# GEOMETRIC SEQUENCES OF DISCS IN THE APOLLONIAN PACKING

DOV AHARONOV AND KENNETH STEPHENSON

*Dedicated to G. M. Goluzin*

This study began innocently enough with a search for extremal configurations of circles in the Rodin and Sullivan "Ring Lemma". This is an elementary geometric lemma which nevertheless plays a key role in recent work on circle packing and its connections to conformal mapping. However, the study evolved into something of a grand tour of various ideas bearing on the geometry of "quads" of circles as we discovered (and rediscovered) intertwining results stretching from antiquity right up to current work on residual sets of Kleinian groups. Our aim here is to share not only our results, but also some of the sights and surprises we encountered on this journey.

Our earliest reference will be to a famous disc configuration of Apollonius of Perga; then we pick up a crucial theorem from Descartes, a little poetry from Frederick Soddy, a loxodromic spiral from Coxeter, and a little on the circle packings of Thurston — a topic that Apollonius himself would surely have enjoyed. Along the way we will see the "shoemaker's knife" and beloved golden ratio of the ancients, Fibonacci sequences, Farey numbers, Soddy's hexlet and bowl of integers, Kleinian limit sets, Mandelbrot's Apollonian Net, Doyle spirals, and the light cone in Lorentz space. We establish the intended sharp bounds for the Ring Lemma — all reciprocal integers, as it happens — and related bounds on "angular resolutions", useful in the computational geometry of graph embeddings. Our work with quads suggests an approach to an important open question on Hausdorff dimensions of residual sets, and we conclude with issues on the dynamics of quad shifts.

Our thanks to Vitaly Bergelson, Shashikant Mulay, and Uri Srebro for references and for valuable conversations on the material presented here. We are aware that some of the connections we have uncovered or rediscovered have been noted by others (some perhaps rediscovered several times); our bibliography is not intended to be exhaustive, particularly as regards the physics literature, but we have tried to provide fair attributions and apologize to anyone whose contributions may have been overlooked.

---

The first author was supported by the Fund for the Promotion of Research at the Technion.

The second author was supported by the National Science Foundation and the Tennessee Science Alliance.

## 1. Introduction

The term **Apollonian Packing** will refer to the infinite pattern of closed discs suggested by Figure 1 in the plane. We treat these as lying on the Riemann sphere, so the exterior of the outer circle is a disc like any other in the pattern — indeed, the packing is homogeneous in that any ordered triple of mutually tangent discs may be mapped to any other by a unique Möbius transformation of the sphere which leaves the overall pattern invariant. The fundamental unit of the packing is a **quad**, a set of four mutually externally tangent discs. Viewed on the sphere, each disc of a quad "fits" in one of the two triangular interstices formed by the other three.

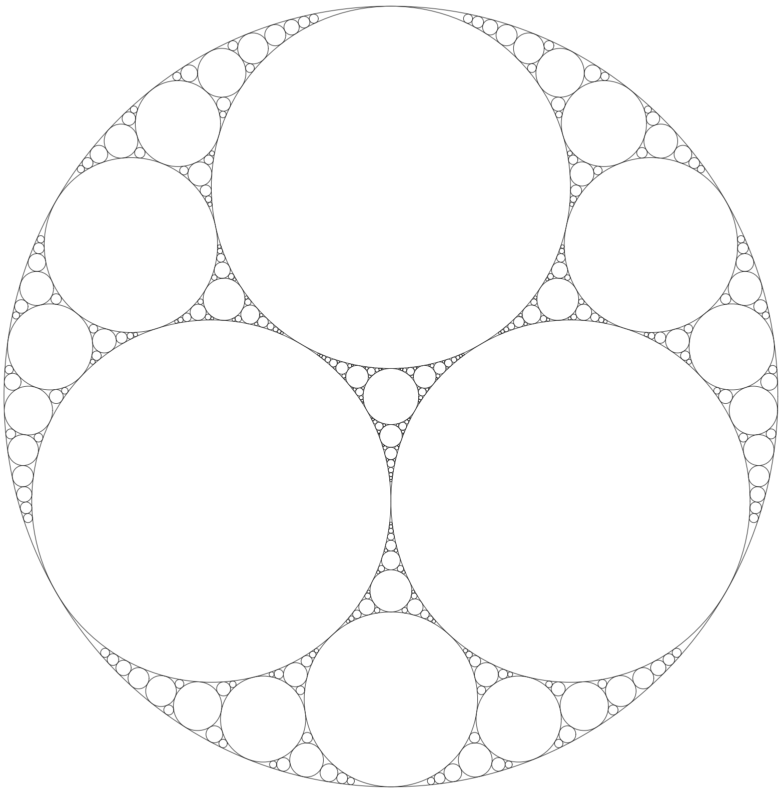

Figure 1. The Apollonian Packing.

In general we refer to a collection P of discs having mutually disjoint interiors is an **A-packing** if (1) each disc fits in an interstice formed by three others, and (2) every point of the sphere lies either in one of the discs or in a triangular interstice formed by three of them. The simplest A-packing consists of a single quad, but by fitting new discs in available interstices it is clear that one can create uncountably many others, even if one identifies those which are Möbius equivalent. The Apollonian packing is the unique (up to Möbius equivalence) A-packing in which every interstice contains a disc, and it is clear that all A-packings may be found within it.

Throughout the paper we will be focusing on three particularly pleasing A-packings, called the **Spiral**, the **Ring**, and the **Knife** patterns and denoted $P_S, P_R$, and $P_K$, respectively. These are displayed, with convenient Möbius normalizations, in Figure 2. Each of these A-packings is generated inductively from an initial quad of discs by using

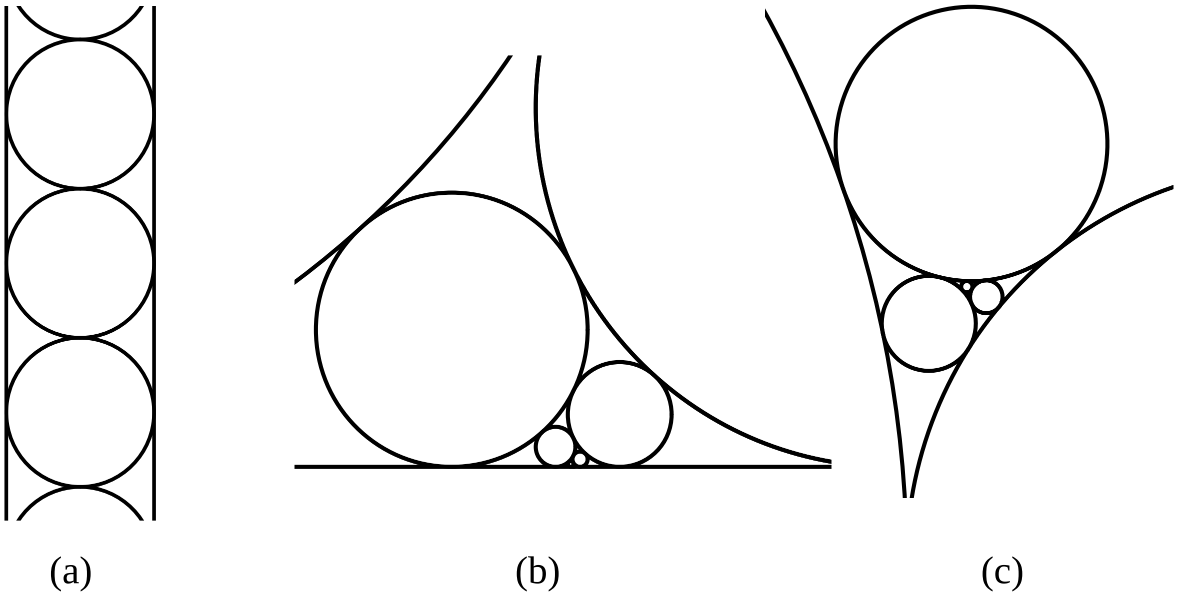


FIGURE 2. The three A-packings.

one of three elegantly simple geometric rules. Say we consider the Spiral $P_S$: Start with a quad of mutually tangent discs in the plane which determines three bounded interstices (curvilinear triangles). Go to the smallest of these and fit the largest possible disc $d_0$ inside it. This forms a new quad ($d_0$ with the three discs bounding the interstice) and it replaces the original interstice by three new smaller ones. Choose the smallest of these and fit the largest possible disc $d_1$ inside it, giving three yet smaller interstices. Choose the smallest of these, fit $d_2$ inside it, and so forth *ad infinitum*. The result is a sequence $\{d_0, d_1, \cdots\}$ of successively tangent discs. We will find, in fact, a group action that extends this construction process in a reverse direction, yielding a doubly infinite sequence $\{\cdots d_{-1}, d_0, d_1, \cdots\}$. These discs, along with the original quad, form the Spiral A-packing $P_S$. The Knife A-packing $P_K$ is obtained in the same way, except that the rule at each stage requires the next disc to be placed in the **largest** of the three interstices from the previous stage, while for the Ring pattern $P_R$, the rule requires that it be placed in the **middle**-sized of the three interstices.

So, what is the interest in these particular A-packings (and what accounts for their names)? We began our study on the practical side, searching for the constants $c(n)$

in the Ring Lemma of [24], which is related to conformal mapping using circle packings. Hansen [11] found the extremal "flowers" and gave an inductive formula for the constants. In [1] the first author brought in Descartes Theorem, leading to explicit expressions for the $c(n)$ and to the realization that they are reciprocal integers. As we show, the packing $P_R$ contains all the extremal flowers for this lemma, hence our **Ring** designation. In admiring the beauties of $P_K$, we were struck by a certain similarity to the "loxodromic" spiral of discs studied by Coxeter in [6]; this happens to be $P_S$, leading to its **Spiral** name. On discovery of a common geometric rule for those constructions, it became clear that $P_K$ was needed to complete the set. Since $P_K$ is the classically famous "arbelos", or "shoemaker's knife", from antiquity, we chose to call this last one the **Knife** pattern.

The construction process for each pattern is pleasing not only for its simplicity, but also because it is asymptotically stable — that is, a given rule will yield the same packing (modulo Möbius transformtions) regardless of the initial quad. The discs are in control. This issue is at the heart of Coxeter's work and explains that feeling of inevitability one gets when contemplating his spiral pattern. The asymptotic stability asserts itself algebraically, numerically, even geometrically. At the end of the day, one finds that by choosing an appropriate initial quad, the radii of the discs $d_j$ in each of these patterns form a geometric sequence. In fact, we will see that these are the only geometric sequences of this type in A-packings, which explains the title of our paper.

So, although we began our study with the best of pragmatic intentions, we became increasingly fascinated by the intrinsic beauty of the geometric and numerical patterns that emerged. We hope to give the reader a window, *via* these A-packings, into the rich web of geometric, number theoretic, algebraic, and recursive properties which we have found.

Briefly, here is how the paper proceeds. In the next section we establish basic definitions and notation, quote the important Descartes Circle Theorem, and give an overview of various mathematical structures within our patterns. We go on to devote a section to each of the A-packings, giving both the geometric intuition and the concrete numerical details for each; the conclusions for geometric sequences are summarized in Theorem 9, Section 6. In Section 7 we apply our understanding of $P_R$ to prove Theorems 10 and 11, giving the sharp values for the constants of the Ring Lemma and related constants and showing their integral nature. Indeed, our continual encounter with integers is a common thread throughout the paper; in Section 7 we investigate these and additional patterns of integers, showing in particular connections to Hausdorff dimensions of limit sets of Kleinian groups. In the last section we look more closely at the quadratic form $q$ appearing in Descartes Circle Theorem. In a final surprise, nature has somehow arranged that $q$ be equivalent to the classical Lorentz form, so the integer reciprocal bends of our packings are identified with integral orbits on the light cone of Minkowski space.

## 2. The Basic Structures

We will be working with discs in the euclidean plane, and like the ancient Greeks we include half-planes and "improper" discs, meaning the exteriors of circles. Radii will be euclidean, with the conventions that a half plane has radius $+\infty$ and an improper disc has radius the negative of that for its bounding circle. With this understanding, the **bend** of a disc (or of its bounding circle) is the reciprocal of its radius; in other words, bends reflect signed curvature.

Our typical notation for a quad of discs is $Q = \langle\langle D_1, D_2, D_3, D_4\rangle\rangle$; the ordering has no intrinsic significance but will be important for notational purposes. We frequently work also with the quads of associated bends, denoted $\langle\langle b_1, b_2, b_3, b_4\rangle\rangle$. (Similar notation may be used for **triples**.)

Given the quad $Q = \langle\langle D_1, D_2, D_3, D_4\rangle\rangle$, consider one of its triples, say $\langle\langle D_2, D_3, D_4\rangle\rangle$. This triple determines two interstices on the sphere, and $D_1$ fills one of them. If disc $D_5$ fills the other, then $Q' = \langle\langle D_2, D_3, D_4, D_5\rangle\rangle$ is also a quad. This transition $Q \to Q'$ will be termed a **quad shift**. Parallel to this, of course, is the shift of bends

$$\langle\langle b_1, b_2, b_3, b_4\rangle\rangle \to \langle\langle b_2, b_3, b_4, b_5\rangle\rangle.$$

We will see that one can generate our A-packings from an initial quad by repeatedly applying the appropriate quad shift to spawn the next disc. The shifts will be denoted $S_K, S_R$, and $S_S$, respectively, for the Knife, Ring, and Spiral A-packings and may be represented schematically as follows:

$$\begin{array}{lrcl} \text{Knife:} & S_K : \langle\langle D_1, D_2, D_3, D_4\rangle\rangle & \mapsto & \langle\langle D_1, D_2, D_4, D_5\rangle\rangle \\ \text{Ring:} & S_R : \langle\langle D_1, D_2, D_3, D_4\rangle\rangle & \mapsto & \langle\langle D_1, D_3, D_4, D_5\rangle\rangle \\ \text{Spiral:} & S_S : \langle\langle D_1, D_2, D_3, D_4\rangle\rangle & \mapsto & \langle\langle D_2, D_3, D_4, D_5\rangle\rangle. \end{array} \tag{1}$$

The key to understanding these shifts, and ultimately the A-packings, is an elegant relationship among the bends of a quad which was observed by Descartes. After establishing that, the remainder of this section introduces a variety of approaches to our three A-packings, loosely characterized as *geometric, recursive, group theoretic, homogeneous,* and *linear*.

### 2.1. Descartes Circle Theorem.

In a letter to Princess Elisabeth of Bohemia in 1643, Descartes sketched a quad and described a beautiful relationship among their bends, hereafter referred to as Descartes Circle Theorem (DCT):

**Theorem 1.** (DCT) *Let $b_1, b_2, b_3$, and $b_4$ be the bends of a quad. Then*

$$\Big(\sum_{j=1}^{4} b_j\Big)^2 = 2\sum_{j=1}^{4} b_j^2. \tag{2}$$

We note that at most one bend can be negative, in which case the others are positive, or at most two can be zero. The DCT holds in all cases. An easy computation from

(2) gives this useful formula for the largest bend (assume that is $b_4$):

$$b_4 = b_1 + b_2 + b_3 + 2\sqrt{b_1 b_2 + b_1 b_3 + b_2 b_3}. \tag{3}$$

The quantity under the radical is always nonnegative for quads of bends.

We picked up this background on the DCT from H. S. M. Coxeter [6], where the reader may find a proof **via** inversion. Coxeter cites additional proofs in the literature, along with extensions to higher dimensions. For a more lyrical view, we highly recommend the poem "The Kiss Precise" by Nobel laureat Frederick Soddy [27]:

Four circles to the kissing come.
The smaller are the benter.

The DCT, also known as Soddy's Formula, underlies all our approaches to the A-packings. Let us begin with the "geometric" viewpoint.

2.2. **The Geometric Structure.** There is much to recommend a naive geometric view of quads, though proofs ultimately rest on computations. Naively, discs can be treated as movable objects, since it is their radii and not their centers that matter. Changes in one radius must have an effect on others if a configuration is to maintain some prescribed tangency pattern — we describe these as "monotonicity" properties.

For instance, consider a quad $Q = \langle\langle D_1, D_2, D_3, D_4 \rangle\rangle$ of proper discs and say $D_4$ is the smallest, innermost, disc. Decreasing the radius of an outer disc forces one to shift the discs slightly and decrease the radius of the innermost disc in order to maintain the quad's tangencies.

**Lemma 2.** *In a quad of discs, the radius (resp. bend) of the innermost disc is monotone increasing in each of the radii (resp. bends) of the others.*

Equally intuitive, but more subtle, is relative monotonicity. Consider $Q$ again: as we have just seen, keeping the radii of two outer discs, say $r_1, r_2$, fixed while decreasing that of the third, $r_3$, causes $r_4$ to decrease. What happens to the ralative sizes of $r_3$ and $r_4$? Intuitively, they become closer in size; as $D_3$ shrinks, it drops ever deeper into the crevasse formed by $D_1$ and $D_2$, pushing $D_4$ ahead of it, and the walls of that crevasse look ever more like parallel lines to them. In fact, we have the

**Lemma 3.** *Let $r_1 \geq r_2 \geq r_3 \geq r_4 > 0$ be the radii of a quad, with $r_1$ and $r_2$ fixed. Then $r_4/r_3 \uparrow 1$ as $r_3 \downarrow 0$. In terms of the bends, $b_4/b_3 \downarrow 1$ as $b_3 \uparrow \infty$.*

Computationally, Lemmas 2 and 3 follow directly from (3).

