Stephen Melczer
smelczer@uwaterloo.ca

An Invitation to Analytic Combinatorics

From One to Several Variables

Submitted Manuscript
September 2020

Springer Nature

Dedicated to the memories of
Philippe Flajolet and Herb Wilf

Foreword

A little over 100 years ago, Hardy and Ramanujan used complex integrals to estimate the number of partitions of a large integer. This gave an inkling of a deep connection between something elementary (counting) and something deep (complex variable theory). The counting problem yields a generating function; coefficients of this generating function can be computed or approximated via Cauchy's integral formula. For many decades Hardy and Ramanujan's work, built on ideas of de Moivre, Bernoulli, Euler, and Dirichlet, stood as a lone *tour de force*. Enumeration continued to appear elementary and isolated from the rest of mathematics.

In the 1950's Hardy and Ramanujan's circle method (integrate over a circle approaching the singularities of the generating function) was systematized somewhat by Hayman. A few decades later, Flajolet and Odlyzko distilled the process of singularity analysis to a set of effective transfer theorems, allowing almost automatic derivation of asymptotics. Notably, over this span of seventy years, none of the analyses exceeded the technical difficulty of Hardy and Ramanujan's.

This book introduces a beast of quite a different color: Analytic Combinatorics in Several Variables (ACSV). Generating functions in more than one variable can be quite powerful. They capture joint distributions of combinatorial features, for example the size of the ground set of a permutation as well as its number of cycles, or the location and orientation of a domino in a random tiling, along with the size of the tiling. Transferring the ideas of the Hardy-Ramanujan analysis to the multivariate setting, however, turns out to summon mathematics from the four corners of the known mathematical world. Starting with a Cauchy integral, this time in several variables, one is led in simple cases to saddle point integration. More complicated cases require some harmonic analysis and singularity theory (inverse Fourier transforms of homogeneous rational forms). Sometimes there is a nontrivial choice of where to put the contour; this invokes algebraic topology and stratified Morse theory. Making all of the steps effective requires computational algebra and homotopy methods.

The story is beautiful but also difficult and eclectic. Melczer finds a way to tell it understandably and simply. Indeed, for Melczer, computation drives understanding. To quote from Knuth, "Science is what we understand well enough to explain to a computer." And so, drawing from the author's expertise in computer algebra, every part of the story is told from the viewpoint of effective computation.

The book contains deep theorems, yes, but it embodies much more: a tutorial in computer algebra, expertly conceived illustrations, and a very rich collection of examples. Among the examples one finds Kronecker coefficients, rational period integrals, and models from statistical physics. The author's background in lattice walk enumeration keeps the book grounded in yet another source of compelling examples. As the title implies, this book succeeds where our own arguably has not, in making this material inviting. Those taking up the invitation to ACSV will be expertly guided into beautiful new territory.

February 2020, *Robin Pemantle and Mark Wilson*

Preface

If you ask a mathematician what they love most about mathematics, certain answers invariably arise: beauty, abstraction, creativity, logical structure, connection (between disciplines and between people), elegance, applicability, and fun. This book can be viewed as a humble attempt to show that combinatorics is the branch of mathematics best situated to embody and illustrate all of these virtues. Our core subject is the large-scale behaviour of combinatorial objects, with a focus on two goals: the calculation of approximate *asymptotic* behaviour of sequences arising in combinatorial contexts, and the derivation of *limit theorems* describing how the parameters of combinatorial objects behave for random objects of large size. We begin with a simple idea, that to study a sequence we should look at its *generating function* (the formal series whose coefficients are the sequence of interest). Algebraic, differential, and functional equations satisfied by the generating function then represent data structures which encode the sequence, with more complicated sequences requiring more complicated encodings. These encodings can be used to computationally manipulate the sequence and, in many cases, determine its underlying properties.

The field of *analytic combinatorics* studies the asymptotic behaviour of univariate sequences by applying techniques from complex analysis to their generating functions. The universality of many properties of analytic functions often allows for an automatic asymptotic computation, with much of the theory now implemented in computer algebra systems. In this sense the study of analytic combinatorics for univariate sequences is somewhat classical; it is captured in glorious detail by the book *Analytic Combinatorics* of Philippe Flajolet and Robert Sedgewick. More recently, in the early twenty-first century, Robin Pemantle and Mark Wilson (later joined by Yuliy Baryshnikov) combined methods from complex analysis, singularity theory, algebraic and differential geometry, and topology to form a theory of *analytic combinatorics in several variables (ACSV)*. The textbook *Analytic Combinatorics in Several Variables*, by Pemantle and Wilson, provides a comprehensive overview of the subject, but its use of advanced constructions coming from the various mathematical disciplines conscripted into their work makes it suitable mainly for researchers with strong mathematical backgrounds across several domains.

The aim of this book is to provide a more accessible introduction to this vast and beautiful area of combinatorics. There are several (potentially overlapping) audiences: mathematicians interested in the behaviour of functions satisfying certain algebraic, differential, or functional equations; combinatorialists interested in learning the theory of analytic combinatorics and analytic combinatorics in several variables; computer scientists interested in the computational aspects of these subjects; and researchers from a variety of domains with an interest in the resulting applications. Our presentation is calibrated for an audience of math and computer science graduate students and researchers, although advanced undergraduates and those from adjacent research areas shouldn't be scared away. Previous knowledge of sequences and series, at the level of an advanced calculus course or first course in analysis, is the main prerequisite. Of all the advanced mathematical topics encountered, a familiarity and comfort with the basics of complex analysis (analytic functions, residues, the Cauchy integral formula) are the most crucial, although the necessary complex analytic background is reviewed in an appendix. Additional knowledge from commutative algebra, singularity theory, algebraic and differential geometry, and topology can help put certain results in context but are not assumed. Applications from numerous mathematical and scientific domains are given, including many illustrations of the various techniques on lattice path enumeration problems. We put a strong focus on computation, and the companion website to this text contains computer algebra code working through the examples contained here.

Because it draws from so many different areas of mathematics, analytic combinatorics in several variables has a reputation of being powerful yet impenetrable. My deepest wish is that this work illuminates the vast riches of ACSV and opens up the area to new researchers across disciplines.

Acknowledgments First and foremost, I owe a great debt to the three architects of analytic combinatorics in several variables, Robin Pemantle, Mark Wilson, and Yuliy Baryshnikov, for their support and guidance over the last several years. This book grew out of my doctoral thesis for the University of Waterloo and the École normale supérieure de Lyon, which was made incalculably better by the supervision of Bruno Salvy and George Labahn, and by the thesis reporters and members of the thesis committee: Jason Bell, Sylvie Corteel, Michael Drmota, Ira Gessel, and Éric Schost. Early versions of this manuscript were used as the basis for short lecture series at the University of Illinois Urbana-Champaign and the Research Institute for Symbolic Computation at JKU Linz, and for graduate courses at the University of Pennsylvania and the Université du Québec à Montréal, and I thank all the students from those classes for their feedback. Russell May gave invaluable comments on the text, helping to improve its presentation and fix many typographical errors. Marcus Michelen, Shaoshi Chen, Chaochao Zhu, Manuel Kauers, Alin Bostan, and Marc Mezzarobba gave feedback on various versions of the manuscript, along with my summer students Keith Ritchie and Andrew Martin. I would also like to thank Persi Diaconis, Jessica Khera, Erik Lundberg, and Armin Straub for suggesting problems which appear here. I originally became interested in analytic combinatorics as an undergraduate student, under projects (some of which inspired problems in this text) supervised by Marni Mishna, Alin Bostan, and Manuel Kauers, and I hope this text

conveys a sense of the wonder I experienced under their supervision. In addition to those mentioned above, some of the work presented here originally appeared in papers coauthored by Mireille Bousquet-Mélou, Julien Courtiel, Éric Fusy, and Kilian Raschel, all of whom taught me a great deal about combinatorics. Thanks also to Peter Paule for encouraging me to write this book and agreeing to be its mathematical editor, and Martin Peters and Leonie Kunz at Springer.

Finally, this book is dedicated to Philippe Flajolet and Herb Wilf, two inspirational mathematicians–on both personal and professional terms–whom I never had a chance to meet. I would like to thank Ruth Wilf for welcoming me to Philadelphia during my time there and sharing stories of Herb. I also thank my family (my parents, Laura, Celia, Harry, and Su) and friends who have supported me over the years.

September 2020, *Stephen Melczer*

Contents

List of Symbols

$\mathbb{C}$	Set of complex numbers
$\mathbb{C}_*$	Set of non-zero complex numbers
$\mathbb{A}$	Set of algebraic numbers (complex roots of polynomials in $\mathbb{Z}[z]$)
$\mathbb{R}$	Set of real numbers
$\mathbb{R}_*$	Set of non-zero real numbers
$\mathbb{R}_{>0}$	Set of positive real numbers
$\mathbb{Q}$	Set of rational numbers
$\mathbb{Z}$	Set of integers
$\mathbb{N}$	Set of natural numbers, including 0
$\mathbb{N}_{>0}$	Set of positive integers
$\Re(z)$	Real part of $z \in \mathbb{C}$
$(a_n)_{n=0}^{\infty}$	Sequence with terms $a_0, a_1, \ldots$
$\mathbb{K}$	An arbitrary field
$\mathbb{K}[z]$	Ring of polynomials with coefficients in $\mathbb{K}$
$\mathbb{K}(z)$	Field of rational functions over $\mathbb{K}$
$\mathbb{K}[[z]]$	Ring of formal power series over $\mathbb{K}$
$\mathbb{K}((z))$	Field of Laurent series over $\mathbb{K}$
$\mathbb{K}^{\mathrm{fra}}((z))$	Field of Puiseux series over $\mathbb{K}$
$[z^n]F(z)$	Coefficient of z^n in (power, Laurent, or Puiseux) series $F(z)$
$f_n = O(g_n)$	States existence of $M, N > 0$ such that $\lvert f_n\rvert \leq M\lvert g_n\rvert$ for all $n \geq N$
$f_n = o(g_n)$	States the limit of f_n/g_n goes to 0 as $n \to \infty$
$f_n \sim g_n$	States the limit of f_n/g_n goes to 1 as $n \to \infty$
$f_n = \tilde{O}(g_n)$	States $f_n = O(g_n \log^k \lvert g_n\rvert)$ for some $k \in \mathbb{N}$
$\Gamma(z)$	The Euler gamma function, defined in the appendix to Chapter 2
$\mathbf{z}$	Vector $(z_1, \ldots, z_d)$
$\mathbf{z}_{\hat{k}}$	Vector $(z_1, \ldots, z_{k-1}, z_{k+1}, \ldots, z_d)$ given by removing kth term of $\mathbf{z}$
$\hat{\mathbf{z}}$	Shorthand for $\mathbf{z}_{\hat{d}} = (z_1, z_2, \ldots, z_{d-1})$
$d\mathbf{z}$	Shorthand for $dz_1\, dz_2 \ldots dz_d$ in integrals
$\mathbf{z}^{\mathbf{i}}$	Monomial $z_1^{i_1} \cdots z_d^{i_d}$
$F_{z_j}(\mathbf{z})$	Partial derivative of $F(\mathbf{z})$ with respect to the variable z_j
$(\nabla F)(\mathbf{z})$	Gradient $(F_{z_1}, \ldots, F_{z_d})$

$(\nabla_{\log} F)(\mathbf{z})$	Logarithmic gradient $(z_1 F_{z_1}, \ldots, z_d F_{z_d})$
$\mathcal{V}(f_1, \ldots, f_r)$	Solutions of the system $f_1(\mathbf{z}) = \cdots = f_r(\mathbf{z})$ for $\mathbf{z} \in \mathbb{C}$
∂_j	Differential operator $\partial/\partial z_j$ or $\partial/\partial\theta_j$, depending on context
$H^{\mathrm{s}}(\mathbf{z})$	Square-free part of polynomial $H(\mathbf{z})$, product of irreducible factors
$\mathcal{V}$	Singularities of meromorphic $F(\mathbf{z})$, when F is understood
$\mathcal{V}_*$	Elements of $\mathcal{V}$ with non-zero coordinates
$\mathcal{V}_{\mathbb{R}}$	Elements of $\mathcal{V}$ with real coordinates
$\mathcal{V}_{>0}$	Elements of $\mathcal{V}$ with positive real coordinates
$\lvert\mathbf{w}\rvert$	The point $(\lvert w_1\rvert, \ldots, \lvert w_d\rvert) \in \mathbb{R}^d_{\geq 0}$ for any $\mathbf{w} \in \mathbb{C}^d$
$T(\mathbf{w})$	The polytorus $\{\mathbf{z} \in \mathbb{C}^d : \lvert z_j\rvert = \lvert w_j\rvert\}$ for any $\mathbf{w} \in \mathbb{C}^d$
$D(\mathbf{w})$	The polydisk $\{\mathbf{z} \in \mathbb{C}^d : \lvert z_j\rvert \leq \lvert w_j\rvert\}$ for any $\mathbf{w} \in \mathbb{C}^d$
$\mathbf{e}^{(j)}$	Elementary basis vector with 1 in position j and all other entries 0
$\mathcal{W}$	Singular set $\mathcal{V}(H)$ together with coordinate hyperplanes $\mathcal{V}(z_1 \cdots z_d)$
$\mathcal{M}$	Domain of analyticity $\mathbb{C}^d \setminus \mathcal{W}$ of Cauchy integral
$\mathcal{M}_{\mathbb{R}}$	Elements of $\mathcal{M}$ with real coordinates
$C_{\mathbf{x}}$	Imaginary fiber $\mathbf{x} + \mathrm{i}\mathbb{R}^d = \{\mathbf{x} + \mathrm{i}\mathbf{y} : \mathbf{y} \in \mathbb{R}^d\}$ defined by $\mathbf{x} \in \mathbb{R}^d$
$\mathrm{sgn}(z)$	The univariate sign function, equal to 0 if $z = 0$ and $z/\lvert z\rvert$ otherwise
$\mathrm{sgn}(\mathbf{z})$	The multivariate sign function, equal to $\mathrm{sgn}(z_1 \cdots z_d)$
$\mathrm{sgn}_{\boldsymbol{\sigma}}(B)$	Sign of component B with respect to $\boldsymbol{\sigma}$
$\boldsymbol{\sigma}_B$	Critical point for bounded component B (see Proposition 8.2)
$B(\boldsymbol{\sigma})$	Bounded component for critical point $\boldsymbol{\sigma}$ (see Definition 8.7)
$N(\mathbf{w})$	Normal cone corresponding to $\mathbf{w}$
$\tau_{\boldsymbol{\sigma}}$	Linking torus
$\nu_{B,\boldsymbol{\sigma}}$	Linking constant of component B and critical point $\boldsymbol{\sigma}$
$\mathcal{O}_{\mathbf{w}}$	The local ring at $\mathbf{w} \in \mathbb{C}^d$
$\mathcal{V}_S$	The flat defined by an index set S
$\mathcal{S}_S$	The stratum defined by an index set S

Chapter 1
Introduction

For it is unworthy of excellent men to lose hours like slaves in the labor of calculation which could safely be relegated to anyone else if the machine were used.
— Gottfried Wilhelm Leibniz

To count is a modern practice, the ancient method was to guess.
— Samuel Johnson

A fundamental problem in mathematics is how to efficiently encode mathematical objects and, from such encodings, determine their underlying properties. As an illustrative example, imagine encoding the elements of different algebraic structures on a computer.

- An integer can be encoded in its binary representation together with a bit indicating its sign.
- A rational number can be encoded as a pair of integers representing its numerator and denominator.
- An algebraic number over the rationals, $\alpha \in \mathbb{A}$, can be encoded by its minimal polynomial $m_\alpha(z) \in \mathbb{Z}[z]$ (defined by a finite list of integers) and an isolating region of the complex plane (defined, for instance, as a disk in the complex plane with rational radius and a centre with rational real and imaginary parts).

Although the first two representations here are somewhat explicit, the third is implicit: we can view the integer polynomial m_α as a 'data structure' storing α, from which information about α can be computed as desired. On the other hand, since the field of complex numbers $\mathbb{C}$ is uncountable its elements cannot even be listed. Thus, there is a nesting of algebraic structures of increasing complexity

$$\mathbb{Z} \subset \mathbb{Q} \subset \mathbb{A} \subset \mathbb{C}$$

going from a ring whose elements can be encoded and manipulated directly to a field where almost nothing can be decided for a general element.

In this text we examine questions of computability (the study of what can be decided) and computational complexity (the study of how fast something can be decided) in enumerative combinatorics, with a focus on the limiting behaviour of sequences and large discrete structures. Examining different encodings of increasingly complicated classes of sequences, we will see how current questions on these topics focus around sequences defined by multivariate rational functions. At the heart of this work is the new field of *analytic combinatorics in several variables* (abbreviated ACSV), and the majority of this text is dedicated to the development and exposition of this theory. In addition to providing tools for the study of univariate sequences, ACSV also allows for an analysis of multidimensional sequences encoded by series expansions of multivariate functions. This allows for striking results, such as limit theorems for various combinatorial objects.

At each stage of our study we focus on effective methods and implementable algorithms, with links to necessary software packages given at the textbook website,

https://melczer.ca/textbook/

This website also contains code working through every computational example in this textbook. The majority of the examples are worked through in the Maple [1] and Sage [51] computer algebra systems, although other software (such as MAGMA [15]) is used occasionally when specific packages are required. Of particular importance for this text is the Maple gfun package of Salvy and Zimmermann [48] and the Sage ore_algebra package of Kauers et al. [33]. Links to the most up to date versions of this software are given on the book website.

1.1 Algorithmic Combinatorics

Given a sequence $(f_n) = f_0, f_1, f_2, \ldots$, the *generating function* of (f_n) is the series

$$F(z) = \sum_{n\geq 0} f_n z^n.$$

We typically consider the case when the f_n are complex numbers, although generating functions can be defined formally over more general rings in a manner described in later chapters. Wilf [53] famously described a generating function as a “clothesline on which we hang up a sequence of numbers for display,” aptly summarizing their role as data structures. Assuming the terms of (f_n) are complex numbers which grow at most exponentially, i.e., there exists $K > 0$ such that $|f_n| < K^n$ for all $n \in \mathbb{N}$, the generating function $F(z)$ can be viewed as the power series expansion of a complex-valued function at the origin. A sequence can thus be encoded by equations satisfied by its generating function.

As an example, consider the following classes of functions which appear frequently in combinatorial contexts.

- A polynomial $F(z) \in \mathbb{Q}[z]$ can be encoded as a finite sequence of rational numbers.
- A rational function $F(z) \in \mathbb{Q}(z)$ can be encoded as a pair of polynomials in $\mathbb{Z}[z]$, corresponding to its numerator and denominator.
- Say that $F(z)$ is *algebraic* if there exists a non-zero polynomial $P(z, y) \in \mathbb{Z}[z, y]$ such that $P(z, F(z)) = 0$. The algebraic functions we consider can be encoded by such polynomials $P(z, y)$ together with initial terms of their power series at the origin to uniquely determine them among the roots of P with respect to y.

- Say that $F(z)$ is *differentially finite (D-finite)* if it satisfies a non-zero linear differential equation with coefficients in $\mathbb{Z}[z]$. A D-finite function can be encoded by an annihilating linear differential equation and a finite set of initial conditions.
- Say that $F(z)$ is *differentially algebraic (D-algebraic)* if there exists $d \in \mathbb{N}$ and a non-zero multivariate polynomial $P(z_0, z_1, \ldots, z_d)$ with coefficients in $\mathbb{Z}$ such that F and its derivatives satisfy $P(F, F', \ldots, F^{(d)}) = 0$. A D-algebraic function can be encoded by the polynomial P and a finite set of initial conditions.

Sequences coming from closely related combinatorial problems often have generating functions which share similar properties—for example, the generating function of any sequence satisfying a linear recurrence relation with constant coefficients is rational, while the generating function of any sequence satisfying a linear recurrence relation with polynomial coefficients is D-finite—and many amazing methods have been developed to go from a combinatorial enumeration problem to an encoding of the relevant generating function.

Given a generating function specified by one of the above encodings, there are many natural questions that can be asked, including

- Can a simple[1] closed form of the term f_n be determined as a function of n?
- Can the asymptotic behaviour of f_n be determined as a function of n directly from this encoding? Can it be determined from any encoding?
- How else can this generating function be encoded, and how can one convert to these other encodings? What is the "simplest" encoding possible?

The asymptotic behaviour of a sequence f_n describes how it behaves as $n \to \infty$. Frequently, the asymptotic behaviour of a sequence can be more revealing than cumbersome exact enumeration formulas and is often easier to determine.

Example 1.1 (Exact vs. Asymptotic Enumeration for Unlabelled Graphs)

Wilf [52, Ex. 3] conjectured in 1982 that the number f_n of unlabelled graphs on n nodes cannot be computed in polynomial time with respect to n. This conjecture is still open, although it has long been known[2] that $f_n \frac{n!}{2^{\binom{n}{2}}} \to 1$ as $n \to \infty$.

1.1.1 Analytic Methods for Asymptotics

Our main tools for asymptotic enumeration come from analysis, and the systematic use of analytic techniques to study the asymptotic behaviour of sequences is known

[1] Of course, the notion of a "simple" closed form expression is subjective, and thus open to interpretation. We do not touch on this delicate topic here, but refer the interested reader to Wilf [52], Stanley [49, Section 1.1], and Pak [40] for interesting meditations on the subject.

[2] Asymptotic behaviour of unlabelled graphs follows from the fact that there are $2^{\binom{n}{2}}$ labelled graphs on n nodes, almost all of which have no automorphisms; see Noy [14, Ch. 6] for details.

as the study of *analytic combinatorics*. The main results of analytic combinatorics illustrate strong links between the singularities of a generating function (roughly, points where the function 'behaves badly') and asymptotics of its coefficients, starting with the *Cauchy integral formula*. When the generating function of (f_n) represents a convergent power series in a neighbourhood of the origin, the Cauchy integral formula implies that for any $n \in \mathbb{N}$ the term f_n can be represented by a complex contour integral

$$f_n = \frac{1}{2\pi \mathrm{i}} \int_\gamma F(z) \frac{dz}{z^{n+1}},$$

where γ is a counter-clockwise circle in the complex plane sufficiently close to the origin. This equality relates the power series coefficients of F to an analytic object, and allows one to determine asymptotics of f_n by applying classical results on the asymptotics of parametrized complex integrals.

Unless the terms f_n decay too fast (faster than any exponential function), a complex-valued function defined by the series $F(z)$ will admit at least one singularity in the complex plane. In Chapter 2 we will see how knowledge of the function $F(z)$ near its singular points closest to the origin, known as *dominant singularities*, directly translates into explicit asymptotic formulas for f_n. The first major result of analytic combinatorics is that the location of the dominant singularities in the complex plane determine the *exponential growth* $\rho = \limsup_{n\to\infty} |f_n|^{1/n}$ of the coefficient sequence, the coarsest measure of its growth. After finding the dominant singularities of F, giving the exponential growth of f_n, full asymptotic behaviour can be deduced by studying F near these points. For most examples encountered in applications it is sufficient to determine the type of each dominant singularity (roughly, if it comes from a division by zero, substituting zero into an algebraic root or logarithm, etc.) together with small amount of additional analytic information which can then be substituted into known formulas.

Example 1.2 (The Catalan Generating Function)

The power series coefficients (c_n) of the *Catalan generating function*

$$C(z) = \frac{1 - \sqrt{1 - 4z}}{2z} = 1 + z + 2z^2 + 5z^3 + \cdots$$

count an incredible number of combinatorial sequences, including the number of rooted binary plane trees (trees where each node has at most two children). Because $z = 1/4$ is a singularity of $C(z)$ due to a square-root vanishing, and there are no other[3] singularities of $C(z)$, the techniques of analytic combinatorics (in particular, Proposition 2.11 in Chapter 2) immediately imply that the nth power series coefficient c_n behaves like $4^n n^{-3/2} \pi^{-1/2}$ as $n \to \infty$. Stanley [50] gives an extensive account of this sequence and 214 (!) different combinatorial objects it counts.

[3] Although there may appear to be a problem at $z = 0$, where the denominator of $C(z)$ is zero, this is not the case: the numerator of $C(z)$ also vanishes at $z = 0$ and the limit of $C(z)$ as z approaches zero exists and is finite.

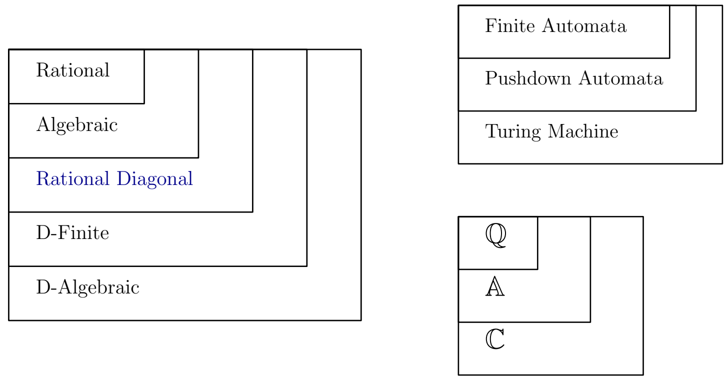


Fig. 1.1 The classes of generating functions, models of computation, and numerical constants discussed here, including the generating function class of *rational diagonals* which is central to analytic combinatorics in several variables.

Methods in analysis are flexible, giving some freedom when setting up calculations, and this flexibility often allows for an automated asymptotic analysis[4]. Automatically determining properties of a generating function from an encoding is the domain of *algorithmic combinatorics*. As we will see, asymptotic expressions for sequences with rational or algebraic generating functions can be determined automatically from the encodings discussed above. In contrast, it is undecidable to take a general D-algebraic function with rational power series coefficients, encoded as above, and determine whether or not its coefficients grow exponentially. Although much is known about the asymptotics of sequences with D-finite generating functions, the decidability of determining such asymptotics is an open problem; D-finite functions thus lie at the boundary of decidability in enumeration.

When discussing decidability results it is interesting to compare our classes of generating functions to families of objects in theoretical computer science. A finite automaton is the simplest model of computation used for pattern recognition, and counting the strings of patterns recognized by a fixed automaton by the number of symbols they contain always results in a rational generating function. Similarly, algebraic generating functions appear when enumerating patterns recognized by certain 'pushdown' automata. Because finite and pushdown automata are relatively simple models of computation, it is not surprising that many properties of rational and algebraic generating functions can be decided. On the other hand, it can be shown [30] that universal Turing machines can be simulated by solving D-algebraic systems! As Turing machines are powerful, it is thus not surprising that D-algebraic functions are hard to get a handle on. The increasingly complicated classes of

[4] We will see several examples in this text. For an early example of an automated system to determine asymptotics of sequences, implemented in CAML and Maple, see the Lambda-Upsilon-Omega (ΛΥΩ) system of Flajolet et al. [26, 27].

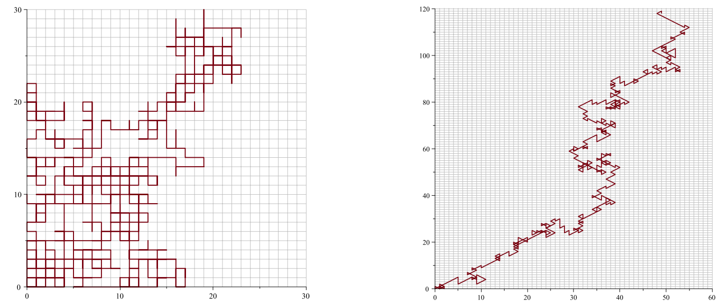


Fig. 1.2 Lattice walks of length 500 restricted to the quadrant $\mathbb{N}^2$ using the step sets $\mathcal{S} = \{(\pm 1, 0), (0, \pm 1)\}$ (left) and $\mathcal{S} = \{(0, -1), (\pm 1, 1)\}$ (right). Note the walks have self-intersections.

generating functions, models of computation, and numerical constants are displayed in Figure 1.1; these connections are further explored in Chapters 2 and 3.

1.1.2 Lattice Path Enumeration

Much of the theory of algorithmic combinatorics appears in the rich field of lattice path enumeration, which forms the main source of combinatorial applications in this textbook[5]. Roughly speaking, a lattice path model is a combinatorial class which encodes the number of ways to "move" on a lattice subject to certain constraints. More precisely, given a dimension $d \in \mathbb{N}$, a finite *set of allowable steps* $\mathcal{S} \subset \mathbb{Z}^d$, and a *restricting region* $\mathcal{R} \subseteq \mathbb{Z}^d$, the *integer lattice path model* taking steps in $\mathcal{S}$ and restricted to $\mathcal{R}$ is the combinatorial class consisting of *walks*: finite sequences of steps beginning at some fixed point of $\mathcal{R}$ which must always stay in $\mathcal{R}$ (see Figure 1.2). We may also restrict the class further by adding other constraints, such as only admitting walks which end in some terminal set.

In addition to a large number of applications, discussed in Chapter 4, lattice path problems form an interesting 'combinatorial playground' to develop new methods for dealing with a wide range of generating function behaviours. For example, the generating functions of two-dimensional models restricted to the quadrant $\mathbb{N}^2$ with step sets $\mathcal{S} \subset \{\pm 1, 0\}^2$ already exhibit a wide variety of behaviour: they admit generating functions which are rational, algebraic, transcendentally D-finite,

[5] Of particular use to us will be the extreme effectiveness of analytic combinatorics in several variables for solving lattice path enumeration problems. It is interesting to note that the development of complex analysis in a single variable was greatly inspired by the study of elliptic functions, while the theory of complex analysis in several variables suffered due to lack of concrete problems. To quote the opening sentence of Blumenthal [13] from 1903, translated from the German, "If, in contrast to the widely and highly developed theory of functions of a single complex variable, the theory of functions of several variables has lagged behind, this is due essentially attributed to the absence of suitable interesting examples which could lead the general theory." The methods of ACSV rely heavily on complex analysis in several variables.

non-D-finite but D-algebraic, and non-D-algebraic. Progress has been made on the study of these quadrant models over the last several decades using tools from the theory of algebraic curves, formal power series approaches to discrete difference equations, probability theory, computer algebra, boundary value problems, potential theory, differential Galois theory, the study of hypergeometric functions, and several branches of complex analysis. Many of the decidability issues discussed above arise naturally in this context.

Example: D-finite Decidability Problems in Lattice Path Enumeration

Consider the sequence a_n counting the number of walks restricted to $\mathbb{N}^2$ which begin at the origin and take n steps in $\mathcal{S} = \{(\pm 1, 0), (0, \pm 1)\}$, illustrated on the left of Figure 1.2. Using results from Chapters 3 and 4, it is possible to show that the generating function $A(z)$ of (a_n) satisfies the linear differential equation

$$\begin{aligned} z^2(4z-1)(4z+1)A'''(z) + 2z(4z+1)(16z-3)A''(t) \\ + 2(112z^2 + 14z - 3)A'(t) + 4(16z+3)A(t) = 0, \end{aligned} \tag{1.1}$$

a result originally conjectured by Bostan and Kauers [17] and proven by Bostan et al. [16]. Using techniques discussed in Chapter 2, it is possible to automatically translate such a differential equation to a linear recurrence relation satisfied by a_n,

$$(n+4)(n+3)a_{n+2} - 4(2n+5)a_{n+1} - 16(n+1)(n+2)a_n = 0. \tag{1.2}$$

Because this is a linear recurrence, the set of all sequences which satisfy (1.2) form a complex vector-space, and because such a sequence is uniquely determined by its first two values this vector space has dimension two. Furthermore, the methods of Chapter 2 allow us to describe a basis of this vector space consisting of two elements, $\Psi_1(n)$ and $\Psi_2(n)$, whose asymptotic expansions

$$\Psi_1(n) = 4^n\, n^{-1} \left(1 - \frac{3}{2n} + \cdots\right) \qquad \text{and} \qquad \Psi_2(n) = (-4)^n\, n^{-3} \left(1 - \frac{9}{2n} + \cdots\right)$$

as $n \to \infty$ can be computed to any fixed accuracy. Because Ψ_1 and Ψ_2 form a basis for all solutions of (1.2), there exist constants $\lambda_1, \lambda_2 \in \mathbb{C}$ such that our lattice path counting sequence, which is one particular solution of (1.2), satisfies

$$a_n = \lambda_1\, \Psi_1(n) + \lambda_2\, \Psi_2(n).$$

Although it is currently unknown how to compute the constants λ_1 and λ_2 directly from (1.1) and initial terms of $A(z)$, we will see how to use *numeric analytic continuation* to rigorously approximate them to any desired accuracy. Heuristically, one could also exploit this recurrence to efficiently compute the term a_N for a large index N, say $N = 1000$, and use this result to approximate $\lambda_1 \approx a_N / \Psi_1(N)$. In this case we can obtain a numerical approximation

$$a_n = (1.273\ldots)\Psi_1(n) + (5.092\ldots)\Psi_2(n) \sim (1.273\ldots)4^n n^{-1}$$

with rigorous error bounds on the numeric constants, giving asymptotics up to a multiplicative constant which can be determined to high accuracy. Among other things, we will see that this asymptotic growth directly implies that $A(z)$ is not algebraic and thus cannot be encoded by a polynomial equation.

In contrast, let b_n count the number of walks restricted to $\mathbb{N}^2$ which begin at the origin and take n steps in $\mathcal{S} = \{(0,-1),(-1,1),(1,1)\}$, illustrated on the right of Figure 1.2. Now the generating function $B(z)$ satisfies a linear differential equation of order 5 whose coefficients are polynomials of degree 16, from which it is possible to show that b_n satisfies a linear recurrence relation of order 6 with polynomial coefficients. This linear recurrence has a basis $\Psi_1(n),\ldots,\Psi_6(n)$ whose expansions as $n \to \infty$ can be computed to any fixed accuracy, where

$$\Psi_1(n) = \frac{3^n}{\sqrt{n}}(1 + \cdots)$$

and all other terms grow exponentially slower than 3^n as $n \to \infty$. As the Ψ_j form a basis there are $\lambda_1,\ldots,\lambda_6 \in \mathbb{C}$ such that the lattice path counting sequence b_n satisfies

$$b_n = \lambda_1\Psi_1(n) + \lambda_2\Psi_2(n) + \cdots + \lambda_6\Psi_6(n). \tag{1.3}$$

Since $\Psi_1(n)$ is the asymptotically dominant basis element it appears that b_n grows like $\lambda_1\, 3^n/\sqrt{n}$. However, perhaps surprisingly, numerically approximating λ_1 shows that it equals zero to many decimal places. In fact, we will see that the number of lattice walks on n steps has exponential growth $2\sqrt{2}$, and its dominant asymptotic behaviour depends on the parity of n.

In general, it is not known whether one can examine decompositions like (1.3) for sequences satisfying linear recurrences with polynomial coefficients and determine which coefficients are identically zero. The issue is that although one can show these constants to be zero to large accuracy, a priori there are no lower bounds on how small they can be when they are non-zero (in fact there is not even a tight characterization of the ring of numbers in which these constants lie). This issue, lying at the heart of effective asymptotic methods for sequences with D-finite generating functions, is discussed in detail in Chapter 2.

Going back to lattice paths, a full and rigorous asymptotic analysis of models restricted to the non-negative quadrant stubbornly resisted several attempts. Bostan et al. [16] gave differential equations satisfied by the generating functions of all quadrant models with step sets $\mathcal{S} \subset \{\pm 1, 0\}^2$, and expressed these generating functions in terms of explicit hypergeometric functions. Although these representations give strong information about the sequences involved, it is still difficult to prove the conjectured asymptotics. For instance, those authors show [16, Conjecture 2] that the constant λ_1 in (1.3) is exactly zero if and only if the integral

$$\int_0^{1/3} \left\{ \frac{(1-3v)^{1/2}}{v^3(1+v^2)^{1/2}} \left[1 + (1-10v^3) \times {}_2F_1\left(\begin{matrix} 3/4, 5/4 \\ 1 \end{matrix} \middle| 64v^4 \right) \right.\right.$$
$$\left.\left. + 6v^3(3-8v+14v^2) \times {}_2F_1\left(\begin{matrix} 5/4, 7/4 \\ 2 \end{matrix} \middle| 64v^4 \right) \right] - \frac{2}{v^3} + \frac{4}{v^2} \right\} dv$$

equals 1, where ${}_2F_1$ denotes the hypergeometric series

$$ {}_2F_1\left(\begin{matrix} a, b \\ c \end{matrix} \middle| z \right) = \sum_{n \geq 0} \frac{(a)_n (b)_n}{(c)_n} \frac{z^n}{n!}, \qquad (x)_n = x(x+1)\cdots(x+n-1).$$

It is not clear how to prove this identity.

Here we will follow work of Melczer and Mishna [37] and Melczer and Wilson [39] to determine asymptotics of these, and many more, lattice path models by encoding their generating functions using a new data structure: *multivariate rational diagonals*. Although a linear differential equation encodes a vector space of solutions, from which the desired generating function is specified by its initial terms, this representation directly encodes the generating function of interest. Asymptotics can then be determined using the methods of analytic combinatorics in several variables.

1.2 Diagonals and Analytic Combinatorics in Several Variables

Diagonal representations encode sequences using multivariate series expansions. Given a d-variate complex function $F(\mathbf{z}) = F(z_1, \ldots, z_d)$ with power series

$$F(\mathbf{z}) = \sum_{(i_1,\ldots,i_d)\in\mathbb{N}^d} f_{i_1,\ldots,i_d} z_1^{i_1} \cdots z_d^{i_d}$$

at the origin, we define the *main diagonal* of $F(\mathbf{z})$ as the univariate function obtained by taking the coefficients of this series where all variable exponents are equal,

$$(\Delta F)(z) = \sum_{n\geq 0} f_{n,\ldots,n} z^n.$$

More generally, one can define the $\mathbf{r}$-*diagonal* of $F(\mathbf{z})$ for any $\mathbf{r} \in \mathbb{R}^d$ as

$$(\Delta_{\mathbf{r}} F)(z) = \sum_{n\geq 0} f_{n\mathbf{r}} z^n = \sum_{n\geq 0} f_{nr_1,\ldots,nr_d} z^n,$$

where the coefficient $f_{n\mathbf{r}}$ is considered undefined when $n\mathbf{r} \notin \mathbb{N}^d$. Our methods also apply to more general series expansions whose exponents contain negative integers.

Although the theory we develop works for more general classes of functions, we focus mainly on series expansions of rational functions so that tools from computer algebra can be used to automate the necessary computations. Rational diagonals are studied extensively in Chapter 3, where we will see that each $\mathbf{r}$-diagonal is D-finite and every algebraic function can be realized as the main diagonal of a bivariate rational function. Because the ring of multivariate rational diagonals lies between the class of algebraic functions, where coefficient asymptotics can be determined automatically, and the ring of D-finite functions, where this problem is still open, they make a prime subject on which to study effective asymptotics. Chapter 3 discusses how diagonals arise naturally in combinatorics (lattice path enumeration, irrational tilings of rectangles, Kronecker coefficients), probability theory (random walk models), number theory (binomial sums such as Apéry's sequence, used in his proof [2] of the irrationality of $\zeta(3)$), physics (statistical mechanics and the Ising model), and many other areas. Just as Bernoulli [11] used linear recurrence relations to approximate roots of univariate polynomials in the early eighteenth century, asymptotic behaviour of the **r**-diagonals of a rational function $1/H(\mathbf{z})$ encodes deep information about the algebraic set defined by $H(\mathbf{z}) = 0$.

1.2.1 The Basics of Analytic Combinatorics in Several Variables

In order to study the asymptotics of rational diagonal coefficient sequences we make use of complex analysis in several variables. Suppose $F(\mathbf{z}) = G(\mathbf{z})/H(\mathbf{z})$ is a multivariate rational function, where G and H are coprime polynomials with integer coefficients. When $H(\mathbf{0})$ is non-zero, F admits a power series expansion

$$F(\mathbf{z}) = \sum_{(i_1,\dots,i_d)\in\mathbb{N}^d} f_{i_1,\dots,i_d} z_1^{i_1} \cdots z_d^{i_d} \tag{1.4}$$

in some open domain of convergence $\mathcal{D}$ around the origin. As in the univariate setting, there is a strong link between the singularities of $F(\mathbf{z})$, which are the elements of the *singular variety* $\mathcal{V} = \{\mathbf{z} \in \mathbb{C}^d : H(\mathbf{z}) = 0\}$, and asymptotics of the **r**-diagonal sequence $f_{n\mathbf{r}}$ as $n \to \infty$. A singularity is called *minimal* if no other point in $\mathcal{V}$ has coordinate-wise smaller modulus. The minimal points are the singularities which are closest to the origin in $\mathbb{C}^d$, similar to dominant singularities in the univariate case, and they play an important role in determining asymptotics.

The study of analytic combinatorics is more delicate in several variables. Although many[6] univariate functions admit a finite number of dominant singularities, in dimension two and greater there will always be an infinite number of minimal points unless $F(\mathbf{z})$ is a polynomial. For a fixed direction **r**, the goal is still to determine a finite number of singularities of $F(\mathbf{z})$ where the local behaviour of $F(\mathbf{z})$ allows one to determine asymptotics of $f_{n\mathbf{r}}$. The fact that this is not always possible is a reflection of the pathologies which can arise when dealing with the singularities of multivariate rational functions.

[6] For instance, any rational, algebraic, or D-finite function has a finite number of singularities in the complex plane.

Critical Points

Fix $\mathbf{r} \in \mathbb{N}^d$ with positive coordinates. The starting point of a multivariate asymptotic analysis is a generalization of the Cauchy integral formula to several variables,

$$f_{n\mathbf{r}} = \frac{1}{(2\pi i)^d} \int_C F(\mathbf{z}) \frac{dz_1 \cdots dz_d}{z_1^{nr_1+1} \cdots z_d^{nr_d+1}}, \tag{1.5}$$

where C is a product of circles sufficiently close to the origin. Using standard integral bounds, we will show that every minimal point $\mathbf{w} \in \mathcal{V}$ gives an upper bound

$$\rho \leq |w_1^{-r_1} \cdots w_d^{-r_d}| \tag{1.6}$$

on the exponential growth $\rho = \limsup_{n\to\infty} |f_{n\mathbf{r}}|^{1/n}$ of the diagonal sequence. To find a set of minimal points where a local singularity analysis of $F(\mathbf{z})$ determines asymptotics, it makes sense to look for the minimal points minimizing this upper bound, as those are the only ones around which the integrand of (1.5) could have the same exponential growth as the diagonal sequence.

Suppose first that the denominator H and its partial derivatives never simultaneously vanish. In Chapter 5 we show that frequently the minimal points giving the best bound in (1.6) satisfy the algebraic system of *smooth critical point equations*

$$H(\mathbf{z}) = 0, \qquad r_1^{-1} z_1 H_{z_1}(\mathbf{z}) = \cdots = r_d^{-1} z_d H_{z_d}(\mathbf{z}),$$

where H_{z_j} represents the partial derivative of H with respect to z_j. When H and its partial derivatives never simultaneously vanish, any solution of this system is called a *critical point* of $F(\mathbf{z})$.

The condition that H and its partial derivatives never simultaneously vanish implies that the singular set $\mathcal{V}$ forms a *manifold*, meaning $\mathcal{V}$ looks like $\mathbb{C}^{d-1}$ around each of its points. When the algebraic set $\mathcal{V}$ is not a manifold, one must partition $\mathcal{V}$ into a collection of manifolds called *strata* and define critical points on each stratum. In Chapter 9 we discuss how the critical points on any stratum can always be specified by a system of polynomial equations. Thus, in practice it is usually easy to characterize the critical points of $F(\mathbf{z})$, which are encoded by algebraic equations, but much more difficult to decide which (if any) are minimal, as minimality is a *semi-algebraic* condition (it relies on inequalities between moduli of variables). When there are minimal critical points where the singular variety is locally a manifold such points must minimize the upper bound (1.6) on the exponential growth ρ, but this may not hold for non-smooth minimal critical points.

Asymptotics

Asymptotic behaviour is determined by deforming the domain of integration in the Cauchy integral (1.4) to be close to the singularities of $F(\mathbf{z})$, without changing the value of the integral, then performing a local singularity analysis. Intuitively, minimal points are those to which the contour C can be easily deformed, as they are the singularities closest to the origin, while critical points are those where such a singularity analysis can be deformed to determine asymptotics. As in the univariate case, the nature of the singular variety at minimal critical points is important to the determination of asymptotics, and now the local geometry of $\mathcal{V}$ plays a large role.

The easiest case is when $\mathcal{V}$ admits a single minimal critical point $\mathbf{w}$, around which $\mathcal{V}$ is a complex manifold. Assuming some minor extra conditions typically satisfied in applications, diagonal asymptotics can be obtained by computing a univariate residue integral followed by an $d-1$ dimensional integral whose domain of integration can be made arbitrarily close to $\mathbf{w}$. When $\mathcal{V}$ has a finite number of such minimal critical points, diagonal asymptotics are determined by applying the so-called saddle-point method to each of the $d-1$ dimensional integrals. Under conditions which can be automatically verified, there are explicit formulas for diagonal asymptotics depending only on the minimal critical points involved and evaluations of the partial derivatives of $G(\mathbf{z})$ and $H(\mathbf{z})$: see, for instance, Theorem 5.1 in Chapter 5.

The next simplest case occurs when $\mathcal{V}$ admits a minimal critical point $\mathbf{w}$ around which $\mathcal{V}$ is the finite union of manifolds whose tangent planes are linearly independent; such a point $\mathbf{w}$ is called a *transverse multiple point*. Under certain conditions which often hold, and which are sufficient for our applications, diagonal asymptotics can again be computed through explicit formulas. Although we discuss how to perform a singularity analysis when $\mathcal{V}$ has a more exotic geometry, our results in these cases rely on more complicated techniques and are thus less explicit. Ultimately, the goal of this program is to automate as much as possible, perhaps proving the decidability of asymptotics for multivariate rational diagonal sequences.

Multivariate Generating Functions and Central Limit Theorems

Instead of simply viewing a multivariate function $F(\mathbf{z})$ as a data structure for a single $\mathbf{r}$-diagonal sequence, we can also view $F(\mathbf{z})$ as a truly multivariate generating function whose coefficients enumerate different parameters on combinatorial objects. In this case we want to know coefficient behaviour for *all* $\mathbf{r}$-diagonals. Our results show that for ‘most’ directions (in a manner to be made precise) asymptotics behave in a uniform manner as $\mathbf{r}$ varies, and in many situations we are even able to prove limit theorems about the coefficients of interest. For example, if for each $n \in \mathbb{N}$ we let $[z_d^n]F(\mathbf{z})$ denote the terms in the series expansion of $F(\mathbf{z})$ containing z_d^n then the results of Chapter 5 give automatically verifiable conditions under which the coefficients of $[z_d^n]F(\mathbf{z})$ approach a multivariate normal distribution as $n \to \infty$. Figure 1.3 illustrates two examples, which we return to in Chapter 5.

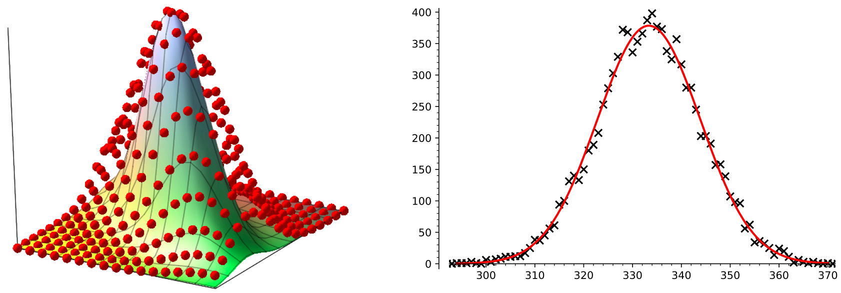


Fig. 1.3 *Left:* The coefficients of z^n in $1/(1-z-xz^2-yz^3)$ approach a bivariate normal distribution as $n \to \infty$ ($n = 200$ is pictured). These coefficients enumerate certain permutations of $\{1, \dots, n\}$ by the number of two and three cycles they contain, part of a family studied by Chung et al. [18] due to a connection to random perfect matchings in bipartite graphs. *Right:* Let $e_{k,n}$ denote the number of pairs of polynomials (r, q) over the finite field of prime order p such that q is monic, $n = \deg q > \deg r$, and Euclid's gcd algorithm applied to (r, q) takes k steps to terminate. The methods of Chapter 5 imply $e_{k,n}$ approaches a normal distribution as $n \to \infty$. Pictured is a frequency count of divisions performed when running Euclid's algorithm on 10 000 random pairs of polynomials with $\deg q = 500$ and $p = 3$, compared to the limiting distribution.

1.2.2 A History of Analytic Combinatorics in Several Variables

Using the binomial theorem, coefficient sequences of multivariate rational functions can often be represented as sums of non-negative terms which are amenable to classical asymptotic techniques covered, for instance, in de Bruijn [19]. Early work on truly multivariate analyses of generating functions came from a probabilistic examination of their coefficients. Starting in the 1970s, Bender [7] derived local and central limit theorems for coefficients in the bivariate case, which Bender and Richmond [8] generalized to any number of variables, followed by related work by those authors and collaborators [10, 29, 9]. Additional work in the 1990s further examined limit theorems for coefficients of bivariate generating functions [28, 25, 21, 22, 23, 31, 32]. Near the end of this probabilistic work, a new approach arose which used multivariate complex residues to determine coefficient asymptotics for functions whose denominators are products of linear factors [35, 12] and for bivariate functions [36]. Although the theory developed in these early approaches was applied in a somewhat ad hoc manner, much of the modern approach was present; for example, the queuing theory work of Kogan and Yakovlev [34] derived asymptotics using residue computations to reduce questions of asymptotic behaviour to an analysis of saddle-point integrals. Egorychev's book [24] on combinatorial identities contains a small number of results using multivariate integrals.

More recently, Pemantle, Wilson, Baryshnikov, and collaborators, have brought together results from several different mathematical disciplines in order to develop a large-scale systematic theory of multivariate asymptotics for combinatorial purposes. Perhaps the first work in this program is Pemantle [41], which studies coefficient sequences of multivariate rational functions whose asymptotic expansions terminate (and thus uniquely determine the sequences of interest). In the early 2000s, Pemantle and Wilson [42] gave a method for determining asymptotics when the singular set $\mathcal{V}$ is a complex manifold and there are minimal critical points. Soon after, those authors extended their results to cover a larger set of singular behaviour [43]. Raichev and Wilson [46, 47] gave a method to determine higher-order asymptotic terms.

These early results follow the 'surgery' approach to ACSV, meaning they use simple and explicit deformations of the multivariate Cauchy residue integral. Later work of Baryshnikov and Pemantle [6] used more complicated deformations of the domain of integration in the multivariate Cauchy integral to extend these results. This approach, extended in the textbook of Pemantle and Wilson [44] and recent papers of Baryshnikov, Melczer, and Pemantle [5, 4, 3], shows how the methods of ACSV fit into the very general framework of stratified Morse theory, allowing the use of advanced homological methods.

Our presentation of analytic combinatorics in several variables focuses more on explicit calculations, including the derivation of effective algorithms working in a large variety of situations. DeVries et al. [20] gave a general algorithm for bivariate asymptotics when $\mathcal{V}$ forms a smooth manifold and Raichev [45] developed a Sage package which helps with some of the computations necessary in an asymptotic analysis. More recently, Melczer and Salvy [38] gave algorithms and complexity bounds for the analysis of multivariate generating functions in any number of variables, with an accompanying Maple implementation. Links to this software can be found on the website for this textbook.

1.3 Organization

This textbook is split into three parts.

Part I

Part I covers the background and motivation for analytic combinatorics in several variables, including a thorough treatment of univariate techniques. We take a computational viewpoint, and our presentation of univariate methods is structured to help guide intuition in the more advanced multivariate setting. Part I begins in Chapter 2, which describes in detail the underpinnings of univariate analytic combinatorics. After some background on generating functions and combinatorial classes, we introduce the main tools of analytic combinatorics through an extended example looking at alternating permutations. This serves as a lead-in to the general theory. The analytic, algebraic, arithmetic, and asymptotic properties of rational, algebraic, D-finite, and D-algebraic power series are also detailed.

Chapter 3 serves as an introduction to the theory of complex analysis in several variables, and the use of multivariate generating functions. After providing the tools necessary for the multivariate singularity analysis which lies at the heart of analytic combinatorics in several variables, we give a full treatment of rational diagonals and their connections to other classes of generating functions. Algebraic tools for the analytic study of multivariate rational functions are described, such as the connection between so-called amoebas and convergent Laurent expansions. Crucially, the singular set of a multivariate rational function encodes not only information about the asymptotic behaviour of the main diagonal sequence of a rational function's power series expansion, but also information about all **r**-diagonals of the power series expansion *and* information about diagonals of other convergent Laurent expansions of the rational function. Thus, a good understanding of these topics helps illustrate why certain behaviours occur when applying the methods of analytic combinatorics in several variables. We show that the different convergent Laurent expansions of a rational function are strongly linked, and study different operators on multivariate expansions which will be useful for our combinatorial applications. Chapter 3 ends by describing several domains of mathematics and the sciences where rational diagonals arise. In addition to showing the importance of rational diagonals, the examples discussed here are used to illustrate the methods of ACSV in later chapters.

Part I ends in Chapter 4, which contains a presentation of the kernel method approach to lattice path enumeration. Beginning with the easy case of unrestricted lattice path models, the mechanics of the kernel method are built up for one-dimensional walks restricted to a half-space and two-dimensional walks restricted to a quadrant. After describing this incredibly effective machinery, which naturally results in rational diagonal expressions for the generating functions involved, we survey results in lattice path enumeration using this approach. We will return to problems in lattice path enumeration multiple times in the text, using the increasingly sophisticated analytic techniques we develop to continually derive more general results.

Part II

Part II covers the theory and applications of ACSV when the singular set of the function under consideration forms a manifold, known as the *smooth case*. It begins in Chapter 5, which describes the basics of analytic combinatorics in several variables and shows how to obtain asymptotics for many rational functions which admit singular varieties that are complex manifolds. After an extended example, which concretely illustrates the methods of ACSV in the smooth case from start to finish, the general theory is developed. Many examples are given and general strategies for applying the tools of analytic combinatorics in several variables are demonstrated.

Chapter 6 describes how to use the methods of ACSV to enumerate lattice path models in a general orthant $\mathbb{N}^d$ when the set of allowable steps is symmetric over every axis. This analysis uses the results of Chapters 4 and 5 and provides the first illustration of a principle that will arise multiple times in the text: structures which exhibit nice combinatorial properties, such as a large number of symmetries, can often be encoded by generating functions which have nice analytic properties. The enumerative results derived are quite general, and illustrate the power of analytic combinatorics in several variables.

Finally, Part II ends in Chapter 7 with a study of effective algorithms for analytic combinatorics in the smooth case. After developing some necessary background on symbolic-numeric computation and methods for encoding solutions of multivariate algebraic systems, we develop rigorous algorithms and complexity results for rational diagonal asymptotics in any dimension, under assumptions which are often satisfied in applications. A large number of examples are worked through using a Maple package of Melczer and Salvy, which implements the theory presented here.

Part III

Part III generalizes the methods of Part II to allow for the analysis of more general rational functions. First, Chapter 8 gives a thorough treatment of rational functions whose denominators are products of real linear factors. Although the singular varieties under consideration may not be manifolds, they will be unions of hyperplanes, allowing for a deep analysis to be performed.

Chapter 9 describes a general theory of analytic combinatorics in several variables when the singular variety is no longer a manifold. We extend the analysis of Chapter 8 to get explicit asymptotics for rational functions whose singular variety looks *locally* like the union of hyperplanes near minimal critical points. We also detail a new approach inspired by Morse theory which gives powerful structure theorems for asymptotics, and provide algebraic criteria that certify when such results hold. This approach has strong implications for the development of effective algorithms and yields one of the most promising techniques for resolving computability questions about sequences with D-finite generating functions.

Chapter 10 uses these new analytic methods to derive deeper and more general results on lattice path enumeration, resulting in enumerative results for lattice paths in the orthant $\mathbb{N}^d$ whose step sets are symmetric over all but one axis. A selection of other lattice path problems are surveyed, giving a large selection of exercises for the reader.

References

1. Maple 2018. *Maplesoft, a division of Waterloo Maple Inc., Waterloo, Ontario*.
2. Roger Apéry. Irrationalité de $\zeta(2)$ et $\zeta(3)$. In *Journées Arith. de Luminy. Colloque International du Centre National de la Recherche Scientifique (CNRS), Centre Universitaire de Luminy, Jun 20-24. Astérisque*, volume 61, pages 11–13, 1979.
3. Y. Baryshnikov, S. Melczer, and R. Pemantle. Asymptotics of multivariate sequences in the presence of a lacuna. *arXiv preprint, arXiv:1905.04187*, 2019. Submitted.
4. Y. Baryshnikov, S. Melczer, and R. Pemantle. Asymptotics of multivariate sequences IV: generating functions with poles on a hyperplane arrangement. In preparation, 2020.
5. Y. Baryshnikov, S. Melczer, and R. Pemantle. Stationary points at infinity for analytic combinatorics. *arXiv preprint, arXiv:1905.05250*, 2020. Submitted.
6. Yuliy Baryshnikov and Robin Pemantle. Asymptotics of multivariate sequences, part III: Quadratic points. *Adv. Math.*, 228(6):3127–3206, 2011.

7. Edward A. Bender. Central and local limit theorems applied to asymptotic enumeration. *J. Combinatorial Theory Ser. A*, 15:91–111, 1973.
8. Edward A. Bender and L. Bruce Richmond. Central and local limit theorems applied to asymptotic enumeration. II. Multivariate generating functions. *J. Combin. Theory Ser. A*, 34(3):255–265, 1983.
9. Edward A. Bender and L. Bruce Richmond. Multivariate asymptotics for products of large powers with applications to Lagrange inversion. *Electron. J. Combin.*, 6:Research Paper 8, 21 pp. (electronic), 1999.
10. Edward A. Bender, L. Bruce Richmond, and S. G. Williamson. Central and local limit theorems applied to asymptotic enumeration. III. Matrix recursions. *J. Combin. Theory Ser. A*, 35(3):263–278, 1983.
11. Daniel Bernoulli. Observationes de seriebus quae formantur ex additione vel subtractione quacunque terminorum se mutuo consequentium. *Commentarii Academiae Scientiarum Imperialis Petropolitanae*, 3:85–100, 1728.
12. Andrea Bertozzi and James McKenna. Multidimensional residues, generating functions, and their application to queueing networks. *SIAM Rev.*, 35(2):239–268, 1993.
13. Otto Blumenthal. Über modulfunktionen von mehreren veränderlichen. (erste hälfte). *Mathematische Annalen*, 56:509–548, 1903.
14. Miklós Bóna, editor. *Handbook of enumerative combinatorics*. Discrete Mathematics and its Applications (Boca Raton). CRC Press, Boca Raton, FL, 2015.
15. Wieb Bosma, John Cannon, and Catherine Playoust. The Magma algebra system. I. The user language. *J. Symbolic Comput.*, 24(3-4):235–265, 1997. Computational algebra and number theory (London, 1993).
16. Alin Bostan, Frédéric Chyzak, Mark van Hoeij, Manuel Kauers, and Lucien Pech. Hypergeometric expressions for generating functions of walks with small steps in the quarter plane. *European J. Combin.*, 61:242–275, 2017.
17. Alin Bostan and Manuel Kauers. Automatic classification of restricted lattice walks. In *21st International Conference on Formal Power Series and Algebraic Combinatorics (FPSAC 2009)*, Discrete Math. Theor. Comput. Sci. Proc., AK, pages 201–215. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2009.
18. Fan Chung, Persi Diaconis, and Ron Graham. Permanental generating functions and sequential importance sampling. *Advances in Applied Mathematics*, page 101916, 2019.
19. N. G. de Bruijn. *Asymptotic methods in analysis*. Bibliotheca Mathematica. Vol. 4. North-Holland Publishing Co., Amsterdam; P. Noordhoff Ltd., Groningen; Interscience Publishers Inc., New York, 1958.
20. Timothy DeVries, Joris van der Hoeven, and Robin Pemantle. Automatic asymptotics for coefficients of smooth, bivariate rational functions. *Online J. Anal. Comb.*, (6):24, 2011.
21. Michael Drmota. Asymptotic distributions and a multivariate Darboux method in enumeration problems. *J. Combin. Theory Ser. A*, 67(2):169–184, 1994.
22. Michael Drmota. A bivariate asymptotic expansion of coefficients of powers of generating functions. *European J. Combin.*, 15(2):139–152, 1994.
23. Michael Drmota. Systems of functional equations. *Random Structures Algorithms*, 10(1-2):103–124, 1997.
24. G. P. Egorychev. *Integral representation and the computation of combinatorial sums*, volume 59 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1984. Translated from the Russian by H. H. McFadden, Translation edited by Lev J. Leifman.
25. P. Flajolet and T. Lafforgue. Search costs in quadtrees and singularity perturbation asymptotics. *Discrete Comput. Geom.*, 12(2):151–175, 1994.
26. Philippe Flajolet, Bruno Salvy, and Paul Zimmermann. Lambda-Upsilon-Omega the 1989 cookbook. Research Report RR-1073, INRIA, 1989.
27. Philippe Flajolet, Bruno Salvy, and Paul Zimmermann. Automatic average-case analysis of algorithms. *Theoret. Comput. Sci.*, 79(1, (Part A)):37–109, 1991. Algebraic and computing treatment of noncommutative power series (Lille, 1988).

28. Philippe Flajolet and Michèle Soria. General combinatorial schemas: Gaussian limit distributions and exponential tails. *Discrete Math.*, 114(1-3):159–180, 1993.
29. Zhicheng Gao and L. Bruce Richmond. Central and local limit theorems applied to asymptotic enumeration. IV. Multivariate generating functions. *J. Comput. Appl. Math.*, 41(1-2):177–186, 1992. Asymptotic methods in analysis and combinatorics.
30. Daniel S. Graça, Manuel L. Campagnolo, and Jorge Buescu. Computability with polynomial differential equations. *Adv. in Appl. Math.*, 40(3):330–349, 2008.
31. Hsien-Kuei Hwang. Large deviations for combinatorial distributions. I. Central limit theorems. *Ann. Appl. Probab.*, 6(1):297–319, 1996.
32. Hsien-Kuei Hwang. Large deviations of combinatorial distributions. II. Local limit theorems. *Ann. Appl. Probab.*, 8(1):163–181, 1998.
33. Manuel Kauers, Maximilian Jaroschek, and Fredrik Johansson. Ore polynomials in Sage. In *Computer algebra and polynomials*, volume 8942 of *Lecture Notes in Comput. Sci.*, pages 105–125. Springer, Cham, 2015.
34. Yaakov Kogan and Andrei Yakovlev. Asymptotic analysis for closed multichain queueing networks with bottlenecks. *Queueing Systems Theory Appl.*, 23(1–4):235–258, 1996.
35. Ben Lichtin. The asymptotics of a lattice point problem associated to a finite number of polynomials. I. *Duke Math. J.*, 63(1):139–192, 1991.
36. Ben Lichtin. The asymptotics of a lattice point problem associated to a finite number of polynomials. II. *Duke Math. J.*, 77(3):699–751, 1995.
37. Stephen Melczer and Marni Mishna. Asymptotic lattice path enumeration using diagonals. *Algorithmica*, 75(4):782–811, 2016.
38. Stephen Melczer and Bruno Salvy. Symbolic-numeric tools for analytic combinatorics in several variables. In *Proceedings of the ACM on International Symposium on Symbolic and Algebraic Computation*, ISSAC '16, pages 333–340, New York, NY, USA, 2016. ACM.
39. Stephen Melczer and Mark C. Wilson. Higher Dimensional Lattice Walks: Connecting Combinatorial and Analytic Behavior. *SIAM J. Discrete Math.*, 33(4):2140–2174, 2019.
40. Igor Pak. Complexity problems in enumerative combinatorics. In *Proc. Int. Cong. of Math.*, volume 3, pages 3139–3166, Rio de Janeiro, 2018.
41. Robin Pemantle. Generating functions with high-order poles are nearly polynomial. In *Mathematics and computer science (Versailles, 2000)*, Trends Math., pages 305–321. Birkhäuser, Basel, 2000.
42. Robin Pemantle and Mark C. Wilson. Asymptotics of multivariate sequences. I. Smooth points of the singular variety. *J. Combin. Theory Ser. A*, 97(1):129–161, 2002.
43. Robin Pemantle and Mark C. Wilson. Asymptotics of multivariate sequences. II. Multiple points of the singular variety. *Combin. Probab. Comput.*, 13(4-5):735–761, 2004.
44. Robin Pemantle and Mark C. Wilson. *Analytic combinatorics in several variables*, volume 140 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2013.
45. Alexander Raichev. New software for computing asymptotics of multivariate generating functions. *ACM Commun. Comput. Algebra*, 45(3-4):183–185, 2011.
46. Alexander Raichev and Mark C. Wilson. Asymptotics of coefficients of multivariate generating functions: improvements for smooth points. *Electron. J. Combin.*, 15(1):Research Paper 89, 17, 2008.
47. Alexander Raichev and Mark C. Wilson. Asymptotics of coefficients of multivariate generating functions: improvements for multiple points. *Online J. Anal. Comb.*, (6):21, 2011.
48. B. Salvy and P. Zimmermann. GFUN: a Maple package for the manipulation of generating and holonomic functions in one variable. *ACM Trans. Math. Softw.*, 20(2):163–177, 1994.
49. Richard P. Stanley. *Enumerative combinatorics. Volume 1*, volume 49 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2012.
50. Richard P. Stanley. *Catalan numbers*. Cambridge University Press, New York, 2015.
51. The Sage Developers. *SageMath, the Sage Mathematics Software System (Version 8.1)*, 2019. `https://www.sagemath.org`.
52. Herbert S. Wilf. What is an answer? *Amer. Math. Monthly*, 89(5):289–292, 1982.
53. Herbert S. Wilf. *generatingfunctionology*. A K Peters, Ltd., Wellesley, MA, third edition, 2006.

Part I
Background and Motivation

Chapter 2
Generating Functions and Analytic Combinatorics

Since there is a great conformity between the Operations in Species, and the same Operations in common Numbers . . . I cannot but wonder that no body has thought of accommodating the lately discover'd Doctrine of Decimal Fractions in like manner to Species . . . especially since it might have open'd a way to more abstruse Discoveries.
— Issac Newton

This book is concerned with the study of sequences and their generating functions. We work over a field $\mathbb{K}$, typically the field of rational numbers $\mathbb{Q}$ or the field of complex numbers $\mathbb{C}$. The basic properties of rings and fields which we use can be found in any standard reference on abstract algebra, such as Dummit and Foote [38], although in most cases one can simply view an abstract ring as the ring of integers $\mathbb{Z}$ and an abstract field as $\mathbb{Q}$ or $\mathbb{C}$.

Definition 2.1 (sequences and formal power series) A *(univariate) sequence* $(a_n) = (a_0, a_1, a_2, \dots)$ over a field $\mathbb{K}$ is a mapping from the natural numbers to $\mathbb{K}$. A *formal power series* over $\mathbb{K}$ in the indeterminate z is any formal expression

$$A(z) = \sum_{n=0}^{\infty} a_n z^n = a_0 + a_1 z + a_2 z^2 + \cdots$$

with *coefficient sequence* (a_n) in $\mathbb{K}$. We write $[z^n]A(z)$ for the term a_n in $A(z)$. If $\mathbb{A}$ is an integral domain (a commutative ring with no zero divisors) with field of fractions $\mathbb{K}$, we write $\mathbb{A}[[z]]$ to denote the elements of $\mathbb{K}[[z]]$ with coefficients in $\mathbb{A}$.

Given two formal power series $A(z) = \sum_{n=0}^{\infty} a_n z^n$ and $B(z) = \sum_{n=0}^{\infty} b_n z^n$ over the same field, we may *define* addition of $A(z)$ and $B(z)$ term-wise as

$$A(z) + B(z) = \sum_{n=0}^{\infty} a_n z^n + \sum_{n=0}^{\infty} b_n z^n = \sum_{n=0}^{\infty} (a_n + b_n)\, z^n,$$

and multiplication of $A(z)$ and $B(z)$ by the *Cauchy product*

$$A(z)B(z) = \left(\sum_{n=0}^{\infty} a_n z^n\right)\left(\sum_{n=0}^{\infty} b_n z^n\right) = \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} a_k b_{n-k}\right) z^n.$$

The word 'formal' in the term formal power series refers to the fact that we do not consider whether a formal series converges, or define what it means to substitute an element of $\mathbb{K}$ into z.

Lemma 2.1 *The collection of formal power series over a field $\mathbb{K}$ in the indeterminate z forms an integral domain under term-wise addition and the Cauchy product, called the* ring of formal power series over $\mathbb{K}$ *and denoted* $\mathbb{K}[[z]]$.

Problem 2.1 asks you to prove Lemma 2.1, and Problems 2.2 to 2.5 ask you to derive basic properties of formal power series: they form a complete metric space, any power series with non-zero constant can be inverted, and if $F(z)$ and $G(z)$ are formal power series such that $G(z)$ has no constant term then the composition $F(G(z))$ is well defined. Because we deal almost exclusively with power series over fields $\mathbb{K} \subset \mathbb{C}$ which define analytic functions, we do not go into much detail on the formal theory. Additional results about $\mathbb{K}[[z]]$ can be found in Stanley [112] or Henrici [65]. When a formal power series over $\mathbb{K} \subset \mathbb{C}$ converges for complex z in a disk around the origin, our formal constructions are consistent with the classical definitions for complex functions.

Example 2.1 (Formal vs. Analytic Power Series)

Let

$$F(z) = \sum_{n\geq 0} z^n \qquad \text{and} \qquad G(z) = \sum_{n\geq 0} n!z^n$$

be formal power series in $\mathbb{Q}[[z]]$. Then

$$(1-z)F(z) = 1 + \sum_{n\geq 1}(1-1)z^n = 1,$$

so that $F(z) = (1-z)^{-1}$ is the multiplicative inverse of $1-z$ in $\mathbb{Q}[[z]]$. Note that

$$\sum_{n\geq 0} z^n = \frac{1}{1-z}$$

as complex-valued functions when $|z| < 1$, so we can apply results from analysis to study the series $F(z)$. Because $n!$ grows faster than z^n for any fixed z the series $G(z)$ does not converge for any $z \neq 0$, but we can still work formally with $G(z)$.

We use power series because their coefficients can encode interesting information, such as the number of objects in a combinatorial structure.

Definition 2.2 (combinatorial classes and generating functions) A *combinatorial class* is a countable collection of objects $\mathcal{C}$ together with a *size function* $|\cdot|: \mathcal{C} \to \mathbb{N}$ such that the inverse image of any natural number under $|\cdot|$ is finite (there are a finite number of objects of any fixed size). The *counting sequence* (c_n) of $\mathcal{C}$ is the sequence $c_n = \#\{x \in \mathcal{C} : |x| = n\}$ counting the number of objects in $\mathcal{C}$ of each size. The *generating function* of $\mathcal{C}$ is the formal power series $C(z) = \sum_{n\geq 0} c_n z^n$, which can be viewed as an element of $\mathbb{Q}[[z]]$ with natural number coefficients. More generally, given a sequence (a_n) over any field $\mathbb{K}$ the *generating function* of (a_n) is the formal power series $A(z) \in \mathbb{K}[[z]]$ whose coefficient sequence is (a_n).

Example 2.2 (Compositions and Partitions)

The class of *integer compositions* consists of the set of all positive integer tuples,

$$\mathcal{C} = \left\{(k_1, \ldots, k_r) : r \in \mathbb{N}_{>0},\ k_j \in \mathbb{N}_{>0}\right\},$$

and the size function that maps a tuple to its sum: $|(k_1, \ldots, k_r)| = k_1 + \cdots + k_r$. The number c_n of objects of size n in $\mathcal{C}$ is the number of ways of writing n as a sum of positive integers. For instance, there are 4 compositions of size 3 since

$$3 = 1 + 1 + 1 = 1 + 2 = 2 + 1.$$

To determine the number of integer compositions of size n, consider the expression

$$1 \odot 1 \odot \cdots \odot 1$$

consisting of n ones and $n - 1$ symbols $\odot$. Replacing each $\odot$ with a plus '+' or a comma ',' gives a tuple of elements adding up to n, and conversely every tuple of size n is obtained in this manner. Thus, $c_n = 2^{n-1}$ and

$$C(z) = \sum_{n\geq 1} 2^{n-1} z^n = z \sum_{n\geq 0} 2^n z^n = \frac{z}{1-2z},$$

either formally or as a convergent power series when $|z| < 1/2$. Similarly, let

$$\mathcal{P} = \left\{(k_1, \ldots, k_r) : r \in \mathbb{N}_{>0},\ k_j \in \mathbb{N}_{>0},\ k_1 \geq k_2 \geq \cdots \geq k_r\right\}$$

be the class of *integer partitions*, containing all non-increasing positive integer tuples where the size of a tuple is again its sum. Although similar to integer compositions, integer partitions are much harder to analyze. In 1748, Euler [40] gave the generating function expression

$$P(z) = \sum_{n\geq 0} p_n z^n = \prod_{n\geq 1} \frac{1}{1-z^n},$$

which follows from the definition of power series multiplication after expanding each term as a geometric series. Using analytic methods, Hardy and Ramanujan [63] famously[1] determined asymptotic behaviour of p_n from a study of $P(z)$.

The theory of generating functions, where a series is studied to understand its coefficients, can be traced back to de Moivre [32], who found correspondences between various properties of a power series and its coefficient sequence; Euler [39], who gave the first explicit definition of formal series and wrestled with their implications; Laplace [75], who created a formal calculus of power series; and Cauchy [20], whose

[1] In one of the rare instances of pure mathematics breaking through to Hollywood, the 2015 film 'The Man Who Knew Infinity' includes a dramatic scene where Ramanujan reveals an approximation of p_{200} to Major MacMahon, who had recently calculated p_{200} exactly by hand after much effort.

work finding analytic expressions for power series coefficients underlies most of analytic combinatorics. The work of many major figures in analysis in the eighteenth and early nineteenth centuries had an impact; Ferraro [44] gives a detailed account of the theory of series during this time period.

Although there are many large and beautiful frameworks which allow one to go from a combinatorial description of a class to a specification of its generating function, we do not discuss these in detail here. Instead, we typically assume that we already have access to a generating function specified in one of the manners discussed in the rest of this chapter. Those wanting to learn more about how to derive generating function expressions can refer to the Symbolic Method of Flajolet and Sedgewick [50] and the Theory of Species described by Bergeron et al. [6]. A nice introduction to such 'generatingfunctionology' can be found in Wilf [125].

2.1 Analytic Combinatorics in One Variable

Unless explicitly stated, we now work only with power series over a field $\mathbb{K} \subset \mathbb{C}$ so that we can apply methods from complex analysis. The appendix to this chapter recalls some basic facts and notation from complex analysis which we make use of. Additional background can be found in any standard reference for complex analysis, such as Henrici [65]. The key idea of analytic combinatorics is to use an analytic function to encode the sequence formed by its power series coefficients.

Definition 2.3 (power series coefficients) A complex-valued function $F(z)$ is *analytic* at $a \in \mathbb{C}$ if $F(z)$ is represented by a convergent power series on some open disk centred at a. Given a function $F(z)$ which is analytic at the origin, we call the coefficients (f_n) of a power series representation

$$F(z) = \sum_{n \geq 0} f_n z^n,$$

valid in a neighbourhood of the origin, the *power series coefficients of F*.

From now on we slightly abuse notation and write $F(z)$ to refer both to the original function and the power series. There is no harm in making this identification, since the original function and the series both define analytic functions which agree on an open disk containing the origin and the series representation is unique. In this way, we also refer to the complex-valued function $F(z)$ as a generating function of (f_n).

We first recall a few basic facts from complex analysis, which follow from the background results discussed in the appendix. Definitions for the complex analytic terms used here can also be found in the appendix.

Deforming Curves of Integration: If C and C' are simple closed curves such that C can be continuously deformed to C' while staying in an open set where $f(z)$ is analytic, then

$$\int_C f(z)dz = \int_{C'} f(z)dz.$$

Cauchy Residue Theorem: Suppose $f(z)$ is meromorphic in a domain Ω and γ is a positively oriented loop in Ω on which f is analytic. Then

$$\frac{1}{2\pi \mathrm{i}} \int_\gamma f(z)dz = \sum_{p\in\Psi} \operatorname*{Res}_{z=p} f(z),$$

where Ψ is the (necessarily finite) set of poles of f inside γ.

Maximum Modulus Integral Bound: If $f(z)$ is continuous on a curve γ, then

$$\left|\int_\gamma f(z)dz\right| \leq \text{length}(\gamma) \times \max_{z\in\gamma} |f(z)|.$$

The methods of analytic combinatorics work because of deep connections between analytic properties of a function and asymptotic properties of its power series coefficients. The basis for this is the Cauchy integral formula, which gives an analytic expression for power series coefficients.

Theorem 2.1 (Cauchy Integral Formula for Coefficients) *Let γ be a positively oriented circle around the origin and $F(z)$ be a complex-valued function analytic inside and on γ. If $F(z)$ is represented in a neighbourhood of the origin by the convergent power series $F(z) = \sum_{n\geq 0} f_n z^n$ then, for any $n \in \mathbb{N}$,*

$$f_n = \frac{1}{2\pi \mathrm{i}} \int_\gamma F(z) \frac{dz}{z^{n+1}}. \tag{2.1}$$

In order to talk about asymptotics we introduce some notation.

Definition 2.4 (asymptotic notation) If f_n and g_n are sequences, and $g_n \neq 0$ for n sufficiently large, then we write

$f_n = O(g_n)$	if there exist $M, N > 0$ such that $\lvert f_n\rvert \leq M\lvert g_n\rvert$ for all $n \geq N$,
$f_n = o(g_n)$	if the limit of f_n/g_n goes to 0 as $n \to \infty$,
$f_n \sim g_n$	if the limit of f_n/g_n goes to 1 as $n \to \infty$,
$f_n = \tilde{O}(g_n)$	if $f_n = O(g_n \log^k \lvert g_n\rvert)$ for some $k \in \mathbb{N}$.

We now work through an example in detail, showing the ease with which complex analysis allows one to derive powerful results, before discussing the general theory.

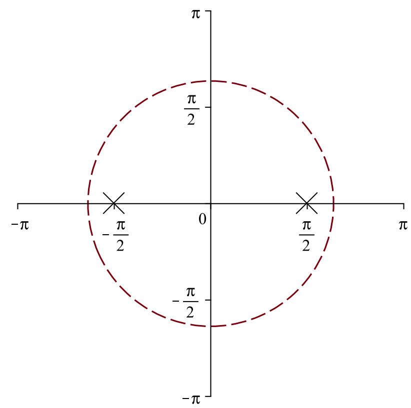


Fig. 2.1 The integral I_n, over the circle $|z| = 2$, grows exponentially slower than the coefficients of tangent. Introducing this integral allows us to determine asymptotics by computing residues at the singularities $z = \pm\pi/2$ of tangent.

2.1.1 A Worked Example: Alternating Permutations

An *alternating permutation*, also known as a zigzag permutation, of size $n = 2m + 1$ is an ordering $(\pi_1, \ldots, \pi_{2m+1})$ of the numbers $1, 2, \ldots, 2m + 1$ such that consecutive numbers alternate between increasing and decreasing:

$$\pi_1 < \pi_2, \quad \pi_2 > \pi_3, \quad \pi_3 < \pi_4, \quad \text{etc.}$$

Let a_n denote the number of alternating permutations of size n, where $a_n = 0$ when n is even. In 1879, André [1] discovered the marvelous fact that if (t_n) denotes the power series coefficients of $\tan z$ then $a_n = n!\, t_n$, meaning

$$\tan z = \sum_{n \geq 0} t_n z^n = \sum_{n \geq 0} \frac{a_n}{n!} z^n$$

for z in a neighbourhood of the origin. See Problem 2.12 for a derivation of the generating function; further historical details can be found in Stanley [111].

To determine asymptotics of a_n it is sufficient to find asymptotics of t_n. Since $\tan z = (\sin z)/(\cos z)$ the tangent function is meromorphic in the complex plane, with simple poles at the roots $\{\pm\pi/2, \pm 3\pi/2, \ldots\}$ of $\cos z$ and no other singularities. In particular, Theorem 2.1 implies

$$t_n = \operatorname*{Res}_{z=0} \left(\frac{\sin z}{z^{n+1} \cos z} \right) = \frac{1}{2\pi i} \int_{|z|=\varepsilon} \frac{\sin z}{\cos z} \frac{dz}{z^{n+1}} \tag{2.2}$$

for any $0 < \varepsilon < \pi/2$. The analysis proceeds in two steps.

Step 1: Introduce an Exponentially Smaller Integral

Let
$$I_n = \frac{1}{2\pi \mathrm{i}} \int_{|z|=2} \frac{\sin z}{\cos z} \frac{dz}{z^{n+1}},$$
and note that $\pi/2 < 2 < 3\pi/2$. Since $\tan z$ is analytic on the compact set $C_2 = \{z \in \mathbb{C} : |z| = 2\}$, there exists $M > 0$ such that $|\tan z| \leq M$ for $|z| = 2$. Thus, the maximum modulus integral bound implies

$$|I_n| \leq \text{length}(C_2) \times \max_{|z|=2} \left| \frac{\tan z}{(2\pi)z^{n+1}} \right| = 2(2\pi) \frac{M}{(2\pi)2^{n+1}} = O(2^{-n}).$$

We will soon show that I_n grows exponentially smaller than t_n, and its introduction allows us to simplify the Cauchy integral through residue computations.

Step 2: Compute Residues

Since $\tan z$ is meromorphic, and in the disk $\{|z| \leq 2\}$ only has poles at 0 and $\pm\pi/2$, the Cauchy residue theorem (Proposition 2.24 in the appendix) implies

$$I_n = \operatorname*{Res}_{z=0} \left(\frac{\sin z}{z^{n+1} \cos z} \right) + \operatorname*{Res}_{z=\pi/2} \left(\frac{\sin z}{z^{n+1} \cos z} \right) + \operatorname*{Res}_{z=-\pi/2} \left(\frac{\sin z}{z^{n+1} \cos z} \right).$$

The residue at the origin is the power series coefficient t_n, so after some simplification, and using the fact that $|I_n| = O(2^{-n})$, we obtain

$$t_n = -\left(\frac{2}{\pi}\right)^{n+1} \operatorname*{Res}_{z=\pi/2} \left(\frac{\sin z}{\cos z} \right) - \left(-\frac{2}{\pi}\right)^{n+1} \operatorname*{Res}_{z=-\pi/2} \left(\frac{\sin z}{\cos z} \right) + O(2^{-n}).$$

Lemma 2.4 of the appendix shows that the residue of any fraction $p(z)/q(z)$ at a point $z = a$ where $q(a) = 0$ and $p(a), q'(a) \neq 0$ equals $p(a)/q'(a)$. Using this fact, or simply computing the Laurent series of $\tan z$ at $z = \pm\pi/2$ in a computer algebra system, shows that the residues of $\tan z$ at $z = \pm\pi/2$ have value -1, and

$$t_n = \left(\frac{2}{\pi}\right)^{n+1} + \left(-\frac{2}{\pi}\right)^{n+1} + O(2^{-n}).$$

Thus, almost magically, we have proven that $a_n \sim 2\left(\frac{2}{\pi}\right)^{n+1} n!$ when n is odd. Recall that $a_n = 0$ when n is even.

Notes

We can come away from this example with several important observations.

1. The only property of the domain of integration $\{|z| = 2\}$ used in our argument was that the circle $\{|z| = 2\}$ separated the poles of $\tan z$ closest to the origin from the remaining poles. In particular, the domain of integration of I_n can be deformed to any circle $\{|z| = \tau\}$ where $\pi/2 < \tau < 3\pi/2$ without changing the value of I_n. Thus, our argument shows the stronger statement that
$$\frac{a_n}{n!} = t_n = 2\left(\frac{2}{\pi}\right)^{n+1} + O\left(\left(\frac{2}{3\pi} + \varepsilon\right)^n\right)$$
for all $\varepsilon > 0$ and n odd.

2. In fact, one can integrate over any circle $\{|z| = \tau\}$ with $\tau \neq (2k+1)\pi/2$ for $k \in \mathbb{N}$ to obtain an expression for t_n as a sum of residues with an error term. As τ grows the residues of $\tan z$ at additional poles will be added to the sum while the error gets exponentially smaller. For example, taking $3\pi/2 < \tau < 5\pi/2$ gives
$$\frac{a_n}{n!} = 2\left(\frac{2}{\pi}\right)^{n+1} + 2\left(\frac{2}{3\pi}\right)^{n+1} + O\left(\left(\frac{2}{5\pi}\right)^n\right)$$
for n odd. Note that we may use an error with exponential growth $2/5\pi$, instead of $2/5\pi + \varepsilon$, because the next asymptotic term is again given by a simple pole. Taking $\tau \to \infty$ even gives a convergent series expansion
$$\frac{a_n}{n!} = 2\left(\frac{2}{\pi}\right)^{n+1} \sum_{k\geq 0} \frac{1}{(2k+1)^{n+1}}$$
for odd n. Proving this infinite series representation requires uniformly bounding $|\tan z|$ away from its poles, for which it is more convenient to integrate over squares than circles; see Problem 2.13 for a proof strategy.

3. In this way, every singularity of $\tan z$ gives an asymptotic contribution to the power series coefficients t_n. The most important part of this contribution, its exponential factor, is determined by the location of the singularity. As each singularity in this example is a simple pole, the sub-exponential factor at each point is a constant.

Instead of simply introducing the integral I_n, it can be instructive to imagine expanding the domain of integration in the Cauchy integral (2.2) away from the origin while staying where the integrand is analytic, as in Figure 2.2. The domain can be expanded freely until it becomes locally 'stuck' in two places near the singularities at $z = \pm\pi/2$. One can compensate for pushing the domain of integration past these singularities by adding a contribution from each, determined by their residues. The domain of integration can then be expanded again, until getting 'stuck' on the next singularities. Since the Cauchy integrand decreases exponentially as the domain of integration moves away from the origin, the contributions from these singularities get successively smaller.

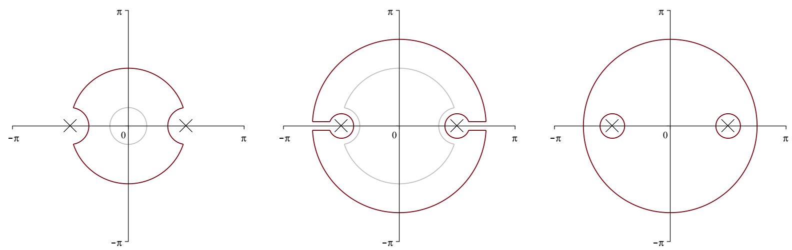


Fig. 2.2 Starting with a small circle at the origin, one can imagine pushing the domain of integration in the Cauchy integral away from the origin: this deformed curve gets ‘stuck’ near the two poles of $\tan z$ closest to the origin. Deforming past these singularities is possible if one adds integrals around each, which can be computed using residues.

This mental picture of sliding around domains of integration will be useful in the multivariate setting, when things are much harder to visualize, and helps guide the analysis in the presence of non-polar singularities such as algebraic branch points.

2.1.2 The Principles of Analytic Combinatorics

We now discuss the general theory of analytic combinatorics, which closely follows what we observed in the preceding example. Because we only consider power series representing analytic functions, the series coefficients of interest can grow at most exponentially. This makes their exponential growth the most significant component of their asymptotic behaviour.

Definition 2.5 (exponential growth) The *exponential growth* of a sequence (f_n) is the constant

$$\rho = \limsup_{n\to\infty} |f_n|^{1/n}.$$

The reason for the limsup is that different subsequences of (f_n) can have different growth; for instance, in combinatorial contexts there may be periodicities in the underlying counting problem. When (f_n) is the coefficient sequence of an analytic function at the origin then ρ is finite and is the coarsest measure of sequence behaviour (it may be zero, in which case f_n decays faster than any exponential function). The classic root test on series convergence immediately implies the following.

Proposition 2.1 *If* $\sum_{n\geq 0} f_n z^n$ *is a power series with finite radius of convergence* $R > 0$ *then the exponential growth of* (f_n) *is* $\rho = 1/R$.

Furthermore, Cauchy’s integral formula links the radius of convergence of a power series and the singularities of the analytic function it represents at the origin. The proof of the following result is Problem 2.16.

Proposition 2.2 *Suppose $F(z)$ is an analytic function represented in a neighbourhood of the origin by a power series with finite radius of convergence R. Then F admits a singularity on the circle $|z| = R$.*

Consequently, the exponential growth of F's power series coefficients is $1/R$, where R is the minimum modulus of a singularity of F. This motivates the first of two Principles of Coefficient Asymptotics[2].

First Principle of Coefficient Asymptotics: The *location* of a function's singularities dictates the exponential growth ρ of its power series coefficients.

Definition 2.6 (dominant singularity) A singularity of minimum modulus is called a *dominant singularity* of F.

The dominant singularities of a function are those which influence dominant asymptotics of its coefficient sequence. In many applications, F admits only a finite number of dominant singularities; for instance, this occurs whenever F is meromorphic.

When the power series coefficients (f_n) of F have exponential growth ρ then $f_n \sim \rho^n \theta(n)$ as $n \to \infty$, where $\theta(n)$ is a function which grows at most sub-exponentially (or decays). As in Section 2.1.1, one can imagine expanding the domain of integration in the Cauchy integral representation of Theorem 2.1 until it gets locally stuck near singularities of F. Each isolated singularity of F will give some asymptotic contribution to the power series coefficients f_n, with the type of the singularity (pole of a given order, algebraic branch point, logarithmic branch point, etc.) determining $\theta(n)$. This leads to the second principle of coefficient asymptotics.

Second Principle of Coefficient Asymptotics: The *nature* of a function's singularities determines the associated sub-exponential growth $\theta(n)$.

The simplest case occurs when the dominant singularities of F are poles. The asymptotic contributions of each singularity can then be determined by a residue calculation analogously to the computation in Section 2.1.1, giving the following.

Proposition 2.3 *Suppose that $F(z)$ is analytic on the circle $|z| = R$ and is analytic in the disk $|z| < R$ except at a finite number of (distinct) polar singularities $\sigma_1, \ldots, \sigma_m$, none of which are zero. Then there exist polynomials $P_1(n), \ldots, P_m(n)$ such that*

$$f_n = \sum_{k=1}^{m} P_k(n)\sigma_k^{-n} + O(R^{-n}),$$

and the degree of P_k is one less than the order of the pole of F at $z = \sigma_k$.

[2] As laid out by Flajolet and Sedgewick [50], the analytic combinatorialist's bible.

Proof If we define the integral

$$I_n = \frac{1}{(2\pi \mathrm{i})^n} \int_{|z|=R} \frac{F(z)}{z^{n+1}} dz$$

then the maximum modulus bound implies $I_n = O(R^{-n})$ and the Cauchy residue theorem gives

$$I_n = \operatorname*{Res}_{z=0} \left(\frac{F(z)}{z^{n+1}} \right) + \sum_{k=1}^{m} \operatorname*{Res}_{z=\sigma_k} \left(\frac{F(z)}{z^{n+1}} \right) = f_n + \sum_{k=1}^{m} \operatorname*{Res}_{z=\sigma_k} \left(\frac{F(z)}{z^{n+1}} \right).$$

As shown in Lemma 2.4 of the appendix, if $z = \sigma_k$ is a pole of order r then

$$\operatorname*{Res}_{z=\sigma_k} \left(\frac{F(z)}{z^{n+1}} \right) = \frac{1}{(r-1)!} \lim_{z \to \sigma_k} \frac{d^{r-1}}{dz^{r-1}} \left((z - \sigma_k)^r F(z) z^{-n-1} \right),$$

which is a polynomial in n of order $r - 1$. □

When $F(z)$ admits non-polar singularities, asymptotics can be computed using more nuanced methods. In the nineteenth century, Darboux [30] investigated coefficients of functions with algebraic singularities; see Section 2.3 below for more details. In the 1930s, Jungen [67] showed that when $F(z)$ behaves near a singularity $z = \rho$ like a sum of terms of the form

$$C(1 - z/\rho)^\alpha \log^r \frac{1}{1 - z/\rho}$$

with $r \in \mathbb{N}$, then each term with $\alpha \in \mathbb{C} \setminus \mathbb{N}$ makes an explicit asymptotic contribution to the power series coefficients of F which begins

$$\rho^{-n} \, n^{-\alpha-1} \, \log^r(n) \, \frac{C}{\Gamma(-\alpha)},$$

where Γ is the Euler gamma function, whose definition is recalled at the end of the appendix. A similar formula is available when $\alpha \in \mathbb{N}$ and $r > 0$, the only other situation in which F has a singularity at ρ. We shall see below that this statement encapsulates the singular behaviour of large classes of generating functions.

Flajolet and Odlyzko [48] coined the term *singularity analysis* for the process of translating local behaviour of a function at its singularities to asymptotic information about its power series coefficients. Those authors gave a modern and uniform treatment of such *transfer theorems,* considering types of singularities beyond those we encounter. See also the survey of Odlyzko [87], or the extensive treatment in Flajolet and Sedgewick [50], for additional details on these methods. Further information on classical methods for determining asymptotics can be found in de Bruijn [31].

2.1.3 The Practice of Analytic Combinatorics

Let $F(z)$ be analytic at the origin with coefficient sequence (f_n). If the dominant singularities of F are known, and they fall into one of the large number of types whose asymptotic contributions have been determined, then one can 'read off' dominant asymptotics of f_n by looking up precomputed transfer theorems. In practice, most of the work in an asymptotic argument thus goes towards the determination and classification of dominant singularities.

Because we deal mainly with power series whose coefficients are non-negative, the following result greatly simplifies the determination of dominant singularities. It has been attributed by various authors to Pringsheim, Borel or Vivanti; see Hadamard [61] and Vivanti [120] for historical context and Titchmarsh [113, Sect. 7.21] for a proof.

Proposition 2.4 (Vivanti-Pringsheim Theorem) *Suppose $F(z)$ is represented at the origin by a series expansion with real coefficients which has a finite radius of convergence $R > 0$. If this series expansion contains only a finite number of negative coefficients then $z = R$ is a singularity of $F(z)$.*

Thus, when the series coefficients f_n of $F(z)$ are non-negative, and do not grow or decay super-exponentially, their dominant asymptotics can usually be determined by the following steps:

1. Find the smallest real $r > 0$ such that $z = r$ is a singularity of $F(z)$;
2. Find all other singularities on the circle $|z| = r$;
3. If F admits a finite number of singularities on this circle, try to determine the asymptotic contribution of each by examining the local behaviour of F (see Propositions 2.11 and 2.17 below);
4. Add the results to get dominant asymptotics of f_n.

Example 2.3 (Asymptotics of Surjections)

A *surjection* of size n is a map from the set $[n] = \{1, 2, \ldots, n\}$ to any set of the form $[r] = \{1, 2, \ldots, r\}$ which covers $[r]$ (all elements of $[r]$ have a preimage). Using the Symbolic Method of Flajolet and Sedgewick [50], it is an easy exercise [50, p. 109] to show that if s_n denotes the number of surjections of size n then the generating function $F(z)$ of $s_n/n!$ satisfies

$$F(z) = \sum_{n\geq 0} \frac{s_n}{n!} z^n = \frac{1}{2 - e^z} = 1 + z + 3(z^2/2!) + 13(z^3/3!) + \cdots .$$

The singularities of $F(z)$ form the set $\{\log 2 + 2\pi i k : k \in \mathbb{Z}\}$, with the smallest real positive singularity equal to $\log 2$. There are no other singularities of modulus $\log 2$ and $F(z)$ has a simple pole at $z = \log 2$ with corresponding residue $-1/2$, so

$$s_n \sim \frac{n!}{2(\log 2)^{n+1}}.$$

We now discuss in detail several important families of generating functions: rational, algebraic, D-finite, and D-algebraic series. Each successive family in this list contains the previous families, and thus represents an increasingly complicated collection of sequences. For instance, although the power series coefficients of a rational generating function can be written as a finite sum of algebraic quantities which can be described explicitly, it is undecidable to determine several basic properties of the coefficients of a general D-algebraic function. We characterize the coefficient sequences corresponding to each class, describe how they can be manipulated, and identify possible asymptotic behaviour.

2.2 Rational Power Series

We begin with the class of rational power series.

Rational Functions and C-finite Sequences

In his pioneering eighteenth-century work, plausibly the first use of generating functions, de Moivre [32, 33] used rational generating functions to study sequences satisfying linear recurrences with constant coefficients.

Definition 2.7 (C-finite sequences and linear recurrences) Given $r \in \mathbb{N}$, a sequence (f_n) over a field $\mathbb{K}$ (not necessarily a subset of $\mathbb{C}$) satisfies a *linear recurrence relation of order r with constant coefficients* over $\mathbb{K}$ if there are constants $c_0, \ldots, c_{r-1} \in \mathbb{K}$ with $c_0 \neq 0$ such that

$$f_{n+r} = c_{r-1} f_{n+r-1} + c_{r-2} f_{n+r-2} + \cdots + c_0 f_n \tag{2.3}$$

for all $n \geq 0$. Any sequence satisfying (2.3) is uniquely determined by its first r terms $f_0, \ldots, f_{r-1}$, called the *initial conditions* of (f_n). A sequence that satisfies a linear recurrence relation over $\mathbb{K}$ is called *constant recursive, or C-finite,* over $\mathbb{K}$.

Remark 2.1 The set of sequences satisfying (2.3) forms a $\mathbb{K}$-vector space of dimension r: if f_n and g_n satisfy (2.3) then so does $f_n + \lambda g_n$ for any $\lambda \in \mathbb{K}$.

De Moivre showed that a sequence is C-finite if and only if its generating function is rational.

Theorem 2.2 *The sequence (f_n) satisfies the recurrence (2.3) if and only if its generating function $F(z)$ satisfies*

$$F(z) = \sum_{n \geq 0} f_n z^n = \frac{p(z)}{1 - c_{r-1} z - \cdots - c_0 z^r}$$

for some polynomial $p(z) \in \mathbb{K}[z]$ *of degree at most* $r - 1$.

The polynomial p depends on the initial conditions of (f_n). Given a rational function F whose numerator has degree greater than or equal to its denominator, polynomial division gives F as the sum of a rational function of this form and a polynomial, which only affects a finite number of coefficients. If $\mathbb{K}$ is not a subset of $\mathbb{C}$, the rational function can be viewed as a multiplicative inverse in the ring of formal power series.

Proof For any $n \geq 0$, the coefficient $[z^{n+r}]F(z)\,(1 - c_{r-1}z - \cdots - c_0 z^r)$ equals

$$[z^{n+r}]\left(\sum_{n\geq 0} f_n z^n\right)(1 - c_{r-1}z - \cdots - c_0 z^r) = f_{n+r} - c_{r-1}f_{n+r-1} - \cdots - c_0 f_n.$$

Thus, $F(z)\,(1 - c_{r-1}z - \cdots - c_0 z^r)$ is a polynomial of degree at most $r - 1$ if and only if (f_n) satisfies (2.3) for all $n \geq 0$. □

Linear recurrence relations have applications in an astounding number of scientific and mathematical fields, including combinatorics, probability, number theory, dynamical systems, theoretical computer science, economics, biology, and more. In computer algebra, fast computation with linear recurrence relations is used as a basis for other algorithms [121, Sect. 12.3]. The text of Everest et al. [41] is dedicated to C-finite sequences and their applications. Bostan et al. [14, Ch. 4] study the complexity of generating terms in C-finite sequences.

Examples of C-finite Sequences

Our first, and most famous, example of a C-finite sequence has transcended academia to appear in popular literature.

Example 2.4 (Virahanka-Fibonacci Numbers)

The *Virahanka-Fibonacci numbers* (v_n) satisfy the recurrence $v_{n+2} = v_{n+1} + v_n$ with initial conditions $v_0 = v_1 = 1$, and have generating function

$$V(z) = \frac{1}{1 - z - z^2} = 1 + z + 2z^2 + 3z^3 + 5z^4 + \cdots .$$

C-finite sequences occur naturally as generating functions of combinatorial classes whose elements are built out of a finite set of fixed objects[3]. Although

[3] For instance, traditional Sanskrit prosody is built from short and long units containing one or two syllables, respectively. As described by the Indian prosodist Virahanka around the 7th century AD, the number of long-short syllable sequences containing n syllables thus satisfies the Virahanka-Fibonacci recurrence $v_{n+2} = v_{n+1} + v_n$ (simply consider whether the final unit has 1 or 2 syllables).

satisfying a C-finite recurrence is a strong condition, some surprisingly complicated sequences can be C-finite.

Example 2.5 (Look-and-Say Digit Sequence)

Conway's look-and-say sequence (l_n), beginning

$$1,\ 11,\ 21,\ 1211,\ 111221,\ 312211,\ 13112221,\ \dots$$

is obtained by starting with 1 and recording the result of reading the previous blocks of digits from left to right (the first term contains 'one 1' giving the second term 11, the second term contains 'two 1s' giving the third term 21). A term of the form $x_1^{m_1} x_2^{m_2} \cdots x_k^{m_k}$ with distinct consecutive bases is followed by $m_1\, x_1\, m_2\, x_2\, \cdots m_k\, x_k$.

The *look-and-say digit sequence* d_n counts the number of digits of appearing in l_n, beginning with the terms

$$1,\ 2,\ 2,\ 4,\ 6,\ 6,\ 8,\ \dots$$

Conway [29] proved the miraculous fact that (d_n) is C-finite, and gave the rational generating function

$$D(z) = \frac{G(z)}{H(z)} = \frac{1 + z + \cdots + 18z^{77} - 12z^{78}}{1 - z + \cdots - 9z^{71} + 6z^{72}}.$$

If the starting term '1' is replaced by any natural number other than the fixed point '22' the resulting generating function will still be rational with denominator H.

Rational generating functions appear often in the study of lattice walks, which will be discussed in Chapter 4. Additional examples come from counting integer partitions whose summands come from any finite set, non-negative solutions to many linear Diophantine equations and inequalities, integer points in convex polyhedral cones and scalings of convex rational polytopes, walks on finite digraphs, words in Coxeter groups, and many other combinatorial objects. See Stanley [112], Bousquet-Mélou [18], and Flajolet and Sedgewick [50] for these and other examples.

The class of C-finite sequences is also closed under several natural operations.

Proposition 2.5 (Closure Properties) *If* (f_n) *and* (g_n) *are C-finite sequences then so are the sequences*

- $(f_n + g_n)_{n\geq 0}$
- $\left(\sum_{k=0}^{n} f_k g_{n-k}\right)_{n\geq 0}$
- $(f_n\, g_n)_{n\geq 0}$
- $(f_{nm+k})_{n\geq 0}$ *for any* $m \in \mathbb{N}$ *and* $0 \leq k < m$

Singh [107] gives a historical account of the Virahanka-Fibonacci numbers focused on the early contributions of Indian poets and mathematicians.

Proof If $F(z)$ and $G(z)$ are rational generating functions of the sequences (f_n) and (g_n), then the generating functions of $(f_n + g_n)$ and $\left(\sum_{k=0}^{n} f_k g_{n-k}\right)$ are the rational functions $F(z)+G(z)$ and $F(z)G(z)$, immediately giving the first two cases of Proposition 2.5. The operation taking $F(z) = \sum_{n\geq 0} f_n z^n$ and $G(z) = \sum_{n\geq 0} g_n z^n$ and returning $(F \odot G)(z) = \sum_{n\geq 0} f_n g_n z^n$ is known as the *Hadamard product* of F and G. C-finiteness of $(f_n g_n)$, and thus rationality of $F \odot G$, follows from Theorem 2.3 below. The Hadamard product of $F(z)$ and the rational function $z^k/(1-z^m)$ has the form $z^k H(z^m)$, where $H(z)$ is rational and equals the generating function of $(f_{nm+k})_{n\geq 0}$. □

ℕ-Rational Series

The following class of rational series arise frequently in combinatorial applications.

Definition 2.8 (ℕ-rational functions) The collection of functions containing the constant 1 and the variable z which is closed under addition, multiplication, and the *pseudo-inverse* operation $f \mapsto 1/(1-zf)$ forms the set of ℕ*-rational functions*.

Series expansions of ℕ-rational functions have deep importance in computer science as the generating functions of regular languages counted by length, equivalent to the class of languages recognized by deterministic finite automata[4], a well studied model of computation (see Rozenberg and Salomaa [95, Ch. 2] for details).

Although every ℕ-rational function is a rational function with natural number power series coefficients, there are rational power series with natural number coefficients which are not ℕ-rational. Of particular interest for asymptotics is the following necessary condition of ℕ-rationality given by Berstel [9], whose proof is Problem 2.9.

Proposition 2.6 *If $F(z)$ is an ℕ-rational function and σ is a dominant singularity of $F(z)$ then $\sigma/|\sigma|$ is a root of unity.*

Example 2.6 (A Non-ℕ-Rational Series)

Consider the rational function

$$F(z) = \frac{2(1+3z)^2}{(1-9z)(1+14z+81z^2)} = \frac{1/2}{1-9ze^{2i\theta}} + \frac{1/2}{1-9ze^{-2i\theta}} + \frac{1}{1-9z},$$

where $\theta = \arccos(1/3)$. Then $F(z)$ is not ℕ-rational by Proposition 2.6 as $e^{2i\theta}$ is not a root of unity. Since the denominator of F has a constant term of 1 the coefficients of $F(z)$ are integers. In fact, basic algebraic manipulations imply $[z^n]F(z) = 9^n\,(1+\cos(2n\theta)) \geq 0$ so $F(z)$ has natural number coefficients.

[4] A *finite automaton* over a finite set $\mathcal{A}$ is a directed multigraph (V, E) whose edges are labelled by elements of $\mathcal{A}$, where one vertex $v \in V$ is chosen as an *initial vertex* and a collection $Q \subset V$ of vertices are chosen as *accepting vertices*. A sequence $(a_1, \ldots, a_r) \in \mathcal{A}^r$ for some $r \in \mathbb{N}$ is *recognized* by the automaton if there is a path in the automaton from v to an element of Q whose consecutive edges have labels $a_1, \ldots, a_r$.

Soittola [109] gave a characterization of $\mathbb{N}$-rational functions based on singularity analysis. Barcucci et al. [5], building on Soittola's work, gave an algorithm which takes a rational function $F(z)$, tests for $\mathbb{N}$-rationality, and, if this test is positive, returns a regular language whose counting sequence is the coefficient sequence of F. In the words of those authors, this allows one to "reverse-engineer a combinatorial problem from a rational generating function."

As a meta-principle, generating functions of combinatorial classes seem to be $\mathbb{N}$-rational. In the proceedings for the 2006 International Congress of Mathematics, Bousquet-Mélou [18] writes that she "never met a counting problem that would yield a rational, but not $\mathbb{N}$-rational, GF." Koutschan [71] gave a Maple implementation of the algorithm of Barcucci et al. and tested "about 60" sequences of combinatorial significance with rational generating functions, finding all functions to be $\mathbb{N}$-rational. As we will soon see, this means certain deep open problems about the behaviour of rational function coefficients can be avoided for combinatorial enumeration problems. Trying to generalize characterizations of $\mathbb{N}$-rationality to coefficient sequences of multivariate rational functions is an open problem, discussed in Chapter 3.

Asymptotics

In addition to characterizing C-finite sequences in terms of their generating functions, de Moivre also studied their closed-form solutions, as did Bernoulli [8] and Euler [40, V. 1 Ch. XIII] around the same time period[5]. The following result gives an exact closed-form representation for terms of sufficiently large index.

Theorem 2.3 *Suppose $F(z) = G(z)/H(z) \in \mathbb{Q}(z)$ is a rational function with G and H coprime polynomials and $H(0) \neq 0$. Let d denote the degree of H and $\sigma_1, \dots, \sigma_m$ be the distinct roots of $H(z)$ in the complex plane. Then there exist polynomials $P_1(n,x), \dots, P_m(n,x)$ in $\mathbb{Q}[n,x]$, whose degrees in x are at most d, such that for all n larger than some fixed natural number the power series coefficients of $F(z)$ satisfy*

$$f_n = \sum_{j=1}^{m} P_j(n, \sigma_j)\, \sigma_j^{-n}. \tag{2.4}$$

The degree of $P_j(n,x)$ in n is one less than the order of the pole of $F(z)$ at $z = \sigma_j$. Conversely, if (f_n) has a representation of the form (2.4) then the generating function $F(z)$ is a rational function.

Proof Theorem 2.3 is Proposition 2.3 specialized and strengthened in the case when $F(z)$ is rational, and their proofs are analogous. In particular, when $R > 0$ is

[5] For instance, Euler advocated the use of partial fraction decomposition to find the general term of a series with rational generating function.

large enough so that all roots of $H(z)$ lie in the disk $|z| < R$ the Cauchy residue theorem and maximum modulus bound imply

$$\left| f_n + \sum_{j=1}^{m} \operatorname*{Res}_{z=\sigma_j} \left(\frac{F(z)}{z^{n+1}} \right) \right| = \left| \frac{1}{2\pi \mathrm{i}} \int_{|z|=R} \frac{F(z)}{z^{n+1}} dz \right| \leq \frac{\max_{|z|=R} |F(z)|}{R^n}. \tag{2.5}$$

Because $F(z)$ is rational, there exists $k > 0$ such that for any R sufficiently large $\max_{|z|=R} |F(z)| \leq R^k$. Thus, whenever $n > k$ the right-hand side of (2.5) goes to zero as $R \to \infty$, giving (2.4). The bounds on the degrees of the P_j follow from Lemma 2.4 in the appendix of this chapter and the fact that each σ_j is an algebraic number of degree at most d.

For any $k \in \mathbb{N}$ and $c, \alpha \in \mathbb{C}$ the generating function of the sequence $f_n = cn^k \alpha^n$ is the rational function $cR_k(\alpha z)$, where $R_k(z)$ is obtained by starting with the geometric series $1/(1-z)$ and alternatively taking the derivative and multiplying by z repeated k times. Distributing the sum in (2.4) then proves the final statement. □

Gourdon and Salvy [58] give algorithms, running in polynomial time in the degree d, for computing the polynomials P_j and determining which terms of the sum dictate dominant asymptotics of f_n. Calculating the P_j can be done symbolically using resultants and other algebraic tools, while determining which roots of $H(z)$ are closest to the origin and dictate asymptotics requires numeric algorithms to separate the algebraic numbers σ_j and their moduli. These algorithms may be viewed as special cases of the machinery we develop in Chapter 7 for the multivariate setting.

Example 2.7 (Virahanka-Fibonacci Asymptotics)

The generating function of the Virahanka-Fibonacci numbers can be written

$$V(z) = \frac{1}{1 - z - z^2} = \frac{1}{\sqrt{5}} \left(\frac{1}{\sigma} \frac{1}{1 - z/\sigma} - \frac{1}{\tau} \frac{1}{1 - z/\tau} \right),$$

where $\sigma = (-1 + \sqrt{5})/2$ and $\tau = (-1 - \sqrt{5})/2$ are the roots of $1 - z - z^2$. Thus, expanding $1/(1 - z/\sigma)$ and $1/(1 - z/\tau)$ as geometric series gives

$$v_n = [z^n]V(z) = \frac{1}{\sigma\sqrt{5}} \sigma^{-n} - \frac{1}{\tau\sqrt{5}} \tau^{-n}.$$

Similarly, the tools developed in Chapter 7 allow one to automatically prove that the look-and-say sequence generating function $D(z)$ admits a single dominant singularity, necessarily real and positive, which is the root $\lambda = .7671198507\ldots$ of the denominator $H(z)$. Since the derivative of H does not vanish at λ this is a simple pole of $D(z)$, and the look-and-say digit counting sequence has asymptotic behaviour

$$d_n \sim \frac{-G(\lambda)}{\lambda\, (\partial H/\partial z)(\lambda)}\, \lambda^{-n} \approx (2.04216\ldots)\,(1.30357\ldots)^n.$$

An automated analysis of this rational generating function was carried out by Gourdon and Salvy [58]. Koutschan [72] showed that $D(z)$ is $\mathbb{N}$-rational.

Equation (2.4) is an exact equality, for n sufficiently large, but not every summand will necessarily contribute to dominant asymptotics. Although this characterization is explicit there are still significant open problems in the area, due to potentially complicated interactions between the algebraic quantities involved.

Open Problem 2.1 (Skolem's Problem) *Is there an algorithm which takes a rational function with integer coefficients, analytic at the origin, and determines whether it has a power series coefficient which is zero? Equivalently, is there an algorithm which takes a C-finite sequence over* $\mathbb{Q}$*, defined by a linear recurrence and initial conditions, and decides whether any term in the sequence is zero?*

In the 1930s, Skolem [108] showed that if a sequence (f_n) of rational numbers[6] is C-finite then the set $\{n : f_n = 0\}$ of indices of zero terms consists of a finite set together with a finite set of arithmetic progressions, and this problem has essentially been open since then. When the set of zero indices is infinite, meaning it contains an arithmetic progression, this can be detected [10]. Currently, all known proofs of Skolem's theorem use p-adic analysis in a manner which does not give an upper bound on the potential indices of sporadic zeroes[7]. Mignotte et al. [84] and Vereshchagin [119] show Skolem's problem is decidable for sequences satisfying C-finite recurrences of orders 3 or 4. As noted in Kenison et al. [70], the crucial open case for recurrences of order 5 consists of sequences (u_n) of the form[8] $u_n = a(\lambda_1^n + \overline{\lambda_1^n}) + b(\lambda_2^n + \overline{\lambda_2^n}) + c\rho^n$ where $|\lambda_1| = |\lambda_2| > |\rho|$ with a, b, c real algebraic numbers and $|a| \neq |b|$. Those authors give a method [70, Thm. 1.2] to determine when C-finite sequences of any order take the value zero on terms whose indices are prime powers.

If (f_n) is a C-finite sequence, so is the sequence (f_n^2). Thus, the problem of detecting zero coefficients over $\mathbb{Q}$ can be reduced to deciding whether every term of a C-finite sequence is positive. This condition can then be relaxed, motivating the following problem which is more directly related to our asymptotic setting.

Open Problem 2.2 (Ultimate Positivity) *Is there an algorithm which takes a rational function with integer coefficients, analytic at the origin, and determines whether its series coefficients are eventually non-negative (i.e., if there are only a finite*

---

[6] Mahler [78] and Lech [76] extended this result to sequences over the field of algebraic numbers, and any field of characteristic zero, respectively. The theorem does not hold in finite characteristic: for instance, if $\mathbb{F}_p$ denotes the finite field of prime order p then the sequence $c_n = (1+x)^n - 1 - x^n$ defined over $\mathbb{K} = \mathbb{F}_p(x)$ satisfies a C-finite recurrence of order 3, but $c_n = 0$ if and only if $n = p^k$ for some $k \in \mathbb{N}$. This counter-example is due to Lech [76].

[7] Additional information on Skolem's theorem can be found in Chapter 2 of Everest et al. [41].

[8] An explicit example of such a sequence, given by Kenison et al. [70], is the sequence defined by $u_{n+5} = -41u_{n+4} + 952u_{n+3} - 178360u_{n+2} - 17673175u_{n+1} + 17850625u_n$ with initial conditions $u_0 = 9,\ u_1 = -281,\ u_2 = 15485,\ u_3 = -1135097,\ u_4 = -30999543.$

number of negative coefficients)? Equivalently, is there an algorithm which takes a C-finite sequence over $\mathbb{Q}$, defined by a linear recurrence and initial conditions, and decides whether its terms are eventually non-negative? What if 'non-negative' is strengthened to 'positive'?

Ouaknine and Worrell [89, 88] show that the Ultimate Positivity problem is decidable for a rational function with square-free denominator (i.e., when the denominator and its derivative are coprime polynomials) and for an arbitrary rational function whose denominator has degree at most 5. Those authors also show that proving Ultimate Positivity for rational functions whose denominators have degree 6 would give a method of computing the so-called *Lagrange constant* for a large collection of transcendental numbers, a significant breakthrough in analytic number theory.

Ultimate Positivity is exactly the setting in which the Vivanti-Pringsheim Theorem can be applied. Of course, when one is dealing with the generating function of a combinatorial class then non-negativity of coefficients is guaranteed. Because of the meta-principle that combinatorial classes always seem to have $\mathbb{N}$-rational generating functions, in practice asymptotics can be effectively decided for combinatorial problems and combinatorialists do not need to worry about these pathological decidability issues. In particular, Proposition 2.6 implies that the dominant singularities of an $\mathbb{N}$-rational function differ by roots of unity, meaning their coefficients have explicit periodic behaviour: there exists a natural number m such that for each $k = 0, \dots, m-1$ the coefficient sub-sequence (f_{mn+k}) has dominant asymptotics of the form $C_k\, n^{\alpha_k}\, \rho_k^n$ where α_k is a computable natural number and C_k and ρ_k are algebraic constants with computable minimal polynomials. Flajolet et al. [49] give a similar characterization for dominant singularities of combinatorial classes obtained from a large collection of recursive 'constructions'.

2.3 Algebraic Power Series

We next examine series satisfying polynomial equations. Throughout this section we assume that $\mathbb{K}$ is an algebraically closed field of characteristic zero, of which the most important cases are $\mathbb{K} = \mathbb{A}$, the field of algebraic numbers over $\mathbb{Q}$, and $\mathbb{K} = \mathbb{C}$. We will see how bivariate polynomials can be used as data structures for algebraic functions, and how algebraic tools can be leveraged to create algorithms for coefficient asymptotics. In particular, the potential singularities of an algebraic function can be characterized, after which asymptotics can be 'read off' from local expansions near the singularities. In order to solve algebraic equations we must consider more general objects than power series.

Laurent and Puiseux Series Expansions

First, we note that to solve equations such as $zF(z) = 1$ it is necessary to introduce series expansions with negative exponents.

Definition 2.9 (formal Laurent series) The *field of formal Laurent series over* $\mathbb{K}$, denoted $\mathbb{K}((z))$, consists of formal series with potentially negative integer exponents bounded from below,

$$\mathbb{K}((z)) = \left\{ \sum_{n\geq N} a_n z^n \,:\, N \in \mathbb{Z},\ a_j \in \mathbb{K} \text{ for all } j \right\},$$

with term-wise addition and Cauchy product.

Problem 2.5 implies that $\mathbb{K}((z))$ is the field of fractions of the ring of formal power series $\mathbb{K}[[z]]$. In a similar fashion, to solve algebraic equations such as $F(z)^R = z$ using formal series we must introduce fractional exponents.

Definition 2.10 (formal Puiseux series) The *field of formal Puiseux series,* denoted $\mathbb{K}^{\text{fra}}((z))$, consists of the set of formal series

$$\mathbb{K}^{\text{fra}}((z)) = \left\{ \sum_{n\geq N} a_n z^{n/R} \,:\, N \in \mathbb{Z},\ R \in \mathbb{N}_{>0},\ a_j \in \mathbb{K} \text{ for all } j \right\}$$

with term-wise addition and Cauchy product.

The variables appearing in a Puiseux series have rational exponents that are bounded from below and have fixed denominators, so expressions such as $\sum_{n\geq 0} z^{-n}$ and $\sum_{n\geq 1} z^{1/n}$ are not Puiseux series. Although $\mathbb{K}^{\text{fra}}((z))$ does not include all series with rational exponents, it is adequate to describe the solutions of any algebraic equation. The following result dates back to Newton [85] and Puiseux [94]; a proof can be found in Walker [122, Sect. 3].

Proposition 2.7 (Puiseux's Theorem) *Let* $\mathbb{K}$ *be an algebraically closed field of characteristic zero. Then* $\mathbb{K}^{fra}((z))$ *is the algebraic closure of* $\mathbb{K}((z))$.

Definition 2.11 (algebraic series and minimal polynomials) A formal series $F(z) \in \mathbb{K}^{\text{fra}}((z))$ is *algebraic* over $\mathbb{K}$ if there exist polynomials $p_0(z), \dots, p_d(z) \in \mathbb{K}[z]$, not all zero, such that

$$p_d(z)F(z)^d + p_{d-1}(z)F(z)^{d-1} + \cdots + p_0(z) = 0.$$

If $F(z)$ is algebraic, a *minimal polynomial* of F is any non-zero polynomial $P(z, y) \in \mathbb{K}[z][y]$ of minimal degree in y with coefficients in $\mathbb{K}[z]$ such that $P(z, F(z)) = 0$; this is unique up to a non-zero multiple of $\mathbb{K}[z]$. The Puiseux series roots of a minimal polynomial of F are called the *conjugates* of F. A series which is not algebraic is called *transcendental*.

Just as the minimal polynomial $x^2 - 2$ of $\sqrt{2}$ over $\mathbb{Q}$ also admits $-\sqrt{2}$ as a root, a general algebraic series $F(z)$ must be encoded by its minimal polynomial together with additional information to distinguish F among its conjugates. Even when F is a power series its conjugates can contain negative and fractional powers.

Example 2.8 (Rooted Binary Trees)

A rooted binary plane tree is a rooted tree (including the empty tree) where every vertex has two (possibly empty) ordered binary plane trees as children. From this recursive definition it can be shown that the generating function $C(z)$ of such trees, counted by their number of vertices, satisfies $P(z, C(z)) = 0$ where

$$P(z, y) = zy^2 - y + 1.$$

By Puiseux's Theorem we know there are two formal series solutions to $P(z, y) = 0$. Let $f(z) = a_r z^r + \cdots$ for some $r \in \mathbb{Q}$ with $a_r \neq 0$ and algebraic coefficients $a_j \in \mathbb{A}$ to be determined, where '$\cdots$' hides terms with strictly larger exponents. Substitution into the equation $P(z, f(z)) = 0$ gives

$$0 = z\,(a_r z^r + \cdots)^2 - (a_r z^r + \cdots) + 1 = \left(a_r^2 z^{2r+1} + \cdots\right) - (a_r z^r + \cdots) + 1. \quad (2.6)$$

In order for this equality to hold, every term on the right hand side of (2.6) must cancel. In particular, there must be at least two terms whose exponent of z is minimal, and these terms must cancel. Thus, either

$$\begin{array}{lllll} & 2r+1 = r < 0 & \text{and} & a_r^2 - a_r = 0 & (z^{2r+1} \text{ and } z^r \text{ are minimal}) \\ \text{or} & 2r+1 = 0 < r & \text{and} & a_r^2 + 1 = 0 & (z^{2r+1} \text{ and } z^0 \text{ are minimal}) \\ \text{or} & r = 0 < 2r+1 & \text{and} & -a_r + 1 = 0 & (z^r \text{ and } z^0 \text{ are minimal}). \end{array}$$

In the first case, $r = -1$ and $a_r = a_{-1} = 1$ since $a_r \neq 0$ by assumption. The second case cannot occur, but the third can when $r = 0$ and $a_r = a_0 = 1$. Thus, we have found the start of two series expansions

$$f_1(z) = z^{-1} + \cdots \qquad \text{and} \qquad f_2(z) = 1 + \cdots$$

such that $P(z, f_1) = P(z, f_2) = 0$. Substituting these starting terms back in (2.6) allows one to repeat the process and obtain further terms in the series expansions. For example, when $r = -1$ and $a_{-1} = 1$ then repeated back substitution shows the next non-zero term is $a_0 z^0$ with $a_0 = -1$, followed by $a_1 z$ with $a_1 = -1$, etc. Puiseux's theorem implies that this process of extending a partial expansion must succeed, giving solutions

$$f_1(z) = z^{-1} - 1 - z - 2z^2 - 5z^3 + \cdots \qquad \text{and} \qquad f_2(z) = 1 + z + 2z^2 + 5z^3 + 14z^4 + \cdots$$

to $P(z, f_1) = P(z, f_2) = 0$.

This computational procedure of extending truncated series solutions can be automated, a process known as the *Newton polygon method* (see, for example, Walker's proof [122] of Puiseux's Theorem). The name comes from the use of a polygon defined by the exponents of monomials in $P(z, y)$ to determine the potential starting indices of solutions. There are several computer algebra packages that can compute Puiseux series solutions of equations, such as the gfun package [101] in Maple.

Example 2.9 (Rooted 3-Ary Trees)

A rooted 3-ary plane tree is a rooted tree where every vertex has three ordered 3-ary plane trees as children. The generating function $F(z)$ of 3-ary plane trees satisfies

$$P(z, F(z)) = 0, \qquad P(z, y) = zy^3 - y + 1.$$

There are now three solutions,

$$f_1(z) = \frac{1}{z^{1/2}} - \frac{1}{2} - \frac{3z^{1/2}}{8} + \cdots, \qquad f_2(z) = \frac{-1}{z^{1/2}} - \frac{1}{2} + \frac{3z^{1/2}}{8} + \cdots,$$

and

$$f_3(z) = 1 + z + 3z^2 + 12z^3 + \cdots$$

satisfying $P(z, f_1) = P(z, f_2) = P(z, f_3) = 0$. Even though the generating function f_3 is a power series, its conjugates f_1 and f_2 have fractional exponents.

When working over the complex numbers there is a useful analytic version of Puiseux's Theorem involving convergent series. When moving from formal to analytic considerations, we account for branch cuts of algebraic functions by considering series in disks with rays removed (the appendix to this chapter contains additional details on branch cuts).

Definition 2.12 (slit disks) A *slit disk* centred at $z = \omega$ is a disk $|z - \omega| < R$ centred at ω with a ray $\{\omega + ta : t > 0\}$ removed, for any $R > 0$ and $a \in \mathbb{C} \setminus \{0\}$.

See Hille [66, Ch. 12] for a proof of the following.

Proposition 2.8 (Puiseux's Theorem, Analytic Version) *Let $P(z, y) \in \mathbb{Q}[z, y]$ be an irreducible polynomial of degree R in y, and let $\omega \in \mathbb{C}$. Then in a sufficiently small slit disk centred at ω there exist R convergent Puiseux series expansions*

$$y_k(z) = \sum_{n \geq N} c_n^{(k)} (1 - z/\omega)^{n/R}, \qquad 1 \leq k \leq R$$

with $P(z, y_k(z)) = 0$ for each $1 \leq k \leq R$, where the branch of $(1 - z/\omega)^{1/R}$ is fixed in the slit disk and all coefficients $c_n^{(k)}$ are algebraic numbers.

The point $z = \omega$ is a singularity of one of the analytic functions defined by these expansions if and only if the corresponding expansion is not a power series.

The coefficients in such a convergent Puiseux series can be determined by applying the Newton polygon method to the polynomial $Q(t, y) = P(\omega(1 - t), y)$ to obtain Puiseux series in t and then substituting $t = 1 - z/\omega$. We consider these expansions to be centred at $z = \omega$ as they converge in a slit disk centred at ω.

Example 2.10 (Puiseux Expansions I)

To find convergent Puiseux series solutions of $P(z, y) = zy^2 - y + 1$ centred at $z = 1/4$, we can find the Puiseux series solutions of

$$P((1-t)/4, y) = (1-t)y^2/4 - y + 1 = 0$$

at $t = 0$, and then substitute $t = 1 - 4z$. The Newton polygon method gives solutions

$$2 + 2t^{1/2} + 2t + 2t^{3/2} + \cdots = 2 + 2\,(1-4z)^{1/2} + 2\,(1-4z) + 2\,(1-4z)^{3/2} + \cdots$$

and

$$2 - 2t^{1/2} + 2t - 2t^{3/2} + \cdots = 2 - 2\,(1-4z)^{1/2} + 2\,(1-4z) - 2\,(1-4z)^{3/2} + \cdots$$

Taking the principal branch of the square-root function $t^{1/2}$, these series converge in the slit disk where $|1 - 4z| < 1$ and z is not a real number larger than $1/4$ (so that $1 - 4z$ is not real and negative).

An algebraic generating function is encoded by a polynomial $P(z, y) \in \mathbb{Z}[z, y]$, but P also encodes other algebraic series. We now discuss how to separate a specific series of interest from the other roots of P.

Definition 2.13 (singular parts and separation orders) Let $P(z, y) \in \mathbb{Z}[z, y]$ and suppose $f(z) = \sum_{n \geq N} c_n z^{n/R}$ satisfies $P(z, f(z)) = 0$. The *singular part of f* is the initial partial sum $f_I(z) = \sum_{n=N}^{\ell} c_n z^{n/R}$, where $\ell \geq N$ is the minimal integer such that f_I is not equal to an initial partial sum of any other Puiseux series expansion $y(z)$ with $P(z, y(z)) = 0$. The *separation order* of $f(z)$ with respect to P is the integer ℓ.

The following result, found in Walsh [123], not only bounds the separation order but also shows how it can be used to determine the smallest field extension of $\mathbb{Q}$ containing all coefficients of a Puiseux series expansion.

Proposition 2.9 *Let $P(z, y) \in \mathbb{Q}[z, y]$ be an irreducible polynomial in y of degree d_y in y and degree d_z in z, and $f(z) = \sum_{n \geq N} c_n z^{n/R}$ satisfy $P(z, f(z)) = 0$. If the separation order of f with respect to P is ℓ then $\ell \leq 2d_z d_y(2d_y - 1)$ and all Puiseux series coefficients $c_N, c_{N+1}, \ldots$ of f lie in the finite extension $\mathbb{Q}(c_N, \ldots, c_\ell)$.*

Knowing that the coefficients of each Puiseux series expansion lie in a finite extension of $\mathbb{Q}$ is necessary for computation.

Example 2.11 (Puiseux Expansions II)

The Puiseux series solutions of

$$P(z, y) = (z + 1/4)y^2 - y + 1$$

begin

$$2 + 4\mathrm{i}z^{1/2} - 8z + \cdots \qquad \text{and} \qquad 2 - 4\mathrm{i}z^{1/2} - 8z + \cdots .$$

Since these series are separated by their truncations at order 2, all coefficients of both Puiseux series expansions lie in $\mathbb{Q}(2, \pm 4\mathrm{i}) = \mathbb{Q}(\mathrm{i})$.

Later, in our arguments on lattice path enumeration, we will need to determine how many Puiseux series solutions of an algebraic equation are fractional power series (i.e., contain no negative exponents). The following result gives an answer.

Proposition 2.10 *Let $P(z, y) = p_0(z) + p_1(z)y + \cdots + p_d(z)y^d \in \mathbb{Q}[z][y]$ be a polynomial of degree d in y such that $p_j(0) \neq 0$ for at least one index j. Then the number of series solutions to $P(z, y) = 0$ in y at $z = 0$ which are fractional power series equals the maximum index j such that $p_j(0) \neq 0$.*

A proof of Proposition 2.10, which involves algebraic manipulations of Puiseux expansions, can be found in Stanley [110, Prop. 6.1.8].

Singularities of Algebraic Curves

We now determine the points where the solution of an algebraic equation can have a singularity, in order to apply the techniques of analytic combinatorics to determine coefficient asymptotics. This requires some algebraic tools.

Definition 2.14 (resultants and discriminants) Given polynomials

$$p(x) = a_0 + a_1 x + \cdots + a_d x^d \qquad \text{and} \qquad q(x) = b_0 + b_1 x + \cdots + b_e x^e$$

over an integral domain R, where $a_d, b_e \neq 0$, the *resultant* of p and q with respect to x is

$$\text{Resultant}_x(p, q) = a_d^e \, b_e^d \prod_{p(\alpha)=q(\beta)=0} (\alpha - \beta),$$

where α and β run through the roots of p and q in an algebraic closure of R. The *discriminant* of $p(x)$ with respect to x is the quantity

$$\text{Disc}_x(p) = \frac{(-1)^{d(d-1)/2}}{a_d} \text{Resultant}_x(p, p'),$$

where $p'(x)$ is the derivative of $p(x)$ with respect to x.

Problem 2.14 asks you to prove that the resultant lies in R, and gives an algorithm for its calculation. Note that the resultant of p and q is zero if and only if p and q

share a root in any algebraic closure of R. Thus, the discriminant $\mathrm{Disc}_x(p)$ is zero if and only if p has a multiple root in any algebraic closure of R. Both the resultant and discriminant can be calculated by any major computer algebra system using efficient algorithms [121, Ch. 6] (the cost is not much greater than polynomial multiplication).

Now, let
$$P(z, y) = p_d(z)y^d + p_{d-1}(z)y^{d-1} + \cdots + p_0(z)$$
be an irreducible polynomial in $\mathbb{Q}[z][y]$ of degree d in y. For fixed $a \in \mathbb{C}$ the polynomial $P(a, y)$ has d solutions for y in the complex plane unless $p_d(a) = 0$ or $P(a, y)$ has a multiple root in y, meaning $\mathrm{Disc}_y(P) \in \mathbb{Q}[z]$ vanishes at $z = a$.

Definition 2.15 (exceptional sets) The *exceptional set* of $P(z, y) \in \mathbb{Z}[z, y]$ is the finite set $\Xi = \Xi(P) = \{z \in \mathbb{C} : p_d(z) = 0 \text{ or } \mathrm{Disc}_y(P)(z) = 0\}$.

For any $a \in \mathbb{C} \setminus \Xi$ the polynomial $P(a, y)$ has d distinct solutions in y, and the implicit function theorem (described in a more general context in Proposition 3.1 of Chapter 3) implies the existence of analytic functions $y_1(z), \ldots, y_d(z)$ defined in a neighbourhood of $z = a$ such that $y_1(a), \ldots, y_d(a)$ are the solutions of $P(a, y) = 0$.

Definition 2.16 (branches and algebraic functions) Each analytic function $y_j(z)$ in the proceeding paragraph defines a *branch* of the curve $P(z, y) = 0$ in any simply connected region of the complex plane where it is analytic. An *algebraic function* on a domain Ω is any function on Ω which is a branch of some algebraic curve.

One immediate consequence of the implicit function theorem is the following.

Lemma 2.2 *If $z = a$ is a singularity of a branch of $P(z, y)$ then $a \in \Xi$.*

Vanishing of $p_d(z)$ corresponds to branches going off to infinity: if $p_d(a) = 0$ then the collection of Puiseux series solutions to $F(z, y)$ at $z = a$ can contain series with negative exponents. Vanishing of $\mathrm{Disc}_y(P)(z)$ corresponds to two branches colliding: if $\mathrm{Disc}_y(P)(a) = 0$ then the collection of Puiseux series solutions to $F(z, y)$ at $z = a$ can contain series with non-integer fractional powers coming in sets of conjugates.

Example 2.12 (Singularities for Plane Trees)

The generating function for the number of rooted binary plane trees is a branch of
$$P(z, y) = zy^2 - y + 1.$$
Then $\mathrm{Disc}_y(P) = 1 - 4z$, so that
$$\Xi = \{z : z = 0 \text{ or } 1 - 4z = 0\} = \{0, 1/4\}.$$
We have already seen that one Puiseux series solution of $P(z, y) = 0$ at the origin is a Laurent series with negative exponents, which corresponds to the vanishing of the leading coefficient of y, and that the series solutions centred at $z = 1/4$ contain non-integer powers due to the vanishing of the discriminant.

The generating function for the number of rooted 3-ary plane trees is a branch of

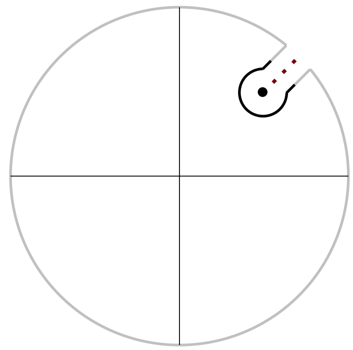


Fig. 2.3 When $F(z)$ behaves like an algebraic branch point near its unique dominant singularity $z = \omega$, the domain of integration in the Cauchy integral can be deformed around the corresponding branch cut $\mathcal{R}$ (dotted) to obtain an open circle and small lip C_1 which can be made arbitrarily close to ω (black) together with a collection of points C_2 whose moduli are bounded above $|\omega|$ (gray). The Puiseux expansion of $F(z)$ is a good approximation when C_1 is sufficiently close to ω, which allows for the transfer of coefficient asymptotics, while the Cauchy integrand is exponentially smaller than ω^{-n} on C_2.

$$P(z, y) = zy^3 - y + 1,$$

which has as its Puiseux series solutions at the origin the power series

$$1 + z + 3z^2 + 12z^3 + 55z^4 + \cdots$$

together with the fractional Puiseux series

$$z^{-1/2} - \frac{1}{2} - \frac{3}{8}z^{1/2} + \frac{1}{2}z + \cdots \qquad \text{and} \qquad -z^{-1/2} - \frac{1}{2} + \frac{3}{8}z^{1/2} + \frac{1}{2}z + \cdots .$$

Here $\text{Disc}_y(P) = z(4-27z)$, so that both the discriminant and the leading coefficient of P vanish at the origin. This helps explain the existence of two series with non-integer powers which are negative.

Asymptotic Behaviour

Suppose $F(z)$ is analytic at the origin and has a single dominant singularity $z = \omega$, where it behaves like an algebraic branch point. The key to determining asymptotics is to deform the domain of integration in the Cauchy integral (2.1) around the singularity ω without crossing the corresponding branch cut; see Figure 2.3. Away from the singularity ω, the domain of integration can be deformed so that the Cauchy integrand in (2.1) is exponentially smaller than ω^{-n}. Near the singularity F behaves like its Puiseux expansion, so coefficient asymptotics of $F(z)$ can be determined from the terms in its Puiseux expansion. In particular, coefficient asymptotics of a summand $(1-z/\omega)^\alpha$ follows from the binomial theorem and Stirling's approximation for the gamma function, and these asymptotic expansions can be summed to give an asymptotic expansion for the coefficients of $F(z)$. A rigorous version of this argument, yielding the following result, can be found in Flajolet and Odlyzko [48].

Proposition 2.11 (Darboux's Theorem) *Suppose $F(z) = \sum_{n\geq 0} f_n z^n$ is analytic in some open disk around the origin except for a single dominant singularity $\omega \in \mathbb{C}\setminus\{0\}$ and points in the ray $\mathcal{R}_\omega = \{t\omega : t \geq 1\}$, such that*

$$F(z) = \sum_{k\geq N} c_k (1 - z/\omega)^{k/R}$$

is a convergent expansion of F in a disk centred at ω with $\mathcal{R}_\omega$ removed. Then each summand $c_k(1 - z/\omega)^{k/R}$ with $k/R \notin \mathbb{N}$ adds an asymptotic contribution

$$c_k \omega^{-n} \,\frac{\Gamma(n - k/R)}{\Gamma(-k/R)\,\Gamma(n + 1)} \sim c_k \omega^{-n} \frac{n^{-k/R-1}}{\Gamma(-k/R)}$$

to f_n, where Γ is the Euler gamma function. In particular, for all $M \geq 0$

$$f_n = \frac{\omega^{-n}}{\Gamma(n + 1)} \left(\sum_{k=N}^{N+M} c_k \,\frac{\Gamma(n - k/R)}{\Gamma(-k/R)} + O\left(n^{-(M+1)/R}\right) \right)$$

where $\frac{\Gamma(n-k/R)}{\Gamma(-k/R)} = 0$ if $k/R \in \mathbb{N}$, and $f_n \sim c_p \omega^{-n} \frac{n^{-p/R-1}}{\Gamma(-p/R)}$ where p is the smallest index such that $c_p \neq 0$ and $p/R \notin \mathbb{N}$. If there are a finite number of dominant singularities, each of this form, then asymptotics of f_n are determined by adding the asymptotic contributions given by each.

Note that Proposition 2.11 only requires the function $F(z)$ to behave locally like an algebraic function near its dominant singularity (in the sense that it has a convergent Puiseux expansion).

