

# Self-Avoiding Walks and Polygons — An Overview

Anthony J Guttmann

*Abstract.* We give a rather personal review of the problem of self-avoiding walks and polygons. After defining the problem, and outlining what is known rigorously, and what is merely conjectured, we highlight the major outstanding problems in the field. We then give several applications in which the author has been involved. These include a study of surface adsorption of polymers, counting possible paths in a telecommunication network and the modelling of biological experiments on polymers, in which a polymer is pulled from a wall. The purpose of the review is to show that the problem is not only of intrinsic mathematical interest, but also has many interesting and useful applications.

## 1. Introduction

The problem of self-avoiding walks is one of deceptive simplicity of definition, hiding malevolent difficulty of solution. The problem was introduced by two theoretical chemists, Orr [31] and Flory [13], as a model of a polymer in dilute solution. It soon became an interesting combinatorial model to mathematicians, and a canonical model of phase transitions, of interest to mathematical physicists. It is also a simple model of a non-Markovian process. Attempts to count the number of SAW have led to the development of new algorithms, with widespread applicability, while many more applications were discovered. These include application to the design of telephone networks, the folding and knotting of biological molecules, and a variety of chemical phenomena. Attempts at a solution have driven several mathematical advances, including developments in stochastic differential equations and probability theory.

Nearly 70 years after the model was proposed, we have a huge amount of numerical information, a substantial amount of exact information — that is to say, results that are universally believed, but remain unproved — and a very small body of rigorous results. In contrast, some other canonical models of phase transitions, such as the Ising model, the Potts model and percolation have either been solved (the Ising model) or much has been rigorously proved. In this short article I will outline the development of the subject, give some applications, and show that we appear to be on the verge of some major breakthroughs, which will result in proofs of much of the exact,

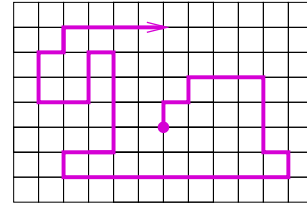


Fig. 1. A self-avoiding walk on the square lattice

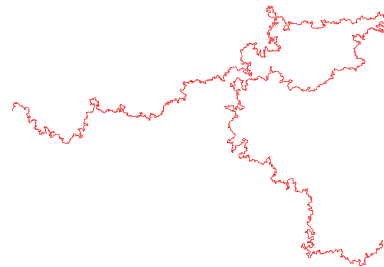


Fig. 2. A typical two-dimensional SAW of  $2^{25}$  steps on the square lattice [courtesy of Nathan Clisby]

but unproved, information that currently exists. Unfortunately all the exact and conjectural information we have applies only to the model on a two-dimensional lattice. In the case of three dimensions, we only have numerical results. Except where otherwise stated, this article will discuss the two-dimensional situation.

## 2. What is Known and What is Not

### 2.1. Self-avoiding walks

A self-avoiding walk (SAW) of length  $n$  on a periodic graph or lattice  $\mathcal{L}$  is a sequence of distinct vertices  $w_0, w_1, \dots, w_n$  in  $\mathcal{L}$  such that each vertex is a nearest neighbour of its predecessor. In Fig. 1 a short SAW on the square lattice is shown, while in Fig. 2 a rather long walk of  $2^{25}$  steps is shown (generated by a Monte Carlo algorithm [8, 9]).

#### 2.1.1. How many self-avoiding walks are there?

Two obvious questions one might ask are (i) how many SAWs are there of length  $n$ , (typically defined up to translations) denoted  $c_n$ , and (ii) how big is a typical  $n$ -step SAW? Indeed, how

might we measure size? A third important, but less obvious question, asks “what is the scaling limit of SAWs?”

Frequently one rather considers the associated generating function

$$C(x) = \sum_{n \geq 0} c_n x^n .$$

To see the difficulty of this problem, the reader is invited to try and calculate the first few terms  $c_n$  on  $\mathbb{Z}^2$ . We take  $c_0$  to be 1, then  $c_1 = 4$  as a one step walk can be in any of 4 directions. Then  $c_2 = 12$ ,  $c_3 = 36$  and  $c_4 = 100$ . It is at the stage of 4-step SAWs that the self-avoiding constraint first manifests itself, and the problem becomes increasingly difficult thereafter.

As we prove below,  $c_n$  grows exponentially. Accordingly, an enormous amount of effort has been expended over the last 50 years in developing efficient methods for counting SAW. For the square lattice, Jensen [20] has extended the known series to 79 step walks, for which he finds  $c_{79} = 10194710293557466193787900071923676$ . Methods for calculating these astonishing numbers are quite complicated (see [14, Chap. 7]), but the best current algorithm still involves a counting problem of exponential complexity, of about  $1.3^n$  (while a direct counting algorithm would have complexity  $2.64^n$ ).

One of the few properties one can readily prove, by virtue of the obvious sub-multiplicative inequality  $c_{n+m} \leq c_n c_m$ , is that the number  $c_n$  grows exponentially. From this inequality it follows that

$$\mu := \lim_{n \rightarrow \infty} c_n^{1/n} = \inf_n c_n^{1/n}$$

exists [29], and further that  $c_n \geq \mu^n$ .

However even the value of this “growth constant”  $\mu$  is difficult to calculate exactly. Only in 2010 was  $\mu$  for one two-dimensional lattice, notably the honeycomb lattice, actually proved by Duminil-Copin and Smirnov [10] to be  $\sqrt{2 + \sqrt{2}}$  (see Sec. 3). For other lattices in two dimensions, and all lattices in higher dimensions, we only have numerical estimates. For example, for the square lattice the best current estimate is  $\mu = 2.638158530323 \pm 2 \times 10^{-12}$ , a result I obtained based on extensive series of Jensen [20].

In fact it is believed that, for dimensionality  $d > 1$  and  $d \neq 4$ ,

$$c_n \sim \text{const.} \times \mu^n n^g .$$

The *critical exponent*  $g$  is believed to depend on the dimension, but not on the details of the lattice. In particular, it is predicted to be a rational number, namely  $11/32$ , in two dimensions. In three dimensions, the best estimate we have is  $g = 0.156957 \pm 0.000009$  given by Clisby [9]. There is no reason to believe that this number is rational.

