LATTICE PATH ENUMERATION

NEIL KOLEKAR

ABSTRACT. In this paper we discuss various lattice path models and derive useful information about algebraic objects that enumerate them. In particular, we prove results on explicit forms, algebraicity, and the D-finiteness of generating functions that enumerate these paths, and occasionally compute closed-form sequences that count them.

1. INTRODUCTION

1.1. **Motivation.** The study of lattice paths and their enumeration comes a long way, from the works of Leibniz and De Moivre, to being a well-studied combinatorial subject worked on by some of the best combinatorialists from the last half-century, such as Banderier, Flajolet, Gessel, Bousquet-Mélou, and many others.

To simply put it, a d-dimensional *lattice path* (or *lattice walk*) is a sequence of *lattice points* in $\mathbb{R}^d$, with d-dimensional vectors that join consecutive points in the sequence (called *steps*). One of the most elementary and well-known examples of lattice paths are *Dyck paths*; these are paths from $(-n, 0)$ to $(n, 0)$ not going below the x-axis where the steps are in $\{\langle 1, 1\rangle, \langle 1, -1\rangle\}$. Dyck paths are enumerated by the Catalan numbers, given by $C_n := \frac{1}{n+1}\binom{2n}{n}$.

There are numerous ways of deriving the enumeration of Dyck paths. One common method is to recursively compute C_n by strong induction on n, whence we approach the recursion

$$C_{n+1} = \sum_{i=0}^{n} C_i \cdot C_{n-i}.$$

As in most enumerative combinatorics problems, we find a generating function that enumerates these lattice paths. The reason for its involvement is that generating functions can be manipulated in ways that combinatorial quantities cannot. As such, the generating function associated with Dyck paths is

$$\sum_{i=0}^{\infty} C_n x^n = \frac{1 - [1 - 4x^2]^{1/2}}{2x}.$$

Unfortunately, most lattice path models cannot be easily enumerated by elementary methods, such as bijections and recursions. However, by the means of manipulation of generating functions, we can extract much information about the sequences that enumerate various lattice path models. This "information" can be things like explicit forms of sequences for enumeration, explicit forms of their corresponding generating functions, determination of algebraicity or transcendence of generating functions, D-finiteness of generating functions, et cetera.

1.2. **Prerequisites.**

Date: July 2023.

1.2.1. *Generating functions as combinatorial objects.* As a final prerequisite, we discuss how generating functions are used as a way to enumerate various combinatorial objects.

One of the key ways generating functions enumerate combinatorial objects is that their coefficients correspond to sequences that enumerate these objects. In particular, the operations of multiplication, exponentiation, and diagonalization have combinatorial significance.

Example (Distributions). Show that the number of solutions in nonnegative integers to $a_1+\ldots+a_k = n$ is $\binom{n+k-1}{k-1}$.

Proof. The generating function for the number of solutions is $f(x) := (1+x+x^2+\ldots)^k = (1-x)^{-k}$, since upon expansion, each monomial $\prod_{i=1}^k x^{a_i}$ has its exponents forming a solution to $a_1+\ldots+a_k = n$ in nonnegative integers. From routine power series manipulation (or simply use of the extended Binomial Theorem), it is easy to see that the coefficient of x^n is f is indeed $\binom{n+k-1}{k-1}$. ∎

1.2.2. *Analytic prerequisites.* Although we defined generating functions in the univariate sense, we can naturally extend this to several variables. The cases of 2 and 3 variables are of particular interest in this extension.

Definition 1.1 (Generating Functions)**.** The *generating function* of a d-dimensional sequence $\{a_{\mathbf{v}}\}_{\mathbf{v}\in\mathbb{Z}^d_{\geq 0}}$ is the d-variate power series

$$F(x_1, \ldots, x_d) = \sum_{\langle j_1,\ldots,j_d\rangle\in\mathbb{Z}^d_{\geq 0}} a_{\langle j_1,\ldots,j_d\rangle} x_1^{j_1}\cdots x_d^{j_d}.$$

In general, the notation $\langle x_1^{j_1}\cdots x_d^{j_d}\rangle F(x_1,\ldots,x_d)$ consistent with the above form of F is a reference to the coefficient $a_{\langle j_1,\ldots,j_d\rangle}$.

Definition 1.2 (Diagonals)**.** Consider the generating function F from the above definition where $d = 2$ (that is, when F is bivariate). Then the diagonal $\triangle F$ of F is the univariate generating function given by

$$\triangle F(x) := \sum_{i\geq 0} a_{\langle i,i\rangle} x^i.$$

We can define algebraicity of Laurent series in a similar way we do for algebraic real numbers.

