EXACT SOLUTIONS TO LATTICE MODELS OF POLYMER ADSORPTION

Colin Brandon Pierce

March 2023

A Thesis submitted to the Faculty of Graduate Studies in Partial Fulfillment of the Requirements for the Degree of Master of Science

Graduate Program in Mathematics and Statistics

York University

Toronto, Ontario

Abstract

Entropy calculations are important in determining the physical properties of a polymeric system. A classical method of modeling polymers is with self-avoiding walks, and entropy may be determined by counting the total number of weighted walks. If directed walks are used, recurrences may be formed and solved to study a variety of physical properties exactly. One solution method is to solve the recurrences with generating functions. Additionally, we may attempt to derive the partition function of the model, which explicitly provides the walk entropy per length. In this thesis, the solution to a variety of polymer models of adsorption are derived with generating functions. The partition functions of these models are then extracted when possible through combinatorial convolution identities or by solving their partial difference equations on the lattice.

Contents

Chapter 1

Introduction

Polymers are a class of large molecules composed of a chain of unit chemicals called *monomers*. Polymers can often be familiar to everyday life with some examples include plastics, clothing fibers, proteins, cellulose, and starch. The joining of monomers is known as *polymerization*, and if the polymer is composed of a single monomer it is called a *homopolymer*, and otherwise, a *copolymer.*

Polymer modeling grew out of an increasing need to understand their molecular properties in the chemical sciences. Initially, it was not even known if 'macromolecules' existed, [1] and the increasing evidence for their existence required a tractable mathematical description to make physical predictions. Access to computers was historically an issue, so it was natural to consider modeling a polymer as a sequence of connected points on a discrete grid. Such a sequence is known as a *lattice path*, and the study of them is a rich field of mathematics with wide-reaching solutions.

A popular lattice path for polymer modeling is the *self-avoiding walk* because of its analogue to how monomers occupy distinct spaces. The two-dimensional self-avoiding random walk is a random sequence of steps in the square lattice:

$$\mathbb{L}_2 = \mathbb{Z} \times \mathbb{Z} \tag{1.1}$$

with the added constraint of no self-intersection. Although much is known about self-avoiding walks, their non-Markovian nature makes them difficult to enumerate exactly, and the total number of walks on the square lattice is known only up to length 79 (a result that required supercomputing to determine.) [2]

Asymptotically, the conjecture for the number of self-avoiding walks per length n is:

$$c_n = A\,\mu^n\,n^{\gamma-1} \tag{1.2}$$

where c_n is the number of walks and A and γ are some (numerical) constants. [3] μ is known as the *growth constant* and provides the scaling behavior for the number of self-avoiding walks. On $\mathbb{L}_2$ it is approximated that $\mu = 2.63815$, [3] and interestingly, if we consider the honeycomb lattice, it is known that $\mu = \sqrt{2+\sqrt{2}}$. [4] In general, however, requiring the exact number of walks requires sophisticated enumeration algorithms and quickly become computationally demanding.

There are, however, subsets of random walks that have a naturally self-avoidant step-set. For example, walks in $\mathbb{L}_2$ that step north or east (which we denote N-E or $\{\uparrow, \rightarrow\}$) are naturally self-avoidant and can be enumerated trivially. Such a walk is known as a *directed path.*

The inclusion of a boundary condition adds modeling possibilities. The random walk with step set N-E on $\mathbb{L}_2$ that starts and ends (and does not cross) the boundary $y = x$ is a classical problem in lattice enumeration known as a *Dyck path.* Such a walk renews when it returns to the boundary – this allows for the derivation of a recurrence relation that can be solved to determine the number of paths at each length. Physically, the line $y = x$ models

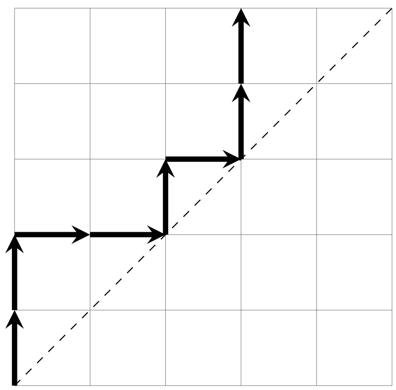

Figure 1.1: The directed path ***NNEENENN*** (of length 8 with 2 renewals) with the boundary $y = x$. Directed walks can be bijected to their step-sequence, which is the basis to enumerating some problems. It is customary to rotate the plane 45 degrees and consider a diagonal lattice with the step-set NE-SE and the condition $h \geq 0$ (for h the height of the walk.)

an impenetrable boundary for a polymer, and its *adsorption* interactions can be modeled by weights picked up by a path each time that it returns to the boundary. Additionally, we may weigh paths that step with a certain pattern to study particular polymer configurations. The great variety of possible boundary conditions and path weights forms a large 'zoo' of polymer models.

The number of directed paths per length can be determined by deriving and expanding its *generating function*:

$$D(t, w_1, w_2, \ldots, w_m) = a_0 + a_1\, t + a_2\, t^2 + \ldots \tag{1.3}$$

where t is the length generating variable, each w_i (for $1 \leq i \leq m$) is a weight determined by the model, and each a_j (for $j \in \mathbb{N}_0$) is found upon expansion. (For example, if we weigh north steps with ν, the path in Figure 1.1 would be the factor ν^5 for the term t^8 in the generating function.) A generating function can also be expressed as:

$$D(t, w_1, w_2, \ldots, w_m) = \sum_{n \geq 0} d_n(w_1, w_2, \ldots, w_m)\, t^n \tag{1.4}$$

where d_n is a function that provides a polynomial of weights per path length n. d_n is known as a *partition function* and can be expressed as:

$$d_n(w_1, w_2, \ldots, w_m) = \sum_{k_1, k_2, \ldots, k_m} p(n, k_1, k_2, \ldots, k_m)\, w_1^{k_1}\, w_2^{k_2} \cdots w_m^{k_m} \tag{1.5}$$

where p (and the summation bounds) must be determined through a variety of methods.

'Solving' our model thus comes down to deriving its generating function or partition function, and can pose difficult mathematical problems. If tractable, exact solutions to discrete models are preferred for a variety of reasons. For example, there are obvious physical limitations to a model defined on discrete points in space and time (especially when considering the simulation capability of modern supercomputers), so a combinatorial understanding is preferred when solvable. Further, enumeration becomes difficult as the length of the polymer grows, so a closed form solution allows for the analysis of scaling behavior. For example, the simple generating function

$$F(t) = \frac{1}{1 - 2dt} \tag{1.6}$$

(with partition function $F_n = (2d)^n$) is convergent up to $t_c = (2d)^{-1}$, so the generating function coefficients grow at rate $O(t_c^{-n})$.

Physically, the configurational entropy of a polymer is the number of paths to each of its endpoints, and in simple models, this is solvable with raised endpoints. Thermodynamic quantities of our model can be then be determined from the radius of convergence of the generating function. For example,

$$F = -\log(t_c) \tag{1.7}$$

would be the *free energy* of our model. Our model weights then allow for the analysis of certain polymer system behavior. For example, a polymer with a tendency to rise but that interacts with a sticky boundary will behave differently depending on how the path is weighted. Further, the derivatives of F with respect to model weights then determine additional thermodynamic quantities of our model.

This thesis will not focus on deriving physical quantities for our models, but will provide the radius of convergence of our generating functions where relevant. Deriving a model's generating function is in general an easier and more intuitive task. We begin by focusing on polymers that interact with a boundary and extend to model other polymer behaviour.

Chapter 2

Generating functions and the modeling of polymer adsorption

2.1 Introduction

The generating function of a polymer model encodes all possible walk lengths and their respective (weighted) coefficients. For ordinary Dyck paths, this can be derived by considering a functional recurrence on its first excursion (this is known as the *wasp-waist* method. [4])

A Dyck path is either length 0 (in which case there is 1 path), or it is a raised Dyck path which then renews (as a Dyck path) upon returning to the boundary. (See Figure 2.1)

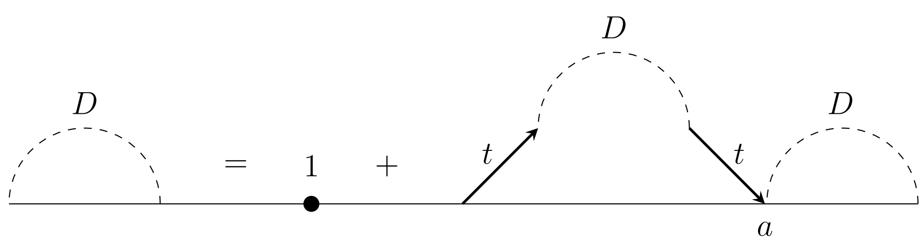

Figure 2.1: Recurrence for Dyck paths

Hence,

$$D(t) = 1 + t^2 D(t)^2 \tag{2.1}$$

which solves to:

$$\begin{aligned} D(t) &= \frac{1-\sqrt{1-4t^2}}{2t^2} = \frac{2}{1+\sqrt{1-4t^2}} \\ &= 1 + t^2 + 2\,t^4 + 5\,t^6 + 14\,t^8 + 42\,t^{10} + \dots \end{aligned} \tag{2.2}$$

(where the *physical root* was selected, i.e., the coefficients of t^{2n} are positive.) Notice that this is the generating function of the *Catalan* numbers. Therefore, the number of Dyck paths of even length is:

$$C_n := [t^{2n}]\, D(t) = \binom{2n}{n} - \binom{2n}{n-1} = \binom{2n}{n}\frac{1}{n+1} = \binom{2n+1}{n}\frac{1}{2n+1} \tag{2.3}$$

(where $[t^{2n}]\, D(t)$ denotes the coefficient of t^{2n} $\forall\, n \in \mathbb{N}$.)

