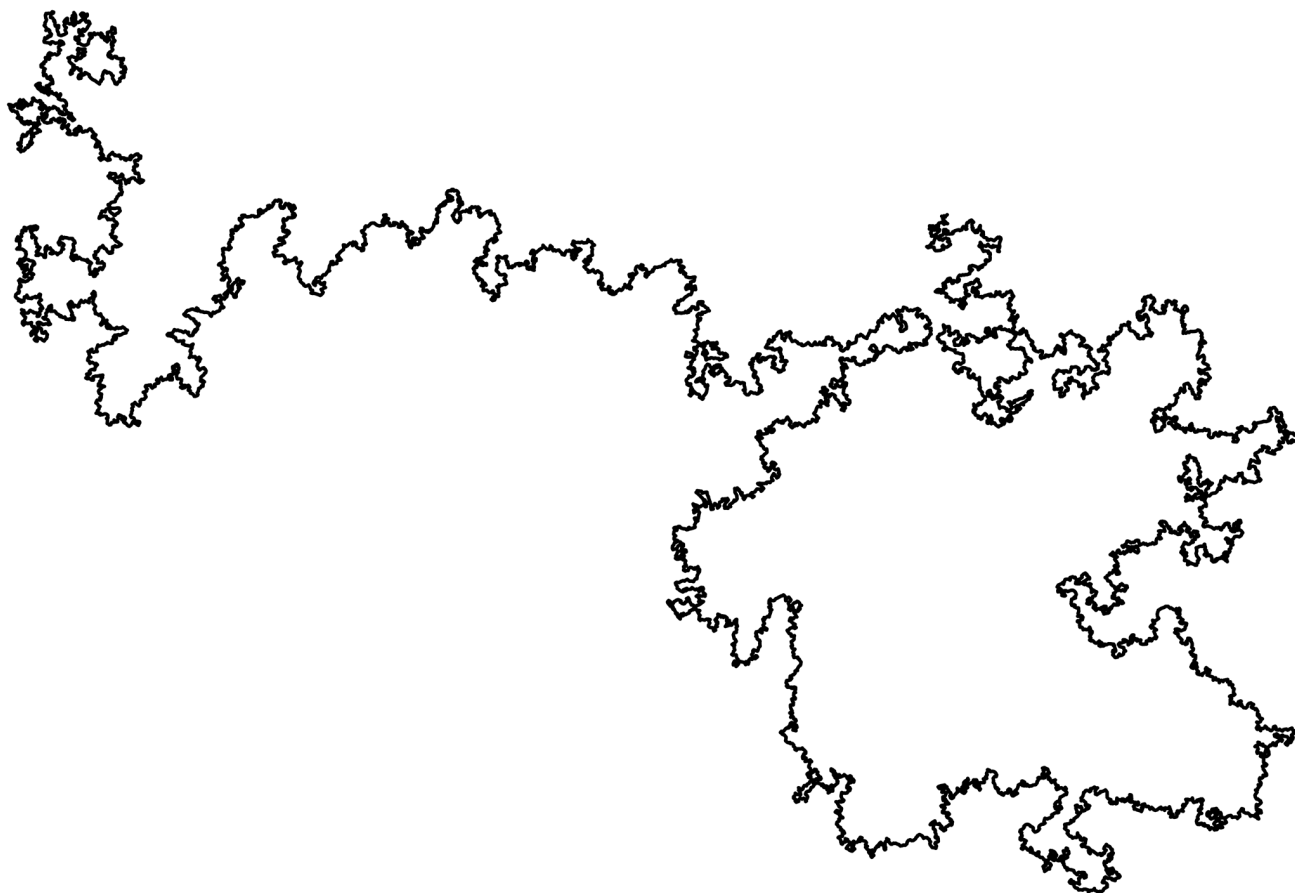

Enumerative Combinatorics of Lattice Polymers



Nathan Clisby

Why Study Lattice Polymers?

There are three principal reasons to study lattice polymers. For physicists (and mathematicians interested in applications), the miraculous phenomenon of “universality” means that *exact* information about real-world polymer systems can be obtained by studying highly idealized models such as lattice polymers. For mathematicians (and

physicists who appreciate mathematical beauty), the physically motivated models are mathematically appealing, and have rich combinatorial structure. The third reason is that it is just a really fun research topic.¹

The most fundamental model is the self-avoiding walk [Flo49, MS93], which is a walk that starts at the origin of a lattice and moves successively to neighboring sites, with the rule that self-intersections are forbidden. An example of a self-avoiding walk on the square lattice is shown in Figure 1, together with a self-avoiding polygon, which is a walk that returns to the origin but is otherwise self-avoiding.

Formally, an n -step self-avoiding walk on \mathbb{Z}^d is a mapping $\omega : \{0, 1, \dots, n\} \rightarrow \mathbb{Z}^d$ with $\omega(0)$ at the origin, steps of

Nathan Clisby is a senior lecturer in the Department of Mathematics at Swinburne University of Technology, Melbourne, Australia. His email address is nclisby@swin.edu.au.

Communicated by Notices Associate Editor Emilie Purvine.

For permission to reprint this article, please contact:

reprint-permission@ams.org.

DOI: <https://doi.org/10.1090/noti2255>

¹*It's possible that the author is somewhat biased.*

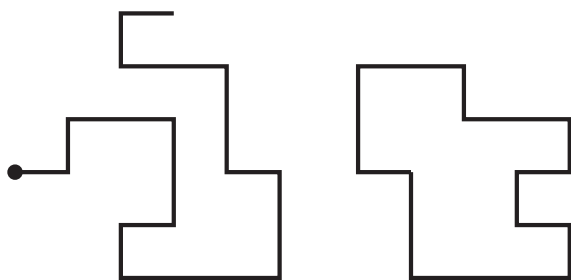


Figure 1. A self-avoiding walk (left) and a self-avoiding polygon (right) on the square lattice.

unit Euclidean length so that $|\omega(i+1) - \omega(i)| = 1$ for each i , and the self-avoidance constraint enforced by $\omega(i) \neq \omega(j)$ for all $i \neq j$. We are interested in the number of walks of length n , which we denote by c_n , the number of polygons of length n , which we denote by p_n , and other properties such as the mean squared distance of the end of an n -step walk from the origin,

$$\mathbb{E}_n(|\omega(n)|^2) = c_n^{-1} \sum_{|\omega|=n} |\omega(n)|^2,$$

which measures the typical size of a walk.