Definition 1.3 (Algebraicity)**.** We say that a Laurent series $F(X) = \{a_n\}_{n\geq -n_0}$ is *algebraic* over the field of characteristic 0 ($\mathbb{K}[X]$) if for some integer d there exists polynomials $A_0(X),\ldots,A_d(X)$ with coefficients in K and not all 0 such that

$$\sum_{i=0}^{d} A_i(X)F(X)^i = 0.$$

Laurent series that are not algebraic are called transcendental.

Definition 1.4 (D-finite functions)**.** A function $f = f(x)$ is called *D-finite* (or holonomic) if there exist polynomials $0 \neq a_0(x),\ldots,a_r(x) \in \mathbb{K}[X]$ such that

$$\sum_{k=0}^{r} a_k(x)f^{(k)}(x) = 0.$$

The following result is a technical fact about D-finiteness.

Lemma 1.5. *If $f(x_1, x_2)$ is a D-finite generating function, then so is its diagonal $\triangle f(x)$.*

Theorem 1.6 (Lagrange inversion). *Let $G(t)$ be any element of $\mathbb{C}[[t]]$. Then the equation $f(y) = yG[f(y)]$ has a unique solution in $\mathbb{C}[[y]]$, and*

$$\langle y^n\rangle (f(y))^n = \frac{k}{n}\langle t^{n-k}\rangle (G(t))^n \qquad \forall n, k > 0 \tag{1.1}$$

$$\langle y^n\rangle \frac{(f(y))^k}{1 - yG'(f(y))} = \langle t^{n-k}\rangle (G(t))^n \qquad \forall n, k \geq 0 \tag{1.2}$$

1.3. **Overview.** In this exposition we discuss:

- Kernel methods in lattice path enumeration: the conventional methods and the algebraic kernel method,
- An application of results on Laurent series factorization to enumerating lattice paths which have a restriction with respect to the line $y = x$,
- Lattice paths under a line of rational slope: a proof of Bizley's formula, the explicit closed-form and generating function for the number of up-right walks under a line of integer slope through the origin, and results on the walks under a line of slope 2/5.

2. Conventional Kernel Method

2.1. **Unrestricted models.** *Unrestricted models* are models in $\mathbb{Z}^d$ where there is not restriction on where the walk can go.

Define

$$W(\mathbf{z}, t) := \sum_{n\geq 0} \left(\sum_{\mathbf{i}\in\mathbb{Z}^d} f_{\mathbf{i},n}\mathbf{z}^{\mathbf{i}} \right) t^n,$$

where $f_{\mathbf{i},n}$ represents the number of walks in the lattice model of interest which consists of n steps and end at a lattice point $\mathbf{i} \in \mathbb{Z}^d$. The kernel method gives us a way of evaluating $W(\mathbf{z}, t)$.
We further define

$$W_n(\mathbf{z}) := \langle t^n\rangle W(\mathbf{z}, t) = \sum_{\mathbf{i}\in\mathbb{Z}^d} f_{\mathbf{i},n}\mathbf{z}^{\mathbf{i}}.$$

It is not difficult to see the recurrence

$$W_{n+1}(\mathbf{z}) = W_n(\mathbf{z}) \left(\sum_{\mathbf{i}\in\mathbb{Z}^d} \mathbf{z}^{\mathbf{i}} \right). \tag{2.1}$$

In fact, multiplying (2.1) by t^{n+1}, followed by summation over all nonnegative n returns

$$W(\mathbf{z}, t) - W_0(\mathbf{z}) = t \left(\sum_{\mathbf{i}\in\mathbb{Z}^d} \mathbf{z}^{\mathbf{i}} \right) W(\mathbf{z}, t),$$

rearranging to

$$\left(1 - t \left(\sum_{\mathbf{i}\in\mathbb{Z}^d} \mathbf{z}^{\mathbf{i}} \right) \right) W(\mathbf{z}, t) = \mathbf{z}^{\mathbf{p}}. \tag{2.2}$$

For the sake of future reference to expressions involved in the above equations, we define the following.

Definition 2.1 (Kernel). We denote $S(\mathbf{z})$ by

$$S(\mathbf{z}) = \sum_{\mathbf{i}\in\mathbb{Z}^d} \mathbf{z}^{\mathbf{i}}.$$

The *kernel* of this lattice path model (not necessarily restricted) is the expression

$$1 - tS(\mathbf{z}).$$

Now equation (2.2) rewrites as

$$K(\mathbf{z},t)\cdot W(\mathbf{z},t)=\mathbf{z}^{\mathbf{P}}. \tag{2.3}$$

Definition 2.2 (unrestricted kernel equation)**.** For unrestricted models, their *kernel equation* is equation 2.3: $K(\mathbf{z},t)\cdot W(\mathbf{z},t)=\mathbf{z}^{\mathbf{P}}$.

