Punctured polygons and polyominoes on the square lattice

Anthony J Guttmann†, Iwan Jensen†, Ling Heng Wong† and Ian G Enting‡

† Department of Mathematics and Statistics, The University of Melbourne, Parkville, Victoria 3010, Australia
‡ CSIRO, Division of Atmospheric Research, Aspendale, Victoria 3195, Australia

E-mail: `tonyg@ms.unimelb.edu.au`, `i.jensen@ms.unimelb.edu.au`, `henryw@ms.unimelb.edu.au` and `ige@dar.csiro.au`

Abstract. We use the finite lattice method to count the number of punctured staircase and self-avoiding polygons with up to three holes on the square lattice. New or radically extended series have been derived for both the perimeter and area generating functions. We show that the critical point is unchanged by a finite number of punctures, and that the critical exponent increases by a fixed amount for each puncture. The increase is 1.5 per puncture when enumerating by perimeter and 1.0 when enumerating by area. A refined estimate of the connective constant for polygons by area is given. A similar set of results is obtained for finitely punctured polyominoes. The exponent increase is proved to be 1.0 per puncture for polyominoes.

1. Introduction

A self-avoiding polygon (SAP) can be defined as a walk on a lattice which returns to the origin and has no other self-intersections. Alternatively, we can define a SAP as a connected sub-graph (of a lattice) whose vertices are of degree two. The history and significance of this problem is nicely discussed in [15]. Generally, SAPs are considered distinct up to a translation, so if there are p_n SAPs of length n there are $2np_n$ returns to the origin (the factor of two arising from the two possible directions in which the loop can be travelled). Staircase polygons are a special case of SAP. A staircase polygon can be viewed as the intersection of two directed walks starting at the origin, taking steps either to the right or down and terminating once the walks meet. We define a punctured SAP as a SAP with one or more holes, with the perimeter of each hole itself being a SAP. In other words a punctured SAP is a SAP enclosing one or more SAPs. To avoid any possible confusion in our definition of punctured polygons, note that the punctures are disjoint—there are no degree-four vertices in the punctured polygons we are considering. A similar definition can be given for punctured staircase polygons. We have shown an example of each case in figure 1. The two principal questions one can ask are 'how many polygons, distinct up to a translation, are there of perimeter $2n$ (including the perimeter of the holes) with k punctures?' and 'how many polygons, distinct up to a translation, are there of area n with k punctures?'

A polyomino is defined as the union of connected cells, where a cell is a unit square with four edges (and its interior). Two cells are said to be *joined* if they share a common edge, and are said to be *connected* if there exists a sequence of cells joining the two cells such that successive pairs of cells are joined. A k-punctured polyomino is a polyomino with k holes. Unlike the situation for punctured polygons, degree-four vertices are permitted, but two
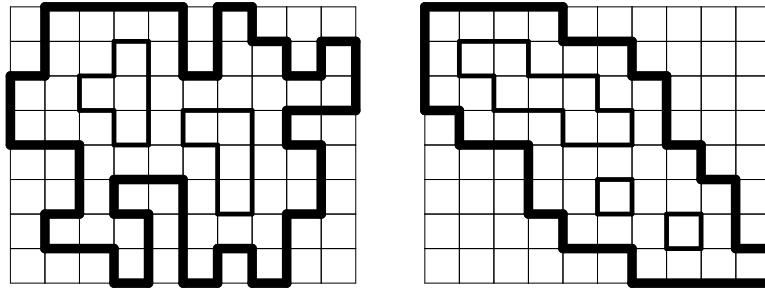
Figure 1. An example of a punctured SAP (left panel) and a punctured staircase polygon (right panel). The thick lines show the perimeter of the enclosing polygon while the medium lines show the perimeter of the holes.
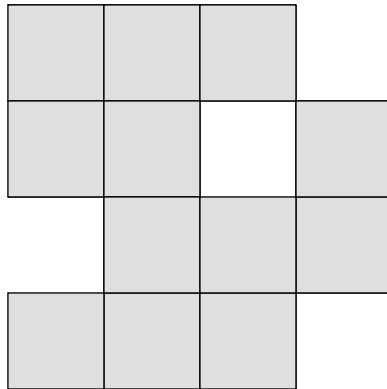
Figure 2. An example of a punctured polyomino that is not a punctured polygon.

holes meeting only at a vertex do count as two holes, rather than one, as they are not joined. Punctured polygons are a proper subset of punctured polyominoes. The difference between them is that punctured polygons do not include those polyominoes where there are two cells touching each other only at a vertex. An example is shown in figure 2.

Reverting to k-punctured polygons, for $k = 0$ we have the simpler—but still unsolved—problem of the number of polygons of perimeter $2n$ or area n. Both these problems have been extensively studied for several decades. The most recent result for polygon perimeters appears to be [17] where polygons of perimeter up to 90 steps are given. In that paper our analysis of the polygon perimeter generating function led us to conclude that

$$P^{(0)}(x) = \sum_n p_{2n}^{(0)} x^n \sim A^{(0)}(x) + B^{(0)}(x)(1 - \mu^2 x)^{2-\alpha}$$

where $p_{2n}^{(0)}$ is the number of SAPs of perimeter $2n$, distinct up to translations. The analysis in [17] yielded a very accurate estimate for the connective constant $\mu = 2.638\,158\,530\,34(10)$ and confirmed the theoretical prediction $\alpha = \frac{1}{2}$ [19]. Furthermore, we obtained estimates for the critical amplitudes $A^{(0)}(x_c) \approx 0.036$ and $B^{(0)}(x_c) \approx 0.234\,913$, where $x_c = 1/\mu^2$. We also concluded that there was no evidence for a non-analytic correction-to-scaling exponent, so that we expect the asymptotic form of the coefficients to behave as

$$p_{2n}^{(0)} \sim \mu^{2n} n^{-\frac{5}{2}} [a_1 + a_2/n + a_3/n^2 + a_4/n^3 + \cdots].$$

The connective constant μ is, of course, the same as that for self-avoiding walks on the same lattice [15].

For polygon areas the most recent published work appears to be [9] in which the first 20 terms of the area generating function were given and analysed. In that work it was found that

$$A^{(0)}(y) = \sum_n a_n^{(0)} y^n \sim C^{(0)}(y) + D^{(0)}(y) \log(1 - \kappa y)$$

where $a_n^{(0)}$ is the number of SAPs of area n, $\kappa \approx 3.970\,87$, and the amplitudes $C^{(0)}$ and $D^{(0)}$ were not estimated. Recently this series has been extended to 26 terms [20], but in this work we have devised a new and exponentially faster algorithm, as a result of which we have extended the series to 42 terms, and we present an analysis of this longer series. The connective constant κ is found to be slightly smaller than that for the related problem of polyominoes [10].

In the following, we refer to the boundary of a polygon and its interior as a *disc* and so we will be discussing punctured discs. An unpunctured disc is a SAP.

For *punctured* discs, the basic problem is, analogously, the calculation of the generating functions

$$P^{(k)}(x) = \sum_n p_{2n}^{(k)} x^n \sim B^{(k)}(x) + C^{(k)}(x)(1 - (\mu^{(k)})^2 x)^{2-\alpha_k} \qquad (1)$$

and

$$A^{(k)}(y) = \sum_n a_n^{(k)} y^n \sim D^{(k)}(y) + E^{(k)}(y)(1 - \kappa^{(k)} y)^{-\beta_k} \qquad (2)$$

where the superscript k refers to the number of holes, or punctures. From the generating function, one wishes to deduce the asymptotic behaviour, believed to be as shown on the rhs of the above equations. The major problem to be investigated is how the behaviour of $P^{(k)}(x)$ and $A^{(k)}(y)$ changes as k is increased. Previous work [25, 26] has been confined to the study of punctured SAPs by area. There it was proved that $\kappa^{(k)} = \kappa^{(0)} = \kappa$, and that if the exponent exists, $\beta_k = \beta_0 + k$. These results apply more generally to punctured surfaces, but in this work we are confining ourselves to two dimensions. As far as we are aware, there has been no previous work on the problem of the perimeter generating function of punctured discs.

The problem is interesting for several reasons. The effect of a change in geometry is a much studied topic in lattice statistics, and our study of the change in perimeter exponent with punctures seems to be entirely new. It has only been possible by the algorithms we have designed and implemented, which are exponentially faster than pre-existing algorithms. A number of related models have been studied previously, such as c-animals [18, 24] and the behaviour of prime knots in polygons [22]. c-animals are lattice animals with exactly c cycles. In [24] it was proved that if the number of such animals $a_n(c) \sim \lambda_c^n n^{\theta_c}$ as $n \to \infty$, c fixed, then $\theta_c = \theta_0 + c$ provided that θ_0 exists. It had been previously proved [28] that $\lambda_c = \lambda_0$. The change in connective constant of c-animals as the number of cycles per vertex changes from zero to non-zero is discussed—among other results—in [18]. Similarly, in a numerical study of knotted polymers [22], it was *conjectured* that the exponent α depends on the number of prime knots n_p that arise in the knot decomposition of a given SAP via the relation $\alpha(n_p) = \alpha(0) + n_p$.

Furthermore, there is considerable pedagogical connection between some of these previous studies and our work here, and also between our work and the study of branched polymers. To sketch this connection, we first remark that the number of polyominoes, also called lattice animals, is just the number of (strongly embedded) site animals. This connection is readily seen by placing a site at the centre of every cell of the polyomino, and joining those sites corresponding to joined cells by a bond. In this way, every lattice animal is mapped to a distinct site animal and vice versa. All other models we are considering map similarly to a different subset of strongly embedded site animals. For example, SAPs map to site animals

whose only cycles are 4-cycles (which may be isolated or joined). Punctured polygons map to site animals with larger cycles, as well as 4-cycles, whereas punctured polyominoes correspond to site animals with more complex restrictions. Thus, this study also complements the earlier studies of *c*-animals. In those studies, the variation of exponent with the number of cycles is considered, whereas in this study we are varying the *types* of allowed cycles.

In order to study these and related systems, when an exact solution cannot be found one has to resort to numerical methods. For many problems the method of series expansions is by far the most powerful method of approximation. For other problems Monte Carlo methods are superior. For the analysis of $P^{(k)}(x)$ and $A^{(k)}(y)$, series analysis is undoubtedly the most appropriate choice. This method consists of calculating the first few coefficients in the expansion of the generating function. Given such a series, using the numerical technique known as differential approximants [13], highly accurate estimates can frequently be obtained for the critical point and exponents, as well as the location and critical exponents of possible non-physical singularities. Other numerical methods are discussed in [17], and those used in this study are described more fully below.

In the next section we will describe the finite lattice method for enumerating punctured polygons. In section 3 we prove the invariance of $\mu^{(k)}$ as k changes, and give an heuristic argument for the exponent shift with k. The results of the analysis of the series are presented in section 4. Results analogous to those known for punctured polygons by area are proved for punctured polyominoes in the appendix.

2. Enumeration of punctured polygons

2.1. Enumeration of punctured self-avoiding polygons

The method used to enumerate punctured SAPs on the square lattice is a generalization of the method devised by Enting [7] for the enumeration of ordinary SAPs. In the following we first describe the original method in some detail and show how simple it is to generalize the method to the enumeration of punctured polygons. The first terms in the series for the generating function can be calculated using transfer matrix techniques to count the number of polygons in rectangles $W + 1$ edges wide and $L + 1$ edges long. The transfer matrix technique involves drawing a line through the rectangle intersecting a set of $W + 2$ edges. For each configuration of occupied or empty edges along the intersection we maintain a (perimeter) generating function for loops to the left of the line cutting the intersection in that particular pattern. Polygons in a given rectangle are enumerated by moving the intersection so as to add one vertex at a time, as shown in figure 3. Since the loops are non-intersecting, each configuration can be represented by an ordered set of edge states $\{n_i\}$, where

$$n_i = \begin{cases} 0 & \text{empty edge} \\ 1 & \text{lower part of loop closed to the left} \\ 2 & \text{upper part of loop closed to the left.} \end{cases}$$

Configurations are read from the bottom to the top. So the configuration along the intersection of the polygon in figure 3 is {0112122}. In passing it is worth noting that there are some major restrictions on the possible configurations. Firstly, since all loop ends are connected to the left of the intersection, every lower loop end must have a corresponding upper end, and it is therefore clear that the total number of '1' is equal to the total number of '2'. Secondly, as we look through the configuration starting from the bottom the number of '1' is never smaller than the number of '2'.
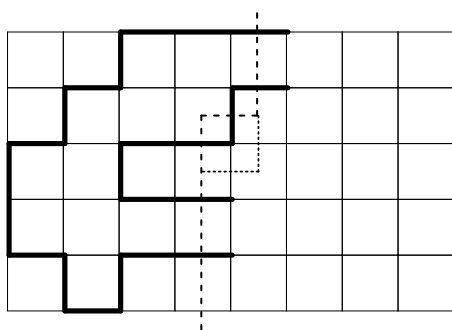
Figure 3. A snapshot of the intersection (dashed line) during the transfer matrix calculation on the square lattice. Polygons are enumerated by successive moves of the kink in the intersection, as exemplified by the position given by the dotted line, so that one vertex at a time is added to the rectangle. To the left of the intersection we have drawn an example of a partially completed polygon.