Despite these accurate estimates, we still cannot even prove the existence of this exponent for  $d < 5$ , let alone establish its value rigorously. For  $d > 4$  the higher dimensionality means that the self-avoiding restriction is less confining than in lower dimensions, and indeed has no effect on the dominant asymptotic behaviour, with the result that the SAW behaves as a random walk. More precisely, Hara and Slade [18, 17] have proved that  $g = 0$  in this case, and that the scaling limit is Brownian motion. In four dimensions the above expression for  $c_n$  must be modified by an additional multiplicative factor  $(\log n)^{1/4}$ , with  $g = 0$ . The appropriately rescaled walk is also expected to have Brownian motion as its scaling limit. These assertions for the four-dimensional case are believed to be true, but no proof exists. Bounds established 50 years ago by Hammersley and Welsh [16] have hardly been improved upon. They proved that, for SAW in dimensionality  $d \geq 2$ ,

$$\mu^n \leq c_n \leq \mu^n e^{\kappa \sqrt{n}} .$$

The lower bound follows immediately from sub-additivity, while the upper bound depends on an *unfolding* of the walk. The number of possible unfoldings can be bounded by the number of partitions of the integer  $n$ , which has the exponential behaviour given above. Note that the existence of a critical exponent implies behaviour  $\mu^n e^{\kappa \log n}$ , which is rather far from the upper bound. A year later, Kesten [23] slightly improved the upper bound to

$$c_n \leq \mu^n e^{\kappa n^{2/(d+2)} \log n} .$$

### 2.1.2. How large is a typical self-avoiding walk?

Another important measure of SAW is the average size of a SAW of length  $n$ , taken uniformly at random. The most common measure is the mean-square end-to-end distance, which is believed to behave as

$$\mathbb{E}_n(|w_n|^2) \sim \text{const.} n^{2\nu} ,$$

(for lattices in dimensions other than 4), where  $\nu$  is another critical exponent. Again, its existence has not been proved for  $d < 5$ , but it is accepted that for two-dimensional lattices  $\nu = 3/4$ . In three dimensions the best numerical estimate is  $\nu = 0.587597 \pm 0.000007$  [7]. In four dimensions it is believed that  $\nu = 1/2$ , and again one expects a multiplicative factor  $(\log n)^{1/4}$ . Finally, for  $d > 4$  it has been proved [18] that  $\nu = 1/2$ . Rigorous results about  $\mathbb{E}_n(|w_n|^2)$  are almost non-existent. It would seem intuitively obvious that

$$cn \leq \mathbb{E}_n(|w_n|^2) \leq Cn^{2-\epsilon} ,$$

but only this year, in an important calculation, has the upper bound been proved by Duminil-Copin and Hammond [11] for two-dimensional SAW.

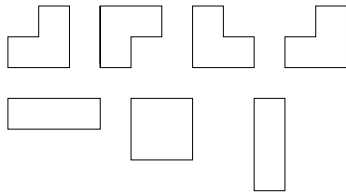


Fig. 3. All seven 8-sided polygons on the square lattice

## 2.2. Self-avoiding polygons

If the end-point of a SAW is adjacent to the origin, an additional step joining the end-point to the origin will produce a self-avoiding circuit. If we ignore knowledge of the origin, and distinguish circuits only by their shape, we refer to *self-avoiding polygons* (SAP). On the square lattice, the first nonzero embedding of a SAP is the unit square, of perimeter 4 and area 1. There are two 6-sided polygons of area 2, and seven 8-sided polygons, shown in Fig. 3, one of which has area 4, and six of which have area 3.

Clearly, SAPs are a subset of SAWs. They are a particularly interesting subset for at least two reasons. Because the conjectured exponents for SAPs (discussed below) are integers or half-integers (which is not the case for SAWs), it is hoped that this means the underlying solution for the SAP case is simpler. Secondly, by including a second parameter, that of area, SAPs can be used to model a range of biological phenomena, such as cell inflation and collapse [12].

Denote by  $p_m$  the number of SAPs of perimeter  $m$ , by  $a_n$  the number of SAPs of area  $n$ , and by  $p_{m,n}$  the number of SAPs of perimeter  $m$  and area  $n$ . We can define two single variable generating functions, for perimeter and area<sup>a</sup> respectively, and a two-variable generating function, as follows:

$$P(x) = \sum_m p_m x^m$$

$$A(q) = \sum_n a_n q^n$$

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

Hammersley [15] proved that the number of SAPs, like SAWs, grows exponentially; more precisely

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

While it is far from obvious, Hammersley also proved that the growth constants  $\mu$  that arise in the polygon case and the walk case are identical.

<sup>a</sup>Clearly the *area* of a polygon is a concept peculiar to the two-dimensional case.

While unproved, a much stronger result is widely believed to hold, namely that

$$p_m \sim \text{const.} \times \mu^m m^{\alpha-3} \quad (1)$$

where  $\alpha$  is a *critical exponent*.<sup>b</sup> The exponent  $\alpha$  is related to the exponent  $\nu$  defined above through the hyper-scaling relation  $d\nu = 2 - \alpha$ . This equation has not been proved, but follows from physical arguments, and of course the assumption that the exponents exist. It therefore follows from the result for  $\nu$  quoted above that in three dimensions  $\alpha = 0.237209 \pm 0.000021$ .

For polygons there is a second growth constant, and exponent, associated with the area generating function. By concatenation arguments it can be readily proved that

$$\lambda = \lim_{n \rightarrow \infty} a_n^{1/n}$$

exists. It is also generally accepted, but not proved, that

$$a_n \sim \text{const.} \times \lambda^n n^\tau.$$

Unfortunately we only have numerical estimates of  $\lambda$  and  $\tau$  [14]. However for two-dimensional lattices  $\tau$  is believed to be  $-1$ , corresponding to a logarithmic singularity of the generating function. That is to say,

$$A(q) \sim \text{const.} \times \log(1 - \lambda q),$$

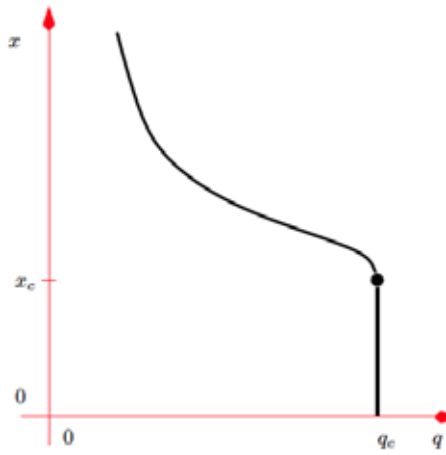
so that  $a_n \sim \text{const} \times \lambda^n / n$ .

