# History and introduction to polygon models, polyominoes and polyhedra

Anthony J Guttmann

∗The genesis and history of self-avoiding polygons, polyominoes and polyhedra embedded in a regular space lattice is discussed, as is the related problem of tilings. We establish some basic notation and describe the connection with statistical mechanics. Throughout, our focus is to outline what is known and what remains unknown, and to discuss how the topic of the book is developed in the subsequent chapters.

## 1 Introduction

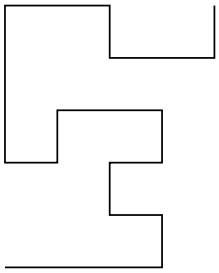
In this book we will primarily be concerned with the properties and applications of self-avoiding polygons (SAP). Two closely related problems are those of polyominoes, and the much broader one of tilings. We will describe and discuss polyominoes, and, in the context of a discussion of SAP, will briefly mention relevant aspects of the subject of tilings. In passing we will also concern ourselves with a discussion of some properties of polyhedra. It will also turn out to be appropriate to discuss self-avoiding walks (SAW), as SAP can usefully be simply related to a proper subset of SAW.

In all cases we shall be considering paths on a regular lattice. This will most often be the two-dimensional square lattice $\mathbb{Z}^2$, or its three-dimensional counterpart, $\mathbb{Z}^3$, called the simple-cubic (sc) lattice, or its higher dimensional analogue $\mathbb{Z}^d$, called the ($d$-dimensional) hyper-cubic lattice. Other two-dimensional lattices, such as the triangular (t) and hexagonal (h) lattices, and other three-dimensional lattices, such as the body-centred cubic (bcc), face-centred cubic (fcc) and tetrahedral or diamond (d) lattices will also be mentioned.

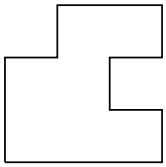
Anthony Guttmann
Department of Mathematics and Statistics, The University of Melbourne, Victoria, Australia e-mail: tonyg@ms.unimelb.edu.au

## *1.1 Definitions and notation*

Let $\mathcal{L}$ be some regular $d$-dimensional lattice. Then an *n-step self-avoiding walk* (SAW) $\omega$ on $\mathcal{L}$ is a sequence of *distinct* points $\omega_0, \omega_1, ..., \omega_n$ in $\mathcal{L}$ such that each point is a nearest neighbour of its predecessor. We assume all walks to begin at the origin ($\omega_0 = 0$) unless stated otherwise. A typical SAW is shown below.



If the end-point of a SAW, $\mathbf{x}$ is adjacent to the origin $\omega_0$, an additional step joining the end-point to the origin will produce a *self-avoiding circuit*, which has been referred to in earlier literature as a *self-avoiding return*. An example is shown below:



The distinction between a self-avoiding circuit and a self-avoiding polygon is that the former is both rooted and directed. An $n$-step SAP can have any of its $n$ vertices as the root, or starting point, and it can be traversed in two directions. Thus if there are $p_n$ SAPs of perimeter $n$, there are $2np_n$ self-avoiding circuits of perimeter $n$. More descriptively, a SAP is a closed, connected, non-intersecting, undirected path on a lattice. We consider two given SAW (or SAP) to be equivalent if they are translates of one another. Thus, on the square lattice, the first non-zero embedding of a SAP is the unit square, of perimeter 4 and area 1. If we denote by $p_m$ the number of SAP of perimeter $m$, by $a_n$ the number of SAP of area $n$, and by $p_{m,n}$ the number of SAP of perimeter $m$ and area $n$, we can define two single variable generating functions, for perimeter and area respectively, and a two-variable generating

function, as follows:

$$P(x) = \sum_m p_m x^m, \tag{1}$$

$$A(q) = \sum_n a_n q^n \tag{2}$$

$$\mathcal{P}(x,q) = \sum_{m,n} p_{m,n} x^m q^n. \tag{3}$$

Turning briefly to SAWs, let $c_n$ be the number of $n$-step SAWs on a lattice $\mathcal{L}$ starting at the origin and ending anywhere. Then the SAW generating function is defined to be

$$C(x) = \sum_n c_n x^n.$$

If one joins an $n$-step and an $m$-step SAW end-to-end, one will either obtain an $n+m$ step SAW, or a non-self-avoiding walk. A moment's thought then yields the inequality $c_{n+m} \leq c_n c_m$. This is a sub-multiplicative inequality, apparently first discussed by Fekete [12], from which follows the existence of the *growth constant*[1] $\mu > 0$, given by

$$\mu = \lim_{n\to\infty} c_n^{1/n} = \inf_n c_n^{1/n}.$$

Kesten [30, 31] has proved the stronger result that $\mu^2 = \lim_{n\to\infty} c_{n+2}/c_n$, and O'Brien [41] has proved that $c_n > c_{n-1}$ for all $n$, yet there is still no proof that $\mu = \lim_{n\to\infty} c_{n+1}/c_n$, for $d = 2,\quad 3,\quad 4$. We note in passing that a number of authors have explicitly stated, or implicitly assumed, this to be true. While it probably is, it must be remembered that no proof exists.

Hammersley [20] similarly proved that $p_m$ grows exponentially with $m$, more precisely that

$$\mu = \lim_{m\to\infty} p_{2m}^{1/2m}.$$

While far from obvious, it *is* true that the growth constants $\mu$ that arise in the polygon case and the walk case are identical [20]. While unproven, a much stronger result is widely believed, notably that

$$p_m \sim \text{const} \times \mu^m m^{\alpha-3}, \tag{4}$$

where $\alpha$ is a *critical exponent*[2]. (Note that $p_{2m+1} = 0$ for SAP on $\mathbb{Z}^d$, as only polygons with even perimeter can exist on those lattices. For such lattices the above asymptotic form is of course only expected to hold for even values of $m$. For so-called *close-packed* lattices, such as the triangular or face centred

[1] Some authors refer to this quantity as the connective constant, but the term *connective constant* originally referred to $\log \mu$. Contemporary usage seems to favour the former meaning.

[2] The notation $a_m \sim b_m$ means that $lim_{n\to\infty} \frac{a_m}{b_m} = 1$.

cubic lattices, polygons of all perimeters greater than two are embeddable, so eqn. (4) stands as stated).

Another quantity of interest for SAW is $c_n(\mathbf{x})$, the number of $n$-step SAWs on $\mathcal{L}$ starting at the origin and ending at $\mathbf{x}$. Then $c_n$ (defined above) and $c_n(\mathbf{x})$ are believed to have the asymptotic behaviour as $n \to \infty$;

$$c_n \sim \text{const} \times \mu^n n^{\gamma-1} \tag{5}$$

$$c_n(\mathbf{x}) \sim \text{const} \times \mu^n n^{\alpha-2} \qquad (\mathbf{x} \text{ fixed} \neq 0) \tag{6}$$

where the growth constant $\mu$ is defined above, and $\gamma$ and $\alpha$ are *critical exponents*. The growth constant depends on the lattice, so changes as one changes from, say the square lattice to the triangular lattice. The critical exponents, by contrast, are expected to depend only on dimensionality, and so do not change from lattice to lattice (for lattices of the same dimensionality).

