ARTICLES

On the Number of Benzenoid Hydrocarbons

Markus Vöge, Anthony J. Guttmann,* and Iwan Jensen

Department of Mathematics and Statistics, The University of Melbourne, Victoria 3010, Australia

Received September 10, 2001

We present a new algorithm which allows a radical increase in the computer enumeration of benzenoids b_h with h hexagons. We obtain b_h up to $h = 35$. We prove that $b_h \sim const.\kappa^h$, prove the rigorous bounds $4.789 \le \kappa \le 5.905$, and estimate that $\kappa = 5.16193016(8)$. Finally, we provide strong numerical evidence that the generating function $\sum b_h z^h \sim A(z) \log(1 - \kappa z)$, estimate $A(1/\kappa)$ and predict the subleading asymptotic behavior. We also provide compelling arguments that the mean-square radius of gyration $\langle R_g^2 \rangle_h$ of benzenoids of size h grows as $h^{2\nu}$, with $\nu = 0.64115(5)$.

1. INTRODUCTION

A *benzenoid* or *planar polyhex* is a special type of hydrocarbon molecule. Its hexagonal system is obtained by deleting all carbon−hydrogen bonds, leaving clusters of hexagons joined at an edge (a carbon−carbon bond). They thus appear as connected geometric figures, being clusters of identical hexagons in the plane, joined at an edge. All internal regions of the cluster are filled with hexagons, that is to say, there are no internal holes. These structures appear to have an independent identity in the chemical and mathematical literature. In the mathematics literature they are discussed as *self-avoiding polygons on the hexagonal lattice.* Furthermore, they are then counted by either *area*—the number of cells, which are each deemed to be of unit area, or by *perimeter*—which is the number of bonds or edges of the perimeter. Thus a single hexagon has area 1 and perimeter 6, while a two-celled polygon clearly has area 2 and perimeter 10.

In the following we will solely be discussing the number of self-avoiding polygons counted by area. (See ref 1 for counts by perimeter.) We will frequently abbreviate this to *polygons.* Even this is an insufficient description, as in the mathematics literature a distinction is made between *fixed* embeddings and *free* embeddings. In the former, polygons are considered distinct up to a translation, that is to say, fixed polygons means "an equivalence class of self-avoiding polygons under translation" while free polygons are considered equivalent under translations, rotations and reflections. More precisely, free polygons refers to "an equivalence class of self-avoiding polygons under translation, rotation and reflection". In Figure 1 a simple example is shown of a hexagonal cluster which has a count of 1 as a free polygon and a count of 12 as a fixed polygon. In the chemistry literature the number of free polygons[2] has been universally considered. This is precisely the number of benzenoids or planar polyhexes.

For many years the enumeration of the number b_h of benzenoids of h cells has been an important topic. The monograph by Gutman and Cyvin[2] provides a comprehensive review of all aspects, with a more up-to-date review in ref 3. Progress has been slow, but incremental, as all previous calculations have been based on direct counting of polygons. As the number of these grows as $b_h \sim 5.16^h$, it is clear that, to obtain one further term, one needs more than 5 times the computer power than one previously had—or 5 times as many computers if one is performing calculations in parallel. Up to 1989, the number of benzenoids up to $h = 12$ was known.[2] Ten years later, this had been improved to $h = 21$,[4] while very recently, the number of benzenoids up to $h = 24$ has been obtained.[5] Thus one extra term per year has been found on average, reflecting a steady 5-fold increase per annum in a combination of computer speed and resources.

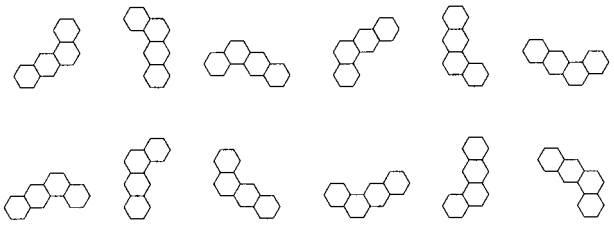

Figure 1. A polygon that counts as one free polygon but as 12 fixed polygons.

In this paper, we present a different algorithm that enables us to obtain b_h for $h \le 35$. This represents an improvement of $5.16^{11} \approx 7 \times 10^7$ over preexisting calculations (or a jump of about 11 years in terms of the traditional algorithms). The algorithm is in fact exponentially faster than direct counting, with both time and memory growing approximately as 1.65^h. Its drawbacks are that it is much more memory intensive than direct counting, for which memory requirements are negligible, as well as being much more difficult to implement.

In this paper we first *prove* some results about b_h. As far as we are aware, no such proofs exist in the chemical literature. Rather, they have been part of the folklore of the discipline for so long that they are accepted without question. We then describe the algorithm and give both the number

* Corresponding author phone: (+61-3)-8344-4316; e-mail: tonyg@ms.unimelb.edu.au.

10.1021/ci010098g CCC: $22.00
Published on Web 03/22/2002

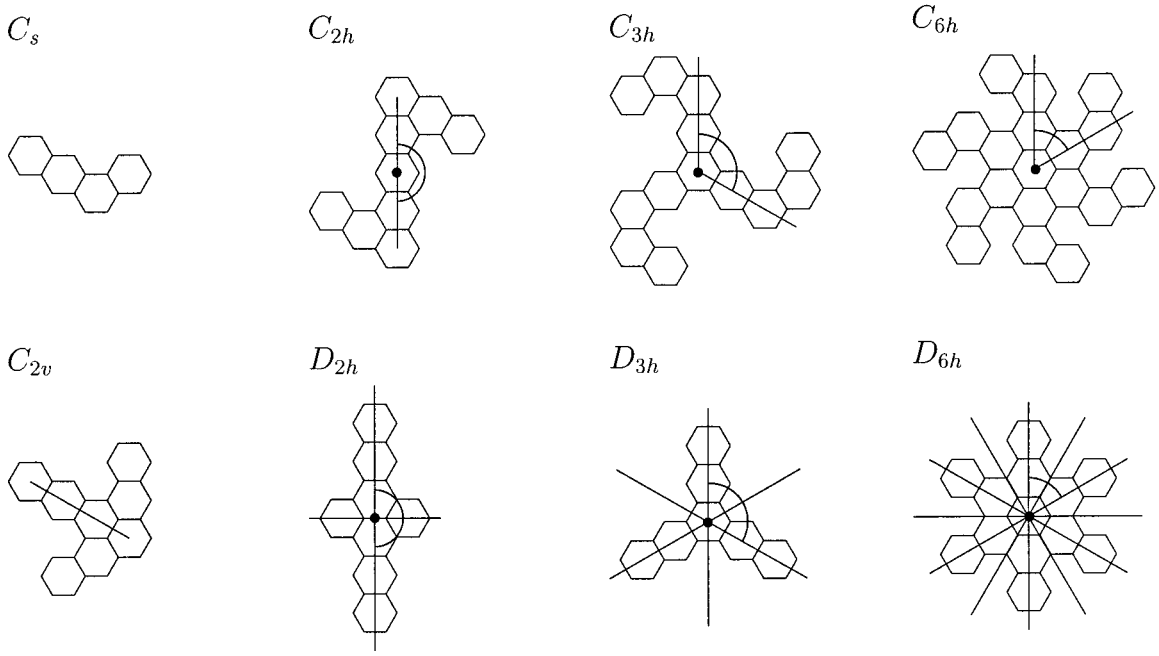

Figure 2. One benzenoid system of each symmetry class is shown. The symmetries are indicated.

of fixed polygons and benzenoids (free polygons) up to size 35. We then apply a range of numerical techniques to the data and thus provide strong evidence that the generating function $G(z) = \sum_{h\geq 1} b_h z^h \sim A(z) \log(1 - \kappa z)$, and, further, that $b_h = const.\kappa^h/h + O(\kappa^h/h^2)$.

We establish rigorous upper and lower bounds on κ and give a very precise (but nonrigorous) estimate of κ. In the following, we will frequently refer to a *polygon* when we mean a benzenoid. Except as necessary, we do not distinguish between fixed and free embeddings, though when the distinction does matter, we are quite precise. The reader may, however, consider "polygon" to be equivalent to "benzenoids" except where otherwise stated.

