Enumeration of lattice paths with respect to a linear boundary

by

Federico Firoozi

B.Math, University of Waterloo, 2020

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Science

in the
Department of Mathematics
Faculty of Science

Declaration of Committee

Name:	**Federico Firoozi**
Degree:	**Master of Science**
Thesis title:	**Enumeration of lattice paths with respect to a linear boundary**
Committee:	**Chair:** Tamon Stephen Professor, Mathematics
	Jonathan Jedwab Supervisor Professor, Mathematics
	Amarpreet Rattan Committee Member Associate Professor, Mathematics
	Marni Mishna Examiner Professor, Mathematics

Abstract

The history of the enumeration of finite lattice paths with respect to a linear boundary is rich with unexpected patterns and symmetries. Let a, b be coprime and let g be a positive integer. We count the number of lattice paths from the startpoint $(0, 0)$ to the endpoint (ga, gb) whose steps are restricted to $\{(1, 0), (0, 1)\}$, with respect to a variable k measuring how much of the path lies above the linear boundary joining the startpoint to the endpoint. A first setting takes $a = 1$ and takes k to be the number of $(0, 1)$ steps lying above the boundary. A 1949 result due to Chung and Feller for the case $b = 1$ shows that the number of paths is independent of k. Huq later showed that the same holds for all b. A second setting instead takes k to be the number of lattice points on the path that lie above the boundary. In this setting, let $N_k(g)$ be the set of lattice paths for fixed a, b; we wish to determine $|N_k(g)|$. Bizley found $|N_0(g)|$ explicitly in 1954. Firoozi, Marwendo, and Rattan recently showed that $|N_k(1)|$ is independent of k. We place both these results in a more general framework by deriving a closed form expression for $|N_k(g)|$, which is significantly more complicated than for the special cases $k = 0$ and $g = 1$. We find for each g that the value $|N_k(g)|$ is constant over each successive set of $a + b$ values of k. Our proof relies on finding an explicit bijection between a subset of $N_k(g)$ and the set $N_{k+1}(g)$. This leads to a recursion for $|N_k(g)|$ whose base case is given by Bizley's result. We use symmetric functions to show that the closed form expression satisfies the recursion.

Keywords: enumeration; lattice paths; flaws; bijection; recursion; symmetric functions

Acknowledgements

The work in this thesis would not have been possible without the immense support I received from my mentors and colleagues.

The main result of this thesis was inspired by a conjecture that was formulated in 2019. I would therefore like to thank my 2019 project partner, Taku Marwendo, for helping to discover this conjecture, and for keeping our spirits high during the project with his humour and high-quality memes. I also want to thank Amarpreet Rattan for his valuable mentorship during that project and in the years that followed; he has helped me to develop crucial intuition (especially with regard to the path enumeration formula), has provided detailed feedback, and has taught me some of the combinatorial techniques used in this thesis.

I owe an enormous debt of gratitude to my supervisor, Jonathan Jedwab, for his exceptional support and guidance during my time in the master's program. Jonathan's mentorship has been instrumental in developing my skills as a researcher; he has taught me to ask better questions, break complicated problems into easier ones, and communicate more effectively. One of the most valuable lessons that I have learned from Jonathan is the value of clear and articulate writing. Thanks to his incredible guidance and detailed feedback, I now ~~cringe every time I see a grammatical error~~ feel much more confident in communicating my ideas with clarity.

I would also like to acknowledge the following people for their contributions to my thesis defense: Shuxing Li and Jingzhou Na for providing valuable feedback during my preparation for the defense; Tamon Stephen for chairing the defense; and Marni Mishna for providing exceptionally detailed feedback regarding the thesis, both during and after the defense.

Finally, I would like to thank my parents, Fariborz Firoozi and Pina Viola, for their incredible support and for always encouraging me to pursue my passions. I would also like to express my appreciation for my wonderful partner, Kelly Wong, for her constant encouragement and care, and for always inspiring me to be the best version of myself.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

The lattice path shown in Figure 1.1 contains exactly five lattice points that lie above the linear boundary joining the startpoint $(0, 0)$ to the endpoint $(8, 6)$.

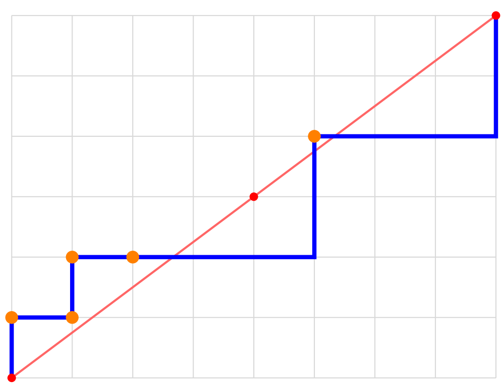


Figure 1.1: Lattice path from $(0, 0)$ to $(8, 6)$.

Let a, b be coprime and let g be a positive integer. Our objective is to count the number of lattice paths from the startpoint $(0, 0)$ to the endpoint (ga, gb), whose steps are restricted to $\{(1, 0), (0, 1)\}$, containing exactly k points that lie above the boundary.

1.1 Basic definitions

We are concerned only with *simple lattice paths*. These are paths in the two-dimensional lattice $\mathbb{Z}^2$ whose steps are restricted to the step set $\{(1, 0), (0, 1)\}$. We henceforth refer to these just as *paths*.

Let p be a path. The *boundary* of p is the line joining its startpoint to its endpoint. The path p *touches* or *contains* the lattice point $(x + i, y + j)$ if p starts at (x, y), and the first $i + j \geq 0$ steps of p consist of i of the $(1, 0)$ steps and j of the $(0, 1)$ steps (in any order).

We consider the points of p to be ordered according to increasing values of $i+j$. A point of p is a *flaw* if it lies strictly above the boundary of p. For example, the path in Figure 1.1 has the five flaws $(0,1),(1,1),(1,2),(2,2),(5,4)$, denoted in orange.

Definition 1.1.1. Let a, b be coprime and let g be a non-negative integer. Let $N(g)$ be the set of all paths from $(0,0)$ to (ga, gb), and $N_k(g)$ be the subset of such paths having exactly k flaws.

Note that the set $N(0) = N_0(0)$ contains only the *empty path* ε, namely the path consisting of no steps. If $g > 0$, then the values that k may take lie in the range $0 \leq k < g(a+b)$. The values $|N_0(g)|$ and $|N_{g(a+b)-1}(g)|$ were found by Bizley in 1954 [7], but no values $|N_k(g)|$ have since been determined for $0 < k < g(a+b) - 1$.

Central objective. For each allowable value of k, find an explicit formula for $|N_k(g)|$ for given a, b.

1.2 Numerical observations

There is a fundamental asymmetry in the definition of a flaw, namely that points of a path lying on the boundary are not flaws. We therefore might expect $|N_k(g)|$ to be non-increasing as k increases. Table 1.1 displays the numerical value of $|N_k(g)|$ for $g = 4$ and $(a, b) = (3, 2)$, obtained by computer program (see Appendix A). We note two apparent properties suggested by these values:

- P1 (Constant on blocks). The value $|N_k(g)|$ is constant on each of the g distinct blocks of $a + b$ consecutive values of k.
- P2 (Strictly decreasing). The value $|N_k(g)|$ is strictly decreasing between successive blocks.

We shall show that properties P1 and P2 both hold for all $g > 0$ and (a, b).

1.3 Overview of results and methods

Table 1.1 displays the value of $|N_k(4)| - |N_{k+1}(4)|$. The values in the table suggest a strategy for achieving our central objective: identify a subset $S_k(g)$ of $N_k(g)$ having cardinality $|N_k(g)| - |N_{k+1}(g)|$, and construct a bijection between $N_k(g) \setminus S_k(g)$ and $N_{k+1}(g)$. We achieve this in our main result (see Theorem 1.3.5). Properties P1 and P2 in Section 1.2 follow as consequences of the main result.

From now on, we regard a, b as fixed coprime integers.

1.3.1 The subset $S_k(g)$ of $N_k(g)$

We introduce some additional vocabulary in order to define the subset $S_k(g)$.

k	$\lvert N_k(4)\rvert$	$\lvert N_k(4)\rvert - \lvert N_{k+1}(4)\rvert$
0	7229	0
1	7229	0
2	7229	0
3	7229	0
4	7229	754
5	6475	0
6	6475	0
7	6475	0
8	6475	0
9	6475	437
10	6038	0
11	6038	0
12	6038	0
13	6038	0
14	6038	586
15	5452	0
16	5452	0
17	5452	0
18	5452	0
19	5452	

Table 1.1: Computer enumeration of $|N_k(4)|$ for $(a, b) = (3, 2)$.

Definition 1.3.1 (No flaws, max flaws)**.** A path in $N(g)$ has *no flaws* if it belongs to $N_0(g)$, and has *max flaws* if it belongs to $N_{g(a+b)-1}(g)$.

Definition 1.3.2 (Path concatenation)**.** Let p_1 and p_2 be paths. The *path concatenation* p_1p_2 is the path which takes all the (ordered) steps of p_1, and then takes all the (ordered) steps of p_2.

Definition 1.3.3 (Boundary points)**.** The *boundary points* of a path $p \in N(g)$ comprise the $\beta + 1$ lattice points of the form (ja, jb) that p contains (where j satisfies $0 \leq j \leq g$). We label the boundary points of p in order as $(0, 0) = B_0, B_1, \ldots, B_\beta = (ga, gb)$. If $\beta > 1$, (so that p contains boundary points other than $(0, 0)$ and (ga, gb)), then p is *boundary point touching (BPT)*.

A path $p \in N(g)$ contains at most $g + 1$ boundary points, because the number of lattice points lying on the boundary is $g + 1$ (see Figure 1.2). These boundary lattice points divide the boundary into g segments of equal length; note that property P1 refers to g distinct equally-sized blocks of consecutive values of k.

We may now define $S_k(g)$.

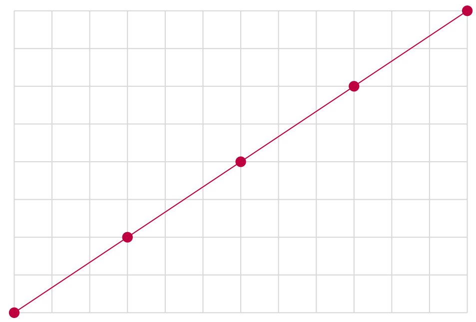


Figure 1.2: A path $p \in N(g)$ contains at most $g+1$ boundary points.

Definition 1.3.4 (Subset $S_k(g)$)**.** Let $S_k(g)$ be the subset of $N_k(g)$ comprising paths of the form p_1p_2 where, for some j satisfying $0 < j \le g$, we have $p_1 \in N_0(g-j)$ and $p_2 \in N_k(j)$ has max flaws. We write $S(g) \coloneqq \biguplus_k S_k(g)$ and $S \coloneqq \biguplus_{k,g} S_k(g)$.

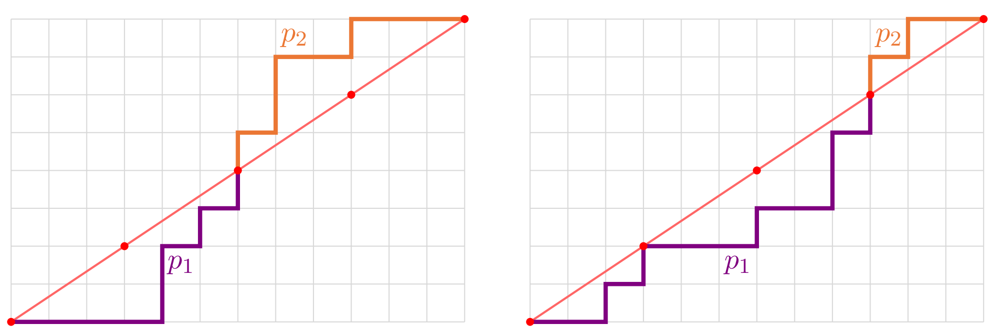
See Figure 1.3 for two example paths in $S(4)$. A path in $S(g)$ is a concatenation of a path p_1 from $(0,0)$ to $\big((g-j)a, (g-j)b\big)$ having no flaws with a path p_2 from $\big((g-j)a, (g-j)b\big)$ to (ga, gb) having max flaws. The condition on p_2 implies that $S_k(g)$ is empty unless $k = j(a+b)-1$ for some j satisfying $0 < j \le g$. Therefore

$$S_k(g) = \varnothing \quad \text{for } k \not\equiv -1 \pmod{a+b}, \tag{1.3.1}$$

and, for each j satisfying $0 < j \le g$,

$$S_{j(a+b)-1}(g) = \Big\{ p_1p_2 : p_1 \in N_0(g-j) \text{ and } p_2 \in N_{j(a+b)-1}(j) \Big\}. \tag{1.3.2}$$

Note that in Definition 1.3.4 the path p_1 may be empty and may be BPT; the path p_2 is non-empty and is not BPT.



(a) A path p_1p_2 in $S_9(4)$, where $p_1 \in N_0(2)$ and $p_2 \in N_9(2)$.

(b) A path p_1p_2 in $S_4(4)$, where $p_1 \in N_0(3)$ and $p_2 \in N_4(1)$.

Figure 1.3: Two example paths in $S(4)$ for $(a,b) = (3,2)$.

1.3.2 Main result and consequences

Theorem 1.3.5 (Main result)**.** *Let* g, k *satisfy* $0 \leq k < g(a+b) - 1$*. Then*

$$|N_k(g) \setminus S_k(g)| = |N_{k+1}(g)|.$$

We shall give a bijective proof of Theorem 1.3.5 in Chapter 3. The following result is a first consequence of Theorem 1.3.5.

Corollary 1.3.6 (Constant on blocks)**.** *For each* j *satisfying* $0 \leq j < g$*, we have*

$$|N_k(g)| = |N_{j(a+b)}(g)| \quad \textit{for all } k \textit{ in the range } j(a+b) \leq k < (j+1)(a+b).$$

Proof. The result follows directly from Theorem 1.3.5 and (1.3.1). □

Corollary 1.3.6 establishes property P1, proving a conjecture due to Firoozi, Marwendo, and Rattan [14]. In view of Corollary 1.3.6, we define

$$\mu_j(g) \coloneqq |N_{j(a+b)}(g)| \quad \text{for each } j \text{ satisfying } 0 \leq j < g.$$

We may then rephrase Corollary 1.3.6 as

$$|N_k(g)| = \begin{cases} \mu_0(g) & \text{if } 0 \leq k < a+b, \\ \mu_1(g) & \text{if } a+b \leq k < 2(a+b), \\ \vdots & \\ \mu_{g-1}(g) & \text{if } (g-1)(a+b) \leq k < g(a+b), \end{cases}$$

or more compactly as

$$|N_k(g)| = \mu_j(g) \quad \text{for all } j,k \text{ satisfying } 0 \leq j < g \text{ and } j(a+b) \leq k < (j+1)(a+b). \tag{1.3.3}$$

Combine (1.3.2) and (1.3.3) to give

$$|S_{j(a+b)-1}(g)| = \mu_0(g-j)\mu_{j-1}(j) \quad \text{for all } j \text{ satisfying } 0 < j < g. \tag{1.3.4}$$

We now observe two further consequences of Theorem 1.3.5.

Corollary 1.3.7 (Recurrence relation)**.** *We have*

$$\mu_{j-1}(g) - \mu_0(g-j)\mu_{j-1}(j) = \mu_j(g) \quad \textit{for each } j \textit{ satisfying } 0 < j < g. \tag{1.3.5}$$

Proof. Since $S_k(g)$ is a subset of $N_k(g)$, we have by Theorem 1.3.5 that

$$|N_k(g)| - |S_k(g)| = |N_{k+1}(g)|.$$

Let j satisfy $0 < j < g$. Take $k = j(a+b) - 1$ and use (1.3.3) and (1.3.4) to give (1.3.5). □

Corollary 1.3.8 (Strictly decreasing)**.** *We have* $\mu_0(g) > \mu_1(g) > \cdots > \mu_{g-1}(g)$.

Proof. This follows from Corollary 1.3.7, noting that $\mu_j(g) > 0$ for $0 \le j < g$ by (1.3.3). □

Corollary 1.3.8 establishes property P2.

1.3.3 The value of $\mu_j(g)$

Recall that our central objective is to find an explicit formula for $|N_k(g)|$ for each allowable value of k and for given (a, b), and that by (1.3.3) it is sufficient to determine the values $\mu_j(g)$. By inspection of Table 1.2, the recurrence relation (1.3.5) for $\mu_j(g)$ has a unique solution for each j, g satisfying $0 \le j < g$, provided the initial values $\mu_0(g)$ are known for all $g > 0$.

$\mu_0(1)$	$\mu_0(2)$	$\mu_0(3)$	$\mu_0(4)$	$\mu_0(5)$	$\cdots$
	$\mu_1(2)$	$\mu_1(3)$	$\mu_1(4)$	$\mu_1(5)$	$\cdots$
		$\mu_2(3)$	$\mu_2(4)$	$\mu_2(5)$	$\cdots$
			$\mu_3(4)$	$\mu_3(5)$	$\cdots$
				$\mu_4(5)$	$\cdots$
					$\cdots$

Table 1.2: The values $\mu_j(g)$ for $0 \le j < g$ can be determined one column at a time using the recurrence relation (1.3.5), provided the values $\mu_0(g)$ in the initial row are known.

The required initial values $\mu_0(g)$ are indeed known, as we describe in Corollary 1.3.12 after introducing some notation involving integer partitions.

