Obtuse Triangular Billiards II: 100 Degrees Worth of Periodic Trajectories

Richard Evan Schwartz *

January 25, 2006

Abstract

We give a rigorous computer-assisted proof that a triangle has a periodic billiard path provided all its angles are at most 100 degrees. One appealing thing about our proof is that the reader can use our software online to see massive visual evidence for our result and also to survey the computer part of the proof to a very fine level of detail.

1 Introduction

1.1 Background

Let $T$ be a triangle—more precisely, a triangular region in the plane—with the shortest edge labelled 1, the next shortest edge labelled 2, and the longest edge labelled 3. A billiard path in $T$ is an infinite polygonal path $\{s_i\} \subset T$, composed of line segments, such that each vertex $s_i \cap s_{i+1}$ lies in the interior of some edge of $T$, say the $w_i$th edge, and the angles that $s_i$ and $s_{i+1}$ make with this edge are complementary. (See [G], [MT] and [T] for surveys on billiards.) The sequence $\{w_i\}$ is the orbit type.

In 1775 Fagnano proved that the combinatorial orbit 123 (repeating) describes a periodic orbit on every acute triangle. It is an exercise to show that 312321 (repeating) describes a periodic orbit on all right triangles. (See [GSV], [H], and [Tr] for some deeper results on right angled billiards.)

* This research is supported by N.S.F. Grant DMS-0305047, by a Guggenheim Fellowship, and by the Ruth M. Davis Endowment.
rational triangle—i.e. a triangle whose angles are all rational multiples of \(\pi\)—has a dense set of periodic billiard paths [BGKT]. (See also [M].) There has been a lot of interest in rational billiards lately, owing to the deep connections it has to many areas of mathematics, such as Teichmüller theory; see e.g. [V] or the surveys mentioned above.

In [GSV] and [HH], some infinite families of periodic orbits, which work for some obtuse irrational triangles, are produced. Aside from these results, very little is known about the obtuse (irrational) case of triangular billiards. One central conjecture is

**Conjecture 1.1 (Triangular Billiards Conjecture)** Every triangle has a periodic billiard path.

I think it is fair to say that this 200 year old problem has been widely regarded as impenetrable.

Pat Hooper and I wrote McBilliards, a graphical user interface which searches for periodic billiard paths in triangles. Operating McBilliards I discovered the following result:

**Theorem 1.2 (100 Degree Theorem)** Let \(T\) be an obtuse triangle whose big angle is at most 100 degrees. Then \(T\) has a stable periodic billiard path.

A periodic billiard path is stable if an open set of triangles has the same combinatorial type of billiard path. (See §4.2.) Pat Hooper has recently shown in [H] that right triangles do not have stable periodic billiard paths.

It is the purpose of this paper to rigorously prove the 100 Degree Theorem. The proof we give is a combination of traditional mathematical analysis and rigorous computation. The whole proof, including the computational part, can be surveyed to a very fine level of detail using McBilliards. Alternatively, one can download McBilliards and then run it as a stand-alone application. See 1.4 below for details.

### 1.2 Proof Outline

Let \(\Delta\) denote the parameter space of obtuse triangles. The point \((x, y) \in \Delta\) represents a triangle with small angles \(x\) and \(y\) radians. Let \(S \subset \Delta\) denote the set of points corresponding to obtuse triangles whose small angles are \(x \leq y\), and whose big angle \(z = \pi - x - y\) satisfies

\[
\frac{\pi}{2} < z < \frac{569\pi}{1024}.
\] (1)
Note that
\[100 \text{ degrees} = \frac{5 \pi}{9} \text{ radians} < \frac{569 \pi}{1024} \text{ radians}.
\]

For aesthetic and computational reasons we want to work as much as possible with numbers which are dyadic rational multiples of \(\pi\). For each word \(W\) let \(O(W) \subset \Delta\) denote the set of triangles for which \(W\) describes a periodic billiard path. We call \(O(W)\) an orbit tile. It suffices to cover \(S\) with orbit tiles.

It is useful to define triangles \(P_2\) and \(P_3\) such that
\[\Delta - S = P_2 \cup P_3.\] (2)

See Figure 2.1 for a fairly accurate picture of \(P_1\) and \(P_2\), as well as the other polygons we presently describe. Using McBilliards you can plot these polygons exactly.

It turns out that there are 4 “trouble spots” in \(S\)—places which are somewhat difficult to cover. All these trouble spots occur along the boundary of \(\Delta\) corresponding to right triangles. Let \(p_n\) denote the point in \(\partial \Delta\) corresponding to a triangle, two of whose angles are \(\pi/2\) and \(\pi/n\). The case \(n = \infty\) corresponds to a degenerate triangle.

- It seems that no neighborhood of \(p_4\) can be covered by a single orbit tile. However, we will cover a neighborhood \(P_4\) of \(p_4\) using 9 orbit tiles. Actually, since we are taking \(x \leq y\) in \(S\), we only need 5 of these tiles to cover \(S \cap P_4\).

- It seems that no neighborhood of \(p_5\) can be covered by a single orbit tile. However, we will cover a neighborhood \(P_5\) of \(p_5\) using 2 orbit tiles.

- In [S1] we proved that no neighborhood of \(p_6\) can be covered by finitely many orbit tiles. However, in [S1] we covered a tiny neighborhood \(P_6\) of \(p_6\) using infinitely many orbit tiles. We will use this result here as a black box, but will give ample pictorial evidence for it. The user of McBilliards can see this evidence in great detail.

- In the remark in §2.2 we will give an easy proof that no neighborhood of \(p_\infty\) can be covered by finitely many orbit tiles. However, we will cover a certain neighborhood \(P_1^1\) by a union of two infinite families of tiles.

\[1\] To be consistent with the above notation we ought to call this neighborhood \(P_\infty\), but we prefer \(P_1\).
Let
\[ S' = S - P_4 - P_5 - P_6 - P_1 = \Delta - \bigcup_{j=1}^{6} P_j. \]  
(3)

To finish the proof of the 100 Degree Theorem we need to cover \( S' \) by orbit tiles. To do this we will produce a list of 221 words \( W_7, ..., W_{221} \), together with 221 regions \( P_7, ..., P_{221} \) such that \( S' \subset P_7 \cup ... \cup P_{221} \) and \( P_j \subset O(W_j) \) for all relevant \( j \). (The reader can survey and plot all these words and polygons using McBilliards. We will explain below how to do this.)

Once we have defined the polygons and words, we have 4 goals:

- Show that \( \Delta \subset \bigcup_{j=1}^{221} P_j \).
- Show that \( P_j \subset O(W_j) \) for \( j = 7, ..., 221 \).
- Show that \( P_4 \) and \( P_5 \) are covered by orbit tiles.
- Show that \( P_1 \) is covered by orbit tiles.

The first item has very little to do with billiards. We just have to show that a certain collection of convex dyadic polygons covers \( \Delta \). We will explain our algorithm for doing this and simply remark that it works. The interested reader can see the polygons using McBilliards and check visually that they indeed form a covering of the relevant region.

### 1.3 Plan of the Paper

- In §2 we describe the regions \( P_1, ..., P_6 \) and (when relevant) the orbit tiles which cover them. We defer the long and boring list of \( P_1, ..., P_{221} \) and \( W_7, ..., W_{221} \) to an appendix. All of this information from §2 and the appendix can be obtained from McBilliards, where it is presented in a much more natural way. However, we would like to have a written record of our result which would survive even if computers do not.
- In §3 we describe how we prove computationally that \( \Delta \) is covered \( \cup P_i \). The reader can see our covering using McBilliards and can visually inspect that our algorithm really works. This chapter has nothing to do with billiards per se.
- In §4 we will develop some basic geometric and combinatorial theory for triangular billiards.
• In §5 we will explain a computational algorithm which verifies an equation of the form
\[ P \subset O(W), \]  
where \( P \) is a polygon with and \( W \) is a word. Our theory works for any kind of polygon, but we work with convex dyadic polygons for computational reasons. This method is fully implemented in McBilliards. Our method in §5 works perfectly for the polygons \( P_{30}, \ldots, P_{221} \).

• In §6 we explain how to deal with the polygons \( P_7, \ldots, P_{29} \). Each of these polygons is special because it shares an edge with the right angle line, the portion of \( \partial \Delta \) which parametrizes right angled triangles.

• In §7 we deal with \( P_4 \) and \( P_5 \). For the most part, our treatment of \( P_4 \) and \( P_5 \) is computer-aided, but we need to intervene occasionally and do some hands-on analysis.

• In §8-9 we will cover \( P_1 \) with infinitely many orbit tiles. This part of the proof is purely traditional, but of course is heavily inspired by computer experimentation. \( P_1 \) is the only polygon which shares an edge with the boundary of our parameter space.

• In §10 we deal with the computational aspects of our calculations. In brief, we reduce everything to a calculation involving huge integers, and then use the BigInteger class in Java, which performs the basic operations with integers of ”arbitrary” size. There is no roundoff error in our calculations.

1.4 How to Use McBilliards

We expect that any reader of this paper would read it in tandem with McBilliards. For this reason, we will take the unusual step of explaining to the user how he operates McBilliards. There are 3 options:

Beginners: Go to the link www.math.brown.edu/~res/Java/App46. This applet is a toy version of McBilliards specifically designed for the 100 Degree Theorem. Applet 46 shows the covering of \( S \) by the polygons discussed above, as well as the word list. You can verify informally that each polygon lies in the appropriate orbit tile by inspecting the unfolding, a geometric object which we will discuss in great detail in this paper.
**Intermediate:** After learning to use Applet 46, go to the McBilliards webpage www.math.brown.edu/~res/Billiards/index.html. From the McBilliards website you can play McBilliards online. Open McBilliards and select the 100 degree window from the more menu. The 100 degree window is something like Applet 46 embedded inside McBilliards. The difference is that you can use all the McBilliards tools to interact with the window—i.e. plotting and analyzing orbit tiles. You can follow all the steps in the computer part of our proof using the 100 degree window. In particular, all the computational tests we make in connection with the 100 Degree Theorem can be launched from this window. Once you learn to use the McBilliards interface, you will see that you can verify our proof of the 100 Degree Theorem down to a very fine detail, in a concrete and visually natural way.

**Advanced 1:** Going back to the McBilliards webpage, you can browse through Pat Hooper’s online documentation for McBilliards, which shows every detail of every class, method, and interface in McBilliards. This is not much fun, since there are hundreds of classes involved. However, the code relevant to the 100 Degree Theorem only takes up a small subset of the total program. To isolate the relevant code, we have put it in files which have the Deg100 prefix, such as Deg100Verifier.java. However, there are some basic classes, such as the complex number class and some graphics classes, which are required to support the code in the Deg100 files.

**Advanced 2:** You can download McBilliards from the McBilliards webpage, and then run it as a stand-alone application assuming that you have a fairly recent version of Java installed on your computer. Here is the procedure:

- Untar the directory Current.tar (which is what you download).
- Enter the directory and type `javac *.java`. This compiles all the Java files.
- Type `java A2`. This launches the application.

The stand-alone version also has some C and C++ plug-ins, which are unrelated to the Deg100 Theorem. If you want to run these, and you have a C/C++ compiler, run the `compile_all` command. At any rate, once you have downloaded the code, you can inspect it as you see fit.
1.5 Discussion

I stopped at 100 degrees just because it is a nice round number. The hard cutoff for our result is 112.5 degrees, or $5\pi/8$ radians. The cutoff arises because there exists an infinite family of orbit tiles which accumulates on every point in the boundary of the parameter space corresponding to a degenerate triangle whose angles are $(\pi - x, x, 0)$, where $x \in [\pi/2, 5\pi/8]$. (This is what we use to cover $P_1$.) To get beyond $5\pi/8$ we would need to cover other neighborhoods of the parameter space boundary with orbit tiles. We have not had any luck doing this. We can see that understanding the structure of the parameter space boundary is probably the key challenge in solving the Triangular Billiards Conjecture.

We think that probably the Triangular Billiards Conjecture is true, but that the structure of obtuse triangular billiards is extremely complicated. Originally we wrote McBilliards with the hope of proving the whole conjecture, and we still have that hope; but the main goal is receding off into the distance, like a mirage that always appears to the driver to be just ahead on the road. On the other hand, the program has revealed a wealth of new and totally unexpected phenomena for irrational billiards, a subject still in its infancy. We hope to report on some of these phenomena in future papers.

1.6 Acknowledgements

I did the initial experiments for this project at the Max Planck Institute in Bonn, during July 2004. I would like to thank the M.P.I. for their hospitality and generous support, and also the Guggenheim Foundation. I would like to thank Mike Boyle, Curt McMullen, Dan Rudolph, Martin Schmoll, Sergei Troubetzkoy, and Sergei Tabachnikov for their encouragement, and also for helpful conversations related to this work. I would especially like to thank Pat Hooper, who is my collaborator on McBilliards, for an infinite number of helpful conversations about triangular billiards, computational geometry, and McBilliards. Pat and I have made about equal contributions to the design and algorithm components of McBilliards, but Pat is really the mastermind behind the sophisticated Java architecture of the web based version.
2 Chopping up the Parameter Space

Here and in an appendix we will give the list of all the words and polygons we use to prove the 100 Degree Theorem.

2.1 The Regions

Figure 2.1 shows a fairly accurate picture of the parameter space $\Delta$ of obtuse triangles, as well as the regions $P_1, \ldots, P_6$ discussed in the introduction. The dotted lines indicate that $P_3$ and $P_4$ continue “behind” $P_2$. (It is easier for our algorithm if these triangles overlap.)

2.2 Covering $P_1$

We introduce the notation

$$\frac{k_1}{k_2} \begin{pmatrix} n_1 \\ n_2 \end{pmatrix} = \frac{\pi}{2} \times \left( \frac{k_1}{2^{n_1}}, \frac{k_2}{2^{n_2}} \right)$$

[example: \(\frac{2}{2} \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \left( \frac{\pi}{8}, \frac{3\pi}{8} \right)\) (5)]

With this notation, the vertices of $P_1$ are

$$P_1 : \begin{bmatrix} 2 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$  (6)
We will prove that $P_1$ is covered by the union of two infinite families of orbit tiles $\{O(A_k)\}_{k=1}^{\infty}$ and $\{O(B_k)\}_{k=1}^{\infty}$. Here

$$A_k = 3w_k3w_k^{-1}; \quad B_k = 3w_{k+1}3w_k^{-1}; \quad w_k = 1(32)^{k+1}1(23)^{k+2} \quad (7)$$

Figure 2.2 shows the tiles $O(A_1), ..., O(A_6)$ and the right hand side shows $O(B_1), ..., O(B_6)$ superimposed over the left hand side. The tiles continue sweeping out to the left, covering $P_1$.

