

# Asymptotics of Polygonal Dual Billiards I

Christopher J. Culter

July 7, 2005

## 1 Introduction

Let  $P$  be the interior of a convex polygon in  $\mathbb{R}^2$  with  $n \geq 3$  sides. We study a dynamical system  $T$  on  $\mathbb{R}^2$  called the *dual billiard* about  $P$ . Given a point  $z$  outside  $P$ , the dual billiard reflects  $z$  across the rightmost vertex of  $P$  as viewed from  $z$ . We prove that  $T$  always has periodic orbits with arbitrarily high even periods, and that when aperiodic orbits exist they are unstable. We develop a theory of *paths* about  $P$  which allows us to prove our new results and to rederive the literature on quasirational polygons in a unified manner. This paper is self-contained.

## 2 Rays

There are  $2n$  rays that extend the edges of  $P$ . Label these rays  $(\alpha_i)_1^{2n}$  in a *clockwise increasing order*; that is, the order in which they are encountered as one traces a sufficiently large circle in the clockwise direction. See Figure 1. By the end of this section we will be able to describe the ordering of  $(\alpha_i)$  intrinsically.

We access the orientation of the plane through the *cross product*

$$z \times w = \Im z^* w$$

for  $z, w \in \mathbb{R}^2$ . Given a point  $z$  and a ray  $R$  with basepoint  $r$  and direction  $dr$ , we say  $z$  is *to the left of  $R$*  if  $z \times dr < r \times dr$  and write  $z \prec R$ . Likewise  $z$  is *to the right of  $R$* ,  $z \succ R$ , iff  $z \times dr > r \times dr$ . The definitions are motivated by the case when  $R$  points up along the imaginary axis; then a point is to the left of  $R$  iff it lies in the left half-plane.

For a set  $Z \subset \mathbb{R}^2$ , we say  $Z \prec R$  if  $z \prec R$  for all  $z \in Z$ , and  $Z \succ R$  iff  $z \succ R$  for all  $z \in Z$ . We are now equipped to describe the direction in which rays wind around  $P$ : a ray  $R$  is called *positive* if  $P \prec R$  and *negative* if  $P \succ R$ . If we have the weaker condition that  $(Z \setminus B) \prec R$  for some bounded set  $B$ , we write  $Z \lesssim R$  and say that  $Z$  is *eventually to the left of  $R$* ; likewise for  $\gtrsim$ . Note that for two rays  $R$  and  $S$ , either  $R \lesssim S$ ,  $R \gtrsim S$ , or  $R$  and  $S$  are colinear. The conditions  $R \lesssim S$  and  $S \gtrsim R$  are essentially the same: they coincide unless  $R$  and  $S$  are antiparallel.

We now wish to state a theorem like

**Theorem? 2.1.** *For all  $i$ ,  $\alpha_{i+1} \gtrsim \alpha_i$  and  $\alpha_i \lesssim \alpha_{i+1}$ .*

This statement requires some adjustments to our notation. To insure that “ $\alpha_{i+1}$ ” makes sense when  $i = 2n$ , we must let  $\alpha$  take indices in  $\mathbb{Z}/2n$ . In fact, for the remainder of the paper, we reserve the letter  $i$  for elements of  $\mathbb{Z}/2n$ . By contrast, we reserve the letter  $j$  for elements of  $\mathbb{Z}$ .

There is a more subtle problem with (2.1). For an example, if we substitute  $i = 0$  then we get  $\alpha_1 \gtrsim \alpha_0$ , which is true enough. But the truth of this statement is not a property of the index 0 any more than it is a property of the index 1. We should in fact regard it as a property of the index  $1/2$ ! Let us develop this idea.

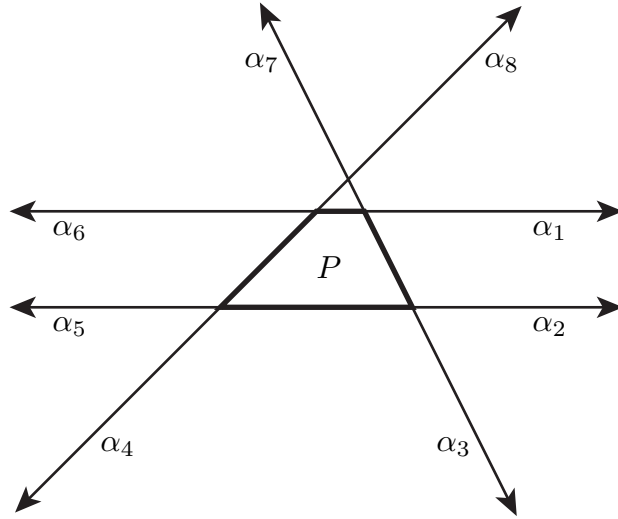


Figure 1: The rays  $\alpha_i$  for a generic trapezoid.

Define the half-step symbols  $\sharp$  and  $\flat$  to act on any number  $t$ :

$$t\sharp = t + 1/2, \quad t\flat = t - 1/2.$$

Let us reserve the letter  $\mu$  for half-integers modulo  $2n$ , that is, elements of the set  $\mathbb{Z}\sharp/2n$ . We reserve the letter  $\nu$  for regular half-integers, elements of  $\mathbb{Z}\sharp$ . Now for any  $\mu$  we may speak of the rays  $\alpha_{\mu\flat}$  and  $\alpha_{\mu\sharp}$ . The correct form of (2.1) is now:

**Theorem 2.2.** *For all  $\mu$ ,  $\alpha_{\mu\sharp} \succ \alpha_{\mu\flat}$  and  $\alpha_{\mu\flat} \prec \alpha_{\mu\sharp}$ .*

**Theorem 2.3.** *For all  $\mu$ , there is no  $i$  such that  $\alpha_i \succ \alpha_{\mu\flat}$  and  $\alpha_i \prec \alpha_{\mu\sharp}$ .*

The above theorems are in fact equivalent to demanding a clockwise increasing order. Together with the choice of  $\alpha_0$ , they fix all  $\alpha_i$ .