Example 2.13 (Asymptotics of 2-Regular Simple Graphs)

A graph is 2-regular if every vertex has degree 2, and simple if the graph contains no loops or multiple edges between any pair of vertices. If g_n denotes the number of 2-regular simple graphs on n vertices where each vertex has a distinct label, then generating function arguments (found in Wilf [125], for instance) imply

$$F(z) = \sum_{n\geq 0} \frac{g_n}{n!} z^n = \frac{e^{-z/2-z^2/4}}{\sqrt{1 - z}}.$$

Because $F(z)$ has a unique dominant singularity at the point $z = 1$, where its Puiseux expansion begins $F(z) = e^{-3/4}(1-z)^{-1/2}+\cdots$, Proposition 2.11 immediately implies

$$\frac{g_n}{n!} \sim e^{-3/4} \frac{n^{-1/2}}{\Gamma(1/2)} = \frac{e^{-3/4}}{\sqrt{n\pi}}.$$

This example is also studied by Flajolet and Odlyzko [48].

Because every algebraic function satisfies the hypotheses of Proposition 2.11, this gives a strong characterization of their asymptotic behaviour.

Corollary 2.1 *Let $F(z) = \sum_{n \geq 0} f_n z^n$ be a branch of an algebraic equation which is analytic at the origin and has dominant singularities $\omega_0\beta, \ldots, \omega_{m-1}\beta$ where $\beta > 0$ and $|\omega_j| = 1$ for each j. Then*

$$f_n = \frac{\beta^n n^s}{\Gamma(s+1)} \sum_{j=0}^{m-1} C_j \omega_j^n + O(\beta^n n^t),$$

where $s \in \mathbb{Q} \setminus \{-1, -2, \ldots\}$, $t < s$, and β, the ω_j, and the C_j are algebraic.

Checking for incompatible asymptotic behaviour of coefficients, or the existence of an infinite number of singularities, gives a powerful test of function transcendence which is not available for constants.

Example 2.14 (A Transcendental Series)

If $F(z) = (1-4z)^{-1/2}$ then a straightforward application of the generalized binomial theorem implies $F(z) = \sum_{n \geq 0} \binom{2n}{n} z^n$ and Proposition 2.11 shows $\binom{2n}{n} \sim 4^n/\sqrt{\pi n}$. The function $G(z) = \sum_{n \geq 0} \binom{2n}{n}^2 z^n$ is also analytic at the origin, but

$$\binom{2n}{n}^2 \sim \frac{16^n}{\pi n}$$

so $G(z)$ is not algebraic by Corollary 2.1. See Problem 2.10 for a generalization.

Automated Asymptotics: A Solvable Connection Problem

Let $P(z, y) \in \mathbb{Q}[z, y]$ and $F(z) = \sum_{n \geq 0} f_n z^n$ be a branch of P which is analytic at the origin, specified by enough power series coefficients to distinguish it from the other branches of P. In order to determine dominant asymptotics of f_n, we must identify the dominant singularities of F and find its local behaviour near each to apply Proposition 2.11. Any singularity of F lies in the (finite) exceptional set Ξ, but we need to be able to determine which elements of Ξ are singularities of F (and not singularities of other branches or spurious points where all branches are analytic).

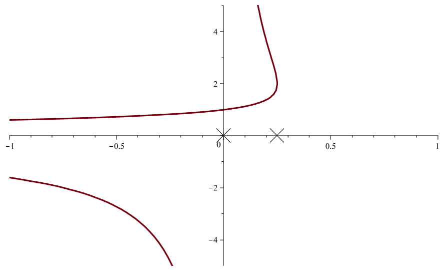


Fig. 2.4 Plot of the real solutions to $zy^2 - y + 1 = 0$ with the values of Ξ marked. At the origin one branch of the solution goes to infinity, while at $z = 1/4$ the two branches collide.

Given $a \in \Xi$ we can use the Newton polygon method to compute series solutions to $P(z, y) = 0$ which are analytic in a neighbourhood of a, potentially minus a or a ray starting at a, and then identify the singular branches at a by finding those with non-integer or negative exponents. Since we know the series expansion of the generating function F at the origin, we are thus trying to solve a *connection problem*: given series expansions of branches of $P(z, y) = 0$ around two distinct points, can we pair up the series in such a way as to be consistent among branches? Since series expansions capture local behaviour of a function, being able to solve this connection problem means capturing global behaviour about branches of P from their local behaviour.

Example 2.15 (Catalan Asymptotics)

Recall again the generating function for the number of rooted binary plane trees, which is the unique power series root of

$$P(z, y) = zy^2 - y + 1$$

at the origin. Near the origin the branches of $P(z, y)$ are represented by the expansions

$$f_1(z) = C(z) = 1 + z + 2z^2 + 5z^3 + \cdots \qquad f_2(z) = z^{-1} - 1 - z - 2z^2 + \cdots$$

while around the point $z = 1/4$ the branches are represented by the expansions

$$g_1(z) = 2 + 2\,(1 - 4z)^{1/2} + \cdots \qquad g_2(z) = 2 - 2\,(1 - 4z)^{1/2} + \cdots .$$

Thus $C(z)$ must have a singularity at $z = 1/4$, but is it locally represented by g_1 or g_2? If $C(z)$ was represented by g_1 near $z = 1/4$ then Proposition 2.11 would imply

$$c_n \sim (2)\,4^n \frac{n^{-3/2}}{\Gamma(-1/2)} = -\frac{4^n}{n^{3/2}\sqrt{\pi}},$$

which is impossible as c_n has non-negative terms. Thus, $C(z)$ corresponds to g_2 and

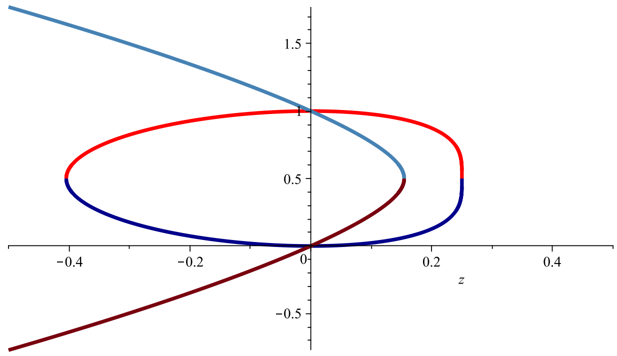


Fig. 2.5 Real parts of branches of the algebraic curve $P(z, y) = y^4 - 2y^3 + (1+2z)y^2 - 2yz + 4z^3 = 0$. The discriminant of P vanishes at $z = 1/4, 0$, and $z = (-1 \pm \sqrt{5})/8$, corresponding to the intersection of at least two of the branches.

$$c_n = (-2)\,4^n \left(\frac{n^{-3/2}}{\Gamma(-1/2)} + O\left(\frac{1}{n}\right) \right) = 4^n \left(\frac{1}{n^{3/2}\sqrt{\pi}} + O\left(\frac{1}{n}\right) \right).$$

In fact, the quadratic formula implies $C(z) = \frac{1-\sqrt{1-4z}}{2z}$, so

$$c_n = (-1/2)[z^{n+1}](1-4z)^{1/2} = \frac{4^n}{\sqrt{\pi}} \frac{\Gamma(n+1/2)}{\Gamma(n+2)},$$

which is simply a restatement of the well known formula $c_n = \frac{1}{n+1}\binom{2n}{n}$.

Such heuristic arguments do not always apply, and more involved techniques are usually required.

Example 2.16 (The Connection Problem for Supertrees)

The generating function $F(z)$ of the combinatorial class of *bicoloured supertrees* is the branch $y = F(z)$ of the algebraic curve

$$P(z, y) = y^4 - 2y^3 + (1 + 2z)y^2 - 2yz + 4z^3$$

whose power series expansion at the origin begins $F(z) = 2z^2 + 2z^3 + \cdots$ (see Flajolet and Sedgewick [50, Ex. VI.10] for details). Figure 2.5 shows the real parts of the branches defined by $P(z, y) = 0$. The set of potential singularities of $F(z)$,

$$\Xi = \left\{0, 1/4, \left(-1 \pm \sqrt{5}\right)/8\right\},$$

is defined by the vanishing of the discriminant. Using only local information about F at the origin (its initial power series terms look like a quadratic function) we can identify F with the dark curved branch in Figure 2.5. Tracing this branch from the origin across the positive real numbers shows that $F(z)$ admits a single dominant singularity at $z = 1/4$.

Tracing along curves can be formally captured by algorithms for real numeric continuation, but we now outline a different approach for combinatorial generating functions which is simpler and more efficient. Because $F(z)$ has non-negative power series coefficients, the Vivanti-Pringsheim theorem implies one of its dominant singularities is real and positive. At the origin, the branches have power series expansions

$$A_1(z) = 1 - 2z^2 + \cdots \qquad A_2(z) = 1 - 2z + \cdots$$
$$A_3(z) = 2z + 2z^2 + \cdots \qquad A_4(z) = F(z) = 2z^2 + 2z^3 + \cdots .$$

Note that all branches are analytic even though the branches intersect: intersection is a necessary condition for singularities of branches not going to infinity, but not a sufficient one. At the first potential positive singularity, $\rho = (\sqrt{5}-1)/8$, the branches have Puiseux expansions

$$B_1(z) = \frac{1}{2} + \frac{\sqrt{5-\sqrt{5}}}{4}\sqrt{1-z/\rho} + \cdots, \quad B_2(z) = \frac{2+\sqrt{10}-\sqrt{2}}{4} - \frac{\sqrt{10}-4\sqrt{2}}{16}(1-z/\rho) + \cdots$$
$$B_3(z) = \frac{1}{2} - \frac{\sqrt{5-\sqrt{5}}}{4}\sqrt{1-z/\rho} + \cdots, \quad B_4(z) = \frac{2-\sqrt{10}+\sqrt{2}}{4} + \frac{\sqrt{10}-4\sqrt{2}}{16}(1-z/\rho) + \cdots,$$

while at $z = 1/4$ they have Puiseux expansions

$$C_1(z) = \frac{1}{2} + \frac{1}{2}(1-4z)^{1/4} + \cdots \qquad C_2(z) = \frac{1}{2} - \frac{1}{2}(1-4z)^{1/4} + \cdots$$
$$C_3(z) = \frac{1}{2} + \frac{i}{2}(1-4z)^{1/4} + \cdots \qquad C_4(z) = \frac{1}{2} - \frac{i}{2}(1-4z)^{1/4} + \cdots$$

Our goal is to determine which of the expansions $B_j(z)$ and $C_k(z)$ correspond to F. In particular, we want to determine whether F corresponds to one of the expansions B_2 and B_3 which are singular at $z = \rho$, or if the dominant singularity of F is $z = 1/4$.

The key is to use the fact that the real branches of $P(z, y)$ can be sorted at real points, and their relative orders do not change between elements of Ξ. To begin, we note that if $r > 0$ is sufficiently small then looking at the initial terms of the series expansions of the A_j shows

$$A_1(r) > A_2(r) > A_3(r) > A_4(r).$$

Similarly, if $s < \rho$ is sufficiently close to ρ then looking at the initial terms of the series expansions of the B_j shows

$$B_2(s) > B_1(s) > B_3(s) > B_4(s).$$

Because the real branches of P can only cross at elements of Ξ, the relative orders of the branches at $z = r$ and $z = s$ must be equal. Since $A_1(r)$ and $B_2(s)$ are largest in the orderings, $A_1(z)$ and $B_2(z)$ are expansions of the same branch; the rest of the branches pair up as (A_2, B_1), (A_3, B_3), and (A_4, B_4). In particular, as the series expansion of F around $z = 0$ is A_4 its series expansion near $z = \rho$ is the power series B_4, and $z = \rho$ is not a singularity of $F(z)$.

Continuing this process, we take $t > \rho$ to be sufficiently close to ρ. Only the branches $B_2(z)$ and $B_4(z)$ are real for $z > \rho$, and examining their initial series terms shows

$$B_2(t) > B_4(t).$$

If $u < 1/4$ is sufficiently close to $1/4$ then examining the initial series terms of the $C_j(z)$ shows that only $C_1(u)$ and $C_2(u)$ are real, and that

$$C_1(u) > C_2(u).$$

Since $F(z)$ is represented by $B_4(z)$ near $z = \rho$, and the relative orders of the real branches do not change between elements of Ξ, we have determined that $z = 1/4$ is the unique dominant singularity of $F(z)$, where it has the convergent expansion

$$F(z) = C_2(z) = \frac{1}{2} - \frac{1}{2}(1 - 4z)^{1/4} + \cdots$$

in a slit disk. Proposition 2.11 then implies the number f_n of bicoloured supertrees of size n has dominant asymptotics

$$f_n = 4^n\, n^{-5/4}\, \frac{-1/2}{\Gamma(-1/4)} + O(4^n n^{-3/2}) = \frac{4^n}{n^{5/4}\, 8\Gamma(3/4)} + O(4^n n^{-3/2}).$$

An analysis of bicoloured supertrees using this approach was carried out in the thesis of Chabaud [21]. DeVries [36] gave an alternative analysis using the multivariate machinery discussed in Chapters 3 and 5.

This *branch sorting algorithm* can be applied to any algebraic function $F(z)$ with non-negative power series coefficients at the origin. As in the example, one sorts the Puiseux series solutions of an annihilating polynomial $P(z, y) = 0$ which take real values between positive elements of the exceptional set Ξ, and uses the fact that their relative orders don't change between potential singularities to link local behaviour of expansions at one point to those at another. Proposition 2.9 allows us to determine which branches are real, and sorting can be done through a lexicographical ranking of series by their coefficients since the lower order terms in a Puiseux series dominate local asymptotic behaviour.

Example 2.17 (Kreweras Lattice Walks)

Let $A(z)$ be the generating function for the number a_n of lattice walks beginning at the origin, staying in $\mathbb{N}^2$, and taking n steps in $\mathcal{S} = \{(-1,0),(0,-1),(1,1)\}$. This lattice path model was studied by Kreweras [74] in the 1960s, who gave a simple formula for walks ending at the origin and a more complicated expression for walks ending anywhere in $\mathbb{N}^2$. Gessel [55] proved that $A(z)$ is algebraic (in fact, he proved the stronger result that the trivariate generating function counting walks by length and x and y endpoint is algebraic). Bousquet-Mélou [17] used the kernel method, which we study in detail in Chapter 4, to derive algebraic equations and enumerate these walks. In particular, the generating function $A(z)$ satisfies $P(z, A(z)) = 0$ where

$$\begin{aligned} P(z,y) &= z^5(3z-1)^3y^6 + 6z^4(3z-1)^3y^5 + z^3(3z-1)(135z^2-78z+14)y^4 \\ &\quad + 4z^2(3z-1)(45z^2-18z+4)y^3 + z(3z-1)(135z^2-26z+9)y^2 \\ &\quad + 2(3z-1)(27z^2-2z+1)y + 43z^2 + z + 2. \end{aligned}$$

Computing the discriminant shows that the exceptional set of P consists of $0, -1, 1/3$, and the two non-real roots of the polynomial $9z^2+3z+1$, so $A(z)$ has a unique dominant singularity at $z = 1/3$. The Newton polygon method shows that there are two Puiseux series solutions of $P(z,y) = 0$ at the origin which have real coefficients,

$$-2z^{-1} - 1 - z - 3z^2 + \cdots \qquad \text{and} \qquad 1 + z + 3z^2 + 7z^3 + \cdots,$$

and two real Puiseux series solutions at $z = 1/3$ which have real coefficients,

$$2\sqrt{2}(1-3z)^{-1/4} - 3 + \cdots \qquad \text{and} \qquad -2\sqrt{2}(1-3z)^{-1/4} - 3 + \cdots.$$

Since $A(z)$ is the real branch of $P(z,y) = 0$ which is larger just to the right of the origin, it is the larger real branch just to the left of $z = 1/3$, and the expansion of $A(z)$ in a slit disk at $z = 1/3$ begins with $2\sqrt{2}(1-3z)^{-1/4}$. Proposition 2.11 then implies $a_n \sim 3^n n^{-3/4}(2\sqrt{2})/\Gamma(1/4)$.

More generally, for series with negative coefficients, rigorous numerical methods can be used. There exist bounds (discussed in the appendix to Chapter 7) on how close the values of the branches to $P(z,y) = 0$ can be at a point $z = a$, depending only on P and a. Since the series expansion of $F(z)$ at the origin is convergent whenever z has smaller modulus than its dominant singularities, one can iterate through the elements of Ξ in order of increasing modulus and numerically approximate F and the branches near each potential singularity to a sufficient accuracy to decide which branch corresponds to F. Full details of both approaches can be found in the PhD thesis of Chabaud [21]. The Sage package of Mezzarobba for numeric analytic continuation of D-finite functions, discussed in Section 2.4 below, can be used for the necessary numeric computations.

Once the dominant singularities of an algebraic generating function are determined, an asymptotic expression for its coefficients can be computed using Proposition 2.11. The decidability issues related to Skolem's Problem for dominant coefficient asymptotics of rational generating functions are still in play for algebraic functions, but such pathological considerations do not typically appear in combinatorial applications.

Examples of Algebraic Series and $\mathbb{N}$-Algebraicity

As briefly seen in the above examples, combinatorial structures which admit a recursive decomposition based on collections of smaller objects *of the same type* are those that tend to have algebraic generating functions. There are a wealth of combinatorial applications, including the enumeration of many types of lattice paths, pattern-avoiding permutations, trees, planar maps on surfaces, polyominoes, dissections of polygons, sequences arising in bioinformatics, and tiling problems. Algebraic functions over a field of characteristic p have connections to p-automatic sequences in theoretical computer science and the combinatorics of words. These examples, and more, can be found in Stanley [110], Bousquet-Mélou [18], and Banderier and Drmota [4]. Algebraic functions also have many closure properties.

Proposition 2.12 *The set of algebraic series forms a field: the sum, difference, product, and quotient (when the denominator is non-zero) of two algebraic series $a(x)$ and $b(x)$ is algebraic. Annihilating polynomials for $a(x)+b(x)$, $a(x)-b(x)$, $a(x)b(x)$, and $a(x)/b(x)$ can be computed from annihilating polynomials for $a(x)$ and $b(x)$.*

Problem 2.15 guides the reader through a proof of Proposition 2.12.

Using resultants to eliminate variables, any component of a solution of a non-trivial algebraic system of equations is also algebraic.

Definition 2.17 ($\mathbb{N}$-algebraic series) A *proper $\mathbb{N}$-algebraic system* is a polynomial system of the form

$$\begin{cases} y_1 &= P_1(z, y_1, \dots, y_d) \\ & \vdots \\ y_d &= P_d(z, y_1, \dots, y_d) \end{cases}$$

where each P_i has natural number coefficients, no constant term, and $[y_i]P_i = 0$. A univariate series $y(z)$ is $\mathbb{N}$-*algebraic* if there exists a constant $c_0 \in \mathbb{N}$ and power series solution $(y_1, \dots, y_d) \in \mathbb{N}[[z]]^d$ of some proper $\mathbb{N}$-algebraic system such that $y(z) = c_0 + y_1(z)$.

Just as $\mathbb{N}$-rational series appear as generating functions when counting words in regular languages, $\mathbb{N}$-algebraic series appear in the study of context-free languages, a superset of regular languages which are recognized by 'pushdown' automata. Chomsky and Schützenberger [23] proved that the generating function counting words in any non-ambiguous context-free language by length is $\mathbb{N}$-algebraic, and conversely any $\mathbb{N}$-algebraic generating function counts the number of words in some non-ambiguous context-free language by length. Flajolet [46] used this enumerative approach to easily prove the ambiguity of several families of context-free languages, a typically hard task. See Chapter 3 of [95] for more details on these topics.

Traces of the modern Symbolic Method, which allows one to convert a combinatorial description of a family of objects into a generating function specification, can be found in the *Delest-Viennot-Schützenberger (DVS) methodology* [34]. The DVS methodology, dating back to Schützenberger's work on formal language theory, states roughly that to enumerate a combinatorial class one should try to: construct a bijection to a context-free language, determine a context-free grammar generating that language, then deduce an algebraic equation for the number of objects under consideration using the rules of the grammar. This approach and the techniques it has inspired have been extremely influential in enumerative combinatorics.

Banderier and Drmota [4] classify the dominant asymptotic behaviour of $\mathbb{N}$-algebraic series, giving a tool for disproving $\mathbb{N}$-algebraicity of generating functions. Subdominant asymptotic terms can be harder to get a handle on, a consequence of the fact that every algebraic power series with integer coefficients is the difference of two $\mathbb{N}$-algebraic series [4, Prop. 2.6]. Currently, a complete classification of $\mathbb{N}$-algebraicity is not known.

Open Problem 2.3 (Decidability of $\mathbb{N}$-Algebraicity) *Is there an algorithm which takes an algebraic series with natural number coefficients, specified by an annihilating polynomial and initial terms, and determines whether it is $\mathbb{N}$-algebraic?*

2.4 D-Finite Power Series

We now turn to a class of generating functions encoded by linear differential equations over a field $\mathbb{K}$ of characteristic zero. Just as polynomials are data structures for algebraic series, differential equations are data structures for these series.

Definition 2.18 (D-finite series and functions) The *derivative* of a formal power series $F(z) = \sum_{n\geq 0} f_n z^n$ is the formal series $\frac{d}{dz}F(z) = \sum_{n\geq 1}(nf_n)z^{n-1}$. We call $F(z)$ *differentially finite (D-finite)* over $\mathbb{K}$ if there exist $a_0(z), \dots, a_r(z) \in \mathbb{K}[z]$ with $a_r(z) \neq 0$ such that

$$a_r(z)\frac{d^r}{dz^r}F(z) + a_{r-1}(z)\frac{d^{r-1}}{dz^{r-1}}F(z) + \cdots + a_0(z)F(z) = 0. \tag{2.7}$$

If $\mathbb{K} \subset \mathbb{C}$ then an analytic function $F(z)$ is called *D-finite* if it satisfies an equation of the form (2.7) when defined. If d is the maximum degree of the coefficient polynomials $a_j(z)$ then we call (2.7) a *D-finite equation of order r and degree d*. Equivalently, $F(z)$ is D-finite if and only if all derivatives $F, F', F'', F''', \dots$ form a finite dimensional vector space over the field of rational functions $\mathbb{K}(z)$. We often write $F^{(r)}$ for the rth derivative $\frac{d^r}{dz^r}F(z)$.

Example 2.18 (D-finite Transcendental Function)

The function $F(z) = e^z = \sum_{n\geq 0} z^n/n!$ is not algebraic (for instance, because its asymptotic growth does not satisfy the conditions of Corollary 2.1) but satisfies $F'(z) - F(z) = 0$, so F is D-finite.

Remark 2.2 Suppose $F(z)$ satisfies a non-homogeneous D-finite equation of the form $\mathcal{L}(F) = c(z)$, where $\mathcal{L}(F) = a_r(z)F^{(r)}(z) + \cdots + a_0(z)F(z)$ and $c(z) \in \mathbb{K}[z]$. Then $c(z)\frac{d}{dz}\mathcal{L}(F) - c'(z)\mathcal{L}(F) = 0$, so any function satisfying a non-homogeneous D-finite equation of order r also satisfies a homogeneous equation of order $r + 1$.

D-finite equations are often manipulated in computer algebra systems by encoding them as elements of a non-commutative ring.

Definition 2.19 (Weyl algebra) The *Weyl algebra* $\mathcal{W}$ over $\mathbb{K}$ is the $\mathbb{K}$-algebra $\mathbb{K}\langle z, \delta\rangle/\langle \delta z - z\delta - 1\rangle$, consisting of $\mathbb{K}$-linear combinations of monomials containing powers of z and δ such that z and δ don't commute but satisfy $\delta z = z\delta + 1$. We let the elements of $\mathcal{W}$ act on $\mathbb{K}[[z]]$ by defining $z \cdot F = zF$ and $\delta \cdot F = F'$; the commutation rule for z and δ was chosen to make this action well defined, since

$$(\delta z) \cdot F(z) = \delta \cdot (zF(z)) = \frac{d}{dz}(zF(z)) = zF'(z) + F(z) = (1 + z\delta) \cdot F(z).$$

Any element $P \in \mathcal{W}$ can be uniquely written $P = p_r(z)\delta^r + p_{r-1}(z)\delta^{r-1} + \cdots + p_0(z)$ for polynomials $p_j(z) \in \mathbb{K}[z]$, meaning

$$P \cdot F(z) = p_r(z)\frac{d^r}{dz^r}F(z) + p_{r-1}(z)\frac{d^{r-1}}{dz^{r-1}}F(z) + \cdots + p_0(z)F(z).$$

D-finite equations over $\mathbb{K}$ are thus in bijection with the elements of $\mathcal{W}$.

Coefficient Properties

The coefficient sequence of a D-finite function satisfies a linear recurrence relation with polynomial coefficients, generalizing the C-finiteness of rational function coefficients.

Definition 2.20 (P-recursive sequences) A sequence (f_n) is *polynomially recursive (P-recursive)* if there exist polynomials $c_0(n), \ldots, c_r(n) \in \mathbb{K}[n]$ such that $c_0(n), c_r(n) \neq 0$ and

$$c_r(n)f_{n+r} + c_{r-1}(n)f_{n+r-1} + \cdots + c_0(n)f_n = 0 \tag{2.8}$$

for all $n \geq 0$. If the coefficient polynomials $c_j(n)$ have highest degree d, then we call (2.8) a *P-recursive relation of order r and degree d*. As in the differential case above, if (f_n) satisfies a non-homogeneous P-recursive relation of order r then it also satisfies a homogeneous P-recursive relation of order $r + 1$.

Definition 2.21 (shift algebra) The *shift algebra* $\mathcal{S}$ over $\mathbb{K}$ is the $\mathbb{K}$-algebra $\mathbb{K}\langle n, S\rangle/\langle Sn-(n+1)S\rangle$, consisting of $\mathbb{K}$-linear combinations of monomials containing powers of n and S, such that n and S don't commute but satisfy $Sn=(n+1)S$. We let the elements of $\mathcal{S}$ act on the set of sequences over $\mathbb{K}$ by defining $n\cdot(f_n)=(nf_n)$ and $S\cdot(f_n)=(f_{n+1})$; the commutation rule for n and S was chosen to make this action well defined, since

$$(Sn)\cdot(f_n)=S\cdot(nf_n)=((n+1)f_{n+1})=((n+1)S)\cdot(f_n).$$

Any element $Q\in\mathcal{S}$ can be uniquely written $Q=q_r(n)S^r+q_{r-1}(n)S^{r-1}+\cdots+q_0(n)$ for polynomials $q_j(n)\in\mathbb{K}[n]$, meaning

$$Q\cdot(f_n)=(q_r(n)f_{n+r}+q_{r-1}(n)f_{n+r-1}+\cdots+q_0(n)f_n)\,.$$

P-recursive relations over $\mathbb{K}$ are thus in bijection with the elements of $\mathcal{S}$.

Both the Weyl algebra and shift algebra are examples of Ore algebras, and are implemented in the Sage ore_algebra package [69].

Proposition 2.13 *Let* $F(z)=\sum_{n\geq 0}f_nz^n\in\mathbb{K}[[z]]$.

1. *If* $F(z)$ *satisfies a D-finite equation of order* r *and degree* d *then* (f_n) *satisfies a P-recursive relation of order at most* $r+d$ *and degree at most* r.
2. *If* (f_n) *satisfies a P-recursive relation of order* r *and degree* d *then* $F(z)$ *satisfies a D-finite equation of order at most* d *and degree at most* $r+d$.

Proof Since $[z^n]F^{(r)}(z)=(n+r)(n+r-1)\cdots(n+1)f_{n+r}$ for $r\in\mathbb{N}$, and

$$[z^n]z^kF(z)=\begin{cases}f_{n-k} & : 0\leq k\leq n\\ 0 & : k>n\end{cases}$$

for $k\in\mathbb{N}$, if $F(z)$ satisfies a differential equation of the form (2.7) then its coefficients satisfy a linear recurrence with polynomial coefficients, where derivatives of F can increase the coefficient degree of the recurrence and both derivatives and multiplications by z can increase the order of the recurrence.

Conversely, let $\partial_z=z\frac{d}{dz}$ be the operator that differentiates by z and then multiplies the result by z, and for any $k\in\mathbb{N}$ let $F_{\leq k}(z)$ be the polynomial $F_{\leq k}(z)=f_0+f_1z+\cdots+f_kz^k$. Then

$$\sum_{n\geq 0}f_{n+k}z^n=\frac{F(z)-F_{\leq k}(z)}{z^k}$$

and repeated differentiation gives

$$\sum_{n\geq 0}\left(n^jf_{n+k}\right)z^n=\partial_z^j\left(\frac{F(z)-F_{\leq k}(z)}{z^k}\right).$$

Thus, if (f_n) satisfies a linear recurrence of the form (2.8) then multiplying by z^n and summing over n gives a linear differential equation satisfied by $F(z)$. Problem 2.11 asks you to prove the stated degree and order bounds. □

Note that it is automatic to translate between D-finite equations satisfied by a series and P-recursive relations satisfied by its coefficient sequence. This is implemented, for example, in the Maple gfun package and the Sage ore_algebra package.

Example 2.19 (Central Binomial Coefficients Squared)

Let $f_n = \binom{2n}{n}^2$ and recall that the asymptotic growth of (f_n) implies the generating function $F(z) = \sum_{n\geq 0} f_n z^n$ is not algebraic. Using Pascal's identity, that $\binom{a-1}{b} + \binom{a-1}{b-1} = \binom{a}{b}$ for all positive integers a and b, light algebraic manipulation shows

$$(n+1)^2 f_{n+1} - 4(2n+1)^2 f_n = 0$$

for all $n \geq 0$. Thus, $F(z)$ is D-finite. Following the proof of Proposition 2.13 translates this to the differential equation

$$z\,(1-16z)\,\frac{d^2}{dz^2}F(z) + (1-32z)\frac{d}{dz}F(z) - 4F(z) = 0.$$

The generating function $F(z)$ can be written in terms of an elliptic integral by solving this differential equation, providing a second path to proving transcendence.

Recall from Section 2.2 that the solutions of a linear recurrence of order r with *constant coefficients* form a $\mathbb{K}$-vector of dimension r (adding two solutions or multiplying by a constant yields another solution). Because (2.8) is still a linear recurrence, its solutions still form a $\mathbb{K}$-vector space, but when the leading coefficient $c_r(n)$ vanishes at non-negative integers the dimension of this vector space may be *greater* than the order r of the recurrence. To account for this, let B be the set of non-negative integer solutions to $c_r(n) = 0$ and

$$I = \{0, \dots, r-1\} \cup \{b + r : b \in B\}.$$

If (f_n) is a solution of the linear recurrence (2.8) then the values of f_n for $n \in I$ uniquely determine (f_n). Furthermore, if $N = \max I$ and $f_0, \dots, f_N$ satisfy (2.8) for $n \leq N - r$ then this finite list always extends to a unique infinite sequence satisfying (2.8). Thus, the finite sequences which satisfy (2.8) for $n \leq N - r$ form a vector space Λ_N whose dimension can be computed by linear algebra and equals the dimension of the vector space of all solutions to (2.8).

Definition 2.22 (generalized initial conditions) The values of f_n when $n \in I$ are *generalized initial conditions* for a sequence (f_n) satisfying (2.8).

Remark 2.3 By Proposition 7.7 in Chapter 7, the elements of B are at most one larger than the maximum absolute value of the coefficients of $c_r(n)$.

Example 2.20 (A Large Solution Space for a Linear Recurrence)

Consider the linear recurrence relation $n(n-2)f_{n+2} - f_{n+1} - (n-1)f_n = 0$ over $\mathbb{K} = \mathbb{Q}$. Here the order $r = 2$ and $B = \{0, 2\}$ so $I = \{0, 1\} \cup \{0 + 2, 2 + 2\} = \{0, 1, 2, 4\}$ with maximum element $N = 4$. In order for $(f_0, \ldots, f_4)$ to be the start of a sequence satisfying this recurrence it must be that

$$\begin{aligned} -f_1 + f_0 &= 0 && (n = 0) \\ -f_3 - f_2 &= 0 && (n = 1) \\ -f_3 - f_2 &= 0 && (n = 2 = N - r) \end{aligned}$$

meaning

$$\underbrace{\begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & -1 & 0 \\ 0 & 0 & -1 & -1 & 0 \end{pmatrix}}_{M} \begin{pmatrix} f_0 \\ f_1 \\ f_2 \\ f_3 \\ f_4 \end{pmatrix} = 0.$$

Since M has rank 2, the rank-nullity theorem implies the vector space of solutions to this matrix equation has dimension 3. A sequence (f_n) satisfying this recurrence is uniquely determined by f_0, f_2, and f_4, and any assignment of values to these terms can be extended to a full sequence by setting $f_1 = f_0$ and

$$f_{n+2} = \frac{f_{n+1} + (n-1)f_n}{n(n-2)}$$

for all $n \notin \{0, 2\}$. Thus, the solution space of this recurrence forms a $\mathbb{Q}$-vector space of dimension 3 (even though the recurrence has order 2). Any assignment $f_0 = f_1 = a$, $f_2 = b$, and $f_4 = c$ for $(a, b, c) \in \mathbb{Q}$ forms a generalized initial condition.

Bostan et al. [14, Ch. 15] study the complexity of generating terms in P-recursive sequences.

Examples of D-finite Functions

Many common functions are easily seen to be D-finite, including exponentials, logarithms, (arc-)sine and (arc-)cosine functions, generalized hypergeometric functions (including common special functions like the Bessel and Airy functions), rational period integrals, and more. Furthermore, the class of D-finite functions contains all algebraic functions.

Proposition 2.14 *If $f(z)$ is an algebraic function of degree d over a field $\mathbb{K}$ of characteristic zero then $f(z)$ is D-finite over $\mathbb{K}$, and is annihilated by a D-finite equation of order at most d.*

Proposition 2.14 has been rediscovered by several authors going back to the 19th century, including Abel, Tannery, Cockle, Harley, and Comtet. Chudnovsky and

Chudnovsky [25] and Bostan et al. [15] have studied this result from a complexity viewpoint, giving explicit bounds on the orders and coefficient degrees of annihilating D-finite equations associated to algebraic functions. In particular, for efficiency reasons it can be desirable to take an annihilating D-finite equation of non-minimal order so as to greatly reduce the degrees of the polynomial coefficients involved.

Proof Let $P(z, y)$ be the minimal polynomial of f and write $P_z(z, y)$ for $(\partial P/\partial z)(z, y)$ and $P_y(z, y)$ for $(\partial P/\partial y)(z, y)$. Differentiating the equation $0 = P(z, f(z))$ with respect to z implies

$$0 = P_z(z, f(z)) + f'(z)P_y(z, f(z)).$$

Because $\mathbb{K}$ has characteristic zero and P_y is a polynomial in y of degree smaller than P, which is the minimal polynomial of $f(z)$, it follows that $P_y(z, f(z)) \neq 0$ and

$$f'(z) = -\frac{P_z(z, f(z))}{P_y(z, f(z))}.$$

If $\mathcal{R} = \mathbb{K}(z)$ denotes the field of rational functions over $\mathbb{K}$, this implies f' lies in the algebraic field extension $\mathcal{R}(f)$ of degree d over $\mathcal{R}$. By induction the $d + 1$ elements $f, f', \dots, f^{(d)}$ all lie in $\mathcal{R}(f)$, meaning they are linearly dependent over $\mathcal{R}$. □

Conversely, it is an interesting problem to prove or disprove algebraicity of a D-finite function specified by a known annihilating D-finite equation and initial conditions. Singer [105] gives an algorithm for deciding when all solutions of a D-finite equation are algebraic, which can be modified to solve this problem, but a need for factorization in non-commutative rings leads to an impractically large runtime. An alternative approach, arising in the context of lattice path enumeration, can be found in Bostan et al. [13].

The class of D-finite functions is also closed under many natural operations.

Proposition 2.15 *Let $F(z) = \sum_{n\geq 0} f_n z^n$ and $G(z) = \sum_{n\geq 0} g_n z^n$ satisfy D-finite equations of orders r and s, and let $H(z)$ be differentiable. Then, when defined,*

(1) The product $F(z)G(z)$ satisfies a D-finite equation of order at most rs;
(2) The sum $F(z) + G(z)$ satisfies a D-finite equation of order at most $r + s$;
(3) $\frac{dF}{dz}(z)$ and $\int F(z)dz$ are D-finite;
(4) If $G(z)$ is algebraic then $F(G(z))$ is D-finite;
(5) The Hadamard product $(F \odot G)(z) = \sum_{n\geq 0}(f_n g_n)z^n$ is D-finite;
(6) H and $1/H$ are both D-finite if and only if H'/H is algebraic.

Proof Item (1) follows from the fact that the products of derivatives $F^{(i)}(z)G^{(j)}(z)$ span the $\mathbb{K}(z)$-vector space of the derivatives of $F(z)G(z)$, while (2) follows from linearity of the derivative. Item (3) follows directly from the definition of D-finiteness and Item (4) follows from the chain rule and basic field theory [110, Thm. 6.4.10]. Item (5) follows from P-recursiveness of the coefficient sequences and an argument analogous to Item (1). Item (6) was given by Harris and Sibuya [64]. □

These closure properties are effective: there are algorithms which take annihilating D-finite equations for F and G and return annihilating equations for their sum, product, composition, and Hadamard product, among other operations. In comparison to Item (4), Singer [106] showed that if $G(z)$ is algebraic of genus at least 1 then $G(F(z))$ is D-finite if and only if $F(z)$ is also algebraic. The Pólya-Carlson theorem [92] implies that a D-finite series with integer coefficients and radius of convergence 1 is in fact rational.

P-recursive sequences, and thus D-finite functions, are central to enumerative combinatorics. In addition to combinatorial classes with rational or algebraic generating functions, transcendental D-finite functions appear in the study of most major combinatorial objects including permutations[9], graphs[10], tableau [56], and, as discussed in detail here, lattice path enumeration. Stanley [110] and Flajolet and Sedgewick [50, Sect. VII.9] contain a large number of additional examples. Salvy [102] gives an excellent survey of D-finite functions from a computer algebra perspective.

Automated Identity Proving and Creative Telescoping

Identities involving P-recursive sequences can be proven by verifying that they hold for a sufficient (finite) number of terms: to prove a P-recursive sequence (c_n) is identically zero it is sufficient to show that $c_k = 0$ when k lies in a finite set of generalized initial conditions. To prove that two P-recursive sequences (a_n) and (b_n) are identical, one can compute a linear recurrence satisfied by $c_n = a_n - b_n$ and verify that $c_n = 0$ for a sufficient number of terms. In fact, given linear recurrence relations for (a_n) and (b_n) one can automatically determine a bound N such that the sequences (a_n) and (b_n) are equal when their terms agree for all $0 \leq n \leq N$. Similarly, to prove two D-finite functions are equal it is sufficient to prove that a finite number of their derivatives (or power series coefficients) are equal at the origin.

Example 2.21 (Sums of Squares)

Suppose we want to prove the identity $\sum_{k=0}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$. If $a_n = \sum_{k=0}^{n} k^2$ then

$$a_{n+1} - a_n = (n+1)^2 \qquad \text{and} \qquad a_{n+2} - a_{n+1} = (n+2)^2,$$

so subtracting $(n+2)^2$ times the first equation from $(n+1)^2$ times the second equation gives a second order linear homogeneous recurrence

---

[9] Many well known families, such as Baxter permutations and permutations with bounded cycle length, have D-finite generating functions. In the 1990s, Noonan and Zeilberger [86] conjectured that the generating function of permutations avoiding any fixed set of patterns was D-finite (see that paper for the definition of pattern avoidance). This was recently shown to be false by Garrabrant and Pak [54], who proved non-D-finiteness for the generating function of permutations avoiding some set of patterns contained in S_{80}. There is some evidence [28] that the generating function of 1324-avoiding permutations may be non-D-finite, which is a large open problem in the area.

[10] D-finite generating functions appear in the enumeration of rooted planar maps of fixed genus [19], k-regular graphs [56] for fixed k, and similar problems.

$$(n+1)^2 a_{n+2} - (2n^2 + 6n + 5)a_{n+1} + (n+2)^2 a_n = 0 \tag{2.9}$$

satisfied by (a_n). Direct substitution shows that the sequence $b_n = n(n+1)(2n+1)/6$ satisfies the same recurrence, and it agrees with a_n for its first two terms, so the two sequences are equal for all n. In general (b_n) will also be specified by a linear recurrence, and one cannot simply substitute. For example, here b_{n+1}/b_n is a rational function so b_n satisfies the first order linear recurrence

$$n(n+1)(2n+1)b_{n+1} - (n+1)(n+2)(2n+3)b_n = 0. \tag{2.10}$$

This implies the sequence $c_n = a_n - b_n$ satisfies a recurrence of order at most $2+1 = 3$, meaning the four sequence shifts $c_{n+3}, c_{n+2}, c_{n+1}$, and c_n lie in a $\mathbb{Q}(n)$−vector space of dimension 3 and thus have a linear dependency. Writing out the c_j sequence in terms of a_j and b_j, then using (2.9) and (2.10) to eliminate all terms except for a_{n+1}, a_n, and b_n gives a system of three homogeneous linear equations in four variables. Computing a non-zero solution of this system shows that c_n satisfies the same second order recurrence,

$$(n+1)^2 c_{n+2} - (2n^2 + 6n + 5)c_{n+1} + (n+2)^2 c_n = 0,$$

as a_n. Note that this recurrence is satisfied when c_n is the difference of *any* two solutions[11] of (2.9) and (2.10): working on the level of recurrences allows for computations to be performed in $\mathbb{Q}(n)$. Since our specified sequences a_n and b_n satisfy $c_0 = a_0 - b_0 = 0$ and $c_1 = a_1 - b_1 = 0$ then $a_n = b_n$ for all n.

A powerful technique for proving identities is the *creative telescoping* framework, introduced by Zeilberger [128, 129], along with Wilf [126], in the 1980s and 1990s based on techniques of Fasenmyer [43] and Gosper [57]. Creative telescoping methods can (among other things) compute integrals of D-finite functions and sums of P-recursive sequences with free parameters by deducing differential equations or recurrence relations they satisfy. Proposed identities can then be verified, or discovered, using the arguments discussed above. The name 'creative telescoping,' apparently coined by van der Poorten [117] in his account of Apéry's proof of the irrationality of $\zeta(3)$, comes from the fact that these methods can be considered a (vast) generalization of the elementary technique by which certain sums can be manipulated to have summands which telescope.

Example 2.22 (A Parametrized Integral)

Consider the parametrized integral

---

[11] It can be shown that the general solution to (2.9) is $A + B\,b_n$ for constants A and B while the general solution to (2.10) is $C\,b_n$ for a constant C, and indeed $A + B\,b_n - C\,b_n = A + (B-C)\,b_n$ is always a solution of (2.9).

$$G(t) = \frac{1}{2\pi\mathrm{i}} \int_\gamma \frac{1}{x - x^2 - t}\, dx,$$

where γ is the positively oriented circle $\{|x| = 1/2\}$ and $|t| \le 1/5$, so $x - x^2 - t$ is never zero for x on γ. The methods of creative telescoping produce the identity

$$(4t-1)\frac{d}{dt}\left(\frac{1}{x - x^2 - t}\right) + \frac{2}{x - x^2 - t} = \frac{d}{dx}\left(\frac{1-2x}{x - x^2 - t}\right),$$

which can easily be verified by differentiation. Integrating this equation with respect to x gives

$$(4t-1)\int_\gamma \frac{d}{dt}\left(\frac{1}{x - x^2 - t}\right) dx + 2\int_\gamma \frac{1}{x - x^2 - t}\, dx = \int_\gamma \frac{d}{dx}\left(\frac{1-2x}{x^2 - x - t}\right) dx = 0,$$

since γ is a closed curve on which $x^2 - x - t$ is not zero. A minor argument shows that the derivative $\frac{d}{dt}$ can be moved outside of the integral with respect to x, so $G(t)$ is D-finite and satisfies $(4t-1)G'(t) + 2G(t) = 0$ with initial condition $G(0) = \frac{1}{2\pi\mathrm{i}}\int_\gamma \frac{dx}{x - x^2} = 1$. In fact, solving this separable differential equation shows $G(t) = (1-4t)^{-1/2}$ is the generating function of the central binomial coefficients. This computation is related to the multivariate diagonals discussed in Chapter 3.

Over the last three decades there has been a tremendous amount of research in the computer algebra community on creative telescoping, extending the framework to new classes of problems and improving efficiency through several 'generations' of algorithms. This work has been some of the most fruitful in modern computer algebra, leading to practical implementations which find application in a vast number of mathematical and scientific fields. A listing of such results is outside the scope of this text, but we briefly return to creative telescoping in Chapter 3 as it allows one to obtain an annihilating D-finite equation satisfied by the diagonal of a multivariate rational function. Detailed accounts of the methods and applications of creative telescoping can be found in recent surveys [91, 73, 27, 22].

Asymptotics of P-Recursive Sequences

Assume now that $\mathbb{K} = \mathbb{A}$ is the field of algebraic numbers. A classic theorem of Fabry [42] states that any D-finite equation over the field of algebraic numbers admits a basis of *formal* series solutions of the form

$$z^\alpha \exp\left(P\left(z^{-1/q}\right)\right) \sum_{j=0}^{d} \left(f_j\left(z^{1/q}\right) \log^j z\right),$$

where the f_j are power series, P is a polynomial, $q \in \mathbb{N}_{>0}$, α is an algebraic number, and $\log z$ is a formal term algebraically independent of z with formal derivative $\frac{d}{dz} \log z = 1/z$. Analogously, a P-recursive relation has a basis of formal series solutions of the form

$$\rho^n n^\alpha \Gamma(n)^\beta \exp\left(P\left(n^{1/q}\right)\right) \sum_{j=0}^{d} \left(f_j\left(n^{-1/q}\right) \log^j n\right),$$

where the f_j are power series, P is a polynomial, $q \in \mathbb{N}_{>0}$, ρ and α are algebraic numbers, and $\beta \in \mathbb{Q}$. Initial terms of these formal series solutions can be computed using the so-called Birkhoff-Trjitzinsky method [11, 12]. A nice summary of the Birkhoff-Trjitzinsky method can be found in Wimp and Zeilberger [127], including a discussion on the difference between formal series solutions and those representing asymptotically meaningful series; an implementation of the Birkhoff-Trjitzinsky method (used for our examples) is contained in the Sage ore_algebra package. We do not delve into the complicated topic of formal versus asymptotic series because, as we will soon see, when dealing with combinatorial generating functions which are analytic at the origin one can always work with D-finite equations and P-recursions whose formal solutions represent meaningful asymptotic expansions.

Example 2.23 (D-Finite and P-Recursive Bases of Solutions)

The number a_n of lattice walks beginning at the origin, staying in $\mathbb{N}^2$, and taking n steps in $\mathcal{S} = \{(\pm 1, 0), (0, \pm 1)\}$ satisfies the P-recursion

$$(n+4)(n+3)a_{n+2} - 4(2n+5)a_{n+1} - 16(n+1)(n+2)a_n = 0.$$

The solutions of this P-recurrence form a two-dimensional $\mathbb{C}$-vector space, and the generalized_series_solutions command of the Sage ore_algebra package determines a basis for this vector space consisting of two elements whose expansions begin

$$4^n\, n^{-1} \left(1 - \frac{3}{2n} + \cdots\right) \qquad \text{and} \qquad (-4)^n\, n^{-3} \left(1 - \frac{9}{2n} + \cdots\right).$$

The generating function $A(z)$ of this sequence satisfies the D-finite equation

$$\begin{aligned} z^2(4z-1)(4z+1)A'''(z) &+ 2z(4z+1)(16z-3)A''(t) \\ &+ 2(112z^2 + 14z - 3)A'(t) + 4(16z+3)A(t) = 0 \end{aligned}$$

whose solutions form a three-dimensional $\mathbb{C}(z)$ vector space. Applying the generalized_series_solutions command now gives three series expansions

$$1 + 2z + 6z^2 + 18z^3 + 60z^4 + \cdots$$
$$z^{-1}(1 + 2z + 4z^2 + 12z^3 + 36z^4 + \cdots)$$
$$z^{-2}(z + 2z^2 + 4z^3 + 12z^4 + \cdots)\log(z) - (1/4 + 3z/2 + 2z^2 + \cdots)$$

corresponding to the elements of a basis for the solution vector space.

Our asymptotic results follow from a careful study of how a D-finite equation determines the singularities of its D-finite function solutions.

Definition 2.23 (singularities of D-finite equations) A point $\zeta \in \mathbb{C}$ is an *ordinary point* of the D-finite equation (2.7) if there exist r linearly-independent solutions of (2.7) which are analytic at $z = \zeta$. A point which is not ordinary is a *singularity* of the D-finite equation (2.7). The *exceptional set* of (2.7) is $\Xi = \{z \in \mathbb{C} : a_r(z) = 0\}$.

A classic argument [124, Thm. 2.2] dating back to Cauchy implies any singularity of (2.7) must be a zero of the leading coefficient $a_r(z)$ and thus lie in the finite algebraic set Ξ, giving the following.

Lemma 2.3 *If $\zeta \notin \Xi$ then any solution of* (2.7) *which is analytic in a slit disk centred at ζ is in fact analytic at ζ.*

To further classify the (potential) singularities of (2.7), we rewrite the equation as

$$\frac{d^r}{dz^r}F(z) + b_{r-1}(z)\frac{d^{r-1}}{dz^{r-1}}F(z) + \cdots + b_0(z)F(z) = 0, \tag{2.11}$$

where $b_j(z) = a_j(z)/a_r(z)$.

Definition 2.24 (regular singular points and indicial polynomials) For any rational function $R(z)$, let $\omega_\zeta(R)$ be the order of the pole of R at $z = \zeta$, equal to 0 if R is analytic at ζ. The point $\zeta \in \Xi$ is a *regular singular point* of (2.11) and, equivalently, (2.7) if

$$\omega_\zeta(b_{r-1}) \leq 1, \qquad \omega_\zeta(b_{r-2}) \leq 2, \qquad \ldots, \qquad \omega_\zeta(b_0) \leq r.$$

A linear differential equation with only regular singular points is called *Fuchsian*. The *indicial polynomial* of (2.7) at a regular singular point ζ is the polynomial

$$I(\theta) = (\theta)_r + \delta_1\,(\theta)_{r-1} + \cdots + \delta_r,$$

where $(\theta)_j = \theta(\theta-1)\cdots(\theta-j+1)$ and $\delta_j = \lim_{z\to\zeta}(z-\zeta)^j\, b_{r-j}(z)$.

In the 1860s, Fuchs [52] published a study of linear differential equations, characterizing those which do not admit solutions with essential singularities. Soon after, Frobenius [51] gave a simplified method for computing a basis of solutions at a regular singular point, which consists of functions locally analytic in a slit disk[12]. The following result can be found in Wasow [124, Ch. 2].

---

[12] In fact, much of this theory seems to be contained in an unpublished manuscript of Riemman from several decades earlier. A historical treatment of early methods in differential equations is given by Gray [60].

Proposition 2.16 *Let ζ be a regular singular point of the D-finite equation (2.7). Then in a slit disk around ζ, equation (2.7) admits a $\mathbb{C}$-linear basis of analytic solutions of the form*

$$(1 - z/\zeta)^{\gamma} \sum_{j=0}^{m} \Big(f_j\,(1 - z/\zeta)\, \log^j (1 - z/\zeta) \Big), \tag{2.12}$$

where γ is a root of the indicial polynomial $I(\theta)$ and the f_j are analytic at the origin. If $I(\theta)$ has distinct roots, no two of which differ by an integer, then there are no logarithmic terms and a linear basis of solutions is given by functions of the form $(1 - z/\zeta)^{\gamma} f\,(1 - z/\zeta)$ where f is analytic at the origin.

Given expansion (2.12) valid in a slit disk, we write $F(z) \sim C(1-z/\zeta)^{\alpha} \log^m(1-z/\zeta)$ if the ratio $C(1 - z/\zeta)^{\alpha} \log^m(1 - z/\zeta)/F(z) \to 1$ whenever $z \to \zeta$ along any path in the slit disk. As with our work on Puiseux expansions, we can transfer a series expansion of the form (2.12) at the dominant singularity of a generating function to asymptotics of its coefficients. The following result is originally due to Jungen [67]; a modern proof can be found in Flajolet and Odlyzko [48].

Proposition 2.17 *Suppose $F(z) = \sum_{n\geq 0} f_n z^n$ is analytic in some open disk around the origin except for a single dominant singularity $\zeta \in \mathbb{C}$ and points in the ray $\mathcal{R}_\zeta = \{t\zeta : t \geq 1\}$. If $F(z)$ has a convergent expansion of the form (2.12) in a disk centred at ζ with $\mathcal{R}_\zeta$ removed and $F(z) \sim C(1 - z/\zeta)^{\alpha} \log^m(1 - z/\zeta)$ with $\alpha \notin \mathbb{N}$ then $f_n \sim C\zeta^{-n} \frac{n^{-\alpha-1}}{\Gamma(-\alpha)} \log^m n$. If there are a finite number of dominant singularities, each satisfying these hypotheses, then asymptotics of f_n are determined by adding the asymptotic contributions given by each.*

Remark 2.4 When $\alpha \in \mathbb{N}$ the identity

$$\frac{d}{dz}(1-z)^{\alpha} \log(1-z)^r = -\alpha(1-z)^{\alpha-1} \log(1-z)^r - r(1-z)^{\alpha-1} \log(1-z)^{r-1}$$

can be used to recursively determine the coefficients of $(1 - z)^{\alpha} \log(1 - z)^r$. Any term in (2.12) with $\alpha \in \mathbb{N}$ and $r = 0$ is a polynomial and may be subtracted from $F(z)$ without changing asymptotic behaviour.

Putting Propositions 2.16 and 2.17 together gives the following.

Proposition 2.18 *Assume that the coefficients $b_j(z)$ of (2.11) are analytic in a disk $|z| < \rho$ except at a unique pole $\zeta \in \Xi$ with $0 < |\zeta| < \rho$. Suppose that ζ is a regular singular point of (2.11) and that $F(z)$ is a solution to (2.11) which is analytic at the origin. If none of the solutions $\alpha_1, \ldots, \alpha_r$ to the indicial equation $I(\theta) = 0$ at ζ differ by an integer then there exist $c_1, \ldots, c_r \in \mathbb{C}$ such that for any τ with $|\zeta| < \tau < \rho$,*

$$[z^n]F(z) = \sum_{j=1}^{r} c_j \Delta_j(n) + O(\tau^{-n}), \tag{2.13}$$

where $\Delta_j(n) = 0$ *when* $\alpha_j \in \mathbb{N}$ *and* $\Delta_j(n)$ *is determined by an effective asymptotic series whose terms can be calculated to any desired order when* $\alpha_j \notin \mathbb{N}$*. If* $\alpha_j \notin \mathbb{N}$ *then the leading term of* $\Delta_j(n)$ *is*

$$\Delta_j(n) = \frac{n^{-\alpha_j - 1}}{\Gamma(-\alpha_j)} \zeta^{-n} \left(1 + O\left(\frac{1}{n}\right)\right).$$

When the roots of $I(\theta)$ *differ by an integer (including the case of multiple roots) then dominant asymptotics of* $[z^n]F(z)$ *are given by a* $\mathbb{C}$*-linear combination of terms of the form* $\zeta^{-n} n^{-\sigma_j} (\log n)^{\ell}$*, where* σ_j *is algebraic and* ℓ *is a non-negative integer. When there are a finite number of dominant singularities, each satisfying these hypotheses, then asymptotics are determined by adding the asymptotic contributions of each.*

The growth of D-finite series coefficients is less restricted than the algebraic case: logarithmic powers may be present and the factor $n^{-1-\alpha_j}$ can be irrational or a negative integer power. Further discussion of Proposition 2.18, its applications, and generalizations to irregular singularities can be found in Flajolet and Sedgewick [50, Sect. VII. 9 and VIII. 7].

Example 2.24 (Central Binomial Coefficients Squared)

As discussed above, the transcendental function $F(z) = \sum_{n \geq 0} \binom{2n}{n}^2 z^n$ satisfies the D-finite equation $z\,(1 - 16z)\,F''(z) + (1 - 32z)F'(z) - 4F(z) = 0$. Since $F(z)$ doesn't decay super-exponentially (it grows) it must have a singularity, and because it is analytic at the origin this singularity is the regular singular point $z = 1/16$. We can compute convergent expansions

$$A(z) = 1 + \frac{1}{4}(1 - 16z) + \frac{9}{64}(1 - 16z)^2 + \cdots$$

$$B(z) = \left(1 + \frac{1}{4}(1 - 16z) + \frac{9}{64}(1 - 16z)^2 + \cdots\right)\log(1 - 16z) + \frac{1}{2}(1 - 16z) + \frac{21}{64}(1 - 16z)^2 + \cdots$$

in a slit disk at $z = 1/16$ of functions $A(z)$ and $B(z)$ which form a basis of solutions for this D-finite equation. The logarithmic factor in B is a reflection of the fact that the indicial polynomial $I(\theta) = \theta^2$ at $z = 1/16$ has a double root. Since A and B form a basis of solutions, we can write $F(z) = \kappa_1 A(z) + \kappa_2 B(z)$ for $\kappa_1, \kappa_2 \in \mathbb{C}$. Furthermore, since A is analytic at $z = 1/16$ while F and B have singularities there, it must be the case that $\kappa_2 \neq 0$ and Proposition 2.17 implies

$$\binom{2n}{n}^2 = \kappa_2 \frac{16^n}{n} \left(1 + O\left(\frac{1}{n}\right)\right).$$

Finding asymptotics of the sequence $\binom{2n}{n}$ with algebraic generating function $G(z) = (1 - 4z)^{-1/2}$ and squaring the result shows $\kappa_2 = 1/\pi$. We discuss the general problem of determining such coefficients in the next subsection.

At an irregular singular point a solution of (2.7) may have an essential singularity, making arguments more difficult and leading to more general asymptotic behaviour. Fortunately, a set of powerful results collectively imply that any D-finite power series with integer coefficients that is analytic at the origin has only regular singularities, and the corresponding indicial polynomial has only rational roots. In fact, the result applies to a larger class of power series.

Definition 2.25 (G-functions) A power series $F(z) = \sum_{n\geq 0} f_n z^n \in \mathbb{Q}[[z]]$ is called a *G-function*[13] if $F(z)$ is D-finite and there exists a constant $C > 0$ such that for all n both $|f_n|$ and the least common denominator of $f_0, \ldots, f_n$ are bounded by C^n.

Proposition 2.19 (André-Chudnovsky-Katz Theorem) *If $F(z)$ is a G-function then a minimal order annihilating D-finite equation for F is Fuchsian, and its indicial polynomial $I(\theta)$ has only rational roots.*

Chudnovsky and Chudnovsky [24, Thm. III] showed that if $F(z)$ is a G-function then a minimal order annihilating D-finite equation $\mathcal{L}$ is *globally nilpotent*, meaning that a certain linear operator derived from $\mathcal{L}$ is nilpotent modulo p for all but a finite number of primes p. A previous result of Katz [68] then restricts the singular behaviour of $\mathcal{L}$, giving Proposition 2.19. André [2, Section VI] extended and clarified these arguments. Note that Proposition 2.19 gives information about all solutions of a D-finite equation from properties of a single solutions, a powerful result. Translating back to coefficient asymptotics gives the following, first used in combinatorial contexts by Garoufalidis [53] and Bostan et al. [16].

Corollary 2.2 *Suppose $F(z)$ is a G-function, for instance a D-finite function with integer coefficients which is analytic at the origin. Then, as $n \to \infty$, the power series coefficients (f_n) of $F(z)$ have an asymptotic expansion given by a sum of terms of the form $C\, n^\alpha\, \zeta^n\, (\log n)^\ell$ where $C \in \mathbb{C}$, $\alpha \in \mathbb{Q}$, $\ell \in \mathbb{N}$, and ζ is algebraic.*

Just as the asymptotic statement in Corollary 2.1 is useful for proving transcendence of functions, Proposition 2.18 and Corollary 2.2 are powerful tools for proving non-D-finiteness. Another useful technique for proving non-D-finiteness is to show that a function has an infinite number of singularities.

Example 2.25 (Number of Alternating Permutations is Not P-Recursive)

In Section 2.1.1 we studied the function $\tan(z) = \sum_{n\geq 0} \frac{a_n}{n!} z^n$, where a_n is the number of alternating permutations of n. Because $\tan(z)$ has a singularity at every $z = \pi/2 + \pi k$ with $k \in \mathbb{Z}$, the tangent function is not D-finite and $a_n/n!$ is not P-recursive. Since the product of P-recursive sequences is P-recursive by Item (5) of Proposition 2.15 above, and $1/n!$ is P-recursive, this implies a_n is not P-recursive.

Example 2.26 (Prime Sequence is Non-D-Finite)

[13] G-functions were introduced by Siegel [104] in his studies on number theory and elliptic integrals.

Flajolet et al. [47] prove non-D-finiteness of several generating functions by examining asymptotics of their power series coefficients, including logarithms, non-integer powers, and the sequence of primes. For instance, if p_n denotes the nth prime number then the prime number theorem allows one to deduce the asymptotic estimate

$$p_n = n \log n + n \log \log n + O(n),$$

and the existence of the log-log term proves that the generating function $P(z) = \sum_{n\geq 0} p_n z^n$ is non-D-finite.

2.4.1 An Open Connection Problem

The interplay between a G-function $F(z)$ and its P-recursive coefficient sequence (f_n) comes into sharp focus when trying to determine asymptotics. First, studying D-finite equations allowed for the asymptotic characterizations in Proposition 2.18 and Corollary 2.2. On the other hand, the Birkhoff-Trjitzinsky method applied to the P-recursion for f_n gives a basis of solutions $\Psi_1(n), \dots, \Psi_r(n)$ for the recurrence whose asymptotic expansions can be computed. To represent f_n in this basis it is most convenient to turn back to the D-finite equation for $F(z)$.

Definition 2.26 (connection coefficients) The constants $\lambda_1, \dots, \lambda_r \in \mathbb{C}$ such that $f_n = \lambda_1 \Psi_1(n) + \cdots + \lambda_r \Psi_r(n)$ are called *connection coefficients* of f_n with respect to the basis $\Psi_1, \dots, \Psi_r$. Similarly, if $\Lambda_1(z), \dots, \Lambda_s(z)$ form a basis of solutions for the D-finite equation satisfied by $F(z)$ then the constants $\lambda_1, \dots, \lambda_s \in \mathbb{C}$ such that $F(z) = \lambda_1 \Lambda_1(z) + \cdots + \lambda_s \Lambda_s(n)$ are called *connection coefficients* of $F(z)$ with respect to the basis $\Lambda_1, \dots, \Lambda_s$.

An exact characterization of connection coefficients for G-functions is currently unknown, although Fischler and Rivoal [45, Thms. 1 and 2] show that such any such constant is the evaluation $g(1)$ of a G-function $g \in \mathbb{Q}(\mathrm{i})[[z]]$ whose radius of convergence can be made arbitrarily large.

Most importantly, to determine even the scale of dominant asymptotics one must figure out which connection coefficients are nonzero. At the generating function level this is a connection problem analogous to the case of algebraic functions: around each element of Ξ local convergent expansions can be computed for a basis of solutions, and one needs to express $F(z)$ as a $\mathbb{C}$-linear combination of these basis elements. If the representation of $F(z)$ in such a basis contains an element whose local series expansion at $\zeta \in \Xi$ contains a non-power-series term, and there is no cancellation with other basis elements, then $z = \zeta$ is a singularity of F.

Numeric approximations of connection coefficients can be rigorously computed to any desired accuracy, allowing one to certify a non-zero connection coefficient and providing heuristic evidence when connection coefficients are zero. Naive techniques for computing such continuations were known in principle perhaps as far back as Frobenius. Efficient algorithms for continuation of D-finite functions, including to a regular singular point, were given by van der Hoeven [115] and outlined in previous work of Chudnovsky and Chudnovsky [26]; the key idea is to use fast evaluation of a D-finite power series inside its radius of convergence[14]. Additional aspects of this problem have been studied by Mezzarobba [80], who also developed a Sage package [81] which can compute hundreds (or thousands) of rigorously certified digits of connection coefficients in seconds when run on a modern computer.

Example 2.27 (A Numeric Connection Solution)

Consider a solution $F(z)$ of the Fuchsian differential equation

$$\begin{aligned} 5z^2(2z-1)(z-3)F^{(4)}(z) \;+\; 2z(59z^2-139z+36)F^{(3)}(z) \\ +\,6(61z^2-80z+6)F''(z) \;+\; 12(25z-11)F'(z)+36F(z)=0, \end{aligned} \tag{2.14}$$

whose coefficient sequence (f_n) satisfies the linear recurrence

$$3(n+3)(5n+4)f_{n+2}-(n+2)(35n+33)f_{n+1}+(10n^2+28n+18)f_n=0.$$

This recurrence has a basis of solutions consisting of two series

$$\Psi_1(n)=\frac{2^n}{n}\left(1-n^{-1}+O\left(n^{-2}\right)\right) \qquad \text{and} \qquad \Psi_2(n)=3^{-n}\left(1+O\left(n^{-2}\right)\right),$$

which can be determined automatically to any order by the Birkhoff-Trjitzinsky method. This means there exists constants λ_1 and λ_2, depending only on the initial conditions f_0 and f_1, such that

$$f_n=\lambda_1\Psi_1(n)+\lambda_2\Psi_2(n)\sim\begin{cases}\lambda_1\,\frac{2^n}{n} & :\lambda_1\neq 0\\ \lambda_2\,3^{-n} & :\lambda_1=0\end{cases}.$$

To say even if f_n increases or decreases, it is necessary to decide whether λ_1 is zero. To this end, we examine a basis of solutions $\Phi_0(z),\dots,\Phi_3(z)$ to the differential equation (2.14) defined by their expansions in a neighbourhood of the origin. We can compute a basis of solutions

$$\begin{aligned} \Phi_0(z)&=1-z^2/2+\cdots & \Phi_1(z)&=z+11z^2/6+\cdots\\ \Phi_2(z)&=z^{-1}+\cdots & \Phi_3(z)&=z^{1/5}\left(z+11z^2/6+\cdots\right), \end{aligned}$$

and matching up initial series terms with $F(z)=f_0+f_1z+\cdots$ implies $F(z)=f_0\Phi_0(z)+f_1\Phi_1(z)$. Note that although we do not know the basis elements explicitly, we can represent F in the basis using only the first two series terms of each element.

The leading coefficient of (2.14) vanishes at $z=1/2$ and $z=3$, corresponding to the potential asymptotic behaviour of f_n. For example, there is a basis of

[14] Only $O(n)$ terms of a D-finite power series are needed to compute n digits of an evaluation inside its radius of convergence, with good bounds on the precise number of necessary terms [83, 82].

solutions $\Gamma_0(z), \ldots, \Gamma_3(z)$ to (2.14) whose expansions at $z = 1/2$ are

$$\begin{aligned}
\Gamma_0(z) &= \log(z-1/2)\,(1 + 2(z-1/2) + \cdots) \\
\Gamma_1(z) &= 1 + (8/25)(1-z/2)^3 + \cdots \\
\Gamma_2(z) &= (z-1/2) + (8/25)(1-z/2)^3 + \cdots \\
\Gamma_3(z) &= (z-1/2)^2 - 2(1-z/2)^3 + \cdots,
\end{aligned}$$

where $\Gamma_0(z)$ contains a logarithmic singularity at $z = 1/2$ and the other three functions are analytic. To determine the singular behaviour of $F(z)$ at $z = 1/2$ we need to express F in terms of the Γ_j basis. Since we have expressed F in the Φ_j basis by examining its behaviour at the origin, it is sufficient to find the change of basis matrix M such that if $a_0, \ldots, a_3, b_0, \ldots, b_3 \in \mathbb{C}$ and

$$a_0\,\Psi_0(z) + \cdots a_3\,\Psi_3(z) = b_0\,\Gamma_0(z) + \cdots b_3\,\Gamma_3(z)$$

then

$$\left(b_0\ b_1\ b_2\ b_3\right)^T = M\left(a_0\ a_1\ a_2\ a_3\right)^T.$$

The analytic_ore_algebra Sage package of Mezzarobba [81] uses numeric analytic continuation to compute a rigorous numeric approximation

$$M \approx \begin{pmatrix} 0.50\ldots & -1.50\ldots & 0.0\ldots & -1.37\ldots \\ 2.14\ldots - i(1.57\ldots) & -2.83\ldots + i(4.71\ldots) & 2.0\ldots & -2.73\ldots + i(4.33\ldots) \\ 0.02\ldots - i(3.14\ldots) & 1.35\ldots - i(9.42\ldots) & -4.0\ldots & 1.05\ldots - i(8.66\ldots) \\ 1.67\ldots - i(6.28\ldots) & -4.44 + i(18.84\ldots) & 8.0\ldots & -4.408 + i(17.32\ldots) \end{pmatrix},$$

where the approximation is computed to 2500 certified digits in around 10 seconds on a modern laptop. The entries of

$$\left(b_0\ b_1\ b_2\ b_3\right) = M\ \left(f_0\ f_1\ 0\ 0\right)^T$$

then determine $F(z)$ in terms of the Γ_j basis. In particular, the coefficient b_0 of Γ_0 is $(0.50\ldots)f_0 + (-1.50\ldots)f_1$, where the dots hide 2498 zeroes, strongly suggesting $b_0 = (1/2)f_0 - (3/2)f_1$. Since $\Gamma(z)$ has a singular expansion beginning with $\log(z - 1/2)$ at $z = 1/2$, and Γ_1, Γ_2, and Γ_3 are analytic at $z = 1/2$, if true this would imply

$$f_n = \left(\frac{3f_1 - f_0}{2}\right) 2^n \left(n^{-1} + \cdots\right) + O\left(3^{-n}\right).$$

Thus, when $3f_1 - f_0$ is far from zero it seems reasonable that the coefficient sequence will grow exponentially and this will be easily detected. When $3f_1 - f_0$ equals zero it seems that $F(z)$ does not have a singularity at $z = 1/2$, and dominant asymptotics are determined by repeating the process with a basis of solutions whose expansions near $z = 3$ are known. When $3f_1 - f_0$ is very close but not equal to zero it seems that the sequence will grow exponentially, but one could be fooled into thinking this was not the case. In general the connection coefficients can be transcendental, making such an analysis difficult.

For this simple example it can be verified directly that

$$\Psi_1(n) = \frac{2^n}{n} \sum_{k\geq 0} \frac{(-1)^k}{n^k} = \frac{2^n}{n+1} \qquad \text{and} \qquad \Psi_2(n) = 3^{-n}$$

form a basis of solutions for the recurrence under consideration, so that

$$f_n = \left(\frac{3f_1 - f_0}{2}\right)\frac{2^n}{n+1} + \left(\frac{3f_0 - 3f_1}{2}\right)3^{-n}.$$

Example 2.28 (A Numeric Connection Solution for Binomial Coefficients)

Repeating this process on the differential equation

$$z\,(1-16z)\,\frac{d^2}{dz^2}F(z) + (1-32z)\frac{d}{dz}F(z) - 4F(z) = 0$$

with initial conditions $F(0) = 1$ and $F'(0) = 4$, which encodes the generating function of the central binomial coefficients squared, gives

$$\binom{2n}{n}^2 = (.3183098861\ldots)\,\frac{16^n}{n}\left(1 + O\left(\frac{1}{n}\right)\right).$$

The numerical approximation corresponds to the connection coefficient $1/\pi$.

Example 2.29 (Simple Lattice Walks in $\mathbb{N}^2$)

Problem 4.4 of Chapter 4 asks you to prove that the generating function $A(z)$ counting the number a_n of lattice walks beginning at the origin, staying in $\mathbb{N}^2$, and taking n steps in $\mathcal{S} = \{(\pm 1, 0), (0, \pm 1)\}$ satisfies the D-finite equation

$$\begin{aligned} z^2(4z-1)(4z+1)A'''(z) &+ 2z(4z+1)(16z-3)A''(t) \\ &+ 2(112z^2 + 14z - 3)A'(t) + 4(16z+3)A(t) = 0, \end{aligned} \tag{2.15}$$

so that (a_n) satisfies the P-recurrence

$$(n+4)(n+3)a_{n+2} - 4(2n+5)a_{n+1} - 16(n+1)(2+n)a_n = 0.$$

This recurrence has a basis of solutions with expansions

$$\Psi_1(n) = 4^n\, n^{-1}\left(1 - \frac{3}{2n} + \cdots\right) \qquad \text{and} \qquad \Psi_2(n) = (-4)^n\, n^{-3}\left(1 - \frac{9}{2n} + \cdots\right),$$

corresponding to the singular points $z = \pm 1/4$ of the D-finite equation (2.15). Performing the same numeric analytic continuation argument gives

$$a_n = (1.273\ldots)\Psi_1(n) + (5.092\ldots)\Psi_2(n) \sim (1.273\ldots)4^n n^{-1}.$$

In Section 6.1.1 of Chapter 6 we use the methods of analytic combinatorics in several variables to provide an asymptotic decomposition of a_n which determines these connection coefficients exactly. In particular, we will see that $a_n \sim (4/\pi)4^n n^{-1}$.

The determination of connection coefficients is a major open problem in enumerative combinatorics. We discuss methods of attacking this connection problem using analytic combinatorics in several variables in the following chapters.

Open Problem 2.4 (Decidability of D-finite Asymptotics) *Given a G-function $F(z)$, specified by an annihilating Fuchsian D-finite equation $\mathcal{L}$ and enough initial conditions to distinguish $F(z)$ as a solution, is it decidable to determine dominant asymptotics of $F(z)$? In particular, is it decidable to determine which singularities of $\mathcal{L}$ are singularities of $F(z)$ and find the corresponding connection coefficients?*

2.5 D-Algebraic Power Series

We end with a brief discussion of functions satisfying certain non-linear differential equations, mainly to highlight how delicate the relevant decidability issues are. As we do not consider such functions later in the text, we do not go into great detail; a more robust overview can be found in the surveys of Rubel [97, 98, 99].

Definition 2.27 (D-algebraic functions) An analytic function or formal series $F(z)$ is *differentially algebraic (D-algebraic)* over a field $\mathbb{K} \subset \mathbb{C}$ if there exists a non-zero polynomial $P(x_0, x_1, \ldots, x_d)$ with coefficients in $\mathbb{K}$ such that F and its derivatives satisfy the *algebraic differential equation (ADE)*

$$P\left(F, F', \ldots, F^{(d)}\right) = 0.$$

For a function or series F, satisfying an ADE is equivalent to the apparently weaker condition that there exists a multivariate polynomial Q with coefficients in $\mathbb{K}$ such that $Q\left(z, F, F', \ldots, F^{(d)}\right) = 0$. For n sufficiently large the coefficients f_n of a D-algebraic function $F(z)$ satisfy a recurrence of the form

$$p(n) f_n = q_n(f_0, \ldots, f_{n-1}),$$

where p and the q_n are polynomials which can be explicitly computed from an annihilating ADE. In particular, a power series solution of an ADE is still uniquely determined by a finite number of initial terms, and given these initial terms the coefficient f_n can be determined in polynomial time with respect to n. A good discussion of the complexity of basic operations with ADEs, such as extending known solutions, can be found in van der Hoeven [116].

Examples of D-Algebraic Series

The family of D-algebraic functions includes all generating function classes so far discussed, and is closed under sum, product, difference, quotient, composition, and compositional inverse, when defined; components of a system of ADEs are themselves D-algebraic. Unlike D-finite functions, analytic D-algebraic functions can have an infinite number of singularities. Examples above such as $\tan(z)$ and the generating function of partitions are non-D-finite but D-algebraic.

Perhaps the most famous combinatorial example of a D-algebraic function is Tutte's derivation [114] of the equation

$$2q^2(1-q)z + (qz + 10H - 6zH')H'' + q(4-q)(20H - 18zH' + 9z^2H'') = 0$$

for the generating function $H(z)$ counting q-coloured rooted triangulations by vertices, for all $q \in \mathbb{N}$. Bergeron and Reutenauer [7] show how a large class of ADEs can be interpreted as the generating functions of certain classes of trees. Solutions of ADEs have been important in theoretical computer science since Shannon's analysis [103] of the Differential Analyzer—an analogue computer designed to solve differential equations—as a model of computation. In fact, initial value problems of (first order) D-algebraic systems can simulate universal Turing machines [59]. That ADEs capture a wide variety of behaviour is not surprising in light of the existence of so-called universal differential equations; for instance, Duffin [37] showed that for any real continuous function $g(z)$ there is a four times continuously differentiable solution of the ADE

$$2F''''(z)F'(z)^2 - 5F'''(z)F''(z)F'(z) + 3F''(z)^3 = 0$$

which approximates $g(z)$ to arbitrary accuracy over the real numbers. The first example of a universal differential equation was given by Rubel [96].

Non-D-Algebraic Functions

Definition 2.28 (hypertranscendence) A non-D-algebraic function is called *hypertranscendental* or *transcendentally transcendental*.

Well known examples of hypertranscendental functions include the Euler gamma function $\Gamma(z)$, the Riemann zeta function, and the 'lacunary' series $\sum_{n\geq 0} z^{2^n}$. Important tools for proving hypertranscendence include theorems related to large gaps between non-zero power series terms [77], new complexity techniques inspired by applications in combinatorics [90], and differential Galois theory [118, 62]. In particular, differential Galois theory gives a powerful collection of tools for proving algebraicity, transcendence, and hypertranscendence of functions by reducing such questions to effectively decidable statements about algebraic groups.

Asymptotics and Decidability

D-algebraic series seem to be on the border between decidability and undecidability, perhaps tipped to the side of undecidability. There are algorithms to determine when an ADE admits a formal power series solution, or a smooth C^∞ solution, but it is undecidable to determine whether there is an analytic solution at the origin; see Denef and Lipshitz [35] for details. Compared to D-finite functions not much is known about D-algebraic asymptotics, and in general not much can be said. Denef and Lipshitz [35] show that it is undecidable to detect whether a D-algebraic series which is analytic at the origin has radius of convergence less than one, or greater than or equal to one. Recast in terms of coefficient asymptotics, this says it is undecidable to determine whether the coefficient sequence of an analytic D-algebraic function exponentially grows. Maillet [79] showed that if $F(z)$ is D-algebraic then its power series coefficients satisfy $|f_n| \leq K(n!)^\alpha$ for some constants $\alpha, K > 0$ and all n. Popken [93] showed that if $F(z)$ is D-algebraic and its power series coefficients f_n are algebraic numbers then $|f_n| \geq \exp\left(-cn(\log n)^2\right)$ for some constant $c > 0$. These results allow for asymptotic proofs of hypertranscendence for extremely fast growing or decaying sequences.

Appendix on Complex Analysis

This appendix summarizes the results from complex analysis needed for our asymptotic arguments. Because we continually introduce new concepts we do not break out definitions as in the rest of this text. Full proofs of the results discussed here can be found in Henrici [65] and Rudin [100].

An *open disk centred at* $a \in \mathbb{C}$ is a set of the form $D = \{z \in \mathbb{C} : |z - a| < r\}$ with $r > 0$. A subset $O \subset \mathbb{C}$ is *open* if for any $a \in O$ there is an open disk centred at a which is contained in O, while a *domain* Ω is a connected open subset of $\mathbb{C}$. A *neighbourhood of* $a \in \mathbb{C}$ is an open subset of $\mathbb{C}$ containing a, and a complex-valued function $f(z)$ is called *analytic at* $z = a$ if there exists a neighbourhood N of a where $f(z)$ is represented by a convergent power series,

$$f(z) = \sum_{n \geq 0} c_n (z - a)^n \qquad (z \in N).$$

We say f is *analytic in the domain* Ω if it is analytic at each point of Ω; this is equivalent to f being complex-differentiable in Ω, i.e., that the limit of $(f(z) - f(a))/(z - a)$ as z approaches a in the complex plane exists for every $a \in \Omega$.

Two analytic functions defined on a domain Ω which agree on any subset $S \subset \Omega$ containing a limit point (including any non-empty open set) must agree on all of Ω. This uniqueness property, sometimes called the *identity theorem*, implies that a non-zero analytic function has a finite number of zeroes in any compact set. If Ω_1 and Ω_2 are non-disjoint domains and $f : \Omega_1 \to \mathbb{C}$ and $g : \Omega_2 \to \mathbb{C}$ are analytic functions which agree on $\Omega_1 \cap \Omega_2 \neq \emptyset$, then one can *analytically continue* f by defining analytic $F : \Omega_1 \cup \Omega_2 \to \mathbb{C}$ with $F(z) = f(z)$ for $z \in \Omega_1$ and $F(z) = g(z)$ for $z \in \Omega_2$. In this case we say $g(z)$ is a *direct analytic continuation* of $f(z)$ to Ω_2.

Example 2.30 (Analytic Continuation of Series)

The function $C(z) = z/(1-2z)$ is analytic in the complex plane, except at the point $z = 1/2$. Since $C(z)$ has the convergent power series representation

$$C(z) = \sum_{n\geq 0} c_n z^n = \sum_{n\geq 1} 2^{n-1} z^n$$

for $|z| < 1/2$, where (c_n) is the counting sequence of integer compositions, we can view this rational function as a direct analytic continuation of the generating function of integer compositions from the disk $\{|z| < 1/2\}$ to $\mathbb{C} \setminus \{1/2\}$.

More generally, we say that g is an *analytic continuation* of f if there exists a sequence of direct analytic continuations on consecutively overlapping domains $\Omega_1, \Omega_2, \dots, \Omega_r$ beginning with f and ending with g. The analytic continuation g is uniquely determined by the sequence of domains Ω_j.

Complex Integration

A *curve* $\gamma : [a, b] \to \mathbb{C}$ is a piecewise continuously differentiable function from the real interval $[a, b]$ into the complex numbers; when $\gamma(t) = \gamma_1(t) + \mathrm{i}\gamma_2(t)$ for real functions γ_1 and γ_2 this states that γ_1 and γ_2 are piecewise continuously differentiable as functions from $[a, b]$ to $\mathbb{R}$. If $f(z)$ is a complex-valued function on γ then the complex path integral of f over γ is

$$\int_\gamma f(z)dz = \int_a^b f(\gamma(t))\gamma'(t)dt,$$

when this integral exists. If $f(z) = u(z) + \mathrm{i}v(z)$ for $u, v : \mathbb{C} \to \mathbb{R}$ then

$$\int_\gamma f(z)dz = \int_\gamma u(z)dz + \mathrm{i}\int_\gamma v(z)dz.$$

The integral is parametrization independent: it depends only on the image $\gamma([a,b])$, not on the actual function γ.

Example 2.31 (A Circular Integral)

The curve defined by $\gamma(t) = e^{it}$ for $t \in [-\pi, \pi]$ describes the unit circle in $\mathbb{C}$. If $n \in \mathbb{Z}$ and $n \neq -1$ then

$$\int_\gamma z^n dz = \int_a^b \gamma(t)^n \gamma'(t) dt = \int_{-\pi}^{\pi} ie^{it(1+n)} dt = \left. \frac{e^{it(1+n)}}{1+n} \right|_{t=-\pi}^{\pi} = 0,$$

while

$$\int_\gamma z^{-1} dz = \int_{-\pi}^{\pi} ie^{it(1-1)} dt = 2\pi i.$$

A curve γ is *closed* if $\gamma(a) = \gamma(b)$ and *simple* if γ is injective except for potentially having $\gamma(a) = \gamma(b)$. Two curves γ and γ' with the same endpoints are *homotopic* in the domain Ω if one can be continuously deformed into the other while staying in Ω. A *loop* in a domain Ω is a simple closed curve which can be continuously deformed to a single point while staying in Ω. A loop divides $\mathbb{C}$ into two regions, one of which, called the *interior* of the loop, is finite. A loop is *positively oriented* if the interior of the loop is on the left when tracing the path of the loop; for example, the loop defined by $\gamma(t) = e^{it}$ as t goes from $-\pi$ to π is positively oriented. The domain Ω is *simply connected* if every simple closed curve is a loop.

The following results are fundamental to complex analysis and form the bedrock of our asymptotic methods. First, deforming a curve of integration does not affect an integral as long as the curve stays in a domain where the integrand is analytic.

Proposition 2.20 (Path Independence of Integration) *Suppose f is analytic in the domain Ω. If γ and γ' are homotopic curves in Ω then*

$$\int_\gamma f(z) dz = \int_{\gamma'} f(z) dz.$$

In particular, if γ is a loop in Ω then $\int_\gamma f(z)dz = 0$.

The relationship between a function's power series coefficients and its analytic behaviour comes from the next result.

Proposition 2.21 (Cauchy Integral Theorem) *Suppose f is analytic in the domain Ω and $D = \{z : |z-a| \leq r\} \subset \Omega$ for some $r > 0$. If $\gamma = \{z : |z-a| = r\}$ is the positively oriented boundary of the disk D, w is in the interior of D, and $n \in \mathbb{N}$, then*

$$f^{(n)}(w) = \frac{1}{2\pi i} \int_\gamma \frac{f(z)}{(z-w)^{n+1}} dz,$$

where $f^{(n)}$ denotes the nth derivative of f, with $f^{(0)}(z) = f(z)$.

The following bound comes from the definition of a complex integral and the triangle inequality.

Proposition 2.22 (Maximum Modulus Integral Bound) *If $f(z)$ is continuous on a curve γ then*

$$\left| \int_\gamma f(z)dz \right| \le \text{length}(\gamma) \cdot \max_{z \in \gamma} |f(z)|,$$

where length(γ) *is the arc length of the curve.*

Singular Behaviour

Roughly speaking, a singularity of a function $f(z)$ is a point near which f behaves badly. Examining local behaviour of a function around such points will be key to our asymptotic methods. For our purposes we need only consider a few common types of singularities.

1. Isolated Singularities

The *open annulus* of radii $0 \le r < R \le \infty$ centred at $z = a$ is the set $A_{r,R}(a) = \{z \in \mathbb{C} : r < |z - a| < R\}$, and a *punctured disk* centred at a is an open annulus centred at a with $r = 0$. We say that $z = a$ is an *isolated singularity* if there exists a punctured disk centred at a where f is analytic, but f is not analytic at a and cannot be made analytic solely by changing (or defining for the first time) the value of $f(a)$.

Proposition 2.23 *Let $a \in \mathbb{C}$, let $0 \le r < R \le \infty$, and suppose f is analytic in the annulus $A_{r,R}(a)$. Then f can be represented by a unique* Laurent expansion

$$f(z) = \sum_{n=-\infty}^{\infty} c_n(z-a)^n = \sum_{n=1}^{\infty} c_{-n}(z-a)^{-n} + \sum_{n=0}^{\infty} c_n(z-a)^n \tag{2.16}$$

for all $z \in A_{r,R}(a)$, with uniform convergence on any compact subset of $A_{r,R}(a)$.

Suppose the expansion (2.16) is valid in a punctured disk around $z = a$. When $c_n = 0$ for $n < 0$ in (2.16) then f is analytic at $z = a$ or can be made analytic there by defining $f(a) = c_0$. Conversely, if there exists $n < 0$ with $c_n \neq 0$ then $z = a$ is an isolated singularity of f. The coefficient c_{-1} of the expansion (2.16) is called the *residue* of f at $z = a$ and denoted $\text{Res}_{z=a} f(z)$.

Note that unlike formal Laurent series, convergent Laurent expansions can have an arbitrary number of terms with negative exponents. There are thus two types of isolated singularities, characterized by the coefficients c_n in (2.16) which vanish.

- If there exists $M > 0$ such that $c_{-M} \neq 0$ but $c_n = 0$ for all $n \le -M$, then we say that a is a *pole of order M*. A pole of order 1 is called a *simple pole*. When f

is analytic or has a pole at $z = a$ we say that f is *meromorphic at* $z = a$, and there exists a punctured disk D centred at a and analytic functions $g(z)$ and $h(z)$ such that $f(z) = g(z)/h(z)$ for $z \in D$. When g vanishes to order s at $a \in \Omega$ and h vanishes to order r at a, i.e., if

$$g(a) = g'(a) = \cdots = g^{(s)}(a) = h(a) = h'(a) = \cdots = h^{(r)}(a) = 0$$

but $g^{(s+1)}(a)$ and $h^{(r+1)}(a)$ are nonzero, then $z = a$ is a singularity of $f(z)$ if and only if $r > s$, in which case $z = a$ is a pole of order $M = r - s$. We say f is *meromorphic in a domain* Ω if it is meromorphic at each point of Ω; a ratio of analytic functions in Ω is always meromorphic in Ω. Near a polar singularity, f behaves like a rational function whose denominator is going to zero. The following result, useful for calculating residues, follows from (2.16) and repeated term-by-term differentiation.

Lemma 2.4 *When $f(z)$ has a pole of order M at $z = a$ then*

$$\operatorname*{Res}_{z=a} f(z) = \frac{1}{(M-1)!} \lim_{z\to a} \frac{d^{M-1}}{dz^{M-1}} (z-a)^M f(z).$$

If $f(z) = g(z)/h(z)$ with $h(a) = 0$ but $h'(a), g(a) \neq 0$ then $z = a$ is a simple pole of $f(z)$ and $\mathrm{Res}_{z=a}\, f(z) = g(a)/h'(a)$.

- If there exist an infinite number of indices $n < 0$ such that $c_n \neq 0$ in (2.16) then f has an *essential singularity* at $z = a$. Behaviour near an essential singularity is more complicated than behaviour near a pole: Picard's Great Theorem states that in any neighbourhood of an essential singularity $f(z)$ takes on every complex value, except possibly one, infinitely many times. An example of an essential singularity is the function $f(z) = \mathrm{e}^{1/z}$ at $z = 0$, which takes on every value but zero infinitely many times in any neighbourhood of the origin.

Example 2.32 (Residues of Tangent)

The function $\tan(z) = \sin(z)/\cos(z)$ is meromorphic in $\Omega = \mathbb{C}$ and its singularities form the set $\mathcal{S} = \{(k + 1/2)\pi : k \in \mathbb{Z}\}$ of zeroes of $\cos(z)$. The numerator $\sin(z)$ and the derivative $\cos(z)' = -\sin(z)$ are non-zero when $z \in \mathcal{S}$, so every singularity of $\tan(z)$ is a simple pole and Lemma 2.4 implies

$$\operatorname*{Res}_{z=(k+1/2)\pi} \tan(z) = \frac{\sin((k+1/2)\pi)}{-\sin((k+1/2)\pi)} = -1.$$

Residue computations can be used to evaluate integrals of meromorphic functions.

Proposition 2.24 (Cauchy Residue Theorem) *Let $f(z)$ be a meromorphic function on a domain $\Omega \subset \mathbb{C}$, and let γ be a positively oriented loop in Ω on which f is analytic. If $a_1, \ldots, a_r$ denote the poles of $f(z)$ interior to γ then*

$$\frac{1}{2\pi \mathrm{i}} \int_\gamma f(z)dz = \sum_{j=1}^{r} \operatorname*{Res}_{z=a_j} f(z).$$

2. Algebraic and Logarithmic Branch Points

A function with an isolated singularity at $z = a$ is not analytic at a but does have a convergent series expansion in some punctured disk around a. When dealing with algebraic and D-finite generating functions we must consider more general singular behaviour, arising from a need to take inverses of non-injective functions.

Example 2.33 (Square-Root Branches)

Because the function $z \mapsto z^2$ is not injective, one must carefully define an inverse function $z \mapsto \sqrt{z}$. Given $z \in \mathbb{C} \setminus \mathbb{R}_{\leq 0}$, the *principal argument* $\mathrm{Arg}(z)$ is obtained by writing $z = r\mathrm{e}^{i\theta}$ for real r and $-\pi < \theta < \pi$, and taking $\mathrm{Arg}(z) = \theta$. The *principal branch* square-root is the analytic function $f : \mathbb{C} \setminus \mathbb{R}_{\leq 0} \to \mathbb{C}$ defined by

$$f(z) = \sqrt{z} = \sqrt{|z|}\, \mathrm{e}^{\mathrm{i}\,\mathrm{Arg}(z)/2},$$

where $\sqrt{|z|}$ is the usual square-root on $\mathbb{R}_{>0}$. One cannot extend $f(z)$ to an analytic function on any punctured disk around the origin as the limit of f does not exist at any negative real number:

$$\lim_{\theta \to \pi} f\left(\varepsilon \mathrm{e}^{\mathrm{i}\theta}\right) = \sqrt{\varepsilon}\mathrm{e}^{\mathrm{i}\pi/2} = \mathrm{i}\sqrt{\varepsilon} \neq -\mathrm{i}\sqrt{\varepsilon} = \sqrt{\varepsilon}\mathrm{e}^{-\mathrm{i}\pi/2} = \lim_{\theta \to -\pi} f\left(\varepsilon \mathrm{e}^{\mathrm{i}\theta}\right)$$

for any $\varepsilon > 0$, while $\varepsilon\mathrm{e}^{\mathrm{i}\pi} = -\varepsilon = \varepsilon\mathrm{e}^{-\mathrm{i}\pi}$. Removing the set $\mathbb{R}_{\leq 0}$ from the domain of f to obtain an analytic function is called making a *branch cut*. In fact, removing any simple curve from the origin to infinity gives a valid branch cut and any continuous choice of the argument function on the branch cut defines an analytic *branch* of the square-root. For any $r \in \mathbb{Q} \setminus \mathbb{Z}$ the branches of z^r are defined by an analogous process.

Example 2.34 (Logarithm Branches)

Similarly, the *principal branch* logarithm is the function defined by

$$\mathrm{Log}(z) = \log|z| + i\mathrm{Arg}(z)$$

for $z \in \mathbb{C}\setminus\mathbb{R}_{\leq 0}$, where $\log|z|$ is the usual logarithm on $\mathbb{R}_{>0}$ and $\mathrm{Arg}(z)$ is the principal argument of z. Different branches of the logarithm can be defined by removing any simple curve from the origin to infinity and selecting any continuous choice of the argument function after taking the branch cut. At any $z \in \mathbb{C}$, all branches of the logarithm which are defined at z differ by some integer multiple of $2\pi\mathrm{i}$.

A *cut disk* around a point $a \in \mathbb{C}$ is a disk around a with a simple curve from a to the boundary of the disk removed. The function $f(z)$ has an *algebraic singularity* or *algebraic branch point* at $a \in \mathbb{C}$ if it can be represented by a convergent series expansion of the form

$$f(z) = \sum_{n \geq N} c_n \left((z-a)^{1/R}\right)^n$$

in a cut disk D centred at a, where $N \in \mathbb{Z}$, $R \in \mathbb{N}$, the map $t \mapsto t^{1/R}$ is defined by a consistent branch in D, and there exists $c_n \neq 0$ where $n/R \notin \mathbb{Z}$. Equivalently, $f(z)$ is not analytic at $z = a$ but $f(z^R + a)$ is meromorphic at the origin for some $R \in \mathbb{N}_{>1}$.

The function $f(z)$ has a *logarithmic singularity* or *logarithmic branch point* at $a \in \mathbb{C}$ if $f(z)$ can be represented by a convergent series expansion of the form

$$(z-a)^N \sum_{j=0}^{d} \left(C_j(z)\ (\log(z-a))^j\right)$$

in a cut disk D centred at a, where $N \in \mathbb{Z}$, $d > 0$, the logarithm is defined by a consistent branch in D, and each $C_j(z)$ is analytic at a. Finally, an *algebro-logarithmic singularity* is a combination of algebraic and logarithmic singularities, that is, a point where $f(z)$ can be locally represented in a cut disk by a series of terms containing natural number powers of logarithms and Rth roots for some positive integer $R > 1$.

3. General Singularities

More generally, let γ be a loop and $f(z)$ be an analytic function on the interior of γ. We say that $a \in \gamma$ is a *singularity* of $f(z)$ if $f(z)$ cannot be analytically continued to a neighbourhood containing a. When discussing singularities of a generating function, the interior of the loop γ should contain the origin. This definition contains the singularities discussed above, and many more.

Example 2.35 (Natural Boundaries)

Problem 2.18 asks you to prove that the function $F(z) = \sum_{n \geq 0} z^{2^n}$ defined for $|z| < 1$ cannot be analytically continued outside the unit circle, meaning every point on the unit circle is a singularity of F (we say the unit circle is a *natural boundary* of F).

The Gamma Function

The *Euler gamma function* $\Gamma(z)$, arising in our asymptotic results on algebraic and D-finite functions, is defined by $\Gamma(z) = \int_0^\infty e^{-t}t^{z-1}dt$ when the real part $\Re(z) > 0$ (so the integral converges) and is uniquely extended to the complex plane with the negative integers removed by analytic continuation; it has a simple pole at each negative integer. Integration by parts shows that $\Gamma(z+1) = z\Gamma(z)$ and $\Gamma(1) = \int_0^\infty e^{-t}dt = 1$, so $\Gamma(n+1) = n!$ for any $n \in \mathbb{N}$. The gamma function can thus be viewed as an analytic extension of the factorial function to non-integers. In fact, the gamma function is the unique analytic extension subject to the additional constraint that $\log(\Gamma(x))$ is convex for $x > 0$. The value $\Gamma(1/2) = \int_0^\infty t^{-1/2}e^{-t}dt = \sqrt{\pi}$ and the recurrence satisfied by $\Gamma(z)$ allows one to calculate the gamma function at any half-integer. Stirling's formula gives an asymptotic expansion $\Gamma(z) \sim z^{z-1/2}e^{-z}\sqrt{2\pi}\left(1 + z^{-1}/12 + z^{-2}/288 + \cdots\right)$ as $|z| \to \infty$ with the argument of z fixed in $(-\pi, \pi)$. Proofs of these results, and further details, can be found in Andrews et al. [3, Ch. 1].

Problems

2.1 Prove Lemma 2.1.

2.2 For $F = \sum_{n\geq 0} f_n z^n$ and $G = \sum_{n\geq 0} g_n z^n$ in $\mathbb{K}[[z]]$, define $d(F,G) = 2^{-\min\{n : f_n \neq g_n\}}$, where $d(F,G) = 0$ if $F = G$. Prove that d satisfies the *strong triangle inequality*,

$$d(F,H) \leq \max\{d(F,G), d(G,H)\}, \quad \text{for all } F, G, H \in \mathbb{K}[[z]].$$

Since d is symmetric, and $d(F,G) \geq 0$ with equality if and only if $F = G$, this implies d defines a *non-archimedean metric* on $\mathbb{K}[[z]]$.

2.3 A sequence of formal power series (F_n) in $\mathbb{K}[[z]]$ is called *Cauchy* if for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that $d(F_n, F_m) < \varepsilon$ whenever $n, m \geq N$. Prove that every Cauchy sequence in $\mathbb{K}[[z]]$ converges to an element of $\mathbb{K}[[z]]$ under the metric d. This implies $\mathbb{K}[[z]]$ forms a *complete* non-archimedean metric space.

2.4 Suppose $F(z) = \sum_{n\geq 0} f_n z^n$ and $G(z) = \sum_{n\geq 0} g_n z^n$ are elements of $\mathbb{K}[[z]]$ with $g_0 = 0$. Prove that the sequence (S_N) in $\mathbb{K}[[z]]$ defined by $S_N = \sum_{k=0}^N f_k G(z)^k$ converges in $\mathbb{K}[[z]]$. We define the *composition* $F(G(z))$ as the limit of this sequence, $F(G(z)) = \lim_{N\to\infty} S_N = \sum_{k\geq 0} f_k G(z)^k$.

2.5 Suppose $F(z) = \sum_{n\geq 0} f_n z^n \in \mathbb{K}[[z]]$ with $f_0 \neq 0$. Show that $F(z) = f_0 - zG(z)$ for some $G \in \mathbb{K}[[z]]$, and prove that the infinite series $\sum_{k\geq 0} z^k G(z)^k / f_0^{k+1}$ converges to $I(z) \in \mathbb{K}[[z]]$ satisfying $F(z)I(z) = 1$.

2.6 Prove that set of formal series in $\mathbb{Q}[[z]]$ which define analytic functions at the origin is dense in $\mathbb{Q}[[z]]$. In other words, show that for any $F \in \mathbb{Q}[[z]]$ there is a sequence of analytic power series whose limit in the metric space $\mathbb{Q}[[z]]$ is F.

2.7 Prove that the coefficient of z^{100} in the rational function

$$F(z) = \frac{1}{(1-z)(1-z^5)(1-z^{10})(1-z^{25})}$$

gives the number of ways to make change for a dollar using pennies (1 cent), nickels (5 cents), dimes (10 cents), and quarters (25 cents), and find that coefficient. Find a recurrence for the number of ways f_n to make change for n cents.

2.8 The *Lagrange inversion formula* states that if $u, R \in \mathbb{Q}[[z]]$ satisfy $u(0) = 0$ and $z = u(z)/R(u(z))$ then $[z^n]u(z) = \frac{1}{n}[t^{n-1}]R(t)^n$ for $n > 0$. Prove this by justifying the calculation

$$n[z^n]u(z) = [z^{n-1}]u'(z) = \frac{1}{2\pi \mathrm{i}}\int \frac{u'(z)}{z^n}dz = \frac{1}{2\pi \mathrm{i}}\int \frac{R(u)^n}{u^n}du = [t^{n-1}]R(t)^n.$$

Start by showing you may assume R and u are polynomials. Use the Lagrange inversion formula to find the coefficients of the power series $y(z)$ satisfying $y = z + zy^2$, which counts rooted trees where each node is a leaf or has two children.

2.9 Let $\mathcal{A} \subset \mathbb{Q}[[z]]$ denote series of rational functions with natural number coefficients such that if σ is a dominant singularity then $\sigma/|\sigma|$ is a root of unity. Prove that for any $a, b \in \mathcal{A}$ the functions $a(z) + b(z)$, $a(z)b(z)$, and $1/(1 - za(z))$ are in $\mathcal{A}$, establishing Proposition 2.6. *Hint:* For the final construction suppose the dominant singularities of $a \in \mathcal{A}$ have modulus $p > 0$ and let $c(z) = 1/(1 - za(z))$. Show that the equation $za(z) = 1$ has a unique root $r \in (0, p)$, and that $z = r$ is a dominant singularity of $c(z)$. Conclude that any dominant singularity θ of $c(z)$ satisfies $\theta a(\theta) = 1 = ra(r)$ so $\theta^k = r^k = |\theta|^k$ for some $k \in \mathbb{N}$.

2.10 Prove that the series $F(z) = \sum_{n\geq 0}\binom{2n}{n}^{\kappa} z^n$ is transcendental over $\mathbb{Q}(z)$ for any natural number $\kappa \geq 2$.

2.11 Prove that if (f_n) satisfies a P-recursive equation of order r and degree d then $F(z) = \sum_{n\geq 0} f_n z^n \in \mathbb{K}[[z]]$ satisfies a D-finite equation of order at most d and degree at most $r + d$.

2.12 Let a_n be the number of alternating permutations (satisfying $\pi_1 > \pi_2 < \pi_3 > \pi_4 < \dots$) when n is odd and zero when n is even. Let

$$T(z) = \sum_{n\geq 0}\frac{a_n}{n!}z^n = \sum_{k\geq 0}\frac{a_{2k+1}}{(2k+1)!}z^{2k+1}.$$

- By considering the location of 1 in a permutation, prove that for all $k \geq 1$

$$a_{2k+1} = \sum_{\substack{1\leq j\leq 2k \\ j \text{ odd}}}\binom{2k}{j}a_j a_{2k-j}$$

- Using this recurrence, show that $T'(z) = T(z)^2 + 1$.

- Solve the differential equation to conclude $T(z) = \tan z$.

2.13 For any positive integer B let $\mathcal{S}_B$ denote the square in the complex plane with corners $\pm B\pi \pm B\pi\mathrm{i}$. Using the hyperbolic trigonometric functions and the fact that

$$|\tan(x + \mathrm{i}y)|^2 = \frac{\tan(x)^2 + \tanh(y)^2}{1 + \tan(x)^2 \tanh(y)^2} = \frac{\cosh(y)^2 - \cos(x)^2}{\cosh(y)^2 + \cos(x)^2 - 1}$$

for all $x, y \in \mathbb{R}$, prove that if $z \in \mathcal{S}_B$ for $B \in \mathbb{N}$ then $|\tan z| \leq \frac{1}{\tanh(\pi)} < 1.1$. Adapt the argument in Section 2.1.1 to prove that for any $n \in \mathbb{N}$ and positive integer B the number a_n of alternating permutations satisfies

$$\frac{a_n}{n!} = 2\left(\frac{2}{\pi}\right)^{n+1} \sum_{k=0}^{B-1} \frac{1}{(2k+1)^{n+1}} + O\left((\pi B)^{-n}\right).$$

2.14 Let A be a commutative ring and

$$P(z) = p_n z^n + \cdots + p_1 z + p_0$$
$$Q(z) = q_m z^m + \cdots + q_1 z + q_0$$

be polynomials of degree n and m in $A[z]$. The *Sylvester matrix* $S(P, Q)$ of P and Q is the $(m + n) \times (m + n)$ matrix obtained by repeating m times the vector $(p_n, \ldots, p_0)$ with each copy shifted once over, then repeating n times the vector $(q_m, \ldots, q_0)$ with each copy shifted once over. For instance, if $n = 4$ and $m = 3$ then

$$S(P, Q) = \begin{pmatrix} p_4 & p_3 & p_2 & p_1 & p_0 & 0 & 0 \\ 0 & p_4 & p_3 & p_2 & p_1 & p_0 & 0 \\ 0 & 0 & p_4 & p_3 & p_2 & p_1 & p_0 \\ q_3 & q_2 & q_1 & q_0 & 0 & 0 & 0 \\ 0 & q_3 & q_2 & q_1 & q_0 & 0 & 0 \\ 0 & 0 & q_3 & q_2 & q_1 & q_0 & 0 \\ 0 & 0 & 0 & q_3 & q_2 & q_1 & q_0 \end{pmatrix}.$$

Define the *resultant* $\mathrm{Resultant}_z(P, Q)$ of P and Q as the determinant of $S(P, Q)$.

a) Show that $\mathrm{Resultant}_z(P, Q)$ lies in A.
b) Prove that if P and Q share a root α then $\mathrm{Resultant}_z(P, Q) = 0$. *Hint:* It is sufficient to prove the existence of a non-zero vector $\mathbf{x}$ such that $S(P, Q)\mathbf{x} = \mathbf{0}$.
c) Suppose that

$$P(z) = a(z - \alpha_1) \cdots (z - \alpha_n) \qquad \text{and} \qquad Q(z) = b(z - \beta_1) \cdots (z - \beta_m)$$

for $\alpha_i, \beta_j \in \mathbb{C}$. Prove that $\mathrm{Resultant}_z(P, Q) = R$ where $R = a^m b^n \prod_{i,j}(\alpha_i - \beta_j)$. *Hint:* Consider $\mathrm{Resultant}_z(P, Q)$ and R to be multivariate polynomials in new variables α_i and β_j. Show that R divides the resultant and consider their degrees.

2.15 Suppose $a(x)$ and $b(x)$ are algebraic series satisfying $P(x, a(x)) = Q(x, b(x)) = 0$ for non-zero $P(x, y), Q(x, y) \in \mathbb{Q}[x, y]$, and let $d = \deg_y(P)$. Prove that

- $y = a(z) + b(z)$ is a root of $\text{Resultant}_z(P(x, y - z), Q(x, y)) \in \mathbb{Q}[y, z]$
- $y = a(x)b(x)$ is a root of $\text{Resultant}_z(z^d P(x, y/z), Q(x, y)) \in \mathbb{Q}[y, z]$
- when $a(0) \neq 0$ then $y = 1/a(x)$ is a root of $\text{Resultant}_z(P(x, y), 1 - yz) \in \mathbb{Q}[y, z]$

2.16 Use Cauchy's integral formula and the maximum modulus integral bound to show that if $f(z)$ is analytic in a disk $|z| \leq R + \varepsilon$ then it is represented at the origin by a power series which converges in $|z| \leq R$. Conclude that a power series admits a singularity on the boundary of its domain of convergence.

2.17 What is wrong with the "proof"

$$1 = \sqrt{1} = \sqrt{(-1)(-1)} = \sqrt{-1}\sqrt{-1} = \left(\sqrt{-1}\right)^2 = -1$$

2.18 Let $F(z) = \sum_{n \geq 0} z^{2^n}$ and $\zeta \in \mathbb{C}$ with $|\zeta| = 1$. Show that $F(z)$ is unbounded in any neighbourhood of ζ inside the unit circle, and conclude that F cannot be analytically continued outside the unit circle.

References

1. Désiré André. Développements de séc x et de tang x. *Comptes rendus de l'Académie des sciences*, 88:965–967, 1879.
2. Yves André. *G-functions and geometry*. Aspects of Mathematics, E13. Friedr. Vieweg & Sohn, Braunschweig, 1989.
3. George E. Andrews, Richard Askey, and Ranjan Roy. *Special functions*, volume 71 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1999.
4. Cyril Banderier and Michael Drmota. Formulae and asymptotics for coefficients of algebraic functions. *Combin. Probab. Comput.*, 24(1):1–53, 2015.
5. E. Barcucci, A. Del Lungo, A. Frosini, and S. Rinaldi. A technology for reverse-engineering a combinatorial problem from a rational generating function. *Adv. in Appl. Math.*, 26(2):129–153, 2001.
6. F. Bergeron, G. Labelle, and P. Leroux. *Combinatorial species and tree-like structures*, volume 67 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1998. Translated from the 1994 French original by Margaret Readdy, with a foreword by Gian-Carlo Rota.
7. François Bergeron and Christophe Reutenauer. Combinatorial resolution of systems of differential equations. III. A special class of differentially algebraic series. *European J. Combin.*, 11(6):501–512, 1990.
8. Daniel Bernoulli. Observationes de seriebus quae formantur ex additione vel subtractione quacunque terminorum se mutuo consequentium. *Commentarii Academiae Scientiarum Imperialis Petropolitanae*, 3:85–100, 1728.
9. Jean Berstel. Sur les pôles et le quotient de Hadamard de séries N-rationnelles. *C. R. Acad. Sci. Paris Sér. A-B*, 272:A1079–A1081, 1971.
10. Jean Berstel and Maurice Mignotte. Deux propriétés décidables des suites récurrentes linéaires. *Bull. Soc. Math. France*, 104(2):175–184, 1976.
11. George D. Birkhoff. Formal theory of irregular linear difference equations. *Acta Math.*, 54(1):205–246, 1930.
12. George D. Birkhoff and W. J. Trjitzinsky. Analytic theory of singular difference equations. *Acta Math.*, 60(1):1–89, 1933.

13. Alin Bostan, Mireille Bousquet-Mélou, Manuel Kauers, and Stephen Melczer. On 3-dimensional lattice walks confined to the positive octant. *Annals of Combinatorics*, 20(4):661–704, 2016.
14. Alin Bostan, Frédéric Chyzak, Marc Giusti, Romain Lebreton, Grégoire Lecerf, Bruno Salvy, and Éric Schost. *Algorithmes Efficaces en Calcul Formel*. Frédéric Chyzak (auto-édit.), Palaiseau, September 2017. 686 pages. Imprimé par CreateSpace. Aussi disponible en version électronique.
15. Alin Bostan, Frédéric Chyzak, Bruno Salvy, Grégoire Lecerf, and Éric Schost. Differential equations for algebraic functions. In *ISSAC 2007*, pages 25–32. ACM, New York, 2007.
16. Alin Bostan, Kilian Raschel, and Bruno Salvy. Non-D-finite excursions in the quarter plane. *J. Combin. Theory Ser. A*, 121:45–63, 2014.
17. Mireille Bousquet-Mélou. Walks in the quarter plane: Kreweras' algebraic model. *Ann. Appl. Probab.*, 15(2):1451–1491, 2005.
18. Mireille Bousquet-Mélou. Rational and algebraic series in combinatorial enumeration. In *International Congress of Mathematicians. Vol. III*, pages 789–826. Eur. Math. Soc., Zürich, 2006.
19. Sean Carrell and Guillaume Chapuy. A simple recurrence formula for the number of rooted maps on surfaces by edges and genus. In *26th International Conference on Formal Power Series and Algebraic Combinatorics (FPSAC 2014)*, Discrete Math. Theor. Comput. Sci. Proc., AT, pages 573–584. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2014.
20. Augustin Louis Cauchy. *Cours d'analyse de l'École royale polytechnique. I. Analyse algébrique*. Debure frères, Paris, 1821.
21. Cyril Chabaud. *Séries génératrices algébriques : asymptotique et applications combinatoires*. PhD thesis, Université Pierre et Marie Curie (UPMC), 2002. Type : Thèse de Doctorat – Soutenue le : 2002-11-22 – Dirigée par : Soria, Michèle.
22. Shaoshi Chen and Manuel Kauers. Some open problems related to creative telescoping. *J. Syst. Sci. Complex.*, 30(1):154–172, 2017.
23. N. Chomsky and M. P. Schützenberger. The algebraic theory of context-free languages. In *Computer programming and formal systems*, pages 118–161. North-Holland, Amsterdam, 1963.
24. D. V. Chudnovsky and G. V. Chudnovsky. Applications of Padé approximations to Diophantine inequalities in values of G-functions. In *Number theory (New York, 1983–84)*, volume 1135 of *Lecture Notes in Math.*, pages 9–51. Springer, Berlin, 1985.
25. D. V. Chudnovsky and G. V. Chudnovsky. On expansion of algebraic functions in power and Puiseux series. I. *J. Complexity*, 2(4):271–294, 1986.
26. David V. Chudnovsky and Gregory V. Chudnovsky. Computer algebra in the service of mathematical physics and number theory. In *Computers in mathematics (Stanford, CA, 1986)*, volume 125 of *Lecture Notes in Pure and Appl. Math.*, pages 109–232. Dekker, New York, 1990.
27. Frédéric Chyzak. *The ABC of Creative Telescoping — Algorithms, Bounds, Complexity*. Habilitation à diriger des recherches (hdr), University Paris-Sud 11, April 2014.
28. Andrew R. Conway and Anthony J. Guttmann. On 1324-avoiding permutations. *Adv. in Appl. Math.*, 64:50–69, 2015.
29. J. H. Conway. The weird and wonderful chemistry of audioactive decay. In Thomas M. Cover and B. Gopinath, editors, *Open Problems in Communication and Computation*, pages 173–188. Springer New York, New York, NY, 1987.
30. G. Darboux. Mémoire sur l'approximation des fonctions de très-grands nombres, et sur une classe étendue de développements en série. *Journal de mathématiques pures et appliquées 3e série*, 4:5–56, 1878.
31. N. G. de Bruijn. *Asymptotic methods in analysis*. Bibliotheca Mathematica. Vol. 4. North-Holland Publishing Co., Amsterdam; P. Noordhoff Ltd., Groningen; Interscience Publishers Inc., New York, 1958.
32. A. de Moivre. *The Doctrine of Chances or a Method of Calculating the Probabilities of Events in Play*. London, 1718.

33. A. de Moivre. *Miscellanea analytica de seriebus et quadraturis*. J. Tomson and J. Watts, London, 1730.
34. M. Delest. Algebraic languages: a bridge between combinatorics and computer science. In *Formal power series and algebraic combinatorics (New Brunswick, NJ, 1994)*, volume 24 of *DIMACS Ser. Discrete Math. Theoret. Comput. Sci.*, pages 71–87. Amer. Math. Soc., Providence, RI, 1996.
35. J. Denef and L. Lipshitz. Decision problems for differential equations. *J. Symbolic Logic*, 54(3):941–950, 1989.
36. Timothy DeVries. A case study in bivariate singularity analysis. In *Algorithmic probability and combinatorics*, volume 520 of *Contemp. Math.*, pages 61–81. Amer. Math. Soc., Providence, RI, 2010.
37. R. J. Duffin. Rubel's universal differential equation. *Proc. Nat. Acad. Sci. U.S.A.*, 78(8, part 1):4661–4662, 1981.
38. David S. Dummit and Richard M. Foote. *Abstract algebra*. John Wiley & Sons, Inc., Hoboken, NJ, third edition, 2004.
39. L. Euler. *Institutiones calculi differentialis cum eius usu in analysi finitorum ac doctrina serierum*. Berolini, 1755.
40. Leonhard Euler. *Introductio in analysin infinitorum. 2 Volumes.* Lausannae: Apud Marcum-Michaelem Bousquet & socios., 1748.
41. Graham Everest, Alf van der Poorten, Igor Shparlinski, and Thomas Ward. *Recurrence sequences*, volume 104 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2003.
42. E. Fabry. *Sur les intégrales des équations différentielles linéaires à coefficients rationnels*. PhD thesis, Thèse de Doctorat ès sciences mathématiques, Faculté des Sciences de Paris N. 544, 1885.
43. Mary Celine Fasenmyer. A note on pure recurrence relations. *Amer. Math. Monthly*, 56:14–17, 1949.
44. Giovanni Ferraro. *The rise and development of the theory of series up to the early 1820s*. Sources and Studies in the History of Mathematics and Physical Sciences. Springer, New York, 2008.
45. Stéphane Fischler and Tanguy Rivoal. On the values of G-functions. *Comment. Math. Helv.*, 89(2):313–341, 2014.
46. Philippe Flajolet. Analytic models and ambiguity of context-free languages. *Theoret. Comput. Sci.*, 49(2–3):283–309, 1987. Twelfth international colloquium on automata, languages and programming (Nafplion, 1985).
47. Philippe Flajolet, Stefan Gerhold, and Bruno Salvy. On the non-holonomic character of logarithms, powers, and the nth prime function. *Electron. J. Combin.*, 11(2):Article 2, 16, 2004/06.
48. Philippe Flajolet and Andrew Odlyzko. Singularity analysis of generating functions. *SIAM J. Discrete Math.*, 3(2):216–240, 1990.
49. Philippe Flajolet, Bruno Salvy, and Paul Zimmermann. Automatic average-case analysis of algorithms. *Theoret. Comput. Sci.*, 79(1, (Part A)):37–109, 1991. Algebraic and computing treatment of noncommutative power series (Lille, 1988).
50. Philippe Flajolet and Robert Sedgewick. *Analytic combinatorics*. Cambridge University Press, Cambridge, 2009.
51. G. Frobenius. Ueber die Integration der linearen Differentialgleichungen durch Reihen. *J. Reine Angew. Math.*, 76:214–235, 1873.
52. L. Fuchs. Zur Theorie der linearen Differentialgleichungen mit veränderlichen Coefficienten. (Ergänzungen zu der im 66$^{\text{sten}}$ Bande dieses Journals enthaltenen Abhandlung). *J. Reine Angew. Math.*, 68:354–385, 1868.
53. Stavros Garoufalidis. G-functions and multisum versus holonomic sequences. *Adv. Math.*, 220(6):1945–1955, 2009.
54. Scott Garrabrant and Igor Pak. Pattern avoidance is not p-recursive. *arXiv preprint, arXiv:1505.06508*, 2015.

55. Ira M. Gessel. A probabilistic method for lattice path enumeration. *J. Statist. Plann. Inference*, 14(1):49–58, 1986.
56. Ira M. Gessel. Symmetric functions and P-recursiveness. *J. Combin. Theory Ser. A*, 53(2):257–285, 1990.
57. R. William Gosper, Jr. Decision procedure for indefinite hypergeometric summation. *Proc. Nat. Acad. Sci. U.S.A.*, 75(1):40–42, 1978.
58. Xavier Gourdon and Bruno Salvy. Effective asymptotics of linear recurrences with rational coefficients. In *Proceedings of the 5th Conference on Formal Power Series and Algebraic Combinatorics (Florence, 1993)*, volume 153, pages 145–163, 1996.
59. Daniel S. Graça, Manuel L. Campagnolo, and Jorge Buescu. Computability with polynomial differential equations. *Adv. in Appl. Math.*, 40(3):330–349, 2008.
60. Jeremy Gray. *Linear differential equations and group theory from Riemann to Poincaré*. Birkhäuser Boston, Inc., Boston, MA, 1986.
61. J. Hadamard. History of science and psychology of invention. *Mathematika*, 1:1–3, 1954.
62. Charlotte Hardouin. Galoisian approach to differential transcendence. In *Galois theories of linear difference equations: an introduction*, volume 211 of *Math. Surveys Monogr.*, pages 43–102. Amer. Math. Soc., Providence, RI, 2016.
63. G. H. Hardy and S. Ramanujan. Asymptotic Formulæ in Combinatory Analysis. *Proc. London Math. Soc. (2)*, 17:75–115, 1918.
64. William A. Harris, Jr. and Yasutaka Sibuya. The reciprocals of solutions of linear ordinary differential equations. *Adv. in Math.*, 58(2):119–132, 1985.
65. Peter Henrici. *Applied and computational complex analysis. Vol. 1*. Wiley Classics Library. John Wiley & Sons, Inc., New York, 1988. Power series—integration—conformal mapping—location of zeros, Reprint of the 1974 original, A Wiley-Interscience Publication.
66. Einar Hille. *Analytic function theory. Vol. II*. Introductions to Higher Mathematics. Ginn and Co., Boston, Mass.-New York-Toronto, Ont., 1962.
67. R. Jungen. Sur les séries de Taylor n'ayant que des singularités algébrico-logarithmiques sur leur cercle de convergence. *Comment. Math. Helv.*, 3(1):266–306, 1931.
68. Nicholas M. Katz. Nilpotent connections and the monodromy theorem: Applications of a result of Turrittin. *Inst. Hautes Études Sci. Publ. Math.*, (39):175–232, 1970.
69. Manuel Kauers, Maximilian Jaroschek, and Fredrik Johansson. Ore polynomials in Sage. In *Computer algebra and polynomials*, volume 8942 of *Lecture Notes in Comput. Sci.*, pages 105–125. Springer, Cham, 2015.
70. George Kenison, Richard Lipton, Joël Ouaknine, and James Worrell. On the skolem problem and prime powers. In *Proceedings of the 45th International Symposium on Symbolic and Algebraic Computation*, ISSAC '20, pages 289–296, New York, NY, USA, 2020. Association for Computing Machinery.
71. Christoph Koutschan. Regular languages and their generating functions: the inverse problem. Master thesis (Diplomarbeit), Friedrich-Alexander-Universität, Erlangen-Nürnberg, Germany, 2005.
72. Christoph Koutschan. Regular languages and their generating functions: the inverse problem. *Theoret. Comput. Sci.*, 391(1-2):65–74, 2008.
73. Christoph Koutschan. Creative telescoping for holonomic functions. In *Computer algebra in quantum field theory*, Texts Monogr. Symbol. Comput., pages 171–194. Springer, Vienna, 2013.
74. G. Kreweras. Sur une classe de problémes liés au treillis des partitions d'entiers. *Cahiers du B.U.R.O.*, 5–105, 1965.
75. Pierre-Simon Laplace. Mémoire sur les suites. *Mémoire de l'Académie royale des sciences de Paris. Mémoires mathématiques et physique (in Œuvre Complètes 10:1–89)*, pages 207–309, 1779.
76. Christer Lech. A note on recurring series. *Ark. Mat.*, 2:417–421, 1953.
77. Leonard Lipshitz and Lee A. Rubel. A gap theorem for power series solutions of algebraic differential equations. *Amer. J. Math.*, 108(5):1193–1213, 1986.
78. K. Mahler. Eine arithmetische eigenschaft der taylor-koeffizienten rationaler funktionen. *Proc. Akad. Wet. Amsterdam*, 38:50–60, 1935.

79. Edmond Maillet. Sur les séries divergentes et les équations différentielles. *Ann. Sci. École Norm. Sup. (3)*, 20:487–518, 1903.
80. Marc Mezzarobba. *Autour de l'évaluation numérique des fonctions D-finies*. Thèse de doctorat, École polytechnique, November 2011.
81. Marc Mezzarobba. Rigorous multiple-precision evaluation of D-finite functions in SageMath. Technical Report 1607.01967, arXiv, 2016.
82. Marc Mezzarobba. Truncation bounds for differentially finite series. *Ann. H. Lebesgue*, 2:99–148, 2019.
83. Marc Mezzarobba and Bruno Salvy. Effective bounds for P-recursive sequences. *J. Symbolic Comput.*, 45(10):1075–1096, 2010.
84. M. Mignotte, T. N. Shorey, and R. Tijdeman. The distance between terms of an algebraic recurrence sequence. *J. Reine Angew. Math.*, 349:63–76, 1984.
85. Isaac Newton. *The method of fluxions and infinite series: with its application to the geometry of curve-lines*. English translation by John Colson from the Latin original. Printed by Henry Woodfall, 1736.
86. John Noonan and Doron Zeilberger. The enumeration of permutations with a prescribed number of "forbidden" patterns. *Adv. in Appl. Math.*, 17(4):381–407, 1996.
87. A. M. Odlyzko. Asymptotic enumeration methods. In *Handbook of combinatorics, Vol. 1, 2*, pages 1063–1229. Elsevier, Amsterdam, 1995.
88. Joël Ouaknine and James Worrell. Positivity problems for low-order linear recurrence sequences. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 366–379. ACM, New York, 2014.
89. Joël Ouaknine and James Worrell. Ultimate positivity is decidable for simple linear recurrence sequences. In *Automata, languages, and programming. Part II*, volume 8573 of *Lecture Notes in Comput. Sci.*, pages 330–341. Springer, Heidelberg, 2014.
90. Igor Pak. Complexity problems in enumerative combinatorics. In *Proc. Int. Cong. of Math.*, volume 3, pages 3139–3166, Rio de Janeiro, 2018.
91. Marko Petkovšek, Herbert S. Wilf, and Doron Zeilberger. $A = B$. A K Peters, Ltd., Wellesley, MA, 1996. With a foreword by Donald E. Knuth, with a separately available computer disk.
92. G. Pólya. Sur Les Series Entieres a Coefficients Entiers. *Proc. London Math. Soc. (2)*, 21:22–38, 1923.
93. J. Popken. *Über arithmetische Eigenschaften analytischer Funktionen*. PhD thesis, University of Groningen. North-Holland Publishing Co., Amsterdam, 1935.
94. V. Puiseux. Recherches sur les fonctions algébriques. *Journal de mathématiques pures et appliquées 1re série*, 15:365–480, 1850.
95. G. Rozenberg and A. Salomaa, editors. *Handbook of formal languages. Vol. 1*. Springer-Verlag, Berlin, 1997. Word, language, grammar.
96. L. A. Rubel. A universal differential equation. *Bull. Amer. Math. Soc.*, 4(3):345–349, 1981.
97. Lee A. Rubel. Some research problems about algebraic differential equations. *Trans. Amer. Math. Soc.*, 280(1):43–52, 1983.
98. Lee A. Rubel. A survey of transcendentally transcendental functions. *Amer. Math. Monthly*, 96(9):777–788, 1989.
99. Lee A. Rubel. Some research problems about algebraic differential equations. II. *Illinois J. Math.*, 36(4):659–680, 1992.
100. W. Rudin. *Real and complex analysis*. McGraw-Hill Book Co., NY, third edition, 1987.
101. B. Salvy and P. Zimmermann. GFUN: a Maple package for the manipulation of generating and holonomic functions in one variable. *ACM Trans. Math. Softw.*, 20(2):163–177, 1994.
102. Bruno Salvy. Linear differential equations as a data structure. *Foundations of Computational Mathematics*, Jan 2019.
103. Claude E. Shannon. Mathematical theory of the differential analyzer. *J. Math. Phys. Mass. Inst. Tech.*, 20:337–354, 1941.
104. Carl L. Siegel. Über einige Anwendungen diophantischer Approximationen [reprint of Abhandlungen der Preußischen Akademie der Wissenschaften. Physikalisch-mathematische Klasse 1929, Nr. 1]. In *On some applications of Diophantine approximations*, volume 2 of *Quad./Monogr.*, pages 81–138. Ed. Norm., Pisa, 2014.

105. M. F. Singer. Algebraic solutions of nth order linear differential equations. In *Proceedings of the Queen's Number Theory Conference, 1979 (Kingston, Ont., 1979)*, volume 54 of *Queen's Papers in Pure and Appl. Math.*, pages 379–420. Queen's Univ., Kingston, Ont., 1980.
106. Michael F. Singer. Algebraic relations among solutions of linear differential equations. *Trans. Amer. Math. Soc.*, 295(2):753–763, 1986.
107. Parmanand Singh. The so-called Fibonacci numbers in ancient and medieval India. *Historia Math.*, 12(3):229–244, 1985.
108. T. Skolem. Ein verfahren zur behandlung gewisser exponentialer gleichungen und diophantischer gleichungen. *Skand. Mat. Kongr., Stockhohn*, 8(163–188), 1934.
109. Matti Soittola. Positive rational sequences. *Theoret. Comput. Sci.*, 2(3):317–322, 1976.
110. Richard P. Stanley. *Enumerative combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999.
111. Richard P. Stanley. A survey of alternating permutations. In *Combinatorics and Graphs*, volume 531 of *Contemp. Math.*, pages 165–196. Amer. Math. Soc., Providence, RI, 2010.
112. Richard P. Stanley. *Enumerative combinatorics. Volume 1*, volume 49 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2012.
113. E. C. Titchmarsh. *The theory of functions*. Oxford University Press, Oxford, 2nd edition, 1939.
114. W. T. Tutte. Chromatic solutions. II. *Canad. J. Math.*, 34(4):952–960, 1982.
115. Joris van der Hoeven. Fast evaluation of holonomic functions near and in regular singularities. *J. Symbolic Comput.*, 31(6):717–743, 2001.
116. Joris van der Hoeven. Computing with D-algebraic power series. *AAECC*, 30(1):17–49, 2019.
117. Alfred van der Poorten. A proof that Euler missed...Apéry's proof of the irrationality of $\zeta(3)$. *Math. Intelligencer*, 1(4):195–203, 1978/79. An informal report.
118. Marius van der Put and Michael F. Singer. *Galois theory of linear differential equations*, volume 328 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2003.
119. Nikolai Vereshchagin. Occurrence of zero in a linear recursive sequence. *Mathematical notes of the Academy of Sciences of the USSR*, 38(2), 1985.
120. G. Vivanti. Sulle serie di potenze. *Annali di Matematica Pura ed Applicata (1867–1897)*, 21(1):193–194, Jan 1893.
121. Joachim von zur Gathen and Jürgen Gerhard. *Modern computer algebra*. Cambridge University Press, Cambridge, third edition, 2013.
122. Robert J. Walker. *Algebraic Curves*. Princeton Mathematical Series, vol. 13. Princeton University Press, Princeton, N. J., 1950.
123. P. G. Walsh. On the complexity of rational Puiseux expansions. *Pacific J. Math.*, 188(2):369–387, 1999.
124. Wolfgang Wasow. *Asymptotic expansions for ordinary differential equations*. Pure and Applied Mathematics, Vol. XIV. Interscience Publishers John Wiley & Sons, Inc., New York-London-Sydney, 1965.
125. Herbert S. Wilf. *generatingfunctionology*. A K Peters, Ltd., Wellesley, MA, third edition, 2006.
126. Herbert S. Wilf and Doron Zeilberger. Rational functions certify combinatorial identities. *J. Amer. Math. Soc.*, 3(1):147–158, 1990.
127. Jet Wimp and Doron Zeilberger. Resurrecting the asymptotics of linear recurrences. *J. Math. Anal. Appl.*, 111(1):162–176, 1985.
128. Doron Zeilberger. Sister Celine's technique and its generalizations. *J. Math. Anal. Appl.*, 85(1):114–145, 1982.
129. Doron Zeilberger. A holonomic systems approach to special functions identities. *J. Comput. Appl. Math.*, 32(3):321–368, 1990.

Chapter 3
Multivariate Series and Diagonals

The residues of this kind occur naturally in several branches of the algebraic analysis and of the infinitesimal analysis. Their consideration provides simple and easy to use methods which apply to a large number of diverse questions, and new formulas which seem to merit the attention of geometers.
— Augustin-Louis Cauchy

They that are ignorant of Algebra cannot imagine the wonders in this kind are to be done by it: and what further improvements and helps advantageous to other parts of knowledge the sagacious mind of man may yet find out, it is not easy to determine.
— John Locke

In this chapter we study multivariate series, because they encode multivariate sequences but also because they can serve as efficient data structures for univariate sequences. For readability we use bold variables to denote multivariate quantities, typically of dimension $d \in \mathbb{N}_{>0}$, and use multi-index notation to denote multivariate quantities, so that

$$\mathbf{z} = (z_1, \dots, z_d) \qquad \text{and} \qquad d\mathbf{z} = dz_1\, dz_2 \dots dz_d,$$

and for $k \in \{1, \dots, d\}$ and $\mathbf{i} \in \mathbb{R}^d$,

$$\mathbf{z}^{\mathbf{i}} = z_1^{i_1} \cdots z_d^{i_d}, \qquad \mathbf{z}_{\widehat{k}} = (z_1, \dots, z_{k-1}, z_{k+1}, \dots, z_d), \qquad \hat{\mathbf{z}} = \mathbf{z}_{\widehat{d}}.$$

Given a (multivariate) differentiable function $F(\mathbf{z})$ we write $F_{z_j}(\mathbf{z})$ to denote the (partial) derivative of F with respect to the variable z_j.

Definition 3.1 (multivariate formal series) A *multivariate formal power series* over a field $\mathbb{K}$ in the variables $\mathbf{z}$ is a formal expression

$$F(\mathbf{z}) = \sum_{\mathbf{i} \in \mathbb{N}^d} f_{\mathbf{i}} \mathbf{z}^{\mathbf{i}} = \sum_{(i_1, \dots, i_d) \in \mathbb{N}^d} f_{i_1, \dots, i_d} z_1^{i_1} \cdots z_d^{i_d},$$

with *coefficients* $f_{\mathbf{i}} \in \mathbb{K}^d$. Analogously to the univariate case, we define the coefficient extraction operator $[\mathbf{z}^{\mathbf{i}}]F(\mathbf{z}) = f_{\mathbf{i}}$ and make the *ring of multivariate formal power series* $\mathbb{K}[[\mathbf{z}]]$ by defining addition termwise

$$\sum_{\mathbf{i} \in \mathbb{N}^d} f_{\mathbf{i}} \mathbf{z}^{\mathbf{i}} + \sum_{\mathbf{i} \in \mathbb{N}^d} g_{\mathbf{i}} \mathbf{z}^{\mathbf{i}} = \sum_{\mathbf{i} \in \mathbb{N}^d} (f_{\mathbf{i}} + g_{\mathbf{i}})\, \mathbf{z}^{\mathbf{i}}$$

and multiplication by the Cauchy product

$$\left(\sum_{\mathbf{i}\in\mathbb{N}^d} f_{\mathbf{i}}\mathbf{z}^{\mathbf{i}}\right)\left(\sum_{\mathbf{i}\in\mathbb{N}^d} g_{\mathbf{i}}\mathbf{z}^{\mathbf{i}}\right) = \sum_{\mathbf{k}\in\mathbb{N}^d}\left(\sum_{\mathbf{i}+\mathbf{j}=\mathbf{k}} f_{\mathbf{i}}g_{\mathbf{j}}\right)\mathbf{z}^{\mathbf{k}}.$$

Note that the Cauchy product is well defined since for each $\mathbf{k} \in \mathbb{N}^d$ there are only a finite number of elements $\mathbf{i}, \mathbf{j} \in \mathbb{N}^d$ such that $\mathbf{i} + \mathbf{j} = \mathbf{k}$.

One could also define $\mathbb{K}[[\mathbf{z}]]$ iteratively, starting with the univariate formal power series ring $\mathbb{K}[[z_1]]$ and taking $\mathbb{K}[[\mathbf{z}]] = (\mathbb{K}[[z_1, \dots, z_{d-1}]])\,[[z_d]]$; our definition is isomorphic to the iterated one, meaning the iterated definition does not depend on the ordering of the variables (compare this to iterated Laurent series in Section 3.3).

3.1 Complex Analysis in Several Variables

As in the univariate case, we deal mainly with multivariate series over $\mathbb{K} \subset \mathbb{C}$ which represent analytic functions. We now recap the basics of complex analysis in several variables, specialized to fit our needs. Standard accounts of complex analysis in several variables can be found in Hörmander [39] and Krantz [44].

Once again, our most basic analytic object is a convergent power series expansion.

Definition 3.2 (polydisks and polytori) Given a point $\mathbf{a} \in \mathbb{C}^d$ and $\mathbf{r} \in \mathbb{R}^d_{>0}$, the *open polydisk* $D_{\mathbf{a}}(\mathbf{r})$ centred at $\mathbf{z} = \mathbf{a}$ of radius $\mathbf{r}$ is defined as a product of disks

$$D_{\mathbf{a}}(\mathbf{r}) = \{\mathbf{z} \in \mathbb{C}^d : |z_1 - a_1| < r_1, \dots, |z_d - a_d| < r_d\},$$

and the *polytorus* $T_{\mathbf{a}}(\mathbf{r})$ centred at $\mathbf{z} = \mathbf{a}$ of radius $\mathbf{r}$ is defined as a product of circles

$$T_{\mathbf{a}}(\mathbf{r}) = \{\mathbf{z} \in \mathbb{C}^d : |z_1 - a_1| = r_1, \dots, |z_d - a_d| = r_d\}.$$

Remark 3.1 The polytorus $T_{\mathbf{a}}(\mathbf{r})$ is a subset of the boundary $\partial D_{\mathbf{a}}(\mathbf{r})$.

For convenience we often drop the subscript when $\mathbf{a} = \mathbf{0}$, and for an arbitrary complex vector $\mathbf{w} \in \mathbb{C}^d$ with non-zero coordinates we let

$$D(\mathbf{w}) = \{\mathbf{z} \in \mathbb{C}^d : |z_1| < |w_1|, \dots, |z_d| < |w_d|\}$$
$$\text{and} \qquad T(\mathbf{w}) = \{\mathbf{z} \in \mathbb{C}^d : |z_1| = |w_1|, \dots, |z_d| = |w_d|\},$$

denote the polydisk and polytorus with radii $\mathbf{r} = (|w_1|, \dots, |w_d|)$.

Definition 3.3 (analytic functions and absolute convergence) A series $\sum_{n\geq 0} c_n$ with complex summands is called *absolutely convergent* if $\sum_{n\geq 0} |c_n|$ converges. A function $F\colon \mathbb{C}^d \to \mathbb{C}$ is *analytic* at the point $\mathbf{a} \in \mathbb{C}^d$ if there exists a radius $\mathbf{r} \in \mathbb{R}^d_{>0}$ such that $F(\mathbf{z})$ is represented by an absolutely convergent power series

$$F(\mathbf{z}) = \sum_{\mathbf{i}\in\mathbb{N}^d} f_{\mathbf{i}}(\mathbf{z} - \mathbf{a})^{\mathbf{i}} = \sum_{\mathbf{i}\in\mathbb{N}^d} f_{\mathbf{i}}(z_1 - a_1)^{i_1} \cdots (z_d - a_d)^{i_d}$$

for $\mathbf{z} \in D_{\mathbf{a}}(\mathbf{r})$. Such a series is said to be *centred* at $\mathbf{a}$.

At first glance it may not be clear how to sum the terms of a multivariate series. Helpfully, the terms of an absolutely convergent series can be rearranged without affecting convergence or changing the value of the series (see Problem 3.1) so the terms of a power series representing an analytic function can be summed in any reasonable order. The following result implies our insistence on absolute convergence is not a large restriction: if a power series converges at $\mathbf{w} \in \mathbb{C}$ then the power series absolutely converges at any point coordinate-wise closer to the centre of the series. For convenience we state the result for series centred at the origin.

Lemma 3.1 (Abel's Lemma) *If* $\mathbf{w} \in \mathbb{C}^d$ *is such that* $\sup_{\mathbf{i}\in\mathbb{N}^d} |f_{\mathbf{i}}\mathbf{w}^{\mathbf{i}}|$ *is finite then* $\sum_{\mathbf{i}\in\mathbb{N}^d} f_{\mathbf{i}}\mathbf{z}^{\mathbf{i}}$ *converges absolutely for any* $\mathbf{z} \in \mathbb{C}^d$ *with* $|z_j| < |w_j|$ *for all* $1 \leq j \leq d$.