Definition 1.3.9 (Integer partition)**.** Let $m_1, m_2, \ldots$ be non-negative integers and let $g = \sum_{i \ge 1} i m_i$. Write

$$\langle 1^{m_1} 2^{m_2} \cdots \rangle \vdash g$$

to denote the integer partition of g having m_i copies of the summand i.

Example 1.3.10. The integer partition $1+1+2+3$ of 7 is written $\langle 1^2 2^1 3^1 \rangle$.

We define the following quantities.

For $i > 0$, let

$$c_i := \frac{1}{i(a+b)} \binom{i(a+b)}{ia}. \tag{1.3.6}$$

For $g > 0$ and an integer partition $\lambda = \langle 1^{m_1} 2^{m_2} \cdots \rangle \vdash g$, let

$$C_\lambda := \prod_{i \geq 1} \frac{c_i^{m_i}}{m_i!}, \tag{1.3.7}$$

$$l(\lambda) := \sum_{i \geq 1} m_i,$$

$$\mathbf{E}_g := \sum_{\lambda \vdash g} (-1)^{g - l(\lambda)} C_\lambda, \tag{1.3.8}$$

$$\mathbf{H}_g := \sum_{\lambda \vdash g} C_\lambda, \tag{1.3.9}$$

and let

$$\mathbf{E}_0 := 1, \quad \mathbf{H}_0 := 1. \tag{1.3.10}$$

In 1954, Bizley found the values of $|N_0(g)|$ and $|N_{g(a+b)-1}(g)|$ explicitly in terms of $\mathbf{H}_g$ and $\mathbf{E}_g$.

Theorem 1.3.11 (Bizley [7])**.** *Let* $g > 0$*. Then*

$$|N_0(g)| = \mathbf{H}_g,$$
$$|N_{g(a+b)-1}(g)| = (-1)^{g+1} \mathbf{E}_g.$$

Using (1.3.3), we obtain the following values for $\mu_0(g)$ and $\mu_{g-1}(g)$.

Corollary 1.3.12 (Known values of $\mu_0(g)$ and $\mu_{g-1}(g)$)**.** *Let* $g > 0$*. Then*

$$\mu_0(g) = \mathbf{H}_g, \tag{1.3.11}$$

$$\mu_{g-1}(g) = (-1)^{g+1} \mathbf{E}_g. \tag{1.3.12}$$

The unique solution $\mu_j(g)$ determined by the recurrence relation (1.3.5) and Corollary 1.3.12 can now be stated explicitly.

Theorem 1.3.13 (Path enumeration formula)**.** *We have*

$$\mu_j(g) = \sum_{k=0}^{j} (-1)^k \mathbf{E}_k \mathbf{H}_{g-k} \quad \textit{for } 0 \leq j < g.$$

Proof. Let $g > 0$. By (1.3.5) and Corollary 1.3.12, for $0 < j < g$ we have

$$\mu_j(g) = \mu_{j-1}(g) + (-1)^j \mathbf{E}_j \mathbf{H}_{g-j}.$$

Therefore (by an implicit induction)

$$\begin{aligned}\mu_j(g) &= \mu_0(g) + \sum_{k=1}^{j}(-1)^k \mathbf{E}_k \mathbf{H}_{g-k} \\ &= \sum_{k=0}^{j}(-1)^k \mathbf{E}_k \mathbf{H}_{g-k}\end{aligned}$$

using (1.3.10) and (1.3.11). □

In Chapter 4, we give an alternative proof of Theorem 1.3.13 which does not assume that the value of $\mu_{g-1}(g)$ is known. This will require some results involving symmetric functions.

1.3.4 Computation using the path enumeration formula

Let $(a, b) = (3, 2)$ and $g = 4$. We illustrate the use of the path enumeration formula Theorem 1.3.13 to calculate the number $|N_k(4)|$ of paths from $(0, 0)$ to $(12, 8)$ having k flaws, for each k satisfying $0 \leq k < 20$. By (1.3.3), it is sufficient to determine $\mu_j(4)$ for each $j = 0, 1, 2, 3$.

We begin by listing the partitions of the integers $1, 2, 3, 4$.

$$\begin{aligned}&\text{Partitions of } 4 : \langle 4^1\rangle,\ \langle 1^1 3^1\rangle,\ \langle 2^2\rangle,\ \langle 1^2 2^1\rangle,\ \langle 1^4\rangle,\\ &\text{Partitions of } 3 : \langle 3^1\rangle,\ \langle 1^1 2^1\rangle,\ \langle 1^3\rangle,\\ &\text{Partitions of } 2 : \langle 2^1\rangle,\ \langle 1^2\rangle,\\ &\text{Partitions of } 1 : \langle 1^1\rangle.\end{aligned}$$

Using (1.3.6), we compute

$$c_1 = 2,\ c_2 = 21,\ c_3 = \frac{1001}{3},\ c_4 = \frac{12597}{2}.$$

Using (1.3.7), we then compute (for example)

$$C_{\langle 1^2 2^1\rangle} = \left(\frac{c_1^2}{2!}\right)\left(\frac{c_2^1}{1!}\right) = 42, \quad C_{\langle 1^3\rangle} = \left(\frac{c_1^3}{3!}\right) = \frac{4}{3}.$$

The full set of C_λ values is

$$\begin{aligned}&C_{\langle 4^1\rangle} = \frac{12597}{2}, \quad C_{\langle 1^1 3^1\rangle} = \frac{2002}{3}, \quad C_{\langle 2^2\rangle} = \frac{441}{2}, \quad C_{\langle 1^2 2^1\rangle} = 42, \quad C_{\langle 1^4\rangle} = \frac{2}{3},\\ &C_{\langle 3^1\rangle} = \frac{1001}{3}, \quad C_{\langle 1^1 2^1\rangle} = 42, \quad C_{\langle 1^3\rangle} = \frac{4}{3},\\ &C_{\langle 2^1\rangle} = 21, \quad C_{\langle 1^2\rangle} = 2,\\ &C_{\langle 1^1\rangle} = 2.\end{aligned}$$

Using (1.3.8) and (1.3.9), we next calculate that (for example)

$$\mathbf{E}_3 = (-1)^{3-1}C_{\langle 3^1\rangle} + (-1)^{3-2}C_{\langle 1^1 2^1\rangle} + (-1)^{3-3}C_{\langle 1^3\rangle} = \frac{1001}{3} - 42 + \frac{4}{3} = 293,$$
$$\mathbf{H}_3 = C_{\langle 3^1\rangle} + C_{\langle 1^1 2^1\rangle} + C_{\langle 1^3\rangle} = \frac{1001}{3} + 42 + \frac{4}{3} = 377.$$

The full set of $\mathbf{E}_k$ and $\mathbf{H}_k$ values is

$$\begin{aligned}
\mathbf{H}_4 &= 7229, & \mathbf{E}_4 &= -5452\\
\mathbf{H}_3 &= 377, & \mathbf{E}_3 &= 293\\
\mathbf{H}_2 &= 23, & \mathbf{E}_2 &= -19\\
\mathbf{H}_1 &= 2, & \mathbf{E}_1 &= 2\\
\mathbf{H}_0 &= 1, & \mathbf{E}_0 &= 1.
\end{aligned}$$

Using Theorem 1.3.13, we determine that

$$\begin{aligned}
&\mu_0(4) = \mathbf{E}_0\mathbf{H}_4 = 1 \cdot 7229 = 7229\\
&\mu_1(4) = \mathbf{E}_0\mathbf{H}_4 - \mathbf{E}_1\mathbf{H}_3 = 1 \cdot 7229 - 2 \cdot 377 = 6475\\
&\mu_2(4) = \mathbf{E}_0\mathbf{H}_4 - \mathbf{E}_1\mathbf{H}_3 + \mathbf{E}_2\mathbf{H}_2 = 1 \cdot 7229 - 2 \cdot 377 - 19 \cdot 23 = 6038\\
&\mu_3(4) = \mathbf{E}_0\mathbf{H}_4 - \mathbf{E}_1\mathbf{H}_3 + \mathbf{E}_2\mathbf{H}_2 - \mathbf{E}_3\mathbf{H}_1 = 1 \cdot 7229 - 2 \cdot 377 - 19 \cdot 23 - 293 \cdot 2 = 5452.
\end{aligned}$$

Using (1.3.3), we may now determine the value of $|N_k(4)|$ for each k satisfying $0 \leq k < 20$. The resulting values agree with the computer enumeration reported in Table 1.1.

Note that we may alternatively use (1.3.12) to compute the value

$$\mu_3(4) = (-1)^{4+1}\mathbf{E}_4 = 5452.$$

Remark 1.3.14. Bizley [7] noted that both $\mathbf{H}_k$ and $\mathbf{E}_k$ are integers because of the enumerations given in Theorem 1.3.11.

Remark 1.3.15. Although the quantity c_i defined in (1.3.6) is not necessarily an integer, it can be shown that ic_i is an integer.

Chapter 2

Background

2.1 Lattice path enumeration problems

In order to place our results in a wider context, we give a very brief review of the lattice path enumeration literature.

The study of lattice path enumeration has a long and rich history spanning hundreds of years [21, 26]. Two historical problems that can be phrased in terms of lattice paths are the 'gambler's ruin problem' [21] and Bertrand's 'ballot problem' [3].

A lattice path problem is usually constrained to lie in d dimensions [24, 29] and specifies a finite step set describing the allowable steps comprising the path [21, 36]. Examples of common small step sets include $\{(1,0),(0,1)\}$ and $\{(1,1),(1,-k)\}$ [22, 26]. We are concerned only with simple paths (those whose step set is $\{(1,0),(0,1)\}$) in the two-dimensional lattice $\mathbb{Z}^2$. Although asymptotic enumeration is a major topic in the study of lattice paths [4, 28, 29], our focus is on exact enumeration.

Lattice path enumeration problems often include a constraint that paths must satisfy with respect to a specified boundary, for example: remaining strictly on one side of the boundary; not crossing the boundary; or touching the boundary a specified number of times. The boundary is often linear [10, 22, 30] or piecewise linear [19, 23]. Linear boundaries of rational slope have been particularly studied [6, 7, 14, 16, 17, 26]. We are concerned with measuring how much of the path lies above a linear boundary having rational slope.

2.2 Methods of analysis

Many methods of analysis have been applied to the study of lattice path enumeration problems. These include generating functions [25, 34], Lagrange inversion [15], the kernel method [4, 6, 28, 29], and symmetric functions [5, 20].

A popular enumeration method is to construct an explicit bijection between two sets of interest. Particular examples include the reflection principle [1, 13, 17] (whose origin is often incorrectly attributed to André [32]), the cycle lemma [11, 12, 30], set partitions [31], and rearrangement of path segments [8, 9]. It is often possible to use multiple methods to solve the same enumeration problem, as demonstrated in [35].

We shall establish our main result (Theorem 1.3.5) by constructing a bijection, and our alternative proof of the path enumeration formula (Theorem 1.3.13) using symmetric functions.

2.3 Two settings

We use a variable k to measure how much of a simple path lies above the linear boundary joining the startpoint $(0, 0)$ to the endpoint (ga, gb) (where a, b are coprime). When the slope b/a of the boundary is an integer (so $a = 1$), we may take k to be the number of $(0, 1)$ steps lying above the boundary (see Figure 2.1), as discussed in Section 2.4. Such steps are called 'flaws' [9, 19, 34, 36].

When the slope b/a of the boundary is rational but non-integer, this definition of k is no longer appropriate because some $(0, 1)$ steps may lie partially above the boundary (see Figure 2.2). In this case, we instead take k to be the number of lattice points of the path that lie above the boundary, as discussed in Section 2.5. We shall use the same name 'flaws' for these lattice points, despite the change of setting.

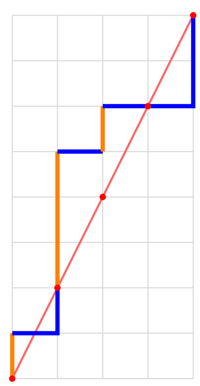
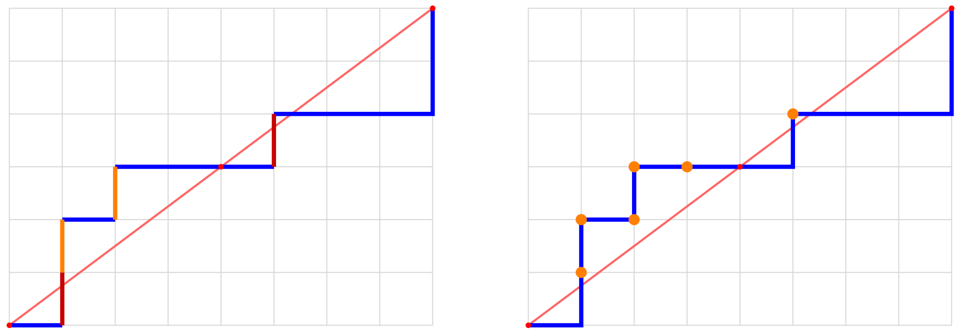


Figure 2.1: A path having 5 of the $(0, 1)$ steps lying above the boundary.

2.4 Boundaries of integer slope

In this section, we take k to be the number of $(0, 1)$ steps of a path from $(0, 0)$ to (g, gb) that lie above the boundary.



(a) This path has two of its $(0,1)$ steps lying partially above the boundary.

(b) The number of lattice points lying above the boundary is unambiguous.

Figure 2.2: A boundary of rational (non-integer) slope.

A classical result states that the number of paths from $(0,0)$ to (g,g) with $k=0$ (known as *Dyck* or *Catalan* paths) equals the g^{th} Catalan number

$$\frac{1}{g+1}\binom{2g}{g}.$$

Chung and Feller's influential 1949 work [10] showed that the same count applies for all k.

Theorem 2.4.1 (Chung-Feller [10, Theorem 2A])**.** *Let k satisfy $0 \leq k \leq g$. Then the number of paths from $(0,0)$ to (g,g) having k of the $(0,1)$ steps lying above the boundary equals*

$$\frac{1}{g+1}\binom{2g}{g}.$$

Theorem 2.4.1 can be proven using bijective methods [35].

Huq generalized Theorem 2.4.1 to paths whose endpoint is (g,gb).

Theorem 2.4.2 (Huq [22, Corollary 5.1.2])**.** *Let k satisfy $0 \leq k \leq gb$. Then the number of paths from $(0,0)$ to (g,gb) having k of the $(0,1)$ steps lying above the boundary equals*

$$\frac{1}{gb+1}\binom{(b+1)g}{g}.$$

Further variations on Theorem 2.4.1 have been found [19, 27, 34].

2.5 Boundaries of rational slope

In this section, we take k to be the number of lattice points of a path from $(0,0)$ to (ga, gb) that lie above the boundary. As in Definition 1.1.1, we let $N_k(g)$ be the appropriate set of lattice paths for fixed coprime a, b.

In 1950, Grossman [18] conjectured an explicit formula for the number $|N_0(g)|$ of paths from $(0,0)$ to (ga, gb) which lie weakly below the boundary (that is, which have no flaws). In 1954, Bizley [7, Eq. (10)] proved Grossman's formula using generating functions. Bizley [7, Eq. (8)] also obtained an explicit formula for the number of paths which lie strictly below the boundary (that is, which have no flaws and are not BPT). Since this second set is in bijection with the set of paths lying strictly above the boundary (via rotation), this result gives the value $|N_{g(a+b)-1}(g)|$. The values $|N_0(g)|$ and $|N_{g(a+b)-1}(g)|$ are stated in Theorem 1.3.11. In the nearly 70 years since Bizley's results were published, the determination of $|N_k(g)|$ has remained an open problem for every intermediate value of k.

However, in 2019, Firoozi, Marwendo, Rattan [14] evaluated $|N_k(g)|$ for the case $g = 1$ and for the case $a = b = 1$.

Theorem 2.5.1 (Evaluation of $|N_k(1)|$ [14])**.** *We have*

$$|N_k(1)| = \frac{1}{a+b}\binom{a+b}{a} \quad \textit{for all } k \textit{ satisfying } 0 \le k < a+b.$$

Theorem 2.5.2 (Evaluation of $|N_k(g)|$ for $a = b = 1$ [14])**.** *Let $a = b = 1$. Then*

$$|N_k(g)| = \sum_{i=0}^{g-\lceil \frac{k+1}{2} \rceil} C_i C_{g-1-i} \quad \textit{for all } k \textit{ satisfying } 0 \le k < 2g,$$

where $C_i = \frac{1}{i+1}\binom{2i}{i}$ is the i^{th} Catalan number.

Firoozi, Marwendo, Rattan [14] conjectured the result we have presented as Corollary 1.3.6 (Constant on blocks), based on numerical experiments together with the results of Theorems 2.5.1 and 2.5.2. Their conjecture was a major inspiration for the formulation of our main result Theorem 1.3.5. The truth of the conjecture is a direct corollary of this result.

The exact values stated in Theorems 2.5.1 and 2.5.2 imply the special cases $g = 1$ and $a = b = 1$ of Corollary 1.3.6, respectively. Theorem 2.5.1 is equivalent to the statement that (1.3.3) holds for $g = 1$, where $\mu_0(1)$ takes the value $\frac{1}{a+b}\binom{a+b}{a}$. This expression for $\mu_0(1)$ is the same as that given by the path enumeration formula, Theorem 1.3.13:

$$\mu_0(1) = \mathbf{E}_0\mathbf{H}_1 = C_{\langle 1 \rangle} = c_1 = \frac{1}{a+b}\binom{a+b}{a}.$$

Similarly, Theorem 2.5.2 is equivalent to the statement that (1.3.3) holds for $a = b = 1$, where $\mu_j(g)$ takes the value

$$\sum_{i=1}^{g-j} C_{i-1} C_{g-i}.$$

However, this expression for $\mu_j(g)$ when $a = b = 1$ does not take the same form as that given by the path enumeration formula (even though both expressions are equal because they count the same set).

