencies of a system—or, to borrow a metaphor from acoustics, observation of the pitches at which a molecule naturally "rings." Once the set of vibra-

Carolyn Gordon and David Webb met in 1985 when both joined the faculty of Washington University in St. Louis; they soon began collaborations that led to a number of joint research papers and to their marriage. Gordon had earlier completed her Ph.D. at Washington University with a thesis in the area of differential geometry. She was a Lady Davis Fellow at the Technion in Haifa, Israel, and taught at Lehigh University before returning to Washington University. Webb received his Ph.D. from Cornell University with a thesis in algebraic K-theory, then taught at the University of Waterloo before moving to Washington University. Gordon and Webb currently reside with their two-year-old daughter in Hanover, New Hampshire, where both are on the faculty of Dartmouth College. Although Gordon and Webb are accustomed to working at close range, much of the research described here was carried out via transatlantic phone calls, electronic mail and faxes while Gordon was attending a conference in Germany. The research described in this article is joint work with Scott Wolpert of the University of Maryland. Address: Department of Mathematics, Bradley Hall Room 301, Dartmouth College, Hanover, NH 03755 Internet: David.L.Webb@dartmouth.edu.

tion frequencies of a system (its *spectrum*) was determined in the laboratory, the information could be applied to astronomical observations, with spectacular consequences: Chemical elements in stellar atmospheres could be identified by their characteristic spectral "fingerprints."

The empirical achievements of spectroscopy generated a number of vexing theoretical questions. For example, from knowledge of the structure of an atom or molecule, how can one predict the discrete set of vibration frequencies of the system? Conversely, given the spectrum of a vibrating system, what can be inferred about the system's structure? The development of quantum mechanics was in part motivated by the need to provide a sound theoretical underpinning for the observations that were turning spectroscopy into a mature empirical science. Sir Arthur Schuster, who coined the word "spectroscopy," articulated the important insight that the study of spectra could be used to obtain important structural information about atoms and molecules, not merely to identify the species being observed. In 1882, Schuster wrote:

> We know a great deal more about the forces which produce the vibrations of sound than about those which produce the vibrations of light. To find out the different tunes sent out by a vibrating system is a problem which may or may not be solvable in certain special cases, but it would baffle the most skillful mathematician to solve the inverse problem and to find out the shape of a bell by means of the sounds which it is capable of sending out. And this is the problem which ultimately spectroscopy hopes to solve in the case of light. In the meantime we must wel-

come with delight even the smallest step in the desired direction.

Schuster's words remain apt over a century later: The relation between a vibrating system and its characteristic vibration frequencies is still poorly understood. This article addresses some small steps in the desired direction, which Schuster would presumably have welcomed.

In 1966 Mark Kac drew the attention of mathematicians to spectral problems by posing a question that is a prototype for those arising in spectral theory: Can one hear the shape of a drum? To consider this question, mathematicians imagine a region D in the Euclidean plane (for example, the interior of a triangle). Suppose that a drumhead is fashioned in the shape of the region D. If one knew all the frequencies of sound the drumhead can emit, could one infer the shape of D? Below, we answer this question negatively by means of an example. It is easy to understand the main result: The example turns out to admit an elementary explanation that imposes few mathematical demands on the reader. This argument, however, furnishes no indication of how one might arrive at the example in the first place. The reader who works through the sections preceding the punch line will, we hope, be rewarded by an indication of some typical features of vibrating systems.

The Vibrating String

To understand the mathematical issues raised by Kac's problem, consider first one of the simplest examples of a vibrating system: a vibrating string. A string of length L can be idealized by the interval $[0,L]$ of all real numbers x between and including 0 and L ($0 \le x \le L$). We represent a possible configura-

tion of th[...]
tion $f(x,t)$[...]
non-nega[...]
terpreted[...]
of the po[...]
points of[...]
tion f m[...]
ditions f[...]
of the tin[...]
string mt[...]
tion, wh[...]
the given[...]

(This is r[...]
is the acc[...]
by Newt[...]
eration] i[...]
right-hai[...]
ture of th[...]
asserts th[...]
graph is[...]

Figure 1. Objects that vibrate, be they drumheads or atoms, have characteristic vibration frequencies. From a knowledge of these vibration frequencies can be gleaned information about the vibrating object. For example, the empirical science of spectroscopy has studied the vibration frequencies of atoms and molecules to provide information about astronomical objects. Yet a full understanding of the relation between a vibrating system and its characteristic vibrations has remained elusive. The authors seek an answer to a question posed by mathematician Mark Kac: Can one hear the shape of a drum? That is, can one infer the shape of a vibrating object by its characteristic frequencies? By exhibiting two different drumheads with identical vibration frequencies, the authors answer Kac's question negatively.

tion of the vibrating string as a function $f(x,t)$ defined for x in $[0,L]$ and any non-negative real number t; $f(x,t)$ is interpreted as the vertical displacement of the point x at time t. Since the endpoints of the string are fixed, the function f must satisfy the boundary conditions $f(0,t) = f(L,t) = 0$ for all values of the time variable t. Such a vibrating string must also satisfy the wave equation, where ∂ (delta) is the change in the given variable:

$$\frac{\partial^2 f}{\partial t^2} = \frac{\partial^2 f}{\partial x^2}$$

(This is reasonable: The left-hand side is the acceleration at the point x, which by Newton's law [force = mass × acceleration] is proportional to the force; the right-hand side is roughly the curvature of the graph of f, and the equation asserts that the more tightly curved the graph is near x, the greater the restoring force felt by a small piece of the string centered at x.)

One way of solving such an equation is first to seek stationary solutions—that is, equations of the form $f(x,t) = g(x)h(t)$. Such a solution represents a basic standing "waveform," given by the graph of g, which has a fixed underlying shape but whose amplitude is varied by a function h of time. Figure 2 depicts a stationary solution at various times. The condition that the ends of the string remain fixed is now expressed by $g(0) = g(L) = 0$.

When a function of the form $f(x,t) = g(x)h(t)$ is substituted into the wave equation, the result is

$$g(x)h''(t) = g''(x)h(t)$$

or equivalently

$$\frac{h''(t)}{h(t)} = \frac{g''(x)}{g(x)}$$

In this equation, the left side does not depend on x, so the right-hand side cannot depend on x either; similarly, the right side does not depend on t, so neither does the left side. Thus both sides of the equation must be some constant. It can be shown that this constant is negative, so we denote it by $-\lambda$ (lambda), where λ is greater than 0. Thus our problem reduces to solving two equations, one involving only the spatial variable x, the other involving only the time variable t:

(1) $g''(x) = -\lambda g(x)$

(2) $h''(t) = -\lambda h(t)$

One can check that the functions $\sin(\sqrt{\lambda}\, x)$ and $\cos(\sqrt{\lambda}\, x)$ are solutions of the spatial equation (1), and, in fact, the general solution turns out to be a combination of these:

$$g(x) = A\,\sin(\sqrt{\lambda}\, x) + B\,\cos(\sqrt{\lambda}\, x)$$

lest

er a
n a
stic
un-
ome
ion,
bly

tion
ems
type
Can
con-
ians
lean
a tri-
fash-
f one
l the
r the
ques-
mple.
esult:
n ele-
s few
ader.
es no
ive at
eader
eced-
be re-
ypical

issues
r first
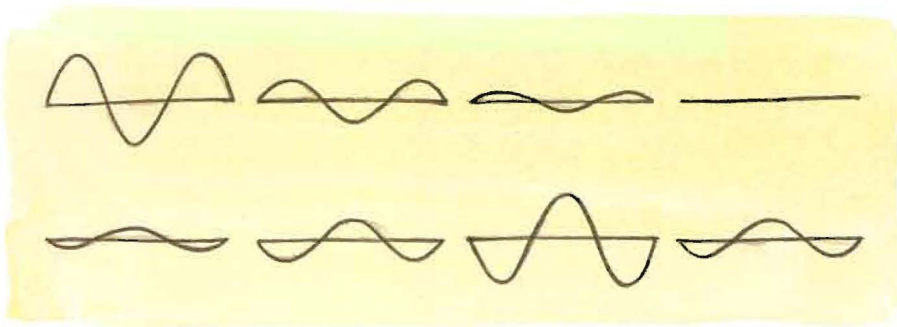f a vi-
ng. A
ed by
bers x
$0 \le x \le$
figura-