2.3. **The Recursive Structure.** Our recursive approach involves sequences of bends obtained from repeated quad shifts. We rely on this Corollary to DCT:

**Corollary 4.** *Suppose $\langle\langle b_1, b_2, b_3, b_4 \rangle\rangle$ and $\langle\langle b_2, b_3, b_4, b_5 \rangle\rangle$ are the bends for quads $Q$ and $Q'$, respectively. Then*

$$b_5 = 2(b_2 + b_3 + b_4) - b_1. \tag{4}$$

*Proof:* The bends these quads have in common are $b_2, b_3, b_4$. By the DCT, each of the two bends which will form a quad with these three must satisfy

$$(x + b_2 + b_3 + b_4)^2 = 2\left(x^2 + b_2^2 + b_3^2 + b_4^2\right)$$
$$\implies x^2 - 2x(b_2 + b_3 + b_4) - 2(b_2b_3 + b_2b_4 + b_3b_4) + b_2^2 + b_3^2 + b_4^2 = 0.$$

The sum of the two roots will be the negative of the coefficient of $x$, implying $b_1 + b_5 = 2(b_2 + b_3 + b_4)$. □

2.4. **The Group Structure.** There are ubiquitous symmetries within the Apollonian packing because its discs turn out to be the orbit of a single disc under the action of a finitely generated Kleinian group. Write $\mathfrak{M}$ for the group of Möbius transformations of the sphere, both orientation-preserving and reversing. ($\mathfrak{M}$ is tightly bound up with circles: for instance, it is generated by reflections (inversions) in circles and its members are characterized by the fact that they carry discs to discs.) The orientation preserving Möbius transformations form a subgroup of index 2 and are routinely identified with the matrix group $\mathrm{SL}(2, \mathbf{C})$.

For the connection with the Apollonian packing, we take our lead from a description in Mandelbrot's well known book The Fractal Geometry of Nature [17]. One can begin with any quad, but for notation we use the quad $Q = \langle\langle D_1, D_2, D_3, D_4\rangle\rangle$ of Figure 3. Each of the four triples in $Q$ has a mutually perpendicular circle and these are labeled $C_1, C_2, C_3, C_4$ in the figure. (In fact, these form a quad themselves, a "dual" to $Q$.)

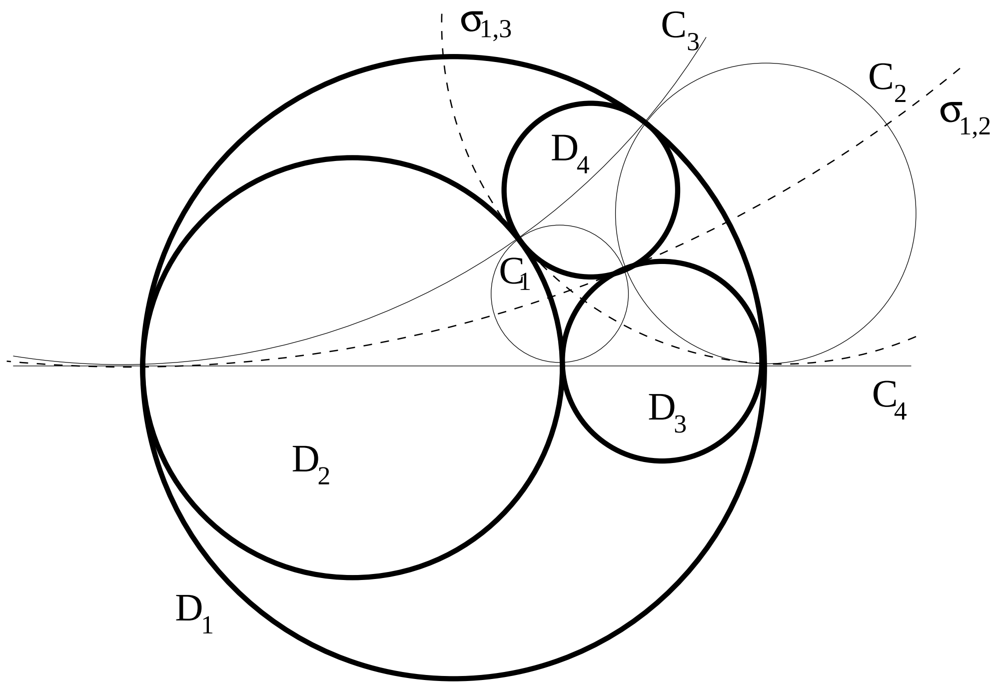


FIGURE 3. Notation for a quad $Q$ and its dual.

Mandelbrot observes that the Apollonian packing $A$ containing $Q$ is the smallest disc packing containing $Q$ which is closed under reflections in the $C_j, j = 1, 2, 3, 4$. These reflections will be denoted by $\phi_j, j = 1, 2, 3, 4$. Indeed, each disc of $A$ is obtained from one of those in $Q$ by applying a unique sequence of the $\phi_j$.

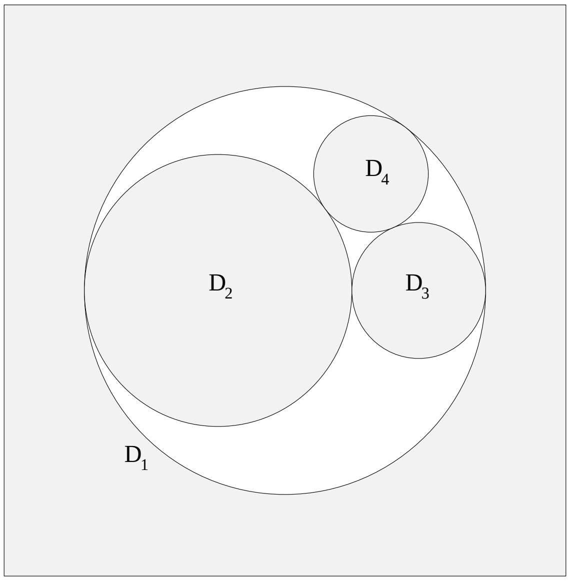


FIGURE 4. The Standard Quad $\mathfrak{Q}$.

The group generated by the $\phi_j$ is not sufficiently rich to generate our quad shifts, due simply to their special ordering, so we exploit additional symmetries of $A$. Namely, for each pair $D_i \neq D_j$ from $Q$, there is a circle $C$ which is orthogonal to $D_i$ and $D_j$ and tangent to the other two discs of $Q$. Write $\sigma_{ij}$ for reflection in this circle, for $1 \leq i, j \leq 4, i \neq j$. In other words, $\sigma_{ij}$ fixes $D_i$ and $D_j$ and interchanges the other two discs; $\sigma_{1,2}$ and $\sigma_{1,3}$ are shown in Figure 3.

To proceed concretely, it is necessary to settle on a particular starting quad. We have chosen the discs

$$D_1 = \{|z - 1/2| \geq 1/2\} \qquad D_2 = \{|z - 1/3| \leq 1/3\}$$
$$D_3 = \{|z - 5/6| \leq 1/6\} \qquad D_4 = \{|z - (5/7 + 2/7i)| \leq 1/7\}$$

shown in Figure 4, with $D_1$ an improper disc. We will denote this in the sequel as the quad $\mathfrak{Q}$. The bends here are small integers, making the computations relatively straightforward. One may confirm the following expressions for the Möbius transformations:

$$\sigma_{12} = \frac{2i\overline{z}}{\overline{z} + 2i} \qquad \sigma_{34} = \frac{2\overline{z}}{5\overline{z} - 2} \qquad \sigma_{13} = \frac{(2+i)\overline{z} - 2}{2\overline{z} - (2-i)} \qquad \sigma_{14} = \frac{4\overline{z} - 4}{3\overline{z} - 4}$$

$$\phi_2 = \frac{(4+i)\overline{z}-4}{4\overline{z}-(4-i)} \qquad \phi_3 = \frac{i\overline{z}}{\overline{z}+i} \qquad \phi_1 = \frac{(6+i)\overline{z}-4}{9\overline{z}-(6-i)}.$$

Our shifts are generated as follows:

$$\begin{array}{ll} \text{Knife:} & g_K(z) = \phi_3 \circ \sigma_{12} = \dfrac{2z}{-iz+2} \\ \text{Ring:} & g_R(z) = \phi_2 \circ \sigma_{12} \circ \sigma_{13} = \dfrac{(4-2i)\overline{z}-4}{(3-2i)\overline{z}-4} \\ \text{Spiral:} & g_S(z) = \phi_1 \circ \sigma_{34} \circ \sigma_{14} \circ \sigma_{12} = \dfrac{(4+2i)z-4}{(7+3i)z-6}. \end{array}$$

Note that under $S_R$ there is one, and under $S_K$ there are two discs which remain invariant; we will call these **guard** discs for later reference.

We denote the group generated by our shift transformations as $\Gamma = \langle g_S, g_R, g_K \rangle$. We leave it to the reader to confirm that the orbit of any disc of $\mathfrak{Q}$, say for instance $D_1$, gives a complete Apollonian packing — that is, one in which every interstice is filled. This specific Apollonian packing will be denoted by $\mathcal{A}$. Any other initial quad would result in a group conjugate to $\Gamma$ and an Apollonian packing which is a Möbius image of $\mathcal{A}$.

2.5. **The Linear Structure.** Another consequence of the DCT is a *linear* structure for quads of bends, exploited by Hirst [13] and Boyd [4]. (Also see Falconer [8].) Our particular concern is with the three quad shifts $S_S, S_R$, and $S_K$. In terms of bends, these are schematically represented as

$$\begin{array}{llll} \text{Knife:} & S_K : \langle\!\langle b_1, b_2, b_3, b_4 \rangle\!\rangle & \mapsto & \langle\!\langle b_1, b_2, b_4, b_5 \rangle\!\rangle \\ \text{Ring:} & S_R : \langle\!\langle b_1, b_2, b_3, b_4 \rangle\!\rangle & \mapsto & \langle\!\langle b_1, b_3, b_4, b_5 \rangle\!\rangle \\ \text{Spiral:} & S_S : \langle\!\langle b_1, b_2, b_3, b_4 \rangle\!\rangle & \mapsto & \langle\!\langle b_2, b_3, b_4, b_5 \rangle\!\rangle. \end{array}$$

If we treat each quad of bends as a real row 4-vector, then by Corollary 4 we can describe the shifts *via* matrices $M_S, M_R$, and $M_K$ in $\mathrm{GL}(4, \mathbf{R})$ given by

$$M_K = \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{bmatrix}, \qquad M_R = \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 2 \end{bmatrix}, \qquad M_S = \begin{bmatrix} 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 2 \end{bmatrix}.$$

One readily verifies using (4) that each of the shifts is represented by multiplying a row vector of bends **on the right** by the corresponding matrix. We have so arranged our shifts that any guard discs (resp. bends) are listed at the beginning while the new disc (resp. bend) appears at the end. Applying a sequence of shifts, for example, one successively multiplies by the appropriate matrices (on the right!) and notes the bends in the resulting fourth coordinate spot.

To generate *all* quads of bends associated with an Apollonian packing one needs both the shift matrices and permutation matrices for interchanging entries. Define $G$ as the subgroup which these generate in $\mathrm{GL}(4, \mathbf{R})$.

2.6. **The Homogenous Structure.** Euclidean scaling of a quad (of discs or of bends) yields another quad. Thus a substantial part of the geometry of a quad is reflected in the *relative proportions* of it bends. In studying a sequence of quads generated by quad shifts, for example, it makes sense to successively rescale the image quads so that, say, the last entry is 1, and then to see if a limit quad emerges. This is what we mean by the **asmyptotic** behavior of a sequence of quads: one looks for the limit and for the appropriate scaling.

Another approach is to identify quads of bends with points of the unit sphere $\mathbf{S}^3$ of $\mathbf{R}^4$. Each of our shifts $S_\bullet$ defines a homeomorphism of $\mathbf{S}^3$ by

$$\Phi_\bullet(\overline{\mathbf{b}}) = \frac{S_\bullet(\overline{\mathbf{b}})}{\|S_\bullet(\overline{\mathbf{b}})\|}.$$

The set of quads of bends is invariant under $\Phi_\bullet$. In this view, applying a sequence of quad shifts has the effect of iterating a homeomorphism of $\mathbf{S}^3$.

In fact, one can go further to real projective space $\mathbf{RP}^3$, since there is no ambiguity about signs with quads of bends — only one of $\overline{\mathbf{v}}$ and $-\overline{\mathbf{v}}$ can be a quad of bends. For $\overline{\mathbf{b}} = \langle\!\langle b_1, b_2, b_3, b_4\rangle\!\rangle$, write $\overline{\mathbf{b}}_h = \langle\!\langle b_1, b_2, b_3, b_4\rangle\!\rangle_h$ to represent $\overline{\mathbf{b}}$ in the homogeneous coordinates of $\mathbf{RP}^3$, meaning that two such vectors are equivalent iff each is a real scalar multiple of the other.

The homogeneity of quads of bends is not accidental. In fact, the DCT defines a symmetric quadratic form, which we will denote by $q$ and refer to as the *Descartes form*:

$$\begin{aligned} q(x_1, x_2, x_3, x_4) &= (x_1^2 + x_2^2 + x_3^2 + x_4^2) - (\tfrac{1}{2})(x_1 + x_2 + x_3 + x_4)^2 \\ &= (\tfrac{1}{2})(x_1^2 + x_2^2 + x_3^2 + x_4^2) - (x_1x_2 + x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4 + x_3x_4). \end{aligned}$$

A nonzero 4-vector $\overline{\mathbf{v}}$ (or its negative) is a quad of bends if and only if $q(\overline{\mathbf{v}}) = 0$. We will study this form in Section 9.

2.7. **The Integral Structure.** We close this section by commenting on the source of the integers which will appear throughout the remainder of the paper. The following is an immediate consequence of Corollary 4.

**Lemma 5.** *If $\overline{\mathbf{b}}$ is a quad of bends having integer entries and $S_\bullet$ is any one of the quad shifts, then $S_\bullet(\overline{\mathbf{b}})$ will also have integer entries.*

In particular, by generating $\mathcal{A}$ using the initial quad $\mathfrak{Q}$ with integer bends, ALL the bends of $\mathcal{A}$ are integers! (That is, all its radii are reciprocal integers.) This is also reflected in the coefficients of the transformations in $\Gamma$, which are Gaussian integers, elements of $\mathbf{Z}[i]$. In §8 especially, we look more closely at this numerical structure.

## 3. The Shoemaker's Knife

We begin our study of the three A-packings with one taken straight from ancient Greek mathematics — the Knife pattern $P_K$. Recall that it is generated from an initial quad by successively placing new discs in interstices, choosing the "largest" interstice each time. This is in most respects the simplest, but nonetheless shows the general features of all.

**Geometric Structure.** Suppose we start with an initial quad $Q = \langle\langle D_1, D_2, D_3, D_4 \rangle\rangle$ whose bends are in nondecreasing order, making $D_4$ the innermost, smallest disc. By Lemma 2, the largest (bounded) interstice is that formed by $\langle\langle D_1, D_2, D_4 \rangle\rangle$. Placing the largest possible disc $D_5$ in this interstice, we obtain the quad $\langle\langle D_1, D_2, D_4, D_5 \rangle\rangle$, automatically again in order of decreasing radii. Repeating, it is clear that this operation is represented by the shift $S_K$, as we anticipated:

$$S_K : \langle\langle D_1, D_2, D_{n-1}, D_n \rangle\rangle \mapsto \langle\langle D_1, D_2, D_n, D_{n+1} \rangle\rangle.$$

Since the initial discs $D_1, D_2$ remain in every quad, we call them **guard** discs for the evolving pattern.

It is clear, of course, that as $n$ grows, the guard discs remain unchanged while the radii of successive discs $D_n$ shrink to zero. Moreover, it seems evident, if one visualizes the first few steps of construction, that the geometry of the successive quads quickly stabilizes and with appropriate scaling, converges to a limit quad. For instance, the bends of the two non-guard discs go to infinity by Lemma 2, so rescaling, say to keep the fourth bend 1, will drive the (rescaled) bends of the guard discs to zero. The limiting quad must therefore have bends $\langle\langle 0, 0, \mu, 1 \rangle\rangle$, and by Lemma 3, $\mu = \mu_K = 1$. Let us confirm this geometric intuition using the various mathematical structures we discussed in the last section and try to verify by each approach that $\mu_K = 1$.