Of great interest is the two-variable generating function  $\mathcal{P}(x, y)$ . From this, we can define the free energy

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

It has been proved [12] 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  $\mathcal{P}(x, q)$ , 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. For  $0 < q < 1$ , the line  $x_c(q)$  is believed to be a line of logarithmic singularities of the generating function  $\mathcal{P}(x, q)$ . The line  $q = 1$ , for  $0 < x < x_c(1)$  is believed to be a line of *finite* essential singularities [12]. At the point  $(x_c, 1)$  we have more complicated behaviour, and this point is called a tricritical point.

<sup>b</sup>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 cubic lattices, polygons of all perimeters greater than two are embeddable, so Eq. (1) stands as stated.



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

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

Here the superscript *(sing)* means the singular part. There is an additional, additive part that is regular in the neighbourhood of  $(x_c, 1)$ .

For self-avoiding polygons, in a series of papers, Richard and co-authors [33, 34, 32] have provided abundant evidence (but no proof) for the surprisingly strong conjecture that

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

Here  $\text{Ai}(x)$  is the Airy function, and  $C(q)$  is a function, independent of  $x$ , that arises as a constant of integration. The exponents are believed to be  $\phi = 2/3$  and  $\theta = 1$ . Here  $A_0$  is a constant known only numerically.

Finally, a veritable treasure trove of rigorous results would be unlocked if we could prove that, in the large size limit, more precisely the scaling limit, that two-dimensional random SAWs are described by one of the SLE processes (Schramm–Loewner Evolution), which, in the past 20 years, have been proved to describe the limit of several discrete models in combinatorics and statistical physics. Indeed, Lawler, Schramm and Werner proved that, if the scaling limit of SAWs exists and is conformally invariant, then this limit has to be  $\text{SLE}_{8/3}$ . This has been checked via simulations by Kennedy [22]. This would in particular imply the conjectured values of the exponents  $g$  and  $\nu$  in the two-dimensional case. We explain this in more detail in the conclusion.

In the next section we give the proof due to Duminil-Copin and Smirnov of the exact growth constant for the honeycomb lattice. In the following section we give three examples of applications of SAWs to other areas of science, and in the conclusion we give more detail of recent developments that we hope point the way to future breakthroughs.

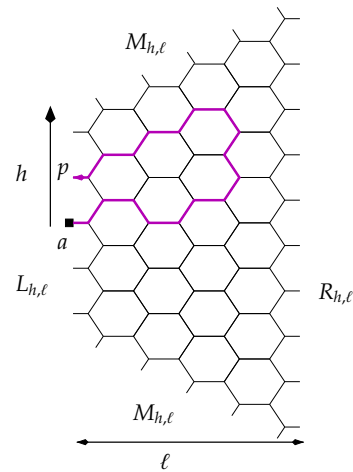


Fig. 4. A trapezoid  $\mathcal{T}$  on the honeycomb lattice

### 3. The Honeycomb Lattice

As mentioned above, a breakthrough was achieved in 2010, when Duminil-Copin and Smirnov [10] proved that the growth constant on the honeycomb lattice is  $\mu = \sqrt{2 + \sqrt{2}}$ , as predicted by Nienhuis [30], using compelling physical arguments from conformal field theory, 30 years previously. The argument is, in hindsight, so simple, and the result so important, that we sketch it here.

We consider SAWs that start from a point  $a$  located on the left side of a trapezoid  $\mathcal{T}$  of width  $\ell$  and height  $h$ , as shown in Fig. 4. For  $p$  a mid-edge of  $\mathcal{T}$ , let  $F(p)$  be the generating function of SAW  $w$  that end at  $p$ , weighted by the number of vertices  $v(w)$  and the number of turns  $T(w)$  (a left turn counts +1, a right turn -1):

$$F(p) \equiv F(p; x, \alpha) := \sum_{w: a \rightsquigarrow p} x^{v(w)} e^{i\alpha T(w)}.$$

For instance, the walk of Fig. 4 visits 17 vertices, makes 10 left turns and 7 right turns, so that its contribution to  $F(p)$  is  $x^{17} e^{3i\alpha}$ . Then, if  $v$  is any vertex of  $\mathcal{T}$  and  $p_1, p_2, p_3$  are the three mid-edges adjacent to it, the following local identity holds:

$$(p_1 - v)F(p_1) + (p_2 - v)F(p_2) + (p_3 - v)F(p_3) = 0, \quad (2)$$

provided  $x = x_c := 1/\sqrt{2 + \sqrt{2}}$ , which is the reciprocal of the conjectured growth constant, and  $\alpha = -5\pi/24$ . (We consider that the honeycomb lattice is embedded in the complex plane  $\mathbb{C}$ , so that  $p_i - v$  is a complex number). This identity is easily proved by grouping as pairs or triplets the SAWs that contribute to its left-hand side, as depicted in Fig. 5. One then checks that the contribution of each group is zero.

If we now sum (2) over all vertices  $v$  of  $\mathcal{T}$ , then due to the terms  $(p_i - v)$ , all terms  $F(p)$  such that  $p$  is not a mid-edge of the border disappear. After

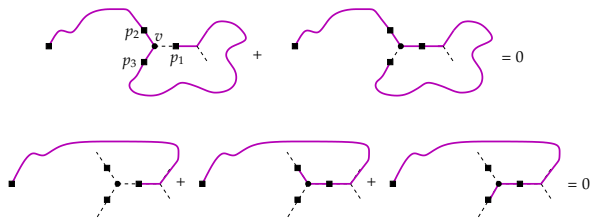


Fig. 5. A very local proof of the local identity (2)

a few more reductions based on symmetries, one is left with

$$\left(\cos \frac{3\pi}{8}\right) L_{h,\ell}(x_c) + \frac{1}{\sqrt{2}} M_{h,\ell}(x_c) + R_{h,\ell}(x_c) = 1,$$

where  $L_{h,\ell}(x)$  (resp.  $R_{h,\ell}(x)$ ,  $M_{h,\ell}(x)$ ) is the generating function of SAWs  $w$  that end on the left side (resp. right side, top or bottom) of  $\mathcal{T}$ , weighted by the number of vertices  $v(w)$ .

By letting  $h$  and then  $\ell$  tend to infinity, Duminil-Copin and Smirnov derived from this identity that the generating function of SAWs diverges at  $x_c$ , but converges when  $x < x_c$ . This means that its radius of convergence is  $x_c$ , so that the growth constant is  $1/x_c = \sqrt{2 + \sqrt{2}}$ .