Figure 2. Snapshots capture a stationary waveform of a vibrating string at various times. The waveform retains the same basic shape as the string oscillates back and forth about the rest position; only the amplitude varies with time.
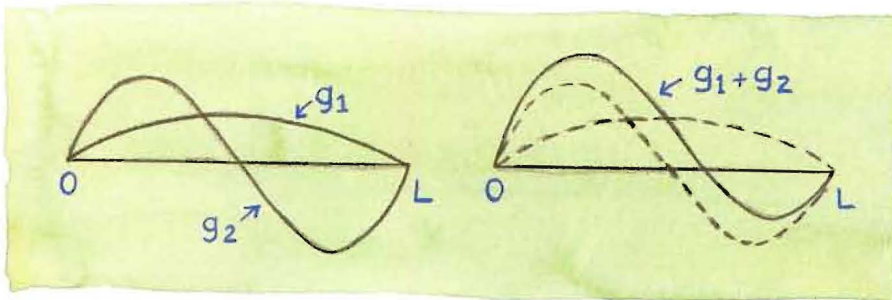

Figure 3. Because the wave equation of a vibrating string is linear, stationary solutions, such as those shown in Figure 2, can be added in a process called superposition to derive a new solution. Here, waves g_1 and g_2 are superposed to form $g_1 + g_2$.
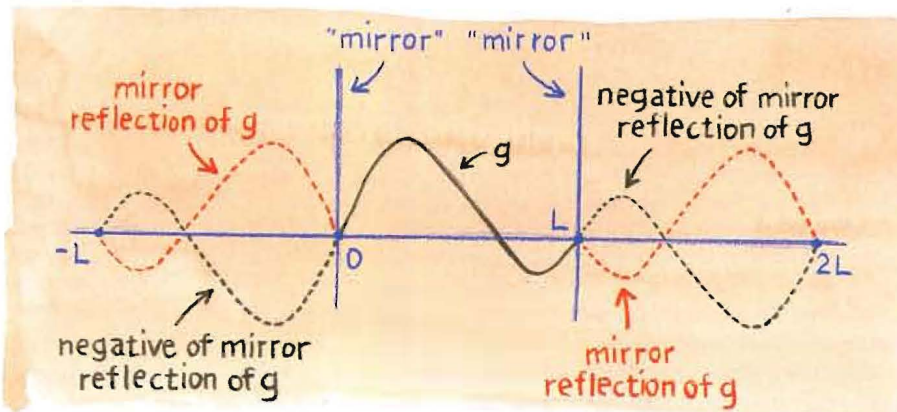

Figure 4. Stationary solutions to the wave equation also obey the reflection principle, which asserts that any solution g can be extended smoothly (that is, without kinks) past the "boundaries" (the string's endpoints, 0 and L) as the negative of the "mirror reflection" of g. The dotted black curves are smooth continuations of the solid black curve g.

where A and B are constants. From the boundary condition $g(0) = 0$, it follows (since $\cos(0) = 1$ and $\sin(0) = 0$) that $B = 0$, so $g(x) = A \sin(\sqrt{\lambda} x)$. The boundary condition $g(L) = 0$ means that the expression $\sqrt{\lambda} L$ must be an integral multiple of π (pi), since the sine is zero only at integral multiples of π. Thus $\sqrt{\lambda} L = n\pi$, so λ must be a number of the form

$$\lambda = \frac{n^2 \pi^2}{L^2}$$

The general solution of equation (2) is

$$h(t) = C \sin(\sqrt{\lambda} t) + D \cos(\sqrt{\lambda} t)$$

the constants C and D are determined by the initial configuration and velocity of the string. For a sinusoidal function of the form $C \sin(kt)$ or $D \cos(kt)$, the frequency is given by $k/2\pi$, and thus h represents a periodic oscillation of frequency $\sqrt{\lambda}/2\pi$. But the same λ as in (1) appears in (2), so it follows that h represents an oscillation of frequency $\sqrt{\lambda}/2\pi = n/2L$. These observations can be summarized as follows: The basic waveform g is given by a sinusoidal function, but its frequency $\sqrt{\lambda}/2\pi$ must be carefully chosen in order that the function be zero at 0 and L. Thus the

frequencies at which the string can vibrate are $1/2L$, $2/2L$, $3/2L$ and so on.

We have thus arrived at the following procedure for determining the possible vibration frequencies of the string. First seek stationary solutions; this will result in a pair of equations, one purely spatial and the other purely temporal. Then solve the spatial equation (1), and use the boundary condition imposed on the waveform g to determine the allowable values of the constant λ; these values in turn determine the frequencies of vibration that can appear in the temporal function h in (2).

This simple example exhibits some important features typical of more complex vibrating systems. First, the wave equation is *linear*. This means that the sum of two solutions is again a solution, and a constant multiple of a solution is again a solution. Thus multiples of stationary solutions can be added together to produce other solutions by *superposition*, and it turns out that the most general motion of the string is such a superposition of the stationary solutions we have just found. Figure 3 shows a solution obtained by adding two solutions. Second, the boundary conditions imposed on the spatial function g determine the *spectrum*—the list of allowable values of λ for which equation (1) has nonzero solutions. The list of possible frequencies consists of the numbers $\sqrt{\lambda}/2\pi$; once a frequency from this list is chosen, the temporal function h is just a periodic oscillation at that frequency. (In the introduction, the term "spectrum" was used to mean the list of frequencies; mathematicians use the term to refer instead to the set of admissible λ for which the spatial equation can be solved. This difference in terminology is harmless, as the list of allowable λ furnishes exactly the same information as the list of frequencies $\sqrt{\lambda}/2\pi$.) Finally, the solutions $g(x) = A \sin(\sqrt{\lambda} x)$ of (1) obey a *reflection principle*: Any function g that solves (1) can be locally extended past the boundary (in such a way that (1) still holds) as the negative of its "mirror reflection" through the boundary. That is, we can smoothly extend the definition of g to the left of the point 0 by setting $g(0 - x) = -g(0 + x)$ for small positive x; similarly, we can smoothly extend g beyond the right-hand boundary point L by $g(L + x) = -g(L - x)$ (*see Figure 4*).

The list of frequencies obtained above is the "overtone series" familiar to musicians; the lowest frequency
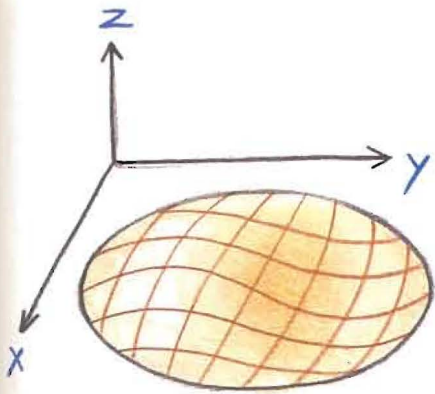
Figure 5. Graph of $z = f(x,y,t)$ for a fixed value of t is a "snapshot" at time t of a vibrating drumhead; the boundary remains fixed. Once again, stationary solutions can be sought for the wave equation, but the spatial equation will not admit to an explicit solution.