**Recursive Structure.** Starting from a Greek perspective, let $Q$ consist of the shaded discs in Figure 5, the others obtained by successive quad shifts. We just show the upper portion of the pattern. portion Interest in this configuration can be traced back more than two thousand years through the Collection of Pappus of Alexandria (*circa* 300 A.D.) to the lost books of Apollonius of Perga, almost 500 years earlier. The half-crescent region in which the sequence of discs lies apparently suggested to them a curved knife blade. (This is but one case in a much broader study of configurations of circles, lines, and points whose beginnings are lost in the mists of time. For some of the ancient texts, with commentary, see [14]; also the classic Heath [12].) We apply a remarkable property of the arbelos proven by the ancients. Label the circles $C_n$ as marked in Figure 5; write $r_n$ for the radius of $C_n$ and $y_n$ for the distance of its center from the diameter $(a, b)$. Then $y_n$ and $r_n$ satisfy this elegant relationship:

$$y_n = 2nr_n. \tag{5}$$

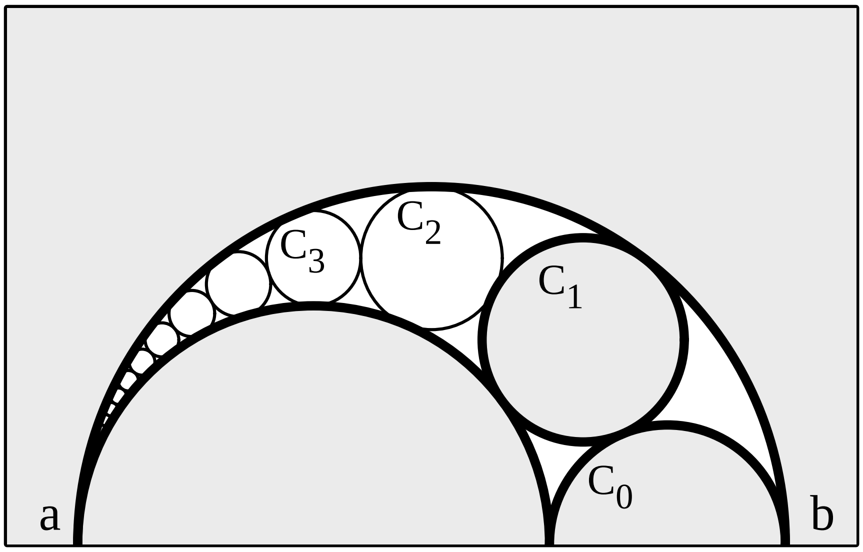


FIGURE 5. The "$\acute{\alpha}\rho\beta\eta\lambda o\zeta$", or shoemaker's knife.

(This, in fact, holds for all such patterns with $C$ and $C_0$ spanning a diameter, independent of their radii. Wilker [30] gives a marvelously short proof based on inversion.) It is clear that $\frac{y_{n+1}}{y_n} \to 1$ as $n \to \infty$. Appealing to (5), we have $\frac{r_{n+1}}{r_n} = \frac{b_n}{b_{n+1}} \to 1$, implying that $\mu_K = 1$.

For later purposes, we also consider a shoemaker's knife generated from another initial quad, that illustrated in Figure 6. The half-plane will be the disc $D_0$. Let $r_j > 0$ denote the radius for $D_j$, $j = 1, 2, 3$, with $r_3$ the smallest. If $d_{ij}$ denotes the distance between the points where $D_i$ and $D_j$ are tangent to $D_0$, then

$$(r_i + r_j)^2 - (r_i - r_j)^2 = d_{ij}^2 \implies d_{ij} = 2\sqrt{r_i r_j}. \tag{6}$$

Since the tangency point of $D_3$ lies between those of $D_1$ and $D_2$, we have by Pythagoras

$$d_{12} = d_{13} + d_{23} \implies \sqrt{r_1 r_2} = \sqrt{r_1 r_3} + \sqrt{r_2 r_3}.$$

Dividing by $\sqrt{r_1 r_2 r_3}$ gives this general formula which we record for later use: *For a quad with $D_0$ a half plane and $r_3$ the smallest radius,*

$$\frac{1}{\sqrt{r_3}} = \frac{1}{\sqrt{r_2}} + \frac{1}{\sqrt{r_1}}. \tag{7}$$

To generate the Knife, treat $D_0$ and $D_1$ as the guard discs and build the chain of discs $\{D_n\}$, starting with $D_2$ and $D_3$. For instance, suppose we set $r_1 = 1$, $r_2 = 1/4$, implying $r_3 = 1/9$. An inductive argument based on (7) gives $r_n = \frac{1}{n^2}$. In terms of bends, the sequence is $b_n = n^2$. Again, therefore, we see that the ratio of successive

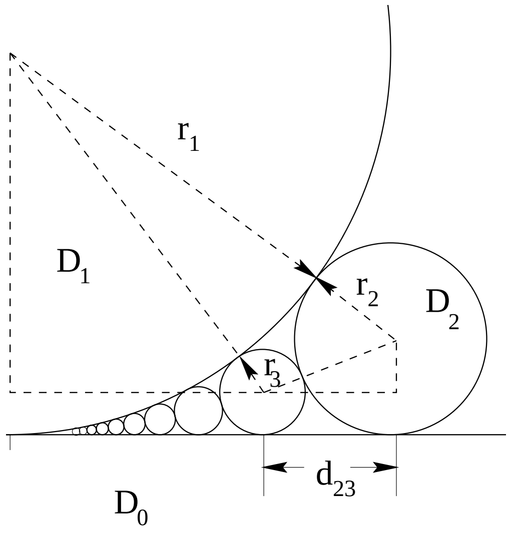


FIGURE 6. Another quad for the shoemaker's knife.

bends converges to $\mu_K = 1$. As an aside, note that

$$1 = d_{12} = \sum_{n=2}^{\infty} d_{n,n+1} = \sum_{n=2}^{\infty} \frac{2}{n(n+1)}.$$

Various summation formulas of similar type may be obtained with other choices for $r_1, r_2$ in the intial quad.

An explicit formula for successive bends in the Knife pattern (at least for certain intial quads) was established by Melzak [19]. Here is a slightly more general formulation.

**Lemma 6.** *Let* $\langle\langle g_1, g_2, b_1, b_2 \rangle\rangle$ *be a quad of bends for generating a shoemaker's knife. In particular* $g_1, g_2$ *represent the bends of the guard discs and successive bends* $b_n$ *are obtained by applying the shift* $S_K$*. Then*

$$b_n = (g_1 + g_2)n^2 - (b_1 - b_2 + 3g_1 + 3g_2)n + 2b_1 - b_2 + 2g_1 + 2g_2, \;\; n \geq 1. \tag{8}$$

*Proof:* By (4), the $b_n$ satisfy for $n \geq 3$ the recursive relationship

$$b_n = 2(g_1 + g_2 + b_{n-1}) - b_{n-2}.$$

The general solution has the form $y(n) = (g_1 + g_2)n^2 + An + B$ for constants $A, B$. Imposing the initial conditions $y(j) = b_j, j = 1, 2$, leads easily to (8). □

It is now evident from (8) that in a shoemaker's knife, $b_{n+1}/b_n \longrightarrow 1 = \mu_K$ as $n \to \infty$.

**Group Structure.** The clearest (most mundane!) shoemaker's knife is the pattern of Figure 2(a), with the tangency point of the two guard discs at infinity and initial bends $\langle\!\langle 0,0,1,1\rangle\!\rangle$. Note that this is the promised geometric sequence, with ratio 1. It is invariant under the shift $S_K$, which in this normalization is simply the translation $T: z \mapsto z+2i$.

Turning instead to the group $\Gamma$ associated with our standard quad $\mathfrak{Q}$, recall that $P_K$ is generated by $g_K(z) = 2z/(-iz+2)$. This transformation, conjugate to T, is parabolic with its single fixed point at $z=0$, and computation gives $g_K'(0) = 1$, implying yet again that $\mu_K = 1$.

Algebraically, we have two methods to determine $\mu_K$, both rather transparent for the Knife. The normalized limit quad of bends clearly has the form $\langle\!\langle 0,0,\mu,1\rangle\!\rangle_h$, for some $\mu$, so by the DCT, $\mu$ satisfies

$$(x+1)^2 = 2(x^2+1) \Longrightarrow (x-1)^2 = 0.$$

On the other hand, these four bends are trivially in geometric sequence, with factor $\mu^{-1}$; that is, the shift is given, before rescaling, by

$$S_K: \ \langle\!\langle 0,0,\mu,1\rangle\!\rangle \mapsto \langle\!\langle 0,0,1,\mu^{-1}\rangle\!\rangle.$$

The recursive relationship provided by (4) gives, for this shift,

$$\mu^{-1} = 2(0+0+1) - \mu \Longrightarrow \mu^2 - 2\mu + 1 = (\mu-1)^2 = 0.$$

**Linear Structure.** The linear structure is not particularly helpful in the Knife case because the matrix $M_K$ has eigenvalue 1 with multiplicity 4 and fails to have a basis of eigenvectors. Indeed, there are two eigenvectors, the quad $\langle\!\langle 0,0,1,1\rangle\!\rangle$ and the non-quad $\langle -1,1,0,0\rangle$.

**Homogeneous Structure.** Let us next move to the sphere $\mathbf{S}^2$ and the homeomorphism $\Phi_K$ defined in §2.6. The quads of bends in a Knife pattern generated from $\overline{\mathbf{b}}$ correspond, after projection to $\mathbf{S}^2$, with the images $\Phi_K^n(\overline{\mathbf{b}})$. What are the dynamics of this homeomorphism? $\Phi_K$ has two fixed points, namely, the two eigenvectors noted above. Numerical experiments suggest that $\Phi_K^n(\overline{\mathbf{v}})$ converges to $\langle\!\langle 0,0,1,1\rangle\!\rangle_h$ not only when we start with $\overline{\mathbf{v}}$ a quad of bends, but for **every** $\overline{\mathbf{v}}$ other than $\langle -1,1,0,0\rangle$. However, we have yet to study this mapping in any detail.

## 4. The Ring

We come next to the A-packing $P_R$, obtained from an initial quad by repeatedly adding the next disc to fill the middle-sized interstice.

In a generic quad with positive bends, Lemma 2 implies that the middle-sized interstice is formed by the triple containing the largest and smallest of the outer discs. As the ring construction proceeds, therefore, only the largest of the original discs remains in all later quads; that is, the pattern evolves with a single guard disc — in the corresponding shift $S_R$, this is the first entry:

$$S_R: \ \langle\!\langle D_1, D_2, D_3, D_4 \rangle\!\rangle \mapsto \langle\!\langle D_1, D_3, D_4, D_5 \rangle\!\rangle.$$

Repeated shifts clearly drive the nonguard discs to zero in radius. Our task, again, is to seek the asymptotic limit of the quads, determining the appropriate scaling for convergence and the limit entries.

**Geometric Structure.** What do we see geometrically? As the nonguard discs necessarily shrink, one must continually rescale and the guard disc takes on the aspect of a half plane, with the other three nestled against it. Starting with any such config-

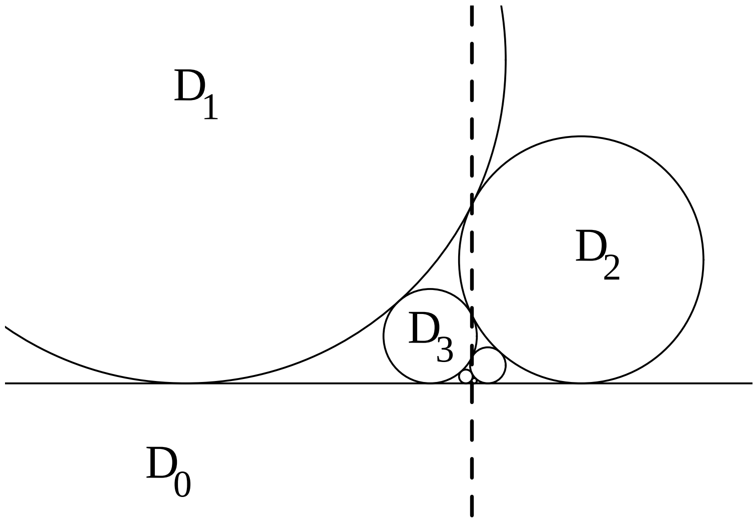


FIGURE 7. A quad for the Ring.

uration, the developmental pattern emerges almost immediately: successively added discs are clearly decreasing in radius. This in turn implies that they form a zig-zag pattern as viewed from the center of the guard disc, since the middle-sized interstice must alternate between the right and the left sides of the latest added disc. Also, this zig-zag yields a nested pattern of tangencies with the guard disc that brackets the new discs ever more tightly, forcing them to converge to a limit point on the boundary of the guard disc. See Figure 7.

In summary, then, normalization consists, at each stage, of scaling and "flip" operations. Assuming the existence of an asymptotic limit $\mu_R$ for the scaling factor, which we will show shortly, we can already conclude that $\mu_R > 1$ and that the limiting quad of bends must be in geometric sequence, $\langle\!\langle 0, 1, \mu_R, \mu_R^2 \rangle\!\rangle_h$.

Elementary monotonicity suggests the existence of $\mu_R$. This is left as an interesting diversion for the reader: you might consider the situation as pictured in Figure 7, with the guard disc $D_0$ as the lower half plane and $D_1$ fixed. Playing Lemmas 2 and 3 against one another, one can show existence and uniqueness of a radius $r_2$ for which $r_2/r_1 = r_3/r_2$. This configuration clearly gives a scaled replicate of itself under $S_R$, implying a limiting quad of bends of the expected geometric form.

**Recursive Structure.** The geometric argument does not suggest the *value* of $\mu_R$. For that, let us apply earlier computations to Figure 7. Each new disc $D_{n+1}$ is tangent to $D_n$ and $D_{n-1}$, so according to (7), the radii $r_n$ satisfy the recurrence relation

$$\frac{1}{\sqrt{r_{n+1}}} = \frac{1}{\sqrt{r_n}} + \frac{1}{\sqrt{r_{n-1}}}.$$

For the sequence $\{r_j\}$ then, one recognizes the familiar pattern of a generalized Fibonacci sequence. It is well known that the increasing terms have an asymptotic ratio, in this case

$$\frac{1}{\sqrt{r_{n+1}}}/\frac{1}{\sqrt{r_n}} = \frac{\sqrt{r_n}}{\sqrt{r_{n+1}}} \longrightarrow \frac{1+\sqrt{5}}{2} = \tau,$$

where $\tau$ is the famous **golden ratio**. Thus $r_n/r_{n+1} \longrightarrow \tau^2$. In particular, $\mu_R = \tau^2 = 1+\tau = (3+\sqrt{5})/2$, and the limiting quad of bends is $\langle\langle 0, 1, \tau^2, \tau^4 \rangle\rangle_h$.

Alternately, there is an explicit formula for successive bends, the parallel to Lemma 6 for the shoemaker's knife case.

**Lemma 7.** *Let $\langle\langle g, b_1, b_2, b_3 \rangle\rangle$ be a quad of bends for generating a Ring under $S_R$, with $g$ the bend of the guard disc. The successive bends $b_n$ are given by the following formula (expressed using the golden ratio $\tau$).*

$$b_n = \alpha(1+\tau)^n + \beta(2-\tau)^n + \gamma(-1)^n - g, \;\; n \geq 1, \tag{9}$$

*where*

$$\begin{aligned}
\alpha &= (1/5)\big[(10-6\tau)g + (8-5\tau)b_1 + (5-3\tau)b_2 + (-3+2\tau)b_3\big],\\
\beta &= (1/5)\big[(4+6\tau)g + (3+5\tau)b_1 + (2+3\tau)b_2 + (-1-2\tau)b_3\big],\\
\gamma &= (1/5)\big[g - b_1 + 3b_2 - b_3\big].
\end{aligned}$$

*Proof:* Define $y(n) = b_n + g, \;\; n \geq 1$. A quad $\langle\langle g, b_n, b_{n+1}, b_{n+2} \rangle\rangle$ shifts to give $\langle\langle g, b_{n+1}, b_{n+2}, b_{n+3} \rangle\rangle$ yielding by (4) the recursive relationship

$$y(n+3) = 2y(n+1) + 2y(n+2) - y(n). \tag{10}$$

This is a linear difference equation to be solved with "initial conditions"

$$y(j) = b_j + g, j = 1, 2, 3.$$

Anticipating a solution to (10) of form $Y(n) = \eta X^n$, some $\eta$ and $X$, we have

$$\begin{aligned}
&\eta X^{n+3} = \eta(2X^{n+1} + 2X^{n+2} - X^n)\\
&\Longrightarrow 0 = X^3 - 2X^2 - 2X + 1 = (X+1)(X^2 - 3X + 1).
\end{aligned}$$

This is an equation we will see again shortly and its solutions are $-1$, $\frac{3+\sqrt{5}}{2} = 1+\tau$, and its conjugate $\frac{3-\sqrt{5}}{2} = 2-\tau$. (Note: $(1+\tau)(2-\tau) = 1$.) The general solution $y(n)$ to (10) has the form

$$y(n) = \alpha(1+\tau)^n + \beta(2-\tau)^n + \gamma(-1)^n, \tag{11}$$

with constants $\alpha, \beta, \gamma$. Imposing the three initial conditions gives a linear system which may be solved for the $\alpha, \beta, \gamma$ given in the statement. Subtract $g$ from (11) to recover the formula for $b_n$. □

Now we can see the value of $\mu_R$. Assuming $\alpha \neq 0$, and noting that $(1+\tau) > 1$ while $(2-\tau) < 1$, formula (9) gives

$$b_{n+1}/b_n \longrightarrow 1+\tau = \frac{3+\sqrt{5}}{2} = \mu_R \text{ as } n \to \infty.$$

Two final observations along these lines. The pattern of bends $\langle\!\langle 0, 1, \mu_R, \mu_R^2 \rangle\!\rangle_h$ implies by DCT that $\mu_R$ satisfy the equation

$$0 = x^4 - 2x^3 - x^2 - 2x + 1 = (x^2 + x + 1)(x^2 - 3x + 1). \tag{12}$$

We saw this second factor in (4); the real roots are $\dfrac{3 \pm \sqrt{5}}{2}$, so again we conclude $\mu_R = \dfrac{3+\sqrt{5}}{2}$. On the other hand, the recurrence relation (4) implies that $\mu_R$ satisfies

$$0 = x^3 - 2x^2 - 2x + 1 = (x+1)(x^2 - 3x + 1). \tag{13}$$

The significance of the other roots to these equations is not clear. The remaining roots of (12) are complex; the root $-1$ of (13) arises again in the linear analysis.