Unfortunately these ideas do not generalise to SAW on the square or triangular lattices, for which we only have accurate numerical estimates for the growth constant  $\mu$ .

#### 4. Applications

One reason that SAWs and SAPs are so extensively studied, apart from their intrinsic mathematical interest, is that they model many problems that arise in other fields. The first such example we will consider extends the proof given above to the situation where the SAW can interact with a surface. The second example considers SAWs crossing a square, with application to telecommunication networks, and the third example models some recent biological experiments where strands of DNA (a polymer) are pulled from a wall with optical tweezers.

##### 4.1. Walks attached to a surface

The interaction of polymers with a surface is scientifically and industrially an important phenomenon. A common example is the adherence of paint to a surface, clearly an industrial process of considerable significance. To model such phenomena requires the inclusion of an interaction term between the polymer and the surface. To achieve this, we add a weight  $y$  to vertices in the surface, as shown in Fig. 6. In physics terms,  $y = e^{-\epsilon/k_B T}$  where  $\epsilon$  is the energy associated with a surface vertex,  $T$  is the absolute temperature and  $k_B$  is Boltzmann's constant. It is known that the growth

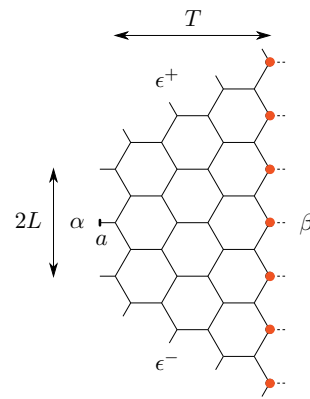


Fig. 6. Trapezoidal domain  $S_{T,L}$  with vertices on the right-hand side wall, shown in bold, carrying a weight  $y$

constant  $\mu = 1/x_c$  for such walks is the same as for the bulk case.

Let  $c_n^+(i)$  be the number of half-plane walks of  $n$ -steps, with  $i$  monomers in the surface, and define the partition function (or generating function) as

$$C_n^+(y) = \sum_{i=0}^n c_n^+(i) y^i.$$

If  $y$  is large, the polymer adsorbs onto the surface, while if  $y$  is small, the walk is repelled by the surface.

**Proposition 1.** For  $y > 0$ ,

$$\mu(y) := \lim_{n \rightarrow \infty} C_n^+(y)^{1/n}$$

exists and is finite. It is a log-convex, non-decreasing function of  $\log y$ , and therefore continuous and almost everywhere differentiable.

For  $0 < y \leq 1$ ,

$$\mu(y) = \mu(1) \equiv \mu.$$

Moreover, for any  $y > 0$ ,

$$\mu(y) \geq \max(\mu, \sqrt{y}).$$

This behaviour implies the existence of a critical value  $y_c$ , with  $1 \leq y_c \leq \mu^2$ , which delineates the transition from the desorbed phase to the adsorbed phase:

$$\mu(y) \begin{cases} = \mu & \text{if } y \leq y_c, \\ > \mu & \text{if } y > y_c. \end{cases}$$

In 1995 Batchelor and Yung [1] extended Nienhuis's [30] work to the adsorption problem just described, and making similar assumptions to Nienhuis conjectured the value of the critical surface fugacity for the honeycomb lattice SAW model, to be  $y_c = 1 + \sqrt{2}$ . In 2012 this was proved by Beaton, Bousquet-Mélou, de Gier, Duminil-Copin and Guttmann [3], and here we will sketch their proof.

Take the same trapezoid as above, now called  $S_{T,L}$ , and add weights to the vertices on the  $\beta$



boundary, as shown in bold in the figure above: Then we find the corresponding identity between generating functions, with  $y^* = 1 + \sqrt{2}$ , to be

$$1 = \cos\left(\frac{3\pi}{8}\right)A_{T,L}(x_c, y) + \cos\left(\frac{\pi}{4}\right)E_{T,L}(x_c, y) + \frac{y^* - y}{y(y^* - 1)}B_{T,L}(x_c, y)$$

where  $y$  is conjugate to the number of visits to the  $\beta$  boundary. The generating functions  $A_{T,L}$ ,  $B_{T,L}$  and  $E_{T,L}$  are two-variable generalisations of those defined in the previous section. To prove the conjecture we need to show that  $y_c = y^*$ .

It is safe to take  $L \rightarrow \infty$  so that the trapezoid becomes a strip. The identity then becomes

$$1 = \cos\left(\frac{3\pi}{8}\right)A_T(x_c, y) + \cos\left(\frac{\pi}{4}\right)E_T(x_c, y) + \frac{y^* - y}{y(y^* - 1)}B_T(x_c, y).$$

It is then straightforward to prove that

- (i)  $E_T(x_c, y) = 0$  for  $0 \leq y < y^*$ ,
- (ii)  $y_c \geq y^*$ ,
- (iii)  $\lim_{T \rightarrow \infty} A_T(x_c, y) = A(x_c, y) = A(x_c)$  is constant for  $0 \leq y < y^*$ .

If we now write

$$\cos(3\pi/8)A(x_c, y) = 1 - \delta$$

then the above identity reduces to

$$B(x_c, y) = \lim_{T \rightarrow \infty} B_T(x_c, y) = \frac{\delta y(y^* - 1)}{y^* - y}$$

and in particular

$$B(x_c, 1) = \delta.$$

**Proposition 1.** *If  $\delta = 0$  then  $y_c = y^*$ .*

The proof uses a decomposition of  $A$  walks in a strip of width  $T$  into  $B$  walks in that same strip, and gives rise to an inequality. In particular, for  $y < y_c = \lim_{T \rightarrow \infty} y_T$ ,

$$0 \leq \alpha x_c + \frac{1}{B_T(x_c, 1)} \frac{y^* - y}{y(y^* - 1)}.$$

If  $B_T(x_c, 1)$  tends to 0, this forces  $y^* \geq y_c$ , otherwise the right-hand side would become arbitrarily large in modulus and negative as  $T \rightarrow \infty$  for  $y^* < y < y_c$ .

Together with  $y^* \leq y_c$ , this establishes  $y_c = y^* = 1 + \sqrt{2}$  and completes the proof of the proposition.

