Diamonds and Dragons: Notes on Dekking's Generalized Dragon Curves

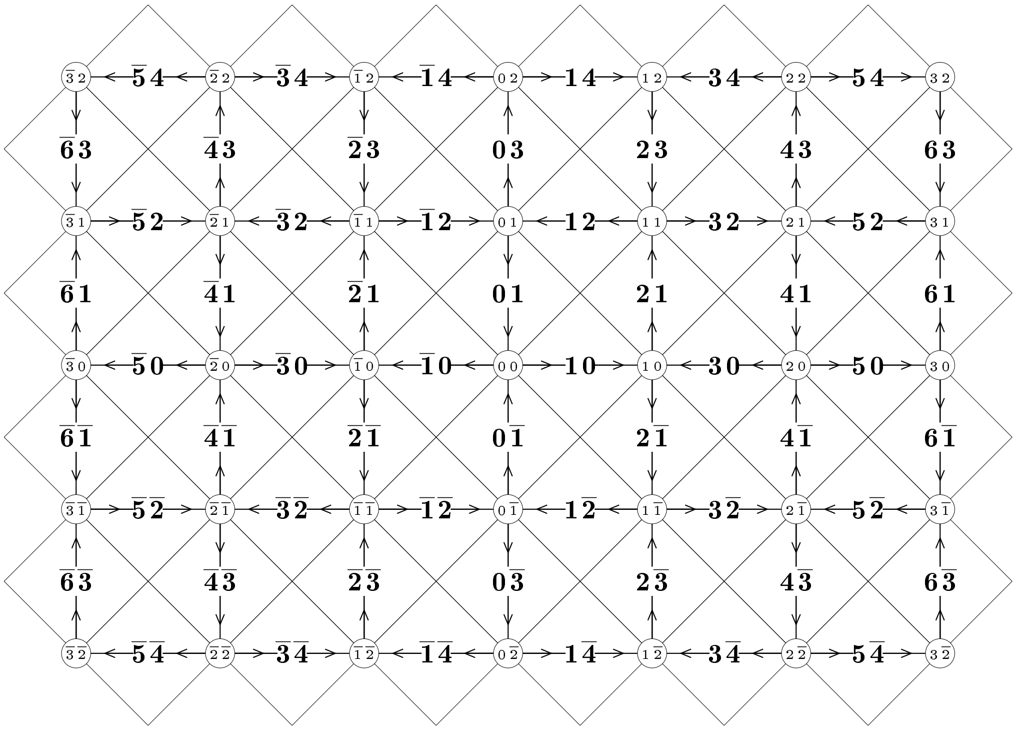
Don Knuth, September 2010

The following notes describe an algorithmic way to look at the intriguing family of curves discovered by F. M. Dekking in the 1970s and discussed briefly in [A]. I've prepared these remarks as grist for the mill while preparing to write some day about "structured recursion" in Chapter 8 of *The Art of Computer Programming*. Dear reader, please become familiar with reference [A] before proceeding further.

∗ — ∗ — ∗

A "zigzag path" in the complex plane begins at the origin and proceeds to take unit steps, advancing by i^{k_0}, i^{k_1}, i^{k_2}, …, where (usually) $k_0 = 0$ and $k_{j+1} = k_j \pm 1$. Thus the path proceeds from 0 to 1 and then to $1 \pm i$, and then to either $\pm i$ or $2 \pm i$, etc. Such a path is equivalent to a walk from 0 in the interesting directed graph below.