Proof Let $C = \sup_{\mathbf{i}\in\mathbb{N}^d} |f_{\mathbf{i}}\mathbf{w}^{\mathbf{i}}|$. The result holds vacuously if $\mathbf{w}$ has a zero coordinate. If $\mathbf{z} \in \mathbb{C}^d$ is such that the maximum of $\frac{|z_j|}{|w_j|}$ for $1 \leq j \leq d$ is some $t < 1$ then

$$\left|f_{\mathbf{i}}\mathbf{z}^{\mathbf{i}}\right| = \left|f_{\mathbf{i}}\mathbf{w}^{\mathbf{i}}\right| \frac{|\mathbf{z}^{\mathbf{i}}|}{|\mathbf{w}^{\mathbf{i}}|} \leq C t^{i_1+\cdots+i_d},$$

so

$$\sum_{\mathbf{i}\in\mathbb{N}^d} \left|f_{\mathbf{i}}\mathbf{z}^{\mathbf{i}}\right| \leq C \sum_{\mathbf{i}\in\mathbb{N}^d} t^{i_1+\cdots+i_d} = \frac{C}{(1-t)^d}$$

is finite. A bounded series with non-negative terms always converges. □

Definition 3.4 (domains, neighbourhoods, and analyticity) A subset $O \subset \mathbb{C}^d$ is called *open* if for all $\mathbf{z} \in O$ there is an open polydisk centred at $\mathbf{z}$ which is contained in O. Analogously to the univariate case, a *domain* in $\mathbb{C}^d$ is an open and connected subset of $\mathbb{C}^d$ and a *neighbourhood* of $\mathbf{a} \in \mathbb{C}^d$ is an open set containing $\mathbf{a}$. We say $F(\mathbf{z})$ is *analytic* on a domain $\Omega \subset \mathbb{C}^d$ if it is analytic at every point of Ω, and *entire* if it is analytic on all of $\mathbb{C}^d$.

The product and sum of two absolutely convergent series are absolutely convergent, and if $f(\mathbf{z})$ and $g(\mathbf{z})$ are represented by absolutely convergent power series on a domain Ω then the series representations for $f(\mathbf{z}) + g(\mathbf{z})$ and $f(\mathbf{z})g(\mathbf{z})$ are obtained by term-wise summation and the Cauchy product, just as for formal series.

Example 3.1 (A Simple Binomial Sum)

The rational function

$$F(x, y) = \frac{1}{1 - x - y}$$

is analytic for any $(x, y) \in \mathbb{C}^2$ such that $x + y \neq 1$. For instance,

$$F(x, y) = \sum_{n\geq 0}(x+y)^n = \sum_{n\geq 0}\left(\sum_{k\geq 0}\binom{n}{k}x^k y^{n-k}\right) = \sum_{(i,j)\in\mathbb{N}^2}\binom{i+j}{i}x^i y^j$$

when (x, y) is sufficiently close to the origin. Note that there is some subtlety in this derivation, hinting at why we discuss absolute convergence: although the first equality holds whenever $|x + y| < 1$, rearrangement is only possible when the sum converges absolutely. For example, the first series converges when $x = 1$ and $y = -1/2$, but many rearrangements of the final series at these values will diverge. Since $\binom{i+j}{i} \leq 2^{i+j}$ for all $i, j \in \mathbb{N}$, the final power series representation is valid, at least, when both $|x|$ and $|y|$ are strictly less than $1/2$.

When $F(\mathbf{z})$ is analytic in a domain $\Omega \subset \mathbb{C}^d$, and $\mathbf{a} \in \Omega$, fixing

$$\hat{\mathbf{z}} = (z_1, \ldots, z_{d-1}) = (a_1, \ldots, a_{d-1}) = \hat{\mathbf{a}}$$

in an absolutely convergent power series representation of $F(\mathbf{z})$ gives an absolutely convergent power series representation of the univariate function $f(z_d) = F(\hat{\mathbf{a}}, z_d)$ for z_d near a_d. An analytic function on a domain is thus analytic in each variable, when the others are fixed[1]. This observation allows us to generalize many important results from complex analysis in a single variable to the multivariate setting by induction. First is the identity lemma for analytic functions.

Lemma 3.2 (Identity Lemma for Analytic Functions) *If $f(\mathbf{z})$ and $g(\mathbf{z})$ are analytic functions in a domain $\Omega \subset \mathbb{C}^d$ and $f(\mathbf{z}) = g(\mathbf{z})$ on a polydisk contained in Ω then $f(\mathbf{z}) = g(\mathbf{z})$ on Ω.*

Thus, as in the univariate case, we can identify a function $F(\mathbf{z})$ which is analytic at the origin with its power series $F(\mathbf{z}) = \sum_{\mathbf{i} \in \mathbb{N}^d} f_{\mathbf{i}} \mathbf{z}^{\mathbf{i}}$ and call $f_{\mathbf{i}}$ the *power series coefficients* of $F(\mathbf{z})$. The Cauchy integral formula also generalizes to several variables by induction.

Theorem 3.1 (Multivariate Cauchy Integral Formula for Coefficients) *Suppose $F(\mathbf{z})$ is analytic on a domain $\Omega \subset \mathbb{C}^d$ and the closure of a polydisk $D_{\mathbf{a}}(\mathbf{r})$ lies in Ω. If $F(\mathbf{z})$ is represented by the power series*

$$F(\mathbf{z}) = \sum_{\mathbf{i} \in \mathbb{N}^d} f_{\mathbf{i}} (\mathbf{z} - \mathbf{a})^{\mathbf{i}}$$

on $D_{\mathbf{a}}(\mathbf{r})$ then, for all $\mathbf{i} \in \mathbb{N}^d$,

$$f_{\mathbf{i}} = \frac{1}{(2\pi \mathrm{i})^d} \int_{T_{\mathbf{a}}(\mathbf{r})} \frac{F(\mathbf{z})}{(\mathbf{z} - \mathbf{a})^{\mathbf{i}+\mathbf{1}}} d\mathbf{z}.$$

---

[1] In fact, the converse holds. In the 1830s, Cauchy defined a multivariate function to be analytic over a domain $\mathcal{D}$ if it was analytic as a univariate function of each variable at every point in $\mathcal{D}$, and this definition was also used by Jordan. Weierstrass, on the other hand, called a multivariate function analytic in a domain $\mathcal{D}$ if it had a multivariate power series representation in the neighbourhood of any point in the domain (Poincaré also used this definition in this doctoral thesis in 1879). Perhaps illustrating the difficulties working in several variables, the two definitions were not shown to be equivalent until work of Hartogs [37] in 1906. See Bottazzini and Gray [18, Chapter 9] for additional historical information on the development of complex analysis in several variables.

This is the most straightforward generalization of the univariate Cauchy integral formula, but not the only natural one in the multivariate setting: there is no multivariate integral formula which generalizes all major properties of the univariate Cauchy integral formula. For example, the domain of integration in Theorem 3.1 is the polytorus $T_{\mathbf{a}}(\mathbf{r})$, not the boundary of the polydisk $D_{\mathbf{a}}(\mathbf{r})$, so this representation cannot be easily modified to work with domains other than polydisks. Aĭzenberg and Yuzhakov [3, Ch. 1] give a detailed survey of different integral representations for multivariate analytic functions, which generalize different properties of the univariate case.

We will also make repeated use of the implicit function theorem, which helps locally parametrize the solutions of an analytic equation. A proof can be found in Hörmander [39, Thm. 2.1.2].

Proposition 3.1 (Implicit Function Theorem) *If $f(\hat{\mathbf{z}}, y)$ is an analytic function at $\mathbf{w} \in \mathbb{C}^d$ and the partial derivative $f_y(\mathbf{w}) \neq 0$ then for $\hat{\mathbf{z}}$ in a neighbourhood of $\widehat{\mathbf{w}}$ there is a unique analytic function $g(\hat{\mathbf{z}})$ with $g(\widehat{\mathbf{w}}) = w_d$ such that $f(\hat{\mathbf{z}}, y) = 0$ if and only if $y = g(\hat{\mathbf{z}})$.*

3.1.1 Singular Sets of Multivariate Functions

One of the most challenging difficulties moving from one to many variables is the diverse behaviour of the singular sets which arise. For our purposes it is sufficient to consider multivariate generalizations of meromorphic functions, whose singularities are defined (at least in a neighbourhood of each point) by explicit equations. We will see that increasing dimension allows rational functions to capture more complicated behaviour, such as that of algebraic series.

Definition 3.5 (multivariate singularities) Suppose $g(\mathbf{z})$ and $h(\mathbf{z})$ are analytic functions in a domain Ω, with $h(\mathbf{z})$ not identically zero, and define $f(\mathbf{z}) = g(\mathbf{z})/h(\mathbf{z})$ when $h(\mathbf{z}) \neq 0$. Since h is analytic and not identically zero, Lemma 3.2 implies $f(\mathbf{z})$ is defined at some point in any polydisk. We say that f has a *singularity* at $\mathbf{a} \in \Omega$ if $|f(\mathbf{z})|$ is unbounded in any neighbourhood of $\mathbf{a}$ in Ω (where defined).

The denominator $h(\mathbf{z})$ must vanish at any singularity, but $h(\mathbf{z})$ may vanish at a point which is not a singularity.

Example 3.2 (Singularities of Multivariate Functions)

The function $e(x, y) = (x + y)/(x - y)$ has a singularity at every point $(x, y) \in \mathbb{C}^2$ where $x = y$. Note that even though the numerator and denominator both vanish at the origin this is still a singularity, since, for instance,

$$e\left(\frac{1}{n} + \frac{1}{n^2}, \frac{1}{n}\right) = 2n + 1$$

is unbounded as $n \to \infty$. On the other hand, the function $f(x, y) = \sin(x - y)/(x - y)$ has no singularities in $\mathbb{C}^2$. Indeed, taking a power series expansion of $\sin(x - y)$ shows

$$f(x, y) = \frac{\sin(x - y)}{x - y} = \sum_{n\geq 0} \frac{(x - y)^{2n}(-1)^n}{n!}$$

for all (x, y) with $x \neq y$. This series is absolutely convergent for all $x, y \in \mathbb{C}^2$, so $|f(x, y)|$ is bounded in any bounded set.

Definition 3.6 (meromorphic functions) A function $f(\mathbf{z})$ is called *meromorphic* at $\mathbf{a} \in \mathbb{C}^d$ if there exists a polydisk $\mathcal{D}$ centred at $\mathbf{a}$ and analytic functions $g(\mathbf{z})$ and $h(\mathbf{z})$ on $\mathcal{D}$, with h not identically zero, such that $f(\mathbf{z}) = g(\mathbf{z})/h(\mathbf{z})$ in $\mathcal{D}$ whenever $h(\mathbf{z}) \neq 0$. Note that $f(\mathbf{z})$ need not be defined at $\mathbf{a}$, and we say that $\mathbf{a}$ is a *singularity* of f if it is a singularity of $g(\mathbf{z})/h(\mathbf{z})$. A function $f(\mathbf{z})$ is *meromorphic* on a domain $\Omega \subset \mathbb{C}^d$ if it is meromorphic at each point of Ω.

Thus, a meromorphic function is one that can locally be written as a fraction of analytic functions, and a singularity of a meromorphic function is a point where this ratio behaves badly.

By far the most important situation for us is when $F(\mathbf{z}) = G(\mathbf{z})/H(\mathbf{z}) \in \mathbb{C}(\mathbf{z})$ is a rational function, defined and analytic at least on the complement of the zero set of $H(\mathbf{z})$ in $\mathbb{C}^d$. The following result shows that if G and H are coprime polynomials then the zero set of $H(\mathbf{z})$ is precisely the set of singularities of $F(\mathbf{z})$.

Proposition 3.2 *Let $F(\mathbf{z}) = G(\mathbf{z})/H(\mathbf{z}) \in \mathbb{C}(\mathbf{z})$ be the ratio of coprime polynomials G and H. Then the singular set of $F(\mathbf{z})$ is the set $\mathcal{V} = \{\mathbf{z} \in \mathbb{C}^d : H(\mathbf{z}) = 0\}$.*

Proof We need to show that if $H(\mathbf{w}) = 0$ for some $\mathbf{w} \in \mathbb{C}$ then $F(\mathbf{z})$ is unbounded in any neighbourhood of $\mathbf{w}$. Without loss of generality, we may assume that $\mathbf{w} = \mathbf{0}$. If $G(\mathbf{0}) \neq 0$ the result is immediate, since the denominator of F approaches zero as $\mathbf{z} \to \mathbf{0}$ while the numerator of F is bounded away from zero. For the same reason, it is sufficient to prove that any neighbourhood of $\mathbf{0}$ contains a point $\boldsymbol{\zeta}$ such that $G(\boldsymbol{\zeta}) \neq 0$ but $H(\boldsymbol{\zeta}) = 0$. Since $G(\mathbf{z})$ and $H(\mathbf{z})$ are coprime as polynomials, they are coprime as elements of $\mathbb{C}(z_1, \dots, z_{d-1})[z_d]$, so the extended Euclidean algorithm implies the existence of polynomials $a, b, c \in \mathbb{C}[\hat{\mathbf{z}}]$ with a non-zero such that

$$a(\hat{\mathbf{z}}) = b(\hat{\mathbf{z}})G(\mathbf{z}) + c(\hat{\mathbf{z}})H(\mathbf{z}). \tag{3.1}$$

We claim that for any $\hat{\mathbf{z}}$ in a sufficiently small neighbourhood $\mathcal{N}$ of the origin there exists $t_{\hat{\mathbf{z}}} \in \mathbb{C}$ such that $H(\hat{\mathbf{z}}, t_{\hat{\mathbf{z}}}) = 0$. Assuming this, if $G(\hat{\mathbf{z}}, t_{\hat{\mathbf{z}}}) = 0$ for each $\hat{\mathbf{z}} \in \mathcal{N}$ then the non-zero polynomial $a(\hat{\mathbf{z}})$ in (3.1) would vanish in a neighbourhood of the origin, a contradiction. Thus, assuming the claim any sufficiently small neighbourhood of the origin contains a point $\boldsymbol{\zeta}$ with $H(\boldsymbol{\zeta}) = 0$ and $G(\boldsymbol{\zeta}) \neq 0$, as desired.

It remains only to prove the existence of $t_{\hat{\mathbf{z}}}$ for $\hat{\mathbf{z}}$ sufficiently close to the origin. Problem 3.2 asks you to show that after an invertible linear change of variables the polynomial $H(0, 0, \dots, 0, z_d)$ is not identically zero. Since $H(\mathbf{z})$ can be considered a polynomial in z_d whose coefficients are polynomials in $z_1, \dots, z_{d-1}$, this implies $p_{\hat{\mathbf{z}}}(z_d) = H(\hat{\mathbf{z}}, z_d)$ is a non-zero polynomial in z_d whenever $\hat{\mathbf{z}}$ is fixed and sufficiently close to the origin. Any root of $p_{\hat{\mathbf{z}}}(z_d)$ can be taken as $t_{\hat{\mathbf{z}}}$, proving the claim. □

Hörmander [39, Thm 6.2.3] shows that Proposition 3.2 holds for ratios of analytic functions, after suitably defining what it means for two analytic functions to be coprime (in fact, other than preliminary results setting up the local algebraic structure of analytic functions, the proof in the meromorphic case is very similar). If $f(\mathbf{z})$ is analytic at a point $\mathbf{w} \in \mathbb{C}^d$ and $f(\mathbf{w}) \neq 0$ then $1/f(\mathbf{z})$ is also analytic at $\mathbf{w}$. Since a ratio of polynomials (or analytic functions) can be locally reduced to a ratio of coprime polynomials (respectively analytic functions), we thus have the following.

Corollary 3.1 *Let $F(\mathbf{z})$ be a meromorphic function. Then $F(\mathbf{z})$ is analytic at $\mathbf{w} \in \mathbb{C}^d$ if and only if $\mathbf{z} = \mathbf{w}$ is not a singularity of $F(\mathbf{z})$.*

When $G(\mathbf{z})$ and $H(\mathbf{z})$ are coprime and $F(\mathbf{z}) = G(\mathbf{z})/H(\mathbf{z})$ is a rational function with singularity $\mathbf{z} = \mathbf{a}$, there are two types of behaviour to consider.

1) If $H(\mathbf{a}) = 0$ but $G(\mathbf{a}) \neq 0$ then $F(\mathbf{z})$ behaves like a univariate function near a pole: the limit of $|F(\mathbf{z})|$ as $\mathbf{z} \to \mathbf{a}$ equals infinity.

2) If $H(\mathbf{a}) = G(\mathbf{a}) = 0$ then F behaves like a univariate function near an essential singularity: for any fixed $C \in \mathbb{C}$ the equation $F(\mathbf{z}) = C$ has a solution[2] in all sufficiently small neighbourhoods of $\mathbf{a}$.

We will see that finding asymptotics of power series coefficients determined by singularities of the first type leads to more explicit and easier-to-apply formulas, although effective methods still exist for the second case. Since any root of a polynomial in at least two variables is not isolated, any singularity of a rational function in at least two variables is not isolated. Similarly, a zero of an analytic function $f(\mathbf{z})$ in at least two variables is never isolated[3], giving the following result.

Proposition 3.3 *If $F(\mathbf{z})$ is a meromorphic function in a domain $\Omega \subset \mathbb{C}^d$ where $d \geq 2$ then no singularity of $F(\mathbf{z})$ in Ω is isolated.*

Proposition 3.3 hints at a major computational difficulty when applying the methods of analytic combinatorics in several variables: unlike univariate functions, which typically have a finite number of dominant singularities that can be checked one-by-one, in the multivariate setting there will always be an infinite number of singularities to sort through.

---

[2] Note that $I(\mathbf{z}) = G(\mathbf{z}) - CH(\mathbf{z})$ and $H(\mathbf{z})$ are coprime, so our proof of Proposition 3.2 shows every neighbourhood of $\mathbf{a}$ contains a point where $I(\mathbf{z})$ is zero and $H(\mathbf{z})$ is non-zero.

[3] This follows from the Weierstrass preparation theorem [39, Thm. 6.1], which describes the zero set of an analytic function $f(\mathbf{z})$ as the vanishing of a polynomial in z_d whose coefficients are analytic functions in $\hat{\mathbf{z}}$.

3.1.2 Domains of Convergence for Multivariate Power Series

In one variable, the domain of convergence of a power series is a disk centred at the point of expansion. In the multivariate setting, things are more complicated.

Definition 3.7 (domains of convergence) The *(open) domain of convergence* of a power series is the set of points $\mathbf{z} \in \mathbb{C}^d$ such that the power series converges absolutely in some neighbourhood of $\mathbf{z}$.

Note that for us a domain of convergence is always an open set. We will make great use of the fact that domains of convergence are convex subsets of $\mathbb{R}^d$ after taking coordinate-wise moduli and then taking logarithms.

Definition 3.8 (Relog map and logarithmic convexity) Let $\mathbb{C}_*$ be the set of non-zero complex numbers. The *real-logarithm (Relog) map* from $\mathbb{C}_*^d$ to $\mathbb{R}^d$ is the function

$$\mathrm{Relog}(\mathbf{z}) = (\log|z_1|,\ \log|z_2|, \ldots,\ \log|z_d|)\,.$$

For notational convenience we occasionally extend the Relog map to all of $\mathbb{C}^d$ by defining $\log 0 = -\infty$. A set $\Omega \subset \mathbb{C}^d$ is called *logarithmically convex (log-convex)* if its image under the Relog map is convex. Equivalently, Ω is log-convex if $\mathbf{a}, \mathbf{b} \in \Omega$ implies

$$\left(|a_1|^t\,|b_1|^{1-t}, \ldots, |a_d|^t\,|b_d|^{1-t}\right) \in \Omega$$

for all $t \in [0, 1]$.

Proposition 3.4 *Suppose* $F(\mathbf{z}) = \sum_{\mathbf{i}\in\mathbb{N}^d} f_{\mathbf{i}}\mathbf{z}^{\mathbf{i}}$ *is a power series centred at the origin with domain of convergence* $\mathcal{D}$*. Then* $\mathcal{D}$ *is logarithmically convex and a point* $\mathbf{z} \in \mathbb{C}^d$ *lies in the closure* $\overline{\mathcal{D}}$ *if and only if the open polydisk* $D(\mathbf{z})$ *is contained in* $\mathcal{D}$*.*

Proof Lemma 3.1 implies that $\mathcal{D}$ is the interior of the set $B = \{\mathbf{z} : \sup_{\mathbf{i}} |f_{\mathbf{i}}\mathbf{z}^{\mathbf{i}}| \text{ finite}\}$, so it is sufficient to prove B is log-convex. Suppose $\mathbf{x}, \mathbf{y} \in \mathbb{C}^d$ satisfy $\sup_{\mathbf{i}} |f_{\mathbf{i}}\mathbf{x}^{\mathbf{i}}| \le C$ and $\sup_{\mathbf{i}} |f_{\mathbf{i}}\mathbf{y}^{\mathbf{i}}| \le C$ for some $C \ge 0$. If $\mathbf{w} = \left(|x_1|^t\,|y_1|^{1-t}, \ldots, |x_d|^t\,|y_d|^{1-t}\right)$ for some $t \in [0, 1]$ then, for any $\mathbf{i} \in \mathbb{N}^d$,

$$\left|f_{\mathbf{i}}\mathbf{w}^{\mathbf{i}}\right| = \left|f_{\mathbf{i}}\mathbf{x}^{\mathbf{i}}\right|^t \left|f_{\mathbf{i}}\mathbf{y}^{\mathbf{i}}\right|^{1-t} \le C,$$

so $\mathbf{w} \in B$ and B is log-convex. If $\mathbf{z} \in \overline{\mathcal{D}}$ and $\mathbf{w} \in D(\mathbf{z})$ then $|f_{\mathbf{i}}\mathbf{w}^{\mathbf{i}}| < |f_{\mathbf{i}}\mathbf{z}^{\mathbf{i}}|$ so $\mathbf{w} \in \mathcal{D}$ by Lemma 3.1, and $D(\mathbf{z}) \subset \mathcal{D}$. □

To apply the methods of analytic combinatorics in several variables we need to link the analytic behaviour of a convergent power series centred at the origin to its domain of convergence. The following result generalizes Proposition 2.2 in Chapter 2 to the multivariate setting; its proof is Problem 3.3.

Proposition 3.5 *Suppose* $F(\mathbf{z})$ *is a meromorphic function, analytic at the origin where it is represented by a power series with domain of convergence* $\mathcal{D}$*. If* $\mathbf{w} \in \mathbb{C}^d$ *lies on the boundary* $\partial\mathcal{D}$ *then there exists a singularity* $\boldsymbol{\zeta} \in \mathbb{C}^d$ *of* $F(\mathbf{z})$ *with the same coordinate-wise modulus as* $\mathbf{w}$*.*

Definition 3.9 (minimal singularities) A singularity of meromorphic $F(\mathbf{z})$ which is on the boundary of the domain of convergence of its power series at the origin is called a *minimal singularity* or *minimal point*. Equivalently, a minimal point $\mathbf{w}$ is a singularity of F such that no other singularity $\mathbf{z}$ of F satisfies $|z_j| \leq |w_j|$ for each $1 \leq j \leq d$ with at least one of the inequalities being strict.

Example 3.3 (Minimal Singularities of a Binomial Sum)

We have already seen that

$$F(x, y) = \frac{1}{1 - x - y} = \sum_{(i,j)\in\mathbb{N}^2} \binom{i+j}{i} x^i y^j$$

for (x, y) in a neighbourhood of the origin. Since

$$\sum_{(i,j)\in\mathbb{N}^2} \binom{i+j}{i} |x|^i |y|^j = \sum_{n\geq 0} (|x| + |y|)^n = \frac{1}{1 - |x| - |y|},$$

with convergence if and only if $|x|+|y| < 1$, the domain of convergence of the power series for $F(x, y)$ at the origin is $\mathcal{D} = \{(x, y) \in \mathbb{C}^2 : |x| + |y| < 1\}$. If $|x| + |y| \leq 1$ and $x + y = 1$ then x and y must be real and positive, so the minimal singularities of $F(x, y)$ form the set $\mathcal{S} = \{(x, 1 - x) : 0 \leq x \leq 1\}$.

The importance of minimal points will become clear in Chapter 5, when the techniques of analytic combinatorics in several variables are discussed. Of particular interest is the following property of minimal points, which helps determine minimal singularities where a local analysis can yield coefficient asymptotics.

Definition 3.10 (logarithmic gradients) Given an analytic function $f(\mathbf{z})$, the *logarithmic gradient* of f is $(\nabla_{\log} f)(\mathbf{z}) = \left(z_1 f_{z_1}, \ldots, z_d f_{z_d}\right)$, where the subscripts indicate partial derivatives.

Proposition 3.6 *Let $F(\mathbf{z}) = G(\mathbf{z})/H(\mathbf{z})$ be a rational function which is analytic at the origin with domain of convergence $\mathcal{D}$, where G and H are coprime polynomials over the complex numbers. Then for any minimal point $\mathbf{w}$ there exists $\boldsymbol{\lambda} \in \mathbb{R}^d_{\geq 0}$ and $\tau \in \mathbb{C}$ such that $(\nabla_{\log} H)(\mathbf{w}) = \tau\boldsymbol{\lambda}$.*

Proof If the partial derivative $H_{z_j}(\mathbf{w})$ is zero for all $1 \leq j \leq d$ then we can take $\tau = 0$ and the conclusion trivially holds. Thus we may assume, without loss of generality, that the partial derivative $H_{z_d}(\mathbf{w}) \neq 0$. Proposition 3.1, the implicit function theorem, implies the existence of an analytic function $g(\hat{\mathbf{z}}) = g(z_1, \ldots, z_{d-1})$ in a neighbourhood $\mathcal{N}$ of $\widehat{\mathbf{w}}$ in $\mathbb{C}^{d-1}$ such that $(\hat{\mathbf{z}}, g(\hat{\mathbf{z}}))$ is a parameterization of the singular set $\mathcal{V} = \{\mathbf{z} \in \mathbb{C}^d : H(\mathbf{z}) = 0\}$ of F near $\mathbf{w}$.

Fix any $1 \leq j \leq d-1$. Then for $\hat{\mathbf{z}} \in \mathcal{N}$ one has $H(\hat{\mathbf{z}}, g(\hat{\mathbf{z}})) = 0$, and differentiating with respect to z_j followed by substituting $\hat{\mathbf{z}} = \widehat{\mathbf{w}}$ yields

$$H_{z_j}(\mathbf{w}) + g_{z_j}(\widehat{\mathbf{w}})H_{z_d}(\mathbf{w}) = 0. \tag{3.2}$$

Furthermore, taking $z_j = w_j e^{\mathrm{i}\theta_j}$ with θ_j real shows that the derivative of $z_d = g(\hat{\mathbf{z}})$ with respect to θ_j as z_j moves around the circle with modulus $|w_j|$ is $g_{\theta_j} = \mathrm{i}w_j g_{z_j}(\widehat{\mathbf{w}})$ at $\theta_j = 0$. By minimality of $\mathbf{w}$ the path of z_d on this curve at $\theta_j = 0$ must be tangent to the circle $\{w_d e^{\mathrm{i}\theta_d} : -\pi \leq \theta \leq \pi\}$, hence there exists $\lambda_j \in \mathbb{R}$ such that $g_{\theta_j} = -\lambda_j \mathrm{i} w_d$. Putting this together with (3.2) implies

$$-\lambda_j \mathrm{i} w_d = \mathrm{i} w_j g_{z_j}(\widehat{\mathbf{w}}) = -\frac{\mathrm{i} w_j H_{z_j}(\mathbf{w})}{H_{z_d}(\mathbf{w})},$$

and defining $\lambda_d = 1$ gives

$$\lambda_j w_d H_{z_d}(\mathbf{w}) = \lambda_d w_j H_{z_j}(\mathbf{w}),$$

as desired. If, instead of taking $z_j = w_j e^{\mathrm{i}\theta_j}$ we take $z_j = w_j e^{x_j}$ for x_j real then the derivative of $z_d = g(\hat{\mathbf{z}})$ with respect to x_j at $x_j = 0$ is $w_j g_{z_j}(\widehat{\mathbf{w}}) = -\lambda_j w_d$. Since $\mathbf{w}$ is minimal the modulus of z_d cannot decrease when the modulus of z_j decreases and the other variables are held constant, meaning $\lambda_j w_d$ must have the same argument as w_d. Thus $\lambda_j \geq 0$, and as $1 \leq j \leq d-1$ was arbitrary the conclusion holds. □

Remark 3.2 Our proof of Proposition 3.6 did not use the hypothesis that H is a polynomial: if $H(\mathbf{z})$ is any analytic function on a neighbourhood U of $\mathbf{w} \in \mathbb{C}^d$, and no $\mathbf{z} \in U$ with $H(\mathbf{z}) = 0$ has coordinate-wise smaller modulus than $\mathbf{w}$, then the conclusion of Proposition 3.6 still holds.

3.2 Diagonals

Since we use multivariate series as a data structure for enumeration problems, we need ways of extracting univariate sequences from multivariate expansions.

Definition 3.11 (main diagonals) If $F(\mathbf{z}) = \sum_{\mathbf{i}\in\mathbb{N}^d} f_{\mathbf{i}}\mathbf{z}^{\mathbf{i}}$ is a (formal or convergent) multivariate power series, the *main diagonal* of $F(\mathbf{z})$ is the univariate power series

$$(\Delta F)(t) = \sum_{n\geq 0} f_{n,\dots,n} t^n$$

defined by the terms of F where all variables have the same exponent.

If the variable appearing in the diagonal is not specified, then for notational convenience we treat it as the final input variable to F. For instance, if $F(x, y)$ is a bivariate function then ΔF is by default treated as a series in the variable y.

Example 3.4 (Central Binomial Coefficients as a Diagonal)

The diagonal

$$\Delta\left(\frac{1}{1-x-y}\right) = \Delta\left(\sum_{(i,j)\in\mathbb{N}^2}\binom{i+j}{i}x^i y^j\right) = \sum_{n\geq 0}\binom{2n}{n}y^n$$

is the generating function of the central binomial coefficients.

Some of the most interesting applications of analytic combinatorics in several variables, such as combinatorial limit theorems, come from linking the behaviours of different univariate sequences defined by the same multivariate series.

Definition 3.12 (r-diagonals) Given $\mathbf{r} \in \mathbb{Q}^d_{\geq 0}$ we define the **r**-*diagonal* of the series $F(\mathbf{z}) = \sum_{\mathbf{i}\in\mathbb{N}^d} f_{\mathbf{i}}\mathbf{z}^{\mathbf{i}}$ to be the univariate power series

$$(\Delta_{\mathbf{r}}F)(t) = \sum_{n\geq 0} f_{nr_1,\ldots,nr_d}t^n$$

whose coefficients form the sequence $[\mathbf{z}^{n\mathbf{r}}]F(\mathbf{z})$, where $f_{nr_1,\ldots,nr_d}$ is (for now) considered to be zero for $n\mathbf{r} \notin \mathbb{N}^d$. When $F(\mathbf{z})$ is a function analytic in a neighbourhood of the origin, the **r**-diagonal $\Delta_{\mathbf{r}}F$ is defined to be the **r**-diagonal of F's power series representation at the origin.

We always use the diagonal notation Δ without a subscript to refer to the main diagonal $\mathbf{r} = \mathbf{1}$, which we give the most attention. By replacing variables by positive integer powers of themselves, one can always reduce[4] to the case $\mathbf{r} \in \mathbb{N}^d$. Although the coefficient $[\mathbf{z}^{n\mathbf{r}}]F(\mathbf{z})$ is not formally defined for $n\mathbf{r} \notin \mathbb{N}^d$, and never is when $\mathbf{r} \notin \mathbb{Q}^d$, we will see in later chapters that asymptotics of **r**-diagonals typically exhibit uniform behaviour as **r** varies smoothly. This will allow us to make *asymptotic statements* about **r**-diagonals for **r** having arbitrary real number coordinates.

We discuss many examples of rational diagonals in Section 3.4 below.

3.2.1 Properties of Diagonals

One of the key characteristics of the diagonal operator is how it interacts with the various classes of generating functions discussed in Chapter 2. Our first result, connecting algebraic functions and bivariate rational diagonals, was known to Pólya [57] in special cases and later proven by Furstenberg [33] and Hautus and Klarner [38].

Proposition 3.7 *Suppose the rational function $F(x, y) = G(x, y)/H(x, y) \in \mathbb{Q}(x, y)$ is analytic at the origin. Then $(\Delta F)(y)$ is an algebraic function, obtained by adding the residues of $\frac{G(x,y/x)}{xH(x,y/x)}$ at its poles $x_1(y), \ldots, x_r(y)$ in x which are fractional power series in y with no constant term (i.e., those which approach zero as $y \to 0$).*

---

[4] For example, $[x^{n/2}y^n]F(x, y) = [x^n y^n]F(x^2, y)$ for any bivariate power series $F(x, y)$.

We give a complex analytic proof of Proposition 3.7; a generalization to series over arbitrary fields is detailed in Stanley [61, Thm. 6.3.3].

Proof Suppose $F(x, y) = \sum_{(i,j)\in\mathbb{N}^2} f_{i,j} x^i y^j$ converges in the polydisk $|x| < R$ and $|y| < S$ for $R, S > 0$. Then for all fixed y with $|y| < RS$ the series

$$x^{-1} F(x, y/x) = \sum_{(i,j)\in\mathbb{N}^2} f_{i,j} x^{i-j-1} y^j$$

converges absolutely and uniformly for all x in the annulus $A_{|y|} = \{|y|/S < |x| < R\}$. If $C_{|y|}$ is a positively oriented circle around the origin staying in $A_{|y|}$ then the Cauchy residue theorem implies

$$\frac{1}{2\pi \mathrm{i}} \int_{C_{|y|}} \frac{F(x, y/x)}{x} dx = \sum_{(i,j)\in\mathbb{N}^2} \left[f_{i,j} y^j \frac{1}{2\pi \mathrm{i}} \int_{C_{|y|}} x^{i-j-1}\, dx \right] = \sum_{j\geq 0} f_{j,j} y^j = (\Delta F)(y),$$

where we may commute the integral and summation by uniform convergence, so

$$(\Delta F)(y) = \frac{1}{2\pi \mathrm{i}} \int_{C_{|y|}} \frac{G(x, y/x)}{xH(x, y/x)} dx = \sum_{i=1}^{r} \operatorname{Res}\left(\frac{G(x, y/x)}{xH(x, y/x)}; x = \rho_i(y) \right),$$

where $\rho_1(y), \ldots, \rho_r(y)$ are the poles of $G(x, y/x)/xH(x, y/x)$ inside $C_{|y|}$. Because the upper radius of $A_{|y|}$ is fixed while the lower radius goes to zero with $|y|$, when $|y|$ is sufficiently small there is always a circle $C_{|y|}$ such that the only poles of F inside $C_{|y|}$ are those that approach the origin as y goes to zero. Each of the poles is algebraic, and thus so is each residue and their sum. □

Bostan et al. [15, Thm. 18] give tight bounds on the algebraic quantities involved. Their results imply that if $F(x, y)$ is a rational function whose numerator and denominator have degrees at most δ in x and in y, then there is a polynomial $P(t, D)$ of degree at most $d^2 4^{\delta+1}$ in t and degree at most 4^δ in D such that $P(t, (\Delta F)(t)) = 0$. See that paper for more precise bounds, and an algorithm to compute such a polynomial P from $F(x, y)$ using $O(8^\delta)$ arithmetic operations with rational numbers.

Example 3.5 (The Algebraic Diagonal of a Binomial Sum)

The generating function $C(y)$ of the central binomial coefficients is the diagonal of the bivariate function $F(x, y) = 1/(1 - x - y)$. The power series expansion of $F(x, y)$ converges in the polydisk $|x| < 1/2$ and $|y| < 1/2$, and

$$x^{-1} F(x, y/x) = \sum_{(i,j)\in\mathbb{N}^2} \binom{i+j}{i} x^{i-1} \left(\frac{y}{x}\right)^j$$

converges absolutely and uniformly whenever $|y| < 1/4$ and $2|y| < |x| < 1/2$. For fixed y with $|y| < 1/6$ all points on the curve $|x| = 3|y|$ satisfy $2|y| < |x| < 1/2$, so

$$C(y) = (\Delta F)(y) = \frac{1}{2\pi i}\int_{|x|=3|y|}\frac{1}{x(1-x-y/x)}dx = \frac{1}{2\pi i}\int_{|x|=3|y|}\frac{1}{x-x^2-y}dx.$$

The denominator $H(x,y) = x - x^2 - y$ has roots

$$\rho_{\pm}(y) = \frac{1 \pm \sqrt{1-4y}}{2},$$

where $\rho_-(y) \to 0$ and $\rho_+(y) \to 1/2$ as $y \to 0$. Thus,

$$C(y) = \operatorname{Res}\left(\frac{1}{x-x^2-y}; x = \rho_-(y)\right) = \frac{1}{1-2\rho_-(y)} = \frac{1}{\sqrt{1-4y}}$$

for y in a neighbourhood of the origin.

Over the rational numbers, or any field of characteristic zero, the diagonal of a rational function in more than two variables need not be algebraic[5].

Example 3.6 (Not All Rational Diagonals are Algebraic)

In Chapter 2 we saw that the function

$$D(z) = \Delta\left(\frac{1}{(1-x-y)(1-w-z)}\right) = \Delta\left[\sum_{i,j\geq 0}\binom{i+j}{i}\binom{k+n}{n}x^i y^j w^k z^n\right] = \sum_{n\geq 0}\binom{2n}{n}^2 z^n$$

is transcendental by examining asymptotics of its coefficients. Problem 3.7 asks you to prove that the trivariate function $F(x,y,z) = 1/(1-x-y-z)$ has transcendental diagonal.

Although the rational diagonal in the last example is not algebraic, it is D-finite. In the 1980s, Christol [22] proved that the diagonal of a rational function is always D-finite. In fact, diagonals satisfy a stronger closure property.

Definition 3.13 (multivariate D-finite functions) Generalizing the univariate setting, if $\mathbb{K}$ is a field of characteristic zero then a power series or analytic function $F(\mathbf{z})$ is called *D-finite* over $\mathbb{K}$ if the $\mathbb{K}(\mathbf{z})$-vector space generated by $F(\mathbf{z})$ and its partial derivatives is finite dimensional. Equivalently, $F(\mathbf{z})$ is D-finite if for each variable z_j

[5] Over a field of prime characteristic, the diagonals of rational [33] and even algebraic [26] functions in any number of variables must be algebraic. There is a beautifully constructed elementary proof of these results using a connection between diagonals, algebraic functions, and certain finite automata taking as input base p expansions of natural numbers; see Lipshitz and Poorten [48] for a survey. Adamczewski and Bell [2] give explicit bounds on the degree and height of the minimal polynomial for the diagonal $(\Delta F \mod p)$ in terms of the prime p, the degree and height of the minimal polynomial of $F(\mathbf{z})$, and the number of variables.

it satisfies a linear differential equation with coefficients in $\mathbb{K}[\mathbf{z}]$ with respect to the partial derivative $\frac{\partial}{\partial z_j}$.

Even though the class of rational (or even algebraic) functions is not closed under taking diagonals, the class of multivariate D-finite functions is closed under taking diagonals.

Theorem 3.2 *If $f(\mathbf{z})$ is D-finite over a field $\mathbb{K}$ then the main diagonal Δf is also D-finite over $\mathbb{K}$.*

See Lipshitz [47] for a proof of Theorem 3.2 using a clever argument enumerating elements in finite-dimensional vector spaces.

An annihilating linear differential equation for the diagonal of a rational (or, more generally, D-finite) function is obtained through the theory of creative telescoping. If $F(\mathbf{z})$ is a rational function whose numerator and denominator have their degree in each variable bounded by δ, Lairez [45] gives an algorithm determining a D-finite equation satisfied by ΔF which runs in complexity $\delta^{O(d)}$. Bostan et al. [16, Thm. 4.2] give a bound on the order and degree of this D-finite equation, both of which are also in $\delta^{O(d)}$. Currently, the most efficient implementation of this creative telescoping machinery for diagonals is a MAGMA package of Lairez [45]. Koutschan [43] maintains a useful Mathematica package, HolonomicFunctions, which calculates an annihilating D-finite equation for the diagonal of a multivariate D-finite function.

Example 3.7 (Creative Telescoping for Binomial Coefficients)

Above we saw that for y sufficiently small the diagonal $G(y)$ of the function $F(x, y) = 1/(1 - x - y)$ is given by the integral

$$G(y) = \frac{1}{2\pi i} \int_\gamma \frac{1}{x - x^2 - y} dx,$$

where γ is a positively oriented circle close to the origin. Example 2.22 of Chapter 2 derived the differential equation $(4y-1)G'(y) + 2G(y) = 0$ for this integral using an identity relating the derivatives of F with respect to x and with respect to y.

Theorem 3.2 restricts the possible asymptotic behaviour of rational diagonal coefficients. Recall from Definition 2.25 in Chapter 2 that $F(z) = \sum_{n\geq 0} f_n z^n \in \mathbb{Q}[[z]]$ is called a *G-function* if f_n and the least common denominator of $f_0, \dots, f_n$ grow at most exponentially with n. In fact, rational diagonals satisfy a stronger property.

Definition 3.14 (global boundedness) A univariate series $F(z) \in \mathbb{Q}[[z]]$ is called *globally bounded* if $F(z)$ represents an analytic function in a neighbourhood of the origin and there exist non-zero $a, b \in \mathbb{Q}$ such that $aF(bz)$ has integer coefficients.

Problem 3.4 asks you to prove that the diagonal of any rational function $F(\mathbf{z}) \in \mathbb{Q}(\mathbf{z})$ that is analytic at the origin is globally bounded, and that every globally bounded function which is analytic at the origin is a G-function. Combining this with Corollary 2.2 in Chapter 2, about the asymptotics of G-functions, then gives the following.

Corollary 3.2 *Suppose $F(\mathbf{z}) \in \mathbb{Q}(\mathbf{z})$ is analytic at the origin. As $n \to \infty$ its diagonal coefficient sequence $[z_1^n \cdots z_d^n]F(\mathbf{z})$ has an asymptotic expansion given by a sum of terms of the form $C\,n^{\alpha}\,\zeta^n\,(\log n)^{\ell}$, where $C \in \mathbb{C}$, $\alpha \in \mathbb{Q}$, $\ell \in \mathbb{N}$, and ζ is algebraic.*

Example 3.8 (A D-Finite Function which is not a Rational Diagonal)

The function $f(z) = e^z = \sum_{n\geq 0} z^n/n!$ is D-finite, but not globally bounded, so it is not the diagonal of a rational function in any number of variables.

Although the results of this section were stated for the main diagonal ΔF, their analogues hold for the $\mathbf{r}$-diagonal $\Delta_{\mathbf{r}}F$ whenever $\mathbf{r}$ has rational coordinates. As previously mentioned, one can immediately reduce to the case $\mathbf{r} \in \mathbb{N}^d$ by taking positive integer powers of each variable. Then, the $\mathbf{r}$-diagonal of a rational (respectively algebraic or D-finite) function $F(\mathbf{z})$ can be written as the main diagonal of a related rational (respectively algebraic or D-finite) function. Problem 3.8 asks you to extend the approach of the following example to the general case.

Example 3.9 (Reduction to Main Diagonal)

Consider the $\mathbf{r} = (2, 3)$ diagonal of the rational function

$$F(x, y) = \frac{1}{1 - x - y} = \sum_{i,j\geq 0} \binom{i+j}{i} x^i y^j.$$

Our first goal is to determine a series expression involving only even powers of x, since these are the only terms which will appear in the $\mathbf{r}$-diagonal. To find such a series, we note that

$$F(-x, y) = \frac{1}{1 + x - y} = \sum_{i,j\geq 0} \binom{i+j}{i} (-1)^i x^i y^j$$

so that

$$G(x, y) = \frac{F(x, y) + F(-x, y)}{2} = \frac{1 - y}{(1 - x - y)(1 + x - y)} = \sum_{i,j\geq 0} \left(\frac{1 + (-1)^i}{2}\right) \binom{i+j}{i} x^i y^j$$
$$= \sum_{i,j\geq 0} \binom{2i+j}{2i} x^{2i} y^j.$$

Similarly, our next goal is to determine a series expression involving only the powers of y in $G(x, y)$ which are divisible by 3. If $\omega = e^{2\mathrm{i}/3}$ is a primitive third root of unity then $1 + \omega^k + \omega^{2k} = 0$ for k an integer not divisible by 3 and $1 + \omega^k + \omega^{2k} = 3$ if 3 divides k. Thus,

$$H(x,y) = \frac{G(x,y) + G(x,\omega y) + G(x,\omega^2 y)}{3} = \sum_{i,j\geq 0} \left(\frac{1+\omega^j+\omega^{2j}}{3}\right)\binom{2i+j}{2i} x^{2i}y^j$$
$$= \sum_{i,j\geq 0} \binom{2i+3j}{2i} x^{2i}y^{3j},$$

and simplifying $H(x,y)$ gives

$$\frac{1-x^2y^3+x^4-y^3-2x^2}{1-x^6+y^6-6x^2y^3+3x^4-2y^3-3x^2} = \sum_{i,j\geq 0} \binom{2i+3j}{2i} x^{2i}y^{3j}.$$

Now that we have an expression involving the terms which appear in the **r**-diagonal, our final step is to convert to a rational function where these terms lie on the main diagonal. Since x always appears to an even power, and y always appears to a power divisible by 3, we can replace x^2 by s and y^3 by t to obtain

$$\frac{1-st+s^2-t-2s}{1-s^3+t^2-6st+3s^2-2t-3s} = \sum_{i,j\geq 0} \binom{2i+3j}{2i} s^i t^j.$$

The main diagonal of this rational function is the **r**-diagonal of $F(x,y)$.

The technique of multiplying power series variables by roots of unity and adding the results to obtain a specific subseries is often referred to as a 'root of unity filter'. Although this process can always be followed to reduce an **r**-diagonal to the main diagonal, the resulting function will typically be more complicated and harder to analyze. We will see later in this book that the methods of analytic combinatorics in several variables allow for asymptotic statements about **r**-diagonals by working directly with **r** as a parameter, so such manipulations are not necessary.

3.2.2 Representing Series Using Diagonals

Since we develop methods to determine asymptotics of diagonal sequences, it is natural to ask which functions can be represented by diagonals. First, we see that any algebraic function can be written as the diagonal of a bivariate rational function, meaning the classes of bivariate rational diagonals and algebraic functions are equal.

Proposition 3.8 *Suppose $P(z,y) \in \mathbb{Q}[z,y]$ with partial derivative $P_y(0,0) \neq 0$. If $F \in \mathbb{Q}[[z]]$ has no constant term and $P(z,F(z)) = 0$ then*

$$F(z) = \Delta\left(\frac{y^2 P_y(zy,y)}{P(zy,y)}\right);$$

i.e., F is the diagonal of a bivariate rational function.

We follow the simple proof of Furstenberg [33], which works for power series over any integral domain. Note that the restriction that $F(z)$ vanish at the origin is mild: if instead $F(0) = a \neq 0$ then the result can be applied to the function $G(z) = F(z) - a$ to obtain a rational function $g(z, y)$ whose diagonal is $G(z)$, and the diagonal of $g(z, y) + a$ is $F(z)$. To work around the restriction that $P_y(0, 0) \neq 0$ one must separate the solutions of $P(z, y) = 0$ near the origin; see Problem 3.9 for an approach suggested by Adamczewski and Bell [2].

Proof Writing $P(z, y) = (y - F(z))u(z, y)$ for some $u \in \mathbb{Q}[[z]][y]$ we see

$$P_y(z, y) = u(z, y) + (y - F(z))u_y(z, y),$$

so that

$$\frac{y^2 P_y(zy, y)}{P(zy, y)} = \frac{y^2}{y - F(zy)} + \frac{y^2 u_y(zy, y)}{u(zy, y)}. \tag{3.3}$$

Since $F(z)$ has no constant term,

$$\frac{y^2}{y - F(zy)} = \frac{y}{1 - F(zy)/y} = \sum_{k \geq 0} y^{1-k} F(zy)^k$$

is a power series whose diagonal is $F(z)$. Because $u(0, 0) = P_y(0, 0) \neq 0$, the second summand in (3.3) is a power series, with a zero diagonal. Taking the diagonal operator in (3.3) then gives the desired result. □

Example 3.10 (The Catalan Generating Function as a Rational Diagonal)

The Catalan generating function $C(z)$ is the solution to $zC(z)^2 - C(z) + 1 = 0$ with $C(0) = 1$. The function $D(z) = C(z) - 1$ has $D(0) = 0$ and satisfies $P(z, D(z)) = 0$ where $P(z, y) = z(y + 1)^2 - (y + 1) + 1$. Since $P_y(0, 0) = -1 \neq 0$,

$$D(z) = \Delta\left(\frac{y(2y^2 z + 2yz - 1)}{y^2 z + 2yz + z - 1}\right)$$

and

$$C(z) = \Delta\left(1 + \frac{y(2y^2 z + 2yz - 1)}{y^2 z + 2yz + z - 1}\right) = \Delta\left(\frac{2y^3 z + 3y^2 z + 2yz - y + z - 1}{y^2 z + 2yz + z - 1}\right).$$

The following generalization of Proposition 3.8 follows from an analogous argument; see Problem 3.10.

Proposition 3.9 *Let* $P(\mathbf{z}, y) \in \mathbb{Q}[\mathbf{z}, y]$ *and suppose* $F(\mathbf{z})$ *is an analytic function at the origin satisfying* $P(\mathbf{z}, F(\mathbf{z})) = 0$ *and* $F(\hat{\mathbf{z}}, 0) = 0$. *If* $P(\mathbf{z}, y) = (y - F(\mathbf{z}))^k u(\mathbf{z}, y)$ *for* $k \in \mathbb{N}$ *and* $u(\mathbf{z}, y)$ *analytic at the origin with* $u(\mathbf{0}, 0) \neq 0$, *then*

$$[\mathbf{z}^{\mathbf{i}}]F(\mathbf{z}) = [\mathbf{z}^{\mathbf{i}} y^{i_d}] \frac{y^2 P_y(\hat{\mathbf{z}}, yz_d, y)}{kP(\hat{\mathbf{z}}, yz_d, y)}.$$

Denef and Lipshitz [27, Thm. 6.2] extend both Proposition 3.8 and Proposition 3.9 by proving that for *any* algebraic function $F(\mathbf{z})$ in d variables there exists a rational function $R(\mathbf{z}, \mathbf{y})$ in $2d$ variables such that $[\mathbf{z}^{\mathbf{i}}]F(\mathbf{z}) = [\mathbf{z}^{\mathbf{i}}\mathbf{y}^{\mathbf{i}}]R(\mathbf{z}, \mathbf{y})$ for all $\mathbf{i} \in \mathbb{N}^d$. Safonov [59] gives an elementary argument that for any algebraic function $F(\mathbf{z})$ in d variables there exists a rational function $R(\mathbf{z}, y)$ in $d + 1$ variables and a unimodular matrix $A \in \mathbb{N}^{d\times d}$ such that if $\mathbf{i} \in \mathbb{N}^d$ and $\mathbf{j} = A\mathbf{i}$ then $[\mathbf{z}^{\mathbf{i}}]F(\mathbf{z}) = [\mathbf{z}^{\mathbf{j}} y^{j_d}]R(\mathbf{z}, y)$.

Example 3.11 (Counting Leaves in Rooted Plane Trees)

A rooted plane tree is a (non-empty) rooted tree where every node is a leaf (has no children) or has a finite sequence of children which are themselves non-empty rooted plane trees. Problem 3.11 asks you to prove that the bivariate generating function $T(u, z) = uz + uz^2 + (u + u^2)z^3 + \cdots$ whose coefficients $t_{k,n} = [u^k z^n]T(z, u)$ give the number of rooted plane trees on n nodes with k leaves satisfies $P(u, z, T(z, u)) = 0$, where

$$P(u, z, y) = y^2 + (z - zu - 1)y + zu.$$

The conditions of Proposition 3.9 hold with $k = 1$, so

$$t_{k,n} = [u^k z^n y^n] \frac{y^2 P_y(u, yz, y)}{P(u, yz, y)} = [u^k z^n y^n] \frac{y(uyz - yz - 2y + 1)}{1 + uyz - uz - yz - y}.$$

In this case one can determine $t_{k,n}$ directly from the equation $P(u, z, T) = 0$ by a clever application of Lagrange inversion (recall Problem 2.8 from Chapter 2). See Flajolet and Sedgewick [28, Sect. III. 3] for further examples of this nature.

Example 3.12 (Counting Bar Graphs)

A *bar graph* is a finite connected union of 1×1 boxes with corners in $\mathbb{Z}^2$ such that lowest edge in each column lies on the x-axis and each column contains all boxes between its highest and lowest box; see Figure 3.1 for an example. An edge between two vertices of $\mathbb{Z}^2$ is called an *external edge* of a bar graph if there is a 1×1 box in the bar graph containing the edge and a 1×1 box in $\mathbb{Z}^2$ not contained in the bar graph which also contains the edge. If $b_{n,k}$ denotes the number of bar graphs with $2n$ horizontal external edges and $2k$ vertical external edges (there are always an even number of both) then Bousquet-Mélou and Rechnitzer [19] show that the bivariate generating function $B(x, y) = \sum_{n,k\geq 0} b_{n,k} x^n y^k$ satisfies $B(0, 0) = 0$ and $P(x, y, B) = 0$, where $P(x, y, z) = z - xy - (x + y + xy)z + xz^2$. The conditions of Proposition 3.9 are satisfied, so we obtain

$$b_{n,k} = [x^n y^k z^k] \frac{z(1 - xyz - 2xz - yz - x)}{1 - xyz - xy - xz - yz - x}.$$

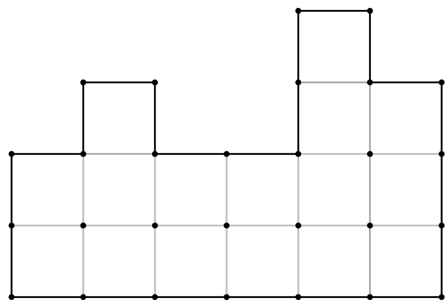


Fig. 3.1 A bar graph with 10 external vertical edges (coloured in gray) and 12 external horizontal edges (coloured in black).

In Chapter 5 we use the methods of analytic combinatorics in several variables to determine asymptotics of $b_{n,k}$ as $n, k \to \infty$.

The fact that any algebraic function can be represented as a bivariate diagonal has asymptotic consequences. For instance, it follows from Proposition 2.11 in Chapter 2 that for any $\alpha \in \mathbb{Q} \setminus \{-1, -2, \dots\}$ there exists a bivariate rational function whose diagonal coefficient sequence has dominant asymptotics of the form $Cn^\alpha \rho^n$ for constants $C, \rho > 0$. Using the methods developed in Chapter 5, for any negative integer α one can construct a trivariate rational function whose main diagonal has dominant asymptotic growth of the form $Cn^\alpha \rho^n$ for constants $C, \rho > 0$.

Although every rational diagonal is a D-finite globally bounded function, it is an open question whether every globally bounded D-finite function can be written as a rational diagonal. This question has been studied for decades.

Example 3.13 (Christol's Open Question)

In 1986, Christol [23, p. 15] noted that the series

$$f(z) = \sum_{n \geq 0} \frac{(1/9)_n (4/9)_n (5/9)_n}{(1/3)_n (1)_n} \frac{z^n}{n!}, \qquad (x)_n = x(x+1)\cdots(x+n-1)$$

a so-called ${}_3F_2$ *hypergeometric function*, is globally bounded and D-finite[6] but is not easily expressed as the diagonal of a rational function. It is still open whether $f(z)$ can be expressed as a rational diagonal. Bostan et al. [14] gave many additional examples of globally bounded hypergeometric functions, some of which were expressed as explicit rational diagonals by Abdelaziz et al. [1] and Bostan and Yurkevich [17].

---

[6] The series $f(z)$ satisfies the D-finite equation $729z^2(z-1)f'''(z) + 81z(37z-21)f''(z) + 9(200z-27)f'(z) + 20f(z) = 0$.

Conjecture 3.1 (Christol [24, Conjecture 4]) Every globally bounded D-finite function can be expressed as the main diagonal of a rational function.

It is also natural to wonder about the relationship between diagonals of rational functions in differing numbers of variables. Given $d \in \mathbb{N}$, let $\mathcal{D}_d$ denote the set of functions which can be obtained as (main) diagonals of d-variate rational functions over $\mathbb{Q}$. If $F(\mathbf{z})$ is any function analytic at the origin, the diagonals of $F(\mathbf{z})$ and $G(\mathbf{z}, t) = F(z_1, \ldots, z_{d-1}, tz_d)$ are equal, giving a sequence of containments $\mathcal{D}_1 \subset \mathcal{D}_2 \subset \mathcal{D}_3 \subset \cdots$. Because $\mathcal{D}_1$ contains univariate rational functions and $\mathcal{D}_2$ is the class of algebraic functions, we have strict containments $\mathcal{D}_1 \subsetneq \mathcal{D}_2 \subsetneq \mathcal{D}_3$, but there is no natural criterion to separate $\mathcal{D}_d$ for $d \geq 3$.

Open Problem 3.1 *Does there exist $d > 3$ such that $\mathcal{D}_3 \neq \mathcal{D}_d$?*

Note that, even when the number of variables d is fixed, there are an infinite number of ways to represent $f \in \mathcal{D}_d$ as a diagonal. Indeed, if $f(z) = (\Delta F)(z)$ for some rational function $F(\mathbf{z})$ then $f(z) = (\Delta G)(z)$ for any $G(\mathbf{z}) = F(\mathbf{z}) + H(\mathbf{z})$ with $H(\mathbf{z})$ a rational function with zero diagonal. One may take, for instance, $H(\mathbf{z}) = z_d h(z_1, \ldots, z_{d-2}, z_{d-1}z_d)$ where $h(w_1, \ldots, w_{d-1})$ is any $(d-1)$-variate rational function that is analytic at the origin. As such representations can have very different properties, characterizing all functions with the same diagonal is an important problem. This can be reduced to characterizing the rational functions with a zero diagonal.

Open Problem 3.2 *Give a useful characterization of the functions $F(\mathbf{z}) \in \mathbb{Q}(\mathbf{z})$ which are analytic at the origin and have zero diagonal.*

The term 'useful characterization' in Open Problem 3.2 is intentionally vague: the characterization should be sufficiently powerful to allow one to select the 'best' representative for a particular purpose, such as an asymptotic analysis. Note that it is decidable to check whether two rational functions $F(\mathbf{z})$ and $G(\mathbf{z})$ have the same diagonal. Creative telescoping algorithms determine an annihilating D-finite equation satisfied by $\Delta(F - G)$, and this equation gives a finite bound N such that the diagonals of F and G are equal if their first N power series coefficients agree.

Example 3.14 (Apéry's Sequence)

We prove that the trivariate rational functions

$$F_1(x, y, z) = \frac{1}{1 - (1 + z)(x + y - xy)}$$
$$F_2(x, y, z) = \frac{1}{1 - x - y - z(1 - x)(1 - y)}$$

have the same diagonal sequence, whose asymptotic expansion is related to Apéry's approach to proving irrationality of the Riemann zeta function at integer arguments (see Section 3.4 below for details). The Magma package of Lairez [45] shows that $(\Delta F_1)(t)$ and $(\Delta F_2)(t)$ both satisfy the differential equation

$$t(t^2 + 11t - 1)f''(t) + (3t^2 + 22t - 1)f'(t) + (t + 3)f(t) = 0,$$

meaning both diagonal sequences satisfy the recurrence relation

$$(n + 2)^2 u_{n+2} = (11n^2 + 33n + 25)u_{n+1} + (n + 1)^2 u_n.$$

Since the first two diagonal terms of F_1 and F_2 are equal, and they satisfy the same recurrence relation of order 2, their diagonal sequences are identical (of course, their non-diagonal power series coefficients are different).

3.3 Multivariate Laurent Expansions and Other Series Operators

In our investigations it will be necessary to consider expansions (and diagonals) of functions which cannot be represented by power series at the origin. Furthermore, non-power series expansions of a rational function centred at the origin can help to give insight into the computations arising in an asymptotic analysis, even when one is only interested in power series expansions. Just as the analysis of an algebraic function encoded by its minimal polynomial P must take into account the other roots of P, our methods to determine asymptotics of a rational function $F(\mathbf{z})$ will naturally pick up properties of all valid series expansions of F at the origin.

Recall from Definition 2.9 in Chapter 2 that the ring of formal Laurent series over the field $\mathbb{K}$ consists of all series with integer exponents bounded from below,

$$\mathbb{K}((z)) = \left\{ \sum_{i \geq q} f_i z^i : q \in \mathbb{Z}, f_i \in \mathbb{K} \right\},$$

equipped with term-wise addition and the Cauchy product, while convergent Laurent series can have an infinite number of terms with negative exponents. We make use of a straightforward multivariate generalization.

Definition 3.15 (multivariate formal Laurent series) The *field of formal iterated Laurent series* in the variables $\mathbf{z}$ is defined inductively as $\mathbb{K}((\mathbf{z})) = \mathbb{K}((z_1, \ldots, z_d)) = \mathbb{K}((z_1, \ldots, z_{d-1}))((z_d))$.

Xin [67, Ch. 2] gives a detailed account of formal iterated Laurent series, and proofs of their basic properties. Unlike multivariate power series, the ordering of the variables matters in this construction.

Example 3.15 (Multivariate Laurent Expansions Depend on Variable Order)

In the field $\mathbb{Q}((x, y)) = \mathbb{Q}((x))((y))$, consisting of a Laurent series in y whose coefficients are Laurent series in x, one can compute the formal expansion

$$\frac{1}{x+y} = \frac{1/x}{1+y/x} = \sum_{n\geq 0}\left((-1)^n x^{-n-1}\right) y^n$$

for the multiplicative inverse of $x + y$. This does not lie in $\mathbb{Q}((y, x)) = \mathbb{Q}((y))((x))$, as it contains terms with arbitrarily large negative exponents in x (note, however, that the coefficient of any *fixed* power of y has only one term with a negative exponent in x). Similarly, in $\mathbb{Q}((y, x))$ one can compute the formal expansion

$$\frac{1}{x+y} = \frac{1/y}{1+x/y} = \sum_{n\geq 0}\left((-1)^n y^{-n-1}\right) x^n$$

which does not lie in $\mathbb{Q}((x, y))$.

Because $\mathbb{K}((\mathbf{z}))$ forms a field, any rational function (or ratio of analytic functions) has an expansion as an iterated Laurent series, making these formal series sufficient for our purposes. Unfortunately, $\mathbb{K}((\mathbf{z}))$ is not closed under composition, even in the univariate case: if $f(z) = 1 + z + z^2 + \cdots$ then $f(1/z)$ is not a well defined formal Laurent series. With this in mind, we define an important subset of Laurent series.

Definition 3.16 (Laurent polynomials) For a variable z, we write $\overline{z} = 1/z$ and $\overline{\mathbf{z}} = (\overline{z}_1, \ldots, \overline{z}_d)$. The *ring of Laurent polynomials* in the variables $\mathbf{z}$, written $\mathbb{K}[\mathbf{z}, \overline{\mathbf{z}}]$, is the subset of $\mathbb{K}((\mathbf{z}))$ consisting of elements with a finite number of non-zero terms.

Since a Laurent polynomial only contains a finite number of terms, if $f(z) \in \mathbb{K}[\mathbf{z}, \overline{\mathbf{z}}]$ and $g \in \mathbb{K}((\mathbf{z}))$ then $f(g(\mathbf{z})) \in \mathbb{K}((\mathbf{z}))$. Note also that the ring of Laurent polynomials does not depend on the ordering of the variables used to construct it.

Example 3.16 (A Multivariate Walk Generating Function)

The ring $\mathbb{Q}[x, \overline{x}][[t]]$ consists of power series in t whose coefficients are Laurent polynomials in x. For instance, in $\mathbb{Q}[x, \overline{x}][[t]]$ one has the expansion

$$\frac{1}{1-t(x+\overline{x})} = \sum_{n\geq 0}(x+\overline{x})^n t^n = \sum_{n\geq 0}\left(\sum_{k=0}^{n}\binom{n}{k} x^{2k-n}\right) t^n. \tag{3.4}$$

This series has a natural interpretation as the multivariate generating function of walks on the integer lattice $\mathbb{Z}$ that begin at the origin and take steps in $\{-1, 1\}$, where the exponent of x marks the endpoint of a walk and the exponent of t marks the number of steps it contains. That the coefficient of t^n is a Laurent polynomial for each n reflects the fact that a walk taking n steps from $\{-1, 1\}$ can only end at a finite number of coordinates. Note that for any $f \in \mathbb{Q}((z))$, we can substitute $x = f(z)$ into both sides of (3.4) and still have a valid equality among elements of $\mathbb{Q}((z))[[t]]$.

Aparicio-Monforte and Kauers [4] discuss constructions of other, more general, rings of formal Laurent series.

3.3.1 Convergent Laurent Series and Amoebas

As usual, we mainly restrict our attention to series representing analytic functions.

Definition 3.17 (multivariate convergent Laurent series) Given domain $\mathcal{D} \subset \mathbb{C}^d$, the set $\mathbb{C}_{\mathcal{D}}\{\mathbf{z}\}$ of *convergent Laurent series on* $\mathcal{D}$ (centred at the origin) consists of series $\sum_{\mathbf{i}\in\mathbb{Z}^d} f_{\mathbf{i}}\mathbf{z}^{\mathbf{i}}$ with $f_{\mathbf{i}} \in \mathbb{C}$ which are absolutely convergent at each point of $\mathcal{D}$ and uniformly convergent on compact subsets of $\mathcal{D}$.

We always consider convergent Laurent expansions centred at the origin. The characterization of multivariate power series domains of convergence discussed above, as well as the Cauchy integral formula, can be extended to convergent Laurent series. Recall the Relog map

$$\mathrm{Relog}(\mathbf{z}) = (\log|z_1|,\ \log|z_2|, \ldots,\ \log|z_d|)\,.$$

Given $\mathbf{x} \in \mathbb{R}^d$, the inverse image $\mathrm{Relog}^{-1}(\mathbf{x})$ of $\mathbf{x}$ under Relog is the polytorus $T(e^{\mathbf{x}})$ of radius $e^{\mathbf{x}} = (e^{x_1}, \ldots, e^{x_d})$. A proof of the following classical result can be found in Pemantle and Wilson [55, Thm. 7.2.2].

Proposition 3.10 *The domain of convergence of the series* $F(\mathbf{z}) = \sum_{\mathbf{i}\in\mathbb{Z}^d} f_{\mathbf{i}}\mathbf{z}^{\mathbf{i}}$ *has the form* $\mathcal{D} = \mathrm{Relog}^{-1}(B)$ *for some (possibly empty) open convex subset* $B \subset \mathbb{R}^d$*, and* F *defines an analytic function on* $\mathcal{D}$*. Conversely, if* $f(\mathbf{z})$ *is an analytic function on* $\mathcal{D} = \mathrm{Relog}^{-1}(B)$ *with* $B \subset \mathbb{R}^d$ *open and convex then there exists a unique series* $F(\mathbf{z}) = \sum_{\mathbf{i}\in\mathbb{Z}^d} c_{\mathbf{i}}\mathbf{z}^{\mathbf{i}} \in \mathbb{C}_{\mathcal{D}}\{\mathbf{z}\}$ *converging to* f*, whose coefficients are given by*

$$c_{\mathbf{i}} = \frac{1}{(2\pi \mathrm{i})^d}\int_{\mathrm{Relog}^{-1}(\mathbf{x})} \frac{f(\mathbf{z})}{\mathbf{z}^{\mathbf{i}+\mathbf{1}}}\,d\mathbf{z} \tag{3.5}$$

for any $\mathbf{x} \in B$*.*

The set of formal expressions $\sum_{\mathbf{i}\in\mathbb{Z}^d} f_{\mathbf{i}}\mathbf{z}^{\mathbf{i}}$ does not have a natural ring structure, which is why we restrict ourselves to iterated Laurent series in the formal case. The set $\mathbb{C}_{\mathcal{D}}\{\mathbf{z}\}$ is, however, a ring when addition is defined term-wise and the multiplication of elements $F, G \in \mathbb{C}_{\mathcal{D}}\{\mathbf{z}\}$ representing analytic functions $f(\mathbf{z})$ and $g(\mathbf{z})$ on $\mathcal{D}$ is defined as the unique element of $\mathbb{C}_{\mathcal{D}}\{\mathbf{z}\}$ converging to the product $f(\mathbf{z})g(\mathbf{z})$; the coefficients of this product can be determined using (3.5).

Amoebas of Laurent Polynomials

There is a nice characterization of the convergent Laurent expansions of a rational function, using an interesting construction from algebraic geometry.

Definition 3.18 (amoebas and their complements) Given a Laurent polynomial $f(\mathbf{z}) \in \mathbb{C}[\mathbf{z}, \overline{\mathbf{z}}]$, the *amoeba* of f is the set

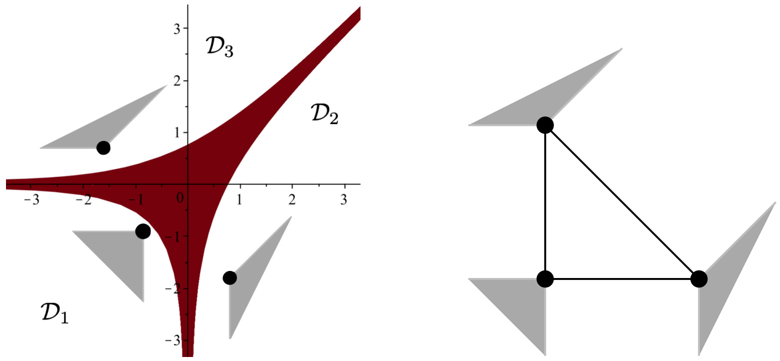


Fig. 3.2 *Left:* The amoeba of $H(x,y) = 1 - x - y$, whose complement in $\mathbb{R}^2$ contains three convex connected components corresponding to convergent Laurent expansions in three domains $\mathcal{D}_1$, $\mathcal{D}_2$, and $\mathcal{D}_3$, pictured with the recession cones of those components. *Right:* The Newton polygon of $H(x,y)$, together with the dual cones at each vertex.

$$\text{amoeba}(f) = \left\{\text{Relog}(\mathbf{z}) : \mathbf{z} \in \mathbb{C}_*^d,\ f(\mathbf{z}) = 0\right\} \subset \mathbb{R}^d,$$

where $\mathbb{C}_* = \mathbb{C} \setminus \{0\}$. The amoeba's *complement* is $\text{amoeba}(f)^c = \mathbb{R}^d \setminus \text{amoeba}(f)$.

Amoebas were introduced to the study of algebraic varieties by Bergman [12], and popularized through the work of Gelfand, Kapranov, and Zelevinsky [35] (who coined the term amoeba as two dimensional amoebas in the plane resemble cellular amoebas with tentacles going off to infinity; see Figure 3.2). The following result, proven in Gelfand, Kapranov, and Zelevinsky [35, Cor. 1.6], shows the connection between amoebas and convergent Laurent expansions.

Proposition 3.11 *If $f(\mathbf{z})$ is a Laurent polynomial then all connected components of the complement* $\text{amoeba}(f)^c$ *are convex subsets of $\mathbb{R}^d$. These real convex subsets are in bijection with the Laurent series expansions of the rational function $1/f(\mathbf{z})$. When $1/f(\mathbf{z})$ has a power series expansion, then it corresponds to the component of $\mathbb{R}^d \setminus$* amoeba(f) *containing all points $(-N, \dots, -N)$ for N sufficiently large.*

Although the introduction of amoebas can seem artificial at first glance, it is actually quite natural. In order to picture the zero set of f, which lives in the extremely hard to visualize space $\mathbb{C}^d$, one takes the absolute value of each coordinate to work over the real numbers. Then, because domains of convergence of Laurent series are logarithmically convex, coordinate-wise logarithms are taken. Amoebas thus represent the 'real shadow' of a complex set, and an understanding of amoebas is crucial to get a good visual picture for power series domains of convergence, singular sets, minimal singularities, and other objects arising in analytic combinatorics in several variables.

Example 3.17 (A Prototypical Amoeba)

The amoeba of $H(x,y) = 1 - x - y$ is drawn in Figure 3.2. As $\mathbb{R}^2 \setminus \text{amoeba}(H)$ has three connected components, there are three convergent Laurent expansions of $F(x, y) = 1/(1 - x - y)$, which converge on three disjoint domains in $\mathbb{C}^2$. In addition to the power series expansion

$$\frac{1}{1-x-y} = \sum_{i,j\geq 0} \binom{i+j}{i} x^i y^j,$$

the binomial theorem implies

$$\frac{1}{1-x-y} = \frac{-1/x}{1-(1-y)/x} = -\sum_{i\geq 0}(1-y)^i x^{-i-1} = \sum_{i,j\geq 0}\binom{i}{j}(-1)^{j+1} y^j x^{-i-1}$$

and

$$\frac{1}{1-x-y} = \frac{-1/y}{1-(1-x)/y} = -\sum_{i\geq 0}(1-x)^i y^{-i-1} = \sum_{i,j\geq 0}\binom{i}{j}(-1)^{j+1} x^j y^{-i-1},$$

when these series absolutely converge. We have already seen that the power series expansion converges on $\mathcal{D}_1 = \{(x, y) : |x| + |y| < 1\}$, and to find the domains of convergence of the Laurent expansions we note

$$\sum_{i,j\geq 0}\binom{i}{j}|y|^j|x|^{-i-1} = \frac{1/|x|}{1-(1+|y|)/|x|}$$

on $\mathcal{D}_2 = \{(x, y) : 1 + |y| < |x|\}$, with divergence outside the closure of $\mathcal{D}_2$, and

$$\sum_{i,j\geq 0}\binom{i}{j}|x|^j|y|^{-i-1} = \frac{1/|y|}{1-(1+|x|)/|y|}$$

on $\mathcal{D}_3 = \{(x, y) : 1 + |x| < |y|\}$, with divergence outside the closure of $\mathcal{D}_3$. Thus, we have found the three convergent Laurent series expansions of $F(x, y)$, valid on the domains $\mathcal{D}_1, \mathcal{D}_2$, and $\mathcal{D}_3$.

Definition 3.19 (minimal points) As for power series expansions, a singularity of a rational (or, more generally, meromorphic) function $F(\mathbf{z})$ which is on the boundary $\partial\mathcal{D}$ of the domain of convergence of one of its convergent Laurent series expansions is called a *minimal singularity* or *minimal point* with respect to that expansion.

Properties of Amoebas

When $F(\mathbf{z}) = G(\mathbf{z})/H(\mathbf{z})$ is a rational function with G and H coprime, the convergent Laurent expansions of F are determined by the connected components of $\mathbb{R}^d \setminus \text{amoeba}(H)$. We now summarize some important properties of amoebas, which help visualize zero sets of multivariate polynomials and properties of Laurent expansions.

1. The Newton Polytope and Amoebas. To state our first collection of results we need a few definitions from the study of convex sets.

Definition 3.20 (definitions from convex analysis)

- The *support* $\mathcal{S}(f) \subset \mathbb{Z}^d$ of a Laurent polynomial $f(\mathbf{z})$ is the finite set of vectors $\mathbf{i} \in \mathbb{Z}^d$ appearing as exponents to monomials $\mathbf{z}^\mathbf{i}$ in f;
- the *Newton polytope* $\mathcal{N}(f) \subset \mathbb{R}^d$ of f is the convex hull of its support $\mathcal{S}(f)$, the smallest convex set in $\mathbb{R}^d$ containing $\mathcal{S}(f)$;
- a *vertex* of $\mathcal{N}(f)$ is a point in $\mathcal{N}(f)$ which does not lie strictly between two other points of the set, also known as an *extreme point*;
- the *dual cone* to a convex set $S \subset \mathbb{R}^d$ at a point $\mathbf{v} \in S$ is the set of vectors $\mathbf{x} \in \mathbb{R}^d$ such that $\mathbf{x} \cdot \mathbf{v} \geq \mathbf{x} \cdot \mathbf{s}$ for all $\mathbf{s} \in S$;
- the *recession cone* of a convex set $B \subset \mathbb{R}^d$ is the set of vectors $\mathbf{x} \in \mathbb{R}^d$ such that $\mathbf{x} + \mathbf{b} \in B$ for all $\mathbf{b} \in B$.

See Figure 3.2 and our examples below for an illustration of these concepts. The following result links the Laurent expansions of a rational function $G(\mathbf{z})/H(\mathbf{z})$ and integer points in the Newton polytope $\mathcal{N}(H)$.

Proposition 3.12 *Let $H \in \mathbb{C}[\overline{\mathbf{z}}, \mathbf{z}]$ be a Laurent polynomial and $\mathbf{B} = \{B_1, \dots, B_r\}$ be the connected components of $\mathbb{R}^d \setminus \text{amoeba}(H)$. Then*

1) *there exists an injective mapping ν from $\mathbf{B}$ to the integer points of $\mathcal{N}(H)$, meaning each connected component of the amoeba complement can be identified with an integer point in the Newton polytope;*
2) *each vertex of $\mathcal{N}(H)$ has a preimage under ν, meaning the number of components in the amoeba complement (and thus Laurent expansions of $1/H$) is at least the number of vertices of $\mathcal{N}(H)$ and at most the number of integer points in $\mathcal{N}(H)$;*
3) *if component $B \in \mathbf{B}$ is the image of $\mathbf{v} \in \mathcal{N}(H)$ under ν, then the dual cone to $\mathcal{N}(H)$ at $\mathbf{v}$ equals the recession cone of B.*

See Forsberg et al. [29] for a proof. Proposition 3.12 allows one to easily compute important information about the amoeba of $H(\mathbf{z})$ directly from the polynomial. In particular, each vertex of the Newton polygon corresponds to a component of the amoeba complement which is unbounded. An integer point $\mathbf{v}$ of the Newton polygon $\mathcal{N}$ that is not a vertex may or may not correspond to a component of the complement; when $\mathbf{v}$ is an interior point of $\mathcal{N}$ it will correspond to a bounded component of the amoeba complement (if any), and when $\mathbf{v}$ is on the boundary of $\mathcal{N}$ it will correspond to an unbounded component of the amoeba complement (if any).

2. Amoeba Boundaries and Contours. Suppose $\mathcal{D}$ is the domain of convergence of a Laurent expansion of rational $F(\mathbf{z}) = G(\mathbf{z})/H(\mathbf{z})$. As in the power series case, if $\mathbf{w} \in \partial\mathcal{D}$ then there is a singularity of F, necessarily minimal, with the same coordinate-wise modulus as $\mathbf{w}$. Note that minimal singularities are precisely those which map to the boundary of amoeba(H) under the Relog map.

Proposition 3.13 *Let $F(\mathbf{z}) = G(\mathbf{z})/H(\mathbf{z})$ be a rational function, where G and H are coprime polynomials over $\mathbb{C}$, and let $\mathcal{V}$ denote the singularities of F (which form the roots of H in $\mathbb{C}^d$). Then for any point $\mathbf{w} \in \partial\mathcal{D} \cap \mathcal{V}$ there exists $\boldsymbol{\lambda} \in \mathbb{R}^d$ and $\tau \in \mathbb{C}$ such that $(\nabla_{\log} H)(\mathbf{w}) = \tau\boldsymbol{\lambda}$. Suppose B is a component of* amoeba$(H)^c$ *with* Relog$(\mathbf{w}) \in \partial B$. *If $\tau \neq 0$ and* $\mathbf{w}$ *has no zero coordinate, then the hyperplane*

$$S_{\mathbf{w}} = \{\mathbf{z} \in \mathbb{R}^d : (\mathbf{z} - \mathrm{Relog}(\mathbf{w})) \cdot \boldsymbol{\lambda} = 0\}$$

is a support hyperplane to B (all points of B lie on one side of $S_{\mathbf{w}}$).

See Figure 3.3 for a visualization.

Proof Existence of $\boldsymbol{\lambda}$ and τ follows from the same proof as Proposition 3.6, which is Proposition 3.13 restricted to power series. For the claim about support hyperplanes, write $w_j = e^{a_j}$ for some $a_j = \log(w_j) \in \mathbb{C}$. If $\tilde{H}(\mathbf{x}) = H(e^{x_1}, \dots, e^{x_d})$ then $(\nabla\tilde{H})(\mathbf{a}) = (\nabla_{\log} H)(\mathbf{w}) = \tau\boldsymbol{\lambda}$ and Taylor's theorem implies that the zero set of $\tilde{H}(\mathbf{x})$ near $\mathbf{a}$ is, up to first order approximation, the hyperplane $\{\mathbf{z} \in \mathbb{C}^d : (\mathbf{z} - \mathbf{a}) \cdot \boldsymbol{\lambda} = 0\}$. The second statement then holds because $\boldsymbol{\lambda}$ has real coordinates, each connected component of amoeba$(H)^c$ is convex, and amoeba(H) is the real part of the zero set of $\tilde{H}$. □

Drawing an amoeba, or its boundary, can be a difficult task[7], and approximation methods are prevalent. Luckily, Proposition 3.13 allows for a natural parameterization of a superset of an amoeba boundary.

Definition 3.21 (amoeba contours) Let $C(H)$ denote the points $\mathbf{z} \in \mathbb{C}_*^d$ such that $H(\mathbf{z}) = 0$ and $(\nabla_{\log} H)(\mathbf{z})$ is a scalar multiple of a vector in $\mathbb{R}^d$. The image of $C(H)$ under the Relog map is called the *contour* of the amoeba of H.

Corollary 3.3 *Let $H \in \mathbb{C}[\overline{\mathbf{z}}, \mathbf{z}]$ be a Laurent polynomial. Then the amoeba boundary satisfies*

$$\partial\, \mathrm{amoeba}(H) \subset \{\mathrm{Relog}(\mathbf{z}) : \mathbf{z} \in C(H)\} \subset \mathrm{amoeba}(H).$$

In two dimensions, the contour can be parametrized by a real variable t by solving

$$xH_x(x, y) - tyH_y(x, y) = 0, \qquad H(x, y) = 0$$

for all complex values of x and y in terms of t, and then taking $(\log|x|, \log|y|)$. Corollary 3.3 was originally given by Mikhalkin [51], and the contour was investigated as

[7] Even determining whether a univariate polynomial with integer coefficients has a root of modulus one, the univariate version of this problem, is NP-hard with respect to the bitsize of the input polynomial coefficients [56].

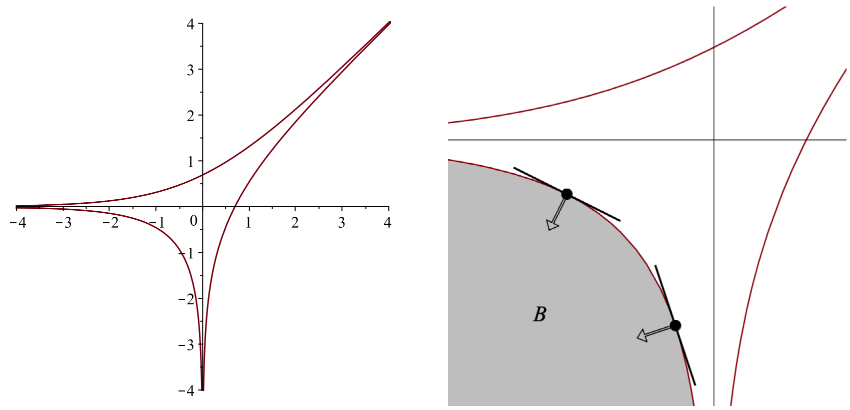


Fig. 3.3 *Left:* The contour of $H(x, y) = 1 - x - y$, parametrized by $\left(\log\left|\frac{t}{1+t}\right|, \log\left|\frac{1}{1+t}\right|\right)$ for $t \in \mathbb{R}$. *Right:* The logarithmic gradients $(\nabla_{\log} H)(\mathbf{w})$ form the normals to support hyperplanes of B at minimal points $\mathbf{w} \in \partial\mathcal{D} \cap \mathcal{V}$.

a tool for computationally drawing amoebas by Theobald [64]. Other approximation methods for amoeba boundaries and contours have been studied in [58, 8, 65, 30].

Example 3.18 (A Prototypical Contour)

The contour of $H(x, y) = 1 - x - y$, shown in Figure 3.3, is determined by solving

$$-x + ty = 1 - x - y = 0$$

for x and y, and then taking the Relog map, giving the real parameterization

$$\left\{\left(\log\left|\frac{t}{1+t}\right|, \log\left|\frac{1}{1+t}\right|\right) : t \in \mathbb{R}\right\}.$$

In this case, the contour is the boundary of amoeba(H).

Example 3.19 (Sketching a Contour)

Figure 3.4 shows points on the contour of the degree four polynomial $Q(x, y) = 1 - x - y - 6xy - x^2y^2$, together with its Newton polygon $\mathcal{N}(Q)$. Four unbounded components of the amoeba complement are visible, corresponding to the four vertices of $\mathcal{N}(Q)$. Furthermore, the contour suggests that there might be a bounded component of the amoeba complement containing the origin, corresponding to the interior integer point of $\mathcal{N}(Q)$.

To see that the origin is not contained in the amoeba, we must show that there is no root of Q with $|x| = |y| = 1$. In order to algebraically test conditions involving the moduli of coordinates, we must work with variables taking real values. If

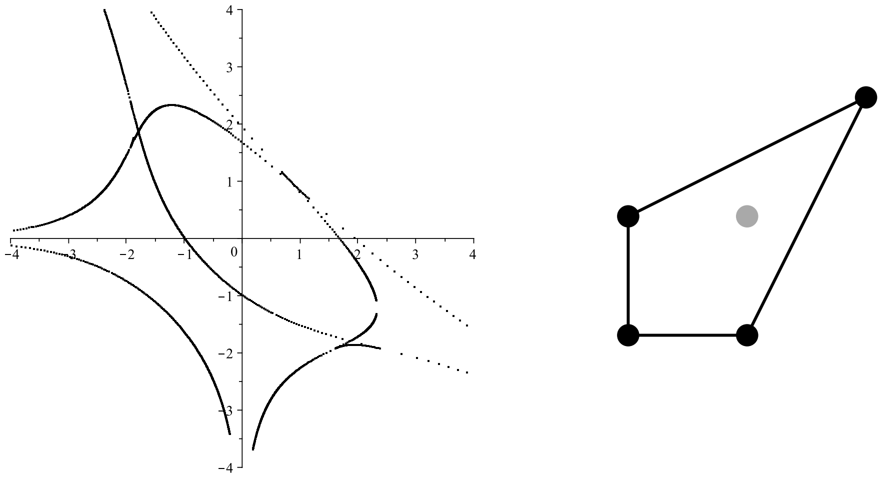


Fig. 3.4 *Left:* Points on the contour of $Q(x, y) = 1 - x - y - 6xy - x^2y^2$, obtained by numerically solving $xQ_x - tyQ_y = Q = 0$ for t incremented by 0.01 between −5 and 5 and taking the Relog map. *Right:* The Newton polytope of $Q(x, y)$.

$$Q(a + \mathrm{i}b, c + \mathrm{i}d) = Q_R(a, b, c, d) + \mathrm{i}\, Q_I(a, b, c, d)$$

for polynomials Q_R and Q_I with real coefficients, then the polynomial Q admits a root with $|x| = |y| = 1$ if and only if the system

$$Q_R(a, b, c, d) = Q_I(a, b, c, d) = a^2 + b^2 - 1 = c^2 + d^2 - 1 = 0$$

has a solution $(a, b, c, d) \in \mathbb{R}^4$. Using a computer algebra package to solve this system shows there is no solution with b or d taking real values.

To conclude that there is a bounded component of the amoeba complement, it is now sufficient to prove that the origin is not contained in any of the four unbounded components of the complement. Using convexity, this can be accomplished by exhibiting points in the amoeba on line segments from the origin to each unbounded component. Such points can be seen directly in Figure 3.4: although points on the contour of Q are not always on the amoeba boundary, they always lie in the amoeba. In Chapter 5 we will see how to easily determine points with explicit algebraic coordinates between the origin and the unbounded components using so-called critical points arising in an asymptotic analysis of $\mathbf{r}$-diagonals for different directions $\mathbf{r}$.

---

3. Amoeba Limit Directions. Finally, we examine how an amoeba goes to infinity.

Definition 3.22 (limit directions of amoebas) A vector $\mathbf{v} \in \mathbb{R}^d$ is called a *limit direction* of amoeba(H) if there exists $\mathbf{x} \in \mathbb{R}^d$ such that $\mathbf{x} + t\mathbf{v} \in$ amoeba(H) for all $t \geq 0$.

Problem 3.14 asks you to prove that if $\mathbf{v}$ is a limit direction of amoeba(H) then there exist distinct vectors $\mathbf{n}$ and $\mathbf{m}$ in $\mathcal{N}(H)$ such that $\mathbf{v}\cdot\mathbf{n} = \mathbf{v}\cdot\mathbf{m} \geq \mathbf{v}\cdot\mathbf{k}$ for all $\mathbf{k} \in \mathcal{N}(H)$. Pictorially, when $\mathbf{v}$ is a limit direction of amoeba(H) then there exists a hyperplane in $\mathbb{R}^d$ with normal $\mathbf{v}$ such that all points of $\mathcal{N}(H)$ lie to one side of the hyperplane and at least two elements of $\mathcal{N}(H)$ lie on the hyperplane.

Limit directions of amoebas help rule out pathological cases for asymptotic analyses. In particular, determining asymptotics of the $\mathbf{r}$-diagonal of $F(\mathbf{z}) = G(\mathbf{z})/H(\mathbf{z})$ can be difficult when amoeba(H) contains a limit direction orthogonal to $\mathbf{r}$.

Example 3.20 (Amoeba Limit Directions)

The amoeba of the polynomial $T(x, y) = 1-x-xy$ has a limit direction $(-1, 1)$ which is orthogonal to the main diagonal direction $(1, 1)$. Expanding $1/T(x, y)$ as a power series using the binomial formula shows that the diagonal coefficient sequence is the trivial sequence $(1, 1, 1, \ldots)$.

3.3.2 Diagonals and Non-Negative Extractions of Laurent Series

Let $F(\mathbf{z}) = \sum_{\mathbf{i}\in\mathbb{Z}^d} f_{\mathbf{i}}\mathbf{z}^{\mathbf{i}}$ be a formal or convergent Laurent series, where $f_{\mathbf{i}}$ is zero for enough indices $\mathbf{i} \in \mathbb{Z}^d$ to make F well defined in the formal case.

Definition 3.23 (diagonals of Laurent expansions) The $\mathbf{r}$-*diagonal* of F for $\mathbf{r} \in \mathbb{Q}^d$ is the univariate series

$$(\Delta_{\mathbf{r}}F)(t) = \sum_{n\geq 0} f_{n\mathbf{r}}t^n,$$

where $f_{n\mathbf{r}} = 0$ if $n\mathbf{r} \notin \mathbb{Z}^d$. We still reserve ΔF for the main diagonal $\mathbf{r} = \mathbf{1}$.

Given a function $f(\mathbf{z})$ over the complex numbers, the diagonal Δf can only be defined after specifying which Laurent expansion of $f(\mathbf{z})$ is under consideration; by Proposition 3.10 this can be done by specifying a point in the domain of convergence of the Laurent expansion. Unless explicitly noted, when given a function analytic at the origin we always consider the *power series expansion* of the function.

Most results discussed above for diagonals of power series expansions of rational functions hold for convergent Laurent series expansions of rational functions. In particular, diagonals of convergent Laurent series with integer coefficients are G-functions, with the corresponding restrictions on asymptotic growth. Proposition 3.10 implies that for $\mathbf{r} \in \mathbb{Z}^d$ the $\mathbf{r}$-diagonal of a convergent Laurent expansion of the rational function $F = G/H$ corresponding to a component $B \subset \text{amoeba}(H)^c$ can be represented as

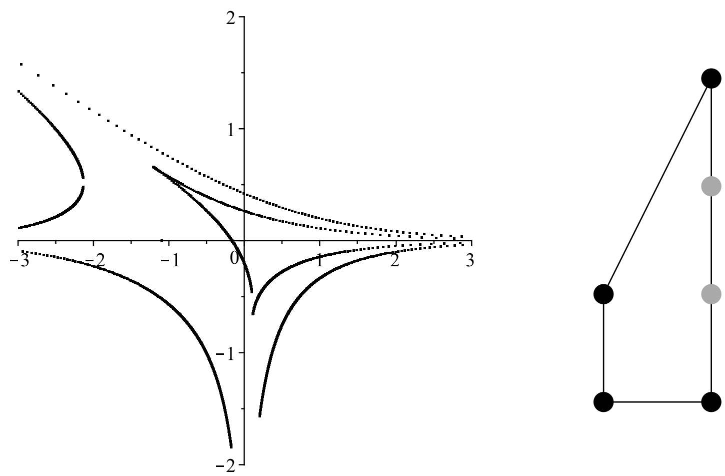


Fig. 3.5 *Left:* Points on the contour of $R(x, y) = 1 - x - y - xy^3$. *Right:* The Newton polygon of R; the non-vertex integer points (shaded gray) do not correspond to convergent Laurent expansions.

$$(\Delta_{\mathbf{r}} F)(t) = \sum_{n\geq 0} f_{n\mathbf{r}} t^n = \sum_{n\geq 0} \left(\frac{1}{(2\pi \mathrm{i})^d} \int_{\mathrm{Relog}^{-1}(\mathbf{x})} \frac{F(\mathbf{z})}{z_1^{r_1 n} \cdots z_d^{r_d n}} \frac{d\mathbf{z}}{z_1 \cdots z_d} \right) t^n$$
$$= \frac{1}{(2\pi \mathrm{i})^d} \int_{\mathrm{Relog}^{-1}(\mathbf{x})} \sum_{n\geq 0} \frac{t^n F(\mathbf{z})}{(z_1^{r_1} \cdots z_d^{r_d})^n} \frac{d\mathbf{z}}{z_1 \cdots z_d}$$
$$= \frac{1}{(2\pi \mathrm{i})^d} \int_{\mathrm{Relog}^{-1}(\mathbf{x})} \frac{\mathbf{z}^{\mathbf{r}-\mathbf{1}} F(\mathbf{z})}{\mathbf{z}^{\mathbf{r}} - t} \, d\mathbf{z},$$

for any $\mathbf{x} \in B$ and $|t| < e^{\mathbf{r}\cdot\mathbf{x}}$. Using this common representation, a D-finite equation satisfied by the $\mathbf{r}$-diagonals of all convergent Laurent expansions of $G(\mathbf{z})/H(\mathbf{z})$ can be computed by the creative telescoping algorithm of Lairez [45].

Example 3.21 (Diagonals and Multiple Laurent Expansions)

The contour of $R(x, y) = 1 - x - y - xy^3$ is shown in Figure 3.5; the amoeba complement $\mathrm{amoeba}(R)^c$ has four components, corresponding to four[8] convergent Laurent expansions of $1/R$. Using the creative telescoping package of Lairez, we can compute the D-finite equation

$$(9t^2 + 6t - 1)(27t^4 + 18t^2 - 1)f''(t) + 6(162t^5 + 135t^4 + 36t^2 - 6t + 1)f'(t)$$
$$+ 6(81t^4 + 81t^3 - 27t^2 - 3t - 4)f(t) = 0$$

[8] Although it appears there may be an additional component containing the origin, this can be ruled out, for instance since $R(1, \mathrm{i}) = 0$ and $(1, \mathrm{i})$ maps to the origin under the Relog map. This illustrates that the contour is a subset of the amoeba boundary.

satisfied by the main diagonal of any convergent Laurent expansion of $1/R$,

$$f(t) = \frac{1}{(2\pi i)^2}\int_\Gamma \frac{dx\,dy}{\left(1-x-y-xy^3\right)(xy-t)}$$

with $\Gamma = \mathrm{Relog}(a,b)$ for some $(a,b) \in \mathrm{amoeba}(R)^c$. This can be converted into an order 6 P-recursive equation satisfied by each diagonal sequence $d_n = [t^n]f(t)$,

$$(n+6)(n+5)d_{n+6} + \cdots + 243(n+1)(n+2)d_n = 0,$$

and, in this instance, even solved in a computer algebra system to get a general solution (as expected with bivariate diagonals, it is a sum of algebraic functions). Taking a power series expansion at the origin shows that the diagonal sequence of *this (power series) expansion* of F begins $1, 2, 6, 23, 90, 357, \ldots$, and these initial terms and the recurrence they satisfy uniquely specify the power series diagonal.

We now turn to the convergent Laurent expansion corresponding to the vertex $(0,1)$ in the Newton polygon of R. Note

$$F(x,y) = \frac{-1/y}{1-\left(1-x-xy^3\right)/y} = -\frac{1}{y}\sum_{k\geq 0}\left(1-x-xy^3\right)^k y^{-k} \tag{3.6}$$

with absolute convergence (at least) whenever $1+|x|+|xy^3| < |y|$; for example, when $|x| = 1/16$ and $|y| = 2$. Expanding $\left((1-x)-xy^3\right)^k y^{-k-1}$ using the binomial theorem shows it contains only monomials of the form

$$x^{k-a+b}y^{2k-3a-1}, \qquad 0 \leq a \leq k, \qquad 0 \leq b \leq a.$$

If $k-a+b = 2k-3a-1$ then $0 \leq k-2a-1 \leq a$, so the exponents of terms appearing on the diagonal are at least $2k-3a-1 \geq (k+1)/2$. Thus, the initial $2N-1$ terms of the series in (3.6) may be summed to determine the initial N terms of its diagonal sequence. In particular, the diagonal sequence of *this (Laurent) expansion* of F begins $0, 1, -3, 10, -39, 156, \ldots$, and these initial terms and the recurrence they satisfy uniquely specify this diagonal.

Problem 3.15 asks you to prove that the diagonal sequences of the two other convergent Laurent expansions of F are identically zero, which can be accomplished by computing initial terms of each sequence and using the above recurrence. In Chapter 5 we will see an easy asymptotic argument, not relying on the calculation of any recurrence, which shows that these final two diagonals must be eventually zero.

We end this section by describing another operation on Laurent series which is used extensively in lattice path enumeration.

Definition 3.24 (non-negative series extraction) Given a field $\mathbb{K}$ and a power series in t whose coefficients are formal iterated Laurent series in $\mathbb{K}$,

$$F(\mathbf{z},t) = \sum_{k\geq 0}\left(\sum_{\mathbf{i}\in\mathbb{Z}^d} f_{\mathbf{i},k}\mathbf{z}^{\mathbf{i}}\right)t^k \in \mathbb{K}((\mathbf{z}))[[t]],$$

the *non-negative series extraction* operator $\left[z_1^{\geq 0}\cdots z_d^{\geq 0}\right]$ takes $F(\mathbf{z},t)$ and returns the power series

$$\left[z_1^{\geq 0}\cdots z_d^{\geq 0}\right] F(\mathbf{z},t) = \sum_{k\geq 0}\left(\sum_{\mathbf{i}\in\mathbb{N}^d} f_{\mathbf{i},k}\mathbf{z}^{\mathbf{i}}\right)t^k = \sum_{(k,\mathbf{i})\in\mathbb{N}^{d+1}} f_{\mathbf{i},k}\mathbf{z}^{\mathbf{i}}t^k \in \mathbb{K}[[\mathbf{z},t]]$$

defined by taking only those terms with positive exponents in the $\mathbf{z}$ variables. The non-negative series extraction of a function $F(\mathbf{z},t)$ having a convergent (instead of formal) Laurent series expansion which is a power series in the t variable is defined analogously.

Recall that the formal Laurent series ring $\mathbb{K}((\mathbf{z}))$ implicitly comes with an ordering of the variables, and even though the image of $F(\mathbf{z},t)$ under $\left[z_1^{\geq 0}\cdots z_d^{\geq 0}\right]$ is a power series it may depend on the underlying ordering. When $F(\mathbf{z},t) \in \mathbb{K}[\mathbf{z},\overline{\mathbf{z}}][[t]]$ is a power series in t whose coefficients are Laurent polynomials in $\mathbf{z}$, the image of $F(\mathbf{z},t)$ under $\left[z_1^{\geq 0}\cdots z_d^{\geq 0}\right]$ is independent of how the variables are ordered.

Certain variants of the kernel method for lattice path enumeration, described in Chapter 4, rely heavily on generating function representations using non-negative series extractions of rational functions. The following result gives a relationship between evaluations of a multivariate function encoded as the non-negative series extraction of a Laurent series and a diagonal extraction.

Proposition 3.14 *Let* $F(\mathbf{z},t) \in \mathbb{K}[\mathbf{z},\overline{\mathbf{z}}][[t]]$. *Then for* $\mathbf{a} \in \{0,1\}^d$,

$$\left[z_1^{\geq 0}\cdots z_d^{\geq 0}\right] F(\mathbf{z},t)\Big|_{\mathbf{z}=\mathbf{a}} = \Delta\left(\frac{F(\overline{z}_1,\ldots,\overline{z}_d, z_1\cdots z_d t)}{(1-z_1)^{a_1}\cdots(1-z_d)^{a_d}}\right). \tag{3.7}$$

The left hand side of (3.7) refers to first taking the non-negative extraction of $F(\mathbf{z},t)$ and then substituting $\mathbf{z} = \mathbf{a}$. When $G(\mathbf{z},t) = \left[z_1^{\geq 0}\cdots z_d^{\geq 0}\right] F(\mathbf{z},t)$ is a multivariate generating function tracking parameters in some combinatorial class, setting $z_j = 0$ in G corresponds to counting objects where the jth parameter is 0, while setting $z_j = 1$ corresponds to counting objects where the jth parameter can take any value. Note that the specialization on the left hand side of (3.7), and the substitution on its right-hand side, are well defined as each coefficient of $F(\mathbf{z},t)$ with respect to t is a Laurent polynomial.

We prove Proposition 3.14 after an example.

Example 3.22 (A Constant Term Extraction using Diagonals)

Let

$$F(x,t) = \frac{1}{1-t(x+\overline{x})} = \sum_{n\geq 0}(x+\overline{x})^n t^n \in \mathbb{Q}[x,\overline{x}][[t]].$$

We have seen above that the coefficient $[x^i t^n]F(x,t)$ for $(i,n) \in \mathbb{Z} \times \mathbb{N}$ counts the number of walks taking n steps in $\{-1, 1\}$ which begin at the origin and end at the point $x = i$. Thus, Proposition 3.14 implies

$$\Delta\left(\frac{1}{1-t(1+x^2)}\right) = \Delta F(\overline{x}, xt) = \left[x^{\geq 0}\right] F(x,t)\Big|_{x=0}$$

counts the number of walks on n steps which begin and end at the origin (note one cannot just substitute $x = 0$ directly into the series for $F(x,t)$ as it contains terms with negative exponents in x), while

$$\Delta\left(\frac{F(\overline{x}, xt)}{(1-x)}\right) = \Delta\left(\frac{1}{(1-x)(1-t(1+x^2))}\right) = \left[x^{\geq 0}\right] F(x,t)\Big|_{x=1}$$

counts the number of walks on n steps beginning at the origin and ending at any point $x \geq 0$. As diagonals of bivariate functions these generating functions are algebraic, and annihilating polynomials can be found using Proposition 3.7.

Any walk beginning and ending at the origin must have even length $2n$ and consist of n steps $+1$ and n steps -1, showing combinatorially that

$$\Delta\left(\frac{1}{1-t(1+x^2)}\right) = \sum_{n\geq 0}\binom{2n}{n}t^{2n} = (1-4t^2)^{-1/2}.$$

Proof (of Proposition 3.14) The right hand side of Equation (3.7) is given by

$$\Delta\left[\left(\sum_{k\geq 0} z_1^k\right)^{a_1} \cdots \left(\sum_{k\geq 0} z_d^k\right)^{a_d} \left(\sum_{k\geq 0}\left(\sum_{\mathbf{i}\in\mathbb{Z}^n} f_{\mathbf{i},k}\, z_1^{k-i_1}\cdots z_d^{k-i_d}\right)t^k\right)\right]$$

$$= \Delta\left[\sum_{k\geq 0}\left(\sum_{\mathbf{j}\in\mathbb{N}^d}\sum_{\mathbf{i}\in\mathbb{Z}^d} f_{\mathbf{i},k}\, z_1^{a_1 j_1 - i_1}\cdots z_d^{a_d j_d - i_d}\right)(z_1\cdots z_d t)^k\right],$$

where these manipulations are valid because enough coefficients $f_{\mathbf{i},k}$ are zero to make $F(\mathbf{z}, t)$ a power series in t whose coefficients are Laurent polynomials in $\mathbf{z}$. If $a_j = 0$ for all $1 \leq j \leq d$ then each $i_j = 0$ in the inner sum for any term on the diagonal. If, however, $a_j = 1$ then any terms with i_j non-negative in the inner sum lie on the diagonal. Evaluating the non-negative series extraction at $z_j = 0$ removes all terms with positive powers of z_j, while evaluating at $z_j = 1$ sums all coefficients with non-negative powers of z_j. □

3.4 Sources of Rational Diagonals

In order to further motivate the study of rational diagonals, and their asymptotics, we end this Chapter by describing several domains of mathematics where they naturally appear. Examples discussed here will be analyzed throughout Parts II and III of this text. Additionally, the theory of lattice path enumeration, our main domain of application, is described in Chapters 4, 6, and 10.

3.4.1 Binomial Sums

Given a rational function $F(\mathbf{z})$ which is analytic at the origin, the binomial theorem implies that the power series coefficients of F can be represented as a sum of binomial coefficients. In fact, the converse is also true, although we need a bit of notation to state the formal result.

Definition 3.25 (rational binomial sums) An *integer indexed sequence of dimension d* is a map $u\colon \mathbb{Z}^d \to \mathbb{Q}$. Any such sequence $u_{i_1,\dots,i_s}$ of dimension $s < d$ can be viewed as a sequence of dimension d by defining $u_{i_1,\dots,i_d} = u_{i_1,\dots,i_s}$, and with this identification one may add and multiply integer indexed sequences of any dimension term by term. The *class of rational binomial sums* is the smallest $\mathbb{Q}$-algebra of integer indexed sequences (closed under term by term addition and multiplication, and multiplication by rational numbers) which

- contains the univariate Kronecker delta sequence defined by $\delta_n = 1$ if $n = 0$ and $\delta_n = 0$ otherwise;
- contains all geometric sequences $g_n = C^n$ for non-zero $C \in \mathbb{Q}$;
- contains the bivariate binomial coefficient sequence $b_{n,k} = \binom{n}{k}$, defined to be zero whenever $k < 0$;
- is closed under affine maps on sequence indices;
- is closed under indefinite summation: whenever $u_{\mathbf{i},k}$ is a binomial sum then so is

$$s_{\mathbf{i},m} = \begin{cases} \sum_{k=0}^{m} u_{\mathbf{i},k} & : m \geq 0 \\ 0 & : m = -1 \\ -\sum_{k=m+1}^{-1} u_{\mathbf{i},k} & : m < -1 \end{cases}$$

  We are typically interested in sequences when their indices have non-negative values, in which case $s_{\mathbf{i},m}$ is defined by regular summation of $u_{\mathbf{i},k}$.

Example 3.23 (Building a Binomial Sum)

Since $b_{n,k} = \binom{n}{k}$ is a binomial sum and the class of binomial sums is closed under affine maps on indices, $b_{n+k,k} = \binom{n+k}{k}$ is a binomial sum. The sequence σ_n which is one when $n \geq 0$ and zero otherwise is the indefinite summation of the Kronecker delta sequence, and thus is a binomial sum. Since the class of binomial sums is closed under multiplication, this implies $\sigma_n \binom{n+k}{k}^2 \binom{n}{k}^2$ is a binomial sum. Finally, since the class of binomial sums is closed under summation, the sequence a_n which is zero if $n < 0$ and

$$a_n = \sum_{k=0}^{n} \binom{n}{k}^2 \binom{n+k}{k}^2$$

for $n \geq 0$ is a binomial sum.

The class of univariate binomial sums supported on the natural numbers is exactly the class of diagonal sequences of rational functions.

Proposition 3.15 *A univariate sequence (u_n), zero whenever $n < 0$, is a rational binomial sum if and only if the generating function $U(z) = \sum_{n \geq 0} u_n z^n$ is the main power series diagonal of a rational function over $\mathbb{Q}$ which is analytic at the origin.*

A proof of Proposition 3.15 can be found in Bostan et al. [16, Thm. 3.5], and a Maple package of Lairez[9] allows one to determine a rational diagonal expression for a given binomial sum. Binomial sums appear in several areas of mathematics, especially number theory.

Example 3.24 (Apéry Numbers as Diagonals)

Apéry [5] proved that the constant $\zeta(3) = \sum_{n=1}^{\infty} \frac{1}{n^3}$ was irrational–an open result for hundreds of years–by constructing two sequences of rational numbers whose ratios converge to $\zeta(3)$ at a rate which implies that $\zeta(3)$ is irrational. Alfred van der Poorten's canonical report [66] on the proof gives the following exercise: "Be the first on your block to prove by a 2-line argument that $\zeta(3)$ is irrational" by determining an algebraic relationship between the two sequences and determining the exponential growth of the binomial sum sequence

$$a_n = \sum_{k=0}^{n} \binom{n}{k}^2 \binom{n+k}{k}^2.$$

The integers a_n are often referred to as the Apéry numbers (see the Online Encyclopedia of Integer Sequences entry A005259) and the Maple package of Lairez shows that their generating function satisfies

$$A(z) = \Delta\left(\frac{1}{1 - t(1+x)(1+y)(1+z)(1+y+z+yz+xyz)}\right).$$

---

[9] The Maple package of Lairez, available at `https://github.com/lairez/binomsums`, contains the command **sumtores** which returns a rational function $R(\mathbf{y}, z) \in \mathbb{Q}(\mathbf{y}, z)$ such that the binomial sum under consideration is the main power series diagonal of $y_1 \cdots y_d R(\mathbf{y}, y_1 y_2 \cdots y_d z)$.

In his seminal paper proving diagonals of rational functions are D-finite, Christol [22] gave two diagonal expressions for the Apéry numbers; Bostan et al. [14, Appendix B] list four additional rational diagonal expressions for the generating function of a_n (two containing 5 variables, one containing 6 variables, and one containing 8 variables, none of which is the representation given here). The singular sets of these rational functions exhibit somewhat different behaviours, with our first set of asymptotic methods in Chapter 5 applying to some, and the others needing the more refined methods of Chapter 9.

Apéry also presented a novel proof of the irrationality of $\zeta(2)$ which relies on asymptotics of the binomial sum sequence

$$c_n = \sum_{k=0}^{n} \binom{n}{k}^2 \binom{n+k}{k}.$$

The c_n are also referred to as Apéry numbers (OEIS entry A005258). Apéry himself noted [6] that the generating function $C(z)$ of (c_n) is the diagonal of two trivariate rational functions

$$C(z) = \Delta\left(\frac{1}{1-(1+z)(x+y-xy)}\right) = \Delta\left(\frac{1}{1-x-y-z(1-x)(1-y)}\right),$$

and the Maple package of Lairez shows

$$C(z) = \Delta\left(\frac{1}{1-z(1+x)(1+y)(xy+y+1)}\right).$$

We use the tools of analytic combinatorics in several variables to easily determine asymptotics of these sequences by hand in Chapter 5, and the algorithms of Chapter 7 are able to rigorously and automatically compute their asymptotics.

3.4.2 Irrational Tilings

As seen in Chapter 2, the smallest collection of univariate rational functions containing 0 and z and closed under addition, multiplication, and $f \mapsto 1/(1-zf)$ form the class of $\mathbb{N}$-rational functions, studied for connections to theoretical computer science and combinatorics. There is a natural multivariate generalization.

Definition 3.26 (multivariate $\mathbb{N}$-rational functions) The set of *d-variate* $\mathbb{N}$-*rational functions* is the smallest set of rational functions containing 0 and $z_1, \ldots, z_d$ which is closed under addition, multiplication, and pseudo-inverse $G \mapsto 1/(1-z_jG)$ for each $j = 1, \ldots, d$.

Garrabrant and Pak [34] give a combinatorial characterization of sequences which are diagonals of $\mathbb{N}$-rational functions.

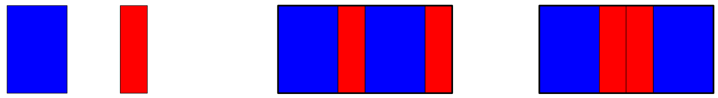


Fig. 3.6 *Left:* Two tiles, of size $1 \times \alpha$ and $1 \times (1-\alpha)$ with $\alpha = 1/\sqrt{2}$. *Right:* Two possible tilings (out of six) of a 1×2 rectangle. There are $\binom{2n}{n}$ possible tilings of a $1 \times n$ rectangle with these tiles.

Definition 3.27 (tile-counting functions) A *tile* is an axis-parallel simply connected closed polygon in the plane of height 1, and a *tiling* of a rectangle R of height 1 with the set of tiles T is a sequence of tiles in T, overlapping only on their boundaries, which cover R (see Figure 3.6). For a fixed set of tiles T and fixed $\varepsilon > 0$, define $f_{T,\varepsilon}(n)$ to be the number of tilings of a $1 \times (n+\varepsilon)$ rectangle using the elements of T for all $n \in \mathbb{N}$. Let $\mathcal{F}$ be the set of all such *tile-counting functions* $f_{T,\varepsilon} : \mathbb{N} \to \mathbb{N}$ as T and ε vary.

Garrabrant and Pak [34, Thm. 1.2] prove the following result.

Proposition 3.16 *The sequence $f(n) \in \mathcal{F}$ if and only if the generating function $\sum_{n \geq 0} f(n) z^n$ is the diagonal of an $\mathbb{N}$-rational function.*

This allows for a combinatorial interpretation of many rational diagonal sequences.

Example 3.25 (Central Binomial Coefficients Enumerate Tilings)

Let $\alpha = 1/\sqrt{2}$. Because α is irrational, tiling a $1 \times n$ rectangle with tiles of size $1 \times \alpha$ and $1 \times (1-\alpha)$, as shown in Figure 3.6, requires exactly n copies of each tile, which can be arranged in any order. Thus, there are $\binom{2n}{n}$ such tilings, and the corresponding generating function is the diagonal of the rational function $F(x, y) = 1/(1 - x - y)$.

By representing the diagonal coefficients of $\mathbb{N}$-rational functions in terms of a restricted class of binomial sums, Garrabrant and Pak [34, Thm. 4.2] strengthen Corollary 3.2 on asymptotics of rational diagonals. In particular, they show that if $F(\mathbf{z})$ is $\mathbb{N}$-rational with main diagonal sequence d_n then there exists $m \in \mathbb{N}$ such that for each $k = 0, \dots, m-1$ the sub-sequence $(d_{mn+k})_{n \geq 0}$ either grows exponentially or is eventually polynomial (when the sequence grows exponentially then its sub-exponential terms can be non-polynomial).

Although there is an algorithm to determine when a univariate function is $\mathbb{N}$-rational, it is currently unknown how to characterize $\mathbb{N}$-rationality in higher dimensions. For example, it is an open question [34, Conj. 4.6] whether the generating function for the Catalan numbers $C_n = \frac{1}{n+1}\binom{2n}{n}$ is the diagonal of an $\mathbb{N}$-rational function (in any number of variables) while as an algebraic function it is the diagonal of a bivariate rational function. The univariate characterization of $\mathbb{N}$-rationality relies heavily on a singularity analysis which does not easily translate into the multivariate case. The techniques developed for the methods of analytic combinatorics in several variables may provide a source of tools to study this problem.

Open Problem 3.3 *Is there an algorithm to determine when a multivariate rational function whose power series at the origin has coefficients in $\mathbb{N}$ is $\mathbb{N}$-rational.*

3.4.3 Period Integrals

The ring of *period numbers*, introduced by Kontsevich and Zagier [42] as "the next most important class in the hierarchy of numbers [after algebraic numbers] according to their arithmetic properties," contains all complex numbers whose real and imaginary parts can be expressed as absolutely convergent integrals of the form

$$\int_\Gamma F(\mathbf{z})d\mathbf{z},$$

where $F(\mathbf{z}) \in \mathbb{Q}(\mathbf{z})$ and $\Gamma \subset \mathbb{R}^n$ is defined by polynomial inequalities with rational coefficients.

Example 3.26 (Period Numbers)

The identities

$$\pi = \int_{x^2+y^2\leq 1} dx\, dy, \qquad \log 2 = \int_1^2 \frac{dx}{x}, \qquad \zeta(3) = \int_{0<x<y<z<1} \frac{dxdydz}{(1-x)yz},$$

show that the constants π, $\log 2$, and $\zeta(3)$ are all period numbers.

The collection of period numbers includes all algebraic numbers, logarithms of algebraic numbers, and all multiple zeta values, but it has been open since the work of Kontsevich and Zagier whether e, Euler's constant γ, and $1/\pi$ are periods (it has been conjectured they are not). It seems to be difficult to find an explicit example of a number which is not a period, although the set of period numbers is countable.

Closely related to period numbers are period integrals of rational functions depending on a parameter.

Definition 3.28 (rational period integrals) The class of *rational period integrals* depending on a parameter t consists of integrals of the form

$$\int_\Gamma F(\mathbf{z}, t)d\mathbf{z},$$

where $F(\mathbf{z}, t) \in \mathbb{Q}(\mathbf{z}, t)$ and Γ is an appropriate domain of integration defined by polynomial inequalities (so that, for example, the integral is absolutely convergent for all values of t in some open subset of the complex plane). Note that one integrand yields multiple period integrals, depending on the domain of integration Γ.

As discussed above, if $F(\mathbf{z}) = \sum_{\mathbf{i}\in\mathbb{N}^d} f_{\mathbf{i}}\mathbf{z}^{\mathbf{i}}$ is a rational function which is analytic at the origin then the multivariate Cauchy integral formula implies

$$(\Delta F)(t) = \frac{1}{(2\pi \mathrm{i})^d} \int_\Gamma \frac{F(\mathbf{z})}{z_1 \cdots z_d - t}\, d\mathbf{z}, \tag{3.8}$$

where Γ is a polytorus sufficiently close to the origin. Since a circle $|z| = \varepsilon$ in the complex plane can be represented by the real algebraic equation $x^2 + y^2 = \varepsilon$, where $z = x + iy$ for real variables x and y, rational diagonals come from rational period integrals and their evaluations at valid rational arguments yield period numbers, up to powers of the (conjecturally not a period number[10]) $1/\pi$. Periods defined by (3.8) for different domains of integration Γ give diagonals of different Laurent expansions of $F(\mathbf{z})$.

Example 3.27 (Periods of Calabi-Yau Threefolds)

Parametrized period integrals defined over cycles on certain algebraic varieties are known to encode important information about the underlying varieties. For instance, such period integrals can be used [53] to count the number of rational curves of fixed degree on hypersurfaces with degree 5 and dimension 3. Much of this theory has been developed for so-called Calabi-Yau threefolds through the use of "mirror symmetry" (see Cox and Katz [25] for details and definitions).

Batyrev and Kreuzer [11] determined a family of Calabi-Yau threefolds, identified by finite sets $P_j \subset \mathbb{Z}^4$, and studied their *principal periods*

$$\overline{\omega}_0(t) = \int_C \frac{1}{1 - t\sum_{\mathbf{v}\in P_j} \mathbf{z}^{\mathbf{v}}} \frac{dz_1 dz_2 dz_3 dz_4}{z_1 z_2 z_3 z_4}$$

where C is a polytorus sufficiently close to the origin. Batyrev and Kreuzer were interested in properties of the Picard-Fuchs differential equations annihilating these integrals: the models break down into 68 classes, of which they were able to guess such equations for the models in 28 classes. Lairez [45] used a fast creative telescoping algorithm to rigorously compute annihilating differential operators for all models[11]. For each polytope P_j, the principal period can be expressed as the diagonal

$$\overline{\omega}_0(t) = \Delta\left(\frac{1}{1 - t(z_1 z_2 z_3 z_4)\sum_{\mathbf{v}\in P_j} \mathbf{z}^{\mathbf{v}}}\right),$$

where the rational function is expanded in the ring $\mathbb{Q}[\mathbf{z}, \overline{\mathbf{z}}][[t]]$.

---

[10] Kontsevich and Zagier call the ring of period numbers extended by powers of $1/\pi$ the *ring of extended periods*. In addition to the connection to Cauchy-type integrals, those authors explain how the ring of extended periods is more natural than the ring of periods from the perspective of motives.

[11] Lairez's complete list of Laurent polynomials $\sum_{\mathbf{v}\in P_j} \mathbf{z}^{\mathbf{v}}$ and their annihilating differential operators can be found at http://pierre.lairez.fr/supp/periods/.

3.4.4 Kronecker Coefficients

Multivariate rational generating functions also appear in representation theory and algebraic combinatorics, although we need to introduce some definitions before seeing an example.

Definition 3.29 (representations of the general linear group) For $r \in \mathbb{N}$ let G_r denote the group of $r \times r$ invertible complex matrices and, for a finite dimensional complex vector space V, let $\mathrm{GL}(V)$ denote the group of linear isomorphisms of V. A group homomorphism $\rho\colon G_r \to \mathrm{GL}(V)$ is called a *representation* of G_r, and a representation is *irreducible* if the only linear subspaces of V that are fixed by all elements of $\rho(G_r) \subset \mathrm{GL}(V)$ are the zero subspace and V itself. An $r \times r$ matrix in G_r is defined by r^2 coordinate variables, and the representation ρ is called a *polynomial representation* if the coordinate functions of ρ are polynomials in these coordinate variables.

Up to isomorphism, the irreducible polynomial representations of G_r have been classified: for each integer partition λ with at most r summands there is an irreducible representation, denoted $V_\lambda(G_r)$, and these form all irreducible polynomial representations of G_r. See Fulton [31, Ch. 8] for details.

Let λ be an integer partition with at most r^2 summands. Using tensor products, the irreducible polynomial representation $V_\lambda(G_{r^2})$ of G_{r^2} can be decomposed as a sum of elements $V_\mu(G_r) \otimes V_\pi(G_r)$, for integer partitions μ and π with at most r summands.

Definition 3.30 (Kronecker coefficients) The multiplicity of $V_\mu(G_r) \otimes V_\pi(G_r)$ in this decomposition is called the *Kronecker coefficient* $k^\lambda_{\mu,\pi}$.

By definition $k^\lambda_{\mu,\pi} \geq 0$, and it can be shown that $k^\lambda_{\mu,\pi} = 0$ unless μ, π, and λ are partitions of the same non-negative integer. As long as r is large enough so that the restriction on the number of summands λ, μ, and π contain is satisfied, the Kronecker coefficient $k^\lambda_{\mu,\pi}$ is independent of r. A detailed account of Kronecker coefficients can be found in Fulton and Harris [32], including many other natural definitions connecting Kronecker coefficients to different algebraic and geometric objects.

Kronecker coefficients have been the subject of intense study in representation theory, algebraic combinatorics, quantum physics, and computer science. For example, positivity of Kronecker coefficients is closely related[12] to the geometric complexity theory approach to resolving the P vs. NP conjecture [21]. From an asymptotic point of view, it is interesting to fix partitions λ, μ, π and study the *dilation sequence* $f_n = k^{n\lambda}_{n\mu,n\pi}$, where multiplying a partition by n corresponds to multiplying each of its summands by n. It follows from work of Meinrenken and Sjamaar [49] that a dilated Kronecker sequence f_n is *quasi-polynomial*, meaning there exists a positive integer r and polynomials $p_1(n), \dots, p_r(n)$ such that $f_n = p_j(n)$ for all $n \equiv j$

[12] It is NP-hard to determine when Kronecker coefficients are positive [40], although it has recently been shown [41] that this result cannot be directly extrapolated to separate the complexity classes P and NP, as was once thought.

mod p. In particular, a dilated Kronecker sequence f_n is C-finite and knowing an asymptotic expansion of $f(n)$ allows one to determine f_n exactly. The complexity of computing Kronecker coefficients is well studied: see, for example, Burgisser and Ikenmeyer [20] or Ikenmeyer et al. [40]. Baldoni and Vergne [9] discuss algorithms for computing dilated Kroencker sequences, and their work has been implemented in a Maple package by Walter[13]. Melczer et al. [50] give asymptotics for a family of Kronecker coefficients which is not a dilated sequence.

Most importantly for this discussion, Mishna et al. [52] give a method to determine a diagonal expression for dilated Kronecker sequences. In fact, their results give rational functions whose **r**-diagonals describe different dilated Kronecker sequences as **r** varies, allowing for unified asymptotic arguments. Combining these diagonal representations with the techniques of analytic combinatorics in several variables is an interesting area for new research.

3.4.5 Positivity Results and Special Functions

Although it is difficult (perhaps undecidable) to determine whether a univariate rational function $F(z) = G(z)/H(z) \in \mathbb{Q}(z)$ has eventually positive Taylor coefficients, in practice it is sometimes possible to decide eventual positivity by looking at asymptotics of the Taylor coefficients $f_n = [z^n]F(z)$. For example, when $H(z)$ has a unique root ρ of smallest modulus then Theorem 2.3 in Chapter 2 implies $f_n \sim Cn^k\rho^{-n}$ for an explicit algebraic number C and non-negative integer k, and f_n is eventually positive if and only if C and ρ are positive real numbers.

Similarly, (eventual) positivity of power series coefficients of multivariate rational functions has been studied at least since work of Friedrichs and Lewy in the 1920s on the discretized time-dependent wave equation in two space dimensions. Those authors asked whether the power series coefficients of

$$F(x, y, z) = \frac{1}{(1-x)(1-y) + (1-x)(1-z) + (1-y)(1-z)}$$

were positive, a result proven and generalized by Szegő [63] using properties of Bessel functions (which has been further generalized by several other authors, including, most recently, Scott and Sokal [60]). Askey and Gasper [7] give a nice discussion on the uses of such results, and discuss the history of Szegő's result. Now that the field of analytic combinatorics in several variables is sufficiently developed, it is possible to generalize asymptotic arguments from the univariate case to the multivariate setting. Baryshnikov et al. [10] use analytic combinatorics to prove eventual positivity for multivariate rational functions whose denominators are linear combinations of elementary symmetric polynomials. The diagonals of many elements in this family have special arithmetic significance as they satisfy D-finite equations with so-called modular parameterization [62]. Of utmost interest [62, Question 1.1]

[13] Available at https://github.com/amsqi/kronecker.

is a determination of necessary conditions such that positivity of the diagonal of F implies positivity of all coefficients of F.

Example 3.28 (Eventual Positivity of Multivariate Rational Functions)

Baryshnikov et al. [10] prove a conjecture of Straub and Zudilin [62] by showing that the diagonal coefficients of

$$F(x, y, z) = \frac{1}{1 - x - y - z + a(xy + xz + yz) + bxyz}$$

are eventually positive when

$$b < \begin{cases} -9a & a \le -3 \\ 2 - 3a + 2(1 - a)^{3/2} & -3 \le a \le 1 \\ -a^3 & a \ge 1 \end{cases}$$

and that the coefficients contain an infinite number of positive and negative terms when the inequality is reversed. See Problem 5.6 in Chapter 5.

3.4.6 The Ising Model and Algebraic Diagonals

The Ising model is an important model of ferromagnetism in statistical physics, introduced by Lenz [46] in the early twentieth century. Roughly speaking, the model considers a magnetic field generated by the spins of particles arranged on a lattice with short range interactions between close electrons; the goal when studying a model is to describe how these short range interactions, possibly in the presence of outside forces, dictate large scale system information for different configurations of spins after certain parameters are fixed. As discussed in Bostan et al. [14, Sec. 30], many properties of the model can be represented by analytic functions encoded as diagonals of explicit d-variate algebraic functions. Such algebraic diagonals can be written as rational diagonals in a larger number of variables, although this process often destroys nice properties of the original algebraic diagonals. There are some results about the theory of analytic combinatorics in several variables in the presence of algebraic singularities [36], though this theory is still in its infancy.

Example 3.29 (Statistics on the Ising Model)

Bostan et al. [13, Appendix C] consider a family of integrals $\Phi_D^{(n)}(w)$ related to the 'n-particle contribution to the diagonal magnetic susceptibility of the Ising model' and give the explicit example

$$\Phi_D^{(3)}(w) = \Delta\left(\frac{1-2w+\sqrt{(1-2w)^2-4w^2t^2}}{2\sqrt{1-t^2}\sqrt{(1-2w)^2-4w^2t^2}} - \frac{1}{2}\right).$$

See that paper for additional examples arising in physics.

3.4.7 Other Sources of Examples

A survey paper by Pemantle and Wilson [54], together with the textbook [55], gives a wide variety of diagonal expressions arising in applications, including examples in the study of trees and graphs, quantum computation, orthogonal polynomials, Gaussian weak and central limit laws, queuing theory, integer solutions to linear equations, tilings of the Aztec Diamond, sequences defined by Riordan arrays, bioinformatics, the study of polytope dilations, convex polyominoes, symmetric Eulerian numbers, and strings with forbidden patterns.

An updated website listing papers which rely on the theory of analytic combinatorics in several variables is maintained at the website

http://ACSVproject.com

Examples from some of these application areas will be worked through or given as problems in Chapters 5 and 9.

Problems

3.1 If (k_n) is a sequence of non-negative integers in which every element of $\mathbb{N}$ appears exactly once, and (f_n) is any sequence over $\mathbb{C}$, then (f_{k_n}) is called a *rearrangement* of (f_n). Prove that if $\sum_{n\geq 0}|f_n|$ converges then $\sum_{n\geq 0}|f_{k_n}|$ converges. Exhibit a sequence (f_n) such that $\sum_{n\geq 0} f_n$ converges but some rearrangement $\sum_{n\geq 0} f_{k_n}$ does not. *Hint:* Use the triangle inequality and the fact that the trailing sums $\sum_{n\geq N}|f_n|$ go to zero as $N \to \infty$.

3.2 Let $H(\mathbf{z})$ be a multivariate polynomial. Show there exist constants $c_1, \dots, c_{d-1}$ such that making the substitution $z_j = w_j - c_j z_d$ for $1 \leq j \leq d-1$ results in a polynomial $H(w_1, \dots, w_{d-1}, z_d)$ that is non-constant when $w_1 = \cdots = w_{d-1} = 0$.

3.3 Prove Proposition 3.3 using the multivariate Cauchy integral formula.

3.4 Show that if $F(\mathbf{z}) \in \mathbb{Q}(\mathbf{z})$ is analytic at the origin then its diagonal $(\Delta F)(z)$ is globally bounded. Prove that any globally bounded function whose coefficients grow at most exponentially is a G-function, so that a rational diagonal is always a G-function.

3.5 (Eisenstein's theorem) Let $F(z)$ be an algebraic series with rational coefficients. Use the fact that $F(z)$ is the diagonal of a bivariate rational function to prove the existence of non-zero integers $a, b \in \mathbb{Z}$ such that $aF(bz)$ has integer coefficients.

3.6 Prove that the function e^x is transcendental without using asymptotic characterizations of algebraicity.

3.7 Using Stirling's formula, or the asymptotic results developed in Chapter 5, prove that the diagonal of the trivariate rational function $F(x, y, z) = 1/(1 - x - y - z)$ is transcendental.

3.8 Following the argument discussed at the end of Section 3.2.1, prove that the **r**-diagonal of a rational function $F(\mathbf{z}) \in \mathbb{Q}(\mathbf{z})$ which is analytic at the origin equals the main diagonal of some function in $\mathbb{Q}(\mathbf{z})$. Prove analogous statements for **r**-diagonals of algebraic and D-finite functions over $\mathbb{Q}$.

3.9 Let $F(z) = z\sqrt{1-z}$ so that $P(z, F(z)) = 0$ where $P(z, y) = y^2 - z^2(1-z)$ satisfies $P(0, 0) = P_y(0, 0) = 0$. Recall the discriminant from Definition 2.14 of Chapter 2, and let $r \in \mathbb{N}$ be the natural number such that the discriminant $D(z)$ of $P(z, y)$ with respect to y can be written $D(z) = z^r d(z)$ where $d(0) \neq 0$. Writing $F(z) = a(z) + z^r b(z)$ where $a(z)$ is a polynomial of degree at most $r \in \mathbb{N}$ and $b(0) = 0$, prove the existence of a polynomial $Q(z, y)$ such that $Q(z, y)$ and $b(z)$ satisfy the conditions of Proposition 3.8. Use this to express $F(z)$ as the diagonal of a bivariate rational function.

3.10 Prove Proposition 3.9 by writing $F(\mathbf{z}) = \sum_{n\geq 1} C_n(\hat{\mathbf{z}})z_d^n$ for analytic $C_n(\hat{\mathbf{z}})$ and explicitly finding the terms in $y^2 P_y(\hat{\mathbf{z}}, yz_d, y)/P(\hat{\mathbf{z}}, yz_d, y)$ where z_d and y have the same exponent.

3.11 Using the recursive nature of a rooted plane tree, prove that the bivariate generating function $T(u, z) = \sum_{n,k\geq 0} t_{k,n} u^k z^n$ counting the number $t_{k,n}$ of binary trees on n nodes with k leaves satisfies the algebraic equation

$$T(u, z) = zu + \frac{zT(u, z)}{1 - T(u, z)}.$$

3.12 (Lucas's theorem) The 'Freshman's dream' identity states that for any polynomial $f(x)$ with integer coefficients, $f(x)^p = f(x^p)$ modulo p. Using this identity and the binomial theorem, show that

$$\binom{n}{m} = \binom{a_0}{b_0} \cdots \binom{a_r}{b_r} \mod p,$$

where $n = a_0 + a_1 p + \cdots + a_r p^r$ and $m = b_0 + b_1 p + \cdots + b_r p^r$ are the base p expansions of n and m (padded with zeroes if necessary to have the same length).

3.13 Using Lucas's Theorem, for any prime p find an algebraic equation satisfied by

$$F(z) = \Delta\left(\frac{1}{(1-w-x)(1-y-z)}\right) = \sum_{n\geq 0} \binom{2n}{n}^2 z^n$$

over the finite field of order p. Recall that $F(z)$ is transcendental over $\mathbb{C}[[z]]$.

3.14 Prove that if $\mathbf{v} \in \mathbb{R}^d$ is a limit direction of amoeba(H) for $H \in \mathbb{Q}[\mathbf{z}, \overline{\mathbf{z}}]$ then there exist distinct vectors $\mathbf{n}$ and $\mathbf{m}$ in $\mathcal{N}(H)$ such that $\mathbf{v} \cdot \mathbf{n} = \mathbf{v} \cdot \mathbf{m} \geq \mathbf{v} \cdot \mathbf{k}$ for all $\mathbf{k} \in \mathcal{N}(H)$. *Hint:* When $t \geq 0$ and $\mathbf{x} + t\mathbf{v}$ lies in amoeba(H) there exists $\boldsymbol{\omega} \in \mathbb{C}_*^d$, with $|\omega_j|$ independent of t, such that $H(\omega_1 e^{v_1 t}, \ldots, \omega_d e^{v_d t}) = 0$. What happens as $t \to \infty$?

3.15 Let $R(x, y) = 1 - x - y - xy^3$. Prove that the diagonal sequences of the two convergent Laurent expansions of $1/R(x, y)$ not discussed in this chapter are identically zero.

References

1. Y. Abdelaziz, C. Koutschan, and J.-M. Maillard. On Christol's conjecture. *J. Phys. A*, 53(20):205201, 16, 2020.
2. Boris Adamczewski and Jason P. Bell. Diagonalization and rationalization of algebraic Laurent series. *Ann. Sci. Éc. Norm. Supér. (4)*, 46(6):963–1004, 2013.
3. I. A. Aĭzenberg and A. P. Yuzhakov. *Integral representations and residues in multidimensional complex analysis*, volume 58 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1983. Translated from the Russian by H. H. McFaden, Translation edited by Lev J. Leifman.
4. Ainhoa Aparicio Monforte and Manuel Kauers. Formal Laurent series in several variables. *Expo. Math.*, 31(4):350–367, 2013.
5. Roger Apéry. Irrationalité de $\zeta(2)$ et $\zeta(3)$. In *Journées Arith. de Luminy. Colloque International du Centre National de la Recherche Scientifique (CNRS), Centre Universitaire de Luminy, Jun 20-24. Astérisque*, volume 61, pages 11–13, 1979.
6. Roger Apéry. Sur certaines séries entières arithmétiques. In *Study group on ultrametric analysis, 9th year: 1981/82, No. 1*, pages Exp. No. 16, 2. Inst. Henri Poincaré, Paris, 1983.
7. Richard Askey and George Gasper. Certain rational functions whose power series have positive coefficients. *Amer. Math. Monthly*, 79:327–341, 1972.
8. Martín Avendaño, Roman Kogan, Mounir Nisse, and J. Maurice Rojas. Metric estimates and membership complexity for archimedean amoebae and tropical hypersurfaces. *Journal of Complexity*, 46:45–65, 2018.
9. V. Baldoni and M. Vergne. Computation of dilated Kronecker coefficients. *J. Symbolic Comput.*, 84:113–146, 2018. With an appendix by M. Walter.
10. Yuliy Baryshnikov, Stephen Melczer, Robin Pemantle, and Armin Straub. Diagonal asymptotics for symmetric rational functions via ACSV. In *29th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms*, volume 110 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages Art. No. 12, 15. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2018.
11. Victor Batyrev and Maximilian Kreuzer. Constructing new Calabi-Yau 3-folds and their mirrors via conifold transitions. *Adv. Theor. Math. Phys.*, 14(3):879–898, 2010.
12. George M. Bergman. The logarithmic limit-set of an algebraic variety. *Trans. Amer. Math. Soc.*, 157:459–469, 1971.

13. A. Bostan, S. Boukraa, G. Christol, S. Hassani, and J.-M. Maillard. Ising n-fold integrals as diagonals of rational functions and integrality of series expansions: integrality versus modularity. *arXiv preprint, arXiv:1211.6031*, 2012.
14. A. Bostan, S. Boukraa, G. Christol, S. Hassani, and J.-M. Maillard. Ising n-fold integrals as diagonals of rational functions and integrality of series expansions. *J. Phys. A*, 46(18):185202, 44, 2013.
15. Alin Bostan, Louis Dumont, and Bruno Salvy. Algebraic diagonals and walks: algorithms, bounds, complexity. *J. Symbolic Comput.*, 83:68–92, 2017.
16. Alin Bostan, Pierre Lairez, and Bruno Salvy. Multiple binomial sums. *J. Symbolic Comput.*, 80(part 3):351–386, 2017.
17. Alin Bostan and Sergey Yurkevich. On a class of hypergeometric diagonals. *arXiv preprint, arXiv:2008.12809*, 2020.
18. Umberto Bottazzini and Jeremy Gray. *Hidden harmony—geometric fantasies*. Sources and Studies in the History of Mathematics and Physical Sciences. Springer, New York, 2013.
19. M. Bousquet-Mélou and A. Rechnitzer. The site-perimeter of bargraphs. *Adv. in Appl. Math.*, 31(1):86–112, 2003.
20. Peter Bürgisser and Christian Ikenmeyer. The complexity of computing Kronecker coefficients. In *20th Annual International Conference on Formal Power Series and Algebraic Combinatorics (FPSAC 2008)*, Discrete Math. Theor. Comput. Sci. Proc., AJ, pages 357–368. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2008.
21. Peter Bürgisser, J. M. Landsberg, Laurent Manivel, and Jerzy Weyman. An overview of mathematical issues arising in the geometric complexity theory approach to VP ≠ VNP. *SIAM J. Comput.*, 40(4):1179–1209, 2011.
22. Gilles Christol. Diagonales de fractions rationnelles et equations différentielles. In *Study group on ultrametric analysis, 10th year: 1982/83, No. 2*, pages Exp. No. 18, 10. Inst. Henri Poincaré, Paris, 1984.
23. Gilles Christol. Fonctions hypergéométriques bornées. *Groupe de travail d'analyse ultramétrique*, 14, 1986-1987.
24. Gilles Christol. Globally bounded solutions of differential equations. In *Analytic number theory (Tokyo, 1988)*, volume 1434 of *Lecture Notes in Math.*, pages 45–64. Springer, Berlin, 1990.
25. David A. Cox and Sheldon Katz. *Mirror symmetry and algebraic geometry*, volume 68 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1999.
26. P. Deligne. Intégration sur un cycle évanescent. *Invent. Math.*, 76(1):129–143, 1984.
27. J. Denef and L. Lipshitz. Algebraic power series and diagonals. *J. Number Theory*, 26(1):46–67, 1987.
28. Philippe Flajolet and Robert Sedgewick. *Analytic combinatorics*. Cambridge University Press, Cambridge, 2009.
29. Mikael Forsberg, Mikael Passare, and August Tsikh. Laurent determinants and arrangements of hyperplane amoebas. *Adv. Math.*, 151(1):45–70, 2000.
30. Jens Forsgå rd, Laura Felicia Matusevich, Nathan Mehlhop, and Timo de Wolff. Lopsided approximation of amoebas. *Math. Comp.*, 88(315):485–500, 2019.
31. William Fulton. *Young tableaux*, volume 35 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 1997. With applications to representation theory and geometry.
32. William Fulton and Joe Harris. *Representation theory*, volume 129 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1991. A first course, Readings in Mathematics.
33. Harry Furstenberg. Algebraic functions over finite fields. *J. Algebra*, 7:271–277, 1967.
34. Scott Garrabrant and Igor Pak. Counting with irrational tiles. *arXiv preprint, arXiv:1407.8222*, 2014.
35. I. M. Gelfand, M. M. Kapranov, and A. V. Zelevinsky. *Discriminants, resultants and multidimensional determinants*. Modern Birkhäuser Classics. Birkhäuser Boston, Inc., Boston, MA, 2008. Reprint of the 1994 edition.