It is widely believed that the following asymptotic behavior occurs for the self-avoiding walk on a regular lattice in any dimension:

$$c_n \sim A\mu^n n^{\gamma-1},$$

$$\mathbb{E}_n(|\omega(n)|^2) \sim Dn^{2\nu},$$

where μ is a lattice-dependent growth constant, and γ and ν are universal critical exponents that only depend on the dimension of the lattice. Simple random walks undergo Brownian motion, and have $\gamma = 1$ and $\nu = \frac{1}{2}$, while walks that move at a steady speed, that is “ballistically,” have $\nu = 1$. For self-avoiding walks there are exact conjectures for $d = 2$ of $\gamma = \frac{43}{32}$ and $\nu = \frac{3}{4}$, and precise numerical estimates from Monte Carlo simulations for $d = 3$ of $\gamma = 1.156953(1)$ and $\nu = 0.5875970(4)$ (the numbers in parentheses indicate the uncertainty in the final digit), while for $d \geq 4$ the self-avoidance constraint is sufficiently weak that critical exponents assume their random walk values $\gamma = 1$ and $\nu = \frac{1}{2}$, with logarithmic corrections when $d = 4$.

Self-avoiding walks are self-similar fractal objects, with fractal dimension $1/\nu$. As $\frac{1}{2} < \nu < 1$ for $d = 2, 3$, self-avoiding walks are visibly more spread out than simple random walks, but do not move ballistically. This can be observed in the image at the start of the article that shows a self-avoiding walk of 1 billion steps on the square lattice, generated using a Markov chain Monte Carlo procedure known as the pivot algorithm.

The self-similarity of two-dimensional self-avoiding walks can be observed in Figure 2. Starting at the top left

corner there is a walk of one billion steps. Then, by following the arrows one has a sequence of subwalks of length $10^8, 10^7, \dots, 10$ steps that are obtained by zooming into the original walk. It is impossible (at least to the author) to distinguish between different length scales until the underlying lattice becomes apparent. This image also serves as a striking “proof by picture” that $\nu = \frac{3}{4}$, as subwalks of length k are scaled by $Ak^{-3/4}$, where A is a fixed constant. If ν was even slightly different from the conjectured value of $\frac{3}{4}$, then the size of the images of these subwalks would noticeably change as k is varied from 10 to 10^9 .

The self-avoiding walk is intimately related to other models in lattice statistical mechanics, as it is the $n \rightarrow 0$ limit of the so-called n -vector model; the famous Ising model corresponds to $n = 1$.

Remarkably, self-avoiding walks can be used to gain exact information about real-world polymer systems due to the phenomenon of universality. Universality is the observation that when a change of state, or phase transition, with a diverging correlation length occurs for a model in statistical mechanics, that sufficiently similar systems in the same “universality class” will have exactly the same kind of behavior in the limit of large system size, and the fine-grained details of the model do not matter.

The key feature that determines the behavior of real polymer chains in a good solvent is the excluded volume constraint, and this property is exactly captured by the self-avoiding walk model. Thus, real polymers, self-avoiding walks on any three-dimensional lattice, weakly avoiding walks that penalize but do not forbid self-intersections, and various semirealistic models of polymers in \mathbb{R}^3 all have the same value of ν . Self-avoiding walks also capture other universal properties, such as the fact that knots are ubiquitous in long polymers; as $n \rightarrow \infty$ the probability of a self-avoiding polygon being knotted approaches 1 exponentially rapidly.

Much is known about self-avoiding walks, but rather less has been proved. Two stand-out results are the proof by Hara and Slade [HS92], via the lace expansion, that for $d \geq 5$ the critical exponents exist and assume random walk values, and the proof by Duminil-Copin and Smirnov [DCS12], via the introduction of a parafermionic observable that satisfies a remarkable discrete holomorphicity condition, that $\mu = \sqrt{2 + \sqrt{2}}$ for the honeycomb lattice. But, although there is overwhelming numerical evidence in favor of the correctness of the asymptotic forms for c_n and $\mathbb{E}_n(|\omega(n)|^2)$ given above, there is as yet no proof of this for $d = 2$ and $d = 3$.

For $d = 2$, it is believed that in the scaling limit, where $n \rightarrow \infty$ and the appropriately scaled lattice spacing approaches zero, that the two-dimensional self-avoiding walk model converges to Schramm-Loewner Evolution

