

Extremal topological and geometric problems for polyominoes

Greg Malen*

Department of Mathematics
Duke University
Durham, North Carolina, U.S.A.
gmalen@math.duke.edu

Erika Berenice Roldan-Roa[†]

Zentrum Mathematik
Technical University of Munich
Garching bei München, Bavaria, Germany
erika.roldan@ma.tum.de

Submitted: Oct 23, 2019; Accepted: Jun 6, 2020; Published: Jun 26, 2020

© The authors. Released under the CC BY license (International 4.0).

Abstract

We give a complete solution to the extremal topological combinatorial problem of finding the minimum number of tiles needed to construct a polyomino with h holes. We denote this number by $g(h)$ and we analyze structural properties of polyominoes with h holes and $g(h)$ tiles, characterizing their efficiency by a topological isoperimetric inequality that relates minimum perimeter, the area of the holes, and the structure of the dual graph of a polyomino. For $h \leq 8$ the values of $g(h)$ were originally computed by Tomas Olivera e Silva in 2015, and for the sequence $h_l = (2^{2l} - 1)/3$ by Kahle and Róldan-Roa in 2019, who also showed that asymptotically $g(h) \approx 2h$. Here we also prove that the polyominoes constructed by Kahle and Róldan-Roa with $h_l = (2^{2l} - 1)/3$ holes and $g(h_l)$ tiles are in fact unique up to isometry with these fixed parameters; that is, having the minimal number of tiles for h_l holes.

Mathematics Subject Classifications: 05A15-16, 05A20, 05B50, 05D99, 57M20

*Supported by NSF DMS 17-13012.

[†]This work was supported by HFSP RGP0051/2017, NSF DMS 17-13012, and NSF DMS-1352386. This project received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 754462.

1 Introduction

A *polyomino* is a planar shape in \mathbb{R}^2 formed by gluing together a finite number of congruent squares along their edges such that it has a connected interior. We refer to the squares on a polyomino as either *squares* or *tiles*.

In this paper, we are interested in the extremal topological problem of finding the minimum number of tiles required for a polyomino to have a specified number of holes. To be precise about the topology, we consider the tiles of a polyomino to be closed. Polyominoes are finite unions of these closed tiles, so they are compact. The *holes* of a polyomino are the bounded, connected components of the polyomino's complement in the plane. For a polyomino A we denote its number of holes and tiles by $h(A)$ and $|A|$, respectively.

Definition 1. For $h \geq 1$, we define the sequence of the minimum number of tiles needed for constructing a polyomino with h holes as

$$g(h) := \min_{h(A)=h} |A|. \tag{1}$$

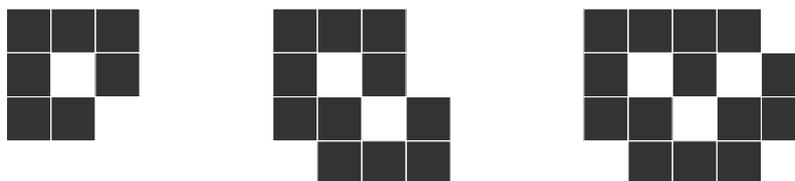


Figure 1: Crystallized polyominoes for $1 \leq h \leq 3$.

Definition 2. A polyomino with h holes is *crystallized* if it has $g(h)$ tiles.

This definition is inspired by the striking resemblance these extremal polyominoes bear to chemical crystals, both visually and in the driving principle of their formation. As the formation of crystals is restricted by the physical properties of the atoms involved, the tiles in these polyominoes are trying to pack densely while maximizing the number of enclosed holes, and these holes impose certain geometrical structures in order to achieve maximum efficiency.

In [1], all polyominoes with $n \leq 28$ tiles were enumerated, up to isometry, by their number of holes. These computations give the values of $g(h)$ for $1 \leq h \leq 8$ and also count the number of crystallized polyominoes for $1 \leq h \leq 8$.

In [4], the related function $f(n)$, defined as the maximum number of holes that a polyomino with n tiles can have, was introduced and its asymptotic behavior was studied. They proved that $g(h)$ is the right inverse of $f(n)$, and that $f(n) = (1/2)n + O(\sqrt{n}) + o(1)$. In the same paper, a sequence of crystallized polyominoes was constructed with $g(h) + h = (2^l + 1)^2 - 1$ for all $l \geq 1$ (see Table 1). We will refer to this sequence of crystallized polyominoes as the K–R sequence, and we denote its elements by S_l for $l \geq 1$.

h	1	2	3	4	5	6	7	8	$(2^{2l} - 1)/3$
$g(h)$	7	11	14	17	19	23	25	28	$[(2^{2l+1} + 3 \cdot 2^{l+1} + 4)/3] - 1$
$ \mathcal{G}_h $	1	4	3	8	1	64	4	37	1^*

Table 1: We denote by $|\mathcal{G}_h|$ the number of polyominoes, up to isometry, that have h holes and $g(h)$ tiles. The table shows all previously known values for $g(h)$ and $|\mathcal{G}_h|$. The values for $1 \leq h \leq 8$ are due to e Silva [1], and the first two rows of the last column are from [4]. *We prove in Theorem 9 that $|\mathcal{G}_{(2^{2l}-1)/3}| = 1$ for all $l \geq 1$.

In [6], the analogous problem for polyiamonds, polyforms constructed with equilateral triangles, was completely solved. In that setting, all crystallized polyiamonds were also found to satisfy a set of optimal geometric and topological structural conditions.

Here, we completely answer these questions for polyominoes. We find all values of $g(h)$, and we characterize the values of h for which optimal structural conditions, which we define precisely in Section 2, are satisfied. In the process of doing so, we are able to prove that the crystallized polyominoes constructed in [4] with $h_l = (2^{2l} - 1)/3$ holes are unique up to isometry.

2 Statement of Main Results

We use the sequences of perfect squares and pronic numbers, N^2 and $N(N+1)$, to produce benchmark values for $g(h)$, tracking the maximum number of holes that can fit in a certain area. Using α to denote a general square or pronic number, let h_α denote the maximum number of holes which can exist in a polyomino A which is contained in the square or pronic rectangle of area α .

Theorem 3. *For any positive integer $N \geq 3$,*

$$h_{N^2} = \begin{cases} \frac{(N-1)^2}{3} - 1 & \text{if } N \equiv 1 \pmod{3} \\ \frac{N(N-2)}{3} & \text{if } N = 2^l + 1 \text{ for } l \geq 1 \\ \frac{N(N-2)}{3} - 1 & \text{else,} \end{cases} \quad (2)$$

and

$$h_{N(N+1)} = \begin{cases} \frac{N(N-1)}{3} - 1 & \text{if } N \equiv 0 \text{ or } 1 \pmod{3} \\ \frac{(N+1)(N-2)}{3} - 1 & \text{if } N \equiv 2 \pmod{3}. \end{cases} \quad (3)$$

For any square or pronic number $\alpha = N^2$ or $\alpha = N(N+1)$ such that $N \leq 5$, the values of h_α and $g(h_\alpha)$ can be determined from the values computed in [1], listed in Table 1. For those first few values it can happen that $h_{N^2} = h_{N(N+1)}$. To avoid this complication, and

because those values are already known, the next theorem gives the values of $g(h_\alpha)$ for $N \geq 6$.

Theorem 4. For $\alpha \in \{N^2, N(N+1) : N \geq 6\}$, $g(h_\alpha) = \alpha - h_\alpha - C$ for

$$C = \begin{cases} 1 & \text{if } \alpha = N^2 \text{ and } N = 2^l + 1 \\ 3 & \text{if } \alpha = N^2 \text{ and } N \equiv 1 \pmod{3} \\ & \text{or } \alpha = N(N+1) \text{ and } N \not\equiv 2 \pmod{3} \\ 4 & \text{if } \alpha = N^2 \text{ and } N \not\equiv 1 \pmod{3}, N \neq 2^l + 1 \\ 5 & \text{if } \alpha = N(N+1) \text{ and } N \equiv 2 \pmod{3}. \end{cases} \quad (4)$$

In the above, C is the number of tiles missing from the boundary in a crystallized polyomino which has the maximum number of holes in a square or pronic rectangle. For instance, the $\alpha = 6^2$ case gives $h_\alpha = 7$ and $C = 4$, and confirms that $g(7) = 36 - 7 - 4 = 25$ from Table 1. We are now able to give the values for $g(h)$ for all $h \geq 1$, in terms of the values of $g(h_\alpha)$.

h	g(h)	h	g(h)	h	g(h)	h	g(h)	h	g(h)
9	30	30	81	51	128	72	175	93	221
10	33	31	83	52	131	73	177	94	223
11	35	32	85	53	133	74	179	95	225
12	38	33	88	54	135	75	182	96	228
13	40	34	90	55	137	76	184	97	230
14	43	35	92	56	140	77	186	98	232
15	45	36	95	57	142	78	188	99	234
16	48	37	97	58	144	79	190	100	236
17	50	38	99	59	146	80	193	101	238
18	53	39	101	60	149	81	195	102	241
19	55	40	104	61	151	82	197	103	243
20	57	41	106	62	153	83	199	104	245
21	59	42	108	63	155	84	201	105	247
22	62	43	110	64	157	85	204	106	249
23	64	44	113	65	160	86	206	107	251
24	67	45	115	66	162	87	208	108	254
25	69	46	117	67	164	88	210	109	256
26	71	47	119	68	166	89	212	110	258
27	74	48	122	69	168	90	215	111	260
28	76	49	124	70	171	91	217	112	262
29	78	50	126	71	173	92	219	113	264

Table 2: Values of $g(h)$ for $9 \leq h \leq 114$. We indicate in bold the values of h if $h = h_\alpha$ for an $\alpha \in \{N^2, N(N+1)\}$. For $h < 9$, see Table 1.

Theorem 5. For $h \geq 1$, let $\alpha = \min\{N^2, N(N + 1) : h \leq h_\alpha\}$. Then

$$g(h) = g(h_\alpha) - 2(h_\alpha - h). \quad (5)$$

Values of $g(h)$ are given in Table 2 for $9 \leq h \leq 113$, with values of h_α in bold. Theorem 5 assures that almost all increments of $g(h)$ from h to $h + 1$ are 2, and implicit in the combination of Theorems 4 and 5 is that the remaining increments from h_α to $h_\alpha + 1$ are always 3.

Then in Theorems 7 and 8 we give structural characterizations of crystallization for all $h \geq 1$. The *dual graph* of a polyomino is the graph whose vertices are indexed by the tiles of the polyomino, with edges between two vertices if and only if the respective tiles in the polyomino share an edge. We say that a polyomino is *acyclic* if its dual graph is a tree, and refer to cycles in the dual graph as *dual cycles*.

The *area of a hole* is defined to be the number of tiles needed to fill it. If the squares are assumed to have unit side length, then this would be equivalent to the Euclidean notion of area. In this case with squares, the holes all have the shape of polyominoes, and thus it is more convenient to discuss area in terms of the tiles which could be used to fill the holes.

A polyomino with h holes and n tiles has *minimal outer perimeter* if it has an outer perimeter equal to $2 \lceil 2\sqrt{n + h} \rceil$.

Definition 6. A polyomino is *efficiently structured* if it is acyclic, each hole has an area of one, and it has minimal outer perimeter.

Theorem 7. *Efficiently structured polyominoes are crystallized.*

However, the converse of Theorem 7 does not hold. We prove in Theorem 8 that there is an exceptional set of crystallized polyominoes which are acyclic with all holes having of area of one, but which fail to attain minimal outer perimeter. As a corollary, this implies that all crystallized polyominoes are in fact acyclic with all holes having an area of one.

For the cases from Theorem 4 in which $C = 4$ or $C = 5$, define

$$\begin{aligned} S &= \{N^2 \mid N \equiv 0 \text{ or } 2 \pmod{3} \text{ and } N \neq 2^l + 1 \text{ for any } l \geq 1\}, \\ R &= \{N(N + 1) \mid N \equiv 2 \pmod{3}\}. \end{aligned}$$

Theorem 8. For $\alpha \in S \cup R$, a crystallized polyomino with $h_\alpha + 1$ holes is acyclic and each of its holes have an area of one, but it does not attain minimal outer perimeter. For all other h , a crystallized polyomino is efficiently structured.