The proof that  $\delta = 0$  is complicated, and unlike most other proofs we have given is almost totally probabilistic. It is unrealistic to give any details, but in essence one first uses renewal theory to show that  $\delta^{-1}$  is the expected height of an *irreducible bridge*, which is a SAW that crosses the strip from left to right, and cannot be expressed

as the concatenation of two or more smaller such bridges. Next one shows that, for irreducible bridges,  $\mathbb{E}[\text{width}] < \infty$  implies that  $\mathbb{E}[\text{height}] < \infty$ . Finally one shows that the assumption that  $\mathbb{E}[\text{height}] < \infty$  leads to a contradiction, from which the desired result that  $\delta = 0$  readily follows.

### 4.2. Walks crossing a square

Some years ago I was asked by a telecommunications engineer to help him with the following problem: His company had a square grid of nodes, connected by wires, and phone-calls could be routed from the bottom left-hand corner to the top right-hand corner of the grid. He wished to know how many such routes there were, as this determined the carrying capacity of the network.

After some discussion we agreed that this was simply the question “how many distinct SAWs are there on a square grid of side-length  $L$  originating at  $(0,0)$  and ending at  $(L,L)$ ?” The problem as stated was first considered by Knuth [27] in 1976, who gave a Monte Carlo estimate for the number of paths for  $L = 10$ , a result we now know exactly. The problem was generalised by Whittington and the author [38] to include a weight  $x$  associated with each step of the walk. This gives rise to a canonical model of a phase transition. For  $x < 1/\mu$  the average length of a SAW grows as  $L$ , while for  $x > 1/\mu$  it grows as  $L^2$ . Here  $\mu$  is the growth constant of unconstrained SAWs on the square lattice, defined above. For  $x = 1/\mu$  numerical evidence, but no proof, was given that the average walk length grows as  $L^{4/3}$ . Let  $c_n(L)$  denote the number of walks of length  $n$ . Clearly  $c_n(L) = 0$  for  $n < 2L$ . We denote the generating function by  $C_L(x) := \sum_n c_n(L)x^n$ . The answer to the original question is  $\sum_n c_n(L)$ .

Subsequently, Madras [28] proved a number of relevant results. In fact, most of Madras’s results were proved for the more general  $d$ -dimensional hypercubic lattice, but here we will quote them in the more restricted two-dimensional setting.

**Theorem 2.** *The following limits,*

$$\mu_1(x) := \lim_{L \rightarrow \infty} C_L(x)^{1/L} \quad \text{and} \quad \mu_2(x) := \lim_{L \rightarrow \infty} C_L(x)^{1/L^2},$$

*are well-defined in  $\mathbb{R} \cup \{+\infty\}$ .*

*More precisely,*

- (i)  $\mu_1(x)$  is finite for  $0 < x \leq 1/\mu$ , and is infinite for  $x > 1/\mu$ . Moreover,  $0 < \mu_1(x) < 1$  for  $0 < x < 1/\mu$  and  $\mu_1(1/\mu) = 1$ .
- (ii)  $\mu_2(x)$  is finite for all  $x > 0$ . Moreover,  $\mu_2(x) = 1$  for  $0 < x \leq 1/\mu$  and  $\mu_2(x) > 1$  for  $x > 1/\mu$ .

In [38] the existence of the limit  $\mu_2(x)$  was proved, and in addition upper and lower bounds on  $\mu_2(x)$  were established.

The average length of a (weighted) walk is defined to be

$$\langle n(x, L) \rangle := \sum_n n c_n(L) x^n / \sum_n c_n(L) x^n. \quad (3)$$

We define  $\Theta(x)$  as follows: Let  $a(x)$  and  $b(x)$  be two functions of some variable  $x$ . We write that  $a(x) = \Theta(b(x))$  as  $x \rightarrow x_0$  if there exist two positive constants  $\kappa_1$  and  $\kappa_2$  such that, for  $x$  sufficiently close to  $x_0$ ,

$$\kappa_1 b(x) \leq a(x) \leq \kappa_2 b(x).$$

**Theorem 3.** For  $0 < x < 1/\mu$ , we have that  $\langle n(x, L) \rangle = \Theta(L)$  as  $L \rightarrow \infty$ , while for  $x > 1/\mu$ , we have  $\langle n(x, L) \rangle = \Theta(L^2)$ .

In [38] it was proved that  $\langle n(1, L) \rangle = \Theta(L^2)$ . The situation at  $x = 1/\mu$  is unknown. In [5] we gave compelling numerical evidence that in fact  $\langle n(1/\mu, L) \rangle = \Theta(L^{1/\nu})$ , where  $\nu = 3/4$ , in accordance with an intuitive suggestion of Madras in [28].

**Theorem 4.** For  $x > 0$ , define  $f_1(x) = \log \mu_1(x)$  and  $f_2(x) = \log \mu_2(x)$ .

- (i) The function  $f_1$  is a strictly increasing, negative-valued convex function of  $\log x$  for  $0 < x < 1/\mu$ .
- (ii) The function  $f_2$  is a strictly increasing, convex function of  $\log x$  for  $x > 1/\mu$ , and satisfies  $0 < f_2(x) \leq \log \mu + \log x$ .

Some, but not all of the above results were previously proved in [38], but these three theorems elegantly capture all that is rigorously known. In [5] an extensive numerical study was described, including exact enumerations up to squares of side 19. For the largest square there are exactly 1 523 344 971 704 879 993 080 742 810 319 229 690 899 454 255 323 294 555 776 029 866 737 355 060 592 877 569 255 844 distinct paths! The number of such paths, as we have seen, grows as  $\lambda^{L^2}$ . In [5] it was also proved that  $1.628 < \lambda < 1.782$  and estimated that  $\lambda = 1.744550 \pm 0.000005$ .

### 4.3. Pulling a polymer from a wall

During the past decade, force has been used as a thermodynamic variable to understand molecular interactions and their role in the structure of biomolecules [35, 21, 37]. By exerting a force in the picoNewton range, one aims to experimentally study and characterise the elastic, mechanical, structural and functional properties of biomolecules [6].

In [24] SAWs were used to model the situation in which a polymer is attached to a surface and pulled from that surface by an applied force. The situation is shown in Fig. 7. Interactions are introduced between neighbouring monomers on the lattice that are not adjacent along the chain. The pulling force is modelled by introducing an energy proportional to the  $x$ -component of the

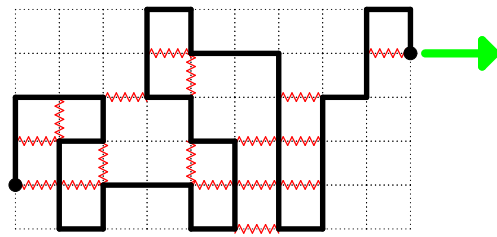


Fig. 7. An interacting self-avoiding walk on the square lattice with one end attached to a surface and subject to a pulling force at the other end. Each step of the walk connecting a pair of monomers is indicated by a thick solid line while interactions between non-bonded nearest neighbour monomers are indicated by jagged lines

end-to-end distance. One end of the polymer is attached to an impenetrable surface while the polymer is being pulled from the other end with a force acting along the  $x$ -axis.