**Remark:** It seems worthwhile to quickly explain why we need an infinite number of orbit tiles to cover $P_1$. Let $T$ be a triangle whose largest angle is $90 + \epsilon$ and whose smallest angle is $\delta$, where $\delta << \epsilon$. Any billiard path $P$ in $T$ must eventually hit the short side of $T$ at a point $x$. But then at least one of the segments $S$ of $P$, incident to $x$, will make an angle comparable to $\epsilon$ with one of the long sides. Tracing $P$ out from $x$ in the direction of this segment, we see that $P$ has to make about $\epsilon/\delta$ bounces, moving roughly away from the short side, before its direction can change enough for it to turn around. Hence $P$ cannot have period much less than $\epsilon/\delta$, a quantity we can make as large as we like.
2.3 $P_2$ and $P_3$

The coordinates for $P_2$ and $P_3$ are:

$$
P_2 : \begin{pmatrix} 0 & 0 & 1 & 1 & 0 & 1 \\
0 & 0 & 1 & 1 & 0 & 0 \end{pmatrix} \\
P_3 : \begin{pmatrix} 0 & 0 & 9 & 455 & 0 & 0 \\
9 & 455 & 0 & 0 & 0 & 0 \end{pmatrix}
$$

2.4 Covering $P_4$

We have coordinates

$$
P_4 : \begin{pmatrix} 7 & 63 & 7 & 65 & 7 & 63 \\
7 & 65 & 7 & 63 & 7 & 63 \end{pmatrix}
$$

Since we are taking $x \leq y$ in $\Delta$ only the left half of $P_4$ lies in $\Delta$. In §7 we cover $P_4 \cap \Delta$ with the orbit tiles corresponding to the following 5 words:

$$
C = (1232313)^2.
$$

$$
D_1 = 231323123231323123231323213231323132313231323132313231
$$

$$
D_2 = 231323132312323132312323123231323132313231323132313231
$$

$$
E_1 = 12323132312323123231323213231323132313231323132313231
$$

$$
E_2 = 1232313231323123231323123231323132313231323132313231
$$

The left hand side of Figure 2.3 shows a close-up $O(C)$ and $O(D_1)$ and $O(D_2)$. Note that $O(C)$ slops over the boundary of $P_4 \cap \Delta$. The boundary here is contained in the line through $p_4$ of slope 1. (See the dotted line in Figure 7.3.) The large tile $O(C)$ is not completely shown. The union of these three tiles covers all of $P_4 \cap \Delta$ except for two line segments. These two line segments are then covered by $O(E_1)$ and $O(E_2)$, as shown on the right hand side of Figure 2.3.
2.5 Covering \( P_5 \)

We have coordinates

\[
P_5 : \begin{array}{c|c|c}
12 & 1641 & 12 & 1637 \\
12 & 2455 & 12 & 2455 \\
12 & 2455 & 12 & 2459 \\
\end{array}
\] (10)

It seems that we cannot cover a neighborhood of \( P_5 \) by a single orbit tile. However, we can find two orbit tiles \( O(F) \) and \( O(G) \) whose union contains a neighborhood of \( p_5 \). The words are

\[
F = 312323132132312313213231321323
\]
\[
G = 1323123231323123132313231323132313231323
\]

Figure 2.4 shows a plot of the tiles \( O(F) \) and \( O(G) \), with \( O(F) \) being the larger one. The two tiles overlap, and share \( p_5 \) as a vertex. The three vertical lines are \( \{ x = \pi/k \} \) for \( k = 4, 5, 6 \).
We will show that $P_5 \subset O(F) \cup O(G)$ in §7.

### 2.6 Covering $P_6$

We have coordinates

$$P_6 : \begin{array}{c|c|c|c|c|c|c} \hline 10 & 345 & 12 & 1380 & 12 & 1352 & 9 & 169 \\ 10 & 679 & 12 & 2712 & 12 & 2740 & 9 & 343 \\ \hline \end{array} \quad (11)$$

Let $P'_6$ denote the region

$$\{(x, y) \in \Delta | \left| x - \frac{\pi}{6} \right| < \frac{1}{175}; \left| (x + y) - \frac{\pi}{2} \right| < \frac{1}{400\sqrt{2}} \}. \quad (12)$$

In [S1] we covered $P'_6$ by a union of two infinite families of orbit tiles.

**Lemma 2.1** $P_6 \subset P'_6$. 

12
We have 
\[
\left| \frac{345\pi}{2048} - \frac{\pi}{6} \right| = 0.005624... < 0.005714... = \frac{1}{175}.
\]
\[
\left| \frac{169\pi}{1024} - \frac{\pi}{6} \right| = 0.005113... < 0.005714... = \frac{1}{175}.
\]
This takes care of the condition on the x coordinate. The region $P_6$ is a parallelogram, with one of the long sides lying in the line $x + y = \pi/2$. To finish our verification we just note that 
\[
\left| \frac{1380\pi}{8192} + \frac{2712\pi}{8192} - \frac{\pi}{2} \right| = 0.00153... < 0.00176... = \frac{1}{400\sqrt{2}}.
\]

Combining our lemma with the result of [S1] we see that $P_6$ is covered by a union of two infinite families of orbit tiles. We call these families \( \{O(Y_k)\}_{k=8}^\infty \) and \( \{O(Z_k)\}_{k=8}^\infty \). The words $Y_k$ are defined for all $k \geq 1$ and the words $Z_k$ are defined for all $k \geq 0$.

We first define the $Y$ family. Let 
\[
A = 3123;\quad B_1 = 23213;\quad B_2 = 23123;\quad C_1 = 213123;\quad C_2 = 123123.
\]
(13) 

We have $Y_k = 2y_k2y_k^{-1}$. For odd indices we have 
\[
y_{2k+1} = AB_1(B_2B_1)^kC_1(B_1B_1)^k;\quad k = 0, 1, 2...
\]
(14) 

For even indices we have 
\[
y_{2k+2} = AB_1(B_2B_1)^kC_2(B_1B_2)^{k+1};\quad k = 0, 1, 2...
\]
(15) 

Now we define the $Z$ family. Define 
\[
A = 123;\quad B = 231;\quad C = 32;\quad D = 213.
\]
(16) 

Next define $E_0$ to be the empty word and 
\begin{itemize}
  \item $E_1 = D.D$;
  \item $E_2 = DA.AD$;
  \item $E_3 = DAD.DAD$;
\end{itemize}
• $E_k = DADA.ADAD$;

and so forth. The decimal points are added to highlight the symmetry of the words. Then $Z_k = 3z_k3z_k^{-1}$, where

$$z_k = ABC3E_kABC$$

(The digit 3 included in the equation is deliberate.) We have started our count at $k = 0$ to keep our notation consistent with $[\text{S1}]$.

The left hand side of Figure 2.5 shows the tiles $O(Y_1), ..., O(Y_4)$. The “tips” of these tiles converge to the point $P(\pi/6)$. The largest tiles $O(Y_1)$ obscures the other tiles. The left vertical grey line indicates the set $y = \pi/6$ and the right grey vertical line indicates the set $y = \pi/5$.

The right hand side of Figure 2.5 shows how the tiles $O(Y_1), ..., O(Y_4)$ and $O(Z_0), ..., O(Z_3)$ interlock and suggests how the neighborhood of $P(\pi/6)$ is filled up. The right hand side shows a more local picture than the left hand side.

Our result in $[\text{S1}]$ starts with the tiles $O(Y_5)$ and $O(Z_8)$ only because we couldn’t easily get good rigorous estimates for the first few tiles. Experimentally, the picture looks pretty much the same wherever we start our count.

![Figure 2.5](image-url)
2.7 Covering $S'$

In the appendix we list $W_7, ..., W_{221}$ and $P_7, ..., P_{221}$. (For the sake of having all the polygons in one place, we also list $P_1, ..., P_6$ again.) In all cases, we list the word first and then the vertices of the convex polygon. To compress our notation for the words, we replace all the length 4 strings by letters. There are 24 allowable strings and we order them lexicographically. So, $a = 1212$, $b = 1213$, ..., $x = 3232$. If the length of the word is not divisible by 4 we will simply list the last two digits of the word at the end. (The reader can also browse through this list online, using McBilliards.)

The first 23 words correspond to polygons which abut $\partial \Delta$. These are listed first, sorted by word length. The remaining words are then listed, sorted by word length. The longest word has length 216.

Our list of polygons is probably irredundant, meaning that the deletion of any polygon on the list destroys the covering property. No polygon has more than 8 vertices, and the coordinates all have the form $x(\pi/2)$ where $x \in [0, 1]$ is a dyadic rational whose denominator is at most $2^{17}$. Our list of words is mildly redundant, but probably not twice as long as the shortest list which could do the job.
3 The Covering Condition

3.1 The Covering Problem

The parameter space $\Delta$ of obtuse triangles has vertices

$$\Delta : \begin{vmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{vmatrix}.$$  \hspace{1cm} (18)

Let $P_1, ..., P_{221}$ be the polygons listed in §2 and the appendix. In this chapter we explain how McBilliards proves that $\Delta \subset \bigcup P_j$.

Let $P_j$ be one of the polygons on our list. Let $e$ be an edge of $P$. We say that $e$ is good if

$$e - \partial \Delta \subset \bigcup_{i \neq j} P_i.$$  \hspace{1cm} (19)

In case $e \in \partial \Delta$ this condition is vacuous. We say that $P_j$ is good if every edge of $P_j$ is good.

**Lemma 3.1** $\Delta \subset \bigcup P_j$ provided that every $P_j$ is good.

**Proof:** If $\Delta$ is not covered by our polygons then $\Delta - \bigcup P_j$ contains some open set $U$ and some point of $\partial U$ is contained in some edge $e$ of some $P_j$. But then $e$ is not good. ♠

To make our problem easier, we scale all our polygons by the constant $2^{27}/\pi$. The result is that all the coordinates of all the polygons are positive integers between 0 and $2^{23}$. Also, given the comments at the beginning of §2.7 we know that all the coordinates are divisible by $2^9$. This fact is useful because we sometimes want to subdivide our edges in half a few time, while retaining the property that the break points are integers. We now are left with the problem of showing that a certain convex integer triangle is covered by 221 other convex integer polygons.

3.2 The Bisection Algorithm

Let $S$ be some segment in the plane, whose endpoints are integers. We call $S$ an integer segment. We say that $S$ is admissible if the midpoint of $S$ also has integer coordinates. In this case, the two segments $S_1$ and $S_2$ formed by bisecting $S$ are also integer segments.
Let $e$ be an edge of $P_i$. To show that a given edge $e$ is covered by our polygons, we perform the following algorithm. We start with a list of edges whose sole member is $e$. At any stage of the algorithm we have a finite list of integer segments. We consider the last segment $S$ on the list.

- If we can show that $S \subset P_j$ for some $j \neq i$ then we omit $S$ from our list. Then we continue.

- If $S$ is admissible and we cannot show that $S \subset P_j$ for some $j \neq i$ then we omit $S$ from our list and append $S_1$ and $S_2$ to the list. Then we continue.

- If $S$ is not admissible and we cannot show that $S \subset P_j$ for some $j \neq i$ then we fail.

- If the list becomes empty we have succeeded in showing that $e$ is good.

The main step in our algorithm involves showing that an integer segment is contained in an integer convex polygon. This problem in turn boils down to checking that each of the endpoints of the segment is contained in the polygon. Showing that an integer point $z$ is contained in an integer polygon $P$ is an integer calculation. We just check the orientations of all the triangles obtained by coning the edges of $P$ to $z$ and see that they all agree. This calculation is done entirely in $\mathbb{Z}$ and produces integers which have roughly 3 times as many digits as the coordinates of $z$ and $P$. We implement our algorithm in Java, using the BigInteger class. The BigInteger class does exact arithmetic on arbitrarily large integers. Here “arbitrarily large” means some huge finite number which depends on the physical characteristics of the computer. We certainly never encounter numbers which have more than 100 digits in our algorithm, and these are small enough for the BigInteger class. We will talk more about the BigInteger class in §6.

The interested reader can see and interact with the cover using McBilliards. In particular, one can re-run our algorithm, either one time at a time or sequentially.


4 Billiard Paths and Defining Functions

In this chapter and the next we develop the machinery needed to establish Equation 4 where we need it. We will also use this material in §8-10 when we deal with the polygon $P_1$.

4.1 Unfoldings

We always work with even length words. Given a word $W = w_1, \ldots, w_{2k}$ we define a sequence $T_1, \ldots, T_{2k}$ of triangles, by the rule that $T_{j-1}$ and $T_j$ are related by reflection across the $w_j$th edge of $T_j$. Here $j = 2, \ldots, 2k$. The set $U(W, T) = \{T_j\}_{j=1}^{2k}$ is known as the unfolding of the pair $(W, T)$. This is a well known construction; see $[T]$. Figure 4.1 shows an example, where $W = (1232313)^2$. We label the top vertices of $U(W, T)$ as $a_1, a_2, \ldots$, from left to right. We label the bottom vertices of $U(W, T)$ as $b_1, b_2, \ldots$, from left to right. This is shown in Figure 4.1. The unfold window in McBilliards draws the unfolding $U(W, T)$ for any given word $W$ and any given triangle $T$.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{unnamed.png}
\caption{Figure 4.1}
\end{figure}

$W$ represents a periodic billiard path in $T$ iff the first and last sides of $U(W, T)$ are parallel and the interior of $U(W, T)$ contains a line segment $L$, called a centerline, such that $L$ intersects the first and last sides at corresponding points. We always rotate the picture so that the first and last sides are related by a horizontal translation. In particular, any centerline of $U(W, T)$ is a horizontal line segment. The unfolding in Figure 4.1 does have a centerline, though it is not drawn. To show that a certain triangle has $W$ as a periodic billiard path we just have to consider the unfolding. After we check that the first and last sides are parallel, and rotate the picture as above, we just have to show that each $a$ vertex lies above each $b$ vertex.

18
4.2 Stability

The word $W$ is stable iff $O(W)$ is an open set, and otherwise unstable. Whether or not a word is stable is a combinatorial condition, checked exactly by the computer. Here is the well-known stability criterion. See [S] for a proof.

**Lemma 4.1** Let $W = w_1, ..., w_{2n}$. Let $n_{dij}$ denote the number of solutions to the equation $w_i = d$ with $i$ congruent to $j$ mod 2. Let $n_d = n_{d0} - n_{d1}$. Then $W$ is stable iff $n_d(W)$ is independent of $d$.

McBilliards has a useful graphical interpretation of Lemma 4.1. The 1-skeleton $H$ of the hexagonal grid has 3 parallel families of edges. Given a word, we can draw a path in $H$ by following the edges as determined by the word: we move along the $d$th family when we encounter the digit $d$. Figure 4.2 shows the path corresponding to the word in §4.1, namely $W = (1232313)^2$. The word is stable iff the path is closed. We call this path the *hexpath*.