The key idea to be exploited here is that each unit segment of the path can be regarded as a diamond-shaped tile of area 1/2, where the segment forms the diagonal of the tile. Each tile is named by complex coordinates in a simple way; namely, if the segment goes from z to z' where $z' = z + i^k$, we associate the tile with the Gaussian integer $z + z'$. For example, the tiles near the origin can be visualized as follows:



Here a pair of boldface integers like '**2 3**' stands for a tile name, in this case $2 + 3i$; negative integers are indicated by overlines. Tile **1 0** is called the unit tile. Similarly, a pair of small integers like '$_{3\,\bar{1}}$' stands for the complex point $3 - i$. Notice that the label $w = a + bi$ of a tile is always an "odd" Gaussian integer, in the sense that $a + b$ is odd; in other words, w is never a multiple of $1 + i$. The midpoint of tile w is point $w/2$ in the complex plane. (When choosing the notation for the present notes I decided to avoid division by 2, and thereby to avoid dealing with half-integers. The reader should however keep in mind that a tile's label is twice the size of its actual position in the plane.)

Moreover, let's say that a Gaussian integer is *triply even* if it is a multiple of $2 + 2i$ — equivalently, if it's a multiple of $(1 + i)^3$. A triply even number $a + bi$ is characterized by having $a \bmod 2 = 0$, $b \bmod 2 = 0$, and $(a + b) \bmod 4 = 0$. Each tile label w can be represented uniquely in the form $e + i^{-k}$, where e is triply even and $0 \le k < 4$. For example, **2 3** is $(2 + 2i) + i^{-3}$; **5 $\bar{\mathbf{2}}$** is $(6 - 2i) + i^{-2}$. The direction of the arc in tile $e + i^{-k}$ is either right, up, left, or down, according as $k = 0$, 1, 2, or 3, respectively; in other words, an arc from z to $z + i^k$ defines a tile of type k.

Using these conventions, a zigzag path α defines a *polyomino*, namely a connected set of tiles, denoted by $p(\alpha)$. We say that the path is *self-avoiding* if no tile is repeated when this polyomino is viewed as a multiset. Notice that a self-avoiding path α might not actually be "simple" in the normal sense of graph theory, because it might go through a point z twice; such duplicity occurs when the Gaussian integer z is an interior point of $p(\alpha)$. The polyomino $p(\alpha)$ does not, in general, characterize the path α; but we can recover α by putting "walls" around the exterior edges and also by erecting a wall of length $\sqrt{2}$ and slope ± 1 at each interior point — thus obtaining a "labyrinth" instead of a polyomino, as illustrated in [A]. In such a labyrinth the path is that of a billiard ball as it bounces off the walls. The intriguing properties of labyrinths are not of interest to us in the present notes, however, so we won't discuss them further.

The simplest way to describe a short zigzag path is to present it as a sequence of s directions or a sequence of $s-1$ folds. For example, $\alpha = 01012$ is the path that goes right, up, right, up, and left; it's also described by the folds $S = DUDD$, as in [A]. This path of length 5 defines an L-shaped pentomino, $p(\alpha) = \{\mathbf{10}, \mathbf{21}, \mathbf{32}, \mathbf{43}, \mathbf{34}\}$, which the reader is encouraged to locate in the diagram above. (Incidentally, only five of the twelve pentomino shapes can actually arise from a zigzag path of length 5; for example, the T clearly can't occur because of its branching structure. More interesting is the fact that the U is also impossible.)

Let $|\alpha|$ be the length of path α, and let $\delta(\alpha)$ be the point at the end of the path. For example, $|01012| = 5$ and $\delta(01012) = 1 + 2i$.

Dekking introduced a fascinating associative operation on zigzag paths, the *folding product* $S * T$ of folding sequences S and T; see [A], where the folding powers S^{*n} and $S^{*\infty}$ are also defined. His folding product can also be regarded as an operation on paths, $\alpha * \beta$. It has many nice properties; for example, $|\alpha * \beta| = |\alpha||\beta|$ and $\delta(\alpha * \beta) = \delta(\alpha)\delta(\beta)$.