In the more complex models that follow, we derive similar kernel equations.

Example (walks starting and ending at origin). The number of walks that both start and end at the origin has its generating function given by

$$E(t)=[\mathbf{z}^0]W(\mathbf{z},t)=\triangle W(\mathbf{z},z_1z_2\cdots z_d t)=\triangle\left(\frac{\mathbf{z}^{\mathbf{P}}}{1-t(z_1\cdots z_d)\cdot S(\mathbf{z})}\right),$$

which is D-finite since diagonals preserve D-finiteness.

2.2. **One-dimensional excursions.** One-dimensional excursions are walks with steps in $\mathbb{Z}$ on the integer number line. Here we exemplify a kernel analysis more interesting than that of restricted models.

Theorem 2.3 (kernel equations for one-dimensional excursions)**.** *The generating function $E(t)$ that enumerates walks with steps in $\mathcal{S}$ which begin and end at the origin is given by*

$$E(t)=t\sum_{i=1}^{m}\frac{r_j'(t)}{r_j(t)},$$

where $r_1(t),\dots,r_m(t)$ are the small roots of $K(x,t):=1-tS(x)$ in x.

Proof. For sufficiently small positive t, the only small roots $r_j(t)$ of $K(x,t)$ lie inside $|x|=1$. In general, a root $r(t)$ of the kernel satisfies

$$S'(r(t))=-\frac{1}{t^2r'(t)} \tag{2.4}$$

by differentiating $S(r(t))=\frac{1}{t}$ with respect to t. By Cauchy's residue theorem,

$$E(t)=\sum_{j=1}^{m}\mathrm{Res}_{x=r_j(t)}\left(\frac{1}{K(x,t)}\right),$$

which evaluates to $t\sum_{i=1}^{m}\frac{r_j'(t)}{r_j(t)}$ upon application of equation (2.4), as desired. ■

2.3. **Walks in the Half-space.** Now we consider walks with step set $\mathcal{S}'\subset\mathbb{Z}^d$ restricted to half-space (i.e. $\mathbb{Z}^{d-1}\times\mathbb{N}$). As a way to count walks that can end at an arbitrary point, or walks ending on the hyperplane $z_d=0$, we can simply project all of the steps onto the dth coordinate. This induces a new *weighted* step set $\mathcal{S}$. Let $H(x,t)\in\mathbb{R}[x][[t]]$ be the power series defined by

$$H(x,t)=\sum_{n\geq 0}\sum_{i\geq 0}h_{i,n}x^it^n,$$

where $h_{i,n}$ enumerates the weighted half-space walks of length n on the steps in $\mathcal{S}$ ending at a point with x-coordinate i.

In a way similar to that in section 2.1, we may compute a recurrence for $H_n(x,t)$:

$$H_{n+1}(x)=S(x)H_n(x)-t\sum_{j=0}^{m-1}S_{<-j}(x)x^j\langle x^j\rangle H(x,t).$$

This induces the kernel equation

$$K(x,t)H(x,t) = 1 - t\sum_{j=0}^{m-1} S_{<-j}(x)x^j\langle x^j\rangle H(x,t),$$

where $K(x,t) = 1 - -tS(x)$ is the kernel for this model.

Example (Dyck paths and prefixes in half-space). Let $\mathcal{S} = \{1,1\}$ be the unweighted step set with characteristic polynomial $S(x) = x^{-1} + x$. By the kernel equation for walks in half-space,

$$\begin{aligned}(1 - t(x^{-1}+x))H(x,t) &= 1 - tx^{-1}\langle x^0\rangle H(x,t)\\ &= 1 - tx^{-1}H(0,t).\end{aligned}$$

We now solve for $H(0,t)$. First we solve the kernel $1 - t(x^{-1}+x)$ in x and obtain the small root

$$r_1(t) = \frac{1 - [1-4t^2]^{1/2}}{2t}$$

and the large root

$$R_1(t) = \frac{1 + [1-4t^2]^{1/2}}{2t}.$$

Since $r_1(t)$ s a power series with no constant term, it is applicable as a substitution for x in the kernel equation for this model. This substitution yields

$$H(0,t) = \frac{1 - [1-4t^2]^{1/2}}{2t^2} = \sum_{k=0}^{\infty} C_k t^{2k}.$$

Conceptually, the number of walks of length $2n$ with step set $\mathcal{S}$ is C_n.