Our central objective in this work is to find an explicit formula for $|N_k(g)|$ for given a, b, for all allowable values of k.

Chapter 3

Proof of main result

For convenience, we restate our main result here.

Theorem 1.3.5. *Let g, k satisfy $0 \leq k < g(a+b) - 1$. Then*

$$|N_k(g) \setminus S_k(g)| = |N_{k+1}(g)|.$$

We shall prove our main result in this chapter by constructing a bijection from $N_k(g) \setminus S_k(g)$ to $N_{k+1}(g)$. We present some intuition for the proof method in Section 3.1 and give an outline of the proof in Section 3.2. We then prove the result in detail in the rest of this chapter.

3.1 Building intuition

We identify various important attributes of paths and describe some basic path operations. We then use these concepts to give a concise proof of Theorem 2.5.1 (which deals with the case $g = 1$).

We begin with some terminology.

Definition 3.1.1 (Elevation)**.** Let (i, j) be a point of a path in $N(g)$. The *elevation* of (i, j) is $ja - ib$.

The elevation of a particular point of a path in $N(g)$ is a measure of the directed perpendicular distance from that point to the path boundary. Note that points on the boundary have zero elevation; points above the boundary have positive elevation; and points below the boundary have negative elevation.

The points of a path closest to the boundary that do not lie on the boundary have special significance.

Definition 3.1.2 (Highest points below, lowest points above)**.** Let p be a path in $N(g)$. The *highest points below the boundary (HPBs)* are those points (if any) of p lying strictly below the boundary which attain the closest elevation to zero. We label the HPBs in order as $H_1, \dots, H_\eta$. The *lowest points above the boundary (LPAs)* are defined analogously and labelled $L_1, \dots, L_\ell$.

See Figure 3.1 for an example of the HPBs and LPAs of a path.

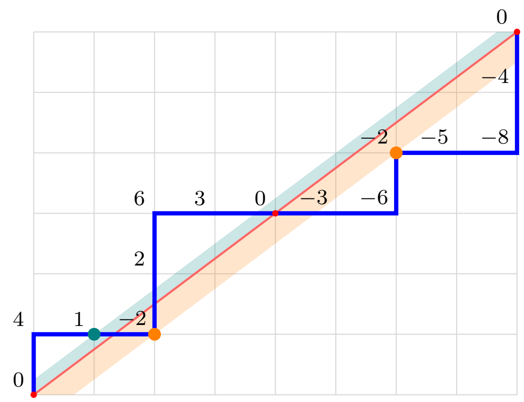


Figure 3.1: A path p in $N_5(2)$ for $(a,b) = (4,3)$, where the elevation of each point of p is marked. The HPBs of p are the points $(2,1), (6,4)$; the (unique) LPA of p is the point $(1,1)$. The set of HPBs and the set of LPAs each impose a (respectively shaded) region which contains no path points in its interior.

Definition 3.1.3 (Path split, subpath)**.** Let $p \in N(g)$ contain the points $(0,0) = R_0$, $R_1, \dots, R_{n-1}$, $R_n = (ga, gb)$ in that order (and possibly contain other points). A *split* of p at the points $R_1, \dots, R_{n-1}$ is a decomposition of p into the n consecutive paths

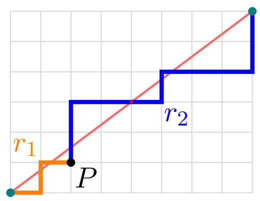
$$p_1 \coloneqq p[R_0, R_1], \quad p_2 \coloneqq p[R_1, R_2], \quad \dots, \quad p_n \coloneqq p[R_{n-1}, R_n],$$

where $p_i = p[R_{i-1}, R_i]$ represents the *subpath* of p between R_{i-1} and R_i. We may then write p as the concatenation $p = p_1 p_2 \cdots p_n$.

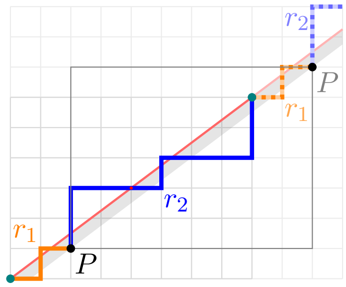
Note that the flaws of a path $p \in N(g)$ are defined in relation to the boundary joining the startpoint $(0,0)$ to the endpoint (ga, gb). However, we consider the flaws of a proper subpath r of p in relation to the boundary of r, not of p.

Remark 3.1.4. If a path p is split at a boundary point of p (other than the startpoint or endpoint) into p_1p_2, each of p_1 and p_2 will have the same 'slope' (that is, the same values of a, b) as p. However, if p is split into r_1r_2 at a point not lying on the boundary of p, then each of r_1 and r_2 will have a slope different from p. In general, a subpath r of p has the same slope as p if and only if r starts and ends at points of p having the same elevation. If r is a proper subpath of $p \in N(g)$ and has the same slope as p, then $r \in N(h)$ for some h satisfying $h < g$ (for the same values of a, b).

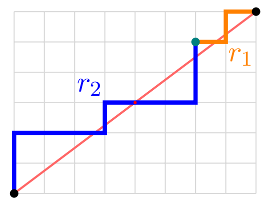
We next consider the operation of cyclically permuting the steps of a path r_1r_2 with respect to the last point P of the r_1 subpath to produce the path r_2r_1: see Figure 3.2 for an example.



(a) The path r_1r_2.



(b) This cyclic permutation sends P to the origin.



(c) The path r_2r_1.

Figure 3.2: Cyclically permuting the steps of r_1r_2 with respect to P.

Lemma 3.1.5. *Let p be a path in $N(g)$ containing a point P, and let f be the mapping that cyclically permutes the steps of p with respect to P.*

(i) *Let the elevation of P be e. Then, under the mapping f, the elevation of each point of p reduces by exactly e.*

(ii) *Let p have $n + 1$ boundary points, and suppose that P is an HPB. Then, under the mapping f, the number of flaws increases by exactly n.*

Proof.

(i) This follows from the definition of elevation.

(ii) This follows from the definition of HPB, noting that the first and last point of p merge to form a single point in the cyclically permuted path $f(p)$. □

See Figure 3.3 for an illustration of the cyclic permutation f used in the proof of Lemma 3.1.5 (using the value $n = 1$ for part *(ii)*).

We now prove Theorem 2.5.1 (which deals with the case $g = 1$) using the concepts of elevation, HPBs, LPAs, and cyclic permutations.

Proof of Theorem 2.5.1. Let k satisfy $0 \leq k < a+b-1$. Since $|N(1)| = \binom{a+b}{a}$, it is sufficient to exhibit a bijection f from $N_k(1)$ to $N_{k+1}(1)$.

Let $p \in N_k(1)$, so that p has startpoint $(0, 0)$ and endpoint (a, b). Since a, b are coprime, by the definition of elevation each of the points of p (apart from the startpoint and endpoint) has a unique elevation.

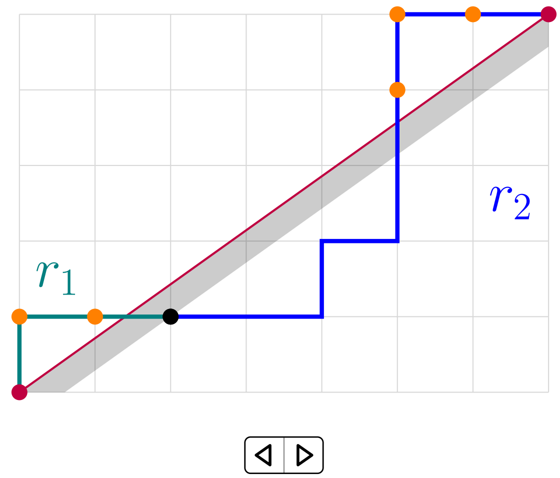


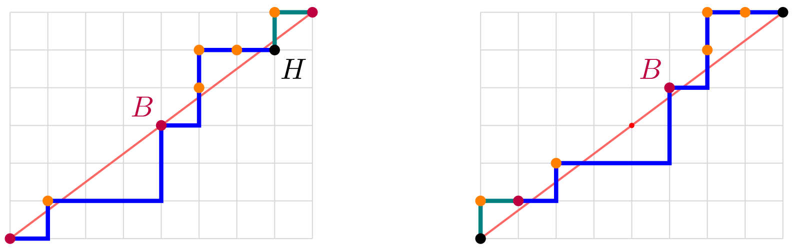
Figure 3.3: Cyclic permutation of a path with respect to its HPB. This animation can be viewed through a JavaScript-enabled PDF reader (such as Adobe Acrobat).

The path p has non-max flaws by assumption, and is not BPT because a, b are coprime. Therefore p has at least one point strictly below the boundary, and by the uniqueness of elevations p has a unique HPB H. Take f to be the mapping that cyclically permutes p with respect to H. By Lemma 3.1.5*(ii)* with $n = 1$, the number of flaws in the resulting path $f(p)$ is $k + 1$ (see Figure 3.3), so f maps $N_k(1)$ to $N_{k+1}(1)$. The map f is invertible: cyclically permute the image $f(p)$ with respect to its unique LPA to recover the original path p. □

The proof of our main result is considerably more involved than might be suggested by the simplicity of the bijection used in the preceding proof. Firstly, we wish to find a bijection from $N_k(g) \setminus S_k(g)$ to $N_{k+1}(g)$. In the case $g = 1$, we did not have to consider the set $S_k(g)$ because it is empty for all k satisfying $0 \leq k < a + b - 1$. Secondly, in general we may not assume that the path $p \in N_k(g) \setminus S_k(g)$ has a unique HPB. Thirdly, cyclic permutation of a path $p \in N_k(g) \setminus S_k(g)$ need not necessarily map to $N_{k+1}(g)$, as we now demonstrate.

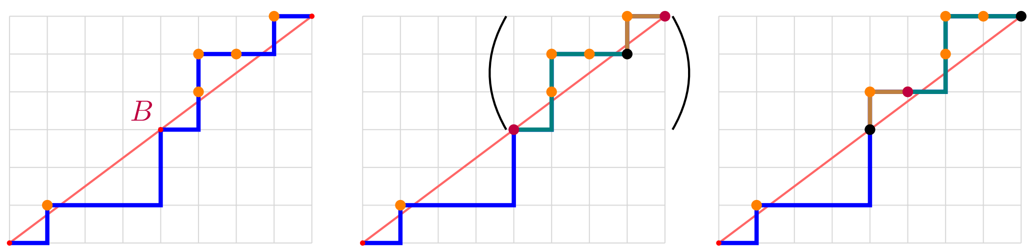
Consider the path $p \in N_5(2)$ shown in Figure 3.4a. Cyclically permute p with respect to its unique HPB H. The resulting path (seen in Figure 3.4b) has 7 flaws, not 6. The reason that the number of flaws of p increases by two is that p touches the boundary at the point B, so it is no longer true that each of the points of p (apart from the startpoint and endpoint) has a unique elevation. Under cyclic permutation, the startpoint and endpoint of p merge to form a single point as before, but additionally the elevation of the boundary point B increases from 0 to create an additional flaw.

We can deal with the example path shown in Figure 3.4b using a modification of the cyclic permutation technique: see Figure 3.5. Split p at the point B, and then apply cyclic



(a) The path p. (b) Cyclic permutation of p.

Figure 3.4: Cyclically permuting the steps of p with respect to H results in two additional flaws.



(a) Split the path p at B. (b) Apply cyclic permutation to the bracketed subpath. (c) Resulting path.

Figure 3.5: Incrementing the flaws of p.

permutation as before but only to the bracketed subpath. This gives a mapping from $N_5(2)$ to $N_6(2)$.

This example illustrates several of the key ideas we shall use for our general mapping: distinguishing paths that are BPT from those that are not; rearranging subpaths of a path; and defining the mapping recursively.

3.2 Proof outline

We shall prove Theorem 1.3.5 by constructing an explicit bijection

$$\phi_{g,k} : N_k(g) \setminus S_k(g) \to N_{k+1}(g)$$

for each k in the range $0 \leq k < g(a+b) - 1$.

We specify three distinct actions for the map $\phi_{g,k}$ by constructing three different bijections. Then, depending on the characteristics of each path p in its domain, $\phi_{g,k}$ will apply exactly one of these actions to p.

We first partition the set $N_k(g) \setminus S_k(g)$ into subsets $X_k(g)$, $Y_k(g)$, $Z_k(g)$ and partition (using a different method) the set $N_{k+1}(g)$ into subsets $\mathcal{X}_{k+1}(g)$, $\mathcal{Y}_{k+1}(g)$, $\mathcal{Z}_{k+1}(g)$. Note that we abuse the standard definition of a set partition by allowing empty sets to appear as partitioning subsets. We then construct an explicit bijection $\phi_{g,k}^X : X_k(g) \to \mathcal{X}_{k+1}(g)$, and similarly $\phi_{g,k}^Y$ and $\phi_{g,k}^Z$, as illustrated in Figure 3.6. These three bijections collectively define the composite map $\phi_{g,k}$.

Since the number k of flaws of a path $p \in N(g)$ is determined, we may define the function

$$\phi_g : N(g) \setminus S(g) \to N(g) \setminus N_0(g), \tag{3.2.1}$$

where

$$\phi_g(p) \coloneqq \phi_{g,k}(p) \quad \text{if } p \in N_k(g) \setminus S_k(g).$$

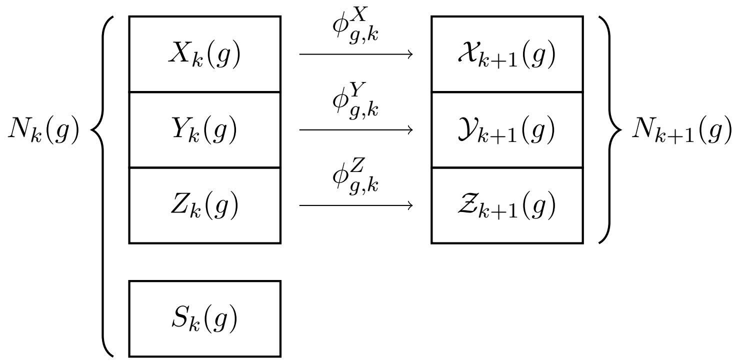


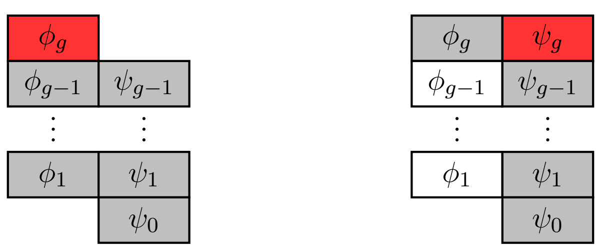
Figure 3.6: The bijection $\phi_{g,k} : N_k(g) \setminus S_k(g) \to N_{k+1}(g)$ is induced by the bijections $\phi_{g,k}^X$, $\phi_{g,k}^Y$, $\phi_{g,k}^Z$ between each partitioning subset of $N_k(g) \setminus S_k(g)$ and the corresponding partitioning subset of $N_{k+1}(g)$.

In order to define the action of ϕ_g, we require an auxiliary map

$$\psi_g : Q(g) \to \mathcal{Q}(g), \tag{3.2.2}$$

where $Q(g)$ and $\mathcal{Q}(g)$ are subsets of $N(g)$.

Both ϕ_g and ψ_g are defined recursively in intertwined fashion, as shown in Figure 3.7.



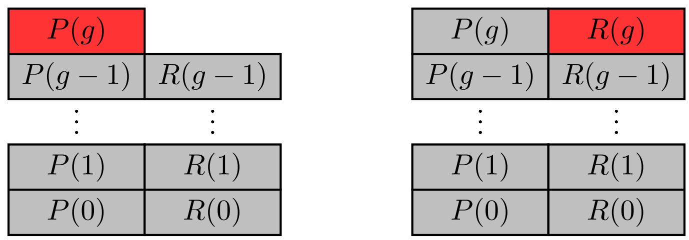
(a) The map ϕ_g is defined using the shaded maps.

(b) The map ψ_g is defined using the shaded maps.

Figure 3.7: Recursive definition of ϕ_g and ψ_g.

Our strategy for proving Theorem 1.3.5 is to show that ϕ_g is a bijection for all g. Represent by $P(g)$ the statement that ϕ_g is a bijection and that some further conditions on ϕ_g hold, and represent by $R(g)$ the statement that ψ_g is a bijection and that some further conditions on ψ_g hold.

We prove by induction on g that $P(g)$ and $R(g)$ hold for all $g \geq 0$, as illustrated in Figure 3.8. We then obtain Theorem 1.3.5 as an immediate consequence since $P(g)$ implies that $\phi_{g,k}$ is a bijection for each k.



(a) Proving $P(g)$ requires all of the shaded statements.

(b) Proving $R(g)$ requires all of the shaded statements.

Figure 3.8: Coupled induction on g used to show that $P(g)$ and $R(g)$ hold.

We construct the maps ϕ_g and ψ_g in Sections 3.3 to 3.5, and then prove $P(g)$ and $R(g)$ by induction in Section 3.6.

3.3 Domain and codomain of ϕ_g and ψ_g

In this section, we use vocabulary from Section 2.5 to define the domain and codomain of the maps $\phi^X_{g,k}$, $\phi^Y_{g,k}$, $\phi^Z_{g,k}$ (collectively giving ϕ_g) and ψ_g.