This generating function is convergent with radius $t_c = \frac{1}{2}$ (since the square root vanishes at $t = \frac{1}{2}$.) This implies that $C_n \sim 2^n$, and therefore, that the Catalan numbers grow like $C_{2n} \sim 4^n$.

2.2 Generating function solutions of polymer models

2.2.1 Counting Dyck path returns

We can model polymer adsorption by weighing paths each time that they return to the boundary. The generating functions of adsorbing models can be derived from recurrences similar to equation (2.1).

Let $D(a, t)$ be the generating function of Dyck paths weighted by an a each time that they return to the boundary. A return-weighted Dyck path is either length 0, or it is a stepped-up Dyck path that then renews (as a return-weighted Dyck path) and picks up an

a. Thus,

$$D(a,t) = 1 + at^2 D(1,t) D(a,t) \tag{2.4}$$

Substituting (2.2) solves the recurrence:

$$\begin{aligned} D(a,t) &= \frac{1}{1 - at^2 D(t)} = \frac{2}{2 - a(1 - \sqrt{1-4t^2})} \\ &= 1 + a\,t^2 + (a^2 + a)\,t^4 + (a^3 + 2a^2 + 2a)\,t^6 + \dots \end{aligned} \tag{2.5}$$

For this generating function, the square root is zero when $t = \frac{1}{2}$ and the denominator is zero when $t = \frac{\sqrt{a-1}}{a}$. Thus, we have that $t_c = \min(\frac{1}{2}, \frac{\sqrt{a-1}}{a})$. The critical value of a is when $\frac{1}{2} = \frac{\sqrt{a-1}}{a}$, that is, when $a_c = 2$. When $a > 2$, t_c is a function of a, and when $a < 2$, t_c is constant.

Physically, t^{2n} represents a polymer with $2n$ monomers, and multiplying by t is the bonding of an additional monomer. The radius of convergence we obtained models the phases of adsorption and desorption in our polymer model (and the model undergoes a state change at the critical value of a.)

Notice that deriving the generating function of our model is not enough to know its partition function, that is, we do not immediately have a function that generates the number of paths (and their weights) per walk length.

We can visually see a few example walks and their weights in the following table:

Half-length of walk	$[t^{2n}]\,D(t,a)$	Dyck paths
$n=1$	a	/\
$n=2$	a^2+a	/\/\ ; //\\
$n=3$	a^3+2a^2+2a	/\/\/\ ; /\//\\ , //\\/\ ; ///\\\ , //\/\\

Table 2.1: Example return-weighted Dyck paths up to $n = 3$

2.2.2 Counting Dyck path double-rises

We can model the dynamics of an adsorbing surface and a polymer's tendency to rise with weights for paths that visit the boundary and raise to a certain height. All Dyck paths necessarily step up once after a boundary visit (unless it is the final step), so we consider the next simplest case which is *double-rises.*

We again solve a recurrence relation to find this model's generating function. The first excursion is either the trivial path (where it has not risen twice), or it is a raised Dyck path (except the trivial path) that has picked up a weight for rising twice. Thus,

$$D(r,t) = 1 + t^2(1 + r(D(t) - 1))D(r,t) \tag{2.6}$$

Which solves to:

$$\begin{aligned} D(r,t) &= \frac{1}{1 - t^2(rD(t) + (1-r))} = \frac{2}{2 - 2t^2 + r\left(2t^2 - 1 + \sqrt{1-4t^2}\right)} \\ &= 1 + t^2 + (r+1)\,t^4 + (4r+1)\,t^6 + (r^2 + 12r + 1)\,t^8 + \dots \end{aligned} \tag{2.7}$$

with $t_c = \min(\frac{1}{2}, \sqrt{\frac{r\sqrt{4r-3}-3r+2}{2(r-1)^2}})$ and $r_c = 3$.

This model may trivially extended to include adsorption by adding a return weight to the recurrence:

$$D(a,r,t) = 1 + at^2(1 + r(D(t) - 1))D(a,r,t) \tag{2.8}$$

Which solves to:

$$\begin{aligned} D(a,r,t) &= \frac{1}{1 - at^2(rD(t) + (1-r))} = \frac{2}{2 - 2at^2 + ar(2t^2 - 1 + \sqrt{1-4t^2})} \\ &= 1 + a\,t^2 + (a^2 + ar)\,t^4 + (a^3 + 2a^2r + 2ar)\,t^6 + \dots \end{aligned} \tag{2.9}$$

with $t_c = \min(\frac{1}{2}, \sqrt{\frac{\sqrt{r^2a(a+4r-4)}+2-r(a+2)}{2a(r-1)^2}})$ and the critical curve $a_c = \frac{4}{r+1}$.

2.2.3 Counting Dyck path peaks

We can model a polymer's tendency to bend by weighing Dyck paths each time that they form a 'peak' (a consecutive up-down step.) The generating function for Dyck paths peaks can again be found by the functional recurrence method. [5] A Dyck path with peaks is either the trivial path, or it has a first excursion where it renews as a Dyck path with peaks. The first excursion is either a (full) Dyck path with peaks, or it is the trivial path (which must pick up a weight since it has stepped up and down.) Hence,

$$D(c,t) = 1 + t^2(c + D(c,t) - 1)\, D(c,t) \tag{2.10}$$

With solution,

$$\begin{aligned} D(c,t) &= \frac{1 - t^2(c-1) - \sqrt{(1+t^2-ct^2)^2 - 4t^2}}{2t^2} \\ &= \frac{2}{1 - t^2(c-1) + \sqrt{(1+t^2-ct^2)^2 - 4t^2}} \\ &= 1 + c\,t^2 + (c^2+c)\,t^4 + (c^3+3c^2+c)\,t^6 + \dots \end{aligned} \tag{2.11}$$

and $t_c = \frac{\sqrt{c+1-2\sqrt{c}}}{|c-1|}$ (with no critical value of c.)

This is the generating function of the *Narayana* numbers – a classical Catalan factorization that enumerates many problems. Notice that, (along with our other models,) substituting $c = 1$ recovers the Catalan generating function. Fortunately, these coefficients have a well-known formula:

$$N_n := \sum_{k=0}^{n} \frac{1}{n}\binom{n}{k}\binom{n}{k-1} c^k = \sum_{k=0}^{n} \frac{1}{n+1}\binom{n+1}{k}\binom{n-1}{k-1} c^k \tag{2.12}$$

and is therefore the partition function of this model (see [6, 7] for applications and extensions of this model.)

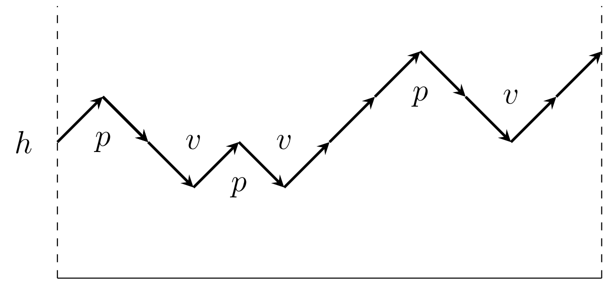

Figure 2.2: There is a one-to-one correspondence between the number of peaks (p) and the number of valleys (v) that a Dyck path has. Further, each peak is followed by a valley (unless it is the final step.)

The generating function of peaks and returns can be obtained by adding a return weight to equation (2.10):

$$D(a,c,t) = \frac{1}{1-at^2(D(c,t)+(c-1))} = \frac{2}{2-a(1+t^2(c-1))-\sqrt{(1+t^2-ct^2)^2-4t^2}}$$
$$= 1 + ac\,t^2 + (a^2c^2+ac)\,t^4 + (a^3c^3+2a^2c^2+a(c^2+c))\,t^6 + \dots \tag{2.13}$$

and has a radius of convergence of $t_c = \sqrt{\min(\frac{a-1}{a^2c-ac+a}, \frac{c+1-2\sqrt{c}}{(c-1)^2})}$ with the critical curve $c_c = \frac{1}{(a-1)^2}$.

We can model polymer oscillation by counting the number of times that a Dyck path changes directions, that is, its number of *peaks* and *valleys*. There is a one-to-one correspondence between the number of peaks and the number of peaks and valleys, that is, a walk with n peaks will have $2n-1$ peaks and valleys (see Figure 2.2).

To obtain the generating function of this model, we first consider equation (2.11) with:

$$D(c^2,t) = 1 + c^2\,t^2 + (c^4+c^2)\,t^4 + (c^6+3c^4+c^2)\,t^6 + \dots \tag{2.14}$$

This weighs a path's peaks and valleys but includes an excess valley weight at the end of each walk. To enumerate peaks and valleys, therefore, we require:

$$D(c,t) \to c^{-1}(D(c^2,t)-1)+1 \tag{2.15}$$

Solving and choosing u to weigh peaks and valleys provides:

$$\begin{aligned} D(u,t) &= \frac{1-t^2(u-1)^2-\sqrt{(1+t^2-u^2t^2)^2-4t^2}}{2ut^2} \\ &= 1+u\,t^2+(u^3+u)\,t^4+(u^5+3u^3+u)\,t^6+\dots \end{aligned} \tag{2.16}$$

2.2.4 Counting Dyck paths with multiple statistics

We may demonstrate the enumeration of large polymer models by counting Dyck paths with all the statistics we have mentioned so far. We also extend to include a weight per rise from the boundary (in addition to double rises.)