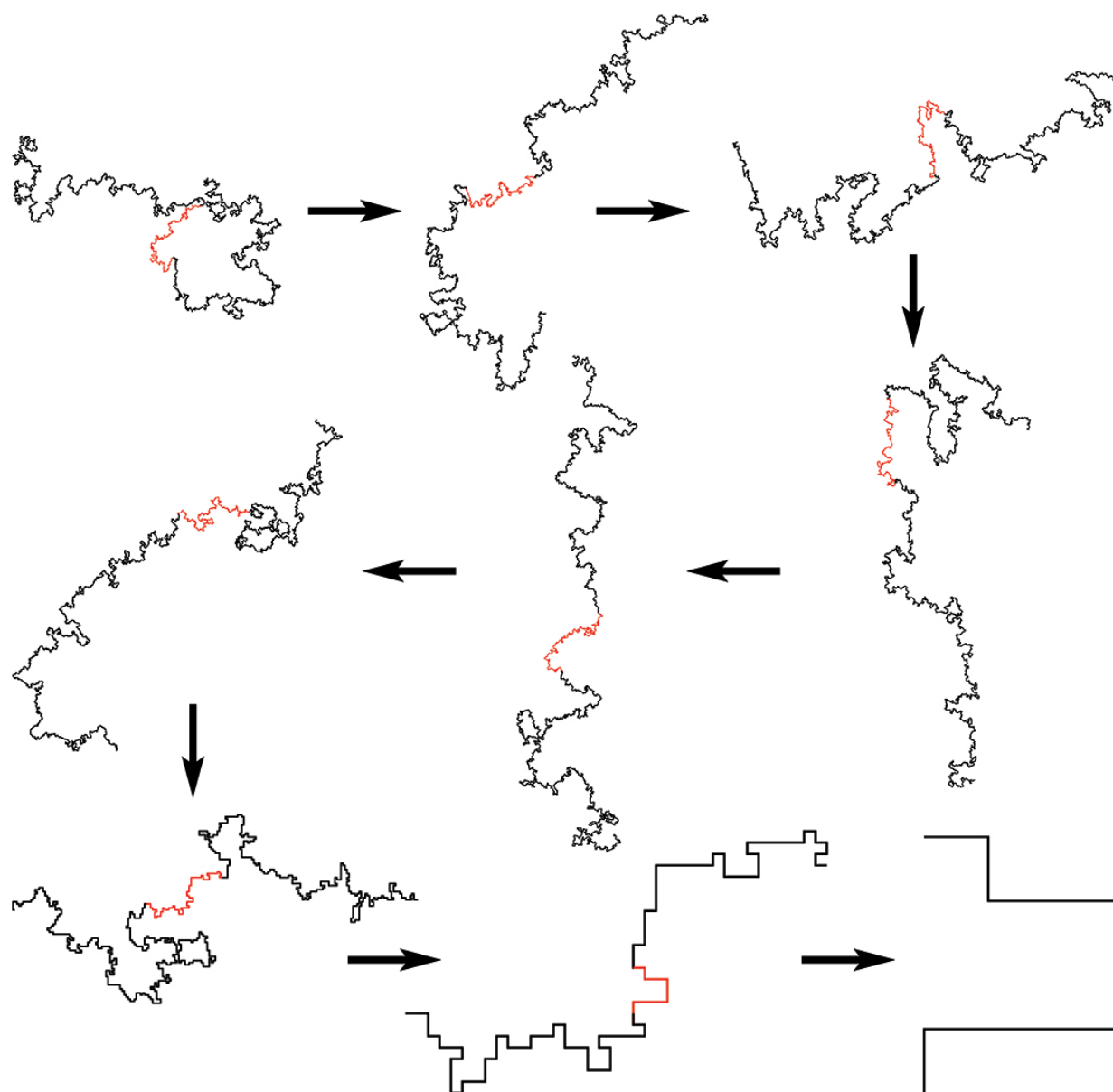


Figure 2. Zooming in to a self-avoiding walk with 10^9 steps on the square lattice. Starting from the top left, subwalks have $10^9, 10^8, \dots, 10$ steps, with each successive subwalk appearing as a red section in the larger walk. The scale factor for the length- k subwalk is $Ak^{-3/4}$, where A is a fixed constant.

(SLE) with parameter $\kappa = \frac{8}{3}$. A proof of this would immediately confirm the conjectured exponent values for the two-dimensional self-avoiding walk, and much else, and would surely be worthy of a Fields Medal.

For $d = 3$, the situation is even more challenging, and the prospects of proving any strong results seem remote. Therefore, current efforts to improve our understanding of three-dimensional lattice polymer systems are focused on experimental mathematics efforts in computer enumeration and Monte Carlo simulation to deepen our qualitative and quantitative understanding of the behavior of these systems.

Despite the large gaps in what has been proved about self-avoiding walks, it has been an extraordinarily rich model in inspiring progress in combinatorics, mathematical physics, probability, and polymer science. Going into detail on these topics would require rather more space than I have available (see [MS93, JvR15] for more of these details) and in this article I will instead narrowly focus on aspects of lattice polymers related to enumerative combinatorics. Although the self-avoiding walk model has not been solved, it has inspired a fabulously diverse ecosystem of combinatorially interesting solvable models. Without even the slightest pretense at comprehensiveness, I

will concentrate on enumerative combinatorial aspects of a select few models of lattice polymers to convey some of the goals and methods of this research area. In addition, one of the most powerful experimental mathematics approaches to the quantitative understanding of self-avoiding walks and related models is to perform exact enumerations on the computer. I will briefly explain the combinatorial and geometric insights that are at the heart of the most efficient known algorithms for the enumeration of self-avoiding walks in two and three dimensions.

I will now give a brief overview of the use of generating functions in the study of lattice polymers.

Generating Functions

For any combinatorial sequence c_n , with $n \geq 0$, we may associate a generating function

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

For lattice models of polymers, typically c_n is the number of walks of length n , and we say that x is the parameter that is conjugate to the length of the walk. Stealing a vivid metaphor from Wilf [Wil94], “a generating function is a clothesline on which we hang up a sequence of numbers for display.” This generating function may be purely formal, but often we can study the behavior of c_n via the analytic properties of $C(x)$; this is the domain of analytic combinatorics.

We may want to keep track of additional properties besides length, in which case we will construct a multivariate generating function. For example, for Dyck paths (defined later) we construct a generating function that weights contacts with a surface

$$D(a, x) = \sum_{m, n \geq 0} d_{m, n} a^m x^n,$$

with $d_{m, n}$ the number of Dyck paths with m contacts and of length n . We can treat a and x as purely formal parameters, or we can give them particular values. In particular, $D(1, x)$ is the generating function for Dyck paths counted by length alone. If we treat a as a fixed constant, then $D(a, x)$ is a function of x , and we can consider the walks to be weighted objects. In this case, there is a direct correspondence between a and so-called Boltzmann weights in statistical mechanics that determine how the lattice polymer will behave as a function of temperature.

When models in statistical mechanics have a phase transition from one state to another, this will manifest as a singularity in the generating function at a critical point. For the generating functions of lattice polymers that we will discuss, this critical point is also the leading singularity of the generating function. For single variable generating functions, such as $C(x)$, the leading singularity, x_c , is a constant, while for multivariate generating functions the

location and nature of the critical point may be a function of the other parameters.

If the leading behavior of the generating function $C(x)$ at x_c is a power-law singularity with exponent $-\gamma$, that is,

$$C(x) \sim A(x) \left(1 - \frac{x}{x_c}\right)^{-\gamma}$$

with $A(x)$ analytic in the vicinity of x_c , then

$$c_n \sim \bar{A} n^{\gamma-1} x_c^{-n} = \bar{A} n^{\gamma-1} \mu^n,$$

where $\mu = 1/x_c$ is the growth constant for the sequence c_n and $\bar{A} = A(x_c)/\Gamma(\gamma)$.

Thus, there is a correspondence between the leading power-law singularity in the generating function and the power-law factor of the asymptotic behavior of the associated sequence. In particular, a pole in the generating function with $\gamma = 1$ will mean that the sequence will grow as a pure exponential with no power-law correction.

More complicated analytical forms and associated asymptotic behavior are possible, including confluent exponents and the occurrence of multiple leading singularities, both of which result in additive power-law corrections for the sequence. Still more exotic behavior may occur, such as stretched exponential asymptotics like

$$c_n \sim A n^{\gamma-1} \mu_1^{n^\sigma} \mu^n,$$

with $\sigma \neq 1$.

Generating functions may have a wide variety of analytical properties. In order of increasing generality, they may be:

- Polynomial, in the case where there are only a finite number of objects. For example, self-avoiding walks in a 4 by 4 grid.
- Rational, where the only singularities are poles.
- Algebraic, so that the generating function is a solution of a polynomial equation.
- Differentiably finite (D-finite), or holonomic, so that the generating function is a solution of a linear differential equation with polynomial coefficients.
- Differentiably algebraic (D-algebraic), so that the generating function is a solution of a polynomial equation involving x , the function, and its derivatives.
- None of the above!