36. Torin Greenwood. Asymptotics of bivariate analytic functions with algebraic singularities. *J. Combin. Theory Ser. A*, 153:1–30, 2018.
37. Fritz Hartogs. Zur Theorie der analytischen Funktionen mehrerer unabhängiger Veränderlichen, insbesondere über die Darstellung derselben durch Reihen, welche nach Potenzen einer Veränderlichen fortschreiten. *Math. Ann.*, 62(1):1–88, 1906.
38. M. L. J. Hautus and D. A. Klarner. The diagonal of a double power series. *Duke Math. J.*, 38:229–235, 1971.
39. Lars Hörmander. *An introduction to complex analysis in several variables*, volume 7 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam, third edition, 1990.
40. Christian Ikenmeyer, Ketan D. Mulmuley, and Michael Walter. On vanishing of Kronecker coefficients. *Comput. Complexity*, 26(4):949–992, 2017.
41. Christian Ikenmeyer and Greta Panova. Rectangular Kronecker coefficients and plethysms in geometric complexity theory. *Adv. Math.*, 319:40–66, 2017.
42. Maxim Kontsevich and Don Zagier. Periods. In *Mathematics unlimited—2001 and beyond*, pages 771–808. Springer, Berlin, 2001.
43. Christoph Koutschan. A fast approach to creative telescoping. *Math. Comput. Sci.*, 4(2-3):259–266, 2010.
44. Steven G. Krantz. *Function theory of several complex variables*. The Wadsworth & Brooks/Cole Mathematics Series. Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA, second edition, 1992.
45. Pierre Lairez. Computing periods of rational integrals. *Math. Comp.*, 85(300):1719–1752, 2016.
46. W Lenz. Beiträge zum verstandnis der magnetischen eigenschaften in festen korpern. *Physikalische Zeitschrift*, 21, 1920.
47. L. Lipshitz. The diagonal of a D-finite power series is D-finite. *J. Algebra*, 113(2):373–378, 1988.
48. Leonard Lipshitz and Alfred J. van der Poorten. Rational functions, diagonals, automata and arithmetic. In *Number theory (Banff, AB, 1988)*, pages 339–358. de Gruyter, Berlin, 1990.
49. Eckhard Meinrenken and Reyer Sjamaar. Singular reduction and quantization. *Topology*, 38(4):699–762, 1999.
50. S. Melczer, G. Panova, and R. Pemantle. Counting partitions inside a rectangle. *Accepted to SIAM Journal on Discrete Mathematics*, 2019.
51. G. Mikhalkin. Real algebraic curves, the moment map and amoebas. *Ann. of Math. (2)*, 151(1):309–326, 2000.
52. M. Mishna, M Rosas, and S. Sundaram. An elementary approach to the quasipolynomiality of the kronecker coefficients. *arXiv preprint, arXiv:1811.10015*, 2018.
53. David R. Morrison. Mirror symmetry and rational curves on quintic threefolds: a guide for mathematicians. *J. Amer. Math. Soc.*, 6(1):223–247, 1993.
54. Robin Pemantle and Mark C. Wilson. Twenty combinatorial examples of asymptotics derived from multivariate generating functions. *SIAM Rev.*, 50(2):199–272, 2008.
55. Robin Pemantle and Mark C. Wilson. *Analytic combinatorics in several variables*, volume 140 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2013.
56. David A. Plaisted. New NP-hard and NP-complete polynomial and integer divisibility problems. *Theoret. Comput. Sci.*, 31(1-2):125–138, 1984.
57. G. Pólya. Sur les séries entières, dont la somme est une fonction algébrique. *L'Enseignement Mathématique*, 22:38–47, 1921.
58. Kevin Purbhoo. A Nullstellensatz for amoebas. *Duke Math. J.*, 141(3):407–445, 2008.
59. K. V. Safonov. On power series of algebraic and rational functions in $\mathbf{C}^n$. *J. Math. Anal. Appl.*, 243(2):261–277, 2000.
60. Alexander D. Scott and Alan D. Sokal. Complete monotonicity for inverse powers of some combinatorially defined polynomials. *Acta Math.*, 213(2):323–392, 2014.
61. Richard P. Stanley. *Enumerative combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999.

62. Armin Straub and Wadim Zudilin. Positivity of rational functions and their diagonals. *J. Approx. Theory*, 195:57–69, 2015.
63. G. Szegö. Über gewisse Potenzreihen mit lauter positiven Koeffizienten. *Math. Z.*, 37(1):674–688, 1933.
64. Thorsten Theobald. Computing amoebas. *Experiment. Math.*, 11(4):513–526 (2003), 2002.
65. Sascha Timme. Fast computation of amoebas, coamoebas and imaginary projections in low dimensions. Master's thesis, Technische Universität Berlin, 2018.
66. Alfred van der Poorten. A proof that Euler missed. . .Apéry's proof of the irrationality of $\zeta(3)$. *Math. Intelligencer*, 1(4):195–203, 1978/79. An informal report.
67. Guoce Xin. *The ring of Malcev-Neumann series and the residue theorem*. PhD thesis, 2004. Thesis (Ph.D.)–Brandeis University.

Chapter 4
Lattice Path Enumeration, the Kernel Method, and Diagonals

And this is the business of the art or doctrine of combinations. Nor is this art or doctrine to be considered merely as a branch of the mathematical sciences, for it has a relation to almost every species of useful knowledge that the mind of man can be employed upon. — Jakob Bernoulli

Primeval man stumbled along with peering eyes, and slow, uncertain footsteps. Now we walk briskly towards our unknown goal. — Arthur Conan Doyle (as J. Stark Munro)

The main objects on which we illustrate the methods of analytic combinatorics in several variables are lattice paths; by studying successively more complicated lattice path models we will obtain a variety of problems whose analyses require techniques touching all areas of ACSV. Lattice path enumeration has a long and colourful history, dating back centuries. Early accounts of what could now be considered lattice path problems arose as far back as the seventeenth century probabilistic work of Pascal and Fermat, including examples analogous to the so-called ballot problem in the work of de Moivre [78] in 1711, although those authors did not pose their questions in terms of lattice paths. An 1878 work of Whitworth [87] uses explicit lattice path terminology (for instance "paces" from an origin) to consider "Arrangements of m things of one sort and n things of another sort under certain conditions of priority," and answered questions posed by the *Educational Times* including the probability of drinking k glasses of wine and k glasses of water in a random order while never drinking more wine than water. Lattice walks were also considered by many to be a recreational topic, as exemplified by an article of Grossman [53] from 1950 entitled "Fun with lattice points" and published in the journal *Scripta Mathematica* aimed at the layperson. An entertaining history of lattice path enumeration can be found in the survey of Humphreys [54].

In modern times, lattice paths appear in many diverse areas of mathematics and the sciences. They are prevalent in probability theory, since sums of discrete random variables are modeled by random walks, and, for similar reasons, are deeply related to statistical methods such as the Kolmogorov-Smirnov goodness-of-fit test [79]. Lattice path models are able to model physical phenomena and find use in statistical mechanics, for instance in the study of polymers in a solution [85]. Additional applications include formal language theory [20], queuing theory [37, 8], the analysis of data structures [31, 43, 44], Liouville quantum gravity [6], the combinatorics of continued fractions [42], the study of other combinatorial structures such as plane partitions, trees, and sequences of Young tableaux, and even mathematical art [55, 56]. Detailed treatments of lattice paths and their applications include the texts of Mohanty [77] and Narayana [79], and the survey of Krattenthaler [64].

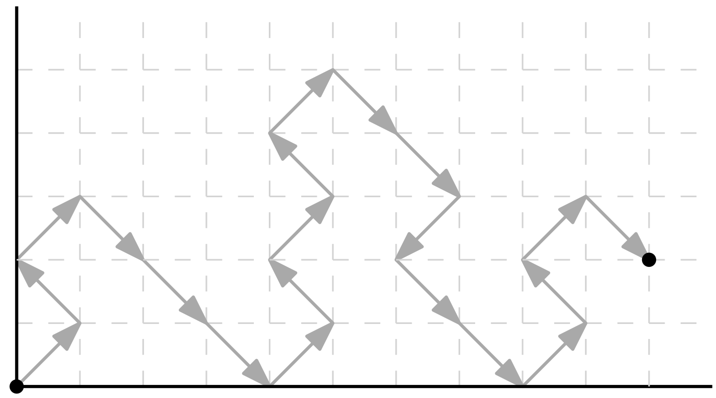


Fig. 4.1 A two-dimensional lattice walk in the quarter-plane of length 20, using the step set $\mathcal{S} = \{(-1,-1), (-1,1), (1,-1), (1,1)\}$.

Formal Setup

Fix a dimension $d \in \mathbb{N}_{>0}$.

Definition 4.1 (lattice path models) Given a finite set of steps $\mathcal{S} \subset \mathbb{Z}^d$, region $\mathcal{R} \subset \mathbb{R}^d$, starting point $\mathbf{p} \in \mathcal{R}$, and terminal set $\mathcal{T} \subset \mathcal{R}$, the *lattice path model taking steps in* $\mathcal{S}$*, starting at* $\mathbf{p}$*, restricted to* $\mathcal{R}$*, and ending in* $\mathcal{T}$ is the combinatorial class of all finite tuples $(\mathbf{s}_1, \dots, \mathbf{s}_r) \in \mathcal{S}^r$ such that $\mathbf{p}+\mathbf{s}_1+\cdots+\mathbf{s}_r \in \mathcal{T}$ and $\mathbf{p}+\mathbf{s}_1+\cdots+\mathbf{s}_k \in \mathcal{R}$ for all $1 \leq k \leq r$. We call the valid sequences in $\mathcal{S}^r$ the *walks* or *paths* of the model: they can be visualized in $\mathbb{Z}^d$ by starting at $\mathbf{p}$ and concatenating the vectors $\mathbf{s}_1, \dots, \mathbf{s}_r$ in order, as in Figure 4.1. The *steps* of a walk are the elements of the tuple defining the walk, and the *size* or *length* of a walk is the number of steps it is composed of. Lattice walks beginning and ending at the same point are often called *excursions*.

We focus mainly on the case when the restricting region $\mathcal{R}$ is an orthant $\mathbb{N}^d$ or a product $\mathbb{Z}^s \times \mathbb{N}^{d-s}$. Many models restricted to other regions can be viewed in this setting, as can other types of lattice paths (for instance, collections of pairwise non-intersecting paths can often be viewed as walks in $\mathbb{N}^d$). By convention we say that all models have a single walk of length zero, which ends at the starting point $\mathbf{p}$.

We also consider weighted walks.

Definition 4.2 (weighted lattice path models) In a weighted lattice path model each step $\mathbf{i} \in \mathcal{S}$ is given a positive real weight $w_{\mathbf{i}} > 0$. In this case, the *weight* of a path $(\mathbf{s}_1, \dots, \mathbf{s}_r) \in \mathcal{S}^r$ is the product of the weights $w_{\mathbf{s}_1} \cdots w_{\mathbf{s}_r}$, and *counting* the number of paths of length n refers to summing the weights of all valid paths of length n.

In probabilistic contexts, one often considers weighted step sets whose weights add to one. Combinatorially, when all weights are positive integers one can imagine having differently coloured copies of steps. Weighted counting is the same as regular unweighted enumeration when every step is given weight one.

The key idea for lattice path enumeration is to work recursively, treating a walk of length n as a walk of length $n-1$ followed by a single step. Unfortunately, knowing the number of walks of length $n-1$ is usually not sufficient to determine how many walks there are of length n: when most of the walks of length $n-1$ end near the boundary of the restricting region $\mathcal{R}$ many of them would leave $\mathcal{R}$ with a single step, but the situation is very different when most walks end far from the boundary. For this reason, we introduce a multivariate generating function $W(\mathbf{z}, t)$ whose coefficients count the number of walks in a model by their length *and* endpoint in $\mathbb{Z}^d$, and use an approach known as the *kernel method* to derive equations satisfied by various generating functions related to the model. The kernel method naturally results in rational diagonal expressions for many lattice path generating functions, to which we apply the methods of ACSV. Our presentation follows a combinatorial approach to the kernel method developed mainly by Bousquet-Mélou and collaborators in a sequence of papers [21, 24, 22, 12] over roughly the last decade. Section 4.2.1 gives some historical remarks on the kernel method.

4.1 Walks in Cones and The Kernel Method

Fix a step set $\mathcal{S} \subset \mathbb{Z}^d$ and, as is the standard in lattice path enumeration, for any variable x write $\overline{x} = 1/x$.

Definition 4.3 (characteristic polynomials) The *characteristic polynomial* of an unweighted lattice path model defined by $\mathcal{S}$ is the Laurent polynomial $S(\mathbf{z}) = \sum_{\mathbf{i}\in\mathcal{S}} \mathbf{z}^{\mathbf{i}}$ whose monomials encode the steps in $\mathcal{S}$. The characteristic polynomial of a weighted lattice path model where each $\mathbf{i} \in \mathcal{S}$ has weight $w_{\mathbf{i}} > 0$ is $S(\mathbf{z}) = \sum_{\mathbf{i}\in\mathcal{S}} w_{\mathbf{i}}\mathbf{z}^{\mathbf{i}}$.

We build up the kernel method by considering successively more complicated restricting regions $\mathcal{R}$. Unless otherwise stated, the walks we consider begin at the origin and can end anywhere in $\mathcal{R}$.

4.1.1 Unrestricted Walks

We begin with unweighted unrestricted walks in $\mathbb{Z}^d$ (so $\mathcal{R} = \mathbb{Z}^d$). Because there are no restrictions on where a walk can go, there are $|\mathcal{S}|^n$ walks on length n on the steps $\mathcal{S}$, and the generating function $C(t)$ counting the number of walks by length is

$$C(t) = \sum_{n\geq 0} |\mathcal{S}|^n t^n = \frac{1}{1 - t|\mathcal{S}|}.$$

In order to illustrate the kernel method, and obtain a more refined analysis of these models, we introduce the multivariate generating function

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

where $f_{\mathbf{i},n}$ denotes the number of walks in the model which have n steps and end at $\mathbf{i} \in \mathbb{Z}^d$. Since $\mathcal{S}$ is finite, $W(\mathbf{z},t)$ lies in the ring $\mathbb{Z}[\mathbf{z},\overline{\mathbf{z}}][[t]]$, meaning W is a power series in t whose coefficients are Laurent polynomials in the $\mathbf{z}$ variables. All series operators such as diagonals will be taken with respect to the expansion of W in $\mathbb{Z}[\mathbf{z},\overline{\mathbf{z}}][[t]]$, following the setup of Section 3.3.2 in Chapter 3. Because there are $|\mathcal{S}|^n$ walks of length n, this series expansion of W converges absolutely whenever $|z_1| = \cdots = |z_d| = 1$ and $|t| < 1/|\mathcal{S}|$.

In order to exploit the recursive nature of a lattice path, for any $n \in \mathbb{N}$ we let

$$W_n(\mathbf{z}) = [t^n]W(\mathbf{z},t) = \sum_{\mathbf{i}\in\mathbb{Z}^d} f_{\mathbf{i},n}\mathbf{z}^{\mathbf{i}}$$

denote the generating function of walks of length n counted by their endpoint. If our walks begin at the point $\mathbf{p} \in \mathbb{Z}^d$ then $W_0(\mathbf{z}) = \mathbf{z}^{\mathbf{p}}$, as by convention there is a single walk of length zero ending at $\mathbf{p}$. Since a walk of length $n+1$ is a walk of length n followed by a single step, it follows that

$$W_{n+1}(\mathbf{z}) = S(\mathbf{z})W_n(\mathbf{z}) \tag{4.1}$$

for all $n \geq 0$, where $S(\mathbf{z}) = \sum_{\mathbf{i}\in\mathcal{S}}\mathbf{z}^{\mathbf{i}}$ is the characteristic polynomial from Definition 4.3. Multiplying (4.1) by t^{n+1} and summing over all $n \in \mathbb{N}$ gives the equation

$$W(\mathbf{z},t) - W_0(\mathbf{z}) = tS(\mathbf{z})W(\mathbf{z},t),$$

so that

$$(1 - tS(\mathbf{z}))\,W(\mathbf{z},t) = \mathbf{z}^{\mathbf{p}}. \tag{4.2}$$

Definition 4.4 (unrestricted kernel equation) Equation (4.2) is known as the *kernel equation* for unrestricted walks, with *kernel* $K(\mathbf{z},t) = 1 - tS(\mathbf{z})$.

In the following sections we derive similar equations for lattice path models restricted to other regions, but the right-hand side of the resulting equations will rely on evaluations and coefficient extractions of (a priori unknown) multivariate generating functions. Here we can simply divide both sides of the kernel equation to obtain

$$W(\mathbf{z},t) = \frac{\mathbf{z}^{\mathbf{p}}}{1 - tS(\mathbf{z})}. \tag{4.3}$$

Specializing $\mathbf{z} = \mathbf{1}$ sums over the possible end locations for a walk of length n, allowing us to recover

$$C(t) = W(\mathbf{1},t) = \frac{1}{1 - tS(\mathbf{1})} = \frac{1}{1 - t|\mathcal{S}|},$$

but further refinements are possible. For instance, the series $E(t) = [\mathbf{z}^{\mathbf{0}}]W(\mathbf{z}, t)$ counts the number of walks which begin and end at the origin. The generating function for such walks is thus given by the main diagonal

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

which is D-finite by Theorem 3.2 of Chapter 3.

Note that (4.3) still holds when each step $\mathbf{i} \in \mathcal{S}$ is given weight $w_{\mathbf{i}} > 0$ and $W(\mathbf{z}, t)$ counts the number of weighted walks marked by endpoint and length, as long as the characteristic polynomial is replaced by its weighted version $S(\mathbf{z}) = \sum_{\mathbf{i}\in\mathcal{S}} w_{\mathbf{i}}\mathbf{z}^{\mathbf{i}}$.

4.1.2 A Deeper Kernel Analysis: One-Dimensional Excursions

When $d = 1$, the generating function counting walks ending at the origin is

$$E(t) = \Delta\left(\frac{x^p}{1 - txS(x)}\right)$$

which, as the diagonal of a bivariate function, is algebraic. Assume now that $p = 0$ and let $-m$ and M with $m, M > 0$ be the smallest and largest elements of $\mathcal{S}$, so that $\mathcal{S}$ contains both a step with negative value and a step with positive value (otherwise there can be no walks ending at the origin). For $|t| < 1/S(1)$ the generating function is given by the integral expression

$$E(t) = \int_{|x|=1} \frac{1}{1 - tS(x)}\frac{dx}{x}, \tag{4.4}$$

which follows either from the diagonal expression or straight from the Cauchy integral formula. The poles of this integrand are the roots of $K(x, t) = 1 - tS(x)$, which we now study. Since $x^m K(x, t) = x^m - tx^m S(x)$ is a polynomial of degree $m + M$ in x, Proposition 2.10 in Chapter 2 implies the kernel $K(x, t)$, considered as a polynomial in x, has m roots $r_1(t), \ldots, r_m(t)$ which are fractional power series in t and M roots $R_1(t), \ldots, R_M(t)$ which approach infinity as t approaches zero.

Definition 4.5 (small and large kernel roots) The roots $r_1(t), \ldots, r_m(t)$ are the *small roots* of $K(x, t)$, while $R_1(t), \ldots, R_M(t)$ are the *large roots* of $K(x, t)$.

Setting $t = 0$ in the equation

$$r_j(t)^m = t\, r_j(t)^m\, S(r_j(t))$$

implies $r_j(0) = 0$; i.e., a small root $r_j(t)$ has no constant term. Although not necessary here, we note that since $r_j(t)$ has no constant term one may substitute $x = r_j(t)$ in any element of $\mathbb{R}[[x, t]]$ to obtain a well-defined Puiseux expansion in t. This will be key to enumerating walks in a half-space.

Proposition 4.1 *Let $\mathcal{S} \subset \mathbb{Z}$ be a step set with smallest element $-m$ for $m > 0$. Then the generating function $E(t)$ enumerating walks on the steps $\mathcal{S}$ which begin and end at the origin has the representation*

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

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

Remark 4.1 The statement of Proposition 4.1 does not require that $\mathcal{S}$ has a positive step. If $\mathcal{S}$ contains only negative steps then there are no excursions of positive length and the expression for $E(t)$ correctly simplifies to 1.

Proof For t non-zero but sufficiently close to the origin only the small roots $r_j(t)$ lie inside the curve $|x| = 1$, so (4.4) and the residue theorem imply that $E(t)$ is a sum of residues at these roots (note $x = 0$ is not a singularity of the integrand as $m > 0$). If $r(t)$ satisfies $1 - tS(r(t)) = 0$ then taking the derivative with respect to t shows $-S(r(t)) - tr'(t)S'(r(t)) = 0$, so that $S'(r(t)) \neq 0$ and

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

Thus, each root of the kernel is a simple zero and

$$E(t) = \sum_{j=1}^{m} \operatorname*{Res}_{x=r_j(t)} \left(\frac{1}{x(1-tS(x))} \right) = t^{-1} \sum_{j=1}^{m} \left(\frac{1}{-r_j(t)\, S'(r_j(t))} \right) = t \sum_{j=1}^{m} \frac{r_j'(t)}{r_j(t)},$$

as desired. □

Example 4.1 (One Dimensional Excursions)

Suppose $\mathcal{S} = \{-2,-1,0,1,2\} \subset \mathbb{Z}$, so that $S(x) = x^{-2}+x^{-1}+1+x+x^2$. As expected, the equation $1 - tS(x) = 0$ has 2 solutions

$$r_1(t) = t^{1/2} + \frac{t}{2} + \frac{5t^{3/2}}{8} + \cdots \qquad r_2(t) = -t^{1/2} + \frac{t}{2} - \frac{5t^{3/2}}{8} + \cdots$$

which approach zero as t approaches zero, and two solutions

$$R_1(t) = t^{-1/2} - \frac{1}{2} - \frac{3t^{1/2}}{8} + \cdots \qquad R_2(t) = -t^{-1/2} - \frac{1}{2} + \frac{3t^{1/2}}{8} + \cdots$$

which approach infinity as t approaches zero. Thus, the generating function of walks beginning and ending at the origin satisfies

$$E(t) = \frac{t\, r_1'(t)}{r_1(t)} + \frac{t\, r_2'(t)}{r_2(t)} = 1 + t + 5t^2 + 19t^3 + 85t^4 + \cdots.$$

Note that although $r_1(t)$ and $r_2(t)$ are not analytic at the origin the generating function $E(t)$ is analytic there. Knowing the minimal polynomial of r_1 and r_2, the minimal polynomial $m(t, y)$ of $E(t)$ can be calculated using resultants, giving

$$m(t, y) = (4 + 5t)(5t - 1)^2(t - 1)^2 y^4 + 2(t - 1)(5t - 2)(5t - 1)y^2 + t.$$

Problem 4.1 asks you to generalize this argument to walks with arbitrary starting point $p \in \mathbb{Z}$. See also Banderier and Flajolet [5] for similar results in this setting.

4.1.3 Walks in a Half-Space

Consider now a weighted step set $\mathcal{S}' \subset \mathbb{Z}^d$ restricted to the half-space $\mathcal{R}' = \mathbb{Z}^{d-1} \times \mathbb{N}$. In order to enumerate walks ending anywhere, or walks ending on the boundary hyperplane $z_d = 0$, it is enough to project each step onto its dth coordinate to obtain a new weighted step set $\mathcal{S} \subset \mathbb{Z}$ and count walks restricted to $\mathcal{R} = \mathbb{N}$. To simplify our presentation we thus restrict ourselves to the case of one-dimensional walks in $\mathbb{N}$ beginning at the origin, losing only a small amount of generality. Now, define

$$H(x, t) = \sum_{n \geq 0} \left(\sum_{i \geq 0} h_{i,n} x^i \right) t^n \in \mathbb{R}[x][[t]]$$

where $h_{i,n}$ counts the number of weighted half-space walks of length n on the steps in $\mathcal{S}$ ending at a point with x-coordinate i. As before, let $-m$ and M be the smallest and largest elements of $\mathcal{S}$, where we assume $m, M > 0$ so that the valid walks are non-trivial and actually interact with the boundary of $\mathcal{R} = \mathbb{N}$. With $H_n(x) = [t^n]H(x, t)$ enumerating the walks of length n by endpoint, our goal is to obtain a recurrence for H_n; this recurrence will not be the same as the unrestricted case as we must take into account walks that try to leave $\mathbb{N}$. Fortunately, using the bivariate function $H(x, t)$ tracking the endpoint of a walk allows us to ensure that walks do not leave $\mathbb{N}$.

Define again the weighted characteristic polynomial

$$S(x) = \sum_{i \in \mathcal{S}} w_i x^i = w_{-m} x^{-m} + \cdots + w_M x^M,$$

where $w_i > 0$ is the real weight associated to the step $i \in \mathcal{S}$. To keep track of walks potentially leaving the half-space, for any integer j with $0 \leq j < m$ we let

$$S_{<-j}(x) = w_{-m} x^{-m} + \cdots + w_{-j-1} x^{-j-1}$$

be the sum of terms in $S(x)$ corresponding to steps moving in a negative direction of magnitude larger than j. As a walk of length $n + 1$ is a walk of length n followed by a single step from $\mathcal{S}$, we have the recurrence

$$H_{n+1}(x) = S(x)H_n(x) - \sum_{j=0}^{m-1} S_{<-j}(x)\ x^j[x^j]H_n(x)$$

for all $n \geq 0$, where the subtracted sum ensures no walks come from outside $\mathcal{R} = \mathbb{N}$. Multiplying by t^{n+1} and summing over all $n \geq 0$ implies

$$K(x,t)H(x,t) = 1 - t\sum_{j=0}^{m-1} S_{<-j}(x)\ x^j[x^j]H(x,t), \tag{4.5}$$

where K is again the kernel $K(x,t) = 1 - tS(x)$. Note that $H_0(x) = 1$ as we assume that walks begin at the origin.

Definition 4.6 (half-space kernel equation) Equation (4.5) is the *kernel equation* for the half-space model defined by $\mathcal{S}$. The coefficient extractions $[x^j]H(x,t)$ are called *sections* of H.

Although the Laurent polynomials $S_{<-j}(x)$ in (4.5) are explicit, the sections of H are unknown series, meaning one can no longer simply solve for the generating function $H(x,t)$. Thankfully, we can obtain explicit generating function expressions using an argument similar to our treatment of one-dimensional excursions. Let $r_1(t), \dots, r_m(t)$ denote the small roots of the kernel $K(x,t)$, discussed in Definition 4.5, which are fractional power series in t with no constant term.

Proposition 4.2 *Let $\mathcal{S} \subset \mathbb{Z}$ be a weighted step set with smallest element $-m$. Then the bivariate generating function $H(x,t)$ enumerating walks restricted to $\mathcal{R} = \mathbb{N}$ by length and endpoint has the representation*

$$H(x,t) = \frac{\prod_{j=1}^m \left(1 - \overline{x}r_j(t)\right)}{1 - tS(x)},$$

where $S(x)$ is the weighted characteristic polynomial of $\mathcal{S}$. In particular, the univariate generating function counting walks in $\mathbb{N}$ using the step set $\mathcal{S}$ is

$$C(t) = H(1,t) = \frac{\prod_{j=1}^m \left(1 - r_j(t)\right)}{1 - tS(1)},$$

and the generating function for the number of walks returning to the boundary $x = 0$ of $\mathcal{R} = \mathbb{N}$ is

$$E(t) = H(0,t) = \frac{(-1)^{m-1}}{w_{-m}t}\prod_{j=1}^m r_j(t).$$

Remark 4.2 We can set $x = 0$ in $H(x,t)$ since $H(x,t)$ contains only non-negative exponents in x. The statement of Proposition 4.2 does not require that $\mathcal{S}$ has a positive step. If $\mathcal{S}$ contains only negative steps then there are no walks of positive length and $H(x,t)$ correctly simplifies to 1 since all roots of $K = 1 - tS(x)$ are small.

Proof The right-hand side of (4.5) is a polynomial in $\overline{x}$ of degree m with constant term one. Substituting $x = r_j(t)$ into (4.5) shows that this polynomial vanishes at each of the m small roots, meaning

$$K(x,t)H(x,t) = \prod_{j=1}^{m} \left(1 - \overline{x} r_j(t)\right).$$

The expressions for $C(t)$ and $E(t)$ follow from substitution. □

Example 4.2 (Dyck Paths and Prefixes)

Let $\mathcal{S} = \{-1, 1\}$ be the unweighted step set ($w_{-1} = w_1 = 1$) with characteristic polynomial $S(x) = \overline{x} + x$. The kernel equation (4.5) becomes

$$(1 - t(\overline{x} + x))H(x,t) = 1 - t\overline{x}\,[x^0]H(x,t) = 1 - t\overline{x}H(0,t). \tag{4.6}$$

Solving the kernel $1 - t(\overline{x} + x) = 0$ for x gives two solutions

$$r_1(t) = \frac{1 - \sqrt{1-4t^2}}{2t} = t + t^3 + 2t^5 + \cdots \qquad R_1(t) = \frac{1 + \sqrt{1-4t^2}}{2t} = t^{-1} - t - t^3 + \cdots,$$

consisting of one small root and one large root. Because $r_1(t)$ is a power series with no constant term, we can substitute $x = r_1(t)$ into (4.6) to obtain

$$H(0,t) = \frac{1 - \sqrt{1-4t^2}}{2t^2} = \sum_{n \geq 0} \frac{1}{n+1}\binom{2n}{n} t^{2n},$$

meaning the number of walks of length $2n$ returning to the origin is the nth Catalan number, and

$$H(x,t) = \frac{1 - t\overline{x}H(0,t)}{1 - t(\overline{x} + x)} = \frac{1 - 2xt - \sqrt{1-4t^2}}{2t(t + tx^2 - x)}.$$

Thus

$$C(t) = H(1,t) = \frac{1}{2t}\left(\sqrt{\frac{1+2t}{1-2t}} - 1\right)$$

has dominant singularities $t = \pm 1/2$, with local expansions

$$C(t) = \sqrt{2}(1-2t)^{-1/2} + \cdots \qquad \text{and} \qquad C(t) = 1 - \frac{1}{\sqrt{2}}(1+2t)^{1/2} + \cdots.$$

Proposition 2.11 in Chapter 2 implies these singularities give contributions $2^n n^{-1/2}\sqrt{2/\pi}$ and $(-2)^n n^{-3/2}/(2\sqrt{\pi})$ to the asymptotics of $c_n = [t^n]C(t)$, so

$$c_n = 2^n n^{-1/2} \frac{\sqrt{2}}{\sqrt{\pi}} \left(1 + O\left(n^{-1}\right)\right).$$

Note that the dominant asymptotic term is not periodic, but higher order asymptotic terms do have periodicity: in fact $c_n = \binom{n}{\lfloor n/2 \rfloor}$ and this periodicity comes from the presence of the floor function.

Walks in $\mathbb{N}$ on the steps $\mathcal{S}$ which begin and end at the origin are commonly referred to as *Dyck paths*, and those without any constraint on endpoint as *Dyck prefixes*. Enumerating Dyck paths is often referred to as the 'ballot problem', since the number of Dyck paths of length $2n$ counts the number of ways two candidates running for an office, Alice and Bob, can split $2n$ votes in such a way that Alice always stays ahead of Bob as the votes are individually added (this problem was studied by Whitworth, Bertrand, André, and others in the nineteenth century). An approach of Knuth to the ballot problem is one of the first combinatorial uses of what has become the kernel method; see Section 4.2.1 below for details.

Since Proposition 4.2 gives an expression for the generating function $H(x, t)$ in terms of explicit algebraic functions, the enumeration of lattice paths restricted to a half-space is essentially a solved problem. Such an expression was originally given by Gessel [49]; see also Bousquet-Mélou and Petkovšek [25, Ex. 3] and Banderier and Flajolet [5] for asymptotic results. An alternative expression for $H(x, t)$ involving a non-negative series expansion of algebraic power series in t and $\overline{x}$ is given by Bostan et al. [12, Prop. 19].

4.1.4 Walks in the Quarter-plane

Next, we consider walks on a weighted step set $\mathcal{S}' \subset \mathbb{Z}^d$ restricted to the quarter-space $\mathcal{R}' = \mathbb{Z}^{d-2} \times \mathbb{N}^2$. As for half-space models, if we are only interested in enumerating the total number of walks in the model, or counting those ending on one or both of the boundary hyperplanes $z_{d-1} = 0$ and $z_d = 0$, we may project a walk onto its last two coordinates and consider the resulting lattice path model restricted to the quadrant $\mathcal{R} = \mathbb{N}^2$. Although lattice path models in a half-space always have algebraic generating functions by Proposition 4.2, it is possible for models in a quadrant to have D-finite, D-algebraic, and even hypertranscendental generating functions. This diversity of behaviour has led to a great deal of attention on quadrant walks in the combinatorial and probabilistic literature, with many natural questions remaining open for decades. We will soon see how rational diagonal representations play a key role in the enumeration of these objects.

Because of the possibility of very pathological behaviour for walks in a quadrant, we put some restrictions on the models we consider.

Definition 4.7 (short step models) Any step set $\mathcal{S} \subset \{\pm 1, 0\}^d$ is called a *short step set*, and a lattice path model with a short step set is called a *short step model*.

The restriction to short steps results in a kernel which is quadratic in each variable, simplifying considerations. In this section we consider only unweighted short step models in the non-negative quadrant. The enumeration of some models with longer steps is discussed in Chapter 10.

The Algebraic Kernel Method

Given a step set $\mathcal{S} \subset \{\pm 1, 0\}^2$ we define the multivariate generating function

$$Q(x, y, t) = \sum_{i,j\geq 0} \left(\sum_{n\geq 0} q_{i,j,n} x^i y^j \right) t^n \in \mathbb{Z}[x, y][[t]],$$

where $q_{i,j,n}$ counts the number of quarter-plane walks of length n on the steps in $\mathcal{S}$ which end at the point (i, j), and recall again the characteristic polynomial

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

Once more the recursive structure of a walk of length $n + 1$ as a walk of length n followed by a single step gives a kernel equation satisfied by $Q(x, y, t)$. Since $Q(x, 0, t)$ and $Q(0, y, t)$ give, respectively, the generating functions of walks ending on the x- and y-axes, and $\mathcal{S}$ contains only unit steps, the kernel equation for short step models in the quarter-plane can be written

$$xy(1 - tS(x, y))Q(x, y, t) = xy - tI(y) - tJ(x) + \varepsilon tQ(0, 0, t), \tag{4.7}$$

where $I(y) = y\left([x^{-1}]S(x, y)\right) Q(0, y, t)$ and $J(x) = x\left([y^{-1}]S(x, y)\right) Q(x, 0, t)$, and the constant

$$\varepsilon = \begin{cases} 1 & : (-1, -1) \in \mathcal{S} \\ 0 & : \text{otherwise} \end{cases}$$

accounts for potentially subtracting walks at the origin twice.

Definition 4.8 (quadrant kernel equation) Equation (4.7) is the *kernel equation* for the quadrant model defined by $\mathcal{S}$, and the Laurent polynomial $K(x, y, t) = 1 - tS(x, y)$ is called the *kernel* of the model.

Our solution of half-space models involved finding the roots of a bivariate kernel in one variable; the additional variable now present in the kernel for quadrant walks complicates this approach as we would need to consider algebraic surfaces defined by $K(x, y, t) = 0$ instead of algebraic curves.

To work around this difficulty, Bousquet-Mélou [21], inspired by probabilistic work of Fayolle et al. [37], developed the so-called *algebraic kernel method*, which does not require finding roots of the kernel. The key is to introduce a group of substitutions which fix the kernel $K(x, y, t)$, and leverage information obtained through application of the group elements to get an expression for $Q(x, y, t)$ as the non-negative series extraction of a rational function.

Because we assume that $\mathcal{S}$ has short steps, there exist unique Laurent polynomials $A_j(y)$ and $B_j(x)$ for $j \in \{-1, 0, 1\}$ such that

$$S(x, y) = xA_1(y) + A_0(y) + \overline{x}A_{-1}(y) = yB_1(x) + B_0(x) + \overline{y}B_{-1}(x).$$

When $\mathcal{S}$ contains a step with negative x-coordinate, and a step with negative y-coordinate, the Laurent polynomials $A_1(y)$ and $B_1(x)$ are non-zero. The transformations

$$\Psi\colon (x, y) \mapsto \left(\overline{x}\,\frac{A_{-1}(y)}{A_1(y)}, y\right) \quad \text{and} \quad \Phi\colon (x, y) \mapsto \left(x, \overline{y}\,\frac{B_{-1}(x)}{B_1(x)}\right)$$

then fix $S(x, y)$, and thus also $K(x, y, t)$.

Definition 4.9 (group of a short-step quadrant model) The *group* $\mathcal{G}$ of the lattice path model determined by $\mathcal{S}$ is the group of transformations of the xy-plane generated by the involutions Ψ and Φ under composition (containing Ψ, Φ, $\Psi \circ \Phi$, $\Phi \circ \Psi$, etc.).

We can view an element $\sigma \in \mathcal{G}$ as a map from Laurent polynomials to iterated Laurent series which takes $f \in \mathbb{C}[\overline{x}, \overline{y}, x, y]$ and returns

$$\sigma \cdot f(x, y) = \sigma\left(f(x, y)\right) = f(\sigma(x, y)) \in \mathbb{C}((x, y)),$$

and extend this to elements $\sum_{n\geq 0} f_n(x, y)t^n \in \mathbb{C}[\overline{x}, \overline{y}, x, y][[t]]$ by defining

$$\sigma \cdot \sum_{n\geq 0} f_n(x, y)t^n = \sigma\left(\sum_{n\geq 0} f_n(x, y)t^n\right) = \sum_{n\geq 0} f_n(\sigma(x, y))t^n \in \mathbb{C}((x, y))[[t]].$$

Example 4.3 (The Group for N-S-E-W Quarter-Plane Walks)

Let $\mathcal{S} = \{(\pm 1, 0), (0, \pm 1)\}$ be the set of cardinal directions. Then

$$S(x, y) = x + (y + \overline{y}) + \overline{x} = y + (x + \overline{x}) + \overline{y}$$

and the kernel equation reads

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

The group $\mathcal{G}$ is generated by the maps

$$\Psi(x, y) = (\overline{x}, y) \quad \text{and} \quad \Phi(x, y) = (x, \overline{y}),$$

so it is the group of order 4 whose elements map (x, y) to one of $(x^{\pm 1}, y^{\pm 1})$.

Remark 4.3 Following the common convention in the lattice path literature, we will not include the stationary step $(0, 0)$ in our step sets. Definition 4.9 implies that adding the stationary step does not change the group of a model, and our arguments below generalize naturally to models obtained by adding the stationary step.

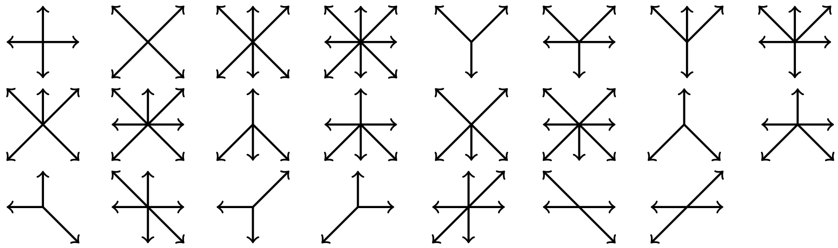


Fig. 4.2 The 23 short step sets defining quarter-plane models with finite group $\mathcal{G}$, up to rotation over the line $y = x$ and with the models trivially isomorphic to half-plane models removed. Each arrow represents an element of the set $\{\pm 1, 0\}^2 \setminus \{(0,0)\}$, and each collection of arrows with common base point defines a step set.

A priori, there are $2^8 - 1$ possible non-empty step sets $\mathcal{S} \subset \{\pm 1, 0\}^2 \setminus \{(0,0)\}$ to consider, but the situation is actually much better. Indeed, if we are interested in models which truly need to be viewed in a quarter-plane then we do not need to consider step sets $\mathcal{S}$

- where every step has at least one negative coordinate, as there are no valid walks of positive length;
- with no steps having positive x-coordinate, or no steps having positive y-coordinate, as they are isomorphic to half-plane models;
- with no steps having negative x-coordinate, or no steps having negative y-coordinate, as they are isomorphic to half-plane models;
- such that $j \geq i$ for all $(i, j) \in \mathcal{S}$, as a sequence of steps whose x-coordinate stays non-negative must have y-coordinates staying non-negative, and the model is isomorphic to a half-plane model;
- such that $i \geq j$ for all $(i, j) \in \mathcal{S}$, as a sequence of steps whose y-coordinate stays non-negative must have x-coordinates staying non-negative, and the model is isomorphic to a half-plane model.

Furthermore, as the quarter-plane is symmetric over the line $y = x$ if $\mathcal{S}$ is a step set and $\mathcal{S}'$ is the step set obtained by reversing the coordinates of the steps in $\mathcal{S}$, then the lattice path models defined by $\mathcal{S}$ and $\mathcal{S}'$ are isomorphic. Removing these half-plane models in disguise, Problem 4.2 asks you to prove there are 79 remaining quarter-plane models up to rotation of $\mathcal{S}$ over the line $y = x$; see Figures 4.2 and 4.4. This classification of short step quarter-plane models was originally given by Bousquet-Mélou and Mishna [24].

Example 4.4 (A One-Dimensional Model in The Plane)

If $\mathcal{S} = \{(-1, 0), (0, 1), (0, -1)\}$ then the first step in $\mathcal{S}$ can never be used, so the quadrant lattice path model defined by $\mathcal{S}$ is equivalent to the one-dimensional model in a half-space counting Dyck prefixes.

Generating Function Representations for Finite Group Models

For each of the 79 quarter-plane models now under consideration the maps Ψ and Φ are explicitly determined by $\mathcal{S}$, and when the group $\mathcal{G}$ is finite it is easily generated in a computer algebra system. Figure 4.2 shows the 23 models with finite group; the remaining 56 models, which have infinite group, are discussed below. When $\mathcal{G}$ is finite, it can often be leveraged to determine the generating function $Q(x, y, t)$.

Example 4.5 (The Generating Function for N-S-E-W Quarter-Plane Walks)

When $\mathcal{S} = \{(\pm 1, 0), (0, \pm 1)\}$ the group $\mathcal{G}$ consists of the maps sending (x, y) to one of $(x^{\pm 1}, y^{\pm 1})$. Since the elements of $\mathcal{G}$ fix the kernel $K(x, y, t)$, applying the group elements to the kernel equation gives the system of equations

$$\begin{aligned}
xy(1 - t(x + \overline{x} + y + \overline{y}))Q(x, y, t) &= xy - tyQ(0, y) - txQ(x, 0)\\
\overline{x}y(1 - t(x + \overline{x} + y + \overline{y}))Q(\overline{x}, y, t) &= \overline{x}y - tyQ(0, y) - t\overline{x}Q(\overline{x}, 0)\\
\overline{x}\,\overline{y}(1 - t(x + \overline{x} + y + \overline{y}))Q(\overline{x}, \overline{y}, t) &= \overline{x}\,\overline{y} - t\overline{y}Q(0, \overline{y}) - t\overline{x}Q(\overline{x}, 0)\\
x\overline{y}(1 - t(x + \overline{x} + y + \overline{y}))Q(x, y, t) &= x\overline{y} - t\overline{y}Q(0, \overline{y}) - txQ(x, 0).
\end{aligned}$$

The crux of the algebraic kernel method is that each of the unknown functions $Q(0, y), Q(x, 0), Q(\overline{x}, 0), Q(0, \overline{y})$ appears exactly twice in this system. Taking an alternating sum of these equations yields, after some algebraic manipulation,

$$xyQ(x, y, t) - \overline{x}yQ(\overline{x}, y, t) + \overline{x}\,\overline{y}Q(\overline{x}, \overline{y}, t) - x\overline{y}Q(x, \overline{y}, t) = \frac{xy - \overline{x}y + \overline{x}\,\overline{y} - x\overline{y}}{1 - t(x + \overline{x} + y + \overline{y})}.$$

Since the generating function $Q(x, y, t)$ lies in $\mathbb{Z}[x, y][[t]]$, every term in $xyQ(x, y, t)$ contains positive powers of x and y. In contrast, when the remaining summands $\overline{x}yQ(\overline{x}, y, t)$, $\overline{x}\,\overline{y}Q(\overline{x}, \overline{y}, t)$, and $x\overline{y}Q(x, \overline{y}, t)$ are expanded in $\mathbb{Z}[\overline{x}, x, \overline{y}, y][[t]]$ every term contains a negative power of x or a negative power of y. Thus, recalling the non-negative series extraction operator from Section 3.3.2 of Chapter 3,

$$Q(x, y, t) = [x^{\geq 0}y^{\geq 0}]\frac{xy - \overline{x}y + \overline{x}\,\overline{y} - x\overline{y}}{xy\,(1 - t(x + \overline{x} + y + \overline{y}))} = [x^{\geq 0}y^{\geq 0}]\frac{(x - \overline{x})(y - \overline{y})}{xy\,(1 - t(x + \overline{x} + y + \overline{y}))}.$$

Proposition 3.14 from Chapter 3 then implies the generating function counting the number of walks in the model is the main diagonal

$$C(t) = Q(1, 1, t) = \Delta\left(\frac{(1 + x)(1 + y)}{1 - txy(\overline{x} + x + \overline{y} + y)}\right).$$

This shows $C(t)$ is D-finite, and the creative telescoping package of Lairez [66] gives the D-finite equation

$$\begin{aligned}
t^2(4t - 1)(4t + 1)C'''(t) &+ 2t(4t + 1)(16t - 3)C''(t)\\
&+ (224t^2 + 28t - 6)C'(t) + (12 + 64t)C(t) = 0.
\end{aligned}$$

In Chapter 3 we determined coefficient asymptotics of $C(t)$ up to a constant factor by analyzing this differential equation. In Chapter 5 we will determine dominant asymptotics exactly using analytic combinatorics in several variables. The number of quarter-plane walks of length n on the steps $\mathcal{S}$ has dominant asymptotic growth $\frac{4}{\pi}\,\frac{4^n}{n}$, so Corollary 2.1 from Chapter 2 implies $C(t)$ is transcendental.

The argument in this example generalizes, forming the algebraic kernel method. Since Ψ and Φ are involutions, when $\mathcal{G}$ is finite it has even order $2k$ and any element $g \in \mathcal{G}$ can be written uniquely in the form

$$g = \Phi \circ \Psi \circ \cdots \circ \Phi \circ \Psi \qquad \text{or} \qquad g = \Psi \circ \Phi \circ \cdots \circ \Phi \circ \Psi,$$

where there are $0 \leq r < 2k$ terms in the correct composition.

Definition 4.10 (sign) The *sign* of $g \in \mathcal{G}$ expressed in this manner is $\mathrm{sgn}(g) = (-1)^r$.

Proposition 4.3 *Assume that the group $\mathcal{G}$ is finite. Then*

$$\sum_{g \in \mathcal{G}} \mathrm{sgn}(g)\, g(xyQ(x, y, t)) = \frac{1}{1 - tS(x, y)} \sum_{g \in \mathcal{G}} \mathrm{sgn}(g)\, g(xy). \tag{4.8}$$

As the elements of $\mathcal{G}$ are bi-rational transformations, the composition $g(xyQ(x, y, t))$ results in an element of $\mathbb{Z}((x, y))[[t]]$; that is, a power series in t whose coefficients are iterated Laurent series in x and y.

Proof Let $X(x, y)$ and $Y(x, y)$ denote the rational functions defined by $(X, y) = \Psi(x, y)$ and $(x, Y) = \Phi(x, y)$. Applying the maps Ψ and Φ successively to (4.7) gives

$$\begin{aligned}
&(\mathrm{id}) && xy(1 - tS(x, y))Q(x, y, t) = xy - tI(y) - tJ(x) + \varepsilon tQ(0, 0, t)\\
&(\Psi) && Xy(1 - tS(X, y))Q(X, y, t) = Xy - tI(y) - tJ(X) + \varepsilon tQ(0, 0, t)\\
&(\Phi \circ \Psi) && XY(1 - tS(X, Y))Q(X, Y, t) = XY - tI(Y) - tJ(X) + \varepsilon tQ(0, 0, t),
\end{aligned}$$

since both Ψ and Φ fix $S(x, y)$. As both $-tI(y)$ and $-tJ(X)$ appear on the right-hand sides of consecutive equations, taking an alternating sum of these three equations cancels those terms. In fact, since Ψ and Φ each fix one coordinate, continuing to compose the group generators in this manner and taking an alternating sum of the resulting equations successively cancels all unknown functions of the form $I(Y')$ and $J(X')$ which appear on the right-hand side. Because the group is finite and of even order, this repeated composition of group elements returns to the identity, resulting in (4.8). □

Definition 4.11 (orbit sum equations) Equation (4.8) is known as the *orbit sum equation* associated to the lattice path model defined by $\mathcal{S}$.

Although it becomes difficult to define the group of a model when $\mathcal{S}$ no longer has small steps, the orbit sum equation can be generalized in several situations [12].

Fig. 4.3 The four step sets defining quadrant models with finite group to which Proposition 4.4 does not apply.

A short examination[1] of (4.8) for the finite group models in Figure 4.2 shows that in 19 cases $xyQ(x, y, t)$ is the only term on the left-hand side which contains only non-negative powers of x and y when expanded in $\mathbb{Z}((x, y))[[t]]$, as in the example above. This gives the following.

Proposition 4.4 *Let $\mathcal{S}$ be one of the 23 step sets defining a model with finite group, displayed in Figure 4.2. If $\mathcal{S}$ is not one of the four models listed in Figure 4.3 then*

$$Q(x, y, t) = [x^{\geq 0} y^{\geq 0}]\left(\frac{1}{1 - tS(x, y)} \sum_{g \in \mathcal{G}} \text{sgn}(g)\, g(xy)\right).$$

When the rational function in Proposition 4.4 has an expansion in $\mathbb{Q}[\overline{x}, \overline{y}, x, y][[t]]$, Proposition 3.14 in Chapter 3 immediately gives a diagonal expression for the generating functions of walks ending anywhere in the quarter plane, returning to the origin, or ending on either boundary axis. Although there are models where the rational function in Proposition 4.4 cannot be expanded with Laurent polynomial coefficients, these models are symmetric over one axis and a more involved argument shows that the same diagonal expression holds (see Proposition 4.8 below). Thus, we obtain the following.

Theorem 4.1 *Let $\mathcal{S}$ be one of the 19 short step sets in Figure 4.2 which is not listed in Figure 4.3. Then for $a, b \in \{0, 1\}$,*

$$Q(a, b, t) = \Delta\left(\frac{O(\overline{x}, \overline{y})}{(1 - x)^a (1 - y)^b (1 - txyS(\overline{x}, \overline{y}))}\right),$$

where O is the orbit sum $O(x, y) = \sum_{g \in \mathcal{G}} \text{sgn}(g) g(xy)$.

In particular, the generating functions $Q(1, 1, t)$ for walks ending anywhere in the quadrant, $Q(1, 0, t)$ and $Q(0, 1, t)$ for walks ending on one of the boundary axes, and $Q(0, 0, t)$ for walks ending at the origin, are D-finite. In fact, Proposition 4.4 implies the multivariate generating function $Q(x, y, t)$ is D-finite, meaning the $\mathbb{Q}(x, y, t)$-vector space spanned by all its partial derivatives is finite-dimensional. The results of Chapter 10 give asymptotics for these models; see Remark 10.3 of Chapter 10. Table 4.1 summarizes the asymptotic behaviour of these models.

In the case of the remaining four models with finite group, with step sets in Figure 4.3, both sides of the orbit sum equation (4.8) are identically zero due to an element $g \in \mathcal{G}$ of negative sign which fixes the product xy.

[1] See the computer algebra worksheets available on the book website.

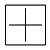
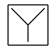
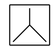
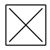
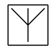
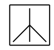
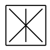
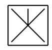
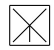
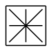
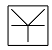
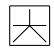
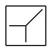
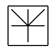
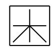
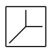
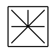
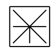
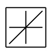
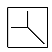
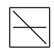
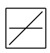
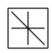
$\mathcal{S}$	Asymptotics	$\mathcal{S}$	Asymptotics	$\mathcal{S}$	Asymptotics
	$\frac{4}{\pi}\frac{4^n}{n}$		$\frac{\sqrt{3}}{2\sqrt{\pi}}\frac{3^n}{\sqrt{n}}$		$\frac{A_n}{\pi}\frac{(2\sqrt{2})^n}{n^2}$
	$\frac{2}{\pi}\frac{4^n}{n}$		$\frac{4}{3\sqrt{\pi}}\frac{4^n}{\sqrt{n}}$		$\frac{B_n}{\pi}\frac{(2\sqrt{3})^n}{n^2}$
	$\frac{\sqrt{6}}{\pi}\frac{6^n}{n}$		$\frac{\sqrt{5}}{3\sqrt{2\pi}}\frac{5^n}{\sqrt{n}}$		$\frac{C_n}{\pi}\frac{(2\sqrt{6})^n}{n^2}$
	$\frac{8}{3\pi}\frac{8^n}{n}$		$\frac{\sqrt{5}}{2\sqrt{2\pi}}\frac{5^n}{\sqrt{n}}$		$\frac{\sqrt{8}(1+\sqrt{2})^{7/2}}{\pi}\frac{(2+2\sqrt{2})^n}{n^2}$
	$\frac{2\sqrt{2}}{\Gamma(1/4)}\frac{3^n}{n^{3/4}}$		$\frac{2\sqrt{3}}{3\sqrt{\pi}}\frac{6^n}{\sqrt{n}}$		$\frac{\sqrt{3}(1+\sqrt{3})^{7/2}}{2\pi}\frac{(2+2\sqrt{3})^n}{n^2}$
	$\frac{3\sqrt{3}}{\sqrt{2}\Gamma(1/4)}\frac{3^n}{n^{3/4}}$		$\frac{\sqrt{7}}{3\sqrt{3\pi}}\frac{7^n}{\sqrt{n}}$		$\frac{\sqrt{570-114\sqrt{6}}(24\sqrt{6}+59)}{19\pi}\frac{(2+2\sqrt{6})^n}{n^2}$
	$\frac{\sqrt{6\sqrt{3}}}{\Gamma(1/4)}\frac{6^n}{n^{3/4}}$		$\frac{3\sqrt{3}}{2\sqrt{\pi}}\frac{3^n}{n^{3/2}}$		$\frac{8}{\pi}\frac{4^n}{n^2}$
	$\frac{4\sqrt{3}}{3\Gamma(1/3)}\frac{4^n}{n^{2/3}}$		$\frac{3\sqrt{3}}{2\sqrt{\pi}}\frac{6^n}{n^{3/2}}$		

Table 4.1 Asymptotics for the 23 D-finite models.

$$A_n = \begin{cases} 24\sqrt{2} & : n \text{ even} \\ 32 & : n \text{ odd} \end{cases}, \quad B_n = \begin{cases} 12\sqrt{3} & : n \text{ even} \\ 18 & : n \text{ odd} \end{cases}, \quad C_n = \begin{cases} 12\sqrt{30} & : n \text{ even} \\ 144/\sqrt{5} & : n \text{ odd} \end{cases}$$

Example 4.6 (A Zero Orbit Model)

The model defined by the final step set $\mathcal{S} = \{(\pm 1, 0), (-1, -1), (1, 1)\}$ in Figure 4.3 is known as *Gessel's model*. Here

$$\Psi = \left(\frac{1}{xy}, y\right) \qquad \text{and} \qquad \Phi = \left(x, \frac{1}{yx^2}\right),$$

and the group $\mathcal{G}$, of order 8, consists of the maps

$$\begin{array}{ccccccccc} & & \left(\frac{1}{xy}, y\right) & \xrightarrow{\Phi} & \left(\frac{1}{xy}, yx^2\right) & \xrightarrow{\Psi} & \left(\frac{1}{x}, yx^2\right) & & \\ & \overset{\Psi}{\nearrow} & & & & & & \overset{\Phi}{\searrow} & \\ (x, y) & & & & & & & & \left(\frac{1}{x}, \frac{1}{y}\right) \\ & \underset{\Phi}{\nwarrow} & & & & & & \underset{\Psi}{\swarrow} & \\ & & \left(x, \frac{1}{yx^2}\right) & \xleftarrow[\Psi]{} & \left(xy, \frac{1}{yx^2}\right) & \xleftarrow[\Phi]{} & \left(xy, \frac{1}{y}\right) & & \end{array}$$

Since $(\Psi \circ \Phi \circ \Psi)(xy) = xy$, and $\Psi \circ \Phi \circ \Psi$ has odd sign, the orbit sum is zero.

Although the above argument does not quite work for the models in Figure 4.3, it turns out that the multivariate generating function $Q(x, y, t)$ for each model is

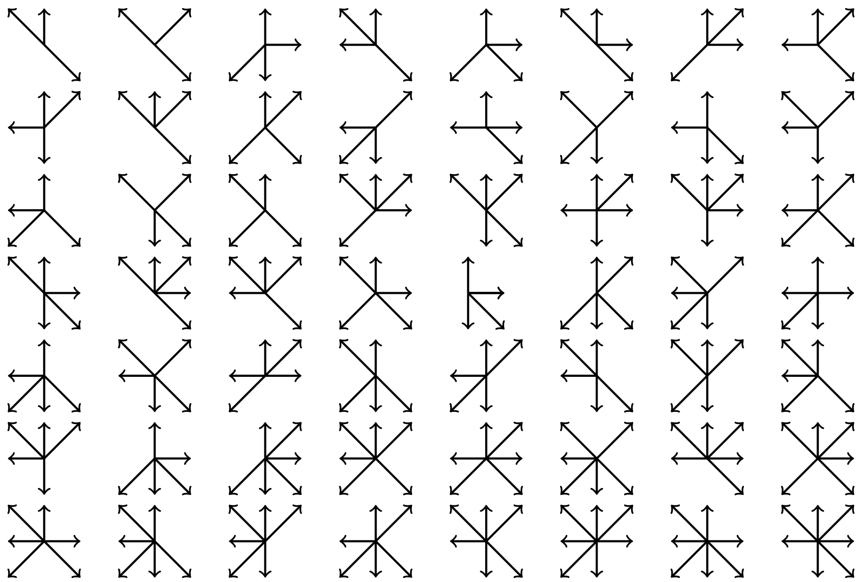


Fig. 4.4 The 56 short step sets defining non-isomorphic quarter plane models with infinite group.

algebraic (and thus also D-finite). The first three models can be analyzed using a variant of the kernel method known as the *obstinate kernel method*, also introduced by Bousquet-Mélou [21]; details can be found in Bousquet-Mélou and Mishna [24]. Algebraicity for the final model, Gessel's model, was open for several years before being proven by a computer algebra approach of Bostan and Kauers [16]. Because these models are algebraic, they can be written as the diagonals of bivariate functions using Proposition 3.8 in Chapter 3, but such representations are less elegant than those obtained directly through the kernel method and Theorem 4.1. Asymptotics can be derived directly from the algebraic equations satisfied by each model's generating function, so we do not consider these four models in any additional detail.

Infinite Group Models

Figure 4.4 shows the 56 short step sets defining quadrant models with infinite group. As can be expected, proving that the group of a model is infinite requires more work than proving finiteness (it is harder to prove the absence of structure than its existence). Since Ψ and Φ are involutions, the group $\mathcal{G}$ they generate is finite if and only if some iteration of the composed map $\Theta = \Phi \circ \Psi$ has finite order: the key is to exhibit some property of Θ which shows no repeated composition Θ^n for integer $n \geq 1$ is the identity. We briefly detail an argument which can be used for this purpose, following the presentation of Bostan et al. [12] which expands on ideas found in Bousquet-Mélou and Mishna [24] and Bostan et al. [17]. First we prove the following result, stated in a general setting as it will be used again in later chapters.

Proposition 4.5 *Let $\mathcal{S} \subset \mathbb{Z}^d$ be a finite set not contained in a half-space (i.e., there does not exist $\mathbf{n} \in \mathbb{Z}^d$ such that $\mathbf{i} \cdot \mathbf{n} \geq 0$ for all $\mathbf{i} \in \mathcal{S}$). If $S(\mathbf{z}) = \sum_{\mathbf{i} \in \mathcal{S}} w_{\mathbf{i}} \mathbf{z}^{\mathbf{i}}$ is a weighted characteristic polynomial with $w_{\mathbf{i}} > 0$ for all $\mathbf{i} \in \mathcal{S}$ then the system*

$$S_{z_1}(\mathbf{z}) = S_{z_2}(\mathbf{z}) = \cdots = S_{z_d}(\mathbf{z}) = 0$$

defined by the partial derivatives of S admits a unique solution in $\mathbb{R}^d_{>0}$, which is the unique minimum of S on $\mathbb{R}^d_{>0}$.

Definition 4.12 (vanishing points) A point where all partial derivatives of S vanish is called a *vanishing point* of S. (These are typically called critical points, but we reserve this name for another concept arising in Chapter 5.)

Proof Let

$$L(\mathbf{x}) = S(e^{\mathbf{x}}) = S(e^{x_1}, \dots, e^{x_d}) = \sum_{\mathbf{i} \in \mathcal{S}} w_{\mathbf{i}} e^{\mathbf{i} \cdot \mathbf{x}}. \tag{4.9}$$

Then for $1 \leq j \leq d$ the chain rule implies $L_{x_j}(\mathbf{x}) = e^{x_j} S_{z_j}(e^{\mathbf{x}})$, so

$$S_{z_1}(\mathbf{z}) = S_{z_2}(\mathbf{z}) = \cdots = S_{z_d}(\mathbf{z}) = 0$$

for $\mathbf{z} \in \mathbb{R}^d_{>0}$ if and only if

$$L_{x_1}(\mathbf{x}) = L_{x_2}(\mathbf{x}) = \cdots = L_{x_d}(\mathbf{x}) = 0,$$

where $\mathbf{x} = (\log z_1, \dots, \log z_d) \in \mathbb{R}^d$. Thus, to characterize the vanishing points of S in $\mathbb{R}^d_{>0}$ it is enough to search for vanishing points of L in $\mathbb{R}^d$. The function $L(\mathbf{x})$ is known as the *Laplace transform* of $S(\mathbf{x})$.

The Laplace transform is useful because the exponential function is *strictly convex*: if $a, b > 0$ and $a + b = 1$ then $e^{at+bs} \leq ae^t + be^s$, with equality if and only if $t = s$. As $L(\mathbf{x})$ is a positive linear combination of exponentials, if $a, b > 0$ with $a + b = 1$ then $L(a\mathbf{x} + b\mathbf{y}) \leq aL(\mathbf{x}) + bL(\mathbf{y})$ with equality if and only $\mathbf{i} \cdot \mathbf{x} = \mathbf{i} \cdot \mathbf{y}$ for all $\mathbf{i} \in \mathcal{S}$. Thus, $L(\mathbf{x})$ is strictly convex unless there exists a non-zero vector $\mathbf{n} \in \mathbb{R}^d$ such that $\mathbf{i} \cdot \mathbf{n} = 0$ for all $\mathbf{i} \in \mathcal{S}$, which cannot occur as $\mathcal{S}$ is not contained in a half-space.

If $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ and $\lambda \in (0, 1)$ this implies

$$L(\lambda\mathbf{y} + (1 - \lambda)\mathbf{x}) < \lambda L(\mathbf{y}) + (1 - \lambda)L(\mathbf{x}),$$

so

$$\frac{L(\mathbf{x} + \lambda(\mathbf{y} - \mathbf{x})) - L(\mathbf{x})}{\lambda} < L(\mathbf{y}) - L(\mathbf{x})$$

and taking $\lambda \to 0$ gives

$$(\mathbf{y} - \mathbf{x})^T \cdot \nabla L(\mathbf{x}) + L(\mathbf{x}) < L(\mathbf{y}).$$

In particular, any vanishing point of L, where its gradient vanishes, is the global minimum of the convex function L on $\mathbb{R}^d$, so L admits at most one vanishing point. Furthermore, for $\mathbf{w} \in \mathbb{R}^d$ and $t \in \mathbb{R}$ we have

$$L(t\mathbf{w}) = \sum_{\mathbf{i}\in\mathcal{S}} w_{\mathbf{i}} e^{t(\mathbf{i}\cdot\mathbf{w})},$$

so $L(t\mathbf{w}) \to \infty$ when $\mathbf{w} \neq \mathbf{0}$ is fixed and t approaches either $\pm\infty$, unless $\mathbf{i}\cdot\mathbf{w} \geq 0$ for all $\mathbf{i} \in \mathcal{S}$ or $\mathbf{i}\cdot\mathbf{w} \leq 0$ for all $\mathbf{i} \in \mathcal{S}$. Since $\mathcal{S}$ is not contained in a half-plane we can conclude $L(\mathbf{z}) \to \infty$ as $|\mathbf{z}| \to \infty$. This shows $L(\mathbf{x})$ admits a global minimizer, which must be its unique vanishing point, and taking coordinate-wise exponentials gives the unique vanishing point of $S(\mathbf{z})$ with positive coordinates. □

This vanishing point with positive coordinates gives a necessary condition for a lattice path model to have finite group.

Proposition 4.6 *Let $\mathcal{S}$ be a two-dimensional small step model not contained in a half-plane and let (a, b) be the unique vanishing point of $S(x, y)$ with positive coordinates guaranteed by Proposition 4.5. If the group of transformations generated by Φ and Ψ is finite then*

$$\frac{S_{xy}(a,b)}{\sqrt{S_{xx}(a,b)\,S_{yy}(a,b)}} = \cos\theta, \tag{4.10}$$

where θ is some rational multiple of π.

Proof Recall that

$$S(x,y) = xA_1(y) + A_0(y) + \overline{x}A_{-1}(y) = yB_1(x) + B_0(x) + \overline{y}B_{-1}(x)$$

for Laurent polynomials $A_j(y)$ and $B_j(x)$, so

$$S_x(x,y) = A_1(y) - \overline{x}^2 A_{-1}(y) \qquad \text{and} \qquad S_y(x,y) = B_1(x) - \overline{y}^2 B_{-1}(x).$$

Since (a, b) is a vanishing point of $S(x, y)$, this implies

$$\overline{a}\,\frac{A_{-1}(b)}{A_1(b)} = a \qquad \text{and} \qquad \overline{b}\,\frac{B_{-1}(a)}{B_1(a)} = b,$$

so $(a, b) = \Psi(a, b) = \Phi(a, b)$ is a fixed point of $\mathcal{G}$. If $\Theta = \Phi \circ \Psi$ then taking a local expansion of Θ centred at (a, b) implies

$$\Theta(a+x, b+y) = (a,b) + (x,y)\cdot J(a,b) + \text{higher order terms}, \tag{4.11}$$

where $J(a, b)$ is the Jacobian matrix of Θ, defined by $J(a,b) = \begin{pmatrix} -1 & -\eta \\ \nu & \nu\eta - 1 \end{pmatrix}$ for

$$\eta = \frac{2S_{xy}(a,b)}{S_{xx}(a,b)} \qquad \text{and} \qquad \nu = \frac{2S_{xy}(a,b)}{S_{yy}(a,b)}.$$

Note that $S_{xx}(a,b) = 2a^{-3}A_{-1}(b)$ and $S_{yy}(a,b) = 2b^{-3}A_{-1}(a)$ are non-zero as $a, b > 0$. Because (a,b) is a fixed point of Θ, Equation (4.11) implies that the iterations Θ^r of the map Θ satisfy

$$\Theta^r(a+x, b+y) = (a,b) + (x,y) \cdot J(a,b)^r + \text{higher order terms}$$

for all $r \in \mathbb{N}$. Thus, if some repeated composition Θ^r is the identity map then $J(a,b)^r$ is the identity matrix, meaning all eigenvalues of $J(a,b)$ are rth roots of unity. By direct calculation the eigenvalues of J satisfy $\lambda^2 - (\eta\nu - 2)\lambda + 1 = 0$, so if the group associated to $\mathcal{S}$ is finite then there exists some rational multiple of π, which we denote θ, such that $\lambda^2 - (\eta\nu - 2)\lambda + 1 = (\lambda - e^{2i\theta})(\lambda - e^{-2i\theta})$. Comparing the coefficients of λ on both sides of this equation implies

$$\frac{S_{xy}(a,b)^2}{S_{xx}(a,b)\,S_{yy}(a,b)} = \frac{\eta\nu}{4} = \frac{e^{2i\theta} + e^{-2i\theta} + 2}{4} = \frac{\cos(2\theta)+1}{2} = \cos^2\theta,$$

as desired. □

Because the left-hand side of (4.10) is an algebraic number, Proposition 4.6 gives an effective method to detect when $\mathcal{S}$ has infinite group. In particular, it is sufficient to do the following:

1. Using resultants, determine the minimal polynomial $M(\lambda)$ of the values of λ in the two solutions of

$$\lambda^2 - (\eta\nu - 2)\lambda + 1 = S_x(a,b) = S_y(a,b) = 0,$$

$$\eta = \frac{2S_{xy}(a,b)}{S_{xx}(a,b)}, \qquad \nu = \frac{2S_{xy}(a,b)}{S_{yy}(a,b)}, \qquad a, b > 0.$$

2. Check whether $M(\lambda)$ is a cyclotomic polynomial (i.e., the minimal polynomial of a root of unity)

Because cyclotomic polynomials can be efficiently computed [1], this gives an automated method of proving that a step set admits an infinite group; it immediately proves that the 51 models in Figure 4.4 not contained in a half-plane have infinite groups. Bostan et al. [17] use a similar approach to show that the generating function $Q(0,0,t)$ of excursions is non-D-finite for these 51 models, giving the following.

Proposition 4.7 *Let $\mathcal{S}$ be a two-dimensional small step model not contained in a half-plane. Unless*

$$c = \frac{S_{xy}(a,b)}{\sqrt{S_{xx}(a,b)\,S_{yy}(a,b)}}$$

can be written $c = \cos\theta$, where θ is some rational multiple of π, the generating function $Q(0,0,t)$ is non-D-finite.

The idea behind Proposition 4.7 is to leverage probabilistic results of Denisov and Wachtel [32] to show that the number of excursions e_n of a model satisfies $e_{2n} \sim$

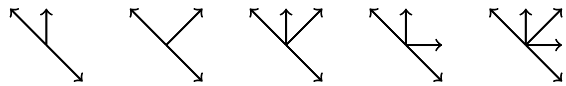


Fig. 4.5 The 5 singular models, whose step sets are contained in a half-plane.

$C\,\rho^n\,n^{\alpha}$, where $\alpha = -1 - \pi/\arccos(-c)$. By Corollary 2.2 of Chapter 2, $Q(0,0,t)$ cannot be D-finite unless c has the stated form (looking only at e_{2n} helps account for periodicities present in some models). Because we focus mainly on models which are D-finite, where asymptotics can be determined by studying rational diagonals, we refer the interested reader to Bostan et al. [17] for details. Bostan et al. [12] generalize this argument to quadrant walks with non-small step sets, closely linking having an infinite group to the non-D-finiteness of excursions; see also Bostan [10].

Example 4.7 (An Infinite Group Model)

Consider the set of steps $\mathcal{S} = \{(-1,-1),(0,-1),(0,1),(1,0),(-1,0)\}$ with characteristic polynomial $S(x,y) = 1/(xy) + 1/y + y + x + 1/x$. Then

$$\Psi(x,y) = \left(\frac{1+\overline{y}}{x}, y\right) \qquad \text{and} \qquad \Phi(x,y) = \left(x, \frac{1+\overline{x}}{y}\right),$$

so

$$\Theta(x,y) = (\Phi\circ\Psi)(x,y) = \Phi\left(\frac{1+\overline{y}}{x}, y\right) = \left(\frac{1+\overline{y}}{x}, \frac{1+x+\overline{y}}{1+y}\right).$$

The eigenvalues of the Jacobian matrix J of Θ at the vanishing point of $S(x,y)$ with positive coordinates $a,b>0$ satisfy

$$\lambda^2 - \left(\frac{1}{(1+a)(1+b)} - 2\right)\lambda + 1 = a^2b - b - 1 = ab^2 - a - 1 = 0,$$

and a resultant calculation implies that at any solution

$$(\lambda^2+3\lambda+1)(\lambda^6+8\lambda^5+28\lambda^4+41\lambda^3+28\lambda^2+8\lambda+1) = 0.$$

Since neither of these irreducible factors is a cyclotomic polynomial, no power of J is the identity and the group of $\mathcal{S}$ is infinite. The work of Denisov and Wachtel [32] implies that the number of excursions has dominant asymptotics $e_n \sim C\,S(a,b)^n\,n^{-\alpha}$ for some constant $C>0$ and irrational

$$\alpha = 1 + \frac{\pi}{\arccos\left(\frac{1}{2\sqrt{(1+a)(1+b)}}\right)} = 2.757\ldots.$$

Although the generating functions enumerating walks returning to the origin are non-D-finite for these infinite group models, the precise nature of the generating functions enumerating walks ending anywhere in the quadrant is still unknown.

Open Problem 4.1 *Determine which lattice path models defined by the step sets in Figure 4.4 have non-D-finite generating function* $Q(1, 1, t)$.

We end our discussion of quadrant walks with the 5 models of Figure 4.4 whose step sets lie in a half-plane; known as *singular models*, they are shown in Figure 4.5. The singular models can be dealt with [76, 71] using another kernel method variant, the *iterated kernel method*. The key is that kernel $K(x, y, t)$ admits a root $y = Y_+(x, t)$ which is a power series in x and t, and a root $x = X_+(y, t)$ which is a power series in y and t, such that Y_+ (respectively X_+) has a lowest-order term which contains positive powers in t and x (respectively t and y). Because of this, X_+ and Y_+ can be repeatedly composed to obtain functions whose lower order terms have increasingly large powers of x and y; each of these compositions is in the orbit of (x, y) under the group of the model, showing the group is infinite. Furthermore, substituting these repeated compositions into the kernel equation and taking an *infinite* alternating series gives an explicit infinite sum representation for the generating function $Q(x, y, t)$. For example, if $\mathcal{S}$ is one of the 3 singular models whose step set is symmetric over the line $y = x$ and $Y_n(x, t) = Y_{n-1}(Y_+(x, t), t)$ for $n \in \mathbb{N}$ with $Y_0 = x$ then the generating function counting all walks satisfies

$$Q(1, 1, t) = \frac{1}{1 - |\mathcal{S}|t} \left(1 - 2 \sum_{n=0}^{\infty} (-1)^n Y_n(1, t) Y_{n+1}(1, t)\right).$$

Elementary arguments show that each summand in this infinite series contributes at least one unique singularity to $Q(1, 1, t)$, so $Q(1, 1, t)$ admits an infinite number of singularities and is thus non-D-finite. Details can be found in Melczer and Mishna [71].

4.1.5 Orthant Walks Whose Step Sets Have Symmetries

Moving on from quadrant walks, we now fix a dimension d and study lattice path models restricted to the orthant $\mathbb{N}^d$ whose step sets have many symmetries. In particular, we consider models defined by a weighted step set $\mathcal{S} \subset \{\pm 1, 0\}^d$, where each $\mathbf{i} \in \mathcal{S}$ is given real weight $w_{\mathbf{i}} > 0$ such that

- Walks on the step set can move forwards and backwards in each direction:

$$\text{For all } j = 1, \ldots, d \text{ there exists } \mathbf{i} \in \mathcal{S} \text{ with } \mathbf{i}_j = 1$$

$$\text{For all } j = 1, \ldots, d \text{ there exists } \mathbf{i} \in \mathcal{S} \text{ with } \mathbf{i}_j = -1$$

- The step set and weighting are symmetric over every axis, or every axis except for one. If there is an axis of non-symmetry we may assume it corresponds to the final coordinate, so if $S(\mathbf{z}) = \sum_{\mathbf{i}\in\mathcal{S}} w_\mathbf{i}\mathbf{z}^\mathbf{i}$ is the weighted characteristic polynomial of the model we have

$$S(z_1, \dots, z_{j-1}, \overline{z}_j, z_{j+1}, \dots, z_d) = S(\mathbf{z})$$

for all $1 \leq j \leq d-1$.

Definition 4.13 (highly and mostly symmetric models) We call a weighted step set which is symmetric over every axis *highly symmetric* and a weighted step set which is symmetric over all but one axis *mostly symmetric*.

Note that while the step sets under consideration are symmetric, the individual walks in the models do not need to satisfy any symmetry conditions. Because we study only short step models, we may write

$$S(\mathbf{z}) = \overline{z}_d A(\hat{\mathbf{z}}) + Q(\hat{\mathbf{z}}) + z_d B(\hat{\mathbf{z}})$$

for Laurent polynomials A, Q, and B which are symmetric in their variables $\hat{\mathbf{z}} = (z_1, \dots, z_{d-1})$, and $\mathcal{S}$ is highly symmetric if and only if $A(\hat{\mathbf{z}}) = B(\hat{\mathbf{z}})$. The symmetries present in such a model will allow us to generalize the algebraic kernel method to this setting and obtain explicit representations of the generating functions involved.

Definition 4.14 (group of a symmetric orthant model) For each $1 \leq j \leq d-1$ define the map

$$\sigma_j(\mathbf{z}) = (z_1, \dots, z_j, \overline{z}_j, z_{j+1}, \dots, z_d),$$

and define the map

$$\gamma(\mathbf{z}) = \left(z_1, \dots, z_{d-1}, \overline{z}_d \frac{A(\hat{\mathbf{z}})}{B(\hat{\mathbf{z}})}\right).$$

The *group* $\mathcal{G}$ of the lattice path model determined by $\mathcal{S}$ is the group of transformations generated by the σ_j and γ. Because the group generators are now commuting involutions, we may write $\mathcal{G}$ explicitly as

$$\mathcal{G} = \left\{\sigma_1^{j_1} \cdots \sigma_{d-1}^{j_{d-1}} \gamma^{j_d} : j_1, \dots, j_d \in \{0, 1\}\right\}.$$

For $\sigma = \sigma_1^{j_1} \cdots \sigma_{d-1}^{j_{d-1}} \gamma^{j_d} \in \mathcal{G}$ the *sign* of σ is $\operatorname{sgn}(\sigma) = (-1)^{j_1 + \cdots + j_d}$.

Imposing our symmetry condition on $\mathcal{S}$ implies $\mathcal{G}$ is a group of order 2^d which is isomorphic to the direct sum of cyclic groups of order 2. Again we may view any element of $\mathcal{G}$ as a map from $\mathbb{C}[\overline{\mathbf{z}}, \mathbf{z}][[t]]$ to $\mathbb{C}((\mathbf{z}))[[t]]$, and the characteristic polynomial $S(\mathbf{z})$ is fixed under the action of all such elements.

Consider the multivariate generating function

$$W(\mathbf{z}, t) = \sum_{\substack{\mathbf{i}\in\mathbb{N}^d \\ n\geq 0}} f_{\mathbf{i},n} \mathbf{z}^\mathbf{i} t^n$$

where $f_{\mathbf{i},n}$ counts the number of weighted walks of length n using the steps in $\mathcal{S}$ which begin at the origin, end at $\mathbf{i} \in \mathbb{N}^d$, and never leave $\mathbb{N}^d$. Problem 4.3 asks you to prove that $W(\mathbf{z}, t)$ satisfies the functional equation

$$\begin{aligned}(z_1 \cdots z_d)W(\mathbf{z}, t) &= (z_1 \cdots z_d) + t(z_1 \cdots z_d)S(\mathbf{z})W(\mathbf{z}, t)\\ &\quad - t \sum_{V \subset [d]} (-1)^{|V|}(z_1 \cdots z_d)S(\mathbf{z})W(\mathbf{z}, t)\Big|_{z_j=0,\, j\in V}\end{aligned} \tag{4.12}$$

where $[d] = \{1, \ldots, d\}$, generalizing (4.7) from the two-dimensional case. Note that (4.12) relies heavily on $\mathcal{S}$ having short steps, but does not use any of our symmetry requirements.

Since the generators of $\mathcal{G}$ commute, and applying a generator to an element of $\mathcal{G}$ negates its sign, we can explicitly determine the orbit sum

$$\sum_{\sigma \in \mathcal{G}} \operatorname{sgn}(\sigma)\sigma(z_1 \cdots z_d) = (z_1 - \overline{z}_1) \cdots (z_{d-1} - \overline{z}_{d-1}) \left(z_d - \overline{z}_d \frac{A(\hat{\mathbf{z}})}{B(\hat{\mathbf{z}})} \right).$$

This gives $W(\mathbf{z}, t)$ as the non-negative series extraction of an explicit rational function.

Theorem 4.2 *If $\mathcal{S}$ is mostly or highly symmetric then the multivariate generating function $W(\mathbf{z}, t)$ tracking endpoint and length satisfies $W(\mathbf{z}, t) = [\mathbf{z}^{\geq 0}]R(\mathbf{z}, t)$, where*

$$R(\mathbf{z}, t) = \frac{(z_1 - \overline{z}_1) \cdots (z_{d-1} - \overline{z}_{d-1}) \left(z_d - \overline{z}_d \frac{A(\hat{\mathbf{z}})}{B(\hat{\mathbf{z}})} \right)}{(z_1 \cdots z_d)(1 - tS(\mathbf{z}))}. \tag{4.13}$$

In the highly symmetric case, $A(\hat{\mathbf{z}}) = B(\hat{\mathbf{z}})$.

Remark 4.4 The non-negative series extraction $[\mathbf{z}^{\geq 0}]R(\mathbf{z}, t)$ works with the expansion of $R(\mathbf{z}, t)$ in $\mathcal{R} = \mathbb{Q}((\mathbf{z}))[[t]]$. The order of the variables in the iterated Laurent series ring is important: if z_d were not the last variable then (4.13) would not necessarily hold in the mostly symmetric case. In the highly symmetric case $R(\mathbf{z}, t)$ has an expansion in $\mathbb{Q}[\overline{\mathbf{z}}, \mathbf{z}][[t]]$ and the order of the variables in the extraction is unimportant.

Proof Equation (4.12) implies

$$(1 - tS(\mathbf{z}))(z_1 \cdots z_d)W(\mathbf{z}, t) = (z_1 \cdots z_d) + \sum_{k=1}^{d} L_k(\mathbf{z}_{\widehat{k}}, t), \tag{4.14}$$

for some $L_k(\mathbf{z}_{\widehat{k}}, t) \in \mathbb{Q}[\mathbf{z}_{\widehat{k}}][[t]]$. Consider $T(\mathbf{z}, t) = \sigma(z_1 \cdots z_d)W(\sigma(\mathbf{z}), t)$ where $\sigma = \sigma_1^{j_1} \cdots \sigma_{d-1}^{j_{d-1}} \gamma^{j_d} \in \mathcal{G}$. When $j_d = 1$ then, because of the ordering of the variables, every term in the expansion of $T(\mathbf{z}, t)$ in $\mathcal{R} = \mathbb{Q}((\mathbf{z}))[[t]]$ will have a negative power of z_d. On the other hand, if $j_d = 0$ and there exists $k \in \{1, \ldots, d-1\}$ such that $j_k = 1$ then every term in the expansion of $T(\mathbf{z}, t)$ in $\mathcal{R}$ will have a negative power of z_k. Thus, $[\mathbf{z}^{\geq 0}]T(\mathbf{z}, t) = 0$ unless $\sigma \in \mathcal{G}$ is the identity element, and

$$[\mathbf{z}^{\geq 0}] \sum_{\sigma \in \mathcal{G}} \operatorname{sgn}(\sigma)\sigma(z_1 \cdots z_d) W(\sigma(\mathbf{z}), t) = \sum_{\sigma \in \mathcal{G}} \operatorname{sgn}(\sigma)[\mathbf{z}^{\geq 0}] \left[\sigma(z_1 \cdots z_d) W(\sigma(\mathbf{z}), t)\right] = (z_1 \cdots z_d) W(\mathbf{z}, t).$$

Since $\mathcal{G}$ fixes $S(\mathbf{z})$, and $(z_1 \cdots z_d)W(\mathbf{z}, t)$ contains only positive powers of the $\mathbf{z}$ variables, to prove Theorem 4.2 it is sufficient to show

$$\sum_{\sigma \in \mathcal{G}} \operatorname{sgn}(\sigma) \left(\sigma \cdot L_k(\mathbf{z}_{\widehat{k}}, t)\right) = 0$$

for each $1 \leq k \leq d$. Fix k and define the sets

$$\mathcal{G}_0 = \left\{\sigma_1^{j_1} \cdots \sigma_{d-1}^{j_{d-1}} \gamma^{j_d} : i_1, \ldots, j_d \in \{0, 1\}, j_k = 0\right\}$$
$$\mathcal{G}_1 = \left\{\sigma_1^{j_1} \cdots \sigma_{d-1}^{j_{d-1}} \gamma^{j_d} : i_1, \ldots, j_d \in \{0, 1\}, j_k = 1\right\}.$$

Because $L_k(\mathbf{z}_{\widehat{k}}, t)$ is independent of z_k, we see $(\sigma_k \sigma) \cdot L_k(\mathbf{z}_{\widehat{k}}, t) = \sigma \cdot L_k(\mathbf{z}_{\widehat{k}}, t)$ for all $\sigma \in \mathcal{G}$. Thus,

$$\begin{aligned}\sum_{\sigma \in \mathcal{G}} \operatorname{sgn}(\sigma) \left(\sigma \cdot L_k(\mathbf{z}_{\widehat{k}}, t)\right) &= \sum_{\sigma \in \mathcal{G}_0} \operatorname{sgn}(\sigma) \left(\sigma \cdot L_k(\mathbf{z}_{\widehat{k}}, t)\right) + \sum_{\sigma \in \mathcal{G}_1} \operatorname{sgn}(\sigma) \left(\sigma \cdot L_k(\mathbf{z}_{\widehat{k}}, t)\right) \\ &= \sum_{\sigma \in \mathcal{G}_0} \left[\operatorname{sgn}(\sigma) + \operatorname{sgn}(\sigma_k \sigma)\right] \left(\sigma \cdot L_k(\mathbf{z}_{\widehat{k}}, t)\right) \\ &= 0,\end{aligned}$$

since $\operatorname{sgn}(\sigma_k g) = -\operatorname{sgn}(g)$ for any $g \in \mathcal{G}$. □

We would like to apply Proposition 3.14 from Chapter 3 to (4.13) in order to obtain a diagonal expression for the generating function $W(\mathbf{1}, t)$ counting the number of walks in the orthant $\mathbb{N}^d$, together with diagonal expressions for walks returning to some or all of the boundary axes. In the highly symmetric case $R(\mathbf{z}, t)$ has an expansion in $\mathbb{Q}[\overline{\mathbf{z}}, \mathbf{z}][[t]]$, and this is valid. In the mostly symmetric case, however, the presence of the $B(\hat{\mathbf{z}})$ term in the denominator means one must expand in the larger ring $\mathcal{R} = \mathbb{Q}((\mathbf{z}))[[t]]$ and Proposition 3.14 does not directly apply. Unfortunately, this requires a more involved formal power series argument.

Proposition 4.8 *Let $\mathcal{S}$ be a mostly or highly symmetric weighted step set. Then, expanding in $\mathcal{R}$, the generating function counting the number of walks of a given length in the lattice path model defined by $\mathcal{S}$ satisfies*

$$W(\mathbf{1}, t) = \Delta\left(\frac{(1 + z_1) \cdots (1 + z_{d-1}) \left(B(\hat{\mathbf{z}}) - z_d^2 A(\hat{\mathbf{z}})\right)}{(1 - z_d) B(\hat{\mathbf{z}})(1 - t z_1 \cdots z_d \overline{S}(\mathbf{z}))}\right),$$

where $\overline{S}(\mathbf{z}) = S(z_1, \ldots, z_{d-1}, \overline{z}_d)$. When $\mathcal{S}$ is highly symmetric then $A(\hat{\mathbf{z}}) = B(\hat{\mathbf{z}})$ and $\overline{S}(\mathbf{z}) = S(\mathbf{z})$, so

$$W(\mathbf{1},t) = \Delta\left(\frac{(1+z_1)\cdots(1+z_d)}{1-tz_1\cdots z_d S(\mathbf{z})}\right).$$

Proof Let

$$\begin{aligned} T(\mathbf{z},t) &= \frac{(1+z_1)\cdots(1+z_{d-1})\left(B(\hat{\mathbf{z}}) - z_d^2 A(\hat{\mathbf{z}})\right)}{(1-z_d)B(\hat{\mathbf{z}})(1-tz_1\cdots z_d\overline{S}(\mathbf{z}))} \\ &= (-1)^{d-1}\frac{(z_1\cdots z_{d-1})^2 R(z_1,\dots,z_{d-1},\overline{z}_d,z_1\cdots z_d t)}{(1-z_1)\cdots(1-z_d)}, \end{aligned}$$

where $R(\mathbf{z},t)$ is the rational function in (4.13). Since $[t^n]R(\mathbf{z},t)$ is a Laurent polynomial in z_d for all $n \in \mathbb{N}$, we are formally justified in substituting $z_d = \overline{z}_d$ into the expansion[2]

$$R(\mathbf{z},t) = \sum_{\mathbf{i}\in\mathbb{Z}^d, n\geq 0} r_{\mathbf{i},n}\mathbf{z}^{\mathbf{i}}t^n$$

of $R(\mathbf{z},t)$ in the ring $\mathcal{R} = \mathbb{Q}((\mathbf{z}))[[t]]$, obtaining

$$T(\mathbf{z},t) = (-1)^{d-1}\sum_{\substack{\mathbf{i}\in\mathbb{Z}^d,\mathbf{j}\in\mathbb{N}^d\\ n\geq 0}} r_{\mathbf{i},n} z_1^{i_1+j_1+n+2}\cdots z_{d-1}^{i_{d-1}+j_{d-1}+n+2} z_d^{-i_d+j_d+n} t^n.$$

Taking the nth diagonal term of this series gives

$$[t^n]\Delta T(\mathbf{z},t) = (-1)^{d-1}\sum_{\mathbf{j}\in\mathbb{N}^d} r_{-j_1-2,\dots,-j_{d-1}-2,j_d,n}$$

and we must show this final sum equals $\sum_{\mathbf{i}\in\mathbb{N}^d} r_{\mathbf{i},n}$ for all $n \geq 0$. Expanding (4.13) as a power series in t implies that for any $n \geq 0$

$$[t^n]R(\mathbf{z},t) = \frac{(z_1-\overline{z}_1)\cdots(z_{d-1}-\overline{z}_{d-1})}{z_1\cdots z_{d-1}} \times \frac{(\overline{z}_d A(\hat{\mathbf{z}}) + Q(\hat{\mathbf{z}}) + z_d B(\hat{\mathbf{z}}))^n}{z_d B(\hat{\mathbf{z}})},$$

so $P_n(\hat{\mathbf{z}}) = [z_d^{\geq 0}][t^n]R(\mathbf{z},t)$ is a Laurent polynomial in $\mathbf{z}$ (any term with a non-negative exponent of z_d is a monomial times non-negative powers of A, B, and Q). In particular, we are formally justified in substituting $z_j = \overline{z}_j$ for $1 \leq j \leq d-1$ into the expansion of $P_n(\hat{\mathbf{z}})$ in $\mathcal{R}$. Since A, B, and Q are unchanged when z_j is replaced by $\overline{z}_j$,

$$\begin{aligned} P_n(\overline{z}_1,\dots,\overline{z}_{d-1}) &= \frac{(\overline{z}_1-z_1)\cdots(\overline{z}_{d-1}-z_{d-1})}{\overline{z}_1\cdots\overline{z}_{d-1}} \times [z_d^{\geq 0}]\frac{(\overline{z}_d A(\hat{\mathbf{z}}) + Q(\hat{\mathbf{z}}) + z_d B(\hat{\mathbf{z}}))^n}{z_d B(\hat{\mathbf{z}})} \\ &= (-1)^{d-1}(z_1\cdots z_{d-1})^2 P_n(\hat{\mathbf{z}}), \end{aligned}$$

[2] Because this expansion of $R(\mathbf{z},t)$ lies in $\mathcal{R}$, many of the coefficients $r_{\mathbf{i},n}$ are zero.

and taking the coefficient of $\mathbf{z}^{\mathbf{i}}$ implies $r_{\mathbf{i},n} = (-1)^{d-1} r_{-i_1-2,\dots,i_{d-1}-2,i_d,n}$ for all $\mathbf{i} \in \mathbb{N}^d$, proving the desired equality for $[t^n]\Delta T(\mathbf{z},t)$. □

In the highly symmetric case, Proposition 4.8 gives a beautiful expression for the generating function as the diagonal of the power series expansion of a simple and explicit multivariate rational function. In the mostly symmetric case, however, the factor of $B(\hat{\mathbf{z}})$ in the denominator means that the resulting rational function may not be analytic at the origin. Although we can work directly with diagonals of Laurent series expansions using the tools from Chapter 3, an easier solution is to use the following alternative encoding.

Proposition 4.9 *Let $\mathcal{S}$ be a mostly or highly symmetric weighted step set. Then the generating function counting the number of walks of a given length in the lattice path model defined by $\mathcal{S}$ satisfies $W(\mathbf{1},t) = \Delta\left(\frac{G(\mathbf{z},t)}{H(\mathbf{z},t)}\right)$, where*

$$G(\mathbf{z},t) = (1+z_1)\cdots(1+z_{d-1})\,(1 - tz_1\cdots z_d\,(Q(\hat{\mathbf{z}}) + 2z_d A(\hat{\mathbf{z}})))$$
$$H(\mathbf{z},t) = (1-z_d)\left(1 - tz_1\cdots z_d\overline{S}(\mathbf{z})\right)(1 - tz_1\cdots z_d\,(Q(\hat{\mathbf{z}}) + z_d A(\hat{\mathbf{z}})))\,.$$

Proof Expanding $R(\mathbf{z},t)$ defined in (4.13) gives

$$(1-\overline{z}_1^2)\cdots(1-\overline{z}_{d-1}^2)\left(1-\overline{z}_d^2\frac{A(\hat{\mathbf{z}})}{B(\hat{\mathbf{z}})}\right)\sum_{n\ge 0} t^n\left(\overline{z}_d A(\hat{\mathbf{z}}) + Q(\hat{\mathbf{z}}) + z_d B(\hat{\mathbf{z}})\right)^n .$$

Since the series

$$(1-\overline{z}_1^2)\cdots(1-\overline{z}_{d-1}^2)\left(\overline{z}_d^2\frac{A(\hat{\mathbf{z}})}{B(\hat{\mathbf{z}})}\right)\sum_{n\ge 0} t^n\left(\overline{z}_d A(\hat{\mathbf{z}}) + Q(\hat{\mathbf{z}})\right)^n$$

contains no positive powers of z_d, we can add it to $R(\mathbf{z},t)$ and obtain $W(\mathbf{z},t)$ as the non-negative series extraction of

$$\frac{(1-\overline{z}_1^2)\cdots(1-\overline{z}_{d-1}^2)\left(1-\overline{z}_d^2 A(\hat{\mathbf{z}})/B(\hat{\mathbf{z}})\right)}{1-tS(\mathbf{z})} + \frac{(1-\overline{z}_1^2)\cdots(1-\overline{z}_{d-1}^2)\left(\overline{z}_d^2 A(\hat{\mathbf{z}})/B(\hat{\mathbf{z}})\right)}{1-t\left(\overline{z}_d A(\hat{\mathbf{z}}) + Q(\hat{\mathbf{z}})\right)}.$$

This simplifies to

$$\frac{(1-\overline{z}_1^2)\cdots(1-\overline{z}_{d-1}^2)\,(1-t\,(2\overline{z}_d A(\hat{\mathbf{z}}) + Q(\hat{\mathbf{z}})))}{(1-t\,(\overline{z}_d A(\hat{\mathbf{z}}) + Q(\hat{\mathbf{z}}) + z_d B(\hat{\mathbf{z}})))\,(1-t\,(\overline{z}_d A(\hat{\mathbf{z}}) + Q(\hat{\mathbf{z}})))},$$

which has an expansion in $\mathbb{Q}[\overline{\mathbf{z}},\mathbf{z}][[t]]$. Proposition 3.14 of Chapter 3 now applies, giving the stated result. □

Not only does the rational function $F(\mathbf{z},t) = G(\mathbf{z},t)/H(\mathbf{z},t)$ in Proposition 4.9 always have an expansion in $\mathbb{Q}[\mathbf{z}][[t]]$, the coefficients of this power series expansion are all non-negative, which will greatly simplify arguments about the singularities of $F(\mathbf{z},t)$. Note that Propositions 4.8 and 4.9 can give different rational diagonal expressions, even in the highly symmetric case (of course, both multivariate rational functions will have the same diagonal). In the highly symmetric case we always use the expression in Propositions 4.8, which is simpler as it makes full use of the symmetries of a model.

Example 4.8 (Two Diagonal Expressions for a Walk Generating Function)

Proposition 4.8 implies the generating function of the quarter-plane model on the steps $\mathcal{S} = \{(\pm 1, 0), (0, \pm 1)\}$ is the diagonal of the rational function

$$\frac{(1+x)(1+y)}{1-t(x^2y+y+y^2x+x)},$$

while Proposition 4.9 implies it is the diagonal of

$$\frac{(1+x)\left(1-ty(1+2xy+x^2)\right)}{(1-y)\left(1-t(x^2y+y+y^2x+x)\right)\left(1-ty(1+xy+x^2)\right)}.$$

Problem 4.5 asks you to prove these two rational functions have equal diagonals.

Propositions 4.8 and 4.9 imply that the generating function $W(\mathbf{1}, t)$ counting walks in a highly or mostly symmetric lattice path model is D-finite. We conclude this subsection by noting that in any dimension $d \geq 2$ there is a model which is symmetric over all but two axes and has non-D-finite generating function. Thus, one cannot obtain rational diagonal expressions for generating functions of all short step models symmetric over $r < d-1$ axes.

Proposition 4.10 *For any dimension $d \geq 2$ there is a step set $\mathcal{S}_d \subset \{\pm 1, 0\}^d$ that is symmetric over all but two axes, such that the generating function counting the number of walks in $\mathbb{N}^2$ starting at the origin and using the steps in $\mathcal{S}_d$ is non-D-finite.*

Proof We give an asymptotic proof. We have shown above that the number of excursions of the two dimensional lattice path model in $\mathbb{N}^2$ defined by the step set $\mathcal{S}_2 = \{(-1,-1), (0,-1), (0,1), (1,0), (-1,0)\}$ has asymptotic growth of the form $e_n \sim C\,\rho^n\,n^\alpha$ for $\alpha \notin \mathbb{Q}$. Probabilistic work of Duraj [34] then implies that the number of walks on $\mathcal{S}_2$ restricted to $\mathbb{N}^2$ and ending anywhere has asymptotic growth of the form $f_n \sim C_2\,\rho^n\,n^\alpha$, so that the generating function counting the number of walks on $\mathcal{S}$ in $\mathbb{N}^2$ is non-D-finite by Corollary 2.2 of Chapter 2.

For $d \geq 3$ let $\mathcal{S}_d = \mathcal{S}_2 \times \{\pm 1\}^{d-1}$. A walk of length n on the steps of $\mathcal{S}_d$ is determined by a walk of length n on the steps $\mathcal{S}_2$ in $\mathbb{N}^2$ and $d-2$ independent walks of length n on the steps $\{\pm 1\}$ on $\mathbb{N}$; i.e., $d-2$ independent Dyck prefixes. Since, as calculated in an example above, the number of Dyck prefixes of length n has dominant asymptotics $d_n \sim (2/\pi)^{1/2} 2^n n^{-1/2}$ the number of walks on $\mathcal{S}_d$ restricted to $\mathbb{N}^d$ has dominant asymptotics of the form $f_n^{(d)} \sim C_d \rho_d^n n^{\alpha_d}$ where $\alpha_d = \alpha - (d-2)/2 \notin \mathbb{Q}$. Thus, the generating function of walks on $\mathcal{S}_d$ in $\mathbb{N}^d$ has non-D-finite generating function by Corollary 2.2 of Chapter 2. □

We have now characterized the generating functions of several large families of lattice path models. We return to these models in Chapters 6 and 10, where we determine asymptotics using analytic combinatorics in several variables.

4.2 Historical Perspective

We end by surveying some of the recent history around the methods and results discussed above.

4.2.1 The Kernel Method

The kernel method has a long and delicate history of which we give only a broad overview. The development can be roughly broken down into two tracks, coming from the probabilistic and combinatorial literature. The queuing theory works of Malyshev [69, 68] in the early 1970s are a good introduction to the probabilistic approach: inspired by the Wiener-Hopf method for solving integral equations—which itself dates back to the 1930s and shares similarities with the kernel method—Malyshev sets up a bivariate kernel equation and examines the solutions of the kernel in one variable. As is typical in this approach, the algebraic solutions of the kernel are analytically continued to a Riemann surface and then studied; a characterization of the solutions is obtained by solving boundary value problems. The survey of Malyshev [70] gives detailed references to early literature in this area. The kernel method also appears[3] around this time in independent work of Kingman [61]. Early applications of the probabilistic kernel method include work by Flatto and collaborators [46, 45, 47] and Fayolle and Iasnogorodski [36], with the latter paper being a clear early exposition of the method. The probabilistic approach was thoroughly detailed and popularized into the wider mathematical community by the textbook of Fayolle et al. [37]. In particular, that text popularized the notion of the group of a walk, which would become central to the algebraic kernel method.

From a combinatorial perspective, the classical kernel method—setting up a functional equation for a bivariate function and substituting a solution of the kernel for one of its variables—is often attributed to the 1968 textbook of Knuth [62], which includes two exercises [62, Ex. 4 and 11, Sect. 2.2.1] and solutions essentially solving the ballot problem with the kernel method[4]. This approach has parallels to work by Brown and Tutte on what came to be known as the 'quadratic method' for solving functional equations arising in planar map enumeration [27, 28], and in statistical mechanical work following from the so-called 'Bethe ansatz' (see, for

[3] Thanks to Alin Bostan for pointing out this reference.

[4] Knuth [62, p. 537] writes "We present here a new method for solving the ballot problem with the use of double generating functions, since this method lends itself to the solution of more difficult problems. . . ".

instance, Slavnov [82]). Such generating function manipulations are common in the enumerative community, and similar techniques appear in a variety of works [30, 29, 18, 67, 80] from the 1970s to 90s before the method was fully distilled. The name 'kernel method' came into use[5] in the late 1990s or early 2000s in the combinatorics community [2, 25], around the same time a full study of the power and flexibility of the method was undertaken. The seminal works on the classical combinatorial kernel method in the early 2000s include Bousquet-Mélou and Petkovšek [25], Banderier et al. [4], and Banderier and Flajolet [5]; work of Prodinger [81] around this time surveys some of the combinatorial applications. As mentioned above, the 'algebraic' kernel method variant was developed by Bousquet-Mélou [21] to deal with kernel equations by examining transformations fixing the kernel instead of solving the kernel. Bousquet-Mélou [19, 21] also developed the 'obstinate' kernel method, where one looks both for roots of the kernel and transformations fixing the kernel. The 'iterated' kernel method, used to show non-D-finiteness of certain generating functions of walk models with infinite group, was introduced by van Rensburg et al. [86]. Bousquet-Mélou and Jehanne [23] gave a generalization solving polynomial functional equations in two variables, and Bostan et al. [12] extended the algebraic kernel method from short step models to those with longer steps.

4.2.2 Recent History of Lattice Paths in Orthants

Through various forms of the ballot problem, lattice path models restricted to a half-space have been studied for centuries. In modern times, Gessel [49] gave general algebraic expressions for half-space models, with a full and complete accounting of the asymptotic behaviour of walks confined to a half-space given by Banderier and Flajolet [5]. Bostan et al. [14] discuss the complexity of enumerating walks and excursions which are either unrestricted or restricted to a half-plane.

Kreweras [65] gave an early study of a lattice path model in the quarter-plane, later revisited in the work of Bousquet-Mélou [21] which introduced the algebraic kernel method; see also Gessel [50]. The probabilistic study of short step models in the quarter-plane followed the development of the probabilistic kernel method discussed above, captured in the text of Fayolle et al. [37]. The lattice path models we consider have a counting sequence whose dominant asymptotics is a finite sum of terms of the form $Cn^\alpha \rho^n$ for constants C, α, and ρ. Using the probabilistic kernel method, Fayolle and Raschel [39] outlined a method which in principle allows one to determine the exponential growth rate ρ for the non-singular models, and determined ρ in many cases; see also the second edition of Fayolle et al. [38, Ch. 11]. The systematic combinatorial study of walks in a quadrant was popularized by Bousquet-Mélou and

[5] The first published use of the term 'kernel method' ("méthode du noyau" in French) occurs in Banderier [2]; this name apparently arose verbally in the research group of Phillipe Flajolet around 1998, who mentioned at the time that the name was in use by probabilists whose identities now seem lost to time. The author thanks Cyril Banderier for recounting some of this history; see also the introduction of Banderier and Wallner [3] for some historical remarks.

Mishna [24], following work of Petkovšek [80], Bousquet-Mélou and Petkovšek [26], and Mishna [75]. Bousquet-Mélou and Mishna proved that the generating functions of 22 of the 23 short step models with finite group in the quarter-plane are D-finite, leaving open only the D-finiteness of Gessel's model.

Using the kernel method and creative telescoping methods, Kauers et al. [58] gave a computer algebra proof of a long standing open question on the number of excursions for Gessel's model in the non-negative quadrant, following an approach discussed by Kauers and Zeilberger [59]; see also Kauers and Zeilberger [60] and Koutschan et al. [63] for related problems. Bostan and Kauers [16] proved that the multivariate generating function $Q(x, y, t)$ tracking endpoint and length for Gessel's model is algebraic by guessing an algebraic equation satisfied by $Q(x, y, t)$ and rigorously proving the equation using algebraic tools; Bostan et al. [9] and Bousquet-Mélou [22] later gave alternate proofs. Similarly, by guessing algebraic and D-finite equations satisfied by the generating functions of each of the 79 non-isomorphic models in the quarter-plane, Bostan and Kauers [15] conjectured[6] the asymptotics for the 23 short step models in the quarter-plane with D-finite generating function shown in Table 4.1. The guessed D-finite equations for these generating functions were later proven by Bostan et al. [13], which also expressed the generating functions in terms of explicit hypergeometric functions. These explicit expressions are not enough to determine dominant asymptotics for every model, due to issues related to the connection problem for D-finite functions. Bostan et al. [17] proved that the generating function counting excursions for the 51 non-singular short step models in the quarter-plane with infinite group are non-D-finite, and a study of the 5 short step singular models carried out by Mishna and Rechnitzer [76] and Melczer and Mishna [71] implies that their generating functions counting walks ending anywhere are non-D-finite.

An extremely fruitful approach to lattice path enumeration is to weight the steps of a model by positive real numbers summing to 1 and then interpret the number of walks of length n ending at a given point as a probability distribution on the set of walks of length n. Most immediately, one can try to use *local* or *central limit theorems* to estimate these probabilities and derive enumerative results. Extending these limit theorems, lattice path models in $\mathbb{Z}^d$ with probabilistically weighted steps can often be rescaled into continuous objects. The most well known example is that the famous *Brownian motion* process can be obtained [57] as a scaling limit of the simple random walk on $\mathbb{Z}$ with steps $\{-1, +1\}$. In higher dimensions, and with varying step sets, walks on $\mathbb{Z}^d$ restricted to various cones can be approximated by multi-dimensional Brownian motion with certain constraints. As a Brownian motion satisfies partial differential equations instead of the partial *discrete* differential equations described by the kernel equation, one can more easily apply the tools of analysis to determine

[6] After determining an algebraic or differential equation for a truncated series, several techniques can be used to give confidence that the generating function under consideration satisfies this equation. In addition to simply computing additional terms of the generating function and verifying that the additional terms also satisfy the equation, Bostan and Kauers [15, Sec. 2.4] give several algebraic and analytic heuristics.

asymptotics[7]. The enumeration of a lattice path model can then be approached by approximating walks by Brownian motions whose asymptotics are well understood; see [35, 83] for classical accounts of the study of random walks. Using this approach, for lattice path models on a large variety of step sets in a wide collection of cones, including orthants in $\mathbb{R}^d$,

- Denisov and Wachtel [32] found the exponential growth ρ and critical exponent α in the dominant asymptotics $Cn^\alpha \rho^n$ for the number of excursions of a model and, when the step set of the model has vector sum zero, for the number of walks ending anywhere in the restricting cone;
- Garbit and Raschel [48] found the exponential growth ρ for the number of walks ending anywhere in the restricting cone;
- Duraj [34] gave a method to determine the exponential growth ρ and critical exponent α for the number of walks ending anywhere in the restricting cone when the step set of the model has vector sum pointing outside the cone.

These results are very general, allowing for enumerative information on the non-singular models studied above, including those with non-D-finite generating function. We note that the process of approximating a discrete walk by Brownian motion makes it extremely difficult to exactly determine the asymptotic constant C and a good bound on the asymptotic error term.

Melczer and Wilson [73] combined the diagonal expressions derived above with the methods of analytic combinatorics in several variables to determine dominant asymptotics, including leading constants, for the quadrant models with transcendental D-finite generating function. Bernardi et al. [7] and Dreyfus et al. [33] examine the behaviour of short step quadrant models with infinite group, proving some of the models admit generating functions which are not only non-D-finite but also hypertranscendental. Bostan et al. [11] examine three-dimensional short step models in the non-negative octant, and Bostan et al. [12] generalize the kernel method to study walks in an orthant with potentially non-short steps; in both cases rational diagonal representations for the transcendental but D-finite generating functions which arise are obtained. Melczer and Mishna [72] and Melczer and Wilson [74] use the techniques of analytic combinatorics in several variables to find asymptotics for highly and mostly symmetric walks in any dimension. Gessel and Zeilberger [51] gave representations for lattice path generating functions in so-called Weyl chambers in arbitrary dimension, which include the diagonal representations of Melczer and Mishna for highly symmetric models in $\mathbb{N}^d$. Tate and Zelditch [84] and Feierl [40, 41] determined asymptotics of walks in Weyl chambers using analytic techniques such as saddle-point integral computations. Zeilberger [89] and Grabiner and Magyar [52] contain other work on generating function expressions for walks in Weyl chambers.

[7] See Bousquet-Mélou [22, Sect. 2] for a nice discussion on partial discrete differential equations.

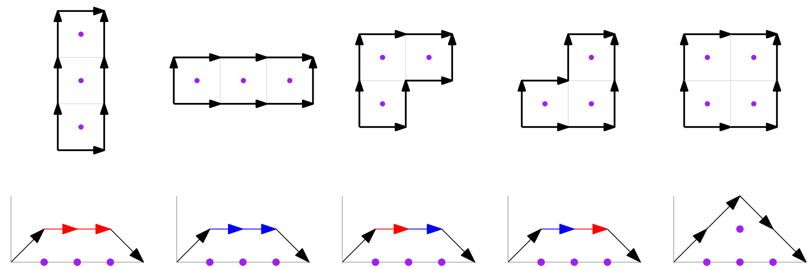


Fig. 4.6 *Top:* The path pairs in $\mathcal{S}_4$: there are four pairs of area three and one pair of area four. *Bottom:* The lattice walks in $\mathfrak{S}_4$: there are four walks of area three and one walk of area four.

Problems

4.1 Generalize Proposition 4.1 to give the generating function of walks returning to the origin when the starting point $p \in \mathbb{N}$ is arbitrary. Carefully consider situations when the integrand in (4.4) may have a pole at $x = 0$.

4.2 Following the argument in Section 4.1.4, prove there are exactly 79 non-isomorphic quarter-plane models which are not isomorphic to half-plane models.

4.3 Prove that the multivariate generating function $W(\mathbf{z}, t)$ tracking the endpoint and length of a lattice walks in $\mathbb{N}^d$ taking short steps satisfies the functional equation (4.12). *Hint:* Decompose a walk of length n as a walk of length $n - 1$ followed by a single step, and use the principle of inclusion-exclusion to keep track of walks leaving the orthant.

4.4 Use Proposition 4.8 and a computer algebra package for creative telescoping to prove that the generating function $A(z)$ for the number of quarter-plane walks on the steps $\mathcal{S} = \{(\pm 1, 0), (0, \pm 1)\}$ satisfies the differential equation

$$z^2(4z-1)(4z+1)A'''(z) + 2z(4z+1)(16z-3)A''(t) \\ + 2(112z^2 + 14z - 3)A'(t) + 4(16z+3)A(t) = 0.$$

4.5 Prove the two rational diagonal expressions for the number of quarter-plane walks on the steps $\mathcal{S} = \{(\pm 1, 0), (0, \pm 1)\}$ given by Propositions 4.8 and 4.9 are equivalent by using a computer algebra package to determine a D-finite equation satisfied by the diagonal of their difference.

4.6 A *path pair* of length n is a pair of paths starting at the origin, consisting of n unit steps to the north or east, meeting again for the first time after n steps; the path pairs of length 4 are shown in Figure 4.6. Let $\mathcal{S}_n$ denote all pairs of paths of length n. An open problem several decades old [88] asks whether the elements of $\mathcal{S}_n$ tile a $2^{n-2} \times 2^{n-2}$ chessboard without overlap (the elements of $\mathcal{S}_n$ are allowed to be rotated). This problem asks you to use the kernel method to show that the elements of $\mathcal{S}_n$ covers the correct number of squares, 4^{n-2}, for the $2^{n-2} \times 2^{n-2}$ chessboard.

1. Let $\mathfrak{S}_n$ denote the class of lattice walks of length n in $\mathbb{Z}$ which start at the origin, take the steps $\mathcal{T} = \{(1,1),(1,-1),(1,0),(1,0)\}$ (there are two different horizontal steps), end at height 0, and touch the x-axis only at their beginning and end. Find a bijection between $\mathcal{S}_n$ and $\mathfrak{S}_n$, such that $\mathbf{p} \in \mathcal{S}_n$ is paired to an element $\mathbf{w} \in \mathfrak{S}_n$ so that the area of $\mathbf{p}$ (number of squares covered) corresponds to area under the walk $\mathbf{w}$ (number of integer points under the walk and above or on the x-axis); see Figure 4.6.

2. Let G be the trivariate generating function $G(y,u,t) = \sum_{j,k,n\geq 0} g_{jkn} y^j u^k t^n$, where g_{jkn} denotes the number of walks on n steps of $\mathcal{T}$ which end at height $y = j$, have area k, and stay above $y = 0$ except at their starting and (potentially) ending point. Note that $[t^n]G(0,u,t)$ is the generating function of the walks in $\mathfrak{S}_n$ by area. Prove that $G(0,u,t) = ut^2 F(0,u,ut)$, where $F(y,u,t)$ satisfies the kernel-like equation

$$F(y,u,t) = 1 + t\left(yu + 2 + \frac{1}{yu}\right)F(yu,u,t) - \frac{t}{yu}F(0,u,t).$$

Hint: G counts only walks which touch the x-axis in their first and (potentially) last steps. F is the generating function of a less restricted class.

3. Using this kernel equation, prove that $F(0,1,t) = \frac{2}{1-2t+\sqrt{1-4t}}$ and thus show that $F(y,1,t) = \frac{2}{1-2t+\sqrt{1-4t}-2ty}$. Furthermore, by differentiating the kernel equation prove that $F_u(0,1,t) = \frac{1-4t+2t^2+(2t-1)\sqrt{1-4t}}{2(1-4t)t^2}$.

4. Using $F(0,1,t)$ and $F_u(0,1,t)$ show that the generating function for the number of boxes filled by all walks of length n is $t^2/(1-4t)$, proving the elements of $\mathcal{S}_n$ cover the right area to tile a $2^{n-2} \times 2^{n-2}$ chessboard.
Hint: What does the partial derivative of G with respect to u represent?

References

1. Andrew Arnold and Michael Monagan. Calculating cyclotomic polynomials. *Math. Comp.*, 80(276):2359–2379, 2011.
2. C. Banderier. *Combinatoire analytique: application aux marches aleéatoires*. Meémoire de DEA, INRIA Rocquencourt, 1998.
3. C. Banderier and M. Wallner. The kernel method for lattice paths below a line of rational slope. In G. E. Andrews, C. Krattenthaler, and A. Krinik, editors, *Lattice Path Combinatorics and Applications*, Developments in Mathematics. Springer International Publishing, 2019.
4. Cyril Banderier, Mireille Bousquet-Mélou, Alain Denise, Philippe Flajolet, Danièle Gardy, and Dominique Gouyou-Beauchamps. Generating functions for generating trees. *Discrete Math.*, 246(1-3):29–55, 2002. Formal power series and algebraic combinatorics (Barcelona, 1999).
5. Cyril Banderier and Philippe Flajolet. Basic analytic combinatorics of directed lattice paths. *Theoret. Comput. Sci.*, 281(1-2):37–80, 2002. Selected papers in honour of Maurice Nivat.

6. O. Bernardi, N. Holden, and X. Sun. Percolation on triangulations: a bijective path to liouville quantum gravity. *arXiv preprint, arXiv:1807.01684*, 2018.
7. Olivier Bernardi, Mireille Bousquet-Mélou, and Kilian Raschel. Counting quadrant walks via Tutte's invariant method. In *Proceedings of FPSAC 2016 (28th International Conference on Formal Power Series and Algebraic Combinatorics), DMTCS proc. BC*, pages 203–214, 2016.
8. W. Böhm. Lattice path counting and the theory of queues. *J. Statist. Plann. Inference*, 140(8):2168–2183, 2010.
9. A. Bostan, I. Kurkova, and K. Raschel. A human proof of Gessel's lattice path conjecture. *Trans. Amer. Math. Soc.*, 369(2):1365–1393, 2017.
10. Alin Bostan. *Computer Algebra for Lattice Path Combinatorics*. Habilitation à diriger des recherches, Université Paris 13, December 2017.
11. Alin Bostan, Mireille Bousquet-Mélou, Manuel Kauers, and Stephen Melczer. On 3-dimensional lattice walks confined to the positive octant. *Annals of Combinatorics*, 20(4):661–704, 2016.
12. Alin Bostan, Mireille Bousquet-Mélou, and Stephen Melczer. Counting walks with large steps in an orthant. *Accepted to Journal of the European Mathematical Society*, 2020.
13. Alin Bostan, Frédéric Chyzak, Mark van Hoeij, Manuel Kauers, and Lucien Pech. Hypergeometric expressions for generating functions of walks with small steps in the quarter plane. *European J. Combin.*, 61:242–275, 2017.
14. Alin Bostan, Louis Dumont, and Bruno Salvy. Algebraic diagonals and walks: algorithms, bounds, complexity. *J. Symbolic Comput.*, 83:68–92, 2017.
15. Alin Bostan and Manuel Kauers. Automatic classification of restricted lattice walks. In *21st International Conference on Formal Power Series and Algebraic Combinatorics (FPSAC 2009)*, Discrete Math. Theor. Comput. Sci. Proc., AK, pages 201–215. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2009.
16. Alin Bostan and Manuel Kauers. The complete generating function for Gessel walks is algebraic. *Proc. Amer. Math. Soc.*, 138(9):3063–3078, 2010. With an appendix by Mark van Hoeij.
17. Alin Bostan, Kilian Raschel, and Bruno Salvy. Non-D-finite excursions in the quarter plane. *J. Combin. Theory Ser. A*, 121:45–63, 2014.
18. Mireille Bousquet-Mélou. A method for the enumeration of various classes of column-convex polygons. *Discrete Math.*, 154(1-3):1–25, 1996.
19. Mireille Bousquet-Mélou. Four classes of pattern-avoiding permutations under one roof: generating trees with two labels. *Electron. J. Combin.*, 9(2):Research paper 19, 31, 2002/03. Permutation patterns (Otago, 2003).
20. Mireille Bousquet-Mélou. Algebraic generating functions in enumerative combinatorics and context-free languages. In *STACS 2005*, volume 3404 of *Lecture Notes in Comput. Sci.*, pages 18–35. Springer, Berlin, 2005.
21. Mireille Bousquet-Mélou. Walks in the quarter plane: Kreweras' algebraic model. *Ann. Appl. Probab.*, 15(2):1451–1491, 2005.
22. Mireille Bousquet-Mélou. An elementary solution of Gessel's walks in the quadrant. *Adv. Math.*, 303:1171–1189, 2016.
23. Mireille Bousquet-Mélou and Arnaud Jehanne. Polynomial equations with one catalytic variable, algebraic series and map enumeration. *J. Combin. Theory Ser. B*, 96(5):623–672, 2006.
24. Mireille Bousquet-Mélou and Marni Mishna. Walks with small steps in the quarter plane. In *Algorithmic probability and combinatorics*, volume 520 of *Contemp. Math.*, pages 1–39. Amer. Math. Soc., Providence, RI, 2010.
25. Mireille Bousquet-Mélou and Marko Petkovšek. Linear recurrences with constant coefficients: the multivariate case. *Discrete Math.*, 225(1-3):51–75, 2000. Formal power series and algebraic combinatorics (Toronto, ON, 1998).
26. Mireille Bousquet-Mélou and Marko Petkovšek. Walks confined in a quadrant are not always D-finite. *Theoret. Comput. Sci.*, 307(2):257–276, 2003. Random generation of combinatorial objects and bijective combinatorics.

27. W. G. Brown and W. T. Tutte. On the enumeration of rooted non-separable planar maps. *Canad. J. Math.*, 16:572–577, 1964.
28. William G. Brown. On the existence of square roots in certain rings of power series. *Math. Ann.*, 158:82–89, 1965.
29. Fan Chung, Ron Graham, John Morrison, and Andrew Odlyzko. Pebbling a chessboard. *Amer. Math. Monthly*, 102(2):113–123, 1995.
30. Robert Cori and Jean Richard. Énumération des graphes planaires à l'aide des séries formelles en variables non commutatives. *Discrete Math.*, 2:115–162, 1972.
31. N. G. de Bruijn, D. E. Knuth, and S. O. Rice. The average height of planted plane trees. In *Graph theory and computing*, pages 15–22. Academic Press, New York, 1972.
32. Denis Denisov and Vitali Wachtel. Random walks in cones. *Ann. Probab.*, 43(3):992–1044, 2015.
33. Thomas Dreyfus, Charlotte Hardouin, Julien Roques, and Michael F. Singer. On the nature of the generating series of walks in the quarter plane. *Invent. Math.*, 213(1):139–203, 2018.
34. Jetlir Duraj. Random walks in cones: the case of nonzero drift. *Stochastic Process. Appl.*, 124(4):1503–1518, 2014.
35. Evgenii B. Dynkin and Aleksandr A. Yushkevich. *Markov processes: Theorems and problems*. Translated from the Russian by James S. Wood. Plenum Press, New York, 1969.
36. Guy Fayolle and Roudolf Iasnogorodski. Two coupled processors: the reduction to a Riemann-Hilbert problem. *Z. Wahrsch. Verw. Gebiete*, 47(3):325–351, 1979.
37. Guy Fayolle, Roudolf Iasnogorodski, and Vadim Malyshev. *Random walks in the quarter-plane*, volume 40 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 1999. Algebraic methods, boundary value problems and applications.
38. Guy Fayolle, Roudolf Iasnogorodski, and Vadim Malyshev. *Random walks in the quarter plane*, volume 40 of *Probability Theory and Stochastic Modelling*. Springer, Cham, second edition, 2017. Algebraic methods, boundary value problems, applications to queueing systems and analytic combinatorics.
39. Guy Fayolle and Kilian Raschel. Some exact asymptotics in the counting of walks in the quarter plane. In *23rd Intern. Meeting on Probabilistic, Combinatorial, and Asymptotic Methods for the Analysis of Algorithms (AofA'12)*, Discrete Math. Theor. Comput. Sci. Proc., AQ, pages 109–124. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2012.
40. Thomas Feierl. Asymptotics for the number of walks in a Weyl chamber of type B. *Random Structures Algorithms*, 45(2):261–305, 2014.
41. Thomas Feierl. Asymptotics for the number of zero drift reflectable walks in a Weyl chamber of type A. *arXiv preprint, arXiv:1806.05998*, 2018.
42. P. Flajolet. Combinatorial aspects of continued fractions. *Discrete Math.*, 32(2):125–161, 1980.
43. P. Flajolet, J. Françon, and J. Vuillemin. Sequence of operations analysis for dynamic data structures. *J. Algorithms*, 1(2):111–141, 1980.
44. Philippe Flajolet. The evolution of two stacks in bounded space and random walks in a triangle. In Jozef Gruska, Branislav Rovan, and Juraj Wiedermann, editors, *Mathematical Foundations of Computer Science 1986*, pages 325–340, Berlin, Heidelberg, 1986. Springer Berlin Heidelberg.
45. L. Flatto and S. Hahn. Two parallel queues created by arrivals with two demands. I. *SIAM J. Appl. Math.*, 44(5):1041–1053, 1984.
46. L. Flatto and H. P. McKean. Two queues in parallel. *Comm. Pure Appl. Math.*, 30(2):255–263, 1977.
47. Leopold Flatto. The longer queue model. *Probability in the Engineering and Informational Sciences*, 3(4):537–559, 1989.
48. Rodolphe Garbit and Kilian Raschel. On the exit time from a cone for random walks with drift. *Rev. Mat. Iberoam.*, 32(2):511–532, 2016.
49. Ira M. Gessel. A factorization for formal Laurent series and lattice path enumeration. *J. Combin. Theory Ser. A*, 28(3):321–337, 1980.
50. Ira M. Gessel. A probabilistic method for lattice path enumeration. *J. Statist. Plann. Inference*, 14(1):49–58, 1986.

51. Ira M. Gessel and Doron Zeilberger. Random walk in a Weyl chamber. *Proc. Amer. Math. Soc.*, 115(1):27–31, 1992.
52. David J. Grabiner and Peter Magyar. Random walks in Weyl chambers and the decomposition of tensor powers. *J. Algebraic Combin.*, 2(3):239–260, 1993.
53. Howard D. Grossman. Fun with lattice points. *Scripta Math.*, 16:207–212, 1950.
54. Katherine Humphreys. A history and a survey of lattice path enumeration. *J. Statist. Plann. Inference*, 140(8):2237–2254, 2010.
55. Veronika Irvine. *Lace Tessellations: A mathematical model for bobbin lace and an exhaustive combinatorial search for patterns*. PhD thesis, University of Victoria, 2016.
56. Veronika Irvine, Stephen Melczer, and Frank Ruskey. Vertically constrained Motzkin-like paths inspired by bobbin lace. *Electron. J. Combin.*, 26(2):Paper 2.16, 2019.
57. Mark Kac. Random walk and the theory of Brownian motion. *Amer. Math. Monthly*, 54:369–391, 1947.
58. Manuel Kauers, Christoph Koutschan, and Doron Zeilberger. Proof of Ira Gessel's lattice path conjecture. *Proc. Natl. Acad. Sci. USA*, 106(28):11502–11505, 2009.
59. Manuel Kauers and Doron Zeilberger. The quasi-holonomic ansatz and restricted lattice walks. *J. Difference Equ. Appl.*, 14(10-11):1119–1126, 2008.
60. Manuel Kauers and Doron Zeilberger. The computational challenge of enumerating high-dimensional rook walks. *Adv. in Appl. Math.*, 47(4):813–819, 2011.
61. J. F. C. Kingman. Two similar queues in parallel. *Ann. Math. Statist.*, 32:1314–1323, 1961.
62. Donald E. Knuth. *The art of computer programming. Vol. 1*. Addison-Wesley, Reading, MA, 1997. Fundamental algorithms, Third edition [of MR0286317].
63. Christoph Koutschan, Manuel Kauers, and Doron Zeilberger. Proof of George Andrews's and David Robbins's q-TSPP conjecture. *Proc. Natl. Acad. Sci. USA*, 108(6):2196–2199, 2011.
64. Christian Krattenthaler. Lattice path enumeration. In *Handbook of enumerative combinatorics*, Discrete Math. Appl. (Boca Raton), pages 589–678. CRC Press, Boca Raton, FL, 2015.
65. G. Kreweras. Sur une classe de problémes liés au treillis des partitions d'entiers. *Cahiers du B.U.R.O.*, 5–105, 1965.
66. Pierre Lairez. Computing periods of rational integrals. *Math. Comp.*, 85(300):1719–1752, 2016.
67. Peter Luksch and Marko Petkovšek. An explicit formula for $|\mathrm{FM}(\mathbf{1}+\mathbf{1}+\mathbf{n})|$. *Order*, 6(4):319–324, 1990.
68. V. A. Malyshev. An analytical method in the theory of two-dimensional positive random walks. *Siberian Mathematical Journal*, 13(6):917–929, 1972.
69. V. A. Malyšev. Wiener-Hopf equations in the quarter-plane, discrete groups and automorphic functions. *Mat. Sb. (N.S.)*, 84 (126):499–525, 1971.
70. V. A. Malyšev. Wiener-Hopf equations and their applications in probability theory. In *Probability theory. Mathematical statistics. Theoretical cybernetics, Vol. 13 (Russian)*, pages 5–35. English translation in J. Math. Sci. 7 (1977), 129–148, 1976.
71. Stephen Melczer and Marni Mishna. Singularity analysis via the iterated kernel method. *Combinatorics, Probability and Computing*, 23(5):861–888, 2014.
72. Stephen Melczer and Marni Mishna. Asymptotic lattice path enumeration using diagonals. *Algorithmica*, 75(4):782–811, 2016.
73. Stephen Melczer and Mark C. Wilson. Asymptotics of lattice walks via analytic combinatorics in several variables. In *Proceedings of FPSAC 2016 (28th International Conference on Formal Power Series and Algebraic Combinatorics), DMTCS proc. BC*, pages 863–874, 2016.
74. Stephen Melczer and Mark C. Wilson. Higher Dimensional Lattice Walks: Connecting Combinatorial and Analytic Behavior. *SIAM J. Discrete Math.*, 33(4):2140–2174, 2019.
75. M. Mishna. Classifying lattice walks in the quarter plane. In *Proceedings of the Nineteenth International Conference on Formal Power Series and Algebraic Combinatorics (FPSAC)*, Tianjin, China, 2007.
76. Marni Mishna and Andrew Rechnitzer. Two non-holonomic lattice walks in the quarter plane. *Theoret. Comput. Sci.*, 410(38-40):3616–3630, 2009.

77. Sri Gopal Mohanty. *Lattice path counting and applications*. Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London-Toronto, Ont., 1979. Probability and Mathematical Statistics.
78. A. De Moivre. De mensura sortis. *Philosophical Trans. London*, 27:213–264, 1711.
79. T. V. Narayana. *Lattice path combinatorics with statistical applications*, volume 23 of *Mathematical Expositions*. University of Toronto Press, Toronto, Ont., 1979.
80. M. Petkovšek. The irrational chess knight. In *Proceedings of FPSAC'98, Toronto*, pages 513–522, 1998.
81. Helmut Prodinger. The kernel method: a collection of examples. *Sém. Lothar. Combin.*, 50:Art. B50f, 19, 2003/04.
82. N. A. Slavnov. *Algebraic Bethe ansatz*, volume arXiv:1804.07350. Lecture notes, Steklov Mathematical Institute, 2019.
83. Frank Spitzer. *Principles of random walk*. Springer-Verlag, New York-Heidelberg, second edition, 1976. Graduate Texts in Mathematics, Vol. 34.
84. Tatsuya Tate and Steve Zelditch. Lattice path combinatorics and asymptotics of multiplicities of weights in tensor powers. *J. Funct. Anal.*, 217(2):402–447, 2004.
85. E. J. Janse van Rensburg. *The statistical mechanics of interacting walks, polygons, animals and vesicles*, volume 18 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford, 2000.
86. E. J. Janse van Rensburg, T. Prellberg, and A. Rechnitzer. Partially directed paths in a wedge. *J. Combin. Theory Ser. A*, 115(4):623–650, 2008.
87. W.A. Whitworth. Arrangements of m things of one sort and n things of another sort under certain conditions of priority. *Messenger of Math.*, 8:105–114, 1878.
88. Wen-Jin Woan, Lou Shapiro, and D. G. Rogers. The Catalan numbers, the Lebesgue integral, and 4^{n-2}. *Amer. Math. Monthly*, 104(10):926–931, 1997.
89. Doron Zeilberger. André's reflection proof generalized to the many-candidate ballot problem. *Discrete Math.*, 44(3):325–326, 1983.

Part II
Smooth ACSV and Applications

Chapter 5
The Theory of ACSV for Smooth Points

We see. . . that the theory of probabilities is at bottom nothing but common sense reduced to calculus: it enables us to appreciate with exactitude that which excellent minds sense by a kind of instinct for which they are often unable to account.
— Pierre-Simon Laplace

Sweet Analytics, 'tis thou has ravish'd me...
– Christopher Marlowe (as Doctor John Faustus)

In this chapter we develop the theory of analytic combinatorics in several variables under a set of assumptions which hold very generally yet simplify our approach. In the easiest cases, these multivariate results can be obtained through clever applications of the univariate methods from Chapter 2 together with a classical 'saddle-point method' for asymptotics of certain parametrized integrals. Dealing with more general situations will require truly multivariate techniques making use of advanced results from several areas of mathematics, but one should always keep the univariate approach in mind: examine the singularities of the generating function under consideration, determine a finite set of those singularities which dictate the exponential growth of coefficients, then perform a local analysis of the function at those points to determine dominant asymptotics. Our results apply both to rational functions and ratios of analytic functions.

We show how to apply our results on a large number of examples in this chapter, and give a detailed treatment of a family of lattice path models in Chapter 6. Algorithms automating the techniques in this chapter are developed in Chapter 7.

A Simplified Asymptotic Statement

Our main results, stated in Theorems 5.2, 5.3, and 5.4, are asymptotic statements holding for convergent Laurent expansions of meromorphic functions. These results are powerful, and very general, but their generality makes stating them slightly cumbersome. We thus begin with a simplified result which covers the most common case appearing in applications. Recall from Definition 3.9 in Chapter 3 that a singularity $\mathbf{w}$ of a meromorphic function $F(\mathbf{z})$ is *minimal* if no other singularity $\mathbf{z}$ of F satisfies $|z_j| \leq |w_j|$ for all $1 \leq j \leq d$ where at least one of the inequalities is strict.

Theorem 5.1 *Let* $\mathbf{r} \in \mathbb{R}^d_{>0}$ *and let* $G(\mathbf{z}), H(\mathbf{z}) \in \mathbb{Q}[\mathbf{z}]$ *be coprime polynomials such that* $F(\mathbf{z}) = G(\mathbf{z})/H(\mathbf{z})$ *admits a power series expansion* $F(\mathbf{z}) = \sum_{\mathbf{i}\in\mathbb{N}^d} f_{\mathbf{i}}\mathbf{z}^{\mathbf{i}}$. *Suppose that the system of polynomial equations*

$$H(\mathbf{z}) = r_2 z_1 H_{z_1}(\mathbf{z}) - r_1 z_2 H_{z_2}(\mathbf{z}) = \cdots = r_d z_1 H_{z_1}(\mathbf{z}) - r_1 z_d H_{z_d}(\mathbf{z}) = 0$$

admits a finite number of solutions, exactly one of which, $\mathbf{w} \in \mathbb{C}_*^d$, is minimal. Suppose further that $H_{z_d}(\mathbf{w}) \neq 0$, that the matrix $\mathcal{H}$ defined by (5.24) and (5.25) below has non-zero determinant, and that $G(\mathbf{w}) \neq 0$. Then, as $n \to \infty$,

$$f_{n\mathbf{r}} = \mathbf{w}^{-n\mathbf{r}} n^{(1-d)/2} \frac{(2\pi r_d)^{(1-d)/2}}{\sqrt{\det(\mathcal{H})}} \frac{-G(\mathbf{w})}{w_d H_{z_d}(\mathbf{w})} \left(1 + O\left(\frac{1}{n}\right)\right) \tag{5.1}$$

when $n\mathbf{r} \in \mathbb{N}^d$. As $\mathbf{r}$ varies in any sufficiently small neighbourhood $\mathcal{N}$ in $\mathbb{R}_{>0}^d$ the solution $\mathbf{w} = \mathbf{w}(\mathbf{r})$ varies smoothly with $\mathbf{r}$. When the above conditions are satisfied for each $\mathbf{w}(\mathbf{r})$ with $\mathbf{r} \in \mathcal{N}$, the hidden constant in the big-O error in (5.1) can be chosen independently of $\mathbf{r}$ in any compact subset of $\mathcal{N}$.

When the zero set of H contains a finite number of points with the same coordinate-wise modulus as $\mathbf{w}$, all of which satisfy the same conditions as $\mathbf{w}$, then an asymptotic expansion of $f_{n\mathbf{r}}$ is obtained by summing the right hand side of (5.1) at each point.

When the power series expansion of $F(\mathbf{z})$ contains only a finite number of negative coefficients, any point $\mathbf{w} \in \mathbb{C}_*^d$ with $H(\mathbf{w}) = 0$ is minimal if and only if $H(tw_1, \ldots, tw_d)$ is non-zero for all $0 \leq t < 1$.

Theorem 5.1 is a special case of Theorem 5.4 below.

Example 5.1 (Asymptotics of Central Binomial Coefficients)

Consider the power series expansion

$$F(x, y) = \frac{1}{1 - x - y} = \sum_{i,j \geq 0} f_{i,j} x^i y^j = \sum_{i,j \geq 0} \binom{i+j}{i} x^i y^j.$$

For $r, s > 0$ the system of equations

$$H(x, y) = sxH_x(x, y) - ryH_y(x, y) = 0$$

becomes

$$1 - x - y = -xs + yr = 0,$$

with solution

$$(x, y) = \left(\frac{r}{r+s}, \frac{s}{r+s}\right).$$

The minimal singularities of $F(x, y)$ were described in Chapter 3, where it was shown that any point (x, y) with $x, y > 0$ and $x + y = 1$ was minimal. Building the matrix $\mathcal{H}$ defined by (5.24) and (5.25), which in the bivariate case is just a constant, shows that the conditions of Theorem 5.1 are satisfied here, and we calculate

$$f_{rn,sn} \sim \left(\frac{r+s}{r}\right)^{rn} \left(\frac{r+s}{s}\right)^{sn} \frac{\sqrt{r+s}}{\sqrt{2rs\pi\, n}}.$$

Section 5.1 goes into great detail on how to derive this expansion.