A Dyck path is either the trivial path or it is a raised Dyck path that picks up a weight s for rising, and a weight a for returning to the boundary. This raised path is either the trivial path (which picks up a c for forming a peak), or it is a raised Dyck path with peaks (that picks up a weight r for double-rising.)

Adding these features together produces the following recurrence:

$$D(a,c,r,s,t) = 1+ast^2(c+r(D(c,t)-1))\,D(a,c,r,s,t) \tag{2.17}$$

With solution:

$$\begin{aligned} D(a,c,r,s,t) &= \frac{1}{1-ast^2(c+r(D(c,t)-1))} \\ &= \frac{2}{2+sat^2(cr-2c+r)+sra(-1+\sqrt{(1+t^2-ct^2)^2-4t^2})} \\ &= 1+asc\,t^2+(a^2s^2c^2+asrc)\,t^4+(a^3c^3s^3+2a^2s^2c^2r+ac^2rs+asrc)\,t^6+\dots \end{aligned} \tag{2.18}$$

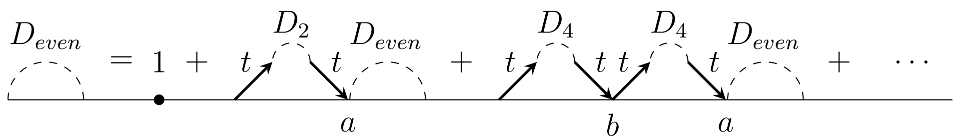

Figure 2.3: Recurrence for alternating Dyck paths returns of length $4n$

This represents the most general model for the parameters we have counted so far, and any of the previous models can be obtained with substitution.

2.2.5 Counting alternating Dyck path returns

We can model a copolymer interacting with a periodic surface by counting Dyck path returns of length 0 (mod 4) with a weight a, and returns of length 2 (mod 4) with a weight b.

The generating function for alternating Dyck paths may be derived by first considering a factorization of Catalan numbers mod 2 and mod 4. [4] Summing terms in a computer algebra system finds that:

$$\begin{aligned} D_4 &:= \sum_{n\geq 0} C_{2n}\, t^{4n} = \frac{\sqrt{2}}{\sqrt{1+\sqrt{1-16t^4}}} \\ &= 1 + 2\,t^4 + 14\,t^8 + 132\,t^{12} + \ldots \end{aligned} \tag{2.19}$$

with $t_c = \frac{1}{2}$.

A walk of length 2 (mod 4) is either a raised walk of length 2 (mod 4) followed by a walk of length 2 (mod 4), or it is a raised walk of length 0 (mod 4) followed by a walk of length 0 (mod 4). Thus,

$$\begin{aligned} & D_2 = t^2(D_2^2 + D_4^2) \\ \implies & D_2 = \frac{8t^2}{(1+\sqrt{1-16t^4})(2+\sqrt{2+2\sqrt{1-16t^4}})} \\ & \quad = \sum_{n\geq 0} C_{2n+1}\, t^{4n+2} = t^2 + 5\,t^6 + 42\,t^{10} + 429\,t^{14} + \ldots \end{aligned} \tag{2.20}$$

with $t_c = \frac{1}{2}$.

The alternating Dyck path generating function of length $4n$ (which we denote D_{even}) can be obtained by considering the geometric recurrence of an arbitrary number of b's before an a (since this can form any sequence of a's and b's of length $4n$.)

Observing our factorization in Figure 2.3, we see that:

$$\begin{aligned} D_{even} &= 1 + at^2 D_2 D_{even} + abt^4 D_4^2 D_{even} + abt^4 D_4^2 bt^2 D_2 D_{even} + \dots \\ &= 1 + at^2 D_2 D_{even} + (\frac{abt^4 D_4^2}{1 - bt^2 D_2}) D_{even} \end{aligned} \tag{2.21}$$

Which solves to:

$$\begin{aligned} D_{even}(a,b,t) &= \frac{1 - bt^2 D_2}{(1 - at^2 D_2)(1 - bt^2 D_2) - abt^4 D_4^2} \\ &= 1 + (ab + a)\, t^4 + (a^2b^2 + 2a^2b + ab^2 + a^2 + 4ab + 5a)\, t^8 + \dots \end{aligned} \tag{2.22}$$

The generating function for paths of length $4n+2$ (which we denote D_{odd}) can be obtained by first denoting $D_p = D_{even}(b,a,t)$ (which counts paths of length $4n$ shifted by two steps.) Any odd length path is an odd length path followed by an even length path (where the latter now ends at length $4n+2$ and has shifted weights.) Therefore,

$$\begin{aligned} D_{odd}(a,b,t) &= bt^2 D_4 D_p + abt^4 D_2 D_4 D_p + a^2bt^6 D_2 D_2 D_4 D_p + \dots \\ &= bt^2 D_4 D_p + \frac{abt^4 D_2 D_4 D_p}{1 - at^2 D_2} \end{aligned} \tag{2.23}$$

Substituting, we obtain:

$$\begin{aligned} D_{odd}(a,b,t) &= \frac{bt^2 D_4 (1 - bt^2 D_2)^2 + abt^4 D_2 D_4 (1 - at^2 D_2)}{(1 - at^2 D_2)^2 (1 - bt^2 D_2) - abt^4 D_4^2 (1 - at^2 D_2)} \\ &= b\, t^2 + (ab^2 + ab + b^2 + 2b)\, t^6 \\ &\quad + (a^2b^3 + 2a^2b^2 + 2ab^3 + a^2b + 7ab^2 + b^3 + 7ab + 7b^2 + 14b)\, t^{10} + \dots \end{aligned} \tag{2.24}$$

Therefore, the full solution is:

$$\begin{aligned}
D_{alt} &= D_{even} + D_{odd} \\
&= \frac{(1-at^2D_2)(1-bt^2+abt^4D_2D_4)+bt^2D_4(1-at^2D_2)^2}{(1-at^2D_2)^2(1-bt^2D_2)-abt^4D_4^2(1-at^2D_2)} \\
&= 1+b\,t^2+(ab+a)\,t^4+(ab^2+ab+b^2+2b)\,t^6+(a^2b^2+2a^2b+ab^2+a^2+4ab+5a)\,t^8+\dots
\end{aligned} \tag{2.25}$$

Chapter 3

Exact enumeration with convolutions

3.1 Introduction

The cases of the Catalan and Narayana numbers are well-known enough to provide the partition functions of Dyck paths and Dyck paths with peaks. This is not generally true of partition functions, and we must attempt to derive them through a variety of methods.

A standard first approach is to search for the terms found by hand in the Online Encyclopedia of Integer Sequences (OEIS). [8] It is sometimes the case that the solution to another combinatorial problem exists that can be bijected to the lattice path model being considered.

Additionally, it is sometimes possible to derive a partition function directly from a generating function. For example, it is a popular exercise in elementary combinatorics books to derive the Catalan numbers from an extended binomial expansion of the generating function. [9, 10] This derivation, however, involves algebraic 'tricks' and quickly becomes unfeasible for multivariate generating functions. An even more rare 'trick' would be to guess and check a formula by inspection, but this is likely to fail for all but the simplest recurrences.

We are instead interested in more methodical approaches to deriving partition functions.

Consider, again, the return-weighted Dyck path generating function (in section 2.2.1). It is fundamentally of the form:

$$D(1,t) = 1 + t^2 D + t^4 D^2 + t^6 D^3 + \cdots = \frac{1}{1 - 1 \cdot t^2 D} \tag{3.1}$$

Each successive power of D is a successive *convolution* of D. Fortunately, there is a pattern for the coefficients of two convoluted generating functions.

If $A(z)$ and $B(z)$ be two arbitrary generating functions:

$$\begin{aligned} A(z) &= a_0 + a_1\, z + a_2\, z^2 + \ldots \\ B(z) &= b_0 + b_1\, z + b_2\, z^2 + \ldots \end{aligned} \tag{3.2}$$

Then,

$$\begin{aligned} A(z)\, B(z) &= a_0 b_0 + (a_0 b_1 + a_1 b_0)\, z + (a_2 b_0 + a_1 b_1 + a_0 b_2)\, z^2 + \ldots \\ \implies [z^n]\, A(z)\, B(z) &= \sum_{k=0}^{n} a_k b_{n-k} \end{aligned} \tag{3.3}$$

This is therefore the partition function of any convoluted generating functions. Ideally, our partition function has no more indices (summation signs) than variables, so we are motivated to find a closed-form to these expressions for simplicity and efficiency. An additional criteria we impose is to avoid solutions that require special functions like the *hypergeometric series* to express. (Such closed-forms are capable of solving more models, but are generally more difficult to work with.)

3.2 Enumeration with convolutions

We demonstrate how convolution identities may be used to derive a model's partition function by considering the expansion of the return-weighted Dyck path generating function. Its solution is of the form:

$$D(a,t) = \sum_{k\geq 0} a^k t^{2k} D(t)^k \tag{3.4}$$

Combinatorial identities exist that are capable of reducing binomial convolutions. Consider the first convolution $D(t)^2 = D(t)\cdot D(t)$. Simplifying finds that:

$$\begin{aligned}[t^{2n}]\, D(t)^2 &= \sum_{k=0}^{n} \binom{2k+1}{k}\frac{1}{2k+1}\binom{2(n-k)+1}{n-k}\frac{1}{2(n-k)+1} \\ &= \binom{2n+2}{n}\frac{2}{2n+2} \quad \text{(by Rothe-Hagan convolution)}\end{aligned} \tag{3.5}$$

Which inspires the guess that:

$$C_n^{(s)} := [t^{2n}]\, D(t)^s = \binom{2n+s}{n}\frac{s}{2n+s} = \binom{2n+s-1}{n} - \binom{2n+s-1}{n-1} \quad \text{for } s\geq 1 \tag{3.6}$$

which may be proved similarly.

