Billiards, Surfaces, and Geometry: A Problem-Centered Approach

Diana Davis

20 Main Street, Exeter NH 03833 USA

2020 *Mathematics Subject Classification.* 37E35: Flows on surfaces

Key words and phrases. billiards, inner billiards, outer billiards, tiling billiards, flat surfaces, interval exchange transformations, problem-based learning

For Margo Angelopoulos, my *eierlegende Wollmilchsau*:
everything I could ever want in a person.

And for Samuel Lelièvre, the golden L of collaborators.

Contents

They did the math

Preface

The study of mathematical billiards is a beautiful field connecting many simple objects in surprising ways: squares, rational numbers, paper folding and unfolding, and symmetries, to name just a few. Its ideas are accessible to students at any level. I've written this book so that more instructors can create, and more students can take, a one-term introductory billiards course. I'm glad you're here.

I have successfully used this book for a course for undergraduates, and a course for advanced high school students. This book has also proved to be a great introduction for a graduate student seeking a working introduction to billiards. The main prerequisite knowledge for this course is high school geometry. Representing transformations of the plane by 2×2 matrices and a familiarity with writing proofs are two other skills that will be helpful. While proof writing is an art that takes some time to master, 2×2 matrix transformations require just a short primer, and you will find this in Appendix A.

I designed this book for a one-semester or one-trimester course. The first four chapters comprise the core of the book, and the fifth chapter contains special topics. There are hands-on activities at the end of the first three chapters, but only the content of the first chapter is required in order to do them, and they do not contain necessary content for later work.

How to learn from this text. Work on the problems! Have you heard of "experiential education," where you go out into the field and learn things by doing them? You can think of this as an "experiential education billiards book." You will find yourself actively doing things – drawing pictures, making color-coded diagrams, cutting things out and folding them up – and these experiences will comprise your billiards education. When mathematicians explicitly compute an example, they call it "getting your hands dirty." In the course of having the experiences outlined by the problems, you will get your hands dirty many times over, building your understanding with each experience.

Interspersed between the problems are paragraphs of explanation, telling you about some of the people who have traveled this journey before you. The people profiled in each section of this book have spent many hours drawing polygons on bits of paper, labeling the edges, drawing line segments, counting and calculating – and through this kind of experience, they have built up humanity's understanding of billiards and flat surfaces, bit by bit. Throughout the book, I've chosen the convention of referring to everyone by their first name, to make them seem as approachable in print as they are in real life.

I hope that you have people on this journey with you: your classmates, your instructor – people you can talk with about your ideas as you work on the problems. Other people are a tremendous resource, as they have ideas and insights that are different from yours, come up with different examples, and ask you questions you've never considered before.

So let's get to it! Approach each problem as an exploration. Get out your colored pens and your ruler, and try things. Draw pictures of several examples, as different from each other as possible. Think about earlier problems and what they might reveal about the one you are working on. *Believe* that you have all the tools necessary to solve the problem, and try lots of things as you claw towards understanding.

The problems in this text. This style of curriculum is integrated: rather than a single problem set with many problems about continued fractions (for example), problems on each topic are sprinkled across many days, gradually increasing in sophistication. This way, students have a chance to discuss and understand each problem on a given topic before moving on to a harder one. For this reason, it's important to leave each class understanding the previous night's homework problems, as the next set of problems usually builds directly on them.

Several of the problems are marked with the word "Challenge," to indicate that their level of difficulty is higher than the rest. These require more time and energy, and more original ideas, than the rest of the problems. Most of the problems in Chapter 5 are also challenging.

If you are teaching this course, work to bring out students' ideas about the problems – work to help them absorb as much as they can each day. If you are working through this book, give yourself time to explore each question.

Materials needed. It is essential to have the following tools readily available to you while you are working:

- a set of colored pens or pencils
- graph paper
- scissors
- tape
- ruler
- tissue paper, string, bagel, thin cord, board, hammer and nails.

How to teach from this text. The way I've used this book for a 50-minute class is as follows:

- First day of class: do §1 problems in class. For homework, students do §2 problems.
- Second day of class: students spend class time discussing their solutions to §2 problems. For homework, students do §3 problems.
- Third day of class: students spend class time discussing their solutions to §3 problems. For homework...

Depending on your students and your class duration, you may need to assign less than a whole section, or more than one section, on each assignment, which is totally fine. I do recommend that you think carefully before assigning two problems with very similar content from two different sections at the same time, as the problems build off of each other, so it may be helpful to discuss the earlier problem with other people before embarking on the next one.

The problems teach all of the material; no lectures are necessary. I wrote the problems to be hard enough that most students will not be able to solve all of the problems on their own, so that students have something to discuss with each other when they get to class. If you read a problem in the book and think, "Hmmm, that seems challenging! I'm not sure how to do that," you're doing it right.

For more guidance on how to run such a class, see Appendix B.

My goal for this book. There are so many beautiful things in the study of billiards and flat surfaces, and I've put all of my favorites in this book: the continued fraction algorithm, the folded-up flat torus, the Arnoux-Yoccoz IET... and the *people* in the billiards and flat surfaces community. I hope that by working on the problems, you'll feel like you really understand something about billiards, and I hope that by reading about the people, you'll feel like you are part of the billiards community.

Diana Davis

Acknowledgements

Thank you to the students who have taken this course over the years, whose insightful questions and ideas improved this book a great deal. Thank you to my department chairs, Frank Morgan at Williams College and Gwyn Coogan at Phillips Exeter Academy, who made it possible for me to teach this course, once and twice respectively. Additional thanks to Gwyn for teaching me how to construct a problem-based curriculum in the first place, and to Frank for giving talks that included pictures of the people who proved the theorems, as I do here.

Thank you to my colleague Samuel Lelièvre, with whom I have spent hundreds of happy hours doing research on billiards and flat surfaces. In March 2023, Samuel read through every problem in this book and provided innumerable suggestions for improvements on every page, notably outlining a complete rewrite of the section on the modular group, giving me the diplotorus layout, and recreating the figures for the Rauzy gasket and the zippered rectangle surface.

Thank you to my Ph.D. advisor, Rich Schwartz, who suggested back in 2009 that I read John Smillie and Corinna Ulcigrai's paper about linear trajectories on the regular octagon, and guided me to learn this beautiful subject and join this wonderful community.

Thank you to my wife, Margo Angelopoulos, who has attended dozens of my talks over the years, wears laser-cut pentagon billiard trajectories on her ears every day, and 3D-printed a pentagon billiard soap dish for each sink in our house. I intended to publish this book before giving birth to our first child, but alas, this did not happen, and Margo put in lots of extra baby time in the seventh and eighth months of our child's life so that I could bring this project to the finish line.

Thank you to my editor at the AMS, Loretta Bartolini, who shepherded me and this project every step of the way. I wrote the first version of this book in 2016, and first talked with Loretta about it in Toronto in 2018. Since then, she has been consistently supportive and very helpful in every aspect all the way through to 2025. Thank you also to the four anonymous referees, who made numerous excellent suggestions that greatly improved this book.

Thank you to Hannah Kushnick for finding many errors, whose correction improved the reading experience of this text.

I acknowledge that some of the problems in this book are similar to examples discussed in my chapter of *Lines in positive genus: an introduction to flat surfaces* [**14**]. I would also like to point out that Problems 4, 15, 16, 23, 39, 61, 78, and 118 take direct inspiration from Sergei Tabachnikov's book *Geometry and Billiards* [**56**].

I gratefully acknowledge funding for my travel since I started writing this book in 2016, so that I could attend conferences and collaborate with colleagues. The sources of funding were as follows: at Swarthmore College, the Mathematics Department; at Brown University, the Mathematics Department and the Chancellor's Professorship; and at Phillips Exeter Academy, the Office of the Dean of Faculty and the Graves Family Teaching and Innovation Award.

Look, this book is essentially a love letter: to the objects of study, and to the people who study them. This book contains all the parts of billiards that I find most beautiful: the way cutting sequences work together with continued fractions, the way you can shear and cut up and reassemble surfaces in totally unexpected ways, the way tiling billiards end up folding into a circle and becoming interval exchange transformations – and all the rest. I've drawn the pictures in the way I imagine them: surfaces with each of their edges in a different bright color, points jumping up and down to move around interval exchange

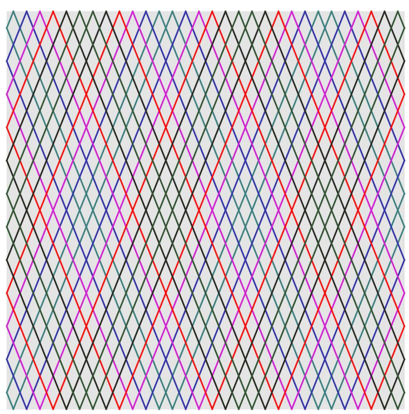
transformations, and shears pushing points around the plane to create rainbow-colored trees. Additionally, this book contains the *people* who are really excited about this stuff, who have been working for years to figure these things out and build up the theory.

I am grateful to have been a member of the billiards and flat surfaces research community. Doing mathematics is hard work, and it's all been tremendously worth it each time I figured something out – a new relationship between familiar things, or a new theorem. After my Ph.D. thesis, all of that work and all of the discoveries have been in collaboration with others, which has made my life rich and satisfying. Working in this community has been a great joy, as I got to know the people in this book, visited many of them at their home universities all around the world, and learned from them.

Thank you all.

Chapter 1

Introduction to billiards in many forms



Six parallel periodic paths on the square billiard table.

In life, *billiards* is a game where a ball bounces around inside a rectangular table. In mathematics, we'll extend the notion of billiards considerably. In this first chapter, we'll meet billiards inside polygonal tables, billiards inside smooth tables, and billiards *outside* of a table. The idea of the first chapter is to introduce all of the big ideas of the course, in their simplest forms. We will understand the simple case very well, and then later when we study more complicated things, we will have a solid background of understanding to build on.

In this chapter, you'll learn to draw beautiful, accurate pictures of periodic billiard paths in a square billiard table. Drawing accurate pictures is an excellent tool that you can use to understand what's going on. I recommend that you draw a picture for every problem, preferably a really big one.

The most powerful tool in the study of billiards in polygons is *unfolding* the billiard table. In its unfolded form, the table becomes a surface, and the path of the ball becomes an infinite line, which sometimes closes up into a periodic path. This opens up the study of linear trajectories on *flat surfaces*, which is a big area of current research and a main object of this course. We'll start with the square torus surface, and later we'll study more complicated surfaces.

Another powerful tool is transforming the *geometric* problem of a billiard path into a *combinatorial* problem about the list of edges that the ball hits. A list of symbols (edge labels) is much simpler than a picture of a path, and these lists (called *bounce sequences* or *cutting sequences*) have a lot of beautiful structure.

Let's get started!

1. What are periodic paths and where can we find them?

1. Consider a ball bouncing around inside a square billiard table, as on the left side of Figure 1. We'll assume that the table has no "pockets" (it's a billiard table, not a pool table!), that the ball is just a point, and that when it hits a wall, it reflects off and the angle of incidence equals the angle of reflection, as in real life.

(a) A billiard path is called *periodic* if it repeats, and the *period* is the number of bounces before repeating. Construct a periodic billiard path of period 2 on the square table.

(b) For which other periods can you construct periodic paths?

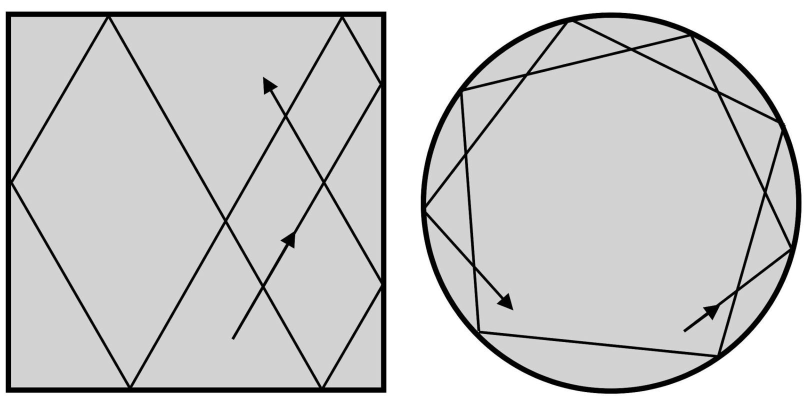


Figure 1. Billiard trajectories on square and circular billiard tables, respectively.

2. Now consider a *circular* billiard table, as on the right side of Figure 1. Again assume that the ball is just a point, and that when it bounces off, the angle of incidence equals the angle of reflection. Note that in a billiard table with curved edges, the ball reflects off of the *tangent line* to the point of impact.

(a) Draw several accurate pictures of billiard trajectories in a circular billiard table.

(b) Consider paths that close up (*periodic* paths), and also paths that don't (*aperiodic* paths). What is the probability that a billiard path in the circular table is periodic?

(c) Describe the behavior in general.

They did the math # 1. Diana Davis

In each section of this book, I'll tell you about someone who has worked on the math in that section, or who is somehow related to that section. The first person is... me, because I wrote the book. Hello! I'm Diana. It's nice to meet you. I love billiards, especially on polygonal billiard tables, and particularly on regular polygons. You can see in the picture for THEY DID THE MATH # 1 that I have a regular octagon billiard table in my office, and I'm pretty excited about it, though when it comes down to it, I'm more of a regular pentagon aficionado [**15**]. I hope you enjoy this book as much as I do.

By the way, are you having trouble solving the first two problems? You might be thinking, "I will try to find a solution online." This approach is unlikely to be fruitful, and is *very* unlikely to fill you with delight. Let me suggest a more delightful way to solve these problems: *paper, scissors, pencil* – and your extremely capable brain! Make a BIG diagram, big enough to really see things clearly. Then make another BIG diagram. This is the way to figure things out!

For Problem 2, here's a suggestion of one way to figure out what's going on in this problem. Trace a large circular object, such as a roll of tape, onto a piece of paper. Use another piece of paper to trace and then cut out the region between a chord and the circle (in pink in Figure 2). Trace along the chord it forms. Then move the paper to the endpoint of the chord you drew, and repeat. Continue until you get a sense of what is going on.

Figure 2. With such simple tools, I made these awesome pictures of billiards on a disk!

For Problem 1, I folded my paper on the diagonal, and then cut off the extra bit to make it square. I cut off a diagonal piece of a sticky note to make a non-isosceles right triangle. The two small angles of

Figure 3. Billiards on the square don't pop as quickly as they do on the disk, but we love them anyway.

the right triangle are the angles that the trajectory will make with horizontal and vertical edges of the square, respectively. (Cutting a piece of paper to get an angle and then using it over and over is easier than measuring with a protractor every time.) I used my triangle to make an accurate billiard bounce (Figure 3), which I extended all the way to the next edge.

My point is: You can do it! If you are having trouble solving later problems in this book, please see the problem map in the "Hints for selected problems" section on page 201.

2. We billiard outside of the box

3. Draw a line on an infinite square grid, and record each time the line crosses a horizontal or vertical edge. We will assume that the direction of travel along a line is always left to right. We could record the line in Figure 4 with the sequence . . . ●●●●●●●●●●●●●● . . ., or we could assign A to horizontal and B to vertical edges, and record it as $\ldots BABBABBABBA\ldots$.

(a) What is the slope of the line in the picture?

(b) Record this *cutting sequence* of As and Bs, for several different lines, including one with slope 3/5. Describe any patterns you notice. What can you predict about the cutting sequence, from the line?

(c) What should you do if the line hits a vertex?

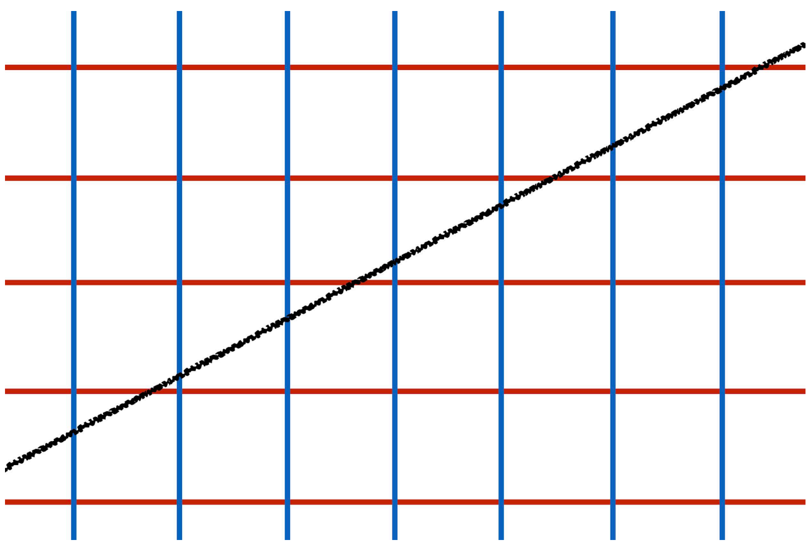


Figure 4. A portion of an infinite square grid for Problem 3.

Caroline Series (THEY DID THE MATH # 2) wrote a series of papers exploring cutting sequences on the square grid and linking them to other areas of mathematics [**50**, **51**]. We will see that cutting sequences are related to group theory and continued fractions; Caroline also explained their relationship with hyperbolic geometry. We will see a little bit of hyperbolic geometry in § 33.

They did the math # 2. Caroline Series

Here are the ways that people typically deal with lines that hit vertices, or billiard trajectories that hit corners of the table:

- *Authoritarian*: Trajectories are not allowed to hit vertices.
- *Minimalist*: If a trajectory hits a vertex, it stops.
- *Indecisive*: The vertex is on both sides: it's ambiguous.
- *Optimistic*: If the ball hits the corner pocket, you win!

In any case, we generally consider trajectories that do not hit vertices.

True story: A few years ago when I was teaching this course, I told my students that we don't let trajectories hit vertices, and they were dissatisfied with my explanation. Then two of them went and played squash together (for real), which is essentially billiards in a cube. The next day, they said: "now we agree, the ball should not be allowed to hit the vertex – when the squash ball hits the corner of the room, it bounces in a totally unpredictable direction!"

To be precise... In fact, while the *cutting sequence* corresponding to a trajectory that hits a vertex is ambiguous, the forward *trajectory* itself is not necessarily ambiguous. For example, on the square billiard table, nearby parallel trajectories continue to be nearby and parallel after two reflections (left side of Figure 5). But on the regular pentagon, two nearby parallel trajectories have very different futures if they hit different sides of a vertex (right side of Figure 5).

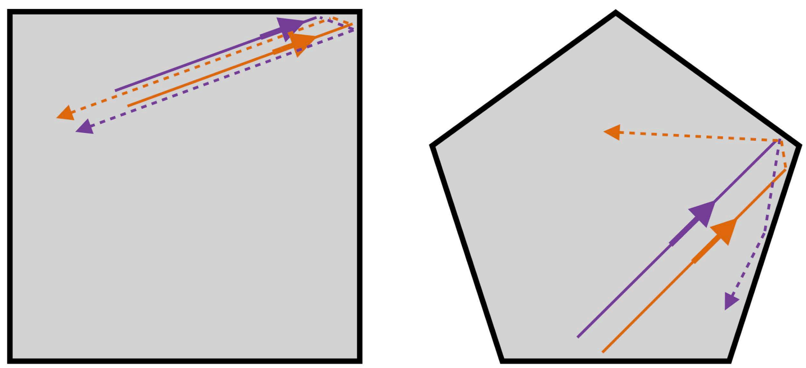


Figure 5. The forward trajectory is ambiguous when the vertex angle does not evenly divide π.

It turns out that if the vertex angle is a divisor of π, the behavior is like the square, and otherwise, the behavior is like the pentagon.[1] In a squash court, the angle between the wall and the floor evenly divides 2π, so perhaps the issue is that the squash ball has a positive radius, and the problem arises when the ball hits both walls simultaneously.

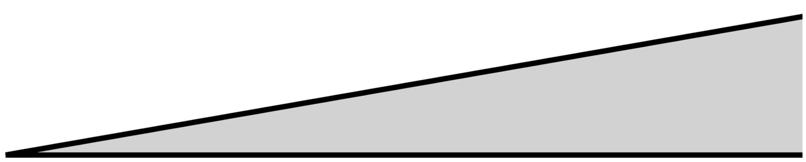


Figure 6. Part of an infinite sector billiard table.

4. Consider a billiard "table" in the shape of an infinite sector with a small vertex angle, say $10°$ as shown in Figure 6. Draw several examples of billiard trajectories in the sector, calculating the angles

[1]Thanks to Barak Weiss (# 24) for pointing out this entire issue.

at each bounce so that your sketch is accurate. Is it possible for a trajectory to bounce infinitely many times within this picture?

5. *Outer billiards.* Though it may seem strange to call it "billiards," we can also define a billiard map on the *outside* of a billiard table. First, choose a starting point p, and a direction, either clockwise or counter-clockwise. Then draw the tangent line from p to the table in that direction to find the point of tangency. Double the vector from p to the point of tangency, and add this to p to get p', as in Figure 7. Repeat to find p'', and so on.

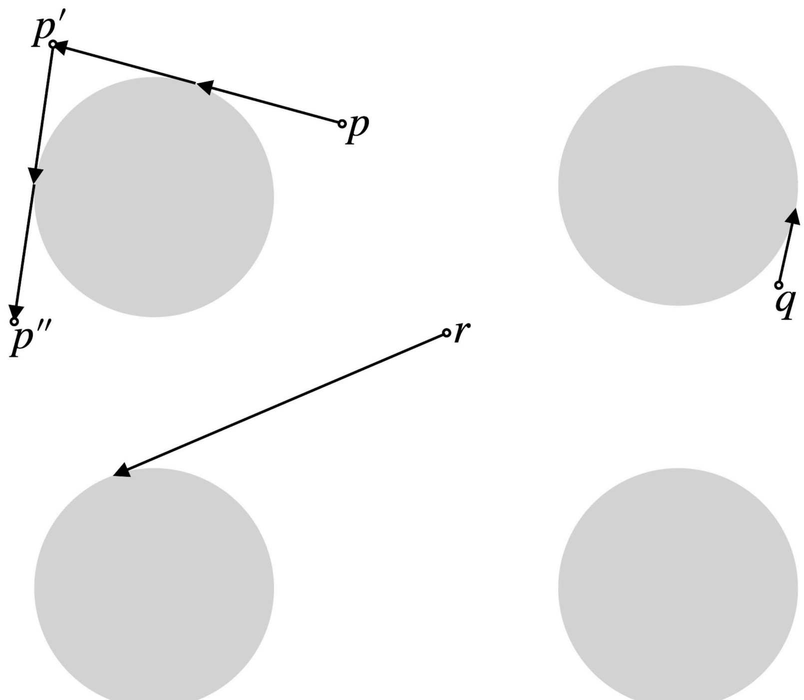


Figure 7. Trace this picture into your notebook so that you can accurately work out the behavior for p, q and r.

(a) Work out the first five or six iterations for the starting points given in Figure 7, and then describe the behavior in general.

(b) What is the probability that p returns to its starting point?

(c) What does the set of *all* the images of p look like? Consider the case when p returns to its starting point, and also when it doesn't.

(d) Can you make a periodic path of period 5? Can you make more than one? If so, how many?

6. *Symmetries of the square.* If you turn a square 90° counter-clockwise, it looks the same as before. We call a 90° counter-clockwise rotation a *symmetry* of the square, because after you do it, you have a square just like the original.

In this problem, we'll find all the symmetries of the square. Of course, if you rotate a square by 90°, it looks identical to the original, so to keep track of the square's orientation, we'll draw an R on it (Figure 8).

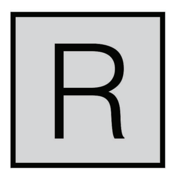


Figure 8. A square. The R is to keep track of its orientation.

(a) Cut out a square and draw an R on one side, as shown, and also hold it up to the light and trace through a backwards R on the back.

(b) How many different symmetries of the square can you find? Record them in a table like Figure 9.

(c) Do you have all of them? If so, explain how you know.

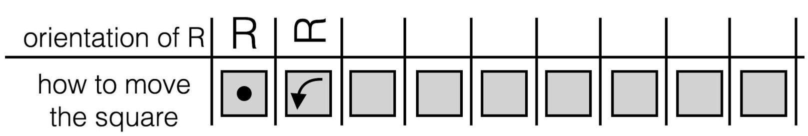


Figure 9. A table for recording symmetries of the square.

3. We unfold

7. A powerful tool for understanding inner billiards is *unfolding* a trajectory into an infinite line, by creating a new copy of the billiard table each time the ball hits an edge [**64**]. Figure 10 shows steps of the unfolding process for a small piece of trajectory of slope ± 2 in the square.

(a) Draw some more steps of the unfolding.

(b) Draw the complete billiard path in the square: keep going until it closes up.

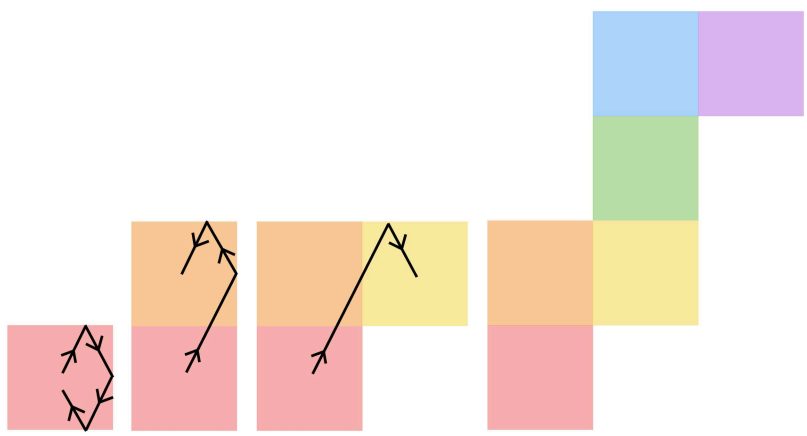


Figure 10. Unfolding the square billiard table for every bounce of a trajectory of slope ± 2.

(c) Use the unfolding to explain why a trajectory with slope 2 yields a *periodic* (repeating) billiard trajectory on the square.

(d) Which other slopes yield a periodic billiard trajectory?

When we say "a trajectory with slope 2," we are assuming that one edge of the square table is horizontal. If our billiard table is tilted, we just rotate it until it does have a horizontal edge. This is one way of reducing our problem (to polygons with a horizontal edge) and making it easier to talk to each other ("slope 2" instead of "with the edge, the trajectory makes an angle whose tangent is 2"). Another way to reduce our work is to only consider trajectories in a small sector of directions; this is what our work in Problem 6 will do for us in the future (Problem 41).

It turns out that billiards on the square are related to number theory, via *continued fractions*. Continued fractions are an efficient (and honestly quite fun) way of expressing real numbers as nested fractions. We'll play with continued fractions for a while to develop our skills, and then see how everything fits together a little later.

8. The *continued fraction expansion* gives an expanded expression of a given number. To obtain the continued fraction expansion for a number, say 15/11, we do the following:

$$\frac{15}{11} = 1 + \frac{4}{11} = 1 + \frac{1}{11/4} = 1 + \frac{1}{1 + 7/4}$$
$$= 1 + \frac{1}{2 + 3/4} = 1 + \cfrac{1}{2 + \cfrac{1}{4/3}} = \mathbf{1} + \cfrac{1}{\mathbf{2} + \cfrac{1}{\mathbf{1} + \cfrac{1}{\mathbf{3}}}}.$$

The idea is to pull off 1s until the number is less than 1, take the reciprocal of what is left, and repeat until the reciprocal is a whole number. Since all the numerators are 1, we can denote the continued fraction expansion compactly by recording only the bolded numbers: $15/11 = [1; 2, 1, 3]$. The semicolon indicates that the initial 1 is outside of the fraction.

(a) Find the continued fraction expansion of $3.14 = 157/50$.

(b) Find the first few steps of the continued fraction expansion of π, and explain why the common approximation 22/7 is a good choice. What is the best fraction to use, if you want a ratio of integers that have 3 or fewer digits?

(c) Find a rational approximation of the number whose continued fraction expansion is $[1; 1, 1, 1, \ldots]$. This number, known as the golden ratio φ, is sometimes called the "most irrational number." Explain.

In part **(a)** you found that the continued fraction expansion of 3.14 is $[3; 7, 7]$. Is this the best approximation for π that is a ratio of integers having three digits or fewer? No, part **(b)** shows that we can find a better rational approximation by using the continued fraction expansion, and truncating it at a convenient point. Indeed, such *convergents* of the continued fraction expansion give the best rational approximations for a given size of denominator.

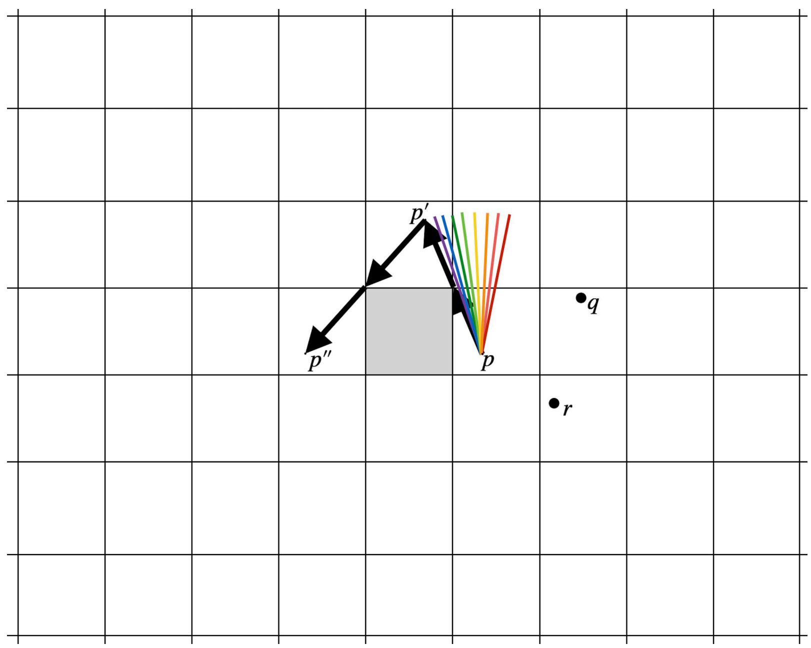


Figure 11. In outer billiards on the square, points bounce around *outside of* the square billiard table. We wonder: how do they move?

9. We can also play outer billiards on polygonal tables. Here, the "tangent line" is always through a vertex − you can think of sweeping a line counter-clockwise until it hits a vertex, as shown in Figure 11.

Find the forward orbits of the points p, q and r in the picture. Can you find any periodic trajectories? Can you find any aperiodic trajectories? *Hint:* be accurate. Consider measuring with your ruler. Actually, don't use a ruler; use symmetry!

The outer billiards system was proposed as a toy model for planetary motion: the table is the sun, and the point is the planet bouncing around it. It is easier to analyze a *discrete* dynamical system, in which a planet jumps from place to place, than a *continuous* dynamical system in which planets move smoothly. It is important to know whether our solar system is stable, or whether the Earth will spin out away from the sun, or what. Related to this, it was for a long time an open problem whether there exists a shape of table, and a point outside the table, such that under the outer billiard map the point eventually bounces off to infinity. The answer is yes: Rich Schwartz (THEY DID THE MATH # 3) showed that the *Penrose kite* has this property [**47**], and Dmitry Dolgopyat and Bassam Fayad showed the same for the half disk [**19**], both in 2009. The picture shows the author celebrating Guy Fawkes Night with Rich in Oxford in 2012.

They did the math # 3. Richard Schwartz

10. *The billiard reflection law, linear case.* We wish to show that, when a billiard trajectory hits the edge of the table, the angle of incidence equals the angle of reflection. We will use the *Fermat principle*: when the ball travels from point A, to the table's edge, and then to B, it follows the (locally) shortest path. We will consider the case when the ball hits a linear edge of the table. Use reflection (or "unfolding") in the edge of Figure 12 to show that the shortest path from A to the edge to B satisfies the billiard reflection law.

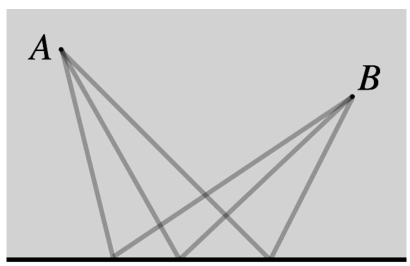


Figure 12. Some options for bouncing off of a linear edge, and a large amount of white space to allow for reflection.

4. We learn to draw accurate pictures

11. Building on our work from Problem 7, here's another way that we can unfold the square billiard table. First, unfold across the top edge of the table, creating another copy in which the ball keeps going (see Figure 13). The new top edge is just a copy of the bottom edge, so we now label them both A to remember that they are the same. Similarly, we can unfold across the right edge of the table, creating another copy of the unfolded table. The new right edge is a copy of the left edge, so we now label them both B. When the trajectory hits the top edge A, it reappears in the same place on the bottom edge A and keeps going. Similarly, when the trajectory hits the right edge B, it reappears on the left edge B.

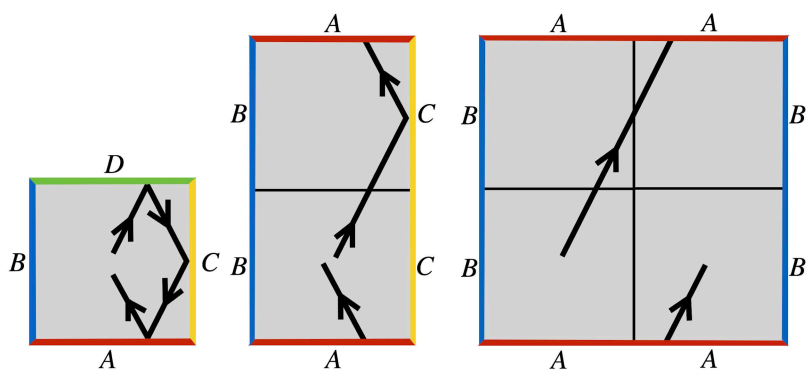


Figure 13. Unfolding the square billiard table, and a billiard trajectory on it.

(a) The partial billiard trajectory shown on the left part of the figure repeats after six bounces. Sketch in the rest of the trajectory in each of the three pictures above. What is the corresponding *cutting sequence* of As and Bs for the trajectory on the surface on the right part of the figure?

(b) When we unfolded the trajectory to a line in Problem 7, we created a new copy of the table every time the trajectory crossed an edge. Explain why, in the picture above, just four copies is enough.

(c) Suppose that you have a rectangular sheet of very stretchy rubber. You tape together the top and bottom edges (edge A) to create a tube, and then you curl the tube around and attach the open ends to each other along their edges (edge B), as shown in Figure 14. Explain. The result is called a *torus*, the surface of a donut.

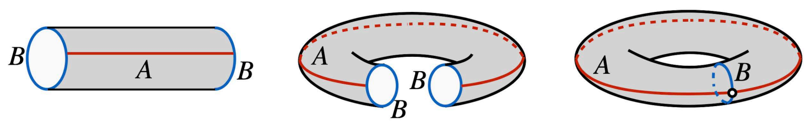


Figure 14. Stretching a flat square into a torus.

The field of mathematics devoted to the study of objects like the square torus that we just constructed is called *flat surfaces*. Hundreds of mathematicians around the world are working on flat surfaces, particularly in France and the United States. It is currently a "hot" field, with many papers posted every week with new results. Amie Wilkinson (They did the math # 4) created a phenomenal animation showing how, as we did with the square in Problem 11, we can make an octagon into a flat surface. It is at 26:00 of her Fields Symposium lecture from 2018, available here: `https://www.youtube.com/watch?v=zjccKzHIniw&t=1560s` The picture shows Jinxin Xue, Aaron Brown, Amie, and Clark Butler hiking in Chile in 2015.

They did the math # 4. Amie Wilkinson

A surface is *flat* if it looks like the Euclidean plane everywhere, except possibly at finitely many "cone points," where the angle must be a multiple of 2π. The torus surface from Problem 11 looks like the plane everywhere, so it is flat; we sometimes call it the *flat torus*. As we will see in Problem 75, the octagon surface from Amie's talk has a single cone point with angle 6π, so it is also a flat surface.

12. Show that the cutting sequence corresponding to a line of slope 1/2 on the square grid is periodic. Which other slopes yield periodic cutting sequences? What can you say about the period, from the slope? Write proofs of your claims.

13. In Problem 11, we ended up with a trajectory of slope 2 on the *square torus* surface. Figure 15 shows some scratchwork for drawing a trajectory of slope 2/5 on the square torus. Starting at the top-left corner, connect the top mark on the left edge to the leftmost mark on the top edge with a line segment, as shown. Then connect the other six pairs with parallel segments, down to the bottom-right corner.

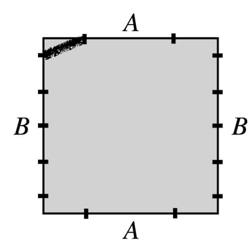


Figure 15. Scratchwork for drawing a trajectory of slope 2/5 on the square torus. Accuracy is *very* important here: look *very* carefully at the spacing of the ticks.

(a) Explain why, on the torus surface, these line segments connect up to form a continuous trajectory. Follow the trajectory along, and write down the corresponding cutting sequence of As and Bs.

(b) Exactly where should you place the tick marks so that all of the segments have the same slope? Prove your claim.

(c) Create an accurate picture for a trajectory of slope 1/2 and then 3/2. *Hint*: make sure that all of your segments look parallel.

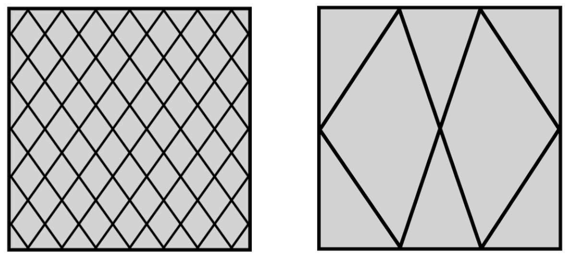


Figure 16. A nice picture of a long periodic trajectory on the square billiard table, plus a "billiard trajectory" where something has gone wrong.

(d) Draw a picture of a *billiard* trajectory with slope $\pm 2/5$.

(e) Something is wrong with the "billiard trajectory" on the right in Figure 16. Explain.

14. Prove that every billiard trajectory on the square with irrational slope is aperiodic.

A note on terminology. In this book, I use the words "path" or "trajectory" to refer to linear motion on a billiard table or a surface. Other authors use the word "geodesic" to describe the same thing. On a surface, a *geodesic* is the (locally) shortest path between two points. For example, on a sphere, the geodesic between any two points is part of a great circle. On a flat surface, geodesics are lines.

15. *The billiard reflection law, curved case.* We proved this law for linear boundaries in Problem 10; now we will prove it for curved boundaries (Figure 17). Prove that, when a billiard ball follows the shortest path in a billiard table, reflecting off a *curved* edge, the angle of incidence equals the angle of reflection. Recall that for a curved boundary, we measure the angle between the trajectory and the tangent line to the point of impact.

Hint: Again, we will use the principle that when the ball travels from point A, to the table's edge, and then to B, it follows the (locally) shortest path. One way is to use the multivariable calculus principle that the gradient vector of the distance function points in the direction

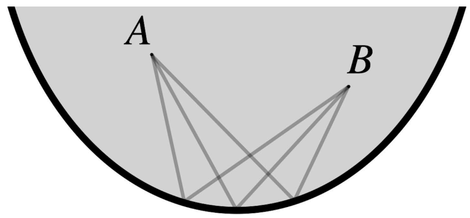


Figure 17. Some options for a billiard trajectory bouncing off of a curved wall.

of greatest increase of the function, and to apply this to both A and B. Another way is to apply an equilibrium tension argument from physics, imagining the boundary of the table as a wire, and the billiard trajectory as an elastic string fixed at A and B that passes through a small ring threaded through the boundary wire.

5. We do a little bit of group theory

16. Consider again (following Problem 4) a billiard table in the shape of an infinite sector, with vertex angle α (Figure 18). Use unfolding to show that any billiard trajectory on such a table makes **(a)** finitely many bounces, and in fact **(b)** at most $\lceil \pi/\alpha \rceil$ bounces. *Hint*: Unfold the sector as many times as you can.

Here the notation $\lceil \cdot \rceil$ is the "ceiling" and means "round up," e.g. $\lceil \pi \rceil = 4$.

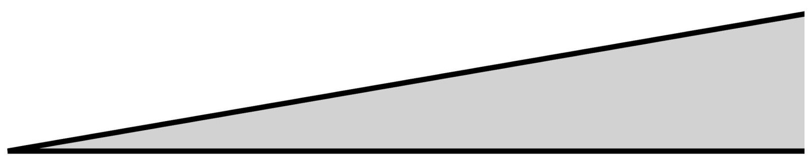


Figure 18. The same infinite sector billiard table as before, encountered now with more tools in our toolbox.

17. In Problem 6, you found the eight symmetries of the square. It turns out that these eight symmetries form a *group*, called the *dihedral group* of the square. For a set of symmetries to be a group, it must have the following properties:

(1) It contains an *identity element*, a symmetry that does nothing;

(2) Each symmetry has an *inverse*, a symmetry that "undoes" its action;

(3) It is *closed*: composing two symmetries (doing one and then the other) yields a symmetry that is also in the group;

(4) Composing symmetries is *associative*, i.e. $a(bc) = (ab)c$ for symmetries a, b, c.

(a) Explain why (1), (2) and (3) hold for the symmetries of the square.

(b) Use your R from Problem 6 to fill in the table in Figure 19, which is known as a *Cayley table*. Do you see any patterns? Prove that they exist.

Note: it is much easier to see patterns if you denote a symmetry by its arrow or dashed line; it is much more difficult to see patterns if you use the oriented R. Use the arrow or dashed line!

(c) Does this group of symmetries commute, i.e. is $ab = ba$ true for every pair of symmetries a, b? If not, is there *any* pair of symmetries that commutes?

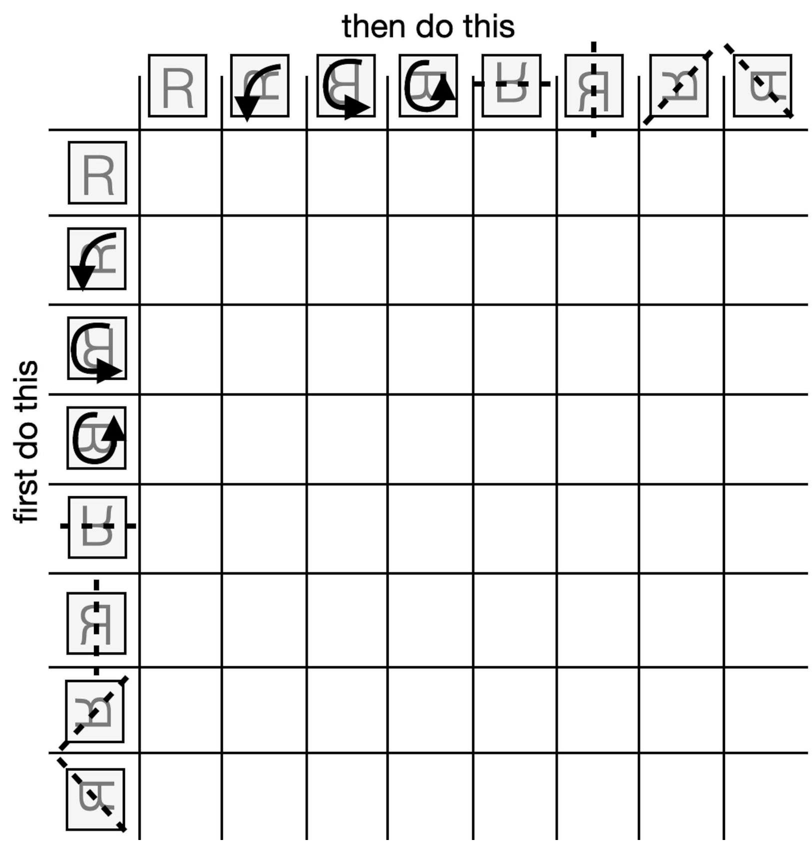


Figure 19. A Cayley table for the symmetries of the square.

18. You will need: colored pens or pencils. Consider again counter-clockwise outer billiards on the square table (Figure 20).

(a) Points p on the blue lines are not allowed, because their images p' are ambiguously defined. Explain.

(b) Points p whose image p' is on a blue line are also not allowed. Explain. These are the *inverse images* of the blue points. Color these points red. (Color them red! I'm specifying the colors here so that you can check your answers with someone else.)

(c) The inverse images of the red lines are also not allowed. Explain. Color these points green. *Hint:* each one has two pieces.

(d) Color the inverse images of the green points black. Keep going. Describe the full set of disallowed points.

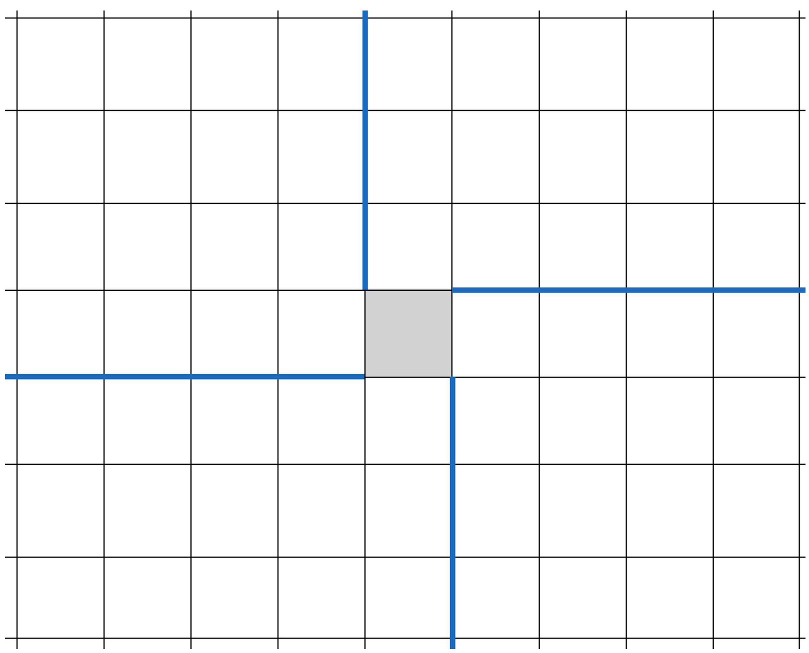


Figure 20. Coloring the disallowed points for outer billiards on the square.