Let's turn now to the key concept of the present notes, namely a binary operation on *tiles* that corresponds closely to this same notion of folding product. If v and w are tiles of types j and k, where $v = d + i^{-j}$ and $w = e + i^{-k}$, we define

$$v * w \;=\; i^{-k}v + ze \;=\; i^{-k}v + z(w - i^{-k}) \;=\; zw + i^{-k}(v - z),$$

where z is a parameter. For example, we always have $v * \mathbf{10} = v$, regardless of z; in other words, the unit tile is a right identity element. Notice also that $v * w$ has type $(j + k) \bmod 4$. The main motivation for this definition is the following fact:

Lemma K. *If $z = \delta(\alpha)$, we have*

$$p(\alpha * \beta) \;=\; \bigcup \{\, v * w \mid v \in p(\alpha), w \in p(\beta) \,\}.$$

Proof. The reader is encouraged to verify this statement — which isn't extremely difficult — in order to get practice with the notation being used here. □

Although the folding product operation is associative on paths, it isn't "quite" associative on tiles, because of the dependence on z. If we exhibit this dependence explicitly by writing $v *_z w$, it turns out that $u *_z (v *_{z'} w) = (u *_z v) *_{zz'} w$. In what follows we shall be working with a fixed value of z, and we will interpret n-ary folding products $w_1 * \cdots * w_n$ by evaluating them from right to left so that $w_1 * w_2 * \cdots * w_n = w_1 * (w_2 * \cdots * w_n)$. In particular, $u * v * w$ will mean $u * (v * w)$, *not* $(u * v) * w$.

Let S be a folding sequence that corresponds to the path α. Then the folding powers S^{*n} correspond to the paths α^{*n}. An immediate corollary of Lemma K is the fact that $z = \delta(\alpha)$ implies

$$p(\alpha^{*n}) \;=\; \bigcup \{\, w_1 * \cdots * w_n \mid w_1, \ldots, w_n \in p(\alpha) \,\}.$$

The same formula holds also with $n = \infty$, if we understand the infinite folding product $w_1 * w_2 * \cdots$ to be defined only when w_j is the unit tile for all large values of j. The unit tile is, of course, always part of $p(\alpha)$.

From now on we shall let z be a fixed Gaussian integer, and we shall study properties of the zigzag paths α for which $\delta(\alpha) = z$. It will be convenient to let $s = |z|^2$; in other words, if $z = a + bi$ we have $s = a^2 + b^2$. Furthermore we let $Z = (2 + 2i)z$.

Given z, there's a natural equivalence relation on the set of all tiles: We write $w \equiv w'$ if $w' - w$ is a Gaussian integer multiple of Z. With this definition it is clear that $v \equiv v'$ and $w \equiv w'$ implies $v * w \equiv v' * w'$; in fact, we have $v * w \equiv v' * w'$ even under the weaker hypothesis that $v \equiv v'$, if w and w' are tiles of the same type k.

The tiles fall into $4s$ equivalence classes under this relation, because there are $|Z|^2 = 8s$ residue classes of Gaussian integers modulo Z and exactly half of them are odd. For example, if $z = 1 + 2i$, we have $s = 5$ and $Z = -2 + 6i$; every tile is equivalent to exactly one of the following twenty tiles near the origin:

$$\mathbf{03},\ \mathbf{23},\ \mathbf{\overline{3}2},\ \mathbf{\overline{1}2},\ \mathbf{12},\ \mathbf{\overline{2}1},\ \mathbf{01},\ \mathbf{21},\ \mathbf{\overline{3}0},\ \mathbf{\overline{1}0},\ \mathbf{10},\ \mathbf{30},\ \mathbf{\overline{2}\overline{1}},\ \mathbf{0\overline{1}},\ \mathbf{2\overline{1}},\ \mathbf{\overline{1}\overline{2}},\ \mathbf{1\overline{2}},\ \mathbf{3\overline{2}},\ \mathbf{\overline{2}\overline{3}},\ \mathbf{0\overline{3}}.$$