As shown above, the existence of the exponential growth term $\mu^n$ is known rigorously, while the existence of sub-dominant terms $n^{\gamma-1}$ and $n^{\alpha-2}$ is believed by all reasonable men (and women), but even the existence of the exponents remains unproved for lattices of dimensionality $d < 5$. Rigorous results concerning the asymptotic behaviour of the properties of SAW and SAP will be given in the next chapter by Whittington.

Indeed, in two dimensions, it is widely accepted, due to the Coulomb gas calculations (see Chapter 14) of Nienhuis [39] that $\gamma = 43/32$ and $\alpha = 1/2$ exactly. In three dimensions only numerical estimates are available[3], notably $\gamma \approx 1.158$ and $\alpha \approx 0.2372$ and there is no reason to expect the exponents to be rational, while in four dimensions and more *mean-field* exponents are expected. That is to say, $\alpha = 0$ and $\gamma = 1$.

In exactly four dimensions there are also confluent logarithmic terms, which vanish in higher dimensions. In four dimensions, one expects $c_n \sim \text{const} \times \mu^n n^{\gamma-1} (\log n)^{1/4}$. Recently, for a hierarchical four dimensional lattice, Brydges and Imbrie [3] proved the presence of the confluent logarithmic term, though no proof is currently known for SAW on a regular lattice. For $d > 4$ Hara and Slade [22, 23] have proved that $\gamma = 1$.

This can be understood heuristically, as with increasing lattice dimensionality, the self-avoiding constraint becomes less significant. If a direction is blocked, the walker can escape to another dimension. Above four dimensions, the self-avoiding constraint is sufficiently weakened that the walks behave like random walks, albeit an exponentially small subset. In four dimensions, which is the so-called *marginal dimensionality*, the self-avoiding constraint is just strong enough to modify the random walk behaviour by the addition of confluent logarithmic terms in the singular part of the generating function. These remarks will be quantified in Chapter 6.

[3] These estimates come from Monte Carlo estimates, series estimates and field theory estimates [36, 2, 34]. The three techniques give results that typically agree up to variations in the last quoted figure.

To prove the existence and value of the critical exponents in two dimensions, the most promising route appears to be based on recent beautiful work by Cardy, Lawler, Schramm, Smirnov, Werner and others [35]. It will be necessary to prove the existence and conformal invariance of the scaling limit for self-avoiding walks. If, as expected, the scaling limit is described by stochastic Löwner evolution, parameterised by 8/3, known as $SLE_{8/3}$ , the existence and value of $\gamma$ for two-dimensional lattices would follow, as would a number of other results. These remarks will be explained in detail in Chapter 15.

The exact value of $\mu$ is not generally known, but for the special case of the two-dimensional hexagonal lattice, it is accepted [39] that $\mu = \sqrt{(2+\sqrt{2})}$. For other two-dimensional lattices we have very accurate numerical estimates, notably $\mu = 2.63815853031..$ for the square lattice, and $\mu = 4.150797226..$ for the triangular lattice. Indeed, for the square lattice note that the estimate above is indistinguishable from the root of the polynomial $581x^4 + 7x^2 - 13$, an observation first made by Conway *et al.* in 1993 [6]. For other lattices, in both two and three dimensions, see the Appendix.

Next we define several measures of the *size* of a $n$-step SAWs and SAPs. For SAWs:

- The *squared end-to-end distance* is

$$R_e^{2,(saw)} \;=\; \omega_n^2 \;. \tag{7}$$

- The *squared radius of gyration* is

$$R_g^{2,(saw)} \;=\; \frac{1}{2(n+1)^2} \sum_{i,j=0}^{n} (\omega_i - \omega_j)^2 \;. \tag{8}$$

- The *squared distance of a monomer from the endpoints* is

$$R_m^{2,(saw)} \;=\; \frac{1}{2(n+1)} \sum_{i=0}^{n} \left[\omega_i^2 \;+\; (\omega_i - \omega_n)^2\right] \;. \tag{9}$$

For SAPs only the second of these three quantities is analogously defined:

- The *squared radius of gyration* of a SAP of perimeter $n$ is

$$R_g^2 \;=\; \frac{1}{2n^2} \sum_{i,j=0}^{n} (\omega_i - \omega_j)^2 \;. \tag{10}$$

The analogue of the end-to-end distance is usually taken to be the *caliper span* of the polygon, which is defined to be the maximum Euclidean distance between any two sites (monomers, in the language of polymers) of the polygon.

If one is considering *rooted* polygons (where one site is nominated as the origin), one could readily define the mean-square distance of a monomer from the origin, but this quantity has not been considered of much interest.

We then consider the mean values $\langle R_e^2 \rangle_n$, $\langle R_g^2 \rangle_n$ and $\langle R_m^2 \rangle_n$ in the probability distribution that gives equal weight to each $n$-step SAW. Very little has been proven rigorously about these mean values, but they are believed to have the leading asymptotic behaviour

$$\langle R_e^{2,(saw)} \rangle_n, \langle R_g^{2,(saw)} \rangle_n, \langle R_m^{2,(saw)} \rangle_n, \langle R_g^{2,(sap)} \rangle_n \ \sim \ \text{const} \times n^{2\nu} \qquad (11)$$

as $n \to \infty$, where $\nu$ is another (universal) critical exponent. Hyperscaling [44] predicts that

$$d\nu \ = \ 2 - \alpha \, . \qquad (12)$$

For SAWs in two dimensions, Coulomb-gas arguments [39, 40] as well as arguments based on stochastic Löwner evolution (SLE) [35] predict that $\nu = 3/4$. Note that this result is not rigorous. In three dimensions we have only numerical estimates $\nu \approx 0.5876$,[4] a result which encompasses a variety of series, Monte Carlo and field theory estimates [36, 2, 34], while at the upper critical dimension of 4, it is believed (but not proved) that $\nu = 1/2$ with logarithmic corrections, notably $\langle R_e^{2,(saw)} \rangle_n \sim const. \times n(\log n)^{1/2}$ while for $d > 4$ the mean-field value $\nu = 1/2$ holds. This is a rigorous result (see the chapter by Clisby and Slade).