Proof. $s=2$ by equation (3.5). Assume s. Then,

$$\begin{aligned}[t^{2n}]\, D(t)^{(s)}D(t) &= \sum_{k=0}^{n} \binom{2(n-k)+s}{n-k}\frac{s}{2(n-k)+s}\binom{2k+1}{k}\frac{1}{2k+1} \\ &= \binom{2n+s+1}{n}\frac{s+1}{2n+s+1} = C_n^{(s+1)} \quad \text{(by Rothe-Hagan convolution)}\end{aligned} \tag{3.7}$$

□

These are known as the *Fuss-Catalan* numbers, and are considered to be a generalization of the Catalan numbers. [10, 11]

Therefore, the partition function of our adsorbing model can be found by expanding:

$$\begin{aligned}
[t^{2n}]\, D(a,t) &= [t^{2n}] \sum_{k\geq 0} \sum_{j\geq 0} a^k C_j^{(k)}\, t^{2j+2k} \\
&= [t^{2n}] \sum_{k\geq 0} \sum_{j\geq 0} a^k C_{j-k}^{(k)}\, t^{2j} \\
&= \sum_{k=0}^{n} \frac{k\, a^k}{2n-k} \binom{2n-k}{n-k} = p_a(n)
\end{aligned} \tag{3.8}$$

Convolution identities allow for the enumeration of many other Dyck path polymer models. For example, we can extract the partition function of the double-rises model in section 2.2.2 with a binomial expansion:

$$\begin{aligned}
[t^{2n}]\, D(a,r,t) &= [t^{2n}] \sum_{k\geq 0} \sum_{\ell=0}^{k} \binom{k}{\ell} t^{2k} a^k r^\ell D(t)^\ell (1-r)^{k-\ell} \\
&= [t^{2n}] \sum_{k\geq 0} \sum_{\ell=0}^{k} \sum_{j\geq 0} \binom{k}{\ell} a^k r^\ell (1-r)^{k-\ell} C_j^{(\ell)}\, t^{2k+2j} \\
&= \sum_{k=0}^{n} \sum_{\ell=0}^{k} \binom{k}{\ell} \binom{2(n-k)+\ell}{n-k} \frac{\ell\, a^k r^\ell (1-r)^{k-\ell}}{2(n-k)+\ell} := f_{a,r}(n)
\end{aligned} \tag{3.9}$$

The case $k = n$ is singular when $\ell = 0$. This candidate partition function can be corrected by defining $C_0^{(0)} = 1$, or alternatively, (to avoid piecewise definitions,) enumerating up to the singular term and adding the missing weights. Enumerating in a computer algebra system with the upper bound $k = n - 1$ demonstrates the pattern that allows this to be done (see Table 3.1).

The only discrepancy between our generating function and our partition function is that our partition function is missing the term 1 at each length. Therefore, we conjecture that the partition function is:

$$p_{a,r}(n) = \sum_{k=0}^{n-1} \sum_{\ell=0}^{k} \binom{k}{\ell} \binom{2(n-k)+\ell}{n-k} \frac{\ell\, a^k r^\ell (1-r)^{k-\ell}}{2(n-k)+\ell} + a^n \tag{3.10}$$

Half-length of walk	$f_{1,r}(n)$	$[t^{2n}]\, D(1,r,t)$
$n=1$	0	1
$n=2$	r	$r+1$
$n=3$	$4r$	$4r+1$
$n=4$	r^2+12r	$r^2+12r+1$
$n=5$	$7r^2+34r$	$7r^2+34r+1$
$n=6$	r^3+32r^2+98r	$r^3+32r^2+98r+1$

Table 3.1: Candidate partition function $f_{1,r}(n)$ for weighted Dyck path double-rises (equation 2.9) up to $n=6$

This may be proved by considering that when $k=n$, a^n weighs the path that repeatedly steps up and down. [12] For any Catalan factorization, this path will always have a coefficient of 1 (with or without weights.)

3.3 Enumerating Dyck path peaks and oscillations

With convolutions in mind, we return to our Dyck path peaks generating function (in section 2.2.3). The convolution of Narayana polynomials is known to be: [12]

$$N_n^{(s)}(c) = \sum_{k=0}^{n} \binom{n-1}{k-1} \binom{n+s}{k} \frac{s}{n+s}\, c^k \tag{3.11}$$

Notice, it must be the case that the convolution of Narayana numbers sums to the convolution of Catalan numbers:

$$
\begin{aligned}
N_n^{(s)}(1) &= \sum_{k=0}^{n} \binom{n-1}{k-1}\binom{n+s}{k}\frac{s}{n+s} \\
&= \frac{s}{n+s}\sum_{k=0}^{n}\binom{n-1}{n-k}\binom{n+s}{k} \\
&= \frac{s}{n+s}\binom{2n+s-1}{n} \\
&= \frac{s}{2n+s}\binom{2n+s}{n} = C_n^{(s)}
\end{aligned}
\tag{3.12}
$$

Therefore, this model's partition function can be extracted. From equation (2.13) we have that:

$$
D(a,c,t) = \frac{1}{1-at^2(D(c,t)+(c-1))} = \sum_{k\geq 0} a^k t^{2k}(D(c,t)+(c-1))^k \tag{3.13}
$$

And so,

$$
\begin{aligned}
[t^{2n}]\, D(a,c,t) &= [t^{2n}] \sum_{k\geq 0}\sum_{\ell=0}^{k}\binom{k}{\ell} a^k t^{2k}(c-1)^{k-\ell} D(c,t)^{\ell} \\
&= [t^{2n}] \sum_{k\geq 0}\sum_{\ell=0}^{k}\binom{k}{\ell} a^k (c-1)^{k-\ell} \sum_{j\geq 0} N_j^{(\ell)}\, t^{2j+2k} \\
&= [t^{2n}] \sum_{k\geq 0}\sum_{\ell=0}^{k}\sum_{j\geq 0}\binom{k}{\ell} a^k (c-1)^{k-\ell} N_{j-k}^{(\ell)}\, t^{2j} \\
&= \sum_{k=0}^{n}\sum_{\ell=0}^{k}\sum_{i=0}^{n-k}\binom{k}{\ell}\binom{n-k-1}{i-1}\binom{n-k+\ell}{i}\frac{\ell\, a^k(c-1)^{k-\ell}c^i}{n-k+\ell} := f_{a,c}(n)
\end{aligned}
\tag{3.14}
$$

Similar to $f_{a,r}$, $f_{a,c}$ is singular for $k=n$ and requires modification to obtain the partition function of our model:

$$
p_{a,c}(n) = \sum_{k=0}^{n-1}\sum_{\ell=0}^{k}\sum_{i=0}^{n-k}\binom{k}{\ell}\binom{n-k-1}{i-1}\binom{n-k+\ell}{i}\frac{\ell\, a^k(c-1)^{k-\ell}c^i}{n-k+\ell} + (ac)^n \tag{3.15}
$$

More generally, we see that any partition function that requires a binomial expansion and convolutions to enumerate will have a singularity that needs to be handled in a similar manner.

We may modify our derived partition functions to model different polymers. As demonstrated in section 2.2.3, mapping $c \to c^2$ (and multiplying by c^{-1} to remove the excess valley) adds weights for the number of valleys and therefore enumerates our model of polymer oscillation and adsorption:

$$p_{a,u}(n) = \sum_{k=0}^{n-1} \sum_{\ell=0}^{k} \sum_{i=0}^{n-k} \binom{k}{\ell} \binom{n-k-1}{i-1} \binom{n-k+\ell}{i} \frac{\ell\, a^k (u^2-1)^{k-\ell} u^{2i-1}}{n-k+\ell} + a^n u^{2n-1} \quad (3.16)$$

(where u weighs peaks and valleys.)

Lastly, we may enumerate our generalized model in section 2.2.4 with a similar expansion. From equation (2.18) we have that:

$$D(a,c,r,s,t) = \frac{1}{1 - ast^2(c - r + rD(c,t))} = \sum_{k \geq 0} a^k s^k t^{2k} (c - r + rD(c,t))^k \quad (3.17)$$

And so,

$$\begin{aligned}
[t^{2n}]\, D(a,c,r,s,t) &= [t^{2n}] \sum_{k\geq 0} \sum_{\ell=0}^{k} \tbinom{k}{\ell} a^k s^k t^{2k} (c-r)^{k-\ell} r^\ell D(c,t)^\ell \\
&= [t^{2n}] \sum_{k\geq 0} \sum_{\ell=0}^{k} \tbinom{k}{\ell} a^k s^k (c-r)^{k-\ell} r^\ell \sum_{j \geq 0} N_j^{(\ell)} t^{2j+2k} \\
&= [t^{2n}] \sum_{k\geq 0} \sum_{\ell=0}^{k} \sum_{j\geq 0} \sum_{i=0}^{k-\ell} \tbinom{k}{\ell} \tbinom{k-\ell}{i} (-1)^{k-\ell-i} a^k s^k c^i r^{k-i} N_{j-k}^{(\ell)} t^{2j} \\
&= \sum_{k\geq 0}^{n-1} \sum_{\ell=0}^{k} \sum_{i=0}^{k-\ell} \sum_{m=0}^{n-k} \tbinom{k}{\ell} \tbinom{k-\ell}{i} \tbinom{n-k-1}{m-1} \tbinom{n-k+\ell}{m} (-1)^{k-\ell-i} \frac{\ell\, a^k s^k c^{i+m} r^{k-i}}{n-k+\ell} + (asc)^n
\end{aligned} \quad (3.18)$$

In general, we see that many models of polymer adsorption with (multivariate) generating functions can be enumerated with convolution identities.