19. Let's gather some data and make some conjectures.

(a) Construct an *accurate* (see Problem 13) picture of a trajectory on the square torus with slope 3/4. Then draw accurate pictures of two more trajectories with slopes of your choice.

(b) For each trajectory, find the corresponding cutting sequence.

(c) What is the relationship between the slope and the cutting sequence?

We discussed in § 2 that we do not allow trajectories to hit vertices. What if you *do* want to allow a trajectory to hit a vertex, and moreover, after it hits the vertex, you want it to keep on going? Well, notice that if you draw a linear trajectory on a piece of paper, at each point on the trajectory, there is π worth of angle on each side of the line (left side of Figure 21). Later, we will see that on the sort of flat surfaces that we are studying, we can have *cone points* with 4π or 6π or other angles around them. When a trajectory hits such a cone point and then continues on through, the way to proceed is to require that there is *at least* π of angle on each side (right side of Figure 21).

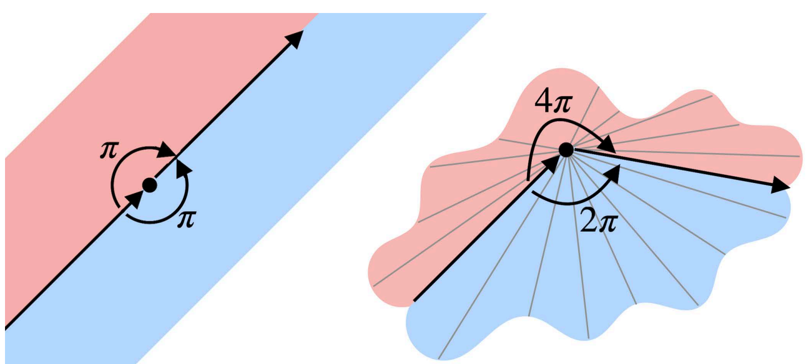


Figure 21. When we decide to allow a trajectory to hit a vertex, we require that at each point of the trajectory, there is *at least* π of angle on each side. (left) In the plane, there is always *exactly* π on each side; (right) for a cone point with e.g. angle 6π, there is considerably more. For a more complete understanding of what it means to have 6π of angle at a vertex than this small cartoon, see Problem 100.

Noelle Sawyer (They did the math # 5) studies these "singular" trajectories and their geodesic continuations [**9**]. The picture shows Naomi Reed and Noelle enjoying breakfast in Austin.

They did the math # 5. Noelle Sawyer

6. We fold up torus trajectories into billiards

20. You will need: tissue paper or other thin paper. We saw that a billiard trajectory on the square table can be *unfolded* to a trajectory on the square torus. Going the other way, a trajectory on the square torus can be *folded* to a billiard trajectory on the square table.

(a) Confirm that each trajectory in Figure 22 is a closed path on the square torus.

(b) Carefully trace the first figure onto a piece of thin paper. Fold it in quarters as indicated by dashed lines, and then hold it up to the light: behold, a billiard trajectory!

Repeat for the second figure.

(c) For each picture, find the corresponding cutting sequence on the square torus, and also the bounce sequence on the square table. Note any observations.

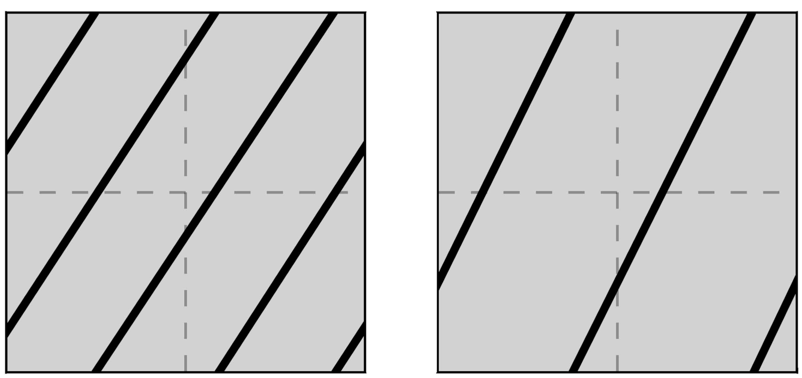


Figure 22. Closed trajectories on the square torus, for folding up into billiard trajectories on the square billiard table.

As previously explained, the study of flat surfaces is a very hot field these days, and many people are proving results about them. Sometimes, people are perfectly satisfied with results about flat surfaces, and they don't fold up their surfaces to get a billiard table back. You will not be one of these people.

In Problem 20, you checked that the two pictured trajectories are, in fact, closed trajectories on the square torus surface. You might wonder: how many closed trajectories *are there* on the square torus? We'll count them later, in Problems 118 and 127. Kasra Rafi (They did the math # 6) has worked on counting closed trajectories on *all possible* surfaces of a given type – all surfaces in a given *stratum* [**24**]. The picture shows the author and Kasra discovering the wonders of the candy cabinet at Oberwolfach in 2014.

They did the math # 6. Kasra Rafi

21. In Problem 8, we constructed a continued fraction algebraically. Figure 23 illustrates the geometric version of the continued fraction algorithm for a number x:

(1) Begin with a $1 \times x$ rectangle, or $p \times q$ if $x = p/q$.

(2) Cut off the largest possible square, as many times as possible. Count how many squares you cut off; this is a_1.

(3) With the remaining rectangle, cut off the largest possible squares; the number of these is a_2.

(4) Continue until there is no remaining rectangle. The continued fraction expansion of x is then $[a_1, a_2, \ldots]$ or possibly $[a_1; a_2, \ldots]$.

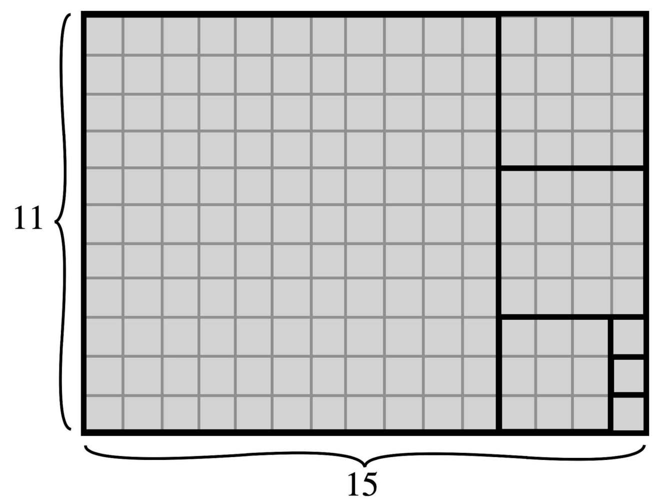


Figure 23. The geometric interpretation of the continued fraction algorithm for 15/11.

(a) Draw the rectangle picture for 5/7 to geometrically compute its continued fraction expansion.

(b) Compute the continued fraction expansion for 5/7 in the way explained in Problem 8, and check that your results agree. Explain why this geometric method is equivalent to the fraction method previously explained, for determining the continued fraction expansion.

22. In Problem 18 you showed that for outer billiards on the square, all of the points on the square grid lines are not allowed. Choose a point p that is *not* on one of the grid lines (refer back to Figure 11). Under the outer billiard map, this point reflects through a sequence of vertices $v_1, v_2, \ldots$ where each v_i is one of the four vertices of the square table. Explain why *every* point that is in the same (open) square as p reflects through that *same* sequence of vertices.

23. Consider a billiard trajectory in the disk, where at each impact the trajectory makes angle α with the tangent line to the circle. (Refer back to Problem 2.)

(a) Find the central angle θ from the circle's center, between each impact point and the next one, as a function of α.

(b) Prove that if $\theta = 2\pi p/q$ for integers p and q, then every billiard orbit is q-periodic and makes p turns around the circle before repeating.

(c) What happens if θ is *not* a rational multiple of π?

24. In Problem 13, we put two marks on edge A and five marks on edge B and connected the marks to create a trajectory with slope $2/5$ on the square torus. Do the same with four marks on edge A and 10 marks on edge B, and explain what you get. *Hint*: Fourteen line segments is a lot, so use a grid to make your picture accurate!

7. Automorphisms come for the torus

25. Prove that a trajectory on the square torus is periodic if and only if its slope is rational.

26. You will need: string, tape.

(a) Mark two dots on a piece of paper, and tape down your piece of string on each dot, leaving a lot of slack in the string. With your pencil, pull out the string until it is taut and trace out all the points the pencil can reach, as shown in Figure 24, to create an ellipse.

(b) Each endpoint of the string is called a *focus* of the ellipse. Show that a billiard trajectory through one focus reflects through the other focus: the string is a billiard path in the ellipse.

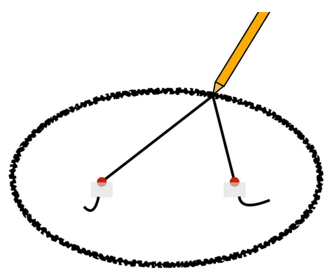


Figure 24. Constructing an ellipse with tape and string.

The reflection property of ellipses is well known, and appears in architecture as the *whispering gallery.* Several U.S. state house rotundas, and the National Statuary Hall at the U.S. Capitol building, have ellipsoidal ceilings, so if you stand at one focus, you can hear someone whisper at the other. Since legislative chambers are often arranged with members of the two political parties on opposite sides, people can actually sit at one focus and listen to what members of the other party are saying at the other focus!

An accessible and impressive example of this is in Grand Central Station in New York City (Figure 25), where although the background noise is very loud, if you speak into one column, someone on the opposite column can hear you.

Figure 25. The whispering gallery in Grand Central Station: the people in opposite corners are talking to each other.

They did the math # 7. Chandrika Sadanand

One way to hear billiards is to stand in a whispering gallery. Another is to imagine that each edge of your billiard table has a different xylophone bar on it, and a billiard trajectory plays an infinite song. You might ask: if you listen to the song, can you reconstruct the shape of the table? Chandrika Sadanand (They did the math # 7) and her collaborators showed that if you know the set of *all possible* such songs – in other words, you know *all possible* bounce sequences on the table – this uniquely determines the shape of the table [**20**]. This sort of thing is known as a "rigidity" result. The picture shows Chandrika and the author in Jerusalem in 2018.

27. An *automorphism* of a surface is an action that takes the surface back to itself, taking nearby points to nearby points. It creates neither holes nor overlaps, and preserves the surface's structure: it is essentially a "symmetry" of the surface. Two types of automorphisms of the square *torus* come from symmetries of the *square*: reflections and rotations, as in Problems 6 and 17.

(a) Explain what a vertical reflection of the square torus looks like on the torus surface. You might think about what it does to the surface, or to a closed path drawn on the surface.

(b) Do the same for a horizontal reflection.

(c) What about diagonal reflections, or rotations?

28. It turns out that there is a third type of automorphism of the square torus that is *not* a symmetry of the square: a *shear*. Figure 26 shows the shear applied to the square at the bottom, and to the surface in three dimensions at the top, where its effect is to twist the torus.

(a) Explain the effect of this shear on the surface, and on a trajectory drawn on that surface.

(b) What 2×2 matrix, applied to the "unit square" $[0,1] \times [0,1]$ shown in the bottom-left picture, gives the parallelogram shown in the bottom-middle picture?

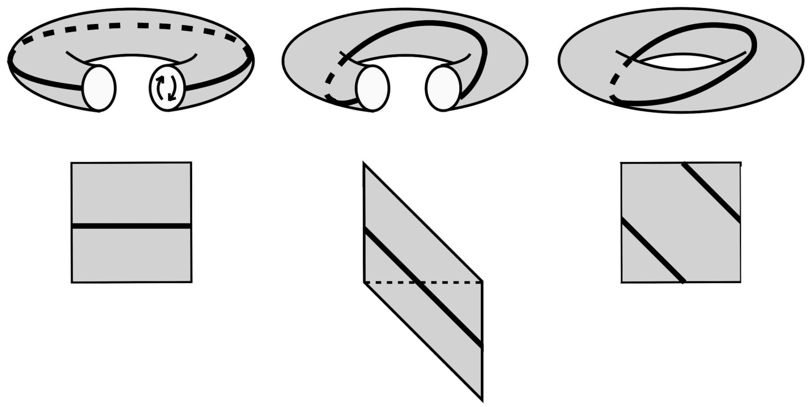


Figure 26. Twisting the 3D torus, shearing the flat torus.

8. Hands-on activities for Chapter 1

29. The pictures in Figure 27 show linear trajectories on the square torus, as usual.

(a) Explain why the purple trajectory (left) is a single trajectory, while the red and blue trajectories (solid and dashed, right) are two different trajectories.

(b) The red and blue trajectories partition the square torus into two pieces. In other words, if the trajectories were walls, the smiley person in the right picture could only explore half of the torus. Justify this statement.

(c) Also explain why the purple trajectory does *not* partition the torus into two pieces – the smiley person in the left picture can explore the whole thing.

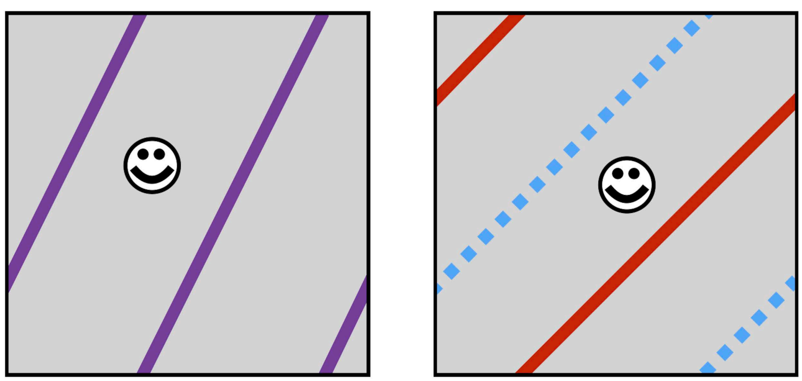


Figure 27. Some closed trajectories on the square torus.

30. *Cutting a bagel into two linked rings.*[2]

You will need: bagel with a large hole in it, serrated knife, tray to catch the crumbs.

1. Draw the red and blue trajectories on your bagel (left side of Figure 28).
2. Cut the bagel: The pointy end of the knife should follow the red trajectory, while the handle follows the blue trajectory. Flip the associated colors halfway through, to keep the handle on the outside.
3. Separate your bagel into linked rings (right side of Figure 28)!

Hint: Sending the knife all the way through is the theoretical construction, which can work in practice, as it did on the bagel in Figure 28 that I made in 2014. These days, I just cut the "skin" of the bagel, rather than sending the knife all the way through, and then use my fingers to tear the soft middle of the bagel in between the skin cuts.

(a) Explain why the procedure above leads to linked rings.

(b) Explain what would have happened if you had cut along the purple trajectory instead.

Figure 28. Cutting a bagel into linked rings, corresponding to parallel closed trajectories on the square torus.

[2]This activity is from George Hart's website: `https://www.georgehart.com/bagel/bagel.html`

Figure 29 shows bagels with trajectories that correspond to slopes 1/2, 2 and 3/2, respectively, on the square torus. Let's make some!

Figure 29. Closed trajectories on bagels.

31. You will need: bagel with a large hole in it, marker. Choose a periodic trajectory, and find a way to mark your bagel to indicate where to draw the trajectory. One method is suggested in Figure 30. Then connect up your marks with smooth curves!

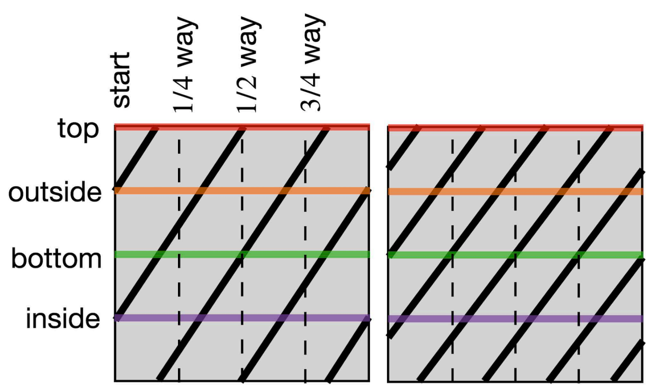


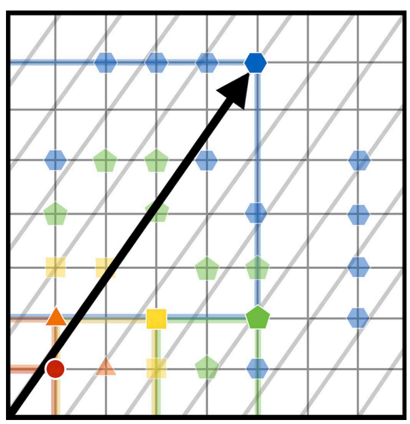
Figure 30. Scratchwork for drawing a trajectory on a bagel.

Connection to knot theory: If the bagel disappears, leaving a trajectory corresponding to slope p/q made out of a piece of string, the result is the (p, q) *torus knot*, meaning that it goes through the center p times and around the outside q times.

Chapter 2

Trajectories, automorphisms, and continued fractions

In Chapter 2, we will explore more about our main protagonists: trajectories, automorphisms, and continued fractions, and we will gradually build a grand unifying theory that unites all three of these ideas.



Thanks to Jaden Sides for the idea behind this picture.

9. We apply symmetries to trajectories

32. Given a trajectory on the square torus, we want to know what happens to that trajectory if we apply an automorphism of the surface. To do this, we can sketch the trajectory before and after applying the automorphism. Do so for each of the eight symmetries of the square in Figure 31, as indicated by the curved arrow or the reflection line, and for the shear. I've done one for you.

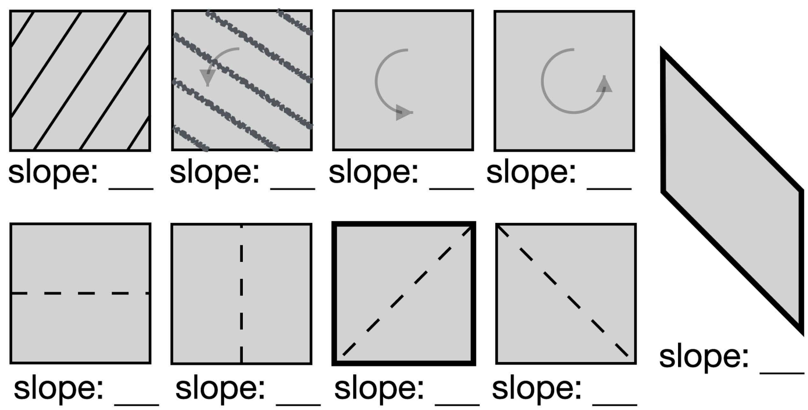


Figure 31. Automorphisms of the square torus, eight of which can be visualized as symmetries of the square. The shear and the flip across the positive diagonal are in bold because we will use them later.

33. (Continuation) Given that the original slope is p/q, for each of the nine symmetries, record the resulting slope after the transformation.

34. An active area of research is to describe all possible cutting sequences on a given surface. On the square torus, that question is: "Which infinite sequences of As and Bs are cutting sequences corresponding to a trajectory?" Let's answer an easier question: How can you tell that a given infinite sequence of As and Bs is *not* a cutting sequence? You have computed many examples of cutting sequences that *do* correspond to a line on the square grid or square torus. Now make up an example of an infinite sequence of As and Bs that *cannot* be a cutting sequence on the square grid or square torus, and justify your answer.

As described above, an active area of research is to describe all possible cutting sequences on a given surface. John Smillie and Corinna Ulcigrai (They did the math # 8) classified all cutting sequences on the *regular octagon* surface, which is created similarly to the square torus [**54**, **55**]. Because cutting sequences are infinite, and most are not periodic, it turns out that there is no finite criterion for deciding whether a given cutting sequence is valid: the algorithm necessarily requires a possibly unbounded number of steps. We will see in the cutting sequence characterization theorem (Problem 79) that the same is true for cutting sequences on the square. The picture shows John and Corinna with the author in Bristol in 2012.

They did the math # 8. John Smillie & Corinna Ulcigrai

35. In Problem 22, we showed that for outer billiards on the square, points in a square move together. Let's explore *how* they move.

(a) Using Figure 32, plot the complete orbit (meaning, until you get back to where you started) of the R and of the winky face under the counter-clockwise outer billiard map. One step is shown for the R. *Hint 1:* To determine the orientation of the image square, you can consider the image of each corner of the square. *Hint 2:* The Rs end up on orange squares, and the winky faces on yellow squares.

(b) Prove that the square of the outer billiard map (this means that you apply it twice) is a *translation* on each individual square.

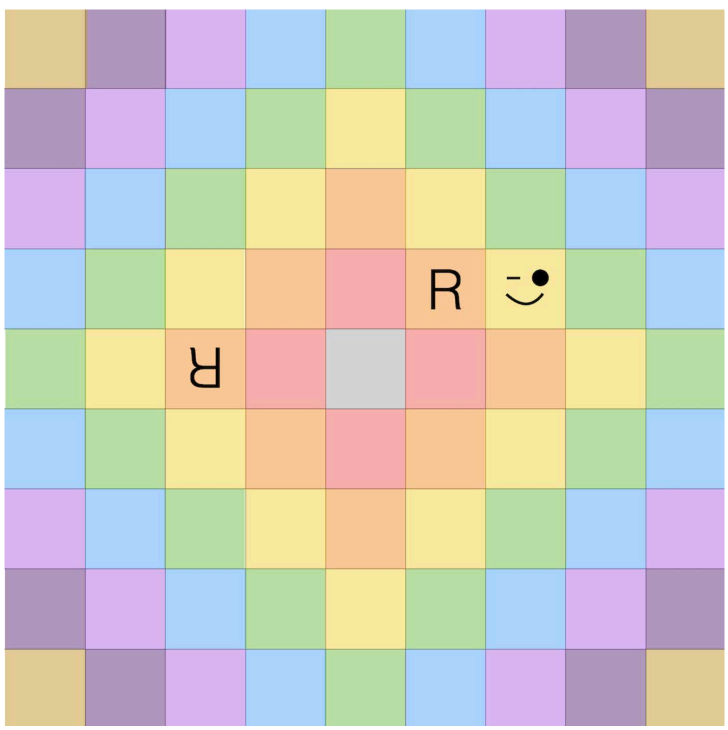


Figure 32. A template for exploring how the outer billiard map on the grey square transforms each colored square: images of square regions form *necklaces* around the billiard table.

10. We dream of an action on cutting sequences

So far, to determine the effect of a surface automorphism on a trajectory lying on that surface, we have drawn a picture of the original trajectory and of the transformed trajectory (Problem 32). It's a great way to understand what's going on, but it's not super efficient. A much more efficient way to write down the effect of the automorphism is to record how it affects the *cutting sequence* corresponding to a trajectory. Then we can act on the cutting sequence – an operation on symbols, not on pictures! – and get the cutting sequence corresponding to the transformed trajectory.

They did the math # 9. Irene Pasquinelli

Irene Pasquinelli achieved this dream: in her master's thesis, and in a subsequent paper with the coauthors pictured in THEY DID THE MATH # 9, she figured out how to determine the effect of an automorphism using only symbolic operations on cutting sequences, for a large class of surfaces [**16**, **39**]. We'll do this for the square torus now.

The picture shows Irene, Corinna Ulcigrai (# 8), and the author in Bristol in 2014.

36. Given a trajectory τ on the square torus, we want to know what happens to that trajectory under an automorphism of the surface.[1] We'll do this by comparing their cutting sequences: the cutting sequence $c(\tau)$ corresponding to the original trajectory τ, and the cutting sequence $c(\tau')$ corresponding to the transformed trajectory τ'. The goal is to figure out how to get $c(\tau')$ directly from $c(\tau)$.

(a) Let τ_2 be the trajectory of slope 2. Sketch a picture of τ_2, and find $c(\tau_2)$.

(b) For each symmetry (1)–(5) below, apply it to τ_2 to get a transformed trajectory τ_2', sketch τ_2', and compute $c(\tau_2')$.

(1) reflection across a horizontal line;
(2) reflection across a vertical line;
(3) reflection across the positive diagonal;
(4) reflection across the negative diagonal;
(5) rotation by 90° counter-clockwise.

(c) Explain how to obtain $c(\tau')$ from $c(\tau)$ for a general trajectory τ, for each of the five symmetries. Prove your answer correct.

37. For each of the five symmetries in the previous question:

(a) Find the 2×2 matrix that performs this symmetry. For the purpose of this question, assume that the square torus is centered at the origin.

(b) Find the determinant of each matrix and give a geometric explanation for why they all turn out to be ± 1.

38. A cutting sequence on the square torus can have blocks of multiple As separated by single Bs, or blocks of multiple Bs separated by single As, but not both. In other words, a sequence of the form $\ldots AA \ldots BB \ldots$ cannot occur. Explain why.

[1] The symbol τ is spelled *tau* and rhymes with "cow."

39. Theorem (billiards in an ellipse). If one segment of a billiard trajectory doesn't pass through the focal segment, then no segments of that trajectory pass through the focal segment, and furthermore all the segments of the trajectory are tangent to the same confocal ellipse.

More precisely: *Consider an ellipse E with foci F_1, F_2. If some segment of a billiard trajectory does not intersect the focal segment F_1F_2 of E, then no segment of this trajectory intersects F_1F_2, and all segments are tangent to the same ellipse E' with foci F_1 and F_2.*

Let's prove it! Steps of the proof below are color-coded in Figure 33.

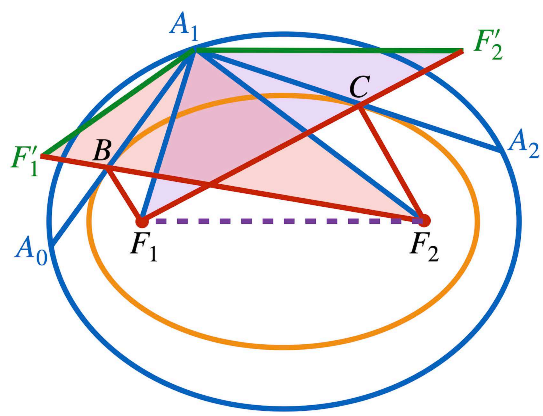


Figure 33. Two confocal ellipses with a finite billiard path, whose markings and colorings follow the steps of Problem 39.

(a) (blue) Consider the billiard trajectory $A_0A_1A_2$ in the larger ellipse E shown in the figure. Explain why $\angle A_0A_1F_1 = \angle A_2A_1F_2$.

(b) (green) Reflect F_1 across $\overline{A_0A_1}$ to create F_1', and reflect F_2 across $\overline{A_1A_2}$ to create F_2'. Explain why $\angle A_0A_1F_1' = \angle A_0A_1F_1$ and $\angle A_2A_1F_2' = \angle A_2A_1F_2$.

(c) Show that $\Delta F_1'A_1F_2$ and $\Delta F_1A_1F_2'$ are congruent.

(d) (red) Mark the intersection of $\overline{F_1'F_2}$ with $\overline{A_0A_1}$ as B, and the intersection of $\overline{F_1F_2'}$ with $\overline{A_1A_2}$ as C. Show that the *string length* $|\overline{F_1B}| + |\overline{BF_2}|$ is the same as the string length $|\overline{F_1C}| + |\overline{CF_2}|$.

(e) Prove the theorem as stated above.

40. Find the continued fraction expansions of $3/2, 5/3, 8/5$, and $13/8$. Describe any patterns you notice, and explain why they occur.
Hint: Also see Figure 34.

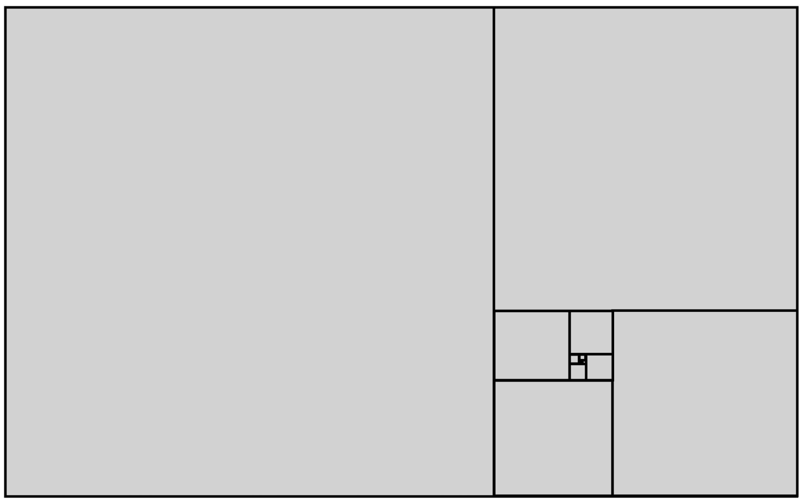


Figure 34. A suggestive picture to accompany Problem 40.

11. The dream comes true

In these problems, we will determine the effect of the shearing automorphism from Problem 28 on a trajectory τ on the square torus, and its corresponding cutting sequence $c(\tau)$. We'll assume that the trajectory is neither vertical nor horizontal, as such trajectories are already easy to understand, and don't work as nicely with our tools.

First, we will apply symmetry (recall Problem 6) to reduce our work to just one set of trajectories:

41. Show that, given a linear trajectory in *any* (non-horizontal, non-vertical) direction on the square torus, we can apply rotations and reflections so that it is going left to right with slope ≥ 1.

Since we have reduced to the case of slopes that are ≥ 1, we will analyze the effect of the vertical shear $\left[\begin{smallmatrix} 1 & 0 \\ -1 & 1 \end{smallmatrix}\right]$, because these slopes work nicely with this shear. Later (in Problems 84 and 88) we will show that every shear can be reduced to this case.

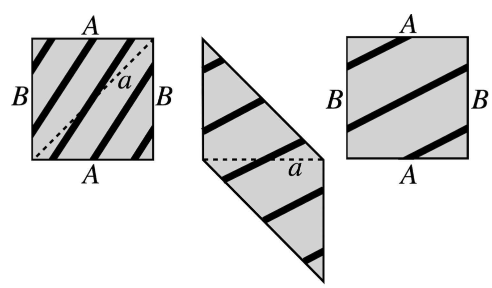


Figure 35. Using an auxiliary edge to understand how the vertical shear transforms a trajectory and its cutting sequence.

As an example, we'll use the trajectory τ with slope $3/2$ (Figure 35), with corresponding cutting sequence $c(\tau) = \overline{BAABA}$ (left picture). We shear it via $\left[\begin{smallmatrix} 1 & 0 \\ -1 & 1 \end{smallmatrix}\right]$, which transforms the square into a parallelogram (middle picture), and then we reassemble the two triangles back into a square torus, while respecting the edge identifications (right picture), yielding the transformed trajectory τ'. The new cutting sequence is $c(\tau') = \overline{BAB}$.

42. Notice that the horizontal edge A in the right picture corresponds to dashed edge a in the left and middle pictures. We can use this *auxiliary edge*, and its corresponding edge crossings, to form an *augmented cutting sequence* $\overline{BAaABA}$, which leads us to the *derived* cutting sequence $\overline{BAA}$:

$$\overline{BAABA} \longrightarrow \overline{BAaABA} \longrightarrow \overline{BaB} \longrightarrow \overline{BAB}.$$

Explain.[2]

They did the math # 10. Priyam Patel

In Problem 41, we showed that we could use symmetries of the square torus surface to reduce our work. Priyam Patel (They did the math # 10) studies such symmetries of surfaces, known as *mapping class groups*, and she also studies curves on surfaces, like the trajectory τ in that problem [**40**]. The picture shows mathematicians

[2]The idea of auxiliary edges and augmented cutting sequences described here comes from John Smillie and Corinna Ulcigrai's paper *Beyond Sturmian sequences: coding linear trajectories in the regular octagon*; see their § 1.2 and Figure 3.

Brandis Whitfield, Noelle Sawyer (# 5), Michelle Chu, Marissa Loving, Aisha Mechery, Priyam, and Cassandra Monroe at a conference on geometry, arithmetic, and groups in Austin in 2022.

43. Perform the geometric process described above for two different trajectories τ of your choice with slope ≥ 1: Using a picture like Figure 36, sketch a trajectory τ, sketch its image as a parallelogram after shearing by $\left[\begin{smallmatrix} 1 & 0 \\ -1 & 1 \end{smallmatrix}\right]$, and then sketch the reassembled square with the new trajectory τ'. For each, record $c(\tau)$ and $c(\tau')$. Try to find the pattern: a rule to get $c(\tau')$ from $c(\tau)$. Then prove your conjecture.

Hint: Apply the "edge marks" technique from Problem 13 on the parallelogram edges to make accurate pictures.

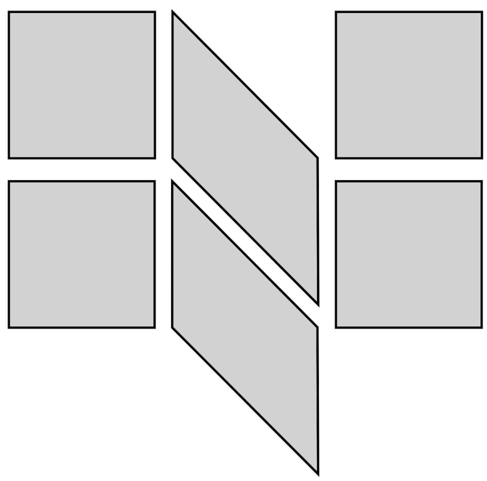


Figure 36. A template for shearing, cutting, and reassembling the square torus, waiting to transform your trajectories.

44. Find the continued fraction expansion of $\sqrt{2} - 1$. Then solve the equation $x = \frac{1}{2+x}$ and explain how these are related.

45. How many billiard paths of period 10 are there on the square billiard table? Of period 12? Construct an accurate sketch of each of them. Does *every* even number have a corresponding periodic billiard path?

46. We have identified the top and bottom edges, and the left and right edges, of a square to obtain a surface: the square torus. If we identify opposite parallel edges of a parallelogram, what surface do we get?

12. We consolidate our gains

We are about to formulate a grand unifying theory relating a trajectory on the square torus, its corresponding cutting sequence, and the continued fraction expansion of its slope. We need these two results (refer to Problem 32):

47. Show that if we apply the flip $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ to the square torus:

(a) The effect on the slope of a trajectory is to take its reciprocal.

(b) The induced effect on the cutting sequence corresponding to a trajectory is to switch As and Bs.

48. Show that if we apply the shear $\begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$ to the square torus:

(a) The effect on the slope of a trajectory is to decrease it by 1.

(b) The induced effect on the cutting sequence corresponding to a trajectory *whose slope is greater than 1* is to remove one A between each pair of Bs. In other words, as you go along the cutting sequence, when you see a B, find the next B, and then between those two Bs, remove exactly one A – and do this for *all* of the Bs (see Figure 37).

...ABAABAAABAABAAABAABAA...

...AB~~A~~AB~~A~~AAB~~A~~AB~~A~~AAB~~A~~AB~~A~~A...

...AB AB AAB AB AAB AB A...

...ABABAABABAABABA...

Figure 37. An example of removing one A between each pair of Bs for part of an infinite cutting sequence.

Let's nail down some proofs of these results, which we have previously conjectured:

49. Show that a trajectory with slope p/q (in lowest terms) on the square billiard table has period $2(p+q)$.

50. Show that the continued fraction expansion of a number terminates (stops) if and only if the number is rational.

We will now begin our study of the multitude of interesting surfaces other than the square torus. Here we go:

51. If we identify opposite parallel edges of a hexagon, what surface do we get? Let's explore this question:

(a) Figure 38 shows one way to figure it out: a hexagon surface is *cut-and-paste equivalent* to a parallelogram surface. This means that you can cut up the pieces of a hexagon surface and reassemble them, respecting the edge identifications, into a parallelogram whose opposite parallel edges are also identified. Explain, and check that the steps in the picture respect the edge identifications.

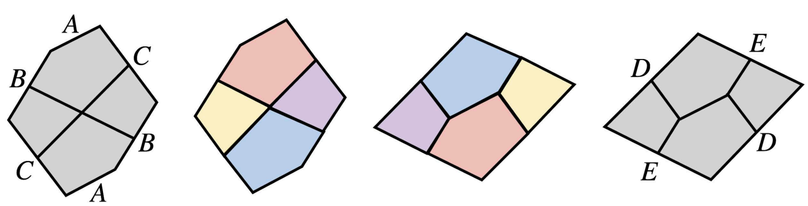


Figure 38. Cutting, reassembling, and pasting a hexagon surface into a parallelogram surface, while respecting the edge identifications.

(b) An alternative approach is to sketch a "movie" of what it looks like to glue identified edges together, assuming that the hexagon is made out of stretchy material. Try this, too.

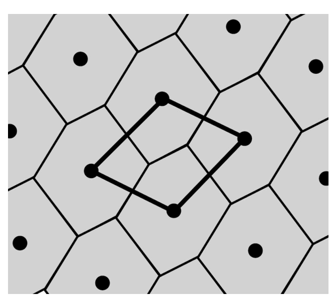


Figure 39. The "random" hexagons from Figure 38 tile the plane by translation. Does this always work?

52. Figure 38 shows a "random" hexagon with three pairs of parallel edges. This hexagon tiles the plane, as shown in Figure 39.

(a) Does a hexagon with three pairs of parallel edges *always* tile the plane? What if you require that, as in Figure 38, the parallel sides are opposite each other?

(b) A polygon is *convex* if each of its angles is less than 180°, or equivalently if every line segment connecting two points of the polygon lies completely within the polygon. Does a non-convex hexagon with three pairs of parallel edges always tile the plane?

They did the math # 11. Maryam Mirzakhani

In Problem 51, we created a surface from an arbitrary hexagon that has three pairs of opposite parallel sides. We could consider the *space* of all possible hexagons, or the space of all of the *surfaces* created by identifying the opposite parallel sides of such hexagons. You might expect that the surface created from a regular hexagon, or other special cases of hexagons, would appear in an identifiable place in the space, and indeed the symmetries of the *surfaces* help us to understand the symmetries of the *space* of surfaces. Maryam Mirzakhani (They did the math # 11) studied spaces of surfaces, and their symmetries [**12**, **36**]. She received the Fields Medal in 2014 and died in 2017.

13. A grand unifying theory emerges

53. Starting with a trajectory on the square torus with positive slope, apply the following algorithm:

(1) If the slope is ≥ 1, apply the shear $\begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$.

(2) If the slope is between 0 and 1, apply the flip $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

(3) If the slope is 0, stop.

An example is shown in Figure 40. (The second line is a continuation of the first.)

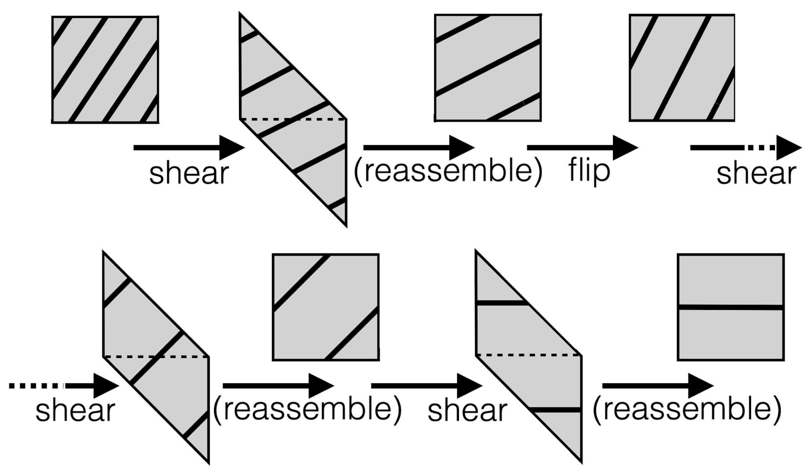


Figure 40. Untwisting a trajectory until it is horizontal.

We can note down the steps we took: shear, flip, shear, shear. We ended with a slope of 0. Work backwards, using this information and your work in Problems 47 and 48, to determine the slope of the initial trajectory. Keep track of each step.

54. (Continuation) Write down the continued fraction expansion for the slope at each step.

55. (Continuation) Write down the cutting sequence for the trajectory at each step.[3]

[3]To see this sort of thing in action on the double pentagon surface, see the video "Cutting Sequences on the Double Pentagon, explained through dance" on Vimeo: `https://vimeo.com/47049144`

56. (Continuation) Formulate a grand unifying theory relating a trajectory on the square torus, its corresponding cutting sequence, and the continued fraction expansion of its slope.

57. We can create a surface by identifying opposite parallel edges of a single polygon, as we have done with the square and hexagon. We'll call such a surface a *translation surface*, since parallel edges are translates of each other, and you can translate the polygon to identify the edges. *Parallel edges* must be parallel and also the same length. *Opposite edges* means that the polygon is on the left side of one of the identified edges, and on the right side of the other.

In a similar way, we can create a surface from two polygons, or from any number of polygons. Some examples are in Figure 41. Edges with the same letter are identified, as with A and B on the square torus. For the surfaces in the middle and on the right, *two* polygons glued together form a single surface.

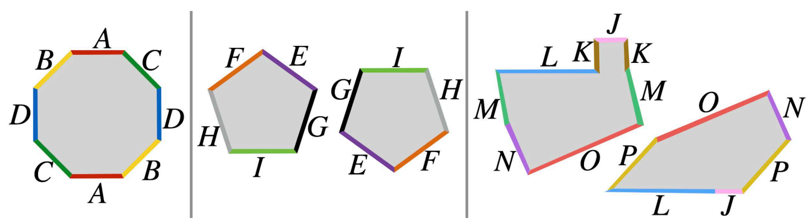


Figure 41. A menagerie of translation surfaces: our newest friends.

(a) Review the part of Amie Wilkinson's talk[4] from 26 to 29 minutes, which shows how to wrap the flat octagon surface (far left) into a curvy surface embedded in 3-space. What is its *genus* – how many holes does it have?

(b) Do your best to repeat her stretching methods for the double pentagon surface (center) to make it into a curvy surface embedded in 3-space.

(c) The octagon surface has 4 edges: A, B, C, and D. How many edges do the other surfaces have?

[4]YouTube: "Dr. Amie Wilkinson - Public Opening of the Fields Symposium 2018," available at `https://www.youtube.com/watch?v=zjccKzHIniw&t=1560s`

Some people love translation surfaces, and other people *really* love translation surfaces. Jayadev Athreya (They did the math # 12) has made many contributions to the field [**5**, **6**, **30**], but his most unique contribution just might be having a double pentagon tattooed on his forearm. The left picture shows the author with Jayadev in Marseille in 2015.

They did the math # 12. Jayadev Athreya

14. We expand from familiar friends to new examples

58. In Problem 53, we gave an algorithm that gradually simplifies a trajectory on the square torus with slope ≥ 1, by untwisting it step by step, until it is a horizontal trajectory. Transform that algorithm into an equivalent algorithm for the *cutting sequence* corresponding to a trajectory. You should translate each of the four sentences ("Starting with...," $1, 2$, and 3) to act purely on sequences of As and Bs. Then apply your algorithm to the cutting sequence $\overline{ABAAB}$ and check that your result at each step is consistent with the pictures in Figure 42.

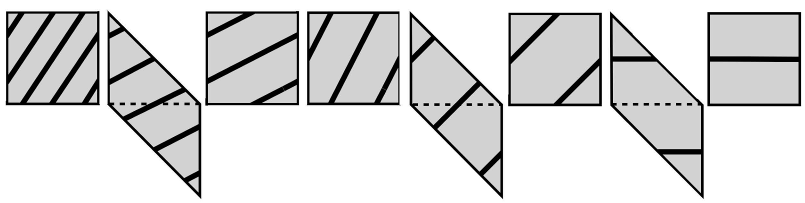


Figure 42. Untwisting a periodic trajectory via shears and flips.

59. Create a translation surface (recall Problem 57) made from *three* polygons, that no one else will think of. How many edges does your surface have?

To count the *faces* of a translation surface, we count how many polygons it's made of. To count its *edges*, it might be easiest to count the edge *labels*, remembering that pairs of opposite parallel edges are identified. Finally, we need to know how to count its *vertices*, which again requires understanding the edge identifications:

60. *Vertex chasing.* To explain how to count the vertices of a surface, we will use the square torus in Figure 43. First, mark any vertex (here, the top left). We want to see which other vertices are the same as this one. The marked vertex is at the left end of edge A, so we also mark the left end of the bottom edge A. We can see that the top and bottom ends of edge B on the left are now both marked, so we mark the top and bottom ends of edge B on the right, as well. Now all of the vertices are marked, so the square torus has just one vertex. (We already knew that – how?)

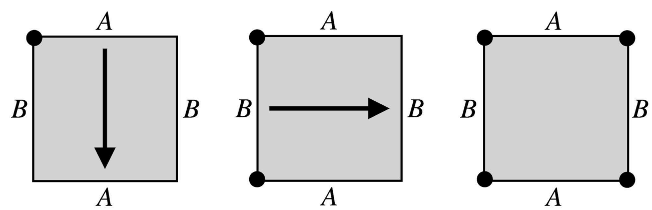


Figure 43. Chasing a vertex on the square torus.

Determine the number of vertices for

(a) a hexagon with opposite parallel edges identified (Problem 51);

(b) each surface in Problem 57; and

(c) your surface created in Problem 59.

61. Let P be a convex quadrilateral that has a 4-periodic inner billiard trajectory that reflects consecutively in all four sides. Prove that P is cyclic: there is a circle containing all four of its vertices. *Hint:* A quadrilateral is cyclic if and only if opposite angles sum to π.

They did the math # 13. Katherine Knox

Is the converse true – given a cyclic quadrilateral, *must it* have a 4-periodic inner billiard trajectory that reflects consecutively in all four sides? Katherine Knox (THEY DID THE MATH # 13), a 7^{th}-grade student participating in the Girls' Angle program in Boston, showed that the answer is no. She proved that a convex quadrilateral has a 4-periodic inner billiard trajectory that reflects consecutively in all four sides if and only if the polygon is cyclic *and* the quadrilateral contains its circumscribing circle's center [**31**].

62. You will need: colored pencils or pens. Consider the counter-clockwise outer billiard map on the *triangular* billiard table, as shown in Figure 44.