Another interesting property of such systems is their average size. Many possible measures exist, and most of these are equivalent. A common measure, and one that is possibly experimentally accessible, is the *mean-square radius of gyration*, $\langle R_g^2 \rangle_h$ of benzenoids of size h. We provide persuasive numerical arguments (though not a proof) that this grows as $h^{2\nu}$, with $\nu = 0.64115(5)$.

2. RIGOROUS RESULTS

We first prove (or outline proofs) of some rigorous results concerning the number of benzenoids. Let B_h denote the number of *fixed* configurations of h hexagons, and let b_h denote the number of *free* configurations of h hexagons. We prove that B_h is asymptotically proportional to b_h. [We say that $f(h)$ is *asymptotically proportional* to $g(h)$ if $f(h)/g(h)$ converges to a finite, nonzero constant.] In this case the constant is exactly 12. Because of this property, we may prove asymptotic results either about B_h or b_h. It turns out that many of our proof techniques, and indeed numerical results, are more readily obtained from the sequence $\{B_h\}$, and we will work primarily with this sequence, transferring our results to statements about $\{b_h\}$ as appropriate.

We first discuss the connection between the two sequences, which requires a careful consideration of the symmetries possessed by different subclasses of benzenoids. There are eight distinct symmetry classes, described below and shown in Figure 2.

We adopt standard group theoretical notation and denote these as follows:

- $C_s^{(h)}$ for benzenoids of area h with no rotational or reflection symmetry,
- $C_{2v}^{(h)}$ for those of area h with one axis of reflection symmetry,
- $C_{2h}^{(h)}$ for those of area h invariant with respect to rotations through π,
- $D_{2h}^{(h)}$ for those of area h with two axes of reflection symmetry and invariant with respect to rotations through π,
- $C_{3h}^{(h)}$ for those of area h and invariant with respect to rotations through $2\pi/3$,
- $D_{3h}^{(h)}$ for those of area h with three axes of reflection symmetry and invariant with respect to rotations through $2\pi/3$,
- $C_{6h}^{(h)}$ for those of area h and invariant with respect to rotations through $\pi/3$, and finally
- $D_{6h}^{(h)}$ for those of area h with six axes of reflection symmetry and invariant with respect to rotations through $\pi/3$.

Note that polygons are classified according to the *highest* symmetry they possess. Thus a polygon with an axis of reflection symmetry, and rotational symmetry through π (such as the benzenoid of two cells), is considered to be in D_{2h}, even though it also possesses the symmetries that are required by C_{2v} and C_{2h}.

In terms of these, the number of benzenoids b_h may be written as

$$b_h = C_s^{(h)} + C_{2v}^{(h)} + C_{2h}^{(h)} + D_{2h}^{(h)} + C_{3h}^{(h)} + D_{3h}^{(h)} + C_{6h}^{(h)} + D_{6h}^{(h)} \quad (1)$$

while the number of *fixed* polygons B_h may be written as

$$B_h = 12C_s^{(h)} + 6C_{2v}^{(h)} + 6C_{2h}^{(h)} + 3D_{2h}^{(h)} + 4C_{3h}^{(h)} + 2D_{3h}^{(h)} + 2C_{6h}^{(h)} + D_{6h}^{(h)} \quad (2)$$

Eliminating $C_s^{(h)}$ we arrive at

$$b_h = \frac{1}{12}(B_h + 6C_{2v}^{(h)} + 6C_{2h}^{(h)} + 9D_{2h}^{(h)} + 8C_{3h}^{(h)} + 10D_{3h}^{(h)} + 10C_{6h}^{(h)} + 11D_{6h}^{(h)}) \quad (3)$$

Our method of counting is to count all terms that appear on the right-hand side of eq 3 and hence to evaluate b_h. We have used a new algorithm, based on transfer matrices together with the *finite lattice method*, or FLM,[6] which we describe below, to count $\{B_h\}$. We have counted the other configurations, those that possess some symmetry, by direct counting—that is to say, a simple back-tracking algorithm. These symmetric configurations are much faster to count, because they are far fewer—in fact exponentially fewer. To take a particular example, consider elements of $\{C_{2v}^{(h)}\}$ (where for simplicity we are considering the situation with h even, and no cells bisected by the axis of symmetry. A more elaborate argument incorporates both these complications). Each element can be defined by an $h/2$ celled benzenoid, joined uniquely to its mirror image. As we prove below, the number of $h/2$-celled benzenoids grows such as $\kappa^{h/2}$, so it immediately follows that h-celled members of $\{C_{2v}^{(h)}\}$ also grow like $\kappa^{h/2}$. That is to say, if $\kappa > 1$ is the growth constant for benzenoids, the growth constant for elements of $\{C_{2v}^{(h)}\}$ is $\sqrt{\kappa}$. Similar arguments apply to elements of $\{C_{2h}^{(h)}\}$, while for elements of $\{D_{2h}^{(h)}\}$ the appropriate growth constant is $\kappa^{1/4}$, for $\{D_{2h}^{(h)}\}$ it is $\kappa^{1/3}$, for $\{D_{3h}^{(h)}\}$ it is $\kappa^{1/6}$, for $\{C_{6h}^{(h)}\}$ it is $\kappa^{1/6}$, and for $\{D_{6h}^{(h)}\}$ it is $\kappa^{1/12}$. Thus we may write $B_h \sim c_1 \times \kappa^h$, $C_{2v}^{(h)} \sim c_2 \times \kappa^{h/2}$, and $C_{2h}^{(h)} \sim c_3 \times \kappa^{h/2}$, where c_1, c_2, and c_3 are constants, with similar results for the remaining terms. Substituting into eq 3 we obtain

$$b_h \sim \frac{1}{12}(c_1\kappa^h + 6c_2\kappa^{h/2} + 6c_3\kappa^{h/2} + O(\kappa^{h/3})) \quad (4)$$

If we write $b_h \sim c_b\kappa^h$, divide through by κ^h, and take the limit as $h \to \infty$, we obtain $c_b = c_1/12$, subject to the proof of our assertion that $B_h \sim \kappa^h$. We prove this immediately below.

Theorem 1: *There is a* growth constant κ *with* $1 < \kappa < \infty$, *for fixed hexagonal polygons, such that*

$$\lim_{h\to\infty} B_h^{1/h} = \kappa = \sup_{h\geq 1} B_h^{1/h} \quad (5)$$

This follows from the observation that if one takes any two (fixed) polygons, one of size h_1 and one of size h_2, one can identify the rightmost, topmost cell of the first polygon, and join it to the leftmost, bottommost cell of the second (so that they share an edge of a hexagon), as shown in Figure 3. In this way one produces a *unique* polygon of size $h_1 + h_2$. However, not all polygons of size $h_1 + h_2$ can be produced in this way. Thus we arrive at the supermultiplicative inequality

$$B_{h_1}B_{h_2} \leq B_{h_1+h_2} \text{ for } h_1, h_2 \geq 1 \quad (6)$$

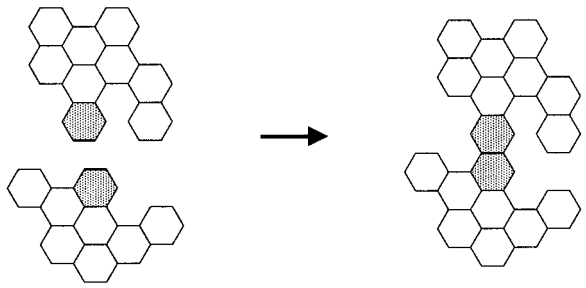

Figure 3. The concatenation of two polygons.

Taking the logarithm and multiplying by -1 yields a subadditive inequality. The final requirement to complete the proof is to show that the sequence $\{B_h^{1/h}\}$ is bounded above. A crude bound, $B_h^{1/h} \leq 2^5$ follows mutatis mutandis from the corresponding argument for square lattice bond animals given in ref 7. We give more details of a much tighter bound below.

Note that (5) is *not* equivalent to the stronger result that

$$\lim_{h\to\infty} B_h/B_{h-1} = \kappa \quad (7)$$

That result is much more difficult to prove. In ref 8, Kesten considered, among other problems, the number p_{2h} of fixed polygons *enumerated by perimeter*. For this very different problem, Kesten proved that

$$\lim_{h\to\infty} p_{2h}/p_{2h-2} = \mu^2 \quad (8)$$

where μ is the analogous growth constant to κ when considering enumeration by perimeter, rather than by area. Unfortunately, Kesten's proof does not translate to the case of polygons enumerated by area, as considered here. It may well be possible to do so, but we have not as yet succeeded.