3.4 Enumerating alternating Dyck path returns

We can attempt to expand equation (2.25) with convolution identities similarly to our other generating functions. Consider first, for simplicity, D_{even} (equation (2.22)). There are many different ways to factorize this model to attempt to enumerate it.

Firstly, factoring the denominator finds that:

$$\begin{aligned} D_{even}(a,b,t) &= \frac{1-bt^2D_2}{1-(a+b)t^2D_2-abt^4(D_4^2-D_2^2)} \\ &= \frac{1-bt^2D_2}{1-(a+b)t^2D_2-abt^4(D_4-D_2)(D_4+D_2)} \\ &= \frac{1-bt^2D_2}{1-(a+b)t^2D_2-abt^4D_{-1}D} \end{aligned} \tag{3.19}$$

where D_{-1} is the generating function of alternating sign Dyck paths:

$$\begin{aligned} D_{-1} &= \frac{-1+\sqrt{1+4t^2}}{2t^2} \\ &= \sum_{n\geq 0}(-1)^n\, C_n\, t^{2n} = 1-t^2+2t^4-5t^6+14t^8+\dots \end{aligned} \tag{3.20}$$

We can expand equation (3.19) to find:

$$\begin{aligned} [t^{2n}]D_{even}(a,b,t) &= (1-bt^2D_2)\sum_{k\geq 0}\sum_{\ell=0}^{k}\binom{k}{\ell}(a+b)^{k-\ell}t^{2(k-\ell)}D_2^{(k-\ell)}a^\ell b^\ell t^{4\ell}D_{-1}^{(\ell)}D^{(\ell)} \\ &= (1-bt^2D_2)\sum_{k\geq 0}\sum_{\ell=0}^{k}\binom{k}{\ell}a^\ell b^\ell(a+b)^{k-\ell}t^{2k+2\ell}D_2^{(k-\ell)}D_{-1}^{(\ell)}D^{(\ell)} \end{aligned} \tag{3.21}$$

which is therefore the partition function of this model. A few facts are in order to consider the tractability of reducing this expansion.

Firstly, it may be shown trivially that: [10]

$$D_{-1}^{(s)} = \left(\frac{-1+\sqrt{1+4t^2}}{2t^2}\right)^s$$
$$= \sum_{n\geq 0}(-1)^n\, C_n^{(s)}\, t^{2n} = \sum_{n\geq 0}\binom{s-n}{n}\frac{s}{s-n}\, t^{2n} \tag{3.22}$$

We can also express D_2 as:

$$D_2 = \frac{(D - D_{-1})}{2} \implies D_2^{(s)} = \frac{1}{2^s}\sum_{i=0}^{s}\binom{s}{i} D^i D_{-1}^{s-i}(-1)^{s-i} \tag{3.23}$$

Thus, to enumerate this model, we require the coefficients to one the following convolutions:

1. $D_2^{(r)}(D_4^2 - D_2^2)^{(s)}$
2. $D_2^{(r)}(D_{-1}D)^{(s)}$
3. $D_{-1}^{(r)}D^{(s)}$

In the first case, the convolution of odd Catalan numbers is not known to be reducible:

$$\sum_{k=0}^{n}\binom{4k+2}{2k+1}\frac{1}{2k+2}\binom{4(n-k)+2}{2(n-k)+1}\frac{1}{2(n-k)+2} \tag{3.24}$$

Although, interestingly, we can show that:

$$C_{2n+1} = \binom{4n+2}{2n+1}\frac{1}{2n+2} = \binom{4n+2}{2n}\frac{2}{4n+2} = \binom{4n}{2n}\frac{4n+1}{(2n+1)(n+1)} = \frac{4n+1}{n+1}C_{2n} \tag{3.25}$$

this convolution is still not reducible to a closed form.

We may, however, find the first convolution of D_2 from its recurrence (in equation 2.20):

$$\begin{aligned}
[t^{2n}]\, D_2^2 &= [t^{2n}]\, t^{-2}\, D_2 - [t^{2n}]\, D_4^2 \\
[t^{2n}]\, D_2^2 &= [t^{2n+2}]\, D_2 - [t^{2n}]\, D_4^2 \\
[t^{2n}]\, D_2^2 &= \frac{1}{2n+3}\binom{4n+6}{2n+2} - 4^n C_n
\end{aligned} \tag{3.26}$$

where $[t^{2n}]\, D_4^2 = 4^n C_n$ is known as the *Shapiro convolution identity.* [13] The convolutions of D_2 therefore form a recurrence relation:

$$D_2^{(n+1)} = D_2^{(n)} - D_4^2 D_2^{(n-1)} \tag{3.27}$$

which requires an increasing variety of unknown convolutions to simplify. Therefore, the convolutions of D_2 do not meet our criteria of a reasonable closed formula.

Our alternative, then, is to convolute $D_{-1}^{(r)} D^{(s)}$. For $r = s = 1$, the coefficients are known: (OEIS A079489)

$$[t^{2n}] D_{-1} D = 2^{2n+1} C_n - C_{2n+1} \tag{3.28}$$

This however cannot be simplified in any reasonable way, and successive convolutions would continually increase the amount of terms in the expression (if reducible). Notice, also, that the latter term of equation (3.28) again requires the coefficients of $D_2^{(s)}$. If we instead evaluate with a computer algebra system, we only find hypergeometric expressions. The case $r \neq s$ is thus also intractable.

We lastly consider the alternate factorization:

$$D_{even} = \frac{1}{1 - at^2\left(\frac{D_2 + bt^2(D_4^2 - D_2^2)}{1 - bt^2 D_2}\right)} \tag{3.29}$$

which, barring the numerator, requires the convolutions of the convolutions of D_2.

Given that D_{even} and D_{odd} are distinct, it follows that this partition function is not readily reducible with convolutions, and we must attempt to find it another way.

Chapter 4

Directed paths and the Bethe Ansatz

4.1 Introduction

The more general solution to a polymer model can be obtained by deriving a partition function that permits raised endpoints, that is, with the start and end points of the walk not restricted to the boundary. Many lattice enumeration methods may solve such a problem, [14] and this may permit, for example, studying a polymer's behavior with one or both endpoints lifted.

One solution method is to solve the model's partial difference equations on a lattice. [15] A recurrence for the number of paths to each point on the lattice may be formulated easily for directed models, and when combined with a boundary condition, forms to create an expression that can be solved to provide a partition function with arbitrary endpoints. We illustrate the method by first enumerating return-weighted Dyck paths. [4]

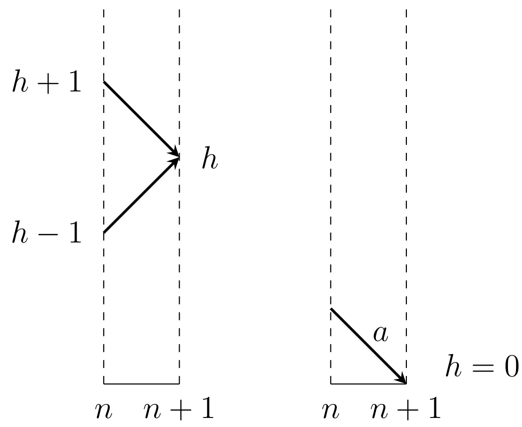

Figure 4.1: Lattice recurrence for Dyck paths

4.2 Lattice enumeration of directed path returns

The partial difference equations for directed paths with a return weight are: (See Figure 4.1)

$$d_n(h) = d_{n-1}(h-1) + d_{n-1}(h+1), \tag{4.1}$$

$$d_n(0) = a\, d_{n-1}(1), \tag{4.2}$$

$$d_0(h) = \delta_{h,0} \tag{4.3}$$

where (4.2) is the boundary condition and (4.3) is the initial condition (for δ the Kronecker delta function.)

These recurrences can be solved with a separation of variables approach:

$$r_n(h) := \Lambda^n\, \Upsilon_h \tag{4.4}$$

Substituting equation (4.4) into equation (4.1) and (4.2) requires that:

$$\Lambda \Upsilon_0 = a \Upsilon_1, \tag{4.5}$$

$$\Upsilon_{h+1} + \Upsilon_{h-1} = \Lambda \Upsilon_h \tag{4.6}$$

If we assume that $\Lambda = \zeta + \bar{\zeta}$ (which represents an up and a down step on the lattice), equation (4.6) is satisfied by either $\Upsilon_h = \zeta^h = e^{ikh}$ or $\Upsilon_h = \bar{\zeta}^h = e^{-ikh}$ (for $k \in [0, 2\pi)$). To satisfy (4.5), we try a *Bethe Ansatz*:

$$\Upsilon_h = A_1 \, \zeta^h + A_2 \, \bar{\zeta}^h \tag{4.7}$$

which assumes that our solution is a linear combination of paths translated upward and downward symmetrically about the n-axis. Our solution is therefore:

$$d_n(h) = \int_C r_n(h) \, dk = \int_C \Lambda^n \, (A_1 \, \zeta^h + A_2 \, \bar{\zeta}^h) \, dk \tag{4.8}$$

for C a unit circle in the complex plane.

It remains to fix Λ, A_1 and A_2. Λ can be fixed by substituting equation (4.8) into (4.1):

$$\Lambda = \zeta + \bar{\zeta} \tag{4.9}$$

which is the *bulk* (no boundaries) solution.