![Figure 4.2](hexpath.png)

There is a canonical map from the set of triangles of the unfolding to the set of vertices of the hexpath: We simply map $T_i$ to the $i$th vertex $v_i$. The edge of $U(T, *)$ between $T_i$ and $T_{i+1}$ corresponds naturally to the midpoint of the edge joining $v_i$ and $v_{i+1}$. The other two edges of $T_i$ correspond naturally to the midpoints of the other two edges of $H$ emanating from $v_i$. We call this correspondence the *angular correspondence*. For any object of the unfolding $X$, we let $\Theta(X)$ denote the point in the plane corresponding to $X$ under the angular correspondence. Below we will give formulas for the angular correspondence and explain its geometric significance. Informally speaking, we would say that the angular correspondence is the Fourier transform of the unfolding.
4.3 Defining Functions

4.3.1 The Goal

Given two points \( p, q \in \mathbb{R}^2 \) we write
\[ p \uparrow q; \quad p \downarrow q; \quad p \uparrow q \]
iff the \( y \) coordinate respectively is greater than, equal, or less than the \( y \) coordinate of \( q \). Suppose that \( p \) and \( q \) are two vertices of our unfolding. In this section we will give the formula for a function \( f = f_{p,q} \) which has the property that \( f = 0 \) iff \( p \downarrow q \). These defining functions are computed purely from the word \( W \). Our sign convention, discussed below, includes the convention that \( f_{a_i, b_j} > 0 \) iff \( a_i \uparrow b_j \). Given the functions and their formulas, we are left with the following problem: If \( Q \subset \Delta \) is some region and we want to show that \( Q \subset O(W) \), we just have to show that \( f_{a_i, b_j} > 0 \) throughout \( Q \), for all pairs \((a_i, b_j)\).

The reader can use the unfolding window in McBilliards to see the formulas for the defining functions for any word and any pair of vertices on the unfolding.

4.3.2 Turning Angles and Turning Pairs

For ease of exposition, assume that \( T \) is a triangle which is not isosceles. The unfolding \( U(W, T) \) has three kinds of edges, depending on the label the edge inherits from \( T \). The edges of \( U(W, T) \) which have type \( j \) are isometric to the \( j \)th edge of \( T \). Here \( j = 1, 2, 3 \). We will frequently refer to an edge of \( U(W, T) \) by the labels on its endpoints. The first edge of \( U(W, T) \) is always \( e(a_1, b_1) \).

Let \( \rho \) denote the positive \( y \)-axis. For each edge \( e \) of \( U(W, T) \) we let \( \theta(e) \) denote the counterclockwise angle through which \( \rho \) must be rotated in order to produce a vector parallel to \( e \). We work mod \( \pi \), so that the direction \( e \) points is irrelevant. We will sometimes use the notation \( \theta(v, w) = \theta(e(v, w)) \). It is easy to see, inductively, that there are integers \( M_e \) and \( N_e \) such that
\[ \theta_e(x, y) = M_e x + N_e y. \tag{20} \]
Here \((x, y) \in \Delta\) is the point on which \( e \) depends. In §4.4 we will explain how \( M_e \) and \( N_e \) are computed. For now we just use them as a black box. We call \((M_e, N_e)\) the turning pair for \( e \). We will explain below how to compute the turning pairs.
4.3.3 The Formula for the Defining Functions

Let $\tilde{U}(W, T)$ be the bi-infinite periodic continuation of $U(W, T)$. For any $d \in \{1, 2, 3\}$ there is an infinite, periodic polygonal path made from type-$d$ edges if $\tilde{U}(W, T)$. The image of this path in $U(W, T)$ is what we call the $d$-spine. Figure 4.3 shows the 3-spine for $U(W, T)$ where $T$ is some triangle and $W = 123231323123232313$.

![Figure 4.3](image)

Let $e_1, \ldots, e_n$ be a complete and irredundant list of the edges which appear in the $d$-spine. We label so that $e_1$ is the leftmost edge. We introduce the function

$$g_d(x, y) = \sum_{i=1}^{n} (-1)^{i-1} \exp(i(M(e_i)x + N(e_i)y)). \quad (21)$$

We say that $p$ and $q$ are $d$-connected if there is a polygonal path of type-$d$ edges connecting $p$ to $q$, and $d$ is as large as possible. Every two points are $d$-connected for some $d \in \{1, 2, 3\}$, and $d$ is unique. Let $e'_1, \ldots, e'_m$ be the set of type-$d$ paths joining $p$ to $q$, ordered from left to right. We define

$$h(x, y) = \sum_{i=1}^{m} (-1)^{i-1} \exp(i(M(e'_i)x + N(e'_i)y)). \quad (22)$$

When $U(W, T)$ is rotated so that the first edge is vertical:

- The translation direction of $U(W, T)$ is parallel to $\pm ig(x, y)$.
- The vector pointing from $p$ to $q$ is parallel to $\pm h(x, y)$.

Therefore, the function

$$f(x, y) = \pm \text{Im} (\overline{g} h) \quad (23)$$

vanishes iff $p \uparrow q$. Here we have set $g = g_d$. 

21
Remark: The appearance of the factor $(-1)^{i-1}$ is at first a bit puzzling. However, we can explain it like this. Recall that we are working mod $\pi$ and thereby ignoring the direction (forwards or backwards) that a given edge (considered instead as a vector) is pointing. The $(-1)^{i-1}$ turns out to be the fudge factor needed to correct for this loss of information.

After some trial and error we found that the sign out in front of Equation 23 is determined as follows: Let $s$ be the number of edges on the list $e_1, \ldots, e_n$ which lie to the left of $e'_1$. (There is a canonical left-to-right order on all the edges of the same type.) Then $(-1)^s \text{Im}(\overline{h}) > 0$ iff the left endpoint of $e_1$ is $a_1$ (respectively $b_1$) and $p \uparrow q$ (respectively $p \downarrow q$). Actually McBilliards uses the above rule as a basis for establishing the following sign conventions:

1. Suppose $p = a_i$ and $q = b_j$. Then $f > 0$ iff $p \uparrow q$.
2. Suppose $p = a_i$ and $q = b_j$ and $i < j$. Then $f > 0$ iff $q \uparrow p$.
3. Suppose $p = b_i$ and $q = b_j$ and $i < j$. Then $f > 0$ iff $q \uparrow p$.

We introduce a shorthand notation for the function $f$. It suffices to list the turning pairs defining $h$ and then the turning pairs defining $g$. For instance, in the example above the defining function for the pair $(a_1, b_4)$ is recorded as

\[
\begin{array}{ccccc}
0 & 1 & 0 & 1 & (+) \\
4 & 1 & 4 & 1 \\
4 & -1 \\
6 & -1 \\
6 & -3 \\
0 & -3
\end{array}
\]

Here $m = 2$ and $n = 6$. The (+) indicates the sign choice. From the notation we read off that

\[
g(x, y) = \exp(i(y)) - \exp(i(4x + y)) + \exp(i(4x - y)) - \ldots - \exp(i(-3y)).
\]

\[
h(x, y) = (+1) \times (\exp(i(y)) - \exp(i(4x + y)));
\]

We call this form 1 for the defining function.

To arrive at a second convenient form for our function we multiple $\overline{g}$ and $h$ together, collect the terms, and use the fact that sine is an odd function. This gives us what we call Form 2 of the defining function:

\[
f(x, y) = \sum_k J_k \sin(A_k x + B_k y); \quad J_k \in \mathbb{N}; \quad A_k, B_k \in \mathbb{Z}. \quad (24)
\]
4.4 Computing the Turning Pairs

Now we explain an algorithm which generates \((M_e, N_e)\). In the end, it boils down to this: There is a suitable real affine transformation \(R\) of the plane such that \((M(e), N(e)) = R(\Theta(e))\). In other words, up to coordinatizing the plane, the angular correspondence above computes the angular pairs. Using the unfold window in McBilliards, one can see the turning pairs computed automatically.

4.4.1 Step 1: Triples

Let \(d\) be the first digit of \(W\). Let \(d_\pm \in \{1, 2, 3\}\) denote the congruence class of \(d - 1\) mod 3. We let \(d_+ \in \{1, 2, 3\}\) denote the congruence class of \(d + 1\) mod 3. Let \(d_0 = d\). Let \(\epsilon \in \{-1, 0, 1\}\). We define

\[
\alpha_0(d_\epsilon) = \epsilon. \quad (25)
\]

Suppose that we have determined \(\alpha_{i-1}(1), \alpha_{i-1}(2)\) and \(\alpha_{i-1}(3)\). Let \(d\) be the \(i\)th digit of \(W\). Define

\[
\alpha_i(d_\epsilon) = \alpha_{i-1}(d_\epsilon) + (-1)^i 2\epsilon. \quad (26)
\]

In this way we produce a triple of labels for each triangle in the unfolding. The unfolding window in McBilliards displays these triples when you click on a triangle of the unfolding. If the plane is suitable coordinatized by variables \((x, y, z)\) such that \(x + y + z = 0\) then the triple associated to \(T_i\) is precisely the coordinates of \(\Theta(T_i)\), the \(i\)th vertex of the hexpath.

4.4.2 Step 2: Edges

Let \(e\) be an edge of \(U(W, T)\). Suppose that \(e\) is the \(d\)th edge of \(T_i\). We define

\[
\beta(e, d_\epsilon) = \alpha_i(d_\epsilon) - (-1)^i \epsilon. \quad (27)
\]

Note that \(e\) could also be an edge of another triangle of \(U(W, T)\). This happens when \(T_{i-1}\) and \(T_i\) are related by a reflection through \(e\). In other words \(d\) is the \(i\)th digit of \(W\). In this situation Equation 27 gives the same answer whether we use \(i-1\) or \(i\) in the formula. This can be seen by comparing Equations 26 and 27.

Let \(e = e(a_1, b_1)\), the leftmost edge of \(U(W, T)\).
Lemma 4.2 We have the general formula

$$\theta(e) - \theta(e) = -\frac{\beta(e, 1)x + \beta(e, 2)y + \beta(e, 3)z}{3}$$  \hspace{1cm} (28)$$

Here $z$ is such that $x + y + z = \pi$.

Proof: Let $e_1 = e$. We first check our formula on the edges of $T_1$. If 1 is the first digit of $W$ then the edge labels of $e_1$ are $(0,0,0)$ and hence both sides of Equation 28 are 0. The edge labels of $e_2$ are $(-1,-1,2)$. In this case Equation 28 gives $\theta(e_2) - \theta(e_1) = \frac{-(x - y + 2z)}{3} = -z$, as it should. The edge labels of $e_3$ are $(1,-2,1)$. In this case Equation 28 gives $\theta(e_3) - \theta(e_1) = \frac{-(x - 2y + z)}{3} = y$, as it should.

Given the simple nature of the formulas in Equation 26 and 27 it suffices to check the induction step for $i = 2$. In other words, we just have to see that Equation 28 works for the edges of $T_2$. Again, we can suppose that 1 is the first digit of $W$. Suppose that 2 is the second digit. Figure 4.4 shows a picture of the situation. One easily checks that Equation 28 holds for all these edges. When the second digit of $W$ is a 3 the verification is similar. ♠

4.4.3 Step 3: Eliminating the third angle

It is useful to have a formula that doesn’t involve the angle $z$. We define

$$M(e) = \frac{\beta(e, 3) - \beta(e, 1)}{3}; \quad N(e) = \frac{\beta(e, 3) - \beta(e, 2)}{3}.$$  \hspace{1cm} (29)$$

Since $z = (-x - y) \mod \pi$ have

$$\theta(e) - \theta(e) = M(e)x + N(e)y.$$  \hspace{1cm} (30)$$
5 The Verification Algorithm

Our goal is to verify that \( P_i \subset O(W_i) \) where \( P_i \) is a given convex dyadic rational polygon and \( O(W_i) \) is the orbit tile of a word \( W_i \). The method we explain in this chapter can and does work just as written for \( i = 30, \ldots, 221 \).

Our verification algorithm tries to produce a cover of \( P \) by convex dyadic squares \( P \subset \bigcup Q_i \), such that \( Q_i \subset O(W_i) \) for all \( i \). (By dyadic rational square we mean a square in \( \Delta \) whose sides are parallel to the coordinate axes and whose vertices have the form \( x(\pi/2) \) where \( x \in [0, 1] \) is a dyadic rational.)

To show that \( Q \subset O(W) \) we need to show that all the associated defining functions \( f_{a,b} \) are positive on \( Q \). We will sometimes write \( f_{ij} = f_{a,b} \) for ease of notation. In the first section we will explain how we do this. In the sections following the first one, we will explain our main algorithm.

5.1 Certificates of Positivity

Let \( Q \) be a dyadic rational square with center \( q \) and radius \( r \). Here \( r \) denotes half the edge length of \( Q \). Suppose that \( f \) is a defining function for a pair of vertices of the unfolding \( U(W,T) \). There are two ways we try to certify that \( f > 0 \) on \( Q \), the gold and the silver. The gold method is nicer.

5.1.1 The Gold Method

Let \( \nabla f = (f_x, f_y) \) be the gradient. From Equation 23 we have

\[
     f_a = \text{Im}(\overline{g}_a h + \overline{h} a) ; \quad a \in \{x, y\}. \tag{31}
\]

We use Equation 24 to get bounds on the second partial derivatives. Using the letters \( a \) and \( b \) to stand arbitrarily for \( x \) and \( y \), we have bounds on the second derivatives:

\[
     |f_{ab}| \leq F_{ab},
\]

where

\[
     F_{xx} = \sum_k A_k^2 |J_k| ; \quad F_{xy} = \sum_k A_k B_k |J_k| ; \quad F_{yy} = \sum_k B_k^2 |J_k|. \tag{32}
\]

We introduce the quantities

\[
     a_x = r(F_{xx} + F_{xy}) ; \quad a_y = r(F_{yx} + F_{yy}). \tag{33}
\]

25
Finally, we define the rectangle

\[ G(q, f) = [f_x(q) - a_x, f_x(q) + a_x] \times [f_y(q) - a_y, f_y(q) + a_y]. \] (34)

Here \( q \) is the center of \( Q \).

It follows from integration that

\[ \nabla f(x, y) \subset G(Q, f); \quad \forall (x, y) \in Q. \] (35)

We say that \( f \) is gold certified if \( G(Q, f) \) is disjoint from the coordinate axes in \( \mathbb{R}^2 \). This is to say that \( G(Q, f) \) is contained in one of the standard quadrants in \( \mathbb{R}^2 \).

If \( f \) is gold certified, then there is some vertex \( v \) of \( Q \) such that throughout \( Q \) the gradient \( \nabla f \) is a positive linear combination of the edges of \( Q \) which emanate from \( Q \). This means that \( f(x, y) > f(v) \) for all \((x, y) \in Q\). Thus, if \( f \) is gold certified and \( f(v) > 0 \) then \( f|_Q > 0 \). We say that we have shown \( f|_Q > 0 \) by the gold method if this situation obtains. Note that the gold method only requires a finite number of computations. The gold method works poorly if \( \nabla f \) points nearly horizontally or vertically in \( Q \).

5.1.2 The Silver Method

Let \( \hat{Q} \) denote the square with the following property: \( Q \) is midscribed in \( \hat{Q} \), as shown in Figure 5.1. Note that \( \hat{Q} \) is not a dyadic rational because its sides are not parallel to the coordinate axes. However, the vertices and center of \( \hat{Q} \) all have the form \( \pi x \), where \( x \) is a dyadic rational.

![Figure 5.1](image-url)
We use all the same notation as in the previous section. We not define the rectangle

\[ S(q, f) = [f_x(q) - 2a_x, f_x(q) + 2a_x] \times [f_y(q) - 2a_y, f_y(q) + 2a_y]. \]  

(36)

It follows from integration that

\[ \nabla f(x, y) \subset S(Q, f); \quad \forall (x, y) \in \hat{Q}. \]  

(37)

We say that \( f \) is silver certified if \( G(Q, f) \) is disjoint from the lines through the origin of slope \( \pm 1 \). This is to say that \( S(Q, f) \) is contained in one of images obtained by rotating the standard quadrants by 45 degrees.