Below we relax many of the conditions of Theorem 5.1, considering more general Laurent expansions and determining an explicit asymptotic expansion of $f_{n\mathbf{r}}$ in decreasing powers of n. In addition, we show that most of our assumptions hold *generically*, meaning they hold for all rational functions with numerator and denominator of fixed degree, except those whose coefficients lie in some fixed algebraic set. Our main results and the necessary surrounding theory are developed in Section 5.2. Section 5.3 illustrates how the theory is put into practice, including (in Section 5.3.3) a characterization of how asymptotics of $f_{n\mathbf{r}}$ varies with $\mathbf{r}$ and an illustration for how this implies a central limit theorem for series coefficients. Section 5.3.4 concludes by proving that most of our assumptions hold generically. We make frequent use of the results and concepts introduced in Chapter 3 throughout.

Setup

As in the univariate case, the analysis begins with the (multivariate) Cauchy integral formula from Theorem 3.1 (for power series expansions) and Proposition 3.10 (for more general Laurent expansions) in Chapter 3. Given a meromorphic function $F(\mathbf{z})$ with absolutely convergent Laurent expansion $F(\mathbf{z}) = \sum_{\mathbf{i}\in\mathbb{Z}^d} f_{\mathbf{i}}\mathbf{z}^{\mathbf{i}}$ in some domain $\mathcal{D} \subset \mathbb{C}^d$, we fix a direction $\mathbf{r} \in \mathbb{R}^d$ and study asymptotics of the coefficient sequence

$$f_{n\mathbf{r}} = \frac{1}{(2\pi i)^d} \int_{T(\mathbf{w})} F(\mathbf{z}) \frac{d\mathbf{z}}{\mathbf{z}^{n\mathbf{r}+\mathbf{1}}}$$

as $n \to \infty$, where $T(\mathbf{w})$ is the polytorus defined by any $\mathbf{w} \in \mathcal{D}$. In most applications, $F(\mathbf{z})$ is a rational function and we study asymptotics of its power series expansion at the origin; in this case $T(\mathbf{w})$ can be taken as any product of circles sufficiently close to the origin. Although $f_{n\mathbf{r}}$ is essentially not defined if $n\mathbf{r} \notin \mathbb{Z}^d$, we will show that asymptotics vary smoothly with $\mathbf{r}$, allowing us to make meaningful asymptotic statements about $f_{n\mathbf{r}}$ even when $\mathbf{r}$ has irrational coordinates.

Our first set of results needs only univariate techniques and the following two facts. The first follows by induction from the univariate complex analysis results discussed in Chapter 2, while the second is a consequence of the triangle inequality.

Deforming Curves of Integration: Let $f(\mathbf{z})$ be a meromorphic function and let $T(\mathbf{a}), T(\mathbf{b}) \subset \mathbb{C}^d$ be tori in $\mathbb{C}^d$ such that $f(\mathbf{z})$ has no singularities with $|z_j|$ between $|a_j|$ and $|b_j|$ for each $1 \le j \le d$. Then

$$\int_{T(\mathbf{a})} f(\mathbf{z})d\mathbf{z} = \int_{T(\mathbf{b})} f(\mathbf{z})d\mathbf{z}.$$

Maximum Modulus Integral Bound: If $f(\mathbf{z})$ is continuous on a contour $C \subset \mathbb{C}^d$ of finite area, then

$$\left|\int_C f(\mathbf{z})d\mathbf{z}\right| \leq \text{area}(C) \times \max_{\mathbf{z}\in C} |f(\mathbf{z})|.$$

We start with a detailed study of the rational function $F(x, y) = 1/(1 - x - y)$, which lays out the steps we must go through in the general case.

5.1 Central Binomial Coefficient Asymptotics

Let

$$F(x, y) = \frac{1}{1 - x - y}.$$

In Chapter 3 we saw that $F(x, y)$ admits three distinct convergent Laurent expansions. Here, we begin by determining coefficient asymptotics in the main diagonal direction $\mathbf{r} = (1, 1)$ of the power series expansion

$$F(x, y) = \sum_{i,j\geq 0} \binom{i+j}{i} x^i y^j,$$

with domain of (absolute) convergence

$$\mathcal{D} = \{(x, y) \in \mathbb{C}^2 : |x| + |y| < 1\}.$$

The Cauchy integral formula implies

$$\binom{2n}{n} = \frac{1}{(2\pi i)^2} \int_{T(a,b)} \frac{1}{1 - x - y} \frac{dxdy}{x^{n+1}y^{n+1}}, \tag{5.2}$$

for any $(a, b) \in \mathcal{D}$, which forms the basis of our analysis. We break our argument into several steps.

Step 1: Bound Exponential Growth

By Corollary 3.2 of Chapter 3, the dominant asymptotic behaviour of a rational diagonal sequence is given by a finite sum of terms of the form $C\, n^\alpha\, \zeta^n\, (\log n)^\ell$, where $C \in \mathbb{C}$, $\alpha \in \mathbb{Q}$, $\ell \in \mathbb{N}$, and ζ is algebraic. As in the univariate case, the first step of the analysis is to determine information about the exponential growth

$$\rho = \limsup_{n\to\infty} |f_{n,n}|^{1/n}.$$

In the univariate setting the exponential growth of a sequence is obtained by finding the minimal modulus of its generating function's singularities and taking the reciprocal. Unfortunately, in dimension two or greater there will always be an infinite number of singularities 'closest' to the origin, with each only giving a bound on the exponential growth.

In our example $|a| + |b| < 1$ whenever $(a, b) \in \mathcal{D}$, so that

$$\left|\frac{1}{1-x-y}\right| \leq \frac{1}{1-|a|-|b|}$$

for all $(x, y) \in T(a, b)$. Thus, the maximum modulus integral bound implies

$$\binom{2n}{n} = \left|\frac{1}{(2\pi \mathrm{i})^2}\int_{T(a,b)} \frac{1}{1-x-y}\,\frac{dxdy}{x^{n+1}y^{n+1}}\right| \leq \frac{|ab|^{-n}}{1-|a|-|b|} \tag{5.3}$$

for all $(a, b) \in \mathcal{D}$. Equation (5.3) gives a family of bounds

$$\limsup_{n\to\infty} \binom{2n}{n}^{1/n} \leq |ab|^{-1} \tag{5.4}$$

on the exponential growth of the central binomial coefficients, one for each pair of points $(a, b) \in \mathcal{D}$. In fact, allowing (a, b) to approach the boundary $\partial\mathcal{D}$ shows that the exponential growth ρ is bounded above by $|ab|^{-1}$ for all $(a, b) \in \overline{\mathcal{D}}$.

It is natural to wonder which points give the best upper bound, and whether that bound is tight. In fact, answering these two questions is usually the hardest step when studying a multivariate generating function. As the upper bound $|ab|^{-1}$ decreases as the coordinates a and b get farther from the origin, the minimum of $|ab|^{-1}$ on

$$\overline{\mathcal{D}} = \{(a, b) \in \mathbb{C}^2 : |a| + |b| \leq 1\}$$

occurs on the boundary of $\mathcal{D}$, where $|a| + |b| = 1$. Thus, we want to minimize $|ab|^{-1}$ subject to $|a| + |b| = 1$. Examining the function $t^{-1}(1-t)^{-1}$ for $0 \leq t \leq 1$ shows the minimum is achieved when $|a| = |b| = 1/2$, where $|ab|^{-1} = 4$.

Returning to the bound in (5.4), we have shown that for every $\varepsilon > 0$ there exists a constant $C_\varepsilon > 0$ such that

$$\binom{2n}{n} \leq C_\varepsilon\,(4+\varepsilon)^n$$

for all $n \in \mathbb{N}$. Using the fact that the central binomial coefficients have an algebraic generating function, we have seen in Chapter 2 that

$$\binom{2n}{n} = \frac{4^n}{\sqrt{\pi\, n}}\left(1 + O\left(\frac{1}{n}\right)\right)$$

as $n \to \infty$, so our upper bound of 4 on the exponential growth is tight.

Step 2: Determine Contributing Singularities

In order to completely determine asymptotics of this diagonal sequence, we will perform a local analysis of $F(x, y)$ near some of its singularities. But which singularities should we study? Because the minimum of $|ab|^{-1}$ on $\overline{\mathcal{D}}$ occurs on the boundary $\partial\mathcal{D}$, and depends only on the moduli of the coordinates, there are singularities of F achieving this minimum. Since the denominator under consideration is $1 - x - y$, the only singularity of $F(x, y)$ with $|x| = |y| = 1/2$ is the point $\boldsymbol{\sigma} = (1/2, 1/2)$.

The Cauchy integrand in (5.2) has exponential growth close to 4^n when (x, y) is near to $\boldsymbol{\sigma}$, and has different exponential growth near any other singularity of $F(x, y)$. The point $\boldsymbol{\sigma}$ is thus the *only* singularity of $F(x, y)$ where the growth of the Cauchy integrand matches our predicted diagonal sequence growth, and therefore the only singularity where a local analysis of the Cauchy integral could possibly capture the asymptotic growth of the diagonal coefficients. Following this logic, we attempt to determine dominant asymptotics by manipulating the Cauchy integral in (5.2) into an integral whose domain stays near this singularity, then replace $F(x, y)$ by a local approximation near $\boldsymbol{\sigma}$.

Step 3: Localize the Cauchy Integral and Compute a Residue

For any $n \in \mathbb{N}$ let

$$I = I_n = \frac{1}{(2\pi i)^2} \int_{|x|=1/2} \left(\int_{|y|=1/4} \frac{1}{1-x-y} \frac{dy}{y^{n+1}} \right) \frac{dx}{x^{n+1}}.$$

Since $\boldsymbol{\sigma} = (1/2, 1/2)$ lies on the boundary of the domain of convergence $\partial\mathcal{D}$, the point $(1/2, 1/4)$ lies in $\mathcal{D}$ and thus $I = \binom{2n}{n}$. In order to obtain an integral where x is restricted to a neighbourhood of $1/2$, we define

$$\mathcal{N} = \left\{ |x| = 1/2 : \arg(x) \in \left(-\frac{\pi}{4}, \frac{\pi}{4} \right) \right\} \quad \text{and} \quad \mathcal{N}' = \{|x| = 1/2\} \setminus \mathcal{N}.$$

Figure 5.1 illustrates the sets $\mathcal{N}$ and $\mathcal{N}'$, and we note that

$$|1 - x| < \underbrace{\left| 1 - \frac{e^{i\pi/4}}{2} \right|}_{\rho} = 0.7368\ldots$$

for $x \in \mathcal{N}$, and $|1 - x| \geq \rho$ for $x \in \mathcal{N}'$. We now compare I to the 'localized' integral

$$I_{\text{loc}} = \frac{1}{(2\pi i)^2} \int_{\mathcal{N}} \left(\int_{|y|=1/4} \frac{1}{1-x-y} \frac{dy}{y^{n+1}} \right) \frac{dx}{x^{n+1}}.$$

Our arguments are local, in the sense that replacing $\mathcal{N}$ with any smaller neighbourhood of $1/2$ will not change our results. For fixed $x \in \mathcal{N}'$,

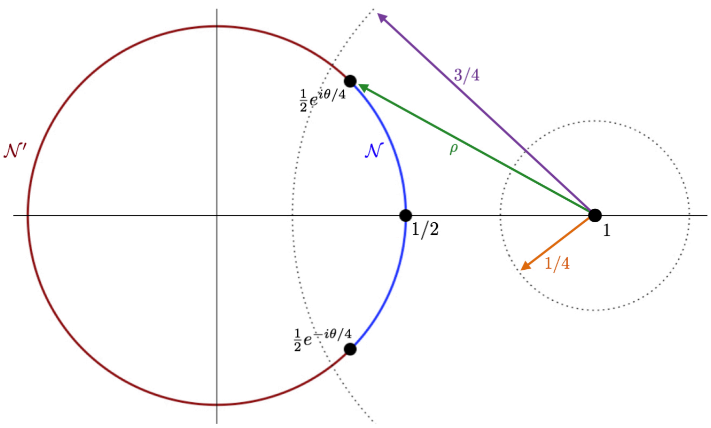


Fig. 5.1 The domains of integration $\mathcal{N}$ and $\mathcal{N}'$, with $|1-x| < \rho$ for $x \in \mathcal{N}$ and $|1-x| \geq \rho > 1/4$ for $x \in \mathcal{N}'$.

$$\begin{aligned}\left|\frac{1}{(2\pi i)}\int_{|y|=1/4}\frac{1}{1-x-y}\frac{dy}{y^{n+1}}\right| &= \left|\frac{1}{(2\pi i)}\int_{|y|=1/4}\frac{1/(1-x)}{1-\frac{y}{1-x}}\frac{dy}{y^{n+1}}\right| \\ &= \left|[y^n]\sum_{j\geq 0}(1-x)^{-(j+1)}y^j\right| \\ &= |1-x|^{-(n+1)} \\ &\leq \rho^{-(n+1)},\end{aligned}$$

where the series expansion is valid as $|y| = 1/4 < |1-x|$ when $x \in \mathcal{N}'$. Thus, the maximum modulus integral bound implies

$$\begin{aligned}|I - I_{\text{loc}}| = \frac{1}{(2\pi i)^2}\int_{\mathcal{N}'}\left(\int_{|y|=1/4}\frac{1}{1-x-y}\frac{dy}{y^{n+1}}\right)\frac{dx}{x^{n+1}} &\leq \frac{\text{length}(\mathcal{N}')}{2\pi}\rho^{-(n+1)}2^{n+1} \\ &= \frac{3}{8\rho\pi}\left(\frac{2}{\rho}\right)^n.\end{aligned}$$

Since $2/\rho \leq 2.72$, this implies we can replace the integral I with I_{loc} in our asymptotic arguments and introduce an error that grows at an exponentially smaller rate than our diagonal sequence. Our next tactic is to introduce the integral

$$I_{\text{out}} = \frac{1}{(2\pi i)^2}\int_{\mathcal{N}}\left(\int_{|y|=3/4}\frac{1}{1-x-y}\frac{dy}{y^{n+1}}\right)\frac{dx}{x^{n+1}}$$

whose domain of integration lies outside the domain of convergence $\mathcal{D}$. For $x \in \mathcal{N}$, the quantity $|1-x|$ is bounded away from $3/4$ so that $1/|1-x-y|$ is bounded when $|y| = 3/4$. The maximum modulus integral bound then implies the existence of a constant C such that

$$|I_{\text{out}}| \leq C\, 2^n \left(\frac{4}{3}\right)^n = O\left(\left(\frac{8}{3}\right)^n\right),$$

which grows exponentially smaller than our diagonal sequence.

Remark 5.1 There are solutions to $1-x-y=0$ with $|x|=1/2$ and $|y|=3/4$, but no solution has $x \in \mathcal{N}$. This is why we introduce and localize to the neighbourhood $\mathcal{N}$.

Finally, define

$$\begin{aligned}\chi &= I_{\text{loc}} - I_{\text{out}} \\ &= \frac{-1}{2\pi \mathrm{i}} \int_{\mathcal{N}} \frac{1}{2\pi \mathrm{i}} \left(\int_{|y|=3/4} \frac{1}{1-x-y} \frac{dy}{y^{n+1}} - \int_{|y|=1/4} \frac{1}{1-x-y} \frac{dy}{y^{n+1}} \right) \frac{dx}{x^{n+1}}.\end{aligned}$$

For any fixed $x \in \mathcal{N}$, the function $F(x,y) = 1/(1-x-y)$ has a unique pole between the curves $\{|y| = 1/4\}$ and $\{|y| = 3/4\}$, at $y = 1-x$. Thus, the (univariate) Cauchy residue theorem discussed in Chapter 2 implies

$$\frac{1}{2\pi \mathrm{i}} \left(\int_{|y|=3/4} \frac{1}{1-x-y} \frac{dy}{y^{n+1}} - \int_{|y|=1/4} \frac{1}{1-x-y} \frac{dy}{y^{n+1}} \right) = -(1-x)^{-(n+1)},$$

for each $x \in \mathcal{N}$, and

$$\chi = \frac{1}{2\pi \mathrm{i}} \int_{\mathcal{N}} \frac{dx}{x^{n+1}(1-x)^{n+1}}. \tag{5.5}$$

This gives a good asymptotic approximation of our diagonal sequence, because

$$\left| \binom{2n}{n} - \chi \right| = |I - (I_{\text{loc}} - I_{\text{out}})| \leq |I - I_{\text{loc}}| + |I_{\text{out}}| = O\left(\left(\frac{2}{\rho}\right)^n\right).$$

Remark 5.2 The emergence of the integral in (5.5) should not be surprising. In fact, the binomial theorem implies

$$\binom{2n}{n} = [x^n]\,(1-x)^{-(n+1)} = \frac{1}{2\pi \mathrm{i}} \int_{|x|=1/2} \frac{dx}{x^{n+1}(1-x)^{n+1}}, \tag{5.6}$$

which is an exact equality instead of an asymptotic approximation. The reason our argument has an (asymptotically negligible) error is that we restrict to the neighbourhood $\mathcal{N}$ of $x = 1/2$ in order to have the integrand of I_{out} bounded when $|y| = 3/4$. In this example, we can define the domain of integration for y in I_{out} to be $|y| = 3/2 + r$ for any $r > 0$ and the restriction of x to the neighbourhood $\mathcal{N}$ of $1/2$ becomes unnecessary; the rest of our argument then yields the exact equality (5.6). The reason we did not take this approach above is because for general rational (or meromorphic) functions one must stick to a local analysis near the singularities of interest. We thus restrict x to the neighbourhood $\mathcal{N}$ and obtain an asymptotically negligible error here as this argument is the one that generalizes. Chapter 8 provides a detailed asymptotic analysis for diagonals of rational functions whose denominators are a product of linear factors, giving exact equalities such as (5.6).

Step 4: Apply the Saddle-Point Method

Parameterizing the domain of integration $\mathcal{N}$ in (5.5) as $\mathcal{N} = \left\{e^{i\theta}/2 : \theta \in (-\pi/4, \pi/4)\right\}$ converts this integral expression for χ into

$$\chi = \frac{4^n}{2\pi} \int_{-\pi/4}^{\pi/4} A(\theta) e^{-n\phi(\theta)} d\theta, \tag{5.7}$$

where

$$A(\theta) = \frac{1}{1 - e^{i\theta}/2} \qquad \text{and} \qquad \phi(\theta) = \log\left(2 - e^{i\theta}\right) + i\theta.$$

This is our first example of a *Fourier-Laplace* integral, whose asymptotics can be calculated using the saddle-point method and related techniques. To determine asymptotics we will replace the *amplitude* function $A(\theta)$ and *phase* $\phi(\theta)$ in (5.7) by the leading terms of their expansions

$$A(\theta) = 2 + 2i\theta - 3\theta^2 + \cdots \qquad \text{and} \qquad \phi(\theta) = \theta^2 + i\theta^3 + \cdots \tag{5.8}$$

at the origin, after which the resulting integral can be computed explicitly. First, we restrict the domain of integration in (5.7) to a small neighbourhood of the origin: if $B_n = n^{-2/5}$ then (5.8) implies

$$\left|e^{-n\phi(\theta)}\right| = e^{-n\Re(\phi)(\theta)} \leq e^{-n^{1/5}+O(1)}$$

whenever $|\theta| \geq B_n$, so that

$$\begin{aligned}\int_{-\pi/4}^{\pi/4} A(\theta)e^{-n\phi(\theta)}d\theta &= \int_{-B_n}^{B_n} A(\theta)e^{-n\phi(\theta)}d\theta + \int_{-\pi/4}^{-B_n} A(\theta)e^{-n\phi(\theta)}d\theta + \int_{B_n}^{\pi/4} A(\theta)e^{-n\phi(\theta)}d\theta \\ &= \int_{-B_n}^{B_n} A(\theta)e^{-n\phi(\theta)}d\theta + O\left(e^{-n^{1/5}}\right).\end{aligned}$$

For $|\theta| \leq B_n$ the expansions in (5.8) imply

$$A(\theta) = 2+O\left(n^{-2/5}\right) \qquad \text{and} \qquad e^{-n\phi(\theta)} = e^{-n\theta^2+O(n^{-1/5})} = e^{-n\theta^2}\left(1 + O\left(n^{-1/5}\right)\right),$$

meaning

$$\chi = \frac{4^n}{2\pi}\left(\int_{-B_n}^{B_n} 2e^{-n\theta^2} d\theta\right)\left(1 + O\left(n^{-1/5}\right)\right). \tag{5.9}$$

Asymptotics of such an integral over the entire real line is easy to determine: making the change of variable $\theta = t/\sqrt{n}$ gives

$$\int_{-\infty}^{\infty} e^{-n\theta^2} d\theta = n^{-1/2} \int_{-\infty}^{\infty} e^{-t^2} dt = \sqrt{\frac{\pi}{n}}.$$

Since

$$\begin{aligned}\int_{B_n}^{\infty} e^{-n\theta^2} d\theta = \int_0^{\infty} e^{-n(B_n+\theta)^2} d\theta = e^{-nB_n^2} \int_0^{\infty} e^{-n\theta^2 - nB_n\theta} d\theta \\ \leq e^{-nB_n^2} \int_0^{\infty} e^{-\theta^2} d\theta \\ = O\left(e^{-n^{1/5}}\right),\end{aligned}$$

we can 'add the tails' back to the domain of integration in (5.9) while introducing an exponentially small error, ultimately yielding

$$\chi = \frac{4^n}{2\pi} \left(\int_{-\infty}^{\infty} 2e^{-n\theta^2} d\theta \right) \left(1 + O\left(n^{-1/5}\right)\right) = \frac{4^n}{\sqrt{\pi n}} \left(1 + O\left(n^{-1/5}\right)\right).$$

Thus, we have obtained an asymptotic estimate

$$\binom{2n}{n} = \frac{4^n}{\sqrt{\pi n}} \left(1 + O\left(n^{-1/5}\right)\right).$$

A more careful analysis can bring the error term $O\left(n^{-1/5}\right)$ to $O(n^{-1})$, and even determine an asymptotic expansion in powers of n^{-1}. The key properties that allow for such an analysis are that A and ϕ are analytic at the origin and

- $\phi(0) = \phi'(0) = 0$ while $\phi''(\theta) \neq 0$ (so that the dominant term in the Taylor series for ϕ at the origin is quadratic, leading to a Gaussian-type integral);
- the real part of ϕ is non-negative on the domain of integration (so that the integral decays away from the origin);
- $\phi'(\theta) \neq 0$ on the domain of integration unless $\theta = 0$ (so that the origin is the only point where local behaviour will dictate asymptotics),

although these can be relaxed to differing degrees (especially for univariate integrals). Good expositions of the asymptotics of univariate saddle-point integrals can be found in de Bruijn [13] and Flajolet and Sedgewick [16, Sect. VIII], which illustrate a deep, rich, and well-developed theory, and motivate some of our choices above. Below we use Proposition 5.3, which gives asymptotics in a generalized multivariate setting.

The miracle which underlies analytic combinatorics in several variables is the fact that, by making a natural choice of singularities to study–those giving the best bound on exponential growth–one ends up with a class of integrals which can be asymptotically approximated[1]. Under the assumptions of this chapter, when dealing with rational functions of d variables the Fourier-Laplace integral expressions ob-

[1] As we will see, this miracle is due to the Cauchy-Riemman equations.

tained are $d - 1$ dimensional. They have no analogue in the univariate case, where one is finished after computing the residue in Step 3.

5.1.1 Asymptotics in General Directions

Suppose we wish to determine asymptotics of the coefficients $f_{rn,sn}$ in a more general direction $\mathbf{r} = (r, s) \in \mathbb{N}^2$. Because changing the direction $\mathbf{r}$ does not affect the underlying rational function $F(x, y)$, its singular set $\mathcal{V}$, or the power series domain of convergence $\mathcal{D}$, we have the modified Cauchy integral representation

$$\binom{rn+sn}{rn} = f_{rn,sn} = \frac{1}{(2\pi i)^2}\int_{T(a,b)} \frac{1}{1-x-y}\,\frac{dxdy}{x^{rn+1}y^{sn+1}} \tag{5.10}$$

for any $(a, b) \in \mathcal{D}$. As in our study of the main diagonal, the maximum modulus integral bound applied to (5.10) gives a bound on the exponential growth

$$\rho = \limsup_{n\to\infty} |f_{rn,sn}|^{1/n} \leq |a|^{-r}|b|^{-s}$$

for all $(a, b) \in \overline{\mathcal{D}}$. In order to minimize this upper bound on exponential growth, we introduce the *height function*

$$h_{\mathbf{r}}(x, y) = h(x, y) = -r\log|x| - s\log|y|$$

and search for the minimum of $h(x, y)$ on $\partial\mathcal{D}$, where $|x| + |y| = 1$. Writing $|x| = t$ for $0 \leq t \leq 1$, we thus search for the minimum of $g(t) = -r\log t - s\log(1-t)$ when $0 \leq t \leq 1$. Since $r, s > 0$ the function $g(t)$ goes to infinity as t approaches 0 or 1, and we may find the minimum of g by solving the equation $g'(t) = 0$. This implies that the minimum of $h_{\mathbf{r}}$ on $\partial\mathcal{D}$ occurs when $(|x|, |y|) = (\frac{r}{r+s}, \frac{s}{r+s})$, and there is precisely one singularity with this coordinate-wise modulus, $(x_*, y_*) = (\frac{r}{r+s}, \frac{s}{r+s})$.

Remark 5.3 Recall the discussion of polynomial amoebas and the Relog map from Section 3.3.1 in Chapter 3. If $p = \log|x|$ and $q = \log|y|$ then

$$h(x, y) = -r\log|x| - s\log|y| = -rp - sq = -(r, s)\cdot(p, q).$$

Thus, minimizing $h(x, y)$ on $\overline{\mathcal{D}}$ corresponds to minimizing the linear function $\tilde{h}(p, q) = -(r, s)\cdot(p, q)$ on the closure of the convex component $B = \mathrm{Relog}(\mathcal{D})$ of the amoeba complement amoeba$(1-x-y)^c$. Pictorially, for any direction (r, s) we want to find the point(s) on ∂B where the support hyperplane to $\overline{B}$ has normal (r, s); see Figure 5.2. Proposition 3.6 from Chapter 3 implies that every point on ∂B is a minimizer of $\tilde{h}$ for some (r, s) when we relax the condition that r and s are natural numbers and allow them to take positive real values.

Following the steps outlined above, we note

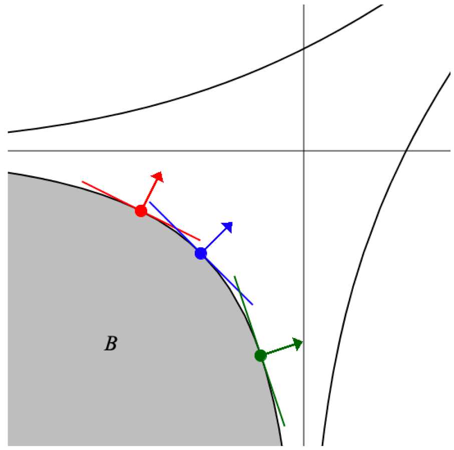


Fig. 5.2 The boundary of amoeba$(1-x-y)$ together with the component B of amoeba$(1-x-y)^c$ corresponding to the power series expansion of $1/(1-x-y)$. Level sets of the height function $\tilde{h}$ for directions (1, 2) on top, (1, 1) in the middle, and (4, 1) on the bottom are shown, together with the corresponding minimizer of $\tilde{h}$ on ∂B.

$$f_{rn,sn} = I = \frac{1}{(2\pi i)^2} \int_{|x|=x_*} \left(\int_{|y|=y_*-\varepsilon} \frac{1}{1-x-y} \frac{dy}{y^{sn+1}} \right) \frac{dx}{x^{rn+1}}$$

for any sufficiently small $\varepsilon > 0$, and introduce the integrals

$$I_{\text{loc}} = \frac{1}{(2\pi i)^2} \int_{\mathcal{N}} \left(\int_{|y|=y_*-\varepsilon} \frac{1}{1-x-y} \frac{dy}{y^{sn+1}} \right) \frac{dx}{x^{rn+1}}$$
$$I_{\text{out}} = \frac{1}{(2\pi i)^2} \int_{\mathcal{N}} \left(\int_{|y|=y_*+\varepsilon} \frac{1}{1-x-y} \frac{dy}{y^{sn+1}} \right) \frac{dx}{x^{rn+1}}$$

where $\mathcal{N} = \{|x| = x_* : \arg(x) \in (-\delta, \delta)\}$ for some sufficiently small $\delta > 0$. The same argument as the main diagonal case shows that both $I - I_{\text{loc}}$ and I_{out} grow exponentially slower than I, and we can approximate $f_{rn,sn}$ with the integral

$$\chi = I_{\text{loc}} - I_{\text{out}} = \frac{x_*^{-rn} y_*^{-sn}}{2\pi} \int_{-\delta}^{\delta} A(\theta) e^{-n\phi(\theta)} d\theta, \tag{5.11}$$

where

$$A(\theta) = \frac{1}{1 - x_* e^{i\theta}} = \frac{r+s}{s} + O(\theta)$$

and

$$\begin{aligned}\phi(\theta) &= r \log\left(x_* e^{i\theta}\right) + s \log\left(1 - x_* e^{i\theta}\right) - r \log(x_*) - s \log(y_*) \\ &= \frac{r(r+s)}{2s}\theta^2 + O(\theta^3).\end{aligned}$$

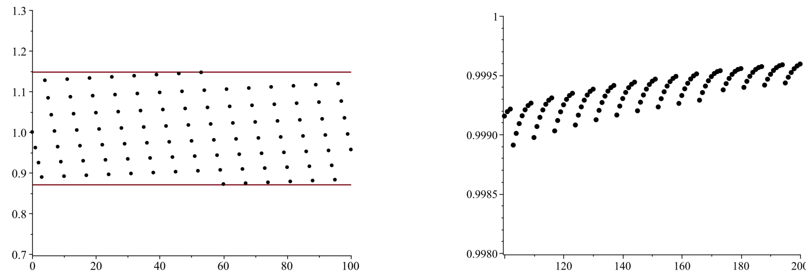


Fig. 5.3 *Left:* Values of the factor $M_n = \left(\frac{\pi}{\pi+1}\right)^{[n\pi]-n\pi}$ arising in our determination of asymptotics near the irrational direction $\mathbf{r} = (\pi, 1)$ for $0 \le n \le 100$, bounded between $\left(\frac{\pi}{\pi+1}\right)^{1/2}$ and $\left(\frac{\pi}{\pi+1}\right)^{-1/2}$. *Right:* The ratio of the coefficient $f_{[n\pi],n}$ and its limit behaviour $M_n A_n$ for $100 \le n \le 200$.

Because χ is a Fourier-Laplace integral satisfying the properties discussed above, the saddle-point method implies

$$\binom{rn+sn}{rn} = f_{rn,sn} \sim \left(\frac{r+s}{r}\right)^{rn} \left(\frac{r+s}{s}\right)^{sn} \left(\frac{r+s}{s}\right) \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-n\left(\frac{r(r+s)}{2s}\theta^2\right)} d\theta$$

$$= \left(\frac{r+s}{r}\right)^{rn} \left(\frac{r+s}{s}\right)^{sn} \frac{\sqrt{r+s}}{\sqrt{2rs\pi\, n}}. \tag{5.12}$$

A close examination of our argument shows that the error term in this approximation varies smoothly with (r, s) and, in particular, our asymptotic result holds uniformly when (r, s) varies in any compact set. This allows us to interpret asymptotics in a direction (r, s) with non-integer coordinates as the limit of asymptotics in rational directions approaching (r, s), in a manner specified precisely by Proposition 5.9 in Section 5.3.3 below.

Example 5.2 ('Asymptotics' in an Irrational Direction)

Although no non-zero multiple of $\mathbf{r} = (\pi, 1)$ contains integer coordinates, when n is large then $n\mathbf{r}$ gets arbitrarily close to vectors with integer coordinates. Let $[n\pi]$ denote the closest integer to $n\pi$ and let A_n denote the asymptotic approximation (5.12) with $r = \pi$ and $s = 1$. Proposition 5.9 below implies

$$f_{[n\pi],n} \to \left(\frac{\pi}{\pi+1}\right)^{[n\pi]-n\pi} A_n$$

as $n \to \infty$, so A_n determines the asymptotic behaviour of the coefficient with integer coordinates closest to $n\mathbf{r}$, up to a bounded factor M_n coming from rounding $n\pi$ to an integer. See Figure 5.3 for an illustration.

In fact, writing

$$\binom{rn+sn}{rn} = \frac{(rn+sn)!}{(rn)!(sn)!} = \frac{\Gamma(rn+sn+1)}{\Gamma(rn+1)\Gamma(sn+1)}$$

extends the coefficients $\binom{rn+sn}{rn}$ to all $(r,s) \in \mathbb{R}^2_{\geq 0}$ using the gamma function $\Gamma(z)$ discussed in the appendix to Chapter 2. The asymptotic behaviour of the gamma function (which can be derived using its integral definition and the saddle-point method) shows that this ratio has asymptotics matching (5.12).

Remark 5.4 If $r < 0$ or $s < 0$ then $h(x,y) = -r\log|x| - s\log|y|$ is unbounded from below on $\overline{\mathcal{D}}$, as $h(x,y)$ approaches negative infinity whenever $x \to 0$ or $y \to 0$. Going back to the bound $|f_{rn,sn}| \leq |a|^{-rn}|b|^{sn}$ for $(a,b) \in \mathcal{D}$ implies $f_{rn,sn}$ decays faster than any exponential function, meaning the sequence is eventually zero due to the constraints on asymptotic growth for rational diagonals. In particular, our asymptotic argument verifies the trivial fact that the power series expansion of $F(x,y)$ contains no terms whose indices are arbitrarily large with negative sign. Note that the set $\mathbb{R}^2_{\geq 0}$ contains all directions $\mathbf{r}$ such that some multiple $t\mathbf{r}$ with $t > 0$ lies in the Newton polygon of $1 - x - y$ at the origin; we generalize this observation in Proposition 5.7.

5.1.2 Asymptotics of Laurent Coefficients

Having studied coefficient asymptotics of the power series expansion of $F(x,y) = 1/(1-x-y)$, we now apply a similar analysis to the other Laurent expansions of F, derived in Chapter 3. Consider, for instance, the convergent Laurent expansion

$$\frac{1}{1-x-y} = \sum_{i,j\geq 0} \binom{i}{j} (-1)^{j+1} y^j x^{-i-1}$$

in the domain $\mathcal{D}_2 = \{(x,y) \in \mathbb{C}^2 : 1 + |y| < |x|\}$ (the final Laurent expansion is the same with x and y swapped). Although the rational function F, and thus its singular set $\mathcal{V}$, are unchanged, the singularities on the boundary of the domain of convergence now form the set $\mathcal{V} \cap \partial\mathcal{D}_2 = \{(1+t, -t) : t > 0\}$. The exponential growth of coefficients along a direction $\mathbf{r} = (r,s) \in \mathbb{R}^2$ is still determined by the minimum of the height function

$$h_{\mathbf{r}}(x,y) = -r\log|x| - s\log|y|$$

on $\overline{\mathcal{D}_2}$. Parameterizing $\mathcal{V} \cap \partial\mathcal{D}_2$ by $(x,y) = (1+t, -t)$ for $t > 0$, we want to minimize

$$g(t) = -r\log(1+t) - s\log(t).$$

From the series expansions

$$g(t) = s\log(t^{-1}) - rt + rt^2/2 + \cdots \qquad (t \to 0^+)$$
$$g(t) = -(r+s)\log(t) - rt^{-1} + rt^{-2}/2 + \cdots \qquad (t \to \infty),$$

we see that $h_{\mathbf{r}}(x, y)$ is unbounded from below, and thus the sequence $f_{nr,ns}$ is eventually zero, whenever $s < 0$ or $r + s > 0$.

Remark 5.5 Recalling Figure 3.2 and the discussion in Section 3.3.1 of Chapter 3, this Laurent expansion of $F(x, y)$ corresponds to the vertex $(1, 0)$ of the Newton polytope $\mathcal{N}(H)$ under the mapping described by Proposition 3.12. The set of directions (r, s) for which $(1, 0) + t(r, s)$ lies in the Newton polytope of $1 - x - y$ are exactly those where $s > 0$ and $r + s < 0$. Again, we return to this observation in Proposition 5.7 below.

Suppose $s > 0$ and $r + s < 0$, so that the minimum of $h_{\mathbf{r}}$ occurs on $\partial\mathcal{D}$. Solving $g'(t) = 0$ gives $(|x|, |y|) = (x_*, y_*) = \left(1 - \frac{s}{r+s}, \frac{s}{r+s}\right) = \left(\frac{r}{r+s}, \frac{s}{r+s}\right)$, with the point (x_*, y_*) being the only singularity on $\partial\mathcal{D}_2$ with this coordinate-wise modulus. Note this is the same formula as in the power series case, with different restrictions on r and s, and that $x_* > 0$ while $y_* < 0$. Below we will see that the most natural algebraic techniques compute a set of potential minimizers of $h_{\mathbf{r}}$ for all domains of convergence, after which one must filter the points to determine which are relevant to the specific Laurent expansion under consideration. This is one reason why a knowledge of Laurent expansions helps provide an understanding of our asymptotic methods, even when one only cares about power series expansions.

If $(x, y) \in \overline{\mathcal{D}_2}$ then $(x, z) \in \mathcal{D}_2$ for any z with $|z| < |y|$, since $1+|z| < 1+|y| \le |x|$. Thus, for any sufficiently small $\varepsilon > 0$,

$$f_{rn,sn} = I = \frac{1}{(2\pi i)^2} \int_{|x|=x_*} \left(\int_{|y|=|y_*|-\varepsilon} \frac{1}{1-x-y} \frac{dy}{y^{sn+1}} \right) \frac{dx}{x^{rn+1}}.$$

Again following the argument above, we introduce the integrals

$$I_{\text{loc}} = \frac{1}{(2\pi i)^2} \int_{\mathcal{N}} \left(\int_{|y|=|y_*|-\varepsilon} \frac{1}{1-x-y} \frac{dy}{y^{sn+1}} \right) \frac{dx}{x^{rn+1}}$$
$$I_{\text{out}} = \frac{1}{(2\pi i)^2} \int_{\mathcal{N}} \left(\int_{|y|=|y_*|+\varepsilon} \frac{1}{1-x-y} \frac{dy}{y^{sn+1}} \right) \frac{dx}{x^{rn+1}},$$

where $\mathcal{N} = \{|x| = x_* : \arg(x) \in (-\delta, \delta)\}$ for some sufficiently small $\delta > 0$. Defining $\mathcal{N}' = \{|x| = x_*\} \setminus \mathcal{N}$, we note that $I - I_{\text{loc}}$ and I_{out} grow slowly:

- If $a \in \mathbb{C}$ with $a = |x_*|$ and $a \neq x_*$ then the radius of convergence of the series $F(a, y)$ in y is larger than $|y_*|$. This shows the existence of $\tau > |y_*|$ such that

$$\int_{|y|=|y_*|-\varepsilon} \frac{1}{1-x-y} \frac{dy}{y^{sn+1}} = [y^{sn}]F(x, y) = O(\tau^{-ns})$$

  for all $x \in \mathcal{N}'$. The maximum modulus bound therefore implies

$$|I - I_{\text{loc}}| = O(x_*^{-rn} \tau^{-ns})$$

  will grow exponentially slower than the coefficient sequence under consideration.

- The maximum modulus bound implies $I_{\text{out}} = O\left(x_*^{-rn}(|y_*| + \varepsilon)^{-sn}\right)$, and $s > 0$.

Thus we may replace I by I_{loc}, subtract I_{out}, and compute a residue to approximate the coefficient $f_{nr,ns}$ by a Fourier-Laplace integral. In fact, we obtain the same Fourier-Laplace integral as in (5.11), the only difference from the power series case being the restrictions on (r, s). Again we may apply a saddle-point argument to obtain

$$\begin{aligned}\binom{-nr-1}{ns}(-1)^{ns+1} = f_{rn,sn} &\sim \left(\frac{r+s}{r}\right)^{rn}\left(\frac{r+s}{s}\right)^{sn}\left(\frac{r+s}{s}\right)\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-n\left(\frac{r(r+s)}{2s}\right)}d\theta \\ &= \left(\frac{r+s}{r}\right)^{rn}\left(\frac{r+s}{s}\right)^{sn}\left(\frac{r+s}{s}\right)\sqrt{\frac{s}{2r(r+s)\pi\, n}}\end{aligned}$$

for $s > 0$ and $r + s < 0$. Note that we *cannot* simplify, for instance, $\sqrt{s/(r+s)} = \sqrt{s}/\sqrt{r+s}$ because s and $r+s$ are negative. Making such invalid simplifications can result in an asymptotic approximation with incorrect sign.

Remark 5.6 Up to taking care with the leading sign, our asymptotic estimate respects the binomial identity $\binom{-nr-1}{ns}(-1)^{ns+1} = -\binom{nr+ns}{ns}$.

5.2 The Theory of Smooth ACSV

We now show that our approach for the central binomial coefficients generalizes to an amazing degree. This section lays out the basics of analytic combinatorics in several variables, culminating in Theorems 5.2 and 5.3, while the next section contains generalizations, further discussion, and results which help apply the theory to real applications. The necessary background on complex analysis in several variables, including singularities, Laurent expansions, polynomial amoebas, and diagonals of analytic functions, was covered in Chapter 3.

Fix a rational function

$$F(\mathbf{z}) = \frac{G(\mathbf{z})}{H(\mathbf{z})}$$

with G and H coprime polynomials over the complex numbers, and let $\mathcal{D}$ denote the domain of convergence of a Laurent expansion

$$F(\mathbf{z}) = \sum_{\mathbf{i}\in\mathbb{Z}^n} f_{\mathbf{i}}\mathbf{z}^{\mathbf{i}}$$

corresponding to a component $B = \text{Relog}^{-1}(\mathcal{D})$ of the amoeba complement $\text{amoeba}(H)^c$. We determine asymptotics of a sequence $f_{n\mathbf{r}}$ for indices $n\mathbf{r} \in \mathbb{Z}^d$ as $n \to \infty$. By deriving asymptotic results which vary smoothly as the direction $\mathbf{r}$ moves in some open set, we will be able to interpret 'asymptotics' of $f_{n\mathbf{r}}$ for $n\mathbf{r} \notin \mathbb{Z}^d$ through a limiting procedure made explicit in Section 5.3.3.

Unless otherwise noted a statement involving $f_{n\mathbf{r}}$ is interpreted to hold when $n \in \mathbb{N}$ and $n\mathbf{r} \in \mathbb{Z}^d$, so that the coefficient is well-defined.

We often assume that $\mathbf{r}$ contains no zero coordinate: if, say, $r_d = 0$ then the $\mathbf{r}$-diagonal of $F(\mathbf{z})$ is the $\hat{\mathbf{r}}$-diagonal of $[z_d^0]F(\mathbf{z})$, equal to the $\hat{\mathbf{r}}$-diagonal of $F(\hat{\mathbf{z}}, 0)$ when the series expansion F contains no terms with a negative exponent of z_d. Recalling the notation $\mathbb{R}_* = \mathbb{R} \setminus \{0\}$ and $\mathbb{C}_* = \mathbb{C} \setminus \{0\}$ from past chapters, this means $\mathbf{r} \in \mathbb{R}_*^d$. We take $F(\mathbf{z})$ to be a rational function in order to visualize our constructions using the amoeba of H, but our ultimate results will hold for ratios of analytic functions.

Our arguments revolve around the singularities of $F(\mathbf{z})$.

Definition 5.1 (singular variety and square-free part) Given complex functions $f_1, \ldots, f_d$, we write $\mathcal{V}(f_1, \ldots, f_d)$ to denote their set of common solutions in $\mathbb{C}^d$. The set of singularities of $F(\mathbf{z})$, denoted $\mathcal{V}$, is known as its *singular variety*; by Proposition 3.2 in Chapter 3 the singular variety $\mathcal{V}$ is the zero set $\mathcal{V}(H)$. Because $\mathcal{V}$ depends only on the irreducible polynomial factors of H, and not on their multiplicity, we let H^{s} denote the *square-free part* of H, equal to the product of its distinct irreducible polynomial factors over the complex numbers. The square-free part of H can be computed in any computer algebra system, either by completely factoring H or through more efficient specialized methods.

The minimal singularities of F, which are those on the boundary of the domain of convergence, form the set $\mathcal{V} \cap \partial\mathcal{D}$.

Remark 5.7 We note that $\mathcal{V} = \mathcal{V}(H) = \mathcal{V}(H^{\mathrm{s}})$. The algebraic properties of H^{s} better capture the geometry of $\mathcal{V}$, which will be useful in our calculations.

This chapter deals with multivariate generating functions where $\mathcal{V}$ is a complex manifold near the singularities of F which dictate coefficient asymptotics.

Definition 5.2 (smooth points) A *smooth point* of $\mathcal{V}$ is an element of $\mathcal{V}$ where at least one partial derivative of H^{s} does not vanish. If every point of $\mathcal{V}$ is smooth then we say $\mathcal{V}$ is *smooth*.

Although not required for most of our arguments, some differential geometry can help motivate and explain our constructions, and the necessary background material can be found in Griffiths and Harris [18, Ch. 0] if desired. For instance, the implicit function theorem (Proposition 3.1 in Chapter 3) implies $\mathcal{V}$ is a complex manifold in some neighbourhood of $\mathbf{w} \in \mathcal{V}$ whenever $\mathbf{w}$ is a smooth point.

Step 1: Bound Exponential Growth

As usual, we begin with the Cauchy integral formula

$$f_{n\mathbf{r}} = \frac{1}{(2\pi i)^d} \int_{T(\mathbf{w})} F(\mathbf{z}) \frac{d\mathbf{z}}{z_1^{nr_1+1} \cdots z_d^{nr_d+1}}, \tag{5.13}$$

where $T(\mathbf{w})$ is the polytorus defined by any $\mathbf{w} \in \mathcal{D}$. Applying the maximum modulus bound to the integral in (5.13) gives a bound

$$|f_{n\mathbf{r}}| = \left| \frac{1}{(2\pi \mathrm{i})^d} \int_{T(\mathbf{w})} F(\mathbf{z}) \frac{d\mathbf{z}}{z_1^{nr_1+1} \cdots z_d^{nr_d+1}} \right| \leq C_{\mathbf{w}} |w_1|^{-nr_1} \cdots |w_d|^{-nr_d}, \tag{5.14}$$

where $C_{\mathbf{w}} = \max_{\mathbf{z} \in T(\mathbf{w})} |F(\mathbf{z})|$ is finite. The part of this upper bound depending on n determines a notion of height for the singular points of F.

Definition 5.3 (height functions) The *height function* of F in the direction $\mathbf{r}$ is the function $h_{\mathbf{r}} \colon \mathbb{C}_*^d \to \mathbb{R}$ defined by

$$h_{\mathbf{r}}(\mathbf{z}) = -\sum_{j=1}^{d} r_j \log |z_j|.$$

Because it is often useful to take logarithms and picture the amoeba of H, we define the *logarithmic height function* $\tilde{h}_{\mathbf{r}} \colon \mathbb{R}^d \to \mathbb{R}$ by $\tilde{h}_{\mathbf{r}}(\mathbf{x}) = h_{\mathbf{r}}(e^{x_1}, \ldots, e^{x_n}) = -\mathbf{r} \cdot \mathbf{x}$. When the direction $\mathbf{r}$ is understood we drop subscripts and write h and $\tilde{h}$.

Equation (5.14) implies an exponential growth bound

$$\limsup_{n\to\infty} |f_{n\mathbf{r}}|^{1/n} \leq |w_1|^{-r_1} \cdots |w_d|^{-r_d} = e^{h_{\mathbf{r}}(\mathbf{w})} \tag{5.15}$$

for every point $\mathbf{w}$ in the closure $\overline{\mathcal{D}}$. When $h(\mathbf{z})$ is unbounded from below on $\overline{\mathcal{D}}$ the coefficient sequence $f_{n\mathbf{r}}$ decays super-exponentially and is thus eventually zero. Otherwise, either the minimum of $h(\mathbf{z})$ on $\overline{\mathcal{D}}$ is achieved and occurs on the boundary $\partial\mathcal{D}$ or, because $\tilde{h}$ is linear and the closed set $\overline{B}$ is convex, the minimum is not achieved and there is a limit direction of amoeba(H) which is normal to $\mathbf{r}$.

Because we need local behaviour of $F(\mathbf{z})$ near some of its singularities to capture the asymptotic growth of its coefficients, we require that $h_{\mathbf{r}}(\mathbf{z})$ achieves its minimum on $\overline{\mathcal{D}}$. In this chapter we work mainly under conditions which allow for easy verification of this assumption. Recall the logarithmic gradient map $\nabla_{\log} f = (z_1 f_{z_1}, \ldots, z_d f_{z_d})$ from Chapter 3. Proposition 3.13 in that chapter states that for any minimal point $\mathbf{w} \in \mathcal{V} \cap \partial\mathcal{D}$ there exists $\boldsymbol{\lambda} \in \mathbb{R}^d$ and $\tau \in \mathbb{C}$ such that $(\nabla_{\log} H^{\mathrm{s}})(\mathbf{w}) = \tau\boldsymbol{\lambda}$, which helps us find minimizers of the height function.

Proposition 5.1 *Let* $\mathbf{r} \in \mathbb{R}_*^d$. *If* $\mathbf{w} \in \mathcal{V} \cap \partial\mathcal{D}$ *and* $(\nabla_{\log} H^{\mathrm{s}})(\mathbf{w}) = \tau\mathbf{r}$ *with* $\tau \neq 0$ *then* $\mathbf{w}$ *is either a minimizer or a maximizer of the map* $h_{\mathbf{r}}(\mathbf{z})$ *on* $\overline{\mathcal{D}}$.

Remark 5.8 If $H(\mathbf{z}) = Q(\mathbf{z})^2$ for some polynomial Q then $(\nabla_{\log} H)(\mathbf{w}) = \mathbf{0}$ for all $\mathbf{w} \in \mathcal{V}$. This is why we work with the square-free part H^{s} instead of H.

Proof When $\tau \neq 0$ then, since $\mathbf{r}$ has no zero coordinate, $\mathbf{w}$ also has no zero coordinate. Proposition 3.13 in Chapter 3 thus implies that $\mathbf{r}$ is the normal vector of a support hyperplane $\mathcal{H}$ of B at $\mathrm{Relog}(\mathbf{w})$, meaning all points in B lie on one side of $\mathcal{H}$. Since the height function is represented by the linear function $\tilde{h}(\mathbf{x}) = -\mathbf{r} \cdot \mathbf{x}$

after taking the Relog map, Relog($\mathbf{w}$) is thus a minimizer or maximizer of $\tilde{h}(\mathbf{x})$ on B, depending on whether $\mathbf{r}$ points away from or towards B at Relog($\mathbf{w}$). The same relationship then holds between $\mathbf{w}$ and $h(\mathbf{z})$ on $\mathcal{D}$. Figure 5.2 above gives a visualization. □

Unfortunately, there may be a minimizer of $h_{\mathbf{r}}(\mathbf{z})$ on $\mathcal{V} \cap \partial\mathcal{D}$ which does not satisfy $(\nabla_{\log} H^{\mathrm{s}})(\mathbf{w}) = \tau\mathbf{r}$, even when $\mathcal{V}$ is smooth. Problems occur when the amoeba of H, which comes from projecting $\mathcal{V}$ to real space through the Relog map, does not accurately reflect the properties of $\mathcal{V}$ back in $\mathbb{C}^d$. Section 5.3 discusses these considerations in more detail, and gives criteria under which the sufficient conditions in Proposition 5.1 are necessary.

Step 2: Determine Contributing Points

Recalling that $\mathbf{w} \in \partial\mathcal{D}$ if and only if the intersection $T(\mathbf{w}) \cap \mathcal{V}$ is non-empty, our next goal is to get a better handle on the elements of $\mathcal{V}$ which correspond to minimizers of $h_{\mathbf{r}}(\mathbf{z})$. For any singularity $\mathbf{w} \in \mathcal{V}$, the vector $(\nabla_{\log} H^{\mathrm{s}})(\mathbf{w})$ is a scalar multiple of $\mathbf{r}$ if and only if the matrix

$$M = \begin{pmatrix} (\nabla_{\log} H^{\mathrm{s}})(\mathbf{w}) \\ \mathbf{r} \end{pmatrix} = \begin{pmatrix} w_1 H^{\mathrm{s}}_{z_1}(\mathbf{w}) & \cdots & w_d H^{\mathrm{s}}_{z_d}(\mathbf{w}) \\ r_1 & \cdots & r_d \end{pmatrix}$$

is rank deficient, where as usual subscripted variables refer to partial derivatives. This happens precisely when all 2×2 minors of M vanish, giving the system of equations

$$\begin{aligned} H^{\mathrm{s}}(\mathbf{w}) &= 0 \\ r_j w_1 H^{\mathrm{s}}_{z_1}(\mathbf{w}) - r_1 w_j H^{\mathrm{s}}_{z_j}(\mathbf{w}) &= 0 \qquad (2 \le j \le d). \end{aligned} \tag{5.16}$$

Note that any point $\mathbf{w} \in \mathcal{V}$ where $H^{\mathrm{s}}_{z_1}(\mathbf{w}) = \cdots = H^{\mathrm{s}}_{z_d}(\mathbf{w}) = 0$ satisfies (5.16). Because H^{s} is square-free, this happens when $\mathbf{w}$ is not a smooth point.

Definition 5.4 (smooth critical points) Equations (5.16) are called the *smooth critical point equations*, and any solution $\mathbf{w} \in \mathbb{C}^d_*$ of (5.16) where some $H^{\mathrm{s}}_{z_j}(\mathbf{w}) \neq 0$ is a *smooth critical point*.

Remark 5.9 (an optional perspective from differential geometry) Suppose $\mathcal{V}$ is smooth and let $\mathcal{V}_* = \mathcal{V} \cap \mathbb{C}^d_*$ denote the submanifold of points in $\mathcal{V}$ whose coordinates are non-zero. If we define the map $\phi(\mathbf{z}) = -\sum_{j=1}^{d} r_j \log z_j$ from $\mathcal{V}_*$ to $\mathbb{C}$, then the points where the differential of this analytic mapping of complex manifolds vanishes (known as 'critical points' in differential geometry settings) are precisely the points where the projection of $\nabla\phi$ to the tangent space of $\mathcal{V}_*$ is zero. Since the tangent space at $\mathbf{w} \in \mathcal{V}$ is the hyperplane whose normal is $(\nabla H^{\mathrm{s}})(\mathbf{w})$, the 'critical points' of $\phi\colon \mathcal{V}_* \to \mathbb{C}$ in the differential geometry sense form the solutions of (5.16) and thus match our definition of critical points.

Although we originally derived (5.16) by minimizing the upper bound (5.15) on exponential growth, the fact that these solutions are critical points of ϕ means we will ultimately reduce to integrals which can be asymptotically approximated by saddle-point methods. This serendipitous property is, in a sense, a result of the Cauchy-Riemann equations. The complex manifold $\mathcal{V}_* \subset \mathbb{C}^d$ defines an underlying real smooth manifold $\mathcal{W}_* \subset \mathbb{R}^{2d}$, obtained by setting $z_j = x_j + iy_j$ for real variables x_j and y_j. Since $h_\mathbf{r}$ is the real part of the map ϕ, the Cauchy-Riemann equations then imply that the critical points of the real-valued smooth map $h_\mathbf{r} \colon \mathcal{W}_* \to \mathbb{R}$ appearing in (5.15) are precisely the critical points of the complex analytic map $\phi \colon \mathcal{V}_* \to \mathbb{C}$. See Section 9.3 of Chapter 9 for more details on this interpretation.

By Proposition 5.1, a smooth critical point $\mathbf{w}$ is either a minimizer or maximizer of $h_\mathbf{r}$ on $\overline{\mathcal{D}}$, depending on whether $\mathbf{r}$ points away from or towards B at $\mathrm{Relog}(\mathbf{w})$, respectively. Only those critical points which are minimizers play a role in our arguments, so we make the following definition.

Definition 5.5 (smooth contributing points) Any smooth critical point $\mathbf{w} \in \mathbb{C}_*^d$ where $\mathbf{r}$ points away from B, i.e., where $\mathbf{x} \cdot \mathbf{r} < \mathrm{Relog}(\mathbf{w}) \cdot \mathbf{r}$ for any (and thus all) $\mathbf{x} \in B$ is called a *smooth contributing point*. Whenever $\mathcal{D}$ is the power series domain of convergence and $\mathbf{r}$ has positive coordinates, Proposition 3.6 in Chapter 3 implies that any minimal smooth critical point is a smooth contributing point.

By definition, a minimal smooth contributing point $\mathbf{w}$ is a minimizer of $h_\mathbf{r}$ on $\overline{\mathcal{D}}$.

Example 5.3 (Smooth Critical Points and the Amoeba Contour)

Figure 5.4 shows a sketch of the contour of $H(x, y) = 1 - x - y - 6xy - x^2y^2$ together with: the components $B_1, \dots, B_5$ of $\mathrm{amoeba}(H)^c$, the images of the smooth critical points under the Relog map when $\mathbf{r} = (1, 1)$, and the resulting support hyperplanes. The component $B_1 = \mathrm{Relog}(\mathcal{D}_1)$ corresponds to the power series expansion of $1/H(x, y)$; there is one critical point on $\partial\mathcal{D}_1$.

- Since $\mathbf{r}$ points away from B_1 at this point, it is a minimizer of $h_\mathbf{r}$ on $\overline{\mathcal{D}_1}$ and is contributing.

The component $B_2 = \mathrm{Relog}(\mathcal{D}_2)$ corresponds to a Laurent expansion of $1/H(x, y)$; there are *three* critical points on $\partial\mathcal{D}_2$.

- The single critical point whose image under Relog lies on the bottom left of B_2 is a maximizer of $h_\mathbf{r}(\mathbf{z})$ on $\overline{\mathcal{D}_2}$, since $\mathbf{r}$ points into B_2 at this point. Thus, this singularity is not contributing.
- There are two critical points in $\mathbb{C}^2$ with the same image $\boldsymbol{\sigma}$ under Relog, which lies on the top right of B_2. Since $\mathbf{r}$ points away from B_2 at $\boldsymbol{\sigma}$ these critical points are both minimizers of $h_\mathbf{r}$ on $\overline{\mathcal{D}_2}$, and thus also contributing points.

The component $B_3 = \mathrm{Relog}(\mathcal{D}_3)$ also has $\boldsymbol{\sigma}$ on its boundary, but $\mathbf{r}$ points into B_3 at $\boldsymbol{\sigma}$ so the corresponding critical points are not contributing. Neither of the remaining domains of convergence have critical points on their boundaries. The

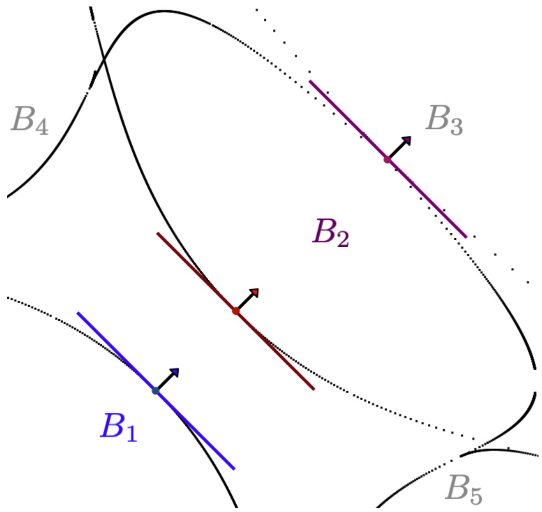


Fig. 5.4 Points on the contour of $H(x, y) = 1 - x - y - 6xy - x^2y^2$, with critical points and supporting hyperplanes shown. The components of the amoeba complement are labelled B_1 to B_5.

height function $h_{\mathbf{r}}$ is unbounded from below on $\mathcal{D}_3, \mathcal{D}_4$, and $\mathcal{D}_5$, meaning the sequence of diagonal coefficients $f_{n,n}$ is eventually zero in the corresponding Laurent expansions of $1/H(x, y)$. The critical point whose image under Relog lies on the bottom left of B_2 is not contributing for any Laurent expansion when $\mathbf{r} = (1, 1)$, but it is a contributing singularity for B_2 in the direction $\mathbf{r} = (-1, -1)$.

Remark 5.10 Although a smooth contributing point $\mathbf{w}$ is a minimizer of $h_{\mathbf{r}}$ on $\overline{\mathcal{D}}$, and a point where all partial derivatives of the function $h_{\mathbf{r}}$ restricted to $\mathcal{V}$ vanish, $h_{\mathbf{r}}$ never admits a local minimizer on $\mathcal{V}$ itself except for the trivial case when $h_{\mathbf{r}}$ is constant on $\mathcal{V}$. Essentially, the real dimension of $\mathcal{V}$ as a manifold is too large to allow h to have a local minimum (this can be formalized by a maximum modulus principle for analytic functions in several variables).

Step 3: Localize the Cauchy Integral and Compute a Residue

Minimal points, lying on $\partial\mathcal{D}$, are the singularities to which the domain of integration in the Cauchy integral (5.13) can be deformed arbitrarily close to. Critical points, on the other hand, are those around which the Cauchy integral can be asymptotically approximated. Contributing points are a subset of critical points where we will be able to introduce additional integrals required for our residue computations while adding only an asymptotically negligible error. Thus, the existence of minimal contributing points suggests that these points will be the ones around which local behaviour of F

will dictate coefficient asymptotics. The analysis is easiest when there are a finite number of such points.

Definition 5.6 (finite and strict minimality) A minimal point $\mathbf{w} \in \mathcal{V} \cap \mathcal{D}$ is called *finitely minimal* if $T(\mathbf{w}) \cap \mathcal{V}$ is finite, and *strictly minimal* if $T(\mathbf{w}) \cap \mathcal{V} = \{\mathbf{w}\}$. In other words, $\mathbf{w}$ is strictly minimal if it is minimal and no other singularity of F has the same coordinate-wise modulus.

Suppose first that $F(\mathbf{z})$ admits a strictly minimal smooth contributing point $\mathbf{w}$. Because $\mathbf{w}$ is a smooth point we may assume, without loss of generality, that $H_{z_d}(\mathbf{w}) \neq 0$. Let $\rho = |w_d|$ and $\mathcal{T} = T(\widehat{\mathbf{w}})$, where we recall the notation $\widehat{\mathbf{w}} = (w_1, \dots, w_{d-1})$. Since $\mathbf{w}$ is contributing, points with $\hat{\mathbf{z}} \in \mathcal{T}$ and $|z_d| = \rho - \varepsilon r_d$ lie in $\mathcal{D}$ for ε sufficiently small. In other words, if $r_d > 0$ we need to slightly decrease the modulus of the dth coordinate of $\mathbf{w}$ to move inside $\mathcal{D}$, while if $r_d < 0$ we need to *increase* the modulus of the dth coordinate to move inside $\mathcal{D}$.

The implicit function theorem (Proposition 3.1 in Chapter 3) together with strict minimality of $\mathbf{w}$ implies the existence of a neighbourhood $\mathcal{N}$ of $\widehat{\mathbf{w}}$ in $\mathcal{T}$, an analytic function $g : \mathcal{N} \to \mathbb{C}$, and $\delta > 0$ sufficiently small such that for $\hat{\mathbf{z}} \in \mathcal{N}$,

(i) $H(\mathbf{z}) = 0$ if and only if $z_d = g(\hat{\mathbf{z}})$;
(ii) $\rho - \delta < |g(\hat{\mathbf{z}})| < \rho + \delta$;
(iii) if $\hat{\mathbf{x}} \in \mathcal{T}$ and $\rho - \delta \le |y| \le \rho + \delta$ then $H(\hat{\mathbf{x}}, y) = 0$ only if $y = g(\hat{\mathbf{x}})$.

To guide our deformations, let

$$\delta_s = \operatorname{sgn}(r_d)\delta = \begin{cases} \delta & \text{if } r_d > 0 \\ -\delta & \text{if } r_d < 0 \end{cases}.$$

As above, we define

$$I = \frac{1}{(2\pi \mathrm{i})^d} \int_{\mathcal{T}} \left(\int_{|z_d| = \rho - \delta_s} F(\mathbf{z}) \frac{dz_d}{z_d^{nr_d+1}} \right) \frac{d\hat{\mathbf{z}}}{\hat{\mathbf{z}}^{n\hat{\mathbf{r}}+\mathbf{1}}}$$

$$I_{\text{loc}} = \frac{1}{(2\pi \mathrm{i})^d} \int_{\mathcal{N}} \left(\int_{|z_d| = \rho - \delta_s} F(\mathbf{z}) \frac{dz_d}{z_d^{nr_d+1}} \right) \frac{d\hat{\mathbf{z}}}{\hat{\mathbf{z}}^{n\hat{\mathbf{r}}+\mathbf{1}}}$$

$$I_{\text{out}} = \frac{1}{(2\pi \mathrm{i})^d} \int_{\mathcal{N}} \left(\int_{|z_d| = \rho + \delta_s} F(\mathbf{z}) \frac{dz_d}{z_d^{nr_d+1}} \right) \frac{d\hat{\mathbf{z}}}{\hat{\mathbf{z}}^{n\hat{\mathbf{r}}+\mathbf{1}}}$$

and

$$\begin{aligned} \chi &= I_{\text{loc}} - I_{\text{out}} \\ &= \frac{-1}{(2\pi \mathrm{i})^d} \int_{\mathcal{N}} \left(\int_{|z_d| = \rho + \delta_s} F(\mathbf{z}) \frac{dz_d}{z_d^{nr_d+1}} - \int_{|z_d| = \rho - \delta_s} F(\mathbf{z}) \frac{dz_d}{z_d^{nr_d+1}} \right) \frac{d\hat{\mathbf{z}}}{\hat{\mathbf{z}}^{n\hat{\mathbf{r}}+\mathbf{1}}}. \end{aligned} \quad (5.17)$$

Our first goal is to show that the coefficients of interest are well-approximated by χ.

Lemma 5.1 *If F admits a strictly minimal smooth contributing point $\mathbf{w} \in \mathbb{C}_*^d$ in the direction $\mathbf{r} \in \mathbb{R}_*^d$ then $|f_{n\mathbf{r}} - \chi| = O(\tau^n)$ for some $\tau < |\mathbf{w}^{-\mathbf{r}}|$.*

Proof Since $\mathbf{w}$ is contributing, $f_{n\mathbf{r}} = I$. Furthermore, since $\mathbf{w}$ is contributing and strictly minimal there exists $\varepsilon > 0$ such that for $\hat{\mathbf{z}} \in \mathcal{T} \setminus \mathcal{N}$ the univariate series, in t, for $F(\hat{\mathbf{z}}, t)$ converges when $|t| = \rho e^{\mathrm{sgn}(r_d)\varepsilon}$. Thus, for any $\hat{\mathbf{z}} \in \mathcal{T} \setminus \mathcal{N}$ the integral

$$\left| \int_{|z_d| = \rho - \delta_s} F(\mathbf{z}) \frac{dz_d}{z_d^{nr_d+1}} \right| = [t^{nr_d}] F(\hat{\mathbf{z}}, t)$$

grows exponentially slower than $\rho^{-nr_d} = |w_d|^{-nr_d}$, and the maximum modulus bound implies that

$$|I - I_{\mathrm{loc}}| = \left| \frac{1}{(2\pi \mathrm{i})^d} \int_{\mathcal{T} \setminus \mathcal{N}} \left(\int_{|z_d| = \rho - \delta_s} F(\mathbf{z}) \frac{dz_d}{z_d^{nr_d+1}} \right) \frac{d\hat{\mathbf{z}}}{\hat{\mathbf{z}}^{n\hat{\mathbf{r}}+\mathbf{1}}} \right|$$

grows exponentially slower than $|\mathbf{w}^{-n\mathbf{r}}|$. Finally, since $\mathbf{w}$ is contributing every point $\mathbf{z}$ in the domain of integration of I_{out} has $h_{\mathbf{r}}(\mathbf{z}) > h_{\mathbf{r}}(\mathbf{w})$, so the integral I_{out} also grows exponentially slower than $|\mathbf{w}^{-n\mathbf{r}}|$. The result then follows from the triangle inequality, as $|f_{n\mathbf{r}} - \chi| = |f_{n\mathbf{r}} - I_{\mathrm{loc}} + I_{\mathrm{out}}| \le |f_{n\mathbf{r}} - I_{\mathrm{loc}}| + |I_{\mathrm{out}}|$. □

The integral χ in Lemma 5.1 can be simplified with a residue computation. Furthermore, we can obtain a similar asymptotic expansion for a finitely minimal point by summing asymptotic contributions of this type.

Corollary 5.1 *Suppose $F(\mathbf{z})$ admits a finitely minimal smooth contributing point $\mathbf{w} \in \mathbb{C}_*^d$ in the direction $\mathbf{r} \in \mathbb{R}_*^d$. Suppose further that $T(\mathbf{w}) \cap \mathcal{V} = \{\mathbf{w}_1, \ldots, \mathbf{w}_r\}$ where each $\mathbf{w}_j$ is a smooth contributing point such that $H_{z_d}^{\mathrm{s}}(\mathbf{w}_j) \neq 0$. For each j let $\mathcal{N}_j$ denote a sufficiently small neighbourhood of $\mathbf{w}_j$ in $T(\mathbf{w})$ such that there exists an analytic parametrization $z_d = g_j(\hat{\mathbf{z}})$ of $\mathcal{V}$ for $\hat{\mathbf{z}} \in \mathcal{N}_j$. Then*

$$f_{n\mathbf{r}} = \sum_{j=1}^{r} \frac{-\operatorname{sgn}(r_d)}{(2\pi \mathrm{i})^{d-1}} \int_{\mathcal{N}_j} \frac{R_j(\hat{\mathbf{z}})}{\hat{\mathbf{z}}^{n\hat{\mathbf{r}}+\mathbf{1}}} d\hat{\mathbf{z}} + O(\tau^n) \tag{5.18}$$

for some $\tau < |\mathbf{w}^{-\mathbf{r}}|$, where

$$R_j(\hat{\mathbf{z}}) = \operatorname*{Res}_{z_d = g_j(\hat{\mathbf{z}})} \frac{G(\mathbf{z})}{z_d^{nr_d+1} H(\mathbf{z})}.$$

Proof Suppose first that $\mathbf{w} = \mathbf{w}_1$ is strictly minimal, so $T(\mathbf{w}) \cap \mathcal{V} = \{\mathbf{w}\}$. Then for any fixed $\hat{\mathbf{z}} \in \mathcal{N}$ the inner integrand of χ defined by (5.17) has a unique pole $z_d = g(\hat{\mathbf{z}})$ with $\rho - \delta_s \le |z_d| \le \rho + \delta_s$. Thus, the (univariate) Cauchy residue theorem implies

$$\frac{1}{2\pi \mathrm{i}} \left(\int_{|z_d| = \rho + \delta_s} F(\mathbf{z}) \frac{dz_d}{z_d^{nr_d+1}} - \int_{|z_d| = \rho - \delta_s} F(\mathbf{z}) \frac{dz_d}{z_d^{nr_d+1}} \right) = -\operatorname{sgn}(r_d) R_1(\hat{\mathbf{z}}),$$

and the result holds. If $\mathbf{w}$ is finitely minimal then existence of the $\mathcal{N}_j$ and g_j follows from the implicit function theorem. Repeating the above argument for the strictly minimal case with $\mathcal{N}$ replaced by the union of the $\mathcal{N}_j$ shows that χ is a sum of residue integrals, each of which simplifies to the stated form. □

Remark 5.11 Corollary 5.1 assumes $H^{\mathrm{s}}_{z_d}(\mathbf{w}_j) \neq 0$ for each $1 \leq j \leq r$, meaning the solutions of $H(\mathbf{z}) = 0$ can be locally parametrized by the same coordinates near each $\mathbf{w}_j$. If, more generally, the partial derivative $H^{\mathrm{s}}_{z_k}(\mathbf{w}_j) \neq 0$ then all occurrences of r_d and z_d in the jth summand of (5.18) should be replaced by r_k and z_k, respectively.

If $\mathbf{w}$ is one of the smooth minimal critical points in Corollary 5.1 then we may write $H(\mathbf{z}) = (z_d - g(\hat{\mathbf{z}}))^p q(\mathbf{z})$ for some positive integer $p \geq 1$ and analytic function $q(\mathbf{z})$ in a neighbourhood of $\mathbf{w}$ with $q(\mathbf{w}) \neq 0$. The integer p is the order of the pole $z_d = g(\hat{\mathbf{z}})$, and may be calculated as the smallest positive integer such that $(\partial^p H/\partial z_d^p)(\mathbf{w}) \neq 0$. If H is square-free then $p = 1$ and

$$R(\hat{\mathbf{z}}) = \operatorname*{Res}_{z_d = g(\hat{\mathbf{z}})} \frac{G(\mathbf{z})}{z_d^{nr_d+1}(z_d - g(\hat{\mathbf{z}}))q(\mathbf{z})} = \frac{G(\hat{\mathbf{z}}, g(\hat{\mathbf{z}}))}{g(\hat{\mathbf{z}})^{nr_d+1} q(\hat{\mathbf{z}}, g(\hat{\mathbf{z}}))} = \frac{G(\hat{\mathbf{z}}, g(\hat{\mathbf{z}}))}{g(\hat{\mathbf{z}})^{nr_d+1} H_{z_d}(\hat{\mathbf{z}}, g(\hat{\mathbf{z}}))}.$$

The general formula is messier, but still explicit. To simplify notation, we write ∂_d^k to denote the partial derivative operator which takes a differentiable function $f(\mathbf{z})$ and returns $\partial_d^k(f) = (\partial^k f/\partial z_d^k)(\mathbf{z})$, where $\partial_d^0(f) = f(\mathbf{z})$.

Lemma 5.2 *If the order of the pole $z_d = g(\hat{\mathbf{z}})$ near $\mathbf{w}$ is p then for all $\hat{\mathbf{z}} \in \mathcal{N}$*

$$R(\hat{\mathbf{z}}) = (-1)^{p+1} g(\hat{\mathbf{z}})^{-nr_d - p} \sum_{k=0}^{p-1} \frac{(-1)^k \; (nr_d + p - 1 - k)_{(p-k-1)}}{k!(p-k-1)!} \, \mathfrak{R}_j(\hat{\mathbf{z}}), \tag{5.19}$$

where $(a)_{(b)} = a(a-1)\cdots(a-b+1)$ denotes the falling factorial for $a \in \mathbb{R}$ and $b \in \mathbb{N}$, and

$$\mathfrak{R}_j(\hat{\mathbf{z}}) = g(\hat{\mathbf{z}})^k \lim_{z_d \to g(\hat{\mathbf{z}})} \partial_d^k\Big((z_d - g(\hat{\mathbf{z}}))^p F(\mathbf{z})\Big).$$

Proof Lemma 2.4 in the appendix to Chapter 2 expresses the residue $R(\hat{\mathbf{z}})$ as the limit of an order $p-1$ derivative. The product rule for derivatives then implies

$$\begin{aligned} R(\hat{\mathbf{z}}) &= \frac{1}{(p-1)!} \lim_{z_d \to g(\hat{\mathbf{z}})} \partial_d^{p-1}\left(z_d^{-nr_d-1}\,(z_d - g(\hat{\mathbf{z}}))^p F(\mathbf{z})\right) \\ &= \frac{1}{(p-1)!} \lim_{z_d \to g(\hat{\mathbf{z}})} \sum_{k=0}^{p-1} \binom{p-1}{k} \frac{(-1)^{p-1-k}\,(nr_d + p - 1 - k)_{(p-1-k)}}{z_d^{nr_d+p-k}} \;\partial_d^k\Big((z_d - g(\hat{\mathbf{z}}))^p F(\mathbf{z})\Big), \end{aligned}$$

from which (5.19) is obtained by algebraic manipulation. □

Step 4: Find Asymptotics using the Saddle-Point Method

To convert the expression (5.18) into a form more suitable for asymptotic approximation, we make the change of coordinates $z_j = w_j e^{i\theta_j}$ for $j = 1, \dots, d-1$. Let $\mathcal{N}' \subset \mathbb{R}^{d-1}$ be the image of $\mathcal{N}$ under this change of variables, which will be a neighbourhood of the origin (and can be taken as any sufficiently small neighbourhood of the origin without affecting our asymptotic statements). If $\boldsymbol{\theta} = (\theta_1, \dots, \theta_{d-1})$ we define the coordinate-wise product $\widehat{\mathbf{w}}e^{i\boldsymbol{\theta}} = \left(w_1 e^{i\theta_1}, \dots, w_{d-1} e^{i\theta_{d-1}}\right)$ to lighten notation. Our next result follows directly from Corollary 5.1 and this change of variables.

Proposition 5.2 *Suppose that* $\mathbf{r} \in \mathbb{R}^d_*$ *and* F *admits a strictly minimal smooth contributing point* $\mathbf{w}$ *such that* $H^{\mathrm{s}}_{z_d}(\mathbf{w}) \neq 0$. *Let* $g(\hat{\mathbf{z}})$ *be the analytic parameterization of* z_d *in terms of* $\hat{\mathbf{z}}$ *in any sufficiently small neighbourhood of* $\mathbf{w}$. *If the order of the pole* $z_d = g(\hat{\mathbf{z}})$ *is* p *then for any sufficiently small neighbourhood* $\mathcal{N}'$ *of the origin in* $\mathbb{R}^{d-1}$ *there exists an* $\varepsilon > 0$ *such that*

$$f_{n\mathbf{r}} = \chi + O\left((|\mathbf{w}^{\mathbf{r}}| + \varepsilon)^{-n}\right),$$

where

$$\chi = \frac{\mathbf{w}^{-n\mathbf{r}}}{(2\pi)^{d-1}} \int_{\mathcal{N}'} A_n(\boldsymbol{\theta})\, e^{-nr_d\phi(\boldsymbol{\theta})} d\boldsymbol{\theta} \tag{5.20}$$

for

$$\begin{aligned} A_n(\boldsymbol{\theta}) &= \frac{(-1)^p \operatorname{sgn}(r_d)}{g\left(\widehat{\mathbf{w}}e^{i\boldsymbol{\theta}}\right)^p}\, \Gamma_n(\widehat{\mathbf{w}}e^{i\boldsymbol{\theta}}) \\ \phi(\boldsymbol{\theta}) &= \log\left(\frac{g\left(\widehat{\mathbf{w}}e^{i\boldsymbol{\theta}}\right)}{g(\widehat{\mathbf{w}})}\right) + i(\hat{\mathbf{r}} \cdot \boldsymbol{\theta})/r_d \end{aligned} \tag{5.21}$$

and

$$\Gamma_n(\mathbf{z}) = \sum_{j=0}^{p-1} \frac{(-1)^j\ (nr_d + p - 1 - k)_{(p-j-1)}}{j!(p-j-1)!}\, g(\hat{\mathbf{z}})^j \lim_{z_d \to g(\hat{\mathbf{z}})} \partial_d^j\Big((z_d - g(\hat{\mathbf{z}}))^p F(\mathbf{z})\Big).$$

The function $\Gamma_n(\mathbf{z})$ *is a polynomial of degree* $p-1$ *in* n, *whose coefficients are analytic functions in* $\hat{\mathbf{z}}$. *If* $z_d = g(\hat{\mathbf{z}})$ *is a simple pole then* $p = 1$ *and*

$$\Gamma_n(\mathbf{z}) = \frac{G(\hat{\mathbf{z}}, g(\hat{\mathbf{z}}))}{H_{z_d}(\hat{\mathbf{z}}, g(\hat{\mathbf{z}}))}$$

is independent of n.

When $\mathbf{w}$ *is finitely minimal and each point* $\mathbf{w}_1, \dots, \mathbf{w}_r$ *of* $T(\mathbf{w}) \cap \mathcal{V}$ *is a smooth contributing point with* $H^{\mathrm{s}}_{z_d}(\mathbf{w}) \neq 0$ *then, up to an error which is exponentially smaller than* $\mathbf{w}^{-n\mathbf{r}}$, *the coefficient sequence* $f_{n\mathbf{r}}$ *is a sum of integrals of the form* (5.20) *with* $\mathbf{w}$ *replaced by* $\mathbf{w}_j$ *and* $g(\hat{\mathbf{z}})$ *replaced by a parameterization* $g_j(\hat{\mathbf{z}})$ *of* z_d *near* $\mathbf{w}_j$.

To compute asymptotics of the integral in (5.20) we use the following result, a multivariate generalization of the saddle-point method. Recall that the Hessian of a differentiable function $g(\boldsymbol{\theta})$ from $\mathbb{R}^k$ to $\mathbb{C}$ is the $k \times k$ matrix whose (i, j)th entry is the second partial derivative $g_{\theta_i\theta_j} = \partial^2 g/\partial\theta_i\partial\theta_j$.

Proposition 5.3 (Asymptotics of Nondegenerate Multivariate Fourier-Laplace Integrals) *Suppose that the functions $P(\boldsymbol{\theta})$ and $\psi(\boldsymbol{\theta})$ from $\mathbb{R}^k$ to $\mathbb{C}$ are analytic in a neighbourhood $\mathcal{N}$ of the origin, and let $\mathcal{H}$ be the Hessian of ψ evaluated at the point $\boldsymbol{\theta} = \mathbf{0}$. If $r \in \mathbb{R}$ is nonzero and*

- *$\psi(\mathbf{0}) = 0$ and $(\nabla\psi)(\mathbf{0}) = \mathbf{0}$;*
- *the origin is the only point of $\mathcal{N}$ where $\nabla\psi$ is 0;*
- *$\mathcal{H}$ is non-singular (has non-zero determinant);*
- *the real part of $r\psi(\boldsymbol{\theta})$ is non-negative on $\mathcal{N}$,*

then for any nonnegative integer M there exist computable constants $K_0, \dots, K_M$ such that

$$\int_{\mathcal{N}} P(\boldsymbol{\theta})\, e^{-nr\psi(\boldsymbol{\theta})} d\boldsymbol{\theta} = \left(\frac{2\pi}{n}\right)^{k/2} \det(r\mathcal{H})^{-1/2} \sum_{j=0}^{M} K_j (rn)^{-j} + O\left(n^{-M-1}\right), \quad (5.22)$$

where the square-root of the determinant is the product of the principal branch square-roots of the eigenvalues of $r\mathcal{H}$ (which all have non-negative real part).

The constant $K_0 = P(\mathbf{0})$ and if $P(\boldsymbol{\theta})$ vanishes to order $L \geq 1$ at the origin then (at least) the constants $K_0, \dots, K_{\lfloor \frac{L}{2} \rfloor}$ are all zero. More precisely, define the differential operator

$$\mathcal{E} = -\sum_{1 \leq i,j \leq k} \left(\mathcal{H}^{-1}\right)_{ij} \partial_i \partial_j$$

where ∂_j denotes differentiation with respect to the variable θ_j and $\mathcal{H}^{-1}$ is the inverse matrix of $\mathcal{H}$. Let

$$\tilde{\psi}(\boldsymbol{\theta}) = \psi(\boldsymbol{\theta}) - (1/2)\boldsymbol{\theta} \cdot \mathcal{H} \cdot \boldsymbol{\theta}^T,$$

which is a scalar function vanishing to order 3 at the origin. Then

$$K_j = (-1)^j \sum_{0 \leq \ell \leq 2j} \left. \frac{\mathcal{E}^{\ell+j}\left(P(\boldsymbol{\theta})\, \tilde{\psi}(\boldsymbol{\theta})^\ell\right)}{2^{\ell+j}\,\ell!(\ell+j)!} \right|_{\boldsymbol{\theta}=\mathbf{0}}. \quad (5.23)$$

Due to the order of vanishing of $\tilde{\psi}$, to determine K_j one only needs to calculate evaluations at $\mathbf{0}$ of the derivatives of P of order at most $2j$ and the derivatives of ψ of order at most $2j + 2$. The hidden constant in the error term $O\left(n^{-M-1}\right)$ of (5.22) varies continuously with the values of the partial derivatives of P and ψ at the origin.

Proof Proposition 5.3 follows from Theorem 2.3 of Pemantle and Wilson [31]. If the final condition of the proposition is strengthened to require that the real part of $r\psi(\boldsymbol{\theta})$ be positive on $\mathcal{N} \setminus \{\mathbf{0}\}$, which holds in the situations we encounter, then the stated conclusions follow from the more elementary Theorem 4.1 of [31]. Hörmander [20, Theorem 7.7.5] gives the explicit formula for higher order terms when A

has compact support (this restriction can be removed when A and ϕ satisfy certain growth conditions which hold in most situations we encounter). Section 5 of [31] shows that the expressions of Hörmander are valid under our stated assumptions. □

We now go through the criteria required to apply Proposition 5.3 to the integral expression in Proposition 5.2. First, we show criticality is equivalent to the gradient of ϕ vanishings.

Lemma 5.3 *Suppose* $\mathbf{w} \in \mathcal{V}$ *is any point such that* $H^{\mathrm{s}}_{z_d}(\mathbf{w}) \neq 0$*, and let* $z_d = g(\hat{\mathbf{z}})$ *be the parameterization of* z_d *on* $\mathcal{V}$ *near* $\mathbf{w}$*. Then the gradient of the function* $\phi(\boldsymbol{\theta})$ *in* (5.21) *equals* $\mathbf{0}$ *at* $\boldsymbol{\theta} = \mathbf{0}$ *if and only if* $\mathbf{w}$ *is a smooth critical point.*

Proof Fix $1 \leq k \leq d-1$. Differentiating the equation $H^{\mathrm{s}}(\hat{\mathbf{z}}, g(\hat{\mathbf{z}})) = 0$ with respect to z_k using the chain rule implies $g_{z_k}(\hat{\mathbf{z}}) = -H_{z_k}(\hat{\mathbf{z}}, g(\hat{\mathbf{z}}))/H_{z_d}(\hat{\mathbf{z}}, g(\hat{\mathbf{z}}))$ for $\hat{\mathbf{z}}$ in a neighbourhood of $\widehat{\mathbf{w}}$. Thus, the derivative of ϕ with respect to θ_k is

$$\phi_{\theta_k}(\boldsymbol{\theta}) = \frac{g_{z_k}\left(\widehat{\mathbf{w}}e^{\mathrm{i}\boldsymbol{\theta}}\right)}{g\left(\widehat{\mathbf{w}}e^{\mathrm{i}\boldsymbol{\theta}}\right)} \mathrm{i}w_k e^{\mathrm{i}\theta_k} + \mathrm{i}\frac{r_k}{r_d} = \frac{-H_{z_k}\left(\widehat{\mathbf{w}}e^{\mathrm{i}\boldsymbol{\theta}}, g\left(\widehat{\mathbf{w}}e^{\mathrm{i}\boldsymbol{\theta}}\right)\right)}{g\left(\widehat{\mathbf{w}}e^{\mathrm{i}\boldsymbol{\theta}}\right) H_{z_d}\left(\widehat{\mathbf{w}}e^{\mathrm{i}\boldsymbol{\theta}}, g\left(\widehat{\mathbf{w}}e^{\mathrm{i}\boldsymbol{\theta}}\right)\right)} \mathrm{i}w_k e^{\mathrm{i}\theta_k} + \mathrm{i}\frac{r_k}{r_d}.$$

Substituting $\boldsymbol{\theta} = \mathbf{0}$ and using $g(\widehat{\mathbf{w}}) = w_d$ shows that vanishing of all partial derivatives of ϕ at the origin is equivalent to $\mathbf{w}$ satisfying the smooth critical point equations. □

Next, we show minimality implies the real part of $r_d\phi$ is non-negative.

Lemma 5.4 *Suppose* $\mathbf{w} \in \mathcal{V}$ *is a minimal contributing point such that* $H^{\mathrm{s}}_{z_d}(\mathbf{w}) \neq 0$*, and let* $z_d = g(\hat{\mathbf{z}})$ *be a parameterization of* z_d *on* $\mathcal{V}$ *near* $\mathbf{w}$*. If* $\phi(\boldsymbol{\theta})$ *is defined by* (5.21) *then the real part* $\Re(r_d\phi(\boldsymbol{\theta})) \geq 0$ *for real* $\boldsymbol{\theta}$*. If* $\mathbf{w}$ *is finitely minimal then* $\Re(r_d\phi(\boldsymbol{\theta})) > 0$ *for* $\boldsymbol{\theta}$ *in any sufficiently small neighbourhood* $\mathcal{N}'$ *of the origin.*

Proof The real part of $r_d\phi$ can be expressed as

$$\Re(r_d\phi) = r_d \log\left|g\left(\widehat{\mathbf{w}}e^{\mathrm{i}\boldsymbol{\theta}}\right)\right| - r_d \log|g(\widehat{\mathbf{w}})|,$$

which is non-negative if and only if $|g(\widehat{\mathbf{w}})|^{r_d} \leq \left|g\left(\widehat{\mathbf{w}}e^{\mathrm{i}\boldsymbol{\theta}}\right)\right|^{r_d}$. Since $g\left(\widehat{\mathbf{w}}e^{\mathrm{i}\boldsymbol{\theta}}\right)$ gives the dth coordinate of a point on $\mathcal{V}$ whose first $d-1$ coordinates lie in $T(\widehat{\mathbf{w}})$, this inequality holds whenever $\mathbf{w}$ is minimal and contributing. If $\mathbf{w}$ is finitely minimal then no other points in any sufficiently small neighbourhood of $\mathbf{w}$ in $\mathcal{V}$ have the same coordinate-wise modulus as $\mathbf{w}$, so the inequality is strict. □

The Hessian of ϕ at the origin may, in general, be singular.

Definition 5.7 (nondegenerate critical points) The critical point $\mathbf{w}$ is *nondegenerate* if the Hessian matrix $\mathcal{H}$ of ϕ at the origin is nonsingular.

We require that all minimal contributing points are nondegenerate. Applying the chain rule, and using that $\mathbf{w}$ is a critical point in the direction $\mathbf{r}$, allows for an explicit determination of the Hessian matrix of ϕ at the origin in terms of the partial derivatives of $H(\mathbf{z})$, ultimately giving the following result.

Lemma 5.5 *Suppose* $\mathbf{w}$ *is a smooth critical point where* $H^{s}_{z_d}(\mathbf{w}) \neq 0$ *and let* $z_d = g(\hat{\mathbf{z}})$ *be the parameterization of* z_d *on* $\mathcal{V}$ *near* $\mathbf{w}$*. If, for* $1 \leq i, j \leq d$*, we define*

$$U_{i,j} = \frac{w_i w_j H^{s}_{z_i z_j}(\mathbf{w})}{w_d H^{s}_{z_d}(\mathbf{w})} \qquad \text{and} \qquad V_i = \frac{w_i H^{s}_{z_i}(\mathbf{w})}{w_d H^{s}_{z_d}(\mathbf{w})} = \frac{r_i}{r_d} \tag{5.24}$$

then the $(d-1)\times(d-1)$ *Hessian matrix* $\mathcal{H}$ *of the function* ϕ *in* (5.21) *evaluated at the origin has* (i, j)*th entry*

$$\mathcal{H}_{i,j} = \begin{cases} V_i V_j + U_{i,j} - V_j U_{i,d} - V_i U_{j,d} + V_i V_j U_{d,d} & : i \neq j \\ V_i + V_i^2 + U_{i,i} - 2V_i U_{i,d} + V_i^2 U_{d,d} & : i = j \end{cases} \tag{5.25}$$

Nondegeneracy of a critical point $\mathbf{w}$ relates behaviour of $h_{\mathbf{r}}$ to the geometry of the singular set $\mathcal{V}$ near $\mathbf{w}$. This regulates the behaviour of nondegenerate critical points.

Lemma 5.6 *There are a finite number of nondegenerate smooth critical points in any direction* $\mathbf{r} \in \mathbb{R}^d_*$*. If* $\mathbf{w} = \mathbf{w}(\mathbf{r}) \in \mathbb{C}^d_*$ *is a nondegenerate smooth critical point in the direction* $\mathbf{r} \in \mathbb{R}^d_*$ *then* $\mathbf{w}$ *is isolated among the set of smooth critical points in the direction* $\mathbf{r}$*, and* $\mathbf{w}(\mathbf{r})$ *varies smoothly as* $\mathbf{r}$ *moves in some sufficiently small neighbourhood of* $\mathbb{R}^d_*$*. If* $H(\mathbf{z})$ *has real coefficients and* $\mathbf{w} \in \mathbb{R}^d_*$ *then when* $\mathbf{r}$ *moves in a sufficiently small neighbourhood each* $\mathbf{w}(\mathbf{r}) \in \mathbb{R}^d_*$*.*

Proof The fact that a nondegenerate critical point is isolated follows directly from the complex Morse lemma, which states that after an analytic change of coordinates ϕ can be written in a $\mathcal{V}_*$-neighbourhood of $\mathbf{w}$ as a sum of squares $\phi(\mathbf{w})+u_1^2+\cdots+u_{d-1}^2$ (see Ebeling [15, Prop. 3.15 and Cor. 3.3]). Bézout's theorem [8, Thm. 8.2] bounds the number of isolated solutions of a polynomial system, which form its 'zero-dimensional component'. Since any nondegenerate smooth critical point is an isolated solution of the smooth critical point equations (5.16), such points are finite in number.

Because $\mathbf{w}$ is a smooth point we may assume, without loss of generality, that $H^{s}_{z_d}(\mathbf{w}) \neq 0$, and we let $z_d = g(\hat{\mathbf{z}})$ be the parameterization of z_d on $\mathcal{V}$ near $\mathbf{w}$. By a multivariate version of the implicit function theorem [21, Thm. 2.1.2], to show that $\mathbf{w}(\mathbf{r})$ varies smoothly with $\mathbf{r}$ it is enough to show that the Jacobian of the smooth critical point system

$$r_k z_1 H_{z_1}(\hat{\mathbf{z}}, g(\hat{\mathbf{z}})) - r_1 z_k H_{z_k}(\hat{\mathbf{z}}, g(\hat{\mathbf{z}})) = 0 \qquad (2 \leq k \leq d)$$

with respect to $\hat{\mathbf{z}}$ is non-singular at $\widehat{\mathbf{w}}$ (to lighten notation we now suppose H is square-free so that $H^{s} = H$, otherwise replace H with H^{s} throughout). The Hessian of the map $\phi(\boldsymbol{\theta}) = r_d \log g(\widehat{\mathbf{w}} e^{i\boldsymbol{\theta}}) - r_d \log g(\widehat{\mathbf{w}}) + i(\boldsymbol{\theta} \cdot \hat{\mathbf{r}})$ at $\boldsymbol{\theta} = \mathbf{0}$ is the Jacobian of the gradient map

$$(\nabla\phi)(\boldsymbol{\theta}) = \frac{i}{L_d(\boldsymbol{\theta})}\Big(r_d L_1(\boldsymbol{\theta}) - r_1 L_d(\boldsymbol{\theta}), \dots, r_d L_{d-1}(\boldsymbol{\theta}) - r_{d-1} L_d(\boldsymbol{\theta})\Big), \tag{5.26}$$

where

$$L_j(\boldsymbol{\theta}) = w_j e^{\mathrm{i}\theta_j} H_{z_j}\left(\widehat{\mathbf{w}}e^{\mathrm{i}\theta}, g\left(\widehat{\mathbf{w}}e^{\mathrm{i}\theta}\right)\right) \qquad (1 \le j \le d-1)$$
$$L_d(\boldsymbol{\theta}) = g\left(\widehat{\mathbf{w}}e^{\mathrm{i}\theta}\right) H_{z_d}\left(\widehat{\mathbf{w}}e^{\mathrm{i}\theta}, g\left(\widehat{\mathbf{w}}e^{\mathrm{i}\theta}\right)\right).$$

The vector in (5.26) is, up to a multiplicative factor and change of coordinates $z_j = w_j e^{\mathrm{i}\theta_j}$, the smooth critical point system. The chain rule implies this change of variables does not affect singularity of the Jacobian. Furthermore, the entries of $(\nabla\phi)(\boldsymbol{\theta})$ vanish when $\boldsymbol{\theta} = \mathbf{0}$ so the factor of $\mathrm{i}/L_d(\boldsymbol{\theta})$ simply multiplies the Hessian of the smooth critical point system by the non-zero constant $\mathrm{i}/(w_d H_{z_d}(\mathbf{w}))$. Thus, the Jacobian determinant of the smooth critical point system at a critical point $\mathbf{w}$ is, up to a non-zero factor, the Hessian determinant of ϕ at $\boldsymbol{\theta} = \mathbf{0}$, which by definition is nonzero when $\mathbf{w}$ is non-degenerate. When all quantities are real, the implicit function theorem implies the critical point $\mathbf{w}(\mathbf{r})$ is also real. □

We are now ready to state the main asymptotic results of this section. Theorems 5.2 and 5.3 follow directly from applying Proposition 5.3 to Proposition 5.2 under our assumptions, taking Lemmas 5.3 to 5.6 into account. The easiest, and most common, situation is when the minimal contributing point is a simple pole. We state this result separately in order to be as explicit as possible.

Theorem 5.2 (Smooth Asymptotics for Simple Poles) *Suppose that the rational function $F(\mathbf{z}) = G(\mathbf{z})/H(\mathbf{z})$ admits a nondegenerate strictly minimal smooth contributing point $\mathbf{w} \in \mathbb{C}_*^d$ in the direction $\mathbf{r} \in \mathbb{R}_*^d$, such that $H_{z_d}(\mathbf{w}) \neq 0$. Then for any nonnegative integer M there exist computable constants $C_0, \ldots, C_M$ such that*

$$f_{n\mathbf{r}} = \mathbf{w}^{-n\mathbf{r}} n^{(1-d)/2} \frac{(2\pi)^{(1-d)/2}}{\sqrt{\det(r_d\mathcal{H})}} \left(\sum_{j=0}^{M} C_j (r_d n)^{-j} + O\left(n^{-M-1}\right)\right), \tag{5.27}$$

where the matrix $\mathcal{H}$ is defined in (5.25), the square-root of the determinant is given by the product of the principal branch square-roots of the eigenvalues of $r_d\mathcal{H}$, and C_j equals the constant K_j defined by (5.23) when

$$P(\boldsymbol{\theta}) = \frac{-\operatorname{sgn}(r_d)\, G\left(\widehat{\mathbf{w}}e^{\mathrm{i}\boldsymbol{\theta}}, g\left(\widehat{\mathbf{w}}e^{\mathrm{i}\boldsymbol{\theta}}\right)\right)}{g\left(\widehat{\mathbf{w}}e^{\mathrm{i}\boldsymbol{\theta}}\right) H_{z_d}\left(\widehat{\mathbf{w}}e^{\mathrm{i}\boldsymbol{\theta}}, g\left(\widehat{\mathbf{w}}e^{\mathrm{i}\boldsymbol{\theta}}\right)\right)}$$

and

$$\psi(\boldsymbol{\theta}) = \log\left(\frac{g\left(\widehat{\mathbf{w}}e^{\mathrm{i}\boldsymbol{\theta}}\right)}{g(\widehat{\mathbf{w}})}\right) + \mathrm{i}(\hat{\mathbf{r}} \cdot \boldsymbol{\theta})/r_d.$$

The leading constant C_0 in this series has the value

$$C_0 = \frac{-\operatorname{sgn}(r_d) G(\mathbf{w})}{w_d\, H_{z_d}(\mathbf{w})},$$

which is nonzero whenever $G(\mathbf{w}) \neq 0$. The asymptotic expansion (5.27) holds uniformly in neighbourhoods $\mathcal{R} \subset \mathbb{R}_*^d$ of $\mathbf{r}$ where there is a smoothly varying nondegenerate strictly minimal contributing point such that H_{z_d} does not vanish. In *other words, if $w(\mathbf{s}) \in \mathbb{C}_*^d$ is the contributing point in direction $\mathbf{s} \in \mathcal{R}$ then there exists $B > 0$ such that*

$$\left| f_{n\mathbf{s}} - \mathbf{w}(\mathbf{s})^{-n\mathbf{s}} n^{(1-d)/2} \frac{(2\pi)^{(1-d)/2}}{\sqrt{\det(s_d \mathcal{H}^{\mathbf{s}})}} \sum_{j=0}^{M} C_j^{\mathbf{s}} (s_d n)^{-j} \right| \le B \left| \mathbf{w}(\mathbf{s})^{-n\mathbf{s}} \right| \; n^{(1-d)/2-M-1}$$

for all $n \in \mathbb{N}$ and $\mathbf{s} \in \mathcal{R}$ with $n\mathbf{s} \in \mathbb{Z}^d$, where $C_j^{\mathbf{s}}$ and $\mathcal{H}^{\mathbf{s}}$ are calculated from the contributing point $w(\mathbf{s})$ as above.

Remark 5.12 Although the constants in (5.27) are defined in terms of partial derivatives of the parameterization $g(\hat{\mathbf{z}})$ for z_d on $\mathcal{V}$, implicitly differentiating the equation $H(\hat{\mathbf{z}}, g(\hat{\mathbf{z}})) = 0$ allows one to determine the partial derivatives of g at $\widehat{\mathbf{w}}$ from the partial derivatives of H at $\mathbf{w}$. Thus, the only information needed to determine the constants $C_0, \dots, C_M$ appearing in Theorem 5.3 are the evaluations at $\mathbf{z} = \mathbf{w}$ of the partial derivatives of $G(\mathbf{z})$ up to order $2M$ and the partial derivatives of $H(\mathbf{z})$ up to order $2M + 2$.

The corresponding asymptotic result for higher order poles is more awkward to state, although all constants are still explicitly computable.

Theorem 5.3 (Smooth Asymptotics) *Suppose that $F(\mathbf{z}) = G(\mathbf{z})/H(\mathbf{z})$ admits a nondegenerate strictly minimal smooth contributing point $\mathbf{w} \in \mathbb{C}_*^d$ in the direction $\mathbf{r} \in \mathbb{R}_*^d$, such that $H_{z_d}^{\mathrm{s}}(\mathbf{w}) \neq 0$, and let p be the smallest integer such that $(\partial_d^p H)(\mathbf{w}) \neq 0$. Then for any nonnegative integer M there exist computable constants $C_0, \dots, C_M$ such that*

$$f_{n\mathbf{r}} = \mathbf{w}^{-n\mathbf{r}} n^{p-1+(1-d)/2} \frac{(2\pi)^{(1-d)/2}}{\sqrt{\det(r_d \mathcal{H})}} \left(\sum_{j=0}^{M} C_j (r_d n)^{-j} + O\left(n^{-M-1}\right) \right). \tag{5.28}$$

The constants C_j are determined by expanding $A_n(\boldsymbol{\theta})$ in Proposition 5.2 as a polynomial in n and applying Proposition 5.3 to the resulting sum of integrals. The leading constant C_0 in this series has the value

$$C_0 = \frac{\operatorname{sgn}(r_d)(-1)^p G(\mathbf{w}) p}{w_d^p \, (\partial_d^p H)(\mathbf{w})},$$

which is nonzero whenever $G(\mathbf{w}) \neq 0$. The asymptotic expansion (5.28) holds uniformly in neighbourhoods $\mathcal{R} \subset \mathbb{R}_*^d$ of $\mathbf{r}$ where there is a smoothly varying contributing point satisfying the same conditions as $\mathbf{w}$.

Remark 5.13 Often H is a pure power of a square-free polynomial, i.e., $H(\mathbf{z}) = Q(\mathbf{z})^p$ for some integer $p > 1$ and square-free polynomial Q. In this case the leading constant C_0 in Theorem 5.3 becomes

$$C_0 = \frac{\operatorname{sgn}(r_d)(-1)^p G(\mathbf{w})}{(p-1)! \left(w_d Q_{z_d}(\mathbf{w}) \right)^p}.$$

This variation is commonly in the literature.

When $\mathbf{w}$ is a finitely minimal critical point, and each point of $T(\mathbf{w}) \cap \mathcal{V}$ satisfies the conditions of Theorem 5.28, then one can add the asymptotic contributions of the minimal contributing points to determine dominant asymptotics.

Corollary 5.2 *Suppose that F admits a finitely minimal smooth contributing point $\mathbf{w} \in \mathbb{C}_*^d$ in the direction $\mathbf{r} \in \mathbb{R}_*^d$, let E be the set of points in $\mathcal{V} \cap T(\mathbf{w})$, and suppose all elements of E satisfy the conditions of Theorem 5.3. For some nonnegative integer M, let $\Phi_{\mathbf{w}}$ denote the right-hand side of (5.28) calculated at $\mathbf{w} \in E$ (equal to (5.27) whenever $\mathbf{w}$ is a simple pole). Then*

$$f_{n\mathbf{r}} = \sum_{\mathbf{w} \in E} \Phi_{\mathbf{w}}$$

gives an asymptotic expansion of $f_{n\mathbf{r}}$ as $n \to \infty$.

Theorem 5.3 gives an asymptotic expansion for the coefficients $f_{n\mathbf{r}}$, but when the numerator G vanishes at all minimal contributing points it is possible that the exponential growth of $f_{n\mathbf{r}}$ is smaller than what is predicted by these points.

Example 5.4 (Vanishing of Asymptotic Terms)

Consider the rational functions

$$A(x, y) = \frac{1}{1 - x - y}, \qquad B(x, y) = \frac{x - 2y^2}{1 - x - y}, \qquad \text{and} \qquad C(x, y) = \frac{x - y}{1 - x - y},$$

all of which admit a single contributing singularity $\boldsymbol{\sigma} = (1/2, 1/2)$ for the main diagonal direction $\mathbf{r} = (1, 1)$. The main power series diagonal of $A(x, y)$ is the generating function of the central binomial coefficients $\binom{2n}{n}$, and Theorem 5.2 with $M = 2$ gives

$$[x^n y^n]A(x, y) = \binom{2n}{n} = 4^n \left(\frac{1}{\sqrt{\pi} n^{1/2}} - \frac{1}{8\sqrt{\pi} n^{3/2}} + \frac{1}{128\sqrt{\pi} n^{5/2}} + O\left(n^{-7/2}\right) \right).$$

The numerator of $B(x, y)$ vanishes at $\boldsymbol{\sigma}$, meaning the constant C_0 in (5.27) is zero, but one can calculate the higher order constants to obtain

$$[x^n y^n]B(x, y) = 4^n \left(\frac{1}{4\sqrt{\pi} n^{3/2}} + \frac{3}{32\sqrt{\pi} n^{5/2}} + O\left(n^{-7/2}\right) \right).$$

Direct inspection shows that the main diagonal of $C(x, y)$ is identically zero. When applied in an automatic manner, Theorem 5.2 allows one to show for any natural number M, in a complexity which is polynomial in M, that

$$[x^n y^n]C(x, y) = O\left(\frac{4^n}{n^M} \right).$$

The fact that the diagonal is identically zero can be recovered from the eventual saddle-point integral by a short argument, see Problem 5.11. Note that $C(x, y)$ admits the minimal contributing point $\left(\frac{r}{r+s}, \frac{s}{r+s}\right)$ in the direction $\mathbf{r} = (r, s)$, and that the numerator of $G(x, y)$ vanishes only when $r = s$. Thus, the exponential growth rate of the coefficient sequence $f_{rn,sn}$ approaches 4^{rn} as $s \to r$, but there is a sharp drop down to zero exactly when $s = r$.

In Section 5.3.4 we show that most of our assumptions, including the condition that the numerator $G(\mathbf{z})$ does not vanish at the smooth critical points of F, hold for all rational functions with numerator and denominator of fixed degree, except those whose coefficients lie in some fixed algebraic set. This helps illustrate why Corollary 5.2 applies to many problems appearing in the mathematical and scientific literature.

Remark 5.14 The arguments above also hold for convergent Laurent expansions of a ratio $F(\mathbf{z}) = P(\mathbf{z})/Q(\mathbf{z})^p$ of analytic functions P and Q with $p \in \mathbb{N}$. Indeed, the numerator of F only enters into the determination of asymptotics for the integral in Proposition 5.2, and Proposition 5.3 only requires analyticity of the functions under consideration. The singular variety $\mathcal{V}$ is now the zero set of analytic $Q(\mathbf{z})$, and we still say $\mathbf{w} \in \mathcal{V}$ is a smooth point if some partial derivative of Q does not vanish at $\mathbf{w}$. The definition of critical points requires only the logarithmic gradient, and we again call a critical point contributing if it is a minimizer of the height function $h_{\mathbf{r}}$ on $\mathcal{D}$. None of the integral manipulations above used that G and H were polynomial, so under these modified definitions our asymptotic results still hold.

Corollary 5.3 *Suppose the hypotheses of Corollary 5.2 hold for a ratio* $F(\mathbf{z}) = P(\mathbf{z})/Q(\mathbf{z})^p$ *where p is a positive integer and for some* $\varepsilon > 0$ *the functions P and Q are analytic on* $\{\mathbf{z} \in \mathbb{C}^d : h_{\mathbf{r}}(\mathbf{z}) > h_{\mathbf{r}}(\mathbf{w}) - \varepsilon\}$. *Then the conclusion of Corollary 5.2 holds when asymptotics are calculated with* $G(\mathbf{z}) = P(\mathbf{z})$, $H(\mathbf{z}) = Q(\mathbf{z})^p$, *and* $H^{\mathrm{s}}(\mathbf{z}) = Q(\mathbf{z})$.

Although we do not discuss what it means for analytic functions P and Q to be coprime, if P and Q are not coprime and one applies Corollary 5.3 then the resulting asymptotic expansion still holds. The only downside is that one may find an apparent minimal critical point by examining $\mathcal{V}(Q)$ which is not actually a singularity of $P(\mathbf{z})/Q(\mathbf{z})^p$. In this case all coefficients C_j in the suspected leading series of the asymptotic expansion will (correctly) be calculated to be zero.

5.3 The Practice of Smooth ACSV

In this section we show how to apply our newly developed asymptotic methods, and give ancillary results which help put the theory into practice. For a rational function $F(\mathbf{z}) = G(\mathbf{z})/H(\mathbf{z})$ the smooth critical point equations (5.16) are defined by polynomial equalities involving the square-free part of H, and are easily manipulated with a computer algebra system. As previously noted, in Section 5.3.4 we show that for 'most' rational functions the singular variety $\mathcal{V}$ is everywhere smooth, there are a finite number of critical points, all of which are nondegenerate, and the numerator does not vanish at any of these points. Thus, the most difficult step in an asymptotic analysis is usually determining which, if any, of the algebraically defined critical points are minimal. We take a full computational perspective on these questions in Chapter 7, where we develop complete algorithms implementing our results.

When $H(\mathbf{z})$ is simple enough, direct arguments can be used to prove minimality.

Example 5.5 (Asymptotics of a Multinomial Expansion)

Consider the power series expansion of the trivariate function

$$F(x, y, z) = \frac{1}{1 - x - y - z} = \sum_{i,j,k\geq 0} \binom{i+j+k}{i, j, k} x^i y^j z^k = \sum_{i,j,k} \frac{(i+j+k)!}{i!j!k!} x^i y^j z^k,$$

whose power series domain of convergence consists of $(x, y, z) \in \mathbb{C}^3$ such that $|x| + |y| + |z| < 1$. Since $H(x, y, z) = 1 - x - y - z$ is square-free, the smooth critical point equations in direction $\mathbf{r} = (a, b, c)$ become

$$\begin{aligned} H(x, y, z) = 1 - x - y - z &= 0 \\ bxH_x(x, y, z) - ayH_y(x, y, z) = -bx + ay &= 0 \\ cxH_x(x, y, z) - ayH_z(x, y, z) = -cx + az &= 0, \end{aligned}$$

so that there is a single smooth critical point

$$(x, y, z) = (x_*, y_*, z_*) = \left(\frac{a}{a+b+c}, \frac{b}{a+b+c}, \frac{c}{a+b+c}\right).$$

If any of a, b, c is non-positive then F admits no minimal critical point for its power series expansion. If $a, b, c \in \mathbb{R}^d_{>0}$ and (x, y, z) satisfy

$$x + y + z = 1, \qquad |x| \leq x_*, \qquad |y| \leq y_*, \qquad |z| \leq z_*,$$

then it must be the case that $(x, y, z) = (x_*, y_*, z_*)$, so we have found a strictly minimal smooth contributing point (since every minimal critical point is contributing for a power series expansion when $\mathbf{r}$ has non-negative coordinates). Equation (5.25) implies the matrix $\mathcal{H}$ in (5.27) equals

$$\mathcal{H} = \begin{pmatrix} \frac{a(a+c)}{c^2} & \frac{ab}{c^2} \\ \frac{ab}{c^2} & \frac{b(b+c)}{c^2} \end{pmatrix}$$

and Theorem 5.2 gives an asymptotic expansion for the power series coefficients $[x^{an} y^{bn} z^{cn}]F(x, y, z)$ with $an, bn, cn \in \mathbb{N}$ that begins

$$\left(\frac{(a+b+c)^{a+b+c}}{a^a b^b c^c}\right)^n n^{-1} \left(\frac{\sqrt{a+b+c}}{2\pi\sqrt{abc}} - \frac{(a+b)(a+c)(b+c)}{24\pi(abc)^{3/2}\sqrt{a+b+c}} n^{-1} + \cdots\right).$$

Note that Corollary 2.1 implies the **r**-diagonal $(\Delta F)(z)$ is non-algebraic for any **r** with positive rational coordinates. If we consider the power series expansion of

$$F_p(x,y,z) = \frac{1}{(1-x-y-z)^p}$$

for $p \in \mathbb{N}$ then Theorem 5.3 gives an asymptotic expansion for the power series coefficients $[x^{an}y^{bn}z^{cn}]F_p(x,y,z)$ with $an, bn, cn \in \mathbb{N}$ that begins

$$\left(\frac{(a+b+c)^{a+b+c}}{a^a b^b c^c}\right)^n n^{p-2} \left(\frac{(a+b+c)^{p-1/2}}{(p-1)!\,2\pi\sqrt{abc}} + O\left(\frac{1}{n}\right)\right).$$

The exponential growth of the coefficient sequence is independent of p as the singular variety $\mathcal{V}$, and thus the location of the minimal contributing point (x_*, y_*, z_*), does not depend on p.

When dealing with Laurent expansions having non-negative coefficients, for instance multivariate generating functions of a combinatorial classes with parameters, we can use this non-negativity to help characterize minimal points.

Definition 5.8 (combinatorial series and functions) A convergent Laurent expansion of meromorphic $F(\mathbf{z})$ is called *combinatorial* if this series expansion contains only a finite number of negative coefficients. If $F(\mathbf{z})$ admits a power series expansion and this power series expansion is combinatorial, the situation we encounter most often, then we call $F(\mathbf{z})$ *combinatorial*. Given $\mathbf{z} \in \mathbb{C}^d$ we write $|\mathbf{z}| = (|z_1|, \ldots, |z_d|)$.

The following result is a multivariate analogue of the Vivanti-Pringsheim Theorem (Proposition 2.4 in Chapter 2) which greatly aids in determining minimality.

Lemma 5.7 *Suppose the Laurent expansion of $F(\mathbf{z})$ with domain of convergence $\mathcal{D}$ is combinatorial. If $\mathbf{w} \in \mathcal{V}_* \cap \partial\mathcal{D}$ is a minimal point of $F(\mathbf{z})$ with non-negative coordinates then the point $|\mathbf{w}|$ with positive coordinates also lies in $\mathcal{V}$ (and is thus also minimal).*

Proof Possibly by adding a Laurent polynomial to $F(\mathbf{z})$, which will not change the singularities of F away from the coordinate axes, we may assume that all Laurent coefficients of F are non-negative. Write the convergent Laurent expansion of F on $\mathcal{D}$ as

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

Since $f_{\mathbf{i}} \ge 0$ for all $\mathbf{i} \in \mathbb{Z}^d$, the triangle inequality implies

$$|F(\mathbf{z})| \le \sum_{\mathbf{i}\in\mathbb{Z}^d} f_{\mathbf{i}}\left|\mathbf{z}^{\mathbf{i}}\right| = F(|\mathbf{z}|)$$

for all $\mathbf{z} \in \mathcal{D}$. Because $F(\mathbf{z})$ is meromorphic its singularities are poles, so if $\mathbf{w}$ lies in $\mathcal{V} \cap \partial\mathcal{D}$ then there exists a sequence $\mathbf{z}^{(n)}$ in $\mathcal{D}$ converging to $\mathbf{w}$ such that $\left|F(\mathbf{z}^{(n)})\right| \to \infty$. But then the sequence $|\mathbf{z}^{(n)}|$ lies in $\mathcal{D}$, converges to $|\mathbf{w}|$, and satisfies $F(|\mathbf{z}^{(n)}|) \to \infty$. Thus $|\mathbf{w}|$ is also a polar singularity of F, as desired. ◻

Remark 5.15 If $F(\mathbf{z}) = G(\mathbf{z})/H(\mathbf{z})$ is a rational function and $B = \text{Relog}(\mathcal{D})$ then Lemma 5.7 implies ∂B, which describes a portion of the amoeba boundary $\partial\,\text{amoeba}(H)$, is contained in the image of $\mathcal{V}_{>0} = \mathcal{V} \cap \mathbb{R}^d_{>0}$ under the Relog map.

Lemma 5.7 gives a simple criterion to prove that a point is minimal when $F(\mathbf{z})$ is combinatorial.

Proposition 5.4 *Fix a convergent Laurent expansion of rational $F(\mathbf{z})$ with domain of convergence $\mathcal{D}$, and pick any $\mathbf{x} \in \mathcal{D}$. A point $\mathbf{w} \in \mathcal{V}$ is minimal if and only if there does not exist $\mathbf{z} \in \mathcal{V}$ with $|\mathbf{z}| = t|\mathbf{w}| + (1-t)|\mathbf{x}|$ and $t \in (0,1)$. If this series expansion is combinatorial then $\mathbf{w} \in \mathcal{V}$ is a minimal point if and only if $|\mathbf{w}| \in \mathcal{V}$ and the line segment from $|\mathbf{x}|$ to $|\mathbf{w}|$ in $\mathbb{R}^d_{>0}$ does not contain an element of $\mathcal{V}$,*

$$\{t|\mathbf{w}| + (1-t)|\mathbf{x}| : 0 < t < 1\} \cap \mathcal{V} = \varnothing.$$

If the Laurent expansion under consideration is a power series expansion, one can take $\mathbf{x} = \mathbf{0}$ and thus restrict attention to the line segment from the origin to $|\mathbf{w}|$.

Proof Since $\mathbf{x} \in \mathcal{D}$ its image $\text{Relog}(\mathbf{x})$ lies in the (open) component $B = \text{Relog}(\mathcal{D})$ of the amoeba complement. If $\mathbf{w}$ is minimal then $\text{Relog}(\mathbf{w}) \in \partial B$ and convexity of B implies no element of $\mathcal{V}$ can lie in the interior of a line segment from $|\mathbf{x}|$ to $|\mathbf{w}|$. If $\mathbf{w}$ is not minimal, then $\text{Relog}(|\mathbf{w}|)$ lies outside of the closed convex set $\overline{B}$, so any continuous path in $\mathbb{R}^d$ from $\text{Relog}(|\mathbf{w}|)$ to the open set B must pass through ∂B. Since the path

$$\log\left(t|\mathbf{w}| + (1-t)|\mathbf{x}|\right), \quad 0 \le t \le 1$$

goes from $\text{Relog}(\mathbf{w})$ to $\text{Relog}(\mathbf{x})$, where the logarithm is taken coordinate-wise, there exists $\mathbf{z} \in \mathcal{V}$ and $t \in (0,1)$ such that $\log\left(t|\mathbf{w}| + (1-t)|\mathbf{x}|\right) = \text{Relog}(\mathbf{z})$. Exponentiating coordinate-wise then gives $|\mathbf{w}| = t|\mathbf{z}| + (1-t)|\mathbf{x}|$.

When $F(\mathbf{z})$ is combinatorial, Lemma 5.7 shows that it is sufficient to consider only the points in $\mathcal{V} \cap \mathbb{R}^d_{>0}$ to determine the minimality of $\mathbf{w}$. ◻

Although Lemma 5.7 can be seen as a multivariate generalization of the Vivanti-Pringsheim Theorem, its conditions are more restrictive as it requires all coefficients of the series expansion to be non-negative, not just those in the direction $\mathbf{r}$ which may be of (combinatorial) interest. As mentioned in Chapter 2, it is still unknown even in the univariate case how to decide when a rational function is combinatorial. In practice, then, one usually applies these results when $F(\mathbf{z})$ is the multivariate generating function of a combinatorial class with parameters, or when the form of $F(\mathbf{z})$ makes combinatoriality easy to prove (for instance, when considering the power series expansion of $F(\mathbf{z}) = \frac{G(\mathbf{z})}{1-J(\mathbf{z})}$ with $J(\mathbf{z})$ a polynomial vanishing at the origin having non-negative coefficients).

Example 5.6 (Lattice Path Asymptotics)

In Chapter 4 we saw that the main diagonal of the power series expansion of

$$F(x, y, t) = \frac{(1+x)(1+y)}{1 - txy(x + y + 1/x + 1/y)}$$

is the generating function for the number of lattice walks starting at the origin, taking the steps $(\pm 1, 0), (0, \pm 1)$, and staying in the first quadrant. The denominator is square-free, and the singular variety $\mathcal{V}$ is smooth as the partial derivative of the denominator with respect to t does not vanish at any points of $\mathcal{V}$. Furthermore, F is combinatorial. The critical point equations imply that there are two critical points,

$$\boldsymbol{\sigma} = (1, 1, 1/4) \qquad \text{and} \qquad \boldsymbol{\varrho} = (-1, -1, -1/4).$$

Proposition 5.4 implies $\boldsymbol{\sigma}$ and $\boldsymbol{\varrho}$ are minimal critical points: if $(x, y, t) \in \mathcal{V}$ with $0 < x, y \leq 1$ and one of the upper inequalities is strict then

$$|t| = \left| \frac{1}{x^2 y + xy^2 + y + x} \right| > 1/4.$$

Furthermore, if $|x| = 1$ and $|y| = 1$ for $x, y \in \mathbb{C}$ then the triangle inequality implies

$$|x^2 y + xy^2 + x + y| = 4$$

only if $xy^2, x, x^2 y$, and y have the same argument. Thus, x^2 and y^2 must be real with modulus 1 and a quick check shows the only other point of $\mathcal{V}$ satisfying this condition is $\boldsymbol{\varrho}$. We have shown these critical points are finitely minimal.

Since the numerator $G(x, y) = (1 + x)(1 + y)$ vanishes (to order 2) when $(x, y, t) = \boldsymbol{\varrho}$, Corollary 5.2 implies only $\boldsymbol{\sigma}$ contributes to the dominant asymptotics of the diagonal sequence. The contributions from each minimal critical point begin

$$\Phi_{\boldsymbol{\sigma}} = 4^n \left(\frac{1}{\pi n} - \frac{6}{\pi n^2} + \frac{19}{2\pi n^3} - \frac{121}{12\pi n^4} + O\left(\frac{1}{n^5}\right) \right)$$
$$\Phi_{\boldsymbol{\varrho}} = (-4)^n \left(\frac{1}{\pi n^3} - \frac{9}{2\pi n^4} + O\left(\frac{1}{n^5}\right) \right).$$

Note that the presence of two minimal critical points leads to periodicity in the higher order asymptotic terms,

$$f_{n,n,n} = \frac{4^n}{\pi n} \left(4 - \frac{6}{n} + \frac{19 + 2(-1)^n}{2n^2} - \frac{63 + 18(-1)^n}{4n^3} + O\left(\frac{1}{n^4}\right) \right).$$

More examples from lattice path enumeration, and additional details on this example, are discussed in Chapter 6.

The algebraic structure of the series terms with non-zero coefficients can also help determine when a minimal point is strictly minimal.

Definition 5.9 (aperiodic series) A power series $P(\mathbf{z}) = \sum_{\mathbf{n}\in\mathbb{N}^d} p_\mathbf{n}\mathbf{z}^\mathbf{n}$ is called *aperiodic* if every element of $\mathbb{Z}^d$ can be written as an integer linear combination of the exponents $\{\mathbf{n} \in \mathbb{N}^d : p_\mathbf{n} \neq 0\}$ appearing in P.

Proposition 5.5 *Suppose $F(\mathbf{z}) = G(\mathbf{z})/H(\mathbf{z})$ is a ratio of functions G and H which are analytic on $\mathbb{C}^d$. If $H(\mathbf{z}) = 1 - P(\mathbf{z})$ for some aperiodic power series P with non-negative coefficients then every minimal point of the power series expansion of $F(\mathbf{z})$ is strictly minimal and has positive real coordinates.*

Proposition 5.5 is often applied when $F(\mathbf{z})$ is a rational function (so P is a polynomial with non-negative coefficients).

Proof Suppose $\mathbf{w}$ is a minimal point and for each $1 \leq j \leq d$ write $w_j = x_j e^{i\theta_j}$ with $x_j > 0$ and $\theta_j \in \mathbb{R}$. Let $p_\mathbf{n}$ denote the coefficient of $z^\mathbf{n}$ in $P(\mathbf{z})$. Then

$$1 = |P(\mathbf{w})| = \left|\sum_{\mathbf{n}\in\mathbb{N}^d} p_\mathbf{n}\mathbf{x}^\mathbf{n} e^{i(\mathbf{n}\cdot\boldsymbol{\theta})}\right| \leq \sum_{\mathbf{n}\in\mathbb{N}^d} p_\mathbf{n}\mathbf{x}^\mathbf{n},$$

where equality holds throughout only if $P(\mathbf{z})$ is a monomial or $\mathbf{n}\cdot\boldsymbol{\theta} = 0$ for all $\mathbf{n} \in \mathbb{N}^d$ such that $p_\mathbf{n} \neq 0$. Thus, either $\boldsymbol{\theta} = \mathbf{0}$ or $P(\mathbf{z})$ is not aperiodic. □

Example 5.7 (Lonesum Matrices)

A *lonesum matrix* is a matrix with entries in $\{0, 1\}$ that is uniquely determined by its row and column sums; the enumeration of lonesum matrices finds application in biology [27] and algebraic statistics [19]. If $B_{n,k}$ denotes the number of $n \times k$ lonesum matrices then Brewbaker [7] proved the generating function expression

$$F(x, y) = \sum_{n,k\geq 0} B_{n,k}\frac{x^n}{n!}\frac{y^k}{k!} = \frac{e^{x+y}}{e^x + e^y - e^{x+y}}.$$

Since the denominator under consideration has the form $H(x, y) = 1 - P(x, y)$ where

$$P(x, y) = 1 - e^x - e^y + e^{x+y} = \sum_{k\geq 1}\sum_{j=1}^{k-1}\frac{x^j y^{k-j}}{j!(k-j)!}$$

is aperiodic, Proposition 5.5 implies all minimal points of the power series expansion of $F(x, y)$ are positive and strictly minimal. Furthermore, if $H(x, y) = 0$ then $y = x - \log(e^x - 1)$ is a decreasing function of $x > 0$ so every point with positive coordinates is minimal (decreasing $x > 0$ and staying in the singular variety $\mathcal{V}$ means increasing y). Manipulating the smooth critical point equations (5.16) in the direction $\mathbf{r} = (r, s)$ shows that (x, y) is a critical point in the direction $\mathbf{r}$ if and only if

$$\frac{r}{s} = \frac{x}{(1-e^x)\log(1-e^{-x})} \qquad \text{and} \qquad \frac{s}{r} = \frac{y}{(1-e^y)\log(1-e^{-y})}.$$

Since the function $f(t) = \frac{t}{(1-e^t)\log(1-e^{-t})}$ has positive derivative for $t > 0$, goes to 0 as $t \to 0$, and goes to infinity as $t \to \infty$, f is a bijection from the positive real line to itself. Thus, in any direction $\mathbf{r} = (r, s)$ with positive coordinates there is a unique critical point (a, b) where $a = a(r, s) = f^{-1}(r/s)$ and $b = b(r, s) = f^{-1}(s/r)$. Because $F(x, y)$ is bivariate the Hessian $\mathcal{H}$ in (5.25) is a constant, and is non-zero at any point with positive coordinates. Finally, since we consider a power series expansion in directions with positive coordinates the minimal critical point (a, b) is always contributing. Corollary 5.3 thus implies

$$B_{rn,sn} = \frac{a(r,s)^{-rn}b(r,s)^{-sn}}{\sqrt{n}} \frac{(rn)!(sn)!}{\sqrt{2\pi sae^{-a}(be^{-b} + ae^{-a} - ab)}} \left(1 + O\left(n^{-1}\right)\right)$$

for all $r, s > 0$. Additional details on this example can be found in Khera et al. [26]. The approach presented here is partially modeled on work of Andrade et al. [12], who studied a related generating function arising in permutation statistics. Problem 5.3 asks you to find asymptotics for this related generating function.

5.3.1 Existence of Minimal Critical Points

Because minimal contributing points are crucial to our asymptotic results, we now study when they do (or do not) exist. First, they may not exist for a somewhat trivial reason: because a minimal contributing point is a minimizer of the height function $h_{\mathbf{r}}$ on $\overline{\mathcal{D}}$, there will be no minimal contributing points when $h_{\mathbf{r}}$ decreases without bound. This can only happen in degenerate cases.

Proposition 5.6 *Consider the convergent Laurent expansion of rational function* $F(\mathbf{z}) = G(\mathbf{z})/H(\mathbf{z})$ *on a domain* $\mathcal{D} \subset \mathbb{C}^d$. *If the height function* $h_{\mathbf{r}}$ *decreases without bound on* $\mathcal{D}$ *then the coefficient sequence* $(f_{n\mathbf{r}})$ *is eventually zero.*

Proof If $h_{\mathbf{r}}$ is unbounded from below then the inequality (5.15) implies $|f_{n\mathbf{r}}|$ goes to zero faster than any exponential function. Since the asymptotic behaviour of a diagonal sequence is constrained by Corollary 3.2 in Chapter 3, this implies $f_{n\mathbf{r}}$ is eventually zero. Alternatively, the conclusion follows from (5.14) and the fact that $C_{\mathbf{w}}$ grows at most polynomially. □

When studying a diagonal sequence which is not eventually zero, Proposition 5.6 implies the height function is bounded on $\overline{\mathcal{D}}$. The following result uses the algebraic properties of polynomial amoebas discussed in Section 3.3.1 of Chapter 3 to help characterize when a diagonal sequence is identically zero.

Proposition 5.7 *Let* $F(\mathbf{z}) = G(\mathbf{z})/H(\mathbf{z})$ *be a ratio of coprime polynomials, let* $\mathcal{N}(H)$ *be the Newton polytope of* H, *and let* $B = \operatorname{Relog}(\mathcal{D})$ *be a connected component* *of* amoeba$(H)^c$ *corresponding to the integer point* $\mathbf{v} \in \mathcal{N}(H)$ *under the mapping described by Proposition 3.12 in Chapter 3. If* $\mathbf{v} + t\mathbf{r} \notin \mathcal{N}(H)$ *for all* $t > 0$ *then the height function* $h_\mathbf{r}(\mathbf{z})$ *is unbounded from below on* $\overline{\mathcal{D}}$*, and* $(f_{n\mathbf{r}})$ *is eventually zero.*

Proof Suppose $h_\mathbf{r}$ is bounded from below on $\overline{\mathcal{D}}$, so that $\tilde{h}_\mathbf{r}(\mathbf{x}) = -\mathbf{r} \cdot \mathbf{x}$ is bounded from below on $\overline{B}$. Suppose also that $\mathbf{y} \in \mathbb{R}^d$ lies in the recession cone $\mathcal{R}$ of B, meaning that $\mathbf{y} + \mathbf{b} \in B$ for all $\mathbf{b} \in B$. Then by induction $n\mathbf{y} + \mathbf{b} \in B$ for all $\mathbf{b} \in B$ and $n \in \mathbb{N}$. If $-\mathbf{r} \cdot \mathbf{y} < 0$ then $-\mathbf{r} \cdot (n\mathbf{y} + \mathbf{b}) = n(-\mathbf{r} \cdot \mathbf{y}) - \mathbf{r} \cdot \mathbf{b} \to -\infty$ as $n \to \infty$, contradicting the fact that $\tilde{h}$ is bounded on B. Thus, it must be the case that $-\mathbf{r} \cdot \mathbf{y} \geq 0$ for all $\mathbf{y} \in \mathcal{R}$. Proposition 3.12 in Chapter 3 characterizes the recession cone as

$$\mathcal{R} = \{\mathbf{x} \in \mathbb{R}^d : \mathbf{x} \cdot \mathbf{n} \leq 0 \text{ for all } \mathbf{n} \in \mathcal{N}(H) - \mathbf{v}\},$$

where $\mathcal{N}(H) - \mathbf{v}$ is the set $\{\mathbf{z} - \mathbf{v} : \mathbf{z} \in \mathcal{N}(H)\}$. Because $\mathcal{N}(H)$ is a closed convex set, if $t\mathbf{r} \notin \mathcal{N}(H) - \mathbf{v}$ for all $t > 0$ then there exists $\mathbf{y} \in \mathbb{R}^d$ such that $\mathbf{r} \cdot \mathbf{y} > 0$ and $\mathbf{n} \cdot \mathbf{y} < 0$ for all $\mathbf{n} \in \mathcal{N}(H) - \mathbf{v}$. But this implies $\mathbf{y} \in \mathcal{R}$ and $-\mathbf{r} \cdot \mathbf{y} < 0$, a contradiction. Thus, when $h_\mathbf{r}$ is bounded from below there exists $t > 0$ such that $t\mathbf{r} \in \mathcal{N}(H) - \mathbf{v}$. □

Unfortunately, even if $h_\mathbf{r}$ achieves its minimum on $\overline{\mathcal{D}}$ this minimum need not be a critical point. In fact, when the series under consideration is not combinatorial it is possible to construct examples with no critical points although the minimum of $h_\mathbf{r}$ on $\overline{\mathcal{D}}$ is achieved and the singular variety $\mathcal{V}$ is smooth.

Example 5.8 (A Diagonal with No Critical Points)

Consider the bivariate function $F(x, y) = G(x, y)/H(x, y) = 1/(2 + y - x(1 + y)^2)$ and let $\mathcal{D}$ be the power series domain of convergence of F. Basic algebra shows that both polynomial systems

$$H = H_x = H_y = 0 \qquad \text{and} \qquad H = xH_x - yH_y = 0$$

have no solutions, so the singular variety $\mathcal{V}$ is smooth but there are no critical points in the main diagonal direction $\mathbf{r} = (1, 1)$. Since the dual cone of the Newton polygon $\mathcal{N}(H)$ at the origin is the quadrant $\mathcal{R} = \{(x, y) \in \mathbb{R}^2 : x, y \leq 0\}$, Proposition 3.12 in Chapter 3 implies the recession cone of $B = \text{Relog}(\mathcal{D})$ is $\mathcal{R}$ and $\overline{B}$ is a closed convex set contained in some translation of the third quadrant of the plane (see Figure 5.5). This implies the linear function $\tilde{h}(p, q) = -p - q$ achieves its minimum for $(p, q) \in \overline{B}$, so the height function $h(x, y)$ achieves its minimum on $\overline{\mathcal{D}}$.

The set of minimal points can be parametrized as

$$\mathcal{V} \cap \partial\mathcal{D} = \left\{ \left(\frac{2 + y}{(1 + y)^2}, y \right) : y \in \left[-2, -\sqrt{3}\right] \cap \left[0, \sqrt{3}\right] \right\},$$

with corresponding logarithmic gradient

$$(\nabla_{\log} H) \left(\frac{2 + y}{(1 + y)^2}, y \right) = -\left(2 + y, 2 + y - \frac{2}{1 + y} \right),$$

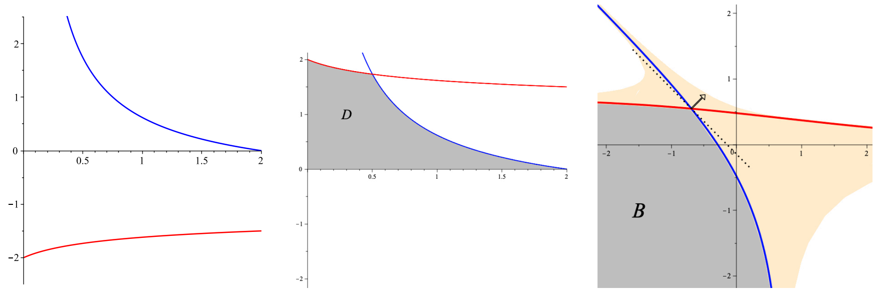


Fig. 5.5 *Left:* A subset of the real points in the smooth singular variety $\mathcal{V} = \mathcal{V}(2 + y - x(1 + y)^2)$. *Center:* Taking coordinate-wise moduli creates an intersection, which causes $h(x, y)$ to achieve its minimum on $\partial\mathcal{D}$. *Right:* The curves in the contour of H corresponding to the real connected components of $\mathcal{V}$ shown on the left, together with amoeba(H) and the component $B = \text{Relog}(\mathcal{D})$ of amoeba$(H)^c$. Although the hyperplane with normal (1, 1) is a supporting hyperplane to B at the intersection, neither branch of the contour has a tangent line perpendicular to (1, 1).

which is indeed never colinear with (1, 1). Since the logarithmic gradient approaches a multiple of (1, 1) as $y \to \infty$, one might say there is a (non-minimal) critical point *at infinity* in this example; we return to this notion in Section 9.3 of Chapter 9.

The problem is that the minimizers $(1/2, \pm\sqrt{3})$ of $h(x, y)$ on $\overline{\mathcal{D}}$ come from two different connected components of $\mathcal{V} \cap \mathbb{R}^2$ which intersect only after taking coordinate-wise moduli, as shown in Figure 5.5. More generally, one can imagine $\mathcal{V}$ wrapping over itself in complex space so that $\mathcal{V}$ is smooth but the boundary of amoeba(H) is not, in such a way that $\mathbf{r}$ is the normal to a supporting hyperplane at B only for 'unnatural' reasons due to a non-smooth point (such an occurrence has been called a *ghost intersection* in the literature).

This difficulty arises when the behaviour of Relog($\mathcal{V}$) in $\mathbb{R}^d$ does not reflect the behaviour of $\mathcal{V}$ in $\mathbb{C}^d$. Luckily, this cannot happen in the combinatorial case, where the points of $\mathcal{V}$ with positive real coordinates map onto the points of ∂B by taking coordinate-wise logarithms.