Finally, we consider the enumeration of crystallized polyominoes, proving that elements of the K–R sequence in [4] are the unique crystallized polyominoes, up to isometry, with $h = (2^{2l} - 1)/3$ for all $l \geq 1$.

Theorem 9. [*K–R Sequence Uniqueness*] For a fixed integer $l \geq 1$, there is only one, up to isometry, crystallized polyomino A with $h(A) = (2^{2l} - 1)/3$.

The rest of the paper is organized as follows. In Section 3, we give background, establish preliminary results, and prove Theorem 7. In Section 4, we exhibit various obstructions to efficient crystallization and prove Theorem 8. In Section 5, we prove Theorems 3, 4, and 9. Finally, in Section 7, we prove Theorem 5 by dismantling the crystallized polyominoes that we construct in Section 6.

3 Background and Preliminary Results

3.1 Basic Definitions and Terminology

In what follows we introduce some terminology and notation. We use the term *arrangement* to refer to any subset of tiles of the regular square tessellation of \mathbb{R}^2 . Polyominoes are themselves arrangements, but arrangements are not required to have connected interior and simply refer to the specified positions of a set of tiles. A space in the square lattice is said to be *filled* if it contains a tile, and *empty* if not.

The *total area* of a polyomino is the number of tiles plus the aggregate area of all its holes, and the *perimeter* $p(A)$ of a polyomino A is defined to be the number of edges that are part of the topological boundary of A . An edge of the perimeter is on the *hole perimeter* of A if it is bounding a hole, and it is on the *outer perimeter* otherwise. The number of edges on the hole perimeter is denoted by $p_h(A)$, and the number of edges on the outer perimeter of A is denoted by $p_o(A)$. As an example, let A be the polyomino in the center of Figure 1, then $p(A) = 30$, $p_o(A) = 18$, and $p_h(A) = 12$.

The *bounding rectangle* of a polyomino A is the smallest rectangle in which A fits. The *boundary layer* of a polyomino is the set of squares that have at least one edge on the outer perimeter, and the *interior* of a polyomino is the set of spaces which are not in the boundary layer. For a polyomino with a rectangular interior, say $(w - 2) \times (l - 2)$, define D_1 to be the arrangement in the boundary layer of the $w \times l$ rectangle in which all spaces except for the four corners are filled. Similarly, define D_2 to be the boundary arrangement in which all but one of the corners are filled. For D_1 and D_2 we suppress the dimensions of the rectangle, as these will always be clear from context.



Figure 2: D_1 and D_2 for a 3×4 rectangular interior.

Remark 10. Observe that $N^2 - (N - r)(N + r) = r^2$, and $N(N + 1) - (N - r)(N + 1 + r) = r^2 + r$. These values track the decrease in area if a square or pronic rectangle is replaced by a narrower rectangle with the same perimeter. In particular the difference is always at least one for squares, and at least two for pronic rectangles.

3.2 Hole Connectivity and the Dual Graph of a Polyomino.

We defined the dual graph of a polyomino in Section 2. Let $b(A)$ be the number of edges in the dual graph of a polyomino A . Since the interior of A must be connected, its dual graph must be connected and therefore have a spanning tree. Thus if A has n tiles, the dual graph has n vertices and its spanning tree has $n - 1$ edges. Therefore

$$b(A) \geq (n - 1). \quad (6)$$

It is important to notice that the dual graph does not capture the topology of a polyomino. In Figure 3 for example, A has no holes but its dual graph contains a cycle, and the dual graph of B is acyclic but B has five holes.

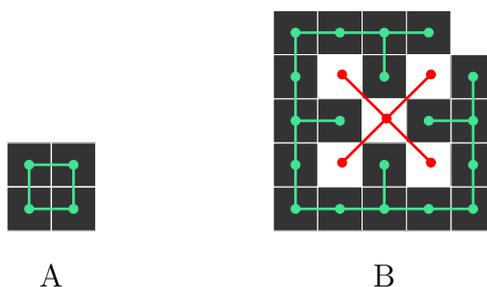


Figure 3: Polyominoes A and B with their dual graphs colored in green, and the hole graph of B colored in red.

Definition 11. The *hole graph* of a polyomino A is the graph whose vertices are indexed by the holes of A , with edges connecting two holes if their boundaries share a common vertex of the polyomino.

We refer to the hole adjacency condition as being *corner adjacent*, as opposed to the *edge adjacent* condition for tiles. A set of holes in a polyomino is said to be *connected* if the corresponding induced subgraph of the hole graph is connected. The hole graph of a polyomino A is necessarily acyclic, as a cycle corresponds to a Jordan curve in the plane which would disconnect the interior of A .

3.3 A Topological Isoperimetric Inequality

For all $n \geq 1$, we denote by $p_{\min}(n)$, the minimum perimeter that a polyomino with n tiles can have. In 1976, F. Harary and H. Harborth [3] proved that the minimum perimeter possible in a polyomino with n tiles is given by

$$p_{\min}(n) = 2 \lceil 2\sqrt{n} \rceil. \quad (7)$$

In [5], it was proven that polyominoes with area $\alpha \in \{N^2, N(N + 1)\}$ that attain minimum perimeter are unique up to isometry. These polyominoes are precisely those with the shape of a square or a pronic rectangle.

It is clear that polyominoes that are not simply connected, that have at least one hole, cannot achieve minimum perimeter as given in (7). The next lemma gives a lower bound for the minimum perimeter that a polyomino can have given its area and its number of holes.

Lemma 12. *If A is a polyomino with n tiles and h holes, then $p_o(A) \geq p_{min}(n + h)$.*

Proof. Let H be the aggregate area of all the holes of A . Then $H \geq h$ because the minimum area that a hole can have is an area of one. Let B be the polyomino obtained by completely covering the holes of A with H tiles. Observe that B has $n + H$ tiles, and by equation (7), $p(B) \geq p_{min}(n + H)$. Using the fact that the function $h(x) = 2 \lceil 2\sqrt{x} \rceil$ is a non-decreasing function we get that $p(B) \geq p_{min}(n + h)$. Then, because $p(B) = p_o(B)$ and A and B have the same outer perimeter, we conclude that

$$p_o(A) \geq p_{min}(n + h). \quad \blacksquare \quad (8)$$

Lemma 12 and equation (7) are the reasons of why we have defined in Section 2 that a polyomino with n tiles and h holes has minimal outer perimeter if its outer perimeter is $2 \lceil 2\sqrt{n + h} \rceil$. In what follows, using techniques introduced in [4], we give an upper bound for the number of holes that a polyomino can have.

Let A be a polyomino with n tiles. Then $4n = 2b(A) + p(A)$ because each square tile has 4 edges and these edges are either on the perimeter of A or connecting two tiles of A . Thus, using $p(A) = p_o(A) + p_h(A)$ we get

$$p_h(A) = 4n - 2b(A) - p_o(A). \quad (9)$$

The minimum number of edges that a hole can have is four, thus from (9) we get

$$h(A) \leq \frac{p_h(A)}{4} = \frac{4n - 2b(A) - p_o(A)}{4}.$$

Then, from inequalities (6) and (8) we get

$$h(A) \leq \frac{4n - 2(n - 1) - p_{min}(n + h)}{4}. \quad (10)$$

This is a topological isoperimetric inequality that bounds the number of holes that a polyomino can have depending on the structure of its dual graph, the area of its holes, and its outer perimeter. The next definition establishes notation for this upper bound as a function of n and h .

Definition 13. For all natural numbers n and h , define

$$M(n, h) = \frac{2n + 2 - p_{min}(n + h)}{4}. \quad (11)$$

From inequality (10), this function immediately gives a necessary condition for the existence of a polyomino, which is intricately tied to efficient structure.

Lemma 14. *If $M(n, h) < h$ for some natural numbers n and h , then a polyomino with n tiles and h holes does not exist. And for all $h \geq 1$, we have $h \leq M(g(h), h)$.*

Lemma 15. *A polyomino with n tiles and h holes is efficiently structured if and only if $h = M(n, h)$.*

Proof of Lemma 15. Let A be a polyomino with n tiles and h holes. If A is acyclic, its dual graph is a tree and $b(A) = n - 1$; if every hole has an area of one, then $p_h(A) = 4h$; and if A has minimal perimeter, then $p_o(A) = p_{min}(n + h)$. Thus if A is efficiently structured, these values give $M(n, h) = h$, as seen in the derivation of inequality (10).

Now suppose A fails to achieve one of the three conditions of efficient structure. If A is not acyclic, then $b(A) > n - 1$. If it does not have minimal outer perimeter, then $p_o(A) > p_{min}(n + h)$. And finally if there is a hole of area at least two, then there is at least one hole with at least six edges, and thus $h \leq (p_h(A) - 2)/4$. By equation (9) and inequality (10), all three issues cause $h < M(n, h)$, by decreasing the numerator in the first two cases and increasing the denominator in the third. So if $M(n, h) = h$, then A is efficiently structured. ■

In [4], it was proven that the crystallized polyominoes defined in the K-R sequence satisfy $M(n, h) = h$, and are therefore efficiently structured.

Corollary 16. *If a polyomino A exists with n tiles and h holes such that $M(n, h) = h + 1/2$, then exactly one of the following three things occurs: 1) A has a single dual cycle; 2) A has a single hole with area two, and the rest have area one; or 3) $p_o(A) = p_{min}(n + h) + 2$.*

This is clear from the construction of $M(n, h)$ and the proof of Lemma 15. To understand how equation (11) changes with respect to the minimal outer perimeter of a polyomino, we observe that

$$p_{min}(n + h + 1) = \begin{cases} p_{min}(n + h) + 2 & \text{if } n + h \text{ is square or pronic} \\ p_{min}(n + h) & \text{else.} \end{cases} \quad (12)$$

Adding a single tile also increases $2n$ to $2n + 2$. Then for a fixed $h \geq 1$, $M(n, h)$ is a non-decreasing function of n such that

$$M(n + 1, h) - M(n, h) = \begin{cases} 0 & \text{if } n + h \text{ is a square or pronic number} \\ 1/2 & \text{else} \end{cases} \quad (13)$$

We now give a proof of Theorem 7, which states that any efficiently structured polyomino is crystallized.

Proof of Theorem 7. Let A be an efficiently structured polyomino with n tiles and h holes. By Lemma 15, this is equivalent to $h = M(n, h)$, and by equation (13) we know that $M(n - 1, h) \leq M(n, h)$. If $M(n - 1, h) < M(n, h)$, then by Lemma 14 there does not exist a polyomino with $n - 1$ tiles and h holes, and thus A is crystallized.

Suppose instead that $M(n-1, h) = M(n, h)$, and assume that there exists a polyomino B with $n-1$ tiles and h holes. Then by equation (13) we get that $n-1+h$ is a square or pronic number, and B is efficiently structured by Lemma 15. By Remark 10, a polyomino with minimal outer perimeter and total area equal to a square and pronic number must be constructed in a square or pronic rectangle, and thus the boundary layer of B is completely filled, which gives a cycle. This contradicts efficient structure for B , and therefore there does not exist a polyomino with $n-1$ tiles and h holes. Thus, A is crystallized. ■

The converse of Theorem 7 is not true. Figure 4 shows crystallized polyominoes which fail to attain minimal outer perimeter and thus are not efficiently structured.

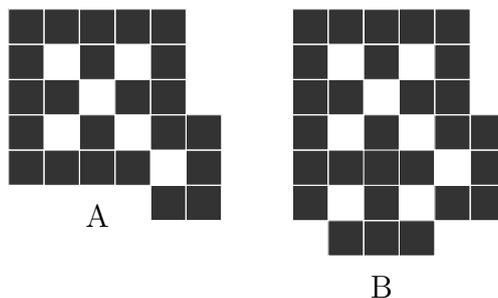


Figure 4: Polyominoes A and B are crystallized polyominoes that are not efficiently structured. Although both are acyclic with each hole having an area of one, neither has minimal outer perimeter. A has 23 tiles and 6 holes and B has 28 tiles and 8 holes. One can see in the calculation that $6 < M(6, 23) = 6.5$, and $8 < M(8, 28) = 8.5$.