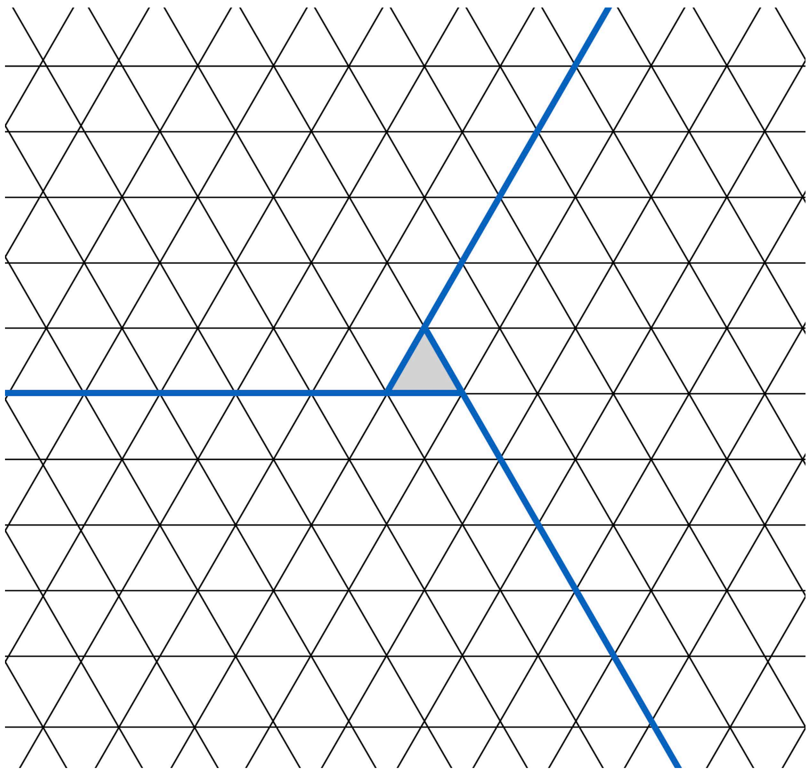


Figure 44. A template for understanding co-moving regions under the outer billiard map on the equilateral triangle.

(a) Explain why points on the thick blue lines are not allowed. Then color the inverse images of the blue lines in red, the inverse images of the red lines in green, the inverse images of the green lines in black, the inverse images of the black lines in purple, etc. (If you use these specified colors, you will be able to check your work with others.)

(b) Identify some *necklaces* of iterated images of triangles, and color each necklace a different color, as we did in Problem 35.

(c) I said to consider a *triangular* billiard triangle, but Figure 44 shows a very special case: an *equilateral* triangle. It turns out that the outer billiards system is "invariant under affine transformations." This means that if you have a picture of an outer billiard orbit, and then you apply a 2×2 matrix transformation (such as $\left[\begin{smallmatrix} 1 & 0 \\ -1 & 1 \end{smallmatrix}\right]$) to the whole system, the orbit is still valid. Thus, once we've understood one triangle, we've understood them all. This is a special property of outer billiards that does not hold for inner billiards. Explain.

This concludes our study of outer billiards. The tables we considered (circle, square, triangle) are quite simple, and the resulting behavior is also simple. Figure 45 shows more exotic necklaces for the outer billiard map on the regular pentagon.

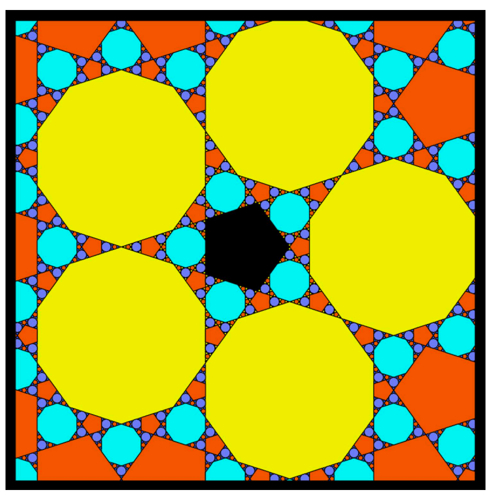


Figure 45. Necklaces for the outer billiard map on the central regular pentagon. I see pentagons, decagons, and... fractals?!

15. We figure out how to ignore trajectories completely

63. Apply the geometric algorithm from Problems 53 and 58 to the trajectory shown in Figure 46, to reduce it to slope 0. Note down the steps you take (shears and flips). Then use this information to work backwards from an ending slope of 0 to determine the slope of the initial trajectory. Show all of your steps.

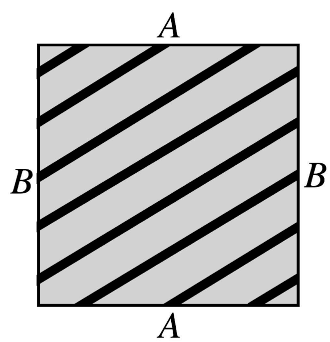


Figure 46. A mystery slope?! Continued fractions to the rescue!

64. (Continuation) Explain how shears and flips on the square torus are related to continued fraction expansions.

65. (Continuation) Find the cutting sequence corresponding to the trajectory above. Apply your algorithm from Problem 58 to it, and check that your results at each step are consistent with each step of your work in Problem 63.

The following problem is, at long last, the payoff for all of our work with continued fractions, shears, flips, and cutting sequences:

66. Find the cutting sequence corresponding to a trajectory on the square torus whose slope has continued fraction expansion $[0; 1, 2, 2]$. *Hint*: you don't need pictures; just use your algorithm and the grand unifying theory (Problem 56).

Problem 66 is an example of abstracting all the way away from trajectories to working with only continued fractions and symbolic cutting sequences. Curt McMullen (They did the math # 14) is a master of plumbing the depths of abstraction in billiards and related areas. He received the Fields Medal in 1998. He was also Maryam Mirzakhani's Ph.D. advisor (# 11). The picture shows Curt sailing with the author in Boston in 2018.

They did the math # 14. Curtis McMullen

Once we've made a surface, the *Euler characteristic* gives us a way of easily determining what kind of surface we obtain, without needing to come up with a clever trick like cutting up and reassembling hexagons into parallelograms (as we did in Problem 51):

Given a surface S made by identifying edges of polygons, with V vertices, E edges, and F faces, its Euler characteristic $\chi(S)$ is[5]

$$\chi(S) = V - E + F.$$

Note that a "face" must be a *simply connected* polygon, without holes.

67. Find the Euler characteristic of each of the surfaces in Figure 47. Comment on any patterns you notice. Can you prove your conjectures?

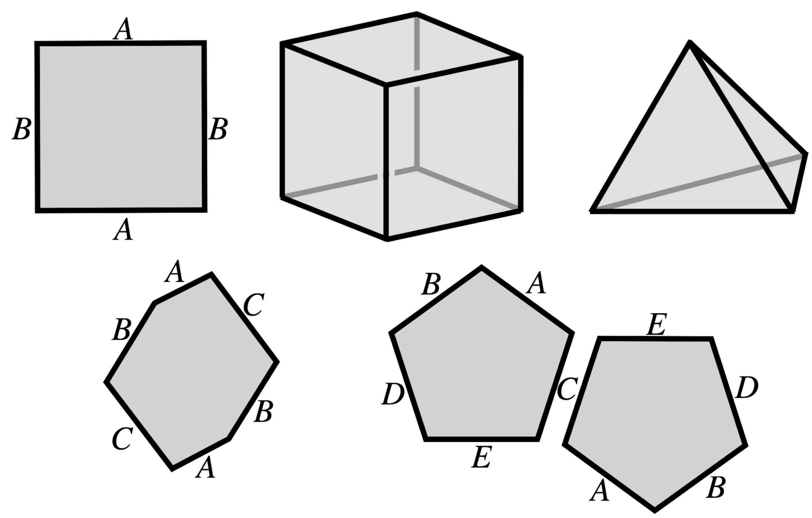


Figure 47. The square torus, the cube, the tetrahedron, the hexagon, and the double pentagon.

68. (Continuation) One of the main goals of the field of *topology* is to classify surfaces by their *genus*, which, informally speaking, is the number of "holes" they have. The surfaces in Figure 48 have genus $1, 2$, and 3, respectively.

[5] *Euler* is pronounced "oiler." χ is spelled *chi* and is pronounced "kye."

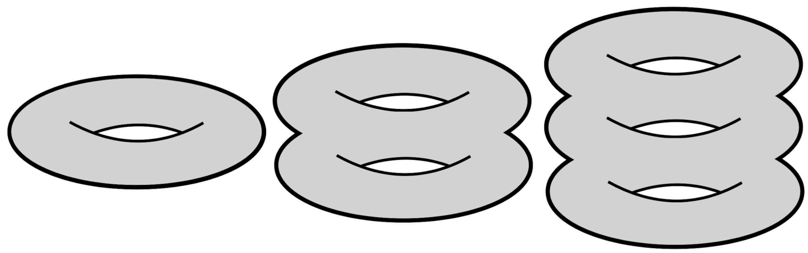


Figure 48. Surfaces with genus 1, 2, and 3, respectively.

We can use the Euler characteristic to determine the genus of a surface: A surface S with genus g has Euler characteristic $\chi(S) = 2 - 2g$. Use this to compute the genus of each of your surfaces from the previous problem, and check that your answers agree with reality.

69. (Challenge) Prove the formula $\chi(S) = 2 - 2g$. One way is to proceed by induction: First, show that $\chi(S) = 2$ for the tetrahedron or some other simplest surface of your choice (base case). Then, show that subdividing by adding a vertex, edge or face maintains the same Euler characteristic. Finally, show that adding a hole decreases the Euler characteristic by 2. (Other methods are also possible.)

16. Hands-on activities for Chapter 2

As Chapter 2 comes to a close, we will bring the flat torus to life in three (!) dimensions.

One way to make a model of a torus from a piece of paper is as follows: Tape the left side to the right side, creating a tube. Then wrap it around to attach the bottom edge to the top edge. To do this, you'll have to flatten the tube. The resulting object looks like a wide bracelet or a paper wallet. It is not very satisfying; the volume inside the torus is zero.

They did the math # 15. Alba Málaga Sabogal

Many people believe that the above description is the *only* way to create a torus that is flat everywhere – that is, it has 2π of angle around every point – out of a piece of paper. But it turns out that we can do better! Along with Pierre Arnoux (# 32) and Samuel Lelièvre (# 30), Alba Málaga Sabogal (They did the math # 15) created the *diplotorus*: a flat torus that encloses a positive volume [**3**] (Figure 49). The picture shows Pierre, Alba and Samuel in Marseille in 2023.

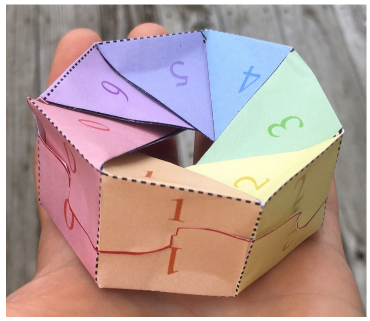


Figure 49. A flat torus, with 2π of angle around every point, enclosing a positive volume. Can you believe it? It's magical!

The idea has been around for a while, and the path it took to get to Pierre, Alba and Samuel was a long and winding road:

- Ulrich Brehm explained how to construct such an object during a talk at Oberwolfach in 1978.
- Then in 1984, Geoffrey Shephard gave another talk about it at Oberwolfach, and brought a model.
- Guy Valette was in the audience for that talk, and made a model of his own when he got home.
- Guy told Robert Ferréol about it, and Robert put it on his web site, where Henry Segerman saw it.
- Henry made a 3D-printed version [**49**], which Pierre, Samuel, Alba, and the author saw at ICERM in 2019. Glen Whitney also brought a paper model of such a torus to ICERM.
- Finally, the idea has made it to *you*!

70. You will need: scissors, perseverance. The picture at the end of this chapter shows a diplotorus layout, which you can print (two-sided!) from the book web page. Cut it out, and then crease it along the indicated lines: the dashed lines should be mountain folds, and the solid lines valley folds. (*Hint*: If you spend a long time making very strong creases on all of the lines, putting the model together will be doable; if your creases are weak or inaccurate, it will be almost impossible.) Then bring the edges with the same numbers and colors together, and attach the flaps via the red slits. The two white disks should coincide at the same point, and the object should look like Figure 49. Behold, a torus that encloses positive volume, and is *flat* (2π of angle around each point) everywhere!

71. Switch the mountain and valley folds, so that the other side of the paper shows. Behold, a closed path on a flat torus (Figure 50)!

Figure 50. A closed path on a truly 3D, truly flat torus.

72. The diplotorus layout looks roughly like a parallelogram, and when you fold it up, you bring the short edges together, and you bring the long edges together. But you do not identify the numbered edges directly across; there is a *twist*. Using Figure 51, show how to cut and paste the diplotorus translation surface into a parallelogram whose opposite parallel edges are identified. (There is more than one correct answer; see Problem 153.)

The famous *Four-Color Theorem* tells us that if you want to color a geographic map in the plane so that regions meeting along an edge are always different colors, you only need at most four colors. It is possible for four regions to all border each other – e.g. Luxembourg, Belgium, France, Germany, as in Figure 52 – but not five.

73. For the torus, the number is seven: given any map on the torus, you need at most seven colors. The coloring on the diplotorus layouts

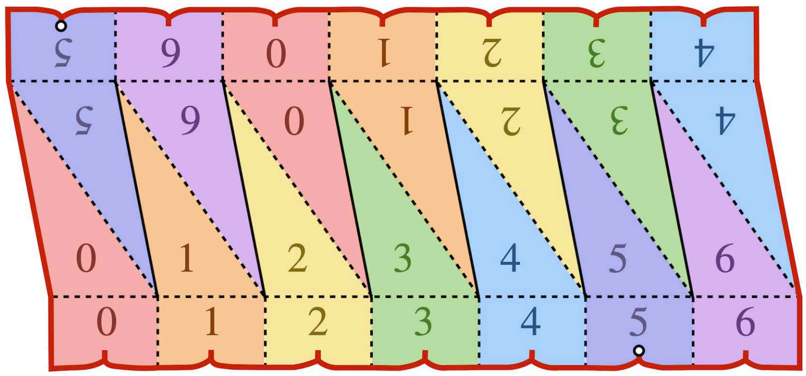


Figure 51. A tiny diplotorus layout, for making into a parallelogram.

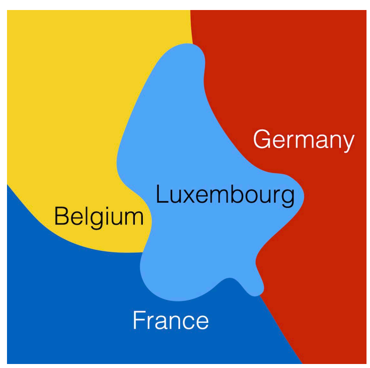


Figure 52. Four countries, four colors needed.

gives an example of seven mutually adjacent regions.[6] Looking at the flat layout in Figure 51 or on the next page, check that the orange region 1 touches regions $0, 2, 3, 4, 5, 6$. Argue that the same is true for each of the other colors.

[6]Thanks to Moira Chas for suggesting this line of inquiry to Samuel, who suggested it to me.

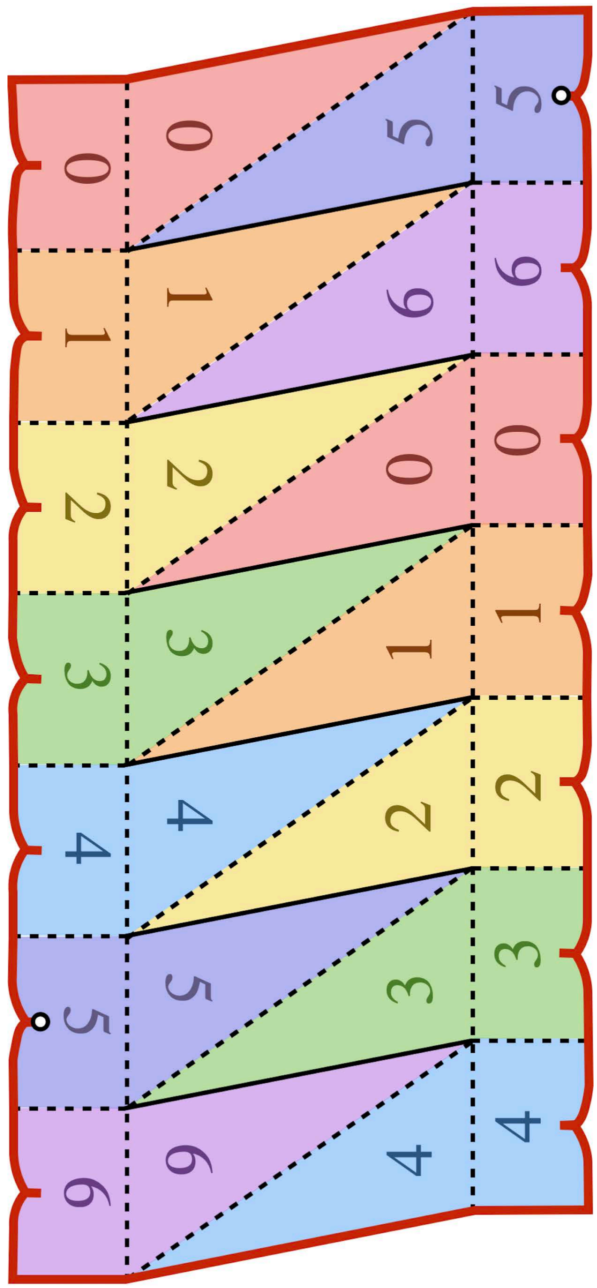


Figure 53. Print out a copy of this picture from the book web site and cut it out. Make mountain folds on the dashed lines, make valley folds on the solid lines, and make your best effort to twist it around so that the white dots coincide. Fun fact: once re-flattened, a diplotorus is completely portable!

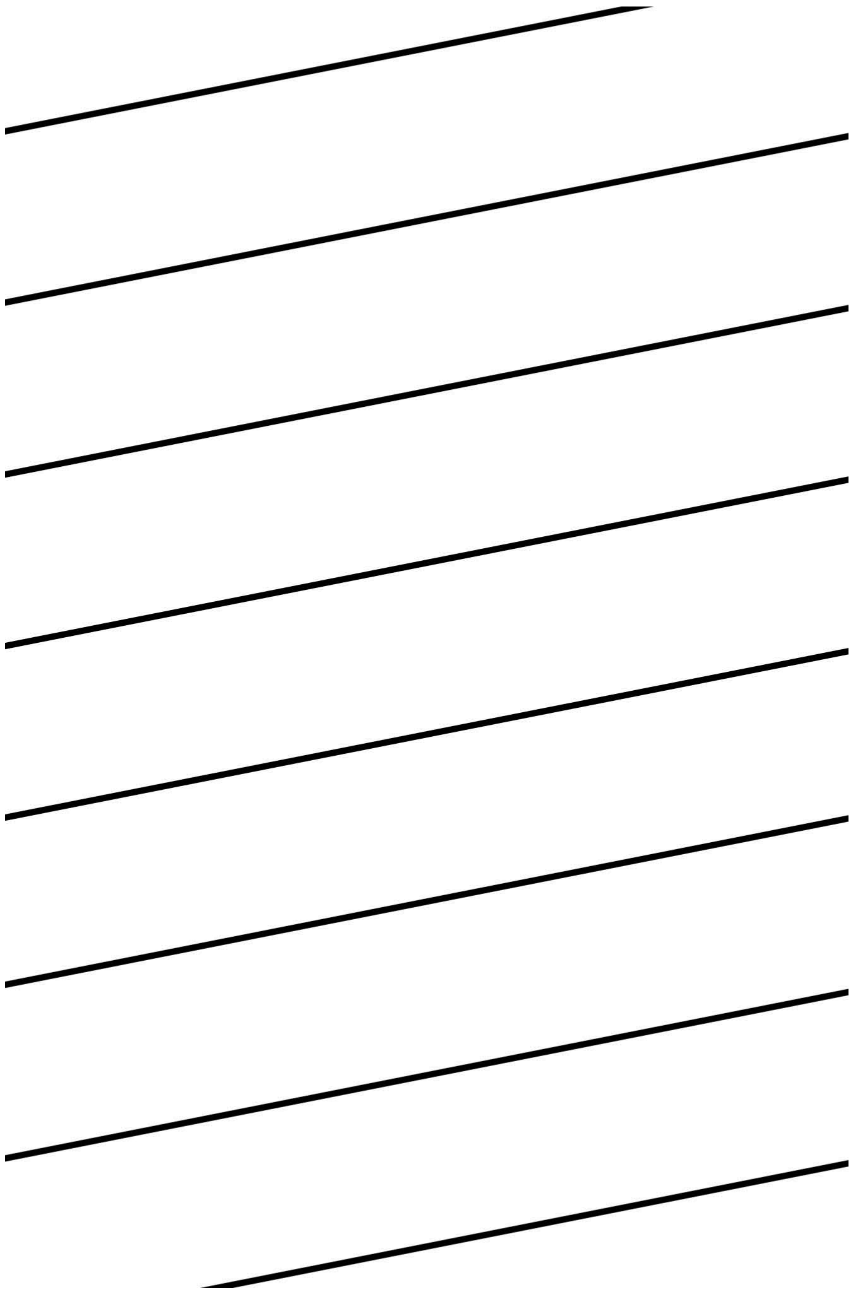
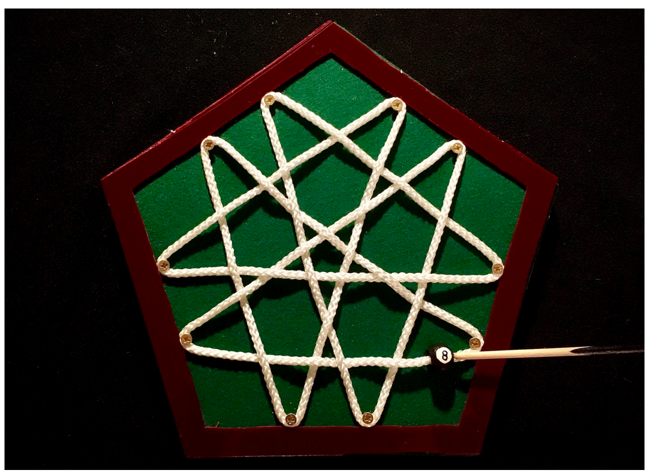


Figure 54. Printed on the back side of the diplotorus, this page yields a periodic trajectory wrapping around the flat torus. Fun fact: this picture has to be exactly the right size in order to work, but its placement on the page is not important. Can you see why? Periodicity for the win!

Chapter 3

Periodicity everywhere



A billiard table trajectory in the style of a Celtic knot

In Chapter 1, we met our main characters; in Chapter 2, we delved deeply into the structure of periodic directions on the square billiard table. We saw how the slopes of trajectories on the square billiard table are connected to continued fractions, and to automorphisms of the square torus surface. In so doing, we saw links between billiards, number theory, and group theory.

In Chapter 3, we will use the tools and insights we gained in our study of the square, as we move further afield to many different types of billiards. We will do billiards on triangular tables, and we will meet yet another kind of billiards. For billiards on non-square tables, the situation is often "the situation is analogous to the square, but not quite as elegant." For other types of billiards, the behavior is often not at all like the square. Exciting!

Here are a few guiding questions you can ask yourself when you meet a new situation in billiards, or in any dynamical system:

- Is a typical trajectory periodic or non-periodic?
- If I change my initial *point* a little bit, do things change a lot, or do they stay basically the same?
- If I change my initial *direction* a little bit, do things change a lot, or do they stay basically the same?
- Can I transform this problem into a situation that I already understand, or into one that is easier to analyze?

17. We meet tiling billiards

We have explored the simplest case of classical billiards – inner billiards on the square – in detail, and understood it deeply. We have explored outer billiards, and understood something about its behavior. Now we will expand our view to a third type of billiards:

Tiling billiards. In this system, a trajectory refracts through a tiling of the plane. The *refraction rule* is that when the trajectory hits an edge of the tiling, it passes through in such a way that the angle of incidence is equal to the angle of reflection, and the trajectory has been reflected across the edge (Figure 55).

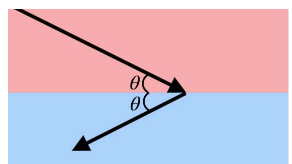


Figure 55. The tiling billiards map: refraction across an edge of a tiling.

74. Sketch some trajectories for the tiling in Figure 56. What kinds of behaviors can you find? Prove that you have found them all.[1]

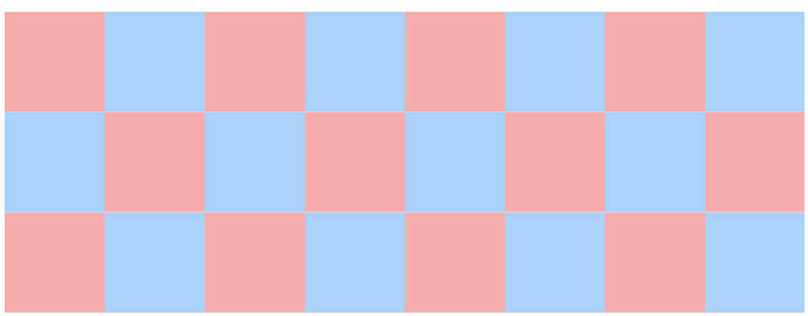
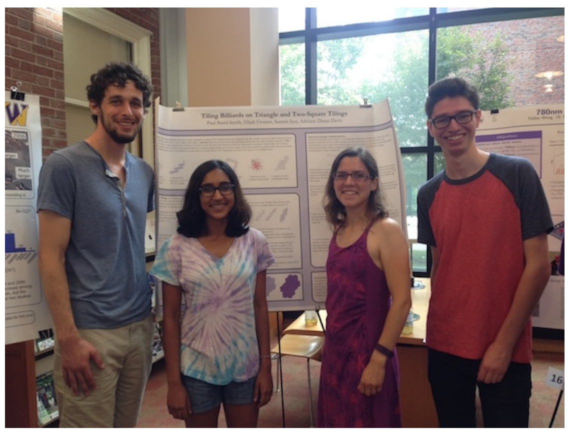


Figure 56. A square grid, for tiling billiards trajectories.

[1] For a beautifully artistic dynamic rendering of tiling billiards, and a preview of # 38, see the video "Refraction Tilings" by Ofir David on YouTube: `https://www.youtube.com/watch?v=t1r1cO1V35I`.

Tiling billiards is motivated by the existence of *metamaterials*, solids that have a negative index of refraction. Typical materials such as water and glass have a positive index of refraction; you have likely worked with these in physics, with *Snell's Law*. The idea here is to create a two-colorable tiling out of materials with opposite indices of refraction, and see what happens as a laser beam refracts around it. The first published results about tiling billiards came from the work of three undergraduate students, along with the author: Elijah Fromm, Sumun Iyer, and Paul Baird-Smith (THEY DID THE MATH # 16), shown with the author at their poster session in 2016 [**7**].



They did the math # 16. Elijah Fromm, Sumun Iyer, and Paul Baird-Smith

Walking around a vertex. We can determine the angle around a vertex by "walking around" it, as shown in Figure 57 for a hexagon surface. The left picture shows that the angle around the black vertex is $3 \cdot \frac{2\pi}{3}$, and the right picture shows the same for the white vertex.

To do this, first choose a vertex (say, the top-left vertex of the hexagon, between edges A and B, marked as black) and walk counter-clockwise around the vertex. In our example, we go from the top end of edge B to the left end of edge A. See that we "come out" on the identified edge A at the bottom of the hexagon, and keep going counter-clockwise: we go from the left end of the bottom edge A to the bottom end of the left edge C. We keep going counter-clockwise

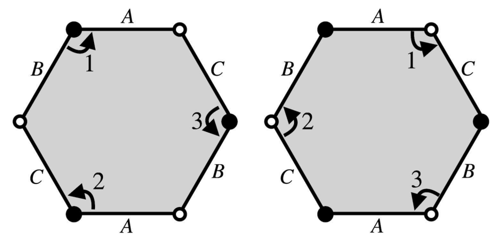


Figure 57. Walking around the black and white vertices.

from the bottom of the right edge C to the top end of the right edge B. We find the identified point on the top end of left edge B, and see that this is where we started! So the angle around the black vertex is $3 \cdot 2\pi/3 = 2\pi$. By the same method, or by symmetry, we can see that the angle around the white vertex is also 2π.

Since the black and white vertices each have 2π of angle around them, all the corners of a hexagon surface come together in a flat plane, as we have already seen in Problem 51 and Figure 39.

75. For each of the surfaces in Figure 58, count its vertices, and then determine the angle around each vertex.

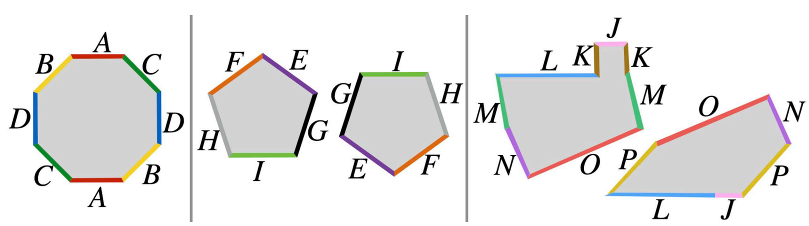


Figure 58. Our friendly menagerie of translation surfaces, coming back around for vertex counting and angle measuring.

76. As mentioned earlier, a surface is called *flat* if it looks like the flat plane everywhere, meaning that there is 2π of angle around every point, *except* possibly at finitely many *cone points* (also known as *singularities*), where the angle around each cone point is a multiple of 2π. For example, the regular octagon surface is flat everywhere except at its single cone point, whose angle is 6π. Prove that every translation surface (Problem 57) is flat.

77. (Challenge) Is the converse true? In other words, is it true that every flat surface can be represented by a collection of polygons, identified along opposite parallel edges? Prove it or find a counterexample.

78. *The Fagnano trajectory.* You have constructed several periodic billiard paths in the square billiard table; other polygons also have periodic paths. A classical theorem says that the *Fagnano trajectory* connecting the feet of the three altitudes of an acute triangle is a 3-periodic billiard trajectory (Figure 59). We will prove this by showing that angles ARP and CRQ are equal; the argument is the same for the other bounces.

(a) Opposite angles of a quadrilateral add up to π if and only if the quadrilateral is *cyclic*. Use this result to show that quadrilaterals $APOR$ and $CROQ$ are cyclic, as the diagram suggests.

(b) Another classic theorem of geometry says that two angles supporting the same circular arc are equal. Use this to show that $\measuredangle PAO = \measuredangle PRO$, and $\measuredangle ORQ = \measuredangle OCQ$.

(c) Use triangles BAQ and BCP to show that $\measuredangle PAO = \measuredangle OCQ$.

(d) Show that $\measuredangle ARP = \measuredangle CRQ$, as desired.

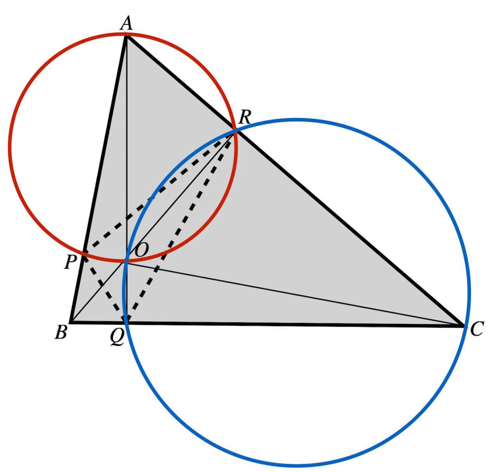


Figure 59. The Fagnano trajectory, with circles identifying two cyclic quadrilaterals.

An active area of research is to *characterize* all possible cutting sequences on a given surface. Now we can do this for the square torus.

Theorem (cutting sequence characterization). Cutting sequences on the square torus are infinite sequences of As and Bs that do not fail under the following algorithm:

(1) If there are multiple Bs separated by single As, switch As and Bs.

(2) If there are multiple As separated by single Bs, remove an A between each pair of Bs.

(3) If the sequence has AA somewhere and BB somewhere else, stop; it fails to be a valid cutting sequence.

79. Earlier in the course, you probably conjectured that a cutting sequence could only have two consecutive numbers of As, such as 2 and 3, between each pair of Bs, e.g., $\overline{BABAAA}$ is not allowed. Use the theorem to prove this conjecture true.

18. Earlier, we unfolded; now, we fold

80. In Problem 74, we saw that for tiling billiards on the square grid, there are only two types of trajectories: those that go to the opposite edge and zig-zag, and those that go to the adjacent edge and make a 4-periodic path. How many types of trajectories are there on the equilateral triangle grid (Figure 60)?

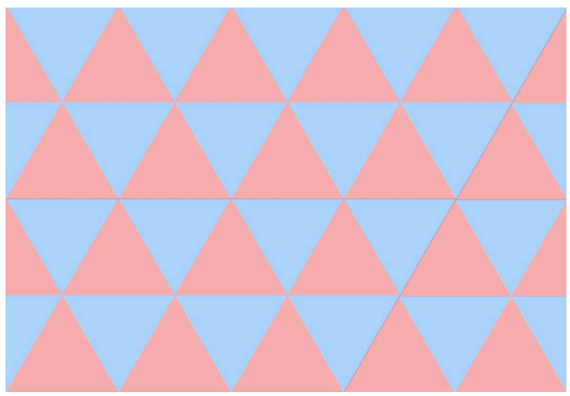


Figure 60. An equilateral triangle grid, for tiling billiards.

In billiards on the square, we *unfolded* a billiard trajectory into a line on the square grid, and into a linear trajectory on the square torus. In an analogous way, *folding* is a powerful technique for understanding tiling billiards trajectories:

81. Consider a tiling billiards trajectory that crosses an edge E of the tiling. Show that, if you fold the tiling along edge E, the two pieces of trajectory that intersect edge E lie on top of each other.

82. Recall the cutting sequence characterization theorem (Problem 79) for trajectories on the square torus.

(a) The vexing part of this characterization is that it doesn't have a step saying, "Stop! Congratulations; you have a valid cutting sequence." It only says, "Keep going; your cutting sequence hasn't proven to be invalid yet." But it turns out that it's the best we can do. Explain why this algorithm *does* stop for a *periodic* cutting sequence.

I left out one technical point of the theorem: It actually characterizes the *closure* of the space of all cutting sequences. Valid cutting sequences are in the interior of the space, and cutting sequences such as $\ldots AAAAABAAAAA\ldots$ are on the boundary of the space.

(b) Explain why the above cutting sequence does not fail in the algorithm, and also explain why it is nonetheless not a valid cutting sequence on the square torus.

(c) Another cutting sequence on the boundary is $\ldots BBBABBB\ldots$. Find yet another example of a cutting sequence on the boundary of the space of cutting sequences.

They did the math # 17. Alex Wright

The *space* of cutting sequences is a rather abstract notion, like the *space* of hexagon surfaces that we discussed in § 12. Typically, the first examples we would think of are on the interior of such a space, and degenerate cases are on the boundary of the space. Alex Wright (THEY DID THE MATH # 17) has studied spaces of translation surfaces, and their orbit closures [**63**]. This picture shows Rodrigo Treviño, the author, and Alex in the Frankfurt airport in 2014.

83. Consider again the 3-periodic Fagnano trajectory from Problem 78. Figure 61 shows a piece of a trajectory that is parallel to the one in the construction and nearby. Continue the new trajectory until it closes up. What is its period?

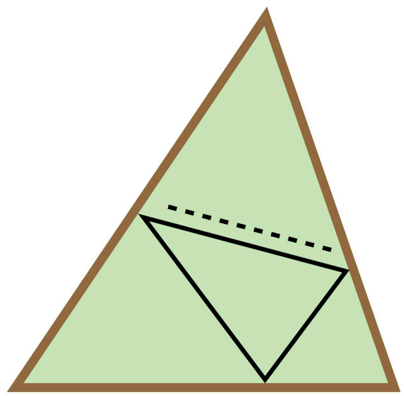


Figure 61. The Fagnano trajectory, and its parallel friend.

Notice that as you follow the dashed trajectory around, initially it says "the solid trajectory is on my right!" and then after a bounce, "the solid trajectory is on my left!" and so on, switching sides at every bounce.

84. In Chapter 2, our strategy for "untwisting" a periodic trajectory on the square torus (see Problem 53) was:

- If the slope is greater than 1, apply a vertical shear, and
- if the slope is less than 1, first flip so that the slope is greater than 1, and *then* apply a vertical shear.

Alternatively, we could say:

- If the slope is greater than 1, apply a vertical shear, and
- if the slope is less than 1, apply a *horizontal* shear.

Explain.

85. You will need: scissors, perseverance. The figure on the next page shows a periodic tiling billiards trajectory on a triangle tiling, which you can print from the book web site. Cut off the white part and then fold along all the edges of the tiling, in such a way that every part of the trajectory lies on a single line. The solid lines should be "valley folds" and the dashed lines should be "mountain folds." *Hint*: Spend a long time making very strong creases on all of the folds. If you have good strong creases everywhere, getting this thing to fold flat will be doable; if your creases are weak or inaccurate, it will be more difficult for you to make it happen.

Flat fold a little patch at first, and then gradually extend it to the whole paper. The result should look like Figure 62. Save your folded paper, as we will use it to prove things in subsequent problems, e.g., Problem 90.

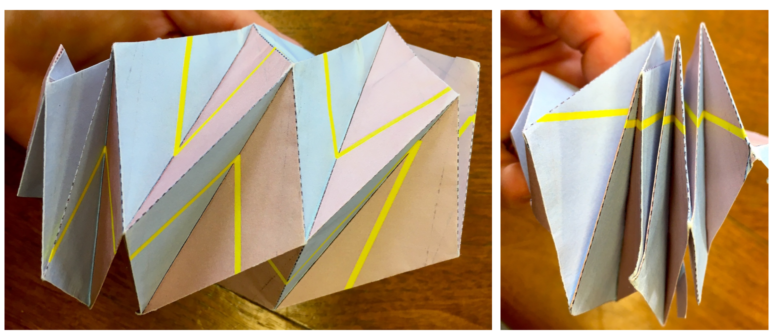


Figure 62. A fully folded tiling billiards trajectory on a triangle tiling: the desired result of Problem 85.

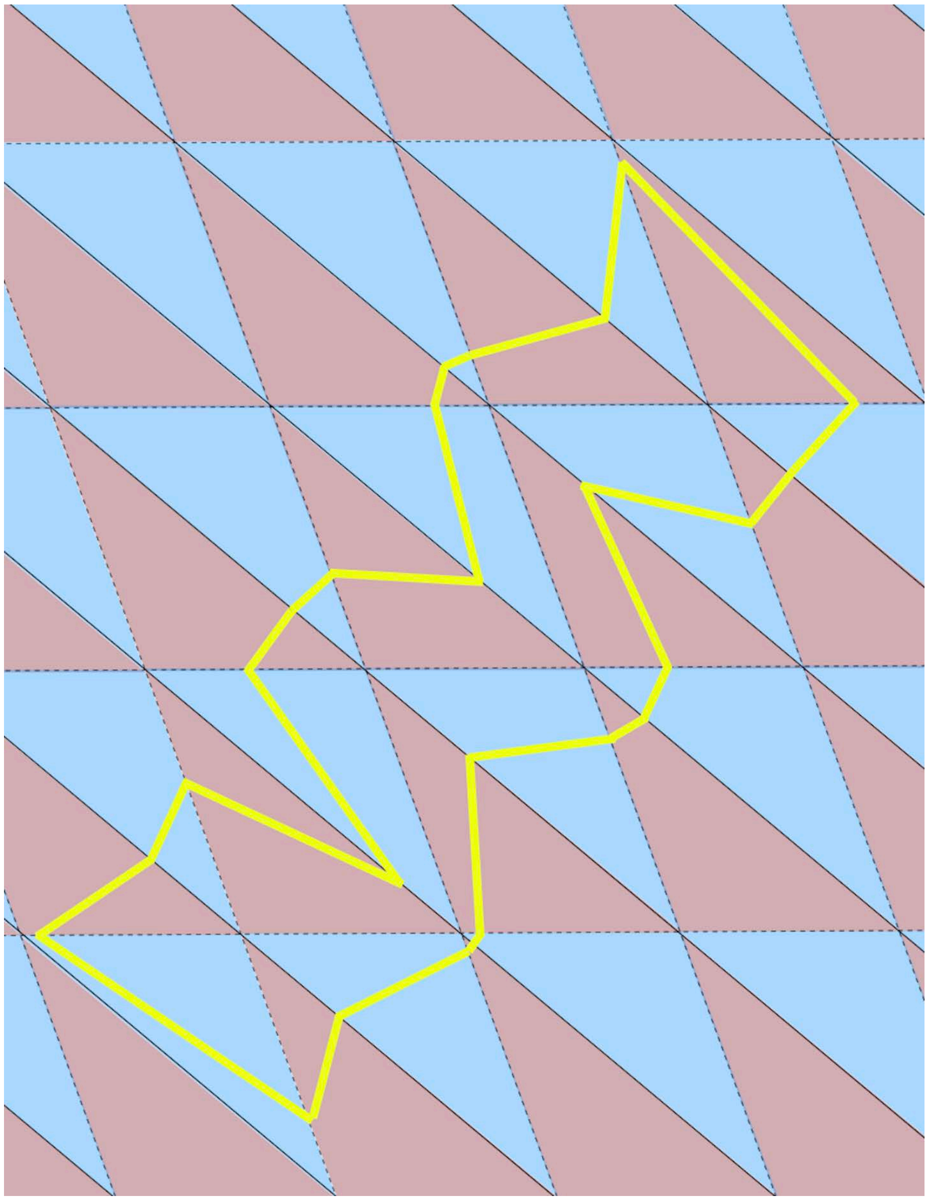


Figure 63. Print out a copy of this picture from the book web site and cut it out. Make mountain folds on the dashed lines, make valley folds on the solid lines, and make your best effort to flat fold the tiling along every edge. It is a challenge, but the payoff is huge in understanding!

19. We meet the biggest open problem in billiards

86. *Lots of triangles have periodic trajectories.*

(a) Explain why the Fagnano trajectory (Problem 78) gives a periodic trajectory in every acute triangle, and only in acute triangles.

(b) Rich Schwartz (# 3) showed me the construction in Figure 64. He calls it "shooting into the corner." Fill in the details, and show that it gives a periodic trajectory for every right triangle.

(c) Find an example of a periodic trajectory in an obtuse triangle.

(d) In fact, the Fagnano trajectory, the shooting into the corner trajectory, and the period-4 solution to **(c)** are all variations on the exact same idea. Explain.[2]

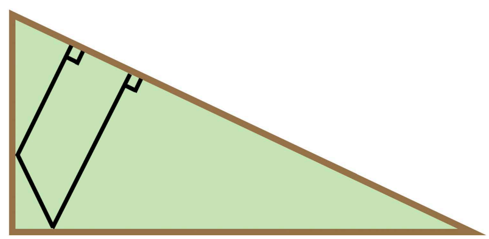


Figure 64. The "shooting into the corner" trajectory.

The biggest open problem in billiards is: *does every triangular billiard table have a periodic trajectory?* The Fagnano trajectory shows that every *acute* triangle has a periodic billiard trajectory, and the "shooting into the corner" construction shows that every *right* triangle has one.

Howard Masur (They did the math # 18) showed that every polygon – and thus every triangle – whose angles are *rational* numbers of degrees has a periodic path [**34**]. Rich Schwartz (# 3) used a computer-aided proof to show that every triangle whose largest angle is less than 100° has a periodic billiard trajectory [**48**], and in 2018 George Tokarsky, Jacob Garber, Boyan Marinov, and Kenneth Moore

[2] Thanks to Alan Bu for pointing this out.

They did the math # 18. Howard Masur

extended that result to 112.3° [**58**]. The problem is open in general for irrational-angled obtuse triangles with an angle larger than 112.3°. It seems that the methods of proof used for the 100° and 112.3° theorems do not work past about 112.5°, so a new idea is needed to move forward.

We have talked a little bit about the space of all possible translation surfaces of a given type. The space is divided into *strata* based on:

- how many cone points the surface has, and
- how many extra multiples of 2π are around each cone point.

(Recall that in Problem 76, we proved that the angle at a cone point of a translation surface is always a multiple of 2π.) We say that the double pentagon surface is in the stratum $\mathcal{H}(2)$ because it has one cone point, with two extra multiples of 2π around it: 6π total, so $2 \cdot 2\pi$ extra. A surface with two cone points, each with angle 4π, is in the stratum $\mathcal{H}(1,1)$.[3] The "$\mathcal{H}$" stands for "holomorphic."

87. For each of the remaining surfaces in Figure 65, identify which stratum it belongs to. Then come up with an example of a surface in $\mathcal{H}(1,1)$.

[3]We read this aloud as "H one one."

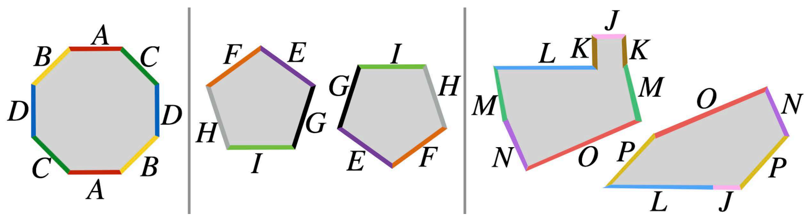


Figure 65. A menagerie of translation surfaces.

Note that a vertex with 2π of angle around it is not really a cone point; we can call it a *marked point* or a *removable singularity*. Depending on how much you care about such points, you can include 0s in your stratum, or not. For example, while we could say that the square torus is in $\mathcal{H}(0)$, we could alternatively note that it doesn't really have any cone points.

88. Here is our dream: to understand the effect of *every* automorphism of the square torus, on the cutting sequence corresponding to a trajectory.

(a) Here is our progress so far (fill in the blanks):

(1) There are three types of automorphisms: rotations, reflections, and shears. We understood the effects of rotations and reflections in Problems __________.

(2) Using rotations and reflections, we reduced our work, now only for shears, to the case of trajectories whose slope is greater than 1, in Problem ______.

(3) We understood the effect of the matrix $\begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$ on a trajectory on the square torus in Problems __________.

By the way, we used $\begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$ because it works nicely with trajectories whose slope is greater than 1: it makes them simpler, a bit like taking the derivative of a polynomial, while $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ makes them more complicated, a bit like taking an integral.

(b) Find the analogous effects on slopes of trajectories, of the matrices $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$, $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$.

89. (Challenge) There is just one more step, to show that every shear can be reduced to the ones we understand. Prove the following:

(4) Every 2×2 matrix with nonnegative integer entries and determinant 1 is a product of powers of the shears $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$.