**Group Structure.** Next, consider the group description of the Ring pattern. The A-packing $P_R$ is generated from the initial quad $\mathfrak{Q}$ by the Möbius transformation $g_R(z)$ in §2.4. This is an anti-conformal transformation known as a "glide-reflection". It has an invariant disc and an invariant axis perpendicular to it, with an action described as "dilation along followed by reflection across the axis." This, of course, is the pattern of scalings and flips observed geometrically. What is the scaling? The composition $g_R \circ g_R$ is the conformal hyperbolic Möbius transformation

$$g_R(g_R(z)) = \frac{(4-4i)z + 4i}{(2-5i)z + 2 + 4i}.$$

A computation shows that its fixed points are

$$\begin{aligned} z_1 &= (1/29)(5/2 - i)(8 + 4\tau i) \approx 0.912832 + 0.282081i \\ z_2 &= (1/29)(5/2 - i)(8 - 4i/\tau) \approx 0.604409 - 0.488977i. \end{aligned}$$

The first of these is attracting; the derivative $(g_R \circ g_R)'(z_1)$ is real and gives the local rate of attraction. Computation shows

$$(g_R \circ g_R)'(z_1) = (\frac{2}{3+\sqrt{5}})^2.$$

The square root of its reciprocal is exactly the scaling factor one would require to study the asymptotic behavior of $g_R$ itself. This, of course, turns out to be $\mu_R = \tau^2$.

**Linear Structure.** In our linear model, the shift for the Ring is represented by the matrix $M_R$, whose characteristic polynomial is

$$x^4 - 3x^3 + 3x - 1 = (x^2 - 1)(x^2 - 3x + 1).$$

This gives these four eigenvalues and associated eigenvectors:

$$\begin{aligned}
\lambda_1 = \frac{3+\sqrt{5}}{2}: &\qquad \overline{\mathbf{v}}_1 = \langle 0, 1, (\frac{3+\sqrt{5}}{2}), (\frac{3+\sqrt{5}}{2})^2 \rangle \\
\lambda_2 = \frac{2}{3+\sqrt{5}}: &\qquad \overline{\mathbf{v}}_2 = \langle 0, (\frac{3+\sqrt{5}}{2})^2, \frac{3+\sqrt{5}}{2}, 1 \rangle \\
\lambda_3 = 1: &\qquad \overline{\mathbf{v}}_3 = \langle -1, 1, 1, 1 \rangle \\
\lambda_4 = -1: &\qquad \overline{\mathbf{v}}_4 = \langle 0, 1, -1, 1 \rangle.
\end{aligned}$$

The largest eigenvalue is $\lambda_1$, which of course is equal to $\mu_R$ — we will see why the linear structure ensures this in a moment. Note that the first two eigenvectors are in fact the same quad of bends, written in different order

$$\begin{aligned}
\overline{\mathbf{v}}_1 =& \langle\langle 0, 1, \mu_R, \mu_R^2 \rangle\rangle \\
\overline{\mathbf{v}}_2 =& \langle\langle 0, \mu_R^2, \mu_R, 1 \rangle\rangle.
\end{aligned}$$

Let $T$ denote the linear transformation $(1/\mu_R)M_R$, and express an arbitrary quad as

$$\overline{\mathbf{b}} = c_1\overline{\mathbf{v}}_1 + c_2\overline{\mathbf{v}}_2 + c_3\overline{\mathbf{v}}_3 + c_4\overline{\mathbf{v}}_4.$$

Clearly $\overline{\mathbf{b}}T^n \longrightarrow c_1\overline{\mathbf{v}}_1$ as $n \to \infty$. Thus, any quad $\overline{\mathbf{b}}$ not in the span of $\{\overline{\mathbf{v}}_2, \overline{\mathbf{v}}_3, \overline{\mathbf{v}}_4\}$ converges asymptotically under $M_R$ to the quad $\overline{\mathbf{v}}_1$. For vectors in this span, the projection on $\overline{\mathbf{v}}_2$ is driven to zero because $\lambda_2 < 1$ and the action on $\overline{\mathbf{v}}_3$ and $\overline{\mathbf{v}}_4$ is clear.

## 5. The Coxeter Spiral

Our third A-packing is $P_S$, obtained from an initial quad by repeatedly filling the smallest interstice — this is the spiral which Coxeter investigated in [6].

In a quad with positive bends, Lemma 2 ensures that the smallest interstice is bounded by the smallest three discs. Thus, at each stage in the spiral development the largest disc is dropped and a new, the smallest, added. In particular, there is no "guard" disc, so the pattern is that of the quad shift $S_S$

$$S_S: \langle\langle D_1, D_2, D_3, D_4 \rangle\rangle \mapsto \langle\langle D_2, D_3, D_4, D_5 \rangle\rangle.$$

The discs clearly shrink quite rapidly (the radius of the smallest being less than one sixth that of the largest at each stage).

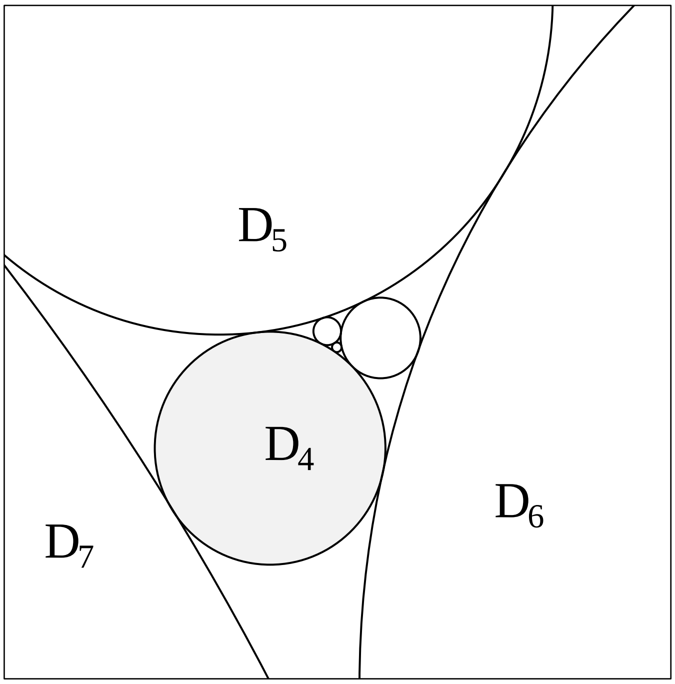


FIGURE 8. Seven discs in the spiral.

**Geometric Structure.** One could anticipate the aymptotic behavior of the discs composing $P_S$ using elementary monotonicity arguments akin to (but more complicated than) those of the previous Ring case. However, there are new considerations here which make the pattern seem even more inevitability. Based on recent rigidity results from the study of circle packings, $P_S$ is in fact completely determined by its combinatorics alone — that is, by the abstract pattern of tangencies among its discs.

Consider any sequence of seven successive discs in the Coxeter spiral pattern, say the $D_1, \cdots, D_7$ as pictured in Figure 8. This represents a string of four successive quads in the construction — indeed, precisely the quads involving $D_4$. In particular, the other six discs are the only ones in $P_S$ tangent to $D_4$. This local pattern holds at every disc, meaning that $P_S$ is a **hexagonal** disc packing, every disc is tangent to six others.

Hexagonal disc packings were characterized in [2]. Indeed, the Coxeter spiral is one of an infinite family called "Doyle" spirals, being one of the so-called "coherent" cases characterized there in Theorem 7. We must refer the reader to that paper for examples and analysis. Suffice it to say that the characterization guarantees that a "flower" such as that shown in Figure 8 has a certain specified pattern of radii. After an appropriate

Möbius transformation and a scaling so that $D_4$ has radius $r_4 = 1$, it is known that the remaining radii have this pattern for some parameters $x, y > 0$:

$$(r_1, r_2, r_3, r_5, r_6, r_7) = (\frac{1}{y}, \frac{1}{x}, \frac{x}{y}, \frac{y}{x}, x, y). \tag{14}$$

Applying Descartes form, §2.6, the bends of the quad $\langle\langle D_4, D_5, D_6, D_7\rangle\rangle$ satisfy

$$q(1, \frac{x}{y}, \frac{1}{x}, \frac{1}{y}) = 0.$$

Algebraic simplification yields

$$y^2 - y[2x\frac{(x+1)^2}{(x-1)^2}] + x^2 = 0. \tag{15}$$

The quads $\langle\langle D_3, D_4, D_5, D_6\rangle\rangle$ and $\langle\langle D_2, D_3, D_4, D_5\rangle\rangle$ are shifts of one another, so (4) applied to their bends simplifies to give

$$y^2 - y[\frac{(x-1)^2}{2}] + x^2 = 0.$$

Identifying the coefficients of $y$ yields the equation

$$\begin{aligned} 0 &= x^4 - 8x^3 - 2x^2 - 8x + 1 \\ &= (x^2 - (4 + 2\sqrt{5})x + 1)(x^2 - (4 - 2\sqrt{5})x + 1). \end{aligned}$$

(We will shortly see other equations reminiscent of this one.) Its two real roots are $(\tau + \sqrt{\tau})^{\pm 2}$, where $\tau$ as usual is the golden ratio. Comparing (14) and Figure 8, we have $x > 1$ and $y > x$; thus set $x = (\tau + \sqrt{\tau})^2$ and solve for $y$ in (15) to verify that $y = (\tau + \sqrt{\tau})^3$.

Define $\mu = \tau + \sqrt{\tau}$. We conclude from (14) that the radii of our seven discs take the form

$$(r_1, \cdots, r_7) = (\mu^{-3}, \mu^{-2}, \mu^{-1}, 1, \mu, \mu^2, \mu^3);$$

in other words, they are geometric as anticipated, with ratio $\mu = \tau + \sqrt{\tau} \approx 2.89005$. This is the value $\mu_S$ of scaling obtained by Coxeter [6]. (It would be interesting to know if Coxeter's spiral is the only Doyle spiral with radii in geometric sequence.)

**Recursive Structure.** Again we can obtain an explicit formula for successive bends in the spiral packings, paralleling Lemmas 6 and 7, by studying the recursive relationship.

**Lemma 8.** *Let $\langle\langle b_1, b_2, b_3, b_4\rangle\rangle$ be a quad of bends for generating a Sprial A-packing under $S_S$. The successive bends $b_n$ are given by a formula expressed using the golden ratio $\tau$,*

$$b_n = \alpha(\tau + \sqrt{\tau})^n + \beta(\tau - \sqrt{\tau})^n + \gamma(\Re(\lambda_3^n)) + \delta(\Im(\lambda_4^n)), \tag{16}$$

*with coefficients $\alpha, \beta, \gamma, \delta$ depending on $b_1, \cdots, b_4$.*

*Proof:* Let $y(n) = b_n$. We have by (4) the recursive relationship

$$y(n+4) = 2y(n+1) + 2y(n+2) + 2y(n+3) - y(n), \forall n. \tag{17}$$

This is a linear difference equation to be solved with initial conditions $y(j) = b_j, j = 1, \cdots, 4$, and again we anticipate solutions of form $Y(n) = \eta X^n$, some $\eta$ and $X$. This leads, as in the proof of Lemma 7, to an equation in $X$. Namely,

$$\begin{aligned} 0 &= X^4 - 2X^3 - 2X^2 - 2X + 1 \\ &= (X^2 - 2\tau X + 1)(X^2 + 2(\tau - 1)X + 1). \end{aligned} \tag{18}$$

There are two real roots, $\lambda_1 = \tau + \sqrt{\tau}$ and $\lambda_2 = \tau - \sqrt{\tau} = 1/(\tau + \sqrt{\tau})$, and two conjugate complex roots of modulus one, $\lambda_3 = (1-\tau) + i(\sqrt{\tau - 1})$ and $\lambda_4 = (1-\tau) - i(\sqrt{\tau - 1})$. The general real solution to (17) is

$$y(n) = \alpha(\lambda_1^n) + \beta(\lambda_2^n) + \gamma(Re(\lambda_3^n)) + \delta(\Im(\lambda_4^n)),$$

with real coefficients. Imposing the initial conditions leads to appropriate values for (16); we leave this computation to the reader. ☐

From (16), after noting that $\lambda_1 > 1$ while $\lambda_2 = 1/\lambda_1 < 1$ and that $|\lambda_3| = |\lambda_4| = 1$, we obtain

$$b_{n+1}/b_n \longrightarrow \tau + \sqrt{\tau} = \mu_S \text{ as } n \to \infty.$$

Alternately, we may proceed in parallel with the earlier cases (the line that Coxeter himself used). Assuming a geometric sequence of bends $\langle\langle \mu^3, \mu^2, \mu^1, 1 \rangle\rangle_h$ implies by DCT that $\mu$ satisfies the equation

$$\begin{aligned} 0 &= x^6 - 2x^5 - x^4 - 4x^3 - x^2 - 2x + 1 \\ &= (1 + x^2)(x^4 - 2x^3 - 2x^2 - 2x + 1). \end{aligned} \tag{19}$$

The linear recurrence relation (4) gives equation (18), where we see again the second factor of (19). The largest root, $\tau + \sqrt{\tau}$, is, of course, our $\mu_S$ yet again. It would be interesting to find the significance of the various missing roots to the above equations in the geometry of the spiral. Coxeter shows, for instance, that the nonreal roots, $\lambda_3, \lambda_4$, of (18) are $\exp(\pm\beta i)$, where $\beta$ is the rotation angle (see below) for the loxodromic transformation $g_S$. In fact, Coxeter shows that $z = \mu_S \exp(\beta i)$ solves the equation

$$z^4 + 4z^3 + 10z^2 + 4z + 1 = (z+1)^4 + (2z)^2 = 0,$$

so reminiscent of others we have seen. What is the significance of the roots $x = \pm i$ of (19)? We don't know, but one might note a difference here with the previous Ring case, were the linear recurrence expression (13) has one root not shared by the DCT equation (12). That extra root turned out to be an eigenvalue of the shift matrix.

**Group Structure.** The A-packing $P_S$ is generated from the initial quad $\mathfrak{Q}$ by the Möbius tranformation $g_S(z)$ in §2.4. This is "loxodromic" with the two fixed points. Temporarily write $g_S$ in the form $(az+b)/(cz+d)$. Then $g_S'(z) = (ad-bc)/(cz+d)^2$ and at fixed points $z$ one finds $cz+d = (a+d)/2 \pm \sqrt{(a-d)^2/2+bc}$. In this instance $ad-bc = 4$, so

$$1/|g_S'(z)| = (1/4)((a+d)/2 \pm \sqrt{(a-d)^2/2+bc}),$$

at the fixed points. Substituting the values of $a, b, c$, and $d$ gives $1/|g_S'(z)| = \tau \pm \sqrt{\tau}$ at the fixed points. The attracting fixed point is the one of interest, and this reciprocal yields our scaling factor $\mu_S = \tau + \sqrt{\tau}$.

**Linear Structure.** The shift for the spiral case is represented by the matrix $M_S$, whose characteristic polynomial is the now familiar expression

$$x^4 - 2x^3 - 2x^2 - 2x + 1.$$

The eigenvalues are the $\lambda_j$ discussed earlier. For each $\lambda_j, j = 1, \cdots, 4$, an associated eigenvector is $\overline{\mathbf{v}}_j = \langle \lambda_j^3, \lambda_j^2, \lambda_j, 1\rangle$. The complex eigenvalues have modulus 1, so the largest eigenvalue is our limit scaling factor $\lambda_1 = \mu_S$. Note that again the first two eigenvectors are quads of bends which, up to scaling, differ only in order. Curiously, the remaining two eigenvectors would also qualify as quads of bends were they not complex — they are zeros of the quadratic form $q$. It is tempting to treat these as quads of circles with complex radii in the setting of inversive geometry.