3. The Quarter-plane Case: Algebraic Kernel Method

The quarter-plane lattice walks are perhaps the most popular and often used model. Here we consider walks on a weighted step set $\mathcal{S}' \subset \mathbb{Z}^d$ restricted to the *quarter-space* $\mathbb{Z}^{d-2}\times\mathbb{N}^2$. Rather than taking an arbitrary unweighted step set $\mathcal{S}$, we restrict our model to a *short step* model, where we add the restriction $\mathcal{S} \subset \{-1,1,0\}^d$. The primary difference between this new model and the models discussed in section 2 is that our kernel becomes a quadratic in each of its three variables.

Definition 3.1. The generating function that allows us to enumerate paths in the quarter-plane models is:

$$Q(x,y,t) = \sum_{i,j,n\geq 0} q_{i,j,n}x^iy^jt^n,$$

where $q_{i,j,n}$ denotes the number of walks in the quarter-plane of length n with steps in $\mathcal{S}$ which starts at the origin and ends at (i,j). We similarly define the characteristic polynomial

$$S(x,y) = \sum_{(i,j)\in\mathcal{S}} x^iy^j \in \mathbb{Z}[x^{-1},y^{-1},x,y].$$

The kernel is the function $K(x,y,t) = 1 - tS(x,y)$.

Theorem 3.2 (conventional kernel equation for quarter-plane models)**.** *The kernel equation for walks in the quarter-plane is given by*

$$xy(1-tS(x,y))Q(x,y,t) = xy - tI(y) - tJ(x) + \mathbf{1}_{(-1,-1)\in\mathcal{S}}tQ(0,0,t),$$

where

$$I(y) := y(\langle x^{-1}\rangle S(x,y))Q(0,y,t)$$

and

$$J(x) := x(\langle y^{-1}\rangle S(x,y))Q(x,0,t).$$

(Here our kernel is $K(t,x,y) = (1 - tS(x,y))$.)

The additional variable in the kernel for quadrant walks is a new conflict because we need to consider algebraic surfaces defined by $K(x,y,t) = 0$. To avoid this conflict, we introduce the *algebraic kernel method*, where we don't have to solve for the kernel. In the algebraic kernel, we utilize a *group model*, where we designate a group to the model. Specifically, we introduce a group $\mathcal{G}$ consisting of substitutions that fix K.

Definition 3.3 (group of short-step quarter-plane model)**.** The group $\mathcal{G}$ of the lattice path model determined by $\mathcal{S}$ is the group of transformations in the xy-plane that are generated by two involutions Ψ and Φ. These involutions are defined as follows: suppose that we can write

$$S(x,y) = xA_1(y) + A_0(y) + xA_{-1}(y) = yB1(x) + B0(x) + yB1(x).$$

There are 79 step sets that determine a unique lattice path model for the quadrant-plane models. Of these models, 23 of them have finite groups and 56 of them have infinite groups. We classify them, derive kernel equations – including a test for determining finiteness/infiniteness of corresponding groups – and overview a result on D-finiteness.

3.1. **Finite Group Models.** We start off with an example of how to find the groups for certain finite group models.

Example (N-S-E-W Quarter-plane walks). Let $\mathcal{S} = \{(\pm 1, 0), (0, \pm 1\}$. Then

$$S(x,y) = x^{-1} + x + y^{-1} + y.$$

By Theorem 3.1,

$$xy(1 - t(x^{-1} + x + y^{-1} + y))Q(x,y,t) = xy - tyQ(0,y) - txQ(x,0).$$

We can easily find $A_i(x)$ and $B_i(x)$ for $i \in \{\pm 1, 0\}$, from which $A_{-1}(x) = A_1(x)$ and $B_{-1}(x) = B_1(x)$. Thus $\Psi(x,y) = (x^{-1}, y)$ and $\Phi(x,y) = (x, y^{-1})$ generate $\mathcal{G}$.

Theorem 3.4 (orbit equation for finite group models)**.** *Suppose that the group $\mathcal{G}$ of a quarter-plane model is finite. Then,*

$$\sum_{g \in \mathcal{G}} sgn(g) g(xyQ(x,y,t)) = \frac{1}{1 - tS(x,y)} \sum_{g \in \mathcal{G}} sgn(g) g(xy).$$

Proof. Let $\Psi(x,y) = (X(x), y)$ and $\Phi(x,y) = (x, Y(y))$. Then, applying id, Ψ, and $\Phi \circ \Psi$ to Theorem 3.1, we obtain

$$\text{id} \Rightarrow xy(1 - tS(x,y))Q(x,y,t) = xy - tI(y) - tJ(x) + \mathbf{1}_{(-1,-1)\in\mathcal{S}} tQ(0,0,t) \tag{3.1}$$

$$\Psi \Rightarrow xy(1 - tS(X,y))Q(X,y,t) = Xy - tI(y) - tJ(X) + \mathbf{1}_{(-1,-1)\in\mathcal{S}} tQ(0,0,t) \tag{3.2}$$

$$\Phi \circ \Psi \Rightarrow XY(1 - tS(X,Y))Q(X,Y,t) = XY - tI(Y) - tJ(X) + \mathbf{1}_{(-1,-1)\in\mathcal{S}} tQ(0,0,t). \tag{3.3}$$