In the next section we develop tools to understand for which values of h a crystallized polyomino with h holes may fail to be efficiently structured.

4 Obstructions to Crystallization

Define $m(h) = \min\{n : M(n, h) \geq h\}$. By equation (10) this is a theoretical lower bound for $g(h)$, and therefore if a polyomino exists with $m(h)$ tiles and h holes, this is an immediate proof that $g(h) = m(h)$.

In this section we establish several obstructions to $g(h)$ hitting the optimal value of $m(h)$. In particular, we aim to show that Theorem 5 is in fact equivalent to stating that for all $h \geq 1$,

$$g(h) = \begin{cases} m(h) + 1 & \text{if } h = h_\alpha + 1 \text{ for square or pronic } \alpha \neq (2^l + 1)^2 \\ m(h) & \text{else.} \end{cases} \quad (14)$$

This asserts that $m(h)$ is the right answer except when trying to fit one too many holes into an optimal shape of minimal outer perimeter, and Theorem 8 will follow by characterizing when $M(m(h)+1, h) = h$. To establish these exceptions at values of $h_\alpha + 1$, we begin by examining the numerical and geometric constraints of $m(h)$.

Lemma 17. For every h , $M(m(h), h) = h$. Furthermore, if $0 < \alpha - (m(h) + h) < 3$ for some square or pronic number α , then $m(h + 1) = m(h) + 3$, and otherwise $m(h + 1) = m(h) + 2$.

Proof. The perimeter of a polyomino is always even, so $2n + 2 - p_{min}(n + h)$ is even, and $M(n, h) = C/2$ for some integer C . Fix $h \geq 1$. For $n = 1$, $p_{min}(1 + h) > 6$ and

$$2n + 2 - p_{min}(n + h) = 4 - p_{min}(1 + h) < 0.$$

Then by equation (13), increasing n by one increases $M(n, h)$ by either 0 or $1/2$, and thus $M(n, h)$ hits every positive multiple of $1/2$ as n increases indefinitely. So $M(m(h), h) = h$ and the inequality in its definition is really equality.

Suppose we have $n = m(h)$ for some h . Then since p_{min} is non-decreasing,

$$\begin{aligned} M(n + 1, h + 1) &= \frac{2n + 2 + 2 - p_{min}(n + h + 2)}{4} \\ &\leq \frac{2n + 2 - p_{min}(n + h)}{4} + \frac{1}{2} \\ &= h + \frac{1}{2}. \end{aligned} \tag{15}$$

So $m(h + 1) \geq m(h) + 2$ for all h . By the same calculation,

$$M(m(h) + 2, h + 1) = h + 1$$

whenever $p_{min}(n + h + 3) = p_{min}(n + h)$. By equation (12) this will be the case unless $0 < \alpha - (m(h) + h) < 3$ for some square or pronic number α , in which case $p_{min}(n + h + 3) = p_{min}(n + h) + 2$, and then

$$\begin{aligned} M(m(h) + 2, h + 1) &= \frac{2n + 4 + 2 - p_{min}(n + h + 3)}{4} \\ &= \frac{(2n + 2 - p_{min}(n + h)) + 2}{4} \\ &= h + 1/2. \end{aligned} \tag{16}$$

Adding a third tile does not cross another square or pronic threshold, so by equation (13) this adds another $1/2$ to the function and $m(h + 1) = m(h) + 3$. ■

Similar to the polyiamond case in [6], this points to each additional hole requiring two extra tiles in general, and then a third tile when the total area of the shape expands past a threshold of N^2 or $N(N + 1)$. To study the structure of crystallization near these thresholds, for a fixed $\alpha \in \{N^2, N(N + 1)\}$ define

$$t_\alpha = \max\{h : m(h) + h \leq \alpha\}. \tag{17}$$

Since $g(h) \geq m(h)$, this is a theoretical upper bound for h_α . We derive equations for this sequence, and characterize these cases by how close $m(t_\alpha) + t_\alpha$ gets to α .

Suppose that $m(h) + h = N^2$, and note that such an h is necessarily $h = t_{N^2}$. Then $M(N^2 - h, h) = h$ by definition of $m(h)$, so

$$\begin{aligned} 4h &= 2(N^2 - h) + 2 - 4N \\ &= 2N^2 - 2h + 2 - 4N \\ \implies h &= \frac{(N - 1)^2}{3}. \end{aligned} \tag{18}$$

This is an integer if and only if $N \equiv 1 \pmod{3}$. It is straightforward to check by similar calculations that the values

$$m(t_\alpha) + t_\alpha = \begin{cases} N^2 - 1 & \text{if } \alpha = N^2 \text{ and } N \not\equiv 1 \pmod{3} \\ N(N + 1) & \text{if } \alpha = N(N + 1) \text{ and } N \not\equiv 2 \pmod{3} \\ N(N + 1) - 2 & \text{if } \alpha = N(N + 1) \text{ and } N \equiv 2 \pmod{3}, \end{cases} \tag{19}$$

along with equation (18) yield that

$$t_{N^2} = \begin{cases} \frac{(N - 1)^2}{3} & \text{if } N \equiv 1 \pmod{3} \\ \frac{N(N - 2)}{3} & \text{else,} \end{cases} \tag{20}$$

and

$$t_{N(N+1)} = \begin{cases} \frac{N(N - 1)}{3} & \text{if } N \equiv 0 \text{ or } 1 \pmod{3} \\ \frac{(N + 1)(N - 2)}{3} & \text{if } N \equiv 2 \pmod{3}. \end{cases} \tag{21}$$

Lemma 17 describing the jumps in $m(h)$ ensures that the values in equation (19) are all maximal for $m(h) + h \leq \alpha$, and so the values of h solved for in these calculations give the values of t_α in (20) and (21).

Observe that Theorem 3 is now equivalent to stating that $h_\alpha = t_\alpha$ if $\alpha = (2^l + 1)^2$, and otherwise $h_\alpha = t_\alpha - 1$. Or in words, only the K-R sequence from [4] is able attain the maximal number of holes theoretically possible in a square or pronic rectangle.

And by the jumps of $m(h)$, Theorem 4 asserts that $g(h_\alpha) = m(h_\alpha)$, which we prove by construction in Section 6. Then the fact that equation (14) is equivalent to Theorem 5 follows immediately from Theorems 3 and 4 and the jumps in $m(h)$. Assuming these three main theorems, for which we provide proofs in Sections 6 and 7, we now prove the structural characterization given in Theorem 8.

Proof of Theorem 8. By equation (14), a crystallized polyomino with h holes will either have $m(h)$ tiles and be efficiently structured by Lemma 15, or it will have $g(h) =$

$m(h) + 1$ tiles. By equation (13), $M(m(h) + 1, h) = h$ when $m(h) + h$ is a square or pronic number, and otherwise $M(m(h) + 1, h) > h$. And by equations (18) and (19), $m(h) + h$ is a square or pronic number precisely when $h = t_\alpha$ for $\alpha \notin S \cup R \cup \{N^2 : N = 2^l + 1\}$.

Therefore, by equation (14), the set $S \cup R$ is exactly the set of α for which

$$g(h_\alpha + 1) = m(h_\alpha + 1) + 1,$$

and $m(h_\alpha + 1) + 1$ tiles and $h_\alpha + 1$ holes is not efficiently structured.

It only remains to show that the efficient condition which fails for $\alpha \in S \cup R$ is minimal outer perimeter. Then let $h = h_\alpha + 1$ for $\alpha \in S \cup R$. By Lemma 17 and equation (14), $g(h_\alpha + 1) - g(h_\alpha) = 3$. But observe that the constant C in Theorem 4 is either four or five for $\alpha \in S \cup R$. Then

$$g(h_\alpha + 1) + (h_\alpha + 1) = (g(h_\alpha) + 3) + h_\alpha + 1 = g(h_\alpha) + h_\alpha + 4,$$

and

$$\alpha - 1 \leq g(h_\alpha) + h_\alpha + 4 \leq \alpha.$$

Therefore

$$p_{min}(g(h_\alpha + 1) + (h_\alpha + 1)) = p_{min}(\alpha).$$

By Remark 10, all bounding rectangles with outer perimeter $p_{min}(\alpha)$ which are not a square or pronic rectangle have area at most $\alpha - 1$. A construction with $g(h_\alpha + 1)$ tiles and $h_\alpha + 1$ holes cannot exist in a shape of area at most $\alpha - 1$, because it either has area α and is too big, or it has area $\alpha - 1$ and by the reasoning in the proof of Theorem 7 it would necessarily fill the entire boundary layer and create a cycle. By definition of h_α , an extra hole cannot fit in the square or pronic rectangle of area α , and therefore such a construction does not achieve minimal outer perimeter.

By equation (13), we know that

$$M(m(h_\alpha) + 1, h_\alpha + 1) = M(m(h_\alpha + 1), h_\alpha + 1) + 1/2 = (h_\alpha + 1) + 1/2.$$

Then by Corollary 16, only one efficient condition can fail. So a crystallized polyomino with $h_\alpha + 1$ holes for $\alpha \in S \cup R$ is acyclic, each of its holes has an area of one, and it fails to achieve minimal outer perimeter, and all other crystallized polyominoes are efficiently structured. ■

We proceed by showing that $g(t_\alpha) > m(t_\alpha)$ whenever $\alpha \neq (2^l + 1)^2$, and determine that this will also require increasing the outer perimeter to get a crystallized polyomino with t_α holes.

4.1 Checkerboard Obstructions

Let the *checkerboard partition* of a rectangle in the square lattice refer to the bipartition of its squares into sets W and B , where the squares alternate between W and B in every row and column like a checkerboard. For convenience we always assume that the top left corner is in W . When a polyomino is forced to use up a sufficient amount of its boundary

layer, then holes in the interior will necessarily be contained in one set of this partition. Recall that D_1 refers to the corner-less rectangular boundary layer, and D_2 has all spaces of it's boundary filled except for a single corner.

Lemma 18 (Checkerboard Lemma). *Let A be an acyclic polyomino with n tiles and h holes, each hole having an area of one. If A has a rectangular interior, then the outermost layer of the interior alternates between holes and tiles, and the set of holes in A is completely contained in the set W in the checkerboard partition of the interior.*

Proof. In general any empty spaces in the interior of a polyomino are contained in the bounded components of its complement in the plane, and are thus part of the holes. So two adjacent empty spaces anywhere in the interior contribute to a hole of area at least two. Then if A has rectangular interior, its boundary contains D_1 , so in the outermost layer of the interior any two adjacent filled spaces form a dual cycle with their two adjacent boundary tiles. These are both contradictions to our assumptions, and hence the spaces in the outermost layer of the interior must alternate.

Furthermore, at least one of the the corners of this layer must be a hole, since filling in all four would connect the boundary sections of D_1 into a cycle (see Figure 5). Then by rotations and reflections, we may assume the top left corner is empty and therefore in the outermost layer of the interior all spaces in W are holes and all spaces in B are tiles.

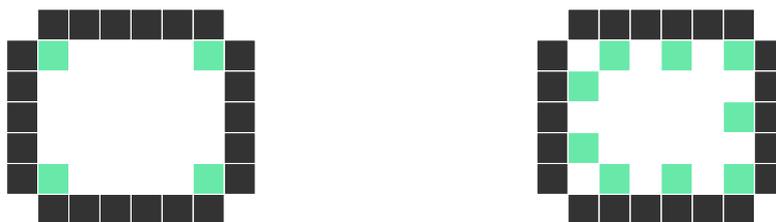


Figure 5: Left: A dual cycle is created if all four interior corners are filled. Right: An alternating interior layer, with holes in W and green tiles in B .