Table 1. The various 'input' states and the 'output' states (with corresponding weights) which arise as the boundary line is moved in order to include one more vertex of the lattice.

Input	Outputs		
'00'	'00'		x^2'12'
'01'/'10'	x'01'		x'10'
'02'/'20'	x'02'		x'20'
'11'/'22'	'$\overline{00}$'		
'21'	'00'		
'12'	'accumulate'		

In table 1 we have listed the possible local 'input' states and the 'output' states which arise as the kink in the intersection is propagated by one step. Some of these update rules are illustrated further in figure 4. The first panel represents the input states '10' and '01' and the possible output states are also '01' and '10'. The second panel represents the input state '11' as part of the configuration {01122}. In this case we connect the two loop ends, but in doing so we see that the upper part of the second loop before the move becomes the lower part of the one remaining loop after the move: that is, the configuration {01122} becomes {00012}. This relabelling of the other loop end when connecting two '1' (or two '2') is denoted by over-lining in table 1. In general, there could be more loops nested in between the two '1' and the corresponding '2' at the other end of the loop. Say, for instance, we had the configuration {11121222} and connected the first two '1', then the new configuration of unconnected loop ends would be {00121212} (drawing a little figure makes this quite clear). The general rule for the relabelling is as follows: when connecting two '1' ('2') we work upward (downward) in the configuration, counting the number of '1' and '2' we pass until the number of '2' ('1') exceeds the number of '1' ('2'). This '2' ('1') is the other end of the inner loop and it should now be changed to a '1' ('2'), thus becoming the lower (upper) end of the outer loop (again drawing a few pictures should make this clearer). The weights corresponding to these configuration transformations are simply calculated by counting the number of steps which have been added to the polygon. Note that the input state '12' is special because connecting the two ends results in a closed loop, so this is only allowed if there are no other loops cut by the intersection and the result is a valid polygon, which is then accumulated in the total count for that particular length. Failure to observe this restriction would result in graphs with disconnected components, either one polygon over another or a polygon within another (this latter case is, of course, of interest when we wish to enumerate punctured polygons). This is illustrated in figure 5 where we show the possible ways a pair of loops can be placed relative to one another (left panels), how the loops can be connected to produce a valid configuration (middle panels) and the ways of connecting loops that lead to invalid graphs containing disconnected components (right panels).
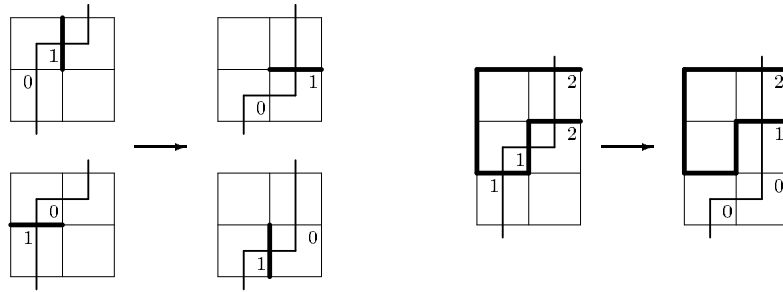
Figure 4. Some of the local configurations which occur as the kink in the intersection is moved one step.
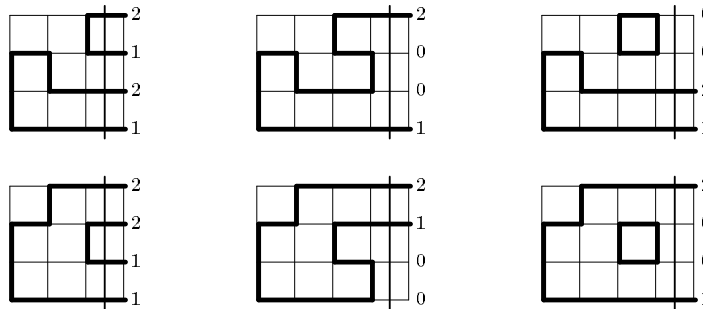
Figure 5. Illustration of how a pair of loops can be placed (left), connected to produce a valid SAP (middle), and connections leading to forbidden graphs (right).

In this figure the invalid SAP in the bottom right panel could result in a valid punctured SAP. Note also that from the input state '00' we can produce the output state '12' only if there are other loops crossing the inter (otherwise we would produce disconnected polygons sitting side by side). We refer the interested reader to [7, 8] for further details regarding the encoding and relabelling of configurations.

Due to the obvious symmetry of the lattice one need only consider rectangles with $L \geqslant W$. In the original approach [7], valid polygons were required to span the enclosing rectangle in the lengthwise direction. So it is clear that polygons with projection on the y-axis $< W$, that is polygons which are narrower than the width of the rectangle, are counted many times. It is, however, easy to obtain the polygons of width exactly W and length exactly L from this enumeration [7]. Any polygon spanning such a rectangle has a perimeter of length at least $2(W + L)$. By adding the contributions from all rectangles of width $W \leqslant W_{\max}$ (where the choice of $W_{\max}$ depends on available computational resources) and length $W \leqslant L \leqslant 2W_{\max} - W + 1$, with contributions from rectangles with $L > W$ counted twice, the number of polygons per vertex of an infinite lattice is obtained correctly up to perimeter $4W_{\max} + 2$.

With the original algorithm the number of configurations required as $W_{\max}$ increased grew asymptotically as $3^{W_{\max}}$ [12]. In a recent improvement of the algorithm [16, 17] valid polygons were required to span the rectangle in *both* directions. In other words we directly enumerate polygons of width exactly W and length L. For each configuration of partially completed polygons we keep track of the current minimum number of steps N_{cur} that have been inserted to the left of the intersection and we calculate the minimum number of additional steps N_{add}

required to produce a valid polygon that spans a rectangle of size at least $W \times W$. If the sum $N_{\text{cur}} + N_{\text{add}} > 4W_{\text{max}} + 2$ the partial generating function for that configuration was discarded because it would make no contribution to the polygon count up to the perimeter lengths we were trying to obtain. Numerical evidence indicated that the computational complexity was reduced significantly. While the number of configurations still grew exponentially as $\lambda^{W_{\text{max}}}$, the value of λ was reduced from $\lambda = 3$ to $\lambda \simeq 2$ with the improved algorithm. Furthermore, for any W we know that contributions will start at $4W$ since the smallest polygons have to span a $W \times W$ rectangle, so for each configuration we need only retain $4(W_{\text{max}} - W) + 2$ terms of the generating functions while in the original algorithm contributions started at $2W + 2$ because the polygons were required to span only the lengthwise direction.

The generalization to enumeration of punctured polygons is obtained by noting that a closed loop is formed whenever we connect a 1-edge to a 2-edge immediately above. If these two edges were the only loop ends in the intersection we would have formed a valid polygon. In other cases we need to ensure that the resulting polygon is a valid punctured polygon. That is, we must ensure that the separate SAP just formed will be completely enclosed within a larger polygon. So we wish to avoid forming separable polygons, as shown in the upper right panel of figure 5, and 'holes within holes'. As it turns out, the rule for a valid '12' closure is simply that we can connect the two loop ends provided there is an *odd* number of loop ends below the loop being closed. To see this consider that as we go through a configuration we note that each time we pass a loop end we go from the outside of the polygon to the inside and vice versa. So, in a configuration all lattice cells between the first and second loop ends will lie inside the finished polygon, lattice cells between the second and third loop end will lie outside the polygon, and so on. Thus we see that by closing a loop which has an odd number of loop ends below it we are closing off a part of the lattice which will lie inside the finished polygon. In particular, we see that we avoid the situation shown in the upper right panel of figure 5 with graphs containing disconnected pieces one over another. Likewise we avoid creating graphs with disconnected pieces sitting side by side.

We also avoid forming holes within holes because closing a loop around the hole would be prohibited since there would be an even number of loop ends below the '1'-edge (except of course when forming a completed punctured polygon). Let us look at the possible edge configurations around a puncture in some detail. First look at the configuration $\{\ldots 1122 \ldots\}$ (where the $\ldots$ are any edge configurations with an even number of edges making the total configuration of edges a valid intersection). The outer '12'-edges cannot be connected (the number of edges below the '1' is even) so we have to connect two other edges on either side of the hole, diminishing the number of edges by two and possibly changing the edge labels on either side of the hole (in which case we end up with one of the subsequent cases). In the top panels of figure 6 we show how the simplest interesting case $\{11112222\}$ leads to only valid punctured polygons. Secondly, in the case $\{\ldots 1121 \ldots\}$, we can connect the two '1'-edges forming a partial loop enclosing the hole. In doing so the matching '2'-edge of the second '1'-edge is changed to a '1'-edge because it now is the new lower edge of the larger loop formed by connecting the two original '1'-edges. The important thing to notice is that the number of edges below this new '1'-edge is even so we cannot connect it to a '2'-edge immediately above and we therefore do not form a hole within a hole. In the middle panel of figure 6 we show the simplest interesting case $\{1111212222\}$ and demonstrate that it leads to valid punctured polygons. Thirdly, the configuration $\{\ldots 2121 \ldots\}$ is not very interesting since connecting the '21'-edges in front of the hole does not result in a loop partially enclosing the hole. Finally, the configuration $\{\ldots 2122 \ldots\}$ is obviously merely the mirror image of the second case. Note that this covers all cases since more complicated configurations would merely correspond to more convoluted loop structures.
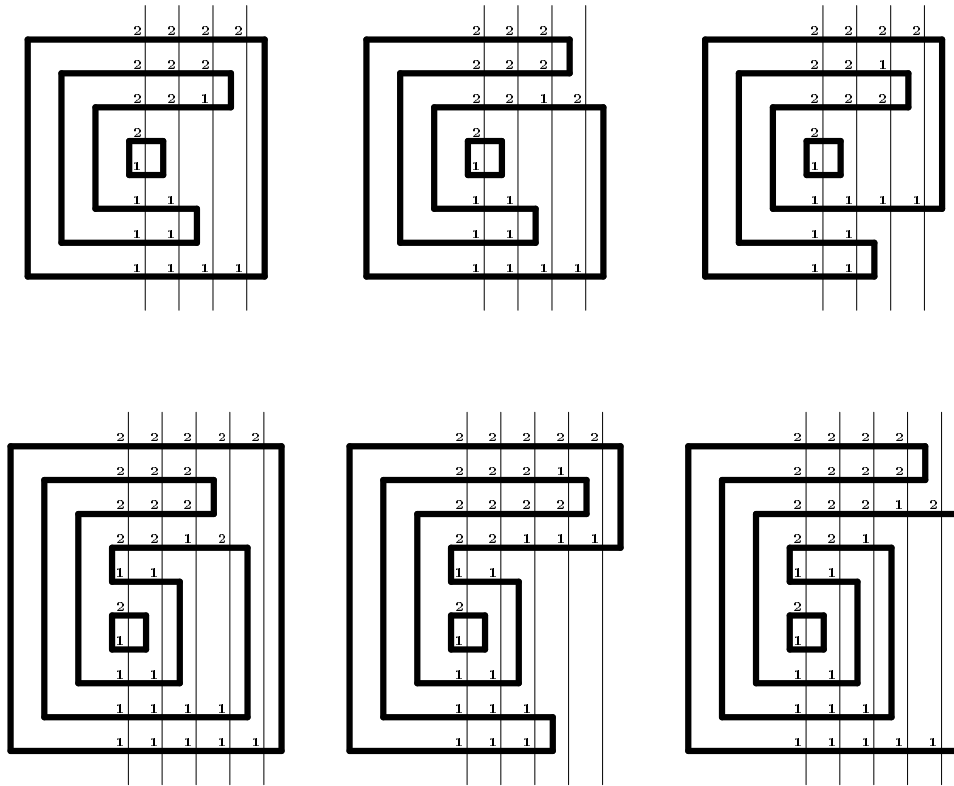
Figure 6. Illustration of how connecting loop ends lead to various cases of punctured SAPs. In the upper (lower) panels we show how the configuration {11112222} ({1111212222}) can be completed given that the first pair of '12'-edges are connected to form a puncture (in the lower panel we further connect the two 1-edges on either side of the puncture). In each panel we also show the intersection line with the numbers giving the labelling of the loop ends.