An alternative approach to proving this result is suggested by the recent work of Madras,[9] who has proved a pattern theorem for lattice animals. It requires three conditions to be satisfied, two of which, translational invariance and a property on weights, are trivially satisfied. Unfortunately the third, which requires certain pattern translates to hold, has not been proved for polygons enumerated by area or perimeter, though it may be possible to do so. The interested reader is referred to ref 9 for details. Thus while we are confident that (7) is true, we cannot as yet provide a proof.

It is widely believed, for benzenoids as well as for a very large class of related lattice objects, including lattice animals, lattice trees, self-avoiding walks, percolation clusters, and self-avoiding polygons, to name but a few, that the asymptotic behavior is in fact

Conjecture 1:

$$b_h \sim A\kappa^h h^\theta \text{ as } h \to \infty$$

or, equivalently

$$B_h \sim 12A\kappa^h h^\theta \text{ as } h \to \infty$$

Note that Theorem 1 implies that $b_h \sim A\kappa^h$, so this conjecture says something about the subdominant asymptotic behavior, which is contained in the term h^θ. For most two-dimensional systems the existence of the subdominant term h^θ has not

been proved, but for many problems, especially those that are conformally invariant,[10] not only is it widely (indeed, universally) believed to be true, but it is equally widely held that the exact value of θ is known. Furthermore θ, unlike κ, displays *universality*. That is to say, κ changes as one changes from the hexagonal lattice to say, the square or triangular lattice, but θ is expected to remain constant.

An additional universal property of such systems, is the exponent characterizing the *linear size* or an equivalent metric property, such as the *mean span*. This is just the average width (or, equivalently, height) of objects of size h (polygons, in this case), averaged over all the objects of size h. We expect the mean span $\langle MS\rangle_h$ to be asymptotically proportional to h^ν, where ν is another universal critical exponent. An equivalent metric is the *mean-square radius of gyration*, $\langle R_g^2\rangle_h$, of benzenoids of size h, which is expected to grow like $h^{2\nu}$. In ref 11 there is a proof of an exponent inequality relating ν and θ, viz.

Theorem 2: $\nu \leq -\theta$, which, with minor modification, holds for (fixed) polygons. The proof, which we outline below, of course, depends on the existence of the exponent ν and the validity of conjecture 1. In this paper we do not estimate ν, but on the grounds of universality we expect it to be the same as for square-lattice polygons, which we have previously estimated[12] to be $\nu = 0.64115(5)$.

Before giving the proof, we remark that for models such as benzenoids, there is a widely accepted argument[13] that θ for a d-dimensional system is given by the Yang-Lee edge singularity exponent in $d-2$ dimensions. As this exponent is known for $d = 0, 1$ it follows that $\theta = -1, -3/2$ for $d = 2, 3$, respectively. As our analysis below shows, this prediction is well borne out by the numerical data.

Indeed, if we assume the existence of θ, we can slightly modify a proof of Madras[11] and prove the following:

Theorem 3: $\theta \leq -1/2$. The proof requires the following lemma:

Lemma 1. *Let κ be a positive number and let $p_1, p_2, \ldots$, be a positive sequence such that $\lim_{h\to\infty} p_h^{1/h} = \kappa$. Also assume that there are numbers $C > 0$ and $s \geq 0$ and integer $k \geq 0$, such that*

$$p_{2h+k} \geq Ch^s p_h^2 \text{ for every } h \geq 1$$

Then

$$p_h \leq \frac{1}{C2^s h^s}\kappa^{h+k} \text{ for every } h \geq 1$$

This lemma is given and proved in ref 11. First, observe that any benzenoid of h-cells has, by definition, area h, if we take a cell as the unit of area. As a space-filling object, it follows that one linear dimension must be at least $\sqrt{h}$. Orient two benzenoid systems, labeled l (for *left*) and r (for *right*), both of size h, so that their vertical dimension is at least $\sqrt{h}$. (This can be done by a rotation, if necessary.) We will consider joining two polygons together to form a polygon system of size $2h+1$, in a manner similar to that discussed and illustrated above in the proof of theorem 1. Move r to the right of l and translate it vertically so that a horizontal line exists, which, when extended to the left, intersects l and r. There are at least $2\sqrt{h}$ choices for this translation. Now translate r to the left until it is just one cell to the right of l. Insert a cell (shown shaded in Figure 4), which joins l and r. This produces a $2h+1$ cell system, which may be produced more than once (because of the permitted rotations). Each polygon may be rotated through $2\pi/3$, allowing up to three possible rotations. Thus up to $3^2 = 9$ possible realizations of this polygon can occur. However not all such $2h+1$ cell systems can be produced in this way. Thus we obtain $9B_{2h+1} \geq 2\sqrt{h}B_h^2$, where the factor $2\sqrt{h}$ arises from the number of possible overlaps, and the factor of 9 arises from the rotational possibilities just discussed.

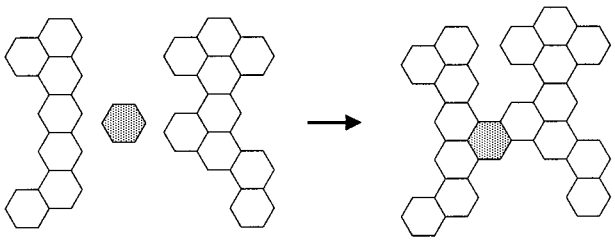

Figure 4. One of the concatenations needed in the proof of $\theta \leq -1/2$.

Then from the above lemma and theorem 1, it follows that

$$B_h \leq \frac{9\kappa^{h+1}}{2\sqrt{2h}}$$

Thus $\theta \leq -1/2$ as claimed.

The proof that $\nu \leq -\theta$ derives from certain properties of the *projection* of the polygon onto one of the lattice axes. Roughly speaking, the x-projection is the width of the polygon (being its projection onto the x-axis), and the y-projection is its height. For *each* polygon of area h there exists a constant K such that

$$\frac{1}{2}(x\text{-projection} + y\text{-projection}) \geq K \times \text{mean span}$$

Then from the above concatenation argument leading to the proof of the bound on θ, we arrive at

$$B_{2h} \geq 2K B_h^2 \langle MS\rangle_h$$

Then from the scaling assumption $\langle MS\rangle_h \sim B_h{}^\nu$ as $h \to \infty$ (where B is a constant, having no connection with B_h) we obtain

$$B_{2h} \geq 2K BB_h^2 h^\nu$$

From the same lemma used above, we arrive at $B_h \leq \kappa^h/2BKh^\nu$. The assumption $B_h \sim \kappa^h h^\theta$, when substituted into the preceding expression, then yields the required inequality.

3. FINITE LATTICE METHOD ALGORITHM

The method used to enumerate polygons on the hexagonal lattice by area is a generalization of the method devised by Enting[6] in his pioneering work on square lattice polygons and the subsequent extension of this approach to hexagonal lattice polygons.[1] We also include the significant enhancements employed in our previous work[14] on square lattice polygons. While all of these papers are concerned with

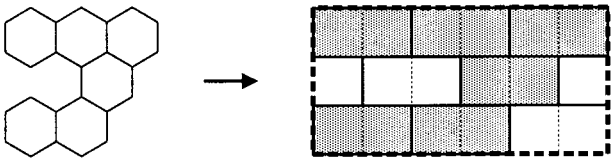

Figure 5. A benzenoid system is shown with the corresponding embedding in the brick-work lattice and the bounding rectangle $Z^1_{6,3}$.

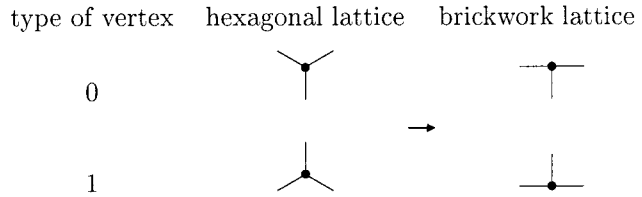

Figure 6. The two types of vertices that appear on the brick-work lattice.

enumeration by *perimeter*, the general method, described in detail in these papers, is the same when enumerating by *area*, and for this reason we shall be brief and only give the absolutely essential information.