As noted above, because polygons are closed, it is almost as natural to ask for their enumeration by *area*, as by perimeter[5]. For specificity, consider polygons on the square lattice. Recall that $a_n$ denotes the number of polygons of area $n$. Then by concatenation arguments, similar to those given above for SAWs, it is possible to prove that

$$\kappa = \lim_{n \to \infty} a_n^{1/n}$$

exists. Further, it is universally believed that

$$a_n \sim \text{const} \times \kappa^n n^\tau . \qquad (13)$$

Only numerical estimates of $\kappa$ and $\tau$ are available. For the hexagonal, square and triangular latices respectively, the best current estimates[6] of $\kappa$ are $\kappa_{hexagonal} = 5.16193016(3)$, $\kappa_{square} = 3.970944(2)$, and $\kappa_{triangular} = 2.9446596(3)$, while $\tau$ is believed to be $-1$, so that the generating function

$$\mathcal{A}(x) = \sum a_n x^n \sim A \log(1 - \kappa x).$$

[4] A newly developed Monte Carlo algorithm by Clisby, briefly described in Chapter 9, is dramatically more powerful, permitting him to estimate $\nu = 0.587597(5)$ and $\mu = 4.684043(12)$.

[5] At least for polygons on two dimensional lattices. The concept of area for higher dimensional lattices is not straightforward

[6] Kindly provided by Iwan Jensen, unpublished

Consequently,

$$a_n \sim const. \times \kappa/n.$$

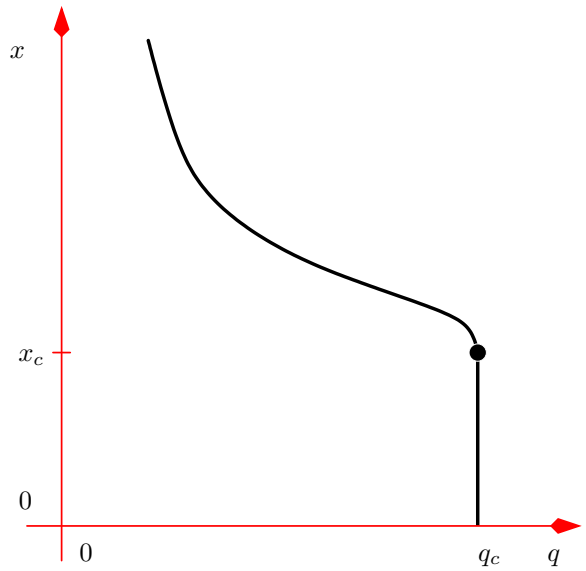
Of great interest is the two-variable generating function

$$\mathcal{P}(x,y) = \sum_m \sum_n p_{m,n} x^m q^n.$$

From this, we can define the free energy

$$\kappa(q) = \lim_{m\to\infty} \frac{1}{m} \log\left(\sum_n p_{m,n} q^n\right).$$

It has been proved [13] that the free energy exists, is finite, log-convex and continuous for $0 < q < 1$. For $q > 1$ it is infinite. The radius of convergence of $P(x,q)$, which we denote $x_c(q)$, is related to the free energy by $x_c(q) = e^{-\kappa(q)}$. This is zero for fixed $q > 1$. A plot of $x_c(q)$ in the $x - -q$ plane is shown (qualitatively) below. In the case of SAP, for $0 < q < 1$, the line $x_c(q)$ is believed to be a line of logarithmic singularities of the generating function $P(x,q)$.. The line $q = 1$, for $0 < x < x_c(1)$ is believed to be a line of *finite* essential singularities [13]. At the point $(x_c, 1)$ we have more complicated behaviour, and this point is called a tricritical point. For the simpler polygon model, that of staircase polygons (see Chapter 3), the figure is qualitatively similar, though for $0 < q < 1$, the line $x_c(q)$ is known to be a line of simple pole singularities of the generating function $P(x,q)$, rather than logarithmic, in the case of SAP. The line $q = 1$, for $0 < x < x_c(1)$ is also known [43] to be a line of *finite* essential singularities. As for SAP, at the point $(x_c, 1)$ we have more complicated tricritical behaviour, and this is discussed in more detail in Chapter 11.




Around the point $(x_c, 1)$ we expect tricritical scaling, so that

$$\mathcal{P}^{(sing)}(x,q) \sim (1-q)^{\theta} F((x_c - x)(1-q)^{-\phi}) \ (x,q) \to (x_c, 1^-).$$

Here the superscript $(sing)$ means the singular part. There is an additional, additive part that is regular in the neighbourhood of $(x_c, 1)$, and so is not relevant to this discussion. For staircase polygons, Prellberg [43] has proved that

$$F(s) = \frac{1}{16}\frac{\mathrm{d}}{\mathrm{d}s} \log \mathrm{Ai}\left(2^{8/3} s\right),$$

where $s = (x_c - x)(1-q)^{-\phi}$. For self-avoiding polygons, in a series of papers, Richard and co-authors [49, 50, 48] have provided abundant evidence for the surprisingly strong conjecture that

$$F(s) = -\frac{1}{2\pi} \log \mathrm{Ai}\left(\frac{\pi}{x_c}\,(4A_0)^{\frac{2}{3}}\, s\right) + C(q).$$

Here $C(q)$ is a function, independent of $x$, that arises as a constant of integration when moving from the rooted to the unrooted SAP scaling function [50]. For both models $\phi = 2/3$, while $\theta = 1/3$ for staircase polygons and $\theta = 1$ for SAP. Here $A_0 = 0.2811506(1)$ is the amplitude of the perimeter generating function, and $x_c$ is the perimeter generating function critical point. This result is described in much greater detail in Chapter 11 by Richard.

## *1.2 The connection with statistical mechanics*

Thus far we have discussed SAW and SAP as a combinatorial problem. However there is an alternative description as a statistical mechanical model. This both sheds light on some aspects of the behaviour of various quantities of interest, described above, such as critical points, critical exponents and critical amplitudes, and also gives us another way to tackle the problems that arise.

In statistical mechanics one starts with a Hamiltonian, which is the energy of a given configuration, taking into account the interactions, and then constructs the *partition function.* This assumes that configurations of a given energy and temperature are distributed according to the Boltzmann distribution. The Ising model and Heisenberg models of ferromagnetism have been known for decades, and were defined by Hamiltonians in which each spin was a classical unit vector of dimension 1 in the case of the Ising model, dimension 2 in the case of the classical planar Heisenberg model, and dimension 3 in the case of the classical Heisenberg model.

In 1968 Stanley [51] proposed the *n-vector model*, in which each spin was represented by a classical, $n$-dimensional vector. Clearly this reproduces the three cases mentioned above when $n = 1$, 2, 3 respectively. Stanley pointed out that the suitably defined limit as $n \to \infty$ reproduces the so-called spherical model, introduced by Berlin and Kac [4]. A more remarkable limit, $n \to 0$ was proposed by de Gennes [8], who showed that in this limit the model described was exactly self-avoiding walks and polygons. This is not at all intuitive. What after all do we mean by the interaction between two zero-component vectors? Nevertheless, proceeding with a curious mixture of formality and non-rigorous argument, a connection can be established. We will sketch the result. More details can be found in [7], and more recent, and perhaps more accessible accounts can be found in Madras and Slade [37] and Hughes [24].