$1/2L$ is the *fundamental* (the pitch of the string), but there is also an infinite sequence of higher overtone frequencies. (Strictly speaking, the pitch is not $1/2L$ measured in vibrations per unit time, since the wave equation actually contains proportionality constants reflecting the string tension and mass density. We have fixed the tension and density once and for all and then have chosen units of measurement so that these constants are unity, enabling us to write the equation in the simpler form stated above; in these units, the pitch is indeed $1/2L$.) We have thus solved the one-dimensional analogue of Kac's question: The shape of a stretched string is captured completely by its length L, and we can recover L from the spectrum as half of the reciprocal of the lowest frequency. Thus one *can* hear the shape of a string.

Of course, the above analysis is an idealization. We are ignoring details such as how the motion is initiated. This oversight is not serious. Although such details affect what one hears, the sound produced is still a combination of the pure wave forms we have determined, added together in various amplitudes or "doses"; the initial conditions merely determine the "dosage" of each pure waveform. An actual string obeys a more complicated equation, since the elastic force is not quite linear. Such effects are extremely small, however. Moreover, much of the interest in spectral questions arises not from the "acoustical" problem we consider here but from the vibrations of atomic and molecular systems. The wave equation governing such quantum-mechanical systems *is* linear, so mathematical problems such as the one Kac stated in acoustical terms are of great interest.

The Vibrating Drumhead

Now consider Kac's question; replace the vibrating string with a vibrating drumhead. Let D be a region in the horizontal x-y plane representing the shape of the drumhead at rest. We consider an idealized drumhead, and we disregard the mechanism initiating the vibration. We represent the drumhead's motion by a function $z = f(x,y,t)$, interpreted as the vertical displacement at time t of the point (x,y) in D as the drumhead vibrates. Thus for fixed t, the graph of the function $z = f(x,y,t)$ defined on D is a "snapshot" at time t of the vibrating drum (*see Figure 5*). As before, the motion of the vibrating membrane is governed by the wave equation, which now takes the form

$$(3) \qquad \frac{\partial^2 f}{\partial t^2} = \Delta f$$

where Δf is a sort of rotation-invariant second derivative of f called the *Laplacian* of f, defined by

$$\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

Since the drumhead is attached to the frame along the perimeter of D, a solution f must satisfy the boundary condition $f(x,y,t) = 0$ for all times t and for all points (x,y) in the boundary of D. As in the case of the vibrating string, we seek stationary solutions of the form $f(x,y,t) = g(x,y)h(t)$, where g depends only on the spatial variables and h depends only on time. The boundary condition is expressed by requiring that g be zero everywhere on the boundary of D. Note that the Laplacian of f involves only the spatial variables x and y. Thus, as before, substitution of such a function f into (3) causes the wave equation to separate into a pair of equations, a spatial equation and a temporal equation:

$$(4) \qquad \Delta g = -\lambda g$$

$$(5) \qquad h''(t) = -\lambda h(t)$$

(compare with equations (1) and (2)). Once the spectrum (the list of the λ for which the spatial equation has a nonzero solution satisfying the boundary conditions) is determined, the solution of the temporal equation is the same as before, and the vibration frequencies are the square roots of the numbers λ occurring in the spectrum, divided by 2π, just as in the case of the vibrating string. Now, however, there is an obstacle: There are almost no regions D for which mathematicians can explicitly solve the spatial equation (4)! We could try to determine the frequencies numerically, but this would not give enough information, since there are infinitely many overtone frequencies, and any numerical calculation could only compute a finite number of
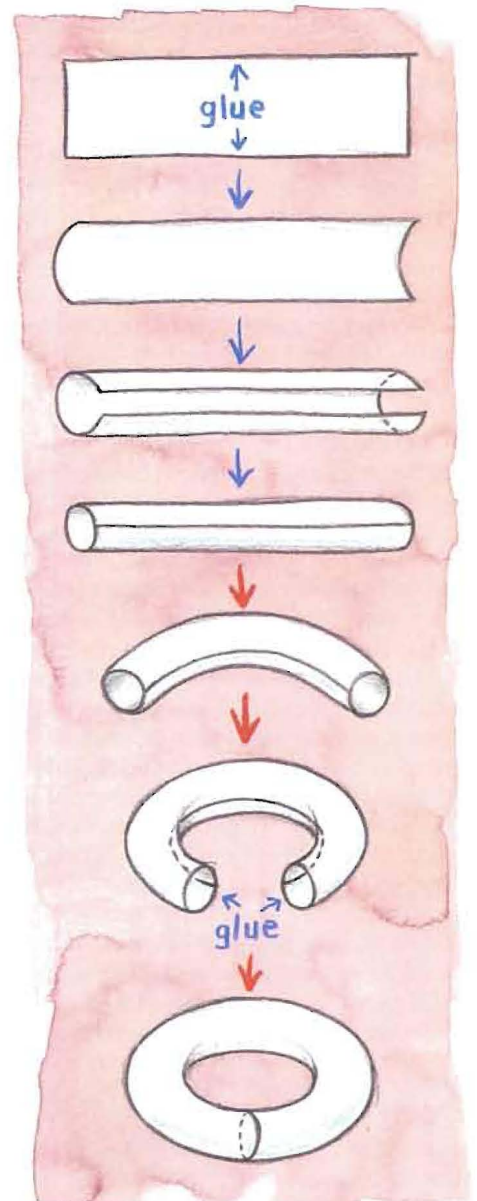

Figure 6. Curved surfaces called manifolds, of which a drum is a specific example, offer an approach to generalization of Kac's question. An example of a manifold, a torus, can be constructed from a parallelogram by gluing together opposite edges.
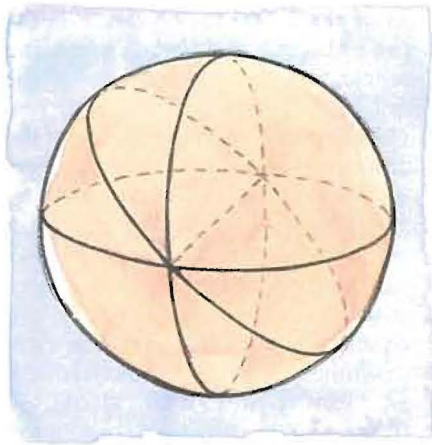
Figure 7. Geodesics are curves that do not deviate from the direction in which they are traveling. Geodesics of a sphere are great circles.

frequencies (and even then only to a finite degree of accuracy). There is no finite set of measurements (like the length of the vibrating string above) that completely captures the shape of a general planar region, so one cannot expect a finite sequence of frequencies to contain much information about the shape of the region.