For solvable models, the goal is to find an explicit expression for the generating function—perhaps as an algebraic expression, or as the solution of a higher order ODE—or at least to find a polynomial-time algorithm for generating coefficients.

The anisotropic generating function for self-avoiding polygons on the square lattice, for which vertical steps have different weights than horizontal steps, is not D-finite [Rec06]. If one then sets the vertical and horizontal

weights to be equal one obtains the standard model for self-avoiding polygons on the square lattice. Therefore, in the absence of a miraculous and unexpected cancellation of singularities, the generating function for self-avoiding polygons is not D-finite. This strongly suggests that it will be, at best, extremely challenging to find exact solutions for self-avoiding walks and self-avoiding polygons on the square lattice.

In this case, our goal is to learn as much about the model as possible by finding efficient algorithms to enumerate the objects, so as to determine as many as possible terms of the sequence c_n .

Consequently, we seek to determine the properties of this sequence via series analysis techniques. The most powerful general technique is the method of differential approximants [Gut89], which takes the finite approximation of the generating function and finds a linear ordinary differential equation with polynomial coefficients that reproduces these initial terms. Features of the model such as x_c and the associated critical exponent can then be estimated by analyzing the singularities of the differential equation.

The application of sophisticated computer enumeration and series analysis techniques to problems in lattice statistical mechanics was pioneered from the late 1950s by Domb and Sykes (see, e.g., [DS61]) and their many collaborators including Fisher, Essam, and Gaunt.

It can be extremely powerful to apply series analysis techniques to series expansions for problems that are suspected of being solvable, or more precisely to have D-finite generating functions. If, for example, it is found that a differential equation involving 100 coefficients is able to exactly reproduce 1000 terms of a sequence that has been found by other means, one would be very confident that the true generating function is in fact the solution of the differential equation. Such a discovery does not constitute a proof, but once the solution is known it is usually much easier to prove that it is correct! Furthermore, when the generating function can be proved to be D-finite, with a priori upper bounds available for the order and degree of the associated differential equation, then this guessing procedure can be made fully rigorous [Zei90].

The Model Spectrum

One characteristic that distinguishes different models of walks is how much information must be maintained about the walk as it gets longer, so as to correctly calculate weights and avoid self-intersections.

Variants of directed walks can on occasion completely forget about their past behavior, while so-called “prudent” walks (defined later) can never completely forget about their past behavior, but can condense that information to a finite set of parameters. In contrast, the most physically relevant model, self-avoiding walks, requires that the walk

remember its entire past behavior. Indeed, each step of a self-avoiding walk of length n depends on every other step, and it is incorrect to think of a self-avoiding walk as being generated step-by-step as it is a truly non-Markovian process.

This makes it straightforward to solve simple variants of directed models, challenging to solve models such as prudent walks that require more information to be retained, and extremely difficult—perhaps even impossible—to solve variants of the self-avoiding walk model.

A unifying principle for exact solutions and efficient enumeration algorithms is that a geometric or combinatorial insight must be gained that allows for independence between different parts of the walk, where the dependence between different parts is mediated by a finite (for solutions) or reduced (for algorithms) set of parameters.

Our goal is to find as much information as possible about the sequences and generating functions of lattice polymer models. In particular, we wish to understand the functional form of the generating function and to determine features such as critical exponents exactly, or to estimate them as accurately as possible. In this way we can gain insight into the phase transitions of the model, and by universality these models may also tell us about real-world systems.

The tools we have at our disposal in this endeavor are, in order of preference, exact solutions, polynomial time algorithms for enumeration, exponential algorithms for exact enumeration, and finally Monte Carlo methods of approximate enumeration. We will touch on each of the exact methods in this article, starting with the exact solution of directed walk models.

Directed Walks

We begin by considering the simplest model of all: directed walks on the square lattice. Walks start at the origin, and are able to move either to the north or east with each step. See Figure 3 for example configurations. The directed nature of the walk means that there is never any possibility for self-intersection, and so choices are completely unrestricted. By convention, we take the number of walks

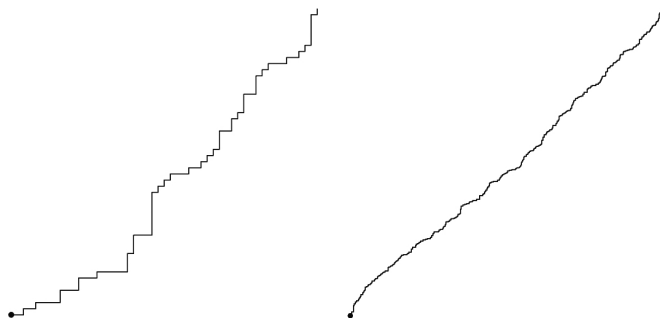


Figure 3. Directed walks on the square lattice of 100 steps (left) and 1000 steps (right).



Figure 5. Example of a Dyck path with 10 steps, and two returns to the surface. The associated weight is a^2x^{10} .

We are introducing this model both because it has interesting combinatorics, and also because it is an extremely simple model of polymer adsorption. We shall see that there is a special value of a where the Dyck path will transform from being typically far from the surface, or “desorbed,” to typically being close to the surface, or “adsorbed.” That is, a phase transition will occur. The contact weight a can be related to experimental quantities used to induce adsorption such as temperature and solvent quality.

To find a functional equation we condition on the first return to the surface. Either a Dyck path has zero steps (contributing weight 1), or it consists of a step above the surface (weight x), then a Dyck path above the surface (where the contact weights are ignored, hence set to 1), then a step down to the surface (weight ax), followed by another Dyck path. This functional equation can be represented pictorially as

$$D(a, x) = 1 + xD(1, x)xD(a, x) \\ = 1 + ax^2D(1, x)D(a, x).$$

If we let $a = 1$ we obtain a quadratic functional equation for $D(1, x)$:

$$D(1, x) = 1 + x^2D(1, x)^2 \\ \Rightarrow D(1, x) = \frac{1 - \sqrt{1 - 4x^2}}{2x^2},$$

where the appropriate branch is chosen to ensure that the solution corresponds to the physical model, which has positive coefficients. We point out that $D(1, x)$ has the expansion

$$D(1, x) = 1 + x^2 + 2x^4 + 5x^6 + \cdots + C_n x^{2n} + \cdots,$$

where C_n is the n th Catalan number, one of the most famous of all sequences. Thus

$$d_{2n} = C_n = \frac{1}{n+1} \binom{2n}{n} \sim \frac{4^n}{\sqrt{n^3\pi}},$$

and by comparing $n^{\gamma-1}$ with $n^{-3/2}$ we observe that $\gamma = -\frac{1}{2}$ and $\mu = 2$ (the same as directed walks) for unweighted Dyck paths.