As pointed out by Enting,[15] there are three conditions that have to be fulfilled to make the FLM a successful technique for a specific problem:

- An underlying connected graph expansion must exist.
- Weights for combining contributions from finite lattices have to be known or calculated.
- Efficient ways of constructing the finite lattice sums must exist: in practice, this means transfer matrix (TM) techniques.

In the following we shall outline how these conditions are realized in the calculation for (fixed) polygons on the hexagonal lattice by area.

We embed the hexagonal lattice in the square lattice as the brick-work lattice, as shown in Figure 5. There are two types of vertices on the hexagonal lattice, type 0 and 1, as shown in Figure 6 with the corresponding vertices on the brick-work lattice.

In the case of polygons, the series expansion *is* the series of polygons enumerated by area on the hexagonal lattice itself. Since polygons are connected, this is, trivially, the required connected graph expansion. This addresses the first condition of the FLM.

Next we address the second condition of the FLM. Consider the rectangles $Z^i_{m,n}$ with length m and height n where the bottom left-hand corner is a vertex of type $i \in \{0, 1\}$. Let $p^i_{m,n}(x) = \sum_h p^i_{m,n,h}x^h$ be the generating function for polygons that fit in the rectangle in such a way that they touch all sides of the finite rectangle. $p^i_{m,n,h}$ is the number of polygons of area h that fit in the rectangle $Z^i_{m,n}$, touching all sides. $Z^i_{m,n}$ is called the bounding rectangle of such a polygon. The bounding rectangle is unique. As a result $p(x)$, the generating function for fixed polygons can be expressed as a sum over the generating functions of polygons on the finite lattices, $p(x) = \sum_{m,n,i,h} p^i_{m,n,h}x^h = \sum_{m,n,i} p^i_{m,n}(x)$. The $p^i_{m,n}(x)$ are polynomials. The minimal degree of $p^i_{m,n}(x)$ increases with m and n. It is at least $n + \max\{0, \lceil m - n - 1)/2 \rceil\}$. If one wants to calculate $p(x)$ up to order $h_{\max}$, only the $p^i_{m,n}(x)$ which have a minimal degree of $h_{\max}$ or less have to be calculated. Hence the generating function can be written as

$$p(x) = \sum_{\substack{1<n<h_{\max} \\ 1\le m\le 2h_{\max}-n+1 \\ i\in\{0,1\}}} p^i_{m,n}(x) + O(x^{h_{\max}+1}) \qquad (9)$$

For this we have to consider the finite lattices $Z^i_{m,n}$ with $1 \le n \le h_{\max}$, $1 \le m \le 2h_{\max} - n + 1$, and $i \in \{0, 1\}$. Hence the weights are one for these lattices and zero otherwise.

To address the third point of Enting's list of conditions, we briefly discuss the TM method which we used to calculate the generating functions on the finite lattices. In ref 6 Enting enumerated polygons on the square lattice by perimeter and outlined the principles of the TM method for enumerating polygons on a lattice in general. Enting and Guttmann[1] used the FLM/TM technique to enumerate polygons on the hexagonal lattice by perimeter. The first terms in the series for the generating function $B(x) = \sum B_h x^h$ can be calculated using TM techniques to count the number of polygons in rectangles of height m and length n. The TM technique involves drawing a boundary line through the rectangle intersecting a set of $m + 1$ edges. Each edge is either empty or occupied by a part of the perimeter of a partially completed polygon. Since the perimeter of a polygon is a self-avoiding closed path, each occupied edge is connected, via a path to the left of the boundary line, to exactly one other occupied edge. We choose to encode the state σ_i of an edge as $\sigma_i = 0$ if the edge at position i is empty, $\sigma_i = 1$ if the edge is occupied *and* is the lower end of a connected path, and $\sigma_i = 2$ if the edge is the upper end of a path. Due to the self-avoidance and the two-dimensionality of the problem this encoding uniquely specifies the connectivity of the edges.

For each configuration of occupied or empty edges along the boundary, we maintain a generating function for partially completed polygons. The generating function is a (truncated) polynomial $p_s(x)$, where $s = \{\sigma_i\}$ is the state vector specifying the configuration. Polygons in a given rectangle are enumerated by moving the boundary so as to add one unit cell at a time. When the boundary line is moved, a given state vector s is transformed into two new state vectors s_1 and s_2 and $x^{k_1}p_s(x)$ is added to $p_{s_1}(x)$ and $x^{k_2}p_s(x)$ is added to $p_{s_2}(x)$, where k_1 and k_2 count the additional edges (if enumerating by perimeter) or unit cells (if enumerating by area) added to the polygon. In the case of enumeration by area, k_1 and k_2 depend on the state vector, that is on whether the added cell is part of the polygon or not. It is quite simple to determine whether a newly added unit cell of the hexagonal lattice belongs to a polygon or not. Moving through a configuration we note that as we reach the first occupied edge we pass from the outside to the inside of a polygon, the next occupied edge takes us to the outside again, and so on. In this fashion all unit cells intersected by the boundary line are uniquely assigned to the interior or exterior of a polygon.

The rules for updating the partial generating functions are described in ref 1 in the case of enumeration by perimeter. The generalization to enumeration by area is quite simple since the encoding and transformations of the configurations are identical. The only change is that the weights assigned to a configuration count the area that has been already inserted. Furthermore, in the TM algorithm we have implemented, we move the intersection line over two vertices in one step instead over one vertex as in ref 1. The updating rules for this case are given in Table 1. At the upper and

Table 1. Update Rules[a]

input	$output_{inside}$		$output_{outside}$	
'00'	*x* '00',	'12'	'00',	*x* '12'
'01'	*x* '01',	'10'	'01',	*x* '10'
'02'	*x* '02',	'20'	'02',	*x* '20'
'10'	*x* '10',	'01'	'10',	*x* '01'
'11'	*x* '11',	'00'*	'11',	*x* '00'*
'12'	*x* '12',	$	'12'	
'20'	*x* '20',	'02'	'20',	*x* '02'
'21'	*x* '21',	'00'	'21',	*x* '00'
'22'	*x* '22',	'00'*	'22',	*x* '00'*

[a] A cell between the two input bonds can be part of the polygon, denoted $output_{inside}$, or not, denoted $output_{outside}$. When the state '00'* is produced, the ends of the loops that have been closed have to be relabeled appropriately. "$" indicates that the polygon can be closed and added to the total count if all other bonds on the intersection line are empty.

lower boundary of the finite lattice the TMs are appropriately modified. The boundary line is drawn either vertically or horizontally through the rectangle. One follows the orientation in which the line intersects fewer edges. In the case of a horizontal boundary the TMs are appropriately modified at both the left and right boundary of the finite lattice.

If one wants to calculate the coefficients B_h up to h_{max}, the maximal number of edges which the boundary line intersects in any of the required finite lattices grows as $2h_{max}/3$.

Every finite lattice falls into one of two cases, $m \geq 2n$ or $m < 2n$. In the first case a vertical boundary is used. In the second case, parenthesized below, a horizontal boundary is used. It is not necessary to calculate every $p^i_{m,n}(x)$ separately. One can consider all the finite lattices $Z^i_{m,n}$ with n fixed, i fixed, and $m \geq 2n$ (m fixed, i fixed, and $n > m/2$) together. That is to say, we can calculate the $p^i_{m,n}(x)$ of all these finite lattices in one sweep by aligning the left (lower) sides of the finite lattices. The boundary is moved from left to right (bottom to top). When a polygon is closed it counts toward the $p^i_{m,n}(x)$ which is determined by its bounding rectangle.

In ref 14 improvements of the FLM/TM were discussed which we have applied to the enumeration of polygons on the hexagonal lattice. The major improvement comes from the calculation of the polygons that span the finite lattices in length and width and not only in length as in previous work. This increases the number of state vectors required to describe all configurations by a factor of almost four, since one has to store a parameter that keeps track of whether the polygon has touched the lower boundary of the rectangle and whether it has touched the upper boundary. This enables one to calculate how many cells are needed to close the polygon (or find a lower bound) such that the resultant polygon touches both the upper and the lower boundaries and stretches to the right boundary and finally that the resulting configuration is connected. If in addition we store the minimal area to the left of the intersection line we can calculate the minimal area (or a lower bound on the area) which every polygon with the particular configuration on the intersection will have. If this is larger than h_{max}, the configuration is discarded. This leads to an exponential reduction in the number of configurations that have to be stored. In the original approach all possible configurations were kept.