Can One Hear the Shape of a Manifold?
When a problem appears intractable, mathematicians often generalize the problem; by relaxing their expectations, they may succeed in solving an easier but related problem, and the techniques used may shed light on the original problem. Thus mathematicians almost immediately asked: Can one hear the shape of a Riemannian manifold? A *manifold* is a curved surface small pieces of which look roughly like small pieces of a Euclidean space. For example, the surface of a doughnut is a manifold: A small enough piece of the surface looks like a (slightly warped) fragment of the Euclidean plane, even though the space as a whole is quite different from the plane. A *Riemannian manifold* is a manifold endowed with a way of measuring distances and angles. (The apparatus for such measurements is somewhat technical. The key idea, however, is that for each point p of the manifold M, there is a *tangent space* to M at p, which can be viewed as the collection of instantaneous velocity vectors of curves in M passing through the point p. This tangent space is an ordinary Euclidean space, so the angle between two curves in M that intersect at p can be defined as the usual Euclidean angle between their instantaneous velocity vectors at p. Similarly, the length of a curve in a manifold can

be defined by integrating the ordinary Euclidean lengths of its instantaneous velocity vectors.) Any Riemannian manifold has a wave equation, so it makes sense to ask: Can one hear the shape of a Riemannian manifold? Of course, if the answer is "yes," this is a harder problem than the original one, since a drumhead is a special case of a Riemannian manifold; it may be, however, that the answer is "no," in which case the more general problem offers wider scope for seeking counterexamples. Two manifolds that have the same spectrum, and hence the same set of vibration frequencies, are said to be *isospectral*.

In 1964, John Milnor of the State University of New York at Stony Brook answered the more general question negatively by exhibiting a pair of isospectral 16-dimensional manifolds. Milnor's examples are constructed by gluing together opposite faces of a cleverly chosen 16-dimensional "parallelogram" to produce *flat tori*. A two-dimensional analogue, a flat 2-torus, can be constructed by taking a parallelogram in the plane, gluing the top and bottom edges together, and then gluing the left and right edges together. Topologically, the resulting surface is the same as the surface of a doughnut. Indeed, the doughnut surface may be constructed by first gluing together the top and bottom edges of a rectangle to form a cylinder, and then bending the ends of the cylinder (formerly the left and right edges of the rectangle) around and gluing them together (*see Figure 6*). Of course, if this procedure is carried out in three-dimensional space, as in Figure 6, the local geometry (lengths and angles) must be deformed a bit during the second step (the red arrows in Figure 5), but topology overlooks such distortions.

A *flat* torus would result if we performed the second gluing without altering the geometry of the surface; this feat cannot be performed in three-dimensional space, but in four-dimensional space there is enough room to do so. The term "flat" refers to the fact that now the local *geometry* (not merely the local topology) is that of the Euclidean plane—locally, lengths and angles measured on the flat torus are the same as they would have been in the original planar parallelogram before the gluing was done. Perhaps the easiest way to visualize a flat torus is to imagine a child's video game in which spaceships flying off the right side of the screen reappear at the same height on the left edge of
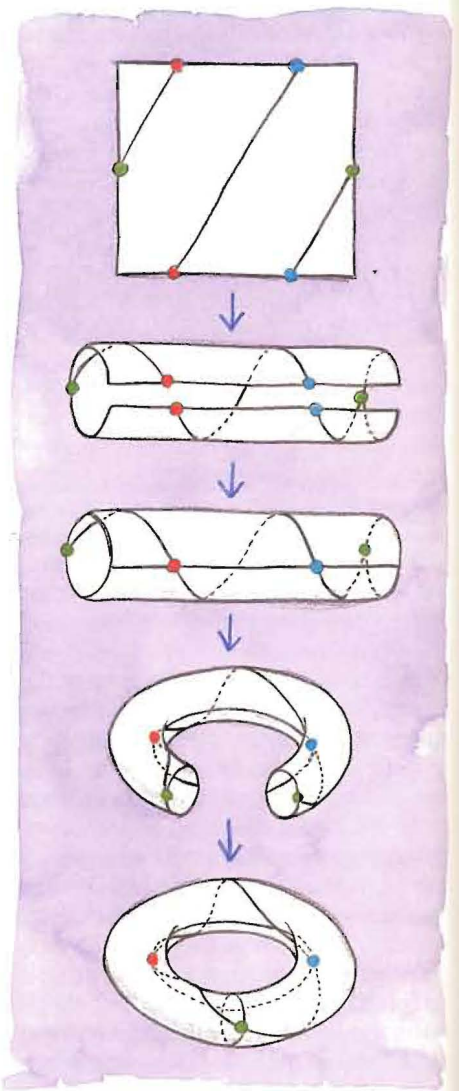

Figure 8. Closed geodesic on a torus "wraps around" and closes up. Note that the torus pictured is not "flat," as the second gluing distorts local geometry; the picture only aids in visualizing how the closed geodesic "sits" in the torus. A closed geodesic on a flat torus is best visualized with a "round trip" of a video game space ship, as in the top image. The list of vibration frequencies of any vibrating manifold is closely related to the list of lengths of its closed geodesics. By constructing two different 16-dimensional tori with identical lists of closed geodesics, John Milnor was able to exhibit isospectral (geometrically different but vibrationally equivalent) manifolds.

the screen, and spaceships flying off the top of the screen reappear at the bottom of the screen. Effectively, the top of the screen has been glued to the bottom and the left edge to the right edge, but no geometric distortion has been introduced, since locally, lengths and angles on the screen are just as they would be in the Euclidean plane.

A *geodesic* on a Riemannian manifold is the natural analogue of a straight line

in the Euclidean plane: a curve that does not deviate from the direction in which it is traveling. Thus someone living in the manifold and traveling along a geodesic at a constant speed would feel no acceleration, since the curve does not "turn." For example, the geodesics on the surface of a sphere are the great circles, the curves in which a plane through the center of the sphere meets the surface (*see Figure 7*). Figure 8 depicts a closed geodesic on a flat 2-torus that arises from a straight line in the plane. Unlike a straight line in Euclidean space, however, a geodesic can "wrap around" and close up, as shown. The list of vibration frequencies of a vibrating manifold is closely related to the list of lengths of *closed geodesics* on the manifold; this is plausible, as one might expect waves to propagate along geodesics. Milnor's two 16-tori were cleverly chosen so that the list of lengths of closed geodesics was the same for each. It can be shown that two flat tori with the same geodesic lengths must be isospectral, so it follows that Milnor's tori both vibrate at exactly the same frequencies.

An Idea from Group Theory
In 1984, Toshikazu Sunada of Tôhoku University realized that an idea from group theory could be brought to bear on the problem of constructing isospectral manifolds. A *group* is a set of objects that can be combined in pairs by an operation (usually written as multiplication or addition) satisfying certain properties. For example, the elements might be integers, and the operation might be addition. Or the elements might be rotations of the Euclidean plane about the origin, and the operation might be composition—performing one rotation and then the other to get a new rotation. A group is required to have an *identity element*, an element that can be combined with any other element without changing the other element. In the case of rotations, the identity element is a rotation through an angle of zero—that is, a rotation that "does nothing" and leaves every point unmoved. Finally, every element x in a group is required to have an *inverse*, an element y such that $xy = yx = e$, where e is the identity element. In the case of rotations, the inverse of a counterclockwise rotation through an angle of d degrees is a clockwise rotation of d degrees, since the second rotation exactly "undoes" the first.

Groups are the natural mathematical incarnation of the idea of symmetry. This fruitful idea appears nearly everywhere in mathematics. Groups have been used to understand problems as diverse as the nature of elementary particles and the solution of algebraic equations. An object exhibits *symmetry* if it can be moved or transformed in a way that leaves it unchanged. As an example, consider the square depicted in Figure 9. We shall consider two geometric operations s and t: s rotates the square counterclockwise by 90 degrees (so it sends vertex 1 to vertex 2, 2 to 3, 3 to 4, and 4 to 1), whereas t reflects, or "flips," the square about the perpendicular bisector of its base (thus t interchanges vertices 1 and 2 as well as vertices 3 and 4). Both of these are symmetries of the square: They leave the shape unchanged, altering only the vertex labels. Other symmetries of the square can be obtained by combining s and t by *composition*, performing one operation, then another in succession. For example, the effect of st (this is read from right to left—first do t, then do s) is to interchange vertices 1 and 3, leaving vertices 3 and 4 fixed. Thus st is a reflection about the diagonal joining vertices 2 and 4 (*see Figure 10*). A little experimentation reveals that the group of symmetries of the square consists of the operations e (the operation that does nothing), s, ss, sss, t, st, sst and $ssst$; the effects of these transformations are shown in Figure 11. Thus the symmetry group of the square has eight elements.