Note that both the terms $-tI(y)$ and $-tJ(X)$ are present in the RHS of both (3.2) and (3.3), so we can take an *alternating sum* of the three equations so both terms cancel. More generally, the maps Ψ and Φ each fix one coordinate, so considering the distinct compositions of group generators in an alternating fashion effectively cancels all functions of the form $I(Y_0)$ and $J(X_0)$, all of which appear on the right side. Since $\mathcal{G}$ is finite and has even order, taking an alternating composition of these generators will eventually return the identity, which gives rise to the desired equation. ■

4. Correlation with Factorization of Laurent Series

The factorization of Laurent series is extremely useful in analysis, to the extent that it gives us a proof of the Lagrange inversion formula (see [Ges80]). As such, [Ges80] notes that each Laurent series $f \in \mathbb{C}[[y, y/t]]$ with constant term 1 has a unique *decomposition* $f(t) = f_- f_0 f_+$, where f_- contains only negative powers of t (and constant term 1), f_+ contains only positive powers of t (and constant term 1), and f_0 is independent of t (and constant term 1). The following technical lemma, a variant of Lagrange inversion, is of significance.

Lemma 4.1 (Gessel). *Let $h(t, y) = \sum_{i=0}^{\infty} a_i t^i$, where $a_0, a_1 \in y\mathbb{C}[[y]]$ and $a_i \in \mathbb{C}[[y]]$. Then the equation $\alpha = h(\alpha, y)$ has a unique solution in $\alpha \in y\mathbb{C}[[y]]$. Define $f := [1 - t^{-1}h(t, y)]^{-1}$. Then in the decomposition of $f = f_- f_0 f_+$, we have $f_- = [1 - t^{-1}\alpha]^{-1}$ and $f_0 = \frac{\alpha}{a_0}$. Furthermore, for $k > 0$,*

$$\alpha^k = \sum_{n=1}^{\infty} \frac{k}{n} \langle t^{n-k} \rangle [h(t, y)]^n.$$

In this section, we give Lemma 4.1 a combinatorial counterpart via lattice paths and use well-known lattice path models to demonstrate examples of the lemma. Thus we have yet another approach to enumerating certain restricted walks.

Here we examine lattice paths over $\mathbb{Z}^2$, and let S be a step set over $\mathbb{Z}^2$. We say that the point (a, b) of a lattice path over $\mathbb{Z}^2$ has *height* $a - b$. Furthermore, the *endpoint* of a path π is the point at which it ends, denoted by $(e_x(\pi), e_y(\pi))$. (In addition, we also allow for the addition of an empty path, which consists of no points or steps.) We can also take products of paths; given two paths π_1 and π_2, their product $\pi_1\pi_2$ is the path π_1 followed by π_2.

Definition 4.2 (enumeration function). For a set P of lattice paths over $\mathbb{Z}^2$, we define its *enumeration function* by

$$\vartheta(P)(t, y) = \sum_{\pi \in P} t^{e_x(\pi) - e_y(\pi)} y^{e_y(\pi)}.$$

Furthermore, we write $\vartheta(p) = \vartheta(\{p\})$ where p is any path.

A basic property of Definition 4.2 is that

$$\vartheta(\pi_1\pi_2) = \vartheta(\pi_1)\vartheta(\pi_2) \tag{4.1}$$

for any two paths π_1 and π_2.

Now for a step set S, let S' be the set of all paths such that each path has its step set being a subset of S. In particular,

$$\vartheta(S') = \sum_{n=0}^{\infty} (\vartheta(S))^n = [1 - \vartheta(S)]^{-1}. \tag{4.2}$$

With the addition of the enumeration function as well as its multiplicative nature, we may define the Laurent series f_-, f_0, f_+ in this context. In order to do so, we need to consider three classes of paths over $\mathbb{Z}^2$ that start at the origin.

Definition 4.3 (sign classes). The sign classes consists of the following three classes of paths, all starting at the origin.

- A *minus-path* is either the empty path or a path such that its endpoint has a negative height less than that of any other point on the path.
- A *zero-path* is a path such that its endpoint has height zero and all other points on the path have nonnegative height.

- A *plus-path* is a path such that its endpoint has a positive height greater than that of any other point on the path.