Now, we can write $D(a, x)$ in terms of $D(1, x)$ to obtain the full solution:

$$D(a, x) = \frac{1}{1 - ax^2D(1, x)} = \frac{2}{2 - a(1 - \sqrt{1 - 4x^2})}.$$

Note that $D(a, x)$ is a function of x^2 , and therefore any coefficients involving an odd power of x must be identically zero. This reflects the fact that there must be an even number of steps to return to the surface.

Our expression for $D(a, x)$ has strikingly different qualitative behavior depending on the value of the weight a , which may be identified with two different universality classes. For $a \leq 2$, the leading singularity is an algebraic singularity at $x_c = \pm \frac{1}{2}$, so $\mu = 2$ and the rate of exponential growth is thus independent of a ; this is the desorbed phase. For $a > 2$, the leading singularity is a pole at $x_c = \pm \frac{1}{a}\sqrt{a-1}$; this is the adsorbed phase.

Dyck paths with contact weights thus provide a qualitative model of the physical adsorption phase transition that is observed with real polymers.

With some extra work [RvR04], it is possible to show that the mean number of contacts for a $2n$ -step weighted Dyck path is to leading order given by

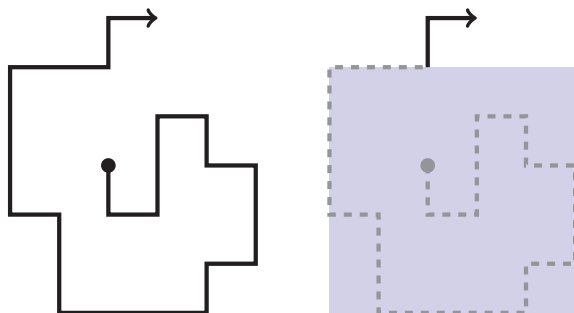
$$\mathbb{E}_{2n}(\text{contacts}) = \begin{cases} \frac{2}{2-a} + o(1), & a < 2, \\ \sqrt{\pi} \sqrt{n} + O(1), & a = 2, \\ \frac{a-2}{a-1} n + O(\sqrt{n}), & a > 2. \end{cases}$$

This conforms with what we might expect: when the polymer is desorbed with $a < 2$, there are on average a finite number of contacts as we would expect the polymer to only approach the surface in the vicinity of the ends. When $a > 2$, the number of contacts is proportional to n , which is consistent with the polymer being adsorbed or stuck to the surface. Finally, precisely at $a = 2$, we have a critical phase where the system changes from being desorbed to adsorbed, and in this case there are $O(n^\phi)$ contacts, with $\phi = \frac{1}{2}$.

For real polymers in two dimensions, it is believed that the exponent ϕ is in fact exactly the same as for weighted Dyck paths, while in three dimensions the question is not fully settled, but the best available evidence is that $\phi = 0.484(4)$ [BOP18].

There is an enormous zoo of directed walk models that are studied for reasons of mathematical interest and physical relevance. These models can be fiendishly difficult to solve, unlike the simple examples presented here. But, for now we move on to a family of models for which steps in all directions are allowed, and thus is closer to the physically relevant self-avoiding walk model.

Prudent walks are walks that make prudent choices to avoid the possibility of self-intersection, by never taking a step towards a site they have previously visited. A short example of a prudent walk is shown in Figure 6, with the region that currently cannot be stepped towards highlighted on the right.



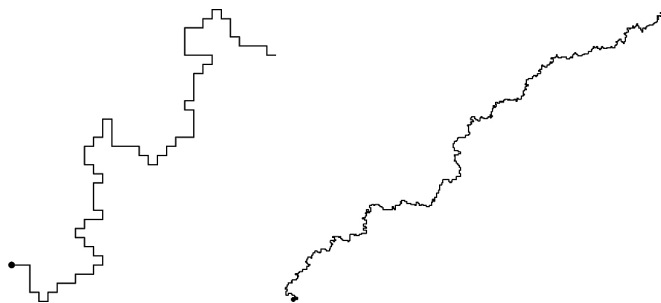
This method of avoidance is significantly less crude than was the case for directed walks, but does not require the arbitrarily long memory of detailed structure that is characteristic of self-avoiding walks. This is illustrated in Figure 6, where on the right the details of the structure of the walk in the shaded region are completely irrelevant to the future choices that may be available—it is safe to forget everything except for the dimensions of the shaded region. The rules for self-avoidance can be encoded by keeping track of the direction of motion, and the bounding rectangle of previously visited sites.

One consequence of the prudent rule is that at each step the end of the walk lies on its bounding rectangle,² and this leads to four natural subcategories of prudent walks: 1-sided prudent walks that are always on the eastern boundary of the rectangle, 2-sided prudent walks that always lie on the northern or eastern boundaries, 3-sided prudent walks that lie on the western, northern, or eastern boundaries, and finally 4-sided prudent walks (or, just prudent walks) that can lie on any boundary. Note that 2-, 3-, and 4-sided prudent walks are not time-reversible.

Two random 4-sided prudent walks are shown in Figure 7. The short walk of 100 steps could almost be mistaken for a self-avoiding walk (compare with the subwalk of 100 steps in Figure 2), but longer walks are very clearly directed.

1-sided prudent walks are more commonly known as partially directed walks, and can move either north, east, or south. We recognize that either the walk never takes an

²The converse is not true—there are walks for which every step lies on the bounding rectangle that are imprudent.



east step, in which case it can take an arbitrary number of north or south steps, or it does take at least one east step, in which case it can take an arbitrary number of north or south steps, then one east step, followed by another walk. Thus,

$$\begin{aligned} R_1(x) &= (1 + 2x + 2x^2 + \cdots) \\ &\quad + (1 + 2x + 2x^2 + \cdots + x)R_1(x) \\ &= \frac{1+x}{1-x} + \frac{(1+x)x}{1-x}R_1(x) \\ \Rightarrow R_1(x) &= \frac{1+x}{1-2x-x^2}. \end{aligned}$$

The leading singularity is a pole at $x_c = -1 + \sqrt{2}$, and so the growth constant is $\mu = 1/x_c = 1 + \sqrt{2} = 2.4142\dots$. Partially directed walks are more numerous than directed walks, but qualitatively they behave in exactly the same way.