A *permutation* of a set X of objects is just a rearrangement of those objects. One useful example of a group is the set of all such permutations; the operation is *composition*—first perform one rearrangement, then the other. For example, if the set X consists of the numbers 1, 2 and 3, there is a permutation α (alpha) usually denoted

$$\begin{vmatrix} 1 & 2 & 3 \\ \downarrow & \downarrow & \downarrow \\ 2 & 1 & 3 \end{vmatrix}$$

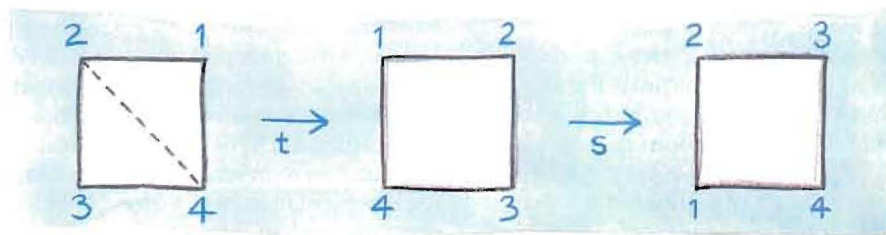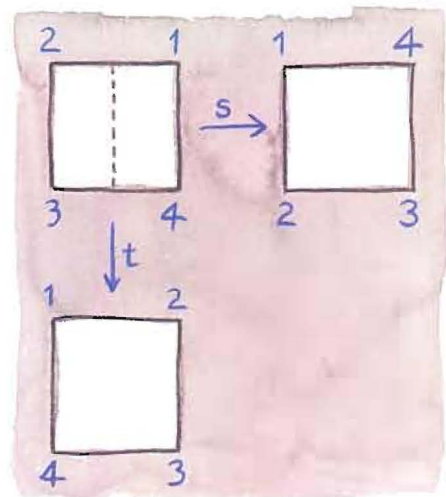that interchanges objects 1 and 2 but

Figure 9. Group theory and its concept of symmetry can be applied to the problem of constructing isospectral manifolds. Here a square shows two of its symmetries: s, *rotation* by 90 degrees counterclockwise; and t, reflection about the dotted line.

leaves 3 unmoved. There is another permutation β (beta) denoted

$$\begin{vmatrix} 1 & 2 & 3 \\ \downarrow & \downarrow & \downarrow \\ 1 & 3 & 2 \end{vmatrix}$$

that simply interchanges 2 and 3 but leaves 1 unmoved. The composite permutation $\alpha\beta$ (first do β, then do α) is

$$\begin{vmatrix} 1 & 2 & 3 \\ \downarrow & \downarrow & \downarrow \\ 2 & 3 & 1 \end{vmatrix}$$

The group G consisting of all permutations of X has six elements: the three listed above as well as

$$\begin{vmatrix} 1 & 2 & 3 \\ \downarrow & \downarrow & \downarrow \\ 3 & 2 & 1 \end{vmatrix} \begin{vmatrix} 1 & 2 & 3 \\ \downarrow & \downarrow & \downarrow \\ 3 & 1 & 2 \end{vmatrix} \begin{vmatrix} 1 & 2 & 3 \\ \downarrow & \downarrow & \downarrow \\ 1 & 2 & 3 \end{vmatrix}$$

(The last permutation serves as the identity element of the group.)

Sunada's technique for constructing isospectral manifolds relies on the no-

Figure 10. *Composition*—the combination of operations—produces other symmetries of the square. In this example, composition st, t is performed first, followed by s. The net effect is to flip the square about the 2–4 diagonal.
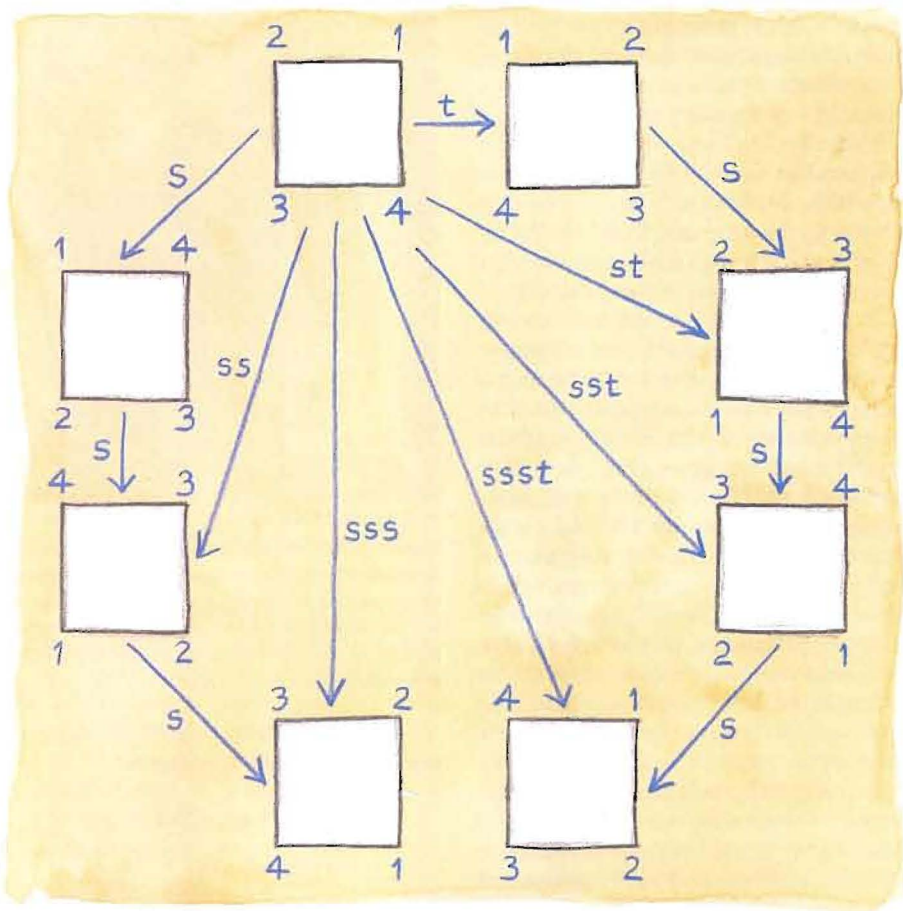
Figure 11. Eight elements—including the identity element, the operation that does nothing—are included in the group of symmetries of a square. The figure depicts the effects of the non-identity elements.

tions of *permutation representations* and *linear representations* of groups. Given a group G, a *permutation representation* of G on a set X is a rule that assigns to each element g of G a permutation $\bar{g}$ of the set X. The permutations must be assigned in such a way that composition of the permutations assigned to any two group elements g and h in G has the same effect as the permutation assigned to the group product of g and h, that is, $\bar{g}\,\bar{h} = \overline{gh}$. A *linear representation* of G of degree n is a rule that associates to each g in G an $n \times n$ matrix $\bar{\bar{g}}$ in such a way that the matrix assigned to a product of any two elements g and h in G is just the product of the matrices assigned to the group product of g and h: $\bar{\bar{g}}\,\bar{\bar{h}} = \overline{\overline{gh}}$. Since the matrices $\bar{\bar{g}}$ and $\bar{\bar{h}}$ multiply in essentially the same way the original group elements g and h did, it might seem that little has been gained. Matrix theory, however, is a very well-developed branch of mathematics that affords some powerful tools; thus viewing group elements as matrices and using the rich theory of matrices may permit one to notice things that might

have escaped detection otherwise. This philosophy of enrichment of structure is a recurrent theme in mathematics.