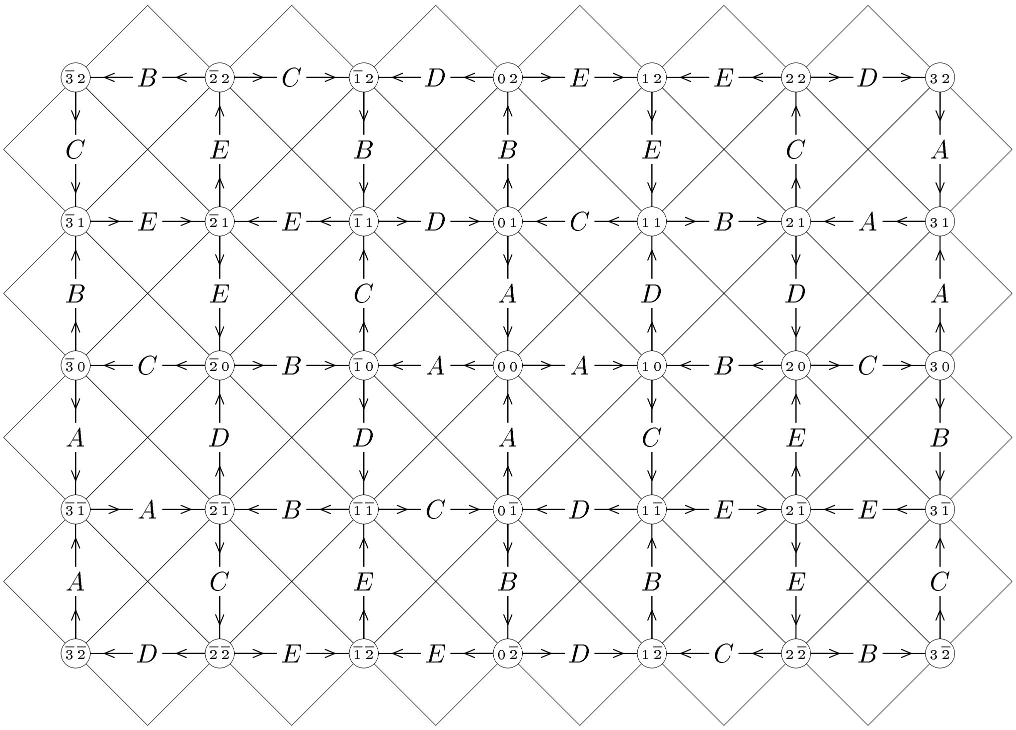
In this particular case there's also another nice set of twenty class representatives, more suitable for computation:

$$\mathbf{\overline{9}0},\ \mathbf{\overline{7}0},\ \mathbf{\overline{5}0},\ \mathbf{\overline{3}0},\ \mathbf{\overline{1}0},\ \mathbf{10},\ \mathbf{30},\ \mathbf{50},\ \mathbf{70},\ \mathbf{90},\ \mathbf{0\overline{9}},\ \mathbf{0\overline{7}},\ \mathbf{0\overline{5}},\ \mathbf{0\overline{3}},\ \mathbf{0\overline{1}},\ \mathbf{01},\ \mathbf{03},\ \mathbf{05},\ \mathbf{07},\ \mathbf{09}.$$

The methods of exercise 7–137 in *The Art of Computer Programming*, Volume 4A, can be used for convenient computation modulo Gaussian integers in general.

Tiles of different types cannot be equivalent to each other, because Z is triply even. Therefore exactly s of the equivalence classes point in each of the four directions. It will be convenient to "crudify" the $\equiv$ relation by allowing tiles w, iw, $-w$, and $-iw$ to be related to each other as well: We say that v is *congruent* to w, and write $v \cong w$, if and only if $v \equiv i^k w$ (modulo Z) for some integer k. It follows that there are s congruence classes. Moreover, we have $v \cong w$ if and only if $v = w * u$ for some tile u.

If, for instance, we call the classes A, B, C, D, and E when $z = 1 + 2i$, the pattern of tiles near the origin is



and the pattern around any Gaussian multiple of $Z/2 = -1 + 3i$ (e.g., the point $\scriptstyle 31$) is the same.