For example, given the matrix $\begin{bmatrix} 3 & 7 \\ 2 & 5 \end{bmatrix}$, we can decompose it as

$$\begin{bmatrix} 3 & 7 \\ 2 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}^2 \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^2 .$$

Once we have proven this last step, we will be able to say that we know the effect of every matrix with determinant 1 on slopes of trajectories, and we could work out the induced effects of $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$ on cutting sequences just as we did for $\begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$.

90. You will need: your folded triangles from Problem 85. Consider a tiling by congruent triangles, created from a tiling by edge-to-edge parallelograms by splitting the parallelograms along parallel diagonals, such as the one that you folded up in Problem 85.

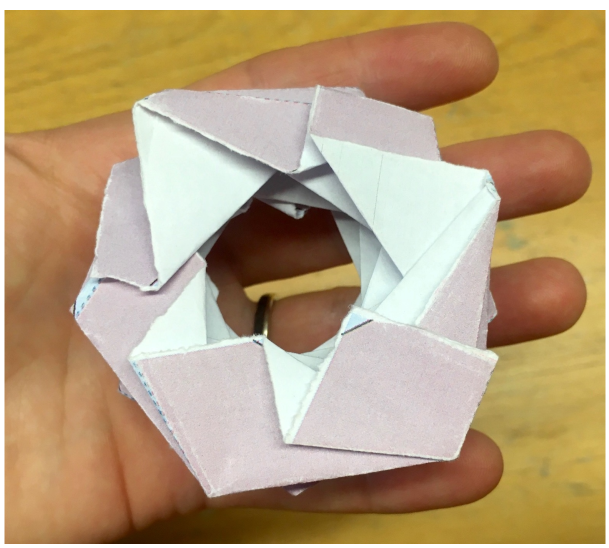


Figure 66. A fully folded tiling billiards trajectory, whose form strongly suggests the presence of a circumscribing circle.

(a) Given two adjacent triangles in the tiling, prove that, if you fold along their shared edge, the circumcenters of the triangles coincide, and thus the two triangles share the same circumscribing circle.

(b) Prove that this result extends globally: if you fold along *all* of the edges of the tiling simultaneously, *all* the triangles, in the folded state, are circumscribed in the same circle (see Figure 66).

(c) Use the above, and the result of Problem 81, to show that for a given tiling billiards trajectory on a triangle tiling, in the folded state, all the pieces of trajectory are contained in a single chord of the circumscribing circle.

20. Families of parallel trajectories

91. Figure 67 shows the Fagnano trajectory in the 40-60-80 triangle (you can print a copy from the book web site). In Problem 83, we showed that there are nearby parallel billiard trajectories of period 6.

(a) In the triangle in the lower right, sketch a period-6 trajectory that is parallel to the given Fagnano trajectory.

(b) How far can you push the period-6 trajectory until it disappears? Add to your picture a period-6 trajectory that is as far as you can make it from the given Fagnano trajectory.

(c) Sketch one of your period-6 trajectories in the shaded triangle that is in the lower left of the picture. Then draw the "unfolding" of your trajectory. The unfolding of the Fagnano trajectory is given.

(d) Imagine the family of *all* possible period-6 trajectories that are parallel to the Fagnano trajectory. Can you sketch *all* of their unfoldings in the picture?

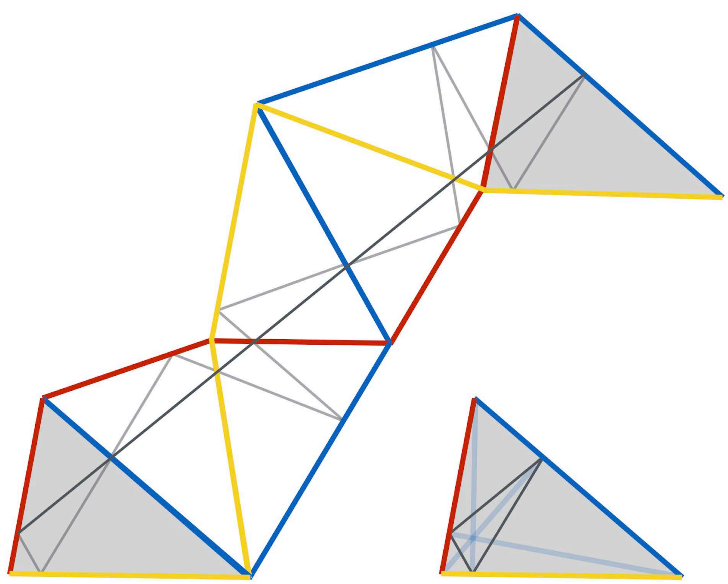


Figure 67. The Fagnano trajectory, and its unfolding.

92. Consider a circle broken into a red arc and a blue arc, taking up 1/3 and 2/3 of the circle respectively, as in Figure 68. The game is to start with any point on the circle, repeatedly rotate it by a 1/3 turn, and each time note down which part of the circle it lands in – say, an A if it lands in the red arc and a B if it lands in the blue arc. Try this for several different starting points, and rotate each of them until you see a pattern.

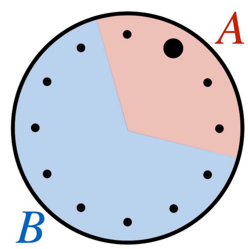


Figure 68. A simple circle rotation.

93. One reason why people like cutting sequences on the square torus is that they have very low *complexity*: The *complexity function* $f(n)$ on a sequence is the number of different "words" of length n in the sequence. In other words, imagine that you have a "window" n letters wide that you slide along the sequence, and you count how many different words appear in the window.

(a) Confirm that the sequence $\overline{ABABB}$ below has complexity $f(n) = n+1$ for $n = 1, 2, 3, 4$ and complexity $f(n) = 5$ for $n \geq 5$.

$\ldots ABABBABABBABABBABABBABABBABABBABABB \ldots$

(b) Explain why a periodic cutting sequence on the square torus with period p has complexity $f(n) = n + 1$ for $n < p$ and complexity $f(n) = p$ for $n \geq p$.

(c) (Challenge) Aperiodic sequences on the square torus are called *Sturmian sequences*. Show that Sturmian sequences have complexity $f(n) = n + 1$.

94. The *defect* of a cone point is 2π minus the cone angle at the cone point. The *total defect* of a surface (or of any polyhedron made from identifying edges of polygons) is the sum of the defects of all of its cone points. A theorem of Descartes says that the total defect of a polyhedron is $2\pi\ \chi(S)$. Check this formula for the cube, the square torus, and the octagon surface, using your answers to Problem 67.

They did the math # 19. William Thurston

Bill Thurston (They did the math # 19) was hugely influential in 20$^{\text{th}}$-century mathematics, particularly in geometry. In addition to his own work, he was the Ph.D. advisor, and the advisor's advisor ("academic grandfather"), of many mathematicians currently working in billiards and related fields. Bill received the Fields Medal in 1982 and died in 2012. One of his later projects was working with Kelly Delp to smooth out the angle defect in polyhedra [**18**]. The picture shows Bill and Kelly at a workshop on mathematics and fashion at Cornell in 2010.

95. Figure 69 shows many trajectories of slope 1/2 on the square torus. The collection of all such parallel trajectories is known as a *foliation* of the surface, "foliation" meaning "leaves." As usual, we care about when a given trajectory crosses a horizontal or vertical edge, and we record such crossings with an A or B, respectively. In this picture, I've added a diagonal of the square, and colored it on both sides: on the bottom to indicate whether an incoming trajectory comes from a red or blue side, and on the top to indicate whether an outgoing trajectory will hit a red or a blue side.

(a) Show how to use just the diagonal (copied larger below) to record the edge crossings of the highlighted trajectory, by translating the points as indicated.

(b) Explain why the dynamics of this system are identical to those of the rotation in Problem 92.

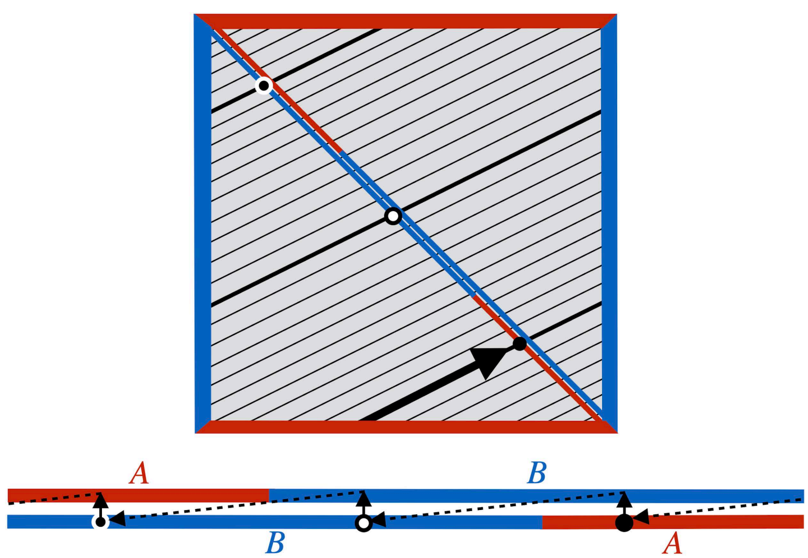


Figure 69. We can record the path taken by a flow on the square torus using its diagonal.

Above, we transformed a problem about trajectories on the square torus into a problem about moving intervals around on a line segment.

They did the math # 20. Jean-Christophe Yoccoz

This sort of system is called an *interval exchange transformation*. In joint work with Pierre Arnoux (# 32), Jean-Christophe Yoccoz (They did the math # 20) came up with the *Arnoux-Yoccoz interval exchange transformation*, which has interesting properties and led to much further research [**4**]. Jean-Christophe received the Fields Medal in 1994 and died in 2016. The picture shows Jean-Christophe with the author in 2014 at Oberwolfach.

21. Interval exchange transformations

96. Figure 70 shows two different tilings of the plane by isosceles right triangles. Consider tiling billiards on each of them.

(a) For each tiling, consider: are there periodic trajectories on the tiling? If so, explain how to construct one and sketch it; if not, prove that periodic trajectories cannot occur.

(b) An *escaping* trajectory eventually leaves a disk of any finite radius. Are there escaping trajectories on either tiling? If so, explain how to construct one and sketch it; if not, prove that escaping trajectories cannot occur.

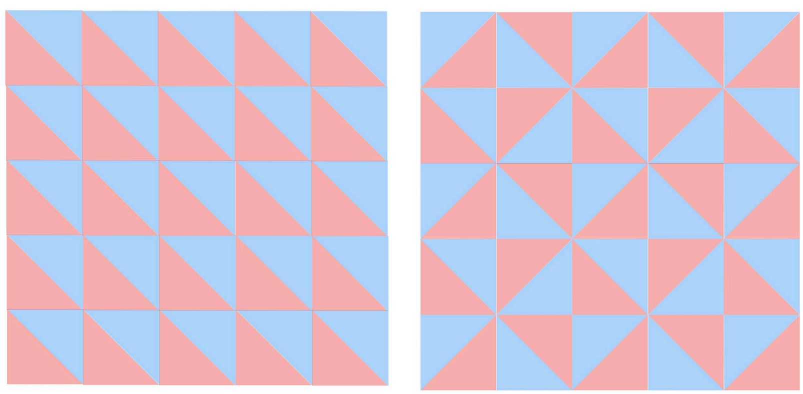


Figure 70. Two tilings by isosceles right triangles.

97. The construction in Problem 95 showed how to represent a trajectory on a surface via the motion of a point on an *interval exchange transformation* (IET). The idea is that you chop up an interval into subintervals, rearrange them, and glue them back together. Then you chop up and rearrange them in the same way – and repeat.

Visually, choose a starting point on the bottom line of the diagram (e.g., line (2) of Figure 71). To apply the transformation, flow *up* to the top line. Note which subinterval (say, C) your point ends up in. Take that subinterval C, with your point stuck in it like a nail in a board, and shift it down to the location of C on the bottom line.

Look at where your point ends up. That's the image point! Then repeat, as many times as you like. An example of the full orbit of one point is shown in line (2) of Figure 71.

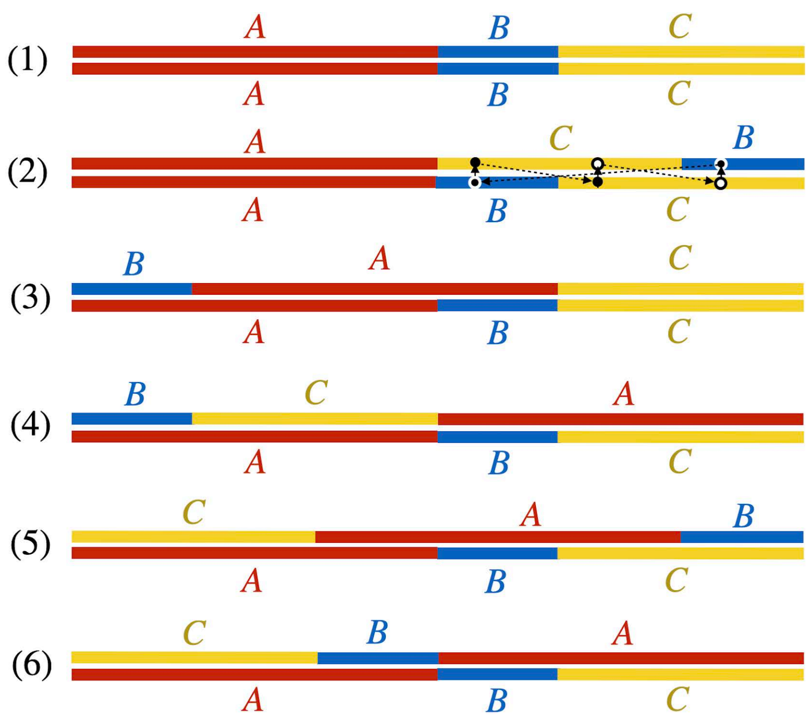


Figure 71. The six ways of rearranging three intervals.

(a) We said in Problem 95 that the dynamics of the IET explored in that problem are identical to those of the rotation in Problem 92. Explain why *every* 2-interval IET is equivalent to a rotation. Is this still the case when the interval lengths are irrational?

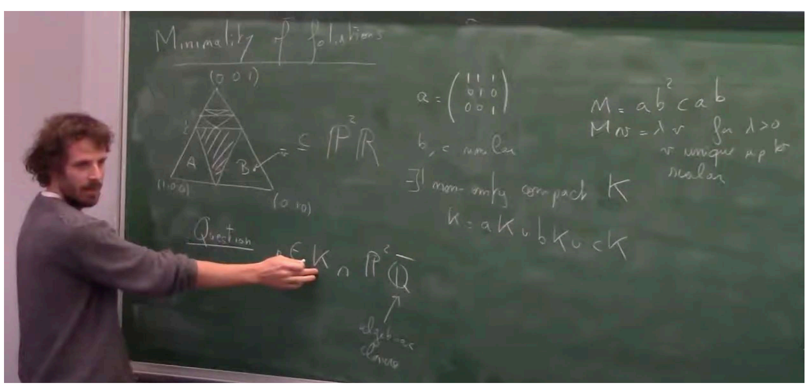
(b) Figure 71 shows the six possible ways of rearranging three intervals. (1) is the identity, and (2) and (3) are the identity on part of the interval and 2-IETs (rotations) on the rest of the interval. Of the remaining three, two of these are also rotations, leaving just one *irreducible* 3-IET. Which one?

98. For the 3-IET that you identified in the previous problem:

(a) Choose a point, mark all the places it goes (its *orbit*), and find the period of its orbit. Does the orbit of every point have the same period?

(b) The interval lengths for the IETs above are $|A| = 1/2$, $|B| = 1/6$, $|C| = 1/3$. Show that the orbit of *every* point is periodic.

Just about everything I know about interval exchange transformations, I learned from Vincent Delecroix (THEY DID THE MATH # 21). With the group pictured in # 16, we had figured out that tiling billiards on triangle tilings are equivalent to orbits on certain IETs – but I knew very little about IETs. At a conference in Marseille in 2017, stretching into the early hours of the morning, Vincent explained to me some essential tools for working with IETs, such as Rauzy diagrams (see § 38). This illustrates a key principle, which is that many of the ideas in mathematics are passed down by oral tradition, one on one, people explaining things to each other and taking notes. In addition to educating colleagues and writing research papers, Vincent writes and maintains software related to exploring translation surfaces, IETs, and other aspects of dynamical systems [**17**]. The picture shows Vincent giving a talk in Warwick in 2017.



They did the math # 21. Vincent Delecroix

99. Each picture in Figure 72 shows a trajectory on the square torus surface, and a parallel trajectory that is slightly shifted, or "perturbed," from the original. Let's consider them to be in the same "family."

(a) For each picture, draw another trajectory that is slightly perturbed from the given ones, and is also in the same family.

(b) If you perturb a trajectory enough, it will eventually hit a vertex. A "singular trajectory" that hits a vertex on both ends is called a *saddle connection*, and forms the boundary of the family of trajectories. Draw in these boundaries for each of the pictures.

(c) The union of such a family of periodic trajectories is called a *cylinder*. Can you guess why this name was chosen?

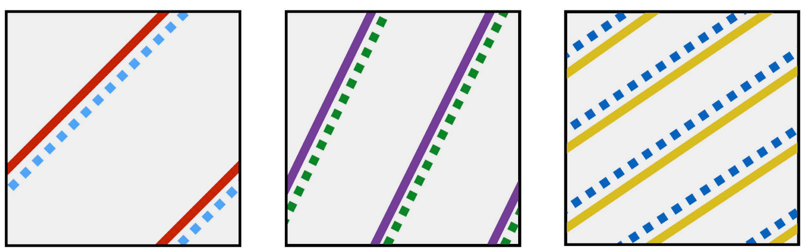


Figure 72. Periodic trajectories on the square torus (solid), and their perturbed friends (dotted).

100. You will need: scissors, tape. We saw that the octagon and double pentagon surfaces each have just one cone point, with 6π of angle around it. What does this even mean? What does it look like? If your birthday is in the first half of the year, use the picture in Figure 74; if it is in the second half of the year, use the picture in Figure 75 – or do both! Print the picture from the book web site, and follow the instructions given in the figure caption. Then think about what it means to have 6π of angle around a vertex. Bring your folded-up, taped figure to class!

The construction in Figure 75 comes from Florent Tallerie [**32**]. This is a small piece of his layout for a genus-2 flat surface, the part at the bottom of the picture in Figure 73.

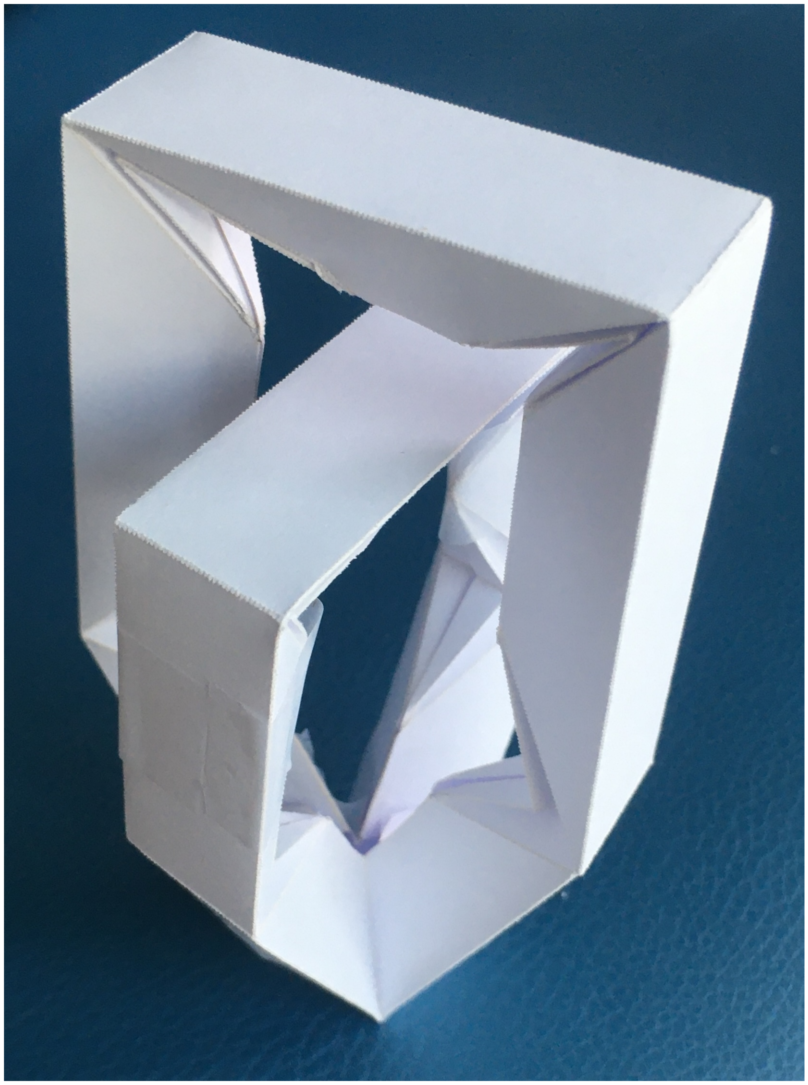


Figure 73. Florent Tallerie's construction of a flat surface of genus 2. The diplotorus layout in Problem 70 solves the problem of creating a flat surface of genus 1, made out of paper and enclosing a positive volume; this construction solves the same problem for genus 2. Problem 94 tells us that the total angle defect is 4π for a genus-2 surface, hence the cone point.

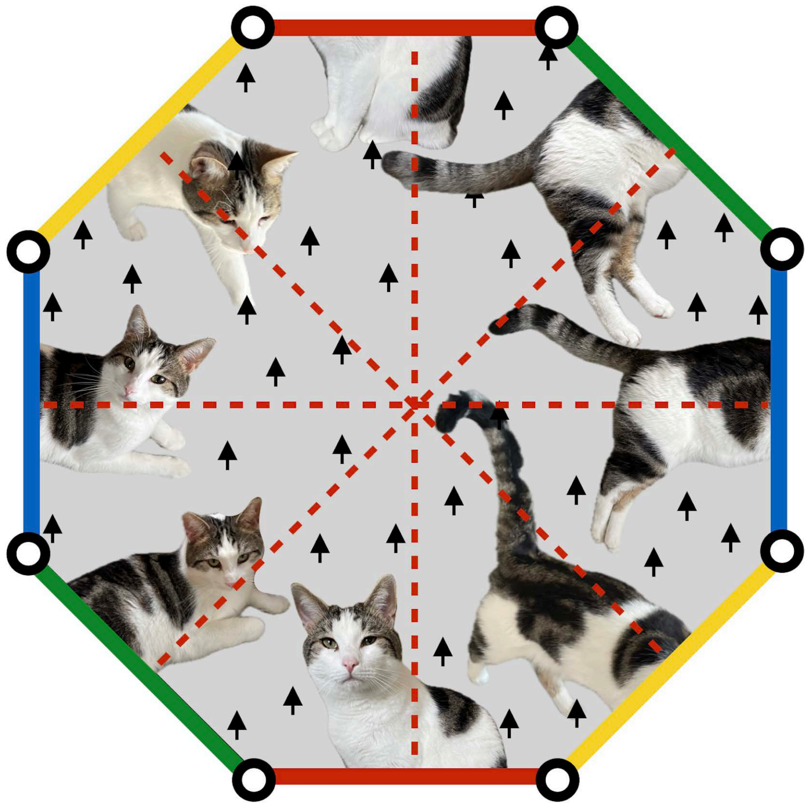


Figure 74. Print out a copy of this picture from the book web site and cut it out. Tear along the dashed lines – tear, don't cut, so that you can remember what is what! Then tape the pieces together along edges with the same color, using the pictures of Spinnaker the cat to remember what is glued to what, and keeping the arrows pointing in the same direction to keep a consistent orientation. Position your tape as close as possible to the vertex, ideally touching the white point. At the end, you should have a "spiral staircase" with 6π of angle around the white point. Can you bring the last two edges together? Imagine circling the white point by walking on your taped-up paper, and convince yourself that you would make three circles around the point.

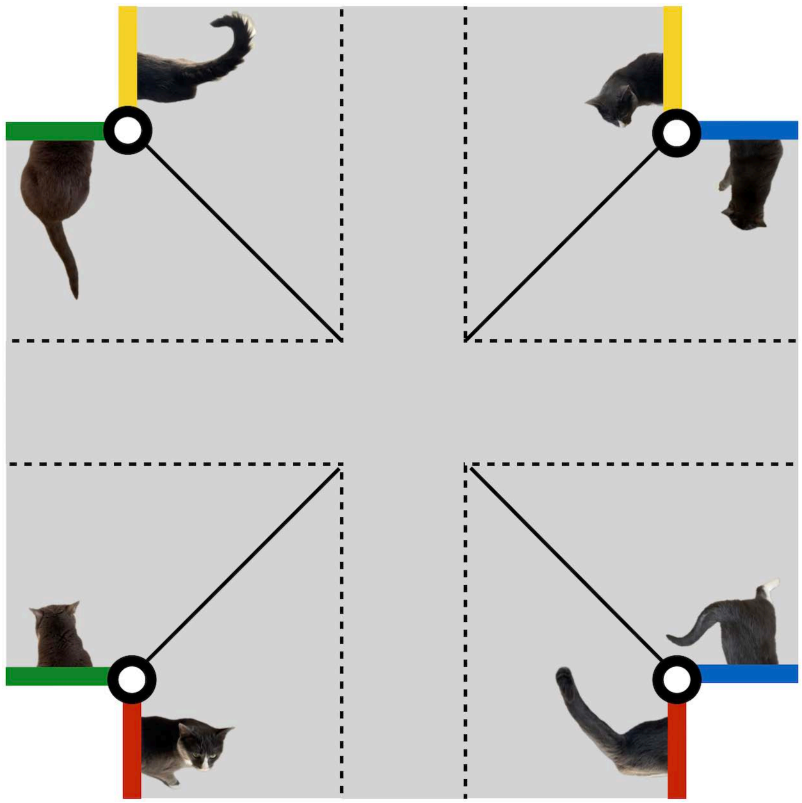


Figure 75. Print out a copy of this picture from the book web site and cut it out. Fold along the indicated edges: dashed lines are "mountain folds" and solid lines are "valley folds." Tape the same-color edges together, using the cats to match identified edges (Jib the cat wishes to remind you that a cat is made from one front half of a cat and one back half of a cat). Position your tape as close as possible to the vertex, ideally touching the white point. Notice that the angle at the white point is 6π! Convince yourself that if you tried to draw a "circle" around the white point, it would take much longer than usual to get back to where you started.

22. We do some computer experiments

101. Let's experiment a bit with tiling billiards trajectories on triangle tilings. Go to the web site https://awstlaur.github.io/negsnel/, coded by Pat Hooper and hosted by Alexander St. Laurent.

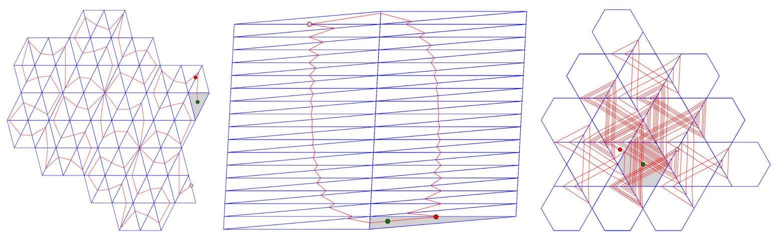


Figure 76. Tiling billiards trajectories, now with live action.

(a) Move the starting point (green) and the direction (red) and see what sort of things you can get. You'll get things like Figure 76.

(b) Click on "Help" at the top and learn how to control the applet with keys.

(c) Click "New" and create a new triangle tiling determined by angles of your choice. Find a really big periodic trajectory. Find a really interesting trajectory. Take a screenshot.

(d) Click on "New" and select some other kind of tiling. Find a really interesting trajectory. Write down the parameters you used. Take a screenshot.

(e) Use the w, a, s, d keys to slightly nudge the direction. Is your trajectory stable or unstable under small perturbations in the direction?

(f) Notice that you can click Edit > Set iterations. Once you get something interesting, increase to more iterations and see what happens when you allow more bounces. (Turn down the iterations when perturbing the trajectory.)

It turns out that programming can be really helpful for figuring out what is going on in a dynamical system. If you have a program that models the system you want to study, you can experiment and get a sense of what is going on. For example, in Problem 101, you probably noticed that the dynamics on some tilings are boring, while the dynamics on other tilings are rich and fascinating. You'd want to spend your time on the latter. Experimentation also leads to conjectures, which you might be able to prove, or to counterexamples, which can stop you from trying to prove something false.

Pat Hooper (THEY DID THE MATH # 22) has written a lot of code for studying billiards and translation surfaces. In collaboration with Vincent Delecroix (# 21) and Julian Rüth, Pat has developed a python package called *sage-flatsurf* that allows people to experiment, compute, and understand far more about flat surfaces than they could with paper, pencil, and brain alone [**17**]. The picture shows the author with Pat in Stony Brook in 2015.

They did the math # 22. W. Patrick Hooper

102. Recall that the union of a family of parallel periodic trajectories is called a *cylinder*, and cylinders are separated by *saddle connections* between cone points (Problem 99). For a translation surface made from polygons, the set of cylinder directions and the set of saddle connection directions coincide. Both cylinders and saddle connections can cross many polygons.

(a) Explain why slopes 2/3 and 5/7 are cylinder directions for the square torus. (Note that the "corner" of the square torus is not a true cone point; we call it a *removable singularity*, as we discussed in Problem 87.)

(b) What are *all* of the cylinder directions for the square torus?

103. So far, we have been studying IETs where you chop up an interval and then rearrange the pieces. Now, suppose that you chop up an interval, *flip each piece*, and then rearrange the pieces. The IET in Figure 77 shows the interval $[0, 1]$ chopped into three pieces (colored red, green and blue).

We flip each piece; the picture represents this by drawing each interval as a triangle, so that you can tell which part of it is which. Then we reassemble the pieces. An IET where every interval is flipped is called a "fully flipped" (or *orientation-reversing*) IET. The picture shows some examples of applying the IET map.

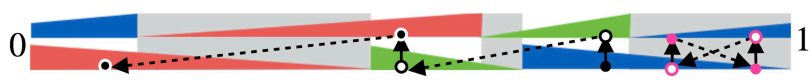


Figure 77. An orientation-reversing 3-IET.

In this picture, we are thinking of 0 and 1 as being equivalent, just like $0 = 2\pi$ on a circle, so that the blue interval is not truly chopped in two, but is just overlapping the break point. So really, this is a *circle exchange transformation* (CET).

(a) For the black point in Figure 77, its first two images are shown. Draw its full orbit. *Hint*: it has period 6.

(b) Notice that some of the intervals overlap their images – with a flip, of course. These regions are shaded in grey. The orbit of one such point is shown in pink. Prove that, for each point in an overlapping region, the period of its orbit is *always* 2 (except that the midpoint is a fixed point).

(c) Argue that (1) the orbit of every point on this IET is periodic, with period $1, 2, 3$, or 6, and (2) if you move your starting point a little bit, the behavior of the point does not change, and (3) these facts do not depend on whether the interval lengths are rational or irrational. These are three reasons to love fully flipped CETs!

Here is another reason to love fully flipped CETs: it turns out that the dynamics of tiling billiards on triangle tilings are equivalent to the dynamics of a fully flipped CET like the one shown in Figure 77! Elijah Fromm, Sumun Iyer, Paul Baird-Smith, and the author discovered this surprising equivalence [**7**]. The periodic nature of fully flipped CETs that you explained in Problem 103(c) gives rise to the abundance of periodic behavior that you noticed in Problem 101(c).

In Problem 90 **(a)**–**(b)**, you proved that when you fold along every edge of a triangle tiling, all of the triangles in this folded state are inscribed in the same circle; this is the circle for the CET. Every time you fold along an edge of the triangle tiling, you are essentially performing a reflection, which gives the CET its fully flipped property. You proved in Problem 90 **(c)** that in the folded state, all of the pieces of trajectory are contained in a single chord of the circumscribing circle. So as the trajectory moves across the tiling, in the folded state the pieces of trajectory go back and forth across the chord, and the triangles dance around the circle according to a CET.

104. The details of the CET described above are tedious to compute, and are beyond the scope of this text. Still, explain why it is plausible that tiling billiards on a triangle tiling is equivalent to the motion of a point on a CET.

They did the math # 23. Pascal Hubert

This surprising connection linked the new field of tiling billiards to the existing field of interval exchange transformations, and to the substantial body of knowledge about the dynamics of IETs. After this, more people got interested in tiling billiards! Two of the people who got interested were Olga Paris-Romaskevich (# 38) and Pascal Hubert (They did the math # 23), who together proved many new results about tiling billiards, including proving interesting conjectures, extending the triangle ideas to quadrilaterals, and applying substantial previous work on IETs by Arnaldo Nogueira [**29**]. The picture shows Pascal and Nicolas Bédaride in Marseille in 2023.

Many people have studied IETs for many decades, as a one-dimensional system without very many pictures. Tiling billiards gives new two-dimensional pictures that represent IETs, which is exciting.

105. Here is a new game: make some number of 1×1 squares going vertically (in Figure 78, six). Then make a big square that goes across all of them, and make some number of those going horizontally (here, one). Then make a big square that goes across all of *them*, and make some number of those going vertically (here, three), and so on. Here we end up with a 7×27 rectangle. Show how to do this to end up with a rectangle whose dimensions are the day and month of your birth. Does every birthday work?

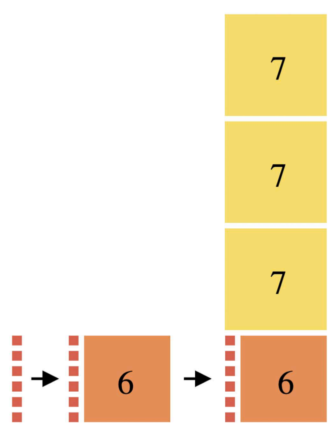


Figure 78. Using squares to make a 7×27 rectangle.

We'll extend the ideas from the problem above in Chapter 4.

106. *Stability under perturbation, part I*
Consider a billiard trajectory in the square billiard table.

(a) If you keep the direction the same, and change your starting *point* a little, what happens? Does the trajectory change a lot, or is it essentially the same?

(b) How about the reverse – if you keep the starting point the same, and change your *direction* a little bit, what happens?

(c) If you keep the starting point and direction the same, and perturb the *table* a little bit so that it is not quite a square, what happens to the trajectory? Consider the case of a rectangle, and also the case of a non-rectangle.

Stay tuned for *Stability under perturbation, part II* in Chapter 4.

23. Hands-on activities for Chapter 3

Celtic knots are a traditional form of decorative art associated with Ireland. They come in many different shapes, some of which are related to... periodic billiards on the square!

Figure 79. Three examples of Celtic knots.

107. All Celtic knots are *alternating*, meaning that if you follow a cord along its journey, it alternates over, under, over, under... as it crosses other parts of the cord.

(a) Check that the knots in Figure 79 are alternating.

(b) Figure 80 shows how to transform a billiard trajectory into a Celtic knot. Do so yourself for the three examples in the bottom row. *Hint:* Draw in the "crossings" first, following the path around to make it alternating, and then fill in the rest of the knot.

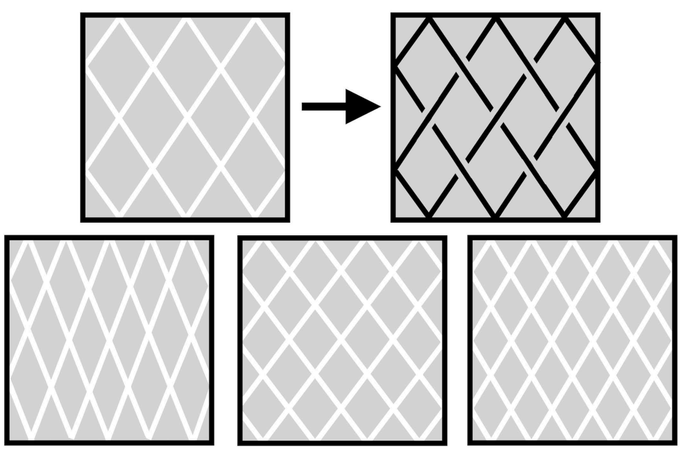


Figure 80. Periodic billiard paths → Celtic knots!

108. You will need: rope. Optional: wooden board, hammer, nails. With a rope, create a Celtic knot based on periodic billiard trajectories. An example is in Figure 81.

Advice: Draw a picture of the desired knot, including the crossings, to help you avoid errors. (Can you weave an alternating knot *without* looking at a diagram of the proper crossings? I have tried many times, but I have always made at least one mistake.) Creating the knot is easiest to do if you have a solid frame, such as a board with nails in it, to hold the cord in place. Mark the board with the crossings, as shown on the right side of Figure 81, so that you will know how to weave your knot as you go.

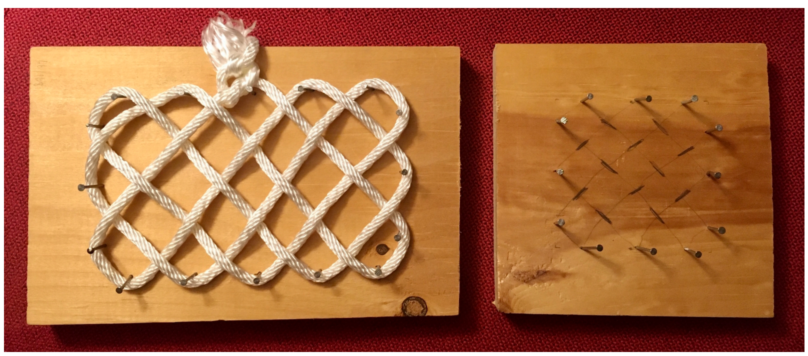
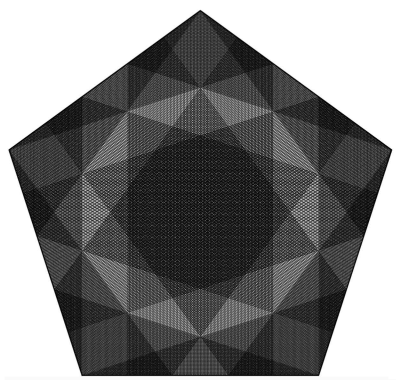


Figure 81. Weaving a Celtic knot based on a billiard trajectory, out of a real piece of rope.

Chapter 4

Cylinders and automorphisms



A long periodic trajectory on the regular pentagon that spends more time in some areas than in others.

In this chapter we build up the full power of cylinders. Our goal is to understand *all* of the periodic trajectories on our surfaces, by understanding how automorphisms act on surfaces and their cylinders, as we did for the square. We will also meet an eclectic menagerie of surfaces designed to do all sorts of interesting things.

24. Twisted cylinders

109. *Stability under perturbation, part II*
Consider a trajectory on a square *outer* billiard table, as in Figures 11 and 32.

(a) If you change your starting point a little, what happens? Does the trajectory change a lot, or it essentially the same?

(b) For inner billiards on the square billiard table, we chose a starting point *and* a direction. Explain why in outer billiards, we don't: the starting point determines the trajectory.

Now consider a *tiling billiards* trajectory on a square grid, as in Figure 56.

(c) If you change your starting *point* a little bit, what happens? Does the trajectory change a lot, or it essentially the same?

(d) If you change your starting *direction* a little bit, what happens?

110. Let's make sure your shearing skills are sharp. Local sheep, beware!

(a) For each of the identified lattice points in Figure 82, draw its image under the horizontal shear $\left[\begin{smallmatrix}1 & 1\\ 0 & 1\end{smallmatrix}\right]$.

(b) Repeat for the vertical shear $\left[\begin{smallmatrix}1 & 0\\ 1 & 1\end{smallmatrix}\right]$.

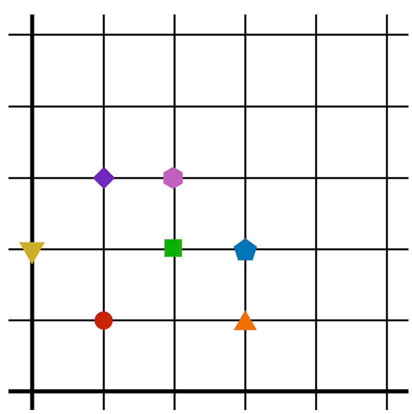


Figure 82. Practice your skills by shearing these points.

111. In Problem 28, we sheared the square torus by applying the matrix $\begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$, which transformed it into a parallelogram, and then we reassembled the pieces back into a square. This action amounted to a twist of the torus surface. Figure 83 shows another way of shearing the square torus (left), this time via the matrix $\begin{bmatrix} 1 & m \\ 0 & 1 \end{bmatrix}$, and reassembling the pieces (right) in such a way that the reassembly respects the edge identifications. The edge identifications are indicated with colors. Explain what is going on.

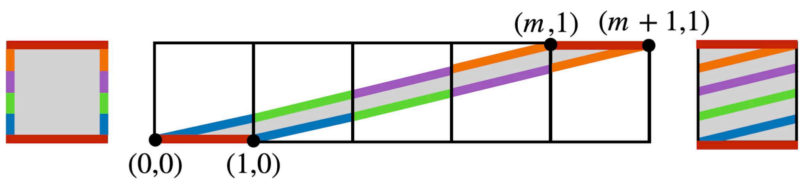
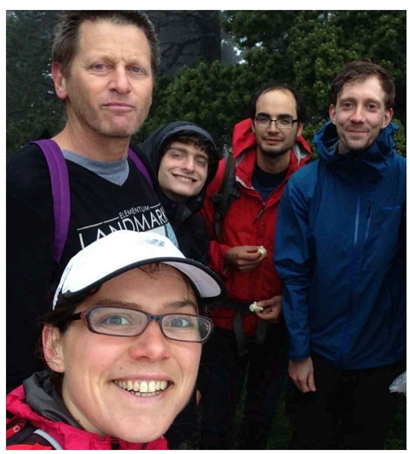


Figure 83. A square torus twisted via $\begin{bmatrix} 1 & m \\ 0 & 1 \end{bmatrix}$, where $m = 4$.



They did the math # 24. Barak Weiss

Barak Weiss (They did the math # 24) introduced me to the idea of twisting a cylinder over and over to see what happens. As we will see in Problem 132, sometimes something happens that is very interesting indeed. The picture shows some mathematicians on a hike in the mountains above Grenoble in 2018: Barak, Fernando Al Assal, Ben Dozier, and René Rühr, with the author.

112. For the square torus, in every cylinder direction there is only one cylinder. For surfaces made from other polygons, there can be multiple cylinders. The double pentagon surface has *two* cylinders in each cylinder direction. Figure 84 shows cylinders on the double pentagon surface in four directions.

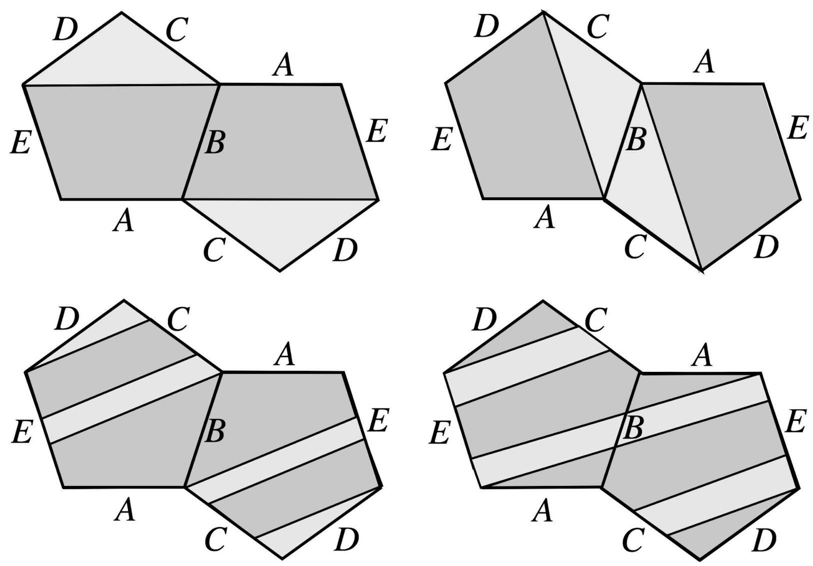


Figure 84. Four different cylinder decompositions of the double pentagon surface.

(a) For each set of cylinders in Figure 84, consider a trajectory on the surface, in the cylinder direction. Write down the cutting sequence for the trajectory in the light cylinder and for the trajectory in the dark cylinder. Think about similarities and differences with our work on the square torus.

(b) Construct a vertical cylinder decomposition of the surface.

(c) The two cylinder decompositions in the top line of the picture are equivalent under a rotation. Is the vertical decomposition from (b) equivalent to any of those shown?

113. Consider the L-shaped surface made of three squares, with edge identifications as shown on the left side of Figure 85. We shear it by the matrix $\left[\begin{smallmatrix}1 & 2\\ 0 & 1\end{smallmatrix}\right]$, as shown.

Show how to reassemble the sheared surface back into the L surface. Make sure that your reassembly respects the edge identifications.

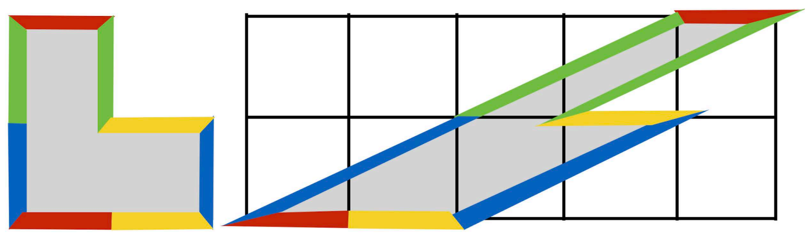


Figure 85. We came, we sheared, we reassembled.

Since we get the same surface back – and thus the sheared version of the L-shaped surface differs from the original only by a cut-and-paste equivalence – we say that $\left[\begin{smallmatrix}1 & 2\\ 0 & 1\end{smallmatrix}\right]$ is an *automorphism* of the L-shaped surface.

25. Let's get illuminated!

Figure 86 shows what happens when you put a candle in a room: the light radiates out in every direction. Look closely at the bottom of the picture: this room has a *mirror* on the wall, so the rays that hit the wall bounce off, following the billiard reflection law.