Now suppose A has holes contained somewhere in B , and let $B' \subseteq B$ be a maximal connected set of holes in A . All spaces in W which are adjacent to B' must be filled, since every hole must have an area of one. And since B' is maximal with respect to being connected, all spaces in B which are corner adjacent to B' must either be filled, or not be contained in the interior of A . But spaces in B' are separated from the boundary by the alternating layer where all spaces of B are filled, so all such spaces in B which are corner adjacent to B' are in the interior of A and must be filled.

The hole graph of a polyomino is acyclic, as noted following Definition 11. Therefore the portion of the hole graph corresponding to B' is a tree. We claim that for any acyclic set of holes, each hole having an area of one, which has all of its corner adjacent spaces filled with tiles, the tiles surrounding the holes form a cycle. This is clearly the case for a single hole, and any tree can be constructed by fixing a root and adding vertices of degree one, one at a time.

So assume that the statement holds for every such acyclic set of at most k holes, each hole with an area of one. A new hole cannot be attached to more than one existing vertex, as this would create a cycle. Thus any new hole requires filling in the seven surrounding tiles which are not the corner adjacent hole it is being attached to, some of which are already filled in A . This adds an indent to the cycle where the new hole was attached, but it remains a cycle, and by induction if A has an acyclic set of holes, each hole with an area of one, with all of its corner adjacent tiles filled, then A has a dual cycle.

This is in contradiction with our assumption of A being acyclic. So all holes of A must be contained in W and all spaces of B must be filled. ■

Recall that it was pointed out in the proof of Theorem 7 that the total area of an efficiently structured polyomino cannot be equal to a square or pronic number. It would need to be constructed in the square or pronic rectangle with that area, and this would require the entire boundary to be filled, creating a dual cycle.

Furthermore, suppose the total area of an efficiently structured polyomino is $\alpha - 1$. Then it must again be constructed in the square or pronic rectangle of area α , since a smaller rectangle with minimal perimeter has total area at most $\alpha - 1$ by Remark 10, and by the same argument a dual cycle would be created. And then it has exactly one empty boundary space, which must be a corner to preserve minimal perimeter, and therefore it has boundary D_2 .

Combining this observation regarding the total area with the Checkerboard Lemma is already sufficiently restrictive to rule out $m(t_\alpha)$ in most cases.

Lemma 19. *If $m(t_\alpha) + t_\alpha = \alpha$, then $g(t_\alpha) > m(t_\alpha)$ and $h_\alpha \leq t_\alpha - 1$.*

Proof. That $g(t_\alpha) > m(t_\alpha)$ follows directly from the preceding observation regarding total area. The only difference is that here we acknowledge that the only time $m(h) + h$ is a square or pronic number is if $h = t_\alpha$ for $\alpha = N^2$ with $N \equiv 1 \pmod{3}$, or $\alpha = N(N + 1)$ with $N \not\equiv 2 \pmod{3}$.

Then if $m(t_\alpha) + t_\alpha = \alpha$, any crystallized polyomino with t_α holes has total area at least $\alpha + 1$. This cannot fit in a square or pronic rectangle of area α , and therefore $h_\alpha \leq t_\alpha - 1$. ■

Lemma 20. *If $2 \mid \alpha$, then $g(t_\alpha) > m(t_\alpha)$ and $h_\alpha \leq t_\alpha - 1$.*

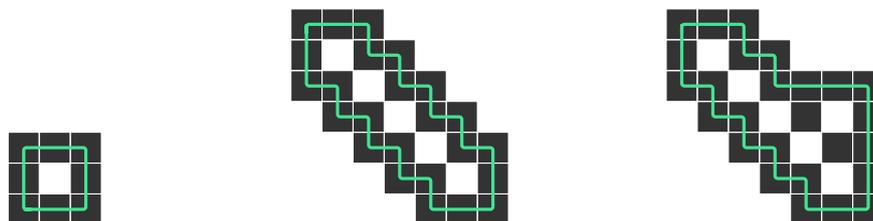


Figure 6: Dual cycle around one hole, around a tree, and then extended around a single extra hole.

Proof. Lemma 19 already covers the case here when $m(t_\alpha) + t_\alpha = \alpha$. Otherwise, by the derivations of equations (20) and (21), if $\alpha = N^2$ then $m(t_\alpha) + t_\alpha = \alpha - 1$, and if $\alpha = N(N + 1)$ then $m(t_\alpha) + t_\alpha = \alpha - 2$.

Case 1: Assume that $2 \mid \alpha = N^2$, and suppose that A is a crystallized polyomino with t_α holes and $m(t_\alpha)$ tiles such that $m(t_\alpha) + t_\alpha = \alpha - 1$. Then A 's bounding rectangle must be an $N \times N$ square by the observation regarding total area, which then also asserts that A must have boundary D_2 .

By the Checkerboard Lemma, the spaces of B in the checkerboard partition of the interior are all filled. So on each of the two parallel sides of even length, one of the corners of the interior must be filled. But three of the corners of the boundary layer are also filled, so by the pigeonhole principle this creates at least one dual cycle of length four in a corner, which contradicts efficient structure. Hence $g(t_\alpha) > m(t_\alpha)$.

Furthermore, if $h_\alpha = t_\alpha$, then there is a polyomino A fitting into the square of area α with t_α holes and $m(t_\alpha) + 1$ tiles. Thus A must have total area α , and all tiles of the boundary layer must be filled. But none of the corner tiles bound any holes, and taking one away will not disconnect A , which would imply that in fact $g(t_\alpha) \leq m(t_\alpha)$. By the preceding argument this is a contradiction, and hence $h_\alpha \leq t_\alpha - 1$.

Case 2: If instead $\alpha = N(N + 1)$ and $m(t_\alpha) + t_\alpha = \alpha - 2$, then similarly by Remark 10 any other rectangle with this perimeter has area at most $\alpha - 2$, and filling a shape of this area would create a dual cycle in the boundary layer, so a pronic rectangle is required. Two tiles on the boundary must be empty, and to maintain minimum perimeter they must be two corners or a corner and one of its adjacent spaces. As before, two interior corners must be filled, and then both of their incident corners in the boundary must be empty to avoid dual cycles. But there are precisely two empty spaces which are not holes or tiles, so the rest of the boundary is filled and the indented corners create a dual cycle, which is again a contradiction. Hence $g(t_\alpha) > m(t_\alpha)$.

Similarly to the previous case, if $h_\alpha = t_\alpha$, then a polyomino A in the pronic rectangle of area α which has t_α holes and $m(t_\alpha) + 1$ tiles has total area $\alpha - 1$, and exactly one empty boundary space. Then if the outermost layer of the interior alternates between tiles and holes, there must again be two interior corners which are filled. At least one of those boundaries must have the corner and its two adjacent boundary spaces filled, creating a dual cycle of length four. And the corner tile there does not bound any holes and can be removed without disconnecting A , which implies that $g(t_\alpha) \leq m(t_\alpha)$, which is a contradiction. Therefore it suffices to show that the outermost layer of the interior alternates.

However, the Checkerboard Lemma no longer applies, since $M(n, h)$ will be $h + 1/2$, which forces exactly one of the three conditions of efficient structure to fail by Corollary 16. Since we are restricting the shape to be contained in a pronic rectangle, the only way to increase the perimeter while filling a total area of $\alpha - 1$ is if the empty space of the boundary is not in a corner. But then A would be acyclic and only have holes of area one, so the Checkerboard Lemma would assure that the outermost layer of the interior

alternates. Otherwise the empty boundary space is in a corner and there is either a single dual cycle or the existence of a hole with area exactly two.

Suppose that there are two adjacent tiles in the outermost layer of the interior, creating a dual cycle of length four with the adjacent tiles of the boundary. Since there cannot be any further cycles or holes of area two, the rest of this layer must alternate, and it has $2(N - 2) + 2(N - 3)$ spaces, which is even. Then surrounding these two adjacent tiles must be two holes, one on either side. And then two tiles, one on the other side of each hole, and then two holes, and so on. Since there are an even number of spaces this will end with either a hole of area two or two tiles placed next to each other creating another cycle, which is a contradiction.

The same is true if we start with a hole of area two, and therefore the spaces in this layer alternate between holes and tiles, which implies by the above that $g(t_\alpha) \leq m(t_\alpha)$, which is a contradiction. Hence $h_\alpha \leq t_\alpha - 1$. ■

Lemmas 19 and 20 prove that $g(t_\alpha) \geq m(t_\alpha) + 1$ and $h_\alpha \leq t_\alpha - 1$ for all odd squares with side length $N \equiv 1 \pmod 3$, all even squares, and all pronic rectangles. This leaves only odd squares which have side lengths $N \equiv 0$ or $2 \pmod 3$.

5 Expansion and Compression

To determine what happens for the remaining odd squares, we develop a technique for lifting efficient arrangements from smaller crystallized polyominoes, and refer to this process as *expansion*. We refer to a set of five tiles arranged to have a central tile with one tile adjacent at each edge as a “plus,” and a set of five holes arranged to have a central hole with one hole adjacent at each corner as an “X” (see Figure 7).

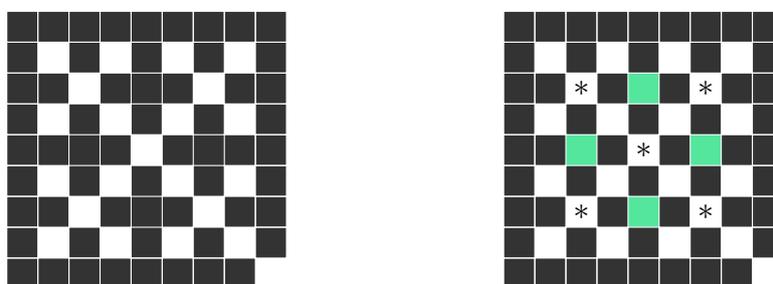


Figure 7: A polyomino with four pluses and five X’s. On the right the central tile of each plus is marked in green, and the central hole of each X with an *.

For an odd integer N , let P_N be the sub-arrangement of tiles and holes in the interior of the $N \times N$ rectangle in which: (a) all spaces in B of the checkerboard partition are filled, and (b) all spaces in the odd rows of W , enumerated from the top down, are empty. The remaining spaces are considered to be undetermined.

P_N is not itself a polyomino since its set of tiles is not connected, but it will play an integral role in the expansion process. Observe that the undetermined spaces of P_N form

a square grid of dimensions $(N - 3)/2 \times (N - 3)/2$. If any undetermined space is filled with a tile it will create a plus, while any hole in these spaces will create an X. We let U_N denote the set of spaces left undetermined by P_N . The process of expansion will be to determine these spaces by using the interior of an $(N + 1)/2 \times (N + 1)/2$ square as a template.

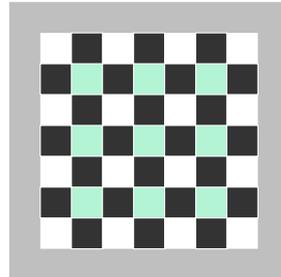


Figure 8: P_9 in the interior of the 9×9 square, with tiles in black, holes in white, and undetermined spaces in green.

Definition 21. Let A be an arrangement of tiles and holes in the $N \times N$ square with boundary D for some $D_1 \subseteq D \subseteq D_2$. The *expansion* of A is the arrangement $E(A)$ in the $(2N - 1) \times (2N - 1)$ square with boundary D , subarrangement P_{2N-1} , and in which the spaces of U_{2N-1} are equivalent to the corresponding spaces in the interior of A .

Definition 22. Conversely, an arrangement A of tiles and holes in the $N \times N$ square which can be written as $A = D \cup P_N \cup U_A$ for some $D_1 \subseteq D \subseteq D_2$ and a specified subarrangement U_A in the spaces of U_N is called *compressible*, and the *compression* of A is $C(A)$, the arrangement in the $(N + 1)/2 \times (N + 1)/2$ square with boundary D and interior equivalent to U_A .

These processes are inverses, with $E(C(A)) = A$ for any compressible A . They also preserve the properties of connectivity in both the hole graph and the dual graph.