Given a permutation representation of a group G on a set X of n elements, it is easy to construct a linear representation of G of degree n. For example, if X has three elements, i, j and k, we label the unit vectors in the directions of the axes of three-dimensional space by i=(1,0,0), j=(0,1,0) and k=(0,0,1). An element g of G can be realized as a permutation $\bar{g}$ of the set X, hence as a rearrangement of the coordinate axes of 3-space. This rearrangement of coordinates determines a *linear transformation* of 3-space (a way of mapping 3-space to itself linearly), hence a 3×3 matrix $\bar{\bar{g}}$. Indeed, a matrix is perhaps best viewed as a way of encoding such a linear transformation relative to a fixed choice of coordinates. This transition from a permutation representation of G to a linear representation of G seems rather innocuous, but an interesting phenomenon can occur: It can happen that two different permutation representations of G can give rise to linear representations

that are essentially the same. This phenomenon—inequivalent permutation representations giving rise to equivalent linear representations—is essential to Sunada's technique. Although a little ingenuity is required to find inequivalent permutation representations that give rise to equivalent linear representations, number theorists had already constructed examples while studying a very different but analogous problem. Although we cannot indicate here why Sunada's method produces isospectral manifolds, we will illustrate by means of an example.

One Can't Hear the Shape of a Drum
Although Sunada's technique had been available for nearly a decade, it was believed that it shed no light directly on Kac's original question, since the spaces manufactured by the method could not be regions in the Euclidean plane. In 1989, however, Pierre Bérard of the Institut Fourier in Grenoble, France, discovered a new proof of Sunada's theorem that permitted wider application of the method. In 1990, the authors and Scott Wolpert used Bérard's discovery to construct a pair of isospectral planar regions that are not geometrically congruent, thereby answering Kac's question: One *cannot* hear the shape of a drum.

The group G we use is one used by Robert Brooks of the University of Southern California and by Peter Buser of the École Polytechnique Fédérale in Lausanne, Switzerland, to construct isospectral surfaces. The group G contains three special elements, α, β and γ, and all the elements of G can be obtained by taking products of α, β and γ. Linear algebra furnishes a natural way in which G permutes the elements of the set $X = \{1, 2, \dots, 7\}$, which we shall describe pictorially; the permutation associated to an element g of G will be written as $\bar{g}$. The way elements of G permute the numbers 1, 2,…,7 can be represented by a *Cayley graph*, a set of points labeled by the numbers 1, 2,…,7, joined together by various edges marked by α, β and γ. The Cayley graph faithfully encodes all the information about the permutations $\bar{\alpha}$, $\bar{\beta}$ and $\bar{\gamma}$. The Cayley graph that encodes our permutation representation is depicted in Figure 12. The edges are color coded for convenience: red = α, green = β and blue = γ. As an example of how the Cayley graph is used, the fact that the points labeled 2 and 6 are joined by an edge marked α expresses the fact
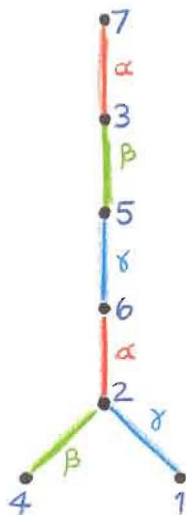
Figure 12. Cayley graphs encode the way a group is applied to a set of numbers through *permutation*. In this example, a group consisting of products of three elements, α, β and γ, permutes the set X (numbers 1 through 7). Red indicates an α permutation, which switches 3 and 7 and 2 and 6. Green is a β permutation, which switches 3 and 5 and 2 and 4. Blue denotes a γ permutation, switching 5 and 6 and 1 and 2. Some numbers remain unmoved by a given permutation. For example, because there is no blue edge emanating from either 3 or 4, the γ permutation leaves these points unmoved.

that the permutation $\bar{\alpha}$ interchanges 2 and 6. The fact that there is no γ edge emanating from the point in the graph labeled 3, on the other hand, expresses that the permutation $\bar{\gamma}$ leaves 3 unmoved. There is another permutation representation of G whose Cayley graph is depicted in Figure 13; it is similar to the graph arising from our first permutation representation, but the edge markings are different.

As described above, one can associate to each permutation representation a linear representation, in this case by 7×7 matrices. The two permutation representations are not the same, since their Cayley graphs differ. It turns out, however, that the two linear representations associated to our permutation representations are equivalent! This means that we can use the permutation representations to construct isospectral regions in the plane by Sunada's method, as follows.

We begin with the "model" triangle T in Figure 14; its edges are labeled by α, β and γ, and are color-coded like the Cayley graphs. We will construct a region D_1 by gluing together along their edges seven copies of the triangle T (labeled by elements 1, 2,...,7 of X) according to the

pattern of the Cayley graph of Figure 12. We begin with the triangle labeled 7, a copy of the model triangle corresponding to the vertex 7 of the Cayley graph in Figure 12. In this graph, 7 is joined to 3 by an edge labeled α, so we reflect the triangle 7 through its α edge and label the resulting "flipped-over" triangle 3. Figure 14 shows the first few steps of this construction, whereas Figure 15 depicts the polygonal region D_1 that results when we have worked our way through the entire Figure 12 Cayley graph in this fashion. Similarly the Cayley graph in Figure 13 gives rise to a plane region D_2, also shown in Figure 15.

Transplantation of Waveforms

We wish to show that the two plane regions in Figure 15 vibrate at precisely the same frequencies. Since they are different shapes, this will mean that one cannot hear the shape of a drum. Bérard's proof of Sunada's theorem furnishes an explicit recipe for "transplanting" a waveform on the drum D_1 to a waveform of the same frequency on D_2. Although the actual details of the recipe are too technical to describe here, once the recipe has been followed, it is easy to verify that what it produces is actually a valid waveform of the same frequency on D_2, as we will show below. A waveform on D_2 can also be transplanted to D_1, so the drums vibrate at precisely the same frequencies.