Our use of half-integers and the  $\sharp, \flat$  symbols is unorthodox but completely natural. So far our only  $2n$ -tuple of geometric objects is  $(\alpha_i)$ . As we define more objects, some will demand to be indexed by  $i$  and others by  $\mu$ ; then the half-step notation will become indispensable.

### 3 The Dual Billiard Map

Let us define the dual billiard map  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ . We leave  $T$  undefined on  $P$ . If  $z \notin P$ , there is a unique positive ray with basepoint  $z$  which is tangent to  $P$ . Let  $v^+(z) \in \partial P$  be the unique point of tangency (if it exists) and let  $Tz$  be the reflection of  $z$  through  $v^+(z)$ :

$$Tz = 2v^+(z) - z.$$

Note that there is a unique point of tangency, in fact a vertex of  $P$ , iff  $z$  is contained in none of the negative  $\alpha_i$ . Thus we leave  $T$  undefined on the negative  $\alpha_i$ . See Figure 2. We define  $v^-(z)$  and the inverse dual billiard map  $T^{-1}$  in the same way, exchanging “positive” with “negative”. Clearly  $T$  and  $T^{-1}$  are inverses on their domains. Let us describe those domains more carefully.

The set  $C = \mathbb{R}^2 \setminus (\alpha_1 \cup \dots \cup \alpha_{2n})$  is exactly the common domain of  $T$  and  $T^{-1}$ , and it has  $2n$  unbounded components. Let  $\alpha_\mu$  be the unbounded component of  $C$  which is  $\succ \alpha_{\mu\flat}$  and  $\prec \alpha_{\mu\sharp}$ . By construction,  $T|_{\alpha_\mu}$  and  $T^{-1}|_{\alpha_\mu}$  are the reflections through some vertices of  $P$ ; call them  $v_\mu^+$  and  $v_\mu^-$  respectively. The definitions of the families  $(v_\mu^+)$  and  $(v_\mu^-)$  are as close as we will come to systematically labelling the vertices of  $P$ :

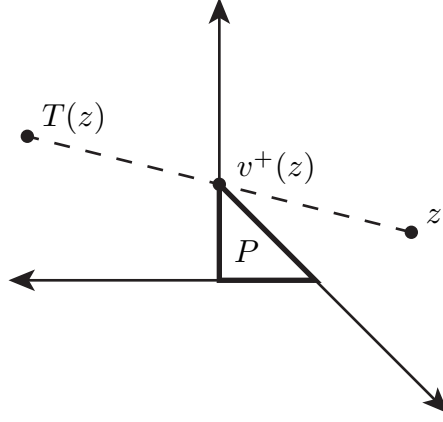


Figure 2: The dual billiard  $T$  about a triangle  $P$ . The points  $v^+(z)$  and  $T(z)$  are shown for a generic  $z$ . If  $z$  were on any of the negative  $\alpha_i$ , then  $T(z)$  is undefined.

**Lemma 3.1.** *For any vertex  $v$  of  $P$ , there is some  $\mu$  such that  $v_\mu^+ = v$ .*

Let us write  $\sigma(i) = +$  if  $\alpha_i$  is positive and  $\sigma(i) = -$  if  $\alpha_i$  is negative.

**Theorem 3.2.** *For all  $i$  we have*

$$v_{ib}^{\sigma(i)} = v_{i\sharp}^{\sigma(i)}$$

and

$$v_{ib}^{-\sigma(i)} \neq v_{i\sharp}^{-\sigma(i)}.$$

The vertex  $v_{ib}^{\sigma(i)} = v_{i\sharp}^{\sigma(i)}$  is maximally distant from the line containing  $\alpha_i$ . The vertices  $v_{ib}^{-\sigma(i)}$  and  $v_{i\sharp}^{-\sigma(i)}$  both lie on  $\alpha_i$ .

*Proof.* PROOF In light of this result, we define  $A_i = v_{ib}^{\sigma(i)} = v_{i\sharp}^{\sigma(i)}$ . □

Define the *shift vector*

$$h_\mu = 2(v_\mu^+ - v_\mu^-);$$

so for all  $z \in \alpha_\mu$ ,  $Tz - T^{-1}z = h_\mu$ . Then for  $z \in T^{-1}\alpha_\mu$  we have

$$T^2z = z + h_\mu.$$

Conversely, if  $T^2z$  is defined and  $z$  is far from  $P$  then  $z$  must lie in  $T^{-1}\alpha_\mu$  for some  $\mu$ . Thus  $T^2$  is locally a translation by one of the vectors  $h_\mu$ .

## 4 Parallel Sides

By convexity, every side of  $P$  is parallel to at most one other side. Let  $n_{\parallel}$  be the number of pairs of parallel sides of  $P$ , and let  $n_{\perp} = n - n_{\parallel}$ . Then we have

$$0 \leq n_{\parallel} \leq n/2 \leq n_{\perp} \leq n.$$

Let  $\theta_i$  be the unit vector which points *opposite* from the ray  $\alpha_i$ ; this convention is used in the next section. If  $\theta_{\mu b} = \theta_{\mu\sharp}$  then we say that the index  $\mu$  is *small* and that both  $\mu b$  and  $\mu\sharp$  are *paired*. If  $\mu$  is not small it is called *large* and if  $i$  is not paired it is called *single*. If  $\mu$  is large (small) then  $\alpha_\mu$  is a generalized cone (halfstrip).