We define the subsets $Q(g)$ and $\mathcal{Q}(g)$ of $N(g)$ that will form the domain and codomain of ψ_g, respectively. Recall that ε is the empty path.

Definition 3.3.1 (Subsets $Q(g)$, $\mathcal{Q}(g)$)**.** Let $Q(0) \coloneqq \{\varepsilon\}$ and $\mathcal{Q}(0) \coloneqq \{\varepsilon\}$. For $g > 0$, let

$$Q(g) \coloneqq \{p \in N(g) : p[B_{\beta-1}, B_\beta] \text{ has at least one flaw}\},$$
$$\mathcal{Q}(g) \coloneqq \{p \in N(g) : p[B_{\beta-1}, B_\beta] \text{ has non-max flaws}\},$$

where the boundary points of $p \in N(g)$ are denoted by $B_0, B_1, \ldots, B_\beta$.

We partition $N_k(g) \setminus S_k(g)$ into the subsets $X_k(g)$, $Y_k(g)$, $Z_k(g)$ that will form the domain of $\phi^X_{g,k}$, $\phi^Y_{g,k}$, $\phi^Z_{g,k}$, respectively.

Definition 3.3.2 (Subsets $X_k(g)$, $Y_k(g)$, $Z_k(g)$)**.** Let g, k satisfy $0 \leq k < g(a+b) - 1$. Partition $N_k(g) \setminus S_k(g)$ into subsets $X_k(g)$, $Y_k(g)$, $Z_k(g)$ comprising those paths p, whose boundary points are $B_0, B_1, \dots, B_\beta$, satisfying the specified conditions:

p belongs to	conditions
$X_k(g)$	$\beta = 1$ (not BPT)
$Y_k(g)$	$\beta > 1$ and $p[B_1, B_\beta] \notin S$
$Z_k(g)$	$\beta > 1$ and $p[B_1, B_\beta] \in S$

We now partition $N_{k+1}(g)$ (using a different method) into the subsets $\mathcal{X}_{k+1}(g)$, $\mathcal{Y}_{k+1}(g)$, $\mathcal{Z}_{k+1}(g)$ that will form the codomain of $\phi^X_{g,k}$, $\phi^Y_{g,k}$, $\phi^Z_{g,k}$, respectively.

Definition 3.3.3 (Subsets $\mathcal{X}_{k+1}(g), \mathcal{Y}_{k+1}(g), \mathcal{Z}_{k+1}(g)$)**.** Let g, k satisfy $0 \le k < g(a+b) - 1$. Partition $N_{k+1}(g)$ into subsets $\mathcal{X}_{k+1}(g)$, $\mathcal{Y}_{k+1}(g)$, $\mathcal{Z}_{k+1}(g)$ comprising those paths p, whose boundary points are $B_0, B_1, \dots, B_\beta$ and whose LPAs are $L_1, \dots, L_\ell$, satisfying the specified conditions:

p belongs to	conditions
$\mathcal{X}_{k+1}(g)$	$p[B_1, B_\beta]$ (possibly empty) has no flaws. $p[L_1, L_\ell] \in \mathcal{Q}(h)$ for some h satisfying $0 \le h < g$
$\mathcal{Y}_{k+1}(g)$	$p[B_1, B_\beta]$ (non-empty) has at least one flaw
$\mathcal{Z}_{k+1}(g)$	$p[B_1, B_\beta]$ (possibly empty) has no flaws. $p[L_1, L_\ell] \notin \mathcal{Q}(h)$ for each h satisfying $0 \le h < g$

3.4 The maps $\phi^X_{g,k}$, $\phi^Y_{g,k}$, $\phi^Z_{g,k}$

In this section, we define the maps

$$\begin{aligned} \phi^X_{g,k} &: X_k(g) \to \mathcal{X}_{k+1}(g), \\ \phi^Y_{g,k} &: Y_k(g) \to \mathcal{Y}_{k+1}(g), \\ \phi^Z_{g,k} &: Z_k(g) \to \mathcal{Z}_{k+1}(g), \end{aligned} \tag{3.4.1}$$

for g, k satisfying $0 \le k < g(a+b) - 1$. These three maps then collectively define $\phi_{g,k} : N_k(g) \setminus S_k(g) \to N_{k+1}(g)$ piecewise as follows:

$$\phi_{g,k} = \begin{cases} \phi^X_{g,k} & \text{on } X_k(g), \\ \phi^Y_{g,k} & \text{on } Y_k(g), \\ \phi^Z_{g,k} & \text{on } Z_k(g). \end{cases}$$

The action that each of $\phi^X_{g,k}$, $\phi^Y_{g,k}$, $\phi^Z_{g,k}$ performs on a path p in its domain involves rearranging and manipulating subpaths of p.

We must therefore specify a path split representation for a typical path in each of the sets $X_k(g)$, $Y_k(g)$, $Z_k(g)$.

Lemma 3.4.1. *Let g, k satisfy $0 \leq k < g(a+b) - 1$, and let $p \in N_k(g) \setminus S_k(g)$ (so that p has non-max flaws). Let $B_0, B_1, \ldots, B_\beta$ be the boundary points of p.*

Case 1: $p \in X_k(g)$**.** *Then $\beta = 1$. Let the HPBs of p be $H_1, \ldots, H_\eta$ (where $\eta \geq 1$), and let*

$$
\begin{aligned}
r_1 &= p[B_0, H_1],\\
r_2 &= p[H_\eta, B_1].
\end{aligned}
$$

Then we may write $p = r_1 s t r_2$, where

1. *st is the unique split (at some HPB H_γ) of $p[H_1, H_\eta]$ such that $s \in Q(h)$ for some h satisfying $0 \leq h < g$ and t has no flaws,*
2. *$r_1 r_2$, $r_1 t r_2$, $r_1 s r_2$ are each not BPT.*

Case 2: $p \in Y_k(g)$**.** *Let*

$$
\begin{aligned}
p_1 &= p[B_0, B_1],\\
p_2 &= p[B_1, B_\beta].
\end{aligned}
$$

Then we may write $p = p_1 p_2$, where $p_2 \in N(h) \setminus S(h)$ for some h satisfying $0 < h < g$.

Case 3: $p \in Z_k(g)$**.** *The subpath $p[B_0, B_1]$ has at least one flaw. Let L be the last LPA of $p[B_0, B_1]$, and let*

$$
\begin{aligned}
r_1 &= p[B_0, L],\\
r_2 &= p[L, B_1],\\
t &= p[B_1, B_{\beta-1}] \text{ (possibly empty)},\\
s &= p[B_{\beta-1}, B_\beta] \text{ (non-empty)}.
\end{aligned}
$$

Then we may write $p = r_1 r_2 t s$, where

1. *t has no flaws and s has max flaws,*
2. *$r_1 s r_2$ is not BPT.*

Proof.

Case 1: $p \in X_k(g)$**.**

1. This follows from Definitions 3.3.1 and 3.3.2, noting that if s is empty then $h = 0$.

2. Since the path $p = r_1 s t r_2$ lies in $X_k(g)$, it is not BPT. The subpath st begins and ends at the same elevation because $st = p[H_1, H_\eta]$, and the subpath s begins and ends at the same elevation because $s \in Q(h)$. Therefore the paths $r_1 r_2$, $r_1 t r_2$, $r_1 s r_2$ are each not BPT.

Case 2: $p \in Y_k(g)$**.** This follows from Definition 3.3.2.

Case 3: $p \in Z_k(g)$**.** The subpath $p[B_0, B_1]$ has at least one flaw, otherwise we would have that $p = p[B_0, B_1]\, p[B_1, B_\beta] \in S$ by Definitions 1.3.4 and 3.3.2, contradicting that $p \notin S_k(g)$.

1. This follows from Definitions 1.3.4 and 3.3.2.

2. The subpath $r_1 r_2 = p[B_0, B_1]$ is not BPT, and the subpath $s = p[B_{\beta-1}, B_\beta]$ begins and ends at the same elevation and has max flaws. Therefore the path $r_1 s r_2$ is not BPT. □

We next specify a path split representation for a typical path in each of the sets $\mathcal{X}_{k+1}(g)$, $\mathcal{Y}_{k+1}(g)$, $\mathcal{Z}_{k+1}(g)$.

Lemma 3.4.2. *Let g, k satisfy $0 \le k < g(a+b) - 1$, and let $\mathbb{p} \in N_{k+1}(g)$ (so that $\mathbb{p}$ has at least one flaw). Let $B_0, B_1, \dots, B_\beta$ be the boundary points of $\mathbb{p}$, and let $L_1, \dots, L_\ell$ be the LPAs of $\mathbb{p}$ (where $\ell \ge 1$).*

Case 1: $\mathbb{p} \in \mathcal{X}_{k+1}(g)$**.** *Let*

$$\begin{aligned} r_2 &= \mathbb{p}[B_0, L_1], \\ \mathbb{s} &= \mathbb{p}[L_1, L_\ell] \text{ (possibly empty)}, \\ r_1 &= \mathbb{p}[L_\ell, B_1], \\ t &= \mathbb{p}[B_1, B_\beta] \text{ (possibly empty)}. \end{aligned}$$

Then we may write $\mathbb{p} = r_2 \mathbb{s} r_1 t$, where

1. *$\mathbb{s} \in \mathcal{Q}(h)$ for some h satisfying $0 \le h < g$ and t has no flaws,*

2. *$r_2 r_1$ is not BPT.*

Case 2: $\mathbb{p} \in \mathcal{Y}_{k+1}(g)$**.** *Let*

$$\begin{aligned} \mathbb{p}_1 &= \mathbb{p}[B_0, B_1], \\ \mathbb{p}_2 &= \mathbb{p}[B_1, B_\beta] \text{ (non-empty)}. \end{aligned}$$

Then we may write $\mathbb{p} = \mathbb{p}_1 \mathbb{p}_2$, where $\mathbb{p}_2 \in N_{k'+1}(h)$ for some h, k' satisfying $0 < h < g$ and $k' \ge 0$.

Case 3: $\mathbb{p} \in \mathcal{Z}_{k+1}(g)$**.** *Then* $\ell > 1$*. Let*

$$\begin{aligned} r_1 &= \mathbb{p}[B_0, L_{\ell-1}], \\ s &= \mathbb{p}[L_{\ell-1}, L_\ell] \textit{ (non-empty)}, \\ r_2 &= \mathbb{p}[L_\ell, B_1], \\ t &= \mathbb{p}[B_1, B_\beta] \textit{ (possibly empty)}. \end{aligned}$$

Then we may write $\mathbb{p} = r_1 s r_2 t$*, where* s *has max flaws and* t *has no flaws.*

Proof.

Case 1: $\mathbb{p} \in \mathcal{X}_{k+1}(g)$**.**

1. This follows from Definition 3.3.3.

2. Since $r_2 \mathbb{s} r_1$ is not BPT, and L_1 and L_ℓ have the same elevation, the path $r_2 r_1$ is not BPT.

Case 2: $\mathbb{p} \in \mathcal{Y}_{k+1}(g)$**.** This follows from Definition 3.3.3.

Case 3: $\mathbb{p} \in \mathcal{Z}_{k+1}(g)$**.** By Definition 3.3.3, we have $\mathbb{p}[L_1, L_\ell] \notin \mathcal{Q}(0)$ and so $\ell > 1$. The statement then follows from Definitions 3.3.1 and 3.3.3. ☐

We now define the composite map $\phi_{g,k} : N_k(g) \setminus S_k(g) \to N_{k+1}(g)$ by specifying each of the maps $\phi^X_{g,k}$, $\phi^Y_{g,k}$, $\phi^Z_{g,k}$ according to their action on a path represented according to Lemma 3.4.1.

Definition 3.4.3 ($\phi^X_{g,k}$, $\phi^Y_{g,k}$, $\phi^Z_{g,k}$)**.** Let g, k satisfy $0 \leq k < g(a+b) - 1$ and let $p \in N_k(g) \setminus S_k(g)$.

Case 1: $p \in X_k(g)$**.** Write $p = r_1 s t r_2$ according to Case 1 of Lemma 3.4.1, where $s \in Q(h)$ for some h satisfying $0 \leq h < g$. Then the map $\phi^X_{g,k} : X_k(g) \to \mathcal{X}_{k+1}(g)$ is given by

$$\phi^X_{g,k}(p) = r_2 \psi_h(s) r_1 t.$$

Case 2: $p \in Y_k(g)$**.** Write $p = p_1 p_2$ according to Case 2 of Lemma 3.4.1, where $p_2 \in N(h) \setminus S(h)$ for some h satisfying $0 < h < g$. Then the map $\phi^Y_{g,k} : Y_k(g) \to \mathcal{Y}_{k+1}(g)$ is given by

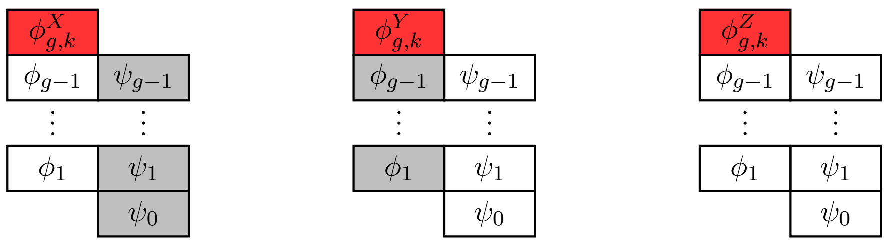
$$\phi^Y_{g,k}(p) = p_1 \phi_h(p_2).$$

Case 3: $p \in Z_k(g)$**.** Write $p = r_1 r_2 ts$ according to Case 3 of Lemma 3.4.1. Then the map $\phi^Z_{g,k} : Z_k(g) \to \mathcal{Z}_{k+1}(g)$ is given by

$$\phi^Z_{g,k}(p) = r_1 s r_2 t.$$

We shall show in Section 3.6 that each of the maps $\phi^X_{g,k}$, $\phi^Y_{g,k}$, $\phi^Z_{g,k}$ has the specified codomain.

We refine Figure 3.7a in Figure 3.9 by illustrating how the maps $\phi^Y_{g,k}$ and $\phi^X_{g,k}$ are recursively defined in terms of the maps $\{\phi_h : 0 < h < g\}$ and the auxiliary maps $\{\psi_h : 0 \leq h < g\}$ (to be defined in Section 3.5), respectively.



(a) Defining $\phi^X_{g,k}$ requires the shaded maps.

(b) Defining $\phi^Y_{g,k}$ requires the shaded maps.

(c) Map $\phi^Z_{g,k}$ is directly defined.

Figure 3.9: Recursive definition of $\phi^X_{g,k}$ and $\phi^Y_{g,k}$. The map $\phi^Z_{g,k}$ is defined directly without the use of recursion.

3.5 The map ψ_g

In this section, we define the auxiliary map $\psi_g : Q(g) \to \mathcal{Q}(g)$. We begin by introducing the reversal of a path.

Definition 3.5.1. Let p be a path whose ordered steps are $s_1, s_2, \ldots, s_n$. The *reversal* of p is the path $\overline{p}$ whose ordered steps are $s_n, \ldots, s_2, s_1$.

Geometrically, $\overline{p}$ is obtained by rotating p by half a revolution. This leads to the following observation.

Remark 3.5.2. For $g \geq 0$, the reversal mapping

$$^{-} : N(g) \to N(g)$$

is a bijection that for all $p = p_1 p_2 \in N(g)$ satisfies $\overline{p_1 p_2} = \overline{p_2}\ \overline{p_1}$.

Under the reversal mapping $^{-}$, a flaw maps to a non-flaw; a boundary point maps to a boundary point; a point below the boundary maps to a flaw. This gives the following counting result.

Remark 3.5.3. Let $p \in N(g)$ have k flaws and $\beta + 1$ boundary points. Then $\overline{p}$ has $g(a+b) - \beta - k$ flaws.

We now specify a path split representation for a typical path in $Q(g)$.

Lemma 3.5.4. *Let $g > 0$ and $q \in Q(g)$ have boundary points $B_0, B_1, \cdots, B_\beta$. Let*

$$\begin{aligned} q_1 &= q[B_0, B_1], \\ q_2 &= q[B_1, B_\beta] \text{ (possibly empty).} \end{aligned}$$

Then we may write $q = q_1 q_2$, where

1. *$q_2 \in Q(h)$ for some h satisfying $0 \leq h < g$,*
2. *if $q_2 \in Q(0)$, then q_1 has at least one flaw.*

Proof. Since $g > 0$, we have that $q = q_1 q_2$ is non-empty. The results follow from Definition 3.3.1 by considering the cases that q_2 is non-empty or empty. □

We now define the map $\psi_g : Q(g) \to \mathcal{Q}(g)$. In view of Definition 3.3.1, we take ψ_0 to be the trivial bijection which maps the empty path to the empty path. We define ψ_g for $g > 0$ according to its action on a path represented according to Lemma 3.5.4.

Definition 3.5.5 (ψ_g)**.** Let $g > 0$. Write $q = q_1 q_2$ according to Lemma 3.5.4, where $q_2 \in Q(h)$ for some h satisfying $0 \leq h < g$. Then the map $\psi_g : Q(g) \to \mathcal{Q}(g)$ in given by

$$\psi_g(q) = \overline{\phi_g\left(\overline{q_1 \psi_h(q_2)}\right)}. \tag{3.5.1}$$

The map ψ_g is defined recursively using the map ϕ_g and the maps $\{\psi_h : 0 \leq h < g\}$, as shown in Figure 3.7b. We shall show in Section 3.6 that the expression (3.5.1) is well-defined and that the map ψ_g has the specified codomain.

We remark that identifying an appropriate map ψ_g was a major milestone in the development of this thesis.