To see how the transplantation works, we record some properties of stationary waveforms on a drumhead that are analogous to properties of the stationary waveforms we saw above on the vibrating string. First, a sum of solutions or a constant multiple of a solution is again a solution. Thus we are free to combine waveforms of a given frequency by superposition. Second, pos-
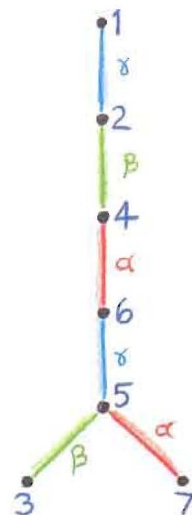

Figure 13. Cayley graph of a second permutation representation of G on the set X shows different effects of the α, β and γ permutations.

sible waveforms (solutions of the spatial equation (4)) obey a reflection principle: A waveform can be locally extended past a boundary edge as the negative of its mirror reflection through the boundary edge. Thus, for example, in Figure 16, a waveform (denoted g) on the left triangle L can be continued across the boundary edge as shown to yield an admissible waveform on the region made up of the two congruent triangles L and R by specifying the value of the function at a point of the right-hand triangle R to be the negative of its value at the corresponding "mirror image" point of L. Thus the function on the right-hand triangle is just $-g$. In order for this notation to make sense, we imagine the region folded along the common edge of both triangles, so that the right-hand triangle is flipped over and folded back on top of the left triangle; in other words, we view the left tri-
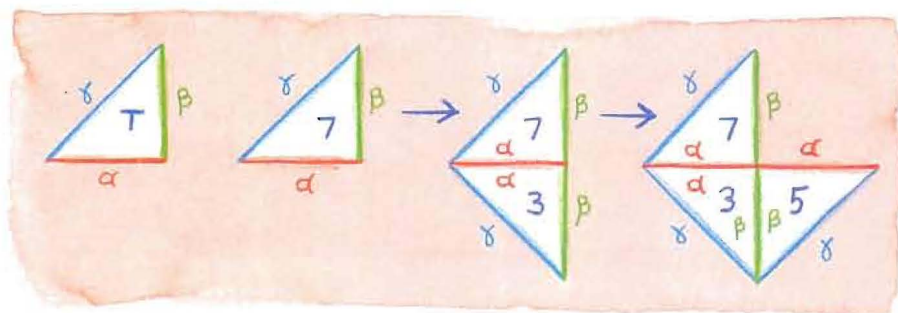

Figure 14. Using the Cayley graph as a guide, it is possible to construct isospectral regions based on the model triangle T. Its edges are labeled α, β and γ, and copies are adjoined to it by gluing new triangles along the appropriate edges. Starting from the top of the Cayley graph in Figure 12, the first triangle is labeled 7. Along its α edge, a "mirror image" triangle labeled 3 is glued. Continuing, a triangle 5 is glued along the β edge of 5. And so on.
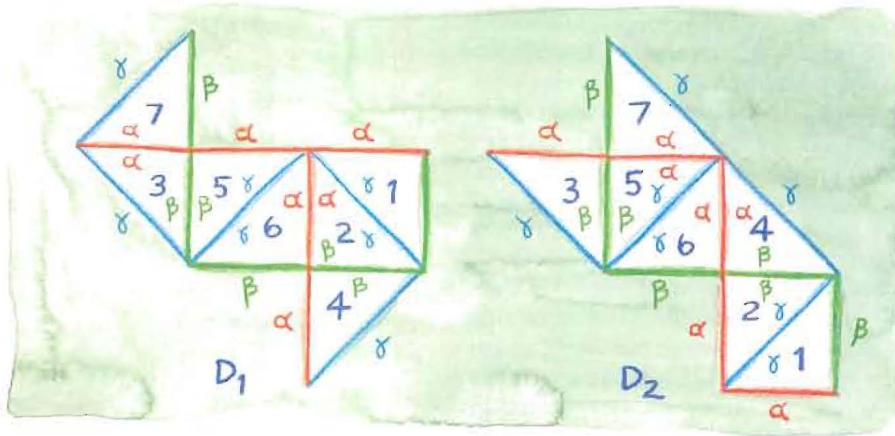
Figure 15. When the Cayley graphs in Figures 12 and 13 are used as patterns for gluing together copies of the model triangle T, two different planar regions, D_1 and D_2, result. We wish to show that these two drumheads of different shape have the same vibration frequencies.

angle L as a copy of the model triangle T, and we identify the right-hand triangle R with L by reflection through the common edge. By successively flipping triangles in this way, we can identify any triangle in D_1 or D_2 with the model triangle T. We can think of the process of relating any triangle in D_1 or D_2 to any other triangle as "origami": We are just folding all the triangles back onto a single model triangle.

Consider now a waveform φ (phi) of a given frequency on the region D_1 of Figure 15. Such a solution of the spatial equation can be viewed as a snapshot of D_1 while vibrating. Consider the portion of the graph of the function φ which lies above just the triangle labeled A in Figure 17; we will denote this piece of the graph of φ by A. Similarly, denote by B the part of the function φ defined on triangle B in Figure 17, and so forth. Each of A, B, C,...,G is a function on a single triangle (indeed, by origami each can be

viewed as a function defined on the model triangle), and the graph of φ is just the graphs of the functions A, B, C,...,G glued together along the interfaces between the triangles. In particular, since the waveform is a smooth function, the values of the functions A (on triangle A) and B (on triangle B) must coincide along their common (red) α edges. Also, notice that since the boundary of D_1 remains fixed during the vibration, the function A is zero on the (blue) γ edge of triangle A, for example.

We now seek to transplant φ to a waveform of the same frequency on D_2. We have seen that we can add and subtract waveforms to produce other waveforms. Consider the function ψ (psi) on the region D_2 described in Figure 17. This function ψ is specified by indicating a function on each of the triangles forming D_2; thus, for instance, on the topmost triangle in D_2 in Figure 17 (the one labeled 7 in D_2 in Figure 15), ψ is the func-
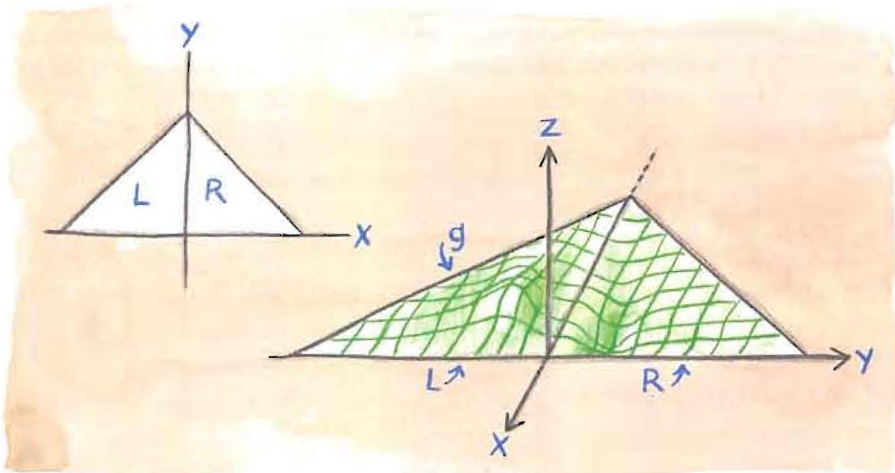
tion $B - C + D$, a superposition of the functions B, $-C$ and D. On each of the individual triangles forming D_2, ψ is certainly a valid waveform (ignoring boundary conditions), since it is a sum of waveforms on the model triangle. To see that ψ is a valid waveform on the whole of D_2 satisfying the boundary conditions, we must check two assertions: First, the seven "pieces" of ψ must fit together smoothly at the interfaces between triangles, and second, the function ψ must be zero on the boundary of D_2.