The constants A_1 and A_2 are now solved for. Substituting (4.8) into the boundary condition obtains:

$$\Lambda^n (A_1 + A_2) = a \, \Lambda^{n-1} (A_1 \, \zeta + A_2 \, \bar{\zeta}) \tag{4.10}$$

From which we express the ratio of the constants as:

$$\frac{A_1}{A_2} = -\frac{\Lambda - a\bar{\zeta}}{\Lambda - a\zeta} = -\frac{\zeta + \bar{\zeta} - a \, \bar{\zeta}}{\zeta + \bar{\zeta} - a \, \zeta} \tag{4.11}$$

This factors our solution as:

$$d_n(h) = A_2(k) \, \Lambda^n \, (\bar{\zeta}^h + \frac{A_1}{A_2} \, \zeta^h) \tag{4.12}$$

$A_2(k)$ can be fixed by bijecting our solution to paths that step in reverse. Taking $k \to -k$ implies that $h \to h_0$ and $\zeta \to \bar{\zeta}$ (where h_0 is the starting height.) This counts the same number of paths and can be written as:

$$d_n(h_0) = A_2(-k)\,\Lambda^n\,(\zeta^{h_0} + \frac{A_2}{A_1}\bar{\zeta}^{h_0}) \tag{4.13}$$

Comparing to equation (4.12) finds:

$$A_2 = C_0\,(\zeta^{h_0} + \frac{A_2}{A_1}\bar{\zeta}^{h_0}) \tag{4.14}$$

And therefore, we can substitute A_2 to obtain the general solution:

$$\begin{aligned} d_n(h_n, h_0) &= C_0 \int_C \Lambda^n(\zeta^{h-h_0} + \frac{A_1}{A_2}\,\zeta^{h+h_0})(\zeta^{h_0-h} + \frac{A_2}{A_1}\,\zeta^{h+h_0})\,dk \\ &= C_0 \int_C \Lambda^n(\zeta^{h-h_0} + \frac{A_1}{A_2}\,\zeta^{h+h_0} + \bar{\zeta}^{h-h_0} + \frac{A_2}{A_1}\,\bar{\zeta}^{h+h_0})\,dk \end{aligned} \tag{4.15}$$

which we simplify by absorbing a factor of two of this symmetric integral into the constant:

$$d_n(h_n, h_0) = C_0 \int_C \Lambda^n(\zeta^{h-h_0} + \frac{A_1}{A_2}\,\zeta^{h+h_0})\,dk \tag{4.16}$$

Notice that equation (4.16) is a Laurent series in ζ about 0. Therefore, by the residue theorem, our solution is the constant term of $d_n(h, h_0)$. This can be represented with a constant term operator as:

$$d_n(h, h_0) = C_0\,\mathrm{CT}[(\zeta + \bar{\zeta})^n\,(\zeta^{h_n-h_0} - \left(\frac{\zeta + \bar{\zeta} - a\,\bar{\zeta}}{\zeta + \bar{\zeta} - a\,\zeta}\right)\zeta^{h_n+h_0})] \tag{4.17}$$

(and readily satisfies our third difference equation (4.3).)

Simplifying, we obtain:

$$d_n(h,h_0) = C_0\,\mathrm{CT}[(\zeta+\bar{\zeta})^n\,(\zeta^{h_n-h_0} - \left(\frac{(1-a)+\zeta^2}{1-\zeta^2(a-1)}\right)\zeta^{h_n+h_0})] \tag{4.18}$$

Walks with lowered endpoints are necessarily even length and occur when $h_n = h_0 = 0$. Thus,

$$d_{2n}(0,0) = C_0\,\mathrm{CT}[(\zeta+\bar{\zeta})^{2n}(1 - \frac{(1-a)+\zeta^2}{1-\zeta^2(a-1)})] \tag{4.19}$$

From which we extract the constant terms algebraically. Simplifying and expanding geometrically obtains:

$$d_{2n}(0,0) = a\,C_0 \sum_{\ell\geq 0}(a-1)^\ell\,\mathrm{CT}[\,(\zeta+\bar{\zeta})^{2n}(1-\zeta^2)\,\zeta^{2\ell}\,] \tag{4.20}$$

And a binomial expansion provides:

$$d_{2n}(0,0) = a\,C_0\sum_{\ell\geq 0}(a-1)^\ell\left(\mathrm{CT}[\sum_{j=0}^{2n}\binom{2n}{j}\zeta^{2n-2j+2\ell}\,] - \mathrm{CT}[\sum_{j=0}^{2n}\binom{2n}{j}\zeta^{2n-2j+2\ell+2}\,]\right) \tag{4.21}$$

which is constant when $j = n+\ell$ and $j = n+\ell+1$ respectively. This fixes j and simplifies to:

$$\begin{aligned} d_{2n}(0;0) &= \sum_{\ell=0}^{n}(a-1)^\ell\left(\binom{2n}{n+\ell} - \binom{2n}{n+\ell+1}\right) \\ &= \sum_{\ell=0}^{n}(a-1)^\ell\frac{2\ell+1}{n+\ell+1}\binom{2n}{n+\ell} \end{aligned} \tag{4.22}$$

where the summation is valid for $\ell \leq n$. (By convention the trivial path does not pick up a weight, so $d_0(0;0) = a$ fixed $C_0 = a^{-1}$.)

The partition function for arbitrary endpoints can be determined similarly. Equation

(4.18) may be factored as:

$$d_n(h,h_0) = \mathrm{CT}[(\zeta+\bar\zeta)^n(\zeta^{h-h_0}-\zeta^{h+h_0+2}) + (a-1)(\zeta+\bar\zeta)^n \frac{\zeta^{h+h_0}-\zeta^{h+h_0+4}}{1-\zeta^2(a-1)}] \tag{4.23}$$

and expanded to obtain the general solution:

$$\begin{aligned} d_n(h,h_0) &= \binom{n}{\frac{n+h-h_0}{2}} - \binom{n}{\frac{n+h+h_0}{2}+1} \\ &+ \sum_{\ell=0}^{(n-h-h_0)/2} \left(\binom{n}{\frac{n+h+h_0}{2}+\ell} - \binom{n}{\frac{n+h+h_0}{2}+\ell+2} \right) (a-1)^{\ell+1} \end{aligned} \tag{4.24}$$

Alternate factorizations can lead to equivalent counting formulas. For example, we can factor expression (4.19) as

$$d_{2n}(0;0) = \mathrm{CT}[\,(\zeta+\bar\zeta)^{2n}\,\frac{\bar\zeta-\zeta}{\zeta+\bar\zeta}\,\frac{1}{1-a\zeta/(\zeta+\bar\zeta)}\,] \tag{4.25}$$

to obtain

$$d_{2n}(0;0) = \sum_{\ell=0}^{2n-1} \frac{\ell\, a^\ell}{n} \binom{2n-\ell-1}{n-1} \tag{4.26}$$

In this way, we may form non-trivial identities by comparing equivalent partition functions:

$$\sum_{\ell=0}^{n} (a-1)^\ell \frac{2\ell+1}{n+\ell+1}\binom{2n}{n+\ell} = \sum_{\ell=0}^{2n-1} \frac{\ell\, a^\ell}{n}\binom{2n-\ell-1}{n-1} = \sum_{\ell=0}^{n} \frac{\ell\, a^\ell}{2n-\ell}\binom{2n-\ell}{n-\ell} \tag{4.27}$$

(where the latter partition function is equation (3.8).)

We can also use substitution to derive formulas that count simpler objects. For example, substituting $a=1$ into equation (4.24) provides an exact formula for the number of Dyck paths with raised endpoints:

$$d_n(h,h_0) = \binom{n}{\frac{n+h-h_0}{2}} - \binom{n}{\frac{n+h+h_0}{2}+1} \tag{4.28}$$

and fixing the starting height at the origin provides:

$$d_n(h,0) = \binom{n}{\frac{n+h}{2}} - \binom{n}{\frac{n+h}{2}+1} = \frac{2h+2}{n+h+2}\binom{n}{\frac{n+h}{2}} \tag{4.29}$$

which is the formula for the number of *ballot paths*. Notice, further, that substituting $h = h_0 = 0$ for walks of even length recovers the Catalan numbers.

4.3 Lattice enumeration of directed path double-rises

Attempting to enumerate our double-rises model (from section 2.2.2) with a Bethe Ansatz provides an instructive result. Consider, first, its recurrences (see Figure 4.2):

$$d_n(h) = d_{n-1}(h-1) + d_{n-1}(h+1) \tag{4.30}$$

$$d_n(2) = r\, d_{n-2}(0) + 2\, d_{n-2}(2) + d_{n-2}(4) \tag{4.31}$$

$$d_0(h) = \delta_{h,0} \tag{4.32}$$

After fixing the bulk, substituting a Bethe Ansatz into equation (4.31) simplifies prematurely to $A_1 = -A_2$. Thus, we see that substituting a Bethe Ansatz with two constants into a raised (simple) boundary condition will have a trivial solution. Alternate boundary conditions (such as one that includes an additional step and adsorbing weight) may enumerate this model, but this becomes complicated to solve because of increasing the number of terms in the recurrence.