If \( f \) is silver certified, then there is some vertex \( v \) of \( \hat{Q} \) such that throughout \( \hat{Q} \) the gradient \( \nabla f \) is a positive linear combination of the edges of \( \hat{Q} \) which emanate from \( \hat{Q} \). This means that \( f(x, y) > f(v) \) for all \( (x, y) \in \hat{Q} \). In particular, this is true for all \( (x, y) \in Q \). Thus, if \( f \) is silver certified and \( f(v) > 0 \) then \( f|_Q > 0 \). We say that we have shown \( f|_Q > 0 \) by the silver method if this situation obtains. Note that the silver method requires a finite number of computations.

The silver method is not as nice as the gold method for the following reason. If \( f|_Q > 0 \) but \( Q \) is quite close to the level set, then it might happen that \( f(v) < 0 \) on the relevant vertex of \( \hat{Q} \). For our purposes, the gold method usually works, and the silver method takes over as a last resort when the gold method fails. The two methods work together beautifully for our purposes.

5.1.3 A Technical Point

The constant \( r \) in the formulas above has the form

\[ r = \frac{\pi}{2} x \]

where \( x \) is some dyadic rational number. When it comes time to do our rigorous computation we will replace \( r \) by the larger

\[ \tilde{r} = 2x \]

because it is a rational quantity. We will then work with the rectangles \( \tilde{G}(Q, f) \) and \( \tilde{S}(Q, f) \), which are defined as above, but with \( \tilde{r} \) in place of \( r \). This replacement makes the functions a bit harder to certify, but helps us reduce the problem to an integer calculation.
5.2 An Inefficient First Try

Here we describe a simple verification algorithm which is too slow to use, but easy to understand. Following this section, we will describe the algorithm we actually do use.

Let $Q$ be a dyadic square and let $W$ be a word. We say that $W$ is good on $Q$ if, for every defining function $f_{ij}$ we can prove that $f_{ij}|Q > 0$ either by the gold method or by the silver method. If $W$ is good on $Q$ then $Q \subset O(W)$.

For our algorithm we start with a list of squares, having the $Q_0$ as its sole member. We have coordinates.

$$Q_0 : \begin{array}{c|c|c|c|c|c|c|c|c} 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \end{array}$$

(38)

That is

$$Q_0 = [0, \frac{\pi}{2}]$$

At any point of the algorithm we have a list of dyadic rational squares. We let $Q$ be the last square on the list. There are several options.

- If $f$ is good on $Q$ we delete $Q$ from our list and add it to our covering.
- If $Q \cap P = \emptyset$ then we delete $Q$ from our list.
- If neither of the above is true, we replace $Q$ on our list by the 4 squares obtained by subdividing $Q$ in half.

If our list ever becomes empty then we have a covering of $P$ by dyadic squares, each of which is contained in $O(W)$. This does the job. One very nice feature of this algorithm is as follows if $P \subset P'$ are two different polygons, and the algorithm works for both $P$ and $P'$, then the covering of $P'$ is obtained from the covering of $P$ simply by adding some more dyadic squares.

The algorithm is inefficient for a variety of reasons. The main reason is that it requires us to evaluate all $O(n^2)$ defining functions for each square on the list (which is not disjoint from $P$.) The algorithm we describe requires $O(n \log(n))$ evaluations for each such square.

5.3 The Tournament

As above, $W$ is a fixed word. Let $Q$ be a dyadic rational square. Say that a player list for $Q$ is a pair $(A, B)$, where both $A$ and $B$ are lists of indices. We
think of $A$ as being a list of some distinguished $a$ vertices and $B$ as being a list of some distinguished $b$ vertices. We say that lists $i < j \in A$ are adjacent if there is no index $k \in A$ such that $i < j < k$. In this section we will make some definitions for $A$ and at the end make the same definitions for $B$.

We say that an $A$-function is a defining function associated to $(a_i, a_j)$, where $i$ and $j$ are adjacent indices in $A$. We say that a vertex $i \in A$ is an $A$-loser if one of the following two situations (when applicable) obtains:

- Let $j > i$ be the index adjacent to $i$. Let $f$ be $A$-function for the pair $(a_i, a_j)$. Then $-f_Q$ can be shown to be positive using either the gold or silver method.
- Let $j < i$ be the index adjacent to $i$. Let $f$ be $A$-function for the pair $(a_i, a_j)$. Then $f_Q$ can be shown to be positive using either the gold or silver method.

One of the situations is not applicable if $i$ is the first or last index in $A$. If $i$ is the only index in $A$ then neither situation is applicable.

If $i \in A$ is an $A$-loser it means that there is another index $j \in A$ such that $a_i \uparrow a_j$ throughout $Q$. In this case any result $a_j \uparrow b_k$ in $Q$ automatically implies that $a_i \uparrow b_k$ in $Q$. If $i$ is not a round loser we call $i$ an $A$-survivor.

We make all the same definitions for the $B$ list, except that we reverse the signs. That is, we say that a vertex $i \in B$ is an $B$-loser if one of the following two situations (when applicable) obtains:

- Let $j > i$ be the index adjacent to $i$. Let $f$ be $B$-function for the pair $(b_i, b_j)$. Then $f_Q$ can be shown to be positive using either the gold or silver method.
- Let $j < i$ be the index adjacent to $i$. Let $f$ be $A$-function for the pair $(b_i, b_j)$. Then $-f_Q$ can be shown to be positive using either the gold or silver method.

We call the following elimination process a round (of a tournament): We consider in order all the $A$-functions $f_1, \ldots, f_m$. We form a new list $A'$ consisting of the $A$-survivors. We call $A$ stable (with respect to $Q$) if $A' = A$. If $A$ is not stable we form a sequence $A \supset A' \supset A'' \ldots$ until the list stabilizes. We call this process the $A$-tournament on $Q$. We call the indices of the final list the $A$-winners. We carry out the same processes for the $B$ list.
5.4 The Improved Algorithm

We start our algorithm with the list consisting of the triple \((Q_0, A_0, B_0)\), where \(Q_0 = [0, \pi/2]^2\) as above, and \(A_0 = B_0 = \{1, 2, 3, \ldots, k\}\) are the complete list of indices. Here \(k\) is half the length of \(W\). During the algorithm we maintain a list of triples like this. At any stage we consider the last triple \((Q, A, B)\) on the list.

If \(Q \cap P = \emptyset\) we discard \((Q, A, B)\) from our list and move on. Otherwise...

- We perform the \(A\)-tournament and \(B\)-tournament to produce triples \((Q, A^*, B^*)\), where \(A^*\) consists of the \(A\)-winners and \(B^*\) consists of the \(B\)-winners.

- For each index \((i, j)\) \(\in A^* \times B^*\) we try to show, using the gold and silver methods, that \(f_{ij}|_Q > 0\). If we succeed for every pair then we add \(Q\) to our covering of \(P\). Otherwise...

- We delete \((Q, A, B)\) from our list and then replace it by the 4 triples \((Q_j, A^*, B^*)\), where \(Q_1, Q_2, Q_3, Q_4\) are the squares obtained by bisecting \(Q\).

If the list ever becomes empty then we have produced a covering of \(P\) by dyadic squares, each of which is contained in \(O(W)\). This is justified by the following

Lemma 5.1 If \(Q\) is added to our cover then \(Q \subset O(W)\).

Proof: Let \((i, j)\) \(\in A_0 \times B_0\) be arbitrary indices. There is a nested sequence of squares \(Q_0 \supset Q_1 \supset \ldots \supset Q_n = Q\) together with a sequence of indices \(i = i_0, \ldots, i_n = i'\) such that \(Q \subset Q_k\) and \(a_{i_k} \uparrow a_{i_{k+1}}\) for all \(k\). Moreover \(i' \in A^*\). The same goes for \(j\) in place of \(i\). Therefore, on \(Q\) we have \(a_i \uparrow a_{i'} \uparrow b_{j'} \uparrow b_j\). ♠

We point out our 3 nice features of our algorithm:

- If \(P \subset P' \subset O(W)\) and the algorithm works for both \(P\) and \(P'\), then the covering produced for \(P'\) is obtained from the covering produced for \(P\) just by adding some squares.
• The gold and silver certificates are inherited. If a defining function $f$ is gold/silver certified on a square $Q$ it is also gold/silver certified on a subsquare $Q'$ of $Q$. We don’t need to recompute the bounds; we just pass along the certificate during the subdivision, assuming the relevant pair of indices gets passed along.

• If $Q$ is one of the squares in our covering, then there is a canonical sequence of squares $Q_0, ..., Q_n = Q$, where $Q_{k+1}$ is one of the 4 squares in the bisection of $Q_k$ for all $k$. The presence of $Q$ in our cover can be completely explained by looking at what happens in $Q_0, ..., Q_n$. We don’t have to look at other “branches” of the algorithm. As we will explain below, McBilliards exploits this feature to produce a nice way for the (tireless) reader to inspect the operation of the algorithm piece by piece. We will discuss this below in some detail.

Figure 5.1 shows the output of our algorithm for $P_{32}$. These squares just barely cover $P_{32}$: The polygon has nearly the same shape.

Figure 5.1
5.5 Surveying the Algorithm

Now we explain how the reader can check the results of the tournament algorithm.

- To run the verify algorithm for a particular word, select the verify single mode on the 100 degree window interface and then click on the desired word. (These words are indexed by little square buttons on the interface.) Be sure to have the trace verify button off.

- Once the picture is plotted on the main McBilliards window, turn on the trace verify button and select your favorite dyadic square that you have just plotted by clicking inside it. Now click on the same word you just clicked.

- With the trace verify mode on, McBilliards re-runs the algorithm, discarding any square which does not contain the selected point. This has the effect of just tracing through the part of the algorithm which deals with the selected square.

- Open up the unfolding window after the selected square has been plotted. Along the bottom of the square you will see three kinds of boxes: the top winners, the bottom winners, and the tournament record. The tournament record consists of a bunch of pairs of the form \((p, q)\), where \(p\) loses to \(q\) on some box which contains the selected one. We call these match boxes.

- If you click on one of these matchboxes, you will see the formulas for the defining function associated to the relevant pair of vertices. You also get to see a graphical display of the gradient and the quadrant which contains the gradient throughout the dyadic square. By moving the point around on the main interface, you can visually check that the gradient remains within the quadrant. Also, you see displayed all the quantities which go into the calculation of the certificates, so you can recompute them yourself from the information.

- If you click on every single match box and make the computations yourself, by hand, you will have given your own proof that the tournament has performed correctly. Finally, you can go through all the pairs of the form (top winner, bottom winner) and make all the same checks.
6 Vanishing on the Right Angle Line

Here we explain how to modify the algorithm in §5 to work for the indices \(i = 7, \ldots, 29\). What makes these polygons special is that they all have an edge on the right angle line.

6.1 Exceptional Pairs

Say that a pair of vertices \((a_i, b_j)\) is exceptional if the associated defining function vanishes along the right angle line. We call such a defining function exceptional as well. For any word \(W\) there is a list \(A\) of \(a\) vertices of \(U(W, \ast)\) and a list \(B\) of \(b\) vertices of \(U(W, \ast)\) such that the set of exceptional pairs of vertices is precisely \(A \times B\). For the words \(W_30, \ldots, P_{229}\) the lists \(A\) and \(B\) are typically (though not always) empty. However, the polygons \(P_30, \ldots, P_{221}\) are all (very) disjoint from the right angle line, and so the lists \(A\) and \(B\) do not concern us. For the words \(W_7, \ldots, W_{29}\) the lists \(A\) and \(B\) are always nonempty and, as we mentioned above, the polygons \(P_7, \ldots, P_{29}\) always have an edge on the right angle line. For this reason, we need to understand what happens with the defining functions associated to vertices in \(A \times B\). It is hard to deal computationally with these defining functions, because they take arbitrarily small positive values on points in the polygons.

Say that a dyadic square is exceptional if it has one or two vertices on the right angle line and at least one vertex in the parameter space \(\Delta\) of obtuse triangles. Figure 6.1 shows a picture of the two kinds of special dyadic squares. Let \(Q\) be an exceptional dyadic square and let \(f\) be an exceptional defining function. Say that \(f\) is certified on \(Q\) if the gold method shows that...
\(\nabla f\) is contained in a quadrant throughout \(Q\) we also insist that \(\nabla f\) points into the obtuse parameter space. In this situation the axis of the quadrant containing \(\nabla f\) is perpendicular to the right angle line, and \(f > 0\) on the portion of \(Q\) which lies in \(\Delta\).

When we run our algorithm for the indices \(i = 7, ..., 29\) we first isolate the lists \(A\) and \(B\). We then run the algorithm as in §5, except that we automatically “pass” any exceptional defining function in the playoffs if the dyadic square in question is exceptional and the defining function is certified on the square. If the algorithm halts, we have a covering of \(P_i\) by a union of dyadic squares and dyadic triangles, each of which is contained in \(O(W_i)\). It only remains to explain how we recognize in advance that a pair of vertices is exceptional. In the next section we will explain some general principles for doing this, and then we will spend the rest of the chapter going through the exceptions one at a time.

While reading our account, the reader may wonder how we know that we have obtained an exhaustive list of exceptional pairs. Actually, it is not necessary for us to do this. We just have to show that the algorithm halts with the exceptional pairs that we have singled out. Given that the algorithm is based on finite precision (though exact) arithmetic, another exceptional pair would cause the algorithm to get hung up, producing a list of ever smaller dyadic squares converging to the right-angle line. Since this does not happen, we know that we have the complete list. The reader who experiments with the unfoldings using McBilliards can see directly that our list is complete.

### 6.2 Using Symmetry

For each of the exceptional words, the unfolding \(U(W, *)\) has bilateral symmetry. The symmetry derives from the fact that we can write \(W = dVdV^{-1}\) where \(V\) is a word having of length

\[
(length(W) - 1)/2.
\]

There are two important features of this symmetry. First, the first and last edges of \(U(W, *)\) are always vertical. This allows us to predict the turning angles of the other edges solely from their turning pairs. (In a minute we will give an example.) Second, we only have to worry about half the vertices when we run our algorithm. In particular, it suffices to deal with the exceptional vertices on the left half of the unfolding.
Figure 6.2 shows the example of $W_{11}$. In this case, the only exceptional pair of vertices is $(a_5, b_1)$.

![Figure 6.2](image)

The vertices $a_5$ and $b_1$ are joined by 2 edges of type 3. The union of these two edges has a line of bilateral symmetry. Call this line $\Lambda_{51}$. The turning pair for $\Lambda_{51}$ is $(-2, -2)$. Mod $\pi$, the angle between the first edge, which is always vertical, and $\Lambda_{51}$, is $-2x - 2y$. But $x + y = \pi/2$ on the right angle line. Hence $\Lambda_{51}$ is vertical for any unfolding with respect to a right triangle. Hence $a_5 \updownarrow b_1$ for all points on the right angle line.