**Lemma 4.1.** For all  $i$ , either  $ib$  or  $i\sharp$  is large.

*Proof.* Otherwise,  $\theta_{i-1} = \theta_i = \theta_{i+1}$ , which violates convexity.  $\square$

We now state a series of easy results which are trivial for  $n = 3$  and thereafter follow by induction:

**Lemma 4.2.** For all  $i$ ,  $\theta_i = -\theta_{i+n}$ .

**Lemma 4.3.** The rays  $\alpha_i, \alpha_{i+n}$  are collinear iff  $i$  is single.

**Lemma 4.4.**  $\mu$  is large iff  $\mu + n$  is large, and  $i$  is single iff  $i + n$  is single.

**Theorem 4.5.** If  $T\alpha_\mu \cap T^{-1}\alpha_t$  is unbounded, then  $\mu \leq t \leq \mu + 2$ . Furthermore, if  $\mu + 1$  is large then  $\mu \leq t \leq \mu + 1$ .

*Proof.* PROVE  $\square$

## 5 Perturbations

We now consider perturbations of  $P$  in the manifold  $\mathbb{P}_n$  of all unlabelled convex  $n$ -gons. Given any  $P' \in \mathbb{P}_n$ , we write  $T'$  for the dual billiard about  $P'$ ; this  $T'$  is considered to be a perturbed version of  $T$ . We similarly define  $T^{-'}$ ,  $v^{+'}$  and  $v^{-'}$ . Now, we cannot define primed versions of everything in sight; for example, we do not attempt to find perturbed rays  $\alpha'_i$  or perturbed regions  $\alpha'_\mu$ . However, we can certainly define perturbed vertices when the perturbation is small.

Let  $P \in U \subset \mathbb{P}_n$  where  $U$  is simply path-connected. To each vertex  $v$  of  $P$  and each perturbation  $P' \in U$  we then assign the corresponding vertex  $v'$  of  $P'$  by following  $v$  along a path in  $U$ . If  $O$  is some function of the vertices of  $P$ , we thus unambiguously define the perturbed object  $O'$ . If  $O_1, O_2$  are objects such that  $O'_1 = O'_2$  for all  $P' \in U$  in some open set  $U$ , then we write  $O_1 \equiv O_2$  and say that the equality holds *identically*. An equality which does not hold identically is called *accidental*. Now, a fundamental example:

**Lemma 5.1.** Let  $(v_1, \dots, v_n)$  be a labelling of the vertices of  $P$  and fix  $c_1, \dots, c_n \in \mathbb{R}$ . Then  $\sum c_j v_j \equiv 0$  iff  $c_j = 0$  for all  $j$ .

*Proof.* The functions  $v_1, \dots, v_n$  constitute local coordinates  $\mathbb{P}_n \rightarrow \mathbb{R}^{2n}$  around  $P$ .  $\square$

So, if a vector is defined to be some linear combination of the vertices of  $P$ , then it is identically zero iff it is zero as a *formal* linear combination.

We now describe how the action of the dual billiard is stable under small perturbations:

**Theorem 5.2.** Let  $v$  be a vertex of  $P$ . If  $v^+(z) = v$ , then  $Tz \equiv 2v - z$ .

*Proof.* We need a  $U$  such that for all  $P' \in U$ ,  $v^{+'}(z) = v'$ . Let  $S_z \subset \mathbb{P}_n$  be the open set of polygons  $P'$  such that no negative ray extending a side of  $P'$  contains  $z$ . Then let  $U$  be a simply path-connected subset of  $S_z$  containing  $P$ . For all  $P' \in U$  we have  $T'z = 2v^{+'}(z) - z = 2v' - z$  as desired.  $\square$

**Corollary 5.3.** For all  $z$ ,  $Tz - T'z \equiv 2[v^+(z) - v^-(z)]$ .

**Corollary 5.4.** For all  $\mu$  and  $z \in T^{-1}\alpha_\mu$ ,  $T^2z \equiv z + h_\mu$ .

## 6 The Strip Map

To find  $T^{-2}z$  and  $T^2z$  we need only know the regions  $T\alpha_s$  and  $T^{-1}\alpha_t$  to which  $z$  belongs. The case  $s = t$  is boring. Interesting dynamics occur when the index changes, crossing some  $i$ ; that is,  $s < i < t$ . We thus define the generalized halfstrip

$$\gamma_i = \bigsqcup_{s < i < t} T\alpha_s \cap T^{-1}\alpha_t.$$

Far from  $P$ , we can supplement the algebraic definition of  $\gamma_i$  with a geometric description:

**Lemma 6.1.** *Modulo a bounded set,*

$$\partial\gamma_i = T\alpha_i \cup T^{-1}\alpha_i.$$

*The central axis of  $\gamma_i$  is tangent to  $P$ .*

**Corollary 6.2.** *If  $P \prec \alpha_i$  then  $\gamma_i \prec \alpha_i$ ; if  $P \succ \alpha_i$  then  $\gamma_i \succ \alpha_i$ .*

**Lemma 6.3.** *If  $i$  is single, then far from  $P$  we have*

$$\begin{aligned} \gamma_i &\cong T\gamma_{ib} \cap T^{-1}\gamma_{i\sharp} \\ &\cong (T\gamma_{ib} \setminus T^{-1}\gamma_{ib})^\circ \\ &\cong (T^{-1}\gamma_{i\sharp} \setminus T\gamma_{i\sharp})^\circ, \end{aligned}$$

where  $^\circ$  is the interior and  $\cong$  denotes equality outside some bounded set.