From the foregoing discussion of the encoding of the state vectors, it is clear that since every occupied edge is uniquely matched, any configuration is just an example of perfectly matched parentheses with gaps. These are well-known in the combinatorics literature and are often called Motzkin paths/words. For our purpose all we need to know is that the number of Motzkin paths of length m denoted M_m behaves for large m as $M_m \sim \text{const.}3^m$. This exponential growth obviously determines the computational complexity of the original approach. The maximal number of bonds intersected by the boundary line grows as $2h_{max}/3$. This implies that the complexity of enumerating benzenoids of size h grows as $3^{2h/3} \approx 2.08^h$, multiplied by some polynomial in h. Thus this approach already provides a dramatic improvement over preexisting direct enumeration algorithms, which have complexity 5.16^h. With the further improvements we have described, it is not possible to give a theoretical analysis of the computational complexity of the improved algorithm, but an empirical analysis suggests that the improvements reduce the complexity to λ^h with $\lambda \approx 1.65$.

In this way we have obtained the coefficients B_h for $h \leq 35$. The calculation took about 10 weeks on a Compaq AlphaServer ES40 and utilized up to 5GB of memory. To minimize memory requirements, all calculations were done using 16 bit integers, *modulo* a prime. This was repeated five times, using five different primes, and the final result reconstructed using the Chinese remainder theorem. Using 32 bit integers would have required only three runs but would have doubled the memory requirements. With more memory still, only one run would be needed, and thus the time taken would be approximately 2 weeks. To calculate the number of benzenoids b_h, we used eq 3 and the counts of those polygons which possess some symmetry. The counts of the latter are given in Table 3. This then leads immediately to the required counts of benzenoids, b_h for $h \leq 35$. The results are shown in Table 2.

4. SERIES ANALYSIS

From the coefficients B_h and b_h given in Table 2, we have the first 35 terms in their respective generating functions, defined by $H(z) = \sum_{h\geq 1} b_h z^h$ and $G(z) = \sum_{h\geq 1} B_h z^h$. Given the expected asymptotic behavior $B_h \sim C\kappa^h h^\theta$, it follows that the expected generating function behavior is $G(z) = \sum_h B_h z^h \sim A(z)(1 - \kappa z)^{-\theta-1}$, with a similar expression for $H(z)$, differing only in the amplitude term $A(z)$, which will be asymptotically equal to $A(z)/12$. Here $C = A(1/\kappa)/\Gamma(\theta + 1)$, and the radius of convergence of the generating function is given by $1/\kappa$. As we alluded to above, we find below that $\theta = -1$, so that the exponent $-\theta - 1 = 0$. This corresponds to a logarithmic singularity, so that in fact

$$G(z) = \sum_h B_h z^h \sim A(z) \log(1 - \kappa z) \text{ as } \kappa \rightarrow 1/z_- \quad (10)$$

We have used two methods to analyze the series studied in this paper. First, to obtain the singularity structure of the generating function we used the numerical method of differential approximants.[16] Very briefly, in this method we approximate the generating function by the solution of a linear, inhomogeneous, ordinary differential equation (ode) with polynomial coefficients. That is to say, we insist that the power series expansion of the solution of the ode agrees, order by order, with the known coefficients of the generating function. One can increase the degree of the polynomials and the order of the underlying differential equation until

Table 2. Number, b_h, of Free Hexagonal Polygons, or Benzenoids, and the Number B_h of Fixed Hexagonal Polygons, Both of Area h Cells

h	b_h	B_h
1	1	1
2	1	3
3	3	11
4	7	44
5	22	186
6	81	813
7	331	3640
8	1435	16590
9	6505	76663
10	30086	358195
11	141229	1688784
12	669584	8022273
13	3198256	38351973
14	15367577	184353219
15	74207910	890371070
16	359863778	4318095442
17	1751594643	21018564402
18	8553649747	102642526470
19	41892642772	502709028125
20	205714411986	2468566918644
21	1012565172403	12150769362815
22	4994807695197	59937663454017
23	24687124900540	296245438278258
24	122238208783203	1466858366128911
25	606269126076178	7275229222292218
26	3011552839015720	36138633393334038
27	14980723113884739	179768675964165939
28	74618806326026588	895425672624735867
29	372132473810066270	4465589678921947602
30	1857997219686165624	22295966620155816954
31	9286641168851598974	111439693993112940196
32	46463218416521777176	557558620919353655115
33	232686119925419595108	2792233438943251452902
34	1166321030843201656301	13995852369729891369431
35	5851000265625801806530	70212003186716473817832

Table 3. Number of Free Hexagonal Polygons of Different Symmetry Classes

h	D_{6h}	C_{6h}	D_{3h}	C_{3h}	D_{2h}	C_{2h}	C_{2c}
1	1	0	0	0	0	0	0
2	0	0	0	0	1	0	0
3	0	0	1	0	1	0	1
4	0	0	1	0	2	1	1
5	0	0	0	0	2	1	9
6	0	0	1	1	3	7	12
7	1	0	1	1	3	7	39
8	0	0	0	0	6	35	61
9	0	0	1	5	7	36	178
10	0	0	4	5	11	169	274
11	0	0	0	0	14	177	796
12	0	0	3	21	21	807	1251
13	2	0	4	26	23	859	3578
14	0	0	0	0	41	3864	5692
15	0	0	3	95	50	4145	16290
16	0	0	12	118	80	18616	26069
17	0	0	0	0	94	20098	74980
18	0	0	6	423	156	90265	120676
19	2	2	19	543	189	97913	348564
20	0	0	0	0	310	440230	563503
21	0	0	10	1923	365	479367	1635175
22	0	0	41	2507	615	2157946	2652445
23	0	0	0	0	748	2357108	7729807
24	0	0	16	8869	1237	10623852	12574028
25	3	8	72	11676	1459	11634624	36784467
26	0	0	0	0	2488	52496803	59975232
27	0	0	28	41410	3023	57618226	176063801
28	0	0	149	54893	5059	260236031	287613473
29	0	0	0	0	5994	286164543	846967739
30	0	0	50	195353	10296	1293603599	1385815492
31	5	32	272	260220	12511	1424824820	4092516853
32	0	0	0	0	21149	6445885963	6705360813
33	0	0	87	929423	25100	7109939186	19852731854
34	0	0	557	1242820	43451	32187437996	32565589836
35	0	0	0	0	52870	35548330099	96642900684

there are no more known coefficients. In practice, it has been found that a first- or second-order ode is usually sufficient to approximate the singularity structure found in problems such as this. One then solves the ode in the standard manner, the critical point being given by the closest zero on the positive real axis of the polynomial multiplying the highest derivative, while the corresponding exponent is obtained from the solution of the appropriate *indicial equation*.[17] A substantial number of such *differential approximants* are constructed, and a statistical procedure is used to estimate the critical point and critical exponent.[16]

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. These are the solutions of the differential equations $zP_M^{(1)}(z)f'(z) + P_N^{(0)}(z)f(z) = Q_L(z)$ and $z^2P_K^{(2)}f''(z) + zP_M^{(1)}(z)f'(z) + P_N^{(0)}(z)f(z) = Q_L(z)$, respectively, where P and Q are polynomials of degree given by their subscripts.

In particular, we used this method to estimate the growth constant κ and the critical exponent θ. As mentioned above, there is a prediction[13] that $\theta = -1$, which we also find numerically. Imposing this conjectured exponent permitted a refinement of the estimate of the growth constant—providing so-called biased estimates.

In Table 4 we have listed estimates for the critical point $1/\kappa$ and exponent $-1 - \theta$ obtained from the first 35 terms of the generating function $G(z)$ for fixed benzenoids. (We chose to analyze this series, rather than $H(z)$, as it appeared to converge marginally faster.) Truncating the series at $O(z^{N_{max}})$ we averaged over those approximants to the series which used exactly N_{max} terms of the series, using approximants such that the difference between N, M, and K did not exceed 2. 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 favorable situations, as for example, in the case of self-avoiding polygons enumerated by perimeter.[14] Based on these estimates we conclude that $1/\kappa = 0.193725984(5)$ and $1 + \theta = 0.00000(1)$.