One starts with a Hamiltonian

$$\mathcal{H} = -\sum_{<i,j>} \boldsymbol{s}_i \cdot \boldsymbol{s}_j - H \sum_i s_i(1), \tag{14}$$

where $\boldsymbol{s}$ is an $n$-component vector, which is located at the vertices of a regular lattice of arbitrary spatial dimension, and the subscript indexes a particular site on the lattice. The first sum is taken over all pairs of sites $(i, j)$ which are nearest neighbours on the lattice. The second sum is a scalar product of a magnetic field $H$, assumed w.l.o.g. to be in the '1' direction, and so couples to spin components in that direction. The vector is assumed to be of magnitude $\sqrt{n}$, so that $|\boldsymbol{s}_i|^2 = n$.

The partition function is obtained by the usual rules of statistical mechanics from

$$Z = \sum_{\text{all configurations}} \exp(-\mathcal{H}/kT). \tag{15}$$

In the above equation, $k$ is Boltzmann's constant and $T$ is the absolute temperature. The *Helmholtz free energy* is defined by

$$F(T,H) = -kT \log Z(T,H),$$

though we usually work in the *thermodynamic limit*,

$$f(T,H) = \lim_{N\to\infty} \frac{1}{N} F(T,H). \tag{16}$$

In the above equation, $N$ refers to the number of spins in the lattice, and should not be confused with the dimensionality of the spin vectors, denoted $n$, whereas the underlying dimensionality of space is denoted $d$. Next, we need to define the *magnetic susceptibility*,

$$\chi(T) = -\frac{\partial^2 f}{\partial H^2}|_{H=0} \tag{17}$$

The Ising model corresponds to the case $n = 1$. The susceptibility then behaves as

$$\chi(T) \sim const(1 - T_c/T)^{-7/4}$$

as $T \to T_c^+$. If one takes a formal series expansion of the susceptibility, after substituting equation (14) into equation (15) and then using equations (16) and (17) to calculate the susceptibility, it can be seen that each term in the expansion corresponds to sums over averages of dot products of the vector spins. These can be interpreted as graphs on the underlying lattice. In the limit as $n \to 0$ all graphs vanish except non-intersecting lattice paths joining site $i$ to site $k$. These are just self-avoiding walks from site $i$ to site $k$. Thus the susceptibility of the $n$ vector model in the $n \to 0$ limit is precisely the generating function for SAW. By the same token, the specific heat exponent for magnetic models can be related to the exponent characterising the generating function for self-avoiding polygons, so that if the specific is written

$$C(T) \sim const(1 - T_c/T)^{-\alpha}$$

as $T \to T_c^+$, we have, for the corresponding polygon generating function,

$$P(x) = \sum p_n x^n \sim const.(1 - x/x_c)^{2-\alpha}.$$

The advantage of this formulation is that one can now bring to bear the full machinery of statistical mechanics to what would otherwise be a combinatorial problem. In statistical mechanics one has *scaling arguments* that link exponents. This is one justification for the hyperscaling relation (12) given above. Other tools from statistical mechanics are also available, such as the *renormalisation group*method, and other methods developed by mathematical physicists working in *field theory*. We won't discuss these techniques in

any detail–each would justify a volume on their own–but will refer to some predictions made by these methods in later chapters.

## *1.3 Some history of the problem*

We conclude this section by a brief discussion of the chronological history of the SAW and SAP problem. The problem of enumerating self-avoiding walks (SAW) was initially proposed as a mathematical problem in a posthumously published paper by Orr in 1947 [42], and shortly thereafter was proposed as a model of long-chain polymers in dilute solution by Flory in 1949 [15]. Self-avoiding polygons (SAP), on the other hand, were first enumerated by Wakefield in 1951 [57]. Wakefield was not studying SAP in their own right, but rather enumerating them as part of a different project, notably the behaviour of the three-dimensional Ising model. In that study, SAP on the simple-cubic lattice contributed to the graphical expansion of the Ising model partition function.

Somewhat later, in 1954 and 1955, Wall and co-workers [54, 55, 56] calculated some properties of two- and three-dimensional SAP which arose as a by-product of their Monte Carlo study of SAW. Some of the SAW configurations they generated by Monte Carlo methods failed due to the coincidence of the end point of trial walks they generated with the origin. Such failures led them to study the probability of these occurrences, and hence they introduced the so-called *probability of initial ring closure*. This is defined as the probability that a SAW ends at a site adjacent to the origin, so that the addition of a single bond produces a self-avoiding circuit. This probability is just $2mp_m/((q-1)c_m)$, where $q$ is the *coordination number* of the lattice, or number of nearest-neighbour sites of a given site. It is $2d$ for a $d$-dimensional hypercubic lattice.

In terms of the notation given above, and assuming the existence of the conjectured sub-dominant terms containing the critical exponents, we have $p_m \sim \text{const} \times \mu^m m^{\alpha-3}$, and $c_m \sim \text{const} \times \mu^m m^{\gamma-1}$. The probability of initial ring closure $p_m^0$ is $p_m^0 = 2mp_m/((q-1)c_m) \sim \text{const} \times m^{\alpha-\gamma-1}$. In two dimensions, from the above-quoted values of $\alpha$ and $\gamma$, we obtain $p_m^0 \sim \text{const}/m^{1.84375}$, while from the best numerical estimates for the exponents in the three-dimensional case, we find $p_m^0 \sim \text{const}/m^{1.921}$. Wall and co-workers in 1955 estimated these exponents to be around 2 in both two and three dimensions, by Monte Carlo methods.

In 1959 Rushbrooke and Eve [47] studied this problem by direct enumeration, and ushered in the computer age of enumeration. They counted polygons up to perimeter 18 on the square lattice, and, more impressively, to order 14 on the SC lattice. The coefficient $p_{14}$ on the sc lattice was counted on the Pegasus computer at Durham University, in a calculation that took 50 hours of CPU time. The exponent estimates obtained from an analysis of

these enumerations were $2.4 \pm 0.2$ and 2.29 for the square and SC lattices respectively. At the same time, and quite independently, Fisher and Sykes [14] estimated these exponents to be $1.75 \pm 0.10$ and $1.810 \pm 0.007$ respectively. These estimates were improved by Hiley and Sykes in 1961 [21] based on further enumerations, and they found exponents $1.805 \pm 0.025$ and $1.92 \pm 0.08$ respectively. Remarkably, the estimate in the three-dimensional case is in precise agreement with current estimates to all quoted digits.