Proposition 5.8 *Consider a combinatorial Laurent expansion of $F(\mathbf{z})$ and let $\mathbf{x}$ be a smooth minimal point in $\mathcal{V}$ with positive real coordinates. Then every point in some neighbourhood of $\mathbf{x}$ in $\mathcal{V} \cap \mathbb{R}^d$ is a minimal point.*

Proof We assume that H is square-free, otherwise replace it with its square-free part in this argument. Proposition 3.13 in Chapter 3 implies that the minimal point $\mathbf{x}$ is critical in some non-zero direction $\mathbf{r} \in \mathbb{R}^d$ and, perhaps by replacing $\mathbf{r}$ with its negation, we may assume that $\mathbf{r}$ points away from B at $\log(\mathbf{x})$. Consider the polynomial system

$$H(\mathbf{z}) = H(t^{r_1} z_1, \dots, t^{r_d} z_d) = 0$$

and assume (without loss of generality) that $H_{z_d}(\mathbf{x}) \neq 0$. At the solution $(\mathbf{z}, t) = (\mathbf{x}, 1)$ the Jacobian of this system with respect to z_d and t is

$$\begin{pmatrix} H_{z_d}(\mathbf{x}) & 0 \\ H_{z_d}(\mathbf{x}) & \mathbf{r} \cdot (\nabla_{\log} H)(\mathbf{x}) \end{pmatrix},$$

which has non-zero determinant as $(\nabla_{\log} H)(\mathbf{x})$ is a non-zero multiple of $\mathbf{r}$. Thus, there exist analytic functions $g(\hat{\mathbf{z}})$ and $T(\hat{\mathbf{z}})$ parameterizing z_d and t near $(\mathbf{x}, 1)$. If no neighbourhood of $\mathbf{x}$ in $\mathcal{V} \cap \mathbb{R}^d$ contains only minimal points then, since $\mathbf{x}$ is minimal, there exists a sequence of non-minimal points $\mathbf{x}^{(n)} \in \mathcal{V} \cap \mathbb{R}^d$ converging to $\mathbf{x}$ such that $H(t_n^{r_1} x_1^{(n)}, \ldots, t_n^{r_d} x_d^{(n)}) = 0$ is satisfied with $t_n = 1$ (since $\mathbf{x}^{(n)} \in \mathcal{V}$) and some $0 < t_n < 1$ (since $\mathbf{x}^{(n)}$ is not minimal). But this contradicts the fact that, for n sufficiently large, $t_n \to 1$ is uniquely determined by $T(x_1^{(n)}, \ldots, x_{d-1}^{(n)})$. □

Corollary 5.4 *Consider a combinatorial Laurent expansion of $F(\mathbf{z})$. If a smooth point $\mathbf{x} \in \mathbb{R}^d_{>0}$ is a local minimizer (or local maximizer) of $h_\mathbf{r}$ on $\partial\mathcal{D}$ then it is a critical point in the direction $\mathbf{r}$.*

Proof If $\mathbf{x}$ is a local minimizer of $h_\mathbf{r}$ on $\partial\mathcal{D}$ then by Proposition 5.8 it is a local minimizer on $\mathcal{V} \cap \mathbb{R}^d$. Thus, the gradient of $\mathbf{z} \mapsto \mathbf{r} \cdot \log(\mathbf{z})$ must be normal to $\mathcal{V}$, giving the critical point equations. If $\mathbf{x}$ is a local maximizer of $h_\mathbf{r}$ then it is a local minimizer of $-h_\mathbf{r}$, and the same argument holds. □

Corollary 5.5 *Consider a combinatorial Laurent expansion of rational $F(\mathbf{z})$, and let $\mathbf{x}$ be a smooth minimal point in $\mathcal{V} \cap \mathbb{R}^d_{>0}$. If a smooth point $\mathbf{w} \in \mathcal{V}$ has the same coordinate-wise modulus as $\mathbf{x}$, then $\mathbf{x}$ is a critical point in some direction $\mathbf{r}$ if and only if $\mathbf{w}$ is a critical point in the direction $\mathbf{r}$.*

Proof Since $\mathbf{x}$ and $\mathbf{w}$ are smooth, minimal, and have no zero coordinates, Proposition 3.13 in Chapter 3 implies the existence of $\mathbf{r}, \boldsymbol{\varrho} \in \mathbb{R}^d$ such that $\mathbf{x}$ and $\mathbf{w}$ are critical in the directions $\mathbf{r}$ and $\boldsymbol{\varrho}$. In other words, there exist $\lambda, \tau \in \mathbb{C}_*$ such that

$$(\nabla_{\log} H)(\mathbf{x}) = \lambda \mathbf{r} \qquad \text{and} \qquad (\nabla_{\log} H)(\mathbf{w}) = \tau \boldsymbol{\varrho}.$$

Since $\mathbf{r}$ and $\boldsymbol{\varrho}$ are uniquely determined (up to non-zero multiple) by $\mathbf{x}$ and $\mathbf{w}$, it is sufficient to show that $\mathbf{x}$ is also critical in the direction $\boldsymbol{\varrho}$. Proposition 3.13 implies that $\boldsymbol{\varrho}$ is the normal to a support hyperplane of $B = \operatorname{Relog}(\mathcal{D})$ at $\operatorname{Relog}(\mathbf{w}) = \log(\mathbf{x})$. This implies $\mathbf{x}$ is a minimizer (or maximizer) of $h_{\boldsymbol{\varrho}}$ on $\overline{\mathcal{D}}$, and Corollary 5.4 then implies $\mathbf{x}$ is critical in the direction $\boldsymbol{\varrho}$, as desired. □

Example 5.9 (Asymptotics of Apéry Numbers)

In Section 3.4 we saw the sequence (a_n) of Apéry numbers, whose generating function can be expressed as the main diagonal of

$$F(x, y, z, t) = \frac{1}{1 - t(1+x)(1+y)(1+z)(1+y+z+yz+xyz)}.$$

This rational function is combinatorial, and solving the critical point equations gives two smooth critical points, one of which,

$$\boldsymbol{\sigma} = \left(1 + \sqrt{2}, \frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 58\sqrt{2} - 82\right),$$

has positive coordinates. Proposition 5.4 implies that $\boldsymbol{\sigma}$ is a minimal critical point, since at any singularity

$$t = \frac{1}{(1+x)(1+y)(1+z)(1+y+z+yz+xyz)},$$

and when x, y, and z are positive and real then decreasing any of their values causes this expression to increase. As $\mathcal{V}$ is smooth and F is combinatorial and admits only one critical point, Corollary 5.5 immediately implies that F is strictly minimal. Thus, Theorem 5.2 implies

$$a_n = \frac{\left(17 + 12\sqrt{2}\right)^n}{n^{3/2}} \frac{\sqrt{48 + 34\sqrt{2}}}{8\pi^{3/2}} \left(1 + O\left(\frac{1}{n}\right)\right).$$

Recall that determining such an asymptotic estimate is a key step in Apéry's proof of the irrationality of $\zeta(3)$.

The generating function of the second sequence of Apéry numbers (c_n) can be written as the main diagonal of

$$F(x, y, z) = \frac{1}{1 - z(1+x)(1+y)(1+y+xy)}.$$

Again F is combinatorial, and an analogous argument shows

$$c_n = \frac{\left(\frac{11}{2} + \frac{5\sqrt{5}}{2}\right)^n}{n} \frac{\sqrt{250 + 110\sqrt{5}}}{20\pi} \left(1 + O\left(\frac{1}{n}\right)\right).$$

We show how these examples can be completely automated in Chapter 7.

Example 5.10 (Asymptotics of Bar Graphs)

Recall from Section 3.2.2 of Chapter 3 the function $F(x, y, z) = \frac{z(1-xyz-2xz-yz-x)}{1-xyz-xy-xz-yz-x}$, whose power series coefficient $b_{n,k} = [x^n y^k z^k]F(x, y, z)$ counts bar graphs with $2n$ horizontal external edges and $2k$ vertical external edges. In a direction $\mathbf{r} = (a, 1, 1)$ the smooth critical point equations have two solutions, the point

$$\mathbf{z}_* = (x_*, y_*, z_*) = \left(\frac{\sqrt{a^2+4} - 2}{a}, \frac{\sqrt{a^2+4} - a}{2}, \frac{\sqrt{a^2+4} - a}{2}\right),$$

which has positive coordinates, and another point whose coordinate-wise modulus is larger than $\mathbf{z}_*$. Although the power series expansion of F contains negative coefficients, if $H(x, y, z)$ is the denominator of F then the power series expansion of $1/H$ is combinatorial and contains the same singular set at F. Thus, $\mathbf{z}_*$ is minimal if the univariate polynomial $p(t) = H(tx_*, ty_*, tz_*)$ has no root $t \in (0, 1)$. A quick analysis of $p(t)$, which factors as $1 - t$ times a quadratic polynomial in t, shows that for any $a > 0$ the only root of $p(t)$ with $t > 0$ is $t = 1$. This proves $\mathbf{z}_*$ is minimal and, since $1/H$ is combinatorial and no other critical point has the same coordinate-wise modulus as $\mathbf{z}_*$, it is strictly minimal. Theorem 5.2 then implies that for all $a > 0$

$$b_{an,n} \sim \left(\frac{\sqrt{a^2+4}-a}{2}\right)^{2n} \left(\frac{\sqrt{a^2+4}-2}{2}\right)^{an} \frac{C_a}{4a^2(a^2+4)\pi n^2},$$

where $C_a = \left((a^2-2a+4)\sqrt{a^2+4} - (a-2)(a^2+4)\right)\sqrt{a^2+4-2\sqrt{a^2+4}}$. See the computer algebra worksheet corresponding to this example for more details.

Remark 5.16 The conclusion of Corollary 5.5 may fail in the non-combinatorial case. For instance, consider the main diagonal of $F(x, y) = 1/(1+2x)(1-x-y)$. Although any point $(-1/2, y)$ with $|y| = 1/2$ is a singularity with the same coordinate-wise modulus as the minimal smooth critical point $(1/2, 1/2)$, none of these points are themselves critical in the main diagonal direction.

Corollary 5.6 *Consider a combinatorial Laurent expansion of $F(\mathbf{z})$. If $\mathbf{w}$ is a smooth minimal critical point in the direction $\mathbf{r}$ then $|\mathbf{w}|$ is a smooth minimal critical point in the direction $\mathbf{r}$. Furthermore, if every point in $\partial\mathcal{D} \cap \mathcal{V}$ is smooth and F admits two distinct minimal contributing points $\mathbf{a}, \mathbf{b} \in \mathbb{R}^d_{>0}$ with positive coordinates in the direction $\mathbf{r}$, then*

$$\mathbf{c}_t = \left(a_1^t b_1^{1-t}, \dots, a_d^t b_d^{1-t}\right)$$

is a minimal contributing point in the direction $\mathbf{r}$ for all $0 \le t \le 1$. In particular, F admits an infinite number of critical points in the direction $\mathbf{r}$, none of which are nondegenerate.

Proof The first statement follows from Lemma 5.7 and Corollary 5.5. To prove the second we note that, by logarithmic convexity of $\mathcal{D}$, the point $\mathbf{c}_t$ lies in $\overline{\mathcal{D}} \cap \mathbb{R}^d_{>0}$ for all $0 \le t \le 1$. Proposition 5.1 implies both $\mathbf{a}$ and $\mathbf{b}$ are minimizers of $h_\mathbf{r}$ on $\overline{\mathcal{D}}$, so $h_\mathbf{r}(\mathbf{a}) = h_\mathbf{r}(\mathbf{b})$. Thus, for any $0 \le t \le 1$,

$$\begin{aligned} h_\mathbf{r}(\mathbf{c}_t) = -\sum_{j=1}^d r_j \log\left(a_j^t b_j^{1-t}\right) &= -t\sum_{j=1}^d r_j \log a_j - (1-t)\sum_{j=1}^d r_j \log b_j \\ &= t h_\mathbf{r}(\mathbf{a}) + (1-t) h_\mathbf{r}(\mathbf{b}) \\ &= h_\mathbf{r}(\mathbf{a}) \end{aligned}$$

is also a minimizer of $h_\mathbf{r}$ on $\overline{\mathcal{D}}$. This implies $\mathbf{c}_t$ is a minimal point for all $0 \le t \le 1$, since minimizers of $h_\mathbf{r}$ occur only on $\partial\mathcal{D}$, and Corollary 5.4 implies each $\mathbf{c}_t$ is a critical point. Since these critical points are not isolated, Lemma 5.6 implies they cannot be nondegenerate. □

5.3.2 Dealing with Minimal Points that are not Critical

Determining minimality of critical points is typically the hardest task in a multivariate singularity analysis. Even when a critical point $\mathbf{w} \in \mathcal{V}$ is known to be minimal, to apply Theorem 5.3 or Corollary 5.2 one must determine if $T(\mathbf{w}) \cap \mathcal{V}$ is finite and, if so, determine its elements. In the combinatorial case, Corollary 5.5 shows that every element of $T(\mathbf{w}) \cap \mathcal{V}$ is critical, greatly simplifying matters from a computational viewpoint as critical points are specified by a set of algebraic equations which usually have a finite number of solutions. Unfortunately, as Remark 7.3 illustrates, this property does not hold in general for non-combinatorial series expansions.

Thankfully, even if $T(\mathbf{w}) \cap \mathcal{V}$ contains non-critical points it turns out only the critical points matter for asymptotic calculations. The requirement of finite minimality arose above because we made essentially univariate deformations of the Cauchy domain of integration, increasing the modulus of each coordinate independently. Using a genuinely multivariate deformation of the Cauchy domain of integration, it is possible to deform around non-critical points in $T(\mathbf{w})$. In this sense, only critical points present true obstructions for deforming the Cauchy domain of integration to points of lower height.

Theorem 5.4 *Consider the convergent Laurent expansion of a rational function $F(\mathbf{z})$ on a domain $\mathcal{D}$. Suppose there exists $\mathbf{w} \in \mathcal{V} \cap \partial\mathcal{D}$ minimizing $h_\mathbf{r}$ on $\overline{\mathcal{D}}$, and that for some $\varepsilon > 0$ all points of the set*

$$\{\mathbf{z} \in \mathcal{V} : h_\mathbf{r}(\mathbf{z}) \geq h_\mathbf{r}(\mathbf{w}) - \varepsilon\}$$

are smooth points of $\mathcal{V}$. Suppose further that the set W of critical points in $\mathcal{V} \cap T(\mathbf{w})$ consists of a finite number of nondegenerate smooth contributing points. Then for every positive integer $M > 0$ an asymptotic expansion of $f_{n\mathbf{r}}$ is obtained by summing the right-hand side of (5.28) in Theorem 5.3 at each $\mathbf{w} \in W$.

The multivariate deformations required to prove Theorem 5.4 require some technical overhead, and most of our applications involve combinatorial rational functions, so we simply sketch two approaches. The first, contained in Baryshnikov and Pemantle [3, Sect. 5], uses Proposition 3.13 in Chapter 3 to construct a vector field describing how to deform the Cauchy domain of integration around non-critical points in $T(\mathbf{w})$. Proposition 3.13 states that for any singularity $\boldsymbol{\zeta} \in \mathcal{V} \cap T(\mathbf{w})$ there exist $\lambda_\zeta \in \mathbb{C}_*$ and non-zero $\mathbf{v}_\zeta \in \mathbb{R}^d$ such that $(\nabla_{\log} H)(\boldsymbol{\zeta}) = \lambda_\zeta \mathbf{v}_\zeta$, and the point $\boldsymbol{\zeta}$ is critical if and only if $\mathbf{v}_\zeta$ is parallel to $\mathbf{r}$. Since $\nabla_{\log} H$ is the normal of the tangent space to the set $\tilde{\mathcal{V}} = \{\mathbf{x} \in \mathbb{C}^d : H(e^{x_1}, \dots, e^{x_d}) = 0\}$, if $\boldsymbol{\zeta}$ is not a critical point then the tangent space to $\tilde{\mathcal{V}}$ at $\log(\boldsymbol{\zeta})$ is a hyperplane whose normal is not parallel to $\mathbf{r}$. Thus, near any non-critical point in $\mathcal{V} \cap T(\mathbf{w})$ one can locally push the Cauchy domain of integration to points of lower height without crossing the singular set. Because the vector $\mathbf{v}_\zeta$ varies smoothly with ζ, it is possible to make a single deformation of the domain of integration to obtain a contour whose points have height bounded below $h_\mathbf{r}(\mathbf{w})$, except in arbitrarily small neighbourhoods of the critical points in $T(\mathbf{w})$. The stated deformation is constructed[2] in [3, Cor. 5.5], after which asymptotics can be determined using the saddle-point methods discussed above [3, Sec. 6].

A second approach, using homological arguments and multivariate residues, is given by Pemantle and Wilson [32, Thm. 9.4.2]. Under the assumptions of Theorem 5.4, one can use advanced topological arguments to write the Cauchy integral as a sum of an (asymptotically negligible) Cauchy integral over a domain whose points have height bounded below $h_\mathbf{r}(\mathbf{w})$ and a 'multivariate residue integral' over a domain of integration Γ lying in $\mathcal{V}$ (this is basically a more general, but less explicit, version of Corollary 5.1 above). Using a 'gradient flow' of the height function h, created essentially by solving a differential equation, it is possible to deform Γ so that all points not lying in arbitrarily small neighbourhoods of critical points have height bounded below $h_\mathbf{r}(\mathbf{w})$. Pemantle and Wilson were motivated by the study of Stratified Morse theory, and more recent work of Baryshnikov et al. [1] shows how to the methods of Stratified Morse characterize asymptotics under conditions which can be computationally verified. We return to this approach in Chapter 9.

Example 5.11 (Non-Finitely Minimal Critical Point)

Consider the power series expansion of the bivariate rational function

$$F(x,y) = \frac{1}{(1+2x)(1-x-y)}$$

in the main diagonal direction $\mathbf{r} = (1,1)$. The singular variety has one non-smooth point $(x,y) = (-1/2, 3/2)$, which is not minimal as it has greater coordinate-wise modulus than $\boldsymbol{\sigma} = (1/2, 1/2)$. The point $\boldsymbol{\sigma}$ is still a minimal critical point, but it is no longer finitely minimal as

$$\mathcal{V} \cap T(\boldsymbol{\sigma}) = \{(1/2, 1/2)\} \quad \cup \quad \{(-1/2, e^{i\theta}/2) : \theta \in (-\pi, \pi)\}.$$

Because the non-critical points in $\mathcal{V} \cap T(\boldsymbol{\sigma})$ have no bearing on the local behaviour of $F(\mathbf{z})$ near $\boldsymbol{\sigma}$, they do not play a role in asymptotics. Theorem 5.4 implies

$$f_{n,n} = \frac{4^n}{\sqrt{\pi n}}\left(\frac{1}{2} - \frac{1}{8n} + \frac{1}{256n^2} + \frac{5}{256n^3} - \frac{819}{65536n^4} + O\left(\frac{1}{n^5}\right)\right).$$

---

[2] Although the hypotheses of [3, Cor. 5.5] state that all minimizers of $h_\mathbf{r}$ on $\overline{\mathcal{D}}$ have the same coordinate-wise modulus, this is not needed under our assumptions. Thanks to Yuliy Baryshnikov for clarification on the constructions from this paper.

5.3.3 Perturbations of Direction and a Central Limit Theorem

Because our asymptotic theorems have error terms which remain uniformly bounded as the direction $\mathbf{r}$ undergoes small perturbations, the coefficients under consideration grow in a predictable manner near fixed directions. In fact, we prove that in many circumstances the series coefficients scale to a normal distribution when viewed in the right manner. This behaviour is well known, having been derived in some of the first major papers on the asymptotic behaviour of multivariate generating functions [4, 5].

To begin, we examine how the exponential growth in our asymptotic formula changes when the direction under consideration varies slightly. Because scaling the direction of interest does not meaningfully change our results (one can recover asymptotics in the original direction by scaling n in the resulting asymptotic formula) we scale our direction so its final coordinate equals one.

Lemma 5.8 *Suppose rational $F(\mathbf{z}) = G(\mathbf{z})/H(\mathbf{z})$ admits a nondegenerate strictly minimal smooth contributing point $\mathbf{w} \in \mathbb{C}_*^d$ in the direction $\mathbf{m} \in \mathbb{R}^d$, scaled so that $m_d = 1$, and suppose $H_{z_d}(\mathbf{w}) \neq 0$. As $\mathbf{r}$ varies in some sufficiently small neighbourhood of $\mathbf{m}$ let $\mathbf{z}(\mathbf{r})$ be the smooth critical point in the direction $\mathbf{r}$ near $\mathbf{z}(\mathbf{r})$, given by Lemma 5.6. If $\mathbf{r} = (\widehat{\mathbf{m}} + \boldsymbol{\varepsilon}, 1)$ where $\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}(n)$ is a sequence of points in $\mathbb{R}^{d-1}$ with each coordinate in $o(n^{-1/3})$ then*

$$\mathbf{z}(\mathbf{r})^{-n\mathbf{r}} \sim \mathbf{w}^{-n\mathbf{m}} \times \widehat{\mathbf{w}}^{-n\boldsymbol{\varepsilon}} \exp\left[-n\frac{\boldsymbol{\varepsilon}^T \mathcal{H}^{-1} \boldsymbol{\varepsilon}}{2}\right],$$

where $\mathcal{H}$ is the non-singular matrix in Lemma 5.5 corresponding to the direction $\mathbf{m}$.

Proof Since $H_{z_d}(\mathbf{w}) \neq 0$, the implicit function theorem yields an analytic parameterization $z_d = g(\hat{\mathbf{z}})$ of z_d on $\mathcal{V}$ near $\mathbf{w}$. If $p(\hat{\mathbf{x}}) = \log(g(e^{x_1}, \dots, e^{x_{d-1}}))$ then $(\hat{\mathbf{x}}, p(\hat{\mathbf{x}}))$ parametrizes $\log(\mathcal{V})$ near the point $\mathbf{a} = \log(\mathbf{w})$. Letting $\mathbf{x}(\mathbf{r}) = \log(\mathbf{z}(\mathbf{r}))$ be the image of the minimal critical point in direction $\mathbf{r}$ after taking coordinate-wise logarithms, it follows that $\mathbf{z}(\mathbf{r})^{-n\mathbf{r}} = \exp[n\tilde{h}_{\mathbf{r}}(\hat{\mathbf{x}}(\mathbf{r}))]$ where $\tilde{h}_{\mathbf{r}}$ is defined by

$$\tilde{h}_{\mathbf{r}}(\hat{\mathbf{y}}) = -\hat{\mathbf{r}} \cdot \hat{\mathbf{y}} - p(\hat{\mathbf{y}})$$

for $\mathbf{y}$ in a neighbourhood of $\mathbf{a}$. Note that

$$p(\log w_1 + \mathrm{i}\theta_1, \dots, \log w_{d-1} + \mathrm{i}\theta_{d-1}) = \log(g(w_1 e^{\mathrm{i}\theta_1}, \dots, w_{d-1} e^{\mathrm{i}\theta_{d-1}}))$$

equals the function ϕ in Lemma 5.5 up to a linear function in the θ_j. If M denotes the Hessian matrix of p with respect to the variables $\hat{\mathbf{x}}$ at the origin then, since $\mathcal{H}$ is the Hessian of ϕ at the origin, the chain rule implies $M = -\mathcal{H}$.

Because $\mathbf{w}$ is a critical point in direction $\mathbf{m}$, the smooth critical point equations imply the gradient of p with respect to $\hat{\mathbf{x}}$ is $(\nabla p)(\hat{\mathbf{a}}) = -\widehat{\mathbf{m}}$. Thus, for any $\hat{\mathbf{y}}$ sufficiently close to $\hat{\mathbf{a}}$ there is a convergent series expansion

$$\tilde{h}_{\mathbf{r}}(\hat{\mathbf{y}}) = \tilde{h}_{\mathbf{r}}(\hat{\mathbf{a}}) + (\widehat{\mathbf{m}} - \hat{\mathbf{r}}) \cdot (\hat{\mathbf{y}} - \hat{\mathbf{a}}) - \frac{(\hat{\mathbf{y}} - \hat{\mathbf{a}})^T M (\hat{\mathbf{y}} - \hat{\mathbf{a}})}{2} + \cdots .$$

In fact, for any $\mathbf{r}$ sufficiently close to $\mathbf{m}$ the logarithmic critical point $\mathbf{x}(\mathbf{r})$ is defined as the unique point approaching $\mathbf{a} = \log(\mathbf{w})$ such that the gradient of $\tilde{h}_{\mathbf{r}}(\mathbf{x})$ with respect to $\mathbf{x}$ vanishes. Vanishing of the gradient is equivalent to the system $r_k = -p_{x_k}(\hat{\mathbf{x}})$ for $1 \leq k \leq d-1$, so the Jacobian of $\hat{\mathbf{r}}$ as a function of $\hat{\mathbf{x}}$ is $-M$. This, in turn, implies the Jacobian of $\hat{\mathbf{x}}$ as a function of $\hat{\mathbf{r}}$ is $-M^{-1}$. In particular, when each coordinate of $\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}(n)$ is $o(n^{-1/3})$ it follows that

$$\hat{\mathbf{x}}(\mathbf{m} + \boldsymbol{\varepsilon}) = \hat{\mathbf{x}}(\mathbf{m}) - M^{-1}\boldsymbol{\varepsilon} + o(n^{-2/3}) = \hat{\mathbf{a}} - M^{-1}\boldsymbol{\varepsilon} + o(n^{-2/3}).$$

Substituting $\hat{\mathbf{y}} = \hat{\mathbf{x}}(\mathbf{m} + \boldsymbol{\varepsilon})$ and $\hat{\mathbf{r}} = \widehat{\mathbf{m}} + \boldsymbol{\varepsilon}$ into the series expansion of $\tilde{h}_{\mathbf{r}}(\hat{\mathbf{y}})$ gives

$$\begin{aligned}\tilde{h}_{\widehat{\mathbf{m}}+\boldsymbol{\varepsilon}}(\hat{\mathbf{x}}(\mathbf{m} + \boldsymbol{\varepsilon})) &= \tilde{h}_{\widehat{\mathbf{m}}+\boldsymbol{\varepsilon}}(\hat{\mathbf{a}}) + \boldsymbol{\varepsilon}^T M^{-1}\boldsymbol{\varepsilon} - \frac{\left(M^{-1}\boldsymbol{\varepsilon}\right)^T M \left(M^{-1}\boldsymbol{\varepsilon}\right)}{2} + o(n^{-1}) \\ &= -(\mathbf{m} \cdot \mathbf{a}) - (\boldsymbol{\varepsilon} \cdot \hat{\mathbf{a}}) + \frac{\boldsymbol{\varepsilon}^T M \boldsymbol{\varepsilon}}{2} + o(n^{-1}),\end{aligned}$$

and the claimed result follows since $M = -\mathcal{H}$ and $\mathbf{a} = \log(\mathbf{w})$. □

From the smoothly varying behaviour of the exponential growth we are able to deduce how dominant asymptotics transition around a fixed direction.

Proposition 5.9 (Smooth Variation of Coefficients) *Consider a Laurent expansion of a rational function $F(\mathbf{z}) = G(\mathbf{z})/H(\mathbf{z})$, and let $\mathbf{m} \in \mathbb{R}^d$ be a direction with $m_d = 1$. Suppose that for all directions $\mathbf{r} = (\hat{\mathbf{r}}, 1)$ in some neighbourhood of $\mathbf{m}$ there is a smoothly varying nondegenerate strictly minimal critical point $\mathbf{w}(\mathbf{r}) \in \mathbb{C}_*^d$ such that $H_{z_d}(\mathbf{w}(\mathbf{r}))$ and $G(\mathbf{w}(\mathbf{r}))$ are non-zero. If $\hat{\mathbf{s}} = \hat{\mathbf{s}}(n)$ is a sequence of vectors in $\mathbb{Z}^{d-1}$ such that each coordinate of $|\hat{\mathbf{s}} - n\widehat{\mathbf{m}}|$ is $o(n^{2/3})$ then*

$$f_{\hat{\mathbf{s}},n} \sim \mathbf{w}^{-n\mathbf{m}} n^{(1-d)/2} \left(\frac{-G(\mathbf{w})(2\pi)^{(1-d)/2}}{w_d H_{z_d}(\mathbf{w})\sqrt{\det \mathcal{H}}}\right) \widehat{\mathbf{w}}^{-(\hat{\mathbf{s}}-n\widehat{\mathbf{m}})} \exp\left[-\frac{(\hat{\mathbf{s}} - n\widehat{\mathbf{m}})^T \mathcal{H}^{-1} (\hat{\mathbf{s}} - n\widehat{\mathbf{m}})}{2n}\right], \quad (5.29)$$

where $\mathcal{H}$ is the non-singular matrix in Lemma 5.5 corresponding to $\mathbf{w}$.

Remark 5.17 If the coordinates of $\hat{\mathbf{s}} - n\widehat{\mathbf{m}}$ are uniformly bounded for all n (for instance, if $\hat{\mathbf{s}}$ is the closest vector to $n\mathbf{m}$ with integer coordinates) then Proposition 5.9 implies the dominant asymptotic behaviour of $f_{\hat{\mathbf{s}},n}$ is the behaviour predicted on the right-hand side of (5.27) by the vector $\mathbf{r} = (\mathbf{m}, 1)$ with potentially non-integer (even irrational) coordinates, up to the bounded multiplicative factor $\widehat{\mathbf{w}}^{-(\hat{\mathbf{s}}-n\widehat{\mathbf{m}})}$.

Proof Theorem 5.2 implies the existence of a constant $B > 0$ such that

$$\left| f_{n\mathbf{r}} - \mathbf{w}(\mathbf{r})^{-n\mathbf{r}} \frac{(2\pi n)^{(1-d)/2}}{\sqrt{\det \mathcal{H}_{\mathbf{w}(\mathbf{r})}}} C_{\mathbf{w}(\mathbf{r})} \right| \leq \mathbf{w}(\mathbf{r})^{-n\mathbf{r}} n^{(1-d)/2-1} B,$$

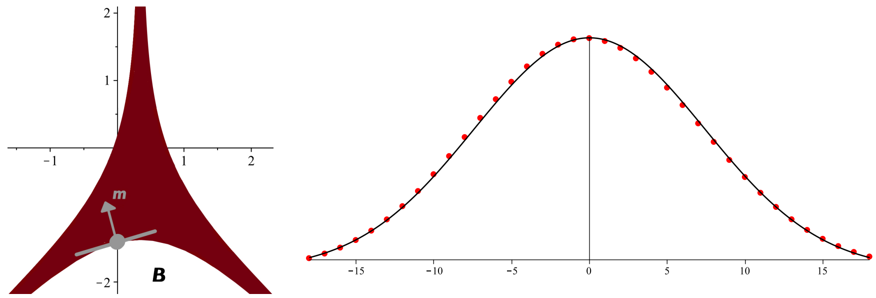


Fig. 5.6 *Left:* The amoeba of $H(x,t) = 1 - t(x + 1 + 2/x)$ with the component B and vector $\mathbf{m} = (-1/4, 1)$ from Proposition 5.10. *Right:* The values of $4^{-n} f_{-n/4+\varepsilon,n}$ with $\varepsilon \in \{-20, \dots, 20\}$ compared to the limiting curve $\nu_n(-n/4+\varepsilon, n)$ for $n = 80$.

where $C_{\mathbf{w}(\mathbf{r})}$ and $\mathcal{H}_{\mathbf{w}(\mathbf{r})}$ approach $-G(\mathbf{w})/w_d H_{z_d}(\mathbf{w})$ and $\mathcal{H}$, respectively, as $\mathbf{r} \to \mathbf{m}$. If $|\hat{\mathbf{s}} - n\widehat{\mathbf{m}}|$ is $o(n^{2/3})$ then $\hat{\mathbf{r}} = \hat{\mathbf{s}}/n = \widehat{\mathbf{m}} + \boldsymbol{\varepsilon}(n)$ where each coordinate of $\boldsymbol{\varepsilon}(n)$ is $o(n^{-1/3})$. Thus, $\hat{\mathbf{r}} \to \widehat{\mathbf{m}}$ as $n \to \infty$ and the result follows from Lemma 5.8. □

Although the term $\widehat{\mathbf{w}}^{-(\hat{\mathbf{s}}-n\widehat{\mathbf{m}})}$ in (5.29) is bounded, its existence prevents us from knowing exact asymptotic behaviour. When $\widehat{\mathbf{w}} = \mathbf{1}$, however, this term disappears and we can be more precise. The following result states that as $n \to \infty$ the coefficients of $f_{\hat{\mathbf{s}},n}$ approach a multivariate normal distribution around the direction $\widehat{\mathbf{m}}$.

Proposition 5.10 (Local Central Limit Theorem) *Consider a combinatorial Laurent expansion of a rational function $F(\mathbf{z}) = G(\mathbf{z})/H(\mathbf{z})$. Suppose that, in some direction $\mathbf{m}$ with $m_d = 1$, there is a nondegenerate strictly minimal smooth contributing point of the form $\mathbf{w} = (\mathbf{1}, t)$ for some $t > 0$. If $H_{z_d}(\mathbf{w})$ and $G(\mathbf{w})$ are non-zero then*

$$\sup_{\hat{\mathbf{s}} \in \mathbb{Z}^{d-1}} n^{(d-1)/2} \left| t^n f_{\hat{\mathbf{s}},n} - C\nu_n(\hat{\mathbf{s}}) \right| \to 0 \tag{5.30}$$

as $n \to \infty$, where $C = -G(\mathbf{w})/\big(w_d H_{z_d}(\mathbf{w})\big)$ and

$$\nu_n(\hat{\mathbf{s}}) = \frac{(2\pi n)^{(1-d)/2}}{\sqrt{\det \mathcal{H}}} \exp\left[-\frac{(\hat{\mathbf{s}} - n\widehat{\mathbf{m}})^T \mathcal{H}^{-1} (\hat{\mathbf{s}} - n\widehat{\mathbf{m}})}{2n} \right]$$

for $\mathcal{H}$ the non-singular matrix in Lemma 5.5 corresponding to $\mathbf{w}$.

A result of the form (5.30) is often called a *local central limit theorem*. Problem 5.7 asks the reader, in essence, to derive 'the' classical local central limit theorem for random variables supported on finite subsets of $\mathbb{Z}^d$ (slightly generalizing Proposition 5.10 from rational to meromorphic functions removes the restriction of finite support); see Durrett [14, Ch. 3] for more traditional derivations of central limit theorems. Before proving Proposition 5.10 we look at a short example, with a more detailed example following the proof.

Example 5.12 (Local Central Limit Theorem for Weighted Walks)

Let $F(x, y) = \frac{1}{1-y(x+1+2/x)}$ be the bivariate generating function

$$F(x, y) = \sum_{i\in\mathbb{Z}, n\geq 0} w_{i,n} x^i y^n$$

where $w_{i,n}$ counts the walks on the integer lattice $\mathbb{Z}$ starting at the origin, ending at $x = i$, and taking steps in the set $\{-1, -1, 0, 1\}$, with two distinct steps in the -1 direction (we have two negative steps to break symmetry). This series expansion of $F(x, y)$ converges in a domain $\mathcal{D}$ containing all points with $x = 1$ and $|y| < 1/4$; the image B of $\mathcal{D}$ under the Relog map is shown on the left of Figure 5.6. Because the hypotheses of Proposition 5.10 are satisfied with $t = 1/4$ and $\mathbf{m} = (-1/4, 1)$, the coefficients of $[y^n]F(x, y) = (x + 1 + 2/x)^n$ behave like a Gaussian curve around their maximum value. The right of Figure 5.6 shows the limit curve compared to actual series coefficients.

Proof (Proposition 5.10) Lemma 5.6 implies there is a non-degenerate smooth critical point $\mathbf{w}(\mathbf{s})$ varying smoothly with $\mathbf{s}$ near $\mathbf{m}$. Since the series under consideration is combinatorial, the points $\mathbf{w}(\mathbf{s})$ all have positive real coordinates, and Proposition 5.8 implies each is minimal. Since $\mathbf{w} = \mathbf{w}(\mathbf{m})$ is strictly minimal, so are the $\mathbf{w}(\mathbf{s})$ when $\mathbf{s}$ is sufficiently close to $\mathbf{m}$.

We begin by bounding the exponential term in ν_n. Because (5.25) expresses the entries of $\mathcal{H}$ by evaluations of derivatives of $H(\mathbf{z})$ at $\mathbf{w} \in \mathbb{R}^d$, the matrix $\mathcal{H}$ is real. The classical *spectral theorem* for real symmetric matrices [22, Thms. 4.1.5 and 4.2.2] then implies the eigenvalues of $\mathcal{H}$ are real and $(\hat{\mathbf{z}}^T\hat{\mathbf{z}})\lambda_{\min} \leq \hat{\mathbf{z}}^T\mathcal{H}\hat{\mathbf{z}} \leq (\hat{\mathbf{z}}^T\hat{\mathbf{z}})\lambda_{\max}$ for all $\hat{\mathbf{z}} \in \mathbb{R}^{d-1}$, where $\lambda_{\min}$ and $\lambda_{\max}$ are the smallest and largest eigenvalues of $\mathcal{H}$, respectively. Since $\mathbf{w}$ is strictly minimal, Lemma 5.4 implies the real part of the function $\phi(\boldsymbol{\theta})$ in (5.21), and thus also the term $\boldsymbol{\theta}^T\mathcal{H}\boldsymbol{\theta}$ in its series expansion at the origin, is strictly positive when $\boldsymbol{\theta} \neq \mathbf{0}$. This implies all eigenvalues of $\mathcal{H}$ are positive, so there exist $C_1, C_2 > 0$ such that for any $\hat{\mathbf{z}} \in \mathbb{R}^{d-1}$ and sufficiently large n

$$\exp\left[-nC_1\hat{\mathbf{z}}^T\hat{\mathbf{z}}\right] \leq \exp\left[-n\frac{\hat{\mathbf{z}}^T\mathcal{H}\hat{\mathbf{z}}}{2}\right] \leq \exp\left[-nC_2\hat{\mathbf{z}}^T\hat{\mathbf{z}}\right] \leq 1. \tag{5.31}$$

Now, fix any $1/2 < p < 2/3$. If every coordinate of $|\hat{\mathbf{s}} - n\widehat{\mathbf{m}}|$ is at most n^p then (5.30) is (5.29) and holds by Proposition 5.9. We thus assume each coordinate of $|\hat{\mathbf{s}} - n\widehat{\mathbf{m}}|$ is at least n^p and prove both $t^n f_{\hat{\mathbf{s}},n}$ and $\nu_n(\hat{\mathbf{s}})$ approach zero faster than $n^{(1-d)/2}$. Under this assumption, (5.31) implies $\nu_n(\hat{\mathbf{s}}) \leq e^{-C_3 n^{1-2p}}$ for some $C_3 > 0$, so $\nu_n(\hat{\mathbf{s}})$ approaches zero faster than any fixed power of n. Similarly, if every coordinate of $|\hat{\mathbf{s}} - n\widehat{\mathbf{m}}|$ is $o(n)$ then $\hat{\mathbf{s}} = n\widehat{\mathbf{m}} + n\boldsymbol{\varepsilon}$ where every coordinate of $\boldsymbol{\varepsilon}$ is $o(1)$; repeating the argument in the proof of Lemma 5.8 then implies that when the coordinates of $|\hat{\mathbf{s}} - n\widehat{\mathbf{m}}|$ are at least n^p then the exponential growth of $t^n f_{\hat{\mathbf{s}},n}$ is at most $e^{-C_4 n^{2p-1}}$ for some $C_4 > 0$. Finally, if each coordinate of $|\hat{\mathbf{s}} - n\widehat{\mathbf{m}}|$ grows at least linearly then $\hat{\mathbf{r}} = \hat{\mathbf{s}}/n$ is bounded away from $\widehat{\mathbf{m}}$. Because the series under consideration is combinatorial, the boundary ∂B of the amoeba complement component corresponding to the series

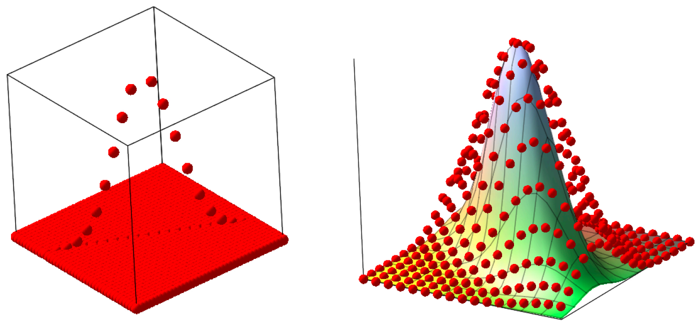


Fig. 5.7 *Left:* The coefficients of $[z^n]P_1(x_1, x_2, z)$ approach a one-dimensional normal distribution as $n \to \infty$. *Right:* The coefficients of $[z^n]P_2(1, x_2, z_3, z)$ approach a two-dimensional normal distribution, with maximum value at the closest integer to $n\mathbf{m}$, as $n \to \infty$.

is a smooth hypersurface near $(\mathbf{0}, \log t)$ with normal $\mathbf{m}$. Thus, any hyperplane through $(\mathbf{0}, \log t)$ with normal $\mathbf{r}$ not parallel to $\mathbf{m}$ intersects the interior of B. Equation (5.15) then implies the exponential growth of $f_{n\mathbf{r}}$ is strictly less than $\mathbf{1}^{\hat{\mathbf{r}}}t = t$. □

Example 5.13 (A Family of Permutations with Restricted Cycles)

For any $d \in \mathbb{N}$ let $\mathcal{F}_d(n)$ be the set of permutations σ on $\{1, \ldots, n\}$ such that $i - d \leq \sigma(i) \leq i + 1$ for all i. Let $\mathcal{F}_d$ be the union of the sets $\mathcal{F}_d(n)$ for all $n \in \mathbb{N}$, and note that every element of $\mathcal{F}_d$, when written in disjoint cycle notation, has cycles of length at most $d+1$. Chung et al. [9] study the combinatorial classes $\mathcal{F}_d$ because of a connection to an algorithm for the generation of random perfect matchings in classes of bipartite graphs. In particular, those authors prove that the number of permutations in $\mathcal{F}_d(n)$ with i_k cycles of length k equals the coefficient $[z_1^{i_1} \cdots z_{d+1}^{i_{d+1}} t^n]P(\mathbf{z}, t)$ in the power series expansion of the rational function

$$P(\mathbf{z}, t) = P(z_1, \ldots, z_{d+1}, t) = \frac{1}{1 - z_1 t - z_2 t^2 - \cdots - z_{t+1} t^{d+1}}.$$

We now show that the joint distribution for the number of cycles satisfies a central limit theorem[3]. The first thing to notice is that the seemingly $(d+1)$-dimensional array of coefficients in $[t^n]P(\mathbf{z}, t)$ is actually d-dimensional, since knowing the coefficient $[\mathbf{z}^{\mathbf{i}} t^n]P(\mathbf{z}, t) \neq 0$ and fixing $i_2, \ldots, i_{d+1}$ and n uniquely determines $i_1 = n - 2i_2 - \cdots - (d+1)i_{d+1}$ (see the left side of Figure 5.7). Thus we set $z_1 = 1$ and examine the coefficients of

$$P(1, \mathbf{z}_{\hat{1}}, t) = \frac{1}{H(z_2, \ldots, z_{d+1}, t)} = \frac{1}{1 - t - z_2 t^2 - \cdots - z_{d+1} t^{d+1}}.$$

To prove a central limit theorem, we start by searching for minimal critical points which satisfy the hypotheses of Proposition 5.10. To that end, set $z_2 = \cdots = z_{d+1} = 1$

[3] Thanks to Persi Diaconis for suggesting this problem.

and consider the roots of $h(t) = H(\mathbf{1}, t) = 1 - t - t^2 - \cdots - t^{d+1}$. Because $h(0)$ is positive, $h(1)$ is negative, and $h'(t)$ is strictly negative for $t > 0$, we see that $h(t)$ has a unique positive root ρ which lies between 0 and 1. At $\boldsymbol{\sigma} = (\mathbf{1}, \rho)$ the logarithmic gradient of the denominator $H(z_2, \dots, z_{d+1}, t)$ equals $-(\rho^2, \dots, \rho^{d+1}, h'(\rho))$, meaning $\boldsymbol{\sigma}$ is a critical point in the rescaled direction

$$\mathbf{m} = \left(\frac{\rho^2}{h'(\rho)}, \dots, \frac{\rho^{d+1}}{h'(\rho)}, 1 \right).$$

Since P is combinatorial, it follows from Proposition 5.4 that $\boldsymbol{\sigma}$ is minimal. Inspection of the smooth critical point equations implies that $\boldsymbol{\sigma}$ is the only critical point in the direction $\mathbf{m}$ so $\boldsymbol{\sigma}$ is strictly minimal by Corollary 5.4 (alternatively, strict minimality follows from Proposition 5.5 and aperiodicity of the denominator). Thus, if the $d \times d$ matrix $\mathcal{H}$ defined by

$$\mathcal{H}_{i,j} = \begin{cases} \frac{\rho^{i+j+1} h''(\rho) - \rho^{i+j}(1+i+j)h'(\rho)}{h'(\rho)^3} & : i \neq j \\ \frac{\rho^{i+j+1} h''(\rho) - \rho^{i+j}(1+i+j)h'(\rho) - \rho^i h'(\rho)^2}{h'(\rho)^3} & : i = j \end{cases}$$

is non-singular then, as $n \to \infty$, Proposition 5.10 implies that the largest coefficients of $[t^n]P(1, \mathbf{z}_{\hat{1}}, t)$ occur at the monomials whose exponents are close to the vector $n\mathbf{m}$. These maximal coefficients approach

$$A_n = \rho^{-n} \frac{-(2\pi n)^{d/2}}{\rho h'(\rho) \sqrt{\det \mathcal{H}}},$$

and

$$\sup_{s_2, \dots, s_{d+1} \in \mathbb{N}} \left| \frac{[z_2^{s_2} \cdots z_{d+1}^{s_{d+1}} t^n] P(1, z_2, \dots, z_t, t)}{A_n} - \nu_n(s_2, \dots, s_{d+1}) \right| \to 0$$

as $n \to \infty$, where $\nu_n(\mathbf{s}) = \exp\left[-\frac{(\mathbf{s} - n\mathbf{m})^T \mathcal{H}^{-1} (\mathbf{s} - n\mathbf{m})}{2n} \right]$. See Figure 5.7 for a plot of actual series coefficients compared to the limiting distribution when $d = 2$. Further details can be found in Melczer [28].

Pemantle and Wilson [32, Sect. 9.6] describe these and additional, ‘weak,’ limit laws holding under weaker assumptions on $F(\mathbf{z})$. Flajolet and Sedgewick [16, Ch. IX] give an in-depth treatment of limit laws for bivariate generating functions. Additional work on limit theorems in combinatorics can be found in Hwang [23, 24], Wallner [34], and the references therein.

5.3.4 Genericity of Assumptions

We end this chapter by showing that our assumptions are satisfied by 'most' rational functions, explaining why they frequently hold in practice[4].

Definition 5.10 (degrees and algebraic sets) The *(total) degree* of a monomial $\mathbf{z}^{\mathbf{n}}$ with $\mathbf{n} \in \mathbb{N}^d$ is the sum $|\mathbf{n}|_1 = n_1 + \cdots + n_d$ of its exponents, and the *(total) degree* of a polynomial is the maximum degree of its monomials. An *algebraic set* in $\mathbb{C}^r$ is any set formed by the common zeroes of a collection of r-variate polynomials with complex coefficients.

The Hilbert basis theorem [10, Sec. 2.5] implies one may always define an algebraic set by the vanishing of a finite collection of polynomials. Problem 5.12 asks you to show that there are $m_\delta = \binom{\delta+d}{d}$ monomials in $\mathbb{C}[\mathbf{z}]$ of degree at most δ.

Definition 5.11 (generic properties of rational functions) A property $\mathcal{P}$ of polynomials in $\mathbb{C}[\mathbf{z}]$ holds *generically* if for every $\delta \in \mathbb{N}$ there exists a proper algebraic subset $C_\delta \subsetneq \mathbb{C}^{m_\delta}$ such that any polynomial of degree δ satisfies $\mathcal{P}$ unless the vector of its coefficients lies in C_δ. A property $\mathcal{P}$ of rational functions holds *generically* if for every pair $\delta_1, \delta_2 \in \mathbb{N}$ there exists a proper algebraic subset $C_{\delta_1,\delta_2} \subsetneq \mathbb{C}^{m_{\delta_1}+m_{\delta_2}}$ such that any rational function with numerator of degree δ_1 and denominator of degree δ_2 satisfies $\mathcal{P}$ unless the vector formed by the coefficients of its numerator and denominator lies in C_{δ_1,δ_2}.

Since the intersection of algebraic sets is algebraic, this definition implies that the conjunction of generic properties is generic. In this section we prove the following.

Proposition 5.11 *Fix* $\mathbf{r} \in \mathbb{R}^d_*$. *Generically, a rational function* $F(\mathbf{z}) = G(\mathbf{z})/H(\mathbf{z})$ *has a smooth singular variety, H is square-free, F admits a finite number of smooth critical points in the direction* $\mathbf{r}$*, each of these critical points is non-degenerate, and the numerator G does not vanish at any of these critical points.*

When at least one of these generically occurring smooth critical points is minimal and contributing, Theorem 5.4 gives an explicit expression for the dominant asymptotic term. We require some algebraic tools before proving Proposition 5.11.

Projective Space, Multivariate Resultants, and Generic Smoothness

The idea behind Proposition 5.11 is that our conditions determine an algebraic system of equations involving both the $\mathbf{z}$ variables and the coefficients of G and H. Consider first the property that the singular variety $\mathcal{V} = \mathcal{V}(H)$ is everywhere smooth; if this property does not hold then the polynomial system

[4] Although a large number of examples that come from combinatorial problems have non-generic behaviour, as we will see in Part III. It is fair to say that some generic properties (like non-degenerate critical points) hold in most applications while others (such as smoothness at all points of the singular variety) merely hold for many applications.

$$H(\mathbf{z}) = H_{z_1}(\mathbf{z}) = \cdots = H_{z_d}(\mathbf{z}) = 0 \tag{5.32}$$

has a solution. If $H(\mathbf{z})$ has degree δ, then we can write

$$H(\mathbf{z}) = \sum_{|\mathbf{i}|_1 \leq \delta} c_{\mathbf{i}} \mathbf{z}^{\mathbf{i}}, \qquad c_{\mathbf{i}} \in \mathbb{C}$$

where, as above, $|\mathbf{i}|_1 = i_1 + \cdots + i_d$ for indices $\mathbf{i} \in \mathbb{N}^d$. Instead of taking a fixed polynomial specified by the coefficients $c_{\mathbf{i}} \in \mathbb{C}$, we now consider a generic polynomial H with degree δ by taking the coefficients $c_{\mathbf{i}}$ as additional variables. In this sense, we consider H to be an element of the polynomial ring $\mathbb{C}[\mathbf{c}][\mathbf{z}]$ with $m_\delta + d$ variables. To show that smoothness is a generic property, one must show that the system (5.32) only has a solution when the coefficient variables $\mathbf{c}$ of H lie in some algebraic set $C_\delta \subset \mathbb{C}^{m_\delta}$ depending only on δ.

Ideally, we would take a solution $(\mathbf{c}, \mathbf{z})$ of this system and project it onto the $\mathbf{c}$ variables to obtain an algebraic equation that must be satisfied by the coefficient variables. Unfortunately, the projection of an algebraic set onto some of its components need not be algebraic.

Example 5.14 (Projections of Algebraic Sets Need Not Be Algebraic)

Consider the algebraic set

$$A = \mathcal{V}(1 - cx) = \{(c, 1/c) : c \in \mathbb{C}_*\} \subset \mathbb{C}^2.$$

Then the projection of A onto its c coordinate is $\mathbb{C}_*$, which is not an algebraic set as any polynomial vanishing on $\mathbb{C}_*$ must vanish on all of $\mathbb{C}$. In particular, the projection is not all of $\mathbb{C}$ but the projection also does not lie in a proper algebraic subset of $\mathbb{C}$.

Essentially, this difficulty comes from the fact that $\mathbb{C}^d$ is not compact so points can 'escape' to infinity. The solution is to work over a different, compact, space.

Definition 5.12 (projective space) The *complex projective space* of dimension n, denoted $\mathbb{P}^n$, consists of all non-zero tuples $(x_0, \ldots, x_n) \in \mathbb{C}^{n+1} \setminus \{\mathbf{0}\}$ modulo the equivalence relation

$$(x_0, \ldots, x_n) \sim (y_0, \ldots, y_n) \text{ if } (x_0, \ldots, x_n) = \lambda(y_0, \ldots, y_n) \text{ for some } \lambda \in \mathbb{C}_*.$$

A point $(x_0, x_1, \ldots, \ldots, x_n) \in \mathbb{C}^{n+1} \setminus \{\mathbf{0}\}$ defines an equivalence class of the relationship $\sim$ which is denoted $(x_0 : x_1 : \cdots : x_n) \in \mathbb{P}^n$.

In order for the zero set of a polynomial $f(x_0, \ldots, x_n)$ to be well defined over $\mathbb{P}^n$, it must depend only on an equivalence class $(x_0 : \cdots : x_n)$. In particular, every monomial in f must have the same degree, so that

$$f(\lambda x_0, \ldots, \lambda x_n) = \lambda^{\deg(f)} f(x_0, \ldots, x_n) = 0$$

whenever $f(x_0, \ldots, x_n) = 0$.

Definition 5.13 (homogeneous polynomials and algebraic sets) A polynomial where every monomial has the same degree is said to be *homogeneous*. We also extend our definition of algebraic sets to different spaces. An *algebraic set*,

- in $\mathbb{P}^n$ is any subset of $\mathbb{P}^n$ formed by the common zeroes of a collection of $(n+1)$-variate complex homogeneous polynomials;
- in the Cartesian product $\mathbb{P}^n \times \mathbb{C}^m$ is any subset of $\mathbb{P}^n \times \mathbb{C}^m$ formed by the common zeroes of a collection of complex polynomials of the form $f(x_0, \ldots, x_n, y_1, \ldots, y_m)$ which are homogeneous in the $\mathbf{x}$ variables (the degree of each monomial with respect to the $\mathbf{x}$ variables only is constant);
- in the Cartesian product $\mathbb{P}^n \times \mathbb{P}^m$ is any subset of $\mathbb{P}^n \times \mathbb{P}^m$ formed by the common zeroes of a collection of complex polynomials of the form $f(x_0, \ldots, x_n, y_0, \ldots, y_m)$ which are homogeneous in the $\mathbf{x}$ variables and the $\mathbf{y}$ variables separately.

Again, the Hilbert basis theorem [10, Sec. 2.5] implies one may always define an algebraic set by the vanishing of a finite collection of polynomials. We introduce projective space for the following property, called *completeness* or *properness*.

Proposition 5.12 *Let $\mathbb{P}^n$ denote complex projective space of dimension n, and let $\mathbb{A}$ denote either complex or projective space of any dimension. If $\pi : \mathbb{P}^n \times \mathbb{A} \to \mathbb{A}$ is the projection map onto $\mathbb{A}$ and $\mathcal{V}$ is any algebraic set in $\mathbb{P}^n \times \mathbb{A}$ then $\pi(\mathcal{V})$ is an algebraic set in $\mathbb{A}$.*

Our proof is based on the one in Mumford [29, Thm. 2.23]. The proof uses the projective Nullstellensatz, a standard result of algebraic geometry which can be found in Mumford [29, Cor. 2.3] and Cox et al. [10, Ch. 8.3].

Proof Suppose first that $\mathbb{A} = \mathbb{C}^m$. Then $\mathcal{V}$ is defined by the vanishing of a finite collection of polynomials $f_1(\mathbf{x}, \mathbf{y}), \ldots, f_r(\mathbf{x}, \mathbf{y})$ on $\mathbb{P}^n \times \mathbb{C}^m$ where each f_j is homogeneous in the $\mathbf{x}$ variables. For fixed $\mathbf{Y} \in \mathbb{C}^m$, the projective Nullstellensatz states that $f_1(\mathbf{x}, \mathbf{Y}) = \cdots = f_r(\mathbf{x}, \mathbf{Y}) = 0$ has a common solution $\mathbf{x} \in \mathbb{P}^n$ if and only if for every $k \in \mathbb{N}$ there exists a monomial of degree k that cannot be written as a $\mathbb{C}[\mathbf{x}]$-linear combination of $f_1(\mathbf{x}, \mathbf{Y}), \ldots, f_r(\mathbf{x}, \mathbf{Y})$. We claim that for every $k \in \mathbb{N}$ the points $\mathbf{Y} \in \mathbb{C}^m$ for which there exists a monomial of degree k that cannot be written as a $\mathbb{C}[\mathbf{x}]$-linear combination of $f_1(\mathbf{x}, \mathbf{Y}), \ldots, f_r(\mathbf{x}, \mathbf{Y})$ form an algebraic set of $\mathbb{C}^m$. Since the arbitrary intersection of algebraic sets is algebraic, this proves the theorem for the case $\mathbb{A} = \mathbb{C}^m$. The case $\mathbb{A} = \mathbb{P}^m$ follows from the case $\mathbb{A} = \mathbb{C}^m$ after noting that $\mathbb{P}^m$ can be written as the union of the sets $U_i = \{\mathbf{y} \in \mathbb{P}^m : y_i = 1\}$ for $0 \leq i \leq d$, each of which is equivalent to $\mathbb{C}^m$.

It remains only to prove our claim. Fix $k \in \mathbb{N}$, let V be the vector space of degree k homogeneous polynomials in $\mathbb{C}[\mathbf{x}]$, and for $1 \leq i \leq r$ let V_i be the vector space of homogeneous polynomials of degree $k - \deg_{\mathbf{x}}(f_i)$ in $\mathbb{C}[\mathbf{x}]$, where $V_i = (0)$ if $k < \deg_{\mathbf{x}}(f_i)$. Then for any $\mathbf{Y} \in \mathbb{C}^m$ the map

$$\begin{aligned} \phi_k : V_1 \times \cdots \times V_r &\to V \\ (g_1, \ldots, g_r) &\mapsto \sum_{i=1}^{r} g_i(\mathbf{x}) f_i(\mathbf{x}, \mathbf{Y}) \end{aligned}$$

is a linear map between the vector spaces $V_1 \times \cdots \times V_r$ and V. After fixing bases of these vector spaces, the map ϕ_k is represented by a matrix $M_{\mathbf{Y}}$ whose entries are complex polynomials in $\mathbf{Y}$. There is a monomial of degree k that cannot be written as a $\mathbb{C}[\mathbf{x}]$-linear combination of $f_1(\mathbf{x}, \mathbf{Y}), \ldots, f_r(\mathbf{x}, \mathbf{Y})$ if and only if the map ϕ_k is not surjective, which happens if and only if all maximal minors of $M_{\mathbf{Y}}$ vanish. The condition of vanishing minors defines an algebraic set in $\mathbb{C}^m$. □

Proposition 5.12 yields a test for proving genericity of properties of homogeneous polynomials. An arbitrary polynomial can be converted into a homogeneous polynomial through the following process.

Definition 5.14 (homogenization) If $f(\mathbf{x}, \mathbf{y})$ is a complex polynomial the *homogenization* of f with respect to $\mathbf{x} = (x_1, \ldots, x_n)$ is the homogeneous polynomial

$$\tilde{f}(x_0, \mathbf{x}, \mathbf{y}) = x_0^{\deg_{\mathbf{x}}(f)} f(x_1/x_0, \ldots, x_n/x_0, \mathbf{y}),$$

where $\deg_{\mathbf{x}}(f)$ denotes the degree of f with respect to the $\mathbf{x}$ variables only.

Corollary 5.7 *Let S be a collection of polynomials in two sets of variables $\mathbf{z}$ and $\mathbf{c}$, of lengths n and m, and let $\tilde{S}$ be the result of homogenizing each polynomial with respect to the $\mathbf{z}$ variables. Suppose there exists $\mathbf{a} \in \mathbb{C}^m$ such that the polynomials in $\tilde{S}$ have no common root in $\mathbb{P}^n$ after the substitution $\mathbf{c} = \mathbf{a}$. Then there exists a proper algebraic subset $C \subsetneq \mathbb{C}^m$ such that $\mathbf{b} \in C$ whenever the polynomials in S have a common root of the form $(\mathbf{w}, \mathbf{b}) \in \mathbb{C}^{n+m}$ for some $\mathbf{w} \in \mathbb{C}^n$.*

Proof The solutions of $\tilde{S}$ over $\mathbb{P}^n \times \mathbb{C}^m$ form an algebraic set, so Proposition 5.12 implies the existence of algebraic $C \subset \mathbb{C}^m$ such that $\mathbf{b} \in C$ if and only if there exists a solution $(w_0, \mathbf{w}, \mathbf{b})$ of $\tilde{S}$. If $(\mathbf{w}, \mathbf{b})$ is a solution of S in $\mathbb{C}^{n+m}$ then $(1, \mathbf{w}, \mathbf{b})$ is a solution of $\tilde{S}$ in $\mathbb{P}^n \times \mathbb{C}^m$, so $\mathbf{b} \in C$. Since $\tilde{S}$ has no solution with $\mathbf{c} = \mathbf{a}$, it follows that $\mathbf{a} \notin C$ and $C \neq \mathbb{C}^m$. □

Remark 5.18 The point of Corollary 5.7 is the properness condition on C, otherwise one can trivially take $C = \mathbb{C}^m$. Working over complex space, instead of projective space, the projection onto the $\mathbf{c}$ variables need not be algebraic, so to prove genericity one must replace the projection by its algebraic closure (the smallest algebraic set it is contained in). As noted in the example above, the algebraic closure can be larger than the projection, and can be all of $\mathbb{C}^m$ even if the projection is missing points.

We have now developed the algebraic framework necessary to start proving Proposition 5.11. For instance, the next lemma implies that generically H is square-free and $\mathcal{V} = \mathcal{V}(H)$ is smooth.

Lemma 5.9 *Generically, the system of equations*

$$H(\mathbf{z}) = H_{z_1}(\mathbf{z}) = \cdots = H_{z_d}(\mathbf{z}) = 0$$

has no solution.

Proof Fix the degree $\delta \geq 1$ of H. Writing

$$H(\mathbf{z}) = \sum_{|\mathbf{i}|_1 \leq \delta} c_{\mathbf{i}} \mathbf{z}^{\mathbf{i}} \in \mathbb{C}[\mathbf{c}][\mathbf{z}]$$

as above, we will use Corollary 5.7 to prove the existence of a proper algebraic set $C_\delta \subsetneq \mathbb{C}^{m_\delta}$ such that the system of equations has a solution only if $\mathbf{c} \in C_\delta$. To show the existence of such a set it is sufficient to exhibit a polynomial H_δ of degree δ such that the homogenization of this system of equations has no solution in $\mathbb{P}^d$. If $H_\delta(\mathbf{z}) = 1 - z_1^\delta - \cdots - z_d^\delta$ then the system of equations becomes

$$z_0^\delta - z_1^\delta - \cdots - z_d^\delta = -\delta z_1^{\delta-1} = \cdots = -\delta z_d^{\delta-1} = 0,$$

which has no solution in projective space as $\mathbf{0} \notin \mathbb{P}^d$ by definition. □

We remark that advanced arguments allow one to be more precise. Given a collection of polynomials $f_1(\mathbf{z}), \ldots, f_r(\mathbf{z})$ of degrees $\delta_1, \ldots, \delta_r$, respectively, write

$$f_j(\mathbf{z}) = \sum_{|\mathbf{i}|_1 \leq \delta_j} c_{j,\mathbf{i}} \mathbf{z}^{\mathbf{i}}$$

for all $1 \leq j \leq r$. Given a polynomial P in the variables $\{u_{j,\mathbf{i}} : |\mathbf{i}|_1 \leq \delta_j, 1 \leq j \leq r\}$ we then let $P(f_1, \ldots, f_r)$ denote the evaluation of P obtained by setting the variable $u_{j,\mathbf{i}}$ equal to the coefficient $c_{j,\mathbf{i}}$.

Proposition 5.13 *For all positive integers* $\delta_0, \ldots, \delta_d$*, there exists an irreducible polynomial* $\text{Res} = \text{Res}_{\delta_0,\ldots,\delta_d} \in \mathbb{Z}[u_{j,\mathbf{i}}]$*, called the* multivariate resultant*, such that* $d+1$ *homogeneous polynomials* $f_0, \ldots, f_d \in \mathbb{C}[z_0, \ldots, z_d]$ *of degrees* $\delta_0, \ldots, \delta_d$ *have a common root in* $\mathbb{P}^d$ *if and only if* $\text{Res}(f_0, \ldots, f_d) = 0$*. The resultant is uniquely defined by the condition* $\text{Res}(z_0^{\delta_0}, \ldots, z_d^{\delta_d}) = 1$.

The techniques used to prove Proposition 5.13, yielding algorithms for calculating the multivariate resultant, are beyond the scope of our work; see Gelfand et al. [17, Ch. 13] or Jouanolou [25] for details on Proposition 5.13 and its proof. For an introductory exposition, and a detailed overview of the computational aspects of the multivariate resultant, see the excellent treatment in Chapter 3 of Cox et al. [11]. Note that the multivariate resultant is defined for *dense* generic polynomial systems: each f_j must contain all monomials of degree at most δ_j, otherwise the resultant may identically vanish or, less pathologically, may become reducible. Refinements of the multivariate resultant to polynomials with different support sets are often studied under the name *sparse resultants* [11, Ch. 7].

The Remaining Properties

Lemma 5.9 implies that generically H is square-free and $\mathcal{V} = \mathcal{V}(H)$ is smooth. We now prove that the remaining properties listed in Proposition 5.11 hold generically.

Numerator does not vanish at critical points

Corollary 5.7 implies that generically G does not vanish at the critical points of H if for every $\delta_1, \delta_2 \in \mathbb{N}$ there exist specific polynomials G and H of degrees δ_1 and δ_2, respectively, such that the homogenization of the polynomial system

$$G(\mathbf{z}) = H(\mathbf{z}) = r_2 z_1 H_{z_1}(\mathbf{z}) - r_1 z_2 H_{z_2}(\mathbf{z}) = \cdots = r_d z_1 H_{z_1}(\mathbf{z}) - r_1 z_d H_{z_d}(\mathbf{z}) = 0$$

has no solution in $\mathbb{P}^d$. If

$$G(\mathbf{z}) = z_1^{\delta_1} \quad \text{and} \quad H(\mathbf{z}) = 1 - z_1^{\delta_2} - \cdots - z_d^{\delta_2}$$

then the system of homogeneous polynomial equations

$$\begin{aligned} G = z_1^{\delta_1} &= 0 \\ \tilde{H} = z_0^{\delta_2} - z_1^{\delta_2} - \cdots - z_d^{\delta_2} &= 0 \\ -\delta_2 r_j z_1^{\delta_2} - \delta_2 r_1 z_j^{\delta_2} &= 0, \qquad j = 2, \ldots, n \end{aligned}$$

has no solution in $\mathbb{P}^d$, as $z_1 = 0$ implies $z_j = 0$ for all j.

Critical points are nondegenerate and finite in number

We prove that generically all critical points are nondegenerate; Lemma 5.6 then implies they are finite in number, finishing off the proof of Proposition 5.11. We may assume H is square-free, as we have already shown this property holds generically.

Consider the matrix $\mathcal{H}$ defined by (5.25) with $\mathbf{w} = \mathbf{z}$. After multiplying every entry of $\mathcal{H}$ by $z_d^3 H_{z_d}^3$ we obtain a polynomial matrix $\tilde{\mathcal{H}}$ whose determinant is an explicit polynomial D in the variables $\mathbf{z}$ and the coefficients of H. Corollary 5.7 implies that to prove all critical points are generically finite it is sufficient to find for each $\delta \in \mathbb{N}$ a polynomial H of degree δ such that the homogenizations of the $d + 1$ equations consisting of $D = 0$ and the smooth critical point equations (5.16) have no common solution in $\mathbb{P}^d$.

Let us take, again, the trusty polynomial $H(\mathbf{z}) = 1 - z_1^{\delta} - \cdots - z_d^{\delta}$. Calculating the quantities in (5.25) and performing some algebraic simplification gives that $\det \tilde{\mathcal{H}} = (-1)^{d-1} \delta^{4(d-1)} (z_1 \cdots z_{d-1})^{\delta} z_d^{\delta(d-1)} \det M$, where M is the $(d-1) \times (d-1)$ matrix

$$M = \begin{pmatrix} z_1^\delta + z_d^\delta & z_2^\delta & z_3^\delta & \cdots & z_{d-1}^\delta \\ z_1^\delta & z_2^\delta + z_d^\delta & z_3^\delta & \cdots & z_{d-1}^\delta \\ z_1^\delta & z_2^\delta & z_3^\delta + z_d^\delta & \cdots & z_{d-1}^\delta \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ z_1^\delta & z_2^\delta & z_3^\delta & \cdots & z_{d-1}^\delta + z_d^\delta \end{pmatrix}$$

whose off diagonal entries are constant down columns. Problem 5.13 asks you to prove that the determinant of M is $z_d^{\delta(d-2)}(z_1^\delta + \cdots + z_d^\delta)$, so the polynomial

$$D = (-1)^{d-1}\ \delta^{4(d-1)}(z_1 \cdots z_d)^\delta\ z_d^{\delta(2d-3)}\ \left(z_1^\delta + \cdots + z_d^\delta\right)$$

is already homogeneous. The smooth critical point equations become

$$\begin{aligned} z_0^\delta - z_1^\delta - \cdots - z_d^\delta &= 0 \\ -\delta r_j z_1^\delta - \delta r_1 z_j^\delta &= 0, \qquad j = 2, \ldots, d \end{aligned}$$

and the homogenized system has no solution in $\mathbb{P}^d$ (the critical point equations imply each z_j^δ is a non-zero scalar multiple of z_1^δ so if the δth powers sum to zero, or any of them is zero, all of them must be zero).

Problems

5.1 Using Theorem 5.2, find dominant asymptotics of the $(1, -2)$-diagonal of $F(x, y) = 1/(1 - x - y)$ when it is expanded as a convergent Laurent series in the domain $\mathcal{D}_3 = \{(x, y) \in \mathbb{C}^2 : 1 + |x| < |y|\}$, proving minimality of any critical points and taking care with the signs and square-free roots involved.

5.2 Figure 3.5 in Section 3.3.2 of Chapter 3 displays amoeba$(1 - x - y - xy^3)$, whose complement in $\mathbb{R}^2$ contains four components. Let $F(x, y) = 1/(1-x-y-xy^3)$ and, for each component of the amoeba complement, determine the directions $\mathbf{r} = (r, s) \in \mathbb{R}^2$ where Proposition 5.7 implies the sequence $[x^{nr} y^{ns}]F(x, y)$ is eventually zero. Are these diagonal sequences identically zero in any other directions?

5.3 Andrade et al. [12] studied the bivariate generating function

$$F(x, y) = \frac{e^{-x-y}}{(1 - e^{-x} - e^{-y})^2},$$

whose coefficients count certain permutation statistics. Give asymptotics of the power series coefficients $[x^{rn} y^{sn}]F(x, y)$ as $n \to \infty$ for all $r, s > 0$.

5.4 Figure 5.4 in Chapter 3 sketches the contour of $Q(x, y) = 1 - x - y - 6xy - x^2y^2$ together with the connected components of amoeba$(Q)^c$. For each of the four directions $\mathbf{r} \in \{\pm 1\}^2$, determine which of the 5 components contains a minimal critical point, and determine which of the 5 convergent Laurent expansions of $1/Q(x, y)$ has an $\mathbf{r}$-diagonal which is eventually zero. Find dominant asymptotics for the main diagonal of the convergent Laurent expansion corresponding to the bounded amoeba component B_2. You may assume that this expansion admits a minimal critical point (we give computational methods for determining minimality in Chapter 7).

5.5 Pantone [30] gave asymptotics for the number of 'singular vector tuples of generic tensors' by analyzing the multivariate generating function

$$F(\mathbf{z}) = \frac{z_1 \cdots z_d}{(1 - z_1) \cdots (1 - z_d) \left(1 - \sum_{i=2}^{d} (i-1) e_i(\mathbf{z})\right)},$$

where $e_i(\mathbf{z})$ is the ith elementary symmetric function

$$e_i(\mathbf{z}) = \sum_{1 \le j_1 < \cdots < j_i \le d} z_{j_1} \cdots z_{j_i}.$$

In particular, the number of 'cubical tensors' in dimension d has as its generating function the main power series diagonal of $F(\mathbf{z})$. Prove that, although the singular variety $\mathcal{V}$ contains non-smooth points, the minimal critical points which determine asymptotics for the number of cubic tensors are determined only by smooth points. Prove that the number of cubic tensors in dimension d is

$$C_d(n) = \frac{(d-1)^{d-1}}{(2\pi)^{(d-1)/2} d^{(d-2)/2} (d-2)^{(3d-1)/2}} (d-1)^{dn} n^{(1-d)/2} \left(1 + O\left(\frac{1}{n}\right)\right).$$

The off-diagonal power series coefficients of $F(\mathbf{z})$ count other (non-cubic) tensors. In three dimensions, determine the directions $\mathbf{r} \in \mathbb{R}^3_{>0}$ where $F(\mathbf{z})$ admits a minimal smooth contributing point.

5.6 Baryshnikov et al. [2] study main diagonal asymptotics for the power series expansion of $1/H_{a,b}(x, y, z)$, where

$$H_{a,b}(x, y, z) = 1 - (x + y + z) + a(xy + xz + yz) + b(xyz)$$

for parameters $a, b \in \mathbb{R}$. Prove that the singular variety of this rational function is smooth unless $4a^3 - 3a^2 + 6ba + b^2 - 4b = 0$. In the smooth cases, determine dominant asymptotics of the main diagonal sequence. You may use without proof the Grace-Walsh-Szegő theorem [6, Thm. 1.1]: if $f(x, y, z)$ is unchanged by permutations of the variables, f is linear in each variable individually, and $f(a, b, c) = 0$ for $|a|, |b|, |c| < r$ then there exists $|d| < r$ with $f(d, d, d) = 0$.

5.7 Let $S(\hat{\mathbf{z}})$ be an aperiodic Laurent polynomial with non-negative coefficients which add up to 1. Prove that Proposition 5.10 applies to the function $F(\mathbf{z}) = 1/(1 - z_d S(\hat{\mathbf{z}}))$ and explicitly determine the limiting function $\nu_n(\hat{\mathbf{s}})$ in (5.30).

5.8 Flajolet and Sedgewick [16, Ex. IX.15] analyze Euclid's gcd algorithm over the finite field of prime order p by deriving the bivariate generating function

$$F(u,z) = \frac{1}{1 - pz - p(p-1)uz} = \sum_{k,n\geq 0} e_{k,n} u^k z^n,$$

where $e_{k,n}$ denotes the number of pairs of polynomials (r, q) over the finite field of order p such that q is monic, $n = \deg q > \deg r$, and Euclid's gcd algorithm applied to (r, q) takes k steps to terminate. Determine, for sufficiently large n, dominant asymptotics of the largest coefficient of $[z^n]F(u, z)$ and the exponent corresponding to this coefficient, then use Proposition 5.10 to show that the distribution of the number of steps in the Euclidean algorithm for a pair of polynomials of degree n satisfies a local central limit theorem.

5.9 Bender and Richmond [5] prove limit theorems for a range of multivariate series, including the power series expansion of the trivariate function

$$R(x,y,z) = \frac{(1 - x + (xy - y - 1)z)(1 - y + (xy - x - 1)z) - xyz}{(1-z)(1 - (x+y+1)z + xyz^2)}$$

which arises from a graph theory application. Determine, for sufficiently large n, dominant asymptotics of the largest coefficient of $r_n(x, y) = [z^n]R(x, y, z)$ and the exponent corresponding to this coefficient, then use Proposition 5.10 to show that the coefficients of $r_n(x, y)$ satisfy a local central limit theorem.

5.10 Ramgoolam et al. [33] use ACSV techniques to study so-called quiver gauge theories. For instance, those authors show that the 'generalized clover quiver class' has multivariate generating function

$$F(\mathbf{z}) = \frac{1}{\prod_{i=1}^{\infty}\left(1 - \sum_{j=1}^{d} z_j^i\right)}.$$

Show that for any direction $\mathbf{r} \in \mathbb{R}^d_{>0}$ the meromorphic function $F(\mathbf{z})$ admits $\boldsymbol{\sigma} = (r_1/|\mathbf{r}|_1, \dots, r_d/|\mathbf{r}|_1)$ as a minimal critical point for its power series expansion and use Corollary 5.3 to determine dominant asymptotics of the $\mathbf{r}$−diagonal, where $|\mathbf{r}|_1 = r_1 + \cdots + r_d$. *Hint:* Write the denominator of F as

$$\left(1 - \sum_{k=1}^{d} z_j\right) \times \prod_{i=2}^{\infty}\left(1 - \sum_{j=1}^{d} z_j^i\right),$$

then show the zeroes of the second factor don't affect minimality of $\boldsymbol{\sigma}$.

5.11 Modify the analysis of Section 5.1 to obtain an integral expression

$$f_{n,n} = \int_{\mathbb{R}} A(\theta) e^{-n\phi(\theta)} d\theta$$

for the power series coefficients of $F(x, y) = (x - y)/(1 - x - y)$ along the main diagonal. What steps in the analysis change with the numerator? Prove that this integral is identically zero.

5.12 Prove that there are $\binom{N+d}{d}$ monomials $\mathbf{z}^{\mathbf{m}}$ of degree at most N. In other words, show that the number of solutions to $a_1 + \cdots + a_d \leq N$ with each $a_i \in \mathbb{N}$ is $\binom{N+d}{d}$.

5.13 Let $M = A + bI_k$ where A is a $k \times k$ matrix whose entries $A_{i,j} = a_j$ are constant down columns, $b \in \mathbb{C}$, and I_k is the $k \times k$ identity matrix. Prove that the determinant of M is $b^{k-1}(a_1 + \cdots + a_k + b)$.

Hint: What are the eigenvalues of M, and what are their multiplicities?

References

1. Y. Baryshnikov, S. Melczer, and R. Pemantle. Stationary points at infinity for analytic combinatorics. *arXiv preprint, arXiv:1905.05250*, 2020. Submitted.
2. Yuliy Baryshnikov, Stephen Melczer, Robin Pemantle, and Armin Straub. Diagonal asymptotics for symmetric rational functions via ACSV. In *29th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms*, volume 110 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages Art. No. 12, 15. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2018.
3. Yuliy Baryshnikov and Robin Pemantle. Asymptotics of multivariate sequences, part III: Quadratic points. *Adv. Math.*, 228(6):3127–3206, 2011.
4. Edward A. Bender. Central and local limit theorems applied to asymptotic enumeration. *J. Combinatorial Theory Ser. A*, 15:91–111, 1973.
5. Edward A. Bender and L. Bruce Richmond. Central and local limit theorems applied to asymptotic enumeration. II. Multivariate generating functions. *J. Combin. Theory Ser. A*, 34(3):255–265, 1983.
6. Julius Borcea and Petter Brändén. The Lee-Yang and Pólya-Schur programs. II. Theory of stable polynomials and applications. *Comm. Pure Appl. Math.*, 62(12):1595–1631, 2009.
7. Chad Brewbaker. A combinatorial interpretation of the poly-Bernoulli numbers and two Fermat analogues. *Integers*, 8:A02, 9, 2008.
8. Peter Bürgisser, Michael Clausen, and M. Amin Shokrollahi. *Algebraic complexity theory*, volume 315 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1997. With the collaboration of Thomas Lickteig.
9. Fan Chung, Persi Diaconis, and Ron Graham. Permanental generating functions and sequential importance sampling. *Advances in Applied Mathematics*, page 101916, 2019.
10. David Cox, John Little, and Donal O'Shea. *Ideals, varieties, and algorithms*. Undergraduate Texts in Mathematics. Springer, New York, third edition, 2007.
11. David A. Cox, John Little, and Donal O'Shea. *Using algebraic geometry*, volume 185 of *Graduate Texts in Mathematics*. Springer, New York, second edition, 2005.
12. Rodrigo Ferraz de Andrade, Erik Lundberg, and Brendan Nagle. Asymptotics of the extremal excedance set statistic. *European J. Combin.*, 46:75–88, 2015.
13. N. G. de Bruijn. *Asymptotic methods in analysis*. Bibliotheca Mathematica. Vol. 4. North-Holland Publishing Co., Amsterdam; P. Noordhoff Ltd., Groningen; Interscience Publishers Inc., New York, 1958.
14. Rick Durrett. *Probability: theory and examples*, volume 31 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, fourth edition, 2010.

15. Wolfgang Ebeling. *Functions of several complex variables and their singularities*, volume 83 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2007. Translated from the 2001 German original by Philip G. Spain.
16. Philippe Flajolet and Robert Sedgewick. *Analytic combinatorics*. Cambridge University Press, Cambridge, 2009.
17. I. M. Gelfand, M. M. Kapranov, and A. V. Zelevinsky. *Discriminants, resultants and multidimensional determinants*. Modern Birkhäuser Classics. Birkhäuser Boston, Inc., Boston, MA, 2008. Reprint of the 1994 edition.
18. Phillip Griffiths and Joseph Harris. *Principles of algebraic geometry*. Wiley-Interscience [John Wiley & Sons], New York, 1978. Pure and Applied Mathematics.
19. Serkan Hoşten and Seth Sullivant. The algebraic complexity of maximum likelihood estimation for bivariate missing data. In *Algebraic and geometric methods in statistics*, pages 123–133. Cambridge Univ. Press, Cambridge, 2010.
20. Lars Hörmander. *The analysis of linear partial differential operators. I*, volume 256 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 1990. Distribution theory and Fourier analysis.
21. Lars Hörmander. *An introduction to complex analysis in several variables*, volume 7 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam, third edition, 1990.
22. Roger A. Horn and Charles R. Johnson. *Matrix analysis*. Cambridge University Press, Cambridge, 1990. Corrected reprint of the 1985 original.
23. Hsien-Kuei Hwang. Large deviations for combinatorial distributions. I. Central limit theorems. *Ann. Appl. Probab.*, 6(1):297–319, 1996.
24. Hsien-Kuei Hwang. Large deviations of combinatorial distributions. II. Local limit theorems. *Ann. Appl. Probab.*, 8(1):163–181, 1998.
25. J.-P. Jouanolou. Le formalisme du résultant. *Adv. Math.*, 90(2):117–263, 1991.
26. Jessica Khera, Erik Lundberg, and Stephen Melczer. Asymptotic enumeration of lonesum matrices. *arXiv preprint, arXiv:1912.08850*, 2019.
27. William Letsou and Long Cai. Noncommutative biology: Sequential regulation of complex networks. *PLOS Computational Biology*, 12(8):e1005089, August 2016.
28. S. Melczer. A limit theorem for permutations with restricted positions. *In preparation*, 2020.
29. David Mumford. *Algebraic geometry. I*. Springer-Verlag, Berlin-New York, 1976. Complex projective varieties, Grundlehren der Mathematischen Wissenschaften, No. 221.
30. Jay Pantone. The asymptotic number of simple singular vector tuples of a cubical tensor. *Online Journal of Analytic Combinatorics*, 12, 2017.
31. Robin Pemantle and Mark C. Wilson. Asymptotic expansions of oscillatory integrals with complex phase. In *Algorithmic probability and combinatorics*, volume 520 of *Contemp. Math.*, pages 221–240. Amer. Math. Soc., Providence, RI, 2010.
32. Robin Pemantle and Mark C. Wilson. *Analytic combinatorics in several variables*, volume 140 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2013.
33. S Ramgoolam, Mark C Wilson, and A Zahabi. Quiver asymptotics: N=1 free chiral ring. *Journal of Physics A: Mathematical and Theoretical*, 53(10):105401, 2020.
34. Michael Wallner. A half-normal distribution scheme for generating functions. *European J. Combin.*, 87:103138, 21, 2020.

Chapter 6
Application: Lattice Walks and Smooth ACSV

The first steps in the path of discovery, and the first approximate measures, are those which add most to the existing knowledge of mankind.
– Charles Babbage

We now apply the asymptotic techniques developed in Chapter 5 to lattice path enumeration problems, using the generating function expressions derived in Chapter 4. As seen in Chapter 4, there is a correspondence between lattice path models having nice combinatorial properties and those whose generating functions admit particularly nice representations as diagonals of rational functions. Because Chapter 5 gives asymptotics in the presence of *smooth* contributing singularities, here we will consider lattice path models whose step sets have lots of symmetries. In addition to the precise results we derive, this chapter serves as an extended illustration of how to apply the powerful machinery of analytic combinatorics in several variables, including the determination of higher-order constants in the resulting asymptotic expansions. Chapter 10 of Part III will return to lattice path enumeration after a theory of analytic combinatorics for non-smooth points is developed in Chapter 9. Several of the asymptotic results in this chapter were originally given by Melczer and Mishna [3] and Melczer [2], on which our presentation is based.

6.1 Asymptotics of Highly Symmetric Orthant Walks

We consider walks in $\mathbb{N}^d$ on a set of steps $\mathcal{S} \subset \{\pm 1, 0\}^d$, where each step $\mathbf{i} \in \mathcal{S}$ is given some positive real weight $w_\mathbf{i} > 0$. As in Chapter 4, we define the weighted characteristic polynomial

$$S(\mathbf{z}) = \sum_{\mathbf{i} \in \mathcal{S}} w_\mathbf{i} \mathbf{z}^\mathbf{i}$$

associated to $\mathcal{S}$. To make sure we are considering walks in the correct dimension, we always make the following assumption.

Standing Assumption: Assume that for each $1 \leq j \leq d$ the set $\mathcal{S}$ contains a step with a positive entry in its jth coordinate (and thus by symmetry also a negative step in its jth coordinate).

Recall from Chapter 4 that $\mathcal{S}$ is called *highly symmetric* if it is symmetric over every axis and *mostly symmetric* if it is symmetric over all but one axis. When $\mathcal{S}$ is highly or mostly symmetric we may assume its axis of non-symmetry, if it exists, lies in its dth coordinate. Thus, we may write

$$S(\mathbf{z}) = \overline{z}_d A(\hat{\mathbf{z}}) + Q(\hat{\mathbf{z}}) + z_d B(\hat{\mathbf{z}})$$

where A, Q, and B are Laurent polynomials which are symmetric in their variables $\hat{\mathbf{z}} = (z_1, \dots, z_{d-1})$ and we again use the notation $\overline{x} = 1/x$ common in lattice path enumeration. Define the multivariate generating function

$$W(\mathbf{z}, t) = \sum_{\substack{\mathbf{i} \in \mathbb{N}^d \\ n \geq 0}} w_{\mathbf{i},n} \mathbf{z}^{\mathbf{i}} t^n,$$

where $w_{\mathbf{i},n}$ counts the number of weighted walks of length n using the steps in $\mathcal{S}$ which begin at the origin, end at $\mathbf{i} \in \mathbb{N}^d$, and never leave $\mathbb{N}^d$. Theorem 4.2 of Chapter 4 gives the representation $W(\mathbf{z}, t) = [\mathbf{z}^{\geq 0}] R(\mathbf{z}, t)$ when

$$R(\mathbf{z}, t) = \frac{(z_1 - \overline{z}_1) \cdots (z_{d-1} - \overline{z}_{d-1}) \left(z_d - \overline{z}_d \frac{A(\hat{\mathbf{z}})}{B(\hat{\mathbf{z}})} \right)}{(z_1 \cdots z_d)(1 - tS(\mathbf{z}))}$$

is expanded in the formal ring $\mathcal{R} = \mathbb{Q}((\mathbf{z}))[[t]] = \mathbb{Q}((z_1)) \cdots ((z_d))[[t]]$. Analytically, this formal expansion corresponds to a convergent Laurent expansion in a domain $\mathcal{D}$ containing any point $(\mathbf{z}, t)$ where $|t|$ is sufficiently smaller than $|z_d|$, which in turn is sufficiently smaller than $|z_{d-1}|$, and so on until $|z_1|$, which is bounded away from zero. Proposition 4.8 of Chapter 4 gives the generating function expression

$$W(\mathbf{1}, t) = \Delta \left(\frac{(1 + z_1) \cdots (1 + z_{d-1}) \left(B(\hat{\mathbf{z}}) - z_d^2 A(\hat{\mathbf{z}}) \right)}{(1 - z_d) B(\hat{\mathbf{z}}) (1 - t z_1 \cdots z_d \overline{S}(\mathbf{z}))} \right) \tag{6.1}$$

for the number of walks of length n starting at the origin and ending anywhere in $\mathbb{N}^d$, where $\overline{S}(\mathbf{z}) = S(z_1, \dots, z_{d-1}, \overline{z}_d)$. Proposition 4.9 of that chapter gives an alternative formula, and similar expressions can be derived for walks ending on some collection of the boundary axes $\{\mathbf{z} \in \mathbb{R}^d : z_j = 0\}$.

Notice that the denominator of the rational function in (6.1) contains multiple factors. If $\mathcal{S}$ is not highly symmetric then $A(\hat{\mathbf{z}}) \neq B(\hat{\mathbf{z}})$, meaning the denominator factor $1 - z_d$ does not divide the numerator and the singular variety under consideration will contain non-smooth points. Although there are situations when only smooth points dictate asymptotics in the mostly symmetric case, these models are better understood after a discussion on non-smooth ACSV, and we return to them in Chapter 10. We will see that the relative weight of steps moving towards or away from the bounding axes plays a deciding role in whether or not smooth points determine asymptotics.

When $\mathcal{S}$ is highly symmetric then $A(\hat{\mathbf{z}}) = B(\hat{\mathbf{z}})$ and $\overline{S}(\mathbf{z}) = S(\mathbf{z})$, so (6.1) becomes

$$W(\mathbf{1}, t) = \Delta\left(\frac{(1 + z_1)\cdots(1 + z_{d-1})(1 + z_d)}{1 - tz_1 \cdots z_d S(\mathbf{z})}\right). \tag{6.2}$$

The rational function in (6.2) now has a smooth singular variety, because the denominator and its partial derivative with respect to t can never simultaneously vanish. The domain of convergence $\mathcal{D}$ is the power series domain of convergence.

6.1.1 Asymptotics for All Walks in an Orthant

We begin by determining asymptotics for the number of walks beginning at the origin and ending anywhere in $\mathbb{N}^d$. Let

$$G(\mathbf{z}) = (1 + z_1)\cdots(1 + z_d) \qquad \text{and} \qquad H(\mathbf{z}, t) = 1 - tz_1 \cdots z_d S(\mathbf{z}),$$

so that we are interested in asymptotics of the *main diagonal* power series coefficients of $F(\mathbf{z}, t) = G(\mathbf{z})/H(\mathbf{z}, t)$. Recall that $\mathbf{z}_{\widehat{k}}$ denotes the vector $\mathbf{z}$ with its kth entry removed. Since we are considering highly symmetric models, for each $1 \le k \le d$ there exist unique Laurent polynomials $A_k(\mathbf{z}_{\widehat{k}})$ and $Q_k(\mathbf{z}_{\widehat{k}})$ such that

$$S(\mathbf{z}) = (\overline{z}_k + z_k)A_k(\mathbf{z}_{\widehat{k}}) + Q_k(\mathbf{z}_{\widehat{k}}), \tag{6.3}$$

where $A_d(\mathbf{z}_{\widehat{d}}) = A(\hat{\mathbf{z}}) = B(\hat{\mathbf{z}})$ as defined above. For any $1 \le k \le d$ the smooth critical point equation

$$tH_t(\mathbf{z}, t) = z_k H_{z_k}(\mathbf{z}, t)$$

from (5.16) of Chapter 5 becomes

$$t(z_1 \cdots z_d)S(\mathbf{z}) = t(z_1 \cdots z_d)S(\mathbf{z}) + tz_k(z_1 \cdots z_d)S_{z_k}(\mathbf{z}),$$

which implies

$$0 = tz_k(z_1 \cdots z_d)S_{z_k}(\mathbf{z}) = t(z_k^2 - 1)(z_1 \cdots z_{k-1}z_{k+1} \cdots z_d)A_k(\mathbf{z}_{\widehat{k}}).$$

This gives a characterization of critical points in the main diagonal direction.

Lemma 6.1 *The point $(\mathbf{x}, t) \in \mathcal{V}$ is a smooth critical point if and only if for each $1 \le k \le d$ either*

- $x_k = \pm 1$, *or*
- *the polynomial $(z_1 \cdots z_{k-1}z_{k+1} \cdots z_d)A_k(\mathbf{z}_{\widehat{k}})$ has a root at $\mathbf{x}_{\widehat{k}}$.*

Note that it is possible to have an infinite set of critical points due to the second condition. This cannot happen for unweighted models in two dimensions (when every weight is one) but does occur for weighted models or when $d \geq 3$.

Example 6.1 (A Curve of Critical Points)

Consider the unweighted highly symmetric model in three dimensions restricted to the non-negative octant defined by the step set

$$\mathcal{S} = \{(-1, 0, \pm 1), (1, 0, \pm 1), (0, 1, \pm 1), (0, -1, \pm 1), (\pm 1, 1, 0), (\pm 1, -1, 0)\}.$$

Some algebraic manipulation shows

$$\begin{aligned} H(x, y, z, t) &= 1 - t(xyz) \sum_{\mathbf{i} \in \mathcal{S}} x^{i_1} y^{i_2} z^{i_3} \\ &= 1 - t(z^2 + 1)(x + y)(xy + 1) - tz(y^2 + 1)(x^2 + 1), \end{aligned}$$

and solving the system of smooth critical point equations gives the two isolated critical points

$$\left(1, 1, 1, \frac{1}{12}\right) \text{ and } \left(-1, -1, -1, \frac{-1}{12}\right)$$

together with a collection of non-isolated critical points defined by

$$\left(x, 1, -1, \frac{1}{4x}\right), \left(x, -1, 1, \frac{1}{4x}\right), \left(1, y, -1, \frac{1}{4y}\right),$$

$$\left(-1, y, 1, \frac{1}{4y}\right), \left(1, -1, z, \frac{1}{4z}\right), \left(-1, 1, z, \frac{1}{4z}\right)$$

for all $x, y, z \in \mathbb{C}_*$. Lemma 5.6 in Chapter 5 immediately implies that the non-isolated critical points are degenerate, which can also be verified by direction calculation (in fact, the Hessian matrix $\mathcal{H}$ in (5.25) has a row of zeroes at each of these points). Thankfully, none of these pathological critical points are minimal: if, for instance, $(x, 1, -1, 1/4x)$ were minimal then Lemma 5.7 in Chapter 5 would imply

$$0 = H\left(|x|, 1, 1, \frac{1}{4|x|}\right) = -\frac{1 + |x|^2}{|x|},$$

which can never occur.

Our observation that the minimal critical points are well behaved in the last example holds more generally.

Proposition 6.1 *The point*

$$\boldsymbol{\sigma} = \left(1, \dots, 1, \frac{1}{S(\mathbf{1})}\right)$$

is a finitely minimal smooth contributing point. All minimal critical points lie in $T(\boldsymbol{\sigma}) \cap \mathcal{V}$*, there are at most* 2^d *critical points in* $T(\boldsymbol{\sigma}) \cap \mathcal{V}$*, and if* $(\mathbf{z}, t)$ *is such a point then* $\mathbf{z} \in \{\pm 1\}^d$.

Proof The point $\boldsymbol{\sigma}$ is critical by Lemma 6.1. Suppose $(\mathbf{x}, t_{\mathbf{x}})$ lies in $D(\boldsymbol{\sigma}) \cap \mathcal{V}$, where we note that any choice of $\mathbf{x}$ uniquely determines $t_{\mathbf{x}}$ on $\mathcal{V}$. Then

$$\left|\sum_{\mathbf{i}\in\mathcal{S}} w_{\mathbf{i}}\mathbf{x}^{\mathbf{i}+\mathbf{1}}\right| = \left|(x_1 \cdots x_d) \sum_{\mathbf{i}\in\mathcal{S}} w_{\mathbf{i}}\mathbf{x}^{\mathbf{i}}\right| = \left|\frac{1}{t_{\mathbf{x}}}\right| \geq S(\mathbf{1}) = \sum_{\mathbf{i}\in\mathcal{S}} w_{\mathbf{i}}.$$

Since $(\mathbf{x}, t_{\mathbf{x}}) \in D(\boldsymbol{\sigma})$ means $|x_k| \leq 1$ for each $1 \leq k \leq d$, and each weight $w_{\mathbf{i}}$ is positive, the only way this can happen is if $|x_k| = 1$ for all $1 \leq k \leq d$ and the monomials $\mathbf{x}^{\mathbf{i}+\mathbf{1}}$ have the same complex argument for all $\mathbf{i} \in \mathcal{S}$.

By symmetry, and the assumption that we take a positive step in each direction, the set $\{\mathbf{x}^{\mathbf{i}+\mathbf{1}} : \mathbf{i} \in \mathcal{S}\}$ contains two elements of the form

$$x_2^{i_2+1} \cdots x_d^{i_d+1} \quad \text{and} \quad x_1^2 x_2^{i_2+1} \cdots x_d^{i_d+1},$$

so x_1^2 is real when these monomials have the same argument. Thus, $x_1 = \pm 1$ and applying the same reasoning to each coordinate gives the stated restriction on $\mathbf{x}$. Since $F(\mathbf{z})$ is combinatorial, Corollary 5.5 in Chapter 5 implies that $(\mathbf{x}, t_{\mathbf{x}}) \in \mathcal{V}$ is a minimal critical point only if $(|\mathbf{x}|, t_{|\mathbf{x}|})$ is also minimal and critical. Since $|\mathbf{x}| = (|x_1|, \dots, |x_d|)$ has non-negative coordinates, for every $1 \leq k \leq d$ the polynomial $A(|\mathbf{x}_{\hat{k}}|) \neq 0$ and Lemma 6.1 implies $(|\mathbf{x}|, t_{|\mathbf{x}|})$ is minimal and critical only if each $|x_k| = 1$. □

Remark 6.1 Our proof of Proposition 6.1 used the explicit representation of critical points in Lemma 6.1 to show any minimal critical point $(\mathbf{x}, t_{\mathbf{x}})$ lies in $T(\boldsymbol{\sigma})$. Alternatively, the smooth critical point equations imply $S_{z_1}(\mathbf{x}) = \cdots = S_{z_d}(\mathbf{x}) = 0$ and one may conclude directly from Proposition 4.5 in Chapter 4, which states that there is a unique positive solution to this system of equations (in this case the point $\mathbf{x} = \mathbf{1}$).

The collection of minimal critical points of $F(\mathbf{z}, t)$ thus form the finite set

$$C = \left\{\left(\mathbf{x}, \frac{1}{S(\mathbf{x})}\right) : \quad \mathbf{x} \in \{\pm 1\}^d, \quad |S(\mathbf{x})| = S(\mathbf{1})\right\}.$$

In order to find asymptotics using Corollary 5.2 from Chapter 5 it remains only to determine the matrix $\mathcal{H}$ whose entries are given in (5.25), noting that for a d-dimensional model the function $F(\mathbf{z}, t)$ has $d + 1$ variables. If $(\mathbf{x}, t) \in \mathcal{V}$ is a critical point with $\mathbf{x} \in \{\pm 1\}^d$ then a direct calculation shows

$$x_i x_j H_{z_i z_j}(\mathbf{x}) = \begin{cases} -1 & : i \neq j \\ -\frac{2x_j A_j(\mathbf{x})}{S(\mathbf{x})} & : i = j \end{cases}$$

while

$$x_j t H_{z_j t}(\mathbf{x}) = -1 \qquad \text{and} \qquad t^2 H_{tt}(\mathbf{x}) = 0$$

for all $0 \leq i \leq j \leq d$, so $\mathcal{H}$ at $(\mathbf{x}, t)$ is the $d \times d$ diagonal matrix

$$\mathcal{H}_{\mathbf{x}} = \frac{2}{S(\mathbf{x})} \begin{pmatrix} x_1 A_1(\mathbf{x}) & 0 & \cdots & 0 \\ 0 & x_2 A_2(\mathbf{x}) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & x_d A_d(\mathbf{x}) \end{pmatrix}.$$

If $(\mathbf{x}, t) \in C$ then $|S(\mathbf{x})| = S(\mathbf{1})$, so

$$\frac{x_k A_k(\mathbf{x})}{S(\mathbf{x})} = \frac{x_k A_k(\mathbf{x})}{(\overline{x}_k + x_k) A_k(\mathbf{x}_{\hat{k}}) + Q_k(\mathbf{x}_{\hat{k}})} = \frac{A_k(\mathbf{1})}{S(\mathbf{1})}$$

since $x_k A_k(\mathbf{x}_{\hat{k}})$ and $Q_k(\mathbf{x}_{\hat{k}})$ must share the same sign. Thus,

$$\mathcal{H}_{\mathbf{x}} = \mathcal{H} = \frac{2}{S(\mathbf{1})} \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_d \end{pmatrix} \tag{6.4}$$

is independent of $\mathbf{x}$, where $a_k = A_k(\mathbf{1})$ is the total weight of the steps in $\mathcal{S}$ that have kth coordinate equal to 1.

Since $G(\mathbf{z}) = (1 + z_1) \cdots (1 + z_n)$ does not vanish at $\boldsymbol{\sigma}$, but vanishes at any other minimal critical point, $\boldsymbol{\sigma}$ is the only point whose asymptotic contribution affects dominant asymptotics of the sequence under consideration. The fact that $\mathcal{H}$ is a diagonal matrix and does not depend on $\mathbf{x}$ helps simplify the calculations for determining higher order asymptotic terms, discussed below. Using the quantities we have computed with Corollary 5.2 from Chapter 5 immediately gives the following.

Theorem 6.1 *Let $\mathcal{S} \subset \{-1, 0, 1\}^d \setminus \{\mathbf{0}\}$ be a set of steps that is symmetric over every axis and moves forwards in each coordinate. Then the number c_n of walks taking n steps in $\mathcal{S}$, beginning at the origin, and never leaving the orthant $\mathbb{N}^2$ has asymptotic expansion*

$$c_n = S(\mathbf{1})^n \, n^{-d/2} \left((a_1 \cdots a_d)^{-1/2} \, \pi^{-d/2} \, S(\mathbf{1})^{d/2} + O\left(\frac{1}{n}\right) \right). \tag{6.5}$$

When every step has weight one then $S(\mathbf{1})$ is simply the number of steps in $\mathcal{S}$, and a_k is the number of steps moving forwards in the kth coordinate.

Example 6.2 (Highly Symmetric Models in Two Dimensions)

There are four non-isomorphic unweighted highly symmetric models in the two-dimensional quarter plane, whose asymptotics are listed in Table 6.1; asymptotics for

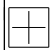
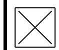
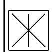
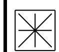
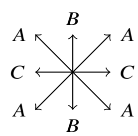
$\mathcal{S}$	Asymptotics	$\mathcal{S}$	Asymptotics
	$4^n\ n^{-1}\ \dfrac{4}{\pi\sqrt{1\times 1}} = \dfrac{4^n}{n}\ \dfrac{4}{\pi}$		$4^n\ n^{-1}\ \dfrac{4}{\pi\sqrt{2\times 2}} = \dfrac{4^n}{n}\ \dfrac{2}{\pi}$
	$6^n\ n^{-1}\ \dfrac{6}{\pi\sqrt{3\times 2}} = \dfrac{6^n}{n}\ \dfrac{\sqrt{6}}{\pi}$		$8^n\ n^{-1}\ \dfrac{8}{\pi\sqrt{3\times 3}} = \dfrac{8^n}{n}\ \dfrac{8}{3\pi}$

Table 6.1 The four highly symmetric models with unit steps in the quarter plane.

these models were originally conjectured by Bostan and Kauers [1]. Corollary 2.1 of Chapter 2 implies that none of these models admits an algebraic generating function.

Example 6.3 (A Weighted Model)

In two dimensions, a weighted lattice path model on the steps $\mathcal{S} = \{\pm 1, 0\}^2 \setminus \{\mathbf{0}\}$ is highly symmetric if and only if each diagonal direction $(\pm 1, \pm 1)$ has the same weight $A > 0$, while the vertical steps $(0, \pm 1)$ both have the same weight $B > 0$, and the horizontal steps $(\pm 1, 0)$ have the same weight $C > 0$:



Under such a weighting,

$$c_n = \left((4A + 2B + 2C)^n\ n^{-1}\ \frac{4A + 2B + 2C}{\pi\sqrt{(2A + B)(2A + C)}} \right) \left(1 + O\left(\frac{1}{n}\right) \right).$$

Theorem 6.1 immediately gives asymptotics for any explicit model, but is also flexible enough to give asymptotics for families of models with varying dimension.

Example 6.4 (Asymptotics for Maximal Step Sets)

Let $\mathcal{S} = \{\pm 1, 0\}^d \setminus \{\mathbf{0}\}$ be the step set containing all non-zero steps, which is highly symmetric when unweighted. Then $S(\mathbf{1}) = |\mathcal{S}| = 3^d - 1$, and $a_k = 3^{d-1}$ for all k, so the total number of walks satisfies

$$c_n = \left((3^d - 1)^n\ n^{-d/2}\ \frac{(3^d - 1)^{d/2}}{3^{d(d-1)/2}\ \pi^{d/2}} \right) \left(1 + O\left(\frac{1}{n}\right) \right).$$

Higher Order Asymptotics

The higher order terms in the asymptotic expansion of c_n are determined by Theorem 5.2 in Chapter 5 through an application of Proposition 5.3; unlike the leading term, the contributions from minimal critical points other than $\boldsymbol{\sigma}$ may play a role. In particular, Theorem 5.2 implies that for any $M \in \mathbb{N}$ the contribution of a minimal critical point $(\mathbf{x}, t) \in C$ to the asymptotics of c_n has the form

$$S(\mathbf{x})^n \; n^{-d/2} \; (2\pi)^{-d/2} \det(\mathcal{H})^{-1/2} \left(\sum_{j=0}^{M} K_j^{\mathbf{x}} n^{-j} + O\left(n^{-M-1}\right) \right),$$

where $\mathcal{H}$ is determined by (6.4). The constant $K_j^{\mathbf{x}}$ satisfies

$$K_j^{\mathbf{x}} = (-1)^j \sum_{0 \le \ell \le 2j} \left. \frac{\mathcal{E}^{\ell+j} \left(P_{\mathbf{x}}(\boldsymbol{\theta}) \, \tilde{\psi}(\boldsymbol{\theta})^{\ell}\right)}{2^{\ell+j} \ell! (\ell+j)!} \right|_{\boldsymbol{\theta}=\mathbf{0}},$$

where

$$P_{\mathbf{x}}(\boldsymbol{\theta}) = \left(1 + x_1 e^{\mathrm{i}\theta_1}\right) \cdots \left(1 + x_d e^{\mathrm{i}\theta_d}\right),$$

$\mathcal{E}$ is the differential operator

$$\mathcal{E} = -\left(\frac{S(\mathbf{1})}{2}\right) \left(\frac{1}{a_1} \partial_1^2 + \cdots + \frac{1}{a_d} \partial_d^2\right), \tag{6.6}$$

and

$$\begin{aligned} \tilde{\psi}(\boldsymbol{\theta}) &= -\log\left(\frac{S(x_1 e^{\mathrm{i}\theta_1}, \ldots, x_d e^{\mathrm{i}\theta_d})}{S(\mathbf{x})}\right) - \left(\boldsymbol{\theta}^t \mathcal{H} \boldsymbol{\theta}\right)/2 \\ &= -\log S(e^{\mathrm{i}\theta_1}, \ldots, e^{\mathrm{i}\theta_d}) + \log S(\mathbf{1}) - \left(\boldsymbol{\theta}^t \mathcal{H} \boldsymbol{\theta}\right)/2. \end{aligned} \tag{6.7}$$

Because $\mathcal{S}$ is highly symmetric $S(e^{\mathrm{i}\theta_1}, \ldots, e^{\mathrm{i}\theta_d})$ is an even function in each variable, which implies its power series expansion at the origin contains only monomials with even exponents. This, in turn, implies the power series expansion of $\tilde{\psi}(\boldsymbol{\theta})$ at the origin also contains only monomials with even exponents. Since, by construction, $\tilde{\psi}$ vanishes to at least order 3 at the origin, it therefore vanishes to order 4 at the origin.

Example 6.5 (Higher Asymptotics Terms of Simple Walks)

Consider the set of cardinal directions $\mathcal{S} = \{(1,0), (0,1), (-1,0), (0,-1)\}$. Then $S(x,y) = x + \overline{x} + y + \overline{y}$ and $A_1(y) = A_2(x) = 1$ so $\mathcal{H}$ is the 2×2 identity matrix multiplied by $1/2$. The set C of critical points consists of two elements,

$$\boldsymbol{\sigma} = (1, 1, 1/4) \qquad \text{and} \qquad \boldsymbol{\varrho} = (-1, -1, -1/4).$$

For $\rho = \pm 1$ let

$$P_\rho(\boldsymbol{\theta}) = \left(1 + \rho e^{\mathrm{i}\theta_1}\right)\left(1 + \rho e^{\mathrm{i}\theta_2}\right).$$

Here $\mathcal{E} = -2(\partial_1^2 + \partial_2^2)$ and

$$\begin{aligned}\tilde{\psi}(\boldsymbol{\theta}) &= -\log\left(e^{\mathrm{i}\theta_1} + e^{-\mathrm{i}\theta_1} + e^{\mathrm{i}\theta_2} + e^{-\mathrm{i}\theta_2}\right) + \log 4 - (\theta_1^2 + \theta_2^2)/4 \\ &= -\log\left(\cos(\theta_1) + \cos(\theta_2)\right) + \log 2 - (\theta_1^2 + \theta_2^2)/4 \\ &= \frac{\theta_1^4}{96} + \frac{\theta_1^2\theta_2^2}{16} + \frac{\theta_2^4}{96} + \cdots .\end{aligned}$$

Thus, for any $M \in \mathbb{N}$ there is an asymptotic expansion

$$c_n = \frac{4^n}{n\pi}\left(\sum_{j=0}^{M}\left[K_j^{(+1)} + (-1)^n K_j^{(-1)}\right] n^{-j} + O\left(n^{-M-1}\right)\right),$$

where

$$K_j^\rho = \sum_{\substack{0 \le \ell \le 2j \\ 0 \le k \le \ell + j}} \frac{(-1)^\ell}{\ell!(\ell+j)!}\binom{\ell+j}{k}\partial_1^{2k}\partial_2^{2\ell+2j-2k}\left(P_\rho(\boldsymbol{\theta})\,\tilde{\psi}(\boldsymbol{\theta})^\ell\right)\Big|_{\boldsymbol{\theta}=\mathbf{0}}.$$

Initial terms can easily be computed in a computer algebra system to obtain

$$c_n = \frac{4^n}{\pi n}\left(4 - \frac{6}{n} + \frac{19 + 2(-1)^n}{2n^2} - \frac{63 + 18(-1)^n}{4n^3} + \cdots\right).$$

6.1.2 Asymptotics for Boundary Returns

It is often of combinatorial interest to determine the number of walks which end on some or all of the boundary hyperplanes of $\mathbb{N}^d$. By permuting coordinates, to enumerate walks returning to r of the boundary hyperplanes it is sufficient to determine the counting sequence (e_n) of walks which return to the boundary axes defined by $z_j = 0$ for $1 \le j \le r$. Applying Proposition 3.14 from Chapter 3 to our representation

$$W(\mathbf{z}, t) = [\mathbf{z}^{\ge 0}]\frac{(z_1 - \overline{z}_1)\cdots(z_{d-1} - \overline{z}_{d-1})(z_d - \overline{z}_d)}{(z_1\cdots z_d)(1 - tS(\mathbf{z}))}$$

for the multivariate generating function $W(\mathbf{z}, t)$ implies that the generating function for e_n, given by $W(\mathbf{z}, t)$ after substituting $z_j = 0$ for $1 \le j \le r$ and $z_j = 1$ otherwise, is the main power series diagonal of

$$F(\mathbf{z},t) = \frac{G(\mathbf{z})}{H(\mathbf{z},t)} = \frac{(1-z_1^2)\cdots(1-z_r^2)(1+z_{r+1})\cdots(1+z_d)}{1-tz_1\cdots z_d S(\mathbf{z})}.$$

As the denominator H is unchanged from above, the set of minimal critical points

$$C = \left\{\left(\mathbf{x}, \frac{1}{S(\mathbf{x})}\right) : \quad \mathbf{x} \in \{\pm 1\}^d, \quad |S(\mathbf{x})| = S(\mathbf{1})\right\}$$

is also unchanged, and asymptotics are still determined by applying Corollary 5.2 from Chapter 5 with the quantities $\mathcal{H}, \mathcal{E}$, and $\tilde{\psi}$ in (6.4), (6.6), and (6.7), respectively. On the other hand, when $r > 0$ the numerator $G(\mathbf{z})$ now vanishes at $\boldsymbol{\sigma} = (\mathbf{1}, 1/S(\mathbf{1}))$ and we must look at higher order terms to determine even the dominant asymptotic growth of e_n.

Although this requires getting down in the computational muck, the vanishing of the numerator actually helps make the computations more tractable. For $(\mathbf{x},t) \in C$ we have $x_j^2 = 1$ for all j, so defining $P_{\mathbf{x}}(\boldsymbol{\theta}) = G(x_1 e^{i\theta_1}, \dots, x_d e^{i\theta_d})$ implies

$$\begin{aligned} P_{\mathbf{x}}(\boldsymbol{\theta}) &= \left(1 - e^{2i\theta_1}\right)\cdots\left(1 - e^{2i\theta_r}\right)\left(1 + x_{r+1}e^{i\theta_{r+1}}\right)\cdots\left(1 + x_d e^{i\theta_d}\right) \\ &= (\theta_1\cdots\theta_r)\Big[(-2i)^r(1+x_{r+1})\cdots(1+x_d) \; + \; \text{higher order terms}\Big]. \end{aligned}$$

Our goal is to determine the smallest $j \in \mathbb{N}$ such that the coefficient

$$K_j^{\mathbf{x}} = (-1)^j \sum_{0\le \ell \le 2j} \frac{\mathcal{E}^{\ell+j}\left(P_{\mathbf{x}}(\boldsymbol{\theta})\,\tilde{\psi}(\boldsymbol{\theta})^\ell\right)}{2^{\ell+j}\ell!(\ell+j)!}\Bigg|_{\boldsymbol{\theta}=\mathbf{0}}$$

is non-zero for some $(\mathbf{x},t) \in C$. Because every term in the power series expansion of $P_{\mathbf{x}}$ at the origin is divisible by $\theta_1\cdots\theta_r$, and because the differential operator $\mathcal{E}$ is a linear combination of $\partial_1^2, \dots, \partial_d^2$, the smallest power of $\mathcal{E}$ which can give a non-zero term when applied to $P_{\mathbf{x}}$ and evaluated at the origin is $\mathcal{E}^r$ (which contains a non-zero multiple of $\partial_1^2\cdots\partial_r^2$ as a summand). Furthermore, the function $\tilde{\psi}(\boldsymbol{\theta})$ vanishes to order 4 at the origin, so for any $\ell \in \mathbb{N}$ the smallest power of $\mathcal{E}$ which can give a non-zero term when applied to $\tilde{\psi}(\boldsymbol{\theta})^\ell$ and evaluated at the origin is $\mathcal{E}^{2\ell}$.

Since the differential operator $\mathcal{E}^{\ell+j}$ has order $2(\ell+j)$, the term

$$\mathcal{E}^{\ell+j}\left(P_{\mathbf{x}}(\boldsymbol{\theta})\,\tilde{\psi}(\boldsymbol{\theta})^\ell\right)\Big|_{\boldsymbol{\theta}=\mathbf{0}}$$

is thus zero unless $2(\ell+j) \ge 2r+4\ell$. The smallest value of j for which this inequality holds is $j = r$, and the inequality holds in this case only when $\ell = 0$. Furthermore, the only term in the power $\mathcal{E}^r$ which can be applied to $P_{\mathbf{x}}(\boldsymbol{\theta})$ to give something non-zero when evaluated at the origin is $\partial_1^2\cdots\partial_r^2$. Since

$$\mathcal{E}^r = \left(-\frac{S(\mathbf{1})}{2}\right)^r \left(\frac{1}{a_1}\partial_1^2 + \cdots + \frac{1}{a_d}\partial_d^2\right)^r = \frac{(-1)^r S(\mathbf{1})^r\, r!}{2^r a_1\cdots a_r}\left(\partial_1^2\cdots\partial_r^2\right) + \cdots$$

and the coefficient of $\theta_1^2 \cdots \theta_r^2$ in the power series expansion of $P_{\mathbf{x}}$ at the origin is $2^r(1+x_{r+1})\cdots(1+x_d)$, it follows that

$$\mathcal{E}^{r+0}\left(P_{\mathbf{x}}(\boldsymbol{\theta})\,\tilde{\psi}(\boldsymbol{\theta})^0\right)\Big|_{\boldsymbol{\theta}=\mathbf{0}} = \frac{(-1)^r S(\mathbf{1})^r\, r!}{2^r a_1\cdots a_r}\times 2^r(1+x_{r+1})\cdots(1+x_d)\left(\partial_1^2\cdots\partial_r^2\right)\left(\theta_1^2\cdots\theta_r^2\right)$$

$$= \frac{(-1)^r S(\mathbf{1})^r(1+x_{r+1})\cdots(1+x_d)\, r!}{a_1\cdots a_r}.$$

If $x_j = -1$ for any $j \geq r+1$ then the term $K_r^{\mathbf{x}}$ is zero. If $x_j = 1$ for all $j \geq r$ then

$$\begin{aligned} K_r^{\mathbf{x}} &= (-1)^r \sum_{0\leq \ell\leq 2r} \frac{\mathcal{E}^{\ell+r}\left(P_{\mathbf{x}}(\boldsymbol{\theta})\,\tilde{\psi}(\boldsymbol{\theta})^\ell\right)}{2^{\ell+r}\ell!(\ell+r)!}\Bigg|_{\boldsymbol{\theta}=\mathbf{0}} \\ &= (-1)^r\, \frac{\mathcal{E}^{r+0}\left(P_{\mathbf{x}}(\boldsymbol{\theta})\,\tilde{\psi}(\boldsymbol{\theta})^0\right)}{2^r r!}\Bigg|_{\boldsymbol{\theta}=\mathbf{0}} \\ &= \frac{S(\mathbf{1})^r 2^{d-r}}{a_1\cdots a_r}. \end{aligned}$$

After this careful set of computations, we have proven the following theorem.

Theorem 6.2 *Let $\mathcal{S} \subset \{-1,0,1\}^d \setminus \{\mathbf{0}\}$ be a set of steps that is symmetric over every axis and moves forwards in each coordinate. Let $\mathcal{A} \subset \{1,\ldots,d\}$ be a set of r elements and let*

$$C_{\mathcal{A}} = \left\{\left(\mathbf{x}, \frac{1}{S(\mathbf{x})}\right) : \quad \mathbf{x}\in\{\pm 1\}^d, \quad x_j = 1 \text{ if } j \notin \mathcal{A}, \quad |S(\mathbf{x})| = S(\mathbf{1})\right\}.$$

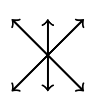
Then the number $e_n^{\mathcal{A}}$ of walks of length n taking steps in $\mathcal{S}$, beginning at the origin, never leaving the orthant $\mathbb{N}^d$, and ending on the intersection of hyperplanes $\{\mathbf{x} : x_j = 0 \text{ for } j \in \mathcal{A}\}$ has asymptotic expansion

$$e_n^{\mathcal{A}} = S(\mathbf{1})^n\, n^{-d/2-r}\, \left(\frac{S(\mathbf{1})^{r+d/2}}{\pi^{d/2}2^r\sqrt{a_1\cdots a_d}\prod_{j\in\mathcal{A}} a_j}\right)\left(\sum_{(\mathbf{x},t)\in C_{\mathcal{A}}} \operatorname{sgn} S(\mathbf{x})^n + O\left(\frac{1}{n}\right)\right). \tag{6.8}$$

Note that the term $\sum_{(\mathbf{x},t)\in C_{\mathcal{A}}} \operatorname{sgn} S(\mathbf{x})^n$ equals $|C_{\mathcal{A}}| > 0$ when n is even, but could be zero when n is odd to account for periodicities.

Example 6.6 (Boundary Walks in Two Dimensions)

Consider the unweighted set of six steps $\mathcal{S} = \{(\pm 1, \pm 1), (0, \pm 1)\}$, visualized as



Then $a_1 = 2$, $a_2 = 3$, and the minimal critical points arising in the ACSV analysis form the set

$$C = \{(1, 1, 1/6),\ (1, -1, 1/6)\}.$$

The number of walks returning to the axis $x = 0$ has dominant asymptotics

$$e_n^{\{1\}} \sim 6^n\ n^{-2}\ \left(\frac{6^2}{\pi\, 2\, \sqrt{a_1 a_2}\, a_1}\right) = 6^n\ n^{-2}\ \left(\frac{3\sqrt{6}}{2\pi}\right),$$

while the number of walks returning to the axis $y = 0$ has dominant asymptotics

$$e_n^{\{2\}} \sim 6^n\ n^{-2}\ \left(\frac{6^2}{\pi\, 2\, \sqrt{a_1 a_2}\, a_2}\right)(1 + (-1)^n) = 6^n\ n^{-2}\ \left(\frac{\sqrt{6}}{\pi}\right)(1 + (-1)^n),$$

and the number of walks returning to the origin $x = y = 0$ has dominant asymptotics

$$e_n^{\{1,2\}} \sim 6^n\ n^{-3}\ \left(\frac{6^3}{\pi\, 2^2\, \sqrt{a_1 a_2}\, a_1 a_2}\right)(1 + (-1)^n) = 6^n\ n^{-3}\ \left(\frac{3\sqrt{6}}{2\pi}\right)(1 + (-1)^n)\,.$$

Note that only even length walks can end at a point with $y = 0$.

6.1.3 Parameterizing the Starting Point

So far we have only considered highly symmetric models whose walks start at the origin in $\mathbb{N}^d$. Problem 6.2 asks you to generalize the methods of Section 4.1.5 to prove that the multivariate generating function $W(\mathbf{z}, t)$ tracking the endpoint and length of highly symmetric walks beginning at $\mathbf{p} \in \mathbb{N}^d$, taking steps in $\mathcal{S}$, and staying $\mathbb{N}^d$ satisfies

$$W(\mathbf{z}, t) = [\mathbf{z}^{\geq 0}]\frac{\left(z_1^{p_1+1} - z_1^{-(p_1+1)}\right)\cdots\left(z_d^{p_d+1} - z_d^{-(p_d+1)}\right)}{(z_1 \cdots z_d)(1 - tS(\mathbf{z}))},$$

where the non-negative extraction occurs in $\mathbb{Q}[\mathbf{z}, \overline{\mathbf{z}}][[t]]$. Thus, the univariate generating function counting the number of walks starting at $\mathbf{p}$ and returning to the set $\{\mathbf{x} \in \mathbb{R}^d : x_1 = \cdots = x_r = 0\}$ is the main power series diagonal of

$$\begin{aligned} F(\mathbf{z}, t) &= \frac{(z_1 \cdots z_d)\left(z_1^{-(p_1+1)} - z_1^{p_1+1}\right)\cdots\left(z_d^{-(p_d+1)} - z_d^{p_d+1}\right)}{(1 - z_{r+1})\cdots(1 - z_d)(1 - tz_1\cdots z_d S(\mathbf{z}))} \\ &= \frac{\mathbf{z}^{-\mathbf{p}}\left(1 - z_1^{2p_1+2}\right)\cdots\left(1 - z_r^{2p_r+2}\right)\ [z_{r+1}]_{p_{r+1}}\cdots[z_d]_{p_d}}{1 - tz_1\cdots z_d S(\mathbf{z})}, \end{aligned} \tag{6.9}$$

where

$$[z_j]_{p_j} = 1 + z_j + \cdots + z_j^{2p_j+1}.$$

Example 6.7 (Walks with Prescribed Start but Free End)

The generating function counting the number of walks beginning at $\mathbf{z} = \mathbf{p}$ and ending anywhere in $\mathbb{N}^d$ is given by

$$\Delta\left(\frac{\mathbf{z}^{-\mathbf{p}}\left(1 + z_1 + \cdots + z_1^{2p_1+1}\right)\cdots\left(1 + z_d + \cdots + z_d^{2p_d+1}\right)}{1 - tz_1\cdots z_d S(\mathbf{z})}\right).$$

For any starting point $\mathbf{p} \in \mathbb{N}^d$, the only minimal critical point where the numerator of this rational function doesn't vanish is $\boldsymbol{\sigma} = (\mathbf{1}, 1/S(\mathbf{1}))$.

To determine asymptotics we again aim to compute the smallest $j \in \mathbb{N}$ such that

$$K_j^{\mathbf{x}} = (-1)^j \sum_{0\le\ell\le 2j} \frac{\mathcal{E}^{\ell+j}\left(P_{\mathbf{x}}(\boldsymbol{\theta})\,\tilde{\psi}(\boldsymbol{\theta})^{\ell}\right)}{2^{\ell+j}\ell!(\ell+j)!}\Bigg|_{\boldsymbol{\theta}=\mathbf{0}}$$

is non-zero for some $(\mathbf{x}, t) \in C$, where the set of minimal critical points C, together with $\tilde{\psi}, \mathcal{H}$, and $\mathcal{E}$ are unchanged from above as the denominator of F is again unchanged, and

$$P_{\mathbf{x}}(\boldsymbol{\theta}) = \mathbf{x}^{-\mathbf{p}}e^{-i\mathbf{p}\cdot\boldsymbol{\theta}}\left(1 - e^{2(p_1+1)i\theta_1}\right)\cdots\left(1 - e^{2(p_r+1)i\theta_r}\right)\left[x_{r+1}e^{i\theta_{r+1}}\right]_{p_{r+1}}\cdots\left[x_{r+1}e^{i\theta_{r+1}}\right]_{p_d}.$$

Since $\tilde{\psi}$ and $\mathcal{E}$ are unchanged from our previous computations, the same arguments as above show that the smallest value of j such that $K_j^{\mathbf{x}}$ could be non-zero is $j = r$. A short computation shows that the coefficient of $\theta_1^2\cdots\theta_r^2$ in the power series expansion of $P_{\mathbf{x}}$ at the origin is

$$\Gamma_{\mathbf{x}} = \mathbf{x}^{-\mathbf{p}}\, 2^r(1+p_1)\cdots(1+p_r)\,[x_{r+1}]_{p_{r+1}}\cdots[x_d]_{p_d},$$

so that

$$K_r^{\mathbf{x}} = (-1)^r\, \frac{\mathcal{E}^{r+0}\left(P_{\mathbf{x}}(\boldsymbol{\theta})\,\tilde{\psi}(\boldsymbol{\theta})^0\right)}{2^r r!}\Bigg|_{\boldsymbol{\theta}=\mathbf{0}} = \frac{S(\mathbf{1})^r}{2^r a_1\cdots a_r}\Gamma_{\mathbf{x}}.$$

If there exists a $j > r$ such that $x_j = -1$ then $[x_j]_{p_j} = 0$, and thus $\Gamma_{\mathbf{x}}$ and $K_r^{\mathbf{x}}$ are zero. Otherwise $[x_j]_{p_j} = [1]_{p_j} = 2p_j + 2$ and, noting that $\mathbf{x}^{-\mathbf{p}} = \mathbf{x}^{\mathbf{p}}$ since $\mathbf{x} \in \{\pm 1\}^d$,

$$K_r^{\mathbf{x}} = \frac{S(\mathbf{1})^r 2^{d-r}}{a_1\cdots a_r}\,\mathbf{x}^{\mathbf{p}}\prod_{j=1}^{d}(1+p_j).$$

Ultimately, we obtain the following generalization of Theorem 6.2.

Theorem 6.3 *Let $\mathcal{S} \subset \{-1,0,1\}^d \setminus \{\mathbf{0}\}$ be a set of steps that is symmetric over every axis and moves forwards in each coordinate. Let $\mathcal{A} \subset \{1,\dots,d\}$ be a set of r elements and let*

$$C_{\mathcal{A}} = \left\{ \left(\mathbf{x}, \frac{1}{S(\mathbf{x})}\right) : \quad \mathbf{x} \in \{\pm 1\}^d, \quad x_j = 1 \text{ if } j \notin \mathcal{A}, \quad |S(\mathbf{x})| = S(\mathbf{1}) \right\}.$$

Then the number $e_n^{\mathcal{A},\mathbf{p}}$ of walks of length n taking steps in $\mathcal{S}$, beginning at $\mathbf{p} \in \mathbb{N}^d$, never leaving the orthant $\mathbb{N}^d$, and ending on the intersection of hyperplanes $\{\mathbf{x} : x_j = 0 \text{ for } j \in \mathcal{A}\}$ has asymptotic expansion

$$e_n^{\mathcal{A},\mathbf{p}} = S(\mathbf{1})^n\; n^{-d/2-r} \left(\frac{S(\mathbf{1})^{r+d/2} \prod_{j=1}^d (1+p_j)}{\pi^{d/2} 2^r \sqrt{a_1 \cdots a_d} \prod_{j \in \mathcal{A}} a_j} \right) \left(\sum_{(\mathbf{x},t) \in C_{\mathcal{A}}} \mathbf{x}^{\mathbf{p}} \;\operatorname{sgn} S(\mathbf{x})^n + O\left(\frac{1}{n}\right) \right).$$

As above, the term $\sum_{(\mathbf{x},t)\in C_{\mathcal{A}}} \mathbf{x}^{\mathbf{p}} \;\operatorname{sgn} S(\mathbf{x})^n$ helps account for underlying periodicities in the model when considering walks returning to the boundary axes.

Example 6.8 (Asymptotics of Walks with Prescribed Start but Free End)

When counting walks beginning at $\mathbf{z} = \mathbf{p}$ and ending anywhere in $\mathbb{N}^d$, the set $C_{\mathcal{A}}$ contains only the point $\boldsymbol{\sigma} = (\mathbf{1}, 1/S(\mathbf{1}))$, and Theorem 6.3 states

$$e_n^{\mathcal{A},\mathbf{p}} = S(\mathbf{1})^n\; n^{-d/2} \left(\frac{S(\mathbf{1})^{d/2} \prod_{j=1}^d (1+p_j)}{\pi^{d/2} \sqrt{a_1 \cdots a_d}} \right) \left(1 + O\left(\frac{1}{n}\right)\right).$$

For any fixed model one may use a computer algebra package to determine the asymptotic series for $e_n^{\mathcal{A},\mathbf{p}}$ to arbitrary precision. Each term will now depend on the starting point $\mathbf{p}$.

Example 6.9 (Higher Order Constants for Simple Walk in Two Dimensions)

Consider the set of cardinal directions $\mathcal{S} = \{(1,0),(0,1),(-1,0),(0,-1)\}$. As above the set C of critical points consists of two elements,

$$\boldsymbol{\sigma} = (1,1,1/4) \qquad \text{and} \qquad \boldsymbol{\varrho} = (-1,-1,-1/4),$$

the differential operator $\mathcal{E} = -2(\partial_1^2 + \partial_2^2)$ and

$$\tilde{\psi}(\boldsymbol{\theta}) = -\log(\cos(\theta_1) + \cos(\theta_2)) + \log 2 - (\theta_1^2 + \theta_2^2)/4.$$

For $\rho = \pm 1$ and $(a,b) \in \mathbb{N}^2$ we now define

$$P_\rho^{a,b}(\boldsymbol{\theta}) = e^{-\mathrm{i}a\theta_1}\left(1+\cdots+\rho^{2a+1}e^{\mathrm{i}(2a+1)\theta_1}\right)\; e^{-\mathrm{i}b\theta_2}\left(1+\cdots+\rho^{2b+1}e^{\mathrm{i}(2b+1)\theta_1}\right)$$
$$= e^{-\mathrm{i}(a\theta_1+b\theta_2)}\left(\frac{1-e^{\mathrm{i}2(a+1)\theta_1}}{1-\rho e^{\mathrm{i}\theta_1}}\right)\left(\frac{1-e^{\mathrm{i}2(b+1)\theta_2}}{1-\rho e^{\mathrm{i}\theta_2}}\right).$$

Then the number of walks on the steps $\mathcal{S}$ beginning at (a,b) and ending anywhere in $\mathbb{N}^2$ has an asymptotic expansion

$$c_n^{a,b} = \frac{4^n}{n\pi}(1+a)(1+b)\left(\sum_{j=0}^{M}\left[K_j^{+1}+(-1)^n K_j^{-1}\right]n^{-j} + O\left(n^{-M-1}\right)\right),$$

where

$$K_j^\rho = \sum_{\substack{0\le \ell\le 2j\\ 0\le k\le \ell+j}} \frac{(-1)^\ell}{\ell!(\ell+j)!}\binom{\ell+j}{k}\partial_1^{2k}\partial_2^{2\ell+2j-2k}\left(P_\rho^{a,b}(\boldsymbol{\theta})\,\tilde{\psi}(\boldsymbol{\theta})^\ell\right)\Big|_{\boldsymbol{\theta}=\mathbf{0}}.$$

Computing the first terms of this expansion gives

$$c_n = \frac{4^n(a+1)(b+1)}{\pi n}\left(4 - \frac{2(2a^2+2b^2+4a+4b+9)}{3n} + \frac{p(a,b)+90(-1)^{a+b+n}}{90n^2} + O\left(\frac{1}{n^3}\right)\right),$$

where $p(a,b)$ is an explicit integer polynomial of degree 4 which can be found in the computer algebra worksheet corresponding to this example.

Problems

6.1 For what dimensions $d \in \mathbb{N}$ does the number of unweighted walks of length n on the maximal step set $\mathcal{S} = \{\pm 1, 0\}^d \setminus \{\mathbf{0}\}$, beginning at the origin and restricted to $\mathbb{N}^d$, have an algebraic generating function?

6.2 Fix $\mathbf{p} \in \mathbb{N}^d$ and a highly symmetric step set $\mathcal{S}$. Prove that the multivariate generating function $W(\mathbf{z},t)$ tracking the endpoint and length of walks beginning at $\mathbf{p}$, taking steps in $\mathcal{S}$, and staying $\mathbb{N}^d$ satisfies the kernel-like functional equation

$$(z_1\cdots z_d)W(\mathbf{z},t) = z_1^{p_1+1}\cdots z_d^{p_d+1} + t(z_1\cdots z_d)S(\mathbf{z})W(\mathbf{z},t)$$
$$- t\sum_{V\subset[d]}(-1)^{|V|}(z_1\cdots z_d)S(\mathbf{z})W(\mathbf{z},t)\Big|_{z_j=0,\, j\in V},$$

generalizing (4.12). Use the kernel method for highly symmetric models presented in Section 4.1.5 of Chapter 4 to show

$$W(\mathbf{z},t) = [\mathbf{z}^{\geq 0}]\frac{\left(z_1^{p_1+1} - z_1^{-(p_1+1)}\right)\cdots\left(z_d^{p_d+1} - z_d^{-(p_d+1)}\right)}{(z_1\cdots z_d)(1-tS(\mathbf{z}))},$$

where the non-negative extraction occurs in $\mathbb{Q}[\mathbf{z},\overline{\mathbf{z}}][[t]]$.

6.3 If $\mathbf{e}^{(j)}$ denotes the jth elementary basis vector with a 1 in position j and all other entries 0, find asymptotics for the number of walks in $\mathbb{N}^d$ which use the step set $\mathcal{S} = \{\pm\mathbf{e}^{(1)}, \dots, \pm\mathbf{e}^{(d)}\}$, begin at the origin, and end anywhere in $\mathbb{N}^d$. What about walks returning to the origin?

6.4 For each of the four highly symmetric step sets $\mathcal{S} \subset \{\pm 1, 0\}^2 \setminus \{\mathbf{0}\}$, find dominant asymptotics for the number of unweighted walks which begin at a point $(a,b) \in \mathbb{N}^2$ and end (i) anywhere in $\mathbb{N}^2$, (ii) on the x-axis, (iii) on the y-axis, and (iv) at the origin.

References

1. Alin Bostan and Manuel Kauers. Automatic classification of restricted lattice walks. In *21st International Conference on Formal Power Series and Algebraic Combinatorics (FPSAC 2009)*, Discrete Math. Theor. Comput. Sci. Proc., AK, pages 201–215. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2009.
2. Stephen Melczer. *Analytic Combinatorics in Several Variables: Effective Asymptotics and Lattice Path Enumeration*. PhD thesis, University of Waterloo and ENS Lyon, 2017. Thesis (PhD).
3. Stephen Melczer and Marni Mishna. Asymptotic lattice path enumeration using diagonals. *Algorithmica*, 75(4):782–811, 2016.

Chapter 7
Automated Analytic Combinatorics

It seems to us obvious. . . to bring out a double set of results, viz. 1st, the numerical magnitudes which are the results of operations performed on numerical data. . . 2ndly, the symbolical results to be attached to those numerical results, which symbolical results are not less the necessary and logical consequences of operations performed upon symbolical data, than are numerical results when the data are numerical.
— Ada Augusta, Countess of Lovelace

After a time many discrepancies occurred, and at one point these discordances were so numerous that I exclaimed, 'I wish to God these calculations had been executed by steam'.
— Charles Babbage

We now turn to the task of automating the asymptotic methods discussed in Chapter 5. A generic rational function has a smooth singular set and admits a finite number of critical points, defined by the polynomial equations (5.16) in Chapter 5. Thus, the bulk of our work involves manipulating the algebraic critical points in order to algorithmically determine which (if any) are minimal. We will make use of *symbolic-numeric algorithms*, using separation bounds on the roots of univariate integer polynomials to determine an accuracy level so that numeric approximation to such accuracy allows us to perform exact calculations.

In order to simplify our presentation we focus on algorithms for asymptotics of power series expansions in positive directions $\mathbf{r} \in \mathbb{Z}^d_{>0}$. The algorithms we develop extend easily to convergent Laurent expansions and general directions $\mathbf{r} \in \mathbb{R}^d_*$.

Our Complexity Model

Our algorithms will take as input multivariate integer polynomials. We study the *bit complexity* of the algorithms, meaning we count the number of additions, subtractions, and multiplications computed by the algorithms when representing all integers in binary. Measuring binary operations gives a more realistic estimate of the time to run an algorithm than counting the number of integer operations, since computer processors perform operations on numbers of bounded size.

The bit complexities of our algorithms will be expressible in terms of the number of variables, the degrees, and the coefficient sizes of their input polynomials. The coefficient size of a polynomial is captured by the following concept.

Definition 7.1 (polynomial heights) The *height* of a multivariate polynomial $P \in \mathbb{Z}[\mathbf{z}]$, denoted $h(P)$, is the base 2 logarithm of the maximum absolute value of its non-zero coefficients. For instance, the height of $P(x, y) = 7 - 11x^5y + 4y^2$ is $h(P) = \log_2 11$. We define the height of the zero polynomial to be zero.

The height of a polynomial gives a bound on the bitsizes of its coefficients and, with the degree of the polynomial, a lower bound on how close its distinct roots can be.

Definition 7.2 (multivariate complexity) Extending from our univariate notation, if $f, g : \mathbb{N}^m \to \mathbb{R}$ are two multivariate sequences then we write

$f = O(g)$ if there exists a constant $M > 0$ such that $|f(\mathbf{n})| \leq M|g(\mathbf{n})|$ for all $\mathbf{n} \in \mathbb{N}^d$.

$f = \tilde{O}(g)$ if $f = O(g \log^k |g|)$ for some natural number k.

When considering the complexity of an algorithm, we often reserve $\delta \geq 2$ for a degree bound on the input polynomials in $\mathbb{Z}[\mathbf{z}] = \mathbb{Z}[z_1, \dots, z_d]$ and define $D = \delta^d$.

The complexities of most of our algorithms are dominated by a factor of the form $\tilde{O}(D^c)$ for some $c \in \mathbb{N}$, and we try to minimize the exponent c. Although the complexity will be singly exponential in the degree and number of variables, this is polynomial in the number coefficients appearing in a generic polynomial of degree δ in d variables, and polynomial in the generic number of critical points. Additional refinements of our algorithms which better capture sparsity in the input polynomials can be found in the paper of Melczer and Salvy [26]. The presentation of this chapter, including the details of the algorithms, is based on the paper of Melczer and Salvy.

7.1 An Overview of Results and Computations

We begin by sketching our main results and viewing some examples, before giving a high-level overview of our algorithms. Recall from Definition 5.8 in Chapter 5 that a function $F(\mathbf{z})$ is *combinatorial* if its power series expansion contains only a finite number of negative terms. Because Lemma 5.7 in Chapter 5 gives an efficient test for minimality when $F(\mathbf{z})$ is combinatorial, it will be significantly easier to determine asymptotics under this assumption. After some additional background, our main result is stated in Theorem 7.1 of Section 7.2, which discusses the complexity and correctness of our main algorithms: Algorithm 1 (which computes dominant asymptotics given the minimal critical points), Algorithm 2 (which determines minimal critical points in the combinatorial case), and Algorithm 3 (which determines minimal critical points in the general case). The necessary algebraic machinery and data structures for the algorithms are developed in Section 7.3, and rely on certain algebraic bounds and algorithms for univariate polynomials discussed in the appendix to this chapter.