The main advantage in defining these classes is that, under the injective map ϑ, minus-paths, zero-paths, and plus-paths are mapped to Laurent series containing only negative powers of t (and constant term 1), series independent of t (and constant term 1), and series containing only positive powers of t (and constant term 1), respectively.

Claim 4.4 (unique decomposition of paths)**.** *Let π be a lattice path over $\mathbb{Z}^2$ starting at the origin. Then π has a unique decomposition $\pi = \pi_-\pi_0\pi_+$, where π_- is a minus-path, π_0 is a zero-path, and π_+ is a plus-path.*

Proof. As a construction, let π_- be the path whose head ■

Theorem 4.5. *Let S be a step set, and let S_-, S_0, and S_+ be sets of minus-, zero-, and plus-paths with their step set being a subset of S. Then $\vartheta(S_-) = [\vartheta(S')]_-$, $\vartheta(S_0) = [\vartheta(S')]_0$, and $\vartheta(S_+) = [\vartheta(S')]_+$.*

Proof. From Claim 4.4 and equation (4.1), we have $\vartheta(S') = \vartheta(S_-)\vartheta(S_0)\vartheta(S_+)$. The desired statement follows from the fact that the decomposition a function in $\mathbb{C}[[y, y/t]]$ is unique. ■

Although Theorem 4.5 on its own seems quite powerful, we can find the explicit decomposition of functions in $y\mathbb{C}[[y, t/y]]$ as a consequence of Lemma 4.1, and then evaluate $\vartheta(S_-)$, $\vartheta(S_0)$, and $\vartheta(S_+)$ using it.

Corollary 4.6 (explicit decomposition)**.** *Let $f = f(t, y) = (1 - t - t^{-1}y - z)^{-1}$, where $z \in y\mathbb{C}[[y]]$. Then,*

(4.3)
$$f_- = [1 - t^{-1} \cdot \frac{1 - z - [(1-z)^2 - 4y]^{1/2}}{2}]^{-1} = 1 + \sum_{k=1}^{\infty}\sum_{n=0}^{\infty}\sum_{j=0}^{\infty} t^{-k}y^{n+k}z^j \frac{k}{2n+j+k}\binom{2n+j+k}{n+k, n, j};$$

$$f_0 = \frac{1 - z - [(1-z)^2 - 4y]^{1/2}}{2y} = 1 + \sum_{n=0}^{\infty}\sum_{j=0}^{\infty} y^n z^j \frac{1}{2n+j+k}\binom{2n+j+1}{n+1, n, j}; \tag{4.4}$$

$$f_+ = [1 - t \cdot \frac{1 - z - [(1-z)^2 - 4y]^{1/2}}{2y}]^{-1} = 1 + \sum_{k=1}^{\infty}\sum_{n=0}^{\infty}\sum_{j=0}^{\infty} t^k y^n z^j \frac{k}{2n+j+k}\binom{2n+j+k}{n+k, n, j}. \tag{4.5}$$

With the assistance of Corollary 4.6, we can compute both the closed-form of the sequences that enumerate certain lattice paths, and also find a closed-form for their generating functions.

Example (Catalan and Ballot numbers). Consider the step set $S = \{(0,1), (1,0)\}$. Now, $\vartheta(S') = [1 - t - t^{-1}y]^{-1}$. By Corollary 4.6, we evaluate $\vartheta(S_0)$ as

$$\vartheta(S_0) = \frac{1 - [1 - 4y]^{1/2}}{2y} = \sum_{n=0}^{\infty} \frac{\binom{2n}{n}}{n+1} y^n = \sum_{n=0}^{\infty} C_n y^n,$$

which returns the Catalan numbers C_n, and we evaluate $\vartheta(S_+)$ as

$$\vartheta(S_+) = \left(1 - t \cdot \frac{1 - z - [1 - 4y]^{1/2}}{2y}\right)^{-1} = 1 + \sum_{k=1}^{\infty}\sum_{n=0}^{\infty} t^k y^n \frac{k}{2n+k}\binom{2n+k}{n} = \sum_{k=0}^{\infty}\sum_{n=0}^{\infty} B_{k,n} t^k y^n,$$

returning the ballot numbers $B_{k,n}$.