## 2 Polyominoes

The modern era of polyominoes began in 1953 when Golomb discussed them in a talk he gave to the Harvard Mathematics Club, which was subsequently published [17]. Martin Gardner added to the popularity of the subject when he discussed Golomb's article in his column in *Scientific American*. Polyominoes became a favourite topic of Gardner, and one he often returned to. As Golomb points out in the preface to the first edition of his book on the subject [16], there are many antecedents, either in the form of particular puzzles, or in discussions of the number of allowable patterns of a particular type in board games, such as *Go*. One notable antecedent appeared in 1907, when a puzzle involving 5-celled polyominoes, (of which there are twelve), was posed in the book *Canterbury Puzzles* [9]. In his book on polyomino puzzles, G. E. Martin [38] also points out that a variety of polyomino puzzle problems appeared in the British journal *Fairy Chess Review* in the 1930s. At that time, they were called *dissection problems*.

Formally, a polyomino (or, as they are sometimes known, a *lattice animal*) of $n$ cells on the square lattice with origin 0 is a connected section graph[7] of the lattice containing the origin and having $n$ vertices. More intuitively, a polyomino is comprised of $n$ connected squares, which must be joined at an edge, and not just at a vertex. A *domino* is an example—indeed, the only example—of a 2-celled polyomino. Polyominoes may be similarly defined on other two-dimensional lattices, such as the triangular or honeycomb lattice. We will assume that we are referring to square lattice polyominoes in the following, unless stated otherwise.

The difference between polyominoes and polygons is that a polyomino may have an internal hole. For example, see fig. **??**. So all polygons are polyominoes, but not vice-versa.

Polyominoes are usually enumerated by *area*—or number of cells. While, like polygons, they can also be enumerated by perimeter, we will see below that this is less easily characterised, as the perimeter generating function has zero radius of convergence. Thus the more common asymptotic form (13) for

[7] A section graph $G^*$ of a graph $G$ is obtained from $G$ by deleting zero or more of its vertices.

lattice objects enumerated by perimeter does not hold when discussing the enumeration of polyominoes by perimeter.

The sorts of results one can prove for polyominoes are not too different to those that can be proved for polygons. If $a_n$ denotes the number of polyomines of area $n$ (or, if one prefers, of $n$ cells), then Klarner [32] proved that the growth constant $\alpha$, defined by

$$\alpha = \lim_{n\to\infty} a_n^{1/n} = \sup_n a_n^{1/n},$$

exists and is non-zero. The proof follows standard concatenation arguments, whereby polyominoes are uniquely concatenated to produce a subset of polyominoes of a size given by the sum of the sizes of the two individual polyominoes. Even before the existence of this limit was proved, it was widely assumed to exist, and indeed Eden, in 1961 [11] undertook a numerical analysis to estimate its value.

Such estimates rely on the enumeration of $a_n$, a computational problem of exponential complexity. As for polygons, the most effective method to date is the finite lattice method (discussed in Chapter 7), which permits enumeration by finite-dimensional transfer matrices. The method was implemented for polyominoes by Conway and Guttmann in 1995 [5], giving rise to $a_n$ for $n \leq 25$, which was dramatically improved by Jensen and Guttmann in 2000 [26], who obtained $a_n$ for $n \leq 46$, then Knuth and Jensen played leap-frog for a few years, extending the series, with Jensen [28] currently holding the record with $a_{56} = 69,150,714,562,532,896,936,574,425,480,218$.

Polyominoes are frequently referred to as *free* or *fixed*. Fixed polyominoes are consider distinct up to a translation. That is to say, fixed polyominoes means "an equivalence class of polyominoes under translation," whereas free polyominoes refers to "an equivalence class of polyominoes under translations, rotations and reflections." Asymptotically, the number of fixed polyominoes of $n$ cells, denoted $B_n$ on the square lattice is 8 times the corresponding number of free polyominoes of $n$ cells, denote $b_n$. While the same distinction can be made for SAP, this is rarely done, and one usually considers fixed SAP, without explicitly saying so.

As for SAW and SAP, it is widely believed, but not proved, that $b_n \sim A\alpha^n n^\theta$ as $n \to \infty$, or, equivalently, that $B_n \sim 8A\alpha^n n^\theta$ as $n \to \infty$.

Like polygons, polyominoes on other two dimensional lattices can be considered. Hexagonal and triangular lattice polyominoes were first introduced by Lunnon [33], and have also been exhaustively enumerated by Jensen [27], who has obtained the counts up to 75-celled animals on the triangular lattice and 46-celled animals on the hexagonal lattice.

As is the case for polygons, the growth constant $\alpha$ will change from lattice to lattice, while the critical exponent $\theta$ is expected to be the same for all two-dimensional lattice, and its value has been predicted by powerful, but non-rigorous physical arguments [45], and is predicted to be $\theta = -1$, corre-

sponding to a logarithmic singularity in the generating function for polyominoes, while $\alpha \approx 4.062570$ [26].

## 3 Polyhedra

A three-dimensional analogue of a polyomino on the square lattice is called a *polycube* on the simple-cubic lattice. Polycubes are composed of unit cubes joined at their faces. Many topologies formed by this rule, including spheres, toroids and objects of higher genus. They are usually counted as *free* polycubes.That is to say a particular polycube is counted only once irrespective of its orientation. Those that are mirror images of one another (so-called *chiral twins*) are usually (but not always) distinct.

Consideration of chiral twins is a slightly subtle distinction from the analogous two-dimensional case, where chirality is accommodated by reflection. But the reflection takes place by flipping the object through the third dimension. By convention, since one can't "see" the fourth dimension, the analogous reflection is forbidden, and that is why chiral twin polycubes are considered distinct.

As with polyominoes, it is customary to distinguish between *fixed* and *free* polycubes. To make this distinction clear, a tower of three cubes counts 1 as a free polycube, but counts 3 as a *fixed* polycube, as it may be oriented parallel to the $x$, $y$ or $z$ axis. Asymptotically, the number of fixed polycubes is 24 times the number of free polycubes.

In three dimensions, the longest series is due to Aleksandrowicz and Barequet [1], who give the first 18 series coefficients. We again have an exact prediction for the critical exponent from the powerful but non-rigorous physical arguments of Parisi and Sourlas [45], which is $\theta = -1.5$. We have analysed the available series coefficients, using the techniques described in Chapter 8, and find $\alpha \approx 8.3479$. (In 1978 Guttmann and Gaunt [19] predicted $\theta = -1.5$ and $\alpha = 8.34 \pm 0.02$, based on much less data).