2-sided prudent walks are considerably more interesting, as walks can move in all four directions and more information needs to be retained to find a suitable functional equation. The principal insight that allows a functional equation to be constructed is that with each step the 2-sided walk may either take a step along either the northern or eastern boundaries, or the walk may *inflate* the boundary by taking an east step from the eastern boundary, or a north step from the northern boundary. When this inflation step occurs the walk can forget about its past behavior, with the exception of the relative location of the north-east corner. The resulting generating function is

$$B_2(x) = \frac{1}{1-2x-2x^2+2x^3} \times \left(1+x-x^3+x(1-x)\sqrt{\frac{1-x^4}{1-2x-x^3}} \right).$$

Despite their increased freedom as compared to 1-sided walks, 2-sided walks nonetheless behave as directed walks moving in the north-east direction. This is reflected in the fact that the leading singularity is a pole at $x_c = 0.4030317\dots$ that is a solution of $1 - 2x - 2x^2 + 2x^3 = 0$. This corresponds to a growth constant of $\mu = 1/x_c = 2.4811943\dots$. As this growth constant is strictly larger than

the value for 1-sided prudent walks, we see that 2-sided prudent walks are exponentially more numerous.

3-sided prudent walks are much more challenging still, but Bousquet-Mélou [BM10] in a *tour de force* was able to find an appropriate functional equation using the conditional independence after an inflationary step occurs, and then to obtain a closed form expression for the generating function as an infinite sum. Bousquet-Mélou explains that besides the length, information about additional “catalytic” parameters must be retained in the construction of the functional relation. For 2-sided walks, a single catalytic variable that tracks the distance of the end to the north-east corner of the bounding rectangle suffices, while for 3- and 4-sided walks it is necessary to retain two and three catalytic variables, respectively. These catalytic variables contain all of the necessary information about the past behavior of a walk so that the future behavior can be correctly modeled.

3-sided walks asymptotically have the same growth rate as 2-sided walks, and after some extra freedom in their initial stage behave in the same manner as 2-sided walks by settling down to follow the diagonal in either the north-east or north-west directions.

Interestingly, in contrast to the other models we have seen, the generating function for 3-sided prudent walks is not D-finite. Bousquet-Mélou showed that the generating function is meromorphic in the disk $|x| < \sqrt{2} - 1$, but with infinitely many poles occurring in the disk. The striking feature is that 3-sided prudent walks have a straightforward, natural definition, but nonetheless have a wickedly complicated solution. Perhaps there is another, more natural, formulation of the solution that captures the simplicity of the underlying problem?

Finally, no explicit expression has been obtained for 4-sided prudent walks. There is a functional equation [BM10] for P_4 , the generating function for prudent walks ending on the northern edge of the bounding rectangle, involving three catalytic variables:

$$\begin{aligned} & \left(1 - \frac{xuvw(1-x^2)}{(u-xv)(v-xu)}\right) P_4(x; u, v, w) \\ &= 1 + xuP_4(x; w, xw, u) + xvP_4(x; w, xw, v) \\ & \quad - \frac{x^2vw}{u-xv} P_4(x; v, xv, w) - \frac{x^2uw}{v-xu} P_4(x; u, xu, w), \end{aligned}$$

where the desired generating function is obtained by setting the catalytic variables to one, $P_4(x) = P_4(x; 1, 1, 1)$. There is a polynomial-time algorithm based on the functional equation to generate series coefficients, and it was recently proved that the limiting large n behavior is exactly the same as 2- and 3-sided prudent walks [PST17].

It is arguable that this means that the model has been “solved,” in the sense that we exactly know the limiting large n behavior, and can quantitatively understand any

features of the model to arbitrary precision. But, it would be far more satisfying to have an explicit solution, and preferably one that captures the simplicity and elegance of the model itself.

A key question is: can the generating function $P_4(x)$ be written in a convenient or illuminating form, or as the solution of a natural equation without catalytic variables?

If progress could be made on this problem it would perhaps give hints as to how to make progress on other difficult lattice polymer problems. And, just maybe, this would also give some information about what a solution of the full two-dimensional self-avoiding walk problem might look like.

We also note that, disappointingly, each of the prudent walk models has $\nu = 1$, and thus behave as directed walks. It is not clear how to devise a solvable model that has $\frac{1}{2} < \nu < 1$.

This takes us to the boundary of lattice polymer models that can be considered solvable, and we now jump to the other side of the fence.

Self-Avoiding Walks on \mathbb{Z}^3

The simplest method of computer enumeration of combinatorial objects is to generate each of the objects that are being counted, which I will refer to as brute force enumeration. Then, the computational complexity of the algorithm is determined completely by the asymptotic growth in number of the objects.

The secret to inventing an effective enumeration algorithm is to find a combinatorial or geometric insight that will allow one to count less numerous objects instead.

The length-doubling algorithm [SBB11], the fastest known method for counting three-dimensional self-avoiding walks, uses a brilliant insight involving inclusion-exclusion to get an exponential improvement in computational complexity.

We start with the observation that every walk of length $2n$ can be uniquely split into two n -step self-avoiding walks at the midpoint. Therefore, c_{2n} is the number of pairs of nonintersecting n -step self-avoiding walks.

We generate the set S_n of all self-avoiding walks of n steps starting at the origin. By definition, $c_n = |S_n|$, so this requires CPU time $O(n^{\nu-1}\mu^n)$. We label all of the sites that the walks visit (there will be $O(n^3)$ sites), and we let B_i be the set of walks that pass through site i , and A_i be the set of pairs of walks that pass through site i . Finally, we count c_{2n} , first through the observation that the number of mutually avoiding pairs is the same as the number of all pairs (c_n^2) minus the number of pairs with at least one intersection:

$$\begin{aligned} c_{2n} &= \text{number of mutually avoiding pairs from } S_n \\ &= c_n^2 - |\cup_i A_i|. \end{aligned}$$

We then apply the standard inclusion-exclusion formula

for $|\cup_i A_i|$, summing over all subsets $V = \{v_1, \dots, v_k\}$ of the set of visited sites, and partitioning by the size of $|V|$:

$$c_{2n} = c_n^2 - \sum_k \sum_{|V|=k} (-1)^{k+1} |A_{v_1} \cap A_{v_2} \cap \dots \cap A_{v_k}|.$$

Finally, we recognize that the number of pairs of walks that pass through a particular set of sites is the square of the number of individual walks that pass through the set. Thus we obtain the final formula:

$$c_{2n} = c_n^2 - \sum_k \sum_{|V|=k} (-1)^{k+1} |B_{v_1} \cap B_{v_2} \cap \dots \cap B_{v_k}|^2.$$

Most sets are empty, because each walk has n steps, and so any collection of more than n sites must be empty. The powerset for each walk has 2^n elements, so this means that there are at most $2^n c_n$ nonempty sets in the sum.