Note that  $\gamma_i$  extends in the opposite direction from  $\alpha_i$ , and thus  $\theta_i$  points along  $\gamma_i$ . Like the rays  $\alpha_i$ , the halfstrips  $\gamma_i$  are also ordered clockwise; this is a subtle point when  $P$  has parallel sides.

Define a strip map  $S_\mu : \gamma_{\mu b} \rightarrow \gamma_{\mu\sharp}$  by  $z \rightarrow T^{2k}z$  where  $k$  is the least nonnegative integer, if it exists, such that  $T^{2k}z \in \gamma_{\mu\sharp}$ . If there is no such  $k$  we leave  $S_\mu z$  undefined. We can characterize the strip map in a completely geometric way:

**Lemma 6.4.** *For all  $z \in \gamma_{\mu b}$ ,  $S_\mu z = z + kh_\mu$  where  $k$  is the least nonnegative integer, if it exists, such that  $z + kh_\mu \in \gamma_{\mu\sharp}$ . If there is no such  $k$  then  $S_\mu z$  is undefined.*

Let  $\tilde{\gamma}_i$  be the open strip which extends  $\gamma_i$ . Geometrically,  $\tilde{\gamma}_i$  is uniquely described by the conditions that  $\alpha_i \subset \partial\tilde{\gamma}_i$  and that the central axis of  $\tilde{\gamma}_i$  is tangent to  $P$ ; roughly speaking,  $\tilde{\gamma}_i$  is twice as wide as  $P$ .

For  $z \in \tilde{\gamma}_{\mu b}$  define  $\tilde{S}_\mu z$  to be the unique point in  $\tilde{\gamma}_{\mu\sharp}$  having the form  $z + kh_\mu$ .

**Lemma 6.5.** *If  $z \in \gamma_{\mu b}$  then  $\tilde{S}_\mu z = S_\mu z$ .*

*Proof.* This follows from the definition of  $\tilde{S}_\mu$  and Lemma ???. □

Let  $w_i$  be the width of  $\tilde{\gamma}_i$ . Let  $\Pi_\mu = \tilde{\gamma}_{\mu b} \cap \tilde{\gamma}_{\mu\sharp}$ . Let  $p_i^+ = |\Pi_{i\sharp}|/w_i$  and  $p_i^- = |\Pi_{ib}|/w_i$ . We can interpret  $p_i^+$  as the length of a side of  $\Pi_{i\sharp}$  parallel to  $\tilde{\gamma}_i$ . If  $\mu$  is small, then  $|\Pi_\mu| = p_{\mu b}^+ = p_{\mu\sharp}^- = +\infty$ .

**Lemma 6.6.** *For all  $\mu$ , we have the vector identity*

$$\frac{\theta_{\mu b}}{w_{\mu b}} - \frac{\theta_{\mu\sharp}}{w_{\mu\sharp}} = \frac{h_\mu}{|\Pi_\mu|}.$$

*Proof.* Assume  $\mu$  is small. Then  $\theta_{\mu\sharp} = \theta_{\mu b}$  and  $w_{\mu\sharp} = w_{\mu b}$ , so the left side is zero. Also,  $|\Pi_\mu| = \infty$ , so the right side is also zero.

Now assume that  $\mu$  is large, so that  $\Pi_\mu$  is a parallelogram. We have

$$p_{\mu b}^+ \theta_{\mu b} - p_{\mu\sharp}^- \theta_{\mu\sharp} = h_\mu$$

and dividing through by  $|\Pi_\mu|$  we get the desired result. □

## 7 The First Return Map

We may consider  $\gamma_0$  to be a “fat surface of section” for the discrete mapping  $T^2$ ; the corresponding Poincaré map we will call the first return map  $F_0$ . Concretely, let  $F$  be the composition of  $2n$  strip maps:

$$F = S_{0b} \circ \cdots \circ S_{0\sharp} : \gamma_0 \rightarrow \gamma_0.$$

We will actually be more interested in an extended first return map: let

$$\tilde{F} = \tilde{S}_{0b} \circ \cdots \circ \tilde{S}_{0\sharp} : \tilde{\gamma}_0 \rightarrow \tilde{\gamma}_0.$$

The extended map has several advantages: it is defined almost everywhere, it is easier to describe, and its symmetries are much more apparent. For these reasons we will tailor the theory of *paths*, which are dual to the first return map, to  $\tilde{F}$  rather than  $F$ . Of course, whenever  $z \in \gamma_0$  we have  $\tilde{F}z = Fz$  anyway.

Before defining paths we take a look at some of the phenomena we wish to describe: FIGURES FIGURES FIGURES!

## 8 Paths

Define a *path* to be a list  $x$  of vectors  $x_0, \dots, x_{2n}$  such that  $x_{\nu\sharp} - x_{\nu b}$  is parallel to  $h_\nu$  for all  $0 < \nu < 2n$ . We place no restriction yet on  $x_0 - x_{2n}$ .

The set of all paths  $\mathcal{P}$  is a vector subspace of  $\mathbb{R}^{2(2n+1)}$ . For each  $0 < \nu < 2n$  define a linear functional  $k_\nu : \mathcal{P} \rightarrow \mathbb{R}$  by

$$x_\nu := k_\nu(x) := \frac{x_{\nu\sharp} - x_{\nu b}}{h_\nu}.$$

For all  $\mu$  there is a unique  $0 < \nu < 2n$  such that  $\mu = \nu + 2n\mathbb{Z}$ , so we may write  $k_\mu := k_\nu$  without causing any confusion.