3.6 The statements $P(g)$ and $R(g)$

In this section, we specify a statement $P(g)$ asserting (among other properties) that $\phi^X_{g,k}$, $\phi^Y_{g,k}$, $\phi^Z_{g,k}$ are bijections for each k, and a statement $R(g)$ asserting (among other properties) that ψ_g is a bijection. We then prove the statements $P(g)$ and $R(g)$ concurrently, by induction on $g \geq 0$ (see Figure 3.8). As noted in Section 3.2, establishing $P(g)$ for each $g > 0$ proves our main result (Theorem 1.3.5).

We still must show that each of the maps $\phi^X_{g,k}$, $\phi^Y_{g,k}$, $\phi^Z_{g,k}$ and ψ_g has the appropriate codomain specified in (3.4.1) and (3.2.2). This will be included in the proofs that each of these maps is a bijection. We now define the statements $P(g)$ and $R(g)$.

Definition 3.6.1 (Statements $P(0)$, $R(0)$)**.** The statement $P(0)$ is defined to be true. The statement $R(0)$ is that $\psi_0 : Q(0) \to \mathcal{Q}(0)$ is a bijection.

Definition 3.6.2 (Statement $P(g)$)**.** Let $g > 0$. The statement $P(g)$ is that the following properties hold for all k satisfying $0 \le k < g(a+b) - 1$.

$P^X_{\text{bij}}(g)$: $\phi^X_{g,k}$ is a bijection.

$P^Y_{\text{bij}}(g)$: $\phi^Y_{g,k}$ is a bijection.

$P^Z_{\text{bij}}(g)$: $\phi^Z_{g,k}$ is a bijection.

$P^X_{\text{elev}}(g)$: Let $p \in X_k(g)$ and let the HPBs of p have elevation $-e$. Then the LPAs of $\phi^X_{g,k}(p)$ have elevation e.

$P_{\text{flaw}}(g)$: Let $p \in N(g) \setminus S(g)$ and $\mathbb{p} = \phi_g(p)$, and write $p = p_1 \cdots p_n$ and $\mathbb{p} = \mathbb{p}_1 \cdots \mathbb{p}_m$ where each path is split at its respective boundary points. Then

(i) if p_1 has at least one flaw, then so does $\mathbb{p}_1$,

(ii) suppose $n > 1$. If $\mathbb{p}_1$ has at least one flaw, then so does p_1.

Definition 3.6.3 (Statement $R(g)$)**.** Let $g > 0$. The statement $R(g)$ is that the following properties hold.

$R_{\text{bij}}(g)$: ψ_g is a bijection.

$R_{\text{flaw}}(g)$: Let $q \in Q(g)$ and $\mathbb{q} = \psi_g(q)$, and write $q = q_1 \cdots q_n$ and $\mathbb{q} = \mathbb{q}_1 \cdots \mathbb{q}_m$ where each path is split at its respective boundary points. Then

(i) $\overline{\mathbb{q}}$ has exactly n more flaws than $\overline{q}$,

(ii) $\mathbb{q}$ has exactly m fewer flaws than q.

$R_{\text{elev}}(g)$: Let $q \in Q(g)$ and let the LPAs of q have elevation e. Then the HPBs of $\psi_g(q)$ have elevation $-e$.

We shall prove the statements $P(g)$ and $R(g)$ by induction on $g \ge 0$, according to the roadmap given below.

Remark 3.6.4. The statements $P(g)$, $R(g)$ are defined differently for $g = 0$ and $g > 0$ (see Definitions 3.6.1 to 3.6.3). Whenever we use the inductive hypothesis that $P(h)$ and $R(h)$

hold for all h satisfying $0 \leq h < g$, we shall be careful to consider the cases $h = 0$ and $h > 0$ separately as necessary.

Proof Roadmap:

I. (Base case). The statement $P(0)$ holds vacuously. The statement $R(0)$ holds trivially by Definition 3.3.1.

II. (Inductive hypothesis). Let $g > 0$. Assume that

$$\text{statements } P(h) \text{ and } R(h) \text{ hold for all } h \text{ satisfying } 0 \leq h < g. \tag{3.6.1}$$

III. (Inductive step for $P(g)$). Lemma 3.6.5. Subject to the inductive hypothesis (3.6.1), statement $P(g)$ holds. This is proven using the following claims.

   (a) Claim 3.6.6. $P^X_{\text{bij}}(g)$ and $P^X_{\text{elev}}(g)$ hold.

   (b) Claim 3.6.7. $P^Y_{\text{bij}}(g)$ holds.

   (c) Claim 3.6.8. $P^Z_{\text{bij}}(g)$ holds.

   (d) Claim 3.6.9. $P_{\text{flaw}}(g)$ holds.

IV. (Inductive step for $R(g)$). Lemma 3.6.10: Subject to the inductive hypothesis (3.6.1), statement $R(g)$ holds. This is proven using Lemma 3.6.5 in addition to the following claims.

   (a) Claim 3.6.11. $R_{\text{bij}}(g)$ holds.

   (b) Claim 3.6.12. $R_{\text{flaw}}(g)$ holds.

   (c) Claim 3.6.13. $R_{\text{elev}}(g)$ holds.

V. (Conclusion). This completes the induction (which implies the main result, Theorem 1.3.5).

Following step II of the Proof Roadmap, let $g > 0$ and assume that the inductive hypothesis (3.6.1) holds. We write

$$X(g) = \bigcupdot_k X_k(g), \qquad Y(g) = \bigcupdot_k Y_k(g), \qquad Z(g) = \bigcupdot_k Z_k(g),$$
$$\mathcal{X}(g) = \bigcupdot_k \mathcal{X}_{k+1}(g), \qquad \mathcal{Y}(g) = \bigcupdot_k \mathcal{Y}_{k+1}(g), \qquad \mathcal{Z}(g) = \bigcupdot_k \mathcal{Z}_{k+1}(g).$$

Lemma 3.6.5. *Subject to the inductive hypothesis* (3.6.1)*, statement $P(g)$ holds.*

We split the proof of Lemma 3.6.5 into Claims 3.6.6 to 3.6.9.

Claim 3.6.6. *The statements $P^X_{\text{bij}}(g)$ and $P^X_{\text{elev}}(g)$ hold.*

Proof. Let k satisfy $0 \le k < g(a+b) - 1$. Let $p \in X_k(g)$ and write $p = r_1 s t r_2$ according to Case 1 of Lemma 3.4.1, where $s \in Q(h)$ for some h satisfying $0 \le h < g$ and t has no flaws. Write $\mathbb{s} = \psi_h(s)$ and $\mathbb{p} = \phi^X_{g,k}(p)$. By Definition 3.4.3,

$$\mathbb{p} = r_2 \mathbb{s} r_1 t. \tag{3.6.2}$$

Let the HPBs of p have elevation $-e$. We shall show in the following sequence of steps that $P^X_{\text{bij}}(g)$ and $P^X_{\text{elev}}(g)$ hold:

(i) $P^X_{\text{elev}}(g)$ holds,

(ii) $\phi^X_{g,k}$ has codomain $\mathcal{X}_{k+1}(g)$,

(iii) $\phi^X_{g,k}$ is one-to-one,

(iv) $\phi^X_{g,k}$ is onto.

(i) $P^X_{\text{elev}}(g)$ holds:

By the definition of r_1, r_2 and Lemma 3.4.1, the path $r_1 r_2$ is not BPT and has a unique HPB with elevation $-e$ occurring at the last point of the r_1 subpath of $r_1 r_2$. By Lemma 3.1.5*(i)*,

the path $r_2 r_1$ has a unique LPA with elevation e occurring at the last point of the r_2 subpath of $r_2 r_1$. (3.6.3)

Since t has no flaws, the path $r_2 r_1 t$ also has a unique LPA with elevation e, occurring at the last point of the r_2 subpath of $r_2 r_1 t$. To show that $P^X_{\text{elev}}(g)$ holds, it is therefore sufficient by (3.6.2) to show that

the elevation of the HPBs (if any) of $\mathbb{s}$ is strictly less than $-e$. (3.6.4)

If $h = 0$, then $\mathbb{s} \in \mathcal{Q}(0)$ is empty. Otherwise $h > 0$, and so $s \in Q(h)$ has at least one flaw by Definition 3.3.1. Since $r_1 s r_2$ is not BPT by Lemma 3.4.1, and the last point of the r_1 subpath of $r_1 s r_2$ has elevation $-e$, the elevation of the LPAs of s is strictly greater than e. Then by $R(h)$ of the inductive hypothesis (with $h > 0$), $R_{\text{elev}}(h)$ holds and so the elevation of the HPBs of $\mathbb{s}$ is strictly less than $-e$. Therefore (3.6.4) holds.

(ii) $\phi^X_{g,k}$ has codomain $\mathcal{X}_{k+1}(g)$:

We shall show that $\mathbb{p} = r_2 \mathbb{s} r_1 t$ is in $\mathcal{X}(g)$ and that $\mathbb{p}$ has $k+1$ flaws.

By the definition of r_1, r_2, the path $r_2 r_1$ is not BPT. Then from (3.6.3) and (3.6.4),

$$r_2 \mathbb{s} r_1 \text{ is not BPT.} \tag{3.6.5}$$

We also find from (3.6.3) and (3.6.4) that the LPAs of $r_2 \mathbb{s} r_1$ are identically the boundary points of $\mathbb{s}$. By the inductive hypothesis, the codomain of ψ_h is $\mathcal{Q}(h)$, and so $\mathbb{s} \in \mathcal{Q}(h)$. It then follows from Definitions 3.3.1 and 3.3.3 that $r_2 \mathbb{s} r_1 \in \mathcal{X}(h^*)$ for some h^*. Since t has no flaws, we then have $\mathbb{p} = r_2 \mathbb{s} r_1 t \in \mathcal{X}(g)$.

It remains to show that $\mathbb{p} = r_2 \mathbb{s} r_1 t$ has $k+1$ flaws. Let s (in isolation) have k' flaws, and let $\mathbb{s} \in \mathcal{Q}(h)$ have $m+1$ boundary points. Since $r_1 r_2$ is not BPT by Lemma 3.4.1, this implies that $r_1 r_2 \mathbb{s}$ has $m+2$ boundary points.

Claim 1. The path $r_1 r_2$ has $k - k'$ flaws.

Claim 2. The path $\mathbb{s}$ has $k' - m$ flaws.

Combining Claims 1 and 2, we see that $r_1 r_2 \mathbb{s}$ has $(k - k') + (k' - m) = k - m$ flaws. Then from Lemma 3.1.5*(ii)* we find that $(r_2 \mathbb{s}) r_1$ has $(k - m) + (m + 1) = k + 1$ flaws. Since t has no flaws, this implies that $\mathbb{p} = r_2 \mathbb{s} r_1 t$ also has $k+1$ flaws, as required.

We now prove Claim 1. We know that $p = r_1 s t r_2 \in X_k(g)$ has k flaws, and that s (in isolation) has k' flaws. Since the s subpath of $p = r_1 s t r_2$ starts and ends at an HPB of p, the path $r_1 t r_2$ (in isolation) has $k - k'$ flaws. Since t has no flaws, the path $r_1 r_2$ also has $k - k'$ flaws, proving Claim 1.

We now prove Claim 2. If $h = 0$, then both s and $\mathbb{s}$ are empty and $m = k' = 0$, and so $\mathbb{s}$ has $0 = k' - m$ flaws. Otherwise $h > 0$, and then by $R_{\text{flaw}}(h)$*(ii)* of the inductive hypothesis, $\mathbb{s}$ has m fewer flaws than s, namely $k' - m$ flaws. This proves Claim 2.

(iii) $\phi^X_{g,k}$ is one-to-one:

Let $p' \in X_k(g)$ satisfy $\phi^X_{g,k}(p') = \phi^X_{g,k}(p)$.

Write $p' = r_1' s' t' r_2'$ according to Lemma 3.4.1, where $s' \in Q(h')$ for some h' satisfying $0 \le h' < g$. By the inductive hypothesis, $\psi_{h'}$ has codomain $\mathcal{Q}(h')$. Let $\mathbb{s}' = \psi_{h'}(s') \in \mathcal{Q}(h')$. Using Definition 3.4.3, we have

$$r_2 \mathbb{s} r_1 t = \phi^X_{g,k}(p) = \phi^X_{g,k}(p') = r_2' \mathbb{s}' r_1' t'. \tag{3.6.6}$$

We know by (3.6.5) that $r_2 \mathbb{s} r_1$ is not BPT, and similarly that $r_2' \mathbb{s}' r_1'$ is not BPT. We therefore conclude from (3.6.6) that $r_2 \mathbb{s} r_1 = r_2' \mathbb{s}' r_1'$ and $t = t'$.

We know by (3.6.3) and (3.6.4) that the subpath $\mathbb{s}$ of $r_2\mathbb{s}r_1$ starts at the first LPA and ends at the last LPA of the path $r_2\mathbb{s}r_1$. The same is true of the $\mathbb{s}'$ subpath of $r_2'\mathbb{s}'r_1'$. Since $r_2\mathbb{s}r_1$ and $r_2'\mathbb{s}'r_1'$ are the same path, they have the same LPAs and so $\mathbb{s} = \mathbb{s}'$ and therefore $r_1 = r_1'$ and $r_2 = r_2'$.

We have seen that ψ_h has codomain $\mathcal{Q}(h)$, and $\psi_{h'}$ has codomain $\mathcal{Q}(h')$. It follows from $\psi_h(s) = \mathbb{s} = \mathbb{s}' = \psi_{h'}(s')$ that $h = h'$. Then $s = s'$, because ψ_h is a bijection by the inductive hypothesis. Therefore $p = r_1str_2 = r_1's't'r_2' = p'$.

(iv) $\phi^X_{g,k}$ is onto:

(Note that we reassign the variable names $p, \mathbb{p}, r_1, r_2, t, s, \mathbb{s}, e, h$ in the rest of this proof.) Let $\mathbb{p} \in \mathcal{X}_{k+1}(g)$. Write $\mathbb{p} = r_2\mathbb{s}r_1t$ according to Lemma 3.4.2, where $\mathbb{s} \in \mathcal{Q}(h)$ for some h satisfying $0 \leq h < g$ and t (possibly empty) has no flaws. Since ψ_h is a bijection by the inductive hypothesis, we may define $s = \psi_h^{-1}(\mathbb{s}) \in Q(h)$. Let $p = r_1str_2$. We shall show that $p \in X_k(g)$ and that $\phi^X_{g,k}(p) = \mathbb{p}$.

Let the LPAs of $\mathbb{p}$ have elevation e.

Claim 1. r_1r_2 is not BPT and has a unique HPB at elevation $-e$ occurring at the last point of the r_1 subpath.

Claim 2. The LPAs (if any) of $s \in Q(h)$ have elevation strictly greater than e.

Since t has no flaws, Claims 1 and 2 imply that

(1) $p = r_1str_2$ is not BPT,

(2) the first and last HPB of p occur at the first and last point of the subpath st.

From *(1)* and Definition 3.3.2, we have $p \in X_{k^*}(g)$ for some k^*. From *(2)* and Definition 3.3.1, the split r_1str_2 of $p \in X_{k^*}(g)$ is consistent with the split described in Lemma 3.4.1 and so by Definition 3.4.3 we have

$$\phi^X_{g,k^*}(p) = r_2\psi_h(s)r_1t = r_2\mathbb{s}r_1t = \mathbb{p} \in \mathcal{X}_{k+1}(g).$$

It then follows from *(ii)* that $\mathbb{p} \in \mathcal{X}_{k^*+1}(g)$, and so $k^* = k$ as required.

We now prove Claim 1. By the definition of r_1, r_2 and Lemma 3.4.2, r_2r_1 is not BPT and has a unique LPA at elevation e occurring at the last point of the r_2 subpath of r_2r_1. Claim 1 now follows from Lemma 3.1.5*(i)*.

We now prove Claim 2. If $h = 0$, then $s \in Q(0)$ is empty. Otherwise $h > 0$, and then by definition of $\mathbb{s}$ in Lemma 3.4.2 the HPBs of $\mathbb{s}$ have elevation $-e'$ for some $e' > e$. Let the elevation of the LPAs of $s \in Q(h)$ be d. Since $h > 0$, we may use $R_{\text{elev}}(h)$ of the inductive hypothesis to show that $d = e'$. Therefore the elevation of the LPAs of s is strictly greater than e, proving Claim 2. $\square$

Claim 3.6.7. *The statement $P^Y_{\text{bij}}(g)$ holds.*

Proof. Let k satisfy $0 \leq k < g(a+b) - 1$. Let $p \in Y_k(g)$ and write $p = p_1 p_2$ according to Lemma 3.4.1, where p_1 is not BPT and $p_2 \in N(h) \setminus S(h)$ for some h satisfying $0 < h < g$. Write $\mathbb{p} = \phi^Y_{g,k}(p)$. By Definition 3.4.3,

$$\mathbb{p} = p_1 \phi_h(p_2). \tag{3.6.7}$$

We shall show in the following sequence of steps that $P^Y_{\text{bij}}(g)$ holds:

(i) $\phi^Y_{g,k}$ has codomain $\mathcal{Y}_{k+1}(g)$,

(ii) $\phi^Y_{g,k}$ is one-to-one,

(iii) $\phi^Y_{g,k}$ is onto.

(i) $\phi^Y_{g,k}$ has codomain $\mathcal{Y}_{k+1}(g)$:

We shall show that $\mathbb{p} = p_1 \phi_h(p_2)$ has $k+1$ flaws and that $\mathbb{p}$ is in $\mathcal{Y}(g)$.