As stated earlier there is very convincing evidence that the critical exponent $\theta = -1$ exactly. This is certainly borne out by our numerical estimate. If we assume this to be true we can obtain a slightly refined estimate for the critical point $1/\kappa$. In Figure 7 we have plotted estimates for the critical exponent $1 + \theta$ against estimates of the critical point $1/\kappa$. Each dot in this figure represents a pair of estimates obtained from a second-order inhomogeneous differential approximant. The order of the inhomogeneous polynomial was varied from 1 to 10. We observe that there is an almost linear relationship between the estimates for $1 + \theta$ and $1/\kappa$ and that for $\theta = -1$ we get $1/\kappa = 0.193725984(3)$ and thus $\kappa = 5.16193016(8)$.

Table 4. Estimates for the Critical Point $1/\kappa$ and Exponent $-1-\theta$ Obtained from First- and Second-Order Inhomogeneous Differential Approximants to the Series for Fixed Polygons B_h[a]

	first-order DA				second-order DA		
N	$1/\kappa$	$-1-\theta$	L	N	$1/\kappa$	$-1-\theta$	L
20	0.193727456(2235)	−0.000599903(879737)	12	20	0.193725721(707	0.000039007(664832)	2
21	0.193726012(847)	−0.000008211(310700)	12	21	0.193725595(3381)	0.000252880(1551463)	4
22	0.193726087(591)	−0.000046729(389321)	10	22	0.193725555(1633)	0.000239905(872254)	4
23	0.193726088(349)	−0.000060360(183821)	12	23	0.193725262(2085)	0.000412881(1159676)	6
24	0.193725788(1373)	0.000104578(758523)	10	24	0.193725939(81)	0.000020898(45224)	7
25	0.193726034(210)	−0.000031553(126857)	10	25	0.193726026(186)	−0.000026792(117701)	5
26	0.193725972(32)	0.000005084(18734)	9	26	0.193725946(64)	0.000023149(49213)	7
27	0.193725982(79)	−0.000000515(54029)	11	27	0.193726008(89)	−0.000018796(64754)	7
28	0.193725990(36)	−0.000006093(25504)	12	28	0.193726015(30)	−0.000026486(28307)	8
29	0.193725992(19)	−0.000007571(15050)	11	29	0.193725994(38)	−0.000009511(29687)	8
30	0.193725988(14)	−0.000004837(12036)	12	30	0.193725988(6)	−0.000004877(5070)	8
31	0.193725989(6)	−0.000005704(5115)	12	31	0.193725992(4)	−0.000007882(3278)	8
32	0.193725987(11)	−0.000003936(10263)	12	32	0.193725989(13)	−0.000005139(9471)	7
33	0.193725985(8)	−0.000002463(7282)	12	33	0.193725984(11)	0.000000671(20669)	7
34	0.193725987(2)	−0.000004259(1720)	10	34	0.193725987(2)	−0.000003343(2085)	8
35	0.193725986(1)	−0.000003014(1159)	9	35	0.193725986(2)	−0.000002860(1873)	8

[a] N terms were used in the approximants, and L is the number of approximants averaged to give the quoted estimates. One standard deviation is shown in parentheses.

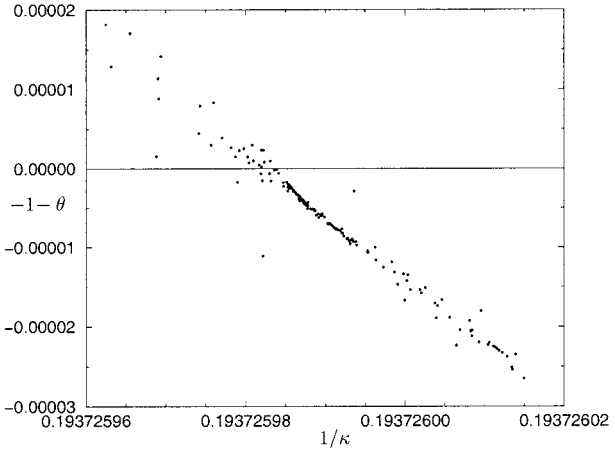

Figure 7. Plot of $-1-\theta$ against $1/\kappa$ as obtained from second-order differential approximants. The linear relationship is clear.

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 $B_h = [z^h]G(z) \approx \kappa^h h^{-1} \sum_{i\geq 0} a_i/h^{f(i)}$. In the most favorable circumstances, if there is no nonanalytic correction term, then $f(i) = i$. In some problems there is a square-root correction term which means $f(i) = i/2$, while a logarithmic correction implies more subtle behavior. In all cases, our procedure is to *assume* a particular form for $f(i)$ and observe how well it fits the data. With the long series we now have at our disposal, it is usually easy to see if the wrong assumption has been made, as if so the sequence of amplitude estimates a_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 refs 14 and 18.

For benzenoids it appears that there are only analytic corrections (that is to say, the expansion of $A(z)$ in the equation for the generating function above has no nonanalytic terms). This means that $f(i) = i$, and is the simplest situation, implying the absence of nonanalytic correction terms.

Table 5. A Fit to the Asymptotic Form of the Coefficients for Fixed Polygons $B_h/\kappa^h \sim h^{-1}[a_1 + a_2/h + a_3/h^2 + a_4/h^3 + a_5/h^4 + \cdots]$[a]

h	a_1	a_2	a_3	a_4	a_5
18	0.2808477	−0.144999	0.029385	−0.01085	0.53583
19	0.2808512	−0.145225	0.034974	−0.07214	0.78748
20	0.2808500	−0.145141	0.032774	−0.04654	0.67593
21	0.2808497	−0.145119	0.032163	−0.03902	0.64132
22	0.2808496	−0.145116	0.032066	−0.03777	0.63521
23	0.2808494	−0.145096	0.031465	−0.02957	0.59334
24	0.2808493	−0.145092	0.031315	−0.02742	0.58177
25	0.2808492	−0.145084	0.031059	−0.02359	0.56032
26	0.2808492	−0.145080	0.030923	−0.02147	0.54787
27	0.2808492	−0.145077	0.030807	−0.01958	0.53631
28	0.2808492	−0.145075	0.030725	−0.01818	0.52740
29	0.2808491	−0.145073	0.030665	−0.01712	0.52041
30	0.2808491	−0.145072	0.030618	−0.01627	0.51455
31	0.2808491	−0.145071	0.030583	−0.01559	0.50976
32	0.2808491	−0.145071	0.030553	−0.01501	0.50548
33	0.2808491	−0.145070	0.030528	−0.01449	0.50148
34	0.2808491	−0.145070	0.030503	−0.01397	0.49740
35	0.2808491	−0.145069	0.030478	−0.01343	0.49302

[a] Estimates of the amplitudes a_1, a_2, a_3, a_4, a_5 are given.

We have argued previously[12] that there is no nonanalytic correction-to-scaling exponent for the square-lattice polygon area generating function. We would expect this to be a universal property, applying to all regular two-dimensional lattices, so it is not surprising that we find the same result here. We conjecture that the asymptotic form for the polygon generating function is as given by eq 10. In evidence of this, we remark that from (10) follows the asymptotic form

$$B_h = \kappa^h h^{-1}[a_1 + a_2/h + a_3/h^2 + a_4/h^3 + a_5/h^4 + \cdots + \mathrm{O}(\exp(-h))] \quad (11)$$

Using the central value of our estimate $\kappa = 5.16193016$ quoted above, we show in Table 5 the estimates of the amplitudes a_1, ..., a_5. From the table we estimate that $a_1 \approx 0.2808491$, $a_2 \approx -0.145065$, $a_3 \approx 0.0304$, $a_4 \approx -0.01$, and $a_5 \approx 0.4$, where in all cases we expect the error to be confined to the last quoted digit. The excellent convergence of all columns is strong evidence that the assumed asymptotic

form is correct. If we were missing a term corresponding to, say, a half-integer correction, the fit would be far worse. This is explained at greater length in ref 18. So good is the fit to the data that if we take the last entry in the table, corresponding to $h = 35$, and use the entries as the amplitudes, then all the coefficients beyond the first are given either exactly (if rounded to the nearest integer), or to the same accuracy as the leading amplitude.