Our algorithms start with polynomials $G, H \in \mathbb{Z}[\mathbf{z}]$ of degrees at most δ and heights at most h (meaning the coefficients of the polynomials have absolute values at most 2^h), and a fixed direction $\mathbf{r} \in \mathbb{Z}^d_{>0}$. Because the constants appearing in the asymptotic expansion (5.27) given by Theorem 5.2 in Chapter 5 are defined by potentially high degree algebraic numbers, it may not be desirable (or even possible) to represent them exactly using radicals. Thus, under verifiable and mostly generic assumptions, our algorithms encode asymptotics by returning three rational functions $A, B, C \in \mathbb{Z}(u)$, a square-free polynomial $P \in \mathbb{Z}[u]$, and a list U of roots of $P(u)$, specified by regions of $\mathbb{C}$ containing exactly one root of P, such that

$$f_{n\mathbf{r}} = (2\pi)^{(1-d)/2} \left(\sum_{u \in U} A(u)\sqrt{B(u)}\, C(u)^n \right) (r_d n)^{(1-d)/2} \left(1 + O\left(\frac{1}{n}\right)\right).$$

In the combinatorial case, Algorithms 1 and 2 return this expansion using a probabilistic method in $\tilde{O}(h\delta^{3+4d})$ bit operations. In the general case, due to the increased difficulty in determining minimal critical points, Algorithms 1 and 3 return this expansion using a probabilistic method in $\tilde{O}\left(h\delta^{12+12d}\right)$ bit operations. In either case, the values of $A(u)$, $B(u)$, and $C(u)$ can be determined to precision $2^{-\kappa}$ at all elements of U in $\tilde{O}(\kappa\delta^d + h\delta^{3+3d})$ bit operations.

The probabilistic aspects of our algorithms come from methods for computing a certain algebraic representation of the smooth critical points, known as a *Kronecker representation*. One can compute this representation deterministically using Gröbner bases (see Section 7.3) but doing so loses the complexity estimates we derive. Melczer and Salvy [26] gave a Maple implementation[1] of the algorithms for the combinatorial case. In practice the probabilistic nature of these algorithms does not have a large impact.

Example 7.1 (Automated Apéry Asymptotics)

As seen in previous chapters, the sequence of Apéry numbers is the main diagonal of the rational function $1/(1 - w(1+x)(1+y)(1+z)(1+y+z+yz+xyz))$. Running the Maple commands

```
> F = 1/(1-w*(1+x)*(1+y)*(1+z)*(1+y+z+y*z+x*y*z)):
> A, U = DiagonalAsymptotics(numer(F),denom(F),[w,x,y,z],u,n);
```

after importing the code of Melczer and Salvy verifies that certain necessary assumptions are satisfied (other than combinatoriality, which can be verified by inspection) and returns the expression

$$f_{n,n,n,n} \sim \left(\frac{u+81}{17u-15}\right)^n \left(\frac{1}{4\pi^{3/2}n^{3/2}}\right) \sqrt{\frac{u+81}{28-24u}}$$

where u is the root of

$$P(u) = u^2 + 162u - 167$$

which is approximately $1.0243\ldots$. The algorithm knows that the root u is real, and determines it to enough precision to uniquely identify it among the roots of P. In this case P is quadratic, so u can be expressed with a square-root to obtain

[1] A link to this package is available on the book website.

$$f_{n,n,n,n} \sim \frac{\sqrt{48+34\sqrt{2}}\left(17+12\sqrt{2}\right)^n}{8\pi^{3/2}n^{3/2}},$$

as was derived by hand in Chapter 5.

Remark 7.1 There is some randomness in our algorithms, so a reader running our examples on their computer may obtain different expressions and polynomials $P(u)$. Of course, all encodings yield the same asymptotic expansions.

Example 7.2 (A Sequence Alignment Problem)

Sequence alignment problems arise in molecular biology when trying to determine evolutionary relationships between species by comparing the 'closeness' of different sequences (see, for instance, Waterman [17, Ch. 39]). Pemantle and Wilson [33] used multivariate generating functions to enumerate certain families of sequence alignments parametrized by two natural numbers k and b. For fixed $k, b \in \mathbb{N}$ our algorithm rigorously computes asymptotics of the sequence of interest. When $k = b = 2$ the sequence of interest is the main power series diagonal of

$$F(x, y) = \frac{x^2y^2 - xy + 1}{1 - (x + y + xy - xy^2 - x^2y + x^2y^3 + x^3y^2)},$$

which is combinatorial. Running our algorithm gives that the main power series diagonal has asymptotic growth

$$\left(\frac{10\,u^4 - 40\,u^3 + 54\,u^2 - 26\,u + 4}{4\,u^4 - 19\,u^3 + 25\,u^2 - 4\,u - 6}\right)^n \frac{\left(4\,u^4 - 14\,u^3 + 14\,u^2 - 2\,u + 2\right)}{\sqrt{n}\sqrt{2\pi}\left(10\,u^4 - 40\,u^3 + 54\,u^2 - 26\,u + 4\right)} \times \sqrt{\frac{10\,u^4 - 40\,u^3 + 54\,u^2 - 26\,u + 4}{4\,u^4 - 16\,u^3 + 20\,u^2 - 8\,u + 4}},$$

where u is the real degree 5 algebraic number defined by

$$P(u) = 2u^5 - 10u^4 + 18u^3 - 13u^2 + 4u - 2 = 0, \qquad u \approx 1.4704\ldots.$$

Note that the roots of P cannot be expressed in radicals[2].

[2] The Galois group of P is the symmetric group S_5.

7.1.1 Surveying the Computations

Before introducing the necessary algebra to fully describe our data structures and algorithms, we discuss the computations to be carried out. At the highest level, Theorem 5.4 in Chapter 5 implies that the following steps determine asymptotics.

DiagonalAsymptotics

INPUT: Polynomials $G(\mathbf{z}), H(\mathbf{z}) \in \mathbb{Z}[\mathbf{z}]$ and direction $\mathbf{r} \in \mathbb{Z}^d_{>0}$
OUTPUT: Asymptotics of the $\mathbf{r}$-power series diagonal of $F(\mathbf{z}) = G(\mathbf{z})/H(\mathbf{z})$

1. If G and H are not coprime, divide each by their greatest common divisor.
2. If $H(\mathbf{0}) = 0$ then Fail.
3. Determine the square-free factorization H^s of H
4. Let C denote the zeroes of the integer polynomial system

$$H^s = z_1 H^s_{z_1} - r_1\lambda = \cdots = z_d H^s_{z_d} - r_d\lambda = 0 \tag{7.1}$$

   in the variables $\mathbf{z}, \lambda$. If C is not finite then Fail.
5. Determine the set $\mathcal{U} \subset C$ of minimal critical points. If $\mathcal{U} = \emptyset$ then Fail.
6. If G vanishes at each element of $\mathcal{U}$, if the determinant of the Hessian $\mathcal{H}$ defined in (5.25) of Chapter 5 is zero at an element of $\mathcal{U}$, or if $\mathcal{U}$ contains non-smooth points then Fail.
7. Sum the leading asymptotic contributions given by Theorem 5.3 in Chapter 5 for each element of $\mathcal{U}$, and return the result.

DiagonalAsymptotics is fleshed out in Algorithm 1. Determining the critical point solutions of (7.1) requires manipulating algebraic quantities, and can be easily implemented using classical algebraic tools discussed in Section 7.3. Verifying the necessary conditions for our methods is also easy, for instance when H is square-free then $(\mathbf{z}, \lambda) \in C$ satisfies $\lambda = 0$ if and only if $\mathbf{z}$ is a non-smooth point of the singular variety or some coordinate of $\mathbf{z}$ is zero. In contrast, the determination of *minimal* critical points can be quite delicate and requires further investigation.

7.1.2 Minimal Critical Points in the Combinatorial Case

Assume first that $F(\mathbf{z})$ is combinatorial and admits a finite number of critical points. Corollary 5.5 from Chapter 5 implies that every minimal critical point has the same coordinate-wise modulus as a critical point with positive coordinates, and Corollary 5.6 implies there can be at most one minimal critical point with positive coordinates. Furthermore, Proposition 5.4 in Chapter 5 implies that minimality can be determined by examining line segments from the origin to critical points with positive coordinates. Because critical points may have large degree algebraic coefficients, testing minimality for such points is done implicitly by introducing the polynomial $H(tz_1, \dots, tz_d) \in \mathbb{Z}[\mathbf{z}, t]$ into the critical point system (7.1) and searching for real solutions with $t \in (0, 1)$.

MinimalCriticalComb

INPUT: Coprime polynomials $G(\mathbf{z}), H(\mathbf{z}) \in \mathbb{Z}[\mathbf{z}]$ and direction $\mathbf{r} \in \mathbb{Z}^d_{>0}$
ASSUMING: $F(\mathbf{z}) = G(\mathbf{z})/H(\mathbf{z})$ is combinatorial
OUTPUT: The set $\mathcal{U}$ of minimal critical points of $F(\mathbf{z})$ in the direction $\mathbf{r}$

1. If H is not square-free then replace H by its square-free part
2. Let $\mathcal{S}$ denote the zeroes of the integer polynomial system

$$H(\mathbf{z}) = z_1 H_{z_1}(\mathbf{z}) - r_1\lambda = \cdots = z_d H_{z_d}(\mathbf{z}) - r_d\lambda = H(tz_1, \dots, tz_d) = 0 \qquad (7.2)$$

   in the variables $\mathbf{z}, \lambda, t$. If $\mathcal{S}$ is not finite or if $\mathcal{S}$ contains a point of the form $(\mathbf{z}, \lambda, t)$ where $\lambda = 0$ and $t = 1$ then Fail.
3. Let $C' = \{\mathbf{z} \in \mathbb{C}^d_* : (\mathbf{z}, \lambda, 1) \in \mathcal{S}$ for some $\lambda \in \mathbb{C}\}$ be the set of critical points.
4. Find $\mathbf{w} \in \mathbb{R}^d_{>0} \cap C'$ such that $t \notin (0, 1)$ for any $(\mathbf{w}, \lambda, t) \in \mathcal{S}$.
   If there is more than one or no such point then Fail.
5. Return the set $\mathcal{U} = \{\mathbf{z} \in C' : \mathbf{z} \in T(\mathbf{w})\}$ of critical points with the same coordinate-wise modulus as $\mathbf{w}$.

MinimalCriticalComb is more fully described in Algorithm 2. Using MinimalCriticalComb to determine minimal critical points, DiagonalAsymptotics gives dominant asymptotics when it returns without failing. This happens whenever the following assumptions are met

(A0) $F(\mathbf{z})$ admits at least one minimal critical point;
(A1) $(\nabla H)(\mathbf{z})$ does not vanish at a root of $H(\mathbf{z})$;
(A2) $G(\mathbf{z})$ is non-zero for at least one minimal critical point;
(A3) all minimal critical points of $F(\mathbf{z})$ are non-degenerate;
(J1) the Jacobian matrix of the system (7.2) in the variables $\mathbf{z}, \lambda, t$ is non-singular at the solutions of (7.2).

Proposition 5.11 from Chapter 5 implies assumptions (A1) to (A3) hold generically. When F is combinatorial then Corollary 5.4 in Chapter 5 and the results of Section 3.3.1 in Chapter 3 imply that F admits a minimal critical point with positive coordinates whenever the coefficients of the linear terms $z_1, z_2, \dots, z_d$ in H are non-zero, which also holds generically. Problem 7.1 asks you to adapt the arguments in Section 5.3.4 of Chapter 5 to prove that (J1) holds generically.

Assumption (A1) implies that the singular variety is smooth, while (A2) implies we can use the explicit expression in Theorem 5.1 for dominant asymptotics. Furthermore, the so-called Jacobian criterion [11, Theorem 16.19] implies that the polynomial system (7.2) is finite whenever its Jacobian matrix has full rank (i.e., is non-singular) at all of its solutions. Thus, (J1) is slightly stronger than requiring that $F(\mathbf{z})$ admit a finite number of critical points: we use (J1) to compute a Kronecker representation of the solutions of this system using Proposition 7.3 below. All of our conditions, except that F is combinatorial, can be verified algorithmically.

Remark 7.2 Asymptotics can still be determined in many situations where (A1) and (A2) do not hold. Assumption (A1) implies H is square-free and the singular variety is everywhere smooth; both conditions hold generically, but are not necessary to apply Theorem 5.3 from Chapter 5. When (A2) doesn't hold the leading term in Theorem 5.3 vanishes but one can try calculating higher-order terms. These situations can be handled by small extensions of our algorithms.

7.1.3 Minimal Critical Points in the General Case

In the combinatorial case, to prove $\mathbf{w} \in \mathcal{V}$ is minimal it is sufficient to check that for each $t \in (0,1)$ the point $t|\mathbf{w}| = (t|w_1|, \ldots, t|w_d|)$ does not lie in $\mathcal{V}$. Thus, one can think of tracing a line segment from the origin to $|\mathbf{w}|$ and stopping if the segment intersects $\mathcal{V}$. In the non-combinatorial case, things are more difficult: to prove that $\mathbf{w} \in \mathcal{V}$ is minimal, it must be determined for each $t \in (0,1)$ whether there exists a point in $\mathcal{V}$ with the same coordinate-wise modulus as $t\mathbf{w}$. One can still imagine tracing a line segment from the origin to $|\mathbf{w}|$, but now each point on the line-segment defines a product of circles which must be checked for elements of $\mathcal{V}$. This, unsurprisingly, leads to an algorithm for minimality which is more computationally expensive.

In order to express the moduli of coordinates algebraically, we convert our d complex variables to $2d$ real variables.

Definition 7.3 (real and imaginary polynomial decomposition) Given a polynomial $f(\mathbf{z}) \in \mathbb{R}[\mathbf{z}]$ we define $f(\mathbf{x} + \mathrm{i}\mathbf{y}) = f(x_1 + \mathrm{i}y_1, \ldots, x_d + \mathrm{i}y_d)$ and let $f^{\Re}(\mathbf{x}, \mathbf{y})$ and $f^{\Im}(\mathbf{x}, \mathbf{y})$ denote the unique polynomials in $\mathbb{R}[\mathbf{x}, \mathbf{y}]$ such that

$$f(\mathbf{x} + i\mathbf{y}) = f^{\Re}(\mathbf{x}, \mathbf{y}) + \mathrm{i} f^{\Im}(\mathbf{x}, \mathbf{y}).$$

Because a polynomial is everywhere differentiable, the Cauchy-Riemann equations [19, Ch. 2.1] state that the partial derivatives of $f^{\Re}$ and $f^{\Im}$ satisfy the equalities $f^{\Re}_{x_j}(\mathbf{x}, \mathbf{y}) = f^{\Im}_{y_j}(\mathbf{x}, \mathbf{y})$ and $f^{\Re}_{y_j}(\mathbf{x}, \mathbf{y}) = -f^{\Im}_{x_j}(\mathbf{x}, \mathbf{y})$ for each $1 \leq j \leq d$, while

$$f_{z_j}(\mathbf{x} + \mathrm{i}\mathbf{y}) = \frac{1}{2}\left(f^{\Re}_{x_j}(\mathbf{x}, \mathbf{y}) + \mathrm{i} f^{\Im}_{x_j}(\mathbf{x}, \mathbf{y})\right) - \frac{\mathrm{i}}{2}\left(f^{\Re}_{y_j}(\mathbf{x}, \mathbf{y}) + \mathrm{i} f^{\Im}_{y_j}(\mathbf{x}, \mathbf{y})\right).$$

Some light algebraic manipulation then shows that the real solutions of the system

$$H^{\Re}(\mathbf{a},\mathbf{b}) = H^{\Im}(\mathbf{a},\mathbf{b}) = 0 \tag{7.3}$$

$$a_j H^{\Re}_{x_j}(\mathbf{a},\mathbf{b}) + b_j H^{\Re}_{y_j}(\mathbf{a},\mathbf{b}) - r_j\lambda_R = 0, \qquad j = 1,\dots,d \tag{7.4}$$

$$a_j H^{\Im}_{x_j}(\mathbf{a},\mathbf{b}) + b_j H^{\Im}_{y_j}(\mathbf{a},\mathbf{b}) - r_j\lambda_I = 0, \qquad j = 1,\dots,d \tag{7.5}$$

in the variables $\mathbf{a}, \mathbf{b}, \lambda_R, \lambda_I$ correspond exactly to all complex solutions of the smooth critical point equations

$$H(\mathbf{z}) = z_1 H_{z_1}(\mathbf{z}) - r_1\lambda = \cdots = z_d H_{z_d}(\mathbf{z}) - r_d\lambda = 0$$

with $\mathbf{z} = \mathbf{a} + \mathrm{i}\mathbf{b}$ and $\lambda = \lambda_R + \mathrm{i}\lambda_I$. Furthermore, if we define the equations

$$H^{\Re}(\mathbf{x},\mathbf{y}) = H^{\Im}(\mathbf{x},\mathbf{y}) = 0 \tag{7.6}$$

$$x_j^2 + y_j^2 - t(a_j^2 + b_j^2) = 0, \qquad j = 1,\dots,d \tag{7.7}$$

then (7.7) encodes a relationship between the coordinate-wise moduli of $\mathbf{x} + \mathrm{i}\mathbf{y}$ and $\mathbf{a} + \mathrm{i}\mathbf{b}$ when $\mathbf{x}, \mathbf{y}, \mathbf{a}$, and $\mathbf{b}$ are real vectors. In particular, Proposition 5.4 in Chapter 5 implies that a point $\mathbf{z} = \mathbf{a} + i\mathbf{b} \in \mathcal{V}$ is minimal if and only if there is no solution to equations (7.6) and (7.7) with $\mathbf{x}, \mathbf{y}, t$ real and $0 \le t < 1$.

The polynomial system defined by (7.3)–(7.7) contains $3d + 4$ equations in the $4d + 3$ variables $\mathbf{a}, \mathbf{b}, \mathbf{x}, \mathbf{r}, \lambda_R, \lambda_I$, and t. Let

- $\mathcal{W} \subset \mathbb{C}^{4d+3}$ denote the complex solutions of the system of equations (7.3)–(7.7)
- $\mathcal{W}_{\mathbb{R}} = \mathcal{W} \cap \mathbb{R}^{4d+3}$ be the real part of $\mathcal{W}$
- $\pi_t : \mathcal{W}_{\mathbb{R}} \to \mathbb{C}$ be the projection map $\pi_t(\mathbf{a}, \mathbf{b}, \mathbf{x}, \mathbf{y}, \lambda_R, \lambda_I, t) = t$
- $C = \{(\mathbf{a},\mathbf{b}) \in \mathbb{R}^{2d}_* : (\mathbf{a}, \mathbf{b}, \mathbf{x}, \mathbf{y}, \lambda_R, \lambda_I, t) \in \mathcal{W}_{\mathbb{R}} \text{ for some } \mathbf{x}, \mathbf{y}, \lambda_R, \lambda_I, t\}$

We now have a strategy for determining minimality: each $(\mathbf{a},\mathbf{b}) \in C$ corresponds to a critical point $\mathbf{z} = \mathbf{a} + \mathrm{i}\mathbf{b}$, and that critical point is minimal if and only if there does not exist $(\mathbf{a}, \mathbf{b}, \mathbf{x}, \mathbf{y}, \lambda_R, \lambda_I, t) \in \mathcal{W}_{\mathbb{R}}$ with $0 \le t < 1$. Unfortunately, even if the set $\mathcal{W}_{\mathbb{R}}$ of real solutions is finite the set $\mathcal{W}$ of complex solutions will be infinite (when non-empty) since there are more variables than equations. Although the real solutions of (7.3)–(7.7) are the only ones which have meaning in our original problem, most algebraic techniques for polynomial system solving work, at least implicitly, over an algebraically closed field.

To get around this problem, we introduce additional equations satisfied by the real solutions we care about which also eliminate spurious complex solutions. Our techniques are inspired by so-called critical point methods[3] for sampling points in real algebraic sets, an influential approach to real polynomial system solving popularized by Grigor'ev and Vorobjov [18] and Renegar [34].

Proposition 7.1 *Let $H \in \mathbb{Q}[\mathbf{z}]$ be a polynomial which does not vanish at the origin, and suppose that the Jacobian matrix of the polynomials in* (7.3)–(7.7) *has full rank*

[3] Although algebraic results are usually most naturally stated over algebraically closed fields, calculus and geometry typically 'respect' real constructions (for instance, there are both real and complex versions of the implicit function theorem). Critical point methods exploit this difference.

at any point in $\mathcal{W}$. The point $\mathbf{z} = \mathbf{a} + i\mathbf{b} \in \mathbb{C}_*^d$ with $\mathbf{a}, \mathbf{b} \in \mathbb{R}_*^d$ is a minimal critical point in the direction $\mathbf{r}$ if and only if $(\mathbf{a}, \mathbf{b})$ satisfies equations (7.3)–(7.5) and there does not exist $(\mathbf{x}, \mathbf{y}, \nu, t) \in \mathbb{R}^{2d+2}$ with $0 < t < 1$ satisfying equations (7.6), (7.7), and

$$(y_j - \nu x_j) H^{\Re}_{x_j}(\mathbf{x}, \mathbf{y}) - (x_j + \nu y_j) H^{\Re}_{y_j}(\mathbf{x}, \mathbf{y}) = 0, \quad j = 1, \dots, d. \tag{7.8}$$

Remark 7.3 If the Jacobian matrix of the polynomial system (7.3)–(7.7) has full rank then the Jacobian matrix of the sub-system (7.3)–(7.5) also has full rank. Thus, the equations (7.3)–(7.5) admit a finite number of complex solutions under the conditions of Proposition 7.1.

Proof For fixed $(\mathbf{a}, \mathbf{b}) \in \mathbb{R}_*^{2d}$ a point $\boldsymbol{\sigma} = (\mathbf{a}, \mathbf{b}, \mathbf{x}, \mathbf{y}, \lambda_R, \lambda_I, t) \in \mathcal{W}_{\mathbb{R}}$ is a local extremum of π_t as a map from $\mathcal{W}_{\mathbb{R}}$ to $\mathbb{R}$ only when the gradient of π_t is perpendicular to the tangent plane of $\mathcal{W}_{\mathbb{R}}$ at $\boldsymbol{\sigma}$. This, in turn, implies the gradient of π_t and the gradients of the polynomials $H^{\Re}, H^{\Im}$, and the $x_j^2 + y_j^2 - t(a_j^2 + b_j^2)$, which define $\mathcal{W}_{\mathbb{R}}$, are linearly dependent. In other words, at any such extremum the matrix

$$J = \begin{pmatrix} \nabla H^{\Re}(\mathbf{x}, \mathbf{y}) \\ \nabla H^{\Im}(\mathbf{x}, \mathbf{y}) \\ \nabla(x_1^2 + y_1^2 - t(a_1^2 + b_1^2)) \\ \vdots \\ \nabla(x_d^2 + y_d^2 - t(a_d^2 + b_d^2)) \\ \nabla(t) \end{pmatrix} = \begin{pmatrix} H^{\Re}_{x_1} & \cdots & H^{\Re}_{x_d} & H^{\Re}_{y_1} & \cdots & H^{\Re}_{y_d} & 0 \\ H^{\Im}_{x_1} & \cdots & H^{\Im}_{x_d} & H^{\Im}_{y_1} & \cdots & H^{\Im}_{y_d} & 0 \\ 2x_1 & \mathbf{0} & 0 & 2y_1 & \mathbf{0} & 0 & -(a_1^2 + b_1^2) \\ \mathbf{0} & \ddots & \mathbf{0} & \mathbf{0} & \ddots & \mathbf{0} & \vdots \\ 0 & \mathbf{0} & 2x_d & 0 & \mathbf{0} & 2y_d & -(a_d^2 + b_d^2) \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

is rank deficient, so a non-trivial linear combination of its rows vanishes. Using the Cauchy-Riemann equations to write $H^{\Im}_{x_j} = -H^{\Re}_{y_j}$ and $H^{\Im}_{y_j} = H^{\Re}_{x_j}$ then implies the existence of $\nu, \lambda_1, \dots, \lambda_d$ such that

$$\begin{aligned} H^{\Re}_{x_j} - \nu H^{\Re}_{y_j} + \lambda_j x_j &= 0 \\ H^{\Re}_{y_j} + \nu H^{\Re}_{x_j} + \lambda_j y_j &= 0 \end{aligned}$$

for each $j = 1, \dots, d$, which simplifies to give (7.8).

For each of the finite number of points $(\mathbf{a}, \mathbf{b}) \in \mathbb{R}_*^{2d}$ satisfying equations (7.3)–(7.5) under our assumptions, the set

$$S = \left\{ (\mathbf{x}, \mathbf{y}, t) \in \mathbb{R}^{2d+1} : t \in [0, 1],\ (\mathbf{a}, \mathbf{b}, \mathbf{x}, \mathbf{y}, t) \text{ satisfy (7.6) and (7.7)} \right\}$$

is compact, since $t \in [0, 1]$ implies $x_j^2 + y_j^2 \le a_j^2 + b_j^2$ for each $j = 1, \dots, d$. Furthermore, S is non-empty because it contains $(\mathbf{a}, \mathbf{b}, 1)$ and has no point with $t = 0$ since $H(\mathbf{0}) \neq 0$. Thus, the continuous function π_t achieves its minimum on the compact set S and such a minimizer must satisfy (7.8) for some $\nu \in \mathbb{R}$ and $0 < t \le 1$. Any solution $(\mathbf{x}, \mathbf{y}, t) \in S$ with $t < 1$ gives a point $\mathbf{x} + i\mathbf{y}$ that has smaller coordinate-wise modulus than $\mathbf{a} + i\mathbf{b}$, meaning $\mathbf{z} = \mathbf{a} + i\mathbf{b}$ is not minimal. Conversely, if $\mathbf{z} = \mathbf{a} + i\mathbf{b}$ is not minimal then there exists $(\mathbf{x}, \mathbf{y}, t) \in S$ with t minimal and $t < 1$. This is a local minimum of π_t on $\mathcal{W}_{\mathbb{R}}$ so there exists $\nu \in \mathbb{R}$ such that (7.8) is satisfied. □

Equations (7.3)–(7.8) now consist of $4d + 4$ equations in $4d + 4$ unknowns, meaning they can have a finite number of complex solutions whose real solutions encode all critical points and help determine minimality. In addition to (A0)–(A3), we work under the following assumption,

(J2) the Jacobian matrix of the system (7.3)–(7.8) in the variables $\mathbf{a}, \mathbf{b}, \mathbf{x}, \mathbf{y}, \lambda_R, \lambda_I, t$ is non-singular at the solutions of (7.3)–(7.8).

Assumption (J2) is slightly stronger than what we need to apply the results of Chapter 5, but this formulation will help us bound the complexity of the algebraic methods we develop below. As noted in Remark 7.3, (J2) implies that F admits only a finite number of critical points. Ultimately, we have obtained the following algorithm.

MinimalCriticalGen

INPUT: Coprime polynomials $G(\mathbf{z}), H(\mathbf{z}) \in \mathbb{Z}[\mathbf{z}]$ and direction $\mathbf{r} \in \mathbb{Z}^d_{>0}$
OUTPUT: The set $\mathcal{U}$ of minimal critical points of $F(\mathbf{z})$ in the direction $\mathbf{r}$

1. Let S be the algebraic set defined by the zeroes of the polynomial system (7.3)–(7.8) in the variables $\mathbf{a}, \mathbf{b}, \mathbf{x}, \mathbf{y}, \lambda_R, \lambda_I, \nu, t$. If S is not finite then Fail.
2. Let $\mathcal{U}$ denote the points $\mathbf{a} + i\mathbf{b} \in \mathbb{C}^d_*$ such that: (1) there exist $\mathbf{x}, \mathbf{y}, \lambda_R, \lambda_R, \nu, t$ with $(\mathbf{a}, \mathbf{b}, \mathbf{x}, \mathbf{y}, \lambda_R, \lambda_I, \nu, t) \in S \cap \mathbb{R}^{4d+4}$, and (2) for any such points $t \notin (0, 1)$.
3. If $\mathcal{U}$ is empty, if it has an element where $\lambda_R = \lambda_I = 0$, or if the elements of $\mathcal{U}$ do not all have the same coordinate-wise modulus then Fail.
4. Return $\mathcal{U}$.

MinimalCriticalGen is more fully described in Algorithm 3. Using MinimalCriticalGen to determine minimal critical points, DiagonalAsymptotics gives dominant asymptotics when it returns without failing. This happens under assumptions (A0)–(A3) and (J2), all of which can be verified algorithmically.

7.2 ACSV Algorithms and Examples

Having described our approach and assumptions, we can now state our main result.

Theorem 7.1 *Let $G(\mathbf{z})$ and $H(\mathbf{z})$ be polynomials in $\mathbb{Z}[\mathbf{z}] = \mathbb{Z}[z_1, \dots, z_d]$ of degrees at most δ and coefficients of absolute value at most 2^h, and suppose that Assumptions (A0)–(A3) hold. Fix a direction $\mathbf{r} \in \mathbb{Z}^d_{>0}$. If $F = G/H$ is combinatorial and Assumption (J1) holds, then Algorithm 2 is a probabilistic algorithm that computes the minimal critical points of F in the direction $\mathbf{r}$ in $\tilde{O}(h\delta^3 D^4)$ bit operations, where $D = \delta^d$. Whether or not $F = G/H$ is combinatorial, if Assumption (J2) holds* *then Algorithm 3 is a probabilistic algorithm that computes the minimal critical points of F in the direction $\mathbf{r}$ in $\tilde{O}\left(h\delta^{12}D^{12}\right)$ bit operations.*

In either case, Algorithm 1 is a probabilistic algorithm that uses these results and $\tilde{O}(h\delta^3 D^4)$ bit operations to compute three rational functions $A, B, C \in \mathbb{Z}(u)$, a square-free polynomial $P \in \mathbb{Z}[u]$ and a list U of roots of $P(u)$, specified by disks in $\mathbb{C}$ containing exactly one root of P, such that

$$[\mathbf{z}^{n\mathbf{r}}]F(\mathbf{z}) = (2\pi)^{(1-d)/2}(r_d n)^{(1-d)/2}\left(\sum_{u\in U} A(u)\sqrt{B(u)}\,C(u)^n + O\left(\frac{1}{n}\right)\right). \tag{7.9}$$

The values of $A(u), B(u)$ and $C(u)$ can be refined to precision $2^{-\kappa}$ at all elements of U in $\tilde{O}(\kappa D + h\delta^3 D^3)$ bit operations.

Remark 7.4 Using techniques which better take into account the multi-homogeneous structure of the polynomial system (7.3)–(7.8) (namely that (7.3)–(7.5) contain only variables which appear with degree at most 2 in the remaining equations) it is possible to reduce the complexity of determining minimal critical points in the general case from $\tilde{O}\left(h\delta^{12}D^{12}\right)$ to $\tilde{O}(h\delta^5 2^{3d} D^9)$ bit operations. This requires even more than the already considerable algebraic framework we discuss here, so we refer to Melczer and Salvy [26] for additional details.

Theorem 7.1 is obtained by encoding the smooth critical points with an efficient algebraic data structure.

Definition 7.4 (Kronecker representations) A *Kronecker representation* or *rational univariate representation* $[P(u), \mathbf{Q}]$ of a finite algebraic set

$$\mathcal{V}(\mathbf{f}) = \{\mathbf{z} : f_1(\mathbf{z}) = \cdots = f_d(\mathbf{z}) = 0\}$$

defined by the polynomial system $\mathbf{f} = (f_1, \ldots, f_d) \in \mathbb{Z}[\mathbf{z}]^d$ consists of

- a new variable u which is an integer linear combination of the $\mathbf{z}$,

  $$u = \kappa_1 z_1 + \cdots + \kappa_d z_d \in \mathbb{Z}[\mathbf{z}],$$

  such that u takes distinct values for $\mathbf{z} \in \mathcal{V}(\mathbf{f})$,
- a square-free polynomial $P \in \mathbb{Z}[u]$,
- polynomials $Q_1, \ldots, Q_d \in \mathbb{Z}[u]$ of degrees less than $\deg(P)$ such that $\mathcal{V}(\mathbf{f})$ is determined by the $\mathbf{z}$-coordinates of the solutions of the system

$$P(u) = 0, \qquad \begin{cases} P'(u)z_1 - Q_1(u) = 0, \\ \qquad\vdots \\ P'(u)z_d - Q_d(u) = 0. \end{cases} \tag{7.10}$$

Kronecker representations are useful because they encode the elements of a finite multi-dimensional algebraic set in terms of a univariate polynomial, and because the polynomials P, Q_j typically have much smaller coefficient size than alternative

Algorithm 1: DiagonalAsymptotics

Input: Polynomials $G(\mathbf{z}), H(\mathbf{z}) \in \mathbb{Z}[\mathbf{z}]$ and direction $\mathbf{r} \in \mathbb{N}^d$
Output: Polynomials A, B, C, P and a set $\mathbf{U}$ of roots of P such that the power series coefficients $f_{n\mathbf{r}}$ of $F = G/H$ satisfy (7.9)

```
/* Verify singularities of F are roots of H                         */
```
if *G and H are not coprime* **then** $G \leftarrow G/\gcd(G, H)$ and $H \leftarrow H/\gcd(G, H)$
if $H(\mathbf{0}) \neq 0$ **then return** FAIL
```
/* Define the set of critical points                                */
```
$C \leftarrow \{H, z_1 H_{z_1} - r_1\lambda, \dots, z_d H_{z_d} - r_d\lambda\}$
$[P, \mathbf{Q}, u] \leftarrow$ Kronecker(C)
```
/* Determine the subset of minimal critical points using Algorithm 2
   or Algorithm 3 below                                             */
```
if *F is known to be combinatorial* **then** $\mathbf{U} \leftarrow$ MinimalCriticalComb$(H, \mathbf{r}, [P, \mathbf{Q}, u])$
else $\mathbf{U} \leftarrow$ MinimalCriticalGen$(H, \mathbf{r}, [P, \mathbf{Q}, u])$
```
/* Find Hessian and numerator at the minimal critical points        */
```
$\widetilde{\mathcal{H}} \leftarrow$ determinant of the polynomial matrix $(z_d H_{z_d})\mathcal{H}$ defined by (5.25) with $\mathbf{w} = \mathbf{z}$
PolySys $\leftarrow C \cup \left\{h - \widetilde{\mathcal{H}}(\mathbf{z}), T - z_1^{r_1} \cdots z_d^{r_d}, g + G(\mathbf{z})\right\}$
$[P, (\mathbf{Q}, Q_{\widetilde{\mathcal{H}}}, Q_T, Q_{-G})] \leftarrow$ Kronecker(PolySys) using Proposition 7.4
```
/* Verify nondegeneracy of critical points and leading order of
   asymptotics then return                                          */
```
if $Q_{\widetilde{\mathcal{H}}}(u) = 0$ *at any* $u \in \mathbf{U}$, *or* $Q_{-G}(u) = 0$ *at all* $u \in \mathbf{U}$ **then return** FAIL
return $(A, B, C, P, \mathbf{U}) = (Q_{-G}/(r_d Q_\lambda),\ (r_d Q_\lambda)^{d-1}(P')^{2-d}/Q_{\widetilde{\mathcal{H}}},\ P'/Q_T,\ P,\ \mathbf{U})$

univariate encodings. Since u is an integer linear combination of the other variables, and the polynomials P and Q_j have integer coefficients, a root of $P(u)$ is real if and only if all coordinates $\mathbf{z}$ at the corresponding element of $\mathcal{V}(\mathbf{f})$ are real.

A Kronecker representation of $\mathcal{V}(\mathbf{f})$ gives an encoding of *all* elements of $\mathcal{V}(\mathbf{f})$, but to talk about *specific* elements of $\mathcal{V}(\mathbf{f})$ it is necessary to introduce additional information.

Definition 7.5 (isolating regions and numeric Kronecker representations) Given a univariate polynomial $P(u)$ and root $w \in \mathbb{C}$ of P, an *isolating disk* for w is a disk $D \subset \mathbb{C}$ of rational radius whose centre has rational real and imaginary parts, such that w is the only root of P in D. If w is a real root of P then an *isolating interval* for w is a finite interval $I \subset \mathbb{R}$ with rational endpoints such that w is the only root of P in I. A *numeric Kronecker representation* $[P(u), \mathbf{Q}, \mathbf{U}]$ of $\mathcal{V}(\mathbf{f})$ is a Kronecker representation $[P(u), \mathbf{Q}]$ of $\mathcal{V}(\mathbf{f})$ together with a sequence $\mathbf{U}$ of isolating intervals for the real roots of the polynomial P and isolating disks for the non-real roots of P.

Determining the elements of $\mathbf{U}$ to sufficient precision allows one to argue about the elements of the underlying algebraic set, including: approximating the solutions to arbitrary accuracy, selecting which solutions have real coordinates in certain intervals, determining which solutions have the same coordinate-wise modulus, and more. In practice, the elements of $\mathbf{U}$ are often stored as floating point approximations whose accuracy is certified to a specific precision.

Section 7.3 discusses the necessary algebraic background, algorithms for computing Kronecker and numeric Kronecker representations, properties of Kronecker

Algorithm 2: MinimalCriticalComb

Input: Polynomial $H(\mathbf{z})$, direction $\mathbf{r} \in \mathbb{N}^d$ and Kronecker representation $[P, \mathbf{Q}, u]$ of the smooth critical points for $1/H$ in the direction $\mathbf{r}$ **assuming** $1/H$ is combinatorial
Output: Set $\mathbf{U}$ of roots of P corresponding to the *minimal* critical points

```
/* Determine the extended critical point system */
```
$\mathcal{S} \leftarrow \left\{H, z_1 H_{z_1} - r_1\lambda, \ldots, z_d H_{z_d} - r_d\lambda, H(tz_1, \ldots, tz_d)\right\}$
$[\tilde{P}, \tilde{\mathbf{Q}}] \leftarrow$ Kronecker$(\mathcal{S})$
```
/* Find the minimal critical point with positive real coordinates */
```
$[\tilde{P}, \tilde{\mathbf{Q}}, \tilde{\mathbf{U}}] \leftarrow$ NumKronecker$(\tilde{P}, \tilde{\mathbf{Q}}, \kappa)$, where $\kappa = \tilde{O}(h\delta^2 D^2)$ is sufficient precision to group the roots of $\tilde{P}$ by the distinct values and signs they give to each $z_j = \tilde{Q}_j/\tilde{P}'$ at the real roots of $\tilde{P}$ using Lemmas 7.4 and 7.5 below
```
/* Find elements with positive real coordinates */
```
$M \leftarrow \{\{u \in \tilde{\mathbf{U}} \cap \mathbb{R} : \tilde{Q}_j(u)/\tilde{P}(u) > 0 \text{ for all } 1 \leq j \leq d\}\}$
```
/* Find unique minimal critical point with positive coordinates */
```
for *each* $j \in \{1, \ldots, d\}$ **do**
  Split apart each set in M by the distinct values its elements give $\tilde{Q}_j(v)/\tilde{P}'(v)$
end
for *each set* $\mathbf{m} \in M$ **do**
  if *one of the values taken by t on the elements of* $\mathbf{m}$ *lies in* $(0, 1)$ **then** $M \leftarrow M \setminus \{\mathbf{m}\}$
end
if *M does not have the form $\{\{u_\zeta\}\}$, or if $Q_\lambda(u_\zeta) = 0$* **then return** fail
```
/* Reduce back to original critical point system */
```
$[P, \mathbf{Q}, \mathbf{U}] \leftarrow$ NumKronecker$(P, \mathbf{Q}, \kappa)$ where $\kappa = \tilde{O}(hD^2)$ is sufficient precision to group the roots of $\tilde{P}$ by the distinct values they give to each $z_j = \tilde{Q}_j/\tilde{P}'$ using Lemma 7.5
Determine the root u_ζ of P corresponding to $\boldsymbol{\zeta}$ from the numerical approximations
```
/* Return the set of minimal critical points */
```
Refine $\mathbf{U}$ to isolating regions of size at most $2^{-\kappa}$, where $\kappa = \tilde{O}(h\delta^3 D^3)$ is sufficient to identify elements with same coordinate-wise modulus using Corollary 7.1
return *set of $u \in \mathbf{U}$ such that $|Q_j(u)/P'(u)| = Q_j(u_\zeta)/P'(u_\zeta)$ for $1 \leq j \leq d$*

representations, and alternative encodings of finite algebraic sets. An algebraic toolbox of methods using the numeric Kronecker representation to decide the properties we require for our asymptotic arguments is developed. Treating the results of Section 7.3 as a black box yields Algorithms 1– 3 implementing Theorem 7.1, whose proof is completed in Section 7.4. The procedures Kronecker($\mathbf{f}$) and NumKronecker($\mathbf{f}, \kappa$) for computing Kronecker and numeric Kronecker representations are discussed in Proposition 7.3 and Proposition 7.5, respectively.

7.2.1 Examples

We now work through some examples, starting by revisiting the Apéry numbers. As usual, worksheets illustrating these examples can be found on the textbook website.

Example 7.3 (Apéry Revisited)

Algorithm 3: MinimalCriticalGen

Input: Polynomial $H(\mathbf{z})$, direction $\mathbf{r} \in \mathbb{N}^d$ and Kronecker representation $[P, \mathbf{Q}, u]$ of the set of smooth critical points in some direction $\mathbf{r}$
Output: Set $\mathbf{U}$ of roots of P corresponding to the *minimal* critical points

```
/* Determine the extended critical point system */
```
$\tilde{S} \leftarrow$ Polynomials in (7.3)–(7.8)
$[\tilde{P}, \tilde{\mathbf{Q}}] \leftarrow$ Kronecker$(\tilde{S})$
```
/* Construct the set of minimal critical points */
```
$[\tilde{P}, \tilde{\mathbf{Q}}, \tilde{\mathbf{U}}] \leftarrow$ NumKronecker$(\tilde{P}, \tilde{\mathbf{Q}}, \kappa)$ where $\kappa = \tilde{O}(h\delta^8 D^8)$ is sufficient precision to group the roots of $\tilde{P}$ by the distinct values they give to each $\tilde{Q}_{a_i}/\tilde{P}'$ and $\tilde{Q}_{b_i}/\tilde{P}'$ using Lemma 7.5 and determine when t lies in $(0, 1)$ using Lemma 7.4
$\mathcal{S} \leftarrow \{\tilde{\mathbf{U}}\}$
for *each* $v \in \{a_1, \dots, a_d, b_1, \dots, b_d\}$ **do**
  Split apart each set in $\mathcal{S}$ by the distinct values its elements give $\tilde{Q}_v(u)/\tilde{P}'(u)$
end
for *each set* $\mathbf{s} \in \mathcal{S}$ **do**
  if *there exists an index j such that $a_j = b_j = 0$ on the elements of* $\mathbf{s}$ **then return** *FAIL*
  if *t has a value in $(0, 1)$ at some element of* $\mathbf{s}$ **then** $\mathcal{S} \leftarrow \mathcal{S} \setminus \{\mathbf{s}\}$
end
$\mathcal{S}_{a,b} \leftarrow$ numerical approximations of $(\mathbf{a}, \mathbf{b})$ determined by elements of $\mathcal{S}$
```
/* Return the minimal critical points */
```
$[P, \mathbf{Q}, \mathbf{U}] \leftarrow$ NumKronecker$(P, \mathbf{Q}, \kappa)$ where $\kappa = \tilde{O}(h\delta^2 D^2)$ is sufficient precision to identify which elements have $(\mathbf{a}, \mathbf{b})$ in $\mathcal{S}_{a,b}$ using Lemma 7.5
return *roots of P defining* $\mathbf{z} = \mathbf{a} + i\mathbf{b}$ *for* $(\mathbf{a}, \mathbf{b}) \in \mathcal{S}_{a,b}$

Let us trace through our algorithms on the combinatorial function

$$\Delta\left(\frac{1}{H(w,x,y,z)}\right) = \Delta\left(\frac{1}{1 - w(1+x)(1+y)(1+z)(1+y+z+yz+xyz)}\right).$$

To begin we form the polynomial system

$$\mathcal{S} = \left\{H(w,x,y,z), wH_w - \lambda, xH_x - \lambda, yH_y - \lambda, zH_z - \lambda, H(tw,tx,ty,tw)\right\}.$$

Using a Gröbner basis method discussed in Section 7.3, we can compute a Kronecker representation $[\tilde{P}, \tilde{\mathbf{Q}}]$ of $\mathcal{V}(\mathcal{S})$ by taking u as a random integer linear combination of $w, x, y,$ and z. For instance, taking $u = w + t$ results in a parametrization where $\tilde{P}$ is a polynomial of degree 14 with integer coefficients at most 17 digits long, and $\tilde{Q}_w, \tilde{Q}_x, \tilde{Q}_y, \tilde{Q}_z, \tilde{Q}_\lambda$ and $\tilde{Q}_t$ are polynomials of degree at most 13 with integer coefficients at most 18 digits long. Each element of $\mathcal{V}(\mathcal{S})$ is encoded by a root of $\tilde{P}(u)$, and the critical points, determined by the elements of $\mathcal{V}(\mathcal{S})$ where $t = 1$, are given by the values of u that are roots of

$$P(u) = \gcd(\tilde{P}, \tilde{P}' - \tilde{Q}_t) = u^2 + 162u - 167.$$

Since P is quadratic, we can find its roots u_1, u_2 exactly to get the critical points

$$\left(\frac{\tilde{Q}_w(u_1)}{\tilde{P}'(u_1)}, \frac{\tilde{Q}_x(u_1)}{\tilde{P}'(u_1)}, \frac{\tilde{Q}_y(u_1)}{\tilde{P}'(u_1)}, \frac{\tilde{Q}_z(u_1)}{\tilde{P}'(u_1)}\right) = \left(-82 + 58\sqrt{2}, 1 + \sqrt{2}, 1/\sqrt{2}, 1/\sqrt{2}\right)$$
$$\approx (0.02\dots, 2.41\dots, 0.707\dots, 0.707\dots)$$
$$\left(\frac{\tilde{Q}_w(u_2)}{\tilde{P}'(u_2)}, \frac{\tilde{Q}_x(u_2)}{\tilde{P}'(u_2)}, \frac{\tilde{Q}_y(u_2)}{\tilde{P}'(u_2)}, \frac{\tilde{Q}_z(u_2)}{\tilde{P}'(u_2)}\right) = \left(-82 - 58\sqrt{2}, 1 - \sqrt{2}, -1/\sqrt{2}, -1/\sqrt{2}\right)$$
$$\approx (-164.02\dots, -0.4\dots, -0.7\dots, -0.7\dots).$$

The numeric Kronecker representation is obtained by approximating the roots of $\tilde{P}(u) = 0$ to a sufficiently high precision; substituting these values into $\tilde{Q}_t(u)/\tilde{P}'(u)$ shows no element of $\mathcal{V}(\mathcal{S})$ has t in $(0, 1)$, meaning the critical point with positive coordinates is minimal (as we have shown in previous chapters by hand).

Since we now care only about the critical points, where $t = 1$, we may replace $\tilde{P}(u)$ by $P(u)$ and reduce the $\tilde{Q}_v$ polynomials accordingly. In particular, the extended Euclidean algorithm can be used to determine the inverse $\tilde{P}'(u)^{-1}$ of $\tilde{P}'(u)$ modulo the polynomial $P(u)$. If we define

$$Q_v(u) = Q_v(u)P'(u)\tilde{P}'(u)^{-1} \mod P(u)$$

for $v \in \{w, x, y, z\}$ then $[P, \mathbf{Q}]$ defines a Kronecker representation

$$P(u) = u^2 + 162u - 167 = 0,$$
$$w = \frac{-164u + 172}{P'(u)},\ x = \frac{2u + 394}{P'(u)},\ y = z = \frac{116}{P'(u)},\ \lambda = -1$$

of the critical points (alternatively, one could recompute a Kronecker representation from scratch using a Gröbner basis computation).

Building the matrix $\mathcal{H}$ in (5.25), multiplying by zH_z, and taking the determinant gives $\det\tilde{\mathcal{H}}$ as a polynomial in x, y, w, and z. Incorporating this polynomial and $T = wxyz$ into the Kronecker representation parametrizes

$$\det\tilde{\mathcal{H}} = \frac{96u - 112}{P'(u)} \qquad \text{and} \quad T = \frac{34u - 30}{P'(u)}$$

at the critical points. Putting everything together, and noting the numerator $G = 1$ in this example, the algorithm returns

$$f_{n,n,n,n} = \left(\sum_{u \in \mathbf{U}} \left(\frac{u + 81}{17u - 15}\right)^n \left(\frac{1}{4\pi^{3/2}n^{3/2}}\right)\sqrt{\frac{u + 81}{28 - 24u}}\right)\left(1 + O\left(\frac{1}{n}\right)\right)$$

where $\mathbf{U} = \{u_1\}$ is the real root $1.0243\dots$ of $\tilde{P}(u) = u^2 + 162u - 167$.

Example 7.4 (Two Critical Points with Positive Coordinates)

The bivariate rational function

$$F(x,y) = \frac{1}{(1-x-y)(20-x-40y)-1} = \frac{1}{1-x-y} \times \frac{1}{20-x-ty-\frac{1}{1-x-y}}$$

is combinatorial and has a smooth singular variety. Computing a Kronecker representation of the algebraic set defined by

$$H(x,y) = xH_x(x,y) - \lambda = yH_y(x,y) - \lambda = H(tx,ty) = 0$$

shows that there are four critical points, given by the solutions where $t = 1$. Two of the critical points,

$$(x_1, y_1) \approx (0.549, 0.309) \qquad \text{and} \qquad (x_2, y_2) \approx (9.997, 0.252),$$

have positive coordinates; since $x_1 < x_2$ while $y_1 > y_2$ it is not clear which of the points (if any) is minimal. Examining the full set of real solutions encoded by the Kronecker representation, not just those where $t = 1$, shows there is a point $t(x_2, y_2) \in \mathcal{V}$ where $t \approx 0.092$, meaning (x_2, y_2) is not minimal. If $t(x_1, y_1) \in \mathcal{V}$ then $t \geq 1$, so (x_1, y_1) *is* a smooth minimal critical point[4]. After verifying there is no other critical point with the same coordinate-wise modulus, we obtain asymptotics

$$C(u_1)^n n^{-1/2}\left(\pi^{-1/2}A(u_1)\sqrt{B(u_1)} + O\left(\frac{1}{n}\right)\right) = (5.884\ldots)^n n^{-1/2}\left(0.054\ldots + O\left(\frac{1}{n}\right)\right),$$

where $A, B, C \in \mathbb{Z}(u)$ and u_1 is a real algebraic number specified by its degree four minimal polynomial.

Example 7.5 (A Non-Combinatorial Series Expansion)

Baryshnikov et al. [1] studied the family of rational functions

$$F_c(x,y,z) = \frac{1}{1-(x+y+z)+cxyz}$$

to determine which values of $c \in \mathbb{R}$ result in eventually positive diagonal sequences. Our algorithms automatically derive such asymptotic results for fixed values of c. For example[5], if $c = 81/8$ and we consider the main diagonal direction then there are three critical points (x, y, z), defined by

[4] In fact there is another point on the line segment from the origin to (x_1, y_1), but this point has the form $t(x_1, y_1)$ where $t \approx 1.709 > 1$.

[5] Our choice of c simplifies the critical points obtained for this example, but our algorithms work for any value of c.

$$x = y = z \in \left\{-2/3, -1/3 \pm \mathrm{i}/\left(3\sqrt{3}\right)\right\}.$$

Computing a Kronecker representation shows that the system (7.3)–(7.8) admits solutions with $0 < t < 1$. In particular, at any root of the system with $0 < t < 1$ either $t = 1/3$ or $t \approx 0.3451$ satisfies $81t^3 + 36t^2 + 4t - 9 = 0$. Algorithm 3 proves that $(-2/3, -2/3, -2/3)$ is the only non-minimal critical point, ultimately showing that the power series diagonal has dominant asymptotics

$$\left(\mathrm{i}81\sqrt{3}/8\right)^n \frac{3\sqrt{3}+3\mathrm{i}}{8\pi n} + \left(-\mathrm{i}81\sqrt{3}/8\right)^n \frac{3\sqrt{3}-3\mathrm{i}}{8\pi n} = \left(\frac{81\sqrt{3}}{8}\right)^n \frac{3\cos\left(\frac{\pi}{6}+\frac{\pi n}{2}\right)}{2n\pi}.$$

Going back to the original motivations of Baryshnikov et al. [1], our algorithm proves that when $c = 81/8$ the diagonal sequence contains an infinite number of negative terms and an infinite number of positive terms.

Our approach to proving minimality using Proposition 7.1 is flexible enough to help even when the direction $\mathbf{r}$ is taken as a parameter. When the argument goes through, proving minimality is reduced to certifying that a univariate polynomial $p_{\mathbf{r}}(t)$ in t whose coefficients are algebraic quantities in $\mathbf{r}$ has no root $t \in (0, 1)$ when $\mathbf{r}$ lies in some range of interest. An invaluable tool for these arguments is Sturm's theorem for real root counting.

Definition 7.6 (Sturm sequences and sign changes) If $f(t)$ is a polynomial in $\mathbb{R}[t]$ of degree δ then the *Sturm sequence* of f is the finite sequence of polynomials $g_0, \dots, g_\delta$ defined by

$$g_0(t) = f(t), \quad g_1(t) = f'(t), \quad g_n = -\mathrm{rem}(g_{n-2}, g_{n-1}) \quad (1 \leq n \leq \delta),$$

where $\mathrm{rem}(g_{n-2}, g_{n-1})$ is the polynomial remainder when dividing g_{n-2} by g_{n-1}. For $t \in \mathbb{R}$ the *sign variation* of f at t, denoted $V_f(t)$, is the number of sign changes between consecutive elements in the sequence $g_0(t), \dots, g_\delta(t)$, ignoring zeroes.

Proposition 7.2 (Sturm's Theorem) *The number of zeroes of $f(t) \in \mathbb{R}[t]$ in the interval $(a, b]$ for any $a, b \in \mathbb{R}$ with $a < b$ equals $V_f(b) - V_f(a)$. In particular, if the Sturm sequence of $f(t)$ has the same number of sign alternations when $t = a$ and $t = b$ then there are no roots of $f(t)$ in $(a, b]$.*

A proof of Proposition 7.2 can be found in [2, Thm 2.50].

Example 7.6 (Distribution of Leaves in Planar Trees)

Recall from Chapter 3 that the number of rooted planar trees on n nodes with k leaves is the coefficient of $u^k z^n y^n$ in the power series expansion of

$$F(u, z, y) = \frac{G(u, z, y)}{H(u, z, y)} = \frac{y(uyz - yz - 2y + 1)}{1 + uyz - uz - yz - y}.$$

Note that the power series of F has negative coefficients away from the terms of combinatorial of interest. Forming the system (7.3)–(7.8) with $\mathbf{r} = (s, 1, 1)$ as a parameter and computing a lexicographic Gröbner basis, as described in Section 7.3, shows that there is a single critical point of F,

$$(u, z, w) = \left(\frac{s^2}{(s-1)^2}, \frac{(s-1)^2}{s}, s \right),$$

and that the value of t at any solution of (7.3)–(7.8) satisfies

$$(t-1)p_1(t)p_2(t)p_3(t)p_4(t) = 0,$$

where p_1, p_2, p_3, p_4 are explicit polynomials in t with coefficients in $\mathbb{Z}[s]$. Minimality of the critical point for $0 < s < 1$ follows from showing that p_1, p_2, p_3, p_4 have no root with $t \in (0, 1)$ for those values of s (note that $s < 1$ since a tree cannot have more leaves than nodes). We prove this for

$$p_1(t) = s^4t^2 - (3s^2 - 2s + 1)t + 1;$$

the arguments for p_2, p_3, and p_4, which are cubic in t, are analogous and given in the worksheet for this example. The Sturm sequence for $p_1(t)$ evaluated at $t = 0$ is

$$\mathbf{v}_0 = \left(1, -3s^2 + 2s - 1, \frac{(5s^2 - 2s + 1)(s-1)^2}{4s^4} \right)$$

while evaluated at $t = 1$ it becomes

$$\mathbf{v}_1 = \left(s(s+2)(s-1)^2, (s-1)(2s^3 + 2s^2 - s + 1), \frac{(5s^2 - 2s + 1)(s-1)^2}{4s^4} \right).$$

Whenever $0 < s < 1$ the signs of the elements of $\mathbf{v}_0$ and $\mathbf{v}_1$ both form the sequence $(+, -, +)$, containing 2 sign changes, so $p_1(t)$ has no root with $0 < t \leq 1$.

Repeating this process with p_2, p_3, and p_4 shows the critical point is minimal, ultimately obtaining asymptotics

$$[u^{sn}z^ny^n]F(u, z, y) = \frac{1}{2\pi n^2} \left(\frac{1}{s^{2s}(1-s)^{2s+2}} \right)^n \left(1 + O\left(\frac{1}{n}\right) \right)$$

for the number of rooted plane trees with n nodes and $k = sn$ leaves.

7.3 Data Structures for Polynomial System Solving

In order to implement our algorithms for asymptotics we need techniques for manipulating the solutions of polynomial systems. We thus give a brief overview of this vast area of mathematics which, out of necessity, is highly tailored to our needs. Those wanting a deeper introduction can turn to the lovely book of Cox et al. [10] or, for an exhaustive look at the topic, to the book series of Mora [28, 29, 30, 31].

7.3.1 Gröbner Bases and Triangular Systems

We start by encoding polynomial systems and their solutions.

Definition 7.7 (ideals and algebraic sets) Given a finite collection of polynomials $f_1(\mathbf{z}), \ldots, f_r(\mathbf{z})$ in $\mathbb{Q}[\mathbf{z}]$, the *ideal of* $\mathbb{Q}[\mathbf{z}]$ *generated by the* f_j is the set

$$(f_1, \ldots, f_r) = \left\{ \sum_{j=0}^{r} g_j(\mathbf{z}) f_j(\mathbf{z}) : g_j(\mathbf{z}) \in \mathbb{Q}[\mathbf{z}] \right\}.$$

Equivalently, $(f_1, \ldots, f_r)$ is the smallest set containing the f_j which is closed under addition of elements in the set and closed under multiplication by any element of $\mathbb{Q}[\mathbf{z}]$. The *algebraic set* or *variety*[6] defined by $f_1(\mathbf{z}), \ldots, f_r(\mathbf{z}) \in \mathbb{Q}[\mathbf{z}]$ is the set of their common solutions

$$\mathcal{V}(f_1, \ldots, f_r) = \{\mathbf{z} \in \mathbb{C}^d : f_1(\mathbf{z}) = \cdots = f_r(\mathbf{z}) = 0\}.$$

If two sets of polynomials $\mathbf{f} = \{f_1, \ldots, f_r\}$ and $\mathbf{g} = \{g_1, \ldots, g_s\}$ generate the same ideal, then each f_j can be written as a $\mathbb{Q}[\mathbf{z}]$-linear combination of the g_k polynomials, and vice-versa. In particular, $\mathbf{f}$ and $\mathbf{g}$ vanish on the same set of points, meaning $\mathcal{V}(f_1, \ldots, f_r) = \mathcal{V}(g_1, \ldots, g_s)$. It thus makes sense to talk about the variety defined by an ideal, and symbolically 'solving' the polynomial system $f_1 = \cdots = f_r = 0$ can be viewed as finding a particularly nice generating set for the ideal $(f_1, \ldots, f_r)$.

Definition 7.8 (zero-dimensional ideals) An ideal I is *zero-dimensional* if $\mathcal{V}(I)$ is finite, and a polynomial system is *zero-dimensional* if its elements generate a zero-dimensional ideal (in other words, the polynomials have a finite number of common roots in $\mathbb{C}^d$).

If all f_j polynomials are linear then one way to determine $\mathcal{V}(f_1, \ldots, f_r)$ is taught in a first class on linear algebra: put the coefficients of the polynomials in a matrix and perform Gaussian elimination. At the polynomial level, this corresponds to taking linear combinations of the f_j in order to determine a new generating set g_k where, for instance, the variable z_1 appears only in g_1, the variable z_2 appears only in g_2, etc. Trying to mirror this construction for non-linear polynomials immediately runs into a problem: if we want to systematically cancel terms in the f_j polynomials to determine a 'simpler' generating set, we need a way of determining which monomials

[6] Some authors reserve the name variety for irreducible algebraic sets, but we do not make this distinction. Unlike in Chapter 5, here we only consider algebraic sets in $\mathbb{C}^d$ (we do not need projective space).

to cancel first. We thus need a way of sorting the monomials which appear, which is facilitated by the following definitions.

Definition 7.9 (total and monomial orderings) A *total order* on a set S is a binary relation $\geq$ on S such that for all $a, b, c \in S$: either $a \geq b$ or $b \geq a$, the order is antisymmetric (if $a \geq b$ and $b \geq a$ then $a = b$), and the order is transitive (if $a \geq b$ and $b \geq c$ then $a \geq c$). A *monomial ordering* is a total order $\geq$ on monomials $\mathbf{z}^{\mathbf{i}}$ with $\mathbf{i} \in \mathbb{N}^d$ such that for all $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{N}^d$: (1) if $\mathbf{z}^{\mathbf{a}} \geq \mathbf{z}^{\mathbf{b}}$ then $\mathbf{z}^{\mathbf{a}+\mathbf{c}} \geq \mathbf{z}^{\mathbf{b}+\mathbf{c}}$ and (2) $\mathbf{z}^{\mathbf{a}} \geq 1$.

Common monomial orders include

lexicographic: $\mathbf{z}^{\mathbf{a}} \geq_{\text{lex}} \mathbf{z}^{\mathbf{b}}$ if the leftmost non-zero entry of $\mathbf{a} - \mathbf{b}$ is positive. Equivalently, monomials are first ordered by their powers of z_1, with ties broken by their powers of z_2, then z_3, etc. For instance, $z_1^3 z_2^2 \geq_{\text{lex}} z_1^2 z_2^3 \geq_{\text{lex}} z_1^2 \geq_{\text{lex}} z_1 z_2^{100}$.

graded lexicographic: $\mathbf{z}^{\mathbf{a}} \geq_{\text{gr}} \mathbf{z}^{\mathbf{b}}$ if $\mathbf{z}^{\mathbf{a}}$ has larger degree than $\mathbf{z}^{\mathbf{b}}$, or if they have the same degree and $\mathbf{z}^{\mathbf{a}} \geq_{\text{lex}} \mathbf{z}^{\mathbf{b}}$. Monomials are sorted by degree, with ties broken by the lexicographic order. For instance, $z_1 z_2^{100} \geq_{\text{gr}} z_1^3 z_2^2 \geq_{\text{gr}} z_1^2 z_2^3 \geq_{\text{gr}} z_1^2$.

reverse graded lex.: $\mathbf{z}^{\mathbf{a}} \geq_{\text{rv}} \mathbf{z}^{\mathbf{b}}$ if $\mathbf{z}^{\mathbf{a}}$ has larger degree than $\mathbf{z}^{\mathbf{b}}$, or if they have the same degree and the rightmost non-zero entry of $\mathbf{a} - \mathbf{b}$ is negative. Monomials are sorted by degree, with ties broken by the reverse ordering to the lexicographic order when the sequence of the variables is also reversed. For instance, $z_2^6 \geq_{\text{rc}} z_1^3 z_2^2 \geq_{\text{rv}} z_1^2 z_2^3 \geq_{\text{rv}} z_1^5$.

Definition 7.10 (leading terms and Gröbner bases) Fix a monomial order $\geq$. The *leading term* of a polynomial $f \in \mathbb{Q}[\mathbf{z}]$ is the term appearing in f which is largest under $\geq$. Given an ideal $I = (f_1, \dots, f_r)$ the leading terms of all polynomials in I form another ideal $\text{LT}(I)$, called the *leading term ideal* of I, which is not necessarily generated by the leading terms of the generating set f_j. A *Gröbner basis* of the ideal I is a finite generating set $g_1, \dots, g_s$ of I such that the ideal generated by the leading terms of the g_j is $\text{LT}(I)$. A *reduced Gröbner basis* is a Gröbner basis such that no leading term of a generating polynomial lies in the ideal generated by the leading terms of the other generating polynomials, and whose polynomials are normalized to have leading terms with coefficient 1.

Gröbner bases were introduced in the 1960s by Buchburger [5, 6], who gave an algorithm for their calculation (they exist for any ideal and ordering) and proved several of their remarkable properties. The reduced Gröbner basis of an ideal under a fixed monomial order is unique, and a Gröbner basis of an ideal $I = (f_1, \dots, f_r)$ with respect to any order contains the polynomial 1 if and only if the polynomial system $f_1 = \cdots = f_r = 0$ has no solution in $\mathbb{C}^d$. Buchburger's algorithm, which reduces to the Euclidean algorithm for univariate polynomials and Gaussian elimination for multivariate linear polynomials, consists of taking two polynomials f, g

from the generating set $\mathbf{f}$ of an ideal, taking the smallest degree monomial linear combination that results in a polynomial with smaller leading term, reducing this new polynomial through multivariate polynomial division with $\mathbf{f}$, then adding the result to the generating set and repeating the procedure until no new generators are found. More efficient algorithms to compute Gröbner bases have been developed, notably the F4 and F5 algorithms of Faugére [13, 14], and most computer algebra systems contain built-in procedures to compute Gröbner bases after one specifies a monomial order.

Example 7.7 (Apéry Critical Points)

The Apéry sequence is the main diagonal of the rational function $1/H(w, x, y, z)$ where $H(w, x, y, z) = 1 - w(1 + x)(1 + y)(1 + z)(1 + y + z + yz + xyz)$, in which case the polynomials in the smooth critical point system (5.16) form the ideal

$$I = (H, xH_x - yH_y, xH_x - zH_z, xH_x - wH_w).$$

The Groebner[Basis] command in Maple, when given the ideal I and lexicographic order where $x > y > z > w$, returns the Gröbner basis

$$G = (w^2 + 164w - 4, 116z - w - 82, 116y - w - 82, 58x - w - 140).$$

Since $I = G$ implies $\mathcal{V}(I) = \mathcal{V}(G)$, the critical points satisfy the system

$$\begin{aligned} w^2 + 164w - 4 &= 0 \\ 116z - w - 82 &= 0 \\ 116y - w - 82 &= 0 \\ 58x - w - 140 &= 0, \end{aligned}$$

and can be determined by solving the first equation for w then substituting the value of w in the remaining equations and solving for the remaining variables.

The fact that a lexicographic Gröbner basis gives a nice set of generators for solving the critical point system in this example is not a coincidence. Given a (finite or infinite) set $S \subset \mathbb{Q}[\mathbf{z}]$ and natural number $0 \leq \ell \leq d-1$, let S_ℓ denote the elements of S containing only the variables $z_{\ell+1}, \ldots, z_d$.

Theorem 7.2 (Elimination Theorem of Lexicographic Gröbner Bases) *Let $I \subset \mathbb{Q}[\mathbf{z}]$ be an ideal and $G = \{g_1, \ldots, g_r\}$ be a Gröbner basis of I with respect to the* lexicographic *monomial order where $z_1 > \cdots > z_d$. Then for any $0 \leq \ell \leq d - 1$ the set I_ℓ is an ideal of $\mathbb{Q}[z_{\ell+1}, \ldots, z_d]$ with lexicographic Gröbner basis G_ℓ.*

Theorem 7.2 says that the finite set G_ℓ, obtained immediately by inspection from G, generates the ideal of elements in I which contain only the variables $z_{\ell+1}, \ldots, z_d$.

Proof That I_ℓ is an ideal follows directly from the fact that I is an ideal. Let $f \in I_\ell$, so that $f \in \mathbb{Q}[z_{\ell+1}, \ldots, z_d]$. Since $f \in I$ and G is a Gröbner basis of I,

$\mathrm{LT}(f)$ is in the ideal generated by the leading terms of G. Because each leading term is a monomial, it follows that $\mathrm{LT}(f)$ is divisible by $\mathrm{LT}(g)$ for some $g \in G$. But then $\mathrm{LT}(f) \in \mathbb{Q}[z_{\ell+1}, \ldots, z_d]$ implies $\mathrm{LT}(g) \in \mathbb{Q}[z_{\ell+1}, \ldots, z_d]$. Because G is a Gröbner basis with respect to the *lexicographic* order where $z_1 > \cdots > z_d$, $\mathrm{LT}(g) \in \mathbb{Q}[z_{\ell+1}, \ldots, z_d]$ implies $g \in \mathbb{Q}[z_{\ell+1}, \ldots, z_d]$: if g contained a term with a variable z_k for $1 \leq k \leq \ell$ then by the definition of the lexicographic order such a variable with smallest index would also appear in its leading term.

We have shown $f \in I_\ell$ implies $\mathrm{LT}(f)$ is in the ideal $\mathrm{LT}(G_\ell)$. Each element of G is in I, so the ideal (G_ℓ) generated by G_ℓ is contained in I_ℓ. Towards a contradiction suppose $I_\ell \neq (G_\ell)$ and pick $f \in I_\ell \setminus (G_\ell)$ with minimal leading term under the lexicographic order. Then $\mathrm{LT}(f) = \mathrm{LT}(g)$ for some $g \in (G_\ell)$. The leading term of $f - g \in I_\ell$ is smaller than $\mathrm{LT}(f)$ under the lexicographic order, so by minimality of f we have $f - g \in (G_\ell)$ which contradicts $f \notin (G_\ell)$. In particular, it must be the case that $I_\ell = (G_\ell)$ and G_ℓ is a Gröbner basis under the lexicographic order. □

This naturally suggests a method to encode an algebraic set in order to easily determine and manipulate its elements.

Definition 7.11 (triangular systems) A (finite) *triangular system*[7] is a system of the form

$$\begin{aligned} P_1(z_1) &= 0 \\ P_2(z_1, z_2) &= 0 \\ &\vdots \\ P_d(z_1, \ldots, z_d) &= 0, \end{aligned}$$

where each $P_k \in \mathbb{Q}[z_1, \ldots, z_k]$ is monic in z_k.

Theorem 7.2 shows a strong connection between triangular systems and lexicographical Gröbner bases: if the ideal I is zero-dimensional then $\mathcal{V}(I)$ can be written as the union of the zeroes of a finite number of triangular systems, and such systems can be easily determined from a lexicographical Gröbner basis of I; see Lazard [24] for additional details on this approach. One can then approximate or, when possible, exactly determine the elements of $\mathcal{V}(I)$ by iteratively solving the sequence of polynomial equations in each triangular system.

Example 7.8 (An Extended Critical Point System)

Consider again the rational function $1/H(w, x, y, z)$ where $H(w, x, y, z) = 1 - w(1 + x)(1 + y)(1 + z)(1 + y + z + yz + xyz)$. Since the symmetries of the main diagonal hide some of the complexities which arise in general, we now consider asymptotics of the $\mathbf{r} = (2, 1, 1, 1)$-diagonal. The polynomials in the extended critical point system (7.2), which allows us to determine minimality, form the ideal

[7] Triangular systems are also known as regular chains, simple systems, or simple extensions (sometimes with small differences in definition, depending on the author).

$$I = (H, wH_w - 2\lambda, xH_x - \lambda, xH_x - \lambda, yH_y - \lambda, zH_z - \lambda, \tilde{H}) \subset \mathbb{Q}[w, x, y, z, \lambda, t]$$

where $\tilde{H}(t, w, x, y, z) = H(tw, tx, ty, tw)$. The reduced lexicographical Gröbner basis G of I with respect to the ordering of variables $\lambda \geq w \geq x \geq y \geq z \geq t$ contains an element of the form $S(t) = (t-1)R(t)$, where $R(t)$ is a polynomial of degree 18 in t. That $S(t)$ factors reflects the fact that we can write $\mathcal{V}(I) = \mathcal{V}(J) \cup \mathcal{V}(K)$ for ideals $J, K \subsetneq \mathbb{Q}[w, x, y, z, \lambda, t]$. The factor $t-1$ corresponds to the actual critical points of H, where $t = 1$, while the real solutions of the factor $R(t)$ encode points on the lines containing the origin and these critical points.

This factorization of $S(t)$ also implies $\mathcal{V}(I)$ is the union of the zeroes of two triangular systems. Replacing $S(t)$ by $t-1$ in G and recomputing the lexicographical Gröbner basis gives the set of polynomials

$$\{t-1,\ 6z^3-7z^2-5z+2,\ y-z,\ 6z^2+4x-3z-3,\ 4482z^2+72w-8355z+2114,\ 2\lambda+1\},$$

which form a triangular system. Similarly, replacing $S(t)$ by $R(t)$ in G and recomputing the lexicographical Gröbner basis gives another set of polynomials

$$\{R(t), U_1(t, z), U_2(t, y), U_3(t, x), U_4(t, w), 2\lambda + 1\},$$

where U_1, U_2, and U_3 are linear in z, y, x, and w, respectively. Unfortunately, the U_j polynomials are extremely large: each is a polynomial of degree 17 whose *smallest* non-zero coefficient is 51 decimal digits long! Our introduction of the Kronecker representation is motivated in large part by wanting to deal with polynomials whose coefficients have a more manageable size.

The complexity of computing Gröbner bases is well studied, yet tricky to pin down precisely. Doubly-exponential lower bounds are known in the worst case [20], yet implementations in computer algebra systems run reasonably well on many applications. As a practical matter, lexicographic Gröbner bases (often desired for their elimination property) are typically larger and take longer to compute than Gröbner bases with respect to other orders[8]. There are algorithms which convert between Gröbner bases with respect to different orders, allowing one to compute a basis with respect to a computationally well-chosen order then convert to a lexicographic basis for its elimination properties. Especially useful for this conversion on zero-dimensional ideals is the FGLM algorithm of Faugére et al. [12]. Perhaps surprisingly, under one natural measurement the most efficient order for a Gröbner basis is the reverse graded lexicographic order [3].

Although useful for encoding finite algebraic sets, triangular systems have some drawbacks (for instance, iteratively solving polynomial equations for each variable can obscure which solutions are have real coordinates). Instead of going deeper into

---

[8] To quote from the deep analysis of Bayer and Stillman [3], "... use of the lexicographic order in computations can unequivocally be discouraged, in favor of more carefully chosen orders."

the theory of triangular systems, we move on to different representations of algebraic sets more suited to our needs.

7.3.2 Univariate Representations

The key to our algorithms is the parametrization of a finite algebraic set by a univariate polynomial in one variable, with the other coordinates being well-chosen rational functions in that variable. We rely heavily on the results of the appendix to this chapter, which summarize important properties of univariate polynomials. With that in mind, let $I \subset \mathbb{Q}[\mathbf{z}]$ be a zero-dimensional ideal and introduce a new variable u satisfying

$$u = \kappa_1 z_1 + \cdots + \kappa_d z_d \tag{7.11}$$

for some $\kappa_j \in \mathbb{Z}$. Geometrically, introducing u corresponds to parameterizing the elements of $\mathcal{V}(I)$ by the level sets of the hyperplane with normal $\boldsymbol{\kappa}$.

Definition 7.12 (separating linear forms) The variable u in (7.11) is called a *separating linear form* if it takes distinct values for each $\mathbf{z} \in \mathcal{V}(I)$.

Because we care mainly about $\mathcal{V}(I)$, and the zero set of a polynomial is unchanged when taking positive powers, we introduce the following concept.

Definition 7.13 (radical of an ideal) The *radical* of an ideal $K \subset \mathbb{Q}[\mathbf{x}]$ is the set $\sqrt{K}$ consisting of $f \in \mathbb{Q}[\mathbf{x}]$ such that $f^n \in K$ for some $n \in \mathbb{N}$.

Remark 7.5 The radical of an ideal K is itself an ideal of $\mathbb{Q}[\mathbf{x}]$, and is as the largest ideal in $\mathbb{Q}[\mathbf{x}]$ such that $\mathcal{V}(K) = \mathcal{V}(\sqrt{K})$. The radical of an ideal can be computed using Gröbner bases [4, Ch. 8.2].

Let $J \subset \mathbb{Q}[\mathbf{z}, u]$ be the radical of the ideal generated by a zero-dimensional ideal I together with the polynomial $u - \boldsymbol{\kappa} \cdot \mathbf{z}$. When u is a separating linear form then the *shape lemma* [15, Prop. 1.6] states that the reduced Gröbner basis of J with respect to the lexicographic ordering where $z_1 \succ \cdots \succ z_d \succ u$ consists of a univariate polynomial $P(u)$ with no repeated roots together with polynomials of the form $z_j - R_j(u)$ for $1 \le j \le d$. Conversely, if $\mathcal{V}(J)$ has a triangular system of this form then u is separating. To determine the elements of $\mathcal{V}(I)$ one simply needs to solve the univariate polynomial $P(u)$ and substitute the result into each R_j. In fact, if the κ_j are picked at random there is a good chance to obtain a separating linear form.

Lemma 7.1 *Suppose a zero-dimensional ideal I is generated by d polynomials in the variables $z_1, \dots, z_d$, each of degree at most δ. If each κ_j in* (7.11) *is chosen uniformly at random from a finite set $S \subset \mathbb{Z}$ then the probability that the variable u defined by* (7.11) *takes distinct values at the elements of $\mathcal{V}(I)$ satisfies*

$$\mathbb{P}[u \textit{ not distinct at elements of } \mathcal{V}(I)] \le \frac{\binom{D}{2}}{|S|},$$

where $D = \delta^d$. In particular, if S has at least $\binom{D}{2}$ elements then some assignment of $\boldsymbol{\kappa} \in S^d$ produces a linear form u which takes distinct values for $\mathbf{z} \in \mathcal{V}(I)$.

Proof Bézout's theorem [38] implies that a polynomial system with a finite number of solutions defined by d equations in d variables, each of degree at most δ, has at most D solutions. Thus, we may write $\mathcal{V}(I) = \{\mathbf{z}_1, \ldots, \mathbf{z}_r\}$ for $r \le D$ with each $\mathbf{z}_j$ distinct. If $u(\mathbf{z}, \boldsymbol{\kappa})$ denotes the right-hand side of (7.11), the linear form u takes distinct values at the elements of $\mathcal{V}(I)$ if and only if the polynomial

$$P(\boldsymbol{\kappa}) = \prod_{i<j} \big(u(\mathbf{z}_i, \boldsymbol{\kappa}) - u(\mathbf{z}_j, \boldsymbol{\kappa})\big)$$

is non-zero. Problem 7.3 asks you to prove the Schwartz-Zippel lemma, stating that the probability $P(\boldsymbol{\kappa})$ vanishes when each κ_j is chosen randomly from a finite set S is at most $\deg P/|S|$. This gives the desired bound. □

The solutions of systems satisfying the shape lemma can be parametrized by a single variable, which is convenient, but the polynomials in these systems can still have extremely large coefficient sizes.

Example 7.9 (Large Coefficient Sizes in a Univariate Representation)

Consider again the extended critical point system forming the ideal

$$I = (H, wH_w - 2\lambda, xH_x - \lambda, xH_x - \lambda, yH_y - \lambda, zH_z - \lambda, \tilde{H})$$

where $H(w, x, y, z) = 1 - w(1+x)(1+y)(1+z)(1+y+z+yz+xyz)$ and $\tilde{H}(t, w, x, y, z) = H(tw, tx, ty, tw)$. Introducing the polynomial $u - t - x$, the reduced lexicographical Gröbner basis with respect to the ordering of variables $\lambda \ge w \ge x \ge y \ge z \ge t \ge u$ has the expected form

$$G = \{P(u), t - R_1(u), w - R_2(u), z - R_3(u), y - R_4(u), x - R_5(u), 2\lambda + 1\},$$

where P has degree 21 and the R_j polynomials have degree 20. Unfortunately, the coefficients of the R_j polynomials are rational numbers whose numerators and denominators have approximately 160 decimal digits!

Effective versions of the *arithmetic Nullstellensätz* [23, Thm. 1] imply that the maximum height of the polynomials P, R_j appearing in a triangular system when u is a separating linear form is bounded by $\tilde{O}(h\delta^{2d})$, and this upper bound appears in practice (such as the last example). We thus turn to the Kronecker representation, which contains polynomials with heights bounded in $\tilde{O}(h\delta^d)$. As the height of a polynomial is the *bitsize* of the coefficients, this will make our computations much more tractable.

7.3.2.1 The Symbolic Kronecker Representation

Recall from Section 7.2 that a Kronecker representation $[P(u), \mathbf{Q}]$ of a finite algebraic set $\mathcal{A} \subset \mathbb{C}^d$ is defined by a separating linear form

$$u = \kappa_1 z_1 + \cdots + \kappa_d z_d \in \mathbb{Z}[\mathbf{z}],$$

square-free polynomial $P \in \mathbb{Z}[u]$, and $\mathbf{Q} \in \mathbb{Z}[u]^d$ of degrees less than $\deg P$ such that $\mathcal{A}$ is determined by the $\mathbf{z}$-coordinates of the solutions of the system

$$P(u) = 0, \qquad \begin{cases} P'(u)z_1 - Q_1(u) = 0, \\ \qquad\qquad \vdots \\ P'(u)z_d - Q_d(u) = 0. \end{cases}$$

Definition 7.14 (Kronecker degrees and heights) The *degree* of a Kronecker representation is the degree of P, and the *height* of a Kronecker representation is the maximum height of its polynomials $P, Q_1, \ldots, Q_d$.

The following result, contained in Schost [39], suggests that the coefficients appearing in Kronecker representations will be much smaller than the coefficients appearing in triangular systems.

Lemma 7.2 *Let* $\mathbf{f} \in \mathbb{Z}[\mathbf{z}]^d$ *be a set of polynomials of degree at most* δ *and heights at most* h, *and let* $Z(\mathbf{f})$ *be the solutions of* $\mathbf{f}$ *where the Jacobian matrix of* $\mathbf{f}$ *is invertible. Then there exists a Kronecker representation of* $Z(\mathbf{f})$ *of degree at most* $D = \delta^d$ *and height* $\tilde{O}(hD)$.

Remark 7.6 The degree bound in Lemma 7.2 follows directly from Bézout's theorem bounding the number of elements of $Z(\mathbf{f}) \subset \mathcal{V}(\mathbf{f})$. A full derivation of the height bound is outside the scope of this text, but we give an intuitive motivation for the (on first glance somewhat contrived) Kronecker representation[9]. Suppose u is a separating linear form for a zero-dimensional ideal I and represent $\mathcal{V}(I)$ by the solutions of a system

$$P(u) = 0, \; z_1 = R_1(u), \; \ldots, \; z_d = R_d(u)$$

for polynomials $P \in \mathbb{Z}[u]$ and $R_j \in \mathbb{Q}[u]$. Since we are interested only in the elements of $\mathcal{V}(I)$, not in the multiplicities of the elements, we may assume P is square-free by replacing it with its square-free factorization if necessary. Thus, $P(u) = (u - u_1) \cdots (u - u_r)$ where the u_j are the distinct roots of P in $\mathbb{C}$. If z_1 takes the value $a_j \in \mathbb{C}$ at the element of $\mathcal{V}(I)$ obtained by setting $u = u_j$ then we can determine the polynomial $R_1(u)$ by interpolation, since knowing $R_1(u_j) = a_j$ implies

[9] Thanks to Éric Schost for providing this explanation.

$$
\begin{aligned}
R_1(u) &= \sum_{j=1}^{d} a_j \prod_{k\neq j}\left(\frac{u-u_k}{u_j-u_k}\right) \\
&= \sum_{j=1}^{d} \frac{a_j}{\prod_{k\neq j}(u_j-u_k)} \prod_{k\neq j}(u-u_k) \\
&= \sum_{j=1}^{d} \frac{a_j}{P'(u_j)} \prod_{k\neq j}(u-u_k).
\end{aligned}
$$

To express $R_1(u)$ as a polynomial with rational coefficients one must put fractions over a common denominator, multiplying the values a_j by different evaluations $P'(u_j)$, which essentially leads to an extra factor of D in the heights of the R_j polynomials. From this perspective it is clearly natural to encode the values of each $z_j/P'(u)$ instead of just z_j, leading to the Kronecker representation.

By Lemma 7.1, a separating linear form can be determined with high probability by randomly selecting the coefficients $\boldsymbol{\kappa}$ in a finite set of integers of sufficient size. For a given choice of $\boldsymbol{\kappa}$ one can compute a Kronecker representation of $\mathbf{f}$, when u is separating, by using a lexicographical Gröbner basis of the radical of $(\mathbf{f})$ to determine polynomials $P(u)$ and $R_j(u)$ such that $\mathcal{V}(\mathbf{f}) = \mathcal{V}(P, z_1 - R_1, \ldots, z_d - R_d)$ and taking $Q_j(u)$ to be the remainder when the polynomial $P'(u)R_j(u)$ is divided by $P(u)$. When u is not separating this can also be detected from a lexicographical Gröbner basis. A more refined analysis can be found in Rouillier [35].

Example 7.10 (A Kronecker Representation)

In the last example we took the ideal

$$
I = (H, wH_w - 2\lambda, xH_x - \lambda, xH_x - \lambda, yH_y - \lambda, zH_z - \lambda, \tilde{H})
$$

with $H(w,x,y,z) = 1-w(1+x)(1+y)(1+z)(1+y+z+yz+xyz)$ and $\tilde{H}(t,w,x,y,z) = H(tw,tx,ty,tw)$, introduced the linear form $u - t - x$, and obtained a Gröbner basis of the form

$$
G = \{P(u), t - R_1(u), w - R_2(u), z - R_3(u), y - R_4(u), x - R_5(u), 2\lambda + 1\}
$$

which implies u is a separating linear form. If we define, for instance, $Q_1(u)$ to be the remainder when $R_1(u)P'(u)$ is divided by the polynomial $P(u)$ then $Q_1(u)$ is a polynomial of degree 20 whose coefficients each have approximately 17 decimal digits. Though somewhat large for manual consideration, 17 decimal digits is an order of magnitude smaller than the 160 digit integers appearing in $R_1(u)$.

Although this procedure determines a Kronecker representation, it relies on Gröbner basis computations whose complexity is not well understood, and moves through intermediate expressions with polynomials of large coefficient size. There are alternative, Gröbner free, approaches to computing a Kronecker representation. The first well-known example is the so-called geometric resolution algorithm of Giusti et al. [16], but for our complexity results we rely on an algorithm of Safey el Din and Schost [36] which uses a 'homotopy-based approach'. To the best of our knowledge neither of these approaches are implemented in a computer algebra system, so our examples (and the ACSV package of Melczer and Salvy [26]) compute Kronecker representations using lexicographic Gröbner bases. The next result follows from Theorem 1 of Safey El Din and Schost [36].

Proposition 7.3 *Let $\mathbf{f} \in \mathbb{Z}[\mathbf{z}]^d$ be a polynomial system and $Z(\mathbf{f})$ be the solutions of $\mathbf{f}$ where the Jacobian matrix of $\mathbf{f}$ is invertible (this implies $Z(\mathbf{f})$ is finite). Suppose the polynomials in $\mathbf{f}$ have degree at most δ and heights at most h. Then there exists an algorithm* Kronecker *that takes $\mathbf{f}$ as input and produces one of the following:*

- *a Kronecker representation of $Z(\mathbf{f})$,*
- *a Kronecker representation of degree less than that of $Z(\mathbf{f})$,*
- fail.

The first outcome occurs with probability at least 21/32 and returns a Kronecker representation of degree $D = \delta^d$ and height $\tilde{O}(hD)$. In any case, the algorithm has bit complexity $\tilde{O}(hD^3)$.

Repeating the algorithm k times, and taking the output with highest degree, allows one to obtain a Kronecker representation of $Z(\mathbf{f})$ with probability $1 - (11/32)^k$. The probabilistic nature of Kronecker comes from randomly sampling a linear form u with coefficients of reasonable size and choosing certain prime numbers at different points in the algorithm. In practice the probabilistic issues are minor and the probability bounds listed are quite pessimistic: see Giusti et al. [16] or Safey El Din and Schost [36] for further discussion. When the Jacobian of $\mathbf{f}$ is invertible at all of its solutions, as it will be under our assumptions (J1) and (J2), Proposition 7.3 gives a Kronecker representation of all solutions of $\mathbf{f}$.

Our next proposition describes how to encode the values of additional polynomials at the solutions of a Kronecker representation. Since the constants appearing in the asymptotic expansion (5.27) of Theorem 5.2 in Chapter 5 are polynomials in the critical points of $F(\mathbf{z})$, this will allow us to determine an expansion of the form (7.9).

Proposition 7.4 *Let $\mathbf{f} \in \mathbb{Z}[\mathbf{z}]^d$ be a zero-dimensional polynomial system containing polynomials of degree at most δ and heights at most h, and suppose $[P(u), \mathbf{Q}]$ is a Kronecker representation of $\mathcal{V}(\mathbf{f})$ determined by Proposition 7.3. Let $q(\mathbf{z}) \in \mathbb{Z}[\mathbf{z}]$ be a polynomial of degree at most ρ and height at most η. Then*

1. *there exists a parametrization $q(u) = Q_q(u)/P'(u)$ of the values taken by q on $\mathcal{V}(\mathbf{f})$, where Q_q is an integer polynomial of degree less than $D = \delta^d$ and height $\tilde{O}((\eta + h\rho)D)$. The polynomial Q_q can be computed in $\tilde{O}((\eta + h\rho)D^3)$ bit operations;*
2. *there exists a polynomial $\Phi_q(T) \in \mathbb{Z}[T]$ of degree at most $D = \delta^d$ and height $\tilde{O}((\eta + h\rho)D^2)$ which vanishes on all values taken by $q(\mathbf{z})$ for $\mathbf{z} \in \mathcal{V}(\mathbf{f})$. The polynomial Φ_q can be computed in $\tilde{O}((\eta + h\rho)D^4)$ bit operations;*

3. *when q has degree 1 (for instance, if q is one of the variables z_j) then there is a polynomial $\Phi_q \in \mathbb{Z}[T]$ vanishing on the values of $q(\mathbf{z})$ for $\mathbf{z} \in \mathcal{V}(\mathbf{f})$ which has degree at most $D = \delta^d$ and height $\tilde{O}((\eta + h)D)$, and Φ_q can be computed in $\tilde{O}((\eta + h)D^3)$ bit operations.*

Proposition 7.4 (1) is proven by adding the polynomial $T - q(\mathbf{z})$ into $\mathbf{f}$, where T is a new variable, and recomputing the Kronecker representation. In order to get the stated height and complexity bound, a more refined method of Safey El Din and Schost [36] must be used which exploits the structure of the new system, since T appears in only one equation. This refinement requires an additional algebraic framework whose setup would take us away from our main focus, so we refer the reader to Melczer and Salvy [26, Prop. 45]. In practice one does not actually need to recompute the entire Kronecker system.

Recall the resultant from Problem 2.14 in Chapter 2. The minimal polynomial Φ_q in Proposition 7.4 (2) divides the resultant of the polynomials $P'(u) - TQ_q(u)$ and $P(u)$ with respect to u, so the stated bounds on its height and degree follow from Lemmas 7.8 and 7.11 in the appendix. The complexity of determining Φ_q follows from a fast algorithm of Kedlaya and Umans [21] for determining the minimal polynomial of Q_q/P' modulo the equation $P(u) = 0$. The improved height bound on Φ_q when q is linear follows from Safey El Din and Schost [36, Lemma 23].

7.3.2.2 The Numeric Kronecker Representation

A Kronecker representation allows us to describe the solutions of a finite algebraic set (for instance, the critical points of a multivariate rational function), but to argue about specific solutions (for instance, the minimal critical points) we need extra information about the roots. Because distinct elements encoded by a Kronecker representation $[P(u), \mathbf{Q}]$ are represented by distinct roots of $P(u)$, it is enough to separate the roots of $P(u)$ in the complex plane. A numeric Kronecker representation is thus a triple $[P(u), \mathbf{Q}, \mathbf{U}]$, where $\mathbf{U}$ contains isolating intervals for the real roots of P and isolating disks for the complex roots of P.

Definition 7.15 (sizes of isolating regions) We say the *size* of an interval is half its length, while the *size* of a disk is its radius.

Using the algorithms and results discussed in the appendix to this chapter, we build up a toolbox of algorithms for manipulating the solutions of a (numeric) Kronecker representation. Our first result, which follows directly from the algorithm PolyRoots described in Lemma 7.12 of the appendix, discusses how to determine the isolating regions $\mathbf{U}$ for the roots of $P(u)$. One can compute the isolating regions to arbitrary accuracy, with a minimal amount of accuracy required to ensure the computed regions separate the (distinct) roots of P.

Proposition 7.5 *Suppose the zero-dimensional system $\mathbf{f} = (f_1, \ldots, f_d) \in \mathbb{Z}[\mathbf{z}]^d$ is given by a Kronecker representation $[P(u), \mathbf{Q}]$ of degree $\mathcal{D}$ and height η. Then there* *exists an algorithm* NumKronecker *which takes* $[P(u), \mathbf{Q}]$ *and* $\kappa > 0$ *and returns a numeric Kronecker representation* $[P(u), \mathbf{Q}, \mathbf{U}]$*, with isolating regions in* $\mathbf{U}$ *of size at most* $2^{-\kappa}$*, in* $\tilde{O}(\mathcal{D}^3 + \mathcal{D}^2\eta + \mathcal{D}\kappa)$ *bit operations.*

Once we have a Kronecker representation, the most basic operation we can perform is to approximate the coordinates of the algebraic points it encodes.

Lemma 7.3 *Given a Kronecker representation* $[P(u), \mathbf{Q}]$ *of degree* $\mathcal{D}$ *and height* η *and* $\kappa \in \mathbb{N}$*, approximations to the solutions of the Kronecker representation with isolating regions for each coordinate of size* $2^{-\kappa}$ *can be determined in* $\tilde{O}(d(\mathcal{D}^3 + \mathcal{D}^2\eta + \mathcal{D}\kappa))$ *bit operations.*

Proof Fix a coordinate z_j and root $v \in \mathbb{C}$ of $P(u) = 0$. Our aim is to evaluate $z_j = Q_j(v)/P'(v)$ to an accuracy of $2^{-\kappa}$. Assume we have approximations $q \approx Q_j(v)$ and $p \approx P'(v)$ such that $|Q_j(v) - q|, |P'(v) - p| \leq 2^{-n}$ for some $n \in \mathbb{N}$. Then

$$\left|\frac{Q_j(v)}{P'(v)} - \frac{q}{p}\right| = \left|\frac{Q_j(v)p - qp + qp - qP'(v)}{P'(v)p}\right| \leq 2^{-n}\left(\frac{1}{|P'(v)|} + \left|\frac{q}{P'(v)p}\right|\right).$$

Suppose now that

$$n \geq 2\eta\mathcal{D} + 2(\mathcal{D}-1)\log_2\mathcal{D} - 2\eta + (\mathcal{D}-1)\log_2\sqrt{\mathcal{D}+1} + 1 = \tilde{O}(\eta\mathcal{D}).$$

Proposition 7.7 (i) in the appendix implies $|v| \leq 2^{\eta} + 1$, so

$$\begin{aligned} |q| \leq |Q_j(v)| + 2^{-n} &\leq 2^{\eta}(1 + |v| + \cdots + |v|^d) + 2^{-a} \\ &\leq 2^{\eta}(\mathcal{D}+1)(2^{\eta}+1)^{\mathcal{D}} + 2^{-n} \\ &= 2^{\tilde{O}(\eta\mathcal{D})}. \end{aligned}$$

Similarly, Proposition 7.7 (iv) in the appendix implies $|P'(v)| \geq 2^{-n+1}$ (in fact our bound on n was chosen so that this inequality would hold) and

$$|p| \geq |P'(v)| - 2^{-n} \geq 2^{-n} = 2^{-\tilde{O}(\eta\mathcal{D})},$$

which means

$$\left|\frac{Q_j(v)}{P'(v)} - \frac{q}{p}\right| \leq 2^{-n+\tilde{O}(\eta\mathcal{D})}.$$

To determine z_j to precision κ it is therefore sufficient to determine $Q_j(v)$ and $P'(v)$ to $\kappa + \tilde{O}(\eta\mathcal{D})$ bits. Lemmas 7.12 and 7.13 in the appendix imply that this can be done at all roots of P in $\tilde{O}(\mathcal{D}^3 + \mathcal{D}^2\eta + \mathcal{D}\kappa)$ bit operations, and doing this for each of the d coordinates gives the stated complexity. □

To determine minimality we need to select the solutions of a Kronecker representation which are positive and real, and test which have real coordinates in various ranges.

Lemma 7.4 *Given a numeric Kronecker representation* $[P(u), \mathbf{Q}, \mathbf{U}]$ *of degree* $\mathcal{D}$ *and height* η*, it can be determined for every real solution to the underlying system* *whether each coordinate is positive, negative, or exactly zero, in* $\tilde{O}(d(\mathcal{D}^3 + \mathcal{D}^2\eta))$ *bit operations.*

Proof Fix a coordinate z_j. The roots of $P(u)$ that correspond to solutions with $z_j = 0$ are exactly those canceling the polynomial $G_j(u) = \gcd(P, Q_j)$. By Lemma 7.10 in the appendix, the gcd G_j has height $\tilde{O}(\mathcal{D} + \eta)$ and can be computed in $\tilde{O}(\mathcal{D}^2 + \eta\mathcal{D})$ bit operations. Using Proposition 7.7 and Lemma 7.12 from the appendix, the roots of $P(u)$ which also cancel $G_j(u)$ can be determined in $\tilde{O}(\mathcal{D}^3 + \mathcal{D}^2\eta)$ bit operations by computing isolating regions of the roots of size $2^{-\tilde{O}(\mathcal{D}^2+\eta\mathcal{D})}$. Proposition 7.7 in the appendix shows that knowing approximations of $Q_j(u)$ and $P'(u)$ to an accuracy of $2^{-\tilde{O}(\eta)}$ allows one to determine their signs when they are real, and Lemma 7.13 in the appendix shows that such approximations can be determined in $\tilde{O}(\eta\mathcal{D}^2)$ bit operations knowing only $\tilde{O}(\eta\mathcal{D})$ bits of the roots of $P(u)$. □

Having obtained high-accuracy approximations of the solutions of a critical point system, we need to determine when specific coordinates of the solutions are exactly equal, and when they share the same modulus. Proposition 7.4 gives a degree and height bound on the minimal polynomial of the values of each coordinate z_j at the solutions of $\mathcal{V}(\mathbf{f})$. Using the root separation bounds in Proposition 7.7 of the appendix and the algorithm for approximating $\mathcal{V}(\mathbf{f})$ discussed in Lemma 7.3 results in the following.

Lemma 7.5 *Given a numeric Kronecker representation* $[P(u), \mathbf{Q}, \mathbf{U}]$ *corresponding to a zero-dimensional system of d polynomials of degrees at most δ and heights at most h, the coordinates of its solutions which are exactly equal can be found in* $\tilde{O}(hD^3)$ *bit operations, where* $D = \delta^d$.

The most complicated, and expensive, operation we must perform is grouping solutions of an algebraic set by coordinate-wise modulus. To see why grouping by modulus is so difficult, let $A \in \mathbb{Z}[u]$ be a polynomial of degree $\mathcal{D}$ and height η, and let $\alpha, \beta \in \mathbb{C}$ be roots of A with $|\alpha| \neq |\beta|$. Proposition 7.7 in the appendix gives a useful lower bound $|\alpha - \beta| \geq 2^{-\tilde{O}(\eta\mathcal{D})}$ on the distance between α and β, but it is much harder to get a good bound on the distance between the moduli $|\alpha|$ and $|\beta|$. One approach to bounding the difference of $|\alpha|$ and $|\beta|$ involves the resultant $R(u) = \mathrm{Res}_T(A(T), T^{\mathcal{D}}A(u/T))$, which vanishes at the products of roots of $A(u)$. In particular, it vanishes at $|\alpha|^2 = \alpha\overline{\alpha}$ and $|\beta|^2 = \beta\overline{\beta}$ so the polynomial $B(T) = R(\sqrt{T})R(-\sqrt{T})$ vanishes at $|\alpha|$ and $|\beta|$. Lemmas 7.8 and 7.11 of the appendix imply that $B(T)$ has degree at most $\mathcal{D}$ and height in $\tilde{O}(\eta\mathcal{D}^2)$, so Proposition 7.7 in the appendix yields $||\alpha| - |\beta|| \geq 2^{-\tilde{O}(\eta\mathcal{D}^3)}$. This is significantly worse than the distance bound between α and β. In fact, a tight separation bound between the moduli of roots of a univariate polynomial doesn't seem to be known[10].

[10] A tight bound in the special case when α and β are both real (i.e., when one is positive and the other negative) is given by Bugeaud et al. [8]. In this situation one can take $f(\delta, h) = O(h\delta + \delta^2)$. See also Bugeaud et al. [7] for some recent results and experiments with low degree polynomials.

Open Problem 7.1 *Under what conditions does there exist a function* $f(\mathcal{D}, \eta) = o(\eta\mathcal{D}^3)$ *such that* $||\alpha| - |\beta|| \geq 2^{-f(\mathcal{D},\eta)}$ *whenever* $|\alpha| \neq |\beta|$ *and* α *and* β *are both complex roots of some polynomial of degree* $\mathcal{D}$ *and height* η*?*

Directly using a moduli separation bound of $2^{-\tilde{O}(\eta\mathcal{D}^3)}$ would increase the cost of our ACSV algorithms. Luckily, in the combinatorial case we only need to determine points with the same coordinate-wise moduli as critical points which themselves are real and positive. This allows for a more efficient algorithm[11].

Lemma 7.6 *Let* $A \in \mathbb{Z}[T]$ *be a square-free polynomial of degree* $\mathcal{D} \geq 2$ *and height* η*, and define* $G(T) = A(\sqrt{T})A(-\sqrt{T})$*. If* $A(\alpha) = 0$ *and* $A(\pm|\alpha|) \neq 0$*, then* $\left|G\left(|\alpha|^2\right)\right| \geq 2^{-b}$ *where* $b = \tilde{O}(\eta\mathcal{D}^2)$ *is given by*

$$b = (\mathcal{D}^2 - 1)\log_2(\mathcal{D} + 1) + (2\eta + \log_2(\mathcal{D} + 1))(\mathcal{D}^2 - 1) \\ + 2\eta\mathcal{D} + 2\mathcal{D}\log_2(\mathcal{D}) + \mathcal{D}\log_2\sqrt{\mathcal{D}^2 + 1}.$$

Proof Again we consider the resultant $R(u) = \mathrm{Res}_T(A(T), T^d A(u/T))$ which vanishes at $|\alpha|^2$ for any root α of $A(u)$. Lemma 7.11 of the appendix implies R has degree at most $\mathcal{D}^2$ and height at most $2\eta\mathcal{D} + \log((2\mathcal{D})!) \leq 2\eta\mathcal{D} + 2\mathcal{D}\log\mathcal{D}$. The polynomial $G(T)$ has degree $\mathcal{D}$ and, by Lemma 7.8 of the appendix, height at most $2\eta + \log(\mathcal{D} + 1)$. The stated bound on $G(|\alpha|^2)$ then follows directly from Proposition 7.7 (iii) in the appendix. □

Corollary 7.1 *Under the same conditions as Lemma 7.6, isolating regions of size* $2^{-\tilde{O}(\eta\mathcal{D}^2)}$ *for the real positive roots* $0 < r_1 < \cdots < r_k$ *of* $A(T)$ *and all roots of moduli exactly* $r_1, \ldots, r_k$ *can be computed in* $\tilde{O}(\eta\mathcal{D}^3)$ *bit operations.*

Proof Suppose α is a root of A and define the polynomial $G(T)$ and constant b as in Lemma 7.6. If $a \in \mathbb{C}$ is such that $|\alpha - a| < 2^{-b-\eta-2}$ then $|\overline{\alpha} - \overline{a}| < 2^{-b-\eta-2}$ and

$$\left||\alpha|^2 - a\overline{a}\right| = |\alpha\overline{\alpha} - \alpha\overline{a} + \alpha\overline{a} - a\overline{a}| \leq |\alpha|2^{-b-\eta-2} + |\overline{a}|2^{-b-\eta-2} \leq 2^{-b}$$

since $|\alpha| \leq 2^h + 1$ by Proposition 7.7 in the appendix. Thus, approximating α to $\tilde{O}(\eta\mathcal{D}^2)$ bits allows one to approximate $|\alpha|^2$ to $\tilde{O}(\eta\mathcal{D}^2)$ bits and, by Lemma 7.13 of the appendix, this is sufficient to compute $G\left(|\alpha|^2\right)$ to accuracy 2^{-b}. Lemma 7.6 in the appendix implies at least one of $\pm|\alpha|$ is a root of A if and only if $\left|G\left(|\alpha|^2\right)\right| < 2^{-b}$, so approximating α to accuracy $\tilde{O}(\eta\mathcal{D}^2)$ bits allows one to decide whether or not at least one of $\pm|\alpha|$ is a root of A; this precision is also sufficient to decide which real roots of A are positive. To determine which of $\pm|\alpha|$ are roots of A, Proposition 7.7 in the appendix implies it is sufficient to evaluate $A(|\alpha|)$ and $A(-|\alpha|)$ to precision $2^{-\tilde{O}(\eta\mathcal{D}^2)}$. All these operations run in the stated complexity. □

Corollary 7.2 *Given a numeric Kronecker representation* $[P(u), \mathbf{Q}, \mathbf{U}]$ *of degree* $\mathcal{D} = \delta^d$ *and height* $\eta = \tilde{O}(hD)$ *corresponding to a zero-dimensional system of* d

---

[11] Thanks to Bruno Salvy for this construction.

polynomials in $\mathbb{Z}[\mathbf{z}]$ *of degrees at most* δ *and heights at most* h*, determining the real solutions of the system and listing the solutions with the same coordinate-wise moduli in the variables* $\mathbf{z}$ *can be done in* $\tilde{O}(hD^4)$ *bit operations.*

7.4 Algorithmic ACSV Correctness and Complexity

We conclude this chapter by proving Theorem 7.1. Correctness of Algorithms 1–3 under our assumptions follows from Theorem 5.4 in Chapter 5, together with the results in Section 5.3 of Chapter 5 characterizing minimality. Thus, it remains only to prove the complexity bounds stated in Theorem 7.1.

Complexity of Algorithm 1 and Numerical Approximation of $A(u), B(u), C(u)$

The Kronecker representation of the set C of critical points can be determined in $\tilde{O}(h\delta^3D^3)$ bit operations by Proposition 7.3, since C contains $d+1$ variables. The entries in the $(d-1)\times(d-1)$ matrix $\widetilde{\mathcal{H}}$ have degrees in $\tilde{O}(\delta)$ and heights in $\tilde{O}(h+\log\delta)$, so a cofactor expansion shows the determinant of $\widetilde{\mathcal{H}}$ has degree in $\tilde{O}(d\delta)$ and height in $\tilde{O}(d(h+\log\delta))$. Proposition 7.4 then implies that the polynomial $Q_{\widetilde{\mathcal{H}}}$ has degree less than δD, height in $\tilde{O}(h\delta^2D)$, and can be determined in $\tilde{O}(h\delta^4D^3)$ bit operations. By assumption the polynomial $G(\mathbf{z})$ has degree at most δ and height at most h, and the polynomial $T(\mathbf{z}) = z_1^{r_1}\cdots z_d^{r_d}$ has degree $O(d)$ and height 1. Thus, Proposition 7.4 implies that the polynomials Q_T and Q_{-G} fall into the degree and height bounds for $Q_{\widetilde{\mathcal{H}}}$, and can be determined in the same complexity. Knowing these bounds, Proposition 7.7 (iii) implies that the roots of $P(u)$ where $Q_{\widetilde{\mathcal{H}}}$ and Q_G vanish can be determined by evaluating these polynomials to $\tilde{O}(h\delta^2D^2)$ bits, which takes $\tilde{O}(h\delta^3D^3)$ bit operations using Lemmas 7.12 and 7.13 in the appendix.

With our degree and height bounds on the polynomials $P(u), P'(u), Q_{\widetilde{\mathcal{H}}}(u)$, $Q_T(u), Q_\lambda(u)$, and $Q_{-G}(u)$, and the knowledge that $Q_{\widetilde{\mathcal{H}}}, Q_T(u)$, and $Q_\lambda(u)$ are non-zero at the roots of $P(u)$, an argument analogous to the one presented in the proof of Lemma 7.3 shows that to determine

$$A(u) = \frac{Q_{-G}(u)}{r_dQ_\lambda(u)}, \qquad B(u) = \frac{(r_dQ_\lambda(u))^{d-1}(P')^{2-d}}{Q_{\widetilde{\mathcal{H}}}(u)}, \qquad C(u) = \frac{P'(u)}{Q_T(u)}$$

at all roots of $P(u)=0$ to κ bits of precision requires $\tilde{O}(\delta D\kappa + h\delta^3D^3)$ bit operations. Note at least $\kappa = \tilde{O}(h\delta^2D^2)$ bits of precision are needed to isolate the roots of $P(u)$.

Complexity of Algorithm 2

By Proposition 7.5, the numeric Kronecker representations can be determined to the required accuracy in $\tilde{O}(h\delta^2D^2)$ bit operations. The most expensive operation is the determination of elements of $\mathbf{U}$ to $\tilde{O}(h\delta^3 D^3)$ bits, in order to group critical points of the same modulus. Determining the roots of $P(u)$ to the required accuracy takes $\tilde{O}(h\delta^3 D^4)$ bit operations by Lemma 7.12 in the appendix, dominating the complexity.

Complexity of Algorithm 3

Determining the numeric Kronecker representation to the required $\tilde{O}(h\delta^8 D^8)$ bits, needed to check which solutions have t in $(0,1)$, can be done in $\tilde{O}(h\delta^{12} D^{12})$ bit operations by Proposition 7.5. This is the most expensive step of the algorithm, the others being simple calculations or computations similar to those discussed in the last two algorithms.

We have now proven Theorem 7.1, validating our algorithms for ACSV.

Appendix on Solving and Bounding Univariate Polynomials

This appendix contains some classical results about univariate polynomials which are used in our construction of the numeric Kronecker representation. Our presentation of height bounds and root separation is inspired by Mignotte [27], and also parallels the appendix of Melczer and Salvy [26]. Our bounds involve the following concepts.

Definition 7.16 (Mahler measure and Euclidean norm) Given a polynomial

$$P(z) = c_D z^D + \cdots + c_0 = c_D(z - r_1)\cdots(z - r_D) \in \mathbb{C}[z]$$

with roots $r_j \in \mathbb{C}$ and leading coefficient $c_D \neq 0$, the *Mahler measure* $M(P)$ and *Euclidean norm* $\|P\|_2$ of P are the quantities

$$M(P) = |c_D| \prod_{j=1}^{D} \max\{1, |r_j|\}, \qquad \|P\|_2 = \left(\sum_{j=1}^{D} |c_j|^2 \right)^{1/2}.$$

Note that $\|P\|_2 \leq 2^{h(P)}\sqrt{\deg(P) + 1}$.

Since the set of roots of a product of polynomials is the union of the roots, the Mahler measure satisfies $M(PQ) = M(P)M(Q)$ for any $P, Q \in \mathbb{C}[z]$. The Mahler measure will be key to the results we derive in this appendix, but it can be difficult to calculate directly as it relies on the roots of P. We thus begin by bounding the Mahler measure by the Euclidean norm, which is easy to calculate.

Proposition 7.6 (Landau's bound) *For any $P \in \mathbb{C}[z]$ the Mahler measure satisfies $M(P) \leq \|P\|_2$, with equality if and only if P is a monomial.*

Our proof of Proposition 7.6 uses the following technical lemma, which Problem 7.9 asks you to establish.

Lemma 7.7 *If* $P(z) \in \mathbb{C}[z]$ *then* $\|(z - r)P(z)\|_2 = \|(\overline{r}z - 1)P(z)\|_2$ *for any* $r \in \mathbb{C}$, *where* $\overline{r}$ *denotes the complex conjugate of* r.

Proof (Proposition 7.6) Rearrange the roots of P so that $r_1, \ldots, r_\ell$ are the roots with moduli larger than 1, meaning $M(P) = |c_D r_1 \cdots r_\ell|$. If

$$Q(z) = c_D \prod_{k=1}^{\ell} (\overline{r_j} z - 1) \prod_{k=\ell+1}^{d} (z - r_j)$$

then the leading coefficient of Q is $c_D \overline{r_1} \cdots \overline{r_\ell}$ and by Lemma 7.7

$$\|P\|_2 = \|Q\|_2 \geq |c_D \overline{r_1} \cdots \overline{r_\ell}| = |c_D r_1 \cdots r_\ell| = M(P).$$

The value $\|Q\|_2$ equals the leading term of Q only if Q is a monomial. □

7.4.1 Polynomial Height Bounds

First, we use the Mahler measure to prove bounds on the heights of univariate polynomials.

Lemma 7.8 *For polynomials* $P_1, \ldots, P_k, P, Q \in \mathbb{Z}[z]$,

$$h(P_1 + \cdots + P_k) \leq \max_i h(P_i) + \log_2 k,$$
$$h(P_1 \cdots P_k) \leq \sum_{i=1}^{k} h(P_i) + \sum_{i=1}^{k-1} \log_2(\deg P_i + 1),$$
$$h(P) \leq \deg P + \log_2 \|PQ\|_2.$$

Proof The first two results follow directly from the definition of polynomial height, so we consider only the last equation. Expanding $P(z) = c_D(z - r_1) \cdots (z - r_D)$ gives Vieta's formulas for the coefficients of P as symmetric functions in its roots,

$$c_k = c_D(-1)^k \sum_{i_1 < \cdots < i_{D-k}} r_{i_1} \cdots r_{i_{D-k}}, \qquad (1 \leq k \leq D).$$

Each summand here is at most $M(P)$, so $|c_k| \leq \binom{D}{k} M(P)$ for all $1 \leq k \leq D$. Since $M(Q) \geq 1$ for any $Q \in \mathbb{Z}[z]$, this implies the maximum of the $|c_k|$ is at most $2^D M(P) \leq 2^D M(P) M(Q) = 2^D M(PQ) \leq 2^D \|PQ\|_2$. □

7.4.2 Polynomial Root Bounds

Our next goal is to establish separation bounds for the roots of an integer polynomial.

Proposition 7.7 *Let $P \in \mathbb{Z}[z]$ be a polynomial of degree $D \geq 2$ and height h. Supposing $P(\alpha) = 0$,*

(i) if $\alpha \neq 0$ then $1/(2^h + 1) \leq |\alpha| \leq 2^h + 1$;
(ii) if $P(\beta) = 0$ and $\alpha \neq \beta$ then $|\alpha - \beta| \geq \sqrt{3} D^{-(D+2)/2} \, \|P\|_2^{1-D}$;
(iii) if $Q(\alpha) \neq 0$ for $Q \in \mathbb{Z}[z]$ then $|Q(\alpha)| \geq (\deg Q + 1)^{1-D} 2^{h(Q)(1-D)} \|P\|_2^{-\deg Q}$;
(iv) if P is square-free then $|P'(\alpha)| \geq 2^{-2Dh+2h} D^{2(1-D)} (D+1)^{(1-D)/2}$.

We split the proof of Proposition 7.7 into three parts.

Proof of Proposition 7.7 (i)

Let $P(z) = c_D z^D + \cdots + c_1 z + p_0$ with $c_D \neq 0$, and let $K = \max_{1 \leq j \leq D-1} |c_j|$. Suppose $P(\alpha) = 0$ and consider the upper bound in Proposition 7.7 (i). We show the stronger result that $|\alpha| \leq 1 + K/|c_D|$, following a proof of Cauchy [9, p. 28]. If $|\alpha| \leq 1$ then the result holds trivially. Otherwise, $|\alpha| > 1$ and

$$|c_D \alpha^D| = |c_{D-1}\alpha^{D-1} + \cdots + c_0| \leq K(|\alpha|^{D-1} + \cdots + |\alpha| + 1) = K \frac{|\alpha|^D - 1}{|\alpha| - 1} \leq K \frac{|\alpha|^D}{|\alpha| - 1},$$

so $|\alpha| \leq 1 + K/|c_D|$ as desired. The lower bound in Proposition 7.7 (i) comes from applying the upper bound to the reverse polynomial $z^D P(1/z)$, whose height is the same as P but whose roots are the reciprocals of the roots of P. □

Proof of Proposition 7.7 (ii)

We start by establishing the following bound on the differences of the roots of P.

Lemma 7.9 *Let $P(z) = c_D z^D + \cdots + c_0 = c_D(z - r_1) \cdots (z - r_D) \in \mathbb{Z}[z]$. Then for any distinct $1 \leq a, b \leq D$,*

$$\sqrt{3} \prod_{i<j} |r_i - r_j| \leq |r_a - r_b| D^{(D+2)/2} \left(\frac{M(P)}{|c_D|} \right)^{D-1}.$$

Proof The constant $\prod_{i<j}(r_i - r_j)$ is the determinant of the Vandermonde matrix

$$V = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ r_1 & r_2 & \cdots & r_D \\ \vdots & \vdots & \cdots & \vdots \\ r_1^{D-1} & r_2^{D-1} & \cdots & r_D^{D-1} \end{pmatrix}.$$

Pick $1 \leq a < b \leq D$ and replace the ath column of this matrix with the ath column minus the bth. Denote the columns of the resulting matrix by $\mathbf{v}_1, \ldots, \mathbf{v}_D$. Note that

$$\prod_{j=1}^{D} \|\mathbf{v}_j\|_2 = \left(\sum_{k=1}^{D-1} \left|r_a^k - r_b^k\right|^2\right) \prod_{k \neq a} \left(1 + |r_k|^2 + \cdots + |r_k|^{2(D-1)}\right)$$
$$\leq \left(\sum_{k=1}^{D-1} \left|\frac{r_a^k - r_b^k}{r_a - r_b}\right|^2\right) |r_a - r_b|^2 \prod_{k \neq a} \max\left(D, D|r_k|^{2(D-1)}\right).$$

If $B = \max(1, |r_a|, |r_b|)$ then

$$B^{-2(D-1)} \left(\sum_{k=1}^{D-1} \left|\frac{r_a^k - r_b^k}{r_a - r_b}\right|^2\right) = \sum_{k=1}^{D-1} \left|\sum_{j=0}^{k-1} B^{-(D-1)} r_a^j r_b^{k-j-1}\right|^2 \leq \sum_{k=1}^{D-1} k^2$$
$$= \frac{D(D-1)(2D-1)}{6}$$
$$\leq D^3/3.$$

Problem 7.8 asks you to prove Hadamard's inequality, which implies $|\det V| \leq \prod_{j=1}^{D} \|\mathbf{v}_j\|_2$ and thus

$$\prod_{i<j} |r_i - r_j| \leq |r_a - r_b|^2 \left(D^{D+2}/3\right) \prod_{1 \leq k \leq D} \max\left(1, |r_k|\right)^{2(D-1)}.$$

The stated inequality then follows from taking square-roots. □

Returning to Proposition 7.7 (ii), recall that the *discriminant* of P is the integer

$$\Delta(P) = c_D^{2D-2} \prod_{i<j} (r_i - r_j)^2 \in \mathbb{Z},$$

which is zero if and only if P has a repeated root. When P has no repeated roots then $\Delta(P)$ is a non-zero integer, so Lemma 7.9 implies

$$\sqrt{3} \leq \sqrt{3}|\Delta(P)| \leq |r_a - r_b|\, D^{(D+2)/2} \left(\frac{M(P)}{|c_D|}\right)^{D-1}. \tag{7.12}$$

If P has no repeated roots this establishes Proposition 7.7 (ii) as $M(P) \leq \|P\|_2$. If P does have repeated roots then the product $\tilde{P}$ of the square-free factors of P is an integer polynomial whose roots are precisely the distinct roots of P, and Proposition 7.7 (ii) follows from establishing (7.12) with P replaced by $\tilde{P}$ and using $M(\tilde{P}) \leq M(P) \leq \|P\|_2$. □

Proof of Proposition 7.7 (iii) and (iv)

If $Q(\alpha) \neq 0$ and $P, Q \in \mathbb{Z}[z]$ then there exists an integer polynomial factor $\tilde{P}$ of P containing α as a root such that $\tilde{P}$ and Q share no roots. Suppose that $\tilde{P}$ and Q have leading coefficients a and b, respectively. Since the resultant of $\tilde{P}$ and Q is a non-zero integer,

$$1 \le |\operatorname{Res}(\tilde{P}, Q)| = |a|^{\deg Q} |b|^{\deg \tilde{P}} \prod_{\substack{\tilde{P}(\gamma)=0 \\ Q(\beta)=0}} |\gamma - \beta| = |a|^{\deg Q} \prod_{\tilde{P}(\gamma)=0} |Q(\gamma)| \,. \quad (7.13)$$

For any $\gamma \in \mathbb{C}$,

$$|Q(\gamma)| \le (\deg Q + 1) 2^{h(Q)} \max(1, |\gamma|)^{\deg Q},$$

so, using $\deg \tilde{P} \le D$ and $M(\tilde{P}) \le M(P) \le \|P\|_2$, we obtain

$$\begin{aligned} |a|^{\deg Q} \prod_{\substack{\tilde{P}(\gamma)=0 \\ \gamma \neq \alpha}} |Q(\gamma)| &\le (\deg Q + 1)^{\deg \tilde{P} - 1} 2^{h(Q)(\deg \tilde{P} - 1)} \left(\frac{M(\tilde{P})}{\max(1, |\alpha|)} \right)^{\deg Q} \\ &\le (\deg Q + 1)^{D-1} 2^{h(Q)(D-1)} \|P\|_2^{\deg Q}. \end{aligned}$$

Combining this bound with (7.13) gives Proposition 7.7 (iii). Proposition 7.7 (iv) is a special instance of Proposition 7.7 (iii) with $Q = P'(z)$. □

7.4.3 Resultant and GCD Bounds

We also require bounds on the resultant and greatest common divisor of two univariate polynomials. Lemma 7.8 directly bounds the greatest common divisor of two polynomials, and an efficient algorithm achieving the following complexity can be found in von zur Gathen and Gerhard [40, Cor. 11.14].

Lemma 7.10 *For $P, Q \in \mathbb{Z}[U]$ of degrees at most D and heights at most h,* $\gcd(P, Q)$ *has height $\tilde{O}(D + h)$ and can be computed in $\tilde{O}(D^2 + hD)$ bit operations.*

The resultant of two polynomials is defined by the determinant of the matrix in Problem 2.14 of Chapter 2. Computing this determinant using a cofactor expansion and applying Lemma 7.8 to the resulting expression gives the following bounds on the resultant.

Lemma 7.11 *For $P, Q \in \mathbb{Z}[T, U]$ let $R = \operatorname{Res}_T(P, Q)$ and*

$$\begin{aligned} \delta &= \deg_T P \deg_U Q + \deg_T Q \deg_U P \\ \eta &= h(P) \deg_T Q + h(Q) \deg_T P + \log_2((\deg_T P + \deg_T Q)!) \\ &\qquad + \log_2(\deg_U P + 1) \deg_T Q + \log_2(\deg_U Q + 1) \deg_T P. \end{aligned}$$

Then $\deg R \le \delta$ and $h(R) \le \eta$. Furthermore, if all coefficients of P and Q as polynomials in T are monomials in U then $h(R) \le h(P) \deg_T Q + h(Q) \deg_T P + \log_2((\deg_T P + \deg_T Q)!)$.

7.4.4 Algorithms for Polynomial Solving and Evaluation

The Kronecker representation encodes the solutions of a multivariate polynomial system using the roots of a square-free univariate polynomial. This is useful, because numerical solvers for univariate polynomials have been studied for hundreds of years, and are implemented in every major computer algebra system. For our complexity analysis we use the following lemma, containing the main results of Sagraloff and Mehlhorn [37] (for real roots) and Mehlhorn et al. [25] (for complex roots).

Lemma 7.12 *Let $A \in \mathbb{Z}[T]$ be a square-free polynomial of degree D and height h. Then there exists an algorithm* PolyRoots *which takes A and a positive integer κ and returns isolating intervals for all real roots of $A(T)$ and isolating disks for all complex roots of $A(T)$, each of size less than $2^{-\kappa}$, in $\tilde{O}(D^3 + D^2 h + D\kappa)$ bit operations.*

In order to determine the solutions specified by a Kronecker representation $[P, \mathbf{Q}]$ we need to evaluate the polynomials Q_j at (approximate) roots of $P(u)$. The following result is contained in Kobel and Sagraloff [22].

Lemma 7.13 *Let $A \in \mathbb{Z}[T]$ be a square-free polynomial of degree D and height h. Then there exists an algorithm* PolyEval *which takes A, a positive integer κ, and a sequence $t_1, \dots, t_m \in \mathbb{C}$ of length $m = O(D)$, and returns approximations $a_1, \dots, a_m \in \mathbb{C}$ such that $|A(t_j) - a_j| < 2^{-\kappa}$ for all $1 \leq j \leq m$ using $\tilde{O}(D(h + \kappa + D \log \max_j |t_j|))$ bit operations. The algorithm needs only $\tilde{O}(h + \kappa + D \log \max_j |t_j|)$ bits of $t_1, \dots, t_m$ to return the approximations $a_1, \dots, a_m$, and if the t_j are real then the approximations a_j are also real.*

Problems

7.1 For any $\delta \in \mathbb{N}_{>0}$, prove that the determinant of the Jacobian of the polynomial system (7.2) is non-zero at the solutions of (7.2) when $H(\mathbf{z}) = 1 - z_1^\delta - \cdots - z_d^\delta$. Using Corollary 5.7 from Chapter 5, conclude that assumption (J1) in the combinatorial case holds generically.

7.2 Algorithms 2 and 3 allow one to prove minimality of critical points with respect to the power series expansion of $F(\mathbf{z})$ using Proposition 5.4. Modify these algorithms to prove minimality with respect to any convergent Laurent expansion of $F(\mathbf{z})$, where the expansion under consideration is specified by a point in its domain of convergence with rational coordinates.

7.3 Prove the Schwartz-Zippel lemma: if $P(\mathbf{z}) \in \mathbb{K}[z_1, \dots, z_d]$ is a polynomial of degree δ over a field $\mathbb{K}$ and S is a finite subset of $\mathbb{K}$, then the probability that P vanishes when each z_j is picked uniformly at random from S satisfies

$$\mathbb{P}\left[P(\mathbf{z}) = 0 : \mathbf{z} \in S^d\right] \leq \frac{\delta}{|S|}.$$

Hint: When $d = 1$ the polynomial P has at most δ zeroes. Working inductively, write a polynomial $P \in \mathbb{K}[\mathbf{z}]$ as a polynomial in the variable z_d whose coefficients are polynomials in $\mathbb{K}[z_1, \dots, z_{d-1}]$.

7.4 Recall that the second Apéry number sequence, related to the irrationality of $\zeta(2)$, is the main power series diagonal of the rational function $1/H(x, y, z)$ where $H(x, y, z) = 1 - (1 + z)(x + y + xy)$. Using a Gröbner basis calculation, determine a Kronecker representation of the extended critical point system

$$\mathcal{S} = \{H, xH_x - \lambda, yH_y - \lambda, zH_z - \lambda, H(tx, ty, tz)\}.$$

Use this to find the minimal critical points and give asymptotics of the diagonal.

7.5 The main power series diagonal of

$$F(x, y, t) = \frac{(1 + x)(1 + y)}{1 - txy(x/y + y/x + xy + 1/(xy) + y + 1/y)}$$

enumerates the number of walks on the six steps $\{(\pm1, \pm1), (0, \pm1)\}$ which begin at the origin and stay in the non-negative quadrant. Using a Gröbner basis calculation, determine a Kronecker representation of the extended critical point system and use this to determine the minimal critical points of F. Do not forget to find all minimal critical points with the same coordinate-wise modulus!

7.6 Modify our algorithms to find the smooth minimal critical points of the rational function

$$F(\mathbf{z}) = \frac{z_1 \cdots z_d}{(1 - z_1) \cdots (1 - z_d)\left(1 - \sum_{i=2}^{d}(i - 1)e_i(\mathbf{z})\right)}$$

for small values of $d \in \mathbb{N}$, where $e_i(\mathbf{z})$ is the ith elementary symmetric function

$$e_i(\mathbf{z}) = \sum_{1 \le j_1 < \cdots < j_i \le d} z_{j_1} \cdots z_{j_i}.$$

Write a program which takes d and returns asymptotics of the diagonal. Note that the denominator of F is not smooth, but all smooth minimal critical points will be zeroes of the denominator's last factor. Try to push your algorithm to the highest value of d possible. Recall that this sequence enumerates singular cubic tensors [32].

7.7 If $M = (m_{i,j})$ is a $D \times D$ positive-definite matrix, prove the inequality $|\det(M)| \le \prod_{k=1}^{D} |m_{k,k}|$. *Hint: Reduce to the case where the diagonal elements have modulus one, then use the arithmetic-geometric mean inequality on the product of M's eigenvalues.*

7.8 Use Problem 7.7 to prove Hadamard's inequality: if M is a $D \times D$ matrix with columns $\mathbf{v}_1, \dots, \mathbf{v}_D$ then $|\det(M)| \le \prod_{k=1}^{D} \|\mathbf{v}_k\|_2$. *Hint: Make use of the matrix $N = MM^*$, where M^* is the conjugate transpose of M.*

7.9 Prove Lemma 7.7 by expanding $\|(z - r)P(z)\|_2^2$ in terms of the coefficients of P.

References

1. Yuliy Baryshnikov, Stephen Melczer, Robin Pemantle, and Armin Straub. Diagonal asymptotics for symmetric rational functions via ACSV. In *29th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms*, volume 110 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages Art. No. 12, 15. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2018.
2. Saugata Basu, Richard Pollack, and Marie-Françoise Roy. *Algorithms in real algebraic geometry*, volume 10 of *Algorithms and Computation in Mathematics*. Springer-Verlag, Berlin, 2003.
3. David Bayer and Michael Stillman. A theorem on refining division orders by the reverse lexicographic order. *Duke Math. J.*, 55(2):321–328, 1987.
4. Thomas Becker and Volker Weispfenning. *Gröbner bases*, volume 141 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1993. A computational approach to commutative algebra, In cooperation with Heinz Kredel.
5. B. Buchberger. *Ein Algorithmus zum Auffinden der Basiselemente des Restklassenringes nach einem nulldimensionalen Polynomideal*. PhD thesis, University of Innsbruck, 1965.
6. B. Buchberger. Ein algorithmisches Kriterium für die Lösbarkeit eines algebraischen Gleichungssystems. *Aequ. Math.*, 4(3):374–383, 1970. English translation in: B. Buchberger, F. Winkler: Groebner Bases and Applications, Proc. of the International Conference "33 Years of Groebner Bases", 1998, RISC, Austria, London Math. Society Lecture Note Series 251, Cambridge Univ. Press, 1998, pp. 535–545.
7. Yann Bugeaud, Andrej Dujella, Wenjie Fang, Tomislav Pejković, and Bruno Salvy. Absolute root separation. *Experimental Mathematics*, 0(0):1–8, 2019.
8. Yann Bugeaud, Andrej Dujella, Tomislav Pejković, and Bruno Salvy. Absolute real root separation. *Amer. Math. Monthly*, 124(10):930–936, 2017.
9. Augustin Louis Cauchy. *Mémoire sur la résolution des équations numériques et sur la théorie de l'élimination*. Chez de Bure Frères, A Paris, 1829.
10. David Cox, John Little, and Donal O'Shea. *Ideals, varieties, and algorithms*. Undergraduate Texts in Mathematics. Springer, New York, third edition, 2007.
11. David Eisenbud. *Commutative algebra*, volume 150 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995. With a view toward algebraic geometry.
12. J. C. Faugère, P. Gianni, D. Lazard, and T. Mora. Efficient computation of zero-dimensional Gröbner bases by change of ordering. *J. Symbolic Comput.*, 16(4):329–344, 1993.
13. Jean-Charles Faugére. A new efficient algorithm for computing Gröbner bases (F_4). *J. Pure Appl. Algebra*, 139(1-3):61–88, 1999. Effective methods in algebraic geometry (Saint-Malo, 1998).
14. Jean-Charles Faugère. A new efficient algorithm for computing Gröbner bases without reduction to zero (F_5). In *Proceedings of the 2002 International Symposium on Symbolic and Algebraic Computation*, pages 75–83. ACM, New York, 2002.
15. Patrizia Gianni and Teo Mora. Algebraic solution of systems of polynomial equations using Groebner bases. In *Applied algebra, algebraic algorithms and error-correcting codes (Menorca, 1987)*, volume 356 of *Lecture Notes in Comput. Sci.*, pages 247–257. Springer, Berlin, 1989.
16. Marc Giusti, Grégoire Lecerf, and Bruno Salvy. A Gröbner free alternative for polynomial system solving. *J. Complexity*, 17(1):154–211, 2001.
17. R. L. Graham, M. Grötschel, and L. Lovász, editors. *Handbook of combinatorics. Vol. 1, 2*. Elsevier Science B.V., Amsterdam; MIT Press, Cambridge, MA, 1995.
18. D. Yu. Grigor'ev and N. N. Vorobjov, Jr. Solving systems of polynomial inequalities in subexponential time. *J. Symbolic Comput.*, 5(1-2):37–64, 1988.
19. Lars Hörmander. *An introduction to complex analysis in several variables*, volume 7 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam, third edition, 1990.

20. Dũng T. Huỳnh. A superexponential lower bound for Gröbner bases and Church-Rosser commutative Thue systems. *Inform. and Control*, 68(1-3):196–206, 1986.
21. Kiran S. Kedlaya and Christopher Umans. Fast polynomial factorization and modular composition. *SIAM J. Comput.*, 40(6):1767–1802, 2011.
22. Alexander Kobel and Michael Sagraloff. On the complexity of computing with planar algebraic curves. *J. Complexity*, 31(2):206–236, 2015.
23. Teresa Krick, Luis Miguel Pardo, and Martín Sombra. Sharp estimates for the arithmetic Nullstellensatz. *Duke Math. J.*, 109(3):521–598, 2001.
24. D. Lazard. Solving zero-dimensional algebraic systems. *J. Symbolic Comput.*, 13(2):117–131, 1992.
25. Kurt Mehlhorn, Michael Sagraloff, and Pengming Wang. From approximate factorization to root isolation with application to cylindrical algebraic decomposition. *J. Symbolic Comput.*, 66:34–69, 2015.
26. S. Melczer and B. Salvy. Effective coefficient asymptotics of multivariate rational functions via semi-numerical algorithms for polynomial systems. *Journal of Symbolic Computation*, 103:234–279, 2021.
27. Maurice Mignotte. *Mathematics for computer algebra*. Springer-Verlag, New York, 1992. Translated from the French by Catherine Mignotte.
28. Teo Mora. *Solving polynomial equation systems. I*, volume 88 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2003. The Kronecker-Duval philosophy.
29. Teo Mora. *Solving polynomial equation systems. II*, volume 99 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2005. Macaulay's paradigm and Gröbner technology.
30. Teo Mora. *Solving polynomial equation systems. Vol. III. Algebraic solving*, volume 157 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2015.
31. Teo Mora. *Solving polynomial equation systems. Vol. IV. Buchberger theory and beyond*, volume 158 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2016.
32. Jay Pantone. The asymptotic number of simple singular vector tuples of a cubical tensor. *Online Journal of Analytic Combinatorics*, 12, 2017.
33. Robin Pemantle and Mark C. Wilson. Twenty combinatorial examples of asymptotics derived from multivariate generating functions. *SIAM Rev.*, 50(2):199–272, 2008.
34. James Renegar. On the computational complexity and geometry of the first-order theory of the reals. I. Introduction. Preliminaries. The geometry of semi-algebraic sets. The decision problem for the existential theory of the reals. *J. Symbolic Comput.*, 13(3):255–299, 1992.
35. Fabrice Rouillier. Solving zero-dimensional systems through the rational univariate representation. *Appl. Algebra Engrg. Comm. Comput.*, 9(5):433–461, 1999.
36. Mohab Safey El Din and Éric Schost. Bit complexity for multi-homogeneous polynomial system solving—Application to polynomial minimization. *J. Symbolic Comput.*, 87:176–206, 2018.
37. Michael Sagraloff and Kurt Mehlhorn. Computing real roots of real polynomials. *J. Symbolic Comput.*, 73:46–86, 2016.
38. Joachim Schmid. On the affine Bezout inequality. *Manuscripta Math.*, 88(2):225–232, 1995.
39. Éric Schost. *Sur la résolution des systèmes polynomiaux à paramétres*. PhD thesis, 2001.
40. Joachim von zur Gathen and Jürgen Gerhard. *Modern computer algebra*. Cambridge University Press, Cambridge, third edition, 2013.

Part III
Non-Smooth ACSV

Chapter 8
Beyond Smooth Points: Poles on a Hyperplane Arrangement

Algebra is a general Method of Computation by certain signs and symbols which have been contrived for the Purpose, and found convenient.
— Colin Maclaurin

Put thyself into the trick of singularity.
— William Shakespeare (as Malvolio)

Let $F(\mathbf{z})$ be a meromorphic function with singular set $\mathcal{V}$. In Chapter 5 we saw how the behaviour of F near a finite number of singularities can dictate coefficient asymptotics *when $\mathcal{V}$ is locally smooth near the points of interest*. This chapter begins our descent into the general theory of analytic combinatorics in several variables by relaxing the condition that $\mathcal{V}$ is smooth. Here, we study rational functions

$$F(\mathbf{z}) = \frac{G(\mathbf{z})}{\prod_{j=1}^{m} \ell_j(\mathbf{z})^{p_j}}$$

where each ℓ_j is a real linear function. The singular set $\mathcal{V}$ of such a function is a union of hyperplanes. If $\boldsymbol{\sigma} \in \mathcal{V}$ lies on exactly one of these hyperplanes then $\mathcal{V}$ is smooth in a neighbourhood of $\boldsymbol{\sigma}$, but $\mathcal{V}$ is not locally smooth near singularities in the intersection of multiple hyperplanes. Generalizing the smooth case of Chapter 5, we will show that if $\boldsymbol{\sigma}$ lies on the intersection of k linearly independent hyperplanes then the asymptotic contribution of $\boldsymbol{\sigma}$ is determined by taking k univariate residues followed by asymptotically approximating a $(d-k)$-dimensional saddle-point integral. Because the singular set of F is so easily characterized, we will be able to carry out the required computations explicitly using linear algebra and standard integral approximations of the form seen in previous chapters.

We will prove that the asymptotic behaviour of a coefficient sequence $[\mathbf{z}^{n\mathbf{r}}]F(\mathbf{z})$ varies smoothly with the direction $\mathbf{r}$ as it stays in different cones of $\mathbb{R}^d$. Directions interior to these cones, defined precisely below, are known as *generic directions*. Theorem 8.2 below shows how to determine asymptotics in generic directions using information directly computable from $F(\mathbf{z})$. Algorithm 4 gives a high-level implementation of this result which, when combined with the computational techniques discussed in Chapter 7, automatically determines asymptotics in generic directions. We discuss asymptotics along non-generic directions in Section 8.3.