Lemma 23. *Let A be a compressible arrangement. Then a pair of adjacent tiles in $C(A)$ expands to a path of length two in A , and a pair of corner adjacent empty spaces in $C(A)$ expands to a path of length two in the hole graph of A .*

Proof. This comes directly from the structure of $P_N \subset A$. Two adjacent interior tiles of $C(A)$ expand to two pluses which intersect in a connecting tile from P_N , and a tile adjacent to the boundary layer of $C(A)$ corresponds to a tile of U_A which is connected to the boundary layer by a tile of P_N (see Figure 9).

Moreover, for any filled space of U_A the four tiles of P_N surrounding it can be thought to extend the four edges of the corresponding tile from $C(A)$. So when adjacent spaces of $C(A)$ are filled, the intermediate tile of P_N is a subdivision of that edge in the expansion. The same holds for empty spaces, where two corner adjacent empty spaces of $C(A)$ create two X's in A which intersect in an empty space of P_N (see Figure 10). ■

While we do not require that an arrangement A is a polyomino in the definitions of expansion and compression, Lemma 23 allows us to determine when the property of being an acyclic polyomino is preserved by these processes.

Lemma 24 (Compression Lemma). *A compressible arrangement A is an acyclic polyomino if and only if $C(A)$ is an acyclic polyomino with each of its holes having an area of one.*

Proof. We prove this by exhibiting the contrapositives in both directions. First Let A be a compressible arrangement. Suppose $C(A)$ has at least two adjacent empty spaces, forming a hole of area at least two. Then the tile separating the corresponding holes of U_A is isolated and A fails to be a polyomino (see Figure 9).



Figure 9: Cycles are preserved by expansion and compression (in green). Two adjacent empty spaces in $C(A)$ disconnect A (in red). Note that $C(A)$ is a polyomino, but A is not.

Next suppose that $C(A)$ has a dual cycle. If the cycle does not use any of the boundary layer, then A contains a cycle by Lemma 23. Otherwise the cycle uses the boundary. But any connected section of the boundary used by the cycle in $C(A)$ is also connected in A since by definition they have the same boundary structure, and this again forms a cycle (see Figure 9).

Now suppose that $C(A)$ is not connected, and thus not a polyomino. Then there must be a connected set of holes which either forms a cycle or is a path that connects to two empty boundary corner spaces at its ends, separating the tiles of $C(A)$ into at least two disjoint pieces (see Figure 10). As in the preceding argument, by Lemma 23 a cycle of holes expands to a cycle of holes in A , which would force A to be disconnected. And a path of holes connecting two empty boundary corners expands to a path connecting two empty boundary corners of A , which is then disconnected. Thus, if A has connected interior and is acyclic, then $C(A)$ is connected, acyclic, and all holes have area one.

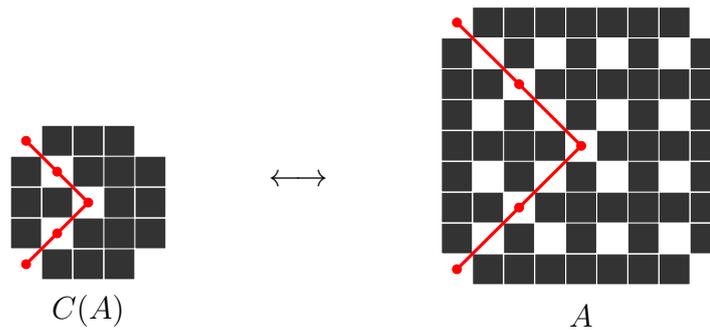


Figure 10: A path of holes connecting two empty corners disconnects both A and $C(A)$.

On the other hand, if A does not have connected interior, then it has either a cycle of holes or a path of holes connecting two empty boundary corners. Any path of holes must alternate between holes in P_N and holes in U_A . Then suppose two holes in U_A are mutually adjacent to a hole in P_N . These holes are either consecutive spaces of U_A along a row or column, which would compress to a hole of area at least two in $C(A)$ (see Figure 9), or they are diagonal from each other and thus corner adjacent in $C(A)$. If no holes of area at least two are created, then a cycle compresses to a cycle and a path connecting two empty boundary corners of A compresses to a path connecting two boundary corners of $C(A)$ (see Figure 10). Thus either a hole of area at least two is created, or $C(A)$ is disconnected.

Similarly, if there is a cycle of tiles in A , then the tiles of the cycle in the interior alternate between P_N and U_A , and two tiles in this cycle from consecutive spaces of U_A compress to adjacent tiles in $C(A)$. Since A and $C(A)$ have the same boundary, this cycle compresses to a cycle in $C(A)$, as in the previous argument for expanding a cycle (see Figure 9). Hence if $C(A)$ is connected, acyclic, and only has holes of area one, then A has connected interior and is acyclic. ■

Observe that if A is compressible, then it cannot have holes of area more than one, and when $D = D_2$ both A and $C(A)$ will necessarily achieve the minimum perimeter for their total area.

Corollary 25. *A compressible acyclic polyomino A with boundary D_2 is efficiently structured if and only if $C(A)$ is efficiently structured.*

Moreover, containing the subarrangement P_N is a necessary condition for crystallization of a polyomino contained in an odd $N \times N$ square with boundary layer $D_1 \subseteq D \subseteq D_2$.

Lemma 26. *Let $N \geq 5$ be an odd positive integer. If A is an efficiently structured polyomino with square $(N - 2) \times (N - 2)$ interior, then all spaces in the odd rows of W in the checkerboard partition of the interior must be holes.*

Proof. For N odd, since we assume the top left corner of the interior is in W , all corners of the interior are in W by parity. Then by the Checkerboard Lemma and by parity, the odd rows of W are precisely those which have holes in the outermost layer of the interior.

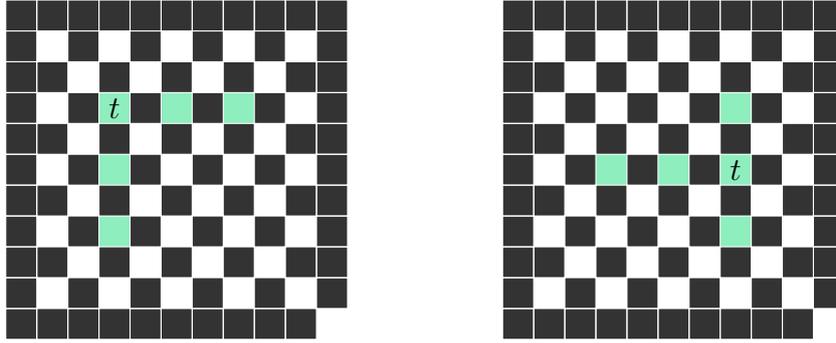


Figure 11: Tiles in odd rows of W cannot connect to the boundary.

If a space in an odd row of W is filled by tile t , then the spaces which are corner adjacent to t in even rows of W must be empty to avoid creating dual cycles. So the only way to connect t to the boundary in order to satisfy the connected interior condition for polyominoes is by filling a path of spaces in the odd rows and columns of W . But as noted above, these rows and columns coincide with the holes in the outermost layer of W , and thus cannot connect to the boundary without creating dual cycles. Therefore they must all be holes. ■

Corollary 27. *Let $N \geq 5$ be an odd integer. If A is an efficiently structured polyomino with square $(N - 2) \times (N - 2)$ interior, then A is compressible.*

This follows immediately from the Checkerboard Lemma and Lemma 26. In particular, for N odd with $N \not\equiv 1 \pmod{3}$, we have that $m(t_{N^2}) + t_{N^2} = N^2 - 1$ and an efficiently structured polyomino with t_{N^2} holes is compressible and has boundary D_2 . This leads to a final obstruction to $g(t_\alpha) = m(t_\alpha)$, and a proof of Theorem 9.

Lemma 28. *Let N be an odd positive integer such that $N \not\equiv 1 \pmod{3}$, and $N \neq 2^l + 1$ for any positive integer l . Then $g(t_{N^2}) > m(t_{N^2})$, and $h_{N^2} \leq t_{N^2} - 1$.*

Proof. Let $N = 2^l p + 1$ for some $l \geq 1$ and some $p > 1$ for which $2 \nmid p$, and suppose that A is a polyomino with t_{N^2} holes and $m(t_{N^2})$ tiles. For $N \not\equiv 1 \pmod{3}$ this gives total area $N^2 - 1$. By Corollary 27, A has boundary D_2 and is compressible. And A is efficiently structured by Lemma 15, so by Corollary 25 $C(A)$ is an efficiently structured polyomino with boundary D_2 , and is itself compressible if $(N + 1)/2$ is odd.

Then let $N_0 = N$, and let $N_i = (N_{i-1} + 1)/2$ be the length of the square for $C^i(A)$, the i -th compression of A . Since $N = 2^l p + 1$, we have that $N_i = 2^{l-i} p + 1$, which is odd for all $i < l$. Therefore l compressions can be performed, with $C^l(A)$ an efficiently structured polyomino with boundary D_2 in an $N_l \times N_l$ square for even $N_l = p + 1 \geq 4$. With D_2 as its boundary layer, the total area of $C^l(A)$ must be $N_l^2 - 1$, but by Lemma 20 there cannot be an efficiently structured polyomino with this area in an even square. Therefore $g(t_{N^2}) > m(t_{N^2})$.

And if $h_{N^2} = t_{N^2}$, then by the preceding argument a crystallized polyomino with h_{N^2} holes must have total area at least N^2 . If it has total area exactly N^2 , then the boundary

layer is completely filled and forms a cycle. As in the proof of Lemma 20, one of the corner tiles is extraneous and can be removed, which is a contradiction. Otherwise the total area is more than N^2 and such a polyomino cannot fit in this square. Therefore $h_{N^2} \leq t_{N^2} - 1$. ■

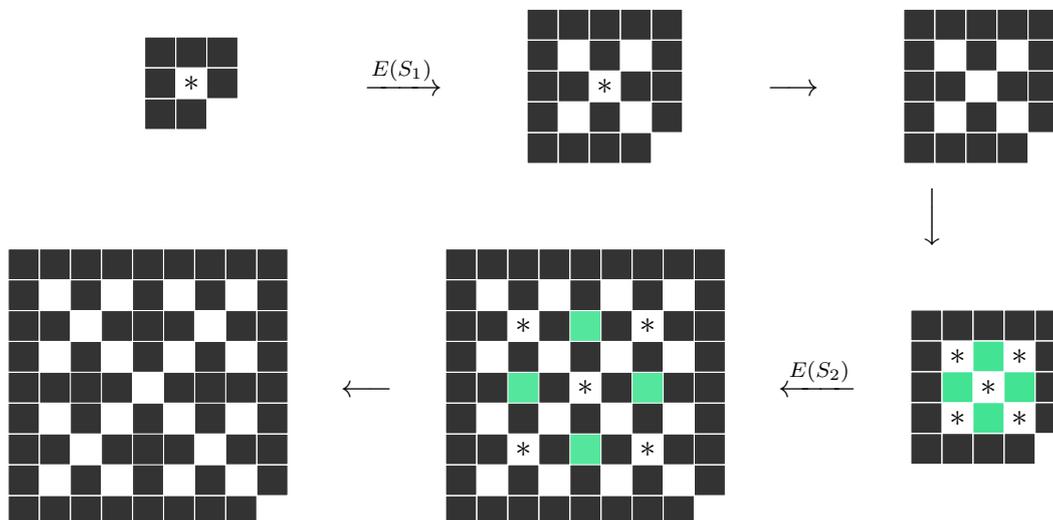


Figure 12: Expansion of S_1 , with $E(S_1) = S_2$ and $E^2(S_1) = E(S_2) = S_3$.

Proof of Theorem 9. For $N = 2^l + 1$, the existence of crystallized polyominoes with $t_{N^2} = (2^{2l} - 1)/3$ holes and $m(t_{N^2}) = [(2^{2l+1} + 3 \cdot 2^{l+1} + 4)/3] - 1$ tiles was proved in [4]. We prove the uniqueness of those constructions by simply compressing them all down to S_1 , the unique crystallized polyomino with a single hole (see Figure 1).