The values of  $x_\mu$  determine  $x$  only up to translation. Thus we define the *center of mass* linear map  $\langle x \rangle : \mathbb{R}^{2(2n+1)} \rightarrow \mathbb{R}^2$  by

$$\langle x \rangle = \frac{1}{2n} \left( \frac{x_0}{2} + x_1 + \cdots + x_{2n-1} + \frac{x_{2n}}{2} \right).$$

We call a path *constant* if  $x_j = \langle x \rangle$  for all  $j$  and thus identify the space of constant paths with  $\mathbb{R}^2$ . A path is *centered* if  $\langle x \rangle = 0$ . Rather than spend a symbol on the space of centered paths, we identify it with  $\mathcal{P}/\mathbb{R}^2$ .

Now, for each  $\mu$  we define a path  $k^\mu$ , dual to the functional  $k_\mu$ , by the conditions  $\langle k^\mu \rangle = 0$  and  $k_{\mu'}^\mu = \delta_{\mu'}^\mu$  where  $\delta$  is the Kronecker delta. It is clear that

$$\mathcal{P} = \mathbb{R}^2 \oplus \bigoplus_{\mu} \mathbb{R}k^\mu$$

and the associated decomposition is

$$x = \langle x \rangle + \sum x_\mu k^\mu.$$

A path  $x$  is *integral* if all  $x_\mu \in \mathbb{Z}$ . The set  $\mathcal{Z}$  of all integral paths is neither a vector space nor a lattice but a sum of the two:

$$\mathcal{Z} = \mathbb{R}^2 \oplus \bigoplus_{\mu} \mathbb{Z}k^\mu.$$

For any set  $S \subset \mathcal{P}$ , we say that  $S$  is *spanned by integral paths* if  $S \subset \text{span}(\mathcal{Z} \cap S)$ . Every set that we have named so far has been spanned by integral paths; this will change in the next section.

We need one more definition to relate paths to the strip maps. A path  $x$  is *valid* if all  $x_j \in \tilde{\gamma}_j$ , and  $x$  is *strongly valid* if all  $x_j \in \gamma_j$ . Define the Cartesian products of strips

$$\tilde{\Gamma} = \prod_{j=0}^{2n} \tilde{\gamma}_j, \quad \Gamma = \prod_{j=0}^{2n} \gamma_j;$$

then  $x$  is valid iff  $x \in \tilde{\Gamma}$  and strongly valid iff  $x \in \Gamma$ . We are motivated by the following theorems:

**Theorem 8.1.** *If  $z \in \text{dom } \tilde{F}$  then  $(z, \tilde{S}_{0\sharp}z, \dots, \tilde{F}z)$  is a valid integral path. Conversely, if  $x$  is a valid integral path then  $x_{2n} = \tilde{F}x_0$ .*

**Theorem 8.2.** *If  $z \in \text{dom } F$  then  $(z, S_{0\sharp}z, \dots, Fz)$  is a strongly valid integral path. Conversely, if  $x$  is a strongly valid integral path then  $x_{2n} = Fx_0$ .*

For a path  $x$  let its *trace*  $[x]$  be the path traced out by  $x$ ; explicitly,

$$[x] = \bigcup_{0 < \nu < 2n} (x_{\nu b} + [0, x_{\nu}]h_{\nu}).$$

If  $x$  is integral then it makes sense to consider its *integral trace*

$$[x]_{\mathbb{Z}} = \bigcup_{0 < \nu < 2n} (x_{\nu b} + \{0, \dots, x_{\nu}\}h_{\nu}).$$

## 9 Closed Paths

A path  $x$  is *closed* if  $x_{2n} = x_0$ ; equivalently, if  $\sum x_{\mu}h_{\mu} = 0$ . Note that the vertices of a closed path may safely be indexed with  $i$  instead of  $j$ . The vector space of closed paths is  $2n$ -dimensional, which is too large to be very useful.

More restrictively, a path  $x$  is *identically closed* if  $\sum x_{\mu}h_{\mu} \equiv 0$ . Denote the space of identically closed paths by  $\mathcal{C}$ . We are motivated by the following:

**Theorem 9.1.** *A path  $x \in \mathcal{Z} \cap \tilde{\Gamma}$  is (identically) closed iff  $x_0$  is an (identically) fixed point of  $\tilde{F}$ .*

*Proof.* This follows directly from (). □

**Corollary 9.2.** *A path  $x \in \mathcal{Z} \cap \Gamma$  is (identically) closed iff  $x_0$  is an (identically) fixed point of  $F$ .*

**Corollary 9.3.** *Every (identically) closed, strongly valid integral path traces out an (identically) periodic orbit of  $T^2$ .*

Closed paths are useful not only for identifying fixed points of  $\tilde{F}$  but also tying orbits together:

**Theorem 9.4.** *Let  $x, y \in \mathcal{Z} \cap \tilde{\Gamma}$  and suppose  $x - y$  is (identically) closed. Then the following holds (identically):*

$$\tilde{F}x - x = \tilde{F}y - y.$$

Clearly all constant paths are identically closed. Fortunately,  $\mathcal{C}$  is much larger than  $\mathbb{R}^2$ . We will identify its elements in sections 11 and 13; for now we simply determine its dimension.

**Lemma 9.5.**  $\dim \mathcal{C} = n + 3$ .

*Proof.* Construct a digraph  $D$  whose edges are  $k^{\mu}$  directed from  $v_{\mu}^{-}$  to  $v_{\mu}^{+}$ . Its cycle space is exactly  $Z_{\mathbb{R}}(D) = \mathcal{C}/\mathbb{R}^2$ . We must show  $\dim Z_{\mathbb{R}}(D) = n + 1$ .