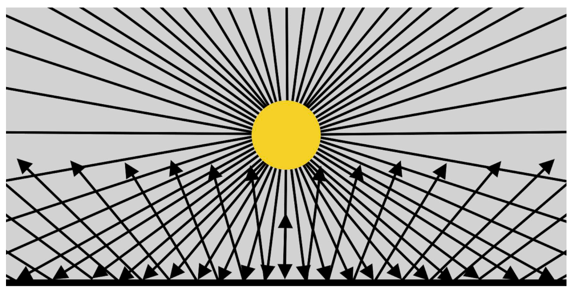


Figure 86. Light from a candle radiates in all directions and reflects off of a mirrored wall.

114. Suppose that you are in a room whose walls are *all* mirrored. You wish to illuminate your entire room with a single candle.

(a) Explain why this problem is easy when the room is convex.

(b) Suppose your mirrored room is an L-shape made of three squares, as shown on the left side of Figure 87, and suppose you place the candle somewhere in the dark square. Does the candle illuminate the whole room? Explain why or why not.

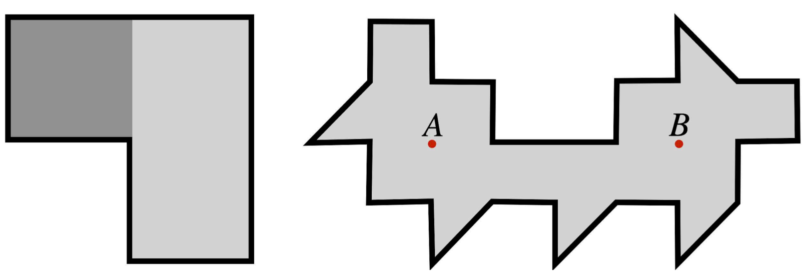


Figure 87. Two interesting examples of mirrored rooms.

The *illumination problem* asks a generalization of the above: for which shapes of mirrored room can you put a candle *anywhere* in the room, and be sure that the light will reach every point? George Tokarsky constructed an example of a polygonal room made of squares and isosceles right triangles to answer this question, shown on the right side of Figure 87 [**57**]. The room contains two points A and B that do not illuminate each other: a candle placed anywhere *other than* points A and B illuminates every point in the room, while a candle placed at A will illuminate every point *except* point B, and vice versa.

Later, Samuel Lelièvre (# 30), Thierry Monteil, and Barak Weiss (# 24) wrote a paper memorably titled "Everything is illuminated" that, in conjunction with a paper by Barak's student Amit Wolecki, shows that all polygons whose angles are rational numbers of degrees are basically like that: every point illuminates every other point, except possibly for a finite collection of points that don't illuminate each other [**33**, **62**].

They did the math # 25. Alex Eskin

To prove their result, Samuel, Thierry and Barak used the "Magic Wand Theorem," the colloquial name for a collection of powerful results from a paper of Alex Eskin (They did the math # 25), Maryam Mirzakhani (# 11), and Amir Mohammadi [**22**, **23**]. In 2020, Alex received the Breakthrough Prize for his work on the Magic Wand Theorem. The picture shows Alex (center) with his wife Anna Smulkowska (left) and mathematician Ursula Hamenstädt (right) in Chicago in 2017.

115. The pictures in Figure 88 show a surface made from a non-regular hexagon.

(a) The first picture shows a foliation (Problem 95) by parallel trajectories. Explain how any trajectory in this direction can be represented by the orbit of a point on an IET.

(b) Using the second picture, draw a foliation in a different direction of your choice. Draw in the diagonal that is closest to perpendicular to your trajectories, and use it to sketch the corresponding IET.

(c) Show that the "top" and "bottom" segments on the diagonal corresponding to a given edge (e.g., edge A) always have the same length.

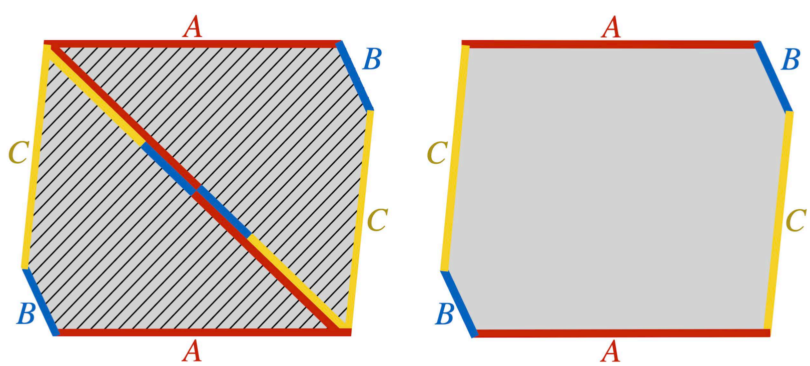


Figure 88. Transforming a surface flow into an IET.

Contextual note. In mathematics, we often care about the *dimension* in which we are working. For example, a torus is a 2D object, and if we look at it as the surface of a bagel, it is a 2D surface *embedded* in 3D space. The family of parallel trajectories in a given direction on the hexagon surface of Figure 88 looks like a 2D system, but we showed above that the behavior of each one can be reduced to the orbit of a point on an IET, which is a 1D system.

116. In Problem 113, we showed that the shear $\left[\begin{smallmatrix}1&2\\0&1\end{smallmatrix}\right]$ is an automorphism of the L-shaped surface made from three squares: You can apply this automorphism, and then rearrange the resulting pieces by translation, while respecting edge identifications, to get back the same surface you started with.

What about 90° rotations? What about the simpler shear $\left[\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right]$? Apply these transformations to the surface using the framework in Figure 89, and determine whether the rotation and the shear are automorphisms of the surface.

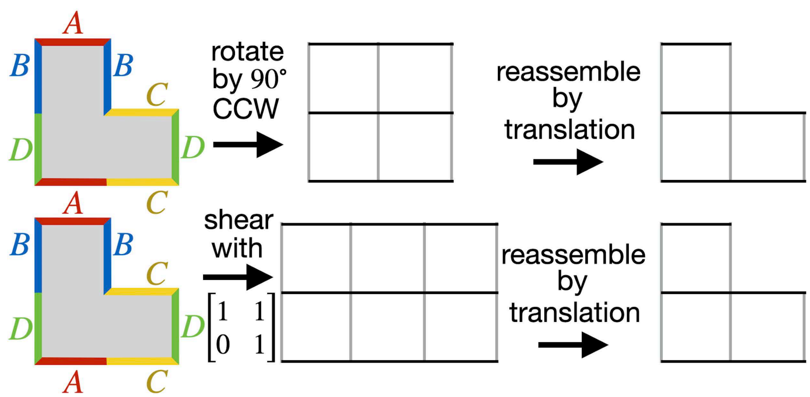


Figure 89. Are this rotation and this shear automorphisms of the L?

117. Do the same for the 2×1 rectangle in Figure 90.

The surfaces in Problems 116–117 are called *square-tiled surfaces*, meaning that they are created by gluing together unit squares, edge to edge. The group of 2×2 matrices with integer entries and determinant 1 is known as the "special (determinant 1) linear group of order 2 (2×2 matrices) with entries in $\mathbf{Z}$ (integers)," and is denoted by $\mathrm{SL}(2,\mathbf{Z})$.

Think back about your experiences shearing and reassembling surfaces in Problems 111, 113, 116, and 117. Given a square-tiled surface and a matrix in $\mathrm{SL}(2,\mathbf{Z})$, what do you think is the probability that the matrix is an automorphism of the surface? Jane Wang and Sunrose Shrestha (They did the math # 26) have studied the

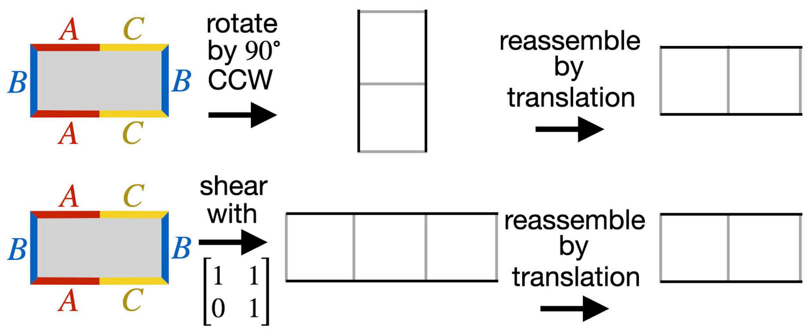


Figure 90. Are they automorphisms of the 2×1 rectangle?

statistics of square-tiled surfaces: exploring what proportion of the square-tiled surfaces in each stratum have various interesting properties [**52**, **53**]. The picture shows billiards enthusiasts (front row) Chandrika Sadanand (# 7), the author, Aaron Calderon, Jane; (back row) Michael Wan, Solly Coles, Samuel Lelièvre (# 30), and Sunrose, in Boston in 2017.

They did the math # 26. Jane Wang & Sunrose Shrestha

118. *Counting periodic trajectories, part I*

One way to count periodic billiard trajectories in the square is to ask how many periodic trajectories it has with length less than L. (Here by *length* we mean the distance along a trajectory in one period, measured using a ruler or perhaps the Pythagorean Theorem.) Of course, periodic trajectories occur in parallel families, which form cylinders (Problem 99); we will count the number of such families.

(a) How long is the trajectory of slope 2? The trajectory of slope 3/4?

(b) Explain why the number of lattice points inside a disk of radius L is approximately πL^2, especially when L is large.

(c) Use the above to show that the number of periodic families of length less than L is approximately $\pi L^2/8$.

26. The tree of periodic directions

119. Figure 91 shows a way of starting with simple vectors and generating more complicated vectors. Here is how we construct this tree (called the *Farey tree*): start with the vector $\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]$ in the lower left. At each step, choose either to add the entries together to get a new x-value (moving right), or to add the entries together to get a new y-value (moving up). Fill in as many entries as you can.

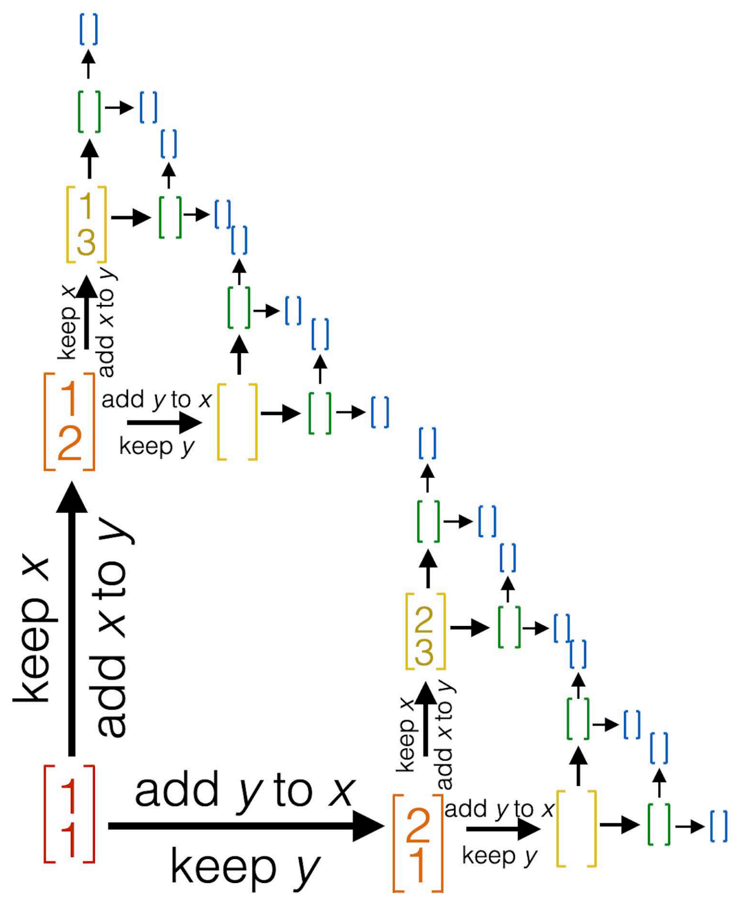


Figure 91. The first five levels of the Farey tree.

The picture shows the first five levels of an infinite binary tree. A *binary tree* means that at each *node* of the tree, you have two choices of where to go – in this case, right or up. I made each level smaller than the previous one so that five levels would fit on the page.

120. (Continuation) Let's explore this tree a bit.

(a) Find $\left[\begin{smallmatrix}1\\2\end{smallmatrix}\right]$, $\left[\begin{smallmatrix}3\\2\end{smallmatrix}\right]$, $\left[\begin{smallmatrix}3\\5\end{smallmatrix}\right]$, and $\left[\begin{smallmatrix}8\\5\end{smallmatrix}\right]$ in the tree. Comment on any patterns.

(b) What vectors appear in this tree? Does your birthday vector $\left[\begin{smallmatrix}\text{month}\\\text{day}\end{smallmatrix}\right]$ appear in the tree? If so, at what level?

(c) For an integer vector $\left[\begin{smallmatrix}p\\q\end{smallmatrix}\right]$, the continued fraction expansion of q/p tells you how to move in the tree to reach $\left[\begin{smallmatrix}p\\q\end{smallmatrix}\right]$. Explain.

121. We have seen that we can often partition a surface into *cylinders*. The boundary of a cylinder is a *saddle connection* – a line segment connecting two cone points with no cone points in its interior – and there are no vertices inside a given cylinder. To construct the saddle connections, draw a line in the cylinder direction through each vertex of the surface until it hits another vertex. The line might pass through many polygons before it reaches its ending vertex. These lines cut the surface up into strips, and then you can follow the edge identifications to see which strips are glued together to form cylinders. Recall Problem 112, where we saw several examples of cylinder decompositions for the double pentagon.

(a) For the surfaces in Figure 92, sketch the *horizontal cylinder decomposition*, by shading each horizontal cylinder a different color, of each of the surfaces below.

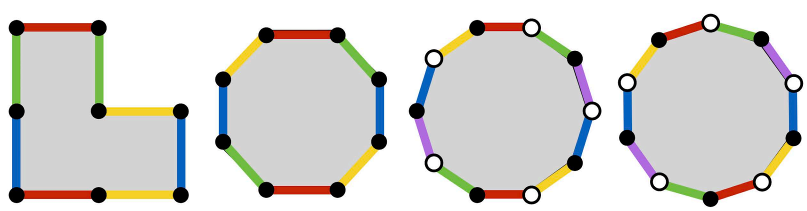


Figure 92. Surfaces desiring a horizontal cylinder decomposition.

Notice that the regular decagon surface has *two* cylinders in one cylinder direction, and *three* cylinders in another direction.

(b) Using the same surfaces, sketch a cylinder decomposition for each one in some non-horizontal direction, preferably not vertical.

We said that a *saddle connection* is a line segment connecting two cone points, with no cone points in its interior. In the decagon surface, some saddle connections connect a vertex to itself (black to black or white to white), while others connect two different vertices (black to white). An automorphism *cannot* map a same-vertex saddle connection to a different-vertex saddle connection, or vice versa.

Speaking of cylinders and automorphisms, let's meet a creature that is legendary in these areas: the eierlegende Wollmilchsau[1] (Figure 93). This surface is interesting because while it is clearly not the square torus, it has some key properties in common with the square torus. We'll explore some of those now.

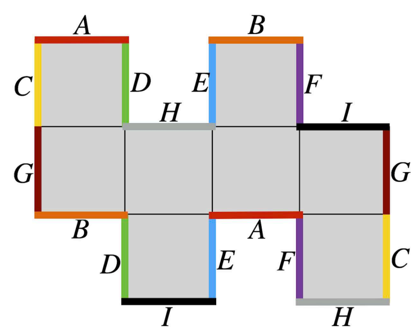


Figure 93. The legend: the eierlegende Wollmilchsau.

122. (a) Show that the surface has two horizontal cylinders and two vertical cylinders, and in each case the cylinder's width (in the cylinder direction) is 4 times its height (perpendicular to the cylinder direction). We say that the cylinders have *modulus* 4. You can think of the cylinder's modulus as its "aspect ratio."

We have previously shown that the square torus has three types of automorphisms: rotations, reflections, and shears. The *group* consisting of all of the automorphisms of a surface is called the *Veech group* of the surface. If we think of the automorphisms in terms of

[1] "EYE-ur-LEEG-un-duh VOLE-milsh-sow"

the 2×2 matrices that perform them, we can say that the Veech group of the square torus is $\mathrm{SL}(2, \mathbf{Z})$.

(b) It turns out that the Veech group of the Eierlegende Wollmilchsau is also $\mathrm{SL}(2, \mathbf{Z})$. To prove this, first recall that in Problem 89, you proved that the matrices $\left[\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right]$ generate $\mathrm{SL}(2, \mathbf{Z})$. Then transform the surface by each of the two generators, and check that you get the same surface back, as we practiced in Problems 113 and 116. Finally, use these results to prove the claim.

They did the math # 27. Gabriela Weitze-Schmithüsen

The eierlegende Wollmilchsau was discovered by Gabriela Weitze-Schmithüsen (They did the math # 27) and Frank Herrlich in 2003 [**25**]. Gabi and Frank gave the surface its catchy name. It translates from German as "egg-laying wool-milk-sow" – an animal that provides eggs, wool, milk and meat, or in other words, everything a person could need. Similarly, this surface provides just about everything you could ever ask for in a surface, including resembling an animal. The picture shows Gabi looking at a banner of herself.

123. Figure 94 shows a 4-IET. An example of the image of one point is shown.

(a) Find the orbit of this point for at least six more iterations. Is its orbit periodic? If so, does the orbit of every point have the same period as this one? (*Hint:* measure carefully! Don't *assume* that everything is periodic!)

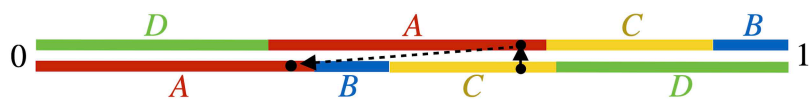


Figure 94. Our first example of a 4-IET.

An IET cuts up an interval of points and reassembles them. So we can think of an IET as a function that maps points between 0 and 1 to points between 0 and 1.

(b) Figure 95 shows the function corresponding to the above 4-IET. Explain.

(c) Use the graph to find the orbit of the same point that you followed in part **(a)**. Note the helpful foliation by lines with slope -1!

124. Show that if the length of every subinterval of an IET is rational, then the orbit of *every* point is periodic.

Interval exchanges are simple to define – just chop up an interval and rearrange the pieces – and even IETs with a small number of intervals can have interesting properties. We have shown that every 2-IET is equivalent to a rotation (Problem 97), and that IETs with rational subinterval lengths have only periodic behavior (Problem 124), but outside of these cases, things can get very interesting indeed.

Jon Chaika (They did the math # 28) has studied many properties of interval exchange transformations, particularly their ergodicity [**10**]. A flow is *ergodic* if, roughly speaking, the amount of time that a point spends in each region is proportional to the region's size.

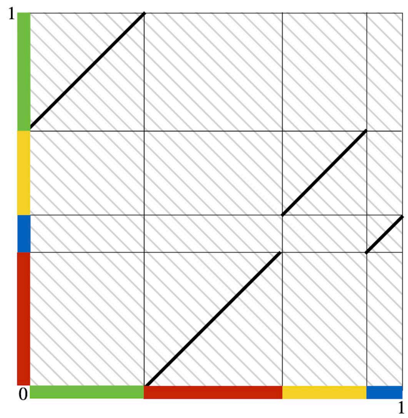


Figure 95. The 4-IET of Figure 94, expressed as a graph.

They did the math # 28. Jon Chaika

For example, in the IET in Problem 123, if the flow is ergodic, then if the interval B has length $1/10$, a point should land in interval B, on average, $1/10$ of the time. The picture shows Evelyn Lamb and Jon, mathematicians who are married to each other, in Moab in 2024.

27. We finally meet Veech

125. You've built up rectangles from squares (Problem 105). You've filled in the binary tree of relatively prime vectors (Problem 119). Now let's look at a third way to generate all of the relatively prime vectors made of positive integers: *shears!*

(a) Start with $(1,1)$ as shown in Figure 96. Apply the horizontal shear $\left[\begin{smallmatrix}1 & 1\\ 0 & 1\end{smallmatrix}\right]$ to the red point to get one new point, $(2,1)$ – draw this in orange. Also apply the vertical shear $\left[\begin{smallmatrix}1 & 0\\ 1 & 1\end{smallmatrix}\right]$ to the red point, to get $(1,2)$ – draw this in orange also.

(b) Now apply the horizontal shear $\left[\begin{smallmatrix}1 & 1\\ 0 & 1\end{smallmatrix}\right]$ to the orange points, and draw these new points in yellow. Do the same for the vertical shear $\left[\begin{smallmatrix}1 & 0\\ 1 & 1\end{smallmatrix}\right]$, applying it to all of the orange points to get new yellow points. You should get four yellow points.

(c) Now apply both the horizontal and vertical shears to the yellow points. Draw these eight new points in green.

(d) Repeat the above for all the green points. Draw the new points in blue. Continue in purple. Mark *all* of the points you get.

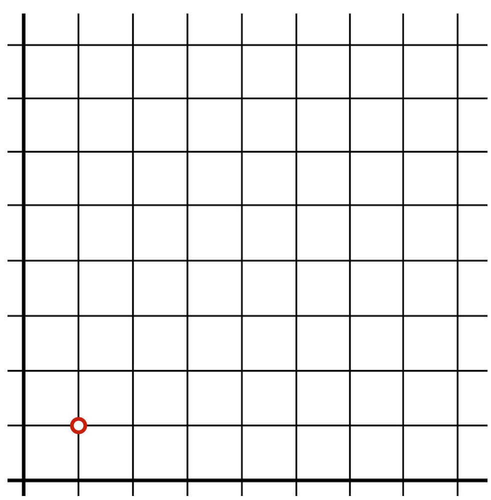


Figure 96. Let us now, at long last, generate our relatively prime vectors using shears.

126. *Making connections, again.*

(a) Explain the connections between the three ways we have seen of generating new points: adding squares (Problem 105), adding vectors (Problem 119), and shearing the plane (Problem 125).

(b) Explain why every point we get in this way is *primitive*, meaning that the greatest common divisor of its components is 1.

(c) We can call this the set of *primitive vectors*, or the set of *visible points*: suppose that you are standing at the origin of an infinite orchard, and there is a tree at every lattice point. Then the points we generated above are the trees that you can see. Explain.

(d) Notice that to reach the point $(5,7)$ in your picture for Problem 125, you applied the transformations

$$(1,1) \xrightarrow{\left[\begin{smallmatrix}1&0\\1&1\end{smallmatrix}\right]} (1,2) \xrightarrow{\left[\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right]} (3,2) \xrightarrow{\left[\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right]} (5,2) \xrightarrow{\left[\begin{smallmatrix}1&0\\1&1\end{smallmatrix}\right]} (5,7).$$

Explain how to use horizontal and vertical shears to implement the continued fraction algorithm for $7/5$. Done in reverse, this is *Euclid's algorithm* for finding the greatest common factor of two numbers: here, 5 and 7.

127. *Counting periodic trajectories, part II*
We can improve on our previous method of counting periodic trajectories (Problem 118) by counting primitive vectors, as these are the directions that give us different billiard trajectories.

Let P be the set of primitive vectors (Problems 125–126). For each natural number k, let kP be the set of primitive vectors multiplied by k, i.e., vectors $[a,b]$ where the greatest common divisor of a and b is k.

(a) Draw the set $2P$ in black on your picture from Problem 125.

(b) Explain why the union of all of the sets $P, 2P, 3P, \ldots$ is every lattice point in the first quadrant, and also show that the sets are disjoint (they have no elements in common). In other words, the sets form a *partition* of the first-quadrant lattice points.

(c) We wish to know the proportion of the integer vectors in the first quadrant that are in P; let's call this proportion x. Show that the proportion of such vectors that are in each set kP is $\frac{x}{k^2}$.

(d) Justify the equation $1 = x\left(\frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \cdots\right)$.

The latter sum is famous; it is known as $\zeta(2)$, as it is the value of the Riemann zeta function for exponent 2. It can be shown that the value is $\frac{\pi^2}{6}$, so the proportion of primitive vectors is $\frac{6}{\pi^2} \approx 61\%$.[2]

For the set of birthday vectors of the form $\left[\begin{smallmatrix}\text{month}\\ \text{day}\end{smallmatrix}\right]$, the probability of primitivity is slightly higher, about 63%.

128. Amazingly, many surfaces made from regular polygons can be sheared, cut up, and reassembled back into the original surface in the same way that we have done with the square, the L, and the Wollmilchsau. One example is the regular octagon surface, shown in Figure 97 sheared by the matrix $\left[\begin{smallmatrix}1 & 2(1+\sqrt{2})\\ 0 & 1\end{smallmatrix}\right]$. The way to reassemble the sheared octagon pieces is indicated with numbers.

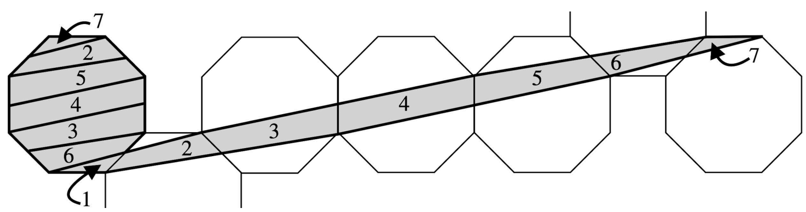


Figure 97. Did you think that something like this could happen to the regular octagon surface? Do you believe that the skinny diagonal thing is really a convex, non-regular octagon? Every day is full of surprises.

(a) By coloring each piece of each edge of the original and sheared octagons as in Problems 111 and 113, show that this reassembly respects the octagon surface's edge identifications. In other words, show that this shear is an automorphism of the octagon surface.

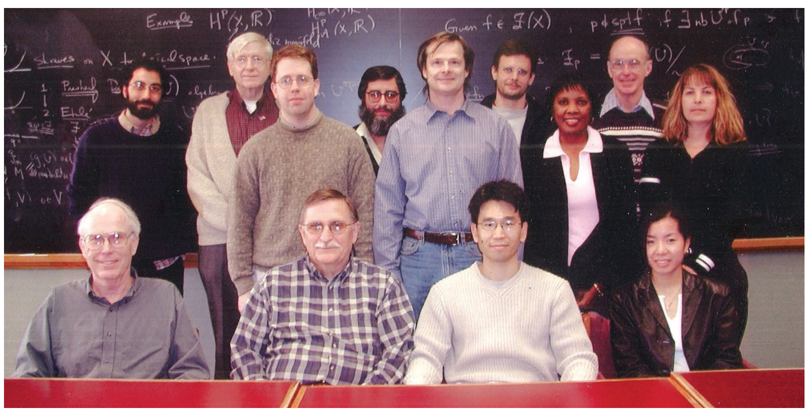
We say that a shear in a cylinder direction *twists* that cylinder, analogous to twisting the dough of a bagel (recall Figure 26). For example, in Problem 111 the shear $\left[\begin{smallmatrix}1 & m\\ 0 & 1\end{smallmatrix}\right]$ twists the square torus's single horizontal cylinder m times.

(b) In Problem 121, you found the octagon surface's two horizontal cylinders. In the shear above, show that the top/bottom cylinder is twisted once, while the middle cylinder is twisted twice.

[2]Thanks to Juan Souto for explaining this proof to me, at a bar in Dublin.

(c) For each of the horizontal cylinders in the regular octagon surface, find its modulus (recall Problem 122.) How is the modulus related to the number of twists?

We previously said that the group of automorphisms of a surface is called its *Veech group* (Problem 122). When the group is particularly nice – to be precise, when the group of automorphisms is a subgroup of $\mathrm{SL}(2, \mathbf{Z})$ whose fundamental domain has finite area in the hyperbolic plane; see § 33 – its Veech group is said to form a *lattice*. A surface whose Veech group forms a lattice is called a *Veech surface* or *lattice surface*.[3] The square torus, the regular octagon surface, and square-tiled surfaces are all examples of Veech surfaces.



They did the math # 29. William Veech

These notions are named for mathematician William Veech (THEY DID THE MATH # 29), who got this field going and then did a lot of tremendous work in it, including coming up with IETs and Veech surfaces, and then proving results about all of their essential properties [**60**]. One of his original examples of a Veech surface was the *double* regular octagon surface, chosen because its cylinders' moduli are equal. The picture shows (back row) Giulio Minervin, William, Brendan Hassett, Fernando Gouvêa, Tim Cochran, Tamas Wiandt,

[3]They have traditionally been called Veech surfaces, but some people think that too many things are named "Veech," and are trying to change the terminology.

Maxine Turner, Frank Jones, and Janie McBane; (front row) Reese Harvey, John Polking, Donghoon Hyeon, and Joungmin Song at Rice University in 2002.

129. *An infinite-area surface*
Consider the infinite staircase surface in Figure 98. It is a square-tiled surface, where edges are identified across, horizontally and vertically, as indicated. The pattern continues forever in both directions [**28**].

(a) How many cone points does the surface have? What is the angle around each one? What is the genus of the surface?

(b) Identify some periodic trajectories on the surface.

(c) Decompose the surface into cylinders in the direction of slope 1/2.

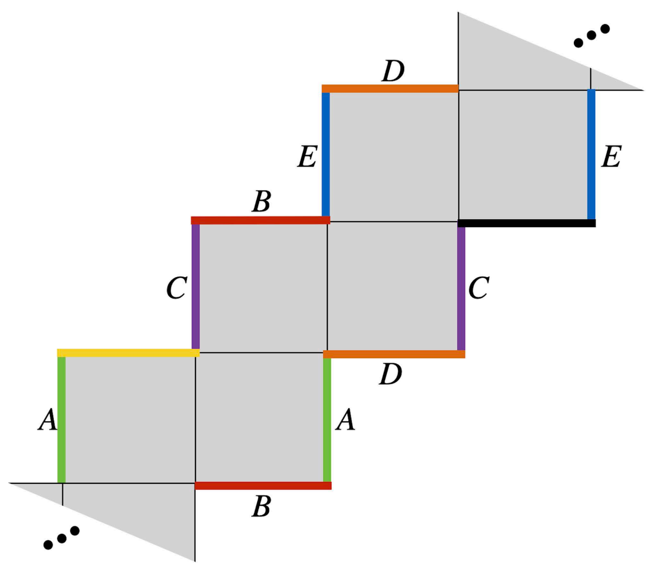


Figure 98. The infinite staircase surface.

28. The modulus miracle

I called the following result the *modulus miracle* when I was a Ph.D. student, because I was absolutely shocked that it was true. Take *any* regular polygon, glue two of them together to make a surface, and *all* of its cylinders have the *same* modulus? Really?!

130. Theorem (modulus miracle). Every horizontal cylinder of a double regular n-gon surface has same modulus, which is $2\cot(\pi/n)$.

(a) Confirm this for the two surfaces in Figure 99, by calculating the modulus ("aspect ratio"; recall Problem 122) for each cylinder, and also the number $2\cot(\pi/n)$.[4]

(b) Explain why this tells us that the horizontal shear $\left[\begin{smallmatrix} 1 & 2\cot(\pi/n) \\ 0 & 1 \end{smallmatrix}\right]$ is always an automorphism of the double regular n-gon surface.

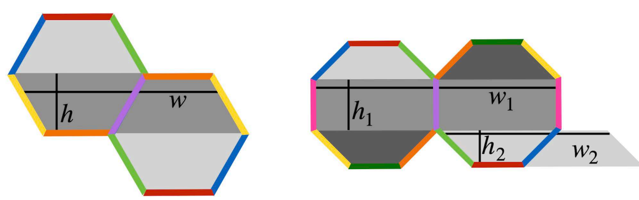


Figure 99. Calculate the moduli of these cylinders. Spoiler alert! In each surface, they are equal.

Thus *all* surfaces made from a double regular polygon have rotation, reflection and shearing symmetries, like the square torus.

The benefit of using a double regular polygon surface instead of a single one is that all of the cylinder moduli are equal. If you do use just a single polygon, like our familiar regular octagon surface, then some of the cylinders have double the modulus of the others (see Problem 128).

131. A particularly nice surface is the "golden L," whose edge identifications and lengths are as shown in Figure 100. The *golden ratio* $\varphi = \frac{1+\sqrt{5}}{2} \approx 1.618$ satisfies the property that when you cut off the largest possible square from a $1 \times \varphi$ rectangle, the leftover rectangle has the same proportions as the original.

[4] For a proof, see [**13**], Proposition 2.4.

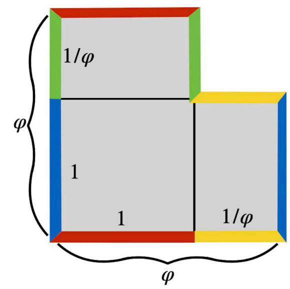
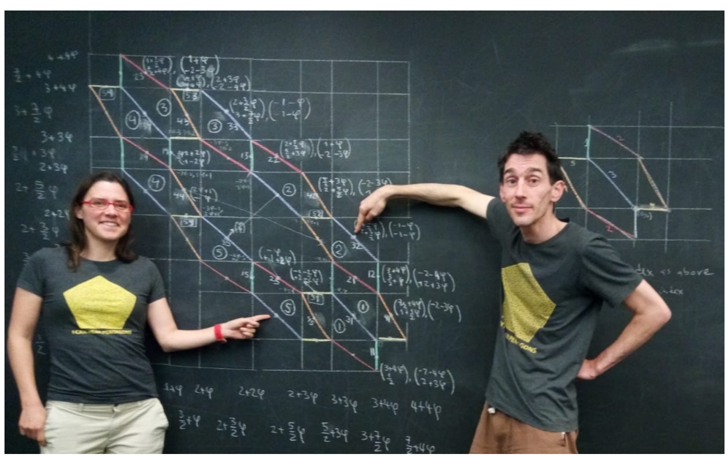


Figure 100. That beautiful surface, the golden L.

(a) Show that the golden ratio satisfies the relation $\varphi = 1 + 1/\varphi$.

(b) Find the continued fraction expansion of φ. (*Hint:* Use part **(a)**.)

(c) Numbers are *commensurable* if they are rational multiples of each other. Are the moduli of the golden L's cylinders commensurable?



They did the math # 30. Samuel Lelièvre

Samuel Lelièvre (THEY DID THE MATH # 30) has studied the golden L surface in detail. In joint work with Jayadev Athreya (# 12) and Jon Chaika (# 28), he studied the *gaps* between slopes of cylinder directions in the golden L [**5**]. It turns out that the golden L and regular pentagon surfaces are closely related; the picture shows

Samuel and the author in Providence in 2019, studying a translation surface made from ten sheared regular pentagons while wearing coordinating T-shirts.

A big question is: "what happens to a trajectory on a surface when you apply an automorphism?" For example, in Problem 32 we explored the effects of rotations, reflections and a shear on a trajectory on the square torus, and determined the effect of each automorphism on the trajectory's slope.

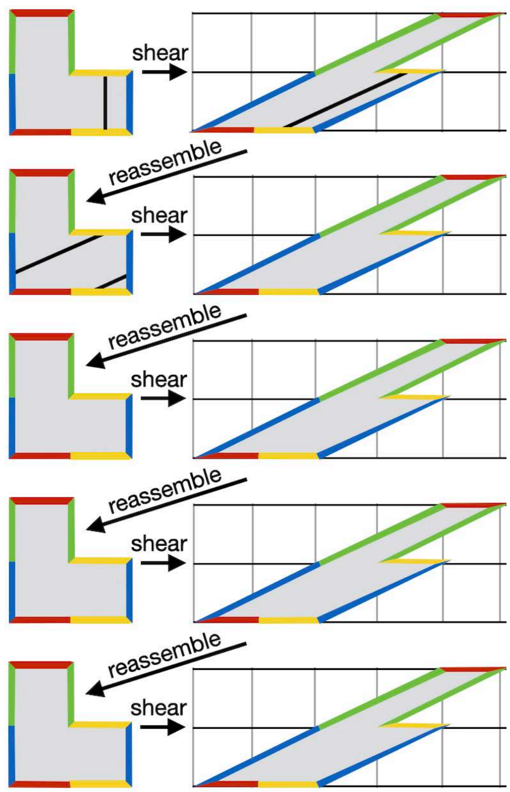


Figure 101. What happens when you repeatedly twist a cylinder that has trajectory in it?

132. Let's see what happens when we apply the horizontal shear $\left[\begin{smallmatrix}1 & 2\\ 0 & 1\end{smallmatrix}\right]$ to the L-shaped table, with a short vertical trajectory on it. Using a picture like Figure 101, sketch the image of this trajectory under five applications of this shear. What would happen if you kept going?

As mentioned before, Barak Weiss (# 24) is the one who suggested to me that interesting things can happen when you twist a surface many times in a cylinder direction. Above, one cylinder fills up, while the other stays empty. Samuel Lelièvre (# 30) and I used this strategy to create the periodic billiard trajectory on the regular pentagon that appears on the first page of this chapter, which has more trajectory in some parts of the table than in others.

133. Our original motivation for studying the square torus was that it was the unfolding of the square billiard table (Problem 11). In fact, we can view *all* regular polygon surfaces as unfoldings of *triangular* billiard tables. For example, let's try unfolding the triangular billiard table with angles $(\pi/2, \pi/8, 3\pi/8)$ until every edge is paired with a parallel, oppositely-oriented edge. In Figure 102, the edges are labeled with numbers, and the orientations are indicated with arrows.

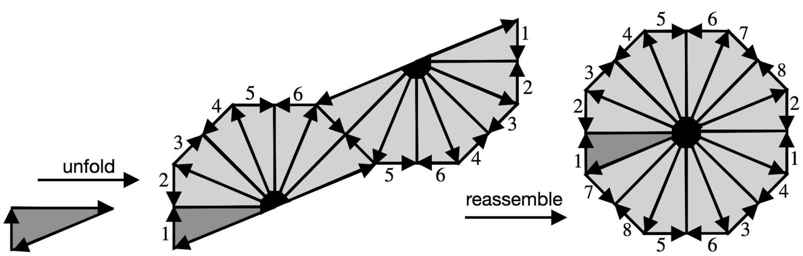


Figure 102. Unfolding a triangular billiard table into the regular octagon surface.

This gives us the regular octagon surface! So the regular octagon surface is the unfolding of the $(\pi/2, \pi/8, 3\pi/8)$ triangle.

(a) Draw the "shooting into the corner" period-six trajectory (Problem 86) in the triangular billiard table (left). Then unfold it to a periodic trajectory on the regular octagon surface (center & right). *Hint:* This trajectory has period 2 on the regular octagon surface, and passes through 6 triangles, including edges 1 and 7.

(b) What triangle unfolds to the double regular pentagon surface? Dissect the double pentagon to figure it out, and then draw the unfolding as above.

134. An *ellipse* with foci F_1, F_2 and string length ℓ (recall Problem 26) consists of all points P satisfying $|F_1P| + |PF_2| = \ell$. Similarly, a *hyperbola* with foci F_1, F_2 and "string length" ℓ consists of all points P satisfying $|F_1P| - |PF_2| = \pm\ell$: see Figure 103.

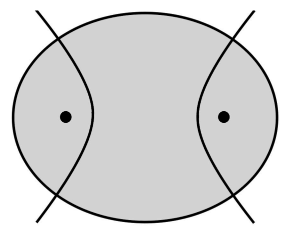


Figure 103. A confocal ellipse and hyperbola, with the shared foci.

In Problem 26, we showed that a trajectory *through* the foci always passes through the foci. In Problem 39, we showed that a trajectory *outside* the focal segment F_1F_2 stays outside and is tangent to an ellipse with the same foci. Show that every segment of a trajectory that passes *between* the foci is tangent to a hyperbola with the same foci. Conclude that *every* segment of such a trajectory will pass between the foci.

This nifty fact will enable us to construct an unilluminable room in Problem 136.

29. The slit torus construction

So far, we have seen a lot of beautiful surfaces that do beautiful things. We have seen that a billiard trajectory with rational slope on a square table is periodic, and a billiard trajectory with irrational slope is aperiodic. It turns out that every aperiodic trajectory on the square billiard table fills up the table evenly – the billiard flow in such a direction is *ergodic*. That's because the square billiard table unfolds to a Veech surface. The *Veech dichotomy* (proved by William Veech, # 29) says that for a given direction on a Veech surface, the billiard flow in that direction is either periodic or ergodic.

When I first learned this, I thought it was obvious. After all, what other possibilities are there? It turns out that there are many other possibilities: surfaces where a trajectory is dense in one region and doesn't touch another region at all, or is half as dense in one region as in another region – or just about anything you can imagine. One nice demonstration of the first possibility is the slit torus.

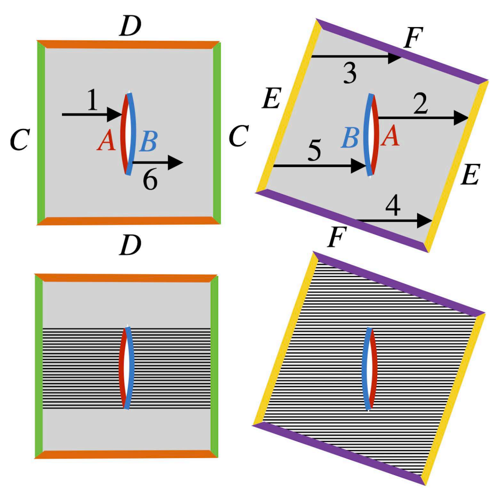


Figure 104. The slit torus construction.

135. The *slit torus* surface is created by joining two square tori along a slit, as shown in Figure 104. One of the tori has horizontal and vertical edges as usual, and the other one is rotated so that its edges have an irrational slope. We cut a vertical *slit* in each one, and identify the left and right edges A and B of one slit to the right and left edges B and A of the other, as shown in Figure 104.

Edges A and B are vertical, but in the picture I have pulled them apart a little bit so that you can see that there is a slit between them.

(a) In the top picture, I have drawn the first six pieces of a horizontal trajectory. Draw the next ten pieces. Do you expect this trajectory to be periodic?

(b) Explain why, over time, a horizontal trajectory through the slit will end up looking like the bottom picture.

They did the math # 31. Moon Duchin

Moon Duchin (They did the math # 31) explained the slit torus construction to me when I was a graduate student. Moon started out working in translation surfaces, and now works on identifying gerrymandering and creating fair districting practices. The picture shows Jane Wang (# 26), Viveka Erlandsson, Justin Lanier, Moon, Solly Coles, Madeline Elyze, Aaron Calderon, Felipe Ramírez, Andre Oliveira, Chandrika Sadanand (# 7), and the author in Somerville during a 2017 billiards research program that Moon organized.

136. *The Penrose unilluminable room.* We can pose the illumination problem as: "Is every mirrored room illuminable from *some* point in the room?" Figure 105 shows a counterexample, a room that cannot be illuminated from *any* point inside [**41**]. The top and bottom are half-ellipses, whose foci are at the black points. Explain why this example works, by explaining which parts of the room are illuminated when the candle is placed **(a)** in the interior of a half-ellipse, **(b)** in the middle part, and **(c)** in one of the rectangular parts.

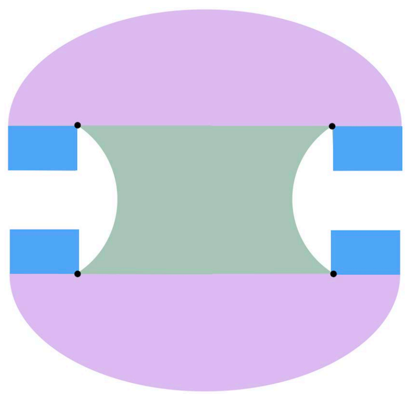


Figure 105. The Penrose unilluminable room.

We have seen that we can generate periodic directions on the square torus in three ways: adding squares (Problem 105), adding vectors (Problem 119), and applying shears (Problem 125). It turns out that applying shears – and more generally, applying automorphisms of the surface – is the method that best generalizes to other surfaces.

137. Figure 106 shows the first quadrant divided into four sectors, each created by neighboring diagonals of the golden L whose corner is at the origin.

(a) The dimensions of the golden L are given in Problem 131. Check that the purple vectors shown spanning diagonals of the golden L are $[1, 0]$, $[\varphi, 1]$, $[\varphi, \varphi]$, $[1, \varphi]$, $[0, 1]$.

(b) Explain why the blue matrix takes the entire first quadrant to the blue sector. Check that its determinant is 1, meaning that it preserves areas and orientations. Repeat for the three other colors.

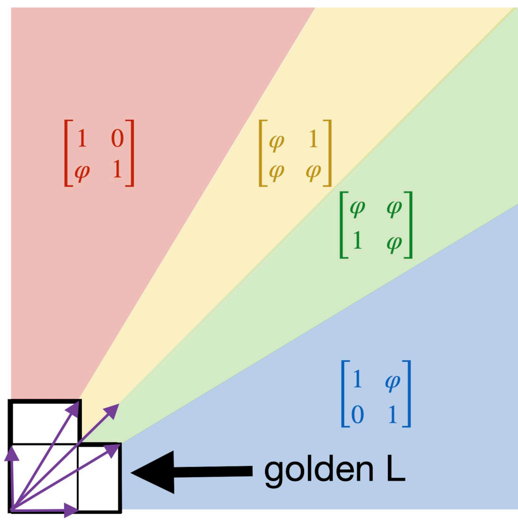


Figure 106. The first quadrant, partitioned into four sectors.

Each of the matrices shown is an automorphism of the golden L. The blue and red matrices are horizontal and vertical shears, respectively. They are known as *parabolic* automorphisms. The green and yellow matrices act similarly to shears in a diagonal direction, but they tend to mix things up more than shears; they are known as *hyperbolic* automorphisms.

138. In Problem 125, we repeatedly applied horizontal and vertical shears to generate *all* of the periodic directions on the square torus. In Problem 126, we explained how applying the two different shears is essentially the continued fraction algorithm in reverse. Similarly, to generate the set of *all* of the periodic directions on the golden L, we start with the vector $\left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$ and repeatedly apply the blue, green, yellow, and red automorphism matrices. People describe this as a “generalized continued fraction algorithm.” Explain.

139. Surfaces with lots of symmetry are rare and precious. For some time, regular polygon surfaces and square-tiled surfaces were the *only* known Veech surfaces. Then William Veech's (# 29) student, Clayton Ward, discovered a larger family of such surfaces, now known as *Ward surfaces* [**61**]. One way to describe a Ward surface is as a regular $2n$-gon with two regular n-gons, where alternating edges of the $2n$-gon are glued to one of the n-gons, and the remaining edges of the $2n$-gon are glued to the other n-gon (left side of Figure 107).