This analysis was all done on the generating function for *fixed* polygons. For benzenoids, or free polygons, the estimates of θ and κ are, as shown above, identical to the corresponding values for fixed polygons, while the amplitudes just need to be divided by 12. Thus

$$b_h = \kappa^h h^{-1}[d_1 + d_2/h + d_3/h^2 + d_4/h^3 + d_5/h^4 + \cdots + \mathrm{O}(\exp(-h))] \quad (12)$$

where $d_i = a_i/12$. The differences between $B_h/12$ and b_h are confined to terms $\mathrm{O}(\exp(-h))$.

5. LOWER BOUND ON κ

A weak lower bound is immediately attainable from eq 5. Using B_{35} we obtain a lower bound of 4.49526.

This bound can be improved using the techniques developed by Rands and Welsh in ref 19. Using concatenation arguments they showed that if we define a sequence p_n such that

$$B_{h+1} = p_{h+1} + p_h B_1 + \cdots + p_2 B_{h-1} + p_1 B_h \quad (13)$$

and construct the generating functions

$$\mathscr{A}(u) = 1 + \sum_{h=1}^{\infty} B_h u^h \quad (14)$$

and

$$\mathscr{P}(u) = \sum_{h=1}^{\infty} p_h u^h \quad (15)$$

then

$$\mathscr{A}(u) = 1 + \mathscr{A}(u)\mathscr{P}(u) \quad (16)$$

and $\mathscr{A}(u)$ is singular when $\mathscr{P}(u) = 1$. The coefficients in $\mathscr{P}(u)$ are obviously constructible to the same order as known for $\mathscr{A}(u)$. If we look at the polynomial P_N obtained by truncating $\mathscr{P}(u)$ at order N, then the unique positive zero, $1/\kappa_N$, of $P_N - 1 = 0$ is a lower bound for κ.

In ref 19 the method was applied to the generating function of *polyominoes* on the hexagonal lattice. Polyominoes are defined to be fixed polygons which are allowed to have internal holes. They are related to coronoids, which will be the subject of a subsequent paper. The application of the method to our *polygon* series up to area 35 leads to a lower bound of 4.7892. It seems likely that more elaborate concatenation schemes, such as that described in ref 7 for square lattice polyominoes, could improve this bound. We discuss this further in ref 20.

Improvements suggested by Rands and Welsh include concatenation in which cells overlap. This construction is not appropriate for polygons, though it is for polyominoes.

6. UPPER BOUND ON κ

A crude upper bound is $B_h \le 2^{4h}$. A polygon of area h is bounded by a self-avoiding walk (SAW) of length at most $2 + 4h$. The first and last steps can be chosen to be fixed. As the coordination number of the hexagonal lattice is 3, the SAW has at most two choices at every step. This immediately gives the above bound.

Obtaining a good upper bound to κ is an involved computation and will be the subject of a separate paper.[20] In this article we merely outline the method and give the result.

W. L. Lunnon[21] found an upper bound for the polyomino growth constant of 6.75. Since polygons are a subset of the polyominoes this gives an immediate upper bound to κ, as using a pattern theorem argument due to Madras,[9] it is easy to prove that the growth constant for polyominoes is strictly greater than that for polygons.

To improve this bound we apply the ideas of Klarner and Rivest[22] which were inspired by Eden.[23] Klarner and Rivest developed a method which involves successive improvement. The idea is that each polygon is converted to a tree on the dual lattice, in this case the triangular lattice. We describe a mapping that associates a unique spanning tree with each polygon. Then, by relaxing the rules for the construction of such a tree, we end up with an algorithm that overcounts the number of polygons. Unlike the generating function for the original problem, we can determine the radius of convergence of the "overcounted" polygons, which therefore provides an upper bound. The algorithm has the advantage of being amenable to systematic improvement.

The mapping from polygon to tree is done by placing a site at the center of each hexagon and joining certain sites if their associated hexagons share an edge. This is done in such a way that cycles are excluded, and all cells included. Thus each polygon is associated with a spanning tree, made up of so-called "twigs" (defined below), chosen from a fixed, finite set. The number of spanning trees (and hence polygons) is bounded above by the number of ways of concatenating the twigs. The cells in a twig are divided into two types, namely *dead* and *living*. Additionally, *forbidden* cells may be associated with a twig, though they are not part of the twig, as shown in Figure 8. Every twig contains at least one dead cell but not necessarily any living cells. Further one of the dead cells is marked by an incoming edge as a root cell. Figure 8 shows the set of basic twigs we use to construct both polygons and larger twigs. Any polygon can be constructed in the following way from the basic twigs. We start with a single living cell that has an incoming bond from below, and we keep a list of living cells. The addition of twigs to the configuration constructed so far proceeds as follows:

- Add a new twig by placing the root cell of the twig over the oldest living cell such that the orientations of both cells coincide.
- The addition is legal if no other cells of the twig overlap with any part of the configuration and no cell of the twig occupies a cell marked as forbidden.

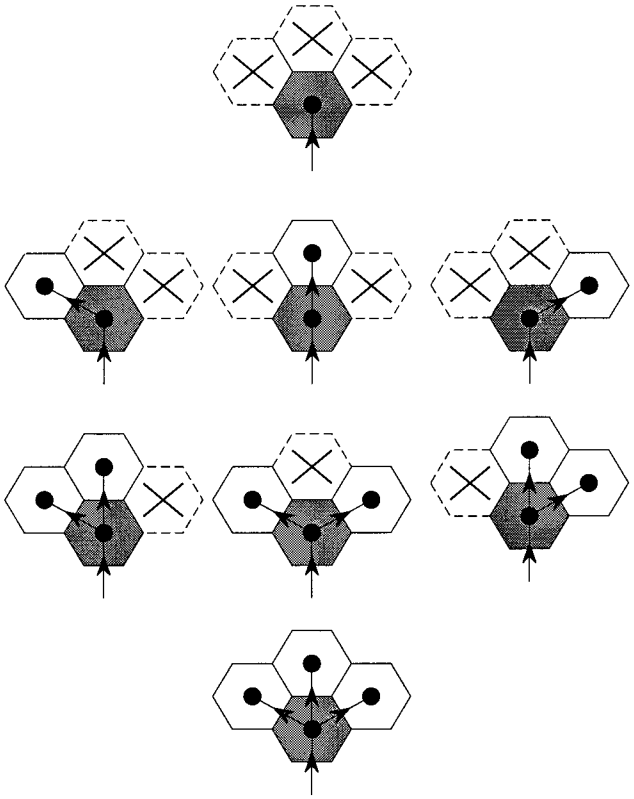

Figure 8. This figure shows the set of basic twigs which we used. The dark cells are dead cells, the white cells are living cells, and the cells marked with a cross are forbidden cells.

•Make the forbidden cells of the twig forbidden cells of the polygon.

•Append the living cells of the twig to the list of living cells.

•Remove the oldest cell from the list of living cells.

•Make the living cell where the twig has been added a dead cell.

This is repeated until no living cells are left. The construction of a polygon of size four is illustrated in Figure 9. In this procedure we did not ensure that the constructed configurations do not contain holes, which is not a problem as we are seeking an upper bound.

We can construct a set of larger twigs (i.e. twigs with more dead cells) from the basic twigs by using almost the same algorithm. A complete set of twigs of size n contains all configurations that can be constructed according to the above rules and that contain exactly n dead cells (and possibly living cells) or contain less than n dead cells and no living cells. One can construct any polygon with such a set of twigs.

The results are systematically improved by increasing the number of dead cells in the set of basic twigs. The calculations become increasingly complex with increasing twig size, requiring exponentially increasing amounts of computer time. The bound we obtain here derives from twigs with 16 dead cells. The C program ran for four weeks on a MacIntosh G3 computer.

From the determination of the twigs we proceeded in the same way as Klarner and Rivest did. We assign to every twig i a weight $w_i = x^{m-1}y^n$, where $m - n$ is the number of live cells of the twig and n is the number of dead cells of the twig.