### 6.3 The Easy Cases

With 6 exceptions, the words $W_7, ..., W_{29}$ have the same analysis as $W_{11}$. That is, they have a single exceptional pair of vertices (on the left) and the spine connecting these vertices has bilateral symmetry. In all these cases, the same analysis as for $W_{11}$ works here word for word. Here we list these cases, together with the exceptional pairs. In the notation we use, the case considered in the previous section is listed as $(11; 5, 1)$. Here are the easy cases:

$$
(7; 5, 1) \quad (9; 5, 10) \quad (10; 5, 1) \quad (11; 5, 1) \quad (12; 8, 13) \quad (13; 1, 11) \\
(14; 5, 1) \quad (17; 5, 1) \quad (19; 19, 3) \quad (20; 5, 1) \quad (22; 1, 23) \quad (24; 27, 5) \\
(25; 5, 33) \quad (26; 42, 31) \quad (27; 38, 7) \quad (28; 5, 45) \quad (29; 48, 11)
$$

(39)

Notice that the pair $(a_5, b_1)$ occurs quite often. The reader can see pictures of all these cases using the unfolding window of McBilliards. In the unfolding window you can select a pair of vertices and see the edges connecting them drawn. In this way you can verify that the path has bilateral symmetry and the line $\Lambda$ of bilateral symmetry has turning pair either $(2, 2)$ or $(-2, -2)$ and hence is vertical when the unfolding is done with respect to a right triangle.

We will treat the remaining cases roughly in order of their complexity.
Figure 6.3 shows $U(W, x)$ for some point $x$. We have highlighted 8 line segments which are all horizontal when $x$ lies on the right angle line. The turning pairs for these segments are all of the form $(k, k)$ for $k \in \{\pm 1, \pm 3, \pm 5\}$. Restricting our attention to the left hand side, we see that the exceptional sets are $A = \{1, 2, 4, 5\}$ and $B = \{8\}$. These are exactly the ones we single out when we run our algorithm.

6.5 $W_{21}$

For $W_{21}$ we have $A = \{22, 23\}$ and $B = \{4\}$. In this case, the pair $(a_4, b_{22})$ has the same kind of bilateral symmetry as for the easy cases. Hence $a_4 \updownarrow b_{22}$ for any unfolding with respect to a right triangle. Finally, the turning pair for the edge connecting $b_{22}$ and $b_{23}$ is $(1, 1)$. Hence, this edge is horizontal for any unfolding with respect to a right triangle.

6.6 $W_{16}$ and $W_{23}$

For $W_{16}$ we have $A = \{4\}$ and $B = \{7, 8, 14, 15\}$. There is an edge of $U(W_{16}, \ast)$ connecting $a_4$ and $b_7$, and this edge has turning pair $(1, 1)$. Hence $(a_4, b_7)$ is an exceptional pair. There is an edge connecting $b_7$ and $b_8$ and this edge has turning pair $(5, 5)$. Hence $b_7 \updownarrow b_8$ on the right angle line. Hence $(a_4, b_8)$ is an exceptional pair. There is a path connecting $b_8$ to $b_{14}$ which has bilateral symmetry. The line of symmetry contains an edge whose turning pair is $(2, 2)$. Hence $b_8 \updownarrow b_{14}$ on the right angle line. Hence $(a_4, b_{14})$ is an exceptional pair. Finally, there is an edge connecting $b_{14}$ to $b_{15}$ which has turning pair $(-1, -1)$. Hence $(a_4, b_{15})$ is an exceptional pair.

For $W_{23}$ we have $A = \{12, 13, 29, 30\}$ and $B = \{9\}$. There is an edge connecting $b_9$ to $a_{12}$ and this edge has turning pair $(-5, -5)$. Hence $(a_{12}, b_9)$
is an exceptional pair. The other 3 pairs are shown to be exceptional just as for $W_{16}$.

6.7 $W_{15}$ and $W_{18}$

For $W_{15}$ we have $A = \{1, 2, 4\}$ and $B = \{8, 9, 11, 12, 13\}$. The same arguments as in the previous section show that $a_1, a_2, a_4$ all lie at the same height when the unfolding is done with respect to a right triangle. The same goes for $b_8, b_9, b_{11}, b_{12}, b_{13}$. Finally, $a_4$ and $b_8$ are connected by an edge whose turning angle is $(5, 5)$. Hence $(a_5, b_8)$ is an exceptional pair. Hence all the pairs listed are exceptional.

For $W_{18}$ we have $A = \{1, 2, 4, 5, 6, 8, 9\}$ and $B = \{14, 16, 17, 18\}$. This case is essentially the same as the case of $W_{15}$ and we omit the details.
7 Special Cases

We still need to deal with the tiny polygons $P_4$ and $P_5$. We will show that $P_5$ is contained in 2 orbit tiles and $P_4$ is contained in 5 orbit tiles. We use essentially the same technique as in the previous chapter. Namely, we show that our verification algorithm halts when we ignore certain special pairs of vertices and then we analyze the special pairs of vertices by hand.

7.1 Covering $P_5$

We have

$P_5 : \begin{array}{c|c|c|c} 12 & 1641 & 12 & 1637 \\ 12 & 2455 & 12 & 2455 & 12 & 2459 \end{array}$ \hfill (40)

This is a tiny triangle whose hypotenuse contains $p_5 = (\pi/5, 3\pi/10)$. Let $H_+$ denote the half-plane given by $x \geq \pi/5$ and let $H_-$ denote the half-plane given by $x \leq \pi/5$. Recall that

$F = 31232313123132313213132321$ $G = 132312323132321321312323132321312312323132321323$

We will show that

$P_5 \cap H_+ \subset O(F); \quad P_5 \cap H_- \subset O(G).$ \hfill (41)

7.1.1 Dealing with $F$

In terms of our listing, we have $F = W_7$, but $P \cap H_+$ is not contained in $P_7$. Indeed $P \cap H_+$ shares a vertex with $O(F)$ and we have to work harder. We have already seen that the pair $(a_5, b_1)$ is exceptional. When we also ignore the pairs $(a_5, b_5)$ and $(a_5, b_6)$ we find that our verification algorithm produces a covering of $P_5 \cap H_+$. We already know from our analysis in the previous chapter that $f_{51} > 0$ on $P_5$. The point here is that the relevant line of bilateral symmetry has turning pair $(-2, -2)$ and hence this line has positive slope throughout $P_5$. This positive slope forces $a_5$ to lie above $b_1$.

It remains to show that $f_{55}$ and $f_{56} > 0$ on $P_5 \cap H_+$. Figure 7.1 shows a picture of $U(F, T)$ when $T$ is the right triangle corresponding to the point $p_5 \in P_5$. 

38
The edge connecting $a_5$ and $b_6$ has turning pair $(-4, 1)$. Points $(x, y) \in P_5 \cap H_+$ have the form

\[ x = \pi/5 + \epsilon; \quad y = 3\pi/10 - \epsilon - \delta. \]

Here $\epsilon$ and $\delta$ are numbers much smaller than $\pi/10$. The turning angle of the edge connecting $a_5$ to $b_5$ is therefore

\[ -\pi/2 - 3\epsilon - \delta. \]

This line has negative slope throughout $P_5 \cap H_+$ and hence $a_5 \uparrow b_5$ there.

The vertices $a_5$ and $b_6$ are connected by a path of length 2 whose line of bilateral symmetry has turning pair $(-3, 2)$. The corresponding turning angle is

\[ -\epsilon - 2\delta. \]

This line has positive slope for $(x, y) \in P_5 \cap H_+$ and hence $a_5 \uparrow b_6$ throughout $P_5 \cap H_+$.

### 7.1.2 Dealing with $G$

In terms of our listing, we have $G = W_{13}$. However, $P_3 \cap H_-$ is not a subset of $P_{13}$ so we have to do more work. When we omit the pairs $(a_1, b_{11})$ and $(a_1, b_{12})$ and $(a_1, b_{13})$ our algorithm produces a covering of $P_5 \cap H_-$. It just remains to show that the defining functions associated to these pairs are positive on $P_3 \cap H_-$. The function $f_{1,11}$ is positive on $P_3$ for the symmetry reason we discussed in the previous chapter.

Here we explain a proof which works for all 3 defining functions at once. When we run our algorithm, each of these omitted defining functions gets certified on a dyadic square which contains $P_5$. We just check that, in all 3 cases, the quadrant which contains the gradients is the $(-, -)$ quadrant. Hence $\nabla f_{ij}$ lies in the $(-, -)$. Also, these functions all vanish at $p_5$. Every
\( p \in P \cap H_- \) can be joined to \( p_5 \) by a path which points from \( p_5 \) into the \((-,-)\) quadrant. Hence \( f_{1j} > 0 \) on \( P \cap H_- \), as desired. Hence \( P \cap H_- \subset O(G) \) as desired.

Now we know that \( P \subset O(F) \cup O(G) \).

### 7.2 Covering \( P_4 \)

We have coordinates

\[
\begin{align*}
P_4 : & \quad 7 & 63 & \quad 7 & 65 & \quad 7 & 63 \\
& \quad 7 & 65 & \quad 7 & 63 & \quad 7 & 63 
\end{align*}
\]

This is a small triangle whose hypotenuse is centered on \( p_4 = (\pi/4, \pi/4) \). Figure 7.3 shows how we cover \( P_4 \cap \Delta \) by 5 regions. The regions \( c, d_1, d_2 \) are meant to be open. The segments \( e_1 \) and \( e_2 \) are meant to be open line segments. The 4 solid lines through \( p_4 \) have slope \(-1, -1/3, 0, \infty\). Compare the left hand side of Figure 2.3. The dotted line is contained in \( \partial \Delta \), and bisects \( P_4 \).

\[\text{Figure 7.3}\]

Let \( C, D_1, D_2, E_1, E_2 \) be the words listed in §2.4. The rest of the chapter is devoted to proving:

- \( c \subset O(C) \).
- \( d_1 \subset O(D_1) \).
- \( d_2 \subset O(D_2) \).
- \( e_1 \in O(E_1) \).
- \( e_2 \in O(E_2) \).
7.2.1 Dealing with C

We have $C = W_{30}$. Let $T$ be the triangle corresponding to the point $p_4$, the right isosceles triangle. Figure 7.4 shows $U(C, T)$.

![Figure 7.4](image)

The defining function $f_{ij}$ vanishes at $p_4$ when $i \in \{1, 2\}$ and $j \in \{4, 5\}$. When we run our algorithm with these vertex pairs excepted, it produces a cover of $P$ by 4 squares. Thus, all the defining functions but the excepted ones are positive on $P$. The algorithm in this case does not also verify that the gradients of the excepted functions lie in the $(-, -)$ quadrant—this isn’t true for $f_{14}$ and $f_{25}$.

In dealing with the 4 exceptional defining functions, we first compute that

$$|f_{xx}|, |f_{xy}|, |f_{yy}| \leq 2^6.$$ 

in all cases. We also note that $P_4$ is contained in a square of radius $2^{-6}$. Hence, both $\partial_x f$ and $\partial_y f$ vary by at most 2 units throughout $P_4$.

- Here is the formula for $f_{15}$.

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>0</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-3</td>
<td>4</td>
<td>-3</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>-3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

We compute that $\nabla f_{15}(p_4) = (-8, -8)$. Hence $\nabla f_{15}$ lies in the $(-, -)$ quadrant throughout $P_4$. Hence $f_{15} > 0$ on the interior of $c$.

- A similar computation to the one above gives $\nabla f_{24}(p_4) = (-8, -8)$. Hence $f_{24} > 0$ on $c$. 

41
Here is the formula for $f_{14}$:

\[
\begin{array}{ccc|ccc}
0 & 1 & 0 & 1 & (-1) \\
4 & 1 & 4 & 1 \\
4 & -3 & & \\
0 & -3 & & \\
\end{array}
\]

We compute that $f_{14}$ vanishes identically along the line $y = \pi/4$. Also, we compute that $\nabla f_{14}(p_4) = (0, -16)$. Hence $\nabla f_{14}$ has positive $y$-coordinate throughout $P_4$. Hence $f_{14} > 0$ on $c$.

The calculation for $f_{25}$ is just like the one for $f_{14}$, but with the roles of $x$ and $y$ switched. Hence $f_{25} > 0$ on $c$.

In summary, all $(a, b)$ defining functions are positive on $c$. We conclude that $c \subset O(C)$.

7.2.2 Dealing with $D_1$

In terms of our listing, $D_1 = W_9$. Here is a picture of $U(D_1, T)$, where $T$ is the right angled isosceles triangle. Taking $i$ and $j$ on the left half of the unfolding, we see that the defining function $f_{ij}$ vanishes at $p_4$ iff $i \in \{5, 6, 7, 8\}$ and $j \in \{1, 2, 3, 4, 10\}$. (The center point by convention counts as a vertex on the left half.) When we run the algorithm with these pairs excepted, it produces a covering of $P_4$ by 3 squares. Once again, the algorithm here does not verify anything about the gradients of the exceptional defining functions.

Figure 7.5

Reflection in a certain edge $e$ swaps $a_6$ and $a_8$. The turning pair for $e$ is $(2, 2)$. Since the leftmost edge stays vertical for all points in the parameter space, $e$ has negative slope throughout $P_4$. Hence $a_6 \uparrow a_8$ throughout $P_4$. This eliminates $a_6$ from consideration.
Figure 7.6 shows $U(D_1, T')$ where $T'$ is a triangle corresponding to a point of $\Delta$ between $e_1$ and the right angle line. (This point isn’t actually in $d_1$, because such points give rise to a picture which looks almost identical to Figure 7.5; we wanted to show the difference dramatically.) Figure 7.6 serves as a reality check to the arguments we give below.

$a_6$ is connected to $a_7$ by an edge whose turning pair is $(0, 2)$. As long as $y < \pi/4$ this edge has positive slope and $a_7 \uparrow a_6$. This condition holds in $d_1$. This eliminates $i = 7$ from consideration. Similar arguments show that $b_2 \uparrow b_1$, $b_3 \uparrow b_2$ and $b_3 \uparrow b_4$ throughout $d_1$. All in all, we just have to deal with the 4 defining functions $f_{ij}$ where $i \in \{5, 8\}$ and $j \in \{4, 10\}$. Here is the analysis:

- $a_8$ and $b_4$ are swapped by reflection in an edge whose turning pair is $(1, 3)$. This edge has positive slope throughout the interior of $d_1$, and vanishes on $e_1$, the line of slope $-1/3$ through $p_4$. Hence $a_8 \uparrow b_4$ throughout $d_1$. Hence $f_{84} > 0$.

- $a_5$ and $b_{10}$ are swapped by reflection in an edge whose turning pair is $(2, 2)$. Hence $f_{5,10} > 0$ on $d_1$.

- $b_4$ and $a_5$ are connected by an edge whose turning pair is $(-2, 4)$. This edge has positive slope in $d_1$. Hence $f_{54} > 0$ in $d_1$.

- $a_8$ and $b_{10}$ are connected by an edge whose turning pair is $(0, 2)$. This line has negative slope in $d_1$. Hence $f_{8,10} > 0$ in $d_1$.

This takes care of all the cases. Hence $d_1 \subset O(D_1)$. 