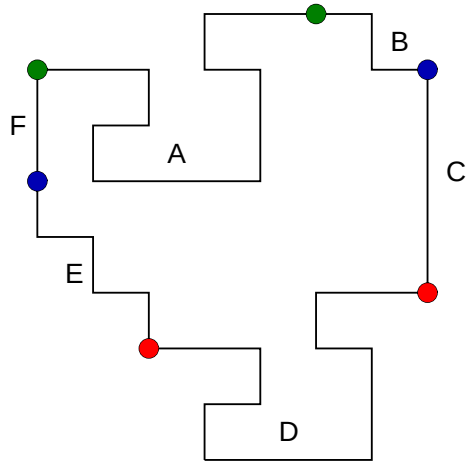
## 4 Tilings

The theory of tilings is a fascinating topic in its own right. Many books have been written on this subject, and we will have little to say about it, except to point out some of the more remarkable results, and indicate some of the open questions. An interesting, and still unanswered question is the necessary and sufficient conditions for a polyomino (or actually a polygon, as we don't want internal holes) to tile the plane. It is unknown if the tiling problem is decidable [46], and it is known that the problem of tiling a *finite* region is NP-complete [52].

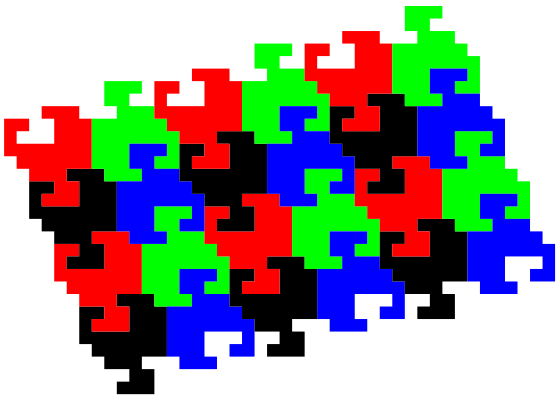
To be more precise, by a *tiling of the plane*, we mean a covering of the Euclidean plane by a countable family of elementary tiles. The tiles may not overlap. If a single tile is used, we refer to the tiling as a *monohedral* tiling. The tile used is called a *prototile*. As a trivial example, it is clear that a $1 \times 1$ cell—indeed, a $k \times k$ cell—can tile $\mathbb{Z}^2$. For a discussion of tilings of the entire plane, the reader is referred to the discussion by Grünbaum and Shephard [18].

It is easy to see that any triangle or rectangle can tile the Euclidean plane, and with a bit more effort one can see that any quadrilateral can tile the plane. Conway [46] has a powerful method that provides a *sufficient*, but not necessary, condition for a prototile to tile the plane. Conway's method requires that if one can divide the boundary up into six segments, labelled sequentially as $A$, $B$, $C$, $D$, $E$, and $F$, such that:
(i) $A$ and $D$ are translates of one another,
(ii) the remaining four segments each possess a 180 deg rotational symmetry about their mid-point, and
(iii) while some segments may be empty, each pair $B - C$ and $E - F$ must be non-empty. Also, $A$ and $D$ can both be empty if at least three of the remaining four segments are non empty.

An example is shown below. Note that segments $A$ and $D$ are translates of one another, while the remaining segments each possess a 180 deg rotational symmetry about their respective mid-points.




To tile the plane with this tile involves using both translates and rotated versions of the basic prototile, as can be seen in the following figure, illustrating the tiling:



If one restricts the movement of tiles to translations only, (that is to say, one forbids rotations and reflections), then there is a necessary and sufficient condition for a prototile to tile the plane. If the boundary can be divided into six consecutive segments, sequentially labeled $A-F$ as above, such that each element of the pairs $A-D$, $B-E$ and $C-F$ are translates of each other, the prototile tiles the plane. (One such pair may be empty, in which case the tiling forms a rectangular lattice, otherwise it forms a hexagonal lattice.) See [46] for a proof.

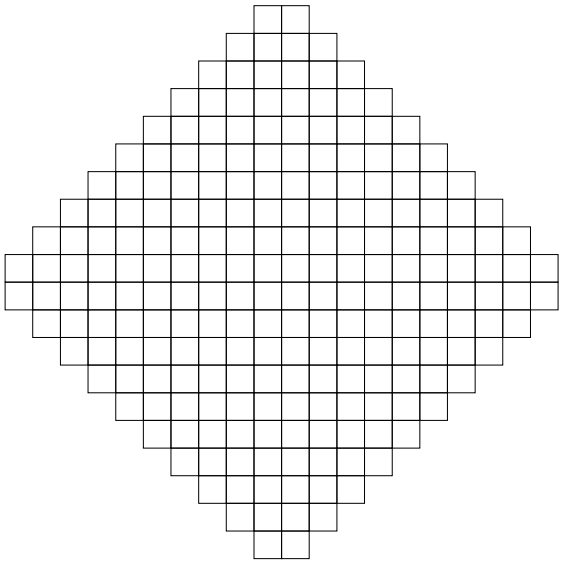
Rhoads [46] has made an interesting study of the effectiveness of Conway's criterion. He shows, among other results, that one needs to go to 9-cell polygons before one finds a polygon that does not satisfy the Conway criterion, but still tiles the plane. More precisely, apart from two 9-cell polygons, all polygons up to order 9 that tile the plane either satisfy the Conway criterion, or two copies form a patch that satisfies the Conway criterion.

A second question that is asked about tilings, once it has been established that a prototile can tile the plane is *In how many ways can the plane be tiled?*. One of the earliest non-trivial results in this direction is due to Kasteleyn [29], who used a Pfaffian to express the number of ways $2jk$ dominoes (or $2 \times 1$ tiles) can tile a $2j \times 2k$ rectangle. The same problem was treated by transfer matrix techniques, discussed in Chapter 7, by Fisher and Temperley [53]. The result is

$$4^{jk} \prod_{l=1}^{j} \prod_{m=1}^{k} \left( \cos^2 \frac{l\pi}{2j+1} + \cos^2 \frac{m\pi}{2k+1} \right).$$

It is far from obvious that this double product should even produce integers! The reader is invited to try it out for small values of $j$ and $k$ and confirm their result by direct enumeration.

We will mention just one other remarkable exact result, the enumeration of *Aztec diamonds.* An Aztec diamond $AZ(k)$ is constructed by reflecting a pyramidal stack about its base. The stack base contains $2k$ cells, the next row $2k-2$ cells, the subsequent row $2k-4$ cells, and the top of the stack contains two cells. The Aztec diamond $AZ(10)$ is shown in the figure below.