Boltzmann weights  $\omega = \exp(-\epsilon/k_B T)$  and  $u = \exp(-F/k_B T)$  conjugate to the nearest neighbour interactions and force, respectively, were introduced, where  $\epsilon$  is the interaction energy,  $k_B$  is Boltzmann's constant,  $T$  the temperature and  $F$  the applied force. For simplicity, we set  $\epsilon = -1$  and  $k_B = 1$ . The relevant finite-length partition functions are

$$Z_N(F, T) = \sum_{\text{all walks}} \omega^m u^x = \sum_{m,x} C(N, m, x) \omega^m u^x, \quad (4)$$

where  $C(N, m, x)$  is the number of interacting SAWs of length  $N$  having  $m$  nearest neighbour contacts and whose end-points are a distance  $x = x_N - x_0$  apart. The partition functions of the constant force ensemble,  $Z_N(F, T)$ , and constant distance ensemble,  $Z_N(x, T) = \sum_m C(N, m, x) \omega^m$ , are related by  $Z_N(F, T) = \sum_x Z_N(x, T) u^x$ . The free energies are evaluated from the partition functions

$$G(x) = -T \log Z_N(x, T) \quad \text{and} \quad G(F) = -T \log Z_N(F, T). \quad (5)$$

Here  $\langle x \rangle = \frac{\partial G(F)}{\partial F}$  and  $\langle F \rangle = \frac{\partial G(x)}{\partial x}$  are the control parameters of the constant force and constant distance ensembles, respectively.

All possible conformations of the SAW were enumerated. The challenge facing exact enumerations is to increase the chain length. Using direct counting algorithms the time required to enumerate all the configurations increases as  $\mu^N$ , where  $\mu$  is the connective constant of the lattice ( $\mu \approx 2.638$  on the square lattice). So even with a rapid increase in computing power only a few more terms can be obtained each decade. In [24] the number of interacting SAWs was calculated using transfer matrix techniques [19]. Combined with parallel processing these algorithms allowed the enumerations to be extended to chain lengths up to 55 steps, roughly doubling the previously available enumerations.

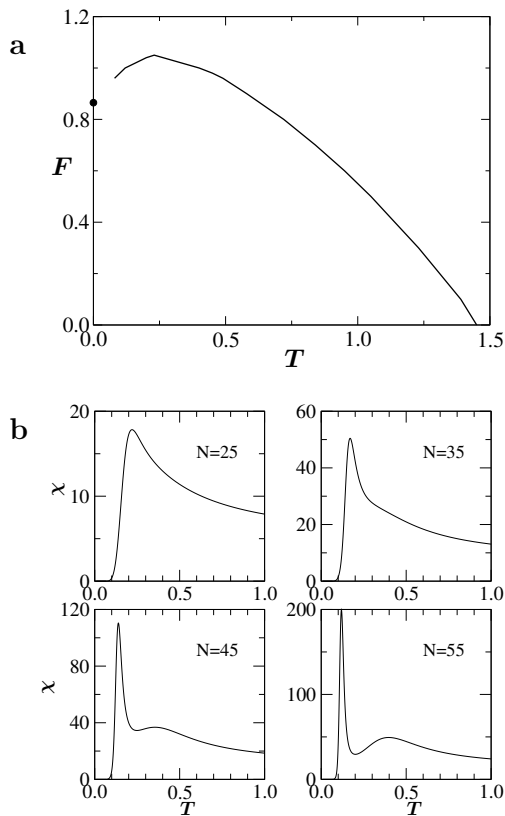


Fig. 8. (a) The force-temperature phase diagram for flexible chains. (b) The fluctuations  $\chi$  in the number of contacts vs. temperature at fixed force  $F = 1.0$  for various values of the chain length  $N$

In Fig. 8(a) we show the force-temperature phase diagram for flexible chains. At low temperature and force the polymer chain is in the collapsed state and as the temperature is increased (at fixed force) the polymer chain undergoes a phase transition to an extended state. The transition temperature as plotted in Fig. 8(a) was found (in the thermodynamic limit  $N \rightarrow \infty$ ) by studying the reduced free energy per monomer. The most notable feature of the phase-diagram is the *re-entrant* behaviour — that is to say, the initial increase of force with temperature, prior to a decrease.

The positive slope  $dF_c/dT$  at  $T = 0$  confirms the existence of re-entrance in the  $F - T$  phase-diagram. The authors pointed out that the value of the transition temperature obtained in the thermodynamic limit and the one obtained from the fluctuations in non-bonded nearest neighbours (which can be calculated exactly for finite  $N = 55$ ) gives almost the same value (within error bars of  $\pm 0.01$ ). The fluctuations are defined as  $\chi = \langle m^2 \rangle - \langle m \rangle^2$ , with the  $k$ 'th moment given by

$$\langle m^k \rangle = \sum_{m,x} m^k C(N, m, x) \omega^m u^x \bigg/ \sum_{m,x} C(N, m, x) \omega^m u^x .$$

In the panels of Fig. 8(b) the emergence of two peaks in the fluctuation curves with increasing  $N$

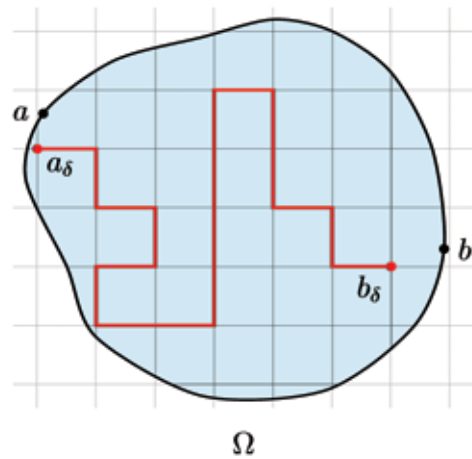


Fig. 9. Discretisation of domain  $\Omega$

at fixed force  $F = 1.0$  are shown. The twin-peaks reflect the fact that in the re-entrant region as one increases  $T$  (with  $F$  fixed) the polymer chain undergoes two phase transitions. The importance of powerful enumeration data is highlighted by the observation that the twin-peaks are not apparent for small values of  $N$ . Many more details and comparison with experiments are given in [24] — our purpose here is just to show the applicability of SAWs to this problem.