In this paper we use the generalization of the algorithm of [17] to count the number of punctured SAPs with up to three holes. Obviously, the smallest hole we can make in a SAP has perimeter 4 so the number of punctured SAPs is obtained correctly up to perimeter $4W_{max} + 2 + 4k$.

The algorithm used for the enumeration of punctured SAPs by area is a simple variation of the algorithm described above. The encoding of configurations along the intersection and the transformation of these configurations as the intersection is moved remain the same. The only change is that the weights are different. In order to count the enclosed area we proceed as follows: a unit of area may be added as the kink is moved to a new lattice cell (in figure 3 this is the cell in which the dotted lines meet). Whether or not a unit of area is added is determined by whether or not this lattice cell is inside or outside the polygon. But we already know from the arguments given above that a lattice cell is inside the polygon if the number of loop ends below the cell is odd (note that any loop end along the vertical edge cutting the horizontal part of the kink is included in the count). So, in the case shown in figure 3 the lattice cell to which the kink is moved lies outside the polygon (there are two loop ends below the kink) and no unit of area is added. Note that a unit of area may be added for any given output configuration. In this case the area generating function is obtained correctly to $2W_{max} + 3k$. The factor $3k$ arises

since it takes three lattice cells to completely surround the simplest puncture, which is just a single cell.

As far as we are aware, series for punctured polygons have not been derived previously. In such circumstances it is even more important than usual to undertake careful tests to ensure that the series are correct. To this end, a second algorithm was implemented, by one of the authors, to independently evaluate the series coefficients. The complete agreement we obtained between the two data sets reassures us as to the correctness of our results.

2.2. Enumeration of punctured staircase polygons

The enumeration of punctured staircase polygons is much simpler. In fact, as we shall demonstrate, it is a problem for which the computational complexity grows only as a polynomial in the number of terms. As stated in the introduction, we can think of staircase polygons as consisting of two non-intersecting directed walks on the square lattice. Punctured staircase polygons naturally will have more than two walks, in fact up to $2k + 2$ walkers can be present in any given column. Due to the restrictions on staircase polygons it follows that a punctured staircase polygon is formed by requiring that the walks be directed and that any two walkers starting at one point join each other later without intersecting other walks. This can be encoded in a transfer matrix calculation as follows: we count the number of punctured polygons in rectangles as before and draw a line through the lattice intersecting the $W + 2$ edges. In this case we need only specify whether or not each edge is part of the polygon or not, so each configuration can be represented by an ordered set of edge states $\{n_i\}$, where

$$n_i = \begin{cases} 0 & \text{empty edge} \\ 1 & \text{part of loop closed to the left.} \end{cases}$$

This uniquely specifies the configuration because the first occupied edge is connected to the last occupied edge and any edges in between are paired: e.g., the second and third (fourth and fifth and so on) edges form a loop to the left and has to be connected to each other later on. The rules for updating the configurations are as follows: from the local input state '00' we can always get the output '00', and the output '11' provided there is an odd number of edges below the kink and also that the edge directly below the kink is empty. The rules for the output '11' ensure that the new pair of walkers lie within the enclosing staircase polygon but not inside an internal staircase polygon (thus preventing holes within holes), and that the lower walk does not intersect other walks. From the local input state '01' and '10' we can always produce the output '01', and the output '10' provided either that the edge directly below the kink is empty or there is an even number of edges below the kink. These rules ensure that the walkers do not intersect other walks except when we close a valid loop. Finally, from the local input state '11' we always get the output state '00'. The weights associated with these updates are obtained in the same way as for the SAP enumeration, whether the enumeration is done by perimeter or area.

As in the previous case, we calculate the number of punctured staircase polygons spanning the rectangles. Adding the contributions from all rectangles of width $W \leqslant W_{\max}$ and length $W \leqslant L \leqslant 2W_{\max} - W + 1$ the number of punctured staircase polygons is obtained correctly up to perimeter $4W_{\max} + 2 + 4k$. Note that for fixed k the maximum number of configurations N_C grows as a polynomial in $W_{\max}$:

$$N_C = \sum_{j=0}^{k} \binom{W_{\max} + 1}{2j + 2}.$$

So in this case the algorithm is of polynomial complexity.

3. Expected behaviour

As mentioned in the introduction, the problem of enumerating SAPs by area has been extensively studied [9, 20]. It has been shown that $\kappa^{(k)} = \kappa^{(0)}$, and that if the exponents exist, $\beta_k = \beta_0 + k$ [26, 27]. That is to say, for k finite, the connective constant for k-punctured discs, by area, is the same as that for unpunctured discs, while the critical exponent increases by one for each puncture. We repeat that the punctures are disjoint.

As far as we are aware, there has been no previous work on the problem of the perimeter generating function of punctured SAPs. We first give a proof that a 1-punctured SAP on the square lattice has the same connective constant as an unpunctured SAP, and indicate how that proof can be generalized to k-punctured SAP.

Before stating and proving the relevant theorem, we require certain preliminary results. For an (unpunctured) SAP of perimeter $2n$ one knows, [15] equation (7.101), that

$$\exp[-b\sqrt{n}]\mu^{2n}/4n \leqslant p_{2n}^{(0)} \leqslant \mu^{2n} \tag{3}$$

where b is a constant. Further, it is obvious that a polygon of perimeter $2n$ has maximum area $n^2/4$, which occurs when the shape is an $n/2 \times n/2$ square. Thus the area of a polygon of perimeter $2n$, denoted $\mathcal{A}_{2n}$, satisfies $\mathcal{A}_{2n} \leqslant n^2/4$.

Now consider 1-punctured polygons of total perimeter $2n$, with inner perimeter $2m$ and outer perimeter $2(n-m)$. It is possible for the inner polygon to be Hamiltonian, in which case its perimeter is greater than that of the surrounding polygon. Indeed, a square polygon of side $2n + 1$, and hence of perimeter $8n + 4$ can contain an internal polygon of perimeter as large as $4n^2$. Hence the semi-perimeter of the inner polygon, m, can range from a minimum value of 2 to a maximum value of $n - 2\sqrt{2n} + 2 < n - \sqrt{n}$ for $n > 1$.

With these preliminaries, we can now state and prove the following theorem.

Theorem. $\lim_{n\to\infty} \frac{1}{2n} \log p_{2n}^{(1)} = \log \mu$, *where μ is the same constant as appears in the corresponding limit for unpunctured SAPs.*

Proof. 1-punctured polygons of total perimeter $2n$ are constructed by placing polygons P of perimeter $2m$ inside polygons Q of perimeter $2n - 2m$. Let $w(P, Q)$ denote the number of ways of placing polygon P inside polygon Q, (which of course depends on both P and Q.) Then

$$p_{2n}^{(1)} = \sum_{m=2}^{n-\sqrt{n}} \sum_{P} \sum_{Q} w(P, Q).$$

We bound this summand by the product of three factors. The first two factors are the number of polygons of perimeter $2m$ and $2n - 2m$, respectively. The third is the number of ways the smaller polygon can be placed inside the surrounding polygon and is clearly less than or equal to the area of the surrounding polygon. Explicitly,

$$p_{2n}^{(1)} \leqslant \sum_{m=2}^{n-\sqrt{n}} (n-m)^2 \mu^{2m} \mu^{2(n-m)}/4 \leqslant \mu^{2n}/4 \sum_{m=2}^{n} (n-m)^2 \leqslant n^3 \mu^{2n}.$$

To obtain a lower bound, consider a 3×3 square polygon with a unit square hole at its centre. This unique realization of $p_{16}^{(1)}$ can be uniquely concatenated with each unpunctured polygon by joining them at a solitary specified edge, and then deleting that edge. For each unpunctured polygon we take the set of left-most vertical edges and choose the bottom edge from this set, and we choose the right-most, top-most vertical edge of $p_{16}^{(1)}$ (a similar operation is shown for polyominoes in figure A.1). The concatenation operation then gives, for each

of the p_{2n-14} unpunctured polygons, a unique member of the set of 1-punctured polygons in which the puncture is a single cell. Thus we have $p_{2n-14}^{(0)} \leqslant p_{2n}^{(1)}$. From the above equations we thus obtain

$$\exp\left[-b\sqrt{n-7}\right]\mu^{2n-14}/4(n-7) \leqslant p_{2n-14}^{(0)} \leqslant p_{2n}^{(1)} \leqslant n^3\mu^{2n}$$

for $n > 8$. The theorem then follows immediately on taking logarithms, dividing through by $2n$ and taking the limit as $n \to \infty$. □

This proof can clearly be extended to 2-punctured polygons, then to 3-punctured polygons etc, by concatenating unpunctured polygons with minimal 2-punctured, 3-punctured etc polygons. In the appendix we prove that k-punctured polyominoes have the same growth constant as unpunctured SAPs by area.

We have been unable to prove a result analogous to the result for the exponent of punctured SAPs by area, but give an argument that depends on certain assumptions that are generally accepted, though not proved. In the case of staircase polygons, however, our assumptions have been proved, and so our result will be rigorously true.

The key results we need are those obtained in [9] to the effect that the mean area of SAPs of perimeter $2n$ is proportional to $n^{1.5}$. This is true both for SAPs and for staircase polygons, and in the latter case it has been proved. More precisely, we need the following result. There exists constants D_1 and D_2 such that the mean area $\bar{A}_{2n}$ of polygons of perimeter $2n$ satisfies

$$D_1 n^{\frac{3}{2}} \leqslant \bar{A}_{2n} = \frac{\sum_Q A_Q}{p_{2n}} \leqslant D_2 n^{\frac{3}{2}}$$

where the sum is taken over all p_{2n} polygons Q of perimeter $2n$.

Further, there exists constants C_1 and C_2 such that the number of polygons of perimeter $2n$ satisfies

$$C_1\mu^{2n}n^{-\frac{5}{2}} \leqslant p_{2n}^{(0)} \leqslant C_2\mu^{2n}n^{-\frac{5}{2}}.$$

For staircase polygons, $\mu = 2$ and the exponent is $\frac{3}{2}$ instead of $\frac{5}{2}$ in the above equation.

As above, let Q denote a polygon of perimeter $2n - 2m$, with area A_Q. The number of ways of placing a given polygon P of perimeter $2m$ inside Q is clearly less than A_Q.

Thus, the number of ways of placing P inside any polygon of perimeter $2n - 2m$ is less than

$$\sum_Q A_Q = p_{2n-2m}^{(0)}\bar{A}_{2n-2m} \tag{4}$$

where the sum is over all polygons Q of perimeter $2n - 2m$.

Hence, the number of ways of placing all of the $p_{2m}^{(0)}$ polygons of perimeter $2m$ inside any polygon of perimeter $2n - 2m$ is less than $p_{2m}^{(0)} p_{2n-2m}^{(0)} \bar{A}_{2n-2m}$, and so

$$p_{2n}^{(1)} < \sum_{m=2}^{n-\sqrt{n}} p_{2m}^{(0)} p_{2n-2m}^{(0)} \bar{A}_{2n-2m} < D_2 C_2^2 \mu^{2n} 2^{-5} S_{2n} \tag{5}$$

where

$$S_{2n} = \sum_{m=2}^{n-\sqrt{n}} m^{-\frac{5}{2}}/(n-m).$$

The last sum may be evaluated in a variety of ways. Using Maple we find that

$$S_{2n} = 0.341\ldots/n + O(n^{-\frac{3}{2}}).$$

Thus we obtain the bound

$$p_{2n}^{(1)} < \text{ const. } \mu^{2n}/n. \tag{6}$$

For staircase polygons the analogous calculation is slightly simpler, and we find

$$S_{2n} \sim 1 + \zeta(3/2) + \text{O}(n^{-\frac{1}{2}})$$

where $\zeta(z)$ is the Riemann zeta function.