In this particular case, the horizontal tiles with fixed imaginary part always form the repeating sequence

$$\ldots,\ C,\ B,\ A,\ A,\ B,\ C,\ D,\ E,\ E,\ D,\ C,\ B,\ A,\ A,\ B,\ C,\ D,\ E,\ E,\ D,\ \ldots$$

and the vertical tiles with fixed real part also follow this sequence. One can show that the analogous property holds, with a sequence of period $2s$, if and only if all prime factors of s are congruent to 1 mod 4.

OK: We've now seen lots of fancy definitions, which have curious mathematical properties. But what's the point? It's time for some real live action.

Theorem S. *Let α be a zigzag path from 0 to z. Then $\alpha^{*\infty}$ is self-avoiding if and only if no two tiles of $p(\alpha)$ are congruent.*

Proof. Let β be the path 0123. Dekking has proved that $\alpha^{*\infty}$ is self-avoiding if and only if $\alpha * \beta$ is self-avoiding. Equivalently, by Lemma K, the tiles

$$\bigl(p(\alpha) * \mathbf{10}\bigr) \cup \bigl(p(\alpha) * \mathbf{21}\bigr) \cup \bigl(p(\alpha) * \mathbf{12}\bigr) \cup \bigl(p(\alpha) * \mathbf{01}\bigr) \tag{$*$}$$

must be distinct, because $p(\beta) = \{\mathbf{10}, \mathbf{21}, \mathbf{12}, \mathbf{10}\}$.

Now if w is any tile of $p(\alpha)$, the tiles $w * \mathbf{10}$, $w * \mathbf{21}$, $w * \mathbf{12}$, $w * \mathbf{01}$ are congruent. But they're also inequivalent, hence distinct, because $\mathbf{10}$, $\mathbf{21}$, $\mathbf{12}$, and $\mathbf{01}$ have different types. If all $|\alpha|$ tiles of $p(\alpha)$ are incongruent, the $4|\alpha|$ tiles of $(*)$ are therefore distinct.

Conversely, if say the third tile of $p(\alpha) * \mathbf{12}$ is the same as the fifth tile of $p(\alpha) * \mathbf{01}$, then the third and fifth tiles of $p(\alpha)$ must be congruent. $\square$

Corollary S. *If α is a zigzag path from 0 to z and $|\alpha| > s = |z|^2$, then $\alpha^{*\infty}$ isn't self-avoiding.* $\square$

A zigzag path α is said to be *plane-filling* if, for all finite sets T of tiles, there is a shift amount w such that every tile of $T + w$ is contained in $p(\alpha^{*\infty})$. For example, the dragon curve $\alpha = 01$ is known to be plane-filling; indeed, $\alpha^{*(2n+3)}$ contains a square array of $(2^n + 1) \times (2^n + 1)$ tiles in that case, when $n \ge 1$.

Theorem P. *Let α be a zigzag path from 0 to z. Then α is plane-filling if and only if α is self-avoiding and $|\alpha| = s = |z|^2$. In other words, α is plane-filling if and only if $p(\alpha)$ contains exactly one tile in each congruence class.*

Proof. Again this is the algebraic equivalent of a theorem proved by Dekking. $\square$

As an example of Theorem P, each plane-filling path with $z = 1 + 2i$ corresponds to a a zigzag path of length 5 from $_{00}$ to $_{12}$ in the previous diagram that runs through tiles of classes A, B, C, D, and E in some order. There are two such paths: 01012 (which gives $ADBCE$) and 01210 (which gives $ADCBE$). When $z = 2 - i$ we have a similar case; the congruence classes and equivalence classes are the same, because $2 - i = -i(1 + 2i)$, but no suitable zigzag path exists. The case $z = -1 + 2i$ is also of interest: Here the path 01212 turns out to be plane-filling, because the congruence classes for $-1 + 2i$ are obtained from those of $1 + 2i$ by negating the real parts of each tile.

Now let's look at a new problem, not considered by Dekking but solved for the special case of the dragon curve by Davis and Knuth in [D]. Suppose α is plane-filling, and suppose we know the coordinates of some tile w. We'd like to decide whether w is part of $p(\alpha^{*\infty})$; and, if so, we'd like to know a sequence of tiles w_1, $\ldots$, w_n in $p(\alpha)$ such that $w = w_1 * \cdots * w_n$.