Given vertices  $w, w'$  adjacent on  $P$ , pick a ray  $\alpha_i$  that contains them. Then  $w, w'$  are connected to  $A_i$  by  $k^{ib}$  and  $k^{i\sharp}$ , so they belong to the same component of  $D$ . Thus  $D$  is connected, and  $\dim Z_{\mathbb{R}}(D) = e - v + k = 2n - n + 1$ . □

**Corollary 9.6.**  $\mathcal{C}$  is spanned by integral paths.

*Proof.* The cycle space has a basis in which each coefficient of each basis vector is 0 or  $\pm 1$ .  $\square$

On the other hand, the space of all closed paths is not necessarily spanned by integral paths. If it is, we say that  $P$  is *rational*. This path-theoretic definition coincides with the usual one:

**Lemma 9.7.** The vertices of  $P$  lie on a lattice iff  $P$  is rational.

*Proof.* AAGH!  $\square$

If  $x$  is closed then  $[x]$  is a polygon and therefore we will sometimes refer to  $x$  itself as a polygon. Furthermore, if  $[x]$  is a simple curve then we can identify its interior, denoted  $(x)$ .

## 10 Necklace Polygons

We can simplify the problem of finding valid paths with a map that halves the dimension and bounds  $\Gamma$ . Define the *reduction map*  $R : \mathbb{R}^{2(2n+1)} \rightarrow \mathbb{R}^{2n+1}$  by coordinates:

$$(Rx)_j = \theta_j \times x_j.$$

Now the set  $R\Gamma = R\tilde{\Gamma}$  is just a product of open intervals, and it is still useful:

**Lemma 10.1.** A path  $x$  is valid iff  $Rx \in R\Gamma$ .

*Proof.* Since  $\tilde{\gamma}_j$  is invariant under translations by  $\mathbb{R}\theta_j$ , we have  $x_j \in \tilde{\gamma}_j$  iff  $\theta_j \times x_j \in \theta_j \times \tilde{\gamma}_j$ .  $\square$

Since  $R$  takes  $\mathcal{P} \cap \Gamma$ , which is unbounded, into a bounded set,  $R$  must annihilate some paths. Let  $\mathcal{N} = \mathcal{P} \cap \ker R$ ; an element of  $\mathcal{N}$  is called a *necklace polygon*. Geometrically, a path  $x$  is a necklace polygon iff every  $x_j$  is a multiple of  $\theta_j$ . It is clear that  $\dim \mathcal{N} = 1$ , since there is one degree of freedom in choosing  $x_0$ , which uniquely determines  $x_1$  and so on. We thus fix a convenient generator of  $\mathcal{N}$ : define  $Q \in \mathcal{N}$  by

$$Q_i = C \frac{\theta_i}{w_i}$$

where

$$C^{-1} = \sum |\Pi_\mu|^{-1}.$$

That  $Q \in \mathcal{N}$  is manifest, and by Lemma 6.6 we have

$$Q_\mu = C |\Pi_\mu|^{-1}.$$

By our choice of  $C$  we now get  $\sum Q_\mu = 1$ .

**Theorem 10.2.** The set  $(\mathcal{P} \cap \Gamma)/\mathcal{N}$  is a bounded subset of the vector space  $\mathcal{P}/\mathcal{N}$ .

*Proof.* Since  $\mathcal{N} = \mathcal{P} \cap \ker R$ , the canonical isomorphism  $\mathcal{P}/\mathcal{N} \cong R\mathcal{P}$  identifies  $(\mathcal{P} \cap \Gamma)/\mathcal{N}$  with  $R(\mathcal{P} \cap \Gamma)$ , which is bounded.  $\square$

This theorem tells us that valid paths never stray far from necklace polygons. Concretely, we have

**Corollary 10.3.** There exists a bounded set of paths  $K$  such that  $(\mathcal{P} \cap \Gamma) \subset K + \mathcal{N}$ .

Now, the space of necklace polygons is typically not spanned by integral paths. If it is, we say that  $P$  is *quasirational*.

**Theorem 10.4.** The following are equivalent:

- $P$  is quasirational.
- The areas  $|\Pi_\mu|$ , for  $\mu$  large, are commensurate.
- The dimension of  $\mathbb{Q}[Q_\mu]$  over  $\mathbb{Q}$  is 1.
- There exists a nonzero integral necklace polygon.

*Proof.* Immediate from the definition of  $N$ . □

**Theorem 10.5.** *If  $P$  is rational then  $P$  is quasirational.*

*Proof.* There should be a path-theoretic proof of this fact, but it is more intuitive to use geometry. If  $P$  is rational then its vertices are rational in some coordinate system, and thus every quantity in sight, including  $|\Pi_\mu|$ , is rational. □

## 11 Balanced Paths

Since  $\mathcal{N}$  need not be spanned by integral paths, it is useful to consider an enlargement of  $\mathcal{N}$  which is. One clue is that  $Q$  is centrally symmetric, or *balanced*; we now expand on this concept.

A closed path  $x$  is *balanced* if all  $x_{\mu+n}h_{\mu+n} = -x_\mu h_\mu$ . We will have nothing to say about general balanced paths. More importantly, a path  $x$  is *identically balanced* if all  $x_{\mu+n}h_{\mu+n} \equiv -x_\mu h_\mu$ . Denote the space of identically balanced paths by  $\mathcal{B}$ .