43
### 7.2.3 Dealing with $D_2$

In terms of our listing, $D_2 = W_{87}$. The analysis of $D_2$ is almost identical to the analysis of $D_1$. We will omit most of the details, but illustrate the main ideas with pictures. Figure 7.7 shows $U(D_2, T)$ and Figure 7.8 shows $U(D_3, T')$. Here $T'$ is a triangle corresponding to a point which lies between the lines $e_1$ and $e_2$. (We have gone outside $d_2$ to get a more dramatic picture.)

![Figure 7.7](image1)

![Figure 7.8](image2)

When we except all the index pairs entailed by Figure 7.7 our algorithm produces a covering of $P_4$ by 4 squares. Using the turning pair arguments, as for $D_1$, we eliminate all the indices except $i \in \{7, 12\}$ and $j \in \{6, 14\}$. Figure 7.8 is a typical picture of the signs of the slopes of the relevant These 4 defining functions have the same analysis as for $D_1$.

### 7.2.4 Dealing with $E_1$

In terms of our listing, $E_1 = W_{107}$. Recall that $e_1$ is the intersection of the line of slope $-1/3$ through $p_4$ with $P_4$. Figure 7.9 shows a picture of $U(E_1, T)$. When we run our algorithm with all the excepted vertices, it produces a
covering of $P_4$ by 47 squares. We also check, during the algorithm, that $\nabla f$ has positive dot product with the vector $(-3, 1)$ throughout $P_4$ whenever $f$ is an exceptional defining function. This shows that all the exceptional defining functions are negative on $e_1$. Hence $e_1 \in O(E_1)$.

Figure 7.9

Remark: Our gradient check is just a small tweak of the silver method. We compute $\nabla f$, then add all the error bounds coming from the second partials, and check that the entire “error box” makes positive dot product with $(-3, 1)$.

7.2.5 Dealing with $E_2$

In terms of our listing, $E_2 = W_{85}$.

Recall that $e_2$ is the intersection of the horizontal line through $p_4$ with $P_4$. Figure 7.10 shows a picture of $U(E_2, T)$.

Figure 7.10

When we run our algorithm with all the excepted vertices, it produces a covering of $P_4$ by 29 squares. We also check, during the algorithm, that $\partial_x f < 0$ throughout $P_4$, whenever $f$ is an exceptional defining function. This shows that all the exceptional defining functions are negative on $e_2$. Hence $e_2 \in O(E_2)$. 

45
8 Overview

Our goal in §8-9 is to cover the polygon $P_1$ by an infinite union of orbit tiles. As we discussed in §2 we are going to cover $P_1$ with two infinite families of orbit tiles. We will deal with the first family in this chapter and the second family in §9. The words in the first family are palindromes in the same sense that the words $W_7, ..., W_{29}$ are palindromes. Accordingly, their unfoldings have bilateral symmetry. At this point the reader can forget essentially everything done in §2-8.

8.1 The Unfoldings

Figures 8.1-8.3 show unfoldings for $A_1$, $A_2$, and $A_3$ respectively, with for various choices of triangle. The pattern continues in the obvious way.
To cover (most of) $P_1$ with our first family of tiles, we break $P_1$ into subregions, each of which is covered by a single tile. Let $N_n$ denote the open triangular sector of $N_0$ bounded by

- The line $y = 3\pi/4$;
- The line through $(0, \pi/2)$ having slope $-(n + 1)/2$.
- The line through $(0, \pi/2)$ having slope $-(n + 2)/2$.

The difference

$$P_1 - \bigcup_{n=1}^{\infty} N_n$$

is an infinite union of line segments of rational slope. We will show that $N_n \subset O(A_n)$ for all $n$. This, all but countably many line segment of $P_1$. We use the second infinite family to cover these line segments. In this chapter we will concentrate on the case $n = 2$, which is sufficiently complex to contain all the ideas in the proof. At the end we will explain how the argument generalizes.

### 8.2 The Top Vertices

By symmetry it suffices to consider the vertices on the right half of the unfolding. We change our labelling scheme somewhat, and start counting our vertices from the center, as in Figure 8.4. Figure 8.4 shows an enlarged version of Figure 8.2. This picture will be our constant companion throughout our analysis.
Given a ray or vector $r$ we let $\theta(r)$ denote the counterclockwise angle through which we must rotate $(0, 1)$ so that it points in the direction of $r$. Unlike in §4 we work mod $2\pi$ rather than mod $\pi$. To simplify we write, for instance, $\theta(a_1b_2) = \theta(-\vec{a}_1\vec{b}_2)$. We let $z$ denote the angle opposite edge 3, so that $x + y + z = \pi$. Here are the angles of importance to us.

$$\theta(a_1a_2) = 6x + \pi; \quad \theta(a_7a_8) = x;$$

(43)

A less obvious computation is:

$$\theta(b_7a_5) = \pi + 3x + 2y$$

(44)

Here is a derivation of the third equation. We rotate $\overrightarrow{b_1a_1}$ by $6x$ to get $\overrightarrow{a_2a_1}$. Then we rotate $\overrightarrow{a_2a_1}$ by $2y$ to get to $\overrightarrow{a_2b_7}$. Then we rotate $\overrightarrow{a_2b_7}$ by $-3x$ to get to $\overrightarrow{a_5b_7}$. We rotate this last ray by $\pi$ to reverse the direction.

The conditions $(x, y) \in N_2$ give rise to the angle constraints

$$x \in (0, \frac{\pi}{12}); \quad y \in \left(\frac{3\pi}{8}, \frac{\pi}{2}\right).$$

(45)

See Figure 8.1. From Equation 44 we now get $\theta(a_1a_2) \in (\pi, \pi + \pi/2)$. But this means that $a_1 \uparrow a_2$. Similarly, Equation 44 tells us that $\theta(a_7a_8) = x \in (0, \pi/2)$. Hence $a_8 \uparrow a_7$.

It remains to compare the heights of the vertices $a_3, ..., a_7$. We are interested in the “fan”, a polygon whose vertices are $b_7$ and $a_3, ..., a_7$. Note that $\overrightarrow{b_7a_5}$ is the line of bilateral symmetry for $F$. 

48
Our constraint \((x, y) \in N_2\) gives
\[
\pi - \pi/4 < 2y < \pi - 3x; \quad 3x < \pi/4. \quad (46)
\]
Combining these bounds with Equation 44 we get
\[
\theta(b_7a_5) \in \left(\frac{7\pi}{4}, 2\pi\right) \quad (47)
\]
In particular, \(b_5a_7\) has positive slope.

Equation 46 guarantees that \(F\) is contained in a halfplane. We can write \(F = F_1 \cup F_2\) where \(F_1\) is the convex hull of \(b_7\) and the odd vertices \(a_3, a_5, a_7\). Then \(F_2\) is a union of 2 small triangles, as shown in Figure 8.5. Given the conditions on \(F\) we see that \(a_7\) is the lowest vertex, amongst the odd vertices of \(F\).
Since the line of symmetry of $F$ has positive slope and $F$ lies in a half-plane, the vertices $a_3, a_5, a_7$ have increasing $X$ coordinates. Moreover, the two triangles $a_3a_4a_5$ and $a_5a_6a_7$ are oriented clockwise. Finally, the line segments connecting $b_7$ to the even vertices are longer than the line segments connecting $b_7$ to the odd vertices. From all that we have said, it follows that each even vertex lies above at least one of the adjacent odd vertices. All in all $a_j \uparrow a_2$ for $j = 3, 4, 5, 6$. We have eliminated all the $a$ vertices except $a_2$ and $a_7$.

8.3 The Bottom Vertices

We now make the same sorts of arguments as above, but for the bottom vertices. This time we can eliminate the vertex $b_j$ if we can show that $b_i \uparrow b_j$ throughout $N_2$.

Since $y < \pi/2$ the line $b_1b_2$ has positive slope. Hence $b_2 \uparrow b_1$. To understand the vertices $b_2, \ldots, b_4$ we consider the “fan” whose vertices are $a_1, b_2, b_3, b_4, b_5, b_6$. This polygon is isometric to the one considered in the previous subsection. The line of symmetry of $F$ is $a_1b_4$. This line has negative slope because of the fact that $3x < \pi/2$. The same argument as above now shows that $b_6 \uparrow b_j$ for $j = 2, 3, 4, 5$.

The angle between $\overrightarrow{b_7b_6}$ and $\overrightarrow{b_7a_5}$ is $4x < \pi/3$. Combining this information

\footnote{To distinguish between the $(x, y)$ coordinates of the parameter space and the $(X, Y)$ coordinates of the Euclidean plane in which we draw the unfoldings we will henceforth use capital letters for the $X$ and $Y$ coordinates of the unfoldings.}
with Equation 47 we see that
\[
\theta(b_7b_6) \in \left(\frac{7\pi}{4}, \frac{5\pi}{2}\right) \equiv \left(-\frac{\pi}{4}, \frac{\pi}{3}\right).
\] (48)

From this we see that \(b_6 \uparrow b_7\). We have eliminated all the \(b\) vertices except \(b_8\) and \(b_8\).

### 8.4 The Remaining Pairs

We have 4 pairs left to analyze. Consider first \((a_7, b_8)\). We have
\[
\theta(b_8a_7) = y \in (0, \pi/2).
\]
Hence \(a_7 \uparrow b_8\).

Now consider \((a_2, b_6)\). We have \(\theta(a_2b_6) = 4x+y \in (\pi/2, \pi)\). Hence \(a_2 \uparrow b_6\).

Now consider \((a_2, b_8)\). Note that \(a_2\) and \(b_8\) are symmetrically located with respect to our favorite line \(b_7a_5\). Thus \(a_2\) and \(b_8\) have the same height iff our line is vertical. From Equation 44 and Equation 47 we see that this happens for a point in closure\(\left(N_n\right)\) iff \(2y + 3x = \pi\). That is, \((x, y)\) has to lie on the right boundary line of \(N_2\). Equation 47 shows that \(a_2 \uparrow b_8\) for \((x, y) \in N_2\).

Now consider \((a_7, b_6)\). Note that \(a_7\) and \(b_6\) have the same height iff the line \(b_7a_4\) is vertical. Essentially the same analysis as we have already done shows that our line has negative slope for \((x, y) \in N_2\), and is vertical for \(2y + 4x = \pi\). Hence \(a_4 \uparrow b_7\). The two points have the same height when \((x, y)\) is in the left boundary of \(N_2\).

In summary \(N_2 \subset O(A_2)\).

### 8.5 The General Case

We deal with the top vertices first. The general versions of Equations 43 and 44 are
\[
\theta(a_1a_2) = (2n + 2)x + \pi; \quad \theta(a_{2n+3}a_{2n+4}) = x; \quad (49)
\]
\[
\theta(b_{2n+3}a_{n+3}) = \pi + (n+1)x + 2y \quad (50)
\]
Equation 49 eliminates \(a_{2n+4}\) and \(a_1\) from consideration.
The conditions \((x, y) \in N_n\) give rise to the angle constraints
\[ x \in (0, \frac{\pi}{4n + 8}); \quad y \in (\frac{3\pi}{4}, \frac{\pi}{2}). \quad (51) \]
For \((x, y) \in N_n\) we have
\[ \pi - \frac{\pi}{4} < 2y < \pi - (n + 1)x. \quad (52) \]
These equations combine together with Equation 50 to show that the line \(b_{2n+3}a_{n+3}\) has positive slope. This line is the center of symmetry of the fan with vertices \(b_{2n+3}, a_3, \ldots, a_{2n+3}\). The same argument as above then shows that \(a_2, \ldots, a_{2n+2}\) lie above \(a_j \uparrow a_{2n+3}\) for \(j = 2, \ldots, 2n + 2\). In this way we eliminate everything but \(a_2\) and \(a_{2n+3}\).

Essentially the same argument eliminates all the \(b\) vertices except \(b_{2n+2}\) and \(b_{2n+4}\). The key point is that the line \(a_1b_{n+2}\), which is the line of symmetry for the fan with vertices \(a_1; b_2, \ldots, b_{2n+2}\), has negative slope. This follows from Equation 51.

The analysis of the edges is the same in the general case. The main points that need to be observed are:

- The points \(a_2\) and \(b_{2n+4}\) have the same height iff \(b_{2n+4}a_{n+3}\) is vertical, and this happens iff \(2y + (n + 1)x = \pi\).

- \(a_{2n+3}\) and \(b_{2n+2}\) have the same height iff the line \(b_{2n+3}a_{n+2}\) is vertical, and this happens iff \(2y + (n + 2)x = \pi\).

All this information assembles together in the same way as in the case \(n = 2\), to show that \(N_n \subset O(A_n)\).
9 The Second Family

9.1 The Unfoldings

Let $N'_n$ denote the open line segment which is the common boundary of $N_n$ and $N_{n+1}$. Then $N_n$ has slope $-(n + 2)/2$, and has endpoints

$$
(0, \frac{\pi}{2}); \quad \left(\frac{\pi}{4n + 8}, \frac{3\pi}{8}\right).
$$

We will show that $N'_n \subset \tau'_n$. We are just trying to show that a single line segment lies in the tile. This is all we need, and it allows us to use an additional relation between the angles of the triangles of interest to us.

We will concentrate on the case $n = 1$ and at the end explain the changes needed for the general case. We go back to our initial convention of labelling the vertices starting from the left. The line $N'_1$ corresponds to triangles whose two acute angles satisfy

$$
3x + 2y = \pi
$$

Figure 9.1

Figure 9.1 shows a picture of $U(W'_1, A)$ for some triangle satisfying Equation 54. The (near) central edge $(a_8, b_8)$ is parallel to both $(a_1, b_1)$ and $(a_{13}, b_{13})$. Indeed the portion of $U(W'_1, A)$ to the left of $(a_8, b_8)$ is isometric to the right half of $U(W_2, A)$ and the portion to the right of $(a_8, b_8)$ is
isometric to the left half of $U(W_1, A)$. This is fitting, because $O(W'_1)$ fits “between” $O(W_1)$ and $O(W_2)$.

### 9.2 Estimates for Rotation Angles

We define $\theta(r)$ as in §8.2. That is, $\theta(r)$ denotes the counterclockwise angle through which $(0, 1)$ must be rotated to produce a vector parallel to $r$. Recall that $x$ is the small angle of our triangles. The goal of this section is to prove:

\[
\begin{align*}
\theta(b_{13}a_{13}) &\in (0, x); \\
\theta(a_{12}, a_{13}) &\in (-x, 0).
\end{align*}
\]  

(55)

**Lemma 9.1** There is some $\epsilon > 0$ such that $\theta(a_{12}, a_{13}) \in [0, \epsilon)$ is impossible.

**Proof:** The conditions in Equation 30 guarantee that the lines $\overline{a_{11}b_{12}}$, $\overline{a_{8}b_{7}}$, and $\overline{a_{5}b_{2}}$ are all parallel to $\overline{a_{12}a_{13}}$. By symmetry, the points $b_{13}$ and $a_3$ are related by a reflection in $\overline{a_{8}b_{7}}$. The point $a_3$ and $b_1$ are related by a reflection in $\overline{a_{2}b_{2}}$.