For $n = 4$, the Ward surface is an octagon and two squares, with edges identified as shown below.

(a) Decompose this surface into horizontal cylinders, and check that their moduli are commensurable.

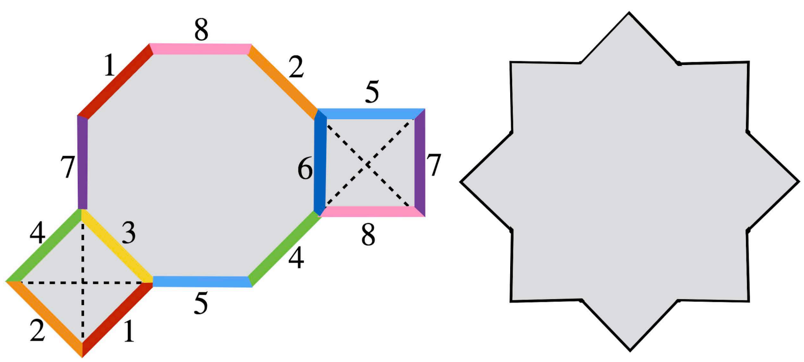


Figure 107. Two views of the same Ward surface.

Ward actually represented this surface as a "flower": you can cut each of the squares into four pieces as shown in the left picture, and glue the eight "petals" around the octagon, as shown on the right side of the figure.

(b) Use the left picture to figure out which edges are identified in the right picture, and write in edge labels to record it.

30. The space of triangles

We have talked here and there about the space of all translation surfaces. The following problem builds understanding about what it means to have a space where each point represents an object.

140. Up to similarity and isometries, a triangle can be uniquely specified by its three angles α, β, γ. There are two restrictions on the angles: $\alpha + \beta + \gamma = \pi$ and $\alpha, \beta, \gamma > 0$. So we can represent the space of all possible triangles by the triangular part of the plane $\alpha+\beta+\gamma = \pi$ that lies in the first octant, as shown in Figure 108. Each *point* of the space represents a *triangle*. So the space of triangles is itself a triangle! It's easier to see the picture if we lay the triangle flat, as shown.

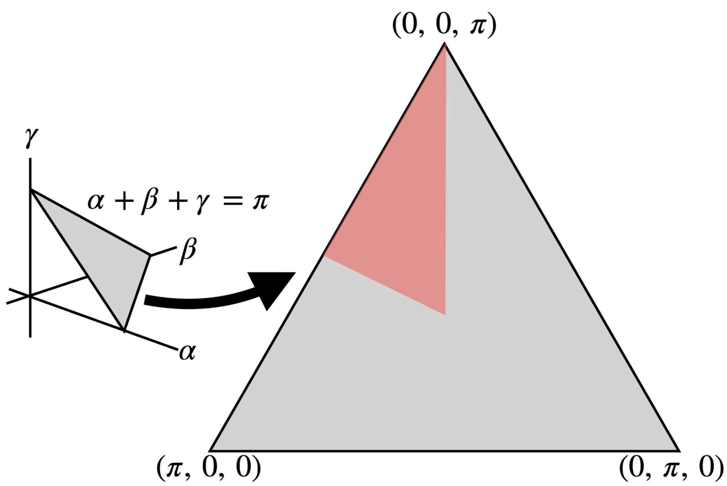


Figure 108. The space of triangles is itself a triangle.

On a LARGE picture of the space of triangles, sketch the following:

(a) the set of right triangles (green),

(b) the set of isosceles triangles (blue),

(c) all triangles with angles $0.12\pi, 0.35\pi, 0.53\pi$ (black dots),

(d) the set of all acute triangles (shaded).

141. In this representation of the space of all triangles, the angles are *marked* – we keep track of which angle is α and which is β, so the $(0.12\pi, 0.35\pi, 0.53\pi)$ triangle is different from the $(0.35\pi, 0.53\pi, 0.12\pi)$ triangle. This is clearly redundant, so we can instead represent the space of triangles with *unmarked* angles. This takes advantage of the *symmetries* of the space of triangles to "fold up" the space so that each triangle is only represented once.

(a) Explain why the space of unmarked triangles is represented by just the red shaded part.

(b) Imagine folding up the space of triangles (grey) along all of its lines of symmetry. Explain why this gives you just the red shaded figure. Triangles with the most symmetry lie at the edges of this smaller space. Explain.

142. (Continuation) *Where can we find the triangles we love?*

(a) We have seen that right triangles with a vertex angle of π/n unfold to (possibly double) regular polygon surfaces. Sketch the set R of these triangles on your picture.

(b) The Ward surface given in Problem 139 is the unfolding of the triangle with angles $\pi/16, \pi/4, 9\pi/16$. Explain. What triangle unfolds to the Ward surface given by a decagon with two pentagons? To the surface with a dodecagon and two hexagons? Mark these triangles, and the rest of the Ward family W, on the diagram.

The sets R and W are *discrete* in the space of triangles: for each triangle t of R or W, it is possible to find a little region in the space of triangles containing t, that does not contain any other point of R or W. The fact that they are discrete makes these and the rest of the Veech surfaces difficult to find!

(c) In the space of triangles, shade in the points that represent triangles that we *know* have a periodic billiard trajectory (see # 18). How much is left?

The zippered rectangle construction.

Figure 109 shows how to create a translation surface out of "zippered" rectangles. The idea is that you glue together some rectangles, and you also make some vertical cuts, like a zipper. As in the slit torus

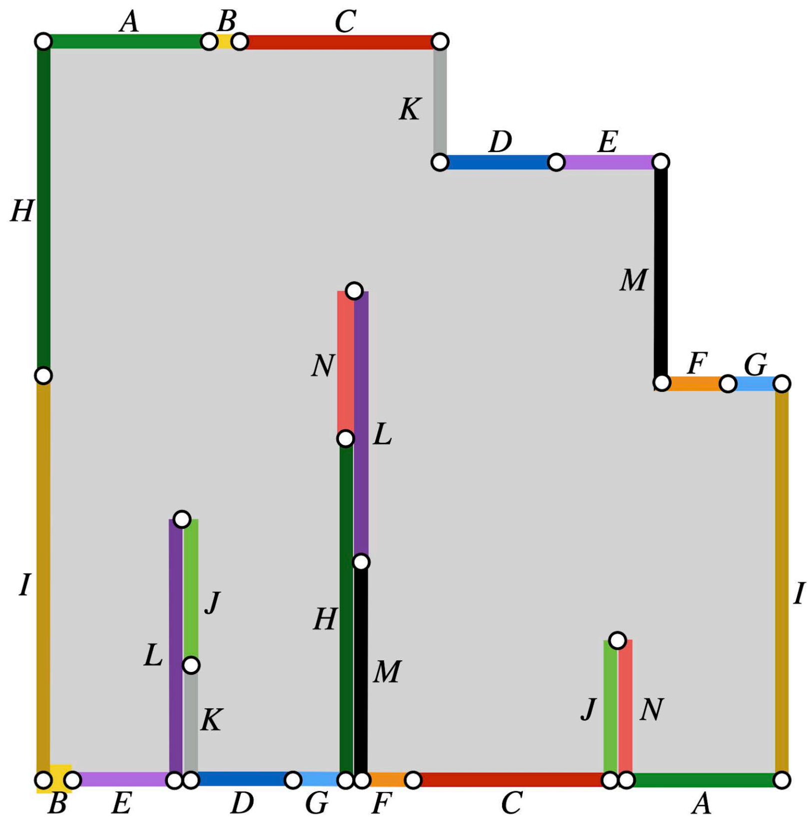


Figure 109. A rather complicated zippered rectangle.

construction (Problem 135), the two edges of each "zipper" are glued to different places.

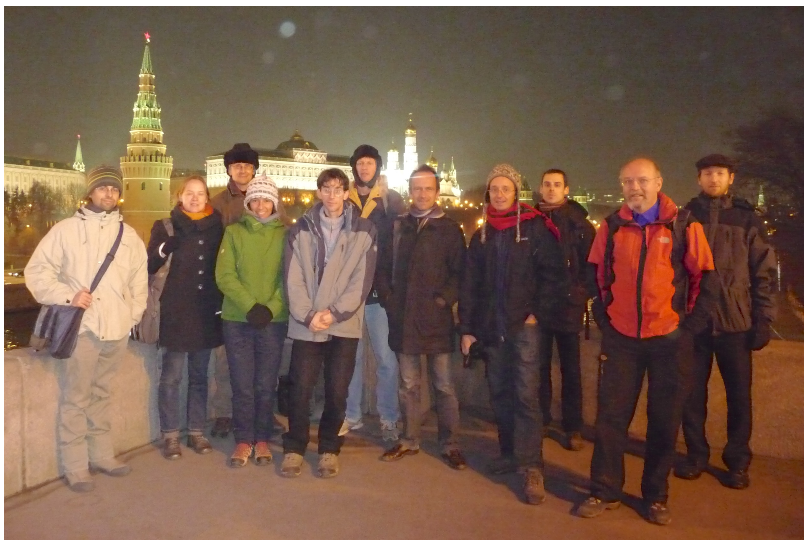
143. For the zippered rectangle surface in Figure 109:

(a) Consider the vertical flow on this surface. Show that its behavior is described by a 7-IET.

(b) Show that the surface has nine vertices: two with 6π of angle around them, and the rest with 2π of angle around them.

Notice that at the bottom of the zippers, the two flaps each have their own vertex point, to indicate that these are typically not identified with the same point. The corners have empty points; you should label them with colors or letters as you identify each vertex of the surface.

(c) Show that the surface has genus 3.



They did the math # 32. Pierre Arnoux

Pierre Arnoux (They did the math # 32) created the surface in Figure 109 to give a geometric realization of his eponymous Arnoux-Yoccoz IET (see Problem 167) [**1**]. The picture shows flat surface enthusiasts Corentin Boissy, Anna Lenzhen, Serge Troubetzkoy, Aslı Yaman, Samuel Lelièvre (# 30), Barak Weiss (# 24), Xavier Bressaud, Pascal Hubert (# 23), Luca Marchese, Pierre, and Alexey Glutsyuk in Moscow in 2012.

144. Recall the slit torus surface in Problem 135. Show that, for the vertical direction, the left part of the surface has a cylinder decomposition but the right part does not. This is another example of behavior that fails to satisfy the Veech dichotomy. What about the horizontal direction?

31. We get a little bit wild

145. *A wild translation surface.*
Figure 110 shows the *Chamanara surface* [**11**]. The edges have lengths $1/2, 1/4, 1/8, \ldots$. Parallel edges of the same length are identified, as shown. The pattern continues all the way into the corners.[5]

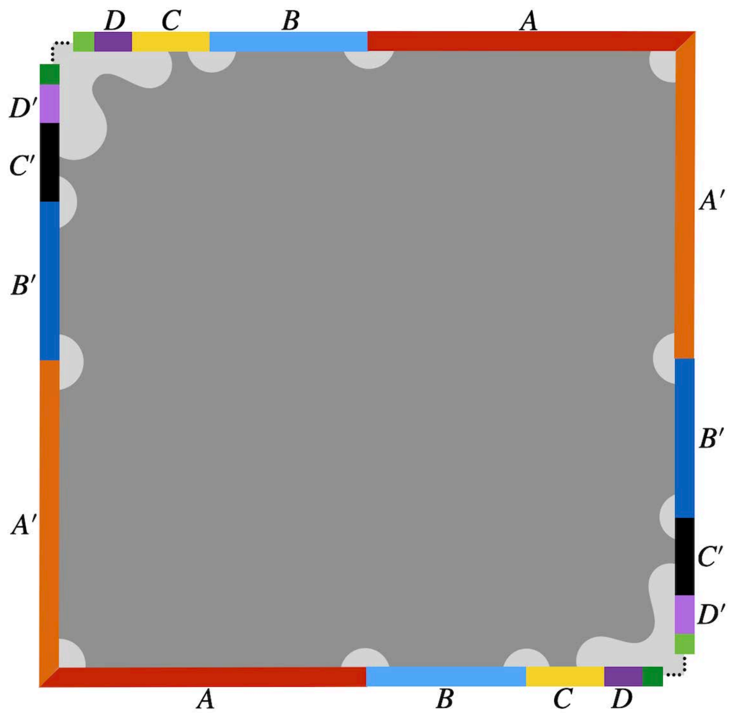


Figure 110. The Chamanara surface, which has infinitely many edges. Here, the surface has a dark blob encroaching on its vertices.

(a) Use vertex chasing (Problem 60) to show that the surface has two vertices.

(b) But wait – how far apart are the two vertices? Find a short path on the surface between the two vertices. How short of a path can you find? Hmmm...

[5] The content of this problem comes from Anja Randecker's thesis [**42**].

(c) Show that the Chamanara surface has a cylinder decomposition in the direction of slope 4, and that all of the cylinders in this direction have modulus 51/4. Indeed, show that for *every* integer n, it has a cylinder decomposition in the direction of slope 2^n.

(d) As you can see, this Chamanara surface has a dark blob that is gradually filling up the surface, avoiding but approaching the vertices, that is growing out towards the corners. Show that (contrary to appearances) the complement of this blob is connected! This makes the surface "wild."

They did the math # 33. Anja Randecker

Anja Randecker (THEY DID THE MATH # 33) studied wild translation surfaces for her Ph.D. thesis [**42**]. She determined that wild translation surfaces had an important property that no one had identified before, so she studied it, and named it *xossiness*: e*x*istence *o*f *s*hort *s*addle connections *i*ntersected *n*ot by *e*ven *s*horter *s*addle connections. The picture shows the author and Anja in Heidelberg in 2022.

In Problem 115, we found that the family of trajectories in a given direction (known as a *foliation*) on a particular translation surface has exactly the same behavior as a certain 3-IET. You might wonder: given *any* IET, can you find a translation surface, and a foliation direction, that matches the IET's behavior? Yes, you can, using a *suspension*.

Given any IET, do the following (Figure 111):

(1) First, for each break point in the top part of your IET, choose a "height" (possibly 0), and draw edges that attain each of the heights (top picture).

(2) Color-code your edges and translate copies of them corresponding to the bottom part of the IET (middle picture).

(3) Finally, make it into a translation surface (bottom picture).

Ta-da! You have a translation surface whose vertical foliation has exactly the same behavior as your IET.

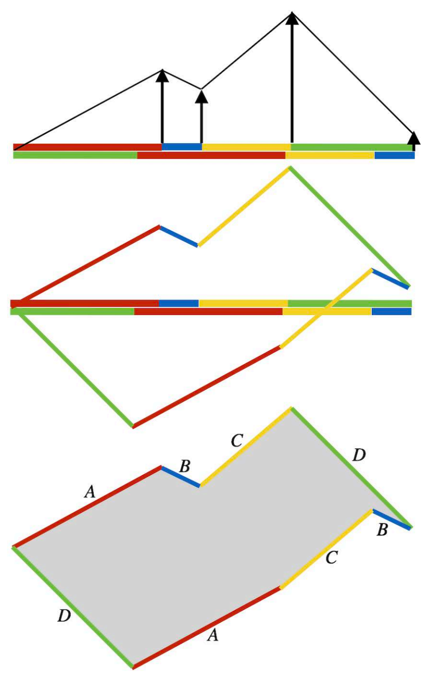


Figure 111. Suspending an IET to create a flat surface.

146. Make up an IET with at least four intervals, different from the above. Suspend it to create a corresponding translation surface, as described above.

147. You and your true love are looking into the same mirror. You are staring into each other's eyes, in the reflection (Figure 112). Are you both looking at the same point on the mirror? Or, considering that people have two eyes, we could ask: is it possible to draw a pair of glasses on the mirror so that each of you sees the other's eyes in the glasses?

Figure 112. Are these cuties looking at the same point on the mirror?

The biggest open problem in the study of Veech surfaces is: *Can we find more Veech surfaces?* and the related question, *Have we found them all yet?* Here are the families of Veech surfaces we have seen so far in this book:

- square-tiled surfaces (Problems 116, 122);
- regular polygons: double regular n-gons for any n, and single regular n-gons for even n (Problem 130);

- Ward surfaces: a regular $2n$-gon, with two regular n-gons glued to it along alternating edges (Problem 139).

For any $m \geq 2$, and any $n \geq 3$, the (m, n) Bouw-Möller surface is created by identifying opposite parallel edges of m semi-regular $2n$-gons. A *semi-regular polygon* is an equiangular polygon with an even number of sides, whose edge lengths alternate between two values, possibly equal and possibly 0. So that the cylinder moduli in Bouw-Möller surfaces are equal, the k^{th} semi-regular $2n$-gon has edge lengths alternating between $\sin \frac{k\pi}{n}$ and $\sin \frac{(k+1)\pi}{n}$ [**26**].

148. (a) Explain why a semi-regular $2n$-gon, half of whose edge lengths are 0, is a regular n-gon.

(b) The $m = 6$, $n = 5$ Bouw-Möller surface is shown in Figure 113. Edge identifications are indicated by numbers (for the reasoning behind the zig-zag edge-numbering system, see § 37). Shade each horizontal cylinder a different color. Does it seem plausible that all of the cylinders have the same modulus?

(c) For the $m = 4$, $n = 3$ Bouw-Möller surface: How many polygons does it have? How many edges does each polygon have, and what are their lengths? Sketch it.

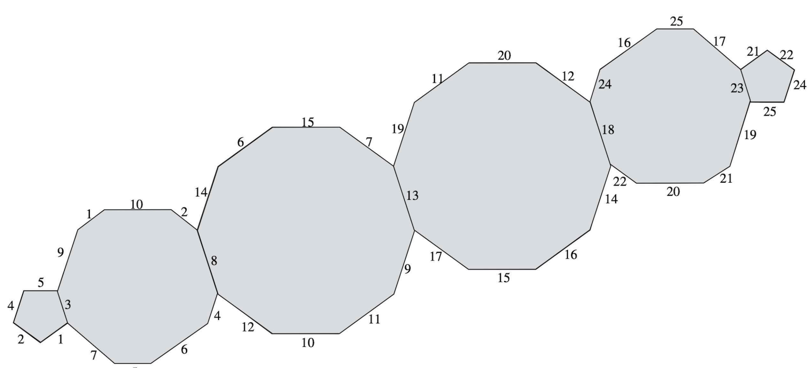


Figure 113. The $m = 6$, $n = 5$ Bouw-Möller surface.

In 2006, Martin Möller (They did the math # 34) and Irene Bouw showed that double regular n-gon surfaces (two polygons) and

They did the math # 34. Martin Möller

Ward surfaces (three polygons) are the simplest examples in a larger family of Veech surfaces with any number $m \geq 2$ of polygons, now called *Bouw-Möller surfaces* [**8**]. Irene and Martin gave an algebraic description of the surfaces, and later, Pat Hooper (# 22) found a polygonal description, as shown in Figure 113 [**26**]. The picture shows Erwan Lanneau, the author, and Martin running in Marseille in 2017.

32. Moving around in the space of surfaces

149. *The rel deformation.* Consider the translation surface in Figure 114, created by suspending an IET as in Problem 146.

(a) Confirm that the surface has two cone points, as suggested by the black and white dots.

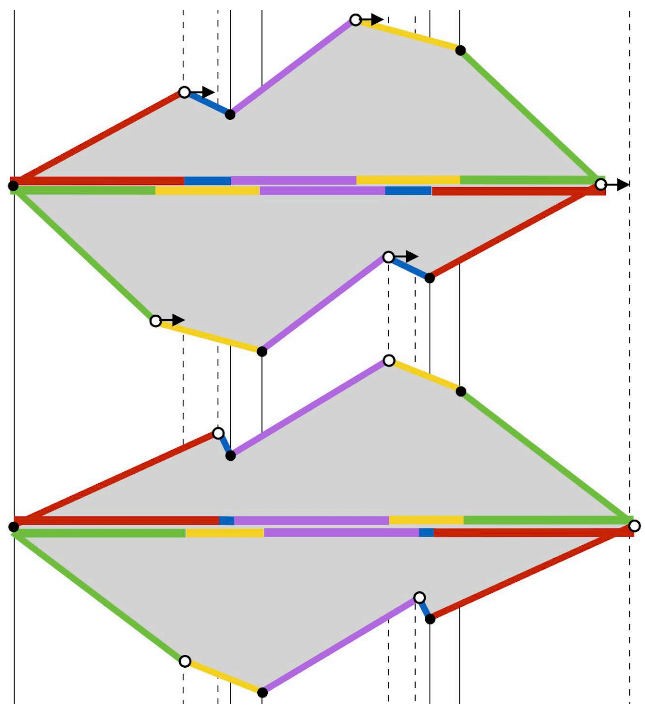


Figure 114. Moving the white vertices relative to the black vertices, a *rel deformation*, yields a nearby surface. Dashed lines pass through white vertices, and solid lines through black vertices.

One way to get a new translation surface "near" the original one is to deform the surface by moving one cone point *relative* to the other. This is known as a *rel deformation* [**27**]. The arrows in the top picture indicate that we will shift the white point slightly to the right. The bottom picture shows the surface after this deformation, along with the associated deformed IET.

(b) Explain what it means to be a “nearby” surface.

(c) How far can you push the white point to the right, and still create a valid translation surface? What about moving the white point in other directions – left, up, down, diagonally, etc.?

(d) Show that both of the above surfaces are in the stratum $\mathcal{H}(1,1)$ (recall Problem 87). The rel deformation is thus a way to move continuously among a family of surfaces in $\mathcal{H}(1,1)$. Explain.

150. Figure 115 shows the regular octagon surface (left), the double pentagon surface (center), and their singular friend (right).

(a) Show how to smoothly deform the regular octagon surface into the double pentagon surface.

(b) In Problem 87, you showed that both of these surfaces are in $\mathcal{H}(2)$. Suppose that we further deform the double pentagon surface into the double square surface (right). Explain why this surface is on the *boundary* of $\mathcal{H}(2)$. What kind of surface is it?

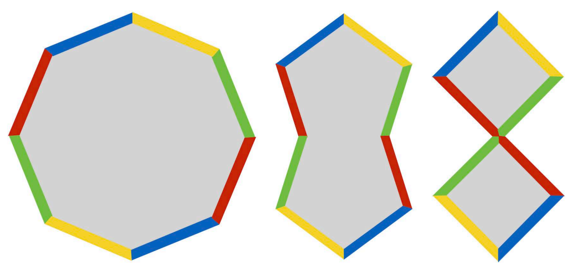


Figure 115. A belt-tightening operation on surfaces in $\mathcal{H}(2)$.

The above examples show that we can move around the space of surfaces in a given stratum. As we move around, most of the surfaces we encounter are like the one in Problem 149: “random” surfaces with no nontrivial automorphisms, or in other words, no rotations, reflections, or shears that preserve the structure of the surface.

On the other hand, the three surfaces in Figure 115 are Veech surfaces, with nice symmetries. But as we move in $\mathcal{H}(2)$ to get from one to the other, the surfaces we encounter in between are typically *not* Veech surfaces. As mentioned in Problem 141, Veech surfaces are discrete: we cannot move continuously among a family of Veech surfaces. This makes them difficult to find. When someone discovers a new family of Veech surfaces, it is a big deal.

In 2016, Curt McMullen (# 14), Ronen Mukamel (# 35), and Alex Wright (# 17) discovered the *gothic* family of Veech surfaces, so named because they look like the floor plan of a Gothic cathedral [**35**]. The edges have slope 0, ∞, and ± 1, and are identified as shown in Figure 116. Its dimensions are as indicated; the lengths a and b determine the surface.

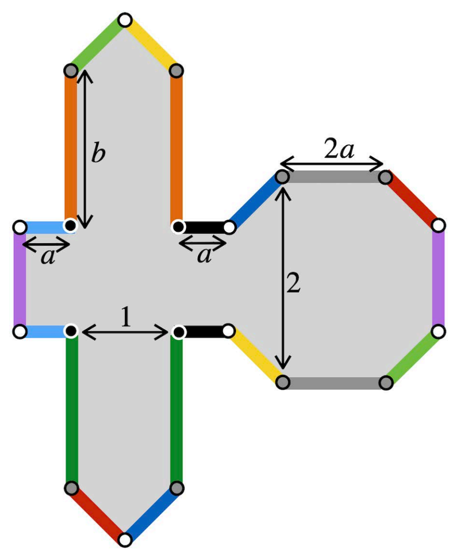


Figure 116. A blueprint for gothic Veech surfaces.

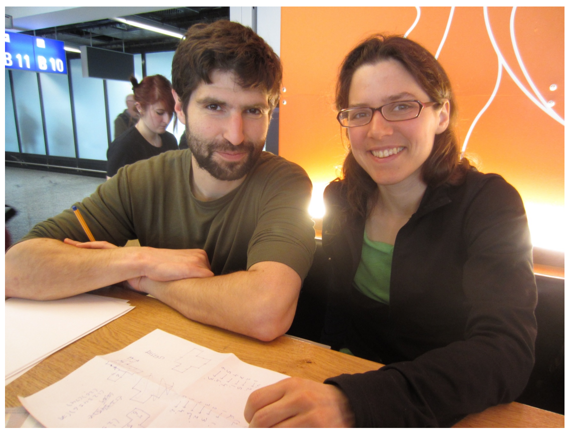
151. Show that each such surface has **(a)** five horizontal cylinders and five vertical cylinders; **(b)** three cone points, as indicated; and **(c)** genus 4.

The real key in showing that members of the gothic family are Veech surfaces is to carefully choose the measurements of a and b. It turns out that it is possible to choose rational numbers x, y and an integer $d \geq 0$ such that when

$$a = x + y\sqrt{d} \quad \text{and} \quad b = -3x - 3/2 + 3y\sqrt{d},$$

the surface is a Veech surface [**35**].

152. What stratum are the gothic surfaces in? Explain why this construction does not give a *continuous* family of gothic lattice surfaces in this stratum.

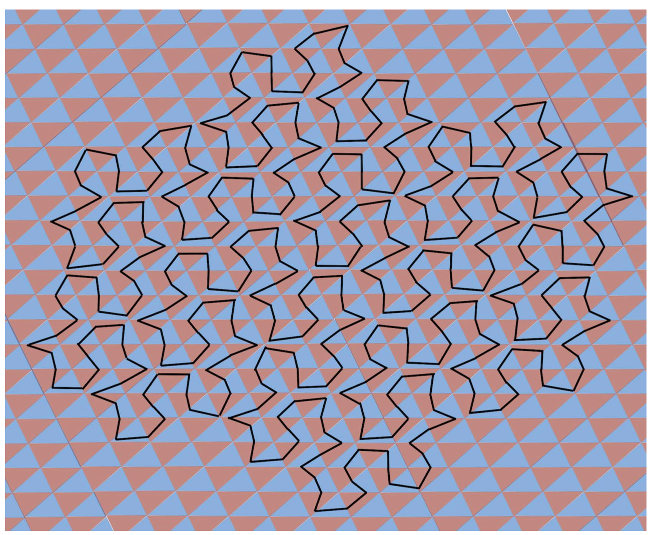


They did the math # 35. Ronen Mukamel

Ronen Mukamel (THEY DID THE MATH # 35) coauthored the result described above. He subsequently took a job working on computational biology and genetics. The picture shows Ronen with the author, pretending to do math in the Frankfurt airport in 2014.

Chapter 5

Further topics and tools



A long periodic tiling billiards trajectory that resembles the Rauzy fractal

Each of the sections in this chapter is a set of problems that explores a single topic or tool. Do you want to learn every single one of these things? Probably not! But if you happen to need one of these ideas, you'll be glad that I've written a set of problems about it. The best way to learn a new piece of mathematics is to work out some problems about it, so... let's get started.

Unlike in the rest of the book, these problems are not scaffolded or spaced out: problems on the given topic go one right after another, so it is key to understand one problem before working on the very next one. The average difficulty of the problems in this chapter is higher than in the rest of the book. Each section is independent of the others, except as noted.

You can do it!

33. The modular group

In Problem 140, we explored the space of all triangles. Now we'll explore the space of all tori. We'll do this by considering the space of surfaces made by gluing opposite parallel edges of parallelograms. In particular, if you can cut and paste one parallelogram surface into another while respecting the edge identifications, we'll consider those to be the same surface.[1]

Given any parallelogram, do the following (Figure 117):

- Translate and rotate the parallelogram until its short edge is on the x-axis, and the parallelogram lies above the x-axis.
- Scale the parallelogram so that its short edge has length 1 and coincides with the segment $[0, 1]$ on the x-axis.

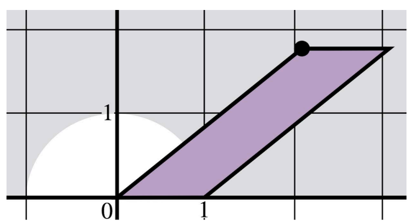


Figure 117. A normalized parallelogram.

153. Given a parallelogram translated, rotated and scaled (or "normalized") as described above:

(a) Show that its upper-left corner uniquely determines its shape.

(b) Show that the upper-left corner always lies outside the unit circle.

Now we want to mod out by cut-and-paste equivalence of parallelogram *surfaces*. To make this happen, we can cut and paste triangles (as suggested by Figure 118), while respecting the surface's edge identifications, to yield an equivalent surface represented by a different parallelogram.

[1]Thanks to Kathryn Lindsey for giving me a personal lecture on this topic in 2016, and to Samuel Lelièvre for helping to make my dreams for this section become reality.

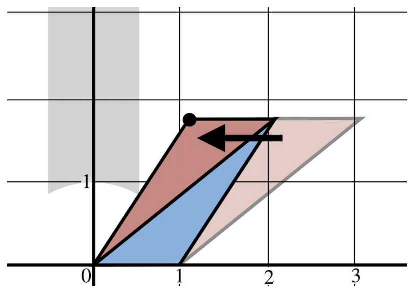


Figure 118. Cutting and pasting a parallelogram surface to yield, in some sense, exactly the same parallelogram surface.

154. Show that, under cut-and-paste equivalence, every normalized parallelogram is equivalent to a parallelogram whose top-left vertex lies in the infinite vertical strip $[-1/2, 1/2] \times [0, \infty]$. Justify the claim that every normalized parallelogram can be represented by a point in the shaded region of Figure 118, which is meant to extend infinitely upward.

The shaded region in Figure 118 is known as a *fundamental domain*. Problem 156 shows that we can choose any region shown in Figure 119 as our fundamental domain. People traditionally choose the one shaded in Figure 118.

155. It is possible that, after a cut-and-paste equivalence, the short side of your parallelogram is no longer the one on the x-axis, so you must switch edges and rescale. Give an example of such a parallelogram.

156. On Figure 119, mark all of the points in the upper halfplane that represent a 2×1 rectangle surface, or any surface equivalent to it under the actions described above. Two such points are marked for you. *Hint:* there is one corresponding point in each of the colored tiles. If you think of this as the hyperbolic plane, the points are reflections of each other across their hyperbolic geodesics boundaries. Can you explain why?

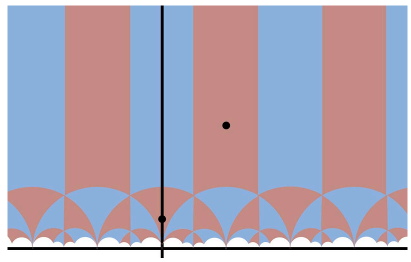


Figure 119. The upper halfplane, partitioned into fundamental domains for the space of parallelogram tori. In reality, infinitely many tiny regions cover the space all the way to the boundary; the picture only shows finitely many.

When working with the square torus, we used its Veech group, which is $SL(2, \mathbf{Z})$: the special (determinant 1) linear group with integer entries. Now we will work with $SL(2, \mathbf{R})$, the special linear group with real-valued entries. The *Iwasawa decomposition* says that any matrix in $SL(2, \mathbf{R})$ can be written as a product of matrices of the form K (compact), A (abelian) and N (nilpotent):

$$K = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \qquad A = \begin{bmatrix} \lambda & 0 \\ 0 & \frac{1}{\lambda} \end{bmatrix}, \qquad N = \begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix}.$$

We wish to understand how the elements of $SL(2, \mathbf{R})$ act on tori. In particular, we want to know how $SL(2, \mathbf{R})$ acts on the space of all tori that we defined in Problem 154. Since every matrix in $SL(2, \mathbf{R})$ can be written as a product of matrices of the form K, A, and N, the problem reduces to understanding the effects of these actions. Our normalization requires that one edge lies on the x-axis, so we ignore rotations, and focus on A (*geodesic flow*) and N (*horocycle flow*).

157. To see a beautifully animated view of the action of these flows on lattices, watch the short video *Shape of Lattices* by Pierre Arnoux and Edmund Harriss: `https://www.youtube.com/watch?v=vLrliPt4Uc0`. Then say which geometric actions described in the video correspond to K, A, and N, respectively.

158. Consider two of the points representing 2×1 rectangles, and the effect of horocycle flow and geodesic flow on them.

(a) Show that, for a point above $y = 1$, horocycle flow acts as indicated by the arrows in Figure 120: a push to the right.

(b) For each of the points, apply a tiny bit of horocycle flow, e.g., $\begin{bmatrix} 1 & 1/10 \\ 0 & 1 \end{bmatrix}$, and then normalize as described at the beginning of this section. What happens to the points?

(c) Do the same for a tiny bit of geodesic flow, e.g., $\begin{bmatrix} 11/10 & 0 \\ 0 & 10/11 \end{bmatrix}$.

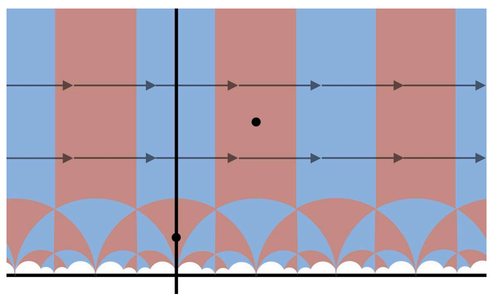


Figure 120. Understanding horocycle and geodesic flows.

They did the math # 36. Marina Ratner

Marina Ratner (THEY DID THE MATH # 36) proved several powerful theorems, which together are known as *Ratner's measure and orbit classification for unipotent flows on homogeneous spaces* [**43**–**45**].

These are key tools in the fields of dynamical systems and ergodic theory, and inspired lots of further work. In her Ph.D. thesis, she studied geodesic flows, and in her later work, she proved several important "rigidity" results about horocycle flows: precisely the two types of flows that we analyzed above. The picture shows Marina (front row) with François Ledrappier, Dmitry Kleinbock, Hillel Furstenberg, and Hee Oh in Banff in 2005. Marina died in 2017.

34. Renormalization

159. *Renormalization and the Rauzy gasket.* Consider a triplet of numbers (a, b, c), where $a, b, c > 0$ and $a + b + c = 1$. You can think of these points as living on the same triangular piece of the plane $x + y + z = 1$ as the space of triangles (Figure 108). Repeatedly perform the following algorithm:

(1) If
    $a > b + c$, subtract $b + c$ from a so that $(a, b, c) \mapsto (a - b - c, b, c)$.
    $b > a + c$, subtract $a + c$ from b so that $(a, b, c) \mapsto (a, b - a - c, c)$.
    $c > a + b$, subtract $a + b$ from c so that $(a, b, c) \mapsto (a, b, c - a - b)$.
    and if none of these are true, STOP.

(2) Rescale the values so that they sum to 1.

(a) Show that $(7/12, 4/12, 1/12) \mapsto (2/7, 4/7, 1/7) \mapsto (2/4, 1/4, 1/4)$.

(b) Let $\alpha \approx 0.54369$ be the real solution to the equation $x + x^2 + x^3 = 1$. Show that

$$(\alpha, \alpha^2, \alpha^3) \mapsto (\alpha^3, \alpha, \alpha^2) \mapsto (\alpha^2, \alpha^3, \alpha) \mapsto (\alpha, \alpha^2, \alpha^3),$$

so that this is a *periodic point*.

For most points, their iterated images eventually fail the condition that one element is greater than the sum of the other two, so the algorithm stops. But there are infinitely many points that can keep going in the algorithm forever; these points form a fractal set known as the *Rauzy gasket*, shown in Figure 121.[2]

Whoa.

The algorithm above is considered a *renormalization* algorithm, because at the end of each step, you "normalize" so that the sum of the coordinates is 1. In Problem 154, we normalized parallelograms. Renormalization algorithms are a powerful tool.

[2]Thanks to Samuel Lelièvre (# 30) for creating the Rauzy gasket picture with me.

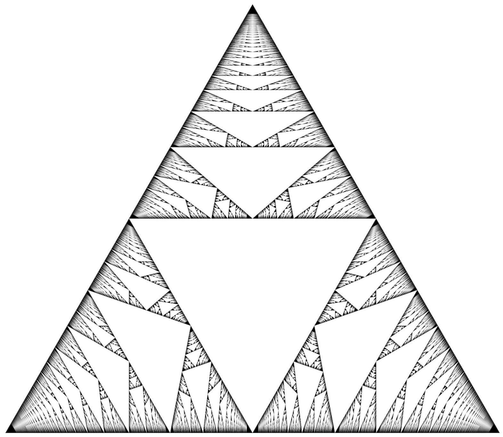
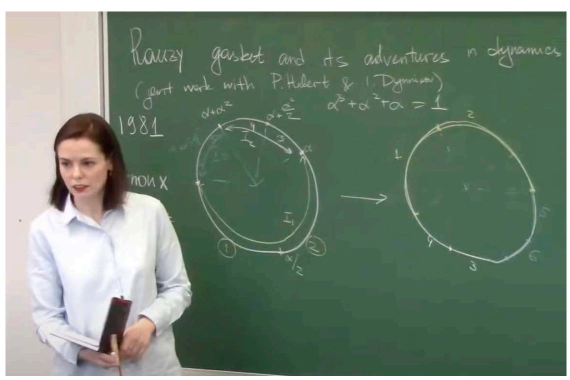


Figure 121. The awe-inspiring Rauzy gasket.

160. The algorithms that we performed on square torus trajectories (Problem 53) and on their corresponding cutting sequences (Problem 58) are also renormalization algorithms. Explain.



They did the math # 37. Alexandra Skripchenko

Sasha Skripchenko (They did the math # 37) studied the Rauzy gasket. In joint work with Pascal Hubert (# 23) and Ivan Dynnikov, she answered a question of Pierre Arnoux (# 32) to show that its Hausdorff dimension is less than 2 [**21**]. The picture shows

Sasha in Warwick in 2018, giving a talk about this work. The picture on the chalkboard is the Arnoux-Yoccoz IET (§ 36).

161. Let

$$M_1 = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad M_2 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \qquad M_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}.$$

(a) Given a triplet of numbers (a, b, c) where $a, b, c > 0$ and $a + b + c = 1$, show that you can implement the algorithm from Problem 159 by multiplying the column vector $\left[\begin{smallmatrix} a \\ b \\ c \end{smallmatrix}\right]$ by suitable *inverses* of M_1, M_2, and M_3.

(b) Consider a finite product $\overline{M}$ of M_1, M_2, and M_3 that includes at least one copy of each of the three (e.g., $\overline{M} = M_1 {M_3}^2 M_2 M_1$), and suppose that (a, b, c) has the property that

$$\overline{M} \left[\begin{smallmatrix} a \\ b \\ c \end{smallmatrix}\right] = \lambda \left[\begin{smallmatrix} a \\ b \\ c \end{smallmatrix}\right]$$

for some real number λ.[3] Show that (a, b, c) is a point in the Rauzy gasket.

(c) Show that *every* point of the Rauzy gasket can be obtained in this way.

(d) Explain why at least one copy of each of M_1, M_2, M_3 is needed.[4]

Hint: Notice that, for example, ${M_1}^{-1}(a, b, c) = (a - b - c, b, c)$, so if ${M_1}^{-1}(a, b, c) = \lambda(a, b, c)$, this is a fixed point of the Rauzy algorithm where the first entry is the largest. You can make a similar argument for a product of M_i's. Given a periodic point in the Rauzy gasket, if we write down the sequence of which entry is largest, we can associate to it a product of M_i's, and our point is an eigenvector of that product.

---

[3]In other words, $\left[\begin{smallmatrix} a \\ b \\ c \end{smallmatrix}\right]$ is an *eigenvector* of $\overline{M}$.

[4]Thanks to Vincent Delecroix (# 21) for explaining how to use the matrices in this problem to work with the Rauzy gasket. In the picture for #21, you can see a picture of the Rauzy gasket, and the matrix M_1 from Problem 161, on the chalkboard.

35. We find the Rauzy fractal in tiling billiards

162. In a *tribonacci* sequence, each term is equal to the sum of the previous *three* terms. Find the first 12 terms of the tribonacci sequence beginning $0, 0, 1, \ldots$.

163. *Substitutions.* Consider a sequence of words made out of two letters, a and b. We use the following substitutions:

$$a \mapsto ab, \qquad b \mapsto a.$$

(a) Compute the first eight terms of the sequence a, ab, aba, $abaab$, ...

(b) Show that the sequence of *lengths* of words is the Fibonacci sequence.

(c) Comment on any patterns you notice.

(d) Using the longest word you created above, plot a "broken line" in the following manner: start in the lower-left corner of a piece of graph paper, and when you read an a, step to the right, and when you read a b, step up. Plot the resulting walk.

Notice that the points stay close to a line of slope $1/\varphi \approx 0.618$.

164. Now consider a sequence of words made out of a, b, and c, with the substitutions

$$a \mapsto ab, \qquad b \mapsto ac, \qquad c \mapsto a.$$

Find the first 6 terms of the sequence a, ab, $abac$, ... and comment on any patterns.

Suppose that you use the resulting sequence to take a three-dimensional "walk" similar to the one in the previous problem, where a, b, and c tell you to take steps in the x-, y-, and z-directions, respectively. It turns out that, as in the previous problem, the points on this walk stay close to a line, now in 3D space. If we project these points in the direction of the line, onto a plane perpendicular to the line, we get a cluster of points that approach the *Rauzy fractal*, shown in Figure 122.

Figure 122. The Rauzy fractal.

165. *Finding the Rauzy fractal in tiling billiards trajectories.* Define $\alpha \approx 0.54369$ as in Problem 159, as the real solution to $x+x^2+x^3=1$. Consider tiling billiards on a triangle tiling with angles

$$\frac{\pi(1-\alpha)}{2} \approx 41.0679888577^\circ,$$
$$\frac{\pi(1-\alpha^2)}{2} \approx 63.396203173^\circ,$$
$$\frac{\pi(1-\alpha^3)}{2} \approx 75.535807969^\circ.$$

(a) Fire up the applet `https://awstlaur.github.io/negsnel/`, select "New Triangle Tiling [angles]", and type in two of the above angles. *Note:* things only get interesting when the angles are irrational, so enter all the digits listed above, to make the angles as close to irrational as possible.

(b) Move the green dot to the circumcenter of the triangle. You will have to approximate this as best you can. You will know when you are doing well because the trajectory will suddenly become very long.

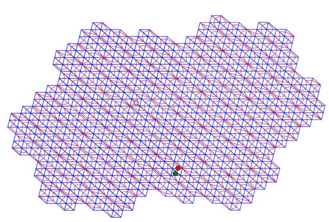


Figure 123. A large tiling billiards trajectory that bears a striking resemblance to the Rauzy fractal.

(c) Move the red and green points to make a trajectory that is as long as you can. If your path does not close up, remember to increase the iterations! Can you find a periodic trajectory larger than the one shown in Figure 123?[5]

As you find longer and longer trajectories, their appearance approaches that of the Rauzy fractal. Can you believe it?

Recall that in Problem 104, we explained that the orbit of a tiling billiards trajectory on a triangle tiling is equivalent to orbit of a point on a certain fully flipped circle exchange transformation (FFCET). For the triangle tiling whose angles are given in Problem 165, the associated FFCET is the Arnoux-Yoccoz IET, which we will see in the next section. Pat Hooper (# 22) suggested in 2016 that we look in this direction; he guessed that if we looked at the triangle tiling associated to the Arnoux-Yoccoz IET, we would probably find something interesting, and he was right.

166. The Rauzy fractal is a "tribonacci shape," in that three smaller copies of it join together to make one large copy of the same shape, as shown in Figure 124. Explain.

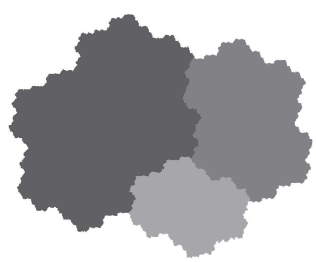
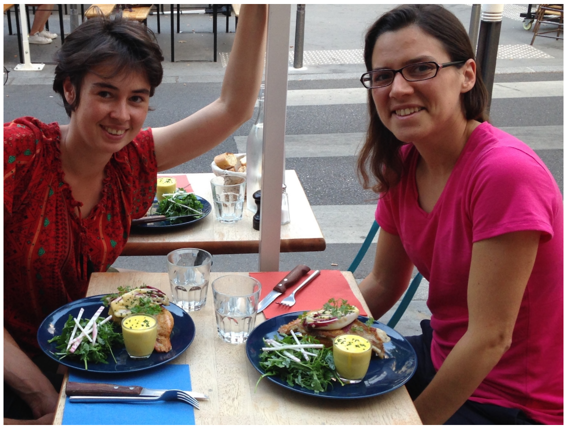


Figure 124. Three Rauzy fractals combine to make... another Rauzy fractal. Fractals out here exhibiting fractal behavior!

Our number $\alpha \approx 0.54369$ is just one point in the Rauzy gasket (Problem 159). It turns out that when you make a triangle tiling based on *any* point in the Rauzy gasket, tiling billiards trajectories

[5]See the second half of the video "Refraction Tilings" by Ofir David on YouTube at `https://www.youtube.com/watch?v=t1r1cO1V35I`.



They did the math # 38. Olga Paris-Romaskevich

passing near the circumcenter *always* give you fractal behavior. Olga Paris-Romaskevich (THEY DID THE MATH # 38) and Pascal Hubert (# 23) proved this, and many other tiling billiards results [**29**, **37**, **38**]. The picture shows Olga with the author in Lyon in 2018, about to partake of a tasty French meal.

36. The Arnoux-Yoccoz IET and arithmetic graphs

167. *The Arnoux-Yoccoz IET.* Define $\alpha \approx 0.54369$ as in Problems 159 and 165, as the real solution to $x + x^2 + x^3 = 1$.

(1) Divide the unit interval into 6 subintervals $I_1, I_2, \ldots I_6$ with consecutive lengths $\alpha/2$, $\alpha/2$, $\alpha^2/2$, $\alpha^2/2$, $\alpha^3/2$, $\alpha^3/2$.