In particular fix $l > 1$, let $N = 2^l + 1$ and $N_i = (N_{i-1} + 1)/2 = 2^{l-i} + 1$, and let A be a compressible polyomino with $(2^{2l} - 1)/3$ holes and $[(2^{2l+1} + 3 \cdot 2^{l+1} + 4)/3] - 1$ tiles. Since $2^l + 1 \not\equiv 1 \pmod{3}$, this has total area $N^2 - 1$, and thus A has boundary D_2 . For $i < l$ all of the N_i are odd, with $N_{l-1} = 3$. Therefore $C^i(A)$ is compressible with boundary D_2 for all $i < l$, and thus $C^{l-1}(A) = S_1$. But compression and expansion are inverses, so if two polyominoes A and A' have that $C^k(A) = C^k(A')$, then $A = A'$. In particular the polyominoes S_l are the unique crystallized polyominoes with $(2^{2l} - 1)/3$ holes, and $S_l = E^{l-1}(S_1)$. ■

The method of expansion and compression then gives an alternate way to construct the sequence S_l . For all other squares and pronic numbers, we give constructions in Section 6 which show that $h_\alpha = t_\alpha - 1$. Assuming these constructions, we use expansion to prove Theorems 3 and 4.

Proof of Theorems 3 and 4. By Lemmas 19 and 20, Given the constructions of Section 6 and the sequence S_l , we have crystallized polyominoes with h_α holes and $\alpha - h_\alpha - C$

tiles for the appropriate C , except when $\alpha = N^2$ for odd N such that $N \not\equiv 1 \pmod 3$ and $N \neq 2^1 + 1$ for any positive integer l .

Consider such an N , which can be written as $N = 2^l p + 1$ for $2, 3 \nmid p$. Then as in the proof of Lemma 28, $p + 1$ is even and we can take a crystallized polyomino A with h_α holes for $\alpha = (p + 1)^2$ and expand it l times to get a crystallized polyomino $C^l(A)$ in an $N \times N$ square. Since $3 \nmid p$, $p + 1 \not\equiv 1 \pmod 3$, and A will be the appropriate construction from Section 6.2 or 6.3, each with boundary D_1 . Then $C^l(A)$ has boundary D_1 and total area $N^2 - 4$. By equation (20), Lemma 17, and Lemma 28, $C^l(A)$ has $h_{N^2} = t_{N^2} - 1$ holes and $N^2 - h_{N^2} - 4$ tiles. ■

6 Constructions of Crystallized Polyominoes

We call a polyomino which is a sequence of pluses which overlap on a non-central tile a *plus tree*. We then refer to the place where such a structure in the interior of a polyomino connects to a section of the boundary layer as its root. The basic elements of our constructions involve choosing an appropriate boundary layer, and filling the interior with efficiently spaced plus trees. A plus tree growing in a certain direction is three spaces wide, and a disjoint plus tree is not allowed to connect to any of its tiles or fill any of the empty spaces between the pluses.

These polyominoes all have $t_\alpha - 1$ holes and will be efficiently structured by construction, and are thus an immediate proof of Theorems 3 and 4. For $N \equiv l \pmod 3$, we denote the k -th element of the even square sequences by $S_{l,k}$, and the k -th element of the pronic rectangle sequences by $R_{l,k}$, where N is the shorter of the two side lengths. All equations for total area are taken from the appropriate cases in equations (19) and (20), and Lemma 17.

6.1 Even Squares, $N \equiv 1 \pmod 3$

The polyomino $S_{1,k}$ is made up of $2k$ vertical plus trees with roots alternating between the bottom and top boundaries from left to right, each with $3k$ pluses, and has boundary D_1 with the top left corner additionally filled. To satisfy the Checkerboard Lemma we also require that the spaces of the outermost layer of the interior alternate between tiles and holes, with a hole in the top left corner.

Each plus tree takes up three of its own columns, and along with the boundary and outermost layer of the interior on each side the total width is $N = 4 + 6k$. Similarly, for a tree rooted at the bottom, count each hash of a plus along with the row above it, then there are two additional rows above and two below, and the height is also $4 + 6k$. This length is always $1 \pmod 3$, so the minimal total area for $(t_{N^2} - 1)$ holes is $N^2 - 3$.

Since N is even the bottom left and top right corners will be indented, and the boundary sections are all connected. Then $S_{1,k}$ achieves total area $(4 + 6k)^2 - 3$, and by con-

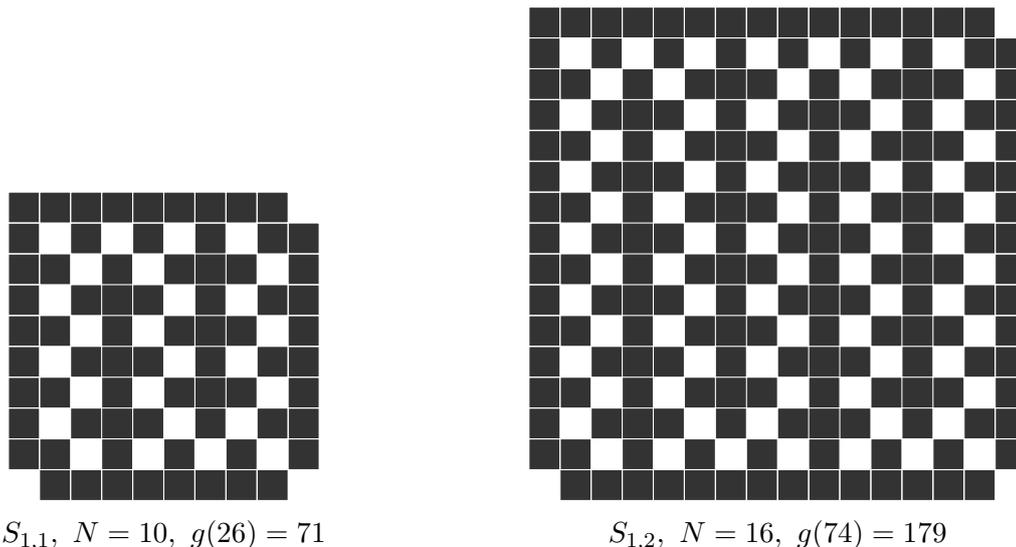


Figure 13: The first two elements of the sequence $S_{1,k}$.

struction it is efficiently structured. Hence $S_{1,k}$ is crystallized for all $k \geq 1$, with

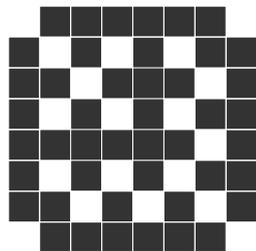
$$\begin{aligned}
 h(S_{1,k}) &= \frac{(N-1)^2}{3} - 1 = \frac{(3+6k)^2}{3} - 1 = 12k^2 + 12k + 2, \\
 |S_{1,k}| &= (4+6k)^2 - h(S_{1,k}) - 3 = 24k^2 + 36k + 11.
 \end{aligned}
 \tag{22}$$

6.2 Even Squares, $N \equiv 2 \pmod{3}$

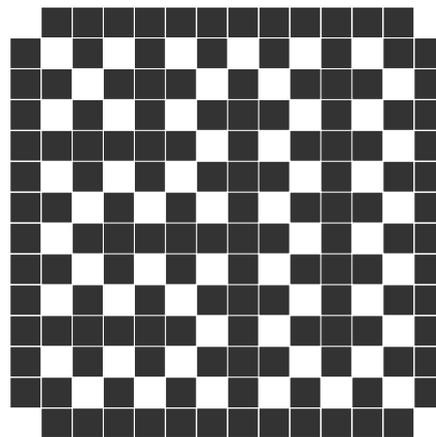
We construct $S_{2,k}$ with boundary D_1 , and make the upper right hand corner equivalent to S_2 to connect the two disjoint boundary sections. To the right of this there are $2k-2$ vertical plus trees, with the roots alternating from the bottom to the top, and below the copy of S_2 there are $2k-2$ horizontal plus trees with two pluses each, with the roots alternating between the left boundary and the leftmost vertical plus tree from the bottom up.

The vertical plus trees, the 2 additional columns to the right boundary and one column separating them from S_2 , which has width five, makes the total width $N = 3(2k-2) + 3 + 5 = 6k + 2$, which is always $2 \pmod{3}$. There are the same number of horizontal plus trees as vertical ones, and the rest of spacing is the same, so the height is also $6k + 2$. For $N \equiv 2 \pmod{3}$, the minimal total area for $t_{N^2} - 1$ holes is $N^2 - 4$. So $S_{2,k}$ is efficiently structured by construction, and thus crystallized for all $k \geq 1$, with

$$\begin{aligned}
 h(S_{2,k}) &= \frac{N(N-2)}{3} - 1 = \frac{(6k+2)6k}{3} - 1 = 12k^2 + 4k - 1, \\
 |S_{2,k}| &= (6k+2)^2 - h(S_{2,k}) - 4 = 24k^2 + 20k + 1.
 \end{aligned}
 \tag{23}$$



$S_{2,1}$, $N = 8$, $g(15) = 45$



$S_{2,2}$, $N = 14$, $g(55) = 137$

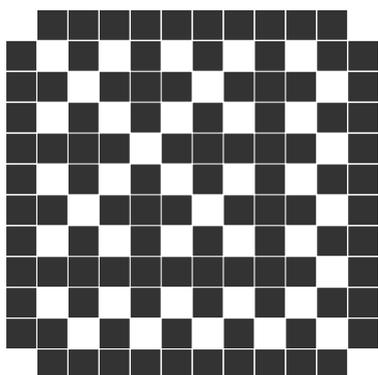
Figure 14: The first two elements of the sequence $S_{2,k}$.

6.3 Even Squares, $N \equiv 0 \pmod{3}$

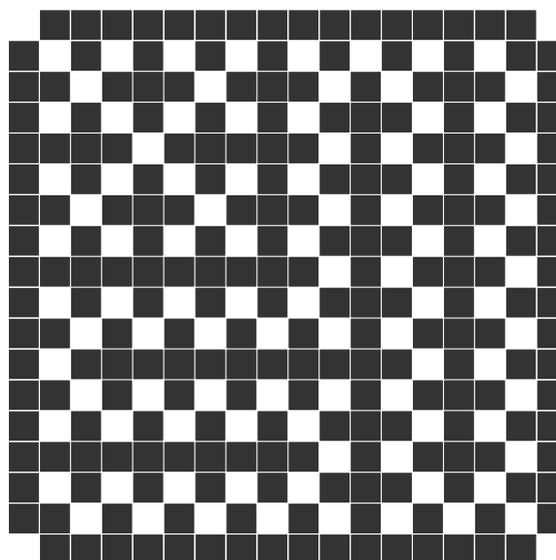
The construction for $N \equiv 0 \pmod{3}$ is the same as the previous case, but with S_3 instead of S_2 in the upper left corner. This increases the side length to $N = 3(2k-2)+3+9 = 6k+6$, and the minimal total area for $t_{N^2} - 1$ holes is $N^2 - 4$. Therefore by construction $S_{0,k}$ is an efficiently structured polyomino, and hence crystallized for all $k \geq 1$, with

$$h(S_{2,k}) = \frac{N(N-2)}{3} - 1 = \frac{(6k+6)(6k+4)}{3} - 1 = 12k^2 + 20k + 7, \quad (24)$$

$$|S_{2,k}| = (6k+6)^2 - h(S_{2,k}) - 4 = 24k^2 + 52k + 25.$$



$S_{0,1}$, $N = 12$, $g(39) = 101$



$S_{0,2}$, $N = 18$, $g(95) = 225$

Figure 15: The first two elements of $S_{0,k}$.