As in the Ring case, let $T$ denote the linear transformation $(1/\mu_S)M_R$, and express an arbitrary vector as

$$\overline{\mathbf{v}} = c_1\overline{\mathbf{v}}_1 + c_2\overline{\mathbf{v}}_2 + c_3\overline{\mathbf{v}}_3 + c_4\overline{\mathbf{v}}_4.$$

Clearly $\overline{\mathbf{v}}T^n \longrightarrow c_1\overline{\mathbf{v}}_1$ as $n \to \infty$. Thus, as in the Ring case, any vector $\overline{\mathbf{v}}$ not in the span of $\{\overline{\mathbf{v}}_2, \overline{\mathbf{v}}_3, \overline{\mathbf{v}}_4\}$, converges asymptotically under $M_R$ to the quad $\overline{\mathbf{v}}_1$, while vectors in that span have easily understood dynamics.

## 6. Summary: Geometric Sequences

With the principal work now completed, the following is an easy deduction:

**Theorem 9.** *Let $\{d_n\}_{-\infty}^{\infty}$ be a sequence of discs from an A-packing $P$ and assume that the following properties hold: (1) each triple of discs $d_{n-1}, d_n, d_{n+1}$ is part of a quad of $P$ and (2) the associated bends $\{b_n\}$ form a geometric sequence, with $b_{n+1} = \mu b_n, \forall n$, for some $\mu \geq 1$.*

*Then one of these conclusions holds:*

*(i) $\mu = 1$, $\{d_n\}$ is part of a Knife A-packing and is an orbit of a parabolic Möbius transformation.*

*(ii) $\mu = \tau^2$, $\{d_n\}$ is part of a Ring A-packing and is an orbit of a glide-reflection Möbius transformation.*

*(iii)* $\mu = \tau + \sqrt{\tau}$, $\{d_n\}$ *is part of a Spiral A-packing and is an orbit of a loxodromic Möbius transformation.*

There are certainly many additional geometric sequences of bends which occur in A-packings but do not satisfy hypothesis (1); it would be interesting to determine if there are assoicated geometric "shifts" and to find the characteristic ratios $\mu$. (See §9.3.) Incidently, the normalizations of the A-packings of Figure 2 were chosen so that each has its bends in the appropriate geometric sequence.

## 7. The Rodin/Sullivan Ring Lemma

The A-packing $P_R$ was called the Ring pattern because of its connection to the Ring Lemma of Rodin and Sullivan [24]. In this section we explain the connection, prove the sharp values for the Ring Lemma constants, and also establish bounds for related angular resolutions.

An $n$-**flower** is a pattern of discs consisting of a **center**, which we will denote $D_0$, and an ordered chain $\{D_1, \cdots, D_n\}$ of tangent **petals**, cyclically tangent to one another and placed counterclockwise about the center. We assume throughout that all bends are nonnegative and that the flower is **univalent**, meaning that the discs have mutually disjoint interiors (this is not always assumed when flowers are used in circle packing). Note that the union of the petals separates $D_0$ from infinity. Each petal $D_j$ has two **neighbors**, $D_{j-1}$ and $D_{j+1}$, where it is convenient to write $D_m = D_k$ if $m = k$ mod $n$.

It is the relationships among the radii and other quantities in these flowers which concern us. Using geometric arguments, Rodin and Sullivan [24] proved that for each $n \geq 3$ there exists a constant $c(n) > 0$ so that any (univalent) $n$-flower with central disc of radius 1 has its petal radii bounded below by $c(n)$. We prove the following sharp form of the Ring Lemma. Recall that the Fibonacci numbers $a_j, j = 1, 2, \cdots$, are defined by $a_1 = a_2 = 1$ and $a_{j+1} = a_j + a_{j-1}, j = 2, 3, \cdots$, and that $\tau$ denotes the golden ratio.

**Theorem 10.** *Let $F_n$ be a univalent $n$-flower of proper discs with central disc of radius 1, $n \geq 3$. Then the radius of each petal is bounded below by the constant*

$$c(n) = \frac{1}{\left[\left(\frac{5-2\sqrt{5}}{5}\right)\left(\frac{3+\sqrt{5}}{2}\right)^n + \left(\frac{5+2\sqrt{5}}{5}\right)\left(\frac{3-\sqrt{5}}{2}\right)^n - 1\right]}, \quad n \geq 3. \tag{20}$$

*This bound is sharp and for each $n$ is a reciprocal integer. In terms of the Fibonacci numbers,*

$$c(n) = \frac{1}{a_{n-2}^2 + a_{n-1}^2 - 1}, \quad n \geq 3. \tag{21}$$

*The ratio $c(n)/c(n+1)$ is monotone increasing in $n$ and converges to* $\dfrac{3+\sqrt{5}}{2} = \tau^2$.

There a companion result regarding angles which it is convenient to prove in parallel. In an $n$-flower, the **flower angle** $\alpha_j$ will denote the angle formed at the center of $D_0$ by the points of tangency with petals $D_j$ and $D_{j+1}, j = 1, \cdots, n$. See Figure 10.

**Theorem 11.** *Let $F_n$ be a univalent $n$-flower of proper discs, $n \geq 3$. Each flower angle is bounded below by the positive constant*

$$b(n) = 2 \ \arctan\left(\frac{1}{a_{n-2}(a_{n-1} + a_{n-3})}\right), \ n \geq 3. \tag{22}$$

*This bound is sharp. In particular, $\tan(b(n)/2)$ is always a reciprocal integer, and asymptotically*

$$\tan(b(n)/2) \sim \frac{1 + \tau^2}{\tau^{2n+1}}.$$

At the end of this section we point out the interest in these angles from the standpoint of graph embeddings.

7.1. **The Extremal Flowers.** The key to the proofs lies with the extremal flowers, which these problems happen to share. Their form was established by Hansen [11] for the Ring Lemma (though there seems to be a monotonicity argument missing in his work). We find that each extremal is a "subflower" of the A-packing $P_R$ of Figure 2(b), meaning that the sharp constants all occur in that one "$\infty$-flower".

Let's begin by noting that with the knowledge that $c(n)$ is positive, a routine limiting argument proves the existence of an extremal $n$-flower for the Ring Lemma — that is, one with central radius 1 and having a petal of radius $c(n)$. (Simply consider a limiting sequence of flowers and use diagonalization to extract a subsequence.) For each $n$ denote this extremal flower by $\widetilde{F}_n$; we assume that its center is the unit disc and the $n^{\text{th}}$ petal has radius $c(n)$. There is no mystery about the extremal situation for $n = 3$: the first two petals are half planes with the center and the third petal between them, so $r_3 = c(3) = 1$.

We use induction to establish a precise description, hence essential uniqueness, for $\widetilde{F}_n, n = 3, 4, \cdots$. It is convenient and reasonable to set $c(2) = \infty$ and $b(2) = \pi$, since one can argue that a unit disc with two tangent half planes is the extremal (essentially the only?) 2-flower.

**Proposition 12.** *For each $n \geq 3$, the extremal flower $\widetilde{F}_n$ for the Ring Lemma satisfies the following description: It has two petals which are half planes (infinite radius). The remaining petals have radii $c(3), c(4), \cdots, c(n)$, and for each $j, 3 < j \leq n$, the neighbors of the petal of radius $c(j)$ are the petals of radii $c(j-1)$ and $c(j-2)$. In particular, the $c(j)$ are all distinct and $\widetilde{F}_n$ is unique up to similarity.*

*The flower $\widetilde{F}_n$ is simultaneously extremal for the angle constant $b(n)$, with extremal angle formed by the petals of radius $c(n)$ and $c(n-1)$.*

The $c(n)$ are decreasing in $n$, and as we see shortly, it is clear that they are distinct. Taking this as given for the moment, the uniqueness of $\widetilde{F}_n$ becomes automatic. Look to Figure 9. The petal $D_3$ of radius $c(3)$ is necessarily tangent to both half planes since $c(3) = 1$. It forms two interstices with the center, and the description implies that the petal of radius $c(4)$ fills one or the other. Once this petal is in place, however, there are no further options; working inductively through decreasing radii, the description allows only one possible location for each of the remaining petals.

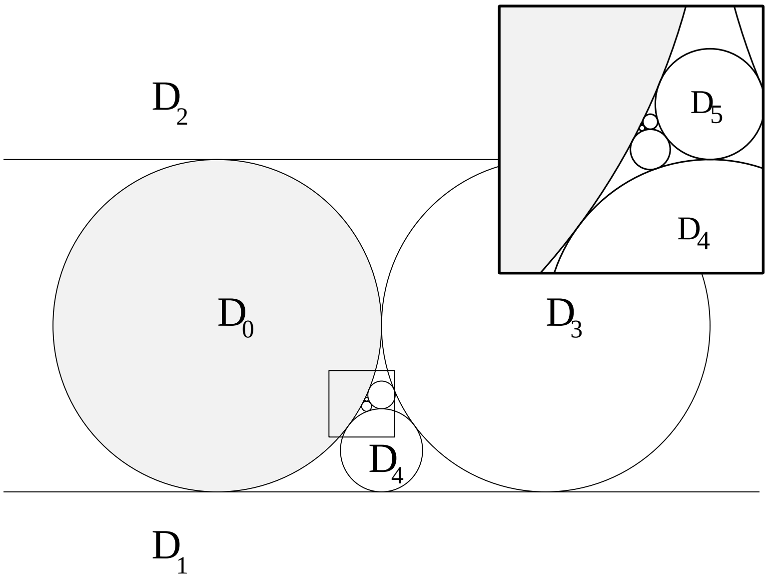


FIGURE 9. The extremal flowers $\widetilde{F}_n$.

*Proof of Proposition:* It is clear that $c(n)$ is strictly decreasing in $n$: filling one of the interstices of $\widetilde{F}_n$ neighboring the petal of radius $c(n)$ gives an $(n+1)$-flower, and the radius of this new petal is clearly less than $c(n)$. (The same reasoning applies to the $b(n)$.)

On closer inspection, we will find that this argument works in reverse. That is, an extremal $(n+1)$-flower is in fact obtained by adding a disc to the smallest interstice of an extremal $n$-flower.

Observe first that an $n$-flower can contain a $j$-flower (with the same center) for some $j < n$. This requires what we will term a **nontrivial tangency** NTT: a petal $D_k$ has a NTT if it is tangent to a nonneighboring petal. Say, for instance, that $D_k$ and $D_m$ share a NTT; by removing the discs between them (those separated from infinity by this tangency), we are left with a proper **subflower**.

Our main computational work lies with the next two "monotonicity" results.

**Lemma 13.** *Let $F$ be a univalent $n$-flower with petals $\{D_1, D_2, \cdots, D_{n-1}, D_n\}$. Suppose that the first and last petals, $D_1$ and $D_n$, have positive bends. Suppose further* *that among all univalent $n$-flowers sharing the intervening $n-2$ petals $\{D_2, \cdots, D_{n-1}\}$ (and center), $F$ has the smallest first petal $D_1$ or $F$ has the smallest flower angle $\alpha_n$. Then either $D_1$ or $D_n$ must have a NTT.*

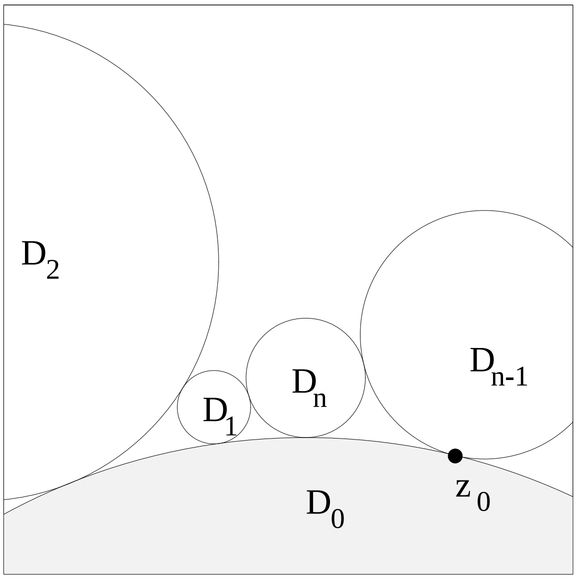


FIGURE 10. The flower in the $z$-plane.

Geometrically this result is rather easy to see. Consider Figure 10. If $D_1$ and $D_n$ were not restrained by NTT's, then one could make $D_1$ smaller by making $D_n$ larger. Imagine increasing $D_n$; its tangency with $D_{n-1}$ forces it to roll towards $D_2$; then $D_1$, squeezed between $D_2$ and $D_n$ must also roll and decrease in size. The effects on $\alpha_n$ are not so intuitive, but not surprising.

*Proof of Lemma 13:* We would like to parametrize all $n$-flowers having the chain of petals $D_2, \cdots, D_{n-1}$. The situation in the $z$-plane is shown in Figure 10. It is best to transform to the $w$-plane using what we will call an "inversion". Namely, we translate an appropriate point $z_0$ to the origin, apply $z \mapsto w = 1/z$, and then rotate the image as convenient. The flower is carried to the $w$-plane, with the image of infinity marked as $w = 0$. We use the same symbols for various quantities in both planes, attaching a prime (') mark in the $w$-plane. "Re-inversion" refers to the inverse mapping.

In the present circumstance, apply an inversion with $z_0$ the point where $D_{n-1}$ is tangent to $D_0$, with the result illustrated in Figure 11. Note that $w = 0$ is necessarily to the left of the vertical line through the center of $D_2'$ and $\sigma$ lies in $(0, 2h]$.

We obtain all possible configurations in the $w$-plane by treating the radius of $D_1'$ as a variable $x$, $0 < x \leq h$. The distances $u$ and $v$ are functions of $x$. The precise domain of $x$ is $[x_1, x_2] \subset (0, 2h]$, where $x_1$ gives $D_n'$ a NTT with some $D_k', 2 \leq k \leq n-2$, and $x_2$ gives $D_1'$ a NTT with some $D_k', 3 \leq k \leq n-1$.

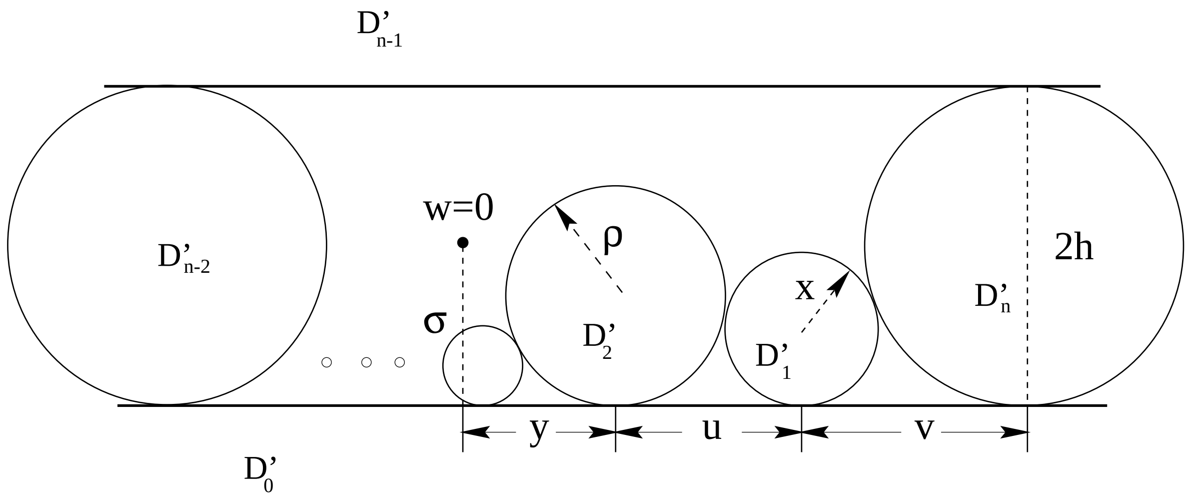


FIGURE 11. Parametrize in the $w$-plane.

Considering radii first, computation shows $u(x) = 2\sqrt{\rho}\sqrt{x}$. If $d(x)$ is the distance from $w = 0$ to the center of $D_1'$, then

$$(d(x))^2 = (\sigma - x)^2 + (y + u(x))^2 = (\sigma^2 + y^2) + 4y\sqrt{\rho}\sqrt{x} + (4\rho - 2\sigma)x + x^2.$$

A computation shows that the radius $r_1 = r_1(x)$ of $D_1$ back in the $z$-plane is $x/(d(x)^2 - x^2)$. It is easiest to work with its bend:

$$\frac{1}{r_1(x)} = \frac{(\sigma^2 + y^2) + 4y\sqrt{\rho}\sqrt{x} + (4\rho - 2\sigma)x}{x} = \frac{(\sigma^2 + y^2)}{x} + \frac{4y\sqrt{\rho}}{\sqrt{x}} + (4\rho - 2\sigma).$$

It is evident that the second derivative with respect to $\sqrt{x}$ is strictly positive for $x > 0$, meaning that the bend has no interior maximum for $x \in [x_1, x_2]$. In other words, $r_1(x)$ achieves its minimum only at an endpoint of this interval, implying that either $D_1$ or $D_n$ has a NTT, as claimed.