(2) Switch the pieces of the same length.

(3) Cut the interval in half and switch the halves.

The construction is illustrated in Figure 125. (Notice that I_1 gets broken into two pieces.) We have seen this IET before, in another form: the vertical flow on the zippered rectangle surface in Problem 143 is the same as this IET. This transformation is *ergodic*, meaning that the orbit of every point fills in the space evenly [**1**].

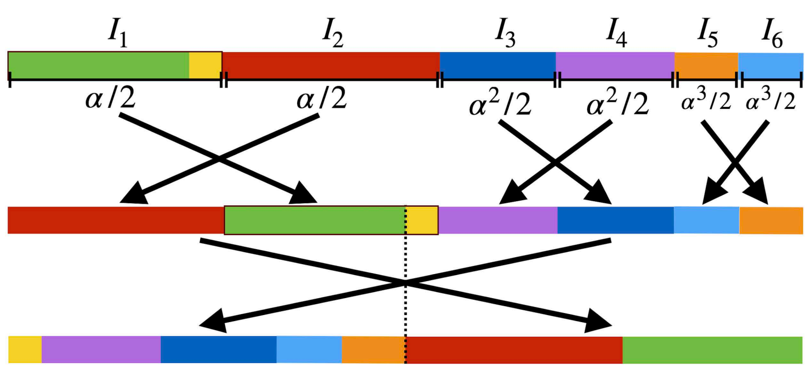


Figure 125. The famous Arnoux-Yoccoz IET.

(a) Choose a point on the interval, and follow its orbit for 10 iterations as it lands in intervals $I_1, \ldots, I_6$. Make a note of the sequence of intervals it ends up in, for use in the next problem. Does it seem plausible that the transformation is ergodic?

Written as a piecewise function, the Arnoux-Yoccoz IET is

$$f(x) = \begin{cases} f_1(x) = x - \frac{\alpha-1}{2} \bmod 1 & \text{if } x \in I_1 \approx [0, 0.27185] \\ f_2(x) = x + \frac{\alpha^3-1}{2} \bmod 1 & \text{if } x \in I_2 \approx [0.27185, 0.54369] \\ f_3(x) = x - \frac{\alpha^3-1}{2} \bmod 1 & \text{if } x \in I_3 \approx [0.54369, 0.69149] \\ f_4(x) = x + \frac{\alpha^2-1}{2} \bmod 1 & \text{if } x \in I_4 \approx [0.69149, 0.83929] \\ f_5(x) = x - \frac{\alpha^2-1}{2} \bmod 1 & \text{if } x \in I_5 \approx [0.83929, 0.91965] \\ f_6(x) = x + \frac{\alpha-1}{2} \bmod 1 & \text{if } x \in I_6 \approx [0.91965, 1] \end{cases}.$$

(b) Check that the orbit of your point under this function matches what you found in part **(a)**.

(c) The Arnoux-Yoccoz IET has seven intervals, but there are only six parts to the piecewise function above. Are you concerned?

We love being able to visualize the behavior of an IET in two dimensions. Efforts to do so that we have seen so far include graphing the associated piecewise function (Problem 123), tiling billiards (e.g., Problem 104), and suspending an IET (Problem 146) or making it part of a zippered rectangle (Problem 143). The form of the function above, with three pairs of related operations, suggests another method, of creating a "walk" in the plane. Let's do that.

168. (Continuation) Let $\mathbf{a}, \mathbf{b}, \mathbf{c}$ be the vectors shown on the left side of Figure 126, chosen so that $\mathbf{a} + \mathbf{b} + \mathbf{c} = \mathbf{0}$. Recalling the piecewise function $f(x)$ from Problem 167, we define a related function $g(z)$ on the plane:

$$g(z) = \begin{cases} g_1(z) = z + \mathbf{a} \\ g_2(z) = z - \mathbf{c} \\ g_3(z) = z + \mathbf{c} \\ g_4(z) = z - \mathbf{b} \\ g_5(z) = z + \mathbf{b} \\ g_6(z) = z - \mathbf{a} \end{cases}.$$

We choose any number $x \in [0, 1]$, and we start at any point z in the plane. We iterate the Arnoux-Yoccoz IET on x, by applying some sequence of functions f_i, for example $f_2, f_5, f_1, \ldots$, to x. At the same time, we apply the corresponding functions g_i, e.g., $g_2, g_5, g_1, \ldots$, to z and its images. This amounts to adding the vectors $\pm\mathbf{a}, \pm\mathbf{b}, \pm\mathbf{c}$, yielding a walk on the triangular grid.

Plot the walk corresponding to the orbit of your point from Problem 167 on an equilateral triangle grid. Does it close up?

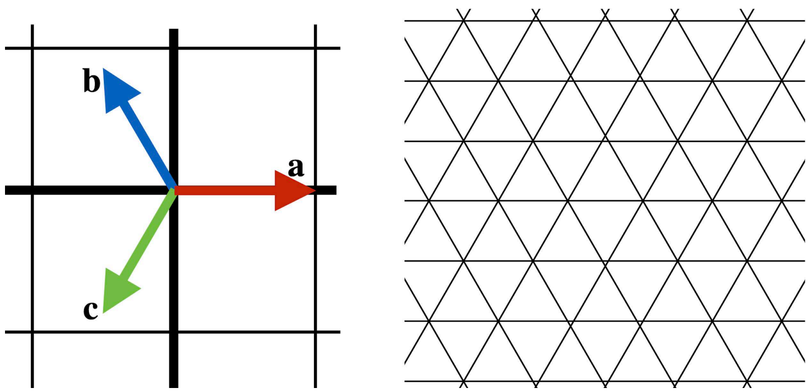


Figure 126. Equilateral triangle vectors, and a grid.

The "walk" shown above is known as an *arithmetic graph*.[6] Rich Schwartz (# 3) has made significant use of arithmetic graphs in his exploration of the behavior of outer billiards on polygons, and in his proof that every triangle whose largest angle is less than 100° has a periodic inner billiard orbit [**47**, **48**].

169. The Arnoux-Yoccoz IET is pictured in Figure 127, with our usual convention of "flow up, then shift when you come down." As the arrows suggest, we are applying a rel deformation (recall Problem 149). We leave the black vertices fixed, and shift the white vertices to the right, as the picture shows.

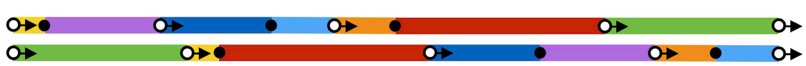


Figure 127. A rel deformation of the Arnoux-Yoccoz IET.

[6] Since "arithmetic" is used as an adjective here, it is pronounced air-ith-MET-ic. This follows the same differential adjective/noun syllable stress pattern as e.g. "I'll reCORD a REcord, and disCOUNT it with a DIScount."

(a) Check that the picture is consistent, e.g., the red subinterval has a black vertex on its left end and a white vertex on its right end, for both copies.

(b) Also check that the picture is consistent at the endpoints (0 and 1) of the interval.

They did the math # 39. Gérard Rauzy

The Rauzy fractal, which we saw in §35 and will imminently see again, is named for Gérard Rauzy (THEY DID THE MATH # 39). Gérard was especially interested in Fibonacci and tribonacci substitutions, and discovered his eponymous fractal [**46**]. The picture shows a group of mathematicians at Pierre Arnoux's (# 32) apartment in Marseille in 2006: (back row) Théodore Tapsoba, Julien Cassaigne, Christian Mauduit, Jun-Ichi Tamura, Shunji Ito, Sébastien Ferenczi, Teturo Kamae, Hiromi Ei; (front row) Hiroko Kamae, Gérard Rauzy, and Geneviève Macquart-Moulin.

It so happens that any walk on the triangular grid corresponding to the orbit of a point on the Arnoux-Yoccoz IET, such as the one you computed in Problem 168, is unbounded. On the other hand, if you change the Arnoux-Yoccoz IET via a rel deformation by some tiny amount r, any walk corresponding to the orbit of a point on the rel-deformed IET is periodic, with larger periods as $r \to 0$ [**27**].

170. (a) Let $r = 0.03$. Write out the rel-deformed Arnoux-Yoccoz IET, which is a modification of the six-part function in Problem 167.

(b) Compute the orbit of the point $x = 0.4$ on the rel-deformed IET. Show that its period is 17.

(c) Show that the arithmetic graph corresponding to the orbit of $x = 0.6$ is as shown in Figure 128. The red point is the starting point.[7]

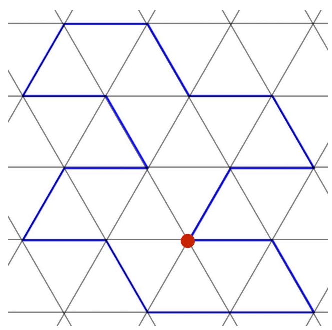


Figure 128. An arithmetic graph corresponding to a walk based on the rel-deformed Arnoux-Yoccoz IET.

The eight shortest arithmetic graphs corresponding to this construction, and the 15$^{\text{th}}$-shortest, are shown in Figure 129 [**27**]. It's the Rauzy fractal again!

[7]Thank you to Pat Hooper (# 22) for helping to make my dreams for this section become reality.

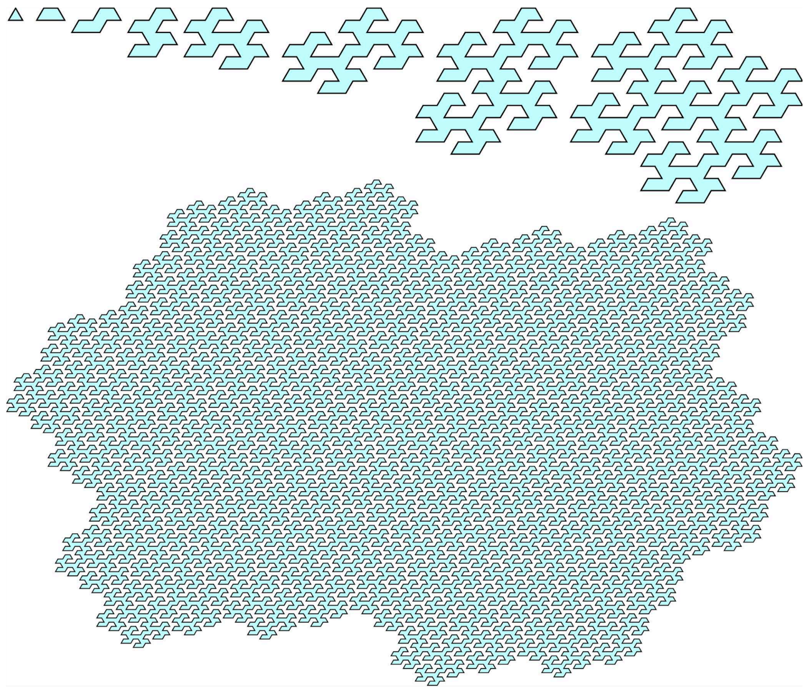


Figure 129. Long walks on the... Rauzy fractal!

37. Transition diagrams

Recall our friend the square torus, with horizontal and vertical edges labeled A and B, respectively. We have explored many ways of understanding linear trajectories on the square torus, including transforming the geometric problem about trajectories into a combinatorial problem about cutting sequences. *Transition diagrams* inject some geometry back into those cutting sequences, via a flow chart of which edge labels can follow which others.

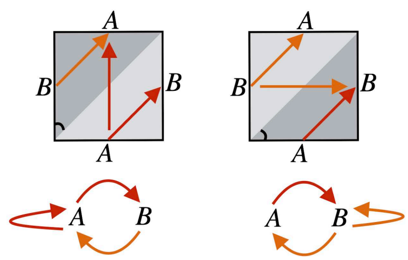


Figure 130. Transition diagrams for the square torus.

In the top-left picture of Figure 130, we have restricted trajectories to those that go left to right with slope ≥ 1. For such trajectories, A can be followed by A or B (red arrows), and B can only be followed by A (orange arrow). We represent this information using the transition diagram shown at the bottom left of the figure. For the top-right picture, we use trajectories with slope between 0 and 1, and the situation is similar: A can only be followed by B (red arrow), while B can be followed by either A or B (orange arrows).

Now that we have this transition diagram, an alternative way to do Problem 38 is to say: "A cutting sequence with AA is valid only on the first transition diagram, and a cutting sequence with BB is valid only on the second transition diagram, so no valid cutting sequence can have both AA and BB."

Now, let's look at a different surface: the double regular pentagon. For the double pentagon, the symmetries of the surface allow us to restrict our attention to trajectories with angle between 0 and $\pi/5$, indicated by the shaded sectors in Figure 131.

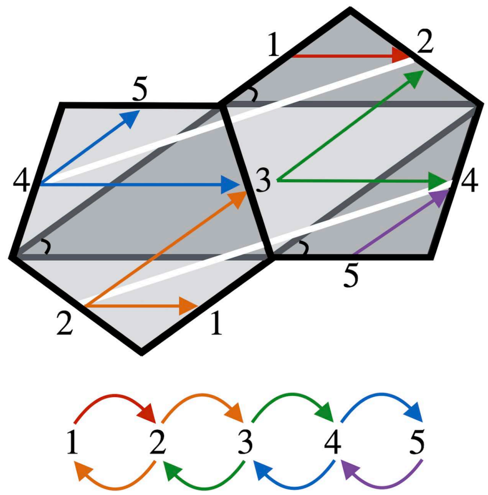


Figure 131. Edge transitions for the sector $[0, \pi/5]$ on the double pentagon.

171. For trajectories within the sector $[0, \pi/5]$ of directions:

(a) First, consider a trajectory that passes through edge 1. Check that edge 1 can only be followed by edge 2 (red arrow), because going to any other edge would require going down (e.g. to edge 4 in the right pentagon), or going too steeply upward (e.g. to edge 3 in the left pentagon). By similar logic, argue that edge 2 can only be followed by edge 1 or edge 3 (orange arrows), and so on.

(b) Work through all five edges, and confirm that the *transition diagram* below the surface accurately reflects the allowed *transitions* for cutting sequences corresponding to such trajectories.

(c) Confirm that the cutting sequence $\overline{2343}$ corresponds to a valid periodic trajectory on the surface (white), and also that it corresponds to a periodic path on the transition diagram.

172. Draw the transition diagram for the double pentagon corresponding to the set of trajectories whose angle is between $\pi/5$ and $2\pi/5$. *Hint*: it is similar in form to the one we computed above.

The purpose of transition diagrams is that they allow us to determine the direction of a trajectory without having to draw a picture. For the square torus, if a cutting sequence is valid on the left transition diagram in Figure 130, its slope is greater than 1, while if it is valid on the right transition diagram, the slope is less than 1.

Similarly, for a cutting sequence corresponding to a trajectory on the double pentagon, you can determine which of the five transition diagrams – one was given, you drew a second, and there are three more – it is valid on, and this will tell you the trajectory's direction. This reduces the geometric problem about trajectories and surfaces to a combinatorial problem about symbols, which is much easier to characterize and check.[8]

173. Ooh, now things get more interesting! Recall the Ward surface from Problem 139. We restrict to angles between 0 and $\pi/8$, as suggested by the darkened sectors in Figure 132.

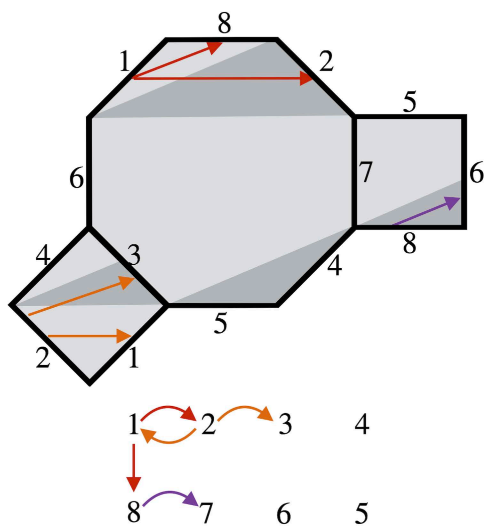


Figure 132. Building a transition diagram for the octagon-square Ward surface.

[8] Everything I know about transition diagrams, I learned from a paper by John Smillie and Corinna Ulcigrai (# 8), *Symbolic coding for linear trajectories in the regular octagon* [**55**], which defined them as above and worked out the example of the square torus in detail; see their § 1.2 and Figure 2.

(a) A few of the arrows on the surface, and on the corresponding transition diagram, are given in Figure 132; you fill in the rest.

(b) The "zig-zag" numbering system looks a little unnatural on the double pentagon surface and on the octagon-square Ward surface, but it sure does make the transition diagram turn out nicely. Explain.

The zig-zag numbering can also make the cutting sequence itself very pleasing. To see an example of this, look at the example of a cutting sequence on the double pentagon given in the Glossary entry for "cutting sequence." Follow along the trajectory and read off the cutting sequence, and see how nice it is!

I came up with the zig-zag numbering scheme myself, and I think it's awesome [**16**]. Before I discovered zig-zag numbering, my transition diagrams for Ward and Bouw-Möller surfaces were a jumbled mess and I couldn't understand what was going on. But with the zig-zag numbering, the transition diagrams have a clear rectangular structure, so I was able to prove theorems about them. This is an example of where the right notation makes a *huge* difference.

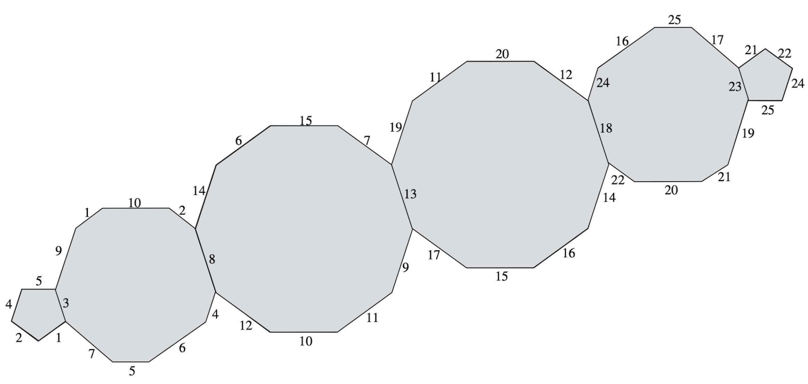


Figure 133. The $m = 6$, $n = 5$ Bouw-Möller surface, with edge identifications indicated by tiny numbers in the zig-zag numbering system.

174. (Challenge) Recall the $m = 6$, $n = 5$ Bouw-Möller surface from Problem 148, shown in Figure 133.

(a) What should the angle restriction be on trajectories for this surface, given its symmetries?

(b) Fill in the rest of the transition diagram in Figure 134. The locations of some of the nodes are given; you figure out where the rest should go, and also the arrows.

Figure 134. A structure for the transition diagram for the $m = 6$, $n = 5$ Bouw-Möller surface, with most of the information left to fill in.

38. Rauzy-Veech induction

Rauzy-Veech induction is a tool for simplifying, and thus better understanding, the behavior of IETs [**59**]. We'll use the example of the 4-IET from Problem 123, shown on the top line of Figure 135.

- First, choose one end of the IET or the other. Typically, people choose the right end, as we do here.
- Of the two subintervals that are at the right end, one is longer than the other; this interval is said to "win," and the other one is said to "lose." For our IET, D wins and B loses.
- Chop off the end of the IET, the length of the losing interval. Follow the path of that losing interval for one iteration: Here, B becomes the right end of D. So in the next line, we replace the right end of D with B, to get a new IET.
- Repeat this process with the new IET, over and over.

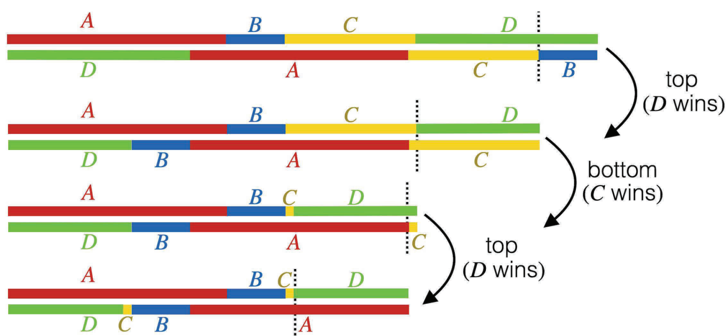


Figure 135. Three steps of Rauzy-Veech induction on a 4-IET.

175. Draw the diagrams for the next three steps of Rauzy induction[9] on the above IET. Notice that the only intervals that change are the ones that are "fighting": for example, in the first step above, D and B are fighting, so the red and yellow intervals A and C stay exactly the same from the first picture to the second. Can you explain why?

[9]Sometimes, people drop the "Veech." Recall the footnote for # 29.

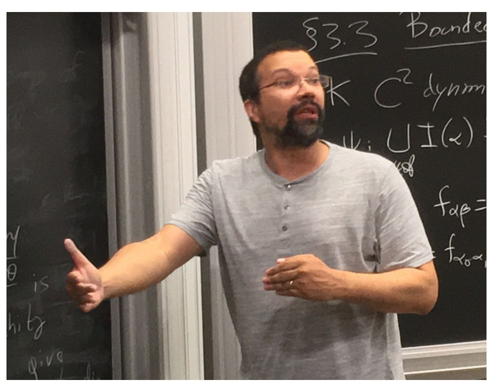
176. In Figure 136, the length ratio of segments A to B is 7:5. Perform Rauzy induction on this IET until both segments are the same length. Then compare with your work in Problems 8 and 21 and explain the relationship between Rauzy induction and the continued fraction algorithm.

Figure 136. A 2-IET for practicing Rauzy-Veech induction.

We want to know if we get back to an IET that is combinatorially "the same" as the one we started with. For the 4-IET in Figure 135, we can keep track of the Rauzy induction steps via the following notation:

$$\begin{pmatrix} A & B & C & D \\ D & A & C & B \end{pmatrix} \xrightarrow[\text{(D wins)}]{\text{top}} \begin{pmatrix} A & B & C & D \\ D & B & A & C \end{pmatrix} \xrightarrow[\text{(C wins)}]{\text{bottom}} \begin{pmatrix} A & B & C & D \\ D & B & A & C \end{pmatrix} \xrightarrow[\text{(D wins)}]{\text{top}} \begin{pmatrix} A & B & C & D \\ D & C & B & A \end{pmatrix}$$

177. Write down the notation for the next three steps of Rauzy induction, corresponding to your diagrams from Problem 175.



They did the math # 40. Carlos Matheus

Carlos Matheus (THEY DID THE MATH # 40) works on dynamical systems and number theory, including IETs, translation surfaces, orbits, strata, and many of the other objects introduced in this book.

I took the content of Problems 176 and 178 from a talk he gave[10] in July 2023 in Marseille, shown in the picture.

178. Starting from the IET represented by $\left(\begin{smallmatrix} A & B & C & D \\ D & C & B & A \end{smallmatrix}\right)$, shown in the middle of the diagram in Figure 137, there are two options: either the top wins, or the bottom wins. Then, starting from each of *those* IETs, there are two options: again, either the top or the bottom wins. And so on! We can work out the entire Rauzy-Veech diagram for all of the options, which ends up looking like the below "butterfly." The right side of the diagram is already done; fill in the grey blanks on the left side to complete it.

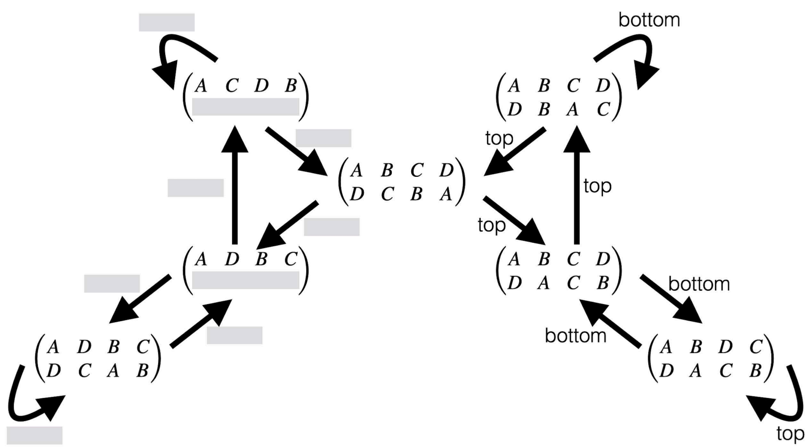


Figure 137. All possible length ratios, recorded in one diagram.

[10] Thank you, Matheus!

Appendix A

A short primer on matrix actions in the plane

Many of the problems in this book ask you to transform the plane using the action of 2×2 matrices. If you're not familiar with this sort of thing, fear not: the following problems are for you.

An introduction to matrices

A1. Given two vectors $[a, b]$ and $[c, d]$, their *dot product* is the scalar value $a \cdot c + b \cdot d$. Confirm that $[1, 3] \bullet [2, 8] = 26$.

A2. A 2×2 matrix stores information in *rows* and *columns*. When we multiply two matrices, each entry in the product matrix is the dot product of the corresponding *row* of the left matrix and the corresponding *column* of the right matrix. Figure 138 shows how to compute the four entries of the matrix product $\left[\begin{smallmatrix} 1 & 3 \\ 5 & 7 \end{smallmatrix}\right] \cdot \left[\begin{smallmatrix} 2 & 4 \\ 8 & 16 \end{smallmatrix}\right]$.

(a) Compute all four entries of this matrix product, and explain how to see each one as a dot product.

(b) We often want to use a 2×2 matrix to transform a vector, which we can think of as a 2×1 matrix. For example, we have already done all of the work to compute the product $\left[\begin{smallmatrix} 1 & 3 \\ 5 & 7 \end{smallmatrix}\right] \cdot \left[\begin{smallmatrix} 2 \\ 8 \end{smallmatrix}\right]$. Explain.

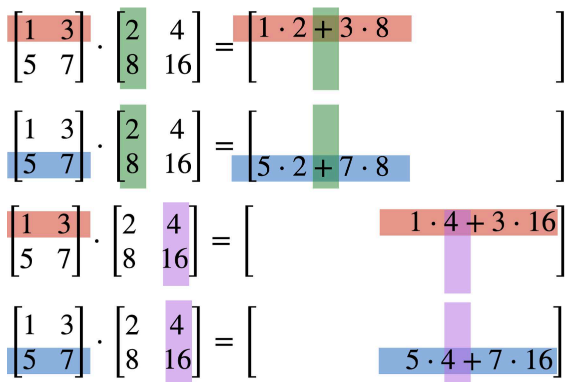


Figure 138. How to multiply 2×2 matrices.

A3. Which are your favorite, rows or columns? Definitely columns. Check this out:

$$\begin{bmatrix} 1 & 3 \\ 5 & 7 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} 1 & 3 \\ 5 & 7 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 7 \end{bmatrix}.$$

Wow, it sure looks like multiplying a 2×2 matrix by $\left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$ gives us the first column of the matrix, and multiplying it by $\left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right]$ gives us the right column. Is this always true? Prove it or find a counterexample.

Terminology: We say that the matrix above "takes $\left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$ to $\left[\begin{smallmatrix} 1 \\ 5 \end{smallmatrix}\right]$ and takes $\left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right]$ to $\left[\begin{smallmatrix} 3 \\ 7 \end{smallmatrix}\right]$."

A4. Make up a 2×2 matrix that takes $\left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$ to $\left[\begin{smallmatrix} 2 \\ 3 \end{smallmatrix}\right]$ *and* takes $\left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right]$ to $\left[\begin{smallmatrix} -1 \\ 2 \end{smallmatrix}\right]$. How many different correct answers are there to this problem?

The geometry of matrix transformations

Algebra is fun and all, but it's time to do some geometry!

A5. Our 2×2 matrices transform the plane. One way to understand *how* a matrix transforms the plane is to apply it to a simple shape and see what happens. Let's apply a couple of our favorite transformations to the "unit square" $[0, 1] \times [0, 1]$ shown in Figure 139.

(a) Apply the matrix transformation $\left[\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right]$ to the unit square in Figure 139. To do this, apply the matrix to each of its four vertices $\left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right]$, $\left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$,

$\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]$, and $\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$, one at a time, and plot the color-coded image points. Shade in the resulting polygon that is the image of the square.

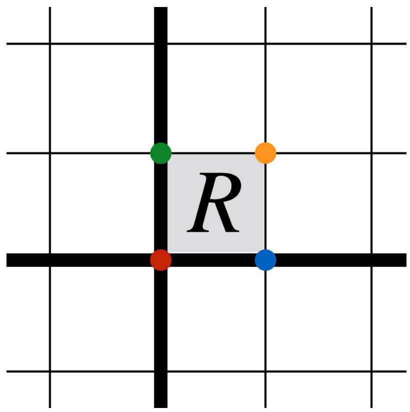


Figure 139. Let's practice transforming points using a matrix.

(b) Repeat the above for the matrix transformation $\left[\begin{smallmatrix}1&0\\1&1\end{smallmatrix}\right]$. Comment on similarities and differences with the previous part.

(c) Repeat the above for the matrix transformation $\left[\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right]$.

(d) As in Problem 6, we have put an R on our square so that we can see how it has moved. Use the color-coded vertices to draw the transformed (possibly stretched-out) image of the R on each of your image polygons. A transformation that flips the R backwards is called an *orientation-reversing* transformation, and otherwise the transformation is *orientation-preserving*. For the transformations $\left[\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right]$, $\left[\begin{smallmatrix}1&0\\1&1\end{smallmatrix}\right]$ and $\left[\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right]$ used above, which ones preserve, and which ones reverse, the orientation?

(e) If you did all of your calculations correctly, you should have found that the red point at the origin is fixed (does not move) under all of the transformations. Can you explain why?

A6. The *inverse* of a transformation M is another transformation M^{-1} that "un-does" the action of M.

(a) Looking at your *transformed* polygon from Problem A5**(a)**, apply the matrix $\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$ to each of its points, and show that the result is back to the original square.

(b) Find a matrix that performs the inverse action of $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$.

(c) Find a matrix that performs the inverse action of $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

For **(b)** and **(c)**, make sure to apply your proposed inverse matrix to your transformed polygon from Problem A5 and check that you get the original unit square back.

A7. Let's transform the plane with more exotic transformations.

(a) Sketch the image of the square shown in Figure 140, under the transformations $\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ and $\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$, respectively.[1]

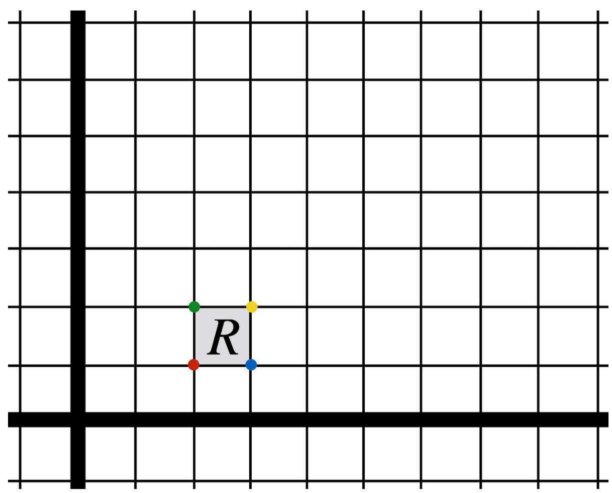


Figure 140. Let's transform a shape that isn't at the origin.

(b) Find the area of the transformed image of the square under each of the two transformations. *Hint*: To find the area of a complicated shape whose vertices are at lattice points, inscribe the shape in a lattice rectangle. Then the area of your shape is the area of that rectangle minus the areas of rectangles and right triangles that are inside the rectangle but outside of your shape.

[1]The matrix $\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$ is known, for historical reasons, as the *cat map*.

A8. For a 2×2 matrix $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$, its *determinant* is the number $ad - bc$.

(a) Compute the determinant of each of the matrix transformations we have studied so far: $\left[\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right]$, $\left[\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right]$, $\left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$, $\left[\begin{smallmatrix} 2 & 0 \\ 0 & 2 \end{smallmatrix}\right]$, and $\left[\begin{smallmatrix} 2 & 1 \\ 1 & 1 \end{smallmatrix}\right]$.

(b) The *magnitude* of the determinant is the area expansion factor of the associated transformation. Check against your answers to Problems A5 and A7 that, for the matrices with determinant ± 1, the area of the transformed polygon is the same as that of the original square, and that for the matrix with determinant 4, the associated transformation expands the area by a factor of 4.

(c) The *sign* of the determinant indicates whether the associated transformation is orientation-preserving $(+)$ or orientation-reversing $(-)$. Check that for the orientation-preserving transformations, the associated matrix determinant is positive, and that for the orientation-reversing transformation it is negative.

A9. Suppose that we want to perform several matrix transformations, one after another:

(a) Apply the transformation $\left[\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right]$ to $(1, 0)$, and plot the result on your graph paper.

(b) Now transform the result by $\left[\begin{smallmatrix} 2 & 1 \\ 1 & 1 \end{smallmatrix}\right]$.

(c) Now transform the result by $\left[\begin{smallmatrix} 1 & -1 \\ 0 & 1 \end{smallmatrix}\right]$.

(d) Now transform the result by $\left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$. Where do you end up?

(e) Starting with the point $(0, 1)$ on a new picture, perform the same series of transformations.

(f) There must be a more efficient way to do this, right? Show how to compute a single 2×2 matrix that performs all four actions at once.

Appendix B

How to teach a problem-centered course

I wrote this book for use in a problem-centered, discussion-based course. For a lot of people, teaching a discussion-based math course is a new idea. *Can* I do it? Do I even *want* to do it?

Yes, and yes!

You will have to give up some things: mostly, control. Students will solve problems using different methods than you expect, and they may explain their reasoning differently than you would have. But by giving up control, you can bring tremendous intellectual vigor to your class. Students wrestling together on a tough problem, and then figuring it out, is magic. Give yourself a chance to experience this magic.

How should the instructor prepare?

Do the problems. Yes, the best way to prepare to teach this class is simple, though not necessarily easy: do the problems. Everything that is required for the course is contained in the problems, and subsequently in the connections made in your students' fertile minds. You do not need to read outside literature to prepare, and your students do not need you to give expository lectures. I promise!

If possible, it's nice if you can do all of the problems before the term starts, or at least be a couple of sections ahead of your students, just so that you will feel more confident in yourself, and know where the story is going. You don't need to *tell* your students where the story is going, just as you don't need to *tell* your students how to do the problems, but doing the problems in advance will help you to *feel* better about things. Problems that look challenging on a first read may turn out to be easier than you expected, or perhaps more difficult. The problems require a lot of drawing pictures, counting, and doing explicit examples, and you will find that if you actually do these things, many patterns and insights become clear.

What if you don't get to all of the problems, or what if there is (*gasp*) a problem that you can't solve? So much the better! You can actually teach a class like this without knowing the material. I know it sounds crazy, and I wouldn't recommend it as a strategy for every day, but the benefit of the instructor not knowing how to solve the problems is that *you will let the students do the work.* How wonderful it is to see a student explain their solution to a problem that you couldn't crack! If there's a tough problem, and none of the students were able to solve it yet, *let them struggle* and work towards figuring it out. Send them up to the board in partners and have them work on it together.

It's a wonderful experience for students to work together on something challenging, and the payoff from the students learning that they *can* solve tough problems is immense. Don't steal all the fun for yourself by explaining *your* solution! I mean it. You want them to rely on themselves and each other – not on you. Don't worry! Your students *will* solve the problem. Their solution method may surprise you!

Add scaffolding if needed. Some of the problems in the book are exploratory in nature, and ask students to make conjectures based on their explorations. For example, Problem 2 asks students to draw some periodic and aperiodic billiard paths in a circular table, and then give the probability that the billiard path is periodic. For many students, this problem is just fine as written. Other students may need the guidance on how to construct examples given at the end of that section. In earlier versions of Problem 5, my students did not do varied enough examples when I gave them blank disks, so I edited the problem to the current version, which gives the students *three* chosen starting points that illustrate very different behaviors. If your experience with your own students suggests that they need more structure for their explorations, or more practice before extending an idea, you may wish to create additional problems for them.

How do I run the class?

Student preparation. Assign one section of problems (typically five problems or so) for each night's homework. Explain to students that preparation for class is essential, as the work they do for homework will form the basis for the next day's discussion. Let them know that they need to bring a written record of their work on each problem, as you will check their notebook. Emphasize that they must, at the very least, draw a picture and write down the given information for each problem, and make an effort towards a solution.

Assess whether the assignment length seems correct for your students: can they complete the homework in a reasonable amount of time, and can they discuss all of the problems during class time? If not, adjust the number of problems you assign.

Classroom preparation. Endeavor to set up your space before class so that students naturally do as much as possible by themselves.

- Choose a classroom with ample board space, ideally lots and lots of board space, as much as you can possibly get.
- Pre-arrange the furniture in groups, with each group next to a large board, or more than one large board.
- Check that each board has several different colors of writing implements and erasers.
- If possible, have a supply of paper, scissors, tape, string, and so on available every day in a visible location, and encourage students to use them.
- If you are using more than one group, have a way of randomly assigning groups when the students enter.

Use of class time. Students should spend class time explaining their solutions to each other, asking each other questions, pointing out errors, discussing, and typically coming to some conclusion on each problem. Depending on your students' emotional and mathematical maturity, they may be able to manage their time and discussions on their own (most college students can do this), or you may need to direct their discussion and help them make good use of their time (young high school students often need this).

Use of incentives. We have two basic goals:

(1) Students should work on the problems before class, so that they are prepared to have substantive discussions, and

(2) students should actively discuss all of the problems within the class period.

For goal (1), check each student's notebook each day. This does not have to be super formal; when you notice that a group is not actively engaged in something, go around to each student and have them show you their work. ("Here is where I did problem 32... here is problem 33...") Remind students of the *purpose* of doing homework, which is to have something to discuss and to prepare their brains to benefit from the discussion. Remind students of the standards, which are to draw a picture and record the information from the problem, and make an effort towards a solution.

For goal (2), sometimes intellectual curiosity is enough. If not, the looming specter of an upcoming test may suffice. More generally, I have had good luck with the following strategy: For a class that meets Monday, Wednesday and Friday, I collect the students' homework that was discussed on Monday, Wednesday and Friday on the following Monday. They have the weekend to make any revisions from what they brought to class. Collecting their solutions incentivizes the students to discuss in class, because they will want to get a good score. I would not grade every problem, but only some selection. I would not collect students' problem sets at the end of the class period in which that problem set is discussed, as this leads students to copy things down without really understanding them.

Thoughts on group size. It is difficult to facilitate a good discussion among six or more students. This is not to say that it is impossible, only that you will have to work hard to do a good job, and the students must be very willing to participate. Groups of four work well. If the problems are difficult, increasing the group size makes it more likely that someone in each group will have solved each problem, but groups of six or more can be tough, as they tend to break into two groups of three students each having their own conversation.

One group, full-class discussion. This can work well when your students are enthusiastic about talking with each other, and with you, about math. If you are going to do this, here are some tips.

Before class, write the problem numbers on the chalkboard, possibly breaking up problems with several parts. For example, if the problems are 3abc, 4, 5abcd, 6abcd, and I have eight students, I would look at the problems and break them into 3abc, 4, 5a, 5b, 5c, 5d, 6ab, 6c, writing each of these at the top of the board, each with plenty of space.[1] This way, each student has an opportunity and an obligation to write up *something*. Arrange your furniture in such a way that most students have a good view of most of the problems, and can see the other students. If you have chalkboards all around the room and you can put everyone around a big seminar table in the middle, that would be great. If not, just do your best.

As the students enter the room, tell them: "Please choose a homework problem and write up your solution." If there are more students than problems, have students join another student who is already at the board. Ditto for any student who finishes writing up their problem: ask them to join another student at the board. Don't let anyone sit down until all of the solutions are up on the boards. The discussions that happen during this time can be the richest part of the class for some students! Students also correct their own mistakes, and other students' mistakes, when they write their homework solutions on the board, saving time later. As students are writing up solutions, take a glance at what students are writing, to make sure that all the solutions are actually going up, and that students understood what

[1] I broke them up this way because of their content: I combined all of 3 because each part is small, and I combined 6a and 6b because 6a does not require explanation.

the problem was asking. If a student is writing up a solution to the wrong problem, or if they have some major misunderstanding, try to nip that in the bud. Focus your efforts on the overarching goal of having students discuss their ideas with each other, and let this guide how much you discuss their solutions with them during this time.

After everything is on the boards, everyone sits down. Have the student(s) who wrote up a solution to the first problem go up to the board and stand next to it. They should first explain what the problem was asking, and then explain their whole solution in detail. You will have to remind them to do these things. The other students should ask questions during the presentation, and also after. You will have to remind and encourage them to do this, too. If another student has something to add, have them go to the board and write it, or draw a picture, so that everyone can see. You will have to remind them to do this. It is much easier to talk than to write or draw, but communicating mathematics works much better through writing or drawing than through talking. Getting in the habit of writing and drawing is a skill that you can help your students to build.

Repeat for each of the problems. You will need to carefully manage class time, hurrying them along if they are taking too long on one problem, and ensuring that they spend adequate time on important problems, to ensure that all problems get explained during class.

Thoughts on assigning students to groups. My students who took classes in the education department told me that studies have shown that "visibly random groupings" promote student learning. Here is how I do it: Let's say I have 24 students and I am putting them into six groups of four. Around the room, I write the numbers 1 through 6 at the top of my six chalkboards. I put four chairs in a semicircle in front of each chalkboard.

I make four little stacks of cards, each containing numbers 1, 2, 3, 4, 5, 6. (If you have five students in each group, you will need two decks of cards.) I shuffle each stack, and then place the four stacks on top of each other. This forms a deck of 24 cards, with numbers 1–6 in random order, four times. I place this deck on the table at the entry door and stand next to it. As the students enter the room, I say: "flip over the next card and look at it." I have to repeat this instruction every time. Students will do weird stuff, like peek at the next card and put it back, or take the card with them. I remind them to *flip* the card, look at it, and put it in the discard pile. Whatever number the student saw, that's their group. Make sure they go towards the board with their assigned number, not just towards their friend.

The benefits of this strategy are that the groups fill up at the same rate, and students know that the groups are random.

If you are only using two groups, you can use colors: have a red group and a black group instead of using the card's number.

The first day of class. An important day for setting the tone! Briefly introduce the course and the syllabus, and then send students to the board to work on the problems in the first section. The goal of the first day is to set the tone: the students will do the work. Do not explain any math! Make them explain it to each other.

After students have solved some problems on the board, go around and have some groups explain their solutions to the whole class. Smile and nod. Do not summarize, point out interesting things, or say any of the dozens of essential thoughts that you feel you must share. This is the students' job! Ditto for the second class.

Keep quiet as much as you possibly can for at least the first three classes. This sets the tone that the students must discuss. If you work hard to keep quiet for the first three classes, the payoff will be great, as your class is likely to have engaging discussions for the rest of the term. They will miss learning from the brilliant comments that you would have shared in the first three days, but in the long run, they will learn much more, as they will get in the habit of having fruitful conversations with each other.

The second day of class. Randomly assign students to groups, and make sure they go to the right place.

Explain to your students that the goal of class time is for students to leave the room understanding each of the problems, and that they should use class time to make that happen. Explain that students are welcome to use the chalkboards, paper, scissors – whatever they need to explore and solve the problems. Let students know that you are available to help them, but that they should ask each other before asking you. Then – start them discussing!

During class, sit calmly in a place where you can see and hear the students. Keep a neutral expression. Listen, watch, but do not react. Look for moments when students are grappling with hard questions, and observe what they do.

Role of the instructor during class time. Encourage students to use the chalkboards, rather than writing on their paper and showing each other. This way, all students can see, and all students can refer to or correct the work. Also, my students have told me that research shows that use of impermanent vertical surfaces promotes learning!

Be available to students by sitting there watching them, but do not jump in. Let them make mistakes and have great ideas.

If you hear a student wondering something and you think that objects might help, grab paper and scissors (or whatever they need) and deliver them to the group.

Watch what students write on the boards, and listen to what they discuss. Make a mental note when they write an error or say something wrong, but do not jump in unless they have truly moved on: I have found that 90% of the time, within 10 minutes they find and correct their own errors. Magical!

If a group calls you over to ask about something, answer them directly; do not try to give them hints to guide them towards your own way of thinking. Consider this: each student in the group tried to solve the problem for homework, and then as a group, they tried to solve it together. That's plenty of thinking time towards this problem. So I ask you *not* to try to socratically lead them towards your own solution. If you want to give them your solution, give it to them!

If it is nearing the end of class and most groups are stuck on a problem that only one group has solved, let that group know that you'd like them to present to the whole class, so that they can prepare a good diagram and think about their explanation. Then get everyone's attention, and have the whole class listen as that group presents their solution.

Troubleshooting if your students are not talking. This may be because you talked too much on the first day of class. Still, here we are, and we must solve this problem. Here are some strategies:

- Make sure that you are checking students' homework. If no one is talking, it is typically because no one has thought about the homework.
- Make your groups smaller. Five students or fewer is ideal. Reduce the group size to two or three until they learn to talk with each other.
- Give each student four pennies. Each time they speak, they must toss a penny on the floor. When they have spent all their pennies, they cannot speak anymore. Now the students who were not speaking are forced to fill the silence. This game does not make for good discussions on the day you play it, but it gets students thinking about their speaking, and can help with subsequent days.
- Use non-random groupings for one day: put the loudest students together in one group, and the quietest students together in another group. Leave the quietest students there and don't talk to them. They will be happy together, and eventually they will start to talk about math.
- Tell each group that for today only, their job is get a complete, correct solution to every problem on the board by the middle of class. Then you will have someone from each group explain each solution to you, and it will *not* be the person who wrote up the solution. Because they are afraid of explaining something that they don't understand or that is incorrect, this strategy is magical for getting students to ask each other questions.
- Send students to the board in partners to work on a new problem. Groups of two facilitate discussion because no one can hide.

What if students don't finish discussing all the problems before the end of class? Reflect on whether you believe that your assignment is really the correct length. If so, it's a matter of getting everyone to be more efficient. If students are working in small groups, remind them that their job is to leave class with a correct solution to each problem, and that they need to manage their time to make this happen. If you are leading a full-class discussion, time management is your job. You will have to decide which problems are worth taking lots of time to discuss, and which ones deserve only a cursory glance.

If you decide that your assignments are actually too long, reduce the number of problems you assign each night.