Theorem F. *If α is plane-filling and u is any tile, there's a unique $v \in p(\alpha)$ such that $u = v * w$.*

Proof. Theorem P tells us that there's a unique $v \in p(\alpha)$ such that $u \cong v$. Suppose u has type l and v has type j; then any solution to $u = v * w$ will have w of type $k = (l - j) \bmod 4$. And indeed the solution must be (and is) $w = e + i^{-k}$, where $e = (u - i^{-k}v)/z$ is triply even. $\square$

By applying this factorization theorem repeatedly, we can decide rather quickly quickly whether any given tile is part of $p(\alpha^{*\infty})$. For example, let $\alpha = 01012$ and let w be the "random" tile $\mathbf{314\,159}$. We find $\mathbf{314\,159} = \mathbf{10} * \mathbf{126\,\overline{93}}$, $\mathbf{126\,\overline{93}} = \mathbf{43} * \mathbf{\overline{11}\,\overline{70}}$, $\mathbf{\overline{11}\,\overline{70}} = \mathbf{10} * \mathbf{\overline{31}\,\overline{10}}$, $\mathbf{\overline{31}\,\overline{10}} = \mathbf{10} * \mathbf{\overline{11}\,10}$, $\mathbf{\overline{11}\,10} = \mathbf{10} * \mathbf{16}$, $\mathbf{16} = \mathbf{32} * \mathbf{30}$, $\mathbf{30} = \mathbf{32} * \mathbf{1\overline{2}}$, $\mathbf{1\overline{2}} = \mathbf{21} * \mathbf{0\overline{1}}$, and now the process stops because $\mathbf{0\overline{1}} = \mathbf{10} * \mathbf{0\overline{1}}$.

The repeated factorization process just illustrated will end with $\mathbf{10}$ if and only if the original tile w is in $p(\alpha^{*\infty})$. Our "random" example failed that test; but all is not lost! Indeed, the folding product of tiles satisfies the important identity

$$v * i^t w \;=\; i^t(v * w). \tag{$**$}$$

Thus if all the tiles of the previous example are multiplied by i, we'll end up with $\mathbf{10}$ instead of $\mathbf{0\overline{1}}$. Namely, $\mathbf{\overline{159}\,314} = \mathbf{10} * \mathbf{93\,126}$, $\mathbf{93\,126} = \mathbf{43} * \mathbf{70\,\overline{-11}}$, etc.; and we find

$$\mathbf{\overline{159}\,314} = \mathbf{10} * \mathbf{43} * \mathbf{10} * \mathbf{10} * \mathbf{10} * \mathbf{32} * \mathbf{32} * \mathbf{21}.$$

We conclude that $\mathbf{\overline{159}\,314}$ belongs to $p(\alpha^{*\infty})$, hence $\mathbf{314\,159}$ belongs to $-ip(\alpha^{*\infty})$.

How does this approach compare to the analysis in [D] of the "simple" dragon curve generated by $\alpha = 01$? Theorem 4 of that paper shows that every tile of the plane lies in exactly one of the four infinite polyominoes $p(\alpha^{*\infty})$, $ip(\alpha^{*\infty})$, $-p(\alpha^{*\infty})$, $-ip(\alpha^{*\infty})$. And in particular the proof of that theorem considers, as an example, the path segment from $33+6i$ to $33+5i$; that's tile $\mathbf{66\,11}$ in our notation. Factorization via Theorem F yields

$$\mathbf{66\,11} \;=\; \mathbf{21} * \mathbf{10} * \mathbf{21} * \mathbf{21} * \mathbf{21} * \mathbf{21} * \mathbf{10} * \mathbf{10} * \mathbf{10} * \mathbf{21} * \mathbf{0}\overline{\mathbf{1}};$$