Now we turn to angles. We work with $\eta$, shown in Figure 12. In the $z$-plane observe by elementary geometry that $\eta$ is half the flower angle $\alpha_n$ formed by $D_1$ and $D_n$. This is equal to the angle marked $\eta$ in the $w$-plane because the angle (at infinity) is preserved under the inversion. This and the auxiliary angles $\phi$ and $\psi$ in the $w$-plane are functions of the variable $x$ used earlier. We have

$$u(x) = 2\sqrt{\rho}\sqrt{x}, \qquad v(x) = 2\sqrt{h}\sqrt{x},$$
$$\cot\phi(x) = (y + 2\sqrt{\rho}\sqrt{x})/\sigma, \qquad \cot\psi(x) = (y + 2\sqrt{\rho}\sqrt{x} + 2\sqrt{h}\sqrt{x})/\sigma.$$

Then

$$\begin{aligned}\cot\eta(x) &= \frac{\cot\phi(x)\cot\psi(x)}{\cot\psi - \cot\phi}\\ &= \frac{(\sigma^2+y^2)+(4y\sqrt{\rho}+2y\sqrt{h})\sqrt{x}+(4\rho+r\sqrt{\rho}\sqrt{h})x}{2\sigma\sqrt{h}\sqrt{x}}\\ &= (2\sigma\sqrt{h})^{-1}\big[\frac{(\sigma^2+y^2)}{\sqrt{x}}+(4y\sqrt{\rho}+2y\sqrt{h})+(4\rho+4\sqrt{\rho}\sqrt{h})\sqrt{x}\big].\end{aligned}$$

Again it is evident that the second derivative is strictly positive for $x > 0$ and $\cot\eta(x)$

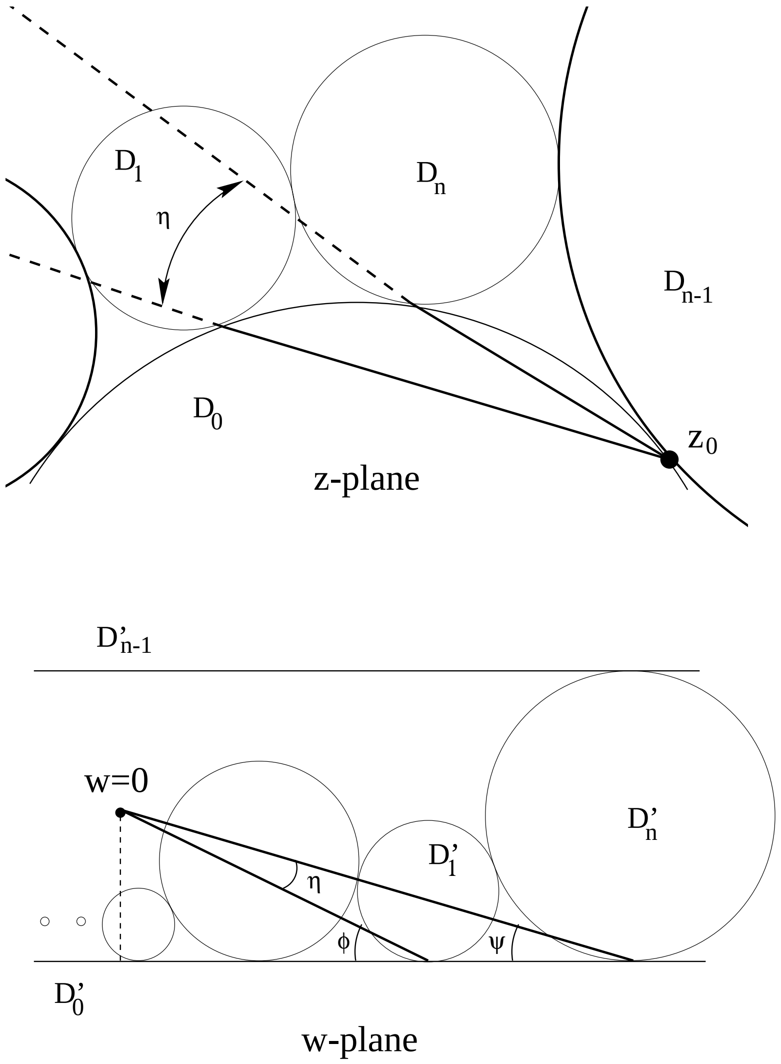


FIGURE 12. Angles in the $z$ and $w$-planes.

has no interior maximum. Thus $\eta(x)$ itself has no interior minimum, so its minimum occurs only when $D_1$ or $D_n$ has a NTT. This completes the proof of Lemma 13. □

Our second monotonicity result involves one further petal, namely a disc $D_{n+1}$ which fills the interstice formed by $D_1$ and $D_n$.

**Lemma 14.** *Let $F$ be a univalent $(n+1)$-flower with petals $\{D_1, D_2, \cdots, D_n, D_{n+1}\}$. Suppose that the petals $D_1$ and $D_n$ have positive bends and are tangent to one another (so $D_{n+1}$ fills their interstice). Suppose further that among all univalent $(n+1)$-flowers with these properties and sharing the intervening $n-2$ petals $\{D_2, \cdots, D_{n-1}\}$ (and center), $F$ has the smallest petal $D_{n+1}$. Then either $D_1$ or $D_n$ must have a NTT in addition to the one they have in common.*

*Proof of Lemma 14:* Because of the hypothesis that $D_1$ and $D_n$ are tangent (a non-trivial tangency in this augmented situation), we can simply update the earlier pictures by adding the last petal in the interstice between $D_1$ and $D_n$ (resp. $D_1', D_n'$). (We won't need to consider angles here.) The computations are identical in spirit to those carried out in the previous lemma; there are no new hurdles (aside from a couple additional variables) and no new insights, so the details are left to the reader. Recall that as $x$ is varied, the disc $D_{n+1}'$ remains nestled in the interstice between $D_1'$ and $D_n'$. □

We may proceed with the Proposition. Consider the radius problem and use induction. The case $n = 3$ is clear. Assume the conclusion for cases $3, \cdots, n$. If $F$ is an extremal $(n+1)$-flower, it is enough to show that $F$ is obtained from $\widetilde{F}_n$ by adding a disc to its smallest interstice, which by Lemma 2 is formed by discs of radius $c(n)$ and $c(n-1)$. For reference purposes, let $\tilde{r}$ be the radius of such a disc.

Assume that $r_{n+1}$ is the smallest petal radius of $F$ and assume without loss of generality that $r_1 \geq r_n$. We will show the inequalities $r_n \geq c(n)$ and $r_1 \geq c(n-1)$. In that case, Lemma 2 will imply $r_{n+1} \geq \tilde{r}$, and the extremal nature of $F$ will require equality. This in turn will imply equality in both inequalities. In particular, $r_n = c(n)$, so our induction hypothesis will give $F \backslash D_{n+1} = \widetilde{F}_n$ and we will be done with the radius problem.

Disc $D_{n+1}$ cannot lie in any proper subflower, since that would give $r_{n+1} \geq c(n)$. Therefore, Lemma 13 implies that $D_n$ has an NTT. This means that $D_n$ does lie in a subflower, giving $r_n \geq c(n)$. Likewise, we see by Lemma 13 that $D_1$ has a NTT. If this is not a tangency with $D_n$, then $D_1$ would lie in a proper subflower which, since it must omit $D_{n+1}$, must also omit $D_n$. In other words, $D_1$ would lie in a $j$-flower for some $j < n$, implying $r_1 \geq c(n-1)$, and we would be done. This leaves only the possibility that $D_1$ and $D_n$ are tangent to one another. However, an application of Lemma 14 tells us that $D_1$ or $D_n$ has yet another NTT. We may assume this involves the larger of the two discs, so again $r_1 \geq c(n-1)$. Therefore, we have shown the inequalities we claimed.

For the flower angle extremals, we simply note that by the law of cosines, $\alpha_j$ is strictly monotone increasing in its petal radii $r_j$ and $r_{j+1}$. Let $F$ be an $n$-flower with extremal flower angle $\alpha_n$. Then $r_1, r_n \geq c(n)$. By Lemma 13, one of these petals $D_1$ or $D_n$ has a NTT. Say it is $D_1$, then $r_1 \geq c(n-1)$. By the noted strict monotonicity of angles, we must then have $r_n = c(n)$, so $F$ is the extremal $\widetilde{F}_n$ for the radius problem, as we needed to prove. $\square$

7.2. **Proofs of the Theorems.** We know from the Proposition that the flowers $\widetilde{F}_n$ are extremals for radii and flower angles. Observe that $\widetilde{F}_n$ is part of our A-packing $P_R$, as normalized in Figure 2(b). Indeed, $\widetilde{F}_3$ is the quad with bends $\langle\langle 1,0,0,1\rangle\rangle$, with the center acting as the guard disc. Each successive quad shift $S_R$ produces the smallest petal of the next extremal flower (the action which originally drew our interest to quad shifts).

The constants $c(n)$ may now be obtained from Lemma 7, using the $\widetilde{F}_3$ as the initial quad. This gives $g = b_3 = 1, b_1 = b_2 = 0$, so (20) follows from (9). It is well known that via difference equations the terms of the Fibonacci sequence are computed as $a_n = \frac{\tau}{2+\tau}\left[\tau^n - (-1/\tau)^n\right]$. An easy computation gives (21). This, of course, explains the reciprocal integer nature of the constants; also, the asymptotic limit of $c(n)/c(n+1)$ is precisely the scaling factor $\mu_R$ associated with the A-packing $P_1$. This completes Theorem 10. $\square$

As for the constants $b(n)$ in the angle problem, the smallest flower angle $\alpha$ is formed by the petals of radius $c(n)$ and $c(n-1)$ in $\widetilde{F}_n$. If $\rho$ is as shown in Figure 13, then $\cot(\alpha_n/2) = 1/\rho$. By the law of cosines,

$$\cos(\alpha) = \frac{(1+r_n)^2 + (1+r_{n-1})^2 - (r_n + r_{n-1})^2}{2(1+r_n)(1+r_{n-1})} = \frac{1 - \dfrac{r_n r_{n-1}}{1+r_n+r_{n-1}}}{1 + \dfrac{r_n r_{n-1}}{1+r_n+r_{n-1}}} = \frac{\cot^2(\alpha/2) - 1}{\cot^2(\alpha/2) + 1}.$$

In particular,

$$\cot(\alpha/2) = \sqrt{\frac{1}{r_n} + \frac{1}{r_{n-1}} + \frac{1}{r_n r_{n-1}}} = \sqrt{b_n + b_{n-1} + b_n b_{n-1}}.$$

Recall that the guard disc in this setting has bend $g = 1$. A further disc placed in the interstice of Figure 13 would, by (3), have bend $b_{n+1} = 1 + b_n + b_{n-1} + 2\sqrt{b_n + b_{n-1} + b_n b_{n-1}}$. Substituting from above and solving, we get

$$\cot(\alpha/2) = \frac{b_{n+1} - b_n - b_{n-1} - 1}{2}. \tag{23}$$

Return now to the setting of Lemma 7. We have started with the quad $\widetilde{F}_3$, so $g = b_3 = 1, b_1 = b_2 = 0$. Recall our notation $y(n) = b_n + g = b_n + 1$ and note that because the coefficient $\gamma$ turns out to be zero, $y(n)$ satisfies the recurrence relation $y(n+1) = 3y(n) - y(n-1)$. That is, the bends satisfy $b_{n+1} = 3b_n - b_{n-1} + 1$. Putting

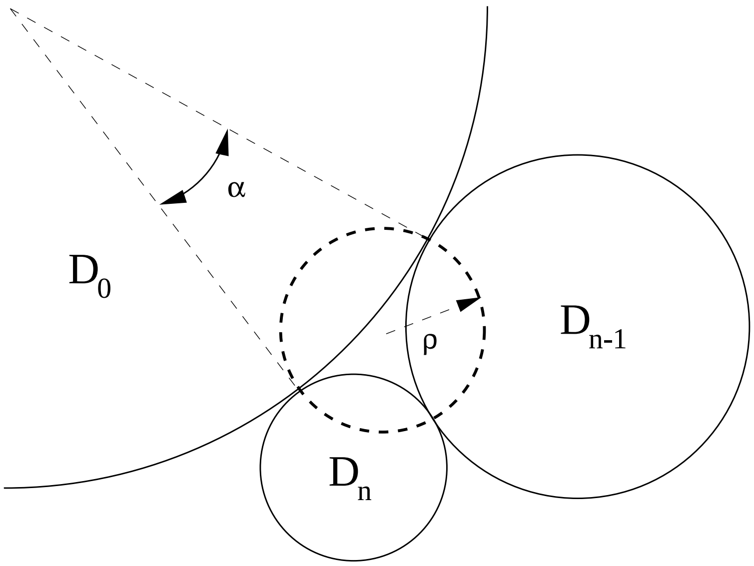


FIGURE 13. The extremal flower angle.

this in (23) and entering the known values of these bends from (21), we obtain

$$\begin{aligned}\cot(\alpha/2) &= b_n - b_{n-1} = (a_{n-2}^2 + a_{n-1}^2 - 1) - (a_{n-3}^2 + a_{n-2}^2 - 1) \\ &= a_{n-1}^2 - a_{n-3}^2 = a_{n-2}(a_{n-1} + a_{n-3}).\end{aligned}$$

This implies the values in (22). The asymptotic estimate follows from the expressions given earlier for the Fibonacci numbers, completing the proof of Theorem 11. □

7.3. **Angular Resolution.** Circle packing has found uses in graph embedding. The reader is referred to [29] for the fundamentals of circle packing and to [21] and [3] for some discussion of computational efficiency. The issue of angular resolution is raised in [16]; it has implications for various uses of embeddings, e.g., effects on rates of convergence for finite element computations (*cf.* [20]).

Briefly, a planar graph $K$ may be embedded in the plane so that it has straight line edges by constructing a disc packing (a circle packing whose circles have mutually disjoint interiors) with the combinatorics of $K$; the centers of the discs and the line segments connecting centers of tangent discs provide the embedded image. Such embeddings have many nice features, as described in [16]: for instance, they simultaneously yield a straight-line embedding of the dual graph whose edges are orthogonal to those of the embedded $K$ and whose faces are convex. (Note: we assume without loss of generality that $K$ is a triangulated graph.)

Our interest is in giving lower bounds on angular resolution, see [16]. The **angular resolution** of a straight-line embedding of a planar graph $K$, is the minimum angle formed by incident edges, and the **angular resolution** of $K$ is the maximum of these angular resolutions over all possible straight-line embeddings of $K$. In a disc packing, the angles of the embedding are flower angles. Thus, if $n$ is the maximal degree of the vertices, then $b(n)$ is a lower bound on the resolution. By Theorem 11, we have

**Theorem 15.** *The resolution $res(K)$ of a planar graph $K$ of degree $n$ satisfies*

$$res(K) \geq 2 \ \arctan(\frac{2+\tau}{\tau^{2n-1}}).$$

*In particular, $res(K) = O(\tau^{2n}) \sim O(.381966^n)$ as $n$ goes to infinity.*

Whether this is sharp among other types of straight-line embeddings is unknown.

## 8. Numerical Patterns in Quads

Configurations of discs having integer bends were apparently first observed by Soddy, who built (quite literally!) his "bowl of integers" almost 60 years ago [25]. He concluded the paper with this hope: "Possibly when a mental arithmetic of reciprocals becomes as natural as that of number we may hope to learn the real secret of the circle and the sphere." The plethora of integers and subtle patterns in the bends of our A-packings would certainly be pleasing to Soddy. We point out some features of these patterns here, first in the plane and then in $\mathbf{R}^3$.

8.1. **The Apollonian Net.** Let $\mathcal{B} \subset \mathbf{R}^4$ denote the set of all quads of bends having integer entries. Lemma 5 informed us that $\mathcal{B}$ is closed under our shifts and their inverses. Let us consider $\mathcal{A}$, built from the quad $\mathfrak{Q}$ of Figure 4, whose bends are $\langle\!\langle -2, 3, 6, 7 \rangle\!\rangle$. Write $b_1, b_2, \cdots$ for the totality of positive bends, ordered by increasing size, and $r_1, r_2, \cdots$ for the corresponding radii.

What can one say about these bends? Which integers occur, in what order, how redundant is the sequence, how redundant are particular quads of bends, and so forth? Note that there is only one obvious symmetry in $\mathcal{A}$, reflection in the $x$-axis. One's first impression is that there are several gaps in this sequence — there is clearly no room for discs of bend 1, 4, or 5 from the beginning, and it seems that at subsequent stages various sizes would be excluded. Support is provided by the fact that the discs of positive bend fill the complement of $D_1$, giving

$$\sum_{j=1}^{\infty} r_j^2 = \sum_{j=1}^{\infty} (1/b_j)^2 = 1/4, \tag{24}$$

considerably less than $\sum_{j=3}^{\infty} 1/j^2 \approx .394934$.