## 5. Conclusion

### 5.1. The scaling limit

One topic we have failed to adequately address is the question of the *scaling limit* of SAW. An intuitive grasp of this concept can perhaps be gained by looking at the first two figures in this article. In the first figure, the effect of the lattice is clear. In the second figure, there is no obvious lattice, and indeed no way to tell that this is not a continuous curve. We formalise this notion as follows: Consider a smooth (enough) closed domain  $\Omega$ , with an underlying square grid, with grid spacing  $\delta$  as shown in Fig. 9. Denote by  $\Omega_\delta$  that portion of the grid contained in  $\Omega$ . Take two distinct points on the boundary of  $\Omega$  labelled  $a$  and  $b$ . Now take the nearest lattice vertex to  $a$ , and label it  $a_\delta$ , and similarly  $b_\delta$  is the label of the nearest lattice vertex to point  $b$ . Consider the set of SAWs  $\omega(\Omega_\delta)$  on the finite domain  $\Omega_\delta$  from  $a_\delta$  to  $b_\delta$ . Recall that  $\delta > 0$  sets the scale of the grid. Now let  $|\omega|$  be the length of a walk  $\omega_\delta \in \omega(\Omega_\delta)$ , and denote the weight of the walk by  $x^{|\omega|}$ . The reason for this is that the walks are of different lengths, making the uniform measure not particularly natural. (There is also a normalising factor, which for simplicity we ignore.)

As we let  $\delta \rightarrow 0$  we expect the behaviour of the walk to depend on the value of  $x$ . For  $x < x_c$  it is possible to prove that  $\omega_\delta$  goes to a straight



line as  $\delta \rightarrow 0$ . (Strictly speaking it converges in distribution to a straight line, with fluctuations  $O(\sqrt{\delta})$ .) For  $x > x_c$  it is expected that  $\omega_\delta$  becomes (again, in distribution) space-filling as  $\delta \rightarrow 0$ . But at  $x = x_c$  it is *conjectured* that  $\omega_\delta$  becomes (in distribution) a random continuous curve, and is conformally invariant. This describes the scaling limit. If this conjecture is correct, a second, pivotal, conjecture by Lawler, Schramm and Werner [25] is that this random curve converges to  $\text{SLE}_{8/3}$  from  $a$  to  $b$  in the domain  $\Omega$ . These two conjectures must be considered the principal open problems in the field. If they could be proved, the existence and value of the critical exponents, as predicted by conformal field theory for two-dimensional walks would be proved.

## 5.2. Schramm–Löwner Evolution

For an approachable discussion of  $\text{SLE}_\kappa$ , see Chapter 15 of [14]. Here we give a very minimal outline. Let  $\mathbb{H}$  denote the upper half-plane. Consider a path  $\gamma$  starting at the boundary and finishing at an internal vertex. Then  $\mathbb{H} \setminus \gamma$  is the complement of this path, and is a slit upper half-plane. It follows from the Riemann Mapping Theorem that it can be conformally mapped to the upper half-plane. Löwner [26] discovered that by specifying the map so that it approaches the identity at infinity, the conformal map so described (actually a family of maps, appropriate to each point on the curve) satisfies a simple differential equation, called the Löwner equation. The mapping can alternatively be defined by a real function. This observation led Schramm to apply the Löwner equation to a conformally invariant measure for planar curves. That is to say, the Löwner equation generates a set of conformal maps, driven by a continuous real-valued function. Schramm's profound insight was to use Brownian motion  $B_t$  as the driving function.<sup>c</sup> So let  $B_t$ ,  $t \geq 0$  be standard Brownian motion on  $\mathbb{R}$  and let  $\kappa$  be a real parameter. Then  $\text{SLE}_\kappa$  is the family of conformal maps  $g_t : t \geq 0$  defined by the Löwner equation

$$\frac{\partial}{\partial t} g_t(z) = \frac{2}{g_t(z) - \sqrt{\kappa} B_t}, \quad g_0(z) = z.$$

This is actually called *chordal*  $\text{SLE}_\kappa$  as it describes paths growing from the boundary and ending on the boundary. If  $\kappa \leq 4$  then the path is almost surely a simple curve, in the upper half plane. Larger values of  $\kappa$  lead to more complicated behaviour.

Hopefully this rather vague description will convey the flavour of this exciting and powerful development in studying not just two-dimensional SAWs, but a variety of other processes, such as percolation, the random cluster

model, and the Ising model. We refer the reader to [4] for greater detail of both SLE and these applications. Despite these remarkable advances, we still have no real idea how to obtain comparable results for the 3-dimensional model.<sup>d</sup>

In this article I have only scratched the surface of this topic. More details on the mathematical aspects can be found in [29] and the recent reviews [2, 4]. More information on numerical aspects and some applications, particularly to the SAP subset can be found in the monograph [14]. Another approach to this problem that has not been discussed is to simplify the problem so that it can be solved (see [14, Chap. 3]). Unfortunately most such simplifications involve rendering the model Markovian, which removes a significant feature.