Let p_2 have k' flaws, and then p_1 has $k - k'$ flaws. Since $h > 0$, by the inductive hypothesis $\phi_{h,k'}$ maps $p_2 \in N_{k'}(h) \setminus S_{k'}(h)$ to a path in $N_{k'+1}(h)$. Therefore $\phi_h(p_2) = \phi_{h,k'}(p_2)$ has $k'+1$ flaws. The number of flaws of $\mathbb{p} = p_1 \phi_h(p_2)$ is then $(k - k') + (k' + 1) = k + 1$. Since $\phi_h(p_2)$ has $k' + 1 > 0$ flaws, by Definition 3.3.3 we obtain $\mathbb{p} = p_1 \phi_h(p_2) \in \mathcal{Y}(g)$.

(ii) $\phi^Y_{g,k}$ is one-to-one:

Let $p' \in Y_k(g)$ satisfy $\phi^Y_{g,k}(p) = \phi^Y_{g,k}(p')$.

Write $p' = p'_1 p'_2$ according to Lemma 3.4.1, where p'_1 is not BPT and $p'_2 \in N(h') \setminus S(h')$ for some h' satisfying $0 < h' < g$. Using Definition 3.4.3, we have

$$p_1 \phi_h(p_2) = \phi^Y_{g,k}(p) = \phi^Y_{g,k}(p') = p'_1 \phi_{h'}(p'_2).$$

Since p_1 and p'_1 are not BPT, it follows that $p_1 = p'_1$ and $\phi_h(p_2) = \phi_{h'}(p'_2)$. Since $h > 0$ and $h' > 0$, by the inductive hypothesis ϕ_h and $\phi_{h'}$ have codomain $N(h)$ and $N(h')$, respectively. It then follows from $\phi_h(p_2) = \phi_{h'}(p'_2)$ that $h = h'$. Since ϕ_h is a bijection by the inductive hypothesis, this gives $p_2 = p'_2$ and so $p = p_1 p_2 = p'_1 p'_2 = p'$.

(iii) $\phi^Y_{g,k}$ is onto:

(Note that we reassign variable names in the rest of this proof.) Let $\mathbb{p} \in \mathcal{Y}_{k+1}(g)$. Write $\mathbb{p} = \mathbb{p}_1\mathbb{p}_2$ according to Lemma 3.4.2, where $\mathbb{p}_1$ is not BPT and $\mathbb{p}_2 \in N_{k'+1}(h)$ for some h, k' satisfying $0 < h < g$ and $k' \geq 0$.

Since $\mathbb{p}_2$ has $k' + 1$ flaws, the path $\mathbb{p}_1$ has $k - k'$ flaws. Since $h > 0$, by the inductive hypothesis $\phi_{h,k'}$ is a bijection from $N_{k'}(h) \setminus S_{k'}(h)$ to $N_{k'+1}(h)$. We may therefore define

$$p_2 = \phi^{-1}_{h,k'}(\mathbb{p}_2) \in N_{k'}(h) \setminus S_{k'}(h),$$

so p_2 has k' flaws. Let $p = \mathbb{p}_1 p_2$, which has $(k - k') + k' = k$ flaws. Since $p_2 \notin S$, we have by Definition 3.3.2 that $p \in Y_k(g)$. Since $\mathbb{p}_1$ is not BPT, the split $\mathbb{p}_1 p_2$ of p is consistent with the split described in Lemma 3.4.1. Therefore by Definition 3.4.3 we have

$$\phi_g(p) = \phi^Y_{g,k}(\mathbb{p}_1 p_2) = \mathbb{p}_1 \phi_h(p_2) = \mathbb{p}_1\mathbb{p}_2 = \mathbb{p},$$

as required. □

Claim 3.6.8. *The statement $P^Z_{\text{bij}}(g)$ holds.*

Proof. Let k satisfy $0 \leq k < g(a+b) - 1$. Let $p \in Z_k(g)$ and write $p = r_1 r_2 t s$ according to Lemma 3.4.1, where t (possibly empty) has no flaws and s (non-empty) has max flaws. Write $\mathbb{p} = \phi^Z_{g,k}(p)$. By Definition 3.4.3,

$$\mathbb{p} = r_1 s r_2 t. \tag{3.6.8}$$

We shall show in the following sequence of steps that $P^Z_{\text{bij}}(g)$ holds:

(i) $\phi^Z_{g,k}$ has codomain $\mathcal{Z}_{k+1}(g)$,

(ii) $\phi^Z_{g,k}$ is one-to-one,

(iii) $\phi^Z_{g,k}$ is onto.

(i) $\phi^Z_{g,k}$ has codomain $\mathcal{Z}_{k+1}(g)$:

We shall show that $\mathbb{p} = r_1 s r_2 t = \phi^Z_{g,k}(p)$ has $k + 1$ flaws and that $\mathbb{p}$ is a member of $\mathcal{Z}(g)$.

The path s has max flaws, and the last point of the r_1 subpath of $r_1 r_2 t s$ lies above the boundary. Therefore the path $\mathbb{p} = r_1 s r_2 t$ has one more flaw than $p = r_1 r_2 t s$ and so has $k + 1$ flaws.

By Lemma 3.4.1, $r_1 s r_2$ is not BPT, and it follows by Definitions 3.3.1 and 3.3.3 that $r_1 s r_2 \in \mathcal{Z}(h)$ for some h. Since t has no flaws, it follows that $r_1 s r_2 t \in \mathcal{Z}(g)$.

(ii) $\phi^Z_{g,k}$ is one-to-one:

Let $p' \in Z_k(g)$ satisfy $\phi^Z_{g,k}(p) = \phi^Z_{g,k}(p')$.

Write $p' = r'_1 r'_2 t' s'$ according to Lemma 3.4.1, where s' has max flaws. Using Definition 3.4.3, we have

$$r_1 s r_2 t = \phi^Z_{g,k}(p) = \phi^Z_{g,k}(p') = r'_1 s' r'_2 t'.$$

Since $r_1 s r_2$ and $r'_1 s' r'_2$ are not BPT by Lemma 3.4.1, it follows that $r_1 s r_2 = r'_1 s' r'_2$ and $t = t'$.

Since s has max flaws, and the last point of subpath r_1 of $r_1 r_2$ is the last LPA of $r_1 r_2$, the subpath s of $r_1 s r_2$ connects the last two LPAs of $r_1 s r_2$. The same is true of the s' subpath of $r'_1 s' r'_2$. Since $r_1 s r_2$ and $r'_1 s' r'_2$ are the same path, they have the same LPAs and so $s = s'$. Therefore $r_1 = r'_1$ and $r_2 = r'_2$, and so $p = r_1 r_2 t s = r'_1 r'_2 t' s' = p'$.

(iii) $\phi^Z_{g,k}$ is onto:

(Note that we reassign variable names in the rest of this proof.) Let $\mathbb{p} \in \mathcal{Z}_{k+1}(g)$. Write $\mathbb{p} = r_1 s r_2 t$ according to Lemma 3.4.2, where t (possibly empty) has no flaws and s (non-empty) has max flaws. Let $p = r_1 r_2 t s$. We shall show that $p \in Z_k(g)$ and that $\phi^Z_{g,k}(p) = \mathbb{p}$.

By Definition 3.3.2, we have $p \in Z_{k^*}(g)$ for some k^*. The split $r_1 r_2 t s$ of $p \in Z_{k^*}(g)$ is consistent with the split described in Lemma 3.4.1 and so by Definition 3.4.3 we have

$$\phi^Z_{g,k^*}(p) = r_1 s r_2 t = \mathbb{p} \in \mathcal{Z}_{k+1}(g).$$

It then follows from *(i)* that $\mathbb{p} \in \mathcal{Z}_{k^*+1}(g)$, and so $k^* = k$ as required. □

Claim 3.6.9. *The statement $P_{\text{flaw}}(g)$ holds.*

Proof. Let $p \in N(g) \setminus S(g)$ and $\mathbb{p} = \phi_g(p)$, and write $p = p_1 \cdots p_n$ and $\mathbb{p} = \mathbb{p}_1 \cdots \mathbb{p}_m$ where each path is split at its respective boundary points.

(i): Suppose p_1 has at least one flaw. We shall show that $\mathbb{p}_1$ has at least one flaw in each of the cases $p \in X(g)$, $p \in Y(g)$, $p \in Z(g)$.

Case $p \in X(g)$**.** Write $p = r_1 s t r_2 \in X(g)$ according to Lemma 3.4.1, where $s \in Q(h)$ for some h satisfying $0 \le h < g$. By Definition 3.4.3,

$$\mathbb{p}_1 \cdots \mathbb{p}_m = \phi_g(p) = r_2 \psi_h(s) r_1 t.$$

By the definition of r_2, none of the interior boundary points of $r_2 \psi_h(s) r_1 t$ is contained in the r_2 subpath, and the last point of this subpath lies above the boundary. Therefore $\mathbb{p}_1$ has at least one flaw.

Case $p \in Y(g)$**.** Using Definition 3.4.3, we have

$$\mathbb{p}_1 \cdots \mathbb{p}_m = \phi_g(p) = p_1 \phi_h(p_2 \cdots p_n)$$

for some h satisfying $0 < h < g$. Since $\mathbb{p}_1$ and p_1 are not BPT by definition, $\mathbb{p}_1 = p_1$ and so $\mathbb{p}_1$ has at least one flaw.

Case $p \in Z(g)$**.** Write $p = r_1 r_2 t s \in Z(g)$ according to Lemma 3.4.1. By Definition 3.4.3,

$$\mathbb{p}_1 \cdots \mathbb{p}_m = \phi_g(p) = r_1 s r_2 t.$$

By the definition of r_1, none of the interior boundary points of $r_1 s r_2 t$ is contained in the r_1 subpath, and the last point of this subpath lies above the boundary. Therefore $\mathbb{p}_1$ has at least one flaw.

(ii): Suppose that $n > 1$ and that $\mathbb{p}_1$ has at least one flaw. We shall show that p_1 has at least one flaw. Since $n > 1$, the path p is BPT and so $p \notin X(g)$ by Definition 3.3.2. Suppose, for a contradiction, that p_1 has no flaws. Since $p = p_1 p_2 \ldots p_n \notin S(g)$ by assumption, this implies that $p_2 \ldots p_n \notin S$. Therefore $p \in Y(g)$ by Definition 3.3.2. By Definition 3.4.3,

$$\mathbb{p}_1 \cdots \mathbb{p}_m = \phi_g(p) = p_1 \phi_h(p_2 \cdots p_n)$$

for some h satisfying $0 < h < g$. Since $\mathbb{p}_1$ and p_1 are not BPT by definition, $\mathbb{p}_1 = p_1$ and so $\mathbb{p}_1$ has no flaws, contrary to assumption. □

Claims 3.6.6 to 3.6.9 collectively establish Lemma 3.6.5, completing step III of the Proof Roadmap. The final step of the Roadmap is step IV, which we prove in Lemma 3.6.10.

Lemma 3.6.10. *Subject to the inductive hypothesis* (3.6.1)*, statement* $R(g)$ *holds.*

We split the proof of Lemma 3.6.10 into Claims 3.6.11 to 3.6.13, and make use of Lemma 3.6.5.

Claim 3.6.11. *The statement* $R_{\text{bij}}(g)$ *holds.*

Proof. Let $q \in Q(g)$ and write $q = q_1 q_2$ according to Lemma 3.5.4, where $q_2 \in Q(h)$ for some h satisfying $0 \leq h < g$. Write $\mathfrak{q} = \psi_g(q)$. By Definition 3.5.5,

$$\mathfrak{q} = \overline{\phi_g\left(\overline{q_1 \psi_h(q_2)}\right)}. \tag{3.6.9}$$

We shall show in the following sequence of steps that $R_{\text{bij}}(g)$ holds:

(i) the expression (3.6.9) for $\mathfrak{q}$ is well-defined,

(ii) ψ_g has codomain $\mathcal{Q}(g)$,

(iii) ψ_g is one-to-one,

(iv) ψ_g is onto.

(i) The expression (3.6.9) for $\mathfrak{q}$ is well-defined:

It is sufficient to show that $\overline{q_1 \psi_h(q_2)} \notin S$, so that $\overline{q_1 \psi_h(q_2)}$ lies in the domain of ϕ_g.

Suppose firstly that $h = 0$, so that q_1 has at least one flaw by Lemma 3.5.4. Since q_1 is not BPT by definition,

$$\overline{q_1} \text{ is not BPT and has non-max flaws (when } h = 0). \tag{3.6.10}$$

Therefore $\overline{q_1 \psi_h(q_2)} = \overline{q_1 \psi_0(q_2)} = \overline{q_1} \notin S$, as required.

We may therefore take $h > 0$. By the inductive hypothesis, $\psi_h(q_2) \in \mathcal{Q}(h)$ and so $\psi_h(q_2)$ has at least one point below the boundary by Definition 3.3.1. Therefore $\overline{\psi_h(q_2)}$ has at least one flaw, and since q_1 is non-empty by definition we find that

$$\overline{q_1 \psi_h(q_2)} = \overline{\psi_h(q_2)}\ \overline{q_1} \notin S,$$

as required.

(ii) ψ_g has codomain $\mathcal{Q}(g)$:

Split $\mathfrak{q} = \overline{\phi_g\left(\overline{q_1 \psi_h(q_2)}\right)}$ at its boundary points into $\mathfrak{q} = \mathfrak{q}_1 \cdots \mathfrak{q}_m$. We shall show that $\mathfrak{q} \in \mathcal{Q}(g)$ by showing that $\overline{\mathfrak{q}_m}$ has at least one flaw (so that $\mathfrak{q}_m$ has non-max flaws).

Equate the two expressions for $\overline{\mathfrak{q}}$ to give

$$\overline{\mathfrak{q}_m} \cdots \overline{\mathfrak{q}_1} = \phi_g\left(\overline{q_1 \psi_h(q_2)}\right). \tag{3.6.11}$$

Suppose firstly that $h = 0$. From (3.6.10) and Definition 3.3.2 we have $\overline{q_1} \in X(g)$. By Lemma 3.6.5, $P^X_{\text{bij}}(g)$ holds and so $\phi_g(\overline{q_1}) \in \mathcal{X}(g)$. Then by (3.6.11), we have

$\overline{\mathbb{q}_m} \cdots \overline{\mathbb{q}_1} = \phi_g\left(\overline{q_1\psi_0(q_2)}\right) = \phi_g(\overline{q_1}) \in \mathcal{X}(g)$ and so $\overline{\mathbb{q}_m}$ has at least one flaw by Definition 3.3.3, as required.

We may therefore take $h > 0$. Let $\mathbb{r} = \psi_h(q_2)$. Split $\mathbb{r}$ at its boundary points into $\mathbb{r} = \mathbb{r}_1 \cdots \mathbb{r}_n$. Then by (3.6.11) we have

$$\overline{\mathbb{q}_m} \cdots \overline{\mathbb{q}_1} = \phi_g\left(\overline{q_1\psi_h(q_2)}\right) = \phi_g\left(\overline{\mathbb{r}_n} \cdots \overline{\mathbb{r}_1}\ \overline{q_1}\right). \tag{3.6.12}$$

By the inductive hypothesis, $\mathbb{r} = \mathbb{r}_1 \cdots \mathbb{r}_n = \psi_h(q_2) \in \mathcal{Q}(h)$ and so $\mathbb{r}_n$ has non-max flaws by Definition 3.3.1. Therefore, $\overline{\mathbb{r}_n}$ has at least one flaw and so $\overline{\mathbb{r}_n} \cdots \overline{\mathbb{r}_1}\ \overline{q_1} \notin S$. By Lemma 3.6.5, we may apply $P_{\text{flaw}}(g)$*(i)* to (3.6.12) to conclude that $\overline{\mathbb{q}_m}$ has at least one flaw, as required.

(iii) ψ_g is one-to-one:

Let $q' \in Q(g)$ satisfy $\psi_g(q) = \psi_g(q')$.

Write $q' = q_1'q_2'$ according to Lemma 3.5.4, where $q_2' \in Q(h')$ for some h' satisfying $0 \leq h' < g$. Using Definition 3.5.5, we have

$$\overline{\phi_g\left(\overline{q_1\psi_h(q_2)}\right)} = \psi_g(q) = \psi_g(q') = \overline{\phi_g\left(\overline{q_1'\psi_{h'}(q_2')}\right)}.$$

By Lemma 3.6.5, we know that ϕ_g is a bijection and so

$$q_1\psi_h(q_2) = q_1'\psi_{h'}(q_2').$$

Since both q_1 and q_1' are not BPT by definition, it follows that $q_1 = q_1'$ and so

$$\psi_h(q_2) = \psi_{h'}(q_2').$$

Since ψ_h has codomain $\mathcal{Q}(h)$ and $\psi_{h'}$ has codomain $\mathcal{Q}(h')$ by the inductive hypothesis, this implies $h = h'$. Then $q_2 = q_2'$, because ψ_h is a bijection by the inductive hypothesis. Therefore, $q = q_1q_2 = q_1'q_2' = q'$.

(iv) ψ_g is onto:

(Note that we reassign variable names in the rest of this proof.) Let $\mathbb{q} \in \mathcal{Q}(g)$. We shall find $q \in Q(g)$ satisfying $\psi_g(q) = \mathbb{q}$.