Additionally, these recurrences are problematic for enumerating this model at arbitrary

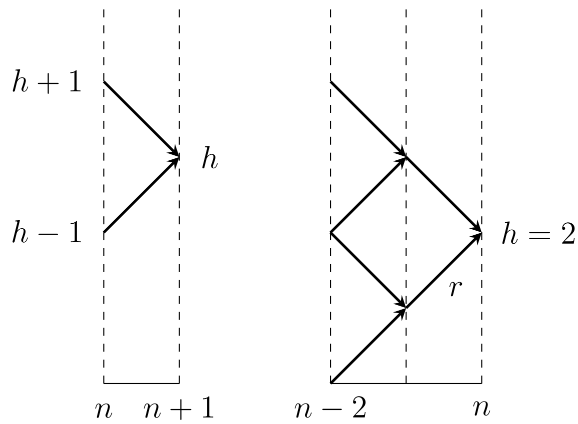

Figure 4.2: Lattice recurrence and boundary condition for Dyck paths with weighted double-rises

endpoint heights. Firstly, notice that we cannot end at an odd length with these recurrences. Further, notice, that these recurrences could not count double-rises that occur at an odd length. For example, a path that steps up and then forms a peak would not pick up a weight for double-rising. If we would attempt to fix this by including additional steps in the recurrences (of either an even or odd amount), we would still not be able to count the double-rises of a path of opposite parity. Despite this, the partition function for double-rises at even lengths could be obtained from these recurrences, and in principle, would enumerate a related model.

4.4 Lattice enumeration of directed path peaks

Since a single-step recurrence could not weigh a peak, we may attempt to enumerate directed path peaks by forming a two-step recurrence (see Figure 4.3). Similar to the recurrences for directed path double-rises, we must note that we are counting paths with peaks at even lengths.

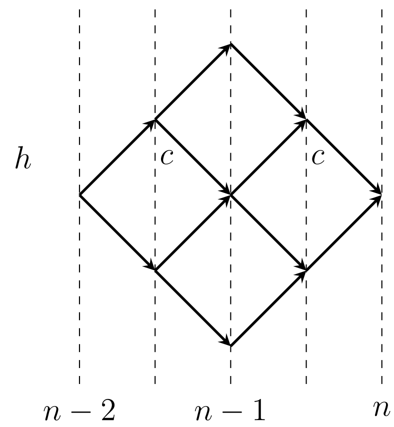

Figure 4.3: Lattice recurrence for Dyck paths with weighted peaks

The (two-step) recurrences for this model are:

$$d_n(h) = d_{n-1}(h-1) + (c+1)\, d_{n-1}(h) + d_{n-1}(h+1), \tag{4.33}$$

$$d_n(0) = ac\, d_{n-1}(0) + a\, d_{n-1}(1) \tag{4.34}$$

$$d_0(h) = \delta_{h,0} \tag{4.35}$$

which we solve with a similar approach to directed path returns. Equation (4.33) fixes our bulk to be:

$$r_n(h) = \Lambda^n (A_1\, \zeta^h + A_2\, \bar{\zeta}^h) = (\zeta + c + 1 + \bar{\zeta})^n (A_1\, \zeta^h + A_2\, \bar{\zeta}^h) \tag{4.36}$$

And the boundary condition fixes the ratio of the constants to be:

$$\frac{A_1}{A_2} = -\frac{\Lambda - ac - a\bar{\zeta}}{\Lambda - ac - a\zeta} \tag{4.37}$$

Our solution is thus:

$$d_n(h) = \mathrm{CT}[(\zeta + c + 1 + \bar{\zeta})^n\, (\zeta^{h_n - h_0} + \frac{A_1}{A_2}\, \zeta^{h_n + h_0})] \tag{4.38}$$

which we extract for walks of even length when $h = h_0 = 0$:

$$d_n(h) = \Lambda^n \left(\frac{a(\zeta^2 - 1)}{\zeta^2 + (1 + c - ac)\zeta + (1 - a)} \right) \tag{4.39}$$

Let

$$\zeta^2 + (1 + c - ac)\zeta + (1 - a) = (\zeta - \alpha)(\zeta - \beta) \tag{4.40}$$

with $\alpha\beta = 1 - a$, and $\alpha + \beta = ac - c - 1$. Expanding the bulk finds that:

$$d_n(0;0) = \sum_{k=0}^{n} \sum_{\ell=0}^{k} \binom{n}{k} \binom{k}{\ell} \zeta^{k-2\ell} (1 + c)^{n-k} \left(\frac{a(\zeta^2 - 1)}{\zeta^2 + (1 + c - ac)\zeta + (1 - a)} \right) \tag{4.41}$$

And expanding the latter term finds that:

$$1 + \frac{A_1}{A_2} = \frac{a(\zeta^2 - 1)}{\zeta^2} \sum_{p \geq 0} \sum_{q \geq 0} \left(\frac{\alpha}{\zeta} \right)^p \left(\frac{\beta}{\zeta} \right)^q \tag{4.42}$$

Thus,

$$d_n(0;0) = \sum_{k=0}^{n} \sum_{\ell=0}^{k} \binom{n}{k} \binom{k}{\ell} \zeta^{k-2\ell} (1 + c)^{n-k} \left(a \sum_{p \geq 0} \sum_{q \geq 0} \alpha^p \beta^q \zeta^{-p-q} - a \sum_{p \geq 0} \sum_{q \geq 0} \alpha^p \beta^q \zeta^{-p-q-2} \right) \tag{4.43}$$

Which we simplify by substituting $p + q = s$ and $p = t$:

$$d_n(0;0) = \sum_{k=0}^{n} \sum_{\ell=0}^{k} \binom{n}{k} \binom{k}{\ell} \zeta^{k-2\ell} (1 + c)^{n-k} a \left(\sum_{s \geq 0} \sum_{t=0}^{s} \alpha^t \beta^{s-t} \zeta^{-s} - \sum_{s \geq 0} \sum_{t=1}^{s-1} \alpha^{t-1} \beta^{s-t-1} \zeta^{-s} \right) \tag{4.44}$$

We now extract the constant term. ζ is constant when $s = k - 2\ell$, so:

$$\begin{aligned} d_n(0;0) =& a\sum_{k=0}^{n}\sum_{\ell=0}^{k/2}\sum_{t=0}^{k-2\ell}\binom{n}{k}\binom{k}{\ell}(1+c)^{n-k}\alpha^t\beta^{k-2l-t} \\ &-a\sum_{k=0}^{n}\sum_{\ell=0}^{k/2-1}\sum_{t=1}^{k-2\ell-1}\binom{n}{k}\binom{k}{\ell}(1+c)^{n-k}\alpha^{t-1}\beta^{k-2l-t-1} \end{aligned} \tag{4.45}$$

(with adjusted upper bounds so that our terms remain positive.) This can be simplified using the finite geometric series formula to obtain:

$$\begin{aligned} d_n(0;0) = & a\sum_{k=0}^{n}\sum_{\ell=0}^{k/2}\binom{n}{k}\binom{k}{\ell}(1+c)^{n-k}\left(\frac{\alpha^{k-2\ell+1}-\beta^{k-2\ell+1}}{\alpha-\beta}\right) \\ &-a\sum_{k=0}^{n}\sum_{\ell=0}^{k/2-1}\binom{n}{k}\binom{k}{\ell}(1+c)^{n-k}\left(\frac{\alpha^{k-2\ell-1}-\beta^{k-2\ell-1}}{\alpha-\beta}\right) \end{aligned} \tag{4.46}$$

4.5 Lattice enumeration of alternating directed path returns

We can attempt to extract the partition function alternating Dyck path returns (from section 2.2.5) with the Bethe Ansatz method.

Consider weights a and b for visits to the hard wall at points of parity $4n$ and $4n+2$ respectively. This provides the following recurrences:

$$d_n(h) = d_{n-1}(h-1) + d_{n-1}(h+1), \tag{4.47}$$

$$d_{4n}(0) = a\, d_{4n-1}(1), \tag{4.48}$$

$$d_{4n+2}(0) = b\, d_{4n+1}(1) \tag{4.49}$$

$$d_0(h) = \delta_{h,0} \tag{4.50}$$

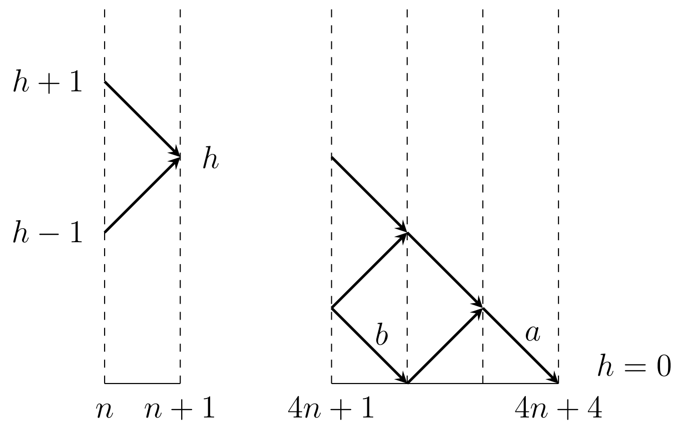

Figure 4.4: Lattice recurrences and boundary condition for alternating Dyck paths

We can attempt to combine the boundary conditions by extending the number of steps in the recurrence (see Figure 4.4)

$$d_{4n+4}(0) = ab\, d_{4n+1}(1) + a\, d_{4n+1}(1) + a\, d_{4n+1}(3) = a(b+1)\, d_{4n+1}(1) + a\, d_{4n+1}(3) \tag{4.51}$$

Equation (4.47) fixes our bulk:

$$r_n(h) = (\zeta + \bar{\zeta})^n \,(A_1\, \zeta^h + A_2\, \bar{\zeta}^h) \tag{4.52}$$

And substituting equation (4.52) into (4.51) fixes the ratio of the constants to be:

$$\frac{A_1}{A_2} = -\frac{(\zeta+\bar{\zeta})^3 - a\bar{\zeta}^3 - a(b+1)\,\bar{\zeta}}{(\zeta+\bar{\zeta})^3 - a\,\zeta^3 - a(b+1)\,\zeta} \tag{4.53}$$