It is clear that identically balanced paths are identically closed. However, they are easier to describe than their more general cousins:

**Lemma 11.1.** *A path  $x$  is identically balanced iff all  $x_\mu = x_{\mu+n}$  AND for all small  $\mu$ ,  $x_\mu = 0$ .*

*Proof.* Recall that for  $\mu$  large,  $h_{\mu+n} \equiv -h_\mu$ ; but for  $\mu$  small,  $h_\mu$  and  $h_{\mu+n}$  are linearly independent combinations of vertices; thus  $x_{\mu+n}h_{\mu+n} \equiv -x_\mu h_\mu \iff x_{\mu+n} = x_\mu = 0$ . □

**Corollary 11.2.** *For  $x \in \mathcal{B}$ ,*

$$x = \langle x \rangle + \sum_{0 < \mu \uparrow < n} x_\mu (k^\mu + k^{\mu+n})$$

This leads to the decompositions:

**Corollary 11.3.**

$$\begin{aligned} \mathcal{B} &= \mathbb{R}^2 \oplus \bigoplus_{0 < \mu \uparrow < n} \mathbb{R}(k^\mu + k^{\mu+n}) \\ \mathcal{Z} \cap \mathcal{B} &= \mathbb{R}^2 \oplus \bigoplus_{0 < \mu \uparrow < n} \mathbb{Z}(k^\mu + k^{\mu+n}) \\ \mathcal{P} &= \mathcal{B} \oplus \bigoplus_{\neg(0 < \mu \uparrow < n)} \mathbb{R}k^\mu \\ \mathcal{Z} &= \mathcal{Z} \cap \mathcal{B} \oplus \bigoplus_{\neg(0 < \mu \uparrow < n)} \mathbb{Z}k^\mu \end{aligned}$$

**Corollary 11.4.**  $\dim \mathcal{B} = n_\perp + 2$ .

**Corollary 11.5.**  $\mathcal{B}$  is spanned by integral paths.

**Corollary 11.6.**  $\mathcal{Z}/\mathcal{B}$  is discrete.

Finally, we state the result which motivated the study of  $\mathcal{B}$  in the first place:

**Theorem 11.7.** *Necklace polygons are identically balanced.*

*Proof.* It suffices to prove that  $Q \in \mathcal{B}$ . Recall that  $Q_\mu = C|\Pi_\mu|^{-1}$  and  $|\Pi_\mu| = |\Pi_{\mu+n}|$ ; thus  $Q_{\mu+n} = Q_\mu$ . Also, for  $\mu$  small we have  $|\Pi_\mu| = \infty$  and thus  $Q_\mu = 0$ .  $\square$

**Corollary 11.8.** *The set  $(\mathcal{P} \cap \Gamma)/\mathcal{B}$  is a bounded subset of the vector space  $\mathcal{P}/\mathcal{B}$ .*

*Proof.* Apply the canonical projection  $\mathcal{P}/\mathcal{N} \rightarrow \mathcal{P}/\mathcal{B}$  to  $(\mathcal{P} \cap \Gamma)/\mathcal{N}$ , which is bounded.  $\square$

We now arrive at a very useful result:

**Theorem 11.9.** *The set  $(\mathcal{Z} \cap \Gamma)/\mathcal{B}$  is finite.*

*Proof.* Since  $(\mathcal{P} \cap \Gamma)/\mathcal{B}$  is bounded and  $\mathcal{Z}/\mathcal{B}$  is discrete, their intersection is finite.  $\square$

**Corollary 11.10.** *The set  $(\mathcal{Z} \cap \Gamma)/\mathcal{C}$  is also finite.*

*Proof.* Apply the canonical projection  $\mathcal{P}/\mathcal{B} \rightarrow \mathcal{P}/\mathcal{C}$ .  $\square$

## 12 Periodicity

We begin with an easy result:

**Lemma 12.1.** *Let  $\mathbb{T}$  be a compact abelian group,  $b \in \mathbb{T}$ , and  $U$  a neighborhood of 0. Then infinitely many positive multiples of  $b$  lie in  $U$ .*

*Proof.* The forward limit set of the sequence  $b, 2b, 3b, \dots$  is nonempty because  $\mathbb{T}$  is compact, and it is closed under subtraction; thus a subgroup of  $\mathbb{T}$ . In particular, some subsequence converges to 0, and so it eventually lies in  $U$ .  $\square$

**Lemma 12.2.**

Recall that either  $0\flat$  or  $0\sharp$  must be large. For convenience, let us assume that  $0\sharp$  is large.

**Theorem 12.3.** *Let  $U$  be a tubular neighborhood of  $\mathcal{N}$  in  $\mathcal{P}$ ; that is,  $U = U + \mathcal{N}$ . Then  $U$  contains infinitely many identically balanced, centered, integral paths.*

*Proof.* This theorem is no more subtle than the fact that in  $\mathbb{R}^2$ , if a strip and a lattice both contain 0 then their intersection is infinite. We compactify the plane by annihilating the strip direction and rolling up all but one of the lattice vectors; this leaves one degree of freedom to rotate around the resulting torus.

We refer to the  $n_\perp$  paths  $k^\mu + k^{\mu+n}$  for  $\mu$  large and  $0 < \mu < n$  collectively as  $b$ . Then we have  $\mathcal{B}/\mathbb{R}^2 = \mathbb{R}[b]$  and  $(\mathcal{Z} \cap \mathcal{B})/\mathbb{R}^2 = \mathbb{Z}[b]$ . Let  $b_0 = k^{0\sharp} + k^{0\sharp+n}$ , and refer to the other  $n_\perp - 1$  paths collectively as  $b_1$ .