To obtain a lower bound, we restrict the inner polygon to be of perimeter 4, that is, a unit square. The number of polygons punctured by a unit square clearly provides a lower bound to the number of 1-punctured polygons. A unit square can be placed anywhere in a polygon Q of perimeter $2n - 4$ except on a boundary site. There are $2n - 4$ boundary sites. The mean area of Q is greater than $D_1(n - 2)^{\frac{3}{2}}$, so we obtain the bound

$$p_{2n}^{(1)} > p_{2n-4}(D_1(n - 2)^{\frac{3}{2}} - 2n + 4) > C_1 \frac{\mu^{2n-4}}{(2n - 4)^{\frac{5}{2}}}(D_1(n - 2)^{\frac{3}{2}} - 2n) > \text{ const. } \mu^{2n}/n. \tag{7}$$

Combining the two bounds gives the result

$$E_1\mu^{2n}/n \leqslant p_{2n}^{(1)} \leqslant E_2\mu^{2n}/n$$

where $E_1 < E_2$ are constants. Accepting the usual asymptotic form that is expected for such models, we conclude that

$$p_{2n}^{(1)} \sim \text{ const. } \mu^{2n}/n. \tag{8}$$

(For staircase polygons the analogous result is $p_{2n}^{(1)} \sim \text{ const. } 4^n$.) Since the number of unpunctured polygons grows like $\mu^n n^{-5/2}$, we see that the exponent is predicted to increase by $\frac{3}{2}$ as the result of a single puncture, while in the case of the area generating function, the exponent is found to increase only by 1. We show in the next section that this prediction is borne out by our numerical calculations.

For punctured staircase polygons, a similar conclusion holds. That is, the connective constant is unaltered at $\mu_{\text{stair}} = 2$, but the exponent increases by 1.5 over its unpunctured counterpart when enumerating 1-punctured staircase polygons by perimeter. In fact, we have been able to calculate the generating function for staircase polygons with a single puncture of perimeter 4, and also with a single puncture of perimeter 6. The generating functions for these special cases are given below, and are precisely in accordance with the more general results given above.

We also note that Cardy [3] recently considered the problem of the number of punctured SAP with k concentric, mutually self-avoiding SAPs surrounding a fixed point of the dual lattice. Thus for $k = 1$ this corresponds to 1-punctured SAPs surrounding a fixed point. In that case Cardy finds [3] for the number of such configurations $b_{2n}^{(1)}$ that $b_{2n}^{(1)} = \mu^{2n}\frac{\ln n}{64\pi^2 n}$. The above calculation may be repeated for Cardy's problem. All that is required is to add a factor $\bar{A}_m$ to the summand in equation (5), as the surrounded point can be anywhere inside the inner polygon. Thus we must multiply by the mean area of that polygon. The sum defining S_{2n} then becomes $S_{2n} = \sum_{m=2}^{n-\sqrt{n}} \frac{1}{m(n-m)} \sim \log n/n$, in agreement with Cardy's result. (Our lower bound for this problem is too weak, but with more effort could be improved.)

4. Analysis of the series

All the series we have investigated are characterized by coefficients that grow exponentially, with sub-dominant term given by a critical exponent. The generic generating function

behaviour is $G(z) = \sum_n g_n z^n \sim A(z)(1 - \sigma z)^{-\xi}$, and hence the coefficients of the generating function $g_n = [z^n]G(z) \sim A(1/\sigma)/\Gamma(\xi)\sigma^n n^{\xi-1}$. Generally speaking the existence of the growth constant σ has been proved, but except for exactly solvable models, such as staircase polygons, the existence of the critical exponent ξ has only been conjectured, although its existence has never been doubted. The radius of convergence of the generating function is usually given by the critical point, which is at $z = 1/\sigma$.

We principally used two methods to analyse all the series studied in this paper. Firstly, to obtain the singularity structure of the generating function we used the numerical method of differential approximants [13]. In particular, we used this method to estimate the growth constant σ and the critical exponent ξ. We were invariably able to conjecture an exact value for ξ, which was always an integer or half-integer for all the problems we investigated. Imposing this conjectured exponent permitted a refinement of the estimate of the growth constant— providing so-called biased estimates.

Once the exact value of the exponent was conjectured, and the growth constant accurately estimated, we turned our attention to the 'fine structure' of the asymptotic form of the coefficients, by fitting the coefficients to the assumed form $g_n = [z^n]G(z) \approx \sigma^n n^{\xi-1} \sum_{i \geqslant 0} c_i/n^{f(i)}$. If there is no non-analytic correction term, then $f(i) = i$, while a square-root correction term means $f(i) = i/2$. For all the series studied, only these two situations were encountered.

In all cases, our procedure is to *assume* a particular form for $f(i)$, and see how it fits the data. With the very long series we now have at our disposal, it is usually easy to see if the wrong assumption has been made—the sequence of amplitude estimates c_i either diverges to infinity or converges to zero. Once the correct assumption is made, convergence is usually rapid and obvious. A detailed demonstration of the method can be found in [5, 17].

As an example of the sort of results we obtained, we show sequences of estimates of the coefficients of the perimeter generating function of unpunctured SAPs in table 2. In that case we conjectured that $f(i) = i$. Because of the large amount of tabular data generated by the method, we have not given this level of detail for the many series investigated here. We show only the results for two series. For the others, we just give our assessment of the apparent convergence of the sequences c_i, and the estimated value of the limits. As the equations involved are linear, the method is easy to implement, and interested readers can readily generate the relevant data themselves. Some subtleties nevertheless exist. For example, for punctured staircase polygons, the perimeter generating function has *two* singularities on the circle of convergence, and so both must be taken into account. We discuss this in more detail in the relevant section below.

As for the first stage of the analysis, the method of differential approximants, we proceeded as follows: estimates of the critical point and critical exponent were obtained by averaging values obtained from first-order $[L/N; M]$ and second-order $[L/N; M; K]$ inhomogeneous differential approximants. For each order L of the inhomogeneous polynomial we averaged over those approximants to the series which used at least the first 80–90% of the terms of the series, and used approximants such that the difference between N, M, and K did not exceed two. These are therefore 'diagonal' approximants. Some approximants were excluded from the averages because the estimates were obviously spurious. The error quoted for these estimates reflects the spread (basically one standard deviation) among the approximants. Note that these error bounds should *not* be viewed as a measure of the true error as they cannot include possible systematic sources of error. However, systematic error can also be taken into account in favourable situations as, for example, in the case of SAPs enumerated by perimeter [17]. Again, in the interests of space, we present only our results, and not the intermediate detail from which our estimates were made. An example in full detail for one of the series investigated in this study can be found in [17]. We turn now to the analysis of all the series.

Table 2. A fit to the asymptotic form $p_{2n}^{(0)} \sim \mu^2 n^{-\frac{5}{2}} [c_0 + c_1/n + c_2/n^2 + c_3/n^3 + \cdots]$ for the number of SAPs enumerated by perimeter; estimates of the amplitudes c_0, c_1, c_2, c_3.

n	c_0	c_1	c_2	c_3
20	0.099 400 85	−0.027 457 05	0.024 763 76	0.118 221 81
21	0.099 401 18	−0.027 475 48	0.025 113 47	0.116 011 07
27	0.099 401 77	−0.027 513 55	0.025 932 11	0.110 118 80
28	0.099 401 79	−0.027 515 10	0.025 972 36	0.109 770 30
29	0.099 401 80	−0.027 516 19	0.026 001 68	0.109 506 67
30	0.099 401 81	−0.027 516 94	0.026 022 73	0.109 310 43
31	0.099 401 82	−0.027 517 45	0.026 037 34	0.109 169 29
32	0.099 401 82	−0.027 517 77	0.026 046 92	0.109 073 54
33	0.099 401 82	−0.027 517 95	0.026 052 54	0.109 015 52
34	0.099 401 82	−0.027 518 02	0.026 055 00	0.108 989 29
35	0.099 401 82	−0.027 518 02	0.026 054 94	0.108 989 93
36	0.099 401 82	−0.027 517 96	0.026 052 85	0.109 013 58
37	0.099 401 82	−0.027 517 85	0.026 049 13	0.109 056 99
38	0.099 401 82	−0.027 517 71	0.026 044 08	0.109 117 57
39	0.099 401 82	−0.027 517 55	0.026 037 96	0.109 193 02
40	0.099 401 82	−0.027 517 36	0.026 030 97	0.109 281 58
41	0.099 401 82	−0.027 517 17	0.026 023 27	0.109 381 60
42	0.099 401 81	−0.027 516 96	0.026 015 00	0.109 491 74
43	0.099 401 81	−0.027 516 75	0.026 006 29	0.109 610 79
44	0.099 401 81	−0.027 516 53	0.025 997 20	0.109 737 96
45	0.099 401 81	−0.027 516 31	0.025 987 85	0.109 871 95

4.1. Staircase polygons

For (unpunctured) staircase polygons, the multi-variable width, height and area generating function is known [2]. As usual, we denote

$$(a)_n = \prod_{i=0}^{n-1} (1 - aq^i).$$

Further, denoting the first two q-Bessel functions as

$$J_0(x, y, q) = \sum_{n \geqslant 0} \frac{(-1)^n x^n q^{\binom{n+1}{2}}}{(q)_n (yq)_n}$$

and

$$J_1(x, y, q) = \sum_{n \geqslant 1} \frac{(-1)^{n-1} x^n q^{\binom{n+1}{2}}}{(q)_{n-1} (yq)_{n-1} (1 - yq^n)}$$

the perimeter and area generating function is simply

$$P(x, y, q) = y \frac{J_1(x, y, q)}{J_0(x, y, q)}$$

where x (y) is the variable conjugate to the horizontal (vertical) semi-perimeter, while q counts the area. No analogous results are known for punctured staircase polygons, although the calculation for SAPs [26] by area can be carried over *mutatis mutandis* to prove that the area generating function for a punctured staircase polygon with k holes has the same radius of convergence as the area generating function for unpunctured staircase polygons. Further, the critical exponent increases by 1 for each puncture, and unlike the case for SAPs, not only can

we prove the existence of a critical exponent for unpunctured staircase polygons, but we know its value. Hence we know the leading term in the asymptotic expansion of the generating function by area for k-punctured staircase polygons.

For the expected behaviour of the perimeter generating function of punctured staircase polygons, the arguments of the preceding section apply directly. The radius of convergence, and hence the connective constant remains unchanged, and the argument given in the preceding section suggests that the critical exponent should increase by 1.5 for each puncture. The results of our analysis, presented below, bear this out.

4.1.1. Area generating function. For unpunctured staircase polygons, the area generating function is given by

$$A(q) = \sum_{n \geqslant 1} a_n^{(0)} q^n = \frac{J_1(1, 1, q)}{J_0(1, 1, q)}.$$

By inspection, this has poles at the zeros of $J_0(1, 1, q)$. The nearest zero is at [1] $1/q = \eta = 2.309\,138\,593\,30$, and there is a simple pole at that point. The next zero is well separated (at $1/q = \lambda = 1.4435\ldots$) and so the asymptotic form of the generating function is

$$A(q) \sim D/(1 - \eta q) + E/(1 - \lambda q) + \cdots$$

and hence

$$a_n^{(0)} = [q^n] A(q) \sim \eta^n (c_0 + \mathrm{O}((\lambda/\eta))^n).$$

Our analysis bears this out, and we estimate $c_0 = 0.128\,815\,79$.

For 1-punctured discs, our analysis, based on more than 100 series coefficients, convincingly suggests the following asymptotic form:

$$A^{(1)}(q) \sim D^{(1)}/(1 - \eta q)^2 + E^{(1)}/(1 - \eta q)^{1.5} + F^{(1)}/(1 - \eta q) + \cdots$$

and hence

$$a_n^{(1)} = [q^n] A^{(1)}(q) \sim \eta^n n \sum_{i \geqslant 0} c_i / n^{i/2}.$$

The sequences of amplitude estimates, assuming this asymptotic form, are shown in table 3. The apparent convergence of the amplitude estimates is, as explained above, our source of evidence for this asymptotic form.