Heuristically speaking, via the magic of the inclusion-exclusion formula the length-doubling algorithm transforms the problem of counting all possible ways that the half-walks can avoid each other to counting all possible ways that the half-walks can intersect.

This transformation results in a significant reduction in computational complexity: neglecting power-law factors, we have changed the computational effort to enumerate c_{2n} from $O(\mu^{2n})$ to $O((2\mu)^n)$. The computational complexity per step decreases from μ (≈ 4.68) for brute force enumeration to $\sqrt{2\mu}$ (≈ 3.06) for the length-doubling algorithm.

This method has been used to enumerate c_n up to $n = 36$ [SBB11].³ Then, series analysis gives estimates of $\gamma = 1.1570(3)$ and $\mu = 4.684040(5)$, which are very accurate, but less so than values available from Monte Carlo simulations using the pivot algorithm.

Self-Avoiding Polygons on \mathbb{Z}^2

The finite lattice method is the most powerful known method for enumerating self-avoiding walks and polygons in two dimensions. We will not go into the details of how it is applied [CJ12], but instead will concentrate on communicating the central idea.

For Dyck paths and prudent walks, earlier parts of the walk are independent of later ones. But, this is not a good strategy for self-avoiding walks or polygons, as each step depends on all others. Instead, we use spatial independence, by dividing configurations with a boundary which ensures that the two sides are independent of each other modulo topological information on how the boundary edges can join up. Thus we build configurations site-by-site, instead of step-by-step.

We will describe the simpler case of self-avoiding polygons on the square lattice, but the same method can be applied to self-avoiding walks with minor technical tweaks.

We start by observing that each polygon has a unique minimum bounding rectangle. Therefore, if we wish to enumerate all self-avoiding polygons of n steps, we can enumerate all polygons that touch each side of a particular rectangle, and then sum over all possible rectangles. (This is where the moniker “finite lattice method” comes from.)

Next, if we build up this set of polygons in a particular rectangle site-by-site, we do not need to remember all of the details of the polygons that are being generated. Instead, we can enumerate all partial configurations to the left of the boundary of the built-up region that are compatible with a given future topology that specifies how occupied edges on the boundary must be joined up. See Figure 8, which shows two configurations that are compatible with a given boundary; the details of the configurations can be safely forgotten, and it is only necessary to keep track of the boundary state.

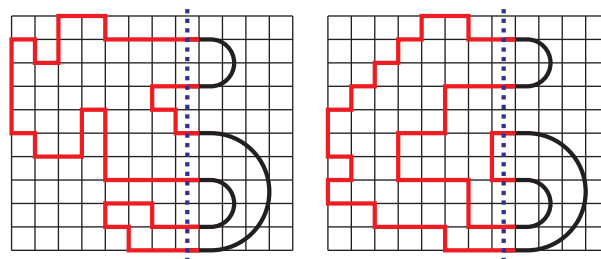


Figure 8. Two partial configurations to the left of the boundary that are compatible with the same boundary topology to the right.

Partial configurations are built up by moving the boundary through the rectangle site-by-site, as shown by the moving boundary in Figure 9.

Thus, the problem of enumerating self-avoiding polygons of n steps has been converted to the problem of enumerating boundary configurations that are compatible with self-avoiding polygons of n steps. The efficiency of the method is determined by the number of boundary configurations, which in turn is determined by the maximum number of occupied edges on the boundary. The occurrence of bad configurations, such as in Figure 9, with $O(n)$ edges concentrated on a particular column means that the number of boundary configurations grows exponentially with n . It can be shown that the algorithmic complexity per step is at most $3^{1/4} \approx 1.316$, but the method of pruning [JG99, CJ12] brings this down to about 1.20.

This is a dramatic improvement over the exponential complexity for the brute force method, which has complexity $\mu \approx 2.638$, and consequently far longer sequences can be generated.

On the square lattice, the finite lattice method has been used to enumerate self-avoiding polygons of up to

³ $c_{36} = 2\,941\,370\,856\,334\,701\,726\,560\,670$.

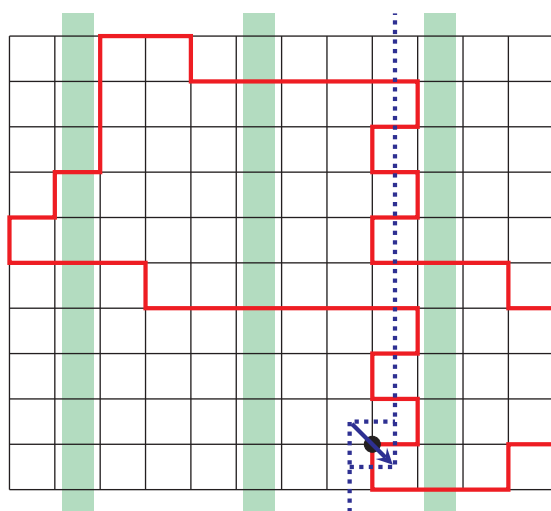


Figure 9. The arrow shows how the boundary is deformed so that a particular site (shown as a solid circle) that is initially on the right-hand side of the boundary ends up on the left-hand side of the boundary. The example configuration is a “bad” polygon configuration with many boundary edges, while a set of good columns is shown in green.

130 steps [CJ12],⁴ and self-avoiding walks of up to 79 steps [Jen13].⁵

Series analysis methods thence allow for spectacularly accurate estimates of critical parameters. For example, analysis of the polygon series gives $\mu = 2.63815853035(2)$. Note that a recently developed transfer-matrix method [JSG16], which involves calculation of eigenvalues of the transfer matrix for different topological sectors, gives an estimate that is even more accurate, of $\mu = 2.63815853032790(3)$.

A major breakthrough was made in a recent paper [Zba19], which combines the transfer matrix method with inclusion-exclusion to derive a subexponential algorithm for enumeration of self-avoiding walks and polygons. The key insight was the recognition that every configuration must have at least one set of “good” columns, each with only $O(\sqrt{n})$ occupied edges, that are spaced at regular intervals in the bounding rectangle; see Figure 9. Then, inclusion-exclusion can be applied to convert the problem of enumerating all polygons to enumerating polygons with any combination of good columns.

Remarkably, the computational effort to enumerate all n -step self-avoiding polygons (or walks) is now $O(r\sqrt{n})$, which is subexponential! The new algorithm has not yet been implemented. Whether it proves to be more efficient in practice for achievable lengths depends crucially on the value of r , which is as yet unknown.