Our solution is thus:

$$d_n(h) = C_0\, \mathrm{CT}[(\zeta+\bar{\zeta})^n \,(\zeta^{h_n - h_0} + \frac{A_1}{A_2}\, \zeta^{h_n + h_0})] \tag{4.54}$$

Which we extract the constant term of for walks of even length with $h = h_0 = 0$:

$$d_{2n}(0,0) = C_0\,\mathrm{CT}[(\zeta+\bar{\zeta})^{2n}(1+\frac{A_1}{A_2})] \tag{4.55}$$

Following a similar procedure to Dyck path returns (and letting $\beta = b+1$), we expand to extract the constant terms:

$$\begin{aligned}
d_{2n}(0,0) &= \mathrm{CT}[\Lambda^{2n-3}\zeta^{-3}(1+\beta\zeta^2-\beta\zeta^4-\zeta^6)\frac{1}{1-a\frac{\zeta^3+\beta\zeta}{\Lambda^3}}] \\
&= \sum_{m\geq 0} a^m\,\mathrm{CT}[\Lambda^{2n-3-3m}(\zeta^2+\beta)^m(\zeta^{-3+m}+\beta\zeta^{-1+m}-\beta\zeta^{1+m}-\zeta^{3+m})] \\
&= \sum_{m\geq 0} a^m\,\mathrm{CT}[\sum_{i=0}^{2n-3-3m}\tbinom{2n-3-3m}{i}\zeta^{2n-3-3m}\cdot \\
&\sum_{j=0}^{m}\tbinom{m}{j}\beta^j\zeta^{2m-2j}(\zeta^{-3+m}+\beta\zeta^{-1+m}-\beta\zeta^{1+m}-\zeta^{3+m})] \\
&= \sum_{m\geq 0}\sum_{i=0}^{2n-3-3m}\sum_{j=0}^{m}\tbinom{m}{j}a^m\beta^j\cdot \\
&\mathrm{CT}[\tbinom{2n-3-3m}{i}(\zeta^{2n-2j-2i-6}+\beta\zeta^{2n-2j-2i-4}-\beta\zeta^{2n-2j-2i-2}-\zeta^{2n-2j-2i})]
\end{aligned} \tag{4.56}$$

This simplifies to:

$$d_{2n}(0,0) = \sum_{m\geq 0}^{\lfloor\frac{2}{3}n-1\rfloor}\sum_{j=0}^{m}\tbinom{m}{j}a^m\beta^j\left(\tbinom{2n-3-3m}{n-j-3}+\beta\tbinom{2n-3-3m}{n-j-2}-\beta\tbinom{2n-3-3m}{n-j-1}-\tbinom{2n-3-3m}{n-j}\right) \tag{4.57}$$

for $n > 3$.

While not immediately apparent when comparing equations (4.19) and (4.25), the constant term factorization that absorbs the bulk from the denominator usually leads to more concise formulas. For example, if we had instead factorized equation (4.55) as:

$$1+\frac{A_1}{A_2} = a\,\frac{1+(b+1)\zeta^2-(b+1)\zeta^4-\zeta^6}{1+3\zeta^2(1+\alpha\,\zeta^2)(1+\beta\,\zeta^2)} \tag{4.58}$$

We would have instead obtained:

$$\begin{aligned}
d_{2n}(0,0) = (&\sum_{k=0}^{n}(-3)^k \sum_{u=0}^{k}\sum_{v=0}^{k}\binom{k}{u}\binom{k}{v}\binom{2n}{n-k-u-v}\alpha^u\beta^v) \\
+(b+1)(&\sum_{k=0}^{n-1}(-3)^k \sum_{u=0}^{k}\sum_{v=0}^{k}\binom{k}{u}\binom{k}{v}\binom{2n}{n-k-u-v-1}\alpha^u\beta^v) \\
-(b+1)(&\sum_{k=0}^{n-2}(-3)^k \sum_{u=0}^{k}\sum_{v=0}^{k}\binom{k}{u}\binom{k}{v}\binom{2n}{n-k-u-v-2}\alpha^u\beta^v) \\
-(&\sum_{k=0}^{n-3}(-3)^k \sum_{u=0}^{k}\sum_{v=0}^{k}\binom{k}{u}\binom{k}{v}\binom{2n}{n-k-u-v-3}\alpha^u\beta^v)
\end{aligned} \tag{4.59}$$

(for $\alpha,\ \beta = \frac{1}{2} - \frac{1}{6}a(b+1) \pm \frac{1}{6}\sqrt{a^2(b+1)^2 - 6a(b-1) - 3}$).

Upon enumeration, a computer algebra system finds that at certain walk lengths, equation (4.57) does not provide the expected polynomial coefficients (when compared to the generating function.) For example, we obtain the correct coefficient for t^8, but for t^6, we obtain $(ab^2 + 2ab + 2a)$ instead of $(ab^2 + ab + b^2 + 2b)$.

This is because the combined boundary condition does not count paths that end with a b, or paths that continue above the hard wall after picking up a b. It may be the case that extending our boundary condition to permit ending at height 2 (which would be an isolated b visit) would enumerate this model correctly (see Figure 4.5). However, this would be complicated to solve because it would increase the number of terms in our boundary condition. Thus, we see that enumerating models that have a complex boundary condition is difficult with the Bethe Ansatz method, and it remains an open problem to derive this model's partition function.

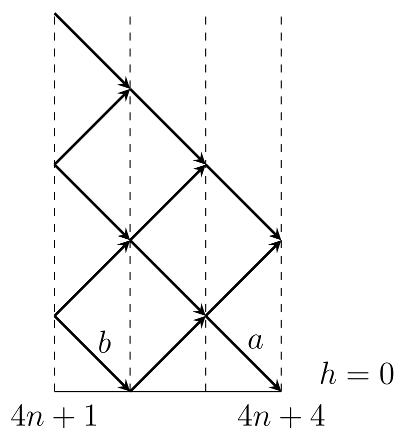

Figure 4.5: Hypothetical extended boundary condition for alternating Dyck path returns

Chapter 5

Conclusion

The focus of this work was deriving the generating function solution of a selection of polymer models, and then attempting to derive its partition function. While this was successful in most cases, it remains to extract the partition function of alternating Dyck paths, and the partition function for peaks at even lengths is likely to have a simpler solution (obtainable by an alternate extraction of the constant term.) Further, the scope of enumeration methods other than the ones demonstrated in this thesis should be explored for polymer modeling, and they may be able to solve the remaining polymer models.

We were also able to extract a novel model of peaks at even length with a Bethe Ansatz. Most generally, we have learned that the method has many constraints for directed paths, and struggles to form solvable expressions for most of the models presented.

Although the majority of the partition functions extracted with convolutions were already known in a combinatorial context, the use of binomial expansions produced novel formulas that were extended to include adsorption. It remains to explore which unsolved models are enumerable with simple convolution identities, and to develop the thermodynamic quantities and phase planes of the ones we have solved.

References

[1] James Mark Harry Allcock Frederick Lampe. *Contemporary Polymer Chemistry*. Pearson Education, 2003, p. 21. ISBN: 978-0-13-065056-6.

[2] Iwan Jensen. *A new transfer-matrix algorithm for exact enumerations: self-avoiding walks on the square lattice*. 2013. DOI: 10.48550/ARXIV.1309.6709. URL: https://arxiv.org/abs/1309.6709.

[3] Gordon Slade Neal Madras. *The Self-Avoiding Walk*. Birkhäuser, 1996.

[4] E.J. Janse van Rensburg. *The Statistical Mechanics of Interacting Walks, Polygons, Animals and Vesicles*. 2nd ed. Oxford University Press, 2015.

[5] Oleg Sofrygin. *Counting the number of Dyck paths according to different types of parameters*. 2003. URL: people.brandeis.edu/~gessel/47a/sofrygin.pdf.

[6] Helmut Prodinger. *A walk in my lattice path garden*. 2021. DOI: 10.48550/ARXIV.2111.14797. URL: https://arxiv.org/abs/2111.14797.

[7] R. Kemp. "On the Average Oscillation of a Stack". In: *Combinatorica* 2.2 (1982), pp. 157–176.

[8] N. J. A. Sloane. *The On-Line Encyclopedia of Integer Sequences*. URL: oeis.org.

[9] Paul Zeitz. *The Art and Craft of Problem Solving*. 2nd ed. Wiley, 2007.

[10] Oren Patashnik Ronald Graham Donald Knuth. *Concrete Mathematics*. 2nd ed. Addison–Wesley, 1994.

[11] Drew Armstrong. *Catalan Numbers: From EGS to BEG*. 2015. URL: https://www.math.miami.edu/~armstrong/Talks/catalan_at_mit.pdf.

[12] Emeric Deutsch. "Dyck path enumeration". In: *Discrete Mathematics* 204 (1999), pp. 167–202. DOI: 10.1016/S0012-365X(98)00371-9.

[13] Alon Regev. *Enumerating triangulations by parallel diagonals*. 2012. DOI: 10.48550/ARXIV.1208.3915. URL: https://arxiv.org/abs/1208.3915.

[14] Katherine Humphreys. "A history and a survey of lattice path enumeration". In: *Journal of Statistical Planning and Inference* 140.8 (2010), pp. 2237–225. DOI: 10.1016/j.jspi.2010.01.020.

[15] R Brak et al. "Lattice Paths and the Constant Term". In: *J. Phys: Conf. Ser.* 42 (2006), pp. 47–58. DOI: 10.1088/1742-6596/42/1/006.