For 2-punctured discs, a similar analysis, based on some 86 series coefficients, convincingly suggests the following asymptotic form:

$$A^{(2)}(q) \sim D^{(2)}/(1 - \eta q)^3 + E^{(2)}/(1 - \eta q)^{2.5} + \cdots$$

and hence

$$a_n^{(2)} = [q^n] A^{(2)}(q) \sim \eta^n n^2 \sum_{i \geqslant 0} a_i / n^{i/2}.$$

For 3-punctured discs our analysis was based on an 89-term series. We found that the above pattern persists, so that the generating function has the following asymptotic form:

$$A^{(3)}(q) \sim D^{(3)}/(1 - \eta q)^4 + E^{(3)}/(1 - \eta q)^{3.5} + \cdots$$

and hence

$$a_n^{(3)} = [q^n] A^{(3)}(q) \sim \eta^n n^3 \sum_{i \geqslant 0} a_i / n^{i/2}.$$

Estimates of the various amplitudes defined above are shown in table 4.

Table 3. A fit to the asymptotic form $a_n^{(1)} \sim \eta^n n[c_0 + c_1/n^{\frac{1}{2}} + c_2/n + c_3/n^{\frac{3}{2}} + c_4/n^2 + \cdots]$ for the number of 1-punctured staircase polygons enumerated by area; estimates of the amplitudes c_0, c_1, c_2, c_3, c_4.

n	c_0 ($\times 10^{-3}$)	c_1 ($\times 10^{-2}$)	c_2 ($\times 10^{-2}$)	c_3 ($\times 10^{-2}$)	c_4 ($\times 10^{-2}$)
94	6.832 79	−1.862 63	−2.632 49	2.959 64	8.605 93
95	6.832 58	−1.861 82	−2.644 19	3.034 65	8.425 57
96	6.832 56	−1.861 73	−2.645 59	3.043 63	8.403 88
97	6.832 39	−1.861 09	−2.654 88	3.103 89	8.257 43
98	6.832 35	−1.860 92	−2.657 38	3.120 14	8.217 73
99	6.832 22	−1.860 40	−2.665 01	3.170 10	8.095 04
100	6.832 17	−1.860 19	−2.668 15	3.190 76	8.044 05
101	6.832 06	−1.859 76	−2.674 57	3.233 25	7.938 64
102	6.832 00	−1.859 53	−2.678 01	3.256 12	7.881 61
103	6.831 91	−1.859 16	−2.683 55	3.293 13	7.788 85
104	6.831 85	−1.858 93	−2.687 08	3.316 90	7.728 99
105	6.831 77	−1.858 60	−2.691 97	3.349 85	7.645 60
106	6.831 71	−1.858 37	−2.695 53	3.374 00	7.584 17
107	6.831 64	−1.858 09	−2.699 88	3.403 64	7.508 42
108	6.831 59	−1.857 86	−2.703 33	3.427 27	7.447 75
109	6.831 53	−1.857 61	−2.707 31	3.454 66	7.377 09
110	6.831 48	−1.857 39	−2.710 63	3.477 58	7.317 69
111	6.831 42	−1.857 16	−2.714 27	3.502 91	7.251 71
112	6.831 37	−1.856 96	−2.717 47	3.525 22	7.193 35
113	6.831 32	−1.856 75	−2.720 77	3.548 32	7.132 66

4.1.2. Perimeter generating function. For unpunctured staircase polygons, the perimeter generating function (ignoring the distinction between height and width) is given by

$$P(x) = \frac{1 - 2x - \sqrt{1 - 4x}}{2}$$

which is, apart from suppression of the first term, the generating function for Catalan numbers. Hence

$$p_{2n}^{(0)} = [x^n]P(x) = \frac{1}{n}\binom{2n - 2}{n - 1} \sim 4^n/n^{\frac{3}{2}} \sum_{i \geqslant 0} c_i/n^i.$$

The values of c_i follow immediately from the exact solution. They are given in table 4.

For 1-punctured discs, our analysis, again based on more than 100 series coefficients, was more equivocal than that of the 1-punctured area generating function. The method of differential approximants clearly identified a singularity at the known critical point, $x_c = \frac{1}{4}$, but almost all approximants had a double root, implying a confluent singularity. The leading exponent was estimated to be -1, implying a pole in the generating function, (and hence immediately lending support to our conjectured change in the critical exponent of $\frac{3}{2}$ as a result of puncturing), but we were unable, from this method, to identify the confluent exponent. Further, a second singularity was identified, of the form const.$(1 + x/x_c)^{6.5}$. At this point we wish to remark on the close similarity between a recently solved model of polygons, the *three-choice polygon* model [6] and 1-punctured staircase polygons. 1-punctured staircase polygons can be thought of as being constructed from two three-choice polygons, with common edges deleted. This geometric similarity is borne out by the fact that the two models have identical singularity distributions, with even the exponents being the same at all non-physical singularities.

Table 4. Amplitude estimates appearing in the asymptotic form of the coefficients of the various models considered. The order of the term associated with the amplitude c_i varies from model to model, and is given in the text for each model. The various connective constants are: $\eta = 2.309\,138\,593\,30$, $\kappa = 3.970\,943\,97$, $\mu = 2.638\,158\,530\,34$, $\tau = 4.062\,591$. All numbers quoted are expected to have errors only in the last quoted digit. k is the number of punctures. The nth coefficient is given by prefactor $\times \sum_{i \geqslant 0} c_i / n^{f(i)}$. See the text for the problem dependent value of $f(i)$.

Model, parameter	k	Prefactor	c_0	c_1	c_2	c_3
Staircase	0	η^n	$0.128\,815\,79$	$O(1.5996^{-n})$		
polygons	1	$\eta^n n$	$0.006\,831$	-0.0185	-0.028	0.04
by area	2	$\eta^n n^2$	7.87×10^{-5}	$-0.000\,43$	0.0010	-0.007
	3	$\eta^n n^3$	6.04×10^{-7}	-5.03×10^{-6}	$0.000\,031$	$-0.000\,22$
Staircase	0	$4^n n^{-\frac{3}{2}} / \sqrt{\pi}$	$\frac{1}{4}$	$\frac{3}{32}$	$\frac{25}{512}$	$\frac{64}{4096}$
polygons	1	4^n	0.0147	-0.19	1.4	
by	2	$4^n n^{\frac{3}{2}}$	8.0×10^{-4}	-2×10^{-2}	0.2	
perimeter	3	$4^n n^3$	3.0×10^{-6}	-1×10^{-3}	0.02	
SAPs	0	κ^n / n	$0.408\,105$	-0.5467	0.626	-3
by	1	κ^n	$0.000\,975$	-0.0097	-0.04	
area	2	$\kappa^n n$	$0.000\,001\,18$	$-0.000\,019$	-0.0001	
	3	$\kappa^n n^2$	1.0×10^{-9}	-2×10^{-8}	-3×10^{-7}	
SAPs	0	$\mu^{2n} n^{-\frac{5}{2}}$	$0.099\,4018$	-0.02751	0.0255	0.12
by	1	μ^{2n} / n	$0.001\,444$	$-0.008\,43$	0.0078	0.026
perimeter	2	$\mu^{2n} n^{\frac{1}{2}}$	$0.000\,011$	-0.0001	0.0005	
	3	$\mu^{2n} n^2$	1×10^{-7}			
Polyominoes	—	τ^n / n	$0.316\,60$	-0.233	0.62	-2.5
Punctured	1	κ^n	$0.009\,22$	-0.107	0.30	
polyominoes	2	$\kappa^n n$	$0.000\,104$	-0.0022	0.009	
	3	$\kappa^n n^2$	$0.000\,0008$	$-0.000\,02$	0.0002	

For 1-punctured discs then, the perimeter generating function is expected to be of the following asymptotic form:

$$P^{(1)}(x) = \sum_n p_{2n}^{(1)} x^n \sim B^{(1)}(x) + C^{(1)}(x)(1 - 4x)^{-1} + D^{(1)}(x)(1 + 4x)^{6.5}$$

and hence

$$p_n^{(1)} = [x^n] P^{(1)}(x) \sim 4^n \sum_{i \geqslant 0} c_i / n^{f(i)} + (-4)^n n^{-7.5} \sum_{i \geqslant 0} d_i / n^i$$

where $B^{(1)}(x)$, $C^{(1)}(x)$ and $D^{(1)}(x)$ are assumed regular in the disc $|4x| \leqslant 1$.

Assuming $f(i) = i$, which implies only analytic correction-to-scaling terms, gave unsatisfactory results. Notably, we observed that the estimates of the amplitude c_1 in the above asymptotic form were steadily increasing, suggesting that the assumed form did not properly account for the correction-to-scaling terms. With $f(i) = i/2$ the amplitude estimates were much more stable. This then implies that the generating function in fact behaves as

$$P^{(1)}(x) \sim B^{(1)}(x) + C^{(1)}(x)(1 - 4x)^{-1} + E^{(1)}(x)(1 - 4x)^{-\frac{1}{2}} + D^{(1)}(x)(1 + 4x)^{6.5}. \tag{9}$$

The amplitude estimates, assuming this asymptotic form, are shown in table 4. The apparent convergence of the amplitude estimates is, as explained above, our source of evidence for this asymptotic form. We did not tabulate the non-physical amplitudes, as they are of little interest to our investigation. However, they need to be included to stabilize estimates of the physical amplitudes. We mention in passing that $d_0 \approx 0.14$.

We also generated series for staircase polygons with two and three punctures, with perimeter 150 and 134 steps, respectively. Our differential approximant analysis lent support to our expectation that the critical exponent increases by 1.5 per puncture. This turned out to be true for the non-physical singularity at $x = -\frac{1}{4}$ as well. A similar analysis to that for the 1-punctured polygons strongly supported the analogous asymptotic forms,

$$P^{(k)}(x) = \sum_n p_{2n}^{(k)} x^n \sim B^{(k)}(x) + C^{(k)}(x)(1 - 4x)^{0.5-1.5k} + D^{(1)}(x)(1 + 4x)^{8-1.5k} \tag{10}$$

for $k > 0$. Hence the asymptotic form of the coefficients is conjectured to be

$$p_{2n}^{(k)} \sim 4^n n^{\frac{3(k-1)}{2}} \sum_{i \geqslant 0} c_i^{(k)} / n^{i/2} + (-4)^n n^{1.5k-9} \sum_{i \geqslant 0} d_i^{(k)} / n^i$$

where the amplitude estimates $c_i^{(k)}$ are given in table 4.

We have also determined the exact generating function for staircase polygons punctured by a single hole of perimeter 4, and also the generating function for staircase polygons punctured by a single hole of perimeter 6. We obtained these generating functions by generating the coefficients using the algorithm discussed, and then searching for an underlying differential equation. As a result we find the perimeter generating functions $P_4^s(x)$ and $P_6^s(x)$ for 1-punctured staircase polygons with a hole of perimeter 4 and 6 respectively:

$$P_4^s(x) = \frac{2x^4 - 16x^3 + 20x^2 - 8x + 1}{2(1 - 4x)} - \frac{1 - 6x + 10x^2 - 4x^3}{2\sqrt{1 - 4x}} \tag{11}$$

and

$$P_6^s(x) = \frac{1 - 26x + 228x^2 - 906x^3 + 1709x^4 - 1378x^5 + 322x^6}{2(1 - 4x)^{\frac{5}{2}}}$$
$$- \frac{32x^6 - 404x^5 + 815x^4 - 586x^3 + 182x^2 - 24x + 1}{2(1 - 4x)^2}. \tag{12}$$

Note that both these exact solutions display the confluent square-root correction that we have found in our numerical investigations in the more general case. In the next section we analyse the analogous generating function for SAPs.

4.2. Self-avoiding polygons

For unpunctured SAPs, the perimeter generating function was recently extended [17] to 90-step polygons, and the asymptotics clearly identified. The polygon generating function is defined to be

$$P^{(0)}(x) = \sum_n p_{2n}^{(0)} x^n \sim A^{(0)}(x) + B^{(0)}(x)(1 - \mu^2 x)^{\frac{3}{2}} \tag{13}$$

where the functions $A^{(0)}$ and $B^{(0)}$ are believed to be regular in the vicinity of $x_c = 1/\mu^2$. We estimated $\mu = 2.638\,158\,530\,34(10)$. From this equation follows the asymptotic form of the coefficients,

$$p_{2n}^{(0)} \sim \mu^{2n} n^{-\frac{5}{2}} [c_1 + c_2/n + c_3/n^2 + c_4/n^3 + \cdots]. \tag{14}$$

We show in table 2 the sequence of estimates of c_i, and in table 4 our estimates of the amplitudes, being the limits of the sequences $\{c_i\}$.