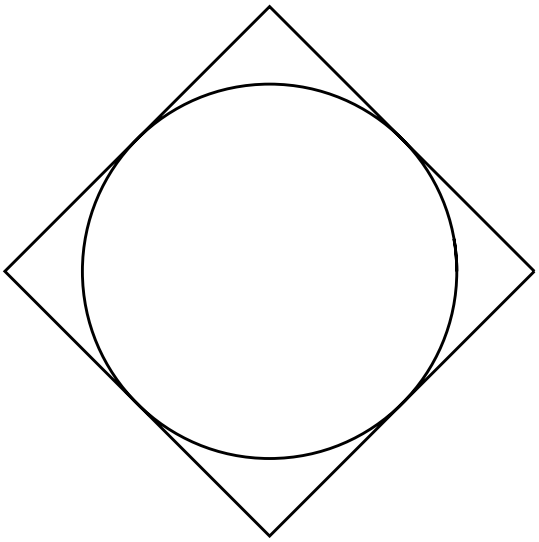
The number of domino tilings of $AZ(n)$ was found by Elkies, Kuperberg, Larsen and Propp [10], and is just

$$2^{(n^2+n)/2}.$$

This relatively simple result might suggest that there is a simple proof. Unfortunately that is not yet the case, despite the fact that the number of different proofs now runs to double figures.

Another typical question that is asked, is *What is the typical shape of an object?*. Some tilings, or objects, display no particular structure, whereas others show quite remarkable properties. For example, the typical large Aztec diamond displays a surprising regularity at its corners, where, as can be seen from figure 4 it is tiled in a regular, brickwork, pattern, switching abruptly to

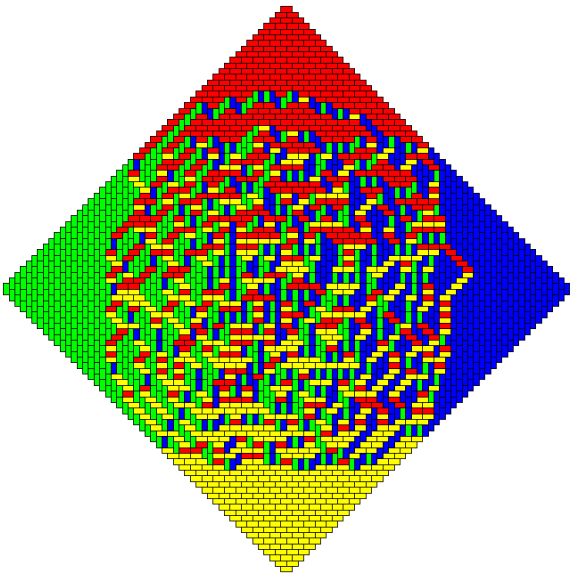
apparent randomness in a circular region, corresponding to the largest possible circle that can be drawn inside the boundary. This geometry is sketched in the figure below:



Jokush, Propp and Shor [25] proved this result in probability. It is known as the *Arctic circle theorem,* The inscribed circle is the Arctic circle, the region at the poles, (and, perhaps disturbingly, at the equator), is "frozen". Hence the name. A typical such tiling is shown in the figure below:

## 5 The rest of the story

In the next chapter, Whittington describes the rigorous results that can be proved for SAP. While the SAP model remains unsolved, a number of simpler polygon models can be exactly solved, and methods for solving these are described in Chapter 3 by Bousquet-Mélou and Brak. In recent years we have gained a better understanding of just why the SAP model is so hard, and these ideas are described in Chapter 4. The key idea is that the solution is not holonomic, and in Chapter 5 Rechnitzer proves this, along the way developing tools that are useful for similar proofs. In Chapter 6 Clisby and Slade develop the lace expansion method that has proved such a powerful tool for both rigorous results and numerical results in higher dimensions ($d \geq 3$).



Much of our knowledge of SAP comes from numerical methods. In two dimensions the finite lattice method (FLM) is, for most applications, the most powerful method, and it is described in Chapter 7 by Enting and Jensen. The result of the application of this method is a series expansion for the generating functions we have defined. Analysing this data to extract critical properties, such as critical exponents, critical points and critical amplitudes is the subject of Chapter 8 by Guttmann and Jensen, while the alternative numerical method, Monte Carlo, is described by Janse van Rensburg in Chapter 9.

Interesting phenomena occur if SAW and SAP are restricted to confined geometries, such as polygons in strips, in slabs and in rectangles. These are described in Chapter 10 by Jensen and Guttmann. The underlying limit distributions of the area of SAPs is the subject of Chapter 11 by Richard. Many physical, chemical and biological properties of polymers can be modelled by including monomer-monomer or monomer-surface interactions, and these are the subject of Chapter 12 by Owczarek and Whittington.

Recent developments that combine powerful physical concepts with advances in probability are the subjest of the last three chapters. In chapter 13 de Gier describes fully packed loop models, in which the lattice is filled with polygons. These models impact on a variety of combinatorial and physical problems. In Chapter 14 Jacobsen describes the theory of conformal invariance, and how it can be used to make amplitude predictions, and many other powerful results, while in Chapter 15 Nienhuis and Kager describe the ideas behind stochastic Löwner evolution, which is believed to describe the scaling limit of a number of processes, including SAP and SAW.