Example (Motzkin numbers). Consider the step set $S = \{(2,0),(1,1),(0,2)\}$. Then, $\vartheta(S) = [1 - t^{-2}y^2 - y - t^2]^{-1}$, so

$$\begin{aligned}\vartheta(S_0) &= \frac{1 - y - (1-2y-3y^2)^{1/2}}{2y^2}\\ &= \sum_{n=0}\sum_{j=0} y^{2n+j}\frac{1}{2n+j+1}\binom{2n+j+1}{n+1,n,j},\\ &= \sum_{n=0}^{\infty} y^n \sum_{m\le\frac{n}{2}} \frac{1}{n+1}\binom{n+1}{m+1,m,n-2m}\end{aligned}$$

returning the Motzkin numbers M_n. (The Motzkin numbers M_n are the number of walks with step set $S = \{(2,0),(1,1),(0,2)\}$ from $(0,0)$ to (n,n) weakly below $y = x$.)

5. Walks under Lines of Rational Slope

5.1. **Integer slopes.** As a starting point, we solve the problem for the case of when the line has integer slope and passes through the origin. We first prove a polynomial functional equation in the generating function $f_k(x)$ that enumerates the paths under $y = kx$ for some positive integer k, hence proving its algebraicity. Furthermore, with the aid of the Lagrange inversion formula, we can evaluate its coefficients in closed-form.

In this section, we use denote $a_{n,k}$ by the number of walks from $(0,0)$ to (n,kn) weakly below the line $y = kx$, where k is a fixed positive integer.

Theorem 5.1 (integer slopes)**.** *Let*

$$f_k(t) := \sum_{n=0}^{\infty} a_{n,k}t^n.$$

Then $f_k(t)$ satisfies the functional equation $t\cdot f_k(t)^k - f_k(t) + 1 = 0$. Furthermore, $a_{n,k} = \frac{\binom{(k+1)n}{n}}{kn+1}$.

Proof. We present a probabilistic proof. In this approach, we use the following random process.

Start at $(1,0)$, and every second flip a coin. Suppose the coin flips heads with probability p and tails with probability $1-p$. For each "head" occurrence, move 1 unit up, and for every "tail" occurrence, move 1 unit to the right. Finally, terminate the process once the endpoint of the path is on the line $y = kx$.

Using Chebyshev's inequality, it can be shown that the probability of the process terminating at some point is 1 when $p \in [\frac{k}{k+1}, 1]$. Thus the probability of the process terminating for $p \in [\frac{k}{k+1}, 1]$ is

$$\sum_{n=0}^{\infty} a_{n,k}p^{nk+1}(1-p)^n = 1.$$

This rewrites as

$$f(p^k(1-p)) = \frac{1}{p}$$

for all $p \in [\frac{k}{k+1}, 1]$. Note that $g : [\frac{k}{k+1}, 1] \mapsto [0, \frac{k^k}{(k+1)^{k+1}}]$ defined by $g(x) = x^k(1-x)$ is a bijective map. Thus, for all $t \in [0, \frac{k^k}{(k+1)^{k+1}}]$, $t \cdot f_k(t)^k - f_k(t) + 1 = 0$. ■

5.2. **The formulae of Bizley and Grossman.** In the 1950s, Bizley and Grossman independently computed two different formulae for the number of up-right paths from $(0,0)$ to (bn, an) weakly above $y = \frac{a}{b}x$. Bizley computed a generating function as the exponential of a power series whose coefficients are explicitly given, and Grossman derived, without proof, a summation over partitions of n. It is not difficult to derive Grossman's formula using casework; however, Bizley had a more instructive proof using the cycle lemma, followed by rewriting as a coefficient extraction from a generating function. We present Bizley's proof.

Theorem 5.2 (Bizley, 1954)**.** *Denote by $f(bn, an)$ the number of up-right paths from $(0,0)$ to (bn, an) weakly below $y = \frac{a}{b}x$. Then*

$$f(bn, an) = \langle t^n \rangle \exp \sum_{j=1}^{\infty} \frac{1}{j(a+b)} \binom{(a+b)j}{aj} t^j$$

Proof. In our proof, we let $f_j(bn, an)$ be the number of up-right paths from $(0,0)$ to (bn, an) weakly below $y = \frac{a}{b}x$ with j *contacts* (not including the origin). We call such a path a $\frac{a}{b}$*-path*. A *highest point* of a path is a point on the path that follows and precedes different steps. Additionally, we say that a *cyclic shift* of a path can be obtained by removing the section of the path from $(0,0)$ to some (r, s) on the path, and moving that section to the other end so that $(0,0)$ is moved to (bn, an). The key idea is that the highest points on the path are not affected by cyclic shifts, and there are exactly j cyclic shifts that transform a $\frac{a}{b}$-path with j highest points to a $\frac{a}{b}$-path with j contacts, vice versa. Thus the number of paths with t highest points is