The area generating function was first studied in [9], where the first 20 terms were given, and the asymptotic form estimated to be

$$A^{(0)}(q) = \sum_n a_n q^n \sim D(q) + E(q) \log(1 - \kappa q) \tag{15}$$

where $\kappa \approx 3.970\,87$, and the logarithm in the above equation was understood to include the possibility of some power of a logarithm other than unity. (Though our analysis below implies that this is not the case.) In this paper we extend the series to 42 terms.

4.2.1. Area generating function. Using our greatly extended 42-term series, our analysis of the unpunctured disc area generating function was carried out by standard methods. We used differential approximants [13] and found unbiased critical point and critical exponent estimates. The unbiased exponent estimate had absolute value less than 10^{-5}, totally supporting our view that it is exactly zero. Assuming this, a biased estimate of the critical point is possible, and in this way we estimate $\kappa = 3.970\,943\,97(9)$.

We then proceeded to seek the asymptotic form of the coefficients by writing

$$a_n = [q^n]A^{(0)}(q) \sim \kappa^n/n \sum_{i \geqslant 0} c_i/n^{f(i)}.$$

Our numerical results were well converged, demonstrating very convincingly that $f(i) = i$. This is the asymptotic form consistent with a pure logarithmic singularity, not raised to any power. Estimates of the amplitudes are given in table 4.

For 1-punctured discs, our analysis, based on the series known to q^{45}, convincingly suggests the following asymptotic form:

$$A^{(1)}(q) \sim D^{(1)}(q) + E^{(1)}(q)/(1 - \kappa q)$$

and hence

$$a_n^{(1)} = [q^n]A^{(1)}(q) \sim \kappa^n \sum_{i \geqslant 0} c_i/n^i.$$

Attempts to fit to alternative forms, corresponding to a confluent logarithm or a non-analytic correction-to-scaling term were unsuccessful, adding to our confidence that the above form is correct. Estimates of the amplitudes are given in table 4.

For 2-punctured discs, our analysis, based on the series known to q^{48}, convincingly suggests the following asymptotic form:

$$A^{(2)}(q) \sim D^{(2)}(q) + E^{(2)}(q)/(1 - \kappa q)^2$$

and hence

$$a_n^{(2)} = [q^n]A^{(2)}(q) \sim \kappa^n n \sum_{i \geqslant 0} c_i/n^i.$$

The conjectured asymptotic form for k-punctured polygons, by area, is thus

$$A^{(k)}(q) \sim D^{(k)}(q) + E^{(k)}(q)/(1 - \kappa q)^k.$$

4.2.2. Perimeter generating function. The results for unpunctured polygons are fully discussed at the beginning of section 4.2. As we found with staircase polygons, the generating function for punctured discs by perimeter was a more challenging numerical analysis problem than either its unpunctured counterpart, or its area counterpart.

We found that the method of differential approximants was not particularly satisfactory. Given that we needed some 100 terms to successfully analyse the (presumably simpler) problem of punctured staircase polygons by perimeter, it is not surprising that for punctured SAPs, for which we have 33 non-zero coefficients (corresponding to perimeters up to 84 steps), the method was not satisfactory. However, it did indicate the presence of a confluent singularity. As we found a confluent square-root singularity for staircase polygons, it is hardly surprising that a confluent singularity is detected for the punctured SAP generating function. In fact an

exponent shift of around 1.7 was seen, compared with the expected value of 1.5. We attribute this to the 'short' series, coupled with the well known deleterious effect of confluent terms in such an analysis. Nevertheless, subsequent analysis of the asymptotic form of the coefficients, assuming an exponent shift of 1.5, together with a square-root confluent term, as found for punctured staircase polygons, gave satisfactory results.

We denote the generating function for k-punctured SAPs, by perimeter, as

$$P^{(k)}(x) = \sum_n p_{2n}^{(k)} x^n \sim B^{(k)}(x) + C^{(k)}(x)(1 - \mu^2 x)^{1.5-1.5k} \tag{16}$$

where the exponent is conjectured. For 1-punctured SAPs, the vanishing of the exponent implies a logarithmic singularity. We fitted the coefficients to the asymptotic form appropriate to $\log(1 - \mu^2 x)$, so that the asymptotic form of the coefficients just involves decreasing integer powers of n. We then found that the estimates of the leading amplitude were monotonically increasing, which implies that the asymptotic form is wrong—too weak. Including a confluent square root singularity, as was found for punctured staircase polygons, stabilized the estimates. Accordingly, we conjecture that the asymptotic form is dominated by a logarithmic singularity, with a sub-dominant square root singularity, so that

$$p_{2n}^{(1)} = [x^n]P^{(1)}(x) \sim \mu^{2n}/n \sum_{i \geqslant 0} c_i/n^{\frac{i}{2}}. \tag{17}$$

Estimates of the amplitudes c_i are given in table 4.

For 2-punctured discs, a similar analysis suggested that the asymptotic form of the generating function is

$$P^{(2)}(x) = \sum_n p_{2n}^{(2)} x^n \sim B^{(2)}(x) + C^{(2)}(x)(1 - \mu^2 x)^{-1.5} \tag{18}$$

again with evidence of a square-root confluent term. As for 1-punctured SAPs, we give estimates of the amplitudes c_i defined by

$$p_{2n}^{(2)} \sim \mu^{2n} n^{\frac{1}{2}} \sum_{i \geqslant 0} c_i/n^{\frac{i}{2}}$$

in table 4.

For 3-punctured discs, a similar analysis suggested that the asymptotic form of the generating function is

$$P^{(3)}(x) = \sum_n p_{2n}^{(3)} x^n \sim B^{(3)}(x) + C^{(3)}(x)(1 - \mu^2 x)^{-3} \tag{19}$$

again with evidence of a square-root confluent term. As for 1-punctured SAPs, we give estimates of the amplitudes c_i defined by

$$p_{2n}^{(3)} \sim \mu^{2n} n^2 \sum_{i \geqslant 0} c_i/n^{\frac{i}{2}}$$

in table 4.

The conjectured asymptotic form for k-punctured polygons, by perimeter, is then

$$P^{(k)}(x) = \sum_n p_{2n}^{(k)} x^n \sim B^{(k)}(x) + C^{(k)}(x)(1 - \mu^2 x)^{1.5-1.5k} \tag{20}$$

where for $k > 0$ we find strong evidence for a square-root correction-to-scaling term.

4.3. Polyominoes

The problem of polyominoes has a long and interesting history, and has been well discussed in the popular scientific literature [10].

The enumeration of square lattice polyominoes to 24 steps [23] was given in 1981, extended to 25 steps in 1995 [4] and currently stands at 28 steps [20].

We have analysed the latest series by the method of differential approximants, and find the generating function behaves as

$$\mathcal{P}(y) = \sum_n a_n y^n \sim G(y) + H(y) \log(1 - \tau y) \tag{21}$$

where a_n is the number of polyominoes of area n. In [4] the estimate $\tau = 4.06265(5)$ was given. The extra terms now available allow us to make the refined estimate $\tau = 4.062591(9)$. Analysis of the asymptotic form of the coefficients is totally consistent with a simple logarithm in the generating function. Thus

$$a_n = [y^n]\mathcal{P}(y) \sim \tau^n \sum_{i \geqslant 0} c_i / n^{i+1}.$$

The amplitude estimates are given in table 4. The leading amplitude is in complete agreement with, but three orders of magnitude more accurate than that given in [11], while the order of the leading term—$O(1/n)$—was predicted by physical arguments back in 1981 [21]. In that paper, the logarithmic singularity in the generating function of strongly embedded site animals was obtained. As discussed in the introduction, these are isomorphic to polyominoes.

As well as extending the polyomino series, Oliveira e Silva has enumerated k-punctured polyominoes [20] for $k \leqslant 6$. Clearly, $k = 0$ polyominoes are just SAPs, and, as we have seen in the previous section, these grow as κ^n where $\kappa = 3.9709 \cdots < \tau$. The arguments in [25, 26] can be modified and applied to show that a k-punctured polyomino has the same growth constant λ as its unpunctured counterpart, and this is done in the appendix.

Thus, the situation is that, for any finite number of punctures,

$$a_n^{(k)} = [y^n]\mathcal{P}^{(k)}(y) \sim \kappa^n$$

but that

$$a_n = [y^n]\mathcal{P}(y) = \sum_{k \geqslant 0} a_n^{(k)} \sim \tau^n.$$

We have analysed the series for k-punctured polyominoes, $k = 0, 1, 2$, and find the asymptotic form of the coefficients to be

$$a_n^{(k)} = [y^n]\mathcal{P}^{(k)}(y) \sim \kappa^n n^{k-1} \sum_{i \geqslant 0} c_i^{(k)} / n^i$$

corresponding to a generating function for k-punctured polyominoes having a kth-order pole, namely

$$\mathcal{P}^{(k)}(y) = \sum_n a_n^{(k)} y^n \sim H^{(k)}(y)(1 - \kappa y)^{-k} \tag{22}$$

where $k = 0$ is to be interpreted as a logarithm. Thus, just as for punctured polygons, it is found that the exponent of k-punctured polyominoes increases by one for each puncture. This is also proved in the appendix.

The coefficients $c_i^{(k)}$ are given in table 4.

5. Conclusion

We have investigated the effect of punctures on SAPs and staircase polygons enumerated both by area and perimeter.

In order to do this, we have developed a new algorithm, exponentially faster than direct counting, whereby we have radically extended a number of series. This extension was necessary in order to probe some rather subtle numerical behaviour.

We found that, in every case, a finite number of punctures does not change the exponential growth factor associated with the unpunctured counterpart of the punctured object being enumerated.

This latter conclusion was also proved for polyominoes. The effect of punctures was also investigated numerically for polyominoes.

Writing $b_n^{(k)}$ for the nth coefficient in the generating function for some k-punctured object, so that

$$b_n^{(k)} \sim (\omega^{(k)})^n n^{\theta(k)}$$

we found $\omega^{(k)} = \omega^{(0)}$ in all cases. We found, further, that $\theta(k) = \theta(0) + k$ if enumerating any of the objects we have considered by area. This can be proved, although subject, in some cases, to the existence of the exponent in question. Subject to certain unproved assumptions we also showed that $\theta(k) = \theta(0) + 3k/2$ if enumerating by perimeter.

We have, for the first time, obtained good numerical estimates of the sub-dominant terms for a range of problems, thus identifying both the nature of the generating function and any correction-to-scaling terms.

We have also obtained an exact solution for the generating functions of staircase polygons, enumerated by perimeter, punctured by a single hole of perimeter 4 and of perimeter 6. These exact solutions provide additional support for our numerically based conjectures of the correction-to-scaling exponent in the general case.

A more accurate estimate of the growth constant for SAPs enumerated by area has been given, complementing our earlier work on the perimeter growth constants [17].

E-mail or WWW retrieval of series

The series for the various generating functions studied in this paper can be obtained via e-mail by sending a request to I.Jensen@ms.unimelb.edu.au or via the World Wide Web on the URL http://www.ms.unimelb.edu.au/~iwan/ by following the instructions.

Appendix A. Growth constants and exponents for k-punctured polyominoes

In this appendix we show that the growth constants are the same for all k-punctured polyominoes. Our method of proof is based on that of van Rensburg and Whittington [25, 26], making the necessary changes for the polyomino problem, and discussing in detail certain special cases. That is to say, if κ denotes the connective constant for SAPs enumerated by area, then this is the growth constant for k-punctured polyominoes for any finite k. Further, if the usual asymptotic form for the number of k-punctured polyominoes is assumed, $s_n^{(k)} \sim C_k n^{-\phi_k} \kappa^n$, and ϕ_0 exists, then $\phi_k = \phi_0 - k$.