On the other hand, The sum of (24) converges to a rational number; a classical theorem of Liouville implies a slow rate of convergence, suggesting a slow rate of growth for $b_j$ and consequent repeats. It turns out, in fact, that the sequence of bends is extremely repetitious. The proof of this goes back to the work of Hirst, Boyd, and others on what is called the **residual set** of $\mathcal{A}$, or the **Apollonian net** in [17].

This turns out to be a topic of considerable interest, so a bit of background is in order. The Apollonian net, denoted $\Lambda$, is the complement in $\mathbf{S}^2$ of the interiors of the discs of $\mathcal{A}$. It may be described in several other ways, as the closure of the union of boundary circles of $\mathcal{A}$, as the accumulation points of the centers of the discs of $\mathcal{A}$, or as the limit set of the Kleinian group $\Gamma$ which generates $\mathcal{A}$ (see §2.4). This last description has driven much of the interest. It has been conjectured, for instance, that this "osculating" packing has a residual set with the minimum Hausdorff dimension among all disc packings. (See [5], [15]) It is known that the dimension of $\Lambda$ and the exponent of convergence of the sequence of reciprocal radii are identical. In particular, both these quantities are conformal invariants. They are not only independent of the particular Apollonian packing considered, but moreover are consequently locally determined at each point of the residual set — the Apollonian net is homogeneous.

Let $\mathfrak{d}$ denote the Hausdorff dimension of $\Lambda$. Numerical estimates of $\mathfrak{d}$ rely on methods of Boyd [5]. The best estimate we are aware of is this, from [18]:

$$\mathfrak{d} \approx 1.305686729.$$

Choose $\sigma$ so that $1 < \sigma < \mathfrak{d}$. Then the $\sigma$-Hausdorff measure of $\Lambda$ is infinite, implying

$$\sum_j (r_j)^\sigma = \sum_j (\frac{1}{b_j})^\sigma = \infty.$$

On the other hand, it is well known that the Riemann Zeta function

$$\zeta(z) = \sum_j (\frac{1}{j})^z$$

converges for $z = \sigma > 1$. Comparing these shows that the sequence $\{b_j\}$ is highly repetitive.

The estimates of $\mathfrak{d}$ essentially reduce to estimates of the "shapes" of interstices of triples in $\mathcal{A}$. The distribution of shapes is then approached probabilistically using random products of matrices (e.g., of our shifts $M_\bullet$) and a result of Furstenburg and Kesten [9] on the asymptotic behavior of such random products. We cannot go into further detail here. However, note that the distribution of shapes is reflected not so much in the individual *radii* (or bends) of the discs of $\mathcal{A}$ but rather in the relative proportions occurring among the *quads* of $\mathcal{A}$. This suggests a switch to the homogeneous setting and a study of random iterations of the maps $\Phi_\bullet$ defined in §2.6. A better understanding of these maps may improve our understanding of $\Lambda$.

8.2. **Packings of Spheres.** One can find even richer patterns of integer bends in $\mathbf{R}^3$. Our focus remains firmly on **quads** of spheres, however; we will not be moving to "quints", the usual generalization of the DCT (see [6] and references therein).

Soddy first encountered reciprocal integer bends within packings of spheres built using "hexlets". When he cut his "bowl of integers" (the 3-dimension physical model is pictured in [25]) with a horizontal plane, he found a pattern of discs with integer bends which is not Apollonian, but rather is the inversion of a hexagonal packing ("penny packing") in the plane in one of its circles. However, it appears that if one cuts his model with a **vertical** plane instead, one finds the Apollonian packing!

The hexlet is another invention (and another poem!) of Soddy. A **hexlet** $H$ consists of a chain of six cyclically tangent spheres, each tangent to all three of a fixed triple of (mutually tangent) spheres. As Soddy puts it in "The Hexlet" [26], the six form a necklet which "Caresses not one but three necks," since it encircles each sphere of the triple.

The easiest hexlet to picture is this standard one, denoted $\mathcal{H}$: Place seven balls of unit radius on a table to form a 6-flower and set another plane on top. The two planes and the center ball form the triple, the six petal balls, the necklet. All other hexlets are easily shown to be Möbius images of $\mathcal{H}$ in $\mathbf{R}^3$.

It is much more interesting, however, to start with an arbitrary triple $\langle\langle S_1, S_2, S_3\rangle\rangle$. Add a fourth sphere, say $N_1$, tangent to all three to form the quad $\langle\langle S_1, S_2, S_3, N_1\rangle\rangle$. There are exactly two spheres simultaneously tangent to all four of these; choose one of them as $N_2$, giving the next quad $\langle\langle S_1, S_2, S_3, N_2\rangle\rangle$. In other words, we have the "hexlet" shift

$$S_H : \langle\langle s_1, s_2, s_3, N_j\rangle\rangle \mapsto \langle\langle s_1, s_2, s_3, N_{j=1}\rangle\rangle. \tag{25}$$

Repeating this process with the new quad, there is only one choice for $N_3$ (since neighbor $N_1$ is already in place). Likewise, we successively choose $N_4, N_5$, and $N_6$. However, an interesting thing happens (in a phenomenon reminiscent of the Doyle construction arguments of [2]): The sphere $N_6$ automatically lands tangent to $N_1$. That is, the construction invariably leads to a hexlet. Put in a different way, the shift $S_H$ of (25) has order six; it is implemented by a elliptic Möbius transformation (of $\mathbf{R}^3$).

Note that each of the six necklet spheres $N_j$ has an "opposite" $N_k$, referring to the sphere three steps around ($k = j + 3 \mod 6$). If we denote the bends of the triple by $b_1, b_2, b_3$, (as in $\mathbf{R}^2$, these are the reciprocals of the signed radii), then the last line of Soddy's poem states this

**Lemma 16.** *In a hexlet $\langle N_1, N_2, \cdots, N_6\rangle$ caressing the triple $\langle\langle S_1, S_2, S_3\rangle\rangle$, the bends $b$ and $b'$ of any two opposite spheres satisfy*

$$b + b' = 2(b_1 + b_2 + b_3). \tag{26}$$

This is evidently true for $\mathcal{H}$, since $1 + 1 = 2(0 + 1 + 0)$. Morley [22] uses this fact and the invariance of relation (26) under inversion to prove the general case.

We now have a notion of **quad shift** suitable for $\mathbf{R}^3$: start with quad $\langle\langle s_1, s_2, s_3, s_4\rangle\rangle$, treat the first three, say, as the triple and $s_4$ as a bead on the necklet, shift to the quad $\langle\langle s_1, s_2, s_3, s_5\rangle\rangle$ where $s_5$ is the bead opposite to $s_4$. The relation (4) holds among their bends. (Note that we get analogues of our earlier shifts simply by dictating the order of entries in the quads.)

It is important to note that any quad $\langle\langle s_1, s_2, s_3, s_4\rangle\rangle$ of spheres has a unique common orthogonal sphere $S$ containing the six tangency points among $s_1, s_2, s_3, s_4$. Consequently, $S$ must also be the common orthogonal to $\langle\langle s_1, s_2, s_3, s_5\rangle\rangle$ and contain its six points of tangency. These facts were observed by Gosset [10], who also computed the bend of $S$. Suppose $\langle\langle s_1, s_2, s_3, s_4\rangle\rangle$ is an arbitrary quad of spheres with bends $b_j, j = 1, \cdots, 4$. Then the bend $b$ of the common orthogonal sphere satisfies

$$2b^2 = (b_1b_2 + b_1b_3 + b_1b_4 + b_2b_3 + b_2b_4 + b_3b_4) - (\tfrac{1}{2})(b_1^2 + b_2^2 + b_3^2 + b_4^2). \tag{27}$$

The right side is precisely the Descartes form $q$. (Note the consistency with the planar case, wherein $b = 0$.)

We are now in position for a recursive construction process. Start with an arbitrary quad of spheres $\langle\langle s_1, s_2, s_3, s_4\rangle\rangle$, $S$ its common orthogonal. Any quad shift yields a new quad and a new sphere $s_5$ orthogonal to $S$. Continue this process *ad infinitum* — carrying out every possible shift of every quad in this growing collection. *The resulting infinite collection of spheres $\{s_j\}$ will all be orthogonal to $S$ and intersect it in a collection of discs which forms an Apollonian packing on $S$.* For a proof, one need only apply a Möbius map to the planar case.

As for integer bends, the identity (26) ensures once more that an initial quad with integer bends leads to all integer bends. Now, however, we have a much wider range of initial quads. For instance, starting with three unit spheres, a quad can be formed using a fourth sphere having any bend less than $3 + 2\sqrt{3}$. So $\langle\langle 1, 1, 1, n\rangle\rangle$ is an integer quad of space bends for $n = 1, 2, \cdots, 6$. These additional patterns of bends may be useful, since, for example, the Hausdorff dimension of the Apollonian net on $S$, being invariant under Möbius transformations, is precisely the value $\mathfrak{d}$ from the original planar case.

8.3. **Farey Numbers.** Let us return to the plane. Figure 14 displays discs from an Apollonian packing which includes the lower half-plane $D_0$. The other discs lie in the upper half-plane $\mathcal{H}$, which may be treated as the hyperbolic plane. We will be concerned only with those discs of the packing which are tangent to $D_0$, termed **horocycles** in $\mathcal{H}$. Denote this collection of these by $\mathcal{D}$. It includes $D_\infty = \{x + iy : y \geq 1\}$, a horocycle tangent to $D_0$ at infinity.

We will focus on the interval $[0, 1]$, and in particular on the tangency points with $D_0$ of the horocycles. Indeed, it is convenient to label each horocycle by its tangency point. In this way, we find what are termed **Farey** numbers, subject to **Farey arithmetic**. Specifically, begin by labeling the horocycles $0 = \frac{0}{1}$ and $1 = \frac{1}{1}$. Their "daughter" (the horocycle in their interstice) then has the label $\frac{0}{1} + \frac{1}{1} = \frac{1}{2}$. Say we next consider the daughter of $\frac{1}{2}$ and $\frac{1}{1}$: that would be the horocycle $\frac{2}{3}$ in Figure 14. The daughter of $\frac{0}{1}$ and $\frac{1}{2}$ is $\frac{1}{3}$. That of $\frac{1}{3}$ and $\frac{1}{2}$ is $\frac{2}{5}$. Continuing in this way one successively labels all the horocycles with tangencies in $[0, 1]$.

The following relationships emerge. If horocycles $\frac{p_1}{q_1}$ and $\frac{p_2}{q_2}$ are tangent to one another, then (a) $p_1q_2 - p_2q_1 = \pm 1$, and (b) the daughter is $\frac{p_1+p_2}{q_1+q_2}$. This is a very appealing way to add fractions! But in this case the results are even correct; that is, the tangency point of the horocycle labeled $\frac{p_1+p_2}{q_1+q_2}$ is in fact the rational number $(p_1 + p_2)/(q_1 + q_2)$, which equals $\frac{p_1}{q_1} + \frac{p_2}{q_2}$.

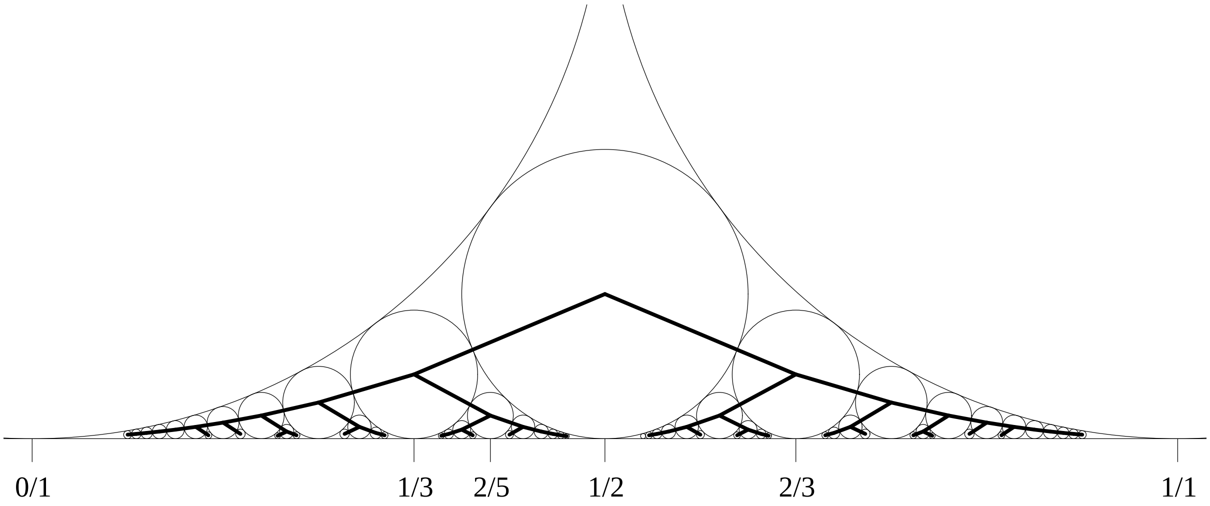


FIGURE 14. The Farey tree.

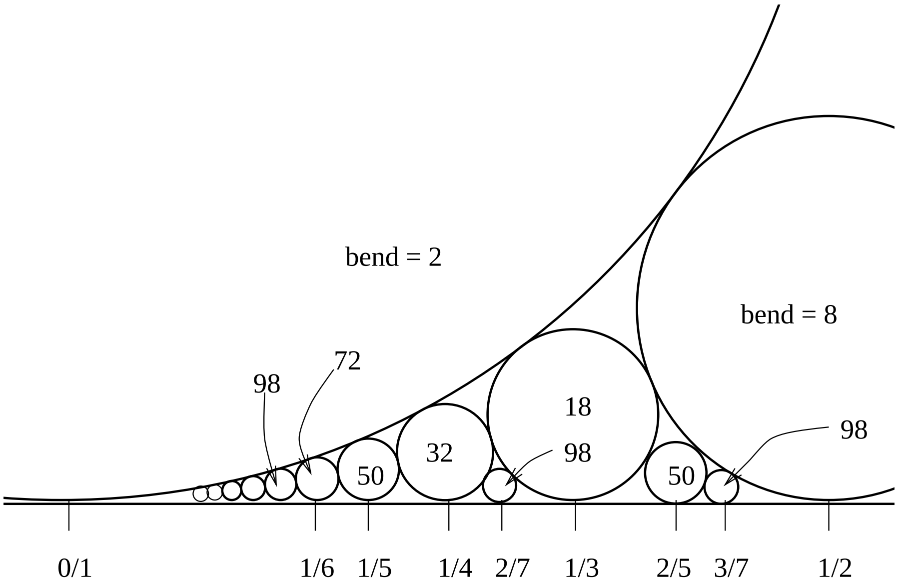


FIGURE 15. Farey numbers.

In Figure 15 we have enlarged the segment $[0, 1/2]$ to show some tangency points and the bends of the horocycles. In fact, *every* rational is a tangency point and thus each has a Farey representation. The recursive way Farey numbers are defined and connections to continued fractions provide a useful hierarchy within the rationals; see [7] for uses in the iteration of circle maps and for an entry into the extensive literature.

Now a few observations about the pattern of bends. All the bends are integers, with the bend of horocycle $\frac{p}{q}$ equal to $2q^2$. By (6), if horocycles $\frac{p_1}{q_1}$ and $\frac{p_2}{q_2}$ are tangent then $|p_1/q_1 - p_2/q_2| = 1/q_1q_2$. Connecting the centers of tangent discs as in Figure 14 gives a pleasing embedding of the so-called **Farey tree**: there's a unique node over each rational; the nodes for a *Farey series* $\mathcal{F}_q$ (irreducible rationals in $(0,1)$ with denominators $\leq q$) are those with imaginary parts $\geq 2q^2$; the angles in this straight-line embedding are bounded below according to Theorem 11; and the length of the edge between the nodes for $\frac{p_1}{q_1}$ and $\frac{p_2}{q_2}$ is $1/2q_1^2 + 1/2q_2^2$.

The discs of $\mathcal{D}$ have a special place in geometry and number theory due to connections with the classical Modular Group. The orientation preserving Möbius transformations preserving $\mathcal{H}$ (the hyperbolic isometries of $\mathcal{H}$) form the subgroup $\mathrm{SL}(2,\mathbf{Z})$ of $\mathrm{SL}(2,\mathbf{C})$. The collection $\mathcal{D}$ is an orbit of discs under $\mathrm{SL}(2,\mathbf{Z})$:

$$\mathcal{D} = \{D = \gamma(D_\infty) : \gamma \in \mathrm{SL}(2,\mathbf{Z})\}.$$

The group $\mathrm{SL}(2,\mathbf{Z})$ is a central object of study in number theory, and this orbit has been particularly important (see, e.g., [23]).