If we relax the construction rules so that the legality of an addition of a twig is not checked (this being the second bullet point above), we can write down the two-variable generating function, $f(x, y) = x/(1 - \sum_i w_i)$. Note that this relaxation allows multiple occupancy of cells and forbidden cells to be occupied. It is this that is responsible for the overcounting of spanning trees (and hence polygons). We are interested in the diagonal terms $a_{n,n}$ of the series expansion of $f(x, y) = \sum_{m,n} a_{m,n} x^m y^n$. These terms represent the configurations that contain no living cells and in which all polygons are included. Therefore we obtain an upper bound for the polygon growth constant if we can find the growth constant for the $a_{n,n}$. Klarner and Rivest describe in ref 22 how one can obtain the growth constants of the diagonal terms of a rational two variable generating function. This requires a change of variable so that the residue theorem can be applied. The diagonal function $f_D(x) = \sum_n a_{n,n} x^n$ can be written as a sum over residues.

In this way, and after substantial computation using a twig size of 16, we obtain the upper bound 5.905. Thus we find

$$4.7892 < \kappa < 5.905 \qquad (17)$$

7. DISCUSSION AND CONCLUSION

We have proved a number of results for benzenoids, including the existence of a growth constant κ. We have established rigorous upper and lower bounds on κ and a very precise numerical estimate of $\kappa = 5.16193016(8)$. We provide compelling arguments that the generating function $\sum B_h x^h \sim A(x) \log(1 - \kappa x)$ and provide a (numerical) expansion of $A(x)$ around $x = 1/\kappa$. This analysis provides very strong evidence for the absence of any nonanalytic correction terms to the proposed asymptotic form for the generating function, which we argue persuasively is dominated by a logarithmic singularity. It is likely that correction terms are exponentially small. Finally we give an asymptotic representation for the coefficients which we believe to be useful for benzenoids of any size at least up to 100.

We have presented an improved algorithm for the enumeration of self-avoiding polygons enumerated by area, and hence of benzenoids, on the hexagonal lattice. The compu-

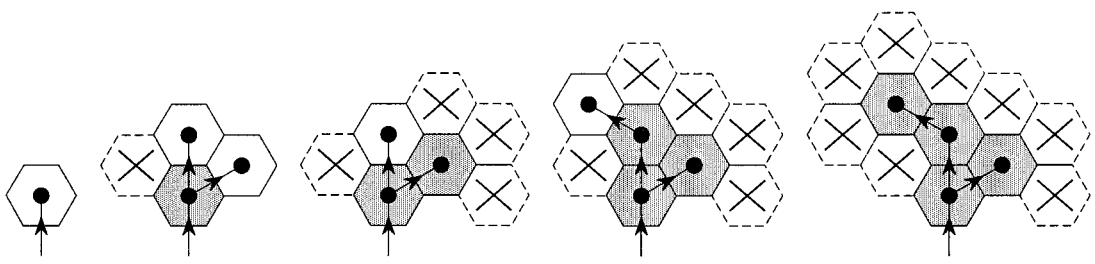

Figure 9. This figure shows an example of a polygon of size four being constructed by the successive addition of twigs.

tational complexity of the algorithm is estimated to be 1.65^h. Implementing this algorithm has enabled us to obtain polygons up to area 35. Decomposing the coefficients into prime factors reveals frequent occurrence of very large prime factors, supporting the widely held view that there is no "simple" formula for the coefficients.

We also discuss the *size* of benzenoids. Many possible measures of size exist, and most of these are equivalent. Accordingly, we focus on the *mean-square radius of gyration*, $\langle R_g^2 \rangle_h$ of benzenoids of size h and provide compelling numerical arguments (though not a proof) that this grows as $h^{2\nu}$, with $\nu = 0.64115(5)$.

A related problem is the enumeration of *coronoid systems.* These are benzenoid systems with one or more internal holes. In the mathematics literature they are referred to as *free polyominoes*. A coronoid system with a single hole is called *a single coronoid system*. The methods described here can also be adapted to the enumeration of coronoid systems, which will be discussed in a subsequent paper.

ACKNOWLEDGMENT

We wish to thank Ivan Gutman for educating us about benzenoids, Neal Madras for helpful discussions (and for providing the tools for our proofs), and Aleks Owczarek and Ivan Gutman for careful comments on the manuscript. A.J.G. and I.J. wish to gratefully acknowledge financial support from the Australian Research Council, while M.V. wishes to thank The University of Melbourne for financial support.

REFERENCES AND NOTES

(1) Enting, I. G.; Guttmann, A. J. Polygons on the honeycomb lattice. *J. Phys. A* **1989**, *22*, 1371-1384.
(2) Gutman, I.; Cyvin, S. J. *Introduction to the Theory of Benzenoid Hydrocarbons*; Springer-Verlag: Berlin, 1989.
(3) Cyvin B. N.; Brunvoll, J.; Cyvin, S. J. Enumeration of Benzenoid Systems and Other Polyhexes. In *Advances in the Theory of Benzenoid Hydrocarbons II*; Gutman, I., Ed.; Springer-Verlag: Berlin, 1992; Vol. 162 .
(4) Caporossi, G.; Hansen, P. Enumeration of polyhex hydrocarbons to h = 21. *J. Chem. Inf. Comput. Science* **1998**, *38*, 610−619.
(5) Brinkmann, G.; Caporossi, G.; Hansen, P. Numbers of benzenoids and fusenes. *Commun. Math. Chem. (MATCH)* **2001**, *43*, 133−134.
(6) Enting, I. G. Generating functions for enumerating self-avoiding rings on the square lattice. *J. Phys. A* **1980**, *13*, 3713−3722.
(7) Whittington, S. G.; Soteros, C. E. Lattice animals: Rigorous results and wild guesses. In *Disorder in Physical Systems*; Grimmett, G. R., Welsh, D. J. A., Eds.; Clarendon, Oxford: 1989.
(8) Kesten, H. On the number of self-avoiding walks. *J. Math. Phys.* **1963**, *4*, 960−969.
(9) Madras, N. A. Pattern theorem for lattice clusters. *Annals Comb.* **1999**, *3*, 357−84.
(10) Henkel M. *Conformal Invariance and Critical Phenomena*; Springer: Berlin, 1999.
(11) Madras, N. A rigorous bound on the critical exponent for the number of lattice trees, animals and polygons. *J. Stat. Phys.* **1995**, *78*, 681−99.
(12) Jensen, I.; Guttmann, A. J. Statistics of lattice animals (polyominoes) and polygons. *J. Phys. A* **2000**, *33*, L257−L263.
(13) Parisi, G.; Sourlas, N. Critical behaviour of branched polymers and the Lee-Yang edge singularity. *Phys. Rev. Lett.* **1981**, *46*, 871−874.
(14) Jensen, I.; Guttmann, A. J. Self-avoiding polygons on the square lattice. *J. Phys. A* **1999**, *32*, 4867−4876.
(15) Enting, I. G. Series Expansions from the Finite Lattice Method. *Nucl. Physics B − Proc. Suppl.* **1996**, *47*, 180−187.
(16) Guttmann, A. J. Analysis of coefficients. In *Phase Transitions and Critical Phenomena*; Domb, C., Lebowitz, J. L., Eds.; Academic Press: New York, 1989; Vol. 13.
(17) Whittaker, E. T.; Watson, G. N. *A course of modern analysis*; The University Press: Cambridge, 1963.
(18) Conway, A. R.; Guttmann, A. J. Square lattice self-avoiding walks and corrections-to-scaling. *Phys. Rev. Lett.* **1996**, *77*, 5284.
(19) Rands, B. M. I.; Welsh, D. J. A. Animals, trees and renewal sequences. *IMA J. Appl. Math.* **1981**, *27*, 1−17.
(20) Vöge, M. Rigorous bounds on growth constants for polygons and polyominoes. Manuscript in preparation.
(21) Lunnon, W. F. Counting hexagonal and triangular polyominoes. In *Graph theory and computing*; Academic Press: New York, 1972; pp 87−100.
(22) Klarner, D. A.; Rivest, R. L. A procedure for improving the upper bound for the number of *n*-ominoes. *Can. J. Math.* **1973**, *25*, 585−602.
(23) Eden, M. A two-dimensional growth process. In *Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability;* University California Press: Berkeley, CA, 1961; Vol. IV; pp 223−239.

CI010098G