hence $\mathbf{66\,11}$ belongs to $-ip(\alpha^{*\infty})$, the "fourth dragon design" as stated in [D]. Moreover, one can show that if $n = (a_t \ldots a_1 a_0)_2$ in binary notation, and if the Gray binary code for n is $g(n) = n \oplus \lfloor n/2 \rfloor = b_t \ldots b_1 b_0$ where $b_j = (a_j + a_{j+1}) \bmod 2$, then the $(n+1)$st tile visited by the dragon curve is $c_0 * c_1 * \cdots * c_t$, where $c_j = \mathbf{10}$ if $b_j = 0$, $c_j = \mathbf{21}$ if $b_j = 1$. This characterization is simpler than the one in [D].

Dekking originally called a plane-filling path α "perfect" if it satisfies what he has more recently called the "carousel property," namely that the entire plane is covered exactly once by $p(\alpha^{*\infty}) \cup ip(\alpha^{*\infty}) \cup -p(\alpha^{*\infty}) \cup -ip(\alpha^{*\infty})$. Algebraically, this means that an arbitrary tile w will reduce via Theorem F to one of the four generalized unit tiles $\{\mathbf{10}, \mathbf{01}, \overline{\mathbf{1}}\mathbf{0}, \mathbf{0}\overline{\mathbf{1}}\}$ after a finite number of steps.

Suppose the plane-filling path α begins with 01. Then Dekking showed that the additional carousel property is equivalent to the condition that, given z, the particular tile $V = (2+2i)z - 2i - 1$ is not in $p(\alpha)$. Indeed, V is the unique solution to the equation $\mathbf{12} = V * \mathbf{12}$; so Theorem F will never reduce V, if V belongs to $p(\alpha)$.

Given a plane-filling path α that begins with 01, consider the directed graph on tiles where $u \to w$ whenever $u = v * w$ for some $v \in p(\alpha)$. This digraph always has loops $w \to w$ when w is one of the four generalized unit tiles; Dekking's theorem says that the carousel property fails if and only if this digraph also contains the loop $\mathbf{12} \to \mathbf{12}$. Then, by (**), there will also be loops $\overline{\mathbf{2}}\mathbf{1} \to \overline{\mathbf{2}}\mathbf{1}$, $\overline{\mathbf{1}}\overline{\mathbf{2}} \to \overline{\mathbf{1}}\overline{\mathbf{2}}$, and $\mathbf{2}\overline{\mathbf{1}} \to \mathbf{2}\overline{\mathbf{1}}$.

Can there be any other loops? I don't know. I don't think $\mathbf{23}$ or $\mathbf{32}$ can loop, but I haven't seen an easy way to rule out the loop $\mathbf{30} \to \mathbf{30}$ when s is sufficiently large.

Nor have I been able to find an algebraic reason why the the plane-filling path $010103032^{*\infty}$ illustrated in [A] covers only 1/16 of the plane, while the "alternating dragon curve" $0103^{*\infty}$ covers 1/8.

Incidentally, if α begins with 03 instead of 01, the corresponding fixed point would be $\mathbf{21}$ instead of $\mathbf{12}$, and the tile $V' = (2-21i)z + 2i - 1$ would be forbidden from $p(\alpha)$. This tile V' is congruent to $\mathbf{21}$, so it can't occur in a self-avoiding α that begins with 01; the potential fixed point at $\mathbf{21}$ is therefore never present unless α begins with 03.

Readers who wish to experiment with these operations on tiles are invited to download the program DRAGON-CALC from the author's website.

References

[A] Addendum to the revised and corrected version of [D], in the book *Selected Papers on Fun and Games* by Donald E. Knuth (Stanford, California: Center for the Study of Language and Information, 2010), 603–614.

[D] Chandler Davis and Donald E. Knuth, "Number representations and dragon curves," *Journal of Recreational Mathematics* **3** (1970), 66–81, 133–149. Revised and corrected in the book *Selected Papers on Fun and Games* by Donald E. Knuth (Stanford, California: Center for the Study of Language and Information, 2010), 571–603.