What if students finish early and there is extra time? Send them to the board in partners to work on a new problem, an artfully chosen problem from the following section. I typically just start them on the most difficult problem. Other options are a problem in which it is difficult to know where to start (so they can figure it out together), a problem in which there is a big paragraph of text to read and interpret (to make it more likely that they will actually do it), or a problem in which they have to work out some examples (because the other person may think to work out examples different from yours).

I would love to hear from you about this curriculum, or this teaching method. Please reach out and tell me how it's going! I love to hear that people have used my materials, and I'd love to know about your experience.

– Diana

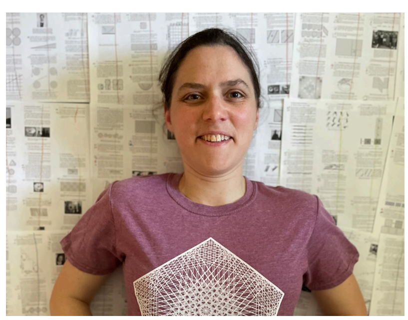


Figure 141. Me with my beloved billiards curriculum.

Image credits

Pictures for THEY DID THE MATH:

- Photograph of Caroline Series is from Wikipedia, by Duncan.Hull - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=91122995.
- Photograph of Amie Wilkinson et al. is courtesy of Amie Wilkinson.
- Photograph of Noelle Sawyer et al. is courtesy of Noelle Sawyer.
- Photograph of Maryam Mirzakhani is from Wikipedia, by Maryeraud9 - Own work, CC BY-SA 4.0, https://en.wikipedia.org/wiki/Maryam_Mirzakhani#/media/File:Maryam_Mirzakhani_Stanford_University_(3x4_cropped).jpg.
- Photograph of Priyam Patel et al. is courtesy of Priyam Patel.
- Photograph of Katherine Knox is courtesy of Kristin Knox and Tom Knox.
- Photographs of Howard Masur and of Alex Eskin et al. are courtesy of Jennifer Taback.
- Photograph of Bill Thurston and Kelly Delp is courtesy of Kelly Delp.

- Screenshots of Vincent Delecroix and Sasha Skripchenko giving talks are courtesy of the Mathematics Department of the University of Warwick.
- Photograph of Jane Wang and Sunrose Shrestha et al. is courtesy of Michael Wan.
- Photograph of Jon Chaika and Evelyn Lamb is courtesy of Evelyn Lamb.
- Photograph of Gabriela Weitze-Schmithüsen is courtesy of SFB-TRR 195.
- Photograph of William Veech et al. is courtesy of Rice University.
- Photographs of Pierre Arnoux et al. and of Gérard Rauzy et al. are courtesy of Pierre Arnoux.
- Photograph of Marina Ratner et al. is courtesy of Anna Ratner, communicated by Nimish Shah.

Other figures:

- The Celtic knot images in Figure 79 are from the license-free library of `shutterstock.com`, with the colors reversed.
- Figures 122 and 124 of the Rauzy fractal are from Wikipedia, by Prokofiev - Own work, CC BY-SA 3.0, `https://commons.wikimedia.org/wiki/File:Rauzy_fractal.png`, the first image being a subset of the full image.
- Figure 128 is courtesy of Pat Hooper.
- Figure 129 of long walks on the Rauzy fractal is reprinted by permission of *Selecta Mathematica*.

Unless noted above, all photographs and figures are by Diana Davis.

Hints for selected problems

As discussed in the Preface on page xiv, the aim of this book is to give you lots of experience working on problems in billiards and related areas, to gradually build up your understanding. Each problem builds on your previous experiences. The map in Figures 144–145 on the next pages suggests how the problems depend on each other: an arrow points from $A \longrightarrow B$ if problem A prepares you to do problem B.

If you are having trouble solving Problem 122 about the eierlegende Wollmilchsau and you need a *hint*, try this:

(1) Find the problem number on the map: ah yes, there it is in green, on the right page in Figure 145, along the left edge, in the middle, in that eierlegende Wollmilchsau shape.

(2) Look at which problem numbers point *towards* the problem: 89, 113, 116, and 121 (also see Figure 142).

(3) Go back and look over your work for each of those four problems. If you couldn't solve one of them, now is the time to do it! Re-do those problems to make sure you fully understand how they work. Look at which problems point towards *those* problems, and review them, too.

(4) Now go try Problem 122 again!

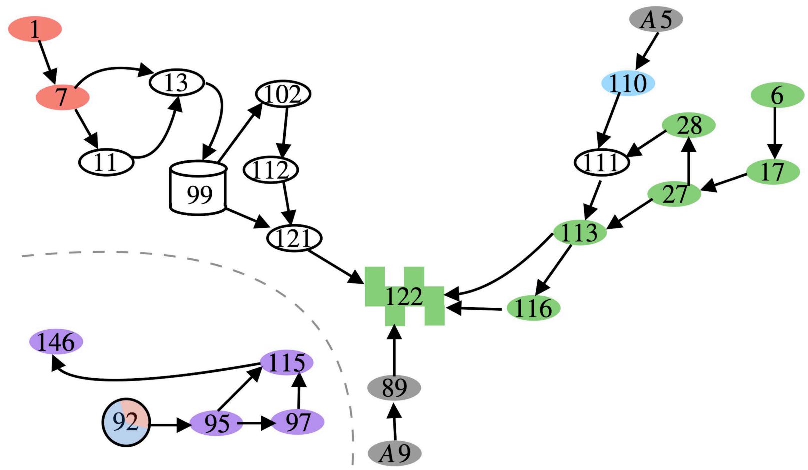


Figure 142. If you want to do Problem 146, do the five problems in the lower left part of the picture. If you want to do Problem 122, doing all of the problems that point towards it will prepare you for that noble quest.

If you are just dipping into this book to learn a few new ideas, the problem map can help you figure out which problems to do. For example, if your advisor wants you to understand how to suspend an interval exchange transformation into a translation surface, you'll want to do Problem 146. Do you need to work *every* problem from 1 to 146, to be able to do it? No! Just find #146 on the map in Figure 144: there it is, in purple, near the bottom of the left page. Problem 115 is pointing towards 146, so you'll need to do that one, and Problems 92, 95 and 97 point towards 115 – and that's it. So to understand suspensions, you just need to do those five problems (see the lower left corner of Figure 142).

You can also use the problem map to decide which problems to omit (Figure 143). For example, if you decide that you want to skip Problem #67 about the Euler characteristic, find it in the middle of the left page in Figure 144, and then use the map to determine that you should consider skipping Problems 68, 69, 94, 143, and 151 as well. Once you've skipped Problem 143, you will lack a bit of the background for 167–170, so proceed with caution!

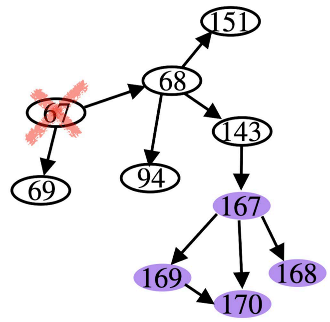


Figure 143. If you skip #67, you may lack the tools for success in these subsequent problems.

This section is called "Hints for *selected* problems" because not every problem has an arrow pointing towards it. Problems 1–6 and 8 each introduce an idea for the first time, as do Problems 74, 92, 105, 119, 140, 153, and 162. For these, read the problem carefully, make a large diagram, and ask a classmate if you need a hint.

Problem 89 is the only pure matrix algebra problem in the book. Problem 157 just asks you to watch a video. No hints here!

Otherwise, every problem builds on the experience of some other problem. I hope that looking at the map on the next two pages inspires you about the inter-relatedness of the ideas in this book, and maybe even about the inter-relatedness of mathematics more generally.

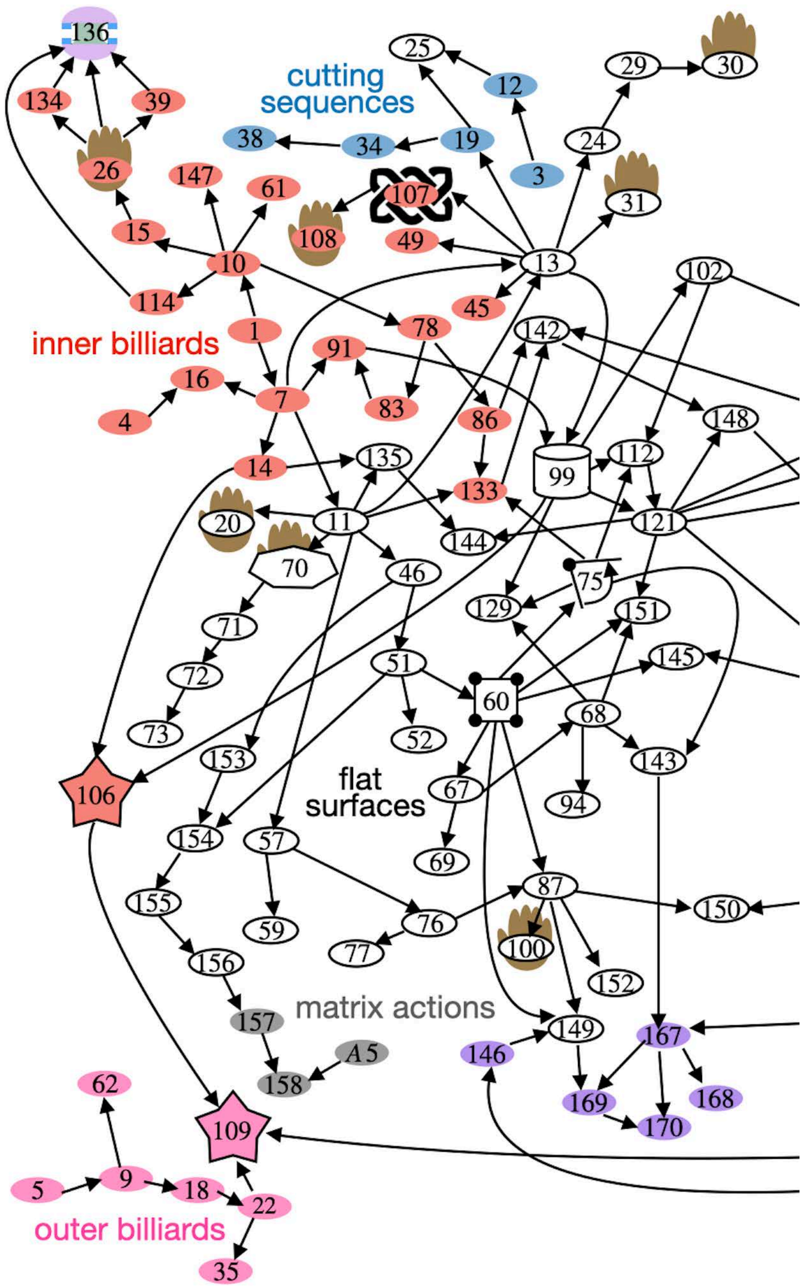


Figure 144. In this nifty diagram, an arrow points from $A \longrightarrow B$ if problem A prepares you to do problem B. It shows how each problem builds on your experiences from doing earlier problems, and how it prepares you for future problems. . .

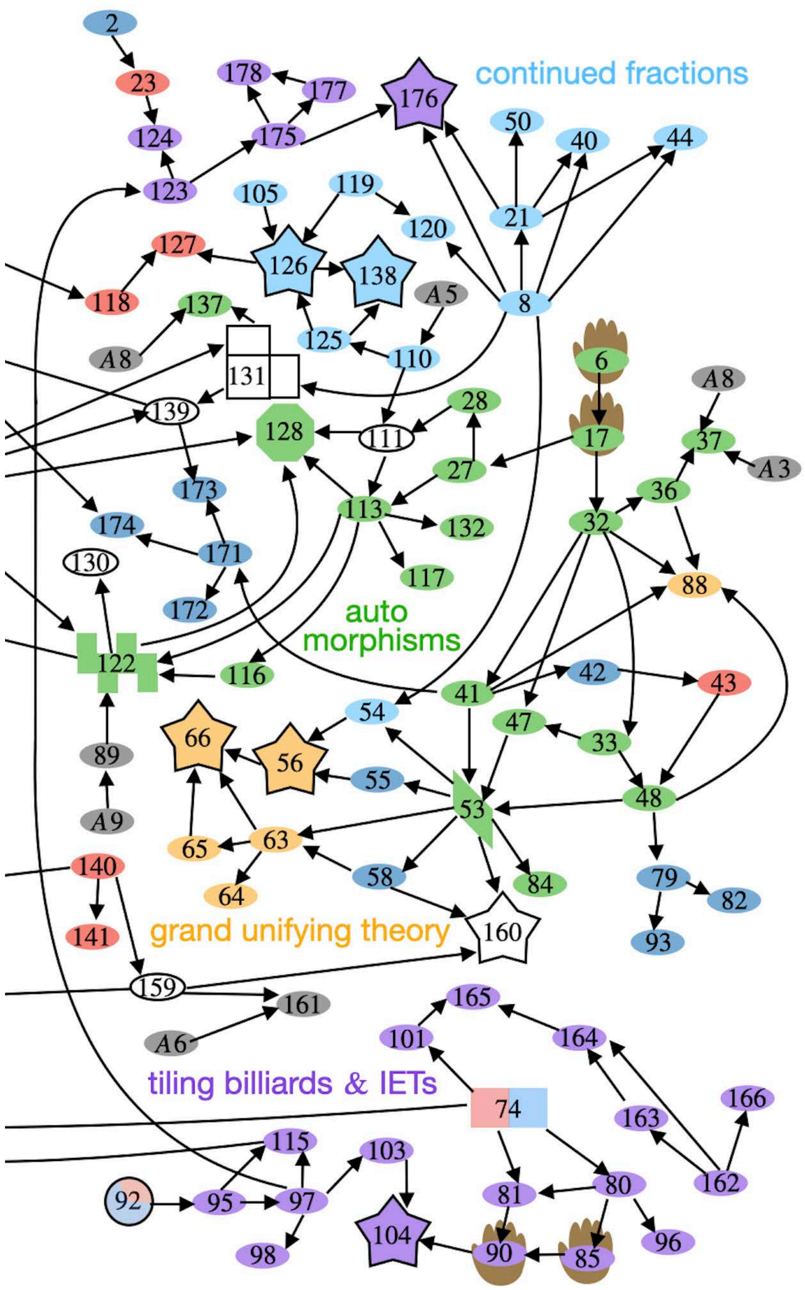


Figure 145. ... and it continues on this page. Problems are color-coded by topic, as indicated. Some problems have special symbols, which suggest their topic. Hands are for hands-on problems; stars are for "synthesis" problems.

Bibliography

[1] Pierre Arnoux, *Un exemple de semi-conjugaison entre un échange d'intervalles et une translation sur le tore*, Bull. de la Société Math. de France 116, 489–500 (1988).

[2] Pierre Arnoux and Štěpán Starosta, *The Rauzy gasket.* In: Further developments in fractals and related fields, Trends Math., Birkhäuser/Springer, New York, 1–23 (2013).

[3] Pierre Arnoux, Samuel Lelièvre, and Alba Marina Málaga Sabogal, *Diplotori: a family of polyhedral flat tori*, in preparation (2021).

[4] Pierre Arnoux and Jean-Christophe Yoccoz, *Construction de difféomorphismes pseudo-Anosov*, C. R. Acad. Sc. Paris Sr. I Math. 292, 75–78 (1981).

[5] Jayadev Athreya, Jon Chaika, and Samuel Lelièvre, *Gap distribution for saddle connections in the golden L*, Cont. Math. 631, 47–62 (2015).

[6] Jayadev Athreya, David Aulicino, and W. Patrick Hooper and with an appendix by Anya Randecker, *Platonic solids and high genus covers of lattice surfaces*, Exp. Math. 31:3, 847-877 (2022).

[7] Paul Baird-Smith, Diana Davis, Elijah Fromm, and Sumun Iyer, *Tiling billiards on triangle tilings, and interval exchange transformations*, J. of the London Math. Soc. (2024).

[8] Irene Bouw and Martin Möller, *Teichmüller curves, triangle groups, and Lyapunov exponents*, Annals of Math. (2006).

[9] Benjamin Call, David Constantine, Alena Erchenko, Noelle Sawyer, and Grace Work, *Unique equilibrium states for geodesic flows on flat surfaces with singularities*, Int. Math. Res. Not. (2022).

[10] Jon Chaika, *Every transformation is disjoint from almost every IET*, Ann. Math. (2012).

[11] Reza Chamanara, *Affine automorphism groups of surfaces of infinite type*, In the Tradition of Ahlfors and Bers, III (Abikoff and Haas, eds.), Cont. Math. 355, 123–145 (2004).

[12] George Csicsery, *Secrets of the surface*, movie (2020).

[13] Diana Davis, *Cutting sequences, regular polygons, and the Veech group*, Geometriae Dedicata, 162(1), 231-261 (2013).

[14] Diana Davis, *Lines in positive genus: an introduction to flat surfaces*. In: Dal'Bo et al. (eds.), *Dynamics Done with Your Bare Hands*, European Mathematical Society (2017).

[15] Diana Davis and Samuel Lelièvre, *Periodic paths on the pentagon, double pentagon and golden L*, Israel Journal of Mathematics (2025).

[16] Diana Davis, Irene Pasquinelli, and Corinna Ulcigrai, *Cutting sequences on Bouw-Möller surfaces: an S-adic characterization*, Ann. Scient. de l'ÉNS 52, 927-1023 (2019).

[17] Vincent Delecroix, W. Patrick Hooper, and Julian Rüth, *sage-flatsurf*, python package, `https://flatsurf.github.io/sage-flatsurf/`, accessed 23 July 2023.

[18] Kelly Delp and William Thurston, *Playing with surfaces: Spheres, monkey pants, and zippergons*, `https://pi.math.cornell.edu/~kdelp/papers/pws.pdf` (2011).

[19] Dmitry Dolgopyat and Bassam Fayad, *Unbounded orbits for semicircular outer billiards*, Ann. Henri Poincaré. 10, 357–375 (2009).

[20] Moon Duchin, Viveka Erlandsson, Christopher Leininger, and Chandrika Sadanand, *You can hear the shape of a billiard table: Symbolic dynamics and rigidity for flat surfaces*, Comment. Math. Helv. (2021).

[21] Ivan Dynnikov, Pascal Hubert, and Alexandra Skripchenko, *Dynamical systems around the Rauzy gasket and their ergodic properties*, Int. Math. Res. Not. 6461–6503 (2023).

[22] Alex Eskin, Maryam Mirzakhani, and Amir Mohammadi, *Isolation, equidistribution, and orbit closures for the* $\mathrm{SL}(2,\mathbf{R})$ *action on Moduli space*, Ann. Math. (2015).

[23] Alex Eskin and Maryam Mirzakhani, *Invariant and stationary measures for the* $\mathrm{SL}(2,\mathbf{R})$ *action on moduli space*, Pub. Math. de l'IHÉS 127, 95–324 (2018).

[24] Alex Eskin, Maryam Mirzakhani, and Kasra Rafi, *Counting closed geodesics in strata*, Inv. Math. 215, 535-607 (2019).

[25] Frank Herrlich and Gabriela Weitze-Schmithüsen, *An extraordinary origami curve*, Math. Nachr. 281, 219–237 (2008).

[26] W. Patrick Hooper, *Grid graphs and lattice surfaces*, Int. Math. Res. Not. 12, 2657–2698 (2013).

[27] W. Patrick Hooper and Barak Weiss, *Rel leaves of the Arnoux–Yoccoz surfaces*, Selecta Math. 24, 875–934 (2018).

[28] W. Patrick Hooper, Pascal Hubert, and Barak Weiss, *Dynamics on the infinite staircase*, Disc. and Cont. Dyn. Sys. 33, 4341–4347 (2013).

[29] Pascal Hubert and Olga Paris-Romaskevich, *Triangle tiling billiards and the exceptional family of their escaping trajectories: circumcenters and Rauzy gasket*, Exp. Math. (2019).

[30] Erica Klarreich, *Mathematicians report new discovery about the dodecahedron*, Quanta Magazine. https://www.quanta magazine.org/mathematicians-report-new-discovery -about-the-dodecahedron-20200831/, accessed 23 July 2023.

[31] Katherine Knox, *Billiard circuits in quadrilaterals*. Am. Math. Mon., 130(8), 755–759 (2023).

[32] Francis Lazarus and Florent Tallerie, *A universal triangulation for flat tori*. Disc. & Comp. Geom. (2023).

[33] Samuel Lelièvre, Thierry Monteil, and Barak Weiss, *Everything is illuminated*, Geom. & Top. 20, 1737–1762 (2019).

[34] Howard Masur, *Closed trajectories for quadratic differentials with an application to billiards*, Duke Math. J. 53, 307–314 (1986).

[35] Curtis McMullen, Ronen Mukamel, and Alex Wright, *Cubic curves and totally geodesic subvarieties of moduli space*, Ann. Math. 185, 957–990 (2017).

[36] Maryam Mirzakhani, *Simple geodesics and Weil-Petersson volumes of moduli spaces of bordered Riemann surfaces*, Inv. Math. 167 (1), 179-222 (2007).

[37] Olga Paris-Romaskevich, *Trees and flowers on a billiard table*, preprint, arXiv:1907.01178 (2019).

[38] Olga Paris-Romaskevich, *Tiling billiards and Dynnikov's helicoid*, Trans. Moscow Math. Soc. 82, 157–174 (2021).

[39] Irene Pasquinelli, *Cutting sequences in Veech surfaces*, master's thesis, available at arXiv:1507.02469 (2015).

[40] Priyam Patel and Nicholas Vlamis, *Algebraic and topological properties of big mapping class groups*. Algebr. Geom. Topol. 18-7, 4109–4142 (2018).

[41] Lionel Penrose and Roger Penrose, *Puzzles for Christmas*, New Scientist 25, 1580–1581 (1958).

[42] Anja Randecker, *Geometry and topology of wild translation surfaces*, KIT Scientific Publishing (2016).

[43] Marina Ratner, *Rigidity of horocycle flows*, Ann. Math. 115, 597–614 (1982).

[44] Marina Ratner, *Horocycle flows, joinings and rigidity of products*, Ann. Math. 118, 277–313 (1983).

[45] Marina Ratner, *The rate of mixing for geodesic and horocycle flows*, Ergodic Th. and Dyn. Sys. 267–288 (1987).

[46] Gérard Rauzy, *Échanges d'intervalles et transformations induites*, Acta Arithmetica 34, 315–328 (1979).

[47] Richard Schwartz, *Unbounded orbits for outer billiards*, J. Mod. Dyn. 3 (2007).

[48] Richard Schwartz, *Obtuse triangular billiards II: 100 degrees' worth of periodic trajectories*, J. Exp. Math., 18, 137–171 (2008).

[49] Henry Segerman, *Hinged flat torus*, YouTube video. `https://www.youtube.com/watch?v=M-m-hKtCQVY`, accessed 23 July 2023.

[50] Caroline Series, *The geometry of Markoff numbers*, Math. Intelligencer 7, 20–29 (1985).

[51] Caroline Series, *The modular surface and continued fractions*, J. London Math. Soc. 2, 31, 69 – 85 (1985).

[52] Sunrose Shrestha, *The topology and geometry of random square-tiled surfaces*, Geom. Ded. (2022).

[53] Sunrose Shrestha and Jane Wang, *Statistics of square-tiled surfaces: symmetry and short loops*, Exp. Math. (2020).

[54] John Smillie and Corinna Ulcigrai, *Geodesic flow on the Teichmüller disk of the regular octagon, cutting sequences and octagon continued fractions maps*, Dynamical Numbers: Interplay between Dynamical Systems and Number Theory, Cont. Math. vol. 532, Amer. Math. Soc., Providence, RI, 29–65 (2010).

[55] John Smillie and Corinna Ulcigrai, *Beyond Sturmian sequences: coding linear trajectories in the regular octagon*, Proc. London Math. Soc. 102: 291–340 (2011).

[56] Sergei Tabachnikov, *Geometry and Billiards*, American Mathematical Society (2005).

[57] George Tokarsky, *Polygonal rooms not illuminable from every point*, Am. Math. Mon. 102, 867–879 (1995).

[58] George Tokarsky, Jacob Garber, Boyan Marinov, and Kenneth Moore, *One hundred and twelve point three degree theorem*, arXiv:1808.06667 (2018).

[59] William Veech, *Gauss measures for transformations on the space of interval exchange maps*, Ann. Math. 115, 201–242 (1982).

[60] William Veech, *Teichmüller curves in moduli space, Eisenstein series and an application to triangular billiards*, Invent. Math. 87, 553–583 (1989).

[61] Clayton Ward, *Calculation of Fuchsian groups associated to billiards in a rational triangle*, Ergodic Th. and Dyn. Sys. 18, 1019–1042 (1998).

[62] Amit Wolecki, *Illumination in rational billiards*, arXiv:1905.09358 (2019).

[63] Alex Wright, *Translation surfaces and their orbit closures: An introduction for a broad audience*, EMS Surv. Math. Sci. 2, 63–108 (2015).

[64] Alexander Zemylakov and Anatolii Katok, *Topological transitivity of billiards in polygons*, Mat. Zametki 18, 291–300 (1975).

Glossary of key terms

angle around a vertex: The total turning angle required to circle around a vertex and return to the initial point. For a *flat surface*, the angle around every vertex must be a multiple of 2π. Figure 146 shows that for the regular octagon surface, the angle around its vertex is 6π. See Problem 75.

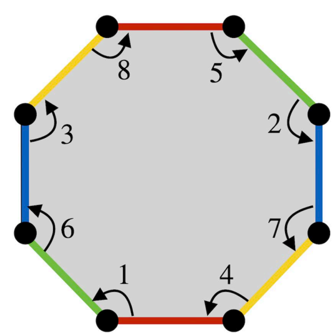


Figure 146. The regular octagon surface has just one vertex, and as you can see from following the arrows in the numbered order, the angle around it is $8 \cdot \frac{3\pi}{4} = 6\pi$.

aperiodic trajectory: A trajectory that is not *periodic*. For example, a trajectory with an irrational slope on the square billiard table or square torus is aperiodic. See Problem 2.

automorphism: An action that takes a surface back to itself, creating neither holes nor overlaps, and preserving the surface's structure. The slogan for an automorphism is "nearby points go to nearby points." In Figure 147, imagine that you have a loop of bread dough with a ribbon of cinnamon swirl around its equator (left). Now imagine breaking the dough apart, giving it a full twist (center), and sticking it back together (right). The cinnamon loop now passes through the central hole. This action is an automorphism, of both the solid dough loop and of its surface, the torus. Some other examples of automorphisms of the torus are reflections and rotations. See Problem 27.

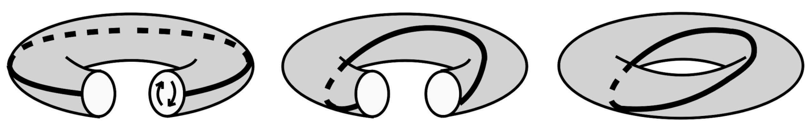


Figure 147. A single full twist is an automorphism of the torus, since nearby points go to nearby points.

cone point: A point of a translation surface such that the angle around it is *not* 2π. Figure 146 shows that the regular octagon surface has a single cone point, whose angle is $8 \cdot \frac{3\pi}{4} = 6\pi$. See Problem 76.

continued fraction: A way of expressing a number as a (possibly infinite) nested fraction. For example, the continued fraction representation of 27/7 is

$$\frac{27}{7} = 3 + \cfrac{1}{1 + \cfrac{1}{6}}$$

which we can express succinctly as $[3; 1, 6]$, the semicolon indicating that the initial "3" is outside of the fraction. See Problem 8.

cut and paste equivalence: Two surfaces are *cut and paste equivalent* if it is possible to cut up and reassemble one into the other, while respecting the edge identifications. Figure 148 shows two examples of pairs of surfaces that are cut and paste equivalent. The left picture shows that it is possible to cut and paste a hexagon surface into a parallelogram surface; the colored edges show that the reassembly respects the edge identifications. The right picture shows two parallelogram surfaces that are evidently cut and paste equivalent, because we have simply swapped the positions of the red and blue triangles, again respecting the edge identifications. We might think of such surfaces as being different, or as being the same, depending on what we care about. See Problems 51 and 153.

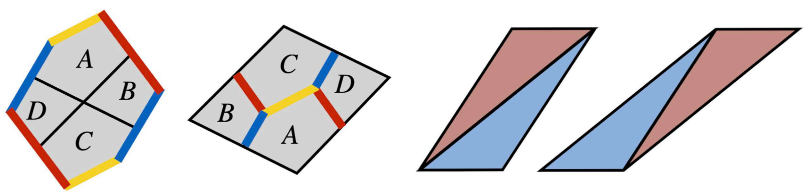


Figure 148. Two examples of pairs of surfaces that are cut and paste equivalent.

cutting sequence: The bi-infinite sequence of edges (or edge *labels*) that a trajectory passes through. Figure 149 gives an example on the double pentagon surface. See Problem 3.

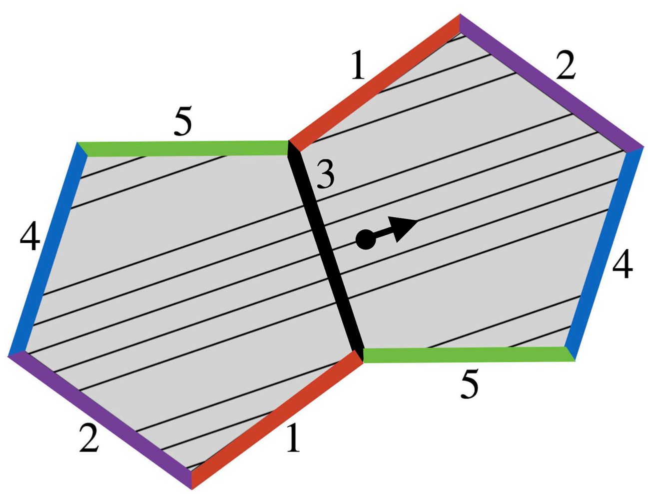


Figure 149. A periodic trajectory on the double pentagon. Starting at the indicated point, its cutting sequence is $\overline{21234543212345432123}$.

cylinder: The union of a maximal family of trajectories on a translation surface that all have the same dynamics. Figure 150 shows a decomposition of the golden L surface into red and blue cylinders. See Problem 99.

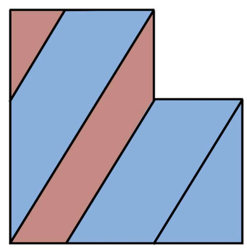


Figure 150. A cylinder decomposition of the golden L in the direction of slope $\phi \approx 1.618$.

ergodic: A flow in a space X is *ergodic* if for any point p in X and for any subset S of X, the amount of the orbit of p that lies in S is proportional to the size of S. Figure 151 shows two long billiard paths on the regular pentagon billiard table. The trajectory on the left appears to exhibit roughly ergodic behavior, as it spends about the same amount of time everywhere, while the trajectory on the right definitely does not, as it visits some parts of the table more often than others, and misses the top left corner completely. See # 28.

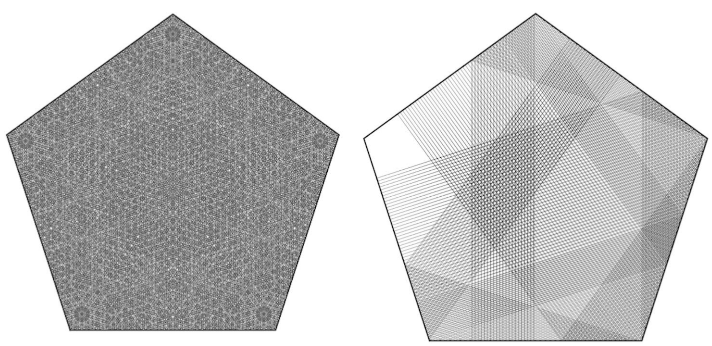


Figure 151. Two long billiard paths, one that fills up the billiard table roughly evenly, and one that definitely does not.

flat surface: A surface that is locally like the flat plane everywhere, except possibly at finitely many *cone points*, where the angle around each cone point is a multiple of 2π. In this book, the flat surfaces we study are all translation surfaces. Problem 77 asks you to say whether *every* flat surface is a translation surface; we won't give away the answer here. See # 4.

genus: When considered as the surface of e.g. a bagel, this is how many "holes" the surface has (Figure 152). See Problem 68.

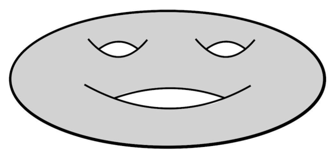


Figure 152. This happy surface has genus 3.

inner billiards: A dynamical system where a point moves linearly inside a domain, and when it hits the boundary of the domain, the angle of incidence between the inbound path and the boundary (or its tangent line) is equal to the angle of reflection between the outbound path and the boundary (Figure 153). See Problem 1.

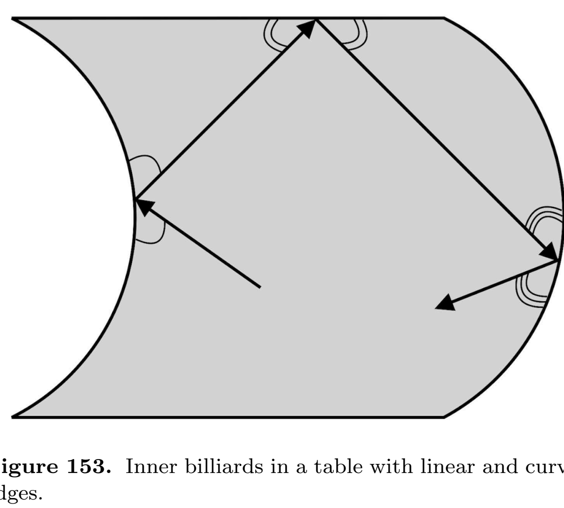


Figure 153. Inner billiards in a table with linear and curved edges.

interval exchange transformation: A dynamical system where the unit interval is cut into pieces that are rearranged, and then glued back together and cut up and rearranged in the same way again. This process is repeated forever. Figure 154 shows a detailed example of the orbit of one point on a 4-IET. See Problem 95.

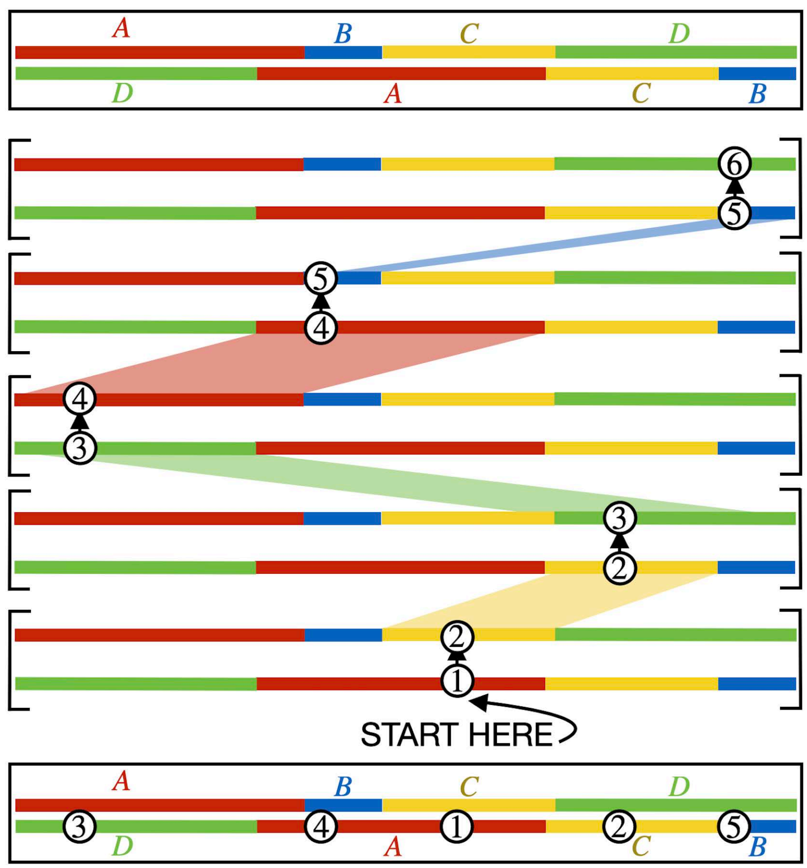


Figure 154. In the box at the top of the picture is a 4-IET. The middle part of the picture shows the orbit of one point, as it flows *up* within the IET, and then shifts *over* to its new location. The box at the bottom of the picture shows a summary: the first five iterations of this point's orbit. Where should you plot the sixth iterate?

oppositely oriented: Parallel edges are *oppositely oriented* if, roughly speaking, the surface lies on the left side of one of the edges and on the right side of the other.

For example, in Figure 155, the parallel edges A *are* oppositely oriented, because if you and your friend each lie on one of the copies with your heads pointing in the same direction (e.g., to the left), one of you has your right arm over the surface (top) and the other has their left arm over the surface (bottom). On the other hand, the parallel edges B are *not* oppositely oriented, because if you and your friend each lie on one of the copies with your heads pointing in the same direction (e.g., up), both of you have your right arm over the surface. See Problem 57.

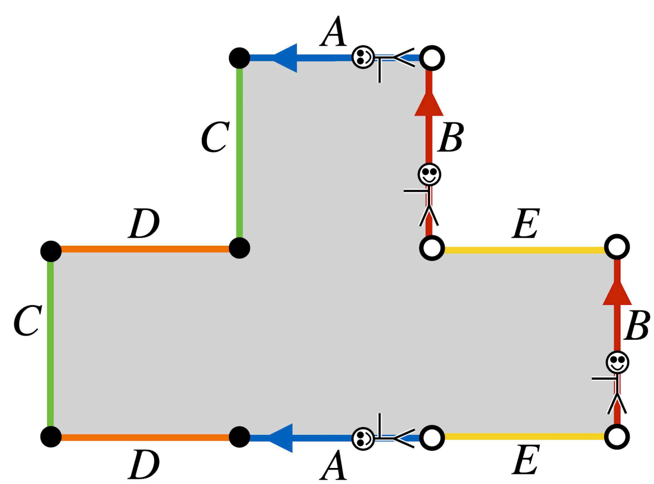


Figure 155. In this surface, parallel edges A are oppositely oriented, while parallel edges B are not. You can determine this by having people lie on the edges with an arm over the surface, as shown.

orbit: The set of images of an input point under repeated applications of a transformation. Figures 154 and 156 each show the first five steps of an orbit, for an interval exchange transformation and for an outer billiard map, respectively. See Problem 98.

outer billiards: A dynamical system where a point moves *outside* of a domain, reflecting through tangent lines to the domain, or through vertices of a polygonal domain. See Problem 5.

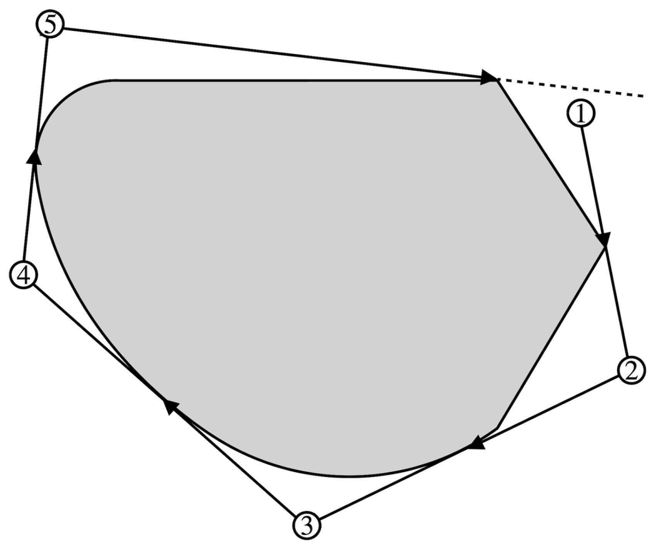


Figure 156. An example of an outer billiard table with both linear and curved edges, and the first five steps of the orbit of a point under the counter-clockwise outer billiard map.

periodic trajectory: A trajectory is *periodic* if returns to its starting point and repeats its path (Figure 157). See Problem 3.

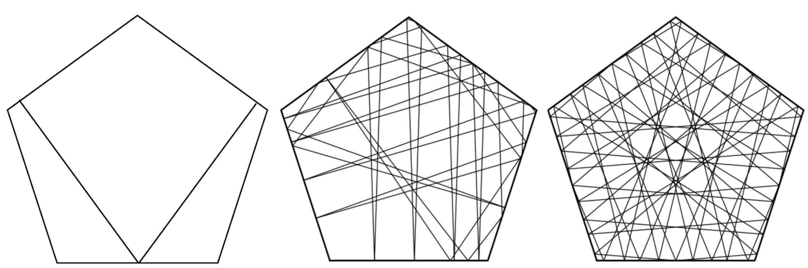


Figure 157. Three examples of periodic billiard paths on the regular pentagon.

saddle connection: A line segment connecting two vertices of a translation surface, with no vertices on its interior (Figure 158). A saddle connection may pass through more than one polygon. See Problem 99.

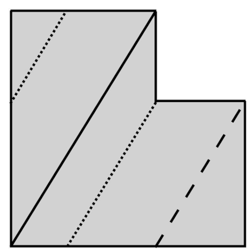


Figure 158. Three saddle connections on the golden L, in the direction of slope ϕ. Notice that these saddle connections bound the cylinders of Figure 150.

shear: For 2×2 matrices, a *horizontal shear* is a matrix of the form $\left[\begin{smallmatrix} 1 & m \\ 0 & 1 \end{smallmatrix}\right]$ and a *vertical shear* is a matrix of the form $\left[\begin{smallmatrix} 1 & 0 \\ m & 1 \end{smallmatrix}\right]$, for a real number m. See Problems A5 and 110.

stratum: Translation surfaces are divided into *strata* based on the angle at each cone point. For a translation surface, the angle at each cone point is a multiple of 2π. For each one, count the *extra* multiples of 2π, and list these in descending order. Figure 159 shows cartoon versions of some the surfaces we study in this book, divided by stratum. See Problem 87.

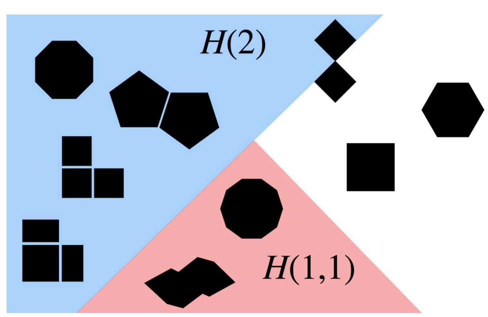


Figure 159. Some of our favorite surfaces, by stratum.

tiling billiards: A dynamical system where a point moves linearly through a tiling of the plane, except that when it hits an edge of the tiling, the angle of incidence between the inbound path and the edge (or its tangent line) is equal to the angle of reflection between the outbound path and the edge (Figure 160).

This system was motivated by the existence of materials with a negative index of refraction, the idea being that one would construct a two-colorable tiling, with the colors corresponding to materials with opposite indices of refraction. We do not require the tiling to be two-colorable, but it so happens that all of the tilings considered in this book are actually two-colorable. See Problem 74.

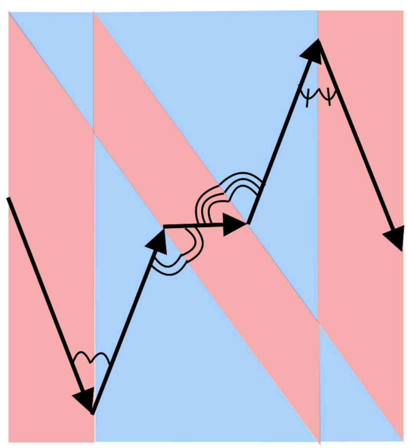


Figure 160. A tiling billiards trajectory on a planar tiling. We can think of the blue and red tiles as being made from materials that have opposite indices of refraction. At each edge crossing, the angle of incidence equals the angle of reflection.

translation surface: A flat surface created from a polygon or collection of polygons by identifying oppositely-oriented parallel edges of the same length. To *identify* edges means to glue them together and make them into the same edge, as though you are taping together a large banner out of many small pieces of printer paper.

For the golden L surface, Figure 161 shows a way to visualize identifying the green and blue pairs of parallel edges. First, you curve the surface in space, and then bring the pairs of green edges together, and the pairs of blue edges together. It remains to do the same for the red and yellow pairs of edges. See Problem 57.

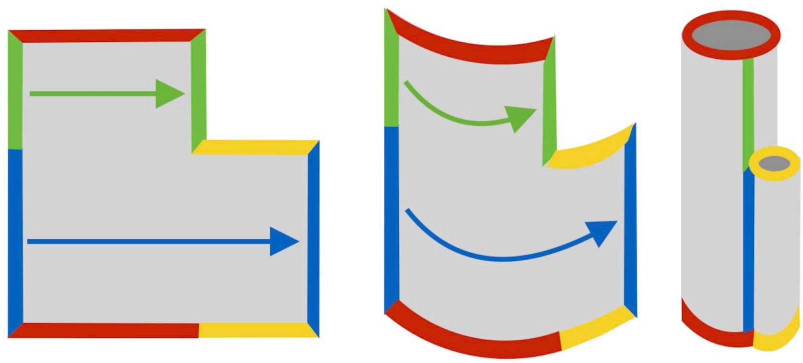


Figure 161. This picture shows how to identify the green and blue pairs of edges of the golden L translation surface. The effect is to roll the surface up into a kind of tube. After further identifying the red and yellow edges, what do you expect the genus of the surface to be?

unfolding: A technique for transforming an inner billiard trajectory into a line on the plane or a trajectory on a flat surface. Sometimes called the *Zemylakov–Katok construction*. The idea is to double the billiard table, imagine that the two copies are joined along the edge of the bounce, and then unfold the second copy like a piece of paper (Figure 162). See Problem 7.

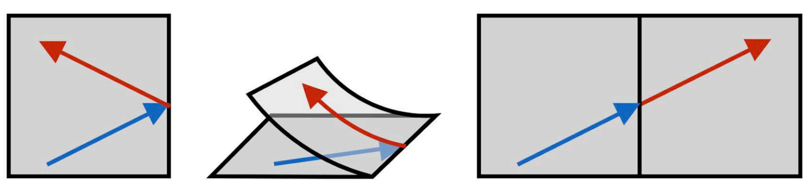


Figure 162. (left) One bounce of a billiard trajectory that we (middle) unfold into (right) a linear trajectory in the plane.

Index of terms

Index of people