$$\frac{1}{j} n(a+b) f_j(bn, an). \tag{5.1}$$

For notational purposes, define

$$F_n(a, b) := \frac{1}{n(a+b)} \binom{(a+b)n}{an}.$$

We sum equation (5.1) over $1 \le j \le k$ to obtain

$$F_n(a, b) = \sum_{j=1}^{n} \frac{1}{j} f_j(bn, an). \tag{5.2}$$

Let $g_j(bn, an)$ be the number of paths from $(0,0)$ to (bn, an) with j highest points. Then,

$$f_j(bn, an) = \sum g_{a_1}(bn, an) \cdot \ldots \cdot g_{a_j}(bn, an), \tag{5.3}$$

where the sum is taken over all j-tuples $(a_1, \ldots, a_j)$ of nonnegative integers with sum n. Equation (5.3) rewrites as

$$f_j(bn, an) = \langle t^n \rangle [P(t)]^j, \tag{5.4}$$

where

$$P(t) := \sum_{i=1}^{\infty} g_i(bn, an) t^n.$$

Hence,

$$F_n(a, b) = \langle t^n \rangle \sum_{i=1}^{n} \frac{1}{i} [P(t)]^i = \langle t^n \rangle \sum_{i=1}^{\infty} \frac{1}{i} [P(t)]^i = -\langle t^n \rangle \log(1 - P(t)). \tag{5.5}$$

Isolating $P(t)$, we obtain

$$P(t) = 1 - e^{-F_1 t - F_2 t^2 - \cdots}.$$

Now $f_j(bn, an) = \langle t^j\rangle(1 - e^{-F_1t-F_2t^2-\cdots})^j$, so

$$f(bn, an) = \sum_{j=1}^{n} \langle t^n\rangle(1 - e^{-F_1t-F_2t^2-\cdots})^j = \langle t^n\rangle e^{F_1t+F_2t^2+\cdots},$$

which proves the desired result. ■

5.3. **Slope of 2/5.** We approach the case of slope 2/5 based on [Cyr20]. A useful lemma when dealing with non-integer slopes is the following.

Lemma 5.3. *Let* $\alpha = \frac{a}{c}$ *and* $\beta = \frac{b}{c}$ *for positive integers* a, b, *and* c *with* $\gcd(a, b, c) = 1$. *Then there is a bijection that maps the up-right walks that begin at* $(0, 0)$ *and stay weakly below* $y = \alpha x + \beta$ *to the walks that begin at* $(0, b)$ *that have step set* $\{(1, a), (1, -c)\}$ *and are above the* x*-axis.*

Proof. Note that the affine transform $(x \quad y)^T \mapsto (x + y \quad ax - cy + b)^T$ has nonzero determinant. Furthermore, the condition $y < \alpha x + \beta$ is mapped to a new condition of the y-coordinate being positive. This is sufficient to imply the conclusion. ■

Now we consider up-right walks under lines of slope 2/5 and y-intercept $c/5$ for integral c. From Lemma 5.3, it suffices to consider paths with step set $\{(1, -2), (1, -5)\}$. Thus our characteristic polynomial is $S(x) = x^{-2} + x^5$. By the kernel equation for walks in half-space, followed by solving a linear system of equations, we have the following.

Theorem 5.4. *Consider walks in* $\mathbb{N}^2$ *with step set* $\{-2, +5\}$. *The number of such walks starting at altitude* 3 *and ending at altitude* 0 *is given by the generating function* $F(x)$, *and altitude* 4 *and ending at altitude* 1 *is given by the generating function* $G(x)$. *Let* $P(z) = z^{-2} + z^5$ *and the kernel* $1 - xP(z) = 0$ *have small roots (for* z*) of* $r_1(x)$ *and* $r_2(x)$. *We have*

$$F(x) = -\frac{r_1r_2(r_1^4 - r_2^4)}{x(r_1 - r_2)}$$

and

$$G(x) = \frac{r_1^6 - r_2^6}{x(r_1 - r_2)}.$$

Using a Computer Algebra System, the following may be derived.

Theorem 5.5. *Denote* A_n *and* B_n *by the number of paths beginning at the origin and ending at* $(5n-1, 2n-1)$ *that respectively stay weakly below* $y = (2/5)x+1/5$ *and stay weakly below* $y = (2/5)x$; *then we have*

$$A_n + B_n = \frac{2}{7n-1}\binom{7n-1}{2n}.$$

Acknowledgements

The author is thankful to Dr. Simon Rubinstein-Salzedo for the opportunity to write this paper, as well as Andrew Lin for his support and valuable feedback throughout the class.

References

[Ges80] Ira Gessel. "A factorization for formal Laurent series and lattice path enumeration". In: (1980).

[Cyr20] Michael Wallner Cyril Banderier. *The kernel method for lattice paths below a line of rational slope.* 2020.