We show that for infinitely many  $x \in \mathbb{Z}[b]$ ,  $Rx \in RU$ , which implies  $x \in U$  since  $U$  is tubular. On  $\mathbb{R}[b]$ , the nullity of  $R$  is 1. Thus  $\mathbb{Z}[Rb]$  need not be a lattice at all, while  $\mathbb{Z}[Rb_1]$  is a full lattice in  $\mathbb{R}[Rb]$ . We show that for infinitely many natural numbers  $k$ , the displaced lattice  $kRb_0 + \mathbb{Z}[Rb_1]$  intersects  $RU$ . Let us define  $\mathbb{T}' = \mathbb{R}[Rb]/\mathbb{Z}[Rb_1]$ ,  $b' = (Rb_0)/\mathbb{Z}[Rb_1]$ , and  $U' = (RU)/\mathbb{Z}[Rb_1]$ .

The theorem reduces to this: given a compact abelian group  $\mathbb{T}'$ , an open neighborhood  $U'$  of 0 and some  $b' \in \mathbb{T}'$ , the sequence  $0, b', 2b', 3b', \dots$  visits  $U'$  infinitely many times. This is clear; its forward limit set is a subgroup of  $\mathbb{T}'$ , and it certainly contains 0.  $\square$

## 13 Axial Polygons

This section completes our taxonomy of paths. Define an *axial polygon* to be a path  $x$  such that each  $x_j$  is on the central axis of  $\tilde{\gamma}_j$ ; equivalently,  $Rx$  is in the center of the box  $R\Gamma$ . Let  $\mathcal{A}$  be the affine space of axial polygons. We already have the most important example: the points  $A_i$  defined in section 3 comprise a path  $A \in \mathcal{A}$ . Since  $A$  is not constant, there is some  $A_\mu \neq 0$ .

Since  $\mathcal{A}$  is the preimage of a point under  $R$ , it is a coset of  $\mathcal{N}$ ,

$$\mathcal{A} = A + \mathcal{N},$$

which tells us everything about  $\mathcal{A}$ . For example, both  $A$  and  $Q$  are manifestly identically closed, so  $\mathcal{A} \subset \mathcal{C}$ , justifying our use of the word ‘‘polygon’’. However,  $\mathcal{A} \not\subset \mathcal{B}$ :

**Lemma 13.1.** *The path  $A$  is not identically balanced.*

*Proof.* If  $P$  has parallel sides, choose a small  $\mu$  and note that  $A_\mu = 1/2 \neq 0$ . Otherwise,  $A_i = A_{i+n}$  for all  $i$  and thus if  $A$  were balanced it would be constant.  $\square$

The above proof does not preclude  $A$  from being accidentally balanced if  $P$  has enough parallel sides; indeed if  $P$  itself is balanced then so is  $A$ .

Since  $A \notin \mathcal{B}$ , we have expanded our census of identically closed paths. Recall that  $\dim \mathcal{C} = n+3$  and  $\dim \mathcal{B} = n_\perp + 2$ . By dimension counting, if  $P$  has no parallel sides then  $\mathcal{C} = \mathcal{B} \oplus \mathbb{R}A$ ; we have explicitly determined  $\mathcal{C}$ . Conversely, if  $n_\parallel > 0$  then there are identically closed paths which we have not described.

Note that  $A$  is half-integral in the sense that  $2A \in \mathcal{Z}$  even though  $A \notin \mathcal{Z}$ . Thus  $\mathcal{Z} \cap \Gamma$  is invariant under reflection through  $A$ , and consequently the first return map  $F$  commutes with the reflection through  $A_0$ . Indeed, every reflection symmetry of  $F$  is generated by an integral or half-integral axial polygon, just as every translation symmetry of  $F$  is generated by an integral necklace polygon. Unless  $P$  is quasirational there can be no more integral or half-integral axial polygons; thus we concentrate on  $A$ .

The reflection symmetry about  $A$  relates pairs of valid paths. To make statements about the dual billiard map  $T$  (remember it?) we should relate pairs of strongly valid paths. That is, suppose  $x$  is integral, closed, and strongly valid. Far from  $P$  we can conclude that  $2A - x$  is valid but *not* strongly valid. However, let us assume for the moment that  $n_\parallel = 0$  and define a new path  $tx$  to be a permutation of  $2A - x$ :

$$(tx)_i = (2A - x)_{i+n}.$$

The suggestively named operator  $t$  is involutive,  $t^2x = x$ , and far from  $P$ ,  $tx$  is strongly valid. Thus  $[x]_\mathbb{Z}$  and  $[tx]_\mathbb{Z}$  are both closed orbits of  $T^2$ ; they are in fact subsets of the same closed orbit of  $T$ .

When  $P$  has parallel sides, we set  $(tx)_i$  to be  $T^{\pm 1}x_{i+n}$ , whichever falls in  $\gamma_i$ ; this definition generalizes the original one. However, this  $t$  is not linear; it is discontinuous on the axis of  $\gamma_i$  for paired  $i$

The next natural question is whether  $t$  has fixed points. Indeed, if  $x$  is a closed, integral, strongly valid path, and  $tx = x$ , then  $[x]_\mathbb{Z}$  is a rare odd-periodic orbit of  $T$ . Let us again specialize to the case  $n_\parallel = 0$ . Then  $tx = x$  iff  $x - A \in \mathcal{B}/\mathbb{R}^2$ ; in particular, if  $x \in \mathcal{A}$ . Is this possible? The existence of integral axial polygons is not only motivated by the case  $n_\parallel = 0$  but dependent on it, since for small  $\mu$ , if  $x$  is axial then  $x_\mu = 1/2 \notin \mathbb{Z}$ .

It is easy to see that  $\mathcal{A} \cap \mathcal{Z}$  is nonempty iff  $\mathcal{A}$  is spanned by integral paths. These conditions then imply that  $P$  is quasirational,  $n_\parallel = 0$ , and there exist odd-periodic orbits of  $T$  arbitrarily far from  $P$ . It would be interesting to see if there are any partial converses.