## References


1. G. Aleksandrowicz and G. Barequet (2008) *Proc. 14th Ann. Int. Computing and Combinatorics Conf.,* Dalian, China Lecture Notes in Computer Science **5092** 100-109 June 2008.
2. Butera P and Comi M (1997) N-vector spin models on the simple cubic and body-centred cubic lattices Phys. Rev B **56** 8212–8240
3. D. Brydges and Y. Imbrie (2003) The Green's function for a hierarchical SAW in four dimensions Commun. Math. Phys **239** 549-584
4. T Berlin and M Kac (1952) The spherical model of a ferromagnet, Phys Rev. **86** 821-835
5. Conway A R and Guttmann A J (1995) On two-dimensional percolation J. Phys A:Math. Gen. **28** 891-904
6. Conway A R, Enting I G and Guttmann A J, Algebraic techniques for enumerating self-avoiding walks on the square lattice (1993), J. Phys. A:Math. Gen **26** 1519-34.
7. M Daoud, J P Cotton, B Farnoux, G Jannink, G Sarma, H Benoit R Dupplesix, C Picot and P-G de Gennes (1975) Solutions of flexible polymers. Neutron experiments and interpretation, Macromolecules **8** 804-818
8. P-G de Gennes, (1972) Exponents for the excluded volume problem as derived by the Wilson method, Phys. Letts A **38** 339-340.
9. H.E. Dudeney, *The Canterbury Puzzles and Other Curious Problems*, (Dover, New York 1958). (Reprint of 1907 edition)
10. N Elkies, G Kuperberg, M Larsen and J. Propp (1992) J. Alg. Comb, **1**, 111-132 and 219-234.
11. M. Eden: A two-dimensional growth process. In *Proc. Fourth Berkeley Symp. Math. Stat. Probab.* vol 4, ed by J. Neyman (Univ. Calif. Press 1961) pp 223-239
12. Fekete, M, Über die Verteilung der Wurzeln bei gewissenalgbraischen Gleichungen mit ganzzahligen Koeffizienten, (1923) Math. Z **17** 228-49
13. Fisher M E, Guttmann A J and Whittington S G Two-dimensional lattice vesicles and polygons (1991) J. Phys A:Math. Gen. **24** 3095-3106.
14. Fisher M E and Sykes M F, Excluded volume problem and the Ising model of ferromagnetism (1959) Phys. Rev. **114** 45-58
15. Flory, P J The configuration of real polymer chains (1949) J. Chem. Phys. **17** 303-10.
16. S. W. Golomb: *Polyominoes*, (Charles Scribners' Sons, 1965)
17. S..W. Golomb (1954) Checkerboards and Polyominoes, American Math. Monthly, **61** No 10 675-82
18. B. Grünbaum and G C Shephard (1989) *Tilings and Patterns: An Introduction* W H Freeman and Co, New York
19. A. J. Guttmann and D. S. Gaunt (1978) On the asymptotic number of lattice animals in bond and site percolation, J. Phys A: Math. Gen, **11** 949-953
20. Hammersley, J M, (1961) On the number of polygons on a lattice, Proc. Cam. Phil. Soc, **57** 516-23.
21. Hiley B J and Sykes M F (1961) Probability of initial ring closure in the restricted random walk model of a macromolecule (1961) J. Chem. Phys. **34** 1531-37
22. T. Hara and G. Slade, (1992) Self-avoiding walk in five or more dimensions. I. The critical behaviour, Commun. Math. Phys., **147** 101–136, (1992)
23. T. Hara and G. Slade, (1992) The lace expansion for self -avoiding walks in five or more dimensions Reviews Math. Phys. **4**, 235-327
24. B D Hughes (1995) *Random walks and random environments, Volume 1: Random walks* (Clarendon Press, Oxford)
25. W Jockusch J. Propp and P Schor Random domino tilings and the arctic circle theorem. Preprint. Available at arxiv.org/abs/math.CO/9801068.
26. Jensen I and Guttmann A J Statistics of lattice animals (polyominoes) and polygons (2000) J. Phys A:Math. Gen. **33** L257-63.

27. Jensen I see www.ms.unimelb.edu.au$\tilde{\text{i}}$wan .
28. I. Jensen: Counting polyominoes: A parallel implementation for cluster computing. Lecture Notes in Computer Science, **2659** 203-212 (2003)
29. P. Kasteleyn (1961) The statistics of dimers on a lattice I, The number of dimer arrangements on a quadratic lattice Physica **27** 1209-1225
30. Kesten, H, On the number of self-avoiding walks, (1963) J. Math. Phys, **4**, 960-9
31. Kesten, H, On the number of self-avoiding walks II, (1964) J. Math. Phys, **5**, 1128-37
32. Klarner D A Cell growth problems (1967) Cand. J. Math. **19** 851-63.
33. W. F. Lunnon: Counting hexagonal and triangular polyominoes (1972) Graph Theory and Computing, ed. R. C Read, Academic Press pp. 87-100.
34. Li B, Madras N and Sokal A D Critical Exponents, hyperscaling and universal amplitude ratios for two-and three-dimensional self-avoiding walks, (1995) J. Stat. Phys. **80** 661-754
35. G.F. Lawler, O. Schramm and W. Werner, in *Fractal Geometry and Applications: A Jubilee of Benoît Mandelbrot*, Part 2, Proceedings of Symposia in Pure Mathematics #72 (American Mathematical Society, Providence RI, 2004), pp. 339–364, math.PR/0204277.
36. MacDonald D, Joseph S, Hunter D L, Moseley L L, Jan N and Guttmann A J (2000) Self-avoiding walks on the simple cubic lattice, (2000) J. Phys. A **33** 5973-83.
37. N Madras and G Slade *The Self-Avoiding Walk*, (Boston:Birkhäuser 1993)
38. G.E. Martin: *Polyominoes: A guide to puzzles and problems in tiling*, (The Mathematical Association of America, 1991)
39. B. Nienhuis, (1982) Exact critical point and critical exponents of O(n) models in two dimensions, Phys. Rev. Lett. **49**, 1062-1065
40. B. Nienhuis, (1984) Critical behaviour of two-dimensional spin models and carge asymmetry in the Coulomb gas J. Stat. Phys. **34**, 731-761
41. O'Brien, G L, Monotonicity of the number of self-avoiding walks (1990) J. Stat. Phys, **59** 969-79
42. Orr, W J C Statistical treatment of polymer solutions at infinite dilution (1947) Trans. of the Faraday soc., **43**, 12-27.
43. Prellberg T Uniform $q$-series asymptotics for staircase polygons, (1995) J. Phys A:Math. Gen. **28** 1289-1304.
44. A. Pelissetto and E. Vicari, Critical phenomena and renormalization-group theory Phys. Reports **368**, 549-727 (2002),
45. G. Parisiand N. Sourlas (1981) Critical behaviour of branched polymers and the Lee-Yang edge singularity Phys. Rev. Letts, **46** 871-874
46. G C Rhoads (2005) Planar tilings by polyominoes, polyhexes, and polyiamonds, J. Comp and Appl. Math. **174** 329-353
47. Rushbrooke and Eve, (1959) On non-crossing lattice polygons J. Chem. Phys. **31**, 1333-4
48. C Richard (2002) Scaling behaviour of two-dimensional polygon models. J. Stat. Phys., **108** 459493
49. C Richard, A J Guttmann and I Jensen (2001) Scaling function and universal amplitude combinations for self-avoiding polygons. J. Phys. A: Math. Gen., 34:L495L501
50. C Richard, I Jensen and A J Guttmann (2004) Scaling function for self-avoiding polygons revisited. J. Stat. Mech.: Th. Exp., page P08007
51. H E Stanley (1968) Dependence of critical properties on dimensionality of spins, Phys Rev Letts, **20** 589-592.
52. D. Schattschneider, (1990) *Visions of Symmetry,* W H Freeman and Co, New York
53. H.N.V. Temperley and M.E. Fisher, (1961) Dimer problem in statistical mechanics–an exact result Phil Mag **6** 1061-1063
54. Wall, F T, Hiller, A L and Wheeler D J, Statistical Computation of Mean Dimensions of Macromolecules I, (1954) J. Chem. Phys, **22** 1036-41.
55. Wall, F T, Hiller, A L and Atchison, W F, Statistical Computation of Mean Dimensions of Macromolecules II, (1955) J. Chem. Phys, **23** 913-921.

56. Wall, F T, Hiller, A L and Atchison, W F, Statistical Computation of Mean Dimensions of Macromolecules III, (1955) J. Chem. Phys, **23** 2314-20.
57. Wakefield, A J (1951) Statistics of the simple cubic lattice Proc. Camb. Phil. Soc., **47** 419-435

# Index