Both of these assertions can be verified by inspecting the relations of the seven functions arising from the structure of D_1, as depicted in Figure 17. For example, let us check that the function $B - C + D$ on the topmost triangle of D_2 in Figure 17 (triangle 7 of D_2 in Figure 15) and the function $A + C + E$ on its neighboring triangle (triangle 5 of D_2 in Figure 15) fit together smoothly along their common interface, the red (α) edge separating the two triangles. Reference to D_1 in Figure 17 shows that triangles A and B share a common red edge, so the functions A and B must be identical on the red edge; similarly, triangles D and E of D_1 in Figure 17 share a common red edge, so the functions D and E must coincide on the red edge. Thus $A + E$ and $B + D$ will agree on the red interface between triangles 7 and 5 of D_2 in Figure 15. But note that the function C is zero on the red edge of the model triangle. Indeed, in D_1 in Figure 17, the red edge of triangle C is a boundary edge, so any waveform of D_1 must be zero on the entire edge, since the boundary stays fixed throughout the vibration. This means that if we imagine the waveform C on triangle 5 of D_2 in Figure 15, in order to continue it smoothly across the red edge to a waveform on triangle 7, we must put the function $-C$ on triangle 7 in accordance with the reflection principle (refer again to Figure 16). Since $B + D$ (on triangle 7 of D_2) agrees with $A + E$ (on triangle 5 of D_2) on their common red edge, and since $-C$ (on triangle 7 of D_2) agrees with C (on triangle 5 of D_2) on the common red edge (where both are zero), it follows that $B - C + D$ (on triangle 7 of D_2) and $A + C + E$ (on triangle 5 of D_2) join together smoothly across the red α edge shared by triangles 7 and 5. It is easy now to check in the same way that the seven pieces of ψ fit together smoothly across all interfaces between triangles of D_2, so the first assertion is verified.

Figure 16. *Transplantation* is used to demonstrate that planar regions D_1 and D_2 have the same vibration frequencies. Transplantation depends on the reflection principle: A waveform on triangle L can be extended to triangle R as the negative of the mirror reflection of the waveform on L.
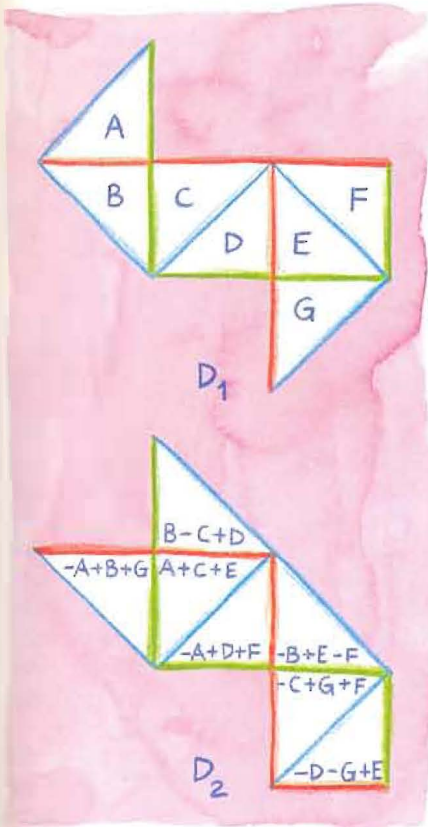
Figure 17. Bérard's proof of Sunada's Theorem provides an explicit recipe for transplanting waveforms on planar region D_1 onto planar region D_2. We can determine that the transplantation provides a valid waveform by examining the interfaces of the seven triangles in D_2 to see if they fit together smoothly and by checking that the waveform is zero on the boundary of D_2.

The second assertion can be checked in a similar fashion. For example, let us verify that ψ is zero on the blue (γ) edge of triangle 7 of D_2 in Figure 15. Reference to Figure 17 shows that triangles C and D of the region D_1 share a common blue edge, so the functions C and D must coincide on that edge; thus $-C + D$ is zero on the blue edge of the model triangle and hence also on the blue edge of triangle 7 of D_2 in Figure 15. Similarly, the blue edge of triangle B in D_1 is a boundary edge; since φ satisfies the boundary condition, it follows that the function B is zero on the blue edge of the model triangle and hence also on the blue edge of triangle 7 of D_2 in Figure 15. We have shown that both $-C + D$ and B are zero on the blue edge of triangle 7 of D_2, so their sum $B - C + D$ is zero on this edge. But $B - C + D$ is by definition the value of ψ on triangle 7 of D_2, so ψ is zero on the blue edge of triangle 7 of D_2, as claimed. In the same way, it is easy to check that ψ is zero on all boundary edges of D_2, so ψ

is a valid waveform for the vibrating drumhead D_2. Thus a waveform φ on D_1 has been explicitly transplanted to a waveform ψ of the same frequency on D_2. Waveforms on D_2 can be transplanted to D_1 similarly, so both drumheads vibrate at the same frequencies.

What *Can* One Hear?

The above example appears to be an instance of failure of the theory rather than success, and hence would seem to undermine our contention that the spectrum of a vibrating object contains geometric information. In fact, it is expected (although not proved) that generically, one *can* hear the shape of a drum, so the phenomenon we have exhibited above is in some sense atypical. We conclude by briefly describing some positive results along these lines. One of the motivating physical questions for the early study of spectral theory was the theory of black-body radiation. It led to the formulation of a problem whose solution constitutes one of the earliest success stories in the subject: Given the spectrum of a vibrating membrane, can one infer the area of the membrane? When the great German mathematician David Hilbert learned of this problem, he predicted that it would not be solved during his lifetime. Barely two years after this pessimistic prediction, however, Hilbert's student Hermann Weyl showed that one can hear the area of a drumhead!
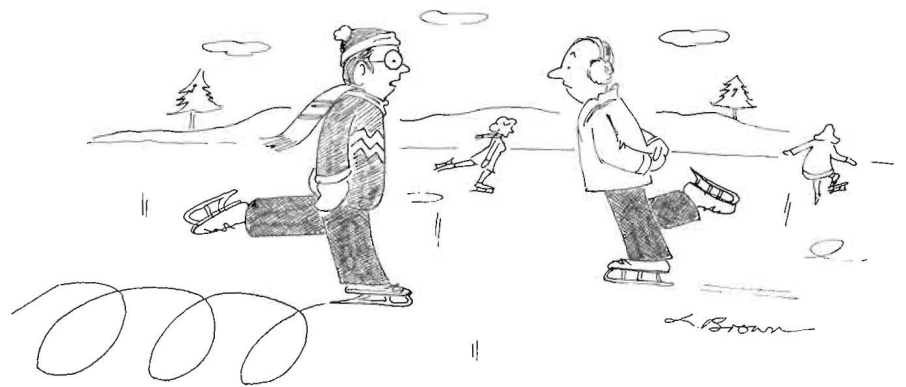
One might call a geometric shape spectrally *solitary* if it has no isospectral companions; a solitary object can thus be fully reconstructed (at least in principle) from its vibration frequencies, which serve as an unambiguous "fingerprint" of the object. Our collaborator, Scott Wolpert, has proved a

beautiful result asserting that in a precise sense, "most" geometric surfaces are spectrally solitary.

Vistas

The subject of spectral geometry, despite some stunning successes, is still poorly understood. There are many natural questions to which the answers are not known. One of the most fascinating questions concerns what information the spectrum contains about the dynamics of a system. Some nonlinear classical mechanical systems exhibit *chaotic* dynamics, in that if one makes a seemingly insignificant modification to the initial conditions, the system's long-time behavior is affected drastically. It is of interest to understand if and how chaos shows up in the spectrum of the associated quantum-mechanical system. Some deep work of Peter Sarnak of Princeton University exploits connections with number theory to show that it is likely that such information is encoded by how regularly the spectral lines are spaced. Another question concerns the "size" of isospectral sets: If a geometric object is not spectrally solitary, how many isospectral companions can it have? Work of Brad Osgood and Ralph Phillips of Stanford University and Peter Sarnak suggests that, in a suitable sense, the set of isospectral companions of a space cannot be very large.

Hilbert's prediction discussed above might lead one to be wary of speculation, but it seems likely that our ignorance of this subject will persist for a long time, despite much effort and some beautiful successes. If Sir Arthur Schuster could return and survey the mathematics of the science he christened, he would probably be pleased—but he would not be satisfied.

"It's a figure eight in binary numbers."