If $\overline{a_{12}a_{13}}$ is vertical or has negative slope, then $a_3$ lies below $b_{13}$. On the other hand, if $\overline{a_{12}a_{13}}$ is vertical has large negative slope then $\overline{a_{2}b_{2}}$ has negative slope. (Here we are using $3x \leq \pi/4$. Compare Equation 46.) But then $b_1$ lies below $a_3$. But then $b_1$ lies below $b_{13}$, a contradiction. ♠

![Figure 9.1](image-url)
Lemma 9.2 There is some $\epsilon > 0$ such $\theta(b_{13}, a_{13}) \in (-\epsilon, 0]$ is impossible.

Proof: Condition 54 guarantees that $\overline{a_{10}b_{12}}$, $\overline{a_{8}b_{8}}$, $\overline{a_{4}b_{2}}$, and $\overline{a_{1}b_{1}}$ are all parallel to $\overline{a_{13}b_{13}}$. Let $a_0$ denote the reflection of $a_2$ through the line $a_1b_1$. Our normalization puts $a_0$ and $a_{12}$ at the same height. The points $a_0$, $a_2$, $a_6$ are successively related to each other by reflections in the lines mentioned above. Likewise, the points $a_{12}$, $b_{11}$, $b_5$ are successively related to each other by reflections in the lines mentioned above. If $\overline{a_{13}b_{13}}$ is either vertical or has sufficiently large negative slope then $b_5$ lies above $a_6$.

The points $a_6$ and $b_3$ are related to each other by a reflection through $\overline{b_2a_7}$. The points $b_3$ and $b_5$ are related to each other by reflection in the line $\overline{a_8b_4}$. If $\overline{a_{13}b_{13}}$ is either vertical or has sufficiently large negative slope then these two last mentioned lines both have negative slope and hence $b_5$ lies below $a_6$. This is a contradiction. ♠

Figure 9.1

In the terminology of §4 consider the 1-spine for $U(W'_1, T)$. This is the path of short edges defined by the vertex sequence:

$$(b_1, a_2, ..., a_7, b_3, ..., b_{11}, a_9, a_{10}, a_{11}, a_{12}, b_{13})$$

Consider what happens when $x \in N'_1$ tends to 0 and the corresponding triangles are scaled so that the edges of the 1-spine have unit length. The
external angles between consecutive segments of the 1-spine converge to 0 and hence the 1-spine converges to a completely horizontal path. But this means that \(\theta(b_{13}a_{13}) \to 0\) and \(\theta(a_{12}a_{13}) \to 0\) in the limit we are taking. From the two lemmas above equation 55 must hold for sufficiently small \(\alpha\). We also know that \(\theta(b_{13}a_{13}) = x + \theta(a_{12}a_{13})\). It now follows from continuity and our two lemmas that Equation 55 holds for all \(x\), when \((x, y) \in N'_1\).

### 9.3 The Top Vertices

We will use the same elimination technique as in §8.2.

![Figure 9.1](image)

1. From Equation 55 we get \(\theta(a_2a_1) \in (x, 2x)\). Hence \(a_1 \uparrow a_2\).

2. We have \(\theta(b_2a_5) = \theta(a_{12}a_{13})\). By Equation 55 we see that \(b_2a_5\) has positive slope. This line happens to be the line of symmetry for the fan with vertices \(b_2; a_3, ...a_7\). The same argument as in §8.2 shows that \(a_j \uparrow a_7\) for \(j = 2, 3, 4, 5, 6\).

3. From Equation 55 we conclude that \(\theta(a_7a_8) \in (0, 6x) \in (0, \pi/2)\). Hence \(a_8 \uparrow a_7\).

4. \(a_7\) and \(a_9\) are related by reflection through \(a_8b_7\), a line with negative slope. Hence \(a_7 \uparrow a_9\).
5. Note that $a_{12}$ and $a_{10}$ are related by a reflection through $b_{12}a_{10}$, a line which has negative slope because it is parallel to $a_{13}b_{13}$. Hence $a_{10} \uparrow a_{12}$.

6. $a_{12}$ and $a_0$, the point defined in the proof of Lemma 9.2, are at the same height. Moreover, $a_0$ and $a_2$ are related by a reflection through the negatively sloped $a_1b_1$. Hence $a_2 \uparrow a_{12}$.

We have eliminated ass the $a$ vertices except $a_9$ and $a_{12}$.

### 9.4 The Bottom Vertices

![Figure 9.1](image)

The same argument as in Item 3 above shows that $b_4 \uparrow b_2$ and that $b_{13} \uparrow b_{12}$.

Considering the fan with vertices $a_8, b_3, ..., b_{11}$, whose line of symmetry $a_8b_7$ has positive slope, we see that $b_3 \uparrow b_j$ for $j = 4, ..., 11$.

We have eliminated all the $b$ vertices except $b_1$, $b_3$, and $b_{13}$. Note that $b_{13}$ and $b_1$ are at the same height, from the way we have normalized. Thus, we just have to consider $b_3$ and $b_{13}$.

Just as in §8 we have 4 pairs left to analyze. The pairs involving $b_{13}$ are easy to handle and we will dispose of them right now.

Consider the pair $(a_{12}, b_{13})$. We have $\theta(b_{13}, a_{12}) \in (y, x+y)$. We also have $3x + 2y = \pi$. Hence $\theta(b_{13}, a_{12}) \in (0, \pi/2)$. Hence $a_{12} \uparrow b_{13}$. 

57
Consider the pair \((a_9, b_{13})\). Since \(b_{13}\) and \(a_9\) are related by reflection through \(b_{12}a_{12}\) and \(\theta(b_{12}, a_{12}) = \theta(a_{12}, a_{13}) \in (-x, 0)\) we have \(a_9 \uparrow b_{13}\).

It remains to consider the pairs \((a_9, b_{3})\) and \((a_{12}, b_{3})\). Given Lemma 9.3 below, the result \(a_{12} \uparrow b_{3}\) implies the result \(a_9 \uparrow b_{3}\). Our strategy is to first prove Lemma 9.3 and then to deal with the pair \((a_{12}, b_{3})\) directly.

### 9.5 Eliminating one of the Pairs

**Lemma 9.3** \(a_9\) lies above the line \(b_{3}a_{12}\).

**Proof:** Let \(\theta_1\) denote the angle \(\angle b_{3}a_{8}a_{9}\). Let \(\theta_2\) denote the angle \(\angle b_{12}a_{9}a_{12}\). The point \(a_9\) lies on \(b_{3}a_{12}\) iff \((\pi - \theta_2) + \theta_1 = \angle a_{8}a_{9}b_{12} = 2y\). Using this fact as a guide, we check signs to determine that \(a_9\) lies above \(b_{3}a_{12}\) provided that \(\pi - \theta_2 + \theta_1 > 2y\).

![Figure 9.1](image)

Using the law of sines we can normalize so that our triangles all have side lengths \(\sin(x), \sin(y), \sin(z)\). Let \(\theta_3 = \angle a_{9}a_{12}b_{12}\). Looking at the triangle with vertices \(a_3, a_8, b_9\) and using the law of sines we get

\[
\theta_3 = \frac{\sin(z)}{\sin(y)} \theta_1. \tag{56}
\]

Using that the sum of the 3 angles in a triangle is \(\pi\), together with Equation 54, we get:
\[
\theta_1 + \theta_3 = \pi - 9x = \pi - 4 \times 3x + 3x = -3\pi + 8y + 3x = 5y - 3z. \tag{57}
\]

The first equation uses Equation 54. Solving for \(\theta_1\) we get:

\[
\theta_1 = \frac{(5y - 3z)\sin(y)}{\sin(y) + \sin(z)}. \tag{58}
\]

Let \(\theta_4 = \angle a_9a_{12}b_{12}\).

From the law of sines we have

\[
\theta_4 = \frac{\sin(z)}{\sin(y)} \theta_2. \tag{59}
\]

We also have

\[
\theta_2 + \theta_4 = \pi - 3x = \pi - 2 \times 3x + 3(\pi - y - z) = \\
\pi - 2(\pi - 2y) + 3\pi - 3y - 3z = 2\pi + y - 3z \tag{60}
\]

(We have complicated this equation so that it readily generalizes.) Solving for \(\theta_2\) we get

\[
\theta_2 = \frac{(2\pi + y - 3z)\sin(y)}{\sin(y) + \sin(z)} \tag{61}
\]
Using Equations 58 and 61 we compute

\[(\pi - \theta_2 + \theta_1) - 2b = \frac{(\pi - 2y)(\sin(z) - \sin(y))}{\sin(y) + \sin(z)}. \quad (62)\]

Note that \(\sin(z) > \sin(y)\). The expression in Equation 62 is positive as long as \(y < \pi/2\), which is certainly our situation. ♠

### 9.6 The Last Pair

Finally we come to the pair \((a_{12}, b_3)\). At this point it is useful to cycle our picture so that \(b_3\) is all the way to the left. See Figure 9.2. Figure 9.2 is cut-and-paste equivalent to Figure 9.1.

Note that \(a_{12}\) lies to the left of both \(b_{14}\) and \(b_{15}\). To see this note that \(a_{14}\) and \(b_{15}\) are related by reflection in the nearly vertical line \(b_{14}a_{17}\) and \(a_{14}\) and \(a_{12}\) are related by reflection in the nearly vertical line \(a_{13}b_{13}\). These lines make an angle of less that \(x\) with the vertical, from Equation 55. The same argument shows that \(b_3\) lies to the left of \(a_{12}\).

Let \(\sigma_1\) and \(\sigma_2\) respectively denote the slopes of \(\overline{b_{15}a_{12}}\) and \(\overline{b_3b_{15}}\) when the picture is rotated so that \(\overline{b_{15}b_{14}}\) is horizontal. Since \(a_{12}\) and \(b_3\) lie to the left of both \(b_{14}\) and \(b_{15}\) the slopes \(\sigma_1\) and \(\sigma_2\) are finite. We will show that that
\(\sigma_1 < \sigma_2\). This, together with the fact that \(b_3\) lies to the left of \(a_{12}\), shows that \(a_{12} \uparrow b_3\), as desired.

Consider the path of 8 vectors \(v_1, \ldots, v_8\) defined by the vertex sequence

\[
(b_{15}, b_{14}, a_{14}, a_{13}, a_{12}, b_{12}, b_{11}, a_8, b_3).
\]  

(63)

In the terminology of §4, this path is part of the 1-spine. The first vector points from \(b_{15}\) to \(b_{14}\), and so forth. These vectors all have the same length, which we normalize to be 1.

Let \(\theta_k\) denote the counterclockwise angle by which \(v_1\) must be rotated to produce \(v_k\). We now calculate these vectors.

Looking at Figure 9.2 have \(\theta_1 = 0\) and

- \(\theta_2 = 6x + \pi = -4y + \pi\).
- \(\theta_3 = 6x - 2z = -4y - 2z\).
- \(\theta_4 = 4x - 2z + \pi = -2y + \pi\).
- \(\theta_5 = 4x - 4z = -2y - 2z\).
- \(\theta_6 = 8x - 4z + \pi = -4y + \pi\).
- \(\theta_7 = 8x - 2z = -4y + 2z\).
- \(\theta_8 = -2z + \pi\).

In working out some of the equalities we used the relations

\[6x = -4y; \quad 2\alpha_j = -2\alpha_{j-1} - 2\alpha_{j+1}.\]

(64)

These relations hold mod \(2\pi\), which is all we care about. The first equation comes from Equation 54. To give an example derivation, we will work out the derivations for \(\theta_4\) and \(\theta_6\):

\[4x - 2z = 4x + 2x + 2y = 6x + 2y = -4y + 2y = -2y.\]

\[8x - 4z = 12x - 4x - 4z = -8y + (4y + 4z) - 4z = -4y.\]

We want to eliminate \(x\) because this is the approach which generalizes to the other words \(W'_n\).
To compute the slope of a point, we divide its $y$ displacement by its $x$-displacement. We set

$$C_k = \sum_{j=1}^{k} \cos(\theta_j); \quad S_k = \sum_{j=1}^{k} \sin(\theta_j).$$

Then $\sigma_1 = S_4/C_4$ and $\sigma_2 = S_8/C_8$. Since $\sigma_1$ and $\sigma_2$ are both finite the terms $C_4$ and $C_8$ never vanish. We compute that

$$\sigma_1 - \sigma_2 = \frac{2\sin(z)}{C_4C_8} (\cos(z) - \cos(y)).$$

The condition $z \in (\pi/2, \pi)$ makes $\cos(z) < 0$. The condition $y \in (0, \pi/2)$ makes $\cos(y) > 0$. Hence $\sigma_1 - \sigma_2 < 0$. Hence $\sigma_1 < \sigma_2$.

This completes our proof that $N'_1 \subset \tau'_1$, the first tile in the second family.

### 9.7 The General Case

For $N_n'$ we have the angle condition

$$(n + 2)x + 2y = \pi.$$  

The proof of Equation 55 works exactly the same way, with the same outcome. Armed with Equation 55 we can use the same arguments as above to eliminate all the pairs of vertices except $(b_3, a_{3n+6})$ and $(b_3, a_{4n+8})$. Figure 9.3 shows the situation for $n = 2$.

![Figure 9.3](image-url)
Lemma 9.3 works in general, with the following changes: Equation 57 becomes

\[
\theta_1 + \theta_3 = \pi - 4 \times (n + 2)x - 3x = 8 y - 3x = 5y + 3x. \tag{68}
\]

Equation 60 becomes

\[
\theta_2 + \theta_4 = 2 \times (n + 2)x - 3x = 2(\pi - 2y) - 3(\pi - y - z) = 2\pi + y - 3z. \tag{69}
\]

In other words, we get the same equations! The rest of the proof is the same.

The analysis of the pair \((b_4, a_{4n+8})\) generalizes in the same way. In general, we consider the path of vectors

\[
(b_{4n+11}, b_{4n+10}, a_{4n+10}, a_{4n+9}, a_{4n+8}, b_{4n+8}, b_{4n+7}, a_{2n+6}, b_3). \tag{70}
\]

The angle sequences we get are

- \(\theta_2 = 2(n + 2)x + \pi = -4y + \pi.\)
- \(\theta_3 = 2(n + 2)x - 2z = -4y - 2z.\)
- \(\theta_4 = 2nx - 2z + \pi = -2y + \pi.\)
- \(\theta_5 = 2nx - 4z = -2y - 2z.\)
- \(\theta_6 = (4n + 4)x - 4z + \pi = -4y + \pi.\)
- \(\theta_7 = (4n + 4)x - 2z = -4y + 2z\)
- \(\theta_8 = -2z + \pi.\)

As above we will show the derivations for \(\theta_4\) and \(\theta_6\).

\[
2nx - 2z = 2nx + 2x + 2y = (2n + 2)x + 2y = -4y + 2y = -2y.
\]

\[
(4n + 4)x - 4z = (4n + 8)x - 4x - 4z = -8y + 4y + 4z - 4z = -4.
\]

The rest of the proof is the same.

63
10 Computational Details

10.1 BigIntegers and BigIntervals

We wrote McBilliards in Java. See www.java.sun.com for information about this language.

The Java programming language has a class called the BigInteger. The BigInteger is an integer, with an “arbitrary” number of base 10 digits. Here “arbitrary” means “subject to the memory limitations of the machine”. Once two BigIntegers are defined, they can be added, subtracted, multiplied, and even exponentiated. If the process of computing the resulting quantity does not exhaust the memory of the machine, then the result is correct. It would probably take integers billions of digits long to exhaust the memory of the machine. In our case we work with integers, all of which have fewer than 200 digits. For this reason, we are convinced that the basic arithmetic operations of the BigInteger class work without fail on the numbers we supply.