A.1. Operations and mappings on punctured polyominoes

Let the set of all k-punctured polyominoes with n cells be denoted by $\Phi_n^{(k)}$, and the set of all polyominoes of n cells be Φ_n. Then, for each n, $\Phi_n = \bigcup_k \Phi_n^{(k)}$. Throughout, let $s_n^{(k)}$ denote the cardinality of $\Phi_n^{(k)}$ and let s_n denote the cardinality of Φ_n. Hence $s_n = \sum_k s_n^{(k)}$.

Following [25, 26] we now define operations on punctured polyominoes, including concatenation, drilling and surgery, which are needed in subsequent proofs. Concatenation allows us to change the size of polyominoes, while drilling and surgery are concerned with changing the number of holes in a polyomino.

A.1.1. Concatenation.
The concatenation mapping defined here is similar to that in [14]. Consider the bounding rectangle $R(P)$ of any polyomino $P \in \Phi_n^{(k)}$. Define the top (bottom) edge of P to be the top (bottom) edge along the east (west) side of $R(P)$.

Now, the concatenation of two polyominoes $P \in \Phi_n^{(h)}$ and $Q \in \Phi_m^{(k)}$ is defined by joining P and Q while superimposing the top edge of P and the bottom edge of Q. The result is an $(h + k)$-punctured polyomino with $m + n$ cells, see figure A.1. Hence we have a map

$$T : \Phi_n^{(h)} \times \Phi_m^{(k)} \longmapsto \Phi_{m+n}^{(h+k)}. \tag{A.1}$$

Lemma A.1. *For all non-negative values of h and k, $s_n^{(h)} s_m^{(k)} \leqslant s_{m+n}^{(h+k)}$ and in particular, $s_{n+1}^{(h)} \geqslant s_n^{(h)}$.*

Proof. For the mapping T defined above, every pair of polyominoes in the domain can be concatenated to form a larger polyomino in the codomain. Conversely, every such polyomino in the codomain can be uniquely broken up into the original ones. However, there are some polyominoes in the codomain which cannot be formed by concatenating two smaller polyominoes. An example is a 2×2 polyomino. Hence, we get the first part of the lemma as

$$|\Phi_n^{(h)} \times \Phi_m^{(k)}| \leqslant |\Phi_{m+n}^{(h+k)}|.$$

Putting $k = 0$ and $m = 1$ and noting that $|\Phi_1^{(0)}| = 1$, we get the second part. $\qquad\square$

A.1.2. Drilling.
Simply creating a hole inside an unpunctured polyomino by removing some interior cells does not allow us to drill certain classes of polyominoes. For instance, a polyomino composed of a single linear sequence of cells cannot be 'drilled' since removing any cell will either disconnect the polyomino or shorten the sequence.

The following definition of drilling differs somewhat from that in [25, 26], although the underlying idea is the same.

First, we drill one hole. Let $P_0 \in \Phi_n^{(0)}$ be the polyomino that we are to puncture. Cover P_0 by a grid system, with each grid square of size $b \times b$ cells for any $b \geqslant 5$ (the minimum size
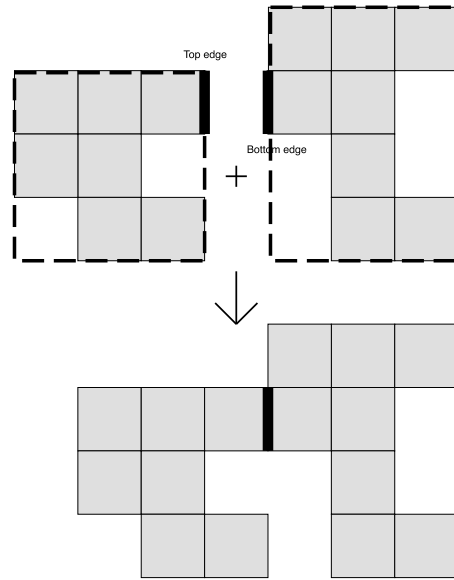
Figure A.1. Concatenation operation.

will be justified later). Say $b = 5$. Pick any grid square G that covers at least one cell of P_0 and drill a hole there, as detailed below.

Step 1. Remove all cells within the selected grid square G. If, after this step, we are left with a 1-punctured polyomino, we are done. If not, go to step 2.

Step 2. Check the corners of G: If there are two disconnected components touching each other only at the corner, connect them by adding a cell at the appropriate corner of G.

Step 3. Put a 1-punctured polyomino with 8 cells at the centre of G.

Step 4. Reconnect the disconnected components outside G to the 1-punctured polyomino by adding linear sequences of cells (non-unique). These steps are illustrated in figure A.2.
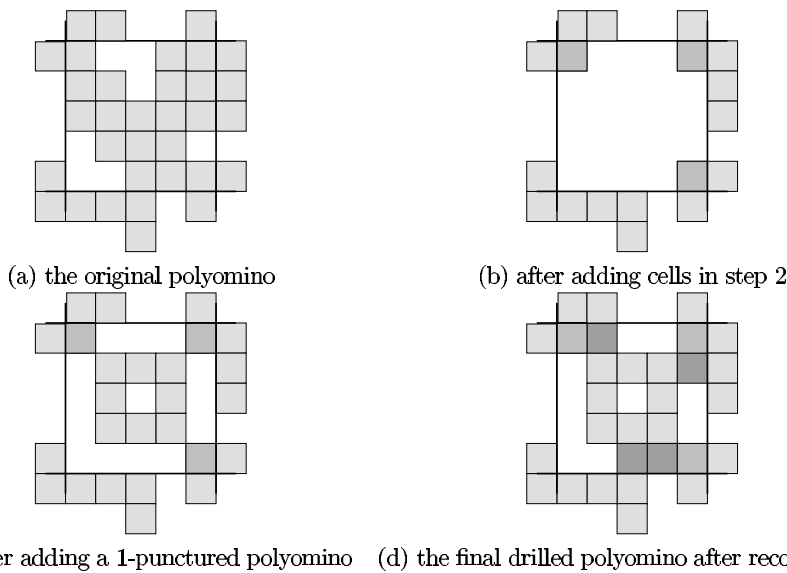
Note that the minimum value of b is 5 because if we have anything less than that, we might create extra holes unexpectedly as the 1-punctured polyomino in step 3 must touch the boundary of the grid square. An example in figure A.3 will illustrate this. In this example the original polyomino has no holes and the drilled polyomino has two holes, whereas with a grid of size 5 (or more), the number of holes in the drilled polyomino is only one.

After this operation, the maximum number of cells that could be removed is $b^2 = 25$ (finishing at step 1). The maximum number of cells that could be added is $b^2 - 2$ (this occurs when there was only one cell in the square before drilling and we end up with a 1-punctured polyomino with $b^2 - 1$ cells after the operation). So, depending on each instance of the drilling operation, we obtain a resulting polyomino $P_1 \in \Phi_j^{(1)}$ where j could be anything from $n - b^2$ to $n + (b^2 - 2)$. The drilling operation thus defines a map

$$D : \Phi_n^{(0)} \longmapsto \bigcup_{j=n-b^2}^{n+(b^2-2)} \Phi_j^{(1)}. \tag{A.2}$$

Theorem A.1. *There exists a real constant C such that for all $b \geqslant 5$,*

$$s_n^{(0)} \leqslant C s_{n+(b^2-2)}^{(1)}. \tag{A.3}$$

(a) the original polyomino   (b) after adding cells in step 2

(c) after adding a 1-punctured polyomino   (d) the final drilled polyomino after reconnections

Figure A.2. Drilling operation: (*a*) the original polyomino. After step 1, all the cells within the grid square are removed. In step 2, we reconnect all the disconnected components around the corner (shown in (*b*)). Then, in step 3, a 1-punctured polyomino is placed at the centre of the grid square (shown in (*c*)). Finally, all the disconnected pieces outside the grid square are reconnected to the punctured polyomino (shown in (*d*)).

(a) the original polyomino   (b) the drilled polyomino
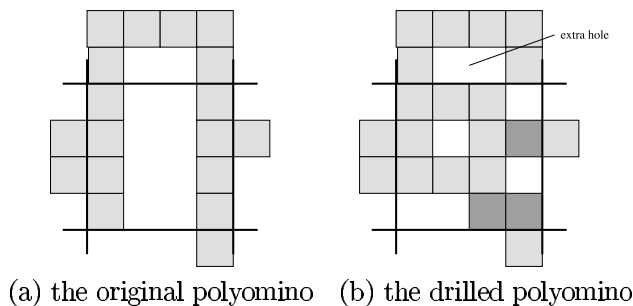
Figure A.3. An illustration of a grid square that is too small.

Proof. Consider the intermediate (possibly disconnected) polyomino after step 1, call it P^i. There are many possible initial polyominoes which gives the same P^i, i.e., all polyominoes with the same configuration outside G. They all map to 'almost' the same resulting polyomino (the uncertainty implied in 'almost' comes from the non-uniqueness of reconnecting in step 4 which will be discussed later). Therefore, the mapping from domain to codomain is M'-to-one where M' is bounded by the number of ways that at most b^2 cells can be connected within G. Let the bound be M.

On the other hand, step 4 of the drilling process is not unique. There might be more than one way to connect those disconnected components to the punctured polyomino. But since G is finite, the number of ways of reconnection is bounded above. Let that bound be C'.

Together with the drilling mapping D defined above, we can write,

$$s_n^{(0)} \leqslant C'M \sum_{j=n-b^2}^{n+(b^2-2)} s_j^{(1)}.$$

By the increasing monotonicity of $s_n^{(h)}$ over n (proven in lemma (A.1)),

$$s_n^{(0)} \leqslant (2b^2-1)C'M s_{n+(b^2-2)}^{(1)}.$$

Set $C = (2b^2-1)C'M$ and the result follows. □

Now consider the drilling of h holes. Denote by $\lceil c \rceil$ the ceiling of c, the smallest integer greater than or equal to c. Now, we could choose h drilling locations from at least $\lceil n/b^2 \rceil$ grid squares. To see this, consider first a polyomino with n cells, where $n \leqslant b^2$. The situation where we have the least number of drilling sites is when n cells fall exactly within one grid square. In this way, we only have one possible grid square where we could drill holes. Similarly, if we have $b^2n + w$ cells, where $n \geqslant 1$ and $0 \leqslant w < b^2$, the minimal number of grid squares is when b^2n cells fall into exactly n grid squares and the other w cells falls into another single grid. Then we have $n + 1$ possible grid squares to drill holes. Therefore, we have at least $\lceil n/b^2 \rceil$ grid squares where we could drill holes.

Letting H be a set of grid squares, let $\Phi_n^{[H]}$ denote the set of all polyominoes with n cells and $|H|$ holes where there is a hole in each grid square of H. Let $s_n^{[H]}$ denote its cardinality. One property of this set is that, for any $n_1, n_2 \in Z^+$ such that $n_1 \leqslant n_2$,

$$s_{n_1}^{[H]} \leqslant s_{n_2}^{[H]}.$$

Theorem A.2. *There exists a real constant K such that for all $b \geqslant 5$,*

$$\binom{\lceil n/b^2 \rceil}{h} s_n^{(0)} \leqslant K s_{n+(b^2-2)h}^{(h)} \qquad \forall b \leqslant \lceil n/b^2 \rceil. \tag{A.4}$$

Proof. Place the polyomino in the grid system. Let A' be the set of all grid squares where we could drill holes. From previous results, we know there are at least $\lceil n/b^2 \rceil$ elements in A'. Truncate the set A' with only the first $\lceil n/b^2 \rceil$ elements and call this set of drilling sites A.