6.4 Pronic Rectangles, $N \equiv 0 \pmod{3}$

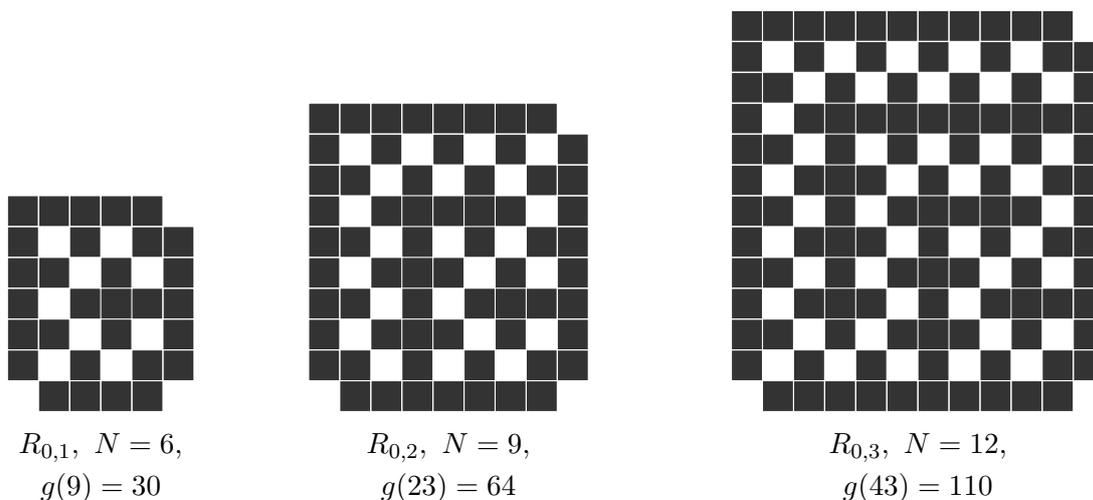


Figure 16: The first three elements of $R_{0,k}$.

Let $R_{0,k}$ have boundary D_1 with the top left corner filled. The interior is then filled with k plus trees whose roots alternate between the right boundary and the bottom boundary. Right-rooted trees move left and then turn straight down, and bottom-rooted trees move up and then turn right, fitting tightly around the previous tree.

The height has k plus trees and two additional rows on top and bottom, for a total of $3k + 4$. The width is one less than the height, since the first tree is rooted on the side and so there are only one column on the right which is not accounted for by the trees, instead of the usual two. Thus the bounding rectangle has dimensions $N \times (N + 1)$ for $N = 3k + 3$, and by construction $R_{0,k}$ has efficient structure with total area $N(N + 1) - 3$. Therefore $R_{0,k}$ is crystallized for all $k \geq 1$, with

$$\begin{aligned}
 h(R_{0,k}) &= \frac{N(N-1)}{3} - 1 = \frac{(3k+3)(3k+2)}{3} - 1 = 3k^2 + 5k + 1, \\
 |R_{0,k}| &= N(N+1) - h(R_{0,k}) - 3 = 6k^2 + 16k + 8.
 \end{aligned}
 \tag{25}$$

6.5 Pronic Rectangles, $N \equiv 0 \pmod{3}$

For $N \equiv 1 \pmod{3}$ we define $R_{1,k}$ to be the same as $R_{0,k}$, but with the initial tree having two pluses. In this case the width increases by two, and the height is the same, so its bounding rectangle is $N \times (N + 1)$ for $N = 3k + 4$. The equation in terms of N is the same as the above, and so $R_{1,k}$ is crystallized for all $k \geq 1$, with

$$\begin{aligned}
 h(R_{1,k}) &= \frac{N(N-1)}{3} - 1 = \frac{(3k+4)(3k+3)}{3} - 1 = 3k^2 + 7k + 3, \\
 |R_{1,k}| &= N(N+1) - h(R_{1,k}) - 3 = 6k^2 + 20k + 14.
 \end{aligned}
 \tag{26}$$

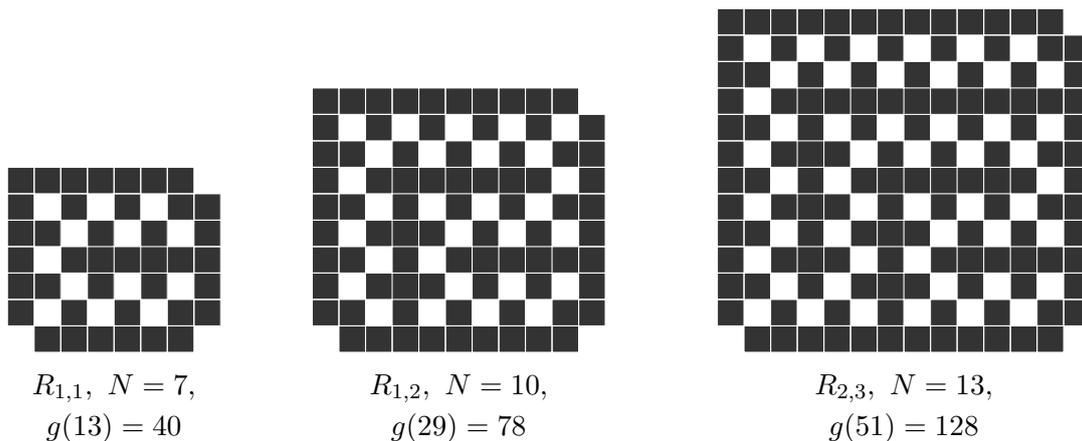


Figure 17: The first three elements of $R_{1,k}$.

6.6 Pronic Rectangles, $N \equiv 2 \pmod 3$

Finally we have the *double spiral* construction for $N \equiv 2 \pmod 3$. In this case, the minimal total area for $t_{N(N+1)} - 1$ holes is $N(N+1) - 5$. We orient $R_{2,k}$ so that the odd length is the width and the even length is the height, with all five empty spaces of the boundary in the bottom row; three from the left corner and two from the right corner. The interior is checkered and filled with two concentric plus tree spirals, one starting on the left of the bottom boundary and one on the bottom of the right boundary. They indent at the bottom left corner, and otherwise turn at right angles as depicted in Figure 18.

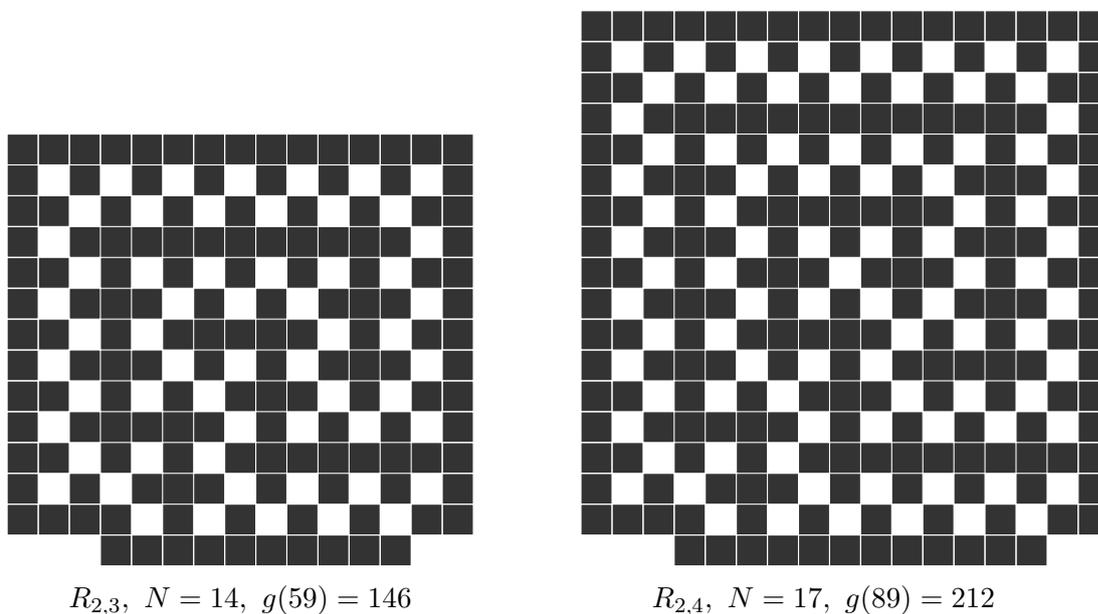


Figure 18: The third and fourth elements of the sequence $R_{2,k}$.

The k -th element of this sequence has k horizontal sections of plus trees and k vertical sections of plus trees, and the extra spaces depend on the parity of k . If k is even, the horizontal trees can be counted with an extra two rows taken up by the bottom rooted plus, and thus with two additional rows at the top and only the boundary row on the bottom, the height is $3k + 5$. And vertically it an extra plus moving left at the center giving two additional columns, with another two additional columns on each side. So the width is $3k + 6$. If k even, the right-rooted spiral can be counted among the k trees, and there is an extra two spaces created by the plus moving down at the center. So the height is $3k + 6$. And then horizontally there is a central hole indicating one extra column, with two additional columns on the sides, making the width $3k + 5$.

So in all cases $N = 3k + 5 \equiv 2 \pmod{3}$, and by construction $R_{2,k}$ is an efficiently structured crystal with total area $N(N + 1) - 5$, and with

$$h(R_{2,k}) = \frac{(N + 1)(N - 2)}{3} - 1 = \frac{(3k + 6)(3k + 3)}{3} - 1 = 3k^2 + 9k + 5, \quad (27)$$

$$|R_{2,k}| = N(N + 1) - h(R_{2,k}) - 5 = 6k^2 + 24k + 20.$$

The first two of this sequence are degenerate with respect to the spiral effect and are depicted below.

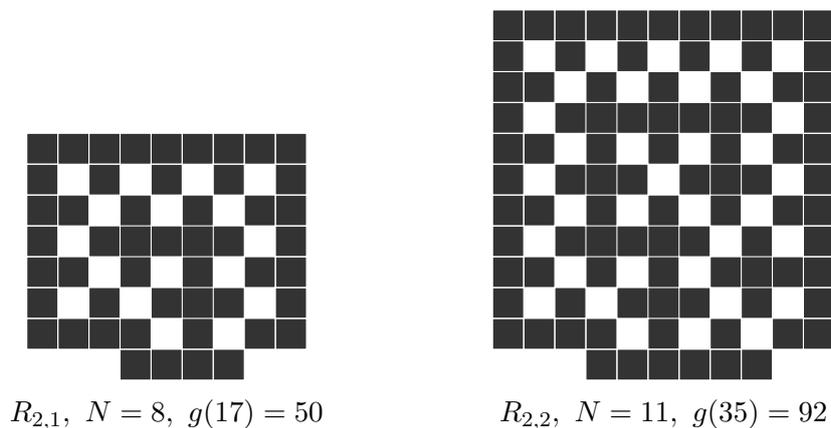


Figure 19: The first two elements of $R_{2,k}$.

7 Dismantling and the Proof of Theorem 5

Given a crystallized polyomino A with h_α holes, we develop algorithms which remove at each step two tiles and one hole. By Lemma 17, this will produce constructions achieving $g(h)$ tiles and h holes for all h for which α is the minimum square or pronic number such that $h \leq h_\alpha$. For $\alpha = N^2$ or $N(N + 1)$, if we can remove at least $\lceil N/3 \rceil$ holes from A in this manner, then the total area of the remaining polyomino is at most $N(N - 1)$ or N^2 , respectively, and we say we have *dismantled* A .

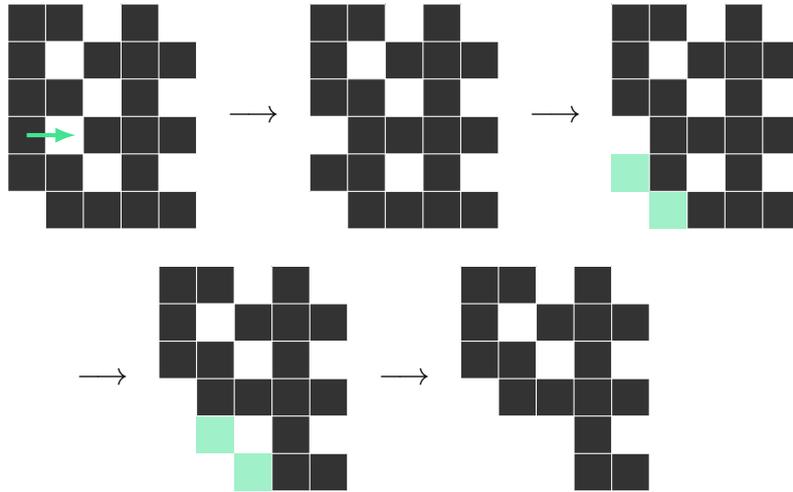


Figure 20: Iterative process for removing two tiles and one hole.