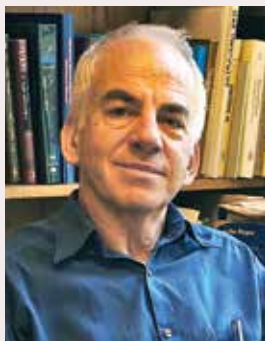
## References

- [1] M. T. Batchelor and C. M. Yung, Exact results for the adsorption of a flexible self-avoiding polymer chain in two dimensions, *Phys. Rev. Lett.* **74** (1995) 2026–2029.
- [2] R. Bauerschmidt, H. Duminil-Copin, J. Goodman and G. Slade, Lectures on self-avoiding walks, arXiv:1206.2092, Clay Mathematics Institute Summer School, Buzios (2010).
- [3] N. R. Beaton, M. Bousquet-Mélou, H. Duminil-Copin, J. de Gier and A. J. Guttmann, The critical fugacity for surface adsorption for SAW on the honeycomb lattice is  $1 + \sqrt{2}$ , arXiv:1109.0358v3 (2012).
- [4] V. Beffara, Schramm–Löwner Evolution and other conformally invariant objects, Clay Mathematics Institute Summer School, Buzios (2010).
- [5] M. Bousquet-Mélou, A. J. Guttmann and I. Jensen, Self-avoiding walks crossing a square, *J. Phys. A: Math. Gen.* **34** (2005) 9159–9181.
- [6] C. Y. Bustamante, Y. R. Chemla, N. R. Forde and D. Izhaky, Mechanical processes in biochemistry, *Annu. Rev. Biochem.* **73** (2004) 705–748.
- [7] N. Clisby, Accurate estimate of the critical exponent  $\nu$  for self-avoiding walks via a fast implementation of the pivot algorithm, arXiv:1002.0494, *Phys. Rev. Lett.* **104** (2010) 055702.
- [8] N. Clisby, Efficient implementation of the pivot algorithm for self-avoiding walks, arXiv:1005.1444, *J. Stat. Phys.* **140** (2010) 349–392.
- [9] N. Clisby, Private communication. Publication in preparation.
- [10] H. Duminil-Copin and S. Smirnov, The connective constant of SAW on the honeycomb lattice is  $\sqrt{2 + \sqrt{2}}$ , *Ann. Math.* **175** (2012) 1653–1665.
- [11] H. Duminil-Copin and A. Hammond, Self-avoiding walk is sub-ballistic, arXiv:1205.0401v1.
- [12] M. E. Fisher, A. J. Guttmann and S. G. Whittington, Two-dimensional lattice vesicles and polygons, *J. Phys. A: Math. Gen.* **24** (1991) 3095–3106.
- [13] P. J. Flory, The configuration of real polymer chains, *J. Chem. Phys.* **17** (1949) 303–310.
- [14] A. J. Guttmann (ed.) *Polygons, Polyominoes and Polycubes*, Lecture Notes in Physics **775** (Springer, The Netherlands, 2009).

<sup>c</sup>It is the only process compatible with both conformal invariance and the so-called domain Markov property.

<sup>d</sup>This is also true of other classical models, such as the Ising model, the Potts model and percolation.

- [15] J. M. Hammersley, The number of polygons on a lattice, *Proc. Camb. Phil. Soc.* **57** (1961) 516–523.
- [16] J. M. Hammersley and D. J. A. Welsh, Further results on the rate of convergence to the connective constant of the hypercubical lattice, *Quart. J. Math. Oxford* (2) **13** (1962) 108–110.
- [17] T. Hara and G. Slade, The lace expansion for self avoiding walk in five or more dimensions, *Reviews Math. Phys.* **4** (1992) 235–327.
- [18] T. Hara and G. Slade, Self avoiding walk in five or more dimensions I. The critical behaviour, *Comm. Math. Phys.* **147** (1992) 101–136.
- [19] I. Jensen, Enumeration of self-avoiding walks on the square lattice, *J. Phys. A* **37** (2004) 5503–5524.
- [20] I. Jensen, Private communication. Publication in preparation.
- [21] M. S. Z. Kellermayer, S. B. Smith, H. L. Granzier and C. Bustamante, Folding-unfolding transitions in single Titin molecules characterized with laser tweezers, *Science* **276** (1997) 1112–1116.
- [22] T. Kennedy, Conformal invariance and stochastic Loewner evolution predictions for the 2D self-avoiding walk — Monte Carlo tests, *J. Stat. Phys.* **114** (2004) 51–78.
- [23] H. Kesten, On the number of self-avoiding walks, *J. Math. Phys.* **4** (1963) 960–969.
- [24] S. Kumar, I. Jensen, J. L. Jacobsen and A. J. Guttmann, Role of conformational entropy in force-induced bio-polymer unfolding, *Phys. Rev. Lett.* **98** (2007) 128101.
- [25] G. F. Lawler, O. Schramm and W. Werner, On the scaling limit of planar self-avoiding walk, *Proc. Symposia Pure Math.* **72** (2004) 339–364.
- [26] K. Löwner, Untersuchungen über schlichte konforme Abbildungen des Einheitskreises, I. *Math. Ann.* **89** (1923) 103–121.
- [27] D. E. Knuth, *Science* **194** (1976) 1235–1242.
- [28] N. Madras, *J. Phys. A: Math. Gen.* **28** (1995) 1535–1547.
- [29] N. Madras and G. Slade, *The Self-avoiding Walk*, Probability and its Applications (Birkhäuser Boston Inc., Boston, MA, 1993).
- [30] B. Nienhuis, Exact critical exponents of the  $O(n)$  model in two dimensions, *Phys. Rev. Lett.* **49** (1982) 1062–1065.
- [31] W. J. C. Orr, Statistical treatment of polymer solutions at infinite dilution, *Trans. Faraday Soc.* **43** (1947) 12–27.
- [32] C. Richard, Scaling behaviour of two-dimensional polygon models, *J. Stat. Phys.* **108** (2002) 459–493.
- [33] C. Richard, A. J. Guttmann and I. Jensen, Scaling function and universal amplitude combinations for self-avoiding polygons, *J. Phys. A: Math. Gen.* **34** (2001) L495–L501.
- [34] C. Richard, I. Jensen and A. J. Guttmann, Scaling function for self-avoiding polygons revisited, *J. Stat. Mech.: Th. Exp.* (2004) P08007.
- [35] M. Rief, M. Gautel, F. Oesterhelt, J. M. Fernandez and H. E. Gaub, Reversible unfolding of individual Titin immunoglobulin domains by AFM, *Science* **276** (1997) 1109–1112.
- [36] I. Rouzina and V. A. Bloomfield, Force-induced melting of the DNA double helix 1. Thermodynamic analysis, *Biophys. J.* **80** (2001) 882–893; Force-induced melting of the DNA double helix 2. Effect of solution conditions, *Biophys. J.* **80** (2001) 894–900.
- [37] L. Tskhovrebova, J. Trinick, J. A. Sleep and R. M. Simmons, Elasticity and unfolding of single molecules of the giant muscle protein Titin, *Nature* **387** (1997) 308–312.
- [38] S. G. Whittington and A. J. Guttmann, *J. Phys. A: Math. Gen.* **23** (1990) 5601–5609.



## Anthony J Guttmann

The University of Melbourne, Australia  
guttman@unimelb.edu.au

Anthony J Guttmann did his PhD at the University of New South Wales, in Sydney, and went on to postdoctoral work at King's College, London. He is currently a professor at The University of Melbourne, and also the Director of MASCOS.