Our basic method is to convert all our calculations into integer calculations and then to use BigIntegers to get the calculations exactly right. Our trick is to multiply the naturally computed quantities of interest to us by a huge integer, namely \(2^{106}\), and then trap these quantities inside an interval of BigIntegers. We then perform a calculation using BigInteger arithmetic, and in the end produce in interval of BigIntegers which contains \(2^{106}\) times the quantity of interest to us.

The only real-valued functions we compute are the ones in Equation 23 and 24. Once we have these quantities, we do make some further algebraic manipulations, as discussed in connection with the gold and silver methods of §5. However, once we have finished with Equations 23 and 24, we have our intervals of BigIntegers and then we manipulate them as discuss below.

We define a BigInterval to be a pair \((L, R)\) of BigIntegers, with \(L \leq R\). There are several basic operations which we can perform on these intervals:

- \((L_1, R_1) + (L_2, R_2) = (L_1 + L_2, R_1 + R_2)\).
- \((L_1, R_1) - (L_2, R_2) = (L_1 - R_2, L_2 - R_1)\).
- \((L_1, R_1) \times (L_2, R_2) = (L_3, R_3)\), where \(L_3 = \min(L_1L_2, L_1R_2, L_2R_1, L_2R_2)\) and \(R_3 = \max(L_1L_2, L_1R_2, L_2R_1, L_2R_2)\).

These operations have the following property: If \(x_j \in (L_j, R_j)\) for \(j = 1, 2\) then \(x_j \times y_j \in (L_1, R_1) \times (L_2, R_3)\). Here \((\cdot)\) is any of the 3 operations just
mentioned. All our calculations boil down to showing that \( x > 0 \) or \( x < 0 \) for some real number \( x \). We do our calculations in such a way as to produce a BigInterval \((L, R)\) such that \( 2^{106}x \in (L, R) \). We would show that \( x < 0 \) by showing that \( R < 0 \) and we would show that \( x > 0 \) by showing that \( L > 0 \).

### 10.2 The Interval Cosine Function

Looking at Equations 23 and 31 we see that we need some way to deal with the sine and cosine functions. When we run our subdivision algorithm, we find that it never produces a dyadic square whose side length is less that \( 2^{18} \). For this reason, we are only evaluating the sine and cosine functions on numbers \(^3\) of the form

\[
\frac{\pi}{2} \frac{k}{2^{20}}.
\]

Using the identities:

\[
\sin(x) = \cos(\pi/2 - x); \quad \cos(x + n\pi) = (-1)^n \cos(x)
\]

we see that it suffices to consider the \( 2^{21} \) values

\[
c_k := 2^{53} \cos(\pi/2 \times \frac{k}{2^{20}}); \quad k = 0, \ldots, 2^{21} - 1.
\]

(There is nothing special about \( 2^{53} \). We like it because it affords about the same precision as a double in C.)

We now explain how we produce a BigInterval \( I_k \) such that \( c_k \in I_k \). Once we have \( I_k \), we evaluate Equations 23 and 31 using the operations discussed above. Producing \( I_k \) is quite easy. The tricky part is proving rigorously that our method really works. We know that there exist packages in Java which perform this task for the elementary functions, but we prefer to work from scratch. We want to stress that it doesn’t really matter how we produce our BigInterval \( I_k \). The important point is the proof that \( c_k \in I_k \). However, it seems worth explaining our simple method.

#### 10.2.1 Producing the Interval

We introduce the routine \texttt{cosBestApprox}. When we evaluate this routine on the pair \((k, 20)\) is produces a BigInteger \( C_k \). We then take

\[
I_k = (C_k - 4, C_k + 4).
\]

\(^3\)Actually we just need \( 2^{18} \) rather than \( 2^{20} \) but we want to give ourselves a little cushion here.
The routine $\text{cosBestApprox}$ essentially computes “the usual” cosine on the relevant point—here $n = 20$ and $k$ is as above—and then rounds to the nearest BigInteger. Our method uses the BigDecimal class, which is just a BigInteger, together with a separate integer which tells where to put the decimal point. Here is our code, all of which can be found online in the file Deg100Trig.java.

```java
public static BigInteger cosBestApprox(int k, int n) {
    double d=Math.PI/2.0;
    d=d*k/Math.pow(2.0,n);
    d=Math.cos(d);
    BigDecimal Y1=new BigDecimal(d);
    BigInteger BIG=getBIG();
    BigDecimal Y2=new BigDecimal(BIG);
    Y1=Y1.multiply(Y2);
    BigInteger X=Y1.toBigInteger();
    return(X);
}
```

The BigInteger BIG is $2^{53}$. Here is the routine which gets it:

```java
public static BigInteger getBIG() {
    BigInteger BIG=new BigInteger("9007199254740992");
    return(BIG);
}
```

10.2.2 Checking that the Method Works

What we actually show is that

$$2^{357}20!c_k \in 2^{357}20!I_k.$$

A huge number like this appears fairly naturally because we want to clear denominators in some Taylor series approximations for cosine.

For $j = 0, 1, ..., 10$ let $L_j$ be the greatest integer less than

$$\frac{2^{400}20!}{2^{400j}(2j)!} \times (\pi/2)^{2j}.$$  \hfill (71)

Let $R_j = L_j + 1$. We compute these 20 integers using Mathematica, which has a reliable arbitrary precision evaluation of the trig functions. The reader can see our values in the file Deg100Trig.java. Consider the sums

66
\[ A_k = L_0 - R_1k^2 + L_2k^4 - R_3k^6 + ... - R_{10}k^{20} \]  

(72)

\[ B_k = R_0 - L_1k^2 + R_2k^4 - L_3k^6 + ... + R_9k^{18} \]  

(73)

Considering the Taylor series for cosine, we easily get that

\[ 2^{357}20!c_k \in [A_k, B_k]. \]  

(74)

To verify that \( c_k \in I_k \) it suffices to check that

\[ 2^{357}20!(C_k - 4) < A_k; \quad B_k < 2^{357}20!(C_k + 4). \]

This is purely a calculation involving BigIntegers. We perform the verification and it works. As a control, we performed the verification using “2” in place of “4” and it failed at some point. The program is contained in the same file as already mentioned. The reader can launch the program right from the 100 Degree window in McBilliards.

**Remark:** We found that \( 2^{357}20! \) worked well for us. This choice yields the following values

- \( A_8 = 193117979382323170336391434868704 \);
- \( A_9 = 1416254196461936667 \);
- \( A_{10} = 8363 \);
- \( A_{11} = 0 \).

This, the choice \( 2^{357}20! \) is well adapted to an approximation based on about 10 terms of the Taylor series.

### 10.3 BigInterval Structures

As one last bit of structure, we define a *BigComplexInterval* to be a structure of the form \( X + iY \) where \( X \) and \( Y \) are BigIntervals. The arithmetic on these objects is just the same as the arithmetic on ordinary complex numbers, except that we substitute the BigInterval operations for the ordinary arithmetic operations on reals. (We never have occasion to do any division, so we are just talking about addition, subtraction, and multiplication.)
Once we have our BigInterval version of sine and cosine, and the BigComplexInterval class, we plug these objects into Equations 23 and 31, wrapping every integer in sight inside a BigInterval. We then perform all the operations described in §5. Our algorithm halts for all 221 polygons and this constitutes our proof of the 100 Degree Theorem.

The reader can run our algorithm and survey its output using McBilliards, as discussed in the paper. In particular, the reader can run the algorithm with or without the BigInterval arithmetic, and see that the output is about the same in both cases. (The output is not exactly the same because we make some convenient but arbitrary cutoffs in the numerical version.)

10.4 Sanity Checks

In order to help insure that we have programmed the computer correctly, we have made 3 additional sanity checks in our calculations.

1. We make sure that our combinatorial method of computing the defining functions, namely Equation 23, is correct. We introduce a straightforward geometric method of computing the defining functions geometrically: We just take the unfolding for the word and the given triangle, rotate it so that it is horizontal, and then measure the difference in heights of the relevant vertices. For each word \( W_i \) we evaluate each defining function on the first vertex of the polygon \( P_i \), using both methods. As long as the geometric method yields a number which is at least .001 we check, up to a tolerance of .000001, that there is a single ratio \( \rho \) such that the ratio of the combinatorial answer to the geometric answer is always \( \rho \). (This ratio depends on the point of evaluation.) In other words, up to a initial rescaling, the two methods agree. We consider this to be extremely strong evidence that we have got Equation 23 correct, and also programmed it correctly into the computer. We do not consider the very small percentage of defining functions which evaluate to a very small number, because the roundoff error interferes with the computation of the ratio.

2. We make sure that our BigInterval versions of our functions yield essentially the same answers as our numerical versions. We make the same evaluations as for the first sanity check, except now we compare the numerical and BigInterval implementations of the combinatorial
method. We check that the first 7 digits of the left endpoint of the Big-Interval version agree with the first 7 digits of $2^{106}$ times the numerical version. In the interest of having the check move along at a steady clip when run from the interface, we only check about 4 percent of the defining functions. This still comes out to a huge number of checks. Unlike the first check, where the point is to verify that all cases of a complicated combinatorial procedure work, here we are just checking a fairly straightforward conversion from ordinary arithmetic operations to BigInterval operations.

3. We make sure that our formula for Equation 31 is correctly implemented. For this purpose we compare the partial derivatives of the defining functions with a crude version of the partial derivatives obtained by taking a difference quotient. Our value of $\Delta x$ and $\Delta y$ in this computation is $2^{-30}$. We check that the two computations of the partial derivatives agree up to a fractional error of .001. By this we mean that $|X_1 - X_2|/|X_1| < .001$. Here $X_1$ and $X_2$ are the two computed versions of the same quantity. We also require $X_1$, which is the difference quotient, to be at least .000001. We test about 1 percent of the defining functions. Given the simple nature of the passage from Equation 23 to Equation 31, this is overwhelming evidence that we have programmed Equation 31 correctly into the computer.

The reader can run our sanity checks, either for individual words or else for all words in sequence, from the 100 Degree window in McBilliards. The code for our sanity checks is contained in the file Deg100 SanityCheck.java. Indeed, all our computer code pertaining to the 100 Degree Theorem can be launched from this window.

We also mention another sanity check. Originally we had programmed McBilliards in C and Tcl. We originally did all the computations for this paper in the C version. (We switched to Java so that the whole proof could be easily accessible right on the web, to someone without specialized computer knowledge; and also because we wanted to make a new and improved McBilliards.)

Perhaps the best sanity check of all is that McBilliards works. This program has many interlocking features, and the interested reader can see that they all fit together in a way which would be extremely unlikely given serious bugs in the program.
11 Appendix

11.1 The First 6 Regions

1.

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

2.

<table>
<thead>
<tr>
<th>7</th>
<th>6</th>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

3.

<table>
<thead>
<tr>
<th>7</th>
<th>6</th>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>9</td>
<td>55</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>455</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

4.

<table>
<thead>
<tr>
<th>7</th>
<th>6</th>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>6</td>
<td>5</td>
<td>7</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>12</td>
<td>1641</td>
<td>12</td>
<td>1637</td>
<td>12</td>
<td>1637</td>
</tr>
<tr>
<td>12</td>
<td>2455</td>
<td>12</td>
<td>2455</td>
<td>12</td>
<td>2459</td>
</tr>
</tbody>
</table>

5.

<table>
<thead>
<tr>
<th>10</th>
<th>345</th>
<th>12</th>
<th>1380</th>
<th>12</th>
<th>1352</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>679</td>
<td>12</td>
<td>2712</td>
<td>12</td>
<td>2740</td>
</tr>
</tbody>
</table>

6.

<table>
<thead>
<tr>
<th>7</th>
<th>6</th>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>70</td>
<td>10</td>
<td>325</td>
<td>9</td>
<td>327</td>
<td>10</td>
</tr>
<tr>
<td>7</td>
<td>61</td>
<td>95</td>
<td>83</td>
<td>355</td>
<td>71</td>
</tr>
</tbody>
</table>

11.2 Tiles Abutting the Right-Angle Line

7. racetza

12 1069 | 10 439 | 11 867 | 14 6559 | 9 199 | 14 6437 |
| 11 1027 | 12 2225| 11 1181| 14 9825 | 9 281 | 14 8781 |

8. mazylodo

14 5711 | 12 1427| 12 1426| 9 163  | 11 504 | 11 452 |
| 11 566  | 13 2734|

9. mazepiton

15 1621 | 8   127| 7   59 | 16 29829| 15 14441| 10 427 |
| 15 13401|

10. ranwstrono

15 11789| 14 6081| 8   95 | 8   91 | 14 5779|
| 15 19379|

11. ranwztno

10 325  | 10 323 | 9   157| 12 1249|
| 9   327 | 10 701 | 9   355 | 7   83 |

12. noorotrowzn

| 9   229 | 13 3539| 14 6771| 14 6721| 12 1803|
| 9   283 | 13 4653| 13 4565| 13 4473| 14 8943|

13. bdtckbkmwach

| 10 411 | 11 839 | 13 3276| 10 391 | 10 385 |
| 11 1191| 13 4916| 10 633 | 10 627 |

14. ranodtwxono

| 11 725  | 12 1467| 10 365 | 15 11646| 10 363 | 12 1421 |
| 11 1200 | 12 2487| 11 1305| 15 21040| 10 661 | 12 2675 |
| 10 605 |

15. mizexprowdce

| 11 737  | 11 725 | 8   85 | 7 37  | 10 289 | 11 727 |
| 8   159  | 11 1293| 8   171| 7 91  | 9   359| 11 1279|

16. bocewawlmzhk

| 8   123  | 12 1781| 6   27 | 14 7350| 7   61 |
| 8   133  | 12 2315| 4   9  | 14 8667| 8   133|

17. razztltwppdno

| 10 387  | 10 393 | 12 1545| 12 1531| 11 759 |
| 10 637  | 10 631 | 10 613 | 12 2437| 12 2435|

18. mzhckpkmwhkzld

| 8   89  | 12 1433| 12 1399| 10 341 | 10 333 | 10 323 | 11 669 |
| 12 2625 | 11 1317| 12 2683| 10 683 | 10 691 | 10 691 | 11 1357 |

19. khowtexphtlcwzekp

| 10 373  | 10 389 | 11 793 | 11 791 | 11 739 |
| 10 651  | 10 635 | 8   155 | 11 1237| 9   325 |

20. rnoerzttwxcwn

| 12 1395 | 12 1401| 10 351 | 13 2799|
| 12 2701 | 12 2695| 11 5371| 13 5373|

21. norlewznernodttwezcn

| 10 431  | 11 877 | 11 879 | 10 429 |
| 10 594  | 11 1171| 11 1153| 9   295 |

22. htmwtdvdrkgkkordtwzewkty

| 11 723  | 12 1481| 11 735 | 11 723 |
| 12 2641 | 12 2601| 11 1313| 11 1325 |
References


[H] W.P. Hooper, Periodic Billiard Paths in Right Triangles are Unstable, Geometriae Dedicata (2006) to appear