Since $\overline{\mathbb{q}}$ has at least one flaw by Definition 3.3.1, we see from (3.2.1) that $\overline{\mathbb{q}}$ lies in the codomain of ϕ_g. Since ϕ_g is a bijection by Lemma 3.6.5, we may then define

$$\mathbb{r} = \phi_g^{-1}(\overline{\mathbb{q}}) \in N(g) \setminus S(g). \tag{3.6.13}$$

Split $\mathbb{r}$ at its boundary points into $\mathbb{r} = \mathbb{r}_1 \cdots \mathbb{r}_n$, so that

$$\phi_g(\mathbb{r}_1 \cdots \mathbb{r}_n) = \phi_g(\mathbb{r}) = \overline{\mathbb{q}}. \tag{3.6.14}$$

Claim 1. We have $\overline{\mathbb{r}_1 \cdots \mathbb{r}_{n-1}} \in \mathcal{Q}(h)$ for some h satisfying $0 \le h < g$.

In view of Claim 1, and since $\psi_h : Q(h) \to \mathcal{Q}(h)$ is a bijection by the inductive hypothesis, we may define $q = \overline{\mathbb{r}_n}\, q'$ where

$$q' = \psi_h^{-1}(\overline{\mathbb{r}_1 \cdots \mathbb{r}_{n-1}}) \in Q(h). \tag{3.6.15}$$

Claim 2. We have $q \in Q(g)$.

Since $\mathbb{r}_n$ is not BPT, $\overline{\mathbb{r}_n}$ is not BPT. Then by Claim 2, the split $\overline{\mathbb{r}_n}\, q'$ of $q \in Q(g)$ is consistent with the split described in Lemma 3.5.4. Therefore by Definition 3.5.5, (3.6.15), and (3.6.14),

$$\psi_g(q) = \overline{\phi_g\left(\overline{\psi_h(q')}\, \mathbb{r}_n\right)} = \overline{\phi_g\Big((\mathbb{r}_1 \cdots \mathbb{r}_{n-1})\, \mathbb{r}_n\Big)} = \overline{\phi_g(\mathbb{r})} = \mathbb{q},$$

as required.

We now prove Claim 1. If $n = 1$, then $\overline{\mathbb{r}_1 \cdots \mathbb{r}_{n-1}}$ is the empty path and so Claim 1 holds with $h = 0$. We may therefore take $n > 1$, so that $\overline{\mathbb{r}_1 \cdots \mathbb{r}_{n-1}} \in N(h)$ for some h satisfying $0 < h < g$. Split $\mathbb{q} \in \mathcal{Q}(g)$ at its boundary points into $\mathbb{q} = \mathbb{q}_1 \cdots \mathbb{q}_m$. Then $\overline{\mathbb{q}_m}$ has at least one flaw by Definition 3.3.1, and by (3.6.14) we have

$$\overline{\mathbb{q}_m} \cdots \overline{\mathbb{q}_1} = \phi_g(\mathbb{r}_1 \cdots \mathbb{r}_n).$$

By Lemma 3.6.5, we may therefore use $P_{\text{flaw}}(g)$ *(ii)* to conclude that $\mathbb{r}_1$ has at least one flaw and so $\overline{\mathbb{r}_1}$ has non-max flaws. Then by Definition 3.3.1, the path $\overline{\mathbb{r}_1 \cdots \mathbb{r}_{n-1}} = \overline{\mathbb{r}_{n-1}} \cdots \overline{\mathbb{r}_1}$ is a member of $\mathcal{Q}(h)$. This proves Claim 1.

We now prove Claim 2. We have $q' \in Q(h)$ by (3.6.15). If $h > 0$, then $q = \overline{\mathbb{r}_n}\, q' \in Q(g)$ by Definition 3.3.1. We may therefore take $h = 0$ so that $n = 1$ by Claim 1. Since $\mathbb{r}_n = \mathbb{r} \notin S$ by (3.6.13) and $\mathbb{r}_n$ is not BPT, $\mathbb{r}_n$ has at least one point below the boundary. Therefore $\overline{\mathbb{r}_n}$ has at least one flaw and is not BPT. Since $q' \in Q(0)$ is empty, $q = \overline{\mathbb{r}_n}\, q' = \overline{\mathbb{r}_n} \in Q(g)$ by Definition 3.3.1. This proves Claim 2. □

Claim 3.6.12. *The statement $R_{\text{flaw}}(g)$ holds.*

Proof. Let $q \in Q(g)$ and $\mathbb{q} = \psi_g(q)$, and write $q = q_1 \cdots q_n$ and $\mathbb{q} = \mathbb{q}_1 \cdots \mathbb{q}_m$ where each path is split at its respective boundary points. Then q has $n+1$ boundary points and $\mathbb{q}$ has

$m+1$ boundary points, and by Remark 3.5.3 parts *(i)* and *(ii)* of $R_{\text{flaw}}(g)$ are equivalent. We therefore prove only part *(i)*.

Let $\overline{q}$ have k flaws. We shall show that $\overline{\mathbb{q}}$ has $k+n$ flaws.

The split $(q_1)(q_2 \dots q_n)$ of $q \in Q(g)$ is consistent with the split described in Lemma 3.5.4, where $q_2 \cdots q_n \in Q(h)$ for some h satisfying $0 \le h < g$. By Definition 3.5.5, we have

$$\overline{\mathbb{q}} = \overline{\psi_g(q)} = \phi_g\left(\overline{\psi_h(q_2 \cdots q_n)}\ \overline{q_1}\right), \tag{3.6.16}$$

which is a valid expression by Claim 3.6.11.

Suppose firstly that $h=0$, so $n=1$. Then $q_2 \cdots q_n \in Q(0)$ is empty and so $\overline{\mathbb{q}} = \phi_g(\overline{q_1})$ by (3.6.16). Since $\overline{q_1} = \overline{q}$ has k flaws, Lemma 3.6.5 implies that $\overline{\mathbb{q}} = \phi_g(\overline{q_1})$ has $k+1 = k+n$ flaws, as required.

We may therefore take $h>0$. Let $\overline{q_1}$ have k' flaws. Since $\overline{q} = \overline{q_2 \cdots q_n}\ \overline{q_1}$, this implies that $\overline{q_2 \cdots q_n}$ has $k-k'$ flaws. Since $h>0$ and $q_2 \dots q_n \in Q(h)$, we may use $R_{\text{flaw}}(h)$*(i)* to show that $\overline{\psi_h(q_2 \cdots q_n)}$ has $(k-k')+(n-1)$ flaws. The number of flaws of $\overline{\psi_h(q_2 \cdots q_n)}\ \overline{q_1}$ is therefore $(k-k'+n-1)+k' = k+n-1$. It then follows from Lemma 3.6.5 and (3.6.16) that $\overline{\mathbb{q}} = \phi_g\left(\overline{\psi_h(q_2 \cdots q_n)}\ \overline{q_1}\right)$ has $(k+n-1)+1 = k+n$ flaws. □

Claim 3.6.13. *The statement $R_{\text{elev}}(g)$ holds.*

Proof. Let $q \in Q(g)$ and let the LPAs of q have elevation e. Write $q = q_1 q_2$ according to Lemma 3.5.4, where $q_2 \in Q(h)$ for some h satisfying $0 \le h < g$. Define $\mathbb{r} = \overline{q_1 \psi_h(q_2)}$. By Definition 3.5.5, we have

$$\overline{\psi_g(q)} = \phi_g\left(\overline{q_1 \psi_h(q_2)}\right) = \phi_g(\mathbb{r}).$$

It is therefore sufficient to show that the LPAs of $\phi_g(\mathbb{r})$ have elevation e.

Suppose firstly that $h=0$. Then $q_2 \in Q(0)$ is empty and so $\mathbb{r} = \overline{q_1}$. Since q_1 is not BPT by definition and $q_1 = q \in Q(g)$ has at least one flaw, $\mathbb{r} = \overline{q_1}$ is not BPT and has non-max flaws and therefore $\mathbb{r} \in X(g)$. Since the LPAs of $q_1 = q$ have elevation e, the HPBs of $\mathbb{r} = \overline{q_1}$ have elevation $-e$. By Lemma 3.6.5, we may apply $P^X_{\text{elev}}(g)$ to $\mathbb{r} \in X(g)$ to conclude that the LPAs of $\phi_g(\mathbb{r})$ have elevation e, as required.

We may therefore take $h>0$ for the remainder of this proof. Split $\mathbb{r}$ at its boundary points into $\mathbb{r} = \mathbb{r}_1 \cdots \mathbb{r}_n = \overline{\psi_h(q_2)}\ \overline{q_1}$, where $n>1$. Since q_1 is not BPT by definition, this implies that

$$\mathbb{r}_1 \cdots \mathbb{r}_{n-1} = \overline{\psi_h(q_2)} \quad \text{and} \quad \mathbb{r}_n = \overline{q_1},$$

and so $q = \overline{\mathbb{r}_n}\, q_2$.

Case 1. $\mathbb{r}_n$ has max flaws.

Then all flaws of $q = \overline{\mathbb{r}_n}\, q_2$ occur within q_2, and so the LPAs of q_2 have elevation e. Since $h > 0$, we may apply $R_{\text{elev}}(h)$ of the inductive hypothesis to $q_2 \in Q(h)$ to show that the HPBs of $\psi_h(q_2)$ have elevation $-e$. Therefore

$$\text{the LPAs of } \mathbb{r}_1 \cdots \mathbb{r}_{n-1} = \overline{\psi_h(q_2)} \text{ have elevation } e. \tag{3.6.17}$$

In particular, $\mathbb{r}_1 \cdots \mathbb{r}_{n-1}$ has at least one flaw and so $\mathbb{r} = (\mathbb{r}_1 \cdots \mathbb{r}_{n-1})\mathbb{r}_n \notin S$.

Since $\mathbb{r} \notin S$ and $\mathbb{r}_n$ has max flaws, there exists a minimum index γ satisfying $1 \leq \gamma < n$ such that $\mathbb{r}_{\gamma+1} \cdots \mathbb{r}_n \in S$. Therefore

$$\mathbb{r}_\gamma \cdots \mathbb{r}_n \in Z(h') \text{ for some } h' \text{ satisfying } 1 < h' \leq g$$

and for all j satisfying $1 \leq j < \gamma$,

$$\mathbb{r}_j \cdots \mathbb{r}_n \in Y(h_j) \text{ for some } h_j \text{ satisfying } 1 < h_j \leq g.$$

By repeated application of Case 2 of Definition 3.4.3, we obtain

$$\phi_g(\mathbb{r}) = (\mathbb{r}_1 \cdots \mathbb{r}_{\gamma-1})\phi_{h'}(\mathbb{r}_\gamma \cdots \mathbb{r}_n). \tag{3.6.18}$$

By definition of γ, we have that $\mathbb{r}_\gamma$ has at least one flaw. Split $\mathbb{r}_\gamma$ at its last LPA into $\mathbb{r}_\gamma = r_1 r_2$. By Case 3 of Definition 3.4.3,

$$\phi_{h'}(\mathbb{r}_\gamma \cdots \mathbb{r}_n) = \phi_{h'}\Big(r_1 r_2(\mathbb{r}_{\gamma+1} \cdots \mathbb{r}_{n-1})\mathbb{r}_n\Big) = r_1 \mathbb{r}_n r_2(\mathbb{r}_{\gamma+1} \cdots \mathbb{r}_{n-1}),$$

and so by (3.6.18)

$$\phi_g(\mathbb{r}) = (\mathbb{r}_1 \cdots \mathbb{r}_{\gamma-1}) r_1 \mathbb{r}_n r_2(\mathbb{r}_{\gamma+1} \cdots \mathbb{r}_{n-1}).$$

Since $\mathbb{r}_n$ has max flaws, it follows using (3.6.17) that the LPAs of $\phi_g(\mathbb{r})$ have elevation e, as required.

Case 2. $\mathbb{r}_n$ has non-max flaws.

The path $q_2 \in Q(h)$ has at least one flaw because $h > 0$, and the path $\mathbb{r}_n$ has a point below the boundary by assumption. We may therefore let the HPBs of $\mathbb{r}_n$ have elevation $-e_1$, and let the LPAs of q_2 have elevation e_2. Since the LPAs of $q = \overline{\mathbb{r}_n}\, q_2$ have elevation e, we have $e = \min\{e_1, e_2\}$.

Since $h > 0$, we may apply $R_{\text{elev}}(h)$ of the inductive hypothesis to show that the HPBs of $\psi_h(q_2)$ have elevation $-e_2$. Therefore

$$\text{the LPAs of } \mathtt{r}_1 \cdots \mathtt{r}_{n-1} = \overline{\psi_h(q_2)} \text{ have elevation } e_2. \tag{3.6.19}$$

Since $\mathtt{r}_n$ has non-max flaws, for all j satisfying $1 \leq j < n$ we have $\mathtt{r}_{j+1} \cdots \mathtt{r}_n \notin S$ and so

$$\mathtt{r}_j \cdots \mathtt{r}_n \in Y(h_j) \text{ for some } h_j \text{ satisfying } 1 < h_j \leq g.$$

By repeated application of Case 2 of Definition 3.4.3, we obtain

$$\phi_g(\mathtt{r}) = \mathtt{r}_1 \cdots \mathtt{r}_{n-1} \phi_{h'}(\mathtt{r}_n) \tag{3.6.20}$$

where $h' = g - h$ satisfies $0 < h' < g$. Since $\mathtt{r}_n$ has non-max flaws and is not BPT, $\mathtt{r}_n \in X(h')$. Since the HPBs of $\mathtt{r}_n$ have elevation $-e_1$, we may apply $P^X_{\text{elev}}(h')$ of the inductive hypothesis ($h' > 0$) to show that the LPAs of $\phi_{h'}(\mathtt{r}_n)$ have elevation e_1. It follows from (3.6.19) and (3.6.20) that the LPAs of $\phi_g(\mathtt{r})$ have elevation $e = \min\{e_1, e_2\}$, as required. □

Claims 3.6.11 to 3.6.13 collectively establish Lemma 3.6.10, completing step IV of the Proof Roadmap. Therefore $P(g)$ and $R(g)$ hold for all $g > 0$, proving Theorem 1.3.5.

Chapter 4

Alternative proof of the path enumeration formula

In this chapter, we give an alternative proof of the path enumeration formula.

Theorem 1.3.13 (Path enumeration formula)**.** *We have*

$$\mu_j(g) = \sum_{k=0}^{j} (-1)^k \mathbf{E}_k \mathbf{H}_{g-k} \quad \textit{for } 0 \le j < g.$$

The proof presented in Chapter 1 assumed that the value of both $\mu_0(g)$ and $\mu_{g-1}(g)$ is known (see Corollary 1.3.12). Here we shall assume that only the value of $\mu_0(g)$ is known. Our proof will require some results involving symmetric functions.

4.1 Alternative proof

We shall prove the path enumeration formula for $\mu_j(g)$ by showing that it satisfies the recurrence relation (1.3.5) and the initial values (1.3.11) for $\mu_0(g)$. Our proof depends on the following identity that we shall establish in Section 4.2.

Theorem 4.1.1 (Sum identity)**.** *Let $g > 0$. Then*

$$\sum_{k=0}^{g} (-1)^k \mathbf{E}_k \mathbf{H}_{g-k} = 0.$$

Alternative proof of Theorem 1.3.13. Let

$$M_j(g) = \sum_{k=0}^{j} (-1)^k \mathbf{E}_k \mathbf{H}_{g-k}.$$

We show that $\mu_j(g) = M_j(g)$ for all j, g satisfying $0 \leq j < g$ by showing that $M_0(g)$ takes the initial values (1.3.11) for $\mu_0(g)$, and that $M_j(g)$ satisfies the recurrence relation (1.3.5) for $\mu_j(g)$.

We have $M_0(g) = (-1)^0 \mathbf{E}_0 \mathbf{H}_g = \mathbf{H}_g$ by (1.3.10).

Set $g = j$ in Theorem 4.1.1 and use (1.3.10) to obtain the relation

$$0 = \sum_{k=0}^{j-1} (-1)^k \mathbf{E}_k \mathbf{H}_{j-k} + (-1)^j \mathbf{E}_j.$$

Using this relation, we calculate

$$\begin{aligned} M_{j-1}(g) - M_0(g-j) M_{j-1}(j) &= \sum_{k=0}^{j-1} (-1)^k \mathbf{E}_k \mathbf{H}_{g-k} - \mathbf{H}_{g-j} \sum_{k=0}^{j-1} (-1)^k \mathbf{E}_k \mathbf{H}_{j-k} \\ &= \sum_{k=0}^{j-1} (-1)^k \mathbf{E}_k \mathbf{H}_{g-k} + (-1)^j \mathbf{E}_j \mathbf{H}_{g-j} \\ &= M_j(g). \end{aligned}$$

□

Note that we can use Theorem 4.1.1 to simplify the expression for $\mu_{g-1}(g)$ given by the path enumeration formula. Let $g > 0$ and take $j = g - 1$ in Theorem 1.3.13 to obtain

$$\mu_{g-1}(g) = \sum_{k=0}^{g} (-1)^k \mathbf{E}_k \mathbf{H}_{g-k} - (-1)^g \mathbf{E}_g \mathbf{H}_0 = (-1)^{g+1} \mathbf{E}_g$$

using Theorem 4.1.1 and (1.3.10). This is the same expression as given in (1.3.12).

It remains to prove Theorem 4.1.1.

4.2 Sum identity

In this section we prove the required sum identity (Theorem 4.1.1) using the algebra of symmetric functions. We begin by reviewing background results from the symmetric function literature.

4.2.1 Symmetric functions

A *symmetric function* over a countable set $X = \{x_1, x_2, \dots\}$ of indeterminates is a formal power series $f(x_1, x_2, \dots)$ of bounded degree with coefficients taken from a commutative ring R that satisfies

$$f(x_1, x_2, \dots) = f(x_{\sigma(1)}, x_{\sigma(2)}, \dots)$$

for every permutation σ of $\mathbb{N} = \{1, 2, \dots\}$. We will always take $R = \mathbb{Q}$.