Pick a subset $H \subseteq A$ such that $|H| = h$ $(\leqslant \lceil n/b^2 \rceil)$ and drill holes one by one in each grid square in H, leading to a series of mappings:

$$M : \Phi_n^{(0)} \longmapsto \Phi_{m_1}^{(1)} \longmapsto \Phi_{m_2}^{(2)} \longmapsto \cdots \longmapsto \Phi_{m_h}^{[H]}$$
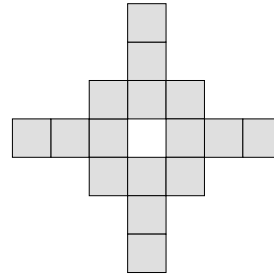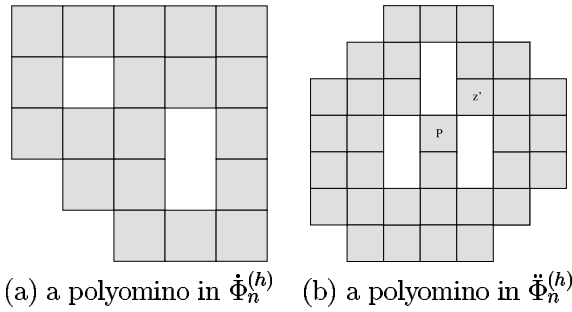
where m_i are appropriate constants depending on each instance of operation and $m_i \leqslant n + (B^2 - 2)h \forall i$ So, from this mapping M, we have

$$s_n^{(0)} \leqslant C^h s_{m_h}^{[H]} \leqslant C^h s_{n+(b^2-2)h}^{[H]}. \tag{A.5}$$

Hence

$$\sum_{H \in A} s_n^{(0)} \leqslant \sum_{H \in A} C^h s_{n+(b^2-2)h}^{[H]}$$

$$\binom{\lceil n/b^2 \rceil}{h} s_n^{(0)} \leqslant C^h \sum_{H \in A} s_{n+(b^2-2)h}^{[H]}$$

$$\binom{\lceil n/b^2 \rceil}{h} s_n^{(0)} \leqslant K s_{n+(b^2-2)h}^{(h)}.$$

The last line arises since $\bigcup_{H \in A} \Phi_n^{[H]} \subseteq \Phi_n^{(|H|)} \equiv \Phi_n^{(h)}$ and put $K = C^h$. □

(a) a polyomino in $\dot{\Phi}_n^{(h)}$ (b) a polyomino in $\ddot{\Phi}_n^{(h)}$

Figure A.4. An illustration of two special cases for the surgery operation. In (*b*), *P* denotes the problem cell whose removal will join three holes together. In this case we choose to remove z'.

Figure A.5. A polyomino in $\dot{\Phi}_n^{(h)}$ where we cannot apply the general surgery operation.

A.1.3. Surgery. The surgery operation removes a linear sequence of cells inside the polyomino and concatenates the sequence to the external boundary of the polyomino. Our objective is to join two holes, thereby reducing the number of holes by one.

We divide our domain into three classes:

(1) the set of polyominoes which have at least one hole with size one cell: denote this set by $\dot{\Phi}_n^{(h)}$ with cardinality $\dot{s}_n^{(h)}$, see figure A.4(*a*);
(2) the set of polyominoes which are not in (1), and have three holes touching each other at corners around a single cell: denote this set by $\ddot{\Phi}_n^{(h)}$ with cardinality $\ddot{s}_n^{(h)}$, see figure A.4(*b*);
(3) the set of polyominoes which are not in (1) nor (2) (the general case): denote this set by $\tilde{\Phi}_n^{(h)} = \Phi_n^{(h)} \backslash (\dot{\Phi}_n^{(h)} \bigcup \ddot{\Phi}_n^{(h)})$ with cardinality $\tilde{s}_n^{(h)}$.

The general case. First, let us look at polyominoes in $\tilde{\Phi}_n^{(h)}$. Consider a polyomino $\alpha_n^{(h)} \in \tilde{\Phi}_n^{(h)}$. Let L be the set of all loop-free sequences† of cells in $\alpha_n^{(h)}$, one end of which must touch the boundary of one hole and the other end must touch the boundary of another hole (or the exterior boundary). Define the length of the sequence to be the number of cells in the sequence. Let the set Z be the set of sequences in L that has minimum length.

Now pick one sequence $z \in Z$ and cut it out. This step will join two holes together (or join one hole and the exterior together). Next, using the definition in section appendix A.1.1, concatenate $\alpha_n^{(h)} \backslash cz \oplus D \oplus z$ (D is a 3×3 polyomino block), and the resulting polyomino has one less hole and nine more cells than the original one. So we have a mapping

$$S_3 : \tilde{\Phi}_n^{(h)} \longmapsto \Phi_{n+9}^{(h-1)}.$$

This mapping is at most n-to-one because for any polyomino in the codomain, we could find at most n locations to connect z back. As a result, we find that

$$\tilde{s}_n^{(h)} \leqslant n s_{n+9}^{(h-1)}. \tag{A.6}$$

Special cases. There are some subclasses of class (1) where we cannot apply the general surgery operation. One example is shown in figure A.5 [27]. In this example, the minimum length of the loop-free sequence connecting two boundaries is two, but when we try to remove any such minimum sequence, we will disconnect the polyomino.

---

† A loop-free sequence is a sequence of cells where each successive pair of cells are joined and no cell appears more than once in the whole sequence.

To deal with class (1) polyominoes, we simply fill such a hole with a single cell. This satisfies our objective of reducing the number of holes by one. So we find a mapping

$$S_1 : \dot{\Phi}_n^{(h)} \longmapsto \Phi_{n+1}^{(h-1)}.$$

The mapping S_1 is n' to 1 where n' is less than n. To see this, consider a polyomino in the codomain. To map back to the domain, we can choose any *interior* cell to remove and the number of choices is obviously less than the total number of cells in the polyomino, so $n' \leqslant n$. Therefore,

$$\dot{s}_n^{(h)} \leqslant n s_n^{(h-1)}. \tag{A.7}$$

Finally, consider polyominoes in $\ddot{\Phi}_n^{(h)}$. The problem with these polyominoes is that when we try to remove the cell between the three holes, we will inevitably join three holes together instead of joining two. One way to get around this is to deliberately choose another sequence (which is also a cell in this class) in the set Z. In particular, we choose to remove the cell $z' \in Z$ such that z' is the top cell (in lexicographic ordering) of Z which is not the problem cell. Similarly to the general case, we find a mapping

$$S_2 : \ddot{\Phi}_n^{(h)} \longmapsto \Phi_{n+9}^{(h-1)}$$

and hence

$$\ddot{s}_n^{(h)} \leqslant n s_{n+9}^{(h-1)}. \tag{A.8}$$

Adding (A.6)–(A.8), we have

$$s_n^{(h)} = \dot{s}_n^{(h)} + \ddot{s}_n^{(h)} + \tilde{s}_n^{(h)} \leqslant n s_{n+9}^{(h-1)} + n s_{n+9}^{(h-1)} + n s_n^{(h-1)}$$

and, hence, the following theorem.

Theorem A.3. *The number of polyominoes of area n with h holes is bounded above by $3n$ times the number of polyominoes of area $(n + 9)$ with one less hole. That is,*

$$s_n^{(h)} \leqslant 3n s_{n+9}^{(h-1)}. \tag{A.9}$$

A.2. Growth constants (by area)

The following theorem proves the existence and equality of all the growth constants for k-punctured polyominoes for all finite k.

Theorem A.4. *There exists a constant β_0 such that for all $h \geqslant 0$*

$$\lim_{n \to \infty} n^{-1} \log(s_n^{(h)}) = \log(\beta_0). \tag{A.10}$$

Proof. We use induction. First, β_0 exists [14]. Assume $\beta_h = \lim_{n \to \infty} n^{-1} \log(s_n^{(h)})$ exists. From the results of the concatenation and surgery operations, we have

$$s_{n-m}^{(h)} s_m^{(1)} \leqslant s_n^{(h+1)} \leqslant 3n s_{n+9}^{(h)}.$$

Choose some value m such that $0 < s_m^{(1)} < \infty$, for example $m = 8$. Take the logarithm, divide by n and take the limit $n \to \infty$. This gives

$$\log(\beta_h) \leqslant \log(\beta_{h+1}) \leqslant \log(\beta_h)$$

and hence $\beta_h = \beta_{h+1}$. Iterating from β_0 gives $\beta_0 = \beta_h$ for all finite h. $\qquad\square$

A.3. Relationships between critical exponents

The following theorem establishes the relationships between critical exponents should they exist.

Theorem A.5. *Assume for all h $s_n^{(h)} \sim C_h n^{-\phi_h} \beta_0^n$ where C_h is a h-dependent constant. Then*

$$\phi_h = \phi_0 - h. \tag{A.11}$$

Proof. From the results of the drilling operation, we have

$$C^{-h} \binom{\lceil n/b^2 \rceil}{h} s_{n-(b^2-1)}^{(0)} \leqslant s_n^{(h)}. \tag{A.12}$$

Since

$$\binom{\lceil n/b^2 \rceil}{h} \sim \frac{1}{h!} \lceil n/b^2 \rceil^h$$

substituting the assumed asymptotic form $s_n^{(h)} \sim C_h n^{-\phi_h} \beta_h^n$ into (A.12), dividing by β_0^n, taking logarithms, dividing by $\log(n)$, letting $n \to \infty$ and using the above result, we get

$$h - \phi_0 \leqslant -\phi_h. \tag{A.13}$$

Next, from the result of the surgery operation, we have

$$s_n^{(h)} \leqslant 3n s_{n+9}^{(h-1)} \leqslant (3n)3(n+9)s_{n+18}^{(h-2)} \leqslant \cdots$$
$$\leqslant (3n)3(n+9)\ldots 3(n+9h)s_{n+9(h-1)}^{(0)} \leqslant [3(n+9h)]^h s_{n+9h}^{(0)}.$$

Again by substituting the assumed asymptotic form, dividing by $\beta_0^{(n)}$, taking logarithms, dividing by $\log(n)$ and letting $n \to \infty$, we get

$$-\phi_h \leqslant h - \phi_0. \tag{A.14}$$

Hence $\phi_h = \phi_0 - h$.　　　　$\square$

References

[1] Bender E A 1974 *Discrete Math.* **8** 219
[2] Bousquet-Mélou M 1996 *Discrete Math.* **154** 1
[3] Cardy J 1999 *Preprint* cond-mat/9911457
[4] Conway A R and Guttmann A J 1995 *J. Phys. A: Math. Gen.* **28** 891
[5] Conway A R and Guttmann A J 1996 *Phys. Rev. Lett.* **77** 5284
[6] Conway A R, Guttmann A J and Delest M P 1997 *Math. Comput. Modeling* **26** 51
[7] Enting I G 1980 *J. Phys. A: Math. Gen.* **13** 3713
[8] Enting I G and Guttmann A J 1989 *J. Phys. A: Math. Gen.* **22** 1371
[9] Enting I G and Guttmann A J 1990 *J. Stat. Phys.* **58** 475
[10] Golomb S 1994 *Polyominoes: Puzzles, Patterns, Problems and Packings* 2nd edn (Princeton, NJ: Princeton University Press)
[11] Guttmann A J 1982 *J. Phys. A: Math. Gen.* **15** 1987
[12] Guttmann A J and Enting I G 1988 *J. Phys. A: Math. Gen.* **21** L165
[13] Guttmann A J 1989 *Phase Transitions and Critical Phenomena* vol 13, ed C Domb and J L Lebowitz (New York: Academic)
[14] Hammersley J M 1961 *Proc. Camb. Phil. Soc.* **57** 516
[15] Hughes B D 1995 Random walks *Random Walks and Random Environments* vol 1 (Oxford: Clarendon)
[16] Jensen I and GuttmannA J 1998 *J. Phys. A: Math. Gen.* **31** 8137
[17] Jensen I and Guttmann A J 1999 *J. Phys. A: Math. Gen.* **32** 4867
[18] Madras N, Soteros C E and Whitington S G 1988 *J. Phys. A: Math. Gen.* **21** 4617
[19] Nienhuis B 1982 *Phys. Rev. Lett.* **49** 1062

[20] Oliveira e Silva T http://www.inesca.pt/~tos/animals.html
[21] Parisi G and Sourlas N 1981 *Phys. Rev. Lett.* **46** 871
[22] Orlandini E, Tesi M C, Janse van Rensburg E J and Whitington S G 1996 *J. Phys. A: Math. Gen.* **29** L299
[23] Redelmeier D H 1981 *Discrete Math.* **36** 191
[24] Soteros C E and Whitington S G 1988 *J. Phys. A: Math. Gen.* **21** 2187
[25] Janse van Rensburg E J 1992 *J. Phys. A: Math. Gen.* **25** 3259
[26] Janse van Rensburg E J and Whittington S G 1989 *J. Phys. A: Math. Gen.* **22** 4939
[27] Janse van Rensburg E J 1999 Private communication
[28] Whitington S G, Torrie G M and Gaunt D S 1983 *J. Phys. A: Math. Gen.* **16** 1695