Consider first the case that a polyomino has a plus tree rooted next to an indented corner, and the tree grows straight along the length of the side boundary. This subarrangement is especially prominent in $S_{1,k}$ in Figure 13, and occurs in all constructions in Section 6 except for the double spiral in 6.6. This can be dismantled by the following process depicted in Figure 20: (a) fill the lowest hole along the side boundary by pushing in the adjacent boundary tile; (b) remove the two boundary tiles adjacent to the indented corner; and (c) remove the indented corner and the boundary tile which is corner adjacent to it. Steps (a) and (b) combine to remove two tiles and one hole, and step (c) also removes two tiles and one hole, and additionally sets up a new indented corner so that the process can be repeated as long as the tree grows along that boundary. Both removals preserve the properties that the polyomino is acyclic and each hole has an area of one.

In the pronic mod 2 case there is not always a long rooted tree, but the tree along the top can be rooted so that this process can be implemented. To do so, as in Figure 21: (a) move the top corner tile to the adjacent interior corner; (b) remove the two tiles that were adjacent to the corner tile; (c) push the boundary tile on the bottom row up to root the tree; (d) remove the two tiles which had been adjacent to the tile that was pushed up.

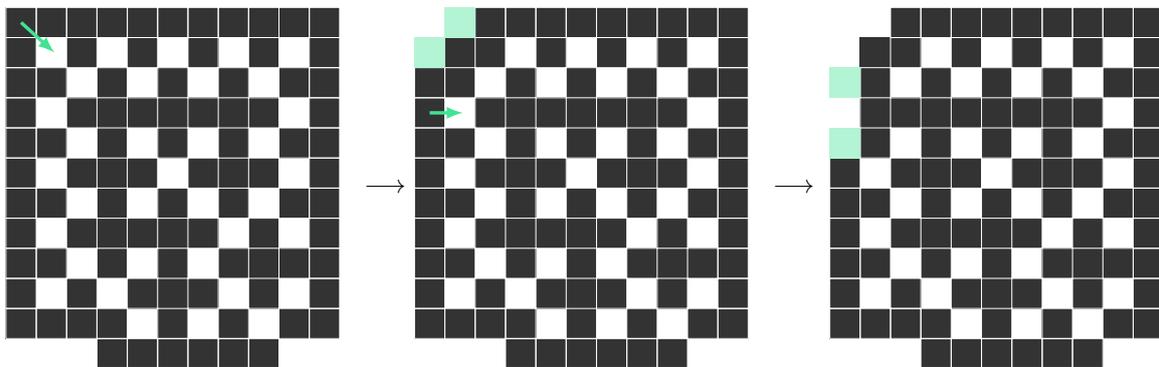


Figure 21: Process for rooting a tree in the double spiral construction.

Dismantling odd squares becomes increasingly difficult upon subsequent expansions. There will always be holes in all four corners by parity, and these holes are always easily removed via methods used in the previous two examples. If the outer corner space is filled, we can apply the first removal from Figure 21, and if it is empty we can apply the second removal from Figure 20.

Removing the corners subtracts 12 total spaces, which suffices to dismantle up to an 11×11 square. For larger squares, after the corners are dealt with we remove the next hole via a complete rearrangement which is a modified version of the double spiral construction, with unique central configurations depending on $N \bmod 3$, as depicted in Figure 22.

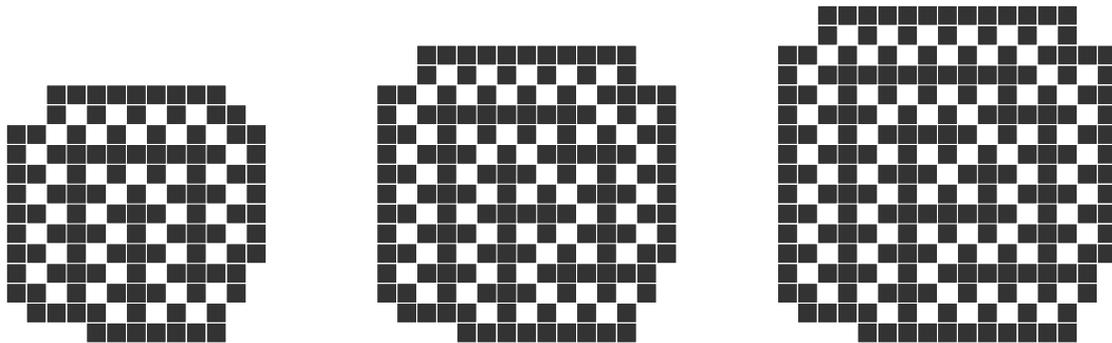


Figure 22: Rearrangements for $N = 13, 15,$ and $17.$

These arrangements generalize straightforwardly for larger N by expanding the boundary and extending each spiral in two directions. For example, the $1 \bmod 3$ construction on the left in Figure 22 extends to $N = 19$ by taking the spiral that stops at the bottom right and extending it left and then up, and the spiral that is pointing up extends right and then down. This adds six to both the width and height, and thus preserves $N \bmod 3$.

The $1 \bmod 3$ construction has 18 empty spaces around the boundary, and the other two have 19 empty spaces. Recall that for $N \equiv 1 \bmod 3$, the construction for h_{N^2} holes has three empty corners, and the constructions for $N \not\equiv 1 \bmod 3$ both have four empty corners. Then another 12 spaces are removed by dismantling the four corner holes, and this rearrangement removes one additional hole, giving the above numbers.

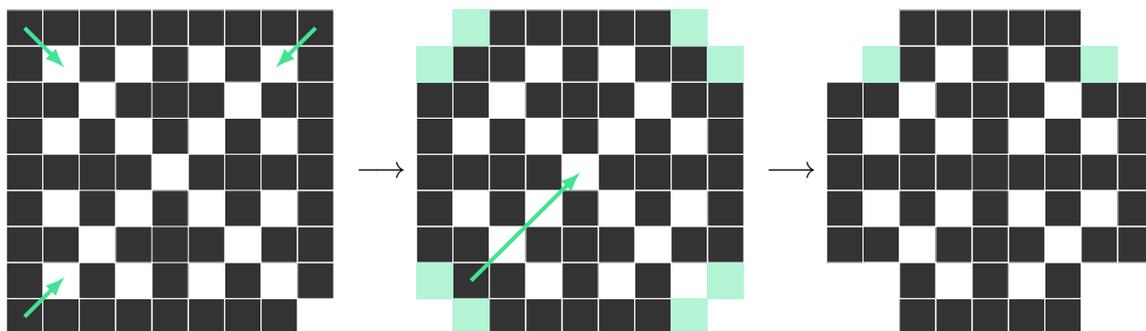


Figure 23: General method to remove five holes from $S_l.$

The only exception is the sequence S_l , which starts with a single empty corner and requires one extra removal. In removing the four corners from S_l , three become indented as in Figure 21. Then we get the additional removal by moving one these three indented corners to the center and removing the other two (see Figure 23). Observe that these rearrangements preserve the acyclic structure of the polyomino and the property that all holes have area one.

Proof of Theorem 5. As noted above, all constructions in Section 6 either have a rooted plus tree growing next to an indented corner along the full side of the polyomino, or have a plus tree which can be rooted as shown in Figure 21, and any crystallized polyomino with h_α holes for α an odd square which is constructed via expansion can be rearranged after some initial removals to have a rooted plus tree growing along a full side of the polyomino. Then, implementing the removal process for holes along a rooted tree clearly suffices to dismantle all of these crystallized polyominoes, as they remove at least $N - 4$ holes, far more than the roughly $N/3$ necessary.

This implies that $g(h) \leq g(h_\alpha) - 2(h_\alpha - h)$ for any h such that α is the minimum square or pronic number with $h \leq h_\alpha$. Let $\alpha' < \alpha$ be the next largest square or pronic number. Then since $m(h_\alpha) = g(h_\alpha)$, Lemma 17 ensures that $g(h_\alpha) - 2(h_\alpha - h) = m(h)$ for all h such that $h + m(h) \geq \alpha'$. When $\alpha' = (2^l + 1)^2$, we have that

$$(h_{\alpha'} + 1) + m(h_{\alpha'} + 1) > \alpha',$$

and otherwise $h_{\alpha'} + 1 = t_{\alpha'}$. Then $m(t_{\alpha'} + 1) = m(t_{\alpha'}) + 3$ by Lemma 17, and by Lemmas 19, 20, and 28 we have that

$$g(t_{\alpha'}) = m(t_{\alpha'}) + 1 = m(t_{\alpha'} + 1) - 2.$$

Therefore $g(h_\alpha) - 2(h_\alpha - h) = g(h)$ for all h such that α is the minimum square or pronic number such that $h \leq h_\alpha$. ■

Observe that these removals will quickly disrupt the property of having minimal outer perimeter. In fact, none of the constructions in Figure 22 are crystallized, since in all three cases $N^2 - 18$ is already less than $N(N - 1)$.

The jumps from h_α to $h_\alpha + 1$ are the only jumps of three for $g(h)$ once $h > 6$. We also demonstrate a constructive procedure to show that three tiles is always sufficient to produce an additional hole when $h > 6$. Whenever there is a plus rooted to a border, the local transformation in Figure 24 maintains all routes of connectivity in the polyomino and does not create any cycles.

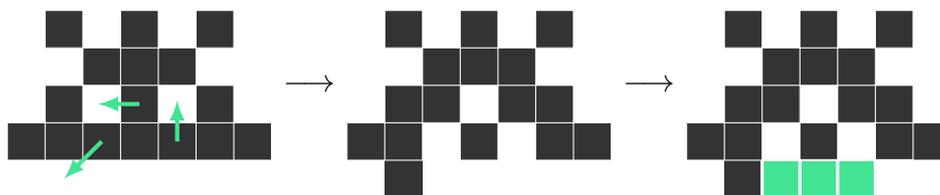


Figure 24: Adding one hole and three tiles to a rooted plus on the boundary.

The $h = 5$ case in a 5×5 square is the largest crystallized polyomino without a rooted plus, and in particular the jump from $g(5) = 19$ to $g(6) = 23$ is the last time there is a jump of more than 3 in $g(h)$.

8 Concluding Remarks and Open Problems

Outside of the K–R sequences, it remains an open problem to enumerate crystallized polyominoes. Expansion can perhaps be leveraged in this pursuit, but in its current formulation it does not capture small changes to the boundary.

There is a subfamily of polyominoes with holes, called *punctured polyominoes* that have been studied before but without asking the extremal problem of maximizing the number of holes within that subfamily. For definitions and results on punctured polyominoes see Chapter 2, 8, and 11 of [2]. An important observation is that punctured polyominoes do not capture the topological structure that we are interested in studying in this paper.

Viewing polyominoes as embedded 2-dimensional cubical complexes and the number of holes as the rank of the first homology group (with integer coefficients), this problem also generalizes immediately to the question of maximizing the $(d - 1)$ -dimensional homology for cubical complexes, such as polycubes in \mathbb{R}^3 .

References

- [1] Tomás Oliveira e Silva. Animal enumerations on the $\{4, 4\}$ Euclidian tiling. <http://sweet.ua.pt/tos/animals/a44.html>, December 2015.
- [2] Anthony J Guttmann. *Polygons, polyominoes and polycubes*, volume 775. Springer, 2009.
- [3] Frank Harary and Heiko Harborth. Extremal animals. *J. Combinatorics Information Syst. Sci.*, 1(1):1–8, 1976.
- [4] Matthew Kahle and Érika Roldán. Polyominoes with maximally many holes. *Geombinatorics*, Vol XXIX(Issue I):5–20, 2019.
- [5] Sascha Kurz. Counting polyominoes with minimum perimeter. *Ars Combin.*, 88:161–174, 2008.
- [6] Greg Malen and Érika Roldán. Polyiamonds attaining extremal topological properties, Parts I and II. *To appear in Geombinatorics*. [arXiv:1906.08447v1](https://arxiv.org/abs/1906.08447v1), 2020.