⁴ $p_{130} = 17076613429289025223970687974244417384681143572320$.

⁵ $c_{79} = 10\,194\,710\,293\,557\,466\,193\,787\,900\,071\,923\,676$.

Future Paths

The study of lattice polymers continues to be a vibrant research area.

On the physics side, there is much interest in gaining deeper understanding of the adsorption transition, the collapse transition (where a polymer changes from an extended state to a scrunched up globule), the behavior of confined polymers, and the influence of knots in the polymer chain.

On the mathematical side, the zoo of combinatorially interesting, solvable models continues to grow.

The biggest open problem in the field is the hoped-for proof that self-avoiding walks converge to Schramm-Loewner Evolution with parameter $\frac{8}{3}$ in the scaling limit.

The search for a better mathematical description of the solutions of models like 4-sided prudent walks is ongoing.

It is an open question as to whether it is possible to invent a natural solvable model for which $\frac{1}{2} < \nu < 1$.

And the quest continues for new combinatorial insights, and improved algorithms, for the enumeration of as-yet-unsolved models like self-avoiding walks and self-avoiding polygons.

ACKNOWLEDGMENTS. The author is grateful for helpful comments from Tony Guttmann, Richard Brak, Doron Zeilberger, and the anonymous referees. The author would also like to thank Emilie Purvine, the handling editor, for many helpful comments and for her support during the submission process.

References

- [BM10] Mireille Bousquet-Mélou, *Families of prudent self-avoiding walks*, J. Combin. Theory Ser. A **117** (2010), no. 3, 313–344, DOI 10.1016/j.jcta.2009.10.001. MR2592903
- [BOP18] C. J. Bradley, A. L. Owczarek, and T. Prellberg, *Universality of crossover scaling for the adsorption transition of lattice polymers*, Phys. Rev. E **97** (2018), 022503.
- [CJ12] Nathan Clisby and Iwan Jensen, *A new transfer-matrix algorithm for exact enumerations: self-avoiding polygons on the square lattice*, J. Phys. A **45** (2012), no. 11, 115202, 15, DOI 10.1088/1751-8113/45/11/115202. MR2902304
- [DCS12] Hugo Duminil-Copin and Stanislav Smirnov, *The connective constant of the honeycomb lattice equals $\sqrt{2 + \sqrt{2}}$* , Ann. of Math. (2) **175** (2012), no. 3, 1653–1665, DOI 10.4007/annals.2012.175.3.14. MR2912714
- [DS61] C. Domb and M. F. Sykes, *Use of series expansions for the Ising model susceptibility and excluded volume problem*, J. Mathematical Phys. **2** (1961), 63–67, DOI 10.1063/1.1724213. MR117976
- [Flo49] P. J. Flory, *The configuration of real polymer chains*, J. Chem. Phys. **17** (1949), 303–310.

- [Gut89] A. J. Guttmann, *Asymptotic analysis of power-series expansions*, Phase transitions and critical phenomena, Vol. 13, Academic Press, London, 1989, pp. 1–234. MR1261226
- [HS92] Takashi Hara and Gordon Slade, *Self-avoiding walk in five or more dimensions. I. The critical behaviour*, Comm. Math. Phys. **147** (1992), no. 1, 101–136. MR1171762
- [Jen13] I. Jensen, *A new transfer-matrix algorithm for exact enumerations: self-avoiding walks on the square lattice*, arXiv:1309.6709 (2013).
- [JG99] Iwan Jensen and Anthony J. Guttmann, *Self-avoiding polygons on the square lattice*, J. Phys. A **32** (1999), no. 26, 4867–4876, DOI 10.1088/0305-4470/32/26/305. MR1718791
- [JSG16] Jesper Lykke Jacobsen, Christian R. Scullard, and Anthony J. Guttmann, *On the growth constant for square-lattice self-avoiding walks*, J. Phys. A **49** (2016), no. 49, 494004, 18, DOI 10.1088/1751-8113/49/49/494004. MR3584385
- [JvR15] E. J. Janse van Rensburg, *The statistical mechanics of interacting walks, polygons, animals and vesicles*, 2nd ed., Oxford Lecture Series in Mathematics and its Applications, vol. 39, Oxford University Press, Oxford, 2015. MR3526233
- [MS93] Neal Madras and Gordon Slade, *The self-avoiding walk*, Probability and its Applications, Birkhäuser Boston, Inc., Boston, MA, 1993. MR1197356
- [PST17] Nicolas Pétrélis, Rongfeng Sun, and Niccolò Torri, *Scaling limit of the uniform prudent walk*, Electron. J. Probab. **22** (2017), Paper No. 66, 19, DOI 10.1214/17-EJP87. MR3698735
- [Rec06] A. Rechnitzer, *Haruspicy 2: The anisotropic generating function of self-avoiding polygons is not D-finite*, J. Comb. Theory Ser. A **113** (2006), 520–546. MR2209707
- [RvR04] A. Rechnitzer and E. J. Janse van Rensburg, *Exchange relations, Dyck paths and copolymer adsorption*, Discrete Appl. Math. **140** (2004), no. 1-3, 49–71, DOI 10.1016/j.dam.2003.08.008. MR2064108
- [SBB11] R. D. Schram, G. T. Barkema, and R. H. Bisseling, *Exact enumeration of self-avoiding walks*, J. Stat. Mech. Theory Exp. **6** (2011), P06019, 8, DOI 10.1088/1742-5468/2011/06/p06019. MR2853997
- [Wil94] Herbert S. Wilf, *generatingfunctionology*, 2nd ed., Academic Press, Inc., Boston, MA, 1994. MR1277813
- [Zba19] Samuel Zbarsky, *Asymptotically faster algorithm for counting self-avoiding walks and self-avoiding polygons*, J. Phys. A **52** (2019), no. 50, 505001, 6, DOI 10.1088/1751-8121/ab52b0. MR4051744
- [Zei90] Doron Zeilberger, *A holonomic systems approach to special functions identities*, J. Comput. Appl. Math. **32** (1990), no. 3, 321–368, DOI 10.1016/0377-0427(90)90042-X. MR1090884



Nathan Clisby

Credits

Article images and author photo are courtesy of Nathan Clisby.



APRIL 16-18

FREE
 — ONLINE —
EVENTS
 FOR ALL AGES











nationalmathfestival.org

Organized by the **Mathematical Sciences Research Institute** (MSRI) in cooperation with the **Institute for Advanced Study** (IAS) and the **National Museum of Mathematics** (MoMath).