A symmetric function is *homogeneous* if all of its terms have equal degree. Let Λ^i be the set of all homogeneous symmetric functions of degree i, together with the additive identity 0, and let Λ be the ring of all (not necessarily homogeneous) symmetric functions.

We define the following homogeneous symmetric functions. Let

$$\begin{aligned} p_i &= \sum_{j\geq 1} x_j^i, \\ e_i &= \sum_{1\leq j_1<\cdots<j_i} x_{j_1}\cdots x_{j_i}, \\ h_i &= \sum_{1\leq j_1\leq\cdots\leq j_i} x_{j_1}\cdots x_{j_i}, \end{aligned}$$

which are each members of Λ^i, and

$$p_0 = 1, \quad e_0 = 1, \quad h_0 = 1,$$

which are each members of Λ^0.

For an integer partition $\lambda = \langle 1^{m_1}2^{m_2}\cdots\rangle$, we let

$$p_\lambda = \prod_{i\geq 1} p_i^{m_i}.$$

For $\lambda \vdash g$, we can see that p_λ belongs to Λ^g.

4.2.2 The algebra of symmetric functions

Both Λ^i and Λ are vector spaces over $\mathbb{Q}$. The set Λ is also an *algebra*: it is a vector space endowed with a bilinear product, namely the product of formal power series. Furthermore [33, pages 286–287], Λ is a *direct sum* of the Λ^i:

$$\Lambda = \bigoplus_{i\geq 0} \Lambda^i = \{f_0 + \cdots + f_n : n \geq 0,\ f_i \in \Lambda^i\}.$$

Hence Λ is a *graded* algebra, meaning that

(i) each element $f \in \Lambda$ may be written as $f = f_0 + \cdots + f_n$ for some $n \geq 0$ and $f_i \in \Lambda^i$,

(ii) given $f_i \in \Lambda^i$, $f_j \in \Lambda^j$, we have $f_i f_j \in \Lambda^{i+j}$.

Let $A = \{a_1, a_2, \ldots\}$ be a (finite or countable) subset of a graded algebra Γ. The subset A is *algebraically independent* if its elements satisfy no non-trivial polynomial equations. The subset A *generates* Γ *as an algebra* if $\Gamma = \mathbb{Q}[a_1, a_2, \ldots]$. The following result describes the algebraic setting for symmetric functions.

Proposition 4.2.1 ([33, pages 286–287 and Corollary 7.7.2]). *Let* **Par** *be the set of all integer partitions. Over* $\mathbb{Q}$*, we have that*

1. *the vector space* Λ^i *has basis* $\{p_\lambda : \lambda \vdash i\}$.
2. *the vector space* Λ *has basis* $\{p_\lambda : \lambda \in \mathbf{Par}\}$.
3. *the set* $\{p_i : i \geq 0\}$ *is algebraically independent and generates* Λ *as an algebra.*

4.2.3 Symmetric function identity

The symmetric functions e_i and h_i satisfy the following identity.

Theorem 4.2.2 (Symmetric function identity [33, page 296, equation (7.13)]). *Let* $g > 0$. *Then*

$$\sum_{k=0}^{g} (-1)^k e_k h_{g-k} = 0.$$

Proof. Following [33, page 296], we take the respective generating functions of e_i and h_i to be the formal power series

$$E(t) = \sum_{i \geq 0} e_i t^i, \tag{4.2.1}$$

$$H(t) = \sum_{i \geq 0} h_i t^i. \tag{4.2.2}$$

Since each of the terms of e_i is a product of i distinct indeterminates, we have that

$$E(t) = (1 + x_1 t)(1 + x_2 t) \cdots = \prod_{i \geq 1} (1 + x_i t).$$

Similarly, each of the terms of h_i is a product of i (not necessarily distinct) indeterminates, and so

$$H(t) = (1 + x_1 t + x_1^2 t^2 + \cdots)(1 + x_2 t + x_2^2 t^2 + \cdots) \cdots = \prod_{i \geq 1} \sum_{k \geq 0} x_i^k t^k.$$

By using $(1 - x)^{-1} = \sum_{k \geq 0} x^k$, we have

$$H(t) = \prod_{i \geq 1} (1 - x_i t)^{-1}.$$

We now find two different representations for the product $H(t)E(-t)$. Firstly, we have that

$$H(t)E(-t) = \prod_{i \geq 1} (1 - x_i t)^{-1} (1 - x_i t) = 1. \tag{4.2.3}$$

Secondly, by the series convolution [2, page 54, equation (1)] of (4.2.1) and (4.2.2), we have that

$$H(t)E(-t) = \left(\sum_{i\geq 0} h_i t^i\right)\left(\sum_{j\geq 0}(-1)^j e_j t^j\right) = \sum_{g\geq 0}\left(\sum_{k=0}^{g}(-1)^k e_k h_{g-k}\right)t^g. \tag{4.2.4}$$

Equating the coefficient of t^g in the expressions (4.2.3) and (4.2.4) gives the required result. □

4.2.4 A specialization of Λ

A *specialization* of Λ [33, Definition 7.8.1] is a map $\nu : \Lambda \to \mathbb{Q}$ satisfying

$$\nu(d_1 f_1 + d_2 f_2) = d_1\nu(f_1) + d_2\nu(f_2), \tag{4.2.5}$$

$$\nu(f_1 f_2) = \nu(f_2)\nu(f_2), \tag{4.2.6}$$

for all $d_1, d_2 \in \mathbb{Q}$ and $f_1, f_2 \in \Lambda$. Since the p_i are algebraically independent and generate Λ by Proposition 4.2.1, a specialization ν is uniquely determined by the values of $\nu(p_i)$ for all $i \geq 0$.

Recall the definition (1.3.6) of c_i. Specify the values

$$\nu(p_i) = ic_i \quad \text{for all } i \geq 0,$$

and take ν to be the specialization that results from (4.2.5) and (4.2.6). We now prove the following result on the values of $\nu(e_i)$ and $\nu(h_i)$.

Proposition 4.2.3. *Let $g \geq 0$. We have*

$$\nu(e_g) = \mathbf{E}_g,$$

$$\nu(h_g) = \mathbf{H}_g.$$

Proof. Since $e_g, h_g \in \Lambda^g$, and $\{p_\lambda : \lambda \vdash g\}$ is a basis for the vector space Λ^g, we may write e_g and h_g each as a (unique) linear combination of the p_λ. It is known [33, Proposition 7.7.6] that

$$h_g = \sum_{\lambda\vdash g} z_\lambda^{-1} p_\lambda, \tag{4.2.7}$$

$$e_g = \sum_{\lambda\vdash g}(-1)^{g-l(\lambda)} z_\lambda^{-1} p_\lambda,$$

where $z_\lambda = \prod_{i\geq 1} i^{m_i} m_i!$. We first compute $\nu(h_g)$. By (4.2.7) and the definition of p_λ, we have

$$\begin{aligned} h_g &= \sum_{\lambda \vdash g} z_\lambda^{-1} p_\lambda \\ &= \sum_{\lambda \vdash g} \prod_{i\geq 1} \frac{{p_i}^{m_i}}{i^{m_i} m_i!}. \end{aligned}$$

Applying ν and using properties (4.2.5) and (4.2.6) then gives

$$\begin{aligned} \nu(h_g) &= \sum_{\lambda \vdash g} \prod_{i\geq 1} \frac{\nu(p_i)^{m_i}}{i^{m_i} m_i!} \\ &= \sum_{\lambda \vdash g} \prod_{i\geq 1} \frac{(ic_i)^{m_i}}{i^{m_i} m_i!} \\ &= \sum_{\lambda \vdash g} \prod_{i\geq 1} \frac{{c_i}^{m_i}}{m_i!} \\ &= \sum_{\lambda \vdash g} C_\lambda \\ &= \mathbf{H}_g. \end{aligned}$$

Similarly, we compute

$$\begin{aligned} \nu(e_g) &= \sum_{\lambda \vdash g} (-1)^{g-l(\lambda)} C_\lambda \\ &= \mathbf{E}_g. \end{aligned}$$

□

4.2.5 Proof of sum identity

Apply the specialization ν of Section 4.2.4 to the symmetric function identity (Theorem 4.2.2), and use properties (4.2.5) and (4.2.6) and Proposition 4.2.3 to prove the sum identity (Theorem 4.1.1).

Chapter 5

Conclusion

We have given in (1.3.3) and Theorem 1.3.13 a closed form expression for $|N_k(g)|$, the number of simple lattice paths having exactly k lattice points lying above the linear boundary joining the startpoint $(0,0)$ to the endpoint (ga, gb). In doing so, we have proved the 2019 'Constant on blocks' conjecture [14].

We propose some open problems for future study.

1. The computation of $|N_k(g)|$ using (1.3.3) and Theorem 1.3.13 is quite involved. Is there a simpler closed form expression for $|N_k(g)|$, perhaps akin to Theorem 2.5.2?

2. Theorem 1.3.13 is proved by solving a recurrence relation, using Corollary 1.3.12 as a base case. Corollary 1.3.12 is predicated on Theorem 1.3.11, which Bizley proved using generating functions [7]. Is there a combinatorial proof of Theorem 1.3.11?

3. Our main result Theorem 1.3.5 is proved by induction on g, using an explicit bijection involving maps ϕ_g, ψ_g and statements $P(g)$, $R(g)$. Is there a shorter proof of this result, for example using different statements, or a different bijection, or an alternative approach such as generating functions?

4. Can Theorem 1.3.13 be generalized to more than two dimensions, using an appropriate definition of flaws to measure how much of a path in higher dimensions lies outside a specified region?

Bibliography

[1] J. Aebly. Démonstration du problème du scrutin par des considérations géométriques. *L'enseignement mathématique*, 23:185–186, 1923.

[2] M. Aigner. *A Course in Enumeration*, volume 238 of *Graduate Texts in Mathematics*. Springer, Berlin, 2007.

[3] D. André. Solution directe du problème résolu par M. Bertrand. *Comptes Rendus de l'Académie des Sciences, Paris*, 105:436–437, 1887.

[4] C. Banderier and P. Flajolet. Basic analytic combinatorics of directed lattice paths. In *Theoret. Comput. Sci.*, volume 281, pages 37–80. Elsevier, 2002. Selected papers in honour of Maurice Nivat.

[5] C. Banderier, M.-L. Lackner, and M. Wallner. Latticepathology and symmetric functions (extended abstract). In *31st International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms*, volume 159 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages Art. No. 2, 16. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2020.

[6] C. Banderier and M. Wallner. The kernel method for lattice paths below a line of rational slope. In *Lattice Path Combinatorics and Applications*, volume 58 of *Dev. Math.*, pages 119–154. Springer, Cham, 2019.

[7] M.T.L. Bizley. Derivation of a new formula for the number of minimal lattice paths from $(0,0)$ to (km,kn) having just t contacts with the line $my = nx$ and having no points above this line; and a proof of Grossman's formula for the number of paths which may touch but do not rise above this line. *J. Inst. Actuar.*, 80:55–62, 1954.

[8] D. Callan. Pair them up! A visual approach to the Chung-Feller theorem. *Coll. Math. J.*, 26(3):196–198, 1995.

[9] Y.-M. Chen. The Chung-Feller theorem revisited. *Discrete Math.*, 308(7):1328–1329, 2008.

[10] K.L. Chung and W. Feller. On fluctuations in coin-tossing. *Proc. Natl. Acad. Sci. U.S.A.*, 35(10):605–608, 1949.

[11] N. Dershowitz and S. Zaks. The cycle lemma and some applications. *European J. Combin.*, 11(1):35–40, 1990.

[12] A. Dvoretzky and T.S. Motzkin. A problem of arrangements. *Duke Math. J.*, 14:305–313, 1947.

[13] W. Feller. *An Introduction to Probability Theory and its Applications. Vol. I*, page 72. John Wiley & Sons, Inc., New York-London-Sydney, third edition, 1968.

[14] F. Firoozi, T. Marwendo, and A. Rattan. Lattice path enumeration with a linear boundary: a conjecture in the spirit of Chung and Feller, 2019. Unpublished manuscript.

[15] I.M. Gessel. A factorization for formal Laurent series and lattice path enumeration. *J. Combin. Theory Ser. A*, 28(3):321–337, 1980.

[16] E. Gorsky, M. Mazin, and M. Vazirani. Rational Dyck paths in the non relatively prime case. *Electron. J. Combin.*, 24(3):Paper No. 3.61, 29, 2017.

[17] I.P. Goulden and L.G. Serrano. Maintaining the spirit of the reflection principle when the boundary has arbitrary integer slope. *J. Combin. Theory Ser. A*, 104(2):317–326, 2003.

[18] H.D. Grossman. Paths in a lattice triangle. *Scr. Math.*, 16:207, 1950.

[19] V.J.W. Guo and X.-X. Wang. A Chung-Feller theorem for lattice paths with respect to cyclically shifting boundaries. *J. Algebraic Combin.*, 50(2):119–126, 2019.

[20] J. Haglund. Catalan paths and q,t-enumeration. In *Handbook of Enumerative Combinatorics*, Discrete Math. Appl. (Boca Raton), pages 679–751. CRC Press, Boca Raton, FL, 2015.

[21] K. Humphreys. A history and a survey of lattice path enumeration. *J. Statist. Plann. Inference*, 140(8):2237–2254, 2010.

[22] A Huq. *Generalized Chung-Feller Theorems for Lattice Paths.* PhD thesis, Brandeis University, 2009.

[23] J. Irving and A. Rattan. The number of lattice paths below a cyclically shifting boundary. *J. Combin. Theory Ser. A*, 116(3):499–514, 2009.

[24] S. Kaparthi and H.R. Rao. Higher-dimensional restricted lattice paths with diagonal steps. *Discrete Appl. Math.*, 31(3):279–289, 1991.

[25] J.H. Kim. Redundant generating functions in lattice path enumeration, 2012.

[26] C. Krattenthaler. Lattice path enumeration. In *Handbook of Enumerative Combinatorics*, Discrete Math. Appl. (Boca Raton), pages 589–678. CRC Press, Boca Raton, FL, 2015.

[27] J. Ma and Y.-N. Yeh. Generalizations of Chung-Feller theorems. *Bull. Inst. Math. Acad. Sin. (N.S.)*, 4(3):299–332, 2009.

[28] S. Melczer. Analytic combinatorics in several variables: Effective asymptotics and lattice path enumeration, 2017.

[29] S. Melczer and M. Mishna. Asymptotic lattice path enumeration using diagonals. In *Proceedings of the 25th International Conference on Probabilistic, Combinatorial and*

Asymptotic Methods for the Analysis of Algorithms, Discrete Math. Theor. Comput. Sci. Proc., BA, pages 313–324. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2014.

[30] T. Nakamigawa and N. Tokushige. Counting lattice paths via a new cycle lemma. *SIAM J. Discrete Math.*, 26(2):745–754, 2012.

[31] M. Renault. Four proofs of the ballot theorem. *Math. Mag.*, 80(5):345–352, 2007.

[32] Marc Renault. Lost (and found) in translation: André's actual method and its application to the generalized ballot problem. *Amer. Math. Monthly*, 115(4):358–363, 2008.

[33] R.P. Stanley. *Enumerative Combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999. With a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin.

[34] J.O. Tirrell. *Orthogonal polynomials, lattice paths, and skew Young tableaux*. ProQuest LLC, Ann Arbor, MI, 2016. Thesis (Ph.D.)–Brandeis University.

[35] E.A. Wolfhagen. An investigation of the Chung-Feller theorem, 2004.

[36] S.H.F. Yan and Y. Zhang. On lattice paths with four types of steps. *Graphs Combin.*, 31(4):1077–1084, 2015.

Appendix A

Code

We illustrate part of the program we used to count lattice paths, written in Python. This program was developed in conjunction with Takudzwa Marwendo in 2019.

For $0 \leq i \leq ga$ and $0 \leq j \leq gb$ and boundary line $ay = bx$, we store the number of partial paths to the point (i, j) having k flaws as the entry `path_num_arr[j][i][k]` of a three-dimensional array. The function `num_array` constructs `path_num_arr` and populates correct values for `path_num_arr[0][i][k]` and `path_num_arr[j][0][k]`, and assigns every other entry of the array the initial value `0`. The function `paths_num` then fills in the values of `path_num_arr[j][i][k]` for $0 < i \leq ga$ and $0 < j \leq gb$, and then returns the array `n_k_list = path_num_arr[gb][ga]`, which holds the values `n_k_list[k]` $= |N_k(g)|$ for all k satisfying $0 \leq k < g(a + b)$.

```
def num_array(ga,gb):
    path_num_arr = []
    for j in range(gb+1):
        path_num_arr.append([])
        for i in range(ga+1):
            path_num_arr[j].append([])
            for k in range(ga+gb):
                if j==0 and k==0:
                    path_num_arr[j][i].append(1)
                elif (i==0 and k==j):
                    path_num_arr[j][i].append(1)
                else:
                    path_num_arr[j][i].append(0)
    return path_num_arr

def paths_num(ga,gb):
    path_count_arr = num_array(ga,gb)
    for j in range(1,gb+1):
        for i in range(1,ga+1):
            is_flaw = 0 #0 if (i,j) not a flaw, 1 if it is
```

```
        if j*ga>gb*i:
            is_flaw = 1
        for k in range(is_flaw, ga+gb):
            path_count_arr[j][i][k]=
                path_count_arr[j-1][i][k-is_flaw]
                +path_count_arr[j][i-1][k-is_flaw]
n_k_list = path_count_arr[gb][ga]
return n_k_list
```