## 9. Descartes Form

The Descartes quadratic form $q$ has been hovering in the background and deserves some attention on its own. In fact, there is quite a bit going on here. We have seen, of course, that $\overline{\mathbf{v}} \in \mathbf{R}^4$ satisfies $q(\overline{\mathbf{v}}) = 0$ iff $\pm\overline{\mathbf{v}}$ consists of bends for some quad of discs. Moreover, our shifts $M_\bullet$ preserve the zero set of $q$. It is easily verified that $q$ is actually **invariant** under the shifts $M_\bullet$, meaning that $q(\overline{\mathbf{v}}M_\bullet) = q(\overline{\mathbf{v}})$. The reader will also have noted that $q$ arose in our consideration of quads of spheres. Thus, even when starting with a "non-quad" vector of bends, the group $G$ of shift and permutation matrices generates the vectors associated with quads, but now these are quads of spheres in an Apollonian packing of some sphere in $\mathbf{R}^3$. In this final section we can take only a brief look at some of the issues relating to $q$. (The authors find that several of these ideas have already been noted by Bo Söderberg in the very nice paper [28].)

**The Lorentz Form.** The matrix representing $q$ is

$$F = (\frac{1}{2}) \begin{bmatrix} 1 & -1 & -1 & -1 \\ -1 & 1 & -1 & -1 \\ -1 & -1 & 1 & -1 \\ -1 & -1 & -1 & 1 \end{bmatrix},$$

so $q(\overline{\mathbf{v}}) = \overline{\mathbf{v}} \cdot F \cdot \overline{\mathbf{v}}^T, \overline{\mathbf{v}} \in \mathbf{R}^4$. A computation shows that $F$ has eigenvalues $1, 1, 1$, and $-1$ and four independent eigenvectors. It is diagonalized by this orthogonal matrix $P$, giving the matrix $L$:

$$P = (\tfrac{1}{2}) \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ -1 & 1 & -1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}, \qquad L = PFP^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}.$$

In particular, $q$ is equivalent to the quadratic form $q_L$ associated to $L$, with $q_L(\overline{\mathbf{u}}) = q(\overline{\mathbf{u}}P)$. The form $q_L$ is the well-known **Lorentz form**; it is nondegenerate, with $\mathrm{rank}(q_L) = 4$ and $\mathrm{sign}(q_L) = 2$ and is associated with Minkowski space. In the traditional space variables $x_1, x_2, x_3$ and time variable $t$, $q_L(x_1, x_2, x_3, t) = x_1^2 + x_2^2 + x_3^2 - t^2$. To distinguish our two settings, use $\mathcal{E}$ for the euclidean 4-space on which $q$ lives, $\mathcal{L}$ for the Minkowski 4-space where $q_L$ lives. Let $T : \mathcal{E} \longrightarrow \mathcal{L}$ denote the linear isomorphism given by $T : \overline{\mathbf{v}} \mapsto \overline{\mathbf{v}}P^{-1}$; so $q_L(\cdot) = q(T(\cdot))$. If $G$ denotes the multiplicative group of linear transformations of $\mathcal{E}$ generated by $M_S, M_R, M_K$ and the permutation matrices (as in §2.5), let $G_L = TGT^{-1}$ be the isomorphic group acting on $\mathcal{L}$.

9.1. **Light Cone.** The zero set of $q_L$ is the "light cone" $Z_L = \{x : x_1^2 + x_2^2 + x_3^2 = t^2\}$. Let us denote by $Z_L^+$ the forward light cone, $Z_L^+ = Z_L \cap \{t > 0\}$, and by $Z^+$ the corresponding points $T^{-1}(Z_L^+) \subset \mathcal{E}$. Note that $\overline{\mathbf{v}} = \langle a, b, c, d\rangle$ in $Z^+$ implies $a + b + c + d > 0$. From this one easily confirms

**Lemma 17.** *The quads of bends $\overline{\mathbf{b}} = \langle\langle b_1, b_2, b_3, b_4\rangle\rangle$ are in one-to-one correspondence under $T$ with the forward light cone $Z_L^+$ in Minkowski space.*

More generally, the forward Lorentz cone $C_L^+ = \{x : x_1^2 + x_2^2 + x_3^2 - t^2 \geq 0, t > 0\}$ is foliated by the nonnegative level sets of $q_L$. Thus $C^+ = T^{-1}(C_L^+)$ is foliated by the level sets of $q$, a foliation preserved by the shifts $M_\bullet$. By (27), each level set $q(\overline{\mathbf{v}}) = c$ represents the bends for quads of spheres which have a mutual orthogonal sphere with bend $\sqrt{c/2}$.

Let us turn now to integral bends. Note that due to the form of $P$, $\overline{\mathbf{v}} \in \mathcal{E}$ will have integer entries iff $T(\overline{\mathbf{v}}) \in \mathcal{L}$ has half-integer entries. We will nonetheless refer to these as "integral", since we could always scale by 2.

First, note the huge variety of integral points on the forward light cone; namely, the points $T(\mathcal{B})$. ($\mathcal{B} \subset \mathcal{E}$ is the set of integral quads, §8.) Furthermore, within $T(\mathcal{B})$ are infinitely many mutually disjoint integral orbits. If $\overline{\mathbf{b}}$ is an integral quad, then $G_L(T(\overline{\mathbf{b}}))$ is an integral orbit in $Z_L^+$. Denoting this orbit by $h(\overline{\mathbf{b}})$, note that $h(\overline{\mathbf{b}}_1) = h(\overline{\mathbf{b}}_2)$ iff $\overline{\mathbf{b}}_1 \in G(\overline{\mathbf{b}}_2)$. Modulo this equivalence relation, we find infinitely many disjoint orbits — those associated with $\overline{\mathbf{b}} = \langle\langle 0, 0, n, n\rangle\rangle$, for example. If we further identify those which are scalar multiples of one another, do we still have infinitely many? There are, of course, patterns of integral orbits within certain level curves of $q_L$ in the Lorentz cone $C^+$. The number is at most countable, of course. Which level curves are these? Will some have multiple orbits? There are many open questions, and whether tools are available to answer them, we don't know.

**The Lorentz Group.** Important to understanding $q_L$ is the associated orthogonal group, the so-called Lorentz group, denoted $\mathrm{O}(1,3)$; this is the subgroup of $\mathrm{GL}(4,\mathbf{R})$ consisting of matrices which leave $q_L$ invariant: that is, $M \in \mathrm{O}(1,3)$ if and only if $MLM^{\mathrm{t}} = L$ or equivalently $q_L(\overline{\mathbf{v}}) = q_L(\overline{\mathbf{v}}M), \overline{\mathbf{v}} \in \mathcal{L}$. Let us mention a few facts about $\mathrm{O}(1,3)$. Because $L^2 = I$, the following implications show that $\mathrm{O}(1,3)$ is closed under transposition:

$$\begin{aligned} B \in \mathrm{O}(1,3) &\Longleftrightarrow BLB^{\mathrm{t}} = L \Longleftrightarrow BLB^{\mathrm{t}}LB = B \\ &\Longleftrightarrow B^{\mathrm{t}}LB = (BL)^{-1}B = L^{-1} = L \Longleftrightarrow B \in \mathrm{O}(1,3). \end{aligned}$$

Every element of $\mathrm{O}(1,3)$ may be expressed as a product of at most 4 "reflections" (relative to the form $q_L$). As a continuous (closed) subgroup of $\mathrm{GL}(4,\mathbf{R})$, the Lorentz group is a real Lie group, and its dimension is 6. It is known that $\mathrm{O}(1,3)$ has four connected components. If $H$ denotes the component containing the identity, then $H$ acts on $C_L^+$; it acts transitively on $Z_L^+$ and more generally, on each nonnegative level set of $q_L$.

We may now carry properties back to the Descartes form $q$. Its orthogonal group $\mathrm{O}(q) \subset \mathrm{GL}(4,\mathbf{R})$ consists of matrices $M$ such that $MFM^{\mathrm{t}} = F$. Since $P$ is orthogonal, one easily verifies that $P(\mathrm{O}(q))P^{-1} = \mathrm{O}(1,3)$. Our group $G$, generated by quad shifts and permutations, is a discrete subgroup of $P^{-1}HP \subset \mathrm{O}(q)$. Because $\mathrm{O}(1,3)$ is closed under transposition, the same holds for $\mathrm{O}(q)$. This helps explain a curiosity we encountered: namely, if one carelessly multiplies a quad by a shift on the wrong side, the result is still a quad. Indeed, one can verify

$$M_K^{\mathrm{t}}M_K = M_R^{\mathrm{t}}M_R = M_S^{\mathrm{t}}M_S = A = \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 2 \\ 2 & 2 & 2 & 13 \end{bmatrix}.$$

It is an interesting exercise, left to the reader, to find the geometric meaning behind these equalities. As a hint, note that if $D_1, D_2$ with bends $b_1, b_2$ are tangent, then the reflection of $D_1$ in $D_2$ has bend $b_1 + 2b_2$. The action of $A$ on a geometric quad is illustrated in Figure 16; the original quad has solid lines, its image, dashed lines.

As members of $\mathrm{O}(q)$, the shift matrices are products of reflections. For example, $M_K$ is the product of two reflections

$$M_K = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & -1 \end{bmatrix}.$$

The others involve additional reflections to interchange rows. Is there a geometric interpretation of these in $\mathcal{E}$? in $\mathcal{L}$? Are they related to the reflections which go into the Möbius transformations $g_\bullet$ of §2.4?

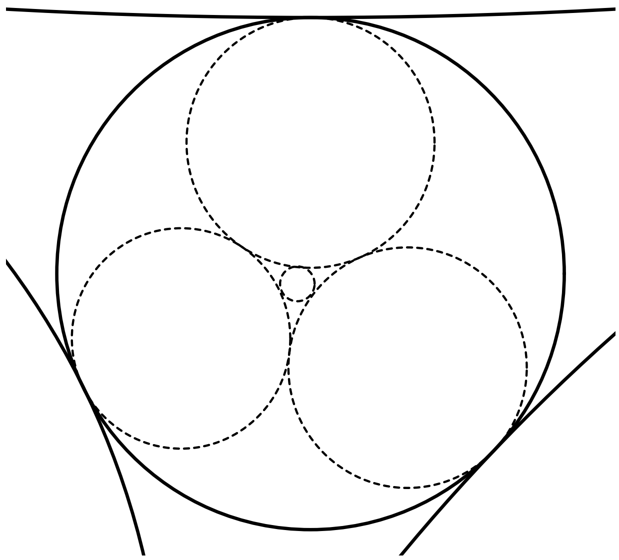

FIGURE 16. Action of $A$.

9.2. **Möbius Actions.** We certainly need a deeper understanding of the various actions of Möbius transformations $\mathfrak{M}$ on quads. For quads of actual discs, the geometric action is clear: each geometric quad is mapped to another. Indeed, $\mathfrak{M}$ acts transitively on the collection of geometric quads. The associated quads of bends are likewise mapped to one another, but this action requires information on the centers. If that is discarded, keeping only the bends, this geometric action is lost. On the other hand, the group $\mathrm{O}(q)$ acts transitively on the set of bends $Z^+$; indeed, it acts on $C^+$, foliated by the level sets of $q$. Since $\mathrm{O}(q)$ is isomorphic as a Lie group to $\mathrm{PSL}(2, \mathbf{C})$, this means that Möbius transformations now provide another action on quads. What is the connection? Is there new geometry in this more abstract setting of bends?

9.3. **Homogeneous Coordinates.** Lastly, let us say a little about ongoing work. In Sections 3, 4, and 5 we described what happens to a quad of bends under repeated applications of one quad shift — namely, the resulting quads seem to converge (in $\mathbf{RP}^3$) to a limit quad of bends. In fact, in each case this same limit quad results generically, that is, for *any* starting vector, whether a quad or not. (See the last paragraphs in § 3,4, and 5.) Moreover, we have evidence that such generic convergence to quads occurs under iteration not only of quad shifts themselves, but in fact, of arbitrary (infinite order) matrices in $G$. In Minkowski space we may identify $\mathbf{RP}^3$ with the ideal boundary of the Lorentz cone $C_L^+$, and what we seem to see there is orbits of infinite cyclic subgroups of $G$ which converge to very special ideal boundary points, namely, integral points in the ideal boundary of $Z_L^+$. We would like to understand this geometry. Analysing other geometric sequences in A-packings could add tantalizing tidbits to this picture.

## References


[1] Dov Aharonov, *The sharp constant in the ring lemma*, Complex Variables, Theory and Appl.

[2] Alan F. Beardon, Tomasz Dubejko, and Kenneth Stephenson, *Spiral hexagonal circle packings in the plane*, Geometriae Dedicata **49** (1994), 39–70.

[3] Philip L. Bowers and Kenneth Stephenson, *A branched Andreev-Thurston theorem for circle packings of the sphere*, Proc. London Math. Soc. (3) **73** (1996), 185–215.

[4] David W. Boyd, *Improved bounds for the disk-packing constant*, Aeq. Math. **9** (1973), 99–106.

[5] ———, *The residual set dimension of the Apollonian packing*, Mathematika **20** (1973), 170–174.

[6] H. S. M. Coxeter, *Loxodromic sequences of tangent spheres*, Aeq. Math. **1** (1968), 104–121.

[7] Predrag Cvitanović, *Circle maps: irrationally winding*, From Number Theory to Physics (M. Waldschmidt, P. Moussa, J.-M. Luck, and C. Itzkyson, eds.), Springer-Verlag, 1992, pp. 631–658.

[8] Kenneth J. Falconer, *The Geometry of Fractal Sets*, Cambridge Tracts in Math., vol. 85, Camb. Univ. Press, Cambridge, New York, 1986.

[9] H. Furstenberg and H. Kesten, *Products of random matrices*, Ann. Math. Stat. **31** (1960), 457–469.

[10] Thorold Gosset, *The hexlet*, February 6, Nature (1937), 251–252, Attached comment by Soddy.

[11] Lowell J. Hansen, *On the Rodin and Sullivan ring lemma*, Complex Variables, Theory and Appl. **10** (1988), 23–30.

[12] T. L. Heath, *History of Greek Mathematics*, Oxford Univ. Press, New York, 1921 (Dover reprint, 1981).

[13] K. E. Hirst, *The Apollonian packing of circles*, Journal Lond. Math. Soc. **42** (1967), 281–291.

[14] Alexander Jones (ed.), *Pappus of Alexandria, book 7 of the collection, part 2. commentary, index, and figures*, vol. 8, Springer-Verlag, Sources in the History of Mathematics and Physical Sciences, 1992.

[15] D. G. Larman, *On the exponent of convergence of a packings of spheres*, Mathematika **13** (1966), 57–59.

[16] S. Malitz and A. Papakostas, *On the angular resolution of planar graphs*, SIAM J. Discrete Math. **7** (1994), 172–183.

[17] Benoit B. Mandelbrot, *The Fractal Geometry of Nature*, W. H. Freeman and Co., San Francisco, 1982.

[18] S. S. Manna and H. J. Herrmann, J. Phys. A **24** (1991).

[19] Z. A. Melzak, *Infinite packings of disks*, Canad. J. Math. **18** (1966), 838–852.

[20] Gary L. Miller and William Thurston, *Separators in two and three dimensions*, Proceedings of the 22nd Annual ACM Symposium on Theory of Computing, ACM, Baltimore, May 1990, pp. 300–309.

[21] Bojan Mohar, *A polynomial time circle packing algorithm*, Discrete Math. **117** (1993), 257–263.

[22] Frank Morley, *The hexlet*, January 9, Nature (1937), 72–73.

[23] S. J. Patterson, *The limit set of a Fuchsian group*, Acta Math. **136** (1976), 241–273.

[24] Burt Rodin and Dennis Sullivan, *The convergence of circle packings to the Riemann mapping*, J. Differential Geometry **26** (1987), 349–360.

[25] F. Soddy, *The bowl of integers and the hexlet*, January 9, Nature (1937), 77–79.

[26] ———, *The Hexlet*, Nature (December 5, 1936), 958, poem.

[27] ———, *The Kiss Precise*, Nature (June 20, 1936), 1021, poem.

[28] Bo Söderberg, *Apollonian tiling, the Lorentz group, and regular trees*, Phys. Rev. A **46** (15 August, 1992), 1859–1866.

[29] Kenneth Stephenson, *Circle packings in the approximation of conformal mappings*, Bulletin, Amer. Math. Soc. (Research Announcements) **23, no. 2** (1990), 407–415.

[30] J. B. Wilker, *Inversive geometry*, The Geometric Vein: The Coxeter Festschrift (C. Davis, B. Grünbaum, and F. A. Sherk, eds.), Springer-Verlag, New York, Heidelberg, Berlin, 1981, pp. 379–442.

Technion, Israel Institute of Technology
*E-mail address*: `dova@techunix.technion.ac.il`

University of Tennessee, Knoxville
*E-mail address*: `kens@math.utk.edu`