Example 8.1 (An Illustration of the Theory)

Consider the power series expansion of the rational function

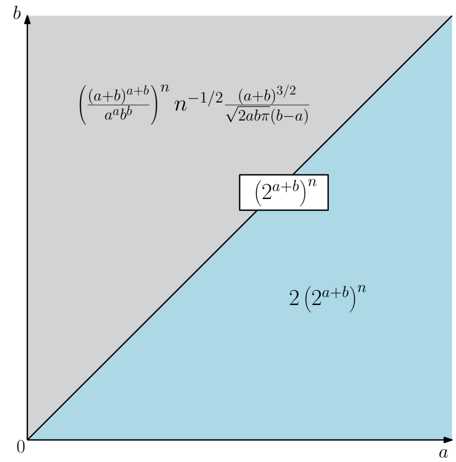


Fig. 8.1 The three asymptotic regimes for the coefficients $[x^{an}y^{bn}]F(x,y)$ of the rational function $F(x,y) = 1/(1-2x)(1-x-y)$.

$$F(x,y) = \frac{1}{(1-2x)(1-x-y)},$$

whose set of singularities $\mathcal{V}$ consists of the union of two hyperplanes. Then $\mathcal{V}$ is locally a hyperplane at all of its points except for $(1/2, 1/2)$, where the two hyperplanes meet, and the theory below implies $\mathbf{r} = (a,b)$ is a non-generic direction if and only if $a = b$. When $a \neq b$ then Theorem 8.2 determines the asymptotic growth of $[x^{an}y^{bn}]F(x,y)$, while in non-generic directions asymptotics are given by Theorem 8.3. Asymptotics for this example are displayed in Figure 8.1.

Early studies of multivariate generating functions with linear denominators were undertaken in queuing theory work of Bertozzi and McKenna [2] and Kogan and Yakovlev [4]; see also Kogan [3].

Example 8.2 (Closed Multiclass Queuing Networks)

Bertozzi and McKenna [2] use multivariate generating functions and residue computations similar to those below to investigate so-called partition functions for queuing networks. For instance, the generating function of a 'closed multiclass queuing network with no infinite servers' has the form

$$F(\mathbf{z}) = \frac{1}{\left(1 - \sum_{j=1}^d \rho_{j,1} z_j\right)\left(1 - \sum_{j=1}^d \rho_{j,2} z_j\right)\cdots\left(1 - \sum_{j=1}^d \rho_{j,m} z_j\right)}$$

for real constants $\rho_{j,k} > 0$ depending on model parameters [2, Eq. (4.18)]. Theorem 8.2 below determines the asymptotics of interest in any generic direction.

Problem 8.6 asks you to determine asymptotics for the two-dimensional case

$$F(x, y) = \frac{1}{(1 - \rho_{1,1}x - \rho_{2,1}y)(1 - \rho_{1,2}x - \rho_{2,2}y)},$$

paying attention to how the values of the weights $\rho_{i,j}$ change asymptotic behaviour.

Our asymptotic argument proceeds in five steps:

1. decompose the Cauchy integral expression for series coefficients as a sum of integrals whose domains of integration are *imaginary fibers* (defined below);
2. identify a finite set of contributing singularities where local behaviour of F determines asymptotics;
3. deform the imaginary fibers to be sufficiently close to contributing singularities;
4. use residue calculations to obtain saddle-point integrals;
5. asymptotically approximate the saddle-point integrals.

Because our singular sets are defined by real linear functions, we are able to carry out the necessary calculations using linear algebra and the techniques of Chapter 5 for functions with smooth contributing singularities. In Chapter 9 we extend our results to rational functions whose singular sets *locally* behave like the union of hyperplanes, before summarizing the theory for general rational functions.

We begin our analysis in this chapter with some preliminary setup. The presentation of this chapter is based on Baryshnikov et al. [1].

8.1 Setup and Definitions

Fix a rational function $F(\mathbf{z}) = G(\mathbf{z})/H(\mathbf{z})$ defined by coprime polynomials $G(\mathbf{z})$ and

$$H(\mathbf{z}) = \prod_{j=1}^{m} \ell_j(\mathbf{z})^{p_j},$$

where each p_j is a positive integer and each

$$\ell_j(\mathbf{z}) = 1 - \mathbf{b}^{(j)} \cdot \mathbf{z} = 1 - b_1^{(j)} z_1 - \cdots - b_d^{(j)} z_d$$

is a real linear function. We seek asymptotics of a convergent Laurent expansion

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

of $F(\mathbf{z})$ in some domain $\mathcal{D} \subset \mathbb{C}^d$. In order to be as explicit as possible in this chapter, we consider only power series expansions and real linear functions. Arbitrary Laurent expansions and complex linear functions are covered by the general theory developed in Chapter 9; Problem 8.3 also discusses some aspects of the general case.

> Unless otherwise specified, in this chapter we take the power series expansion of F and thus consider only directions $\mathbf{r}$ with positive coordinates.

Because G and H are coprime, Proposition 3.2 from Chapter 3 implies that the singular set of F is the union of hyperplanes

$$\mathcal{V} = \{\mathbf{z} : H(\mathbf{z}) = 0\}.$$

To discuss the local geometry of $\mathcal{V}$ we make the following definitions.

Definition 8.1 (flats, strata, and arrangements) For any $S = \{k_1, \ldots, k_s\} \subset \{1, \ldots, m\}$, the *flat* defined by the linear functions $\ell_{k_1}, \ldots, \ell_{k_s}$ is the set

$$\mathcal{V}_S = \mathcal{V}_{k_1,\ldots,k_s} = \{\mathbf{z} \in \mathbb{C}^d : \ell_{k_1}(\mathbf{z}) = \cdots = \ell_{k_s}(\mathbf{z}) = 0\}$$

of their common solutions, with corresponding *real flat*

$$\mathcal{V}'_S = \mathcal{V}_S \cap \mathbb{R}^d = \{\mathbf{z} \in \mathbb{R}^d : \ell_{k_1}(\mathbf{z}) = \cdots = \ell_{k_s}(\mathbf{z}) = 0\}.$$

The *stratum* defined by S is the flat $\mathcal{V}_S$ minus any other flats it strictly contains,

$$\mathcal{S}_S = \mathcal{V}_S \setminus \bigcup_{\mathcal{V}_T \subsetneq \mathcal{V}_S} \mathcal{V}_T.$$

The *dimension* of $\mathcal{V}_S$ is its dimension as a linear space, and the dimension of $\mathcal{S}_S$ is defined to be the dimension of $\mathcal{V}_S$. A *hyperplane arrangement* $\mathcal{A}$ defined by $\mathcal{V}$ is the collection of maximal subsets of $\{1, \ldots, m\}$ corresponding to distinct non-empty flats. The collection of strata formed by the elements of a hyperplane arrangement of $\mathcal{V}$ form a *stratification* of $\mathcal{V}$.

Remark 8.1 Our definition of a hyperplane arrangement is necessary as different index sets can give rise to the same flat. For instance, when three pairs of lines intersect at a common point $\mathbf{p}$ then the flat defined by any two of them is $\{\mathbf{p}\}$. The study of hyperplane arrangements is a well established field in combinatorics, detailed, for instance, in Orlik and Terao [5].

Remark 8.2 (a perspective from differential geometry) Each stratum $\mathcal{S}_S$ forms a complex manifold, and the dimension of $\mathcal{S}_S$ according to Definition 8.1 equals the dimension of $\mathcal{S}_S$ as a manifold. A stratification of $\mathcal{V}$ partitions $\mathcal{V}$, which is not locally a manifold near any point lying in at least two distinct flats, into complex manifolds. Because flats are algebraic objects (defined by linear equalities) they are easy to define and manipulate, but our eventual integral expressions will require the strata, which are *semi-algebraic* (defined by linear equalities and in-equalities). For our analytic arguments to work properly we also need the strata of $\mathcal{V}$ to 'fit together nicely'. In this chapter we simply require independence between the linear factors at any common solution. In Section 9.3 of Chapter 9 we discuss the more general concept of a *Whitney stratification* for an algebraic set.

Definition 8.2 (simple functions) The function $F(\mathbf{z})$ is called *simple* if any collection of coefficient vectors $\mathbf{b}^{(k_1)}, \ldots, \mathbf{b}^{(k_s)}$ is linearly independent whenever $\ell_{k_1}(\mathbf{z}) = \cdots = l_{k_s}(\mathbf{z}) = 0$ has a solution (i.e. if any subset of hyperplanes comprising $\mathcal{V}$ that share a common point have linearly independent normals).

When F is simple then $\mathcal{V}$ defines only one hyperplane arrangement, consisting of subsets $S \subset \{1, \ldots, m\}$ such that $\mathcal{V}_S$ is non-empty, and the dimension of any stratum $\mathcal{S}_S$ equals $d - |S|$. We will see that if F is not simple then we can algorithmically decompose F as the sum of simple functions.

As in previous chapters, our analysis begins with a multivariate Cauchy integral

$$[\mathbf{z}^{n\mathbf{r}}]F(\mathbf{z}) = \frac{1}{(2\pi i)^d} \int_{\mathcal{T}} F(\mathbf{z}) \frac{d\mathbf{z}}{\mathbf{z}^{n\mathbf{r}+\mathbf{1}}}, \tag{8.1}$$

where $\mathcal{T}$ is a product of circles in the domain of convergence $\mathcal{D}$. Again the growth of the Cauchy integrand is captured by the height function

$$h_{\mathbf{r}}(\mathbf{z}) = h(\mathbf{z}) = -\sum_{j=1}^{d} r_j \log |z_j|.$$

In Chapter 5, when we considered only smooth points, we searched for the *smooth critical points* of $\mathcal{V}$ by checking when the gradient $(\nabla H)(\mathbf{z})$ was a multiple of $(\nabla h_{\mathbf{r}})(\mathbf{z})$. The smooth critical points gave potential minimizers of the height function on $\overline{\mathcal{D}}$, and thus helped determine the singularities of interest for an asymptotic analysis. Now that we consider non-smooth singularities we must refine our notion of critical points: if $\mathbf{z}$ is a root of at least two of the ℓ_j then $(\nabla H)(\mathbf{z}) = 0$, so any non-smooth point trivially satisfies the smooth critical point equations (5.16) derived in Chapter 5. We thus define critical points relative to each stratum, as each stratum defines a smooth set.

Definition 8.3 (critical points) For any $S = \{k_1, \ldots, k_s\} \subset \{1, \ldots, m\}$ such that $\mathbf{b}^{(k_1)}, \ldots, \mathbf{b}^{(k_s)}$ are linear independent, the *critical points of the flat* $\mathcal{V}_S$ *in the direction* $\mathbf{r}$ are the points $\mathbf{z} \in \mathbb{R}_*^d$ with non-zero real coordinates where the matrix

$$N = \begin{pmatrix} -\nabla \ell_{k_1} \\ \vdots \\ -\nabla \ell_{k_s} \\ -\nabla h \end{pmatrix} = \begin{pmatrix} & \mathbf{b}^{(k_1)} & \\ & \vdots & \\ & \mathbf{b}^{(k_s)} & \\ r_1/z_1 & \cdots & r_d/z_d \end{pmatrix}$$

is rank deficient. If $s = d$ then every point in $\mathcal{V}_S$ is critical (linear independence of the $\mathbf{b}^{(j)}$ implies s cannot be greater than d). Otherwise, the critical points of $\mathcal{V}_S$

are the solutions of the system defined by the ℓ_j together with the determinants of all $(s+1) \times (s+1)$ matrix minors of N. We call a critical point on the flat $\mathcal{V}_S$ a *critical point of the stratum* $\mathcal{S}_S$ when it also lies in $\mathcal{S}_S$ (in general it may only be a limit point of the stratum). The *set of critical points* of F, denoted Ω, consists of all critical points on any flat defined by linearly independent collections of the ℓ_j. Each $\boldsymbol{\sigma} \in \Omega$ lies in a unique stratum, which we call *the stratum of* $\boldsymbol{\sigma}$ and denote $S(\boldsymbol{\sigma})$.

Remark 8.3 Up to a reordering of variables and the ℓ_k, we may assume that N contains pivots in its first s diagonal entries. For $j = 1, \ldots, d-s$, let N_j denote the $(s+1) \times (s+1)$ matrix constructed from the first s columns of N together with its $(s+j)$th column. By definition, at any critical point the determinants of the N_j matrices vanish. Conversely, if all determinants of the N_j vanish then each of the last $d-s$ columns of N lie in the dimension s linear space spanned by the first s columns. Thus, the critical points on $\mathcal{V}_S$ are the real solutions of the *critical point equations for* $\mathcal{V}_S$,

$$\ell_{k_1} = \cdots = \ell_{k_s} = \det N_1 = \det N_{d-s} = 0. \tag{8.2}$$

If $d = s$ then (8.2) simply states $\ell_{k_1} = \cdots = \ell_{k_s} = 0$.

Remark 8.4 When $S = \{k\}$ contains a single element then (8.2) becomes the smooth critical point equations (5.16) from Chapter 5 with $H^s = \ell_k$.

Our definition of critical points will be a posteriori validated by showing that, after suitable residue computations, the Cauchy integral for series coefficients becomes a finite sum of integrals near critical points which can be asymptotically approximated.

A critical point defined by a flat may not lie on the stratum defined by the flat, a pathological case which restricts some of the necessary computations. Luckily, as the following definition suggests, we will see that this typically does not occur.

Definition 8.4 (generic directions) The direction $\mathbf{r}$ is *generic* if for each flat $\mathcal{V}_S$ any critical point $\boldsymbol{\sigma}$ of $\mathcal{V}_S$ lies on the stratum $\mathcal{S}_S$ (i.e., does not lie in a subflat of $\mathcal{V}_S$).

With this setup out of the way, we are now able to work through our main arguments. We split the analysis into two cases, first dealing with asymptotics in generic directions and then discussing non-generic directions. The general theory is developed alongside a detailed running example.

8.2 Asymptotics in Generic Directions

Returning to the Cauchy integral, and recalling that we consider the power series expansion of $F(\mathbf{z})$, the coefficients of interest are given by

$$f_{n\mathbf{r}} = [\mathbf{z}^{n\mathbf{r}}]F(\mathbf{z}) = \frac{1}{(2\pi i)^d} \int_{\mathcal{T}} \frac{G(\mathbf{z})}{\prod_{j=1}^{m} \ell_j(\mathbf{z})^{p_j}} \frac{d\mathbf{z}}{\mathbf{z}^{n\mathbf{r}+\mathbf{1}}} \tag{8.3}$$

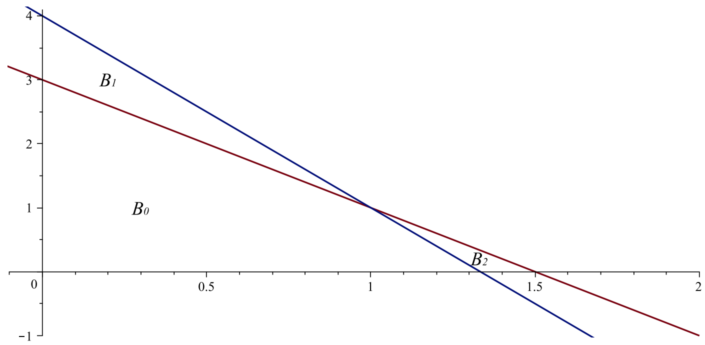


Fig. 8.2 The real part of the hyperplane arrangement $l_1(x, y)l_2(x, y) = 0$ in our running example, with the bounded components B_0, B_1, and B_2 of $\mathcal{M}_{\mathbb{R}}$ marked.

whenever $n\mathbf{r} = (nr_1, \ldots, nr_d)$ is a positive integer vector and $\mathcal{T}$ is a product of circles sufficiently close to the origin. For most of this section our results include as a hypothesis that F is simple. In sub-Section 8.2.6 below we discuss how non-simple functions can be decomposed into a finite sum of simple functions, to which our main results can then be applied.

To simplify notation we write the Cauchy integrand in (8.3) as

$$\omega = \frac{G(\mathbf{z})}{\mathbf{z}^{n\mathbf{r}+\mathbf{1}} \; \prod_{j=1}^{m} \ell_j(\mathbf{z})^{p_j}},$$

and let $\mathcal{W} = \mathcal{V}(z_1 \cdots z_d H)$ denote the singular set of ω, equal to $\mathcal{V}$ with the coordinate hyperplanes added. The domain of analyticity of ω is $\mathcal{M} = \mathbb{C}^d \setminus \mathcal{W}$, so that the domain of integration $\mathcal{T}$ in (8.3) can deformed without changing the value of the Cauchy integral as long as the deformations stay in $\mathcal{M}$. Because $H(\mathbf{z})$ is the product of real linear functions, the real part $\mathcal{M}_{\mathbb{R}} = \mathcal{M} \cap \mathbb{R}^d$ gives a good approximation to $\mathcal{M}$ for asymptotic purposes, and is more easily visualized. Similarly, we let $\mathcal{V}_{\mathbb{R}}$ denote the real elements of $\mathcal{V}$. The hyperplanes comprising $\mathcal{V}_{\mathbb{R}}$ split $\mathcal{M}_{\mathbb{R}}$ into convex polyhedra, and we let $\mathcal{B}$ denote the collection of all connected components of $\mathcal{M}_{\mathbb{R}}$.

Our running example in this section is the following.

Example 8.3 (Running Example in the Generic Case)

Let $F(x, y) = \frac{1}{\ell_1(x,y)\ell_2(x,y)}$, where

$$\ell_1 = 1 - \frac{2x + y}{3} \qquad \text{and} \qquad \ell_2 = 1 - \frac{3x + y}{4}.$$

Then given a direction $\mathbf{r} = (r_1, r_2)$, the coefficient

$$[x^{r_1 n} y^{r_2 n}]F(x, y) = \frac{1}{(2\pi \mathrm{i})^2} \int_{\mathcal{T}} \frac{1}{\left(1 - \frac{2x+y}{3}\right)\left(1 - \frac{3x+y}{4}\right)} \frac{dxdy}{x^{r_1 n+1} y^{r_2 n+1}} \qquad (8.4)$$

where

$$\mathcal{T} = \{(\varepsilon_1 e^{i\theta_1}, \varepsilon_2 e^{i\theta_2} : -\pi < \theta_1, \theta_2 \leq \pi\}$$

is any product of circles with sufficiently small radii $\varepsilon_1, \varepsilon_2 > 0$. In particular, no zero (a, b) of $H(x, y) = \ell_1(x, y)\ell_2(x, y)$ can satisfy $|a| \leq \varepsilon_1$ and $|b| \leq \varepsilon_2$, so in this example it is necessary and sufficient that

$$\frac{2\varepsilon_1 + \varepsilon_2}{3} < 1 \qquad \text{and} \qquad \frac{3\varepsilon_1 + \varepsilon_2}{4} < 1.$$

Below we take $\varepsilon_1 = \varepsilon_2 = \varepsilon$ for $\varepsilon < 1$. The bounded components of $\mathcal{M}_{\mathbb{R}}$ form the set $\mathcal{B} = \{B_0, B_1, B_2\}$ whose elements are displayed in Figure 8.2.

8.2.1 Step 1: Express the Cauchy Integral as Sum of Imaginary Fibers

We first convert the Cauchy integral into a sum of integrals over particularly convenient domains of integration.

Definition 8.5 (imaginary fiber) An *imaginary fiber* is any set of the form

$$C_{\mathbf{x}} = \mathbf{x} + i\mathbb{R}^d = \{\mathbf{x} + i\mathbf{y} : \mathbf{y} \in \mathbb{R}^d\} \subset \mathbb{C}^d,$$

with *basepoint* $\mathbf{x} \in \mathbb{R}^d$.

As the linear functions ℓ_j have real coefficients, if $\ell_j(\mathbf{x}) \neq 0$ for $\mathbf{x} \in \mathbb{R}^d$ then

$$\ell_j(\mathbf{x} + i\mathbf{y}) = \ell_j(\mathbf{x}) + i\ell_j(\mathbf{y}) \neq 0$$

for any $\mathbf{x} + i\mathbf{y} \in C_{\mathbf{x}}$. The real domain of analyticity $\mathcal{M}_{\mathbb{R}}$ of ω consists of convex connected components, and given two imaginary fibers with basepoints $\mathbf{x}$ and $\mathbf{y}$ in the same component of $\mathcal{M}_{\mathbb{R}}$ there is a continuous deformation between the fibers which stays in $\mathcal{M}$, defined by

$$g(t) = t\mathbf{x} + (1 - t)\mathbf{y} + i\mathbb{R}^d$$

for $0 \leq t \leq 1$. This makes imaginary fibers useful for our purposes, as they can be easily deformed without crossing $\mathcal{M}$. Because the Cauchy integrand decays exponentially in n at points sufficiently far from the origin, if $\mathbf{x}$ and $\mathbf{y}$ lie in the same component of $\mathcal{M}_{\mathbb{R}}$ then $\int_{C_{\mathbf{x}}} \omega = \int_{C_{\mathbf{y}}} \omega$. The Cauchy integral can be written as a sum of integrals over imaginary fibers by 'stretching' circles along the imaginary axis.

Example 8.3 Continued (Deforming to Imaginary Fibers))

Because $\ell_1(x, y)$ and $\ell_2(x, y)$ have real coefficients, the circles $|x| = \varepsilon$ and $|y| = \varepsilon$ which make up $\mathcal{T}$ can be deformed along their imaginary axes without crossing $\mathcal{M}$, and thus without changing the value of the Cauchy integral in (8.4). As shown in

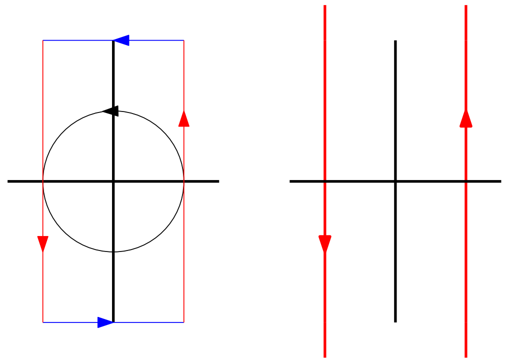


Fig. 8.3 Moving only in the imaginary axis, one can deform a circle around the origin into a rectangle of arbitrarily large height. Because the horizontal lines can be moved arbitrarily far from the x-axis, deforming the domain of integration from a product of circles to the sum of oriented imaginary fibers introduces no error when n is sufficiently large.

Figure 8.3, each of these circles can be deformed into rectangles of fixed width and arbitrarily large height. Because the Cauchy integrand decays exponentially as the moduli of x and y increase, we can replace each of the circles comprising $\mathcal{T}$ by two oriented imaginary fibers.

Ultimately, the product of circles $\mathcal{T}$ becomes four imaginary fibers, so that

$$[x^{nr_1} y^{nr_2}]F(x,y) = \frac{1}{(2\pi i)^2}\left[\int_{C_{\varepsilon,\varepsilon}} \omega - \int_{C_{-\varepsilon,\varepsilon}} \omega - \int_{C_{\varepsilon,-\varepsilon}} \omega + \int_{C_{-\varepsilon,-\varepsilon}} \omega\right], \qquad (8.5)$$

where the sign in front of each integral keeps track of orientation as $\mathbb{R}^2$ keeps its standard orientation in each imaginary fiber.

The general case follows in the same manner. To keep track of the necessary signs we make the following definitions.

Definition 8.6 (signs of points) The *sign* function on $\mathbb{R}$ is defined by

$$\operatorname{sgn}(z) = \begin{cases} -1 & : z < 0 \\ 0 & : z = 0 \\ 1 & : z > 1 \end{cases}.$$

Given $\mathbf{z} \in \mathbb{R}^d$, the *sign* of $\mathbf{z}$ is $\operatorname{sgn}(\mathbf{z}) = \operatorname{sgn}(z_1 \cdots z_d)$.

Proposition 8.1 *Consider the power series expansion of $F(\mathbf{z})$ and let $\mathbf{r} \in \mathbb{R}^d_{>0}$. Whenever n is sufficiently large and $n\mathbf{r}$ has integer coordinates,*

$$[\mathbf{z}^{n\mathbf{r}}]F(\mathbf{z}) = \sum_{\boldsymbol{\tau} \in \{\pm 1\}^d} \frac{\operatorname{sgn}(\boldsymbol{\tau})}{(2\pi i)^d} \int_{\varepsilon\boldsymbol{\tau} + i\mathbb{R}^d} \omega. \qquad (8.6)$$

Proof Because the Cauchy domain of integration $\mathcal{T}$ in (8.1) is a product of d circles, one in each variable, the general argument follows from the one-dimensional case. For any $K > 0$, let

$$\mathcal{C}_K = \{z : \Re(z) \in [-\varepsilon, \varepsilon] \text{ and } \Im(z) = \pm K\}$$

and

$$\mathcal{D}_K = \{z : \Re(z) = \pm\varepsilon \text{ and } \Im(z) \in [-K, K]\}$$

denote the horizontal and vertical sides of the rectangle in Figure 8.3, respectively, oriented according to the arrows in the figure. As we can deform the circle $|z| = \varepsilon$ into $\mathcal{C}_K \cup \mathcal{D}_K$ while staying in $\mathcal{M}$, we have

$$\int_{|z|=\varepsilon} F(z)\frac{dz}{z^{nr+1}} = \int_{\mathcal{C}_K} F(z)\frac{dz}{z^{nr+1}} + \int_{\mathcal{D}_K} F(z)\frac{dz}{z^{nr+1}},$$

and the stated result holds if the integral over $\mathcal{C}_K$ goes to zero as $K \to \infty$. In fact, each point in $\mathcal{C}_K$ has modulus at least K, so

$$\left|\int_{\mathcal{C}_K} F(z)\frac{dz}{z^{nr+1}}\right| \leq \kappa \, \frac{\max_{z\in\mathcal{C}_K} |F(z)|}{K^{nr}} \tag{8.7}$$

for a constant $\kappa > 0$. Since F is a rational function and $\mathcal{C}_K$ is bounded away from the singular set of F for all $K > 0$, when n is sufficiently large the denominator of the upper bound in (8.7) grows faster than its numerator. Thus, the upper bound in (8.7) goes to zero as $K \to \infty$. □

8.2.2 Step 2: Determine the Contributing Singularities

The main benefit to working with integrals over imaginary fibers is that one can make deformations in the real space $\mathcal{M}_{\mathbb{R}}$ instead of the complex space $\mathcal{M}$. On the imaginary fiber $\mathcal{C}_{\mathbf{x}}$ with real basepoint $\mathbf{x} \in \mathcal{M}_{\mathbb{R}}$ the modulus of the Cauchy integrand ω satisfies

$$|\omega| = O\left(|x_1|^{-r_1 n} \cdots |x_d|^{-r_d n}\right) = O\left(e^{h_{\mathbf{r}}(\mathbf{x})}\right)$$

as $n \to \infty$, since $|x_j + \mathrm{i}y_j| \geq |x_j|$ for all $y_j \in \mathbb{R}$. Because $\int_{\mathcal{C}_{\mathbf{x}}} \omega$ depends only on the component $B \subset \mathcal{M}_{\mathbb{R}}$ in which $\mathbf{x}$ lies, each point in B gives an upper bound on the exponential growth of $\int_{\mathcal{C}_{\mathbf{x}}} \omega$.

Remark 8.5 When each coordinate of $\mathbf{r}$ is positive and $\mathbf{x}$ lies in an unbounded component of $\mathcal{M}_{\mathbb{R}}$ then taking $\mathbf{x} \to \infty$ proves that $\int_{\mathcal{C}_{\mathbf{x}}} \omega = 0$ for all sufficiently large n.

The minimizers of this upper bound on $|\omega|$ on each flat are given by critical points.

Lemma 8.1 *Fix a real flat L defined by the vanishing of linearly independent $\ell_{k_1}, \ldots, \ell_{k_s}$, and suppose $\mathbf{r}$ has positive coordinates. Then $\boldsymbol{\sigma} \in \mathbb{R}^d$ is a critical point of F, in the sense of Definition 8.3, if and only if $\boldsymbol{\sigma}$ is a local minimizer of h on L. Consequently, each real flat of $\mathcal{V}$ admits exactly one critical point in every orthant of $\mathbb{R}^d$ where it is bounded, and no critical points in any other orthant.*

Proof The Hessian of $h(\mathbf{z})$ is a diagonal matrix with entries r_j/z_j^2, so h is continuous and strictly convex on the intersection of L with any orthant. Thus, the gradient of h restricted to L is zero at $\mathbf{w} \in L$ if and only if $\mathbf{w}$ is a minimizer of h on L, and there can be at most one such point in each orthant.

As we assume the coefficient vectors $\mathbf{b}^{(k_j)}$ are linearly independent, up to a reordering of variables and the ℓ_{k_j} we may also assume that the matrix M with rows $\mathbf{b}^{(k_1)}, \ldots, \mathbf{b}^{(k_s)}, \mathbf{e}^{(s+1)}, \ldots, \mathbf{e}^{(d)}$ has full rank, where $\mathbf{e}^{(j)}$ is the jth elementary basis vector with a 1 in position j and all other entries 0. If $\mathbf{z}' = (z_{s+1}, \ldots, z_d)$ then L is parametrized as

$$L = \left\{ M^{-1} \begin{pmatrix} \mathbf{1} \\ \mathbf{z}' \end{pmatrix} : \mathbf{z}' \in \mathbb{R}^{d-s} \right\}.$$

To lighten notation we write this parametrization $\mathbf{z} = g(\mathbf{z}')$, so that the derivative of $z_j = g_j(\mathbf{z}')$ with respect to z_a equals $M^{-1}_{j,a}$ for all $s + 1 \leq a \leq d$. Then the height function restricted to L has the parametrization

$$h(\mathbf{z}') = -\sum_{j=1}^{d} r_j \log(g_j(\mathbf{z}')), \tag{8.8}$$

and setting its gradient equal to zero gives the system

$$0 = -\sum_{j=1}^{d} \frac{r_j}{z_j} M^{-1}_{j,a} \qquad (s + 1 \leq a \leq d). \tag{8.9}$$

If $\mathbf{v} = -(r_1/z_1, \ldots, r_d/z_d)$ then (8.9) implies the existence of $\boldsymbol{\lambda} \in \mathbb{R}^{d-s}$ such that $\mathbf{v}M^{-1} = (\boldsymbol{\lambda}, \mathbf{0})$, and therefore $\mathbf{v} = M(\boldsymbol{\lambda}, \mathbf{0})$. The gradient of $h(\mathbf{z}')$ thus vanishes if and only if $\mathbf{v}$ can be written as a linear combination of $\mathbf{b}^{(k_1)}, \ldots, \mathbf{b}^{(k_s)}$, which happens if and only if $\mathbf{z} = g(\mathbf{z}')$ is a critical point in the sense of Definition 8.3. There can be no minimizer of the height function in any orthant of $\mathbb{R}^d$ where L is unbounded, and there must be a minimizer in every orthant where it is bounded as the height function approaches infinity as $\mathbf{z}$ approaches any coordinate axis. □

Example 8.3 Continued (Finding Critical Points))

Each of the four real basepoints $(\pm\varepsilon, \pm\varepsilon)$ of the imaginary fibers $C_{\pm\varepsilon,\pm\varepsilon}$ lie in distinct components of $\mathcal{M}_{\mathbb{R}}$. Because $(\varepsilon, -\varepsilon)$, $(-\varepsilon, \varepsilon)$ and $(-\varepsilon, -\varepsilon)$ are in unbounded components of $\mathcal{M}_{\mathbb{R}}$ (see Figure 8.4 below) the integrals of ω over these domains are zero, and

$$[x^{r_1 n} y^{r_2 n}]F(x,y) = \frac{1}{(2\pi \mathrm{i})^2} \int_{(\varepsilon,\varepsilon)+\mathrm{i}\mathbb{R}^2} \omega. \tag{8.10}$$

The singular set $\mathcal{V}$ can be decomposed into the flats

$$\mathcal{V}_1 = \mathcal{V}\left(1 - \frac{2x+y}{3}\right), \quad \mathcal{V}_2 = \mathcal{V}\left(1 - \frac{3x+y}{4}\right), \quad \mathcal{V}_{1,2} = \mathcal{V}(l_1, \ell_2) = \{(1,1)\},$$

comprising two lines and a point, corresponding to strata

$$\mathcal{S}_1 = \mathcal{V}_1 \setminus \mathcal{V}_{1,2}, \qquad \mathcal{S}_2 = \mathcal{V}_2 \setminus \mathcal{V}_{1,2}, \qquad \mathcal{S}_{1,2} = \mathcal{V}_{1,2}.$$

The critical points on $\mathcal{V}_1$ are those where the matrix

$$N = \begin{pmatrix} -\nabla \ell_1 \\ -\nabla h \end{pmatrix} = \begin{pmatrix} 2/3 & 1/3 \\ r_1/x & r_2/y \end{pmatrix}$$

is rank deficient. Solving $\ell_1 = \det N = 0$ gives the critical point

$$\boldsymbol{\sigma}_1 = \left(\frac{3\hat{r}_1}{2}, 3\hat{r}_2\right),$$

where $\hat{\mathbf{r}} = \mathbf{r}/(r_1 + r_2)$ is the vector with unit coordinate sum in the direction defined by $\mathbf{r}$. Similarly, $\mathcal{V}_2$ admits only the critical point

$$\boldsymbol{\sigma}_2 = \left(\frac{4\hat{r}_1}{3}, 4\hat{r}_2\right).$$

The flat $\mathcal{V}_{1,2}$ consists of the single point $\boldsymbol{\sigma}_{12} = (1,1)$, which is trivially a critical point. The direction $\mathbf{r}$ is non-generic precisely when one of $\boldsymbol{\sigma}_1$ or $\boldsymbol{\sigma}_2$ equals $(1,1)$, which happens if and only if $\mathbf{r}$ is a multiple of $(2,1)$ or $(3,1)$.

Remark 8.6 If $\boldsymbol{\sigma}$ is a critical point on the flat defined by linearly independent $\ell_{k_1}, \ldots, \ell_{k_s}$ then there exist unique real numbers $a_1, \ldots, a_s$ such that

$$-(\nabla h)(\boldsymbol{\sigma}) = \sum_{j=1}^{s} a_j(-\nabla \ell_{k_j}) = \sum_{j=1}^{s} a_j \mathbf{b}^{(k_j)}. \tag{8.11}$$

Existence of a critical point $\boldsymbol{\sigma}$ such that some coefficient a_j in (8.11) is zero is equivalent to the direction $\mathbf{r}$ being non-generic.

Lemma 8.1 characterizes the minimizers of the height function on each *flat*, but we really want the minimizers of the height function on the closure of each (bounded) *component* of $\mathcal{M}_{\mathbb{R}}$. In fact, pairing each bounded component of $\mathcal{M}_{\mathbb{R}}$ with the critical point that minimizes the height function on its closure gives a bijection.

Proposition 8.2 *Suppose* $\mathbf{r}$ *is a generic direction with positive coordinates and* F *is simple. For bounded component* $B \in \mathcal{B}$ *let* $\boldsymbol{\sigma}_B$ *be the point minimizing* h *on* $\overline{B}$*. Then* *the map $B \mapsto \boldsymbol{\sigma}_B$ gives a bijection between the set of critical points Ω and the set of bounded components in $\mathcal{B}$.*

Proof The minimum $\boldsymbol{\sigma}_B$ of the strictly convex function h on any bounded $B \in \mathcal{B}$ occurs on some flat defining part of the boundary of B, so Lemma 8.1 implies $\boldsymbol{\sigma}_B$ is always a critical point. Now, suppose $\boldsymbol{\sigma}$ is the critical point on a flat $\mathcal{V}_{k_1,\ldots,k_s}$ in some fixed orthant, and write

$$-(\nabla h)(\boldsymbol{\sigma}) = \sum_{j=1}^{s} a_j \mathbf{b}^{(k_j)},$$

where each $a_j \in \mathbb{R}$ is non-zero since $\mathbf{r}$ is generic. Because the $\mathbf{b}^{(k_j)}$ are linearly independent, there exists a unique component $B \subset \mathcal{M}_{\mathbb{R}}$ with $\boldsymbol{\sigma} \in \overline{B}$ such that B has non-empty intersection with the set

$$A = \{\mathbf{z} : a_j \ell_j(\mathbf{z}) > 0 \text{ for all } j\}.$$

In fact the polyhedron B is contained in A, so for any $\mathbf{z} \in \overline{B}$

$$(\nabla h)(\boldsymbol{\sigma}) \cdot (\mathbf{z} - \boldsymbol{\sigma}) = \sum_{j=1}^{s} a_j \left(1 - \mathbf{z} \cdot \mathbf{b}^{(k_j)}\right) = \sum_{j=1}^{s} a_j \ell_{k_j}(\mathbf{z}) \geq 0.$$

In particular, h increases as it moves from $\boldsymbol{\sigma}$ towards any arbitrarily close $\mathbf{z} \in \overline{B}$, so $\boldsymbol{\sigma}$ is a local (and thus global) minimizer of the strictly convex function h on $\overline{B}$. The fact that h has a minimum on B also implies B is bounded.

If $C \neq B$ is another component of $\mathcal{M}_{\mathbb{R}}$ with $\boldsymbol{\sigma} \in \overline{C}$ then there exists some $1 \leq j \leq s$ such that $a_j \ell_{k_j}(\mathbf{z}) < 0$ for $\mathbf{z} \in C$. Let $\mathbf{z} \in \overline{C}$ be any vector such that $\ell_{k_i}(\mathbf{z}) = 0$ for $i \neq j$ and $\ell_{k_j}(\mathbf{z}) < 0$. Then

$$(\nabla h)(\boldsymbol{\sigma}) \cdot (\mathbf{z} - \boldsymbol{\sigma}) = a_j (1 - \mathbf{b}^{(k_j)} \cdot \mathbf{z}) < 0,$$

so, for sufficiently small $\varepsilon > 0$, the vector $\mathbf{v} = \boldsymbol{\sigma} + \varepsilon(\mathbf{z} - \boldsymbol{\sigma})$ lies in $\overline{C}$ and satisfies $h(\mathbf{v}) < h(\boldsymbol{\sigma})$. In other words, $\boldsymbol{\sigma}$ is not a minimizer of h on $\overline{C}$.

Thus, we have shown that each minimizer of h on the closure of a bounded component of $\mathcal{M}_{\mathbb{R}}$ is a critical point, and that each critical point is the minimizer on the closure of a unique bounded component. □

Definition 8.7 The *height* of a bounded component $B \in \mathcal{B}$ is the height of the critical point $\boldsymbol{\sigma}_B$, which is the minimum of h on $\overline{B}$. Given critical point $\boldsymbol{\sigma} \in \Omega$, we write $B(\boldsymbol{\sigma})$ for the bounded component in $\mathcal{B}$ on which the minimum of h occurs at $\boldsymbol{\sigma}$.

Since Proposition 8.1 expresses the series coefficients of interest as a sum of integrals over imaginary fibers near the origin, the critical points which are minimizers for bounded components 'closer' to the origin play a special role.

Definition 8.8 (contributing singularities) Given $\boldsymbol{\sigma}$ on the real flat $\mathcal{V}_{k_1,\ldots,k_s}$ defined by $\ell_{k_1}, \ldots, \ell_{k_s}$, we define the real cone

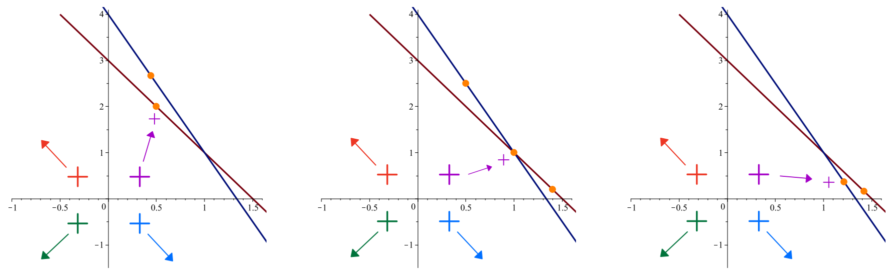


Fig. 8.4 Positions of the contributing points (dots) in direction $\mathbf{r} = (r_1, r_2)$ when $\hat{r}_1 < 2/3$ (left), $2/3 < \hat{r}_1 < 3/4$ (centre), and $3/4 < \hat{r}_1$ (right). If B_0 is the bounded component of $\mathcal{M}_{\mathbb{R}}$ containing the origin then we move the imaginary fiber with basepoint near the origin in B_0 to an imaginary fiber whose basepoint is close to the contributing singularity on ∂B_0. The other imaginary fibers can be pushed off to infinity without crossing $\mathcal{V}$ or the coordinate axes.

$$N(\boldsymbol{\sigma}) = \left\{ \sum_{j=1}^{s} a_j \mathbf{b}^{(k_j)} : a_j > 0 \right\} \subset \mathbb{R}^d.$$

The point $\boldsymbol{\sigma} \in \mathbb{R}_*^d$ is called a *contributing singularity*, or *contributing point*, when $-(\nabla h)(\boldsymbol{\sigma}) \in N(\boldsymbol{\sigma})$.

Remark 8.7 Recall the logarithmic gradient $\nabla_{\log} f = (z_1 f_{z_1}, \ldots, z_d f_{z_d})$ from previous chapters. Then $-(\nabla h)(\boldsymbol{\sigma}) \in N(\boldsymbol{\sigma})$ is equivalent to $\mathbf{r}$ being in the positive real span of the logarithmic gradients $\nabla_{\log} \ell_{k_j}$.

Since $\ell_j(\mathbf{0}) > 0$ for each $1 \le j \le m$, our proof of Proposition 8.2 immediately gives the following.

Corollary 8.1 *Let $\boldsymbol{\sigma} \in \mathbb{R}_*^d$ lie on the real flat defined by linearly independent functions $\ell_{k_1}, \ldots, \ell_{k_s}$, and suppose $\boldsymbol{\sigma}$ is contained in an orthant O of $\mathbb{R}^d$. If $\mathbf{r}$ is a nongeneric direction with positive coordinates then $\boldsymbol{\sigma}$ is a contributing singularity if and only if $\boldsymbol{\sigma}$ is the unique minimizer of $h(\mathbf{z})$ on the bounded polyhedron*

$$\left\{ \mathbf{z} \in \mathbb{R}^d : \ell_{k_j}(\mathbf{z}) \ge 0 \textit{ for } j = 1, \ldots, s \right\} \cap O$$

defined by the $\ell_{k_j}(\mathbf{z})$ and the coordinate axes.

Remark 8.8 If $\boldsymbol{\sigma}$ is a minimal point (on the boundary of the domain of convergence $\mathcal{D}$) lying in a unique hyperplane in $\mathcal{V}$ then Definition 8.8 matches Definition 5.5 of Chapter 5 for a minimal smooth contributing point.

Example 8.3 Continued (Finding Contributing Points)

Let B_0 be the bounded component of $\mathcal{M}_{\mathbb{R}}$ containing the origin. The cones $N(\boldsymbol{\sigma}_1)$ and $N(\boldsymbol{\sigma}_2)$ are rays which contain $-(\nabla h)(\boldsymbol{\sigma}_1)$ and $-(\nabla h)(\boldsymbol{\sigma}_2)$, respectively, meaning the critical points $\boldsymbol{\sigma}_1$ and $\boldsymbol{\sigma}_2$ are always contributing singularities. Furthermore,

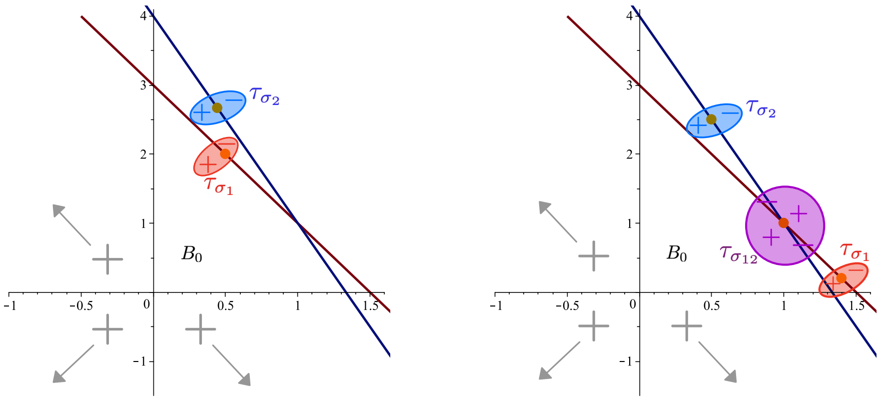


Fig. 8.5 The signed fibers (+ and −) around each contributing singularity for $\hat{r}_1 \in (0, 2/3)$, left, and $\hat{r}_1 \in (2/3, 3/4)$, right. The situation when $\hat{r}_1 \in (3/4, 1)$ is analogous to that of $\hat{r}_1 \in (0, 2/3)$. Note that the signs in all bounded components not containing the origin cancel. Each $\boldsymbol{\tau}_{\boldsymbol{\sigma}_j}$ is a linking torus, defined later in this section.

$$N(\boldsymbol{\sigma}_{12}) = \left\{ \left(\frac{8a + 9b}{12}, \frac{4a + 3b}{12} \right) : a, b > 0 \right\}$$

so that $(-\nabla h)(\boldsymbol{\sigma}_{12}) \in N(\boldsymbol{\sigma}_{12})$ if and only if $\hat{r}_1 \in (2/3, 3/4)$. In other words, $\boldsymbol{\sigma}_{12}$ is contributing if and only if $\mathbf{r}$ is a generic direction and $\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2 \notin \overline{B_0}$. Figure 8.4 shows the contributing points in different situations.

8.2.3 Step 3: Express the Cauchy Integral as Sum of Local Contributing Integrals

We have now split the Cauchy integral describing the coefficients of interest into a finite sum of integrals over imaginary fibers, and identified contributing singularities where local behaviour of F will dictate coefficient asymptotics. We next show that the expression in Proposition 8.1 for a sum of integrals over fibers near the origin can be replaced by an explicit sum of integrals over fibers near each contributing point. This is the most difficult step of our argument, and the hardest step to generalize to arbitrary rational functions.

Example 8.3 Continued (Preparing for Residue Computations)

There are three different situations, depending on whether $\hat{r}_1$ lies in $(0, 2/3)$, $(2/3, 3/4)$, or $(3/4, 1)$. In the first and third cases there are two contributing points, both of which are smooth, while in the second case there are three contributing points. The different situations are illustrated in Figure 8.5.

As in Figure 8.2 above, let B_0 be the bounded component of $\mathcal{M}_{\mathbb{R}}$ containing the origin, and B_1 and B_2 be the other bounded components of $\mathcal{M}_{\mathbb{R}}$ in the first quadrant. Suppose, for instance, that $\hat{r}_1 \in (0, 2/3)$ so $\boldsymbol{\sigma}_1$ and $\boldsymbol{\sigma}_2$ are the only contributing points, with $\boldsymbol{\sigma}_1$ on the boundary of B_0. One can imagine moving the basepoint of the Cauchy integral in (8.10) to be arbitrarily close to $\boldsymbol{\sigma}_1$ while staying in B_0,

$$[x^{r_1 n} y^{r_2 n}]F(x, y) = \frac{1}{(2\pi \mathrm{i})^2} \int_{(\varepsilon,\varepsilon)+\mathrm{i}\mathbb{R}^2} \omega = \frac{1}{(2\pi \mathrm{i})^2} \int_{\boldsymbol{\sigma}_1-(\varepsilon,\varepsilon)+\mathrm{i}\mathbb{R}^2} \omega.$$

In order to reduce to a residue computation, we add and subtract an integral over a signed fiber with basepoint arbitrarily close to $\boldsymbol{\sigma}_1$ in B_1,

$$[x^{r_1 n} y^{r_2 n}]F(x, y) = \frac{1}{(2\pi \mathrm{i})^2} \left[\int_{\boldsymbol{\sigma}_1-(\varepsilon,\varepsilon)+\mathrm{i}\mathbb{R}^2} \omega - \int_{\boldsymbol{\sigma}_1+(\varepsilon,\varepsilon)+\mathrm{i}\mathbb{R}^2} \omega\right] + \frac{1}{(2\pi \mathrm{i})^2} \int_{\boldsymbol{\sigma}_1+(\varepsilon,\varepsilon)+\mathrm{i}\mathbb{R}^2} \omega.$$

We now have an extra integral with basepoint in B_1, but this can be moved arbitrarily close to the contributing point $\boldsymbol{\sigma}_2$, after which we add and subtract an integral over a signed fiber with basepoint in an unbounded component of $\mathcal{M}_{\mathbb{R}}$,

$$\begin{aligned}
[x^{r_1 n} y^{r_2 n}]F(x, y) = {} & \frac{1}{(2\pi \mathrm{i})^2} \left[\int_{\boldsymbol{\sigma}_1-(\varepsilon,\varepsilon)+\mathrm{i}\mathbb{R}^2} \omega - \int_{\boldsymbol{\sigma}_1+(\varepsilon,\varepsilon)+\mathrm{i}\mathbb{R}^2} \omega\right] \\
& + \frac{1}{(2\pi \mathrm{i})^2} \left[\int_{\boldsymbol{\sigma}_2-(\varepsilon,\varepsilon)+\mathrm{i}\mathbb{R}^2} \omega - \int_{\boldsymbol{\sigma}_2+(\varepsilon,\varepsilon)+\mathrm{i}\mathbb{R}^2} \omega\right] \\
& + \frac{1}{(2\pi \mathrm{i})^2} \int_{\boldsymbol{\sigma}_2+(\varepsilon,\varepsilon)+\mathrm{i}\mathbb{R}^2} \omega.
\end{aligned}$$

As $\boldsymbol{\sigma}_2 + (\varepsilon, \varepsilon)$ lies in an unbounded component of $\mathcal{M}_{\mathbb{R}}$, the final integral here is zero. We show below that each of the differences of integrals can be reduced to a one-dimensional integral by a residue computation.

When $\hat{r}_1 \in (3/4, 1)$ the analysis the same after replacing $\boldsymbol{\sigma}_1$ by $\boldsymbol{\sigma}_2$ and B_1 by B_2. When $\hat{r}_1 \in (2/3, 3/4)$ the analysis is similar except the integral with basepoint near the origin is moved to the contributing singularity $\boldsymbol{\sigma}_{12}$ on the intersection $\mathcal{V}_1 \cap \mathcal{V}_2$. Now signed integrals over fibers with basepoints in three components of $\mathcal{M}_{\mathbb{R}}$ are added: one has a basepoint in the unbounded component of $\mathcal{M}_{\mathbb{R}}$ adjacent to $\boldsymbol{\sigma}_{12}$ while the other two have basepoints in the bounded components B_1 and B_2. Adding signed integrals over fibers with basepoints near the smooth contributing points cancels out the integrals whose fibers intersect the bounded components B_1 and B_2.

Recall also that the Cauchy integrals appearing in Proposition 8.1 are signed depending on the orthant of their basepoint. As all contributing singularities were in the first quadrant here, the leading sign was positive.

In order to properly talk about adding and subtracting imaginary fibers, without worrying about fibers with basepoints in unbounded components of $\mathcal{M}_{\mathbb{R}}$, we introduce the notion of a chain of integration.

Definition 8.9 (chains of integration) A *chain of integration* in $\mathcal{M}$ is a formal integer linear combination of imaginary fibers $a_1\mathcal{C}_{\mathbf{x}_1} + \cdots + a_k\mathcal{C}_{\mathbf{x}_k}$, where each $\mathbf{x}_j \in \mathcal{M}_{\mathbb{R}}$ and $a_j \in \mathbb{Z}$. Integration over such a chain is defined by linearity,

$$\int_{a_1\mathcal{C}_{\mathbf{x}_1}+\cdots+a_k\mathcal{C}_{\mathbf{x}_k}} f(\mathbf{z})d\mathbf{z} = a_1\int_{\mathcal{C}_{\mathbf{x}_1}} f(\mathbf{z})d\mathbf{z} + \cdots + a_k\int_{\mathcal{C}_{\mathbf{x}_k}} f(\mathbf{z})d\mathbf{z}$$

for any function f where each integral is defined; by definition any integral over the empty chain 0 equals zero. We write

$$a_1\mathcal{C}_{\mathbf{x}_1} + \cdots + a_k\mathcal{C}_{\mathbf{x}_k} \doteq 0 \tag{8.12}$$

if and only if for every bounded component $B \subset \mathcal{M}_{\mathbb{R}}$ the integer coefficients a_j of all imaginary chains in (8.12) with basepoints in B sum to zero (i.e., after identifying all basepoints in the same components of $\mathcal{M}_{\mathbb{R}}$, and simplifying the linear combination, the only remaining fibers have basepoints in unbounded regions of $\mathcal{M}_{\mathbb{R}}$). Then $\doteq$ gives an equivalence relation on chains by setting $\kappa_1 \doteq \kappa_2$ if and only if $\kappa_1 - \kappa_2 \doteq 0$.

Because $\int_{\mathcal{C}_{\mathbf{x}}} \omega = 0$ whenever $\mathbf{x}$ lies in an unbounded component of $\mathcal{M}_{\mathbb{R}}$, the integral $\int_{\kappa} \omega$ depends only on the equivalence class $[\kappa]$ of a chain κ under $\doteq$.

Definition 8.10 (linking tori) Given component $B \subset \mathcal{M}_{\mathbb{R}}$, we define the class $[\mathcal{C}_B]$ of $\mathcal{C}_B$ under $\doteq$ as $[\mathcal{C}_{\mathbf{x}}]$ for any $\mathbf{x} \in B$. If B is an unbounded component then $[\mathcal{C}_B] = [0]$. For any $\mathbf{x} \in \mathcal{V}_{\mathbb{R}}$ the set of *adjacent components* to $\mathbf{x}$ is the collection

$$\mathrm{Adj}(\mathbf{x}) = \left\{B \in \mathcal{B} : \mathbf{x} \in \overline{B}\right\}$$

of components of $\mathcal{M}_{\mathbb{R}}$ with $\mathbf{x}$ in their closure. Given $\boldsymbol{\sigma}$ in the strata $\mathcal{S}_{k_1,\ldots,k_s}$ and $B \in \mathrm{Adj}(\boldsymbol{\sigma})$, the *sign of B with respect to* $\boldsymbol{\sigma}$ is

$$\mathrm{sgn}_{\boldsymbol{\sigma}}(B) = \mathrm{sgn}\left(\ell_{k_1}(\mathbf{x})\cdots\ell_{k_s}(\mathbf{x})\right)$$

for any $\mathbf{x} \in B$. A *linking torus* around a critical point $\boldsymbol{\sigma} \in \Omega$ is any chain of integration $\boldsymbol{\tau}_{\boldsymbol{\sigma}}$ such that

$$\boldsymbol{\tau}_{\boldsymbol{\sigma}} \doteq \sum_{B\in\mathrm{Adj}(\boldsymbol{\sigma})} \mathrm{sgn}_{\boldsymbol{\sigma}}(B)\mathcal{C}_B,$$

well defined up to a selection of basepoints in each bounded component of $\mathrm{Adj}(\boldsymbol{\sigma})$ and up to adding imaginary fibers with basepoints in unbounded components of $\mathcal{M}_{\mathbb{R}}$.

Remark 8.9 This definition is highly tailored to our needs. More generally, one would work in the framework of relative homology on the space $\mathcal{M} \subset \mathbb{C}^d$ with respect to the height function h (see, for instance, Baryshnikov et al. [1]).

We will prove that the Cauchy integral describing the coefficients of interest can be written as the sum of integrals over the linking tori for each contributing point

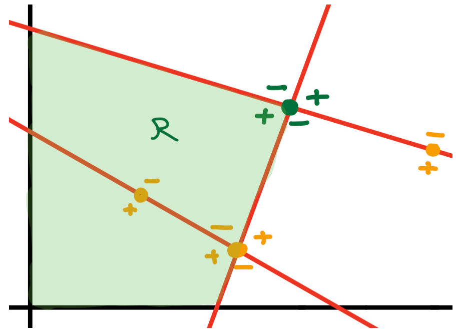


Fig. 8.6 Proving linear independence of the linking tori for Lemma 8.2, with the critical points and signs of imaginary fibers for each linking torus displayed. The hyperplanes containing $\boldsymbol{\sigma}_j$ (the dark point) form the shaded bounded region $\mathcal{R}$ in the first quadrant. The signs of the imaginary fibers around each critical point $\boldsymbol{\sigma} \neq \boldsymbol{\sigma}_j$ in $\overline{\mathcal{R}}$ whose basepoints lie in $\mathcal{R}$ sum to zero.

of F. First, we show that the chain defined by any imaginary fiber is a unique integer sum of linking tori.

Lemma 8.2 *Suppose F is simple and $\mathbf{r}$ is generic. For any component $B \in \mathcal{B}$ there exist unique integers $\nu_{B,\boldsymbol{\sigma}}$ such that*

$$C_B \doteq \sum_{\boldsymbol{\sigma}\in\Omega} \nu_{B,\boldsymbol{\sigma}}\, \boldsymbol{\tau}_{\boldsymbol{\sigma}}. \tag{8.13}$$

Furthermore, $\nu_{B,\boldsymbol{\sigma}} = 0$ unless B is bounded, B and $\boldsymbol{\sigma}$ lie in the same orthant of $\mathbb{R}^d$, and $h(\boldsymbol{\sigma}_B) \leq h(\boldsymbol{\sigma})$.

Proof If B is unbounded the result holds with all $\nu_{B,\boldsymbol{\sigma}} = 0$. Otherwise, let $\boldsymbol{\sigma} = \boldsymbol{\sigma}_B$ be the critical point minimizing h on $\overline{B}$. Since

$$\boldsymbol{\tau}_{\boldsymbol{\sigma}} \doteq \sum_{B'\in\mathrm{Adj}(\boldsymbol{\sigma})} \mathrm{sgn}_{\boldsymbol{\sigma}}(B')C_{B'},$$

it follows that

$$\mathrm{sgn}_{\boldsymbol{\sigma}}(B)C_B \doteq \boldsymbol{\tau}_{\boldsymbol{\sigma}} - \sum_{B'\in\mathrm{Adj}(\boldsymbol{\sigma})\setminus\{B\}} \mathrm{sgn}_{\boldsymbol{\sigma}}(B')C_{B'}. \tag{8.14}$$

For any component $B' \in \mathrm{Adj}(\boldsymbol{\sigma}) \setminus \{B\}$ the minimizer of h on B' is not $\boldsymbol{\sigma}$, but $\boldsymbol{\sigma}$ lies on $\overline{B'}$. Thus, either there exists $\boldsymbol{\sigma}' \in \Omega \cap \overline{B'}$ with $h(\boldsymbol{\sigma}') < h(\boldsymbol{\sigma})$ or B' is unbounded. In other words, up to sign, the fiber C_B equals $\boldsymbol{\tau}_{\boldsymbol{\sigma}}$ plus an integer linear combination of imaginary fibers defined by components B' of smaller height. Continually iterating the decomposition of (8.14) on each $C_{B'}$ thus eventually leaves only imaginary fibers with basepoints in unbounded components of $\mathcal{M}_{\mathbb{R}}$, all of which lie in the same orthant of $\mathbb{R}^d$.

Uniqueness of the $\nu_{B,\boldsymbol{\sigma}}$ follows from linear independence of the $\boldsymbol{\tau}_{\boldsymbol{\sigma}}$, which we now prove. Suppose there exist $\boldsymbol{\sigma}_1, \ldots, \boldsymbol{\sigma}_c \in \Omega$ and $a_1, \ldots, a_c \in \mathbb{Z}$ such that

$$a_1 \boldsymbol{\tau}_{\boldsymbol{\sigma}_1} + \cdots + a_c \boldsymbol{\tau}_{\boldsymbol{\sigma}_c} \doteq 0. \tag{8.15}$$

Replacing each torus in (8.15) by a signed sum of imaginary fibers gives

$$\sum_{B \in \mathcal{B}} c_B C_B \doteq 0$$

for integers c_B which can be written in terms of the a_j. By definition, (8.15) holds if and only if $c_B = 0$ for all bounded B.

Given a critical point $\boldsymbol{\sigma}$ the *codimension* of $\boldsymbol{\sigma}$, denoted codim$(\boldsymbol{\sigma})$, is the codimension of the stratum $S(\boldsymbol{\sigma})$ which contains it (so a critical point which lies on the intersection of k hyperplanes and no proper subflat has codimension k). We will go iteratively through the critical points $\boldsymbol{\sigma}_j$ to prove that each a_j in (8.15) equals zero. To that end, let $\boldsymbol{\sigma}_j$ be any of the critical points of minimal codimension among those for which we have not yet shown that $a_j = 0$. Let O denote the orthant containing $\boldsymbol{\sigma}_j$ with the hyperplanes $\mathcal{V}_k$ containing $\boldsymbol{\sigma}_j$ removed. Since $\boldsymbol{\sigma}_j$ is a critical point, the orthant O contains a bounded connected component $\mathcal{R} \subset \mathbb{R}^d$ and, since each $c_B = 0$,

$$\sum_{\substack{B \in \mathcal{B} \\ B \subset \mathcal{R}}} c_B = 0; \tag{8.16}$$

see Figure 8.6. Consider (8.16) as an expression in the coefficients a_k from (8.15). If we know that some $a_k = 0$ then a_k does not affect (8.16). Otherwise,

- If $\boldsymbol{\sigma}_k \neq \boldsymbol{\sigma}_j$ lies outside of $\overline{\mathcal{R}}$ then a_k does not appear in (8.16).
- If $\boldsymbol{\sigma}_k \neq \boldsymbol{\sigma}_j$ lies inside of $\mathcal{R}$ then the contribution of a_k to (8.16) is zero, since $\boldsymbol{\tau}_{\boldsymbol{\sigma}_k}$ is the signed sum of $2^{\text{codim}(\boldsymbol{\sigma}_k)}$ imaginary fibers, each of which lies in $\mathcal{R}$ with an equal number of positive and negative signs appearing.
- If $\boldsymbol{\sigma}_k \neq \boldsymbol{\sigma}_j$ lies on the boundary of $\mathcal{R}$ then suppose the smallest flat containing both $\boldsymbol{\sigma}_k$ and $\boldsymbol{\sigma}_j$ has codimension N. Note $N \leq \text{codim}(\boldsymbol{\sigma}_j) \leq \text{codim}(\boldsymbol{\sigma}_k)$ by our selection of $\boldsymbol{\sigma}_j$, so $N < \text{codim}(\boldsymbol{\sigma}_k)$ or else $\boldsymbol{\sigma}_j$ and $\boldsymbol{\sigma}_k$ would be distinct critical points on a single stratum, contradicting Lemma 8.1. Again the contribution of a_k to (8.16) is zero, since $\boldsymbol{\tau}_{\boldsymbol{\sigma}_k}$ is the sum of $2^{\text{codim}(\boldsymbol{\sigma}_k)}$ signed imaginary fibers, of which $2^{\text{codim}(\boldsymbol{\sigma}_k)-N}$ have basepoints in $\mathcal{R}$, and an equal number of positive and negative signs appear among those basepoints in $\mathcal{R}$.

Thus, each torus $\boldsymbol{\tau}_{\boldsymbol{\sigma}_k}$ for $k \neq j$ is a signed sum of imaginary fibers whose contributions to (8.16) cancel out, and

$$0 = \sum_{\substack{B \in \mathcal{B} \\ B \subset \mathcal{R}}} c_B = \pm a_j.$$

Repeating this argument eventually implies that all $a_j = 0$. □

Definition 8.11 (linking constants) For any $B \in \mathcal{B}$ and $\boldsymbol{\sigma} \in \Omega$ the constant $\nu_{B,\boldsymbol{\sigma}}$ in (8.13) is called the *linking constant* of B and $\boldsymbol{\sigma}$.

Applying Lemma 8.2 to Proposition 8.1 proves that the series coefficients of interest can be written as an integer linear combination of integrals over linking tori. In Chapter 9 we will see how Stratified Morse Theory provides the foundation for a similar result when $H(\mathbf{z})$ is not necessarily the product of linear functions. What is special about having only linear factors in the denominator, in addition to our simplified arguments, is that we can identify the integer linking constants.

The key step is the following lemma, which allows us to connect how imaginary fibers in adjacent components of $\mathcal{M}_{\mathbb{R}}$ are represented in terms of linking tori. Since Proposition 8.1 starts with a sum of imaginary fibers near the origin, this will allow us to move through the bounded components of $\mathcal{M}_{\mathbb{R}}$ and determine which linking tori appear.

Lemma 8.3 *Suppose F is simple and let B and B' be two components of $\mathcal{M}_{\mathbb{R}}$ in the same orthant O of $\mathbb{R}^d$ which are separated by a unique hyperplane. If $\boldsymbol{\sigma}_* \in \Omega$ does not lie on this separating hyperplane then $\nu_{B,\boldsymbol{\sigma}_*} = \nu_{B',\boldsymbol{\sigma}_*}$.*

If B is an unbounded component of $\mathcal{M}_{\mathbb{R}}$ then there is no well-defined critical point $\boldsymbol{\sigma}_B$, but to unify notation we write $h(\boldsymbol{\sigma}_B) = -\infty$.

Proof Under our hypotheses there exist $j \in \{1, \dots, m\}$ and $\varepsilon_i \in \{\pm 1\}$ for all $1 \le i \le m$ such that

$$\begin{aligned} B &= \{\mathbf{x} \in O : \varepsilon_i \ell_i(\mathbf{x}) > 0 \text{ for all } 1 \le i \le m\} \\ B' &= \{\mathbf{x} \in O : \varepsilon_i \ell_i(\mathbf{x}) > 0 \text{ for all } i \ne j, \text{ and } \varepsilon_j \ell_j(\mathbf{x}) < 0\}. \end{aligned}$$

If both B and B' are unbounded then the desired result holds as both integers are zero. If B and B' are bounded then one of the critical points $\boldsymbol{\sigma}_B$ and $\boldsymbol{\sigma}_{B'}$ must lie on the hyperplane $\mathcal{V}_j$: if not then the line through $\boldsymbol{\sigma}_B$ and $\boldsymbol{\sigma}_{B'}$ intersects $\mathcal{V}_j$ at some point in the closure of B and B' which, by strict convexity of h, has smaller height than either $\boldsymbol{\sigma}_B$ or $\boldsymbol{\sigma}_{B'}$ and contradicts their minimality. If B is bounded but B' is unbounded then the same argument with $\boldsymbol{\sigma}_{B'}$ replaced with any point of B' with arbitrarily small height shows $\boldsymbol{\sigma}_B \in \mathcal{V}_j$. Thus, without loss of generality we may assume that B is bounded and $\ell_j(\boldsymbol{\sigma}_B) = 0$. Because B and B' are separated only by $\mathcal{V}_j$ we have $\operatorname{sgn}_{\boldsymbol{\sigma}_B}(B) = -\operatorname{sgn}_{\boldsymbol{\sigma}_B}(B')$ and manipulating the sum for $\boldsymbol{\tau}_{\boldsymbol{\sigma}_B}$ yields

$$\operatorname{sgn}_{\boldsymbol{\sigma}_B}(B)\,[C_B - C_{B'}] \doteq \boldsymbol{\tau}_{\boldsymbol{\sigma}_B} - \sum_{B'' \in \mathbf{A} \setminus \{B, B'\}} \operatorname{sgn}_{\boldsymbol{\sigma}_B}(B'') C_{B''}, \tag{8.17}$$

where $\mathbf{A} = \operatorname{Adj}(\boldsymbol{\sigma}_B)$ contains the components of $\mathcal{M}_{\mathbb{R}}$ with $\boldsymbol{\sigma}_B$ in their closure.

We prove that the coefficient of $\boldsymbol{\tau}_{\boldsymbol{\sigma}_*}$ appearing in the right-hand side of (8.17) is zero for any $\boldsymbol{\sigma}_* \in \Omega \setminus \mathcal{V}_j$ by induction on the maximum of $h(\boldsymbol{\sigma}_B)$ and $h(\boldsymbol{\sigma}_{B'})$. The base case occurs when B and B' are unbounded and both heights are negative infinity. Otherwise, all elements of $\mathbf{A} \setminus \{B, B'\}$ have height less than $h(\boldsymbol{\sigma}_B)$ and can be grouped into pairs of components separated only by the hyperplane $\mathcal{V}_j$. The imaginary fibers

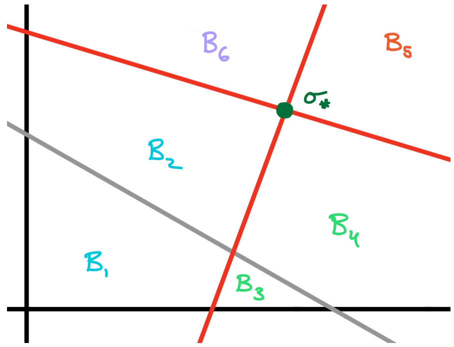


Fig. 8.7 The hyperplanes through $\boldsymbol{\sigma}$ divide the positive orthant into four regions. The integers $\nu_{B,\boldsymbol{\sigma}}$ take a common value when B differs between adjacent components of $\mathcal{M}_{\mathbb{R}}$ in one of these regions. For instance, in this example $\nu_{B_1,\boldsymbol{\sigma}} = \nu_{B_2,\boldsymbol{\sigma}}$ and $\nu_{B_3,\boldsymbol{\sigma}} = \nu_{B_4,\boldsymbol{\sigma}}$.

corresponding to any such pair of components (B_1, B_2) appear with opposite sign on the right-hand side of (8.17), and by the induction hypothesis $\nu_{B_1,\boldsymbol{\sigma}_*} = \nu_{B_2,\boldsymbol{\sigma}_*}$. Thus, the coefficient of $\boldsymbol{\tau}_{\boldsymbol{\sigma}_*}$ on the right-hand side of (8.17) is zero, which means that $\nu_{B,\boldsymbol{\sigma}_*} = \nu_{B',\boldsymbol{\sigma}_*}$. □

Lemma 8.3 immediately gives us the diagonal sequence under consideration as an explicit sum of integrals over linking tori, the most important theoretical result of this chapter.

Theorem 8.1 *Consider the power series expansion of simple $F(\mathbf{z})$ and let $\mathbf{r}$ be a generic direction with positive coordinates. If χ denotes the set of contributing singularities of $F(\mathbf{z})$ then*

$$[\mathbf{z}^{n\mathbf{r}}]F(\mathbf{z}) = \sum_{\boldsymbol{\sigma}\in\chi} \frac{\operatorname{sgn}(\boldsymbol{\sigma})}{(2\pi i)^d} \int_{\tau_{\boldsymbol{\sigma}}} \frac{G(\mathbf{z})}{\mathbf{z}^{n\mathbf{r}+\mathbf{1}} \prod_{j=1}^{m} \ell_j(\mathbf{z})^{p_j}}\, d\mathbf{z}, \tag{8.18}$$

where $\tau_{\boldsymbol{\sigma}}$ is the linking torus from Definition 8.10.

Proof Let O be an orthant of $\mathbb{R}^d$ and B_0 be the component of $\mathcal{M}_{\mathbb{R}} \cap O$ whose closure contains the origin. By Proposition 8.1 it is sufficient to show that

$$C_{B_0} \doteq \sum_{\boldsymbol{\sigma}\in\chi\cap O} \boldsymbol{\tau}_{\boldsymbol{\sigma}}.$$

Lemma 8.2 implies $\nu_{B_0,\boldsymbol{\sigma}} = 0$ unless $\boldsymbol{\sigma} \in O$. Thus, we need to prove that for any critical point $\boldsymbol{\sigma} \in \Omega \cap O$

$$\nu_{B_0,\boldsymbol{\sigma}} = \begin{cases} 1 & : \boldsymbol{\sigma} \text{ is a contributing singularity} \\ 0 & : \text{otherwise} \end{cases}.$$

Suppose $\boldsymbol{\sigma}$ lies on the flat $\mathcal{V}_{k_1,\ldots,k_s}$ and no proper subflat. Then the union of the hyperplanes $\mathcal{V}_{k_1}, \ldots, \mathcal{V}_{k_s}$ divides O into 2^s connected regions,

$$\mathcal{O} \setminus \left(\mathcal{V}_{k_1} \cup \cdots \cup \mathcal{V}_{k_s}\right) = \mathcal{O}_1 \cup \cdots \cup \mathcal{O}_{2^s},$$

and each component B' of $\mathcal{M}_{\mathbb{R}} \cap \mathcal{O}$ lies in one of the $\mathcal{O}_j$; see Figure 8.7 for an example. Lemma 8.3 states that the linking constant $\nu_{B',\boldsymbol{\sigma}}$ depends only on the region $\mathcal{O}_j$ in which B' is contained, and each of these regions contain a component adjacent to $\boldsymbol{\sigma}$. On the region containing the component $B = B(\boldsymbol{\sigma})$, Equation (8.14) shows the linking constant has the common value $\text{sgn}_{\boldsymbol{\sigma}}(B)$, which equals 1 since $\boldsymbol{\sigma}$ is a contributing singularity. On every other region the component adjacent to $\boldsymbol{\sigma}$ has smaller height than $\boldsymbol{\sigma}$, meaning the linking constant on that component (and thus every other component in the same region) is zero by Lemma 8.2. Corollary 8.1 therefore implies that B_0 lies in the same region $\mathcal{O}_j$ as $B(\boldsymbol{\sigma})$ if and only if $\boldsymbol{\sigma}$ is a contributing singularity. □

8.2.4 Step 4: Compute Residues

We now have an exact expression for the coefficients of interest as a sum of integrals over imaginary fibers. The final steps of the analysis are concerned with the asymptotic evaluation of the integrals appearing in (8.18). After using residue computations to reduce to saddle-point integrals, we obtain asymptotics through the use of Proposition 5.3 in Chapter 5.

The simplest case is when a contributing singularity lies on a flat defined by the intersection of d linear factors; this is known as a *complete intersection*. In the complete intersection case taking residues allows one to exactly determine the Cauchy integral, and no saddle-point integral approximations are necessary.

Example 8.3 Continued (Asymptotics in the Complete Intersection Case)

Let $\hat{r}_1 \in (2/3, 3/4)$, so that $\boldsymbol{\sigma}_{12} = (1,1)$ is the contributing singularity of maximum height (and the one on the unique bounded component of $\mathcal{M}_{\mathbb{R}}$ which is adjacent to the origin). Then, up to an exponentially decaying error coming from the contributing singularities of lower height, the sequence $\delta_n = [x^{r_1 n} y^{r_2 n}]F(x,y)$ equals

$$\delta_n = \frac{1}{(2\pi i)^2} \sum_{(\kappa_1,\kappa_2)\in\{\pm 1\}^2} \kappa_1\kappa_2 \int_{\mathbf{1}-\varepsilon\boldsymbol{\kappa}+i\mathbb{R}^2} \frac{1}{\left(1-\frac{2x+y}{3}\right)\left(1-\frac{3x+y}{4}\right)} \frac{dx\,dy}{x^{r_1 n+1} y^{r_2 n+1}} \tag{8.19}$$

for any $\varepsilon > 0$ sufficiently small. Making the substitution $(a,b) = (1-x, 1-y)$ in (8.19) to move $\boldsymbol{\sigma}_{12}$ to the origin gives

$$\delta_n = \frac{1}{(2\pi i)^2} \sum_{(\kappa_1,\kappa_2)\in\{\pm 1\}^2} \kappa_1\kappa_2 \int_{\varepsilon\boldsymbol{\kappa}+i\mathbb{R}^2} \frac{1}{\left(\frac{2a+b}{3}\right)\left(\frac{3a+b}{4}\right)} \frac{da\,db}{(1-a)^{r_1 n+1}(1-b)^{r_2 n+1}}. \tag{8.20}$$

Let

$$M = \begin{pmatrix} \mathbf{b}^{(1)} \\ \mathbf{b}^{(2)} \end{pmatrix} = \begin{pmatrix} 2/3 & 1/3 \\ 3/4 & 1/4 \end{pmatrix}$$

so that $M^{-1} = \begin{pmatrix} -3 & 4 \\ 9 & -8 \end{pmatrix}$. Making the change of variables $\begin{pmatrix} p \\ q \end{pmatrix} = M \begin{pmatrix} a \\ b \end{pmatrix}$ in (8.20) then implies

$$\delta_n = \frac{1}{(\det M)(2\pi \mathrm{i})^2} \sum_{(\kappa_1,\kappa_2)\in\{\pm 1\}^2} \kappa_1\kappa_2 \int_{\varepsilon\boldsymbol{\kappa}+\mathrm{i}\mathbb{R}^2} \frac{1}{pq} \frac{dp\,dq}{(1+3p-4q)^{r_1 n+1}(1-9p+8q)^{r_2 n+1}}. \tag{8.21}$$

Repeating the argument of Proposition 8.1, in reverse, implies that this signed sum of integrals equals the integral over a product of circles,

$$\delta_n = \frac{1}{(\det M)(2\pi \mathrm{i})^2} \int_{\mathcal{T}_\varepsilon \times \mathcal{T}_\varepsilon} \frac{1}{pq} \frac{dp\,dq}{(1+3p-4q)^{r_1 n+1}(1-9p+8q)^{r_2 n+1}}$$

where $\mathcal{T}_\varepsilon = \{\varepsilon e^{\mathrm{i}\theta} : -\pi < \theta \le \pi\}$. Applying the univariate residue theorem (twice) thus gives

$$[x^{r_1 n} y^{r_2 n}]F(x,y) = \delta_n + O(\tau^n) = \frac{1}{\det M} + O(\tau^n) = 12 + O(\tau^n),$$

for some $0 < \tau < 1$.

When $\boldsymbol{\sigma}$ lies in a flat defined by less than d equations, one first computes a sequence of residues and then must deal with an additional integral which is amenable to a saddle-point analysis.

Example 8.3 Continued (Asymptotics in the Non-Complete Intersection Case)

Consider the main diagonal direction $r_1 = r_2 = 1$ so $\boldsymbol{\sigma}_1 = (3/4, 3/2) \in \mathcal{V}_1$ is the contributing singularity of maximum height (and the one on the unique bounded component of $\mathcal{M}_\mathbb{R}$ which is adjacent to the origin). For $\varepsilon > 0$ sufficiently small, the points $(3/4, 3/2) \pm \varepsilon(3/2, 0)$ lie in the two components of $\mathcal{M}_\mathbb{R}$ whose closure contains $\boldsymbol{\sigma}_1$. Thus, up to an error decreasing exponentially faster than $(8/9)^n$, the diagonal sequence $[x^n y^n]F(x,y)$ equals

$$\delta_n = \frac{1}{(2\pi \mathrm{i})^2} \sum_{\kappa\in\{\pm 1\}} \kappa \int_{(3/4,3/2)-\kappa\varepsilon(3/2,0)+\mathrm{i}\mathbb{R}^2} \frac{\left(1-\frac{3x+y}{4}\right)^{-1}}{\left(1-\frac{2x+y}{3}\right)} \frac{dx\,dy}{x^{n+1}y^{n+1}}. \tag{8.22}$$

Let M be the matrix

$$M = \begin{pmatrix} \mathbf{b}^{(1)} \\ 0 \quad\quad 1 \end{pmatrix} = \begin{pmatrix} 2/3 & 1/3 \\ 0 & 1 \end{pmatrix}.$$

Making the change of variables $\begin{pmatrix} p \\ q \end{pmatrix} = M\left(\boldsymbol{\sigma}_1 - \begin{pmatrix} x \\ y \end{pmatrix}\right)$ in (8.22) implies

$$\delta_n = \frac{1}{(2\pi \mathrm{i})^2} \sum_{\kappa \in \{\pm 1\}} \kappa \int_{(\kappa\varepsilon,0)+\mathrm{i}\mathbb{R}^2} \frac{16}{p(1+18p-2q)} \frac{dp\,dq}{(3/4-3p/2+q/2)^{n+1}(3/2-q)^{n+1}}. \quad (8.23)$$

Thus, if $\tilde{\omega}$ is the integrand in (8.23) then there exists $\tau \in (0, 8/9)$ such that

$$[x^n y^n]F(x,y) = \frac{1}{(\det M)(2\pi \mathrm{i})^2}\left[\int_{(\varepsilon,0)+\mathrm{i}\mathbb{R}^2} \tilde{\omega} - \int_{(-\varepsilon,0)+\mathrm{i}\mathbb{R}^2} \tilde{\omega}\right] + O(\tau^n). \quad (8.24)$$

As one deforms the domain of integration $(\varepsilon, 0)+\mathrm{i}\mathbb{R}^2$ into $(-\varepsilon, 0)+\mathrm{i}\mathbb{R}^2$ by moving the basepoint $(\varepsilon, 0)$ on the line segment to $(-\varepsilon, 0)$, the only singularities of $\tilde{\omega}$ encountered are those lying in the set

$$\Gamma = \{(0, is) : s \in \mathbb{R}\} = \{0\} \times \mathrm{i}\mathbb{R}.$$

The difference of integrals in (8.24) thus equals the integral over the tubular domain

$$\mathcal{T}_\varepsilon \times \mathrm{i}\mathbb{R} = \{(\varepsilon e^{i\theta}, is) : -\pi < \theta \le \pi, s \in \mathbb{R}\}$$

around Γ, and the (univariate) residue theorem implies

$$\begin{aligned} [x^n y^n]F(x,y) &= \frac{3}{2(2\pi \mathrm{i})^2} \int_{\mathcal{T}_\varepsilon \times \mathrm{i}\mathbb{R}} \frac{16}{p(1+18p-2q)} \frac{dp\,dq}{(3/4-3p/2+q/2)^{n+1}(3/2-q)^{n+1}} + O(\tau^n) \\ &= \frac{3}{4\pi \mathrm{i}} \int_{\mathrm{i}\mathbb{R}} \frac{16}{(1-2q)} \frac{dq}{(3/4+q/2)^{n+1}(3/2-q)^{n+1}} + O(\tau^n). \end{aligned}$$

Unlike the complete intersection case, asymptotics of this remaining integral must be calculated using the saddle-point method. Making a final change of variables $q = iw$ gives

$$[x^n y^n]F(x,y) = \frac{3}{4\pi} \int_{\mathbb{R}} \frac{128}{(1-2iw)(9+4w^2)} e^{-n\log(9/8+w^2/2)} dw + O(\tau^n),$$

and Proposition 5.3 of Chapter 5 implies

$$[x^n y^n]F(x,y) = \left(\frac{8}{9}\right)^n n^{-1/2} \frac{16}{\sqrt{\pi}} \left(1 + O\left(\frac{1}{n}\right)\right).$$

We now give the general argument. Let $\boldsymbol{\sigma}$ be a contributing singularity on the flat $\mathcal{V}_S$ with $S = \{k_1, \ldots, k_s\}$, and to save space define

$$I_{\boldsymbol{\sigma}} = \frac{\operatorname{sgn}(\boldsymbol{\sigma})}{(2\pi \mathrm{i})^d} \int_{\boldsymbol{\tau}_{\boldsymbol{\sigma}}} \frac{G(\mathbf{z})}{\mathbf{z}^{\mathbf{nr}+\mathbf{1}} \prod_{j=1}^m \ell_j(\mathbf{z})^{p_j}}.$$

As we assume the $\mathbf{b}^{(k_j)}$ are linearly independent, up to a reordering of variables and the ℓ_{k_j} we may also assume that the matrix

$$M = \begin{pmatrix} \mathbf{b}^{(k_1)} \\ \vdots \\ \mathbf{b}^{(k_s)} \\ \mathbf{e}^{(s+1)} \\ \vdots \\ \mathbf{e}^{(d)} \end{pmatrix} \tag{8.25}$$

is non-singular, where once again $\mathbf{e}^{(j)}$ denotes the jth elementary basis vector. Let

$$\tilde{G}(\mathbf{z}) = \frac{G(\mathbf{z})}{z_1 \cdots z_d \prod_{j \notin S} \ell_j(\mathbf{z})^{p_j}},$$

so that the Cauchy integrand ω from (8.3) can be written

$$\omega = \frac{\tilde{G}(\mathbf{z})}{\ell_{k_1}(\mathbf{z})^{p_{k_1}} \cdots \ell_{k_s}(\mathbf{z})^{p_{k_s}} \mathbf{z}^{n\mathbf{r}}}.$$

Finally, given $\boldsymbol{\kappa} \in \{\pm 1\}^s$ let

$$\mathbf{v}_{\boldsymbol{\kappa}} = M^{-1} \left(\kappa_1 \mathbf{e}^{(1)} + \cdots + \kappa_s \mathbf{e}^{(s)} \right).$$

By definition, the dot product $\mathbf{v}_{\boldsymbol{\kappa}} \cdot \mathbf{b}^{(k_j)} = \kappa_j$ for all $j = 1, \ldots, s$. If $\mathbf{x}_{\boldsymbol{\kappa}} = \boldsymbol{\sigma} - \varepsilon \mathbf{v}_{\boldsymbol{\kappa}}$ then $\ell_{k_j}(\mathbf{x}_{\boldsymbol{\kappa}}) = \varepsilon \kappa_j$ for $j = 1, \ldots, s$, so for $\varepsilon > 0$ sufficiently small the set $\{\mathbf{x}_{\boldsymbol{\kappa}} : \boldsymbol{\kappa} \in \{\pm 1\}^s\}$ consists of 2^s real points, one in each of the 2^s components of $\mathcal{M}_{\mathbb{R}}$ adjacent to $\boldsymbol{\sigma}$. Note that $\mathbf{x_1}$ is contained in the component $B(\boldsymbol{\sigma})$ on which $\boldsymbol{\sigma}$ minimizes h.

If we assume that $p_{k_1} = \cdots = p_{k_s} = 1$, so that each linear factor in the denominator of F appears only to first order, then

$$I_{\boldsymbol{\sigma}} = \frac{\operatorname{sgn}(\boldsymbol{\sigma})}{(2\pi \mathrm{i})^d} \sum_{\boldsymbol{\kappa} \in \{\pm 1\}^s} \operatorname{sgn}(\boldsymbol{\kappa}) \int_{C_{\mathbf{x}_{\boldsymbol{\kappa}}}} \frac{\tilde{G}(\mathbf{z})}{\ell_{k_1}(\mathbf{z}) \cdots \ell_{k_s}(\mathbf{z}) \mathbf{z}^{n\mathbf{r}}} d\mathbf{z},$$

and making the change of variables $\mathbf{w} = M(\boldsymbol{\sigma} - \mathbf{z})$ yields

$$I_{\boldsymbol{\sigma}} = \frac{\operatorname{sgn}(\boldsymbol{\sigma})}{|\det M| (2\pi \mathrm{i})^d} \sum_{\boldsymbol{\kappa} \in \{\pm 1\}^s} \operatorname{sgn}(\boldsymbol{\kappa}) \int_{\varepsilon(\boldsymbol{\kappa}, \mathbf{0}) + \mathrm{i}\mathbb{R}^d} \frac{\tilde{G}\left(\boldsymbol{\sigma} - M^{-1}\mathbf{w}\right)}{w_1 \cdots w_s \left(\boldsymbol{\sigma} - M^{-1}\mathbf{w}\right)^{n\mathbf{r}}} d\mathbf{w}.$$

Replacing the alternating sum of integrals with an integral over a product of circles, and applying the residue theorem s times, gives

$$
\begin{aligned}
I_{\boldsymbol{\sigma}} &= \frac{\operatorname{sgn}(\boldsymbol{\sigma})}{|\det M|\,(2\pi\mathrm{i})^d} \int_{\mathcal{T}_\varepsilon^s \times \mathrm{i}\mathbb{R}^{d-s}} \frac{\tilde{G}\left(\boldsymbol{\sigma} - M^{-1}\mathbf{w}\right)}{w_1 \cdots w_s \left(\boldsymbol{\sigma} - M^{-1}\mathbf{w}\right)^{n\mathbf{r}}}\, d\mathbf{w} \\
&= \frac{\operatorname{sgn}(\boldsymbol{\sigma})}{|\det M|\,(2\pi\mathrm{i})^{d-s}} \int_{\mathrm{i}\mathbb{R}^{d-s}} \left.\frac{\tilde{G}\left(\boldsymbol{\sigma} - M^{-1}\mathbf{w}\right)}{\left(\boldsymbol{\sigma} - M^{-1}\mathbf{w}\right)^{n\mathbf{r}}}\right|_{w_j=0,\, j\le s} dw_{s+1}\cdots dw_d \\
&= \frac{\operatorname{sgn}(\boldsymbol{\sigma})\boldsymbol{\sigma}^{-n\mathbf{r}}}{|\det M|\,(2\pi)^{d-s}} \int_{\mathbb{R}^{d-s}} A_{\boldsymbol{\sigma}}(\mathbf{y}) e^{-n\phi_{\boldsymbol{\sigma}}(\mathbf{y})} d\mathbf{y},
\end{aligned} \tag{8.26}
$$

where $y_j = (-\mathrm{i})w_{s+j}$ for $1 \le j \le d-s$ and

$$
A_{\boldsymbol{\sigma}}(\mathbf{y}) = \tilde{G}\left(\boldsymbol{\sigma} - \mathrm{i}M^{-1}\begin{pmatrix}\mathbf{0}\\ \mathbf{y}\end{pmatrix}\right), \quad \phi_{\boldsymbol{\sigma}}(\mathbf{y}) = \sum_{j=1}^{d} r_j \log\left(\frac{\sigma_j - \mathrm{i}\left(M^{-1}\begin{pmatrix}\mathbf{0}\\ \mathbf{y}\end{pmatrix}\right)_j}{\sigma_j}\right). \tag{8.27}
$$

Note that the real part of $\phi_{\boldsymbol{\sigma}}$ is the height function h, and that by construction the gradient of $\phi_{\boldsymbol{\sigma}}$ vanishes at the origin. When $d = s$, i.e., in the case of a complete intersection, (8.26) is taken to mean

$$
I_{\boldsymbol{\sigma}} = \frac{\tilde{G}(\boldsymbol{\sigma})}{|\det M|}\,\boldsymbol{\sigma}^{-n\mathbf{r}} = \frac{G(\boldsymbol{\sigma})}{\sigma_1\cdots\sigma_d \prod_{j\notin S} \ell_j(\boldsymbol{\sigma})^{p_j}\,|\det M|}\,\boldsymbol{\sigma}^{-n\mathbf{r}}.
$$

Combining this calculation with Theorem 8.1 gives the following.

Proposition 8.3 *Consider the power series expansion of simple $F(\mathbf{z})$ and suppose $\mathbf{r}$ is a generic direction with positive coordinates. If $\mathbf{p} = \mathbf{1}$ and χ denotes the set of contributing singularities of $F(\mathbf{z})$ then*

$$
[\mathbf{z}^{n\mathbf{r}}]F(\mathbf{z}) = \sum_{\boldsymbol{\sigma}\in\chi} I_{\boldsymbol{\sigma}}, \tag{8.28}
$$

where $I_{\boldsymbol{\sigma}}$ is defined in (8.26) *using the quantities in* (8.25) *and* (8.27).

For general $\mathbf{p} \in \mathbb{N}^m$, when the linear factors in the denominator appear to higher order,

$$
I_{\boldsymbol{\sigma}} = \frac{1}{|\det M|\,(2\pi\mathrm{i})^d} \int_{\mathcal{T}_\varepsilon^s \times \mathrm{i}\mathbb{R}^{d-s}} \frac{\tilde{G}\left(\boldsymbol{\sigma} - M^{-1}\mathbf{w}\right)}{w_1^{p_{k_1}} \cdots w_s^{p_{k_s}} \left(\boldsymbol{\sigma} - M^{-1}\mathbf{w}\right)^{n\mathbf{r}}}\, d\mathbf{w}, \tag{8.29}
$$

and the iterated residue obtained by integrating over the product of circles in the variables $w_1, \dots, w_s$ is

$$
\begin{aligned}
R_{\boldsymbol{\sigma}}(w_{s+1}, \dots, w_d) =& \frac{1}{(p_{k_1}-1)!\cdots(p_{k_s}-1)!} \\
&\times \left(\frac{\partial^{p_{k_1}+\cdots+p_{k_s}-s}}{\partial w_1^{p_{k_1}-1}\cdots\partial w_s^{p_{k_s}-1}}\right) \left.\frac{\tilde{G}\left(\boldsymbol{\sigma} - M^{-1}\mathbf{w}\right)}{\left(\boldsymbol{\sigma} - M^{-1}\mathbf{w}\right)^{n\mathbf{r}}}\right|_{w_1=\cdots=w_s=0}.
\end{aligned} \tag{8.30}
$$

Repeated differentiation shows

$$R_{\boldsymbol{\sigma}} = P_{\boldsymbol{\sigma}}(n) \times \left(\boldsymbol{\sigma} - M^{-1}\mathbf{w}\right)^{-n\mathbf{r}}\Big|_{w_1=\cdots=w_s=0},$$

where $P_{\boldsymbol{\sigma}}(n)$ is a polynomial

$$P_{\boldsymbol{\sigma}}(n) = \sum_{j=1}^{p_{k_1}+\cdots+p_{k_s}-s} a_{j,\boldsymbol{\sigma}}(w_{s+1},\dots,w_d)\, n^j, \tag{8.31}$$

in n whose coefficients are rational functions in $w_{s+1},\dots,w_d$ which are analytic at the origin. The leading coefficient of $P_{\boldsymbol{\sigma}}(n)$ is

$$a_{p_{k_1}+\cdots+p_{k_s}-s,\boldsymbol{\sigma}} = \prod_{j=1}^{s} \frac{v_j(\mathbf{w})^{p_{k_j}-1}}{(p_{k_j}-1)!}\tilde{G}\left(\boldsymbol{\sigma} - M^{-1}\mathbf{w}\right)\Bigg|_{w_1=\cdots=w_s=0},$$

where $\mathbf{v}(\mathbf{w}) = \left(\frac{r_1}{\sigma_1-(M^{-1}\mathbf{w})_1} \ \cdots \ \frac{r_d}{\sigma_d-(M^{-1}\mathbf{w})_d}\right)\cdot M^{-1}$. Thus, we obtain the following.

Proposition 8.4 *Consider the power series expansion of simple $F(\mathbf{z})$, and suppose $\mathbf{r}$ is a generic direction with positive coordinates. If χ denotes the set of contributing singularities for $F(\mathbf{z})$ then*

$$[\mathbf{z}^{n\mathbf{r}}]F(\mathbf{z}) = \sum_{\boldsymbol{\sigma}\in\chi} I_{\boldsymbol{\sigma}}, \tag{8.32}$$

where, for $\boldsymbol{\sigma}$ on the flat defined by the vanishing of $\ell_{k_1},\dots,\ell_{k_s}$ appearing in the denominator of $F(\mathbf{z})$ to multiplicities $p_{k_1},\dots,p_{k_s}$,

$$I_{\boldsymbol{\sigma}} = \sum_{j=0}^{p_{k_1}+\cdots+p_{k_s}-s} n^j \int_{\mathbb{R}^{d-s}} A_{j,\boldsymbol{\sigma}}(\mathbf{y})e^{-n\phi_{\boldsymbol{\sigma}}(\mathbf{y})}d\mathbf{y}$$

for effective rational functions $A_{j,\boldsymbol{\sigma}} = a_{j,\boldsymbol{\sigma}}(-\mathrm{i}y_1,\dots,-\mathrm{i}y_{d-s})$ determined by (8.30) and (8.31) and $\phi_{\boldsymbol{\sigma}}$ determined by (8.25) and (8.27).

Remark 8.10 When $d = s$ the residue R_n depends only on n. Thus, if $\boldsymbol{\sigma}$ lies on a complete intersection then $I_{\boldsymbol{\sigma}} = \boldsymbol{\sigma}^{-n\mathbf{r}}P_{\boldsymbol{\sigma}}(n)$, where $P_{\boldsymbol{\sigma}}(n)$ is a polynomial in n of degree $p_{k_1}+\cdots+p_{k_d}-d$.

8.2.5 Step 5: Determine Asymptotics

Recall Proposition 5.3 from Chapter 5, which determines an asymptotic expansion for integrals of the form appearing in Propositions 8.3 and 8.4. The following result provides the only missing details we need for asymptotics, and can be viewed as an analogue of Lemma 5.6 in Chapter 5 which holds beyond the smooth case.

Lemma 8.4 *Fix an orthant $O \subset \mathbb{R}^d$. For any real flat $L = \mathcal{V}'_{k_1,\dots,k_s}$ bounded in O, the critical point $\boldsymbol{\sigma}(\mathbf{r})$ of L in O for the direction $\mathbf{r}$ varies smoothly with $\mathbf{r}$. As $\mathbf{s}$ varies over a neighbourhood of $\mathbf{r}$ in $\mathbb{R}^d_*$ the values of $\boldsymbol{\sigma}(\mathbf{s})$ cover a neighbourhood of $\boldsymbol{\sigma}(\mathbf{r})$ in L. Furthermore, if A is the $d \times (d-c)$ submatrix of M^{-1} consisting of its $d-s$ rightmost columns, and D is the diagonal matrix with entries $\sqrt{r_1}/\sigma_1, \dots, \sqrt{r_d}/\sigma_d$, then the Hessian of $\phi_{\boldsymbol{\sigma}}$ defined by (8.27) at the origin equals*

$$\mathcal{H}_{\boldsymbol{\sigma}} = (DA)^T(DA), \tag{8.33}$$

which has strictly positive determinant.

Proof Recall the explicit parametrization (8.9) for the critical points on the flat L, derived in the proof of Lemma 8.1. Differentiating the right-hand side of (8.9) with respect to z_b and substituting $\mathbf{z}' = \boldsymbol{\sigma}'$ gives

$$\sum_{j=1}^{d} \frac{r_j}{\sigma_j^2} M^{-1}_{j,a} M^{-1}_{j,b},$$

so the Jacobian of the system (8.9) with respect to $\mathbf{z}'$ at $\boldsymbol{\sigma}'$ equals $\mathcal{H}_{\boldsymbol{\sigma}} = (DA)^T(DA)$. Because of the form of $\mathcal{H}_{\boldsymbol{\sigma}}$, the Cauchy-Binet formula (see Problem 8.9) implies its determinant is a sum of the squares of the determinants of the $(d-s) \times (d-s)$ submatrices of DA. Since the final $d-s$ rows of DA form a diagonal matrix Λ with non-zero entries $\sqrt{r_{s+1}}/\sigma_{s+1}, \dots, \sqrt{r_d}/\sigma_d$, this in turn implies

$$\det \mathcal{H}_{\boldsymbol{\sigma}} \geq (\det \Lambda)^2 = \frac{r_{s+1} \cdots r_d}{\sigma_{k+1}^2 \cdots \sigma_d^2} > 0,$$

and $\mathcal{H}_{\boldsymbol{\sigma}}$ is non-singular. Thus, the critical point $\boldsymbol{\sigma}(\mathbf{r})$ varies smoothly with $\mathbf{r}$ and when $\mathbf{s}$ varies over a neighbourhood of some fixed $\mathbf{r}$ then $\boldsymbol{\sigma}(\mathbf{s})$ covers a neighbourhood of $\boldsymbol{\sigma}(\mathbf{r})$ in L. Since the system (8.9) is formed from the gradient of $h(\mathbf{z}')$ in (8.8), the Jacobian of (8.9) equals the Hessian of $h(\mathbf{z}')$. The function $\phi(\mathbf{y})$ in (8.27) equals $-h(\mathbf{z}')$ after a substitution of the form $z_j = C_j - \mathrm{i}y_j$ for constants C_j. This implies the Hessian of $\phi(\mathbf{y})$ is $-(-\mathrm{i})(-\mathrm{i})\mathcal{H}_{\boldsymbol{\sigma}} = \mathcal{H}_{\boldsymbol{\sigma}}$. □

Theorem 8.2 *Consider the power series expansion of simple $F(\mathbf{z})$, suppose $\mathbf{r}$ is a generic direction with positive coordinates, and let χ denote the set of contributing singularities for $F(\mathbf{z})$. Then there exist asymptotic series $\Phi_{\boldsymbol{\sigma}}(n)$ such that*

$$[\mathbf{z}^{n\mathbf{r}}]F(\mathbf{z}) = \sum_{\boldsymbol{\sigma} \in \chi} \Phi_{\boldsymbol{\sigma}}(n). \tag{8.34}$$

In particular, if $\boldsymbol{\sigma}$ lies on the flat $\mathcal{V}_S$ with $S = \{k_1, \dots, k_s\}$ then, for any positive integer K, there exist effective constants $C_j^{\boldsymbol{\sigma}}$ such that

$$\Phi_{\boldsymbol{\sigma}}(n) = \boldsymbol{\sigma}^{-n\mathbf{r}}\, n^{p_{k_1}+\cdots+p_{k_s}-(s+d)/2} \left(\sum_{j=0}^{K} C_j^{\boldsymbol{\sigma}} n^{-j} + O\left(n^{-K-1}\right) \right). \tag{8.35}$$

If $G(\boldsymbol{\sigma}) \neq 0$ *then the leading asymptotic term of* $\Phi_{\boldsymbol{\sigma}}$ *is*

$$C_0^{\boldsymbol{\sigma}} = \prod_{j=1}^{s} \frac{v_j^{p_{k_j}-1}}{(p_{k_j}-1)!} \times \frac{G(\boldsymbol{\sigma})}{\prod_{j \notin S} \ell_j(\boldsymbol{\sigma})^{p_j}} \times \frac{\det(\mathcal{H}_{\boldsymbol{\sigma}})^{-1/2}}{\sigma_1 \cdots \sigma_d \,|\det M|\,(2\pi)^{(d-s)/2}} \neq 0,$$

where M *is defined in* (8.25), $\mathcal{H}_{\boldsymbol{\sigma}}$ *is defined in* (8.33), *and* $\mathbf{v} = \begin{pmatrix} \frac{r_1}{\sigma_1} & \cdots & \frac{r_d}{\sigma_d} \end{pmatrix} \cdot M^{-1}$. *The implied constant in the error term of* (8.35) *can be uniformly bounded as* $\mathbf{r}$ *varies in any connected compact set which does not contain a non-generic direction.*

Proof As stated above, this is an application of Proposition 5.3 from Chapter 5 to Proposition 8.4. In particular, if $\boldsymbol{\sigma} \in \chi$ and $\phi_{\boldsymbol{\sigma}}(\mathbf{y})$ is defined by (8.27) then

- $\phi_{\boldsymbol{\sigma}}$ and its gradient vanish at the origin by direct inspection;
- Lemma 8.4 shows that the Hessian matrix $\mathcal{H}_{\boldsymbol{\sigma}}$ is non-singular;
- the real part of $\phi_{\boldsymbol{\sigma}}$ is

$$\Re\left(\phi_{\boldsymbol{\sigma}}\right) = \sum_{j=1}^{d} r_j \log \left| \frac{\sigma_j - \mathrm{i}\left(M^{-1}\begin{pmatrix}\mathbf{0}\\ \mathbf{y}\end{pmatrix}\right)_j}{\sigma_j} \right|,$$

  which is non-negative on $\mathbb{R}^{d-s}$, and equal to 0 precisely when $\mathbf{y} = \mathbf{0}$.
- because $\phi_{\boldsymbol{\sigma}}(\mathbf{y})$ has strictly negative real part when $\mathbf{y}$ is bounded away from the origin, which is an isolated critical point, we can restrict each integral in Proposition 8.4 to a neighbourhood where the origin is the only critical point of $\phi_{\boldsymbol{\sigma}}$ while introducing only an exponentially negligible error.

This verifies all the necessary conditions to apply Proposition 5.3. □

Figure 8.8 presents our current asymptotic picture for our running Example 8.3.

Remark 8.11 When $\boldsymbol{\sigma}$ is a contributing singularity there exist $a_1, \ldots, a_s > 0$ with

$$-(\nabla h)(\boldsymbol{\sigma}) = a_1 \mathbf{b}^{(k_1)} + \cdots + a_s \mathbf{b}^{(k_s)},$$

so that $\begin{pmatrix}\mathbf{a} & \mathbf{0}\end{pmatrix} M = -(\nabla h)(\boldsymbol{\sigma})$. Thus, for $1 \leq j \leq s$ the constants v_j in Theorem 8.2 are $v_j = a_j > 0$. When there is a single contributing singularity $\boldsymbol{\sigma}$ of maximum height, and $G(\boldsymbol{\sigma}) \neq 0$, then Theorem 8.2 gives dominant asymptotics of the coefficient sequence,

$$[\mathbf{z}^{n\mathbf{r}}]F(\mathbf{z}) \sim \boldsymbol{\sigma}^{-n\mathbf{r}}\, n^{p_{k_1}+\cdots+p_{k_s}-(s+d)/2} \times \prod_{j=1}^{s} \frac{v_j^{p_{k_j}-1}}{(p_{k_j}-1)!} \times \frac{G(\boldsymbol{\sigma}) \det(\mathcal{H}_{\boldsymbol{\sigma}})^{-1/2}}{\sigma_1 \cdots \sigma_d \prod_{j \notin S} \ell_j(\boldsymbol{\sigma}) \,|\det M|\,(2\pi)^{(d-s)/2}}.$$

In the simple pole case, when each $p_j = 1$, this simplifies further to

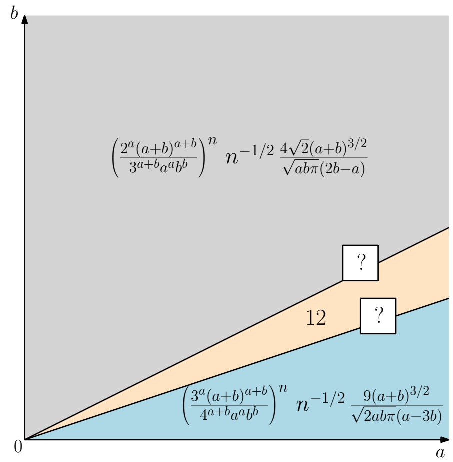


Fig. 8.8 The three asymptotic regimes for the coefficients $[x^{an}y^{bn}]F(x,y)$ of the rational function $F(x,y) = 1 \Big/ \left(1 - \frac{2x+y}{3}\right)\left(1 - \frac{3x+y}{4}\right)$. Asymptotics in non-generic directions have not yet been derived and are thus denoted by question marks.

$$[\mathbf{z}^{n\mathbf{r}}]F(\mathbf{z}) \sim \boldsymbol{\sigma}^{-n\mathbf{r}}\, n^{(d-s)/2} \frac{G(\boldsymbol{\sigma})\det(\mathcal{H}_{\boldsymbol{\sigma}})^{-1/2}}{\sigma_1 \cdots \sigma_d \prod_{j \notin S} \ell_j(\boldsymbol{\sigma})\, |\det M|\, (2\pi)^{(d-s)/2}}.$$

Unfortunately, when the numerator of F vanishes at the contributing singularities determining dominant asymptotics can be more difficult.

Example 8.4 (Asymptotics With Vanishing Numerator)

Returning to an example from Chapter 5, let

$$B(x,y) = \frac{x - 2y^2}{1 - x - y} \qquad \text{and} \qquad C(x,y) = \frac{x-y}{1-x-y}.$$

In the main diagonal direction $\mathbf{r} = (1,1)$ the functions B and C both admit the single contributing singularity $\boldsymbol{\sigma} = (1/2, 1/2)$, but while $[x^n y^n]B(x,y)$ has the expected exponential growth of 4^n, even though its numerator vanishes at $\boldsymbol{\sigma}$, the sequence $[x^n y^n]C(x,y)$ is identically zero.

Given a contributing singularity $\boldsymbol{\sigma}$, the saddle-point integral $I_{\boldsymbol{\sigma}}$ in (8.32) can be written as a series in n^{-1}, and we can determine any number of terms in the series using Proposition 5.3 of Chapter 5.

Remark 8.12 When $\boldsymbol{\sigma}$ lies on a complete intersection then $I_{\boldsymbol{\sigma}} = \boldsymbol{\sigma}^{-n\mathbf{r}} P_{\boldsymbol{\sigma}}(n)$, where $P_{\boldsymbol{\sigma}}(n)$ is a polynomial in n of degree $p_{k_1} + \cdots + p_{k_d} - d$. Thus, to determine whether $I_{\boldsymbol{\sigma}}$ is identically zero one simply needs to compute the first $p_{k_1} + \cdots + p_{k_d} - d$ terms of its asymptotic expansion using Proposition 5.3.

Algorithm 4: GenericHyperplaneAsymptotics

Input: Rational function $F(\mathbf{z}) = G(\mathbf{z})/\prod_{j=1}^{m} \ell_j(\mathbf{z})^{p_j}$ with linear $\ell_j(\mathbf{z}) = 1 - \mathbf{b}^{(j)} \cdot \mathbf{z}$, and direction $\mathbf{r} = (r_1, \ldots, r_d)$
Output: Asymptotic expansion of $[z^{n\mathbf{r}}]F(\mathbf{z})$ as $n \to \infty$, uniform in compact set around $\mathbf{r}$

```
if matrix with rows b^(j) not full rank then
    run Algorithm 5 (SimpleDecomp) on F and apply GenericHyperplaneAsymptotics to
     to each summand in the result
end
compute the flats defined by the ℓ_j(z)
find the critical points on each flat by solving (8.2)
sort the critical points using height function h_r and set σ equal to point of max height
while no contributing points have been found do
    compute the cone N(σ) from Definition 8.8
    if (−∇h)(σ) on boundary of N(σ) then
        return FAIL, "r not a generic direction"
    end
    if (−∇h)(σ) ∉ N(σ) then
        set σ to next lowest height critical point and repeat loop
    end
    if G(σ) ≠ 0 then
        σ is a contributing singularity
        repeat above steps to identify contributing singularities of the same height as σ
    else if G in polynomial ideal generated by the factors ℓ_j with ℓ_j(σ) = 0 then
        replace F by sum of rational functions using (8.36) and analyze each summand
    else
        return FAIL
    end
end
return sum of contributions (8.35) from each contributing singularity of highest height
```

There is not currently an effective characterization for vanishing of the integral $I_{\boldsymbol{\sigma}}$. On the other hand, Lemma 8.4 implies that a critical point $\boldsymbol{\sigma}(\mathbf{s})$ covers a neighbourhood of its flat L when $\mathbf{s}$ moves over a neighbourhood in $\mathbb{R}^d$. Thus, if $G(\boldsymbol{\sigma}(\mathbf{s})) = 0$ vanishes for all $\mathbf{s}$ in some open set of $\mathbb{R}^d$ then G vanishes identically on L. In particular, this implies that if L is defined by the vanishing of $\ell_1, \ldots, \ell_s$ then there exist polynomials $Q_1, \ldots, Q_s$ such that

$$G(\mathbf{z}) = Q_1(\mathbf{z})\ell_1(\mathbf{z}) + \cdots + Q_s(\mathbf{z})\ell_s(\mathbf{z}).$$

In other words,

$$\frac{G(\mathbf{z})}{\prod_{1\le j\le m}^{m} \ell_j(\mathbf{z})^{p_j}} = \frac{Q_1(\mathbf{z})}{\prod_{j=1}^{m} \ell_j(\mathbf{z})^{p_j}/\ell_{k_1}(\mathbf{z})} + \cdots + \frac{Q_s(\mathbf{z})}{\prod_{j=1}^{m} \ell_j(\mathbf{z})^{p_j}/\ell_{k_s}(\mathbf{z})}, \qquad (8.36)$$

and this can be detected automatically by a Gröbner basis computation (of the kind discussed in Chapter 7). The analysis at any such critical point $\boldsymbol{\sigma}(\mathbf{s})$ can then be applied to the summands, each of which has a denominator which vanishes to lower order at $\boldsymbol{\sigma}(\mathbf{s})$. If each $p_{k_j} = 1$ then the hyperplane arrangement defined by any

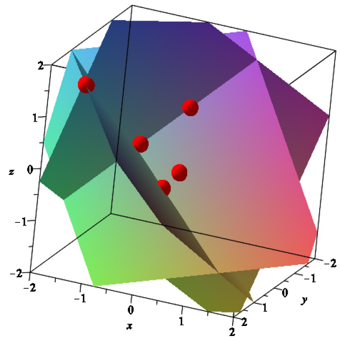


Fig. 8.9 Plot of singular hyperplanes and contributing singularities for our three-dimensional example (note only 5 of the 7 contributing singularities are visible from this angle).

summand does not contain the flat L and $I_{\sigma(\mathbf{s})}$ is identically zero. Thus, for most directions either $G(\sigma) \neq 0$ at all contributing singularities σ or I_σ is identically zero for some contributing singularity in a manner that we are able to detect.

Algorithm 4 lists the computations needed to determine asymptotics. Implementations of Algorithm 4 are linked on the book website.

Example 8.5 (A Three-Dimensional Example)

Consider the three-dimensional example

$$F(x, y, z) = \frac{x + y + z}{(1 - 2x - 3y - 4z)(1 + 4x + 3y + 2z)(1 + x - y + z)},$$

and label the denominator factors ℓ_1, ℓ_2, and ℓ_3 from left to right. Using an implementation of Algorithm 4, we find that there are seven contributing singularities in the direction $\mathbf{r} = (1, 1, 2)$: one each on the hypersurfaces $\mathcal{V}_1$, $\mathcal{V}_2$, and $\mathcal{V}_3$, two points (necessarily in different orthants) on the flat $\mathcal{V}_{1,2}$, and one point on each of the flats $\mathcal{V}_{1,3}$ and $\mathcal{V}_{2,3}$. The unique contributing point of highest height is $\sigma = (1/8, 1/12, 1/8)$ on $\mathcal{V}_1$, giving dominant asymptotics $[x^n y^n z^{2n}]F(x, y, z) = \frac{6144^n}{n}\left(\frac{\sqrt{2}}{14\pi} + O(1/n)\right)$.

8.2.6 Dealing with Non-Simple Arrangements

Our main results above require the factors ℓ_j in the denominator to be linearly independent. Luckily, when dependencies between the factors do exist such relations can be used to decompose the rational function of interest into a sum of functions where this is no longer a problem. The key observation is that from a linear dependence

$$a_1 \ell_{k_1}(\mathbf{z}) + \cdots + a_s \ell_{k_s}(\mathbf{z}) = 0 \tag{8.37}$$

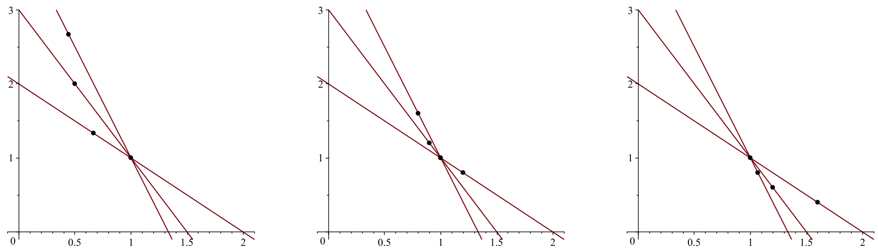


Fig. 8.10 Plots of the linearly dependent lines ℓ_1, ℓ_2, ℓ_3 together with the critical points for $\mathbf{r} = (1, 2)$ on the left, $\mathbf{r} = (3, 2)$ in the centre, and $\mathbf{r} = (8, 2)$ on the right.

with non-zero constants one can divide by any $a_j \ell_{k_j}(\mathbf{z})$ and then divide by any product of linear factors in the denominator to obtain

$$\frac{1}{\boldsymbol{\ell}^{\mathbf{q}}} = \frac{1}{\ell_1^{q_1} \cdots \ell_m^{q_m}} = \sum_{\substack{1 \le i \le s \\ i \ne j}} \frac{(-a_i/a_j)}{\boldsymbol{\ell}^{\mathbf{q}_{i,j}}}, \tag{8.38}$$

for any $\mathbf{q} \in \mathbb{N}^m$, where $\mathbf{q}_{i,j} = \mathbf{q} + \mathbf{e}^{(k_j)} - \mathbf{e}^{(k_i)}$ is one larger than $\mathbf{q}$ in its jth coordinate and one smaller than $\mathbf{q}$ in its ith coordinate.

Example 8.6 (A Non-Simple Rational Function)

Consider the rational function $F(x, y) = 1/(\ell_1 \ell_2 \ell_3)$, where

$$\ell_1(x, y) = 1 - \frac{2x + y}{3} \qquad \text{and} \qquad \ell_2(x, y) = 1 - \frac{3x + y}{4}$$

are the same as our running example, and

$$\ell_3(x, y) = 1 - \frac{x + y}{2}.$$

As the three lines defined by ℓ_1, ℓ_2, and ℓ_3 meet at the point $(x, y) = (1, 1)$, these functions are linearly dependent. Figure 8.10 shows the arrangement of critical points for three different directions $\mathbf{r} = (p, q)$. When $0 < p/q < 1$ or $3 < p/q$ then $F(x, y)$ admits a smooth minimal critical point and we do not need to worry about the linear dependency to determine dominant asymptotics. However, if $1 < p/q < 3$ then the only minimal critical point is the point $(1, 1)$ where the three lines defined by the ℓ_j intersect. Comparing coefficients in the equation $a\ell_1 + b\ell_2 - \ell_3 = 0$, where a and b are unknown constants, gives a linear system whose solution $(a, b) = (3, -2)$ implies $\ell_3 = 3\ell_1 - 2\ell_2$, so

$$1 = \frac{3\ell_1}{\ell_3} - \frac{2\ell_2}{\ell_3},$$

and thus

$$F(x,y) = \frac{\frac{3\ell_1}{\ell_3} - \frac{2\ell_2}{\ell_3}}{\ell_1\ell_2\ell_3} = \frac{3}{\ell_2\ell_3^2} - \frac{2}{\ell_1\ell_3^2}.$$

Finding coefficient asymptotics of $F(x, y)$ in the direction $\mathbf{r} = (p, q)$ is therefore reduced to finding asymptotics of the two summands in this expression, both of which are simple rational functions whose asymptotics are determined by Theorem 8.2 and Algorithm 4. If $1 < p/q < 2$ then

$$[x^{pn}y^{qn}]\frac{1}{\ell_2\ell_3^2} \sim 4(3q-p)n \quad \text{and} \quad [x^{pn}y^{qn}]\frac{1}{\ell_1\ell_3^2} \sim 12(q-p)n,$$

so

$$[x^{pn}y^{qn}]F(x,y) \sim 12(p-q)n.$$

The coefficients of $1/\ell_2\ell_3^2$ have the same behaviour when $2 < p/q < 3$, but the coefficients of $1/\ell_1\ell_3^2$ decay exponentially in this case, meaning

$$[x^{pn}y^{qn}]F(x,y) \sim 12(3q-p)n$$

when $2 < p/q < 3$. Any direction where $p/q = 2$ is a non-generic direction for $1/\ell_1\ell_3^2$, and will require the methods of Section 8.3 below.

To show that any rational function of interest can be decomposed into a sum of simple functions, we need to introduce some notation. The following terminology is borrowed from the theory of matroids.

Definition 8.12 (supports, broken circuits, and χ-independence) The *support* of a rational function

$$F(\mathbf{z}) = \frac{G(\mathbf{z})}{\ell_1(\mathbf{z})^{q_1}\cdots\ell_m(\mathbf{z})^{q_m}}$$

written with coprime numerator G and linear denominator factors ℓ_j is

$$\sup(F) = \left\{\ell_j(\mathbf{z}) : q_j > 0\right\}.$$

A minimal linearly dependent set of factors $\{\ell_{k_1}, \ldots, \ell_{k_s}\}$ is called a *circuit*; note that any proper subset of a circuit is linearly independent. A *broken circuit* is the independent set obtained by removing the element of largest index from a circuit. A collection of the linear factors is χ-*independent* if it does not contain a broken circuit.

Any dependent set contains a broken circuit, so a χ-independent set is also independent. In our last example the only dependent set is $\{\ell_1, \ell_2, \ell_3\}$, so the unique broken circuit is $\{\ell_1, \ell_2\}$. The χ-independent sets are then $\{\ell_1, \ell_3\}$ and $\{\ell_2, \ell_3\}$ together with the sets containing a single linear factor (but the independent set $\{\ell_1, \ell_2\}$ is not χ-independent). In our example we decomposed $F(x, y)$ into a sum of simple rational functions whose supports were precisely the χ-independent sets of size two. The following result shows that such a decomposition is always possible.

Algorithm 5: SimpleDecomp

Input: Rational function $F(\mathbf{z}) = G(\mathbf{z})/\prod_{j=1}^{m} \ell_j(\mathbf{z})^{q_j}$
Output: Collection $F_1(\mathbf{z}), \dots, F_r(\mathbf{z})$ of rational functions with $\sup(F_j) \subset \sup(F)$ such that $F = F_1 + \cdots + F_r$ and each F_j has independent support

Set $S \leftarrow F(\mathbf{z})$
while *there exists summand* $\tilde{F}$ *of* S *with broken circuit* $\{\ell_{j_1}, \dots, \ell_{j_{s-1}}\}$ *in its support* **do**
  apply (8.38) to $\tilde{F}$ with $j_s > j_{s-1}$ such that $\{\ell_{j_1}, \dots, \ell_{j_s}\}$ is dependent
  replace $\tilde{F}$ by the result of the previous step in the sum S
end

Proposition 8.5 *Let $\ell_1(\mathbf{z}), \dots, \ell_m(\mathbf{z})$ be any linear functions. Then the set of rational functions*

$$\mathcal{L} = \left\{ \frac{1}{\ell_{j_1}(\mathbf{z}) \cdots \ell_{j_s}(\mathbf{z})} : \{\ell_{j_1}, \dots, \ell_{j_s}\} \text{ is } \chi\text{-independent} \right\}$$

is linearly independent over the complex numbers. Furthermore, the span of the rational functions

$$\left\{ \frac{1}{\ell_{j_1}(\mathbf{z})^{p_1} \cdots \ell_{j_s}(\mathbf{z})^{p_s}} : \{\ell_{j_1}, \dots, \ell_{j_s}\} \text{ is } \chi\text{-independent and } \sum_{i=1}^{s} p_j = M \right\}$$

over the complex numbers contains the inverses of all products of the ℓ_j with M (not necessarily distinct) factors.

Proof Linear independence of the set $\mathcal{L}$ follows from well-known results in the study of hyperplane arrangements [5, Thms. 3.43, 3.126, 5.89].

To prove the second conclusion, suppose $1/\boldsymbol{\ell}^{\mathbf{q}}$ is a rational function whose support contains a broken circuit $\{\ell_{j_1}, \dots, \ell_{j_{s-1}}\}$, with the factors ordered by increasing index. Then by the definition of a broken circuit there exists $j_s > j_{s-1}$ such that $\{\ell_{j_1}, \dots, \ell_{j_{s-1}}, \ell_{j_s}\}$ is linearly independent, and there exist constants a_j such that (8.37) holds. Equation (8.38) with $j = s$ expresses $1/\boldsymbol{\ell}^{\mathbf{q}}$ as a linear combination of rational functions of the form $1/\boldsymbol{\ell}^{\mathbf{q}'}$, where each $\mathbf{q}'$ is lexicographically smaller than $\mathbf{q}$. This process can be repeated on each summand that still contains a broken circuit, continuing until no factor contains a broken circuit. The iteration must terminate as each step yields rational functions whose denominator exponents are lexicographically smaller than the previous step. □

Algorithm 5 implements our proof.

8.3 Asymptotics in Non-Generic Directions

In this section we discuss asymptotics in non-generic directions, under the assumption that F is simple (if not F can be decomposed using Algorithm 5). To begin we express the series coefficients of interest in terms of certain 'negative-moment Gaussian integrals'. Unfortunately, asymptotics of such integrals do not follow from Proposition 5.3 in Chapter 5. We give asymptotics when the integrals to be approximated are one-dimensional.

Example 8.3 Continued (Residue Integrals in Non-Generic Directions)

Consider again the function

$$F(x, y) = \frac{1}{\left(1 - \frac{2x+y}{3}\right)\left(1 - \frac{3x+y}{4}\right)}.$$

The coefficient sequence

$$[x^{r_1 n} y^{r_2 n}]F(x, y) = \frac{1}{(2\pi i)^2} \int_{\mathcal{T}} \frac{1}{\left(1 - \frac{2x+y}{3}\right)\left(1 - \frac{3x+y}{4}\right)} \frac{dxdy}{x^{r_1 n+1} y^{r_2 n+1}} \tag{8.39}$$

is non-generic in any direction $\mathbf{r}$ with $\hat{r}_1 \in \{2/3, 3/4\}$, where its asymptotic behaviour can differ from what happens when $0 < \hat{r}_1 < 2/3$ (exponential decay), $2/3 < \hat{r}_1 < 3/4$ (the limit is the constant 12), or $3/4 < \hat{r}_1 < 1$ (exponential decay, determined by a different critical point).

To be precise, we fix the non-generic direction $\mathbf{r} = (2, 1)$. Since Proposition 8.1 above does not require $\mathbf{r}$ to be generic, if we again define

$$M = \begin{pmatrix} \mathbf{b}^{(1)} \\ \mathbf{b}^{(2)} \end{pmatrix} = \begin{pmatrix} 2/3 & 1/3 \\ 3/4 & 1/4 \end{pmatrix}$$

and make the change of variables $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} - M \begin{pmatrix} p \\ q \end{pmatrix}$, the manipulations of Section 8.2.2 still imply that

$$[x^{2n} y^n]F(x, y) = \frac{12}{(2\pi i)^2} \int_{(\varepsilon,\varepsilon)+i\mathbb{R}^2} \tilde{\omega} \tag{8.40}$$

for any $\varepsilon > 0$ sufficiently small, where

$$\tilde{\omega} = \frac{1}{pq} \frac{dpdq}{(1 + 3p - 4q)^{2n+1}(1 - 9p + 8q)^{n+1}}.$$

That we are in a non-generic direction is reflected in the fact that the critical points $\boldsymbol{\sigma}_1$ and $\boldsymbol{\sigma}_{1,2}$ are equal, and in the fact that the linear combination

$$(-\nabla h_{\mathbf{r}})(\boldsymbol{\sigma}_{1,2}) = 3\,\mathbf{b}^{(1)} + 0\,\mathbf{b}^{(2)} \tag{8.41}$$

contains a zero coefficient. Theorem 8.1 no longer applies: because the coefficient of $\mathbf{b}^{(2)}$ in (8.41) is zero, adding the necessary Cauchy integrals over imaginary fibers to (8.40) in order to reduce to a residue computation in p and q results in an error which grows too quickly.

To see why we get into trouble, note that the exponential growth of $\tilde{\omega}$ with respect to n is captured by the modified height function

$$\tilde{h}(p,q) = -2\log(1+3p-4q) - \log(1-9p+8q),$$

which satisfies $(\nabla\tilde{h})(0,0) = (3,0)$. Thus, for $\varepsilon > 0$ sufficiently small $\tilde{h}(-\varepsilon,\varepsilon)$ is less than $\tilde{h}(0,0) = 0$, meaning we can add the Cauchy integral over the imaginary fiber with basepoint $(-\varepsilon,\varepsilon)$ and introduce only a negligible error. In particular, there exists $\tau \in (0,1)$ such that

$$[x^{2n}y^n]F(x,y) = \frac{12}{(2\pi\mathrm{i})^2}\left[\int_{(\varepsilon,\varepsilon)+\mathrm{i}\mathbb{R}^2}\tilde{\omega} - \int_{(-\varepsilon,\varepsilon)+\mathrm{i}\mathbb{R}^2}\tilde{\omega}\right] + O(\tau^n).$$

Unfortunately, since the second coordinate of $(\nabla\tilde{h})(0,0)$ is zero this argument does not allow us to show $\tilde{h}(\varepsilon,-\varepsilon)$ is less than 0 (which is, in fact, not true). We can therefore take a residue in the p variable, but attempting to take a residue in the q variable would introduce an error growing faster than the sequence of interest. Taking a residue in the p variable gives

$$\begin{aligned}[x^{2n}y^n]F(x,y) &= \frac{6}{\pi\mathrm{i}}\int_{\varepsilon+\mathrm{i}\mathbb{R}}\frac{1}{q}\,\frac{1}{(1-4q)^{2n+1}(1+8q)^{n+1}}\,dq + O(\tau^n)\\ &= \frac{6}{\pi\mathrm{i}}\int_{\varepsilon+\mathrm{i}\mathbb{R}}\frac{1}{q(1-4q)(1+8q)}e^{-n[2\log(1-4q)+\log(1+8q)]}dq + O(\tau^n).\end{aligned}$$

Making the substitution $y = \mathrm{i}q$ then yields

$$[x^{2n}y^n]F(x,y) = \frac{-6}{\pi\mathrm{i}}\int_{\mathbb{R}+\mathrm{i}\varepsilon}\frac{A(y)}{y}e^{-n\phi(y)}dy + O(\tau^n), \tag{8.42}$$

where

$$A(y) = \frac{1}{(1-4\mathrm{i}y)(1+8\mathrm{i}y)}, \qquad \phi(y) = 2\log(1-4\mathrm{i}y) + \log(1+8\mathrm{i}y).$$

The integral in (8.42) would be a standard saddle-point integral, with an asymptotic expansion determined by Proposition 5.3 of Chapter 5, if the factor of y was not present in the denominator of the integrand. In generic directions, when there is a contributing point determining asymptotics which is a root of k of the linear denominator factors we are able to take k successive residues. The existence of the pathological denominator term in (8.42) is a reflection of the fact that $\boldsymbol{\sigma}_1$ lies on the zero set of two linear factors, but we are only able to take one residue.

The general argument is similar to this example. First, we must extend our definition of contributing points to account for non-generic directions.

Definition 8.13 (general contributing points) A point $\mathbf{w} \in \mathcal{V}_{k_1,\dots,k_t}$ is now called *contributing* when

$$(-\nabla h)(\mathbf{w}) = \sum_{j=1}^{t} a_j \mathbf{b}^{(k_j)} \tag{8.43}$$

and each a_j is non-negative (in generic directions each a_j will be strictly positive).

Fix a direction $\mathbf{r}$ and suppose that F admits a unique contributing singularity $\boldsymbol{\sigma}$ of highest height, which lies in some flat $\mathcal{V}_{1,\dots,t}$ and no proper subflat. If $\mathbf{r}$ is not generic but each coefficient in (8.43) is positive (i.e., if $\mathbf{r}$ is not generic because of a contributing point of lower height) then the arguments of Section 8.2 still provide dominant asymptotic behaviour by analyzing the saddle-point integral $I_{\boldsymbol{\sigma}}$ corresponding to $\boldsymbol{\sigma}$. Thus, we may assume

$$(-\nabla h)(\boldsymbol{\sigma}) = a_1 \mathbf{b}^{(1)} + \cdots + a_t \mathbf{b}^{(s)} + 0\,\mathbf{b}^{(s+1)} + \cdots + 0\,\mathbf{b}^{(t)}$$

for some $s < t$ and $a_j > 0$. Recall the matrix M with rows $\mathbf{b}^{(1)}, \dots, \mathbf{b}^{(t)}, \mathbf{e}^{(t+1)}, \dots, \mathbf{e}^{(d)}$, where we assume that the variables and ℓ_j have been permuted so that M has full rank, and let

$$\tilde{G}(\mathbf{z}) = \frac{G(\mathbf{z})}{z_1 \cdots z_d \prod_{j>s} \ell_j(\mathbf{z})^{p_j}}.$$

For any integer $\kappa > 0$ let $\mathbf{0}_\kappa$ denote the zero vector of length κ. The following result shows that, as in our example, the coefficients of interest are given by an integral of the same form as the generic case, except for an extra monomial in the denominator of the integrand which makes the integrand singular at the origin.

Proposition 8.6 *Under the assumptions of the previous paragraph,*

$$[\mathbf{z}^{n\mathbf{r}}]F(\mathbf{z}) = \frac{n^{p_1+\cdots+p_s-s}\; i^{p_{s+1}+\cdots+p_t}}{|\det M|(2\pi)^{d-s}} \int_{\mathbb{R}^{d-s}+i\varepsilon(\mathbf{1}_{t-s},\mathbf{0})} \frac{R(\mathbf{y})\left(\boldsymbol{\sigma} + iM^{-1}\binom{\mathbf{0}}{\mathbf{y}}\right)^{-n\mathbf{r}}}{y_1^{p_{s+1}}\cdots y_{t-s}^{p_t}} d\mathbf{y}\left(1 + O\left(\frac{1}{n}\right)\right),$$

where $\mathbf{y} = (y_1, \dots, y_{d-s})$ *and*

$$R(\mathbf{y}) = \prod_{j=1}^{s} \frac{v_j(\mathbf{w})^{p_j-1}}{(p_j-1)!} \tilde{G}\left(\boldsymbol{\sigma} - M^{-1}\mathbf{w}\right)\Bigg|_{\substack{w_j=0,\ 1\le j\le s \\ w_j=-iy_{j-s},\ s+1\le j\le d}}$$

for $\mathbf{v}(\mathbf{w}) = \left(\frac{r_1}{\sigma_1-(M^{-1}\mathbf{w})_1} \;\cdots\; \frac{r_d}{\sigma_d-(M^{-1}\mathbf{w})_d}\right)\cdot M^{-1}$.

Proof Since $\boldsymbol{\sigma}$ is the unique contributing point of highest height, Proposition 8.1 implies that the coefficients of interest are given, up to an exponentially negligible error term, by

$$I = \frac{1}{(2\pi i)^d} \int_{\boldsymbol{\sigma}-\varepsilon\mathbf{m}+i\mathbb{R}^d} \frac{\tilde{G}(\mathbf{z})}{\ell_1(\mathbf{z})^{p_1}\cdots\ell_t(\mathbf{z})^{p_t}} \frac{d\mathbf{z}}{\mathbf{z}^{n\mathbf{r}}}$$

for $\mathbf{m} = M^{-1}\left(\mathbf{e}^{(1)} + \cdots + \mathbf{e}^{(t)}\right)$. Substituting $\mathbf{w} = M(\boldsymbol{\sigma} - \mathbf{z})$ gives

$$I = \frac{1}{|\det M|(2\pi i)^d} \int_{\varepsilon(\mathbf{1}_t,\mathbf{0})+i\mathbb{R}^d} \frac{\tilde{G}(\boldsymbol{\sigma} - M^{-1}\mathbf{w})}{w_1^{p_1}\cdots w_t^{p_t}\left(\boldsymbol{\sigma} - M^{-1}\mathbf{w}\right)^{n\mathbf{r}}}\, d\mathbf{w}. \tag{8.44}$$

Because $a_1, \ldots, a_s > 0$ in (8.43) we can add or subtract the integrals obtained by replacing the domain of integration in (8.44) by any of the $2^s - 1$ imaginary fibers with basepoints $(\pm\mathbf{1}_s, \mathbf{1}_{t-s}, \mathbf{0}) \neq (\mathbf{1}_t, \mathbf{0})$ and introduce only an exponentially negligible error. Taking a signed sum of these integrals corresponds to taking the residue of the integrand of I in $w_1, \ldots, w_s$ at the origin. As in the previous sections of this chapter, this residue is a polynomial of degree $p_1 + \cdots + p_s - s$, whose leading term after substituting $y_j = iw_{j+s}$ for $1 \le j \le d - s$ becomes $R(\mathbf{y})$. □

Determining asymptotics in a non-generic direction thus reduces to a study of integrals of the form

$$\int_{\mathbb{R}^b + i(\mathbf{1}_a,\mathbf{0})} \frac{A(\mathbf{y})}{y_1^{m_1}\cdots y_a^{m_a}} e^{-n\phi(\mathbf{y})} d\mathbf{y}$$

where $\mathbf{m} \in \mathbb{N}^a$, $a < b$, and A and ϕ are analytic at the origin. Here we study the case $t = d$ and $s = d - 1$, where Proposition 8.6 reduces to a one-dimensional integral of the form

$$\frac{i^{p_d}}{|\det M|(2\pi)} \int_{\mathbb{R}+i\varepsilon} \frac{A(y)}{y^{p_1}} e^{-n\phi(y)}.$$

The following result states that A and ϕ can be replaced by the leading terms of their power series expansions at the origin while introducing an error term which grows slower than the sequence of interest.

Lemma 8.5 *Let* $I = \int_{\mathbb{R}+i\varepsilon} y^{-k} A(y) e^{-n\phi(y)}$ *for some* $\varepsilon > 0$ *and* $r, k \in \mathbb{N}$. *If*

- *A is a rational function with no zeroes in the strip $\mathbb{R} + (-2\varepsilon, 2\varepsilon)i \subset \mathbb{C}$;*
- *$\phi(y)$ is a finite sum of terms $r_j \log(p_j + iq_j y)$ for real $r_j > 0$ and $p_j, q_j \neq 0$;*
- *$A(y) = a + O(y)$ and $\phi(y) = by^2 + O(y^3)$ at the origin, for some $b > 0$;*
- *$\phi'(y) = 0$ only if $y = 0$,*

then

$$I = a \int_{\mathbb{R}+i\varepsilon} y^{-k} e^{-nby^2} dy + o\left(n^{-(k+1)/2}\right). \tag{8.45}$$

Problem 8.4 asks you to prove Lemma 8.5, and Problem 8.5 asks you to determine asymptotics for the integral in (8.45). Combining Proposition 8.6 and Lemma 8.5 then gives the following.

Theorem 8.3 *Suppose $F(\mathbf{z})$ is simple and $\mathbf{r}$ is a non-generic direction with a unique contributing singularity $\boldsymbol{\sigma}$ of maximal height, lying on the flat $\mathcal{V}_{1,\ldots,d}$. If*

$$-(\nabla h_{\mathbf{r}})(\boldsymbol{\sigma}) = a_1\,\mathbf{b}^{(1)} + \cdots + a_{d-1}\,\mathbf{b}^{(d-1)} + 0\,\mathbf{b}^{(d)}$$

with each $a_j > 0$ *then as* $n \to \infty$

$$[\mathbf{z}^{n\mathbf{r}}]F(\mathbf{z}) = \boldsymbol{\sigma}^{-\mathbf{r}}\, n^{p_1+\cdots+p_{d-1}+(p_d+1)/2-d}\,(C + o(1)),$$

where

$$C = \prod_{j=1}^{d-1} \frac{a_j^{p_j-1}}{(p_j-1)!}\,\frac{G(\boldsymbol{\sigma})}{\prod_{j>d}\ell_j(\boldsymbol{\sigma})^{p_j}}\,\frac{(\mathbf{q}^T\mathbf{q}/2)^{(p_d-1)/2}}{2(\sigma_1\cdots\sigma_d)\,|\det M|\,\Gamma\left(\frac{p_d+1}{2}\right)},$$

M is the matrix with rows $\mathbf{b}^{(1)},\ldots,\mathbf{b}^{(d)}$*, and* $\mathbf{q}$ *is the rightmost column of the matrix*

$$Q = \begin{pmatrix} \sqrt{r_1}/\sigma_1 & 0 & \mathbf{0} & 0 \\ 0 & \sqrt{r_2}/\sigma_2 & \mathbf{0} & 0 \\ \mathbf{0} & \mathbf{0} & \ddots & \mathbf{0} \\ 0 & 0 & \mathbf{0} & \sqrt{r_d}/\sigma_d \end{pmatrix} M^{-1}.$$

In the simple pole case, when $\mathbf{p} = \mathbf{1}$*, the leading asymptotic term is half what it would be if* $\mathbf{r}$ *was generic (i.e., if* $-(\nabla h_{\mathbf{r}})(\boldsymbol{\sigma})$ *was a positive linear combination of* $\mathbf{b}^{(1)},\ldots,\mathbf{b}^{(d)}$*).*

Proof Under these hypotheses Proposition 8.6 implies the coefficients of interest are determined up to a factor of $1 + O(1/n)$ by a univariate integral of the form

$$I = \boldsymbol{\sigma}^{-\mathbf{r}}\, n^{p_1+\cdots+p_d-d}\int_{\mathbb{R}+\mathrm{i}\varepsilon}\frac{A(y)}{y^{p_d}}e^{-\psi(y)}dy,$$

where rational $A(y)$ and log-linear $\psi(y)$ are analytic functions at the origin with

$$A(y) = \tilde{G}(\boldsymbol{\sigma})\prod_{j=1}^{d-1}\left(\begin{pmatrix}\frac{r_1}{\sigma_1} & \cdots & \frac{r_d}{\sigma_d}\end{pmatrix}\cdot M^{-1}\right)_j^{p_j-1} + O(y) = \tilde{G}(\boldsymbol{\sigma})\prod_{j=1}^{d-1} a_j^{p_j-1} + O(y)$$

$$\psi(y) = \log\left(\frac{\left(\boldsymbol{\sigma} + \mathrm{i}M^{-1}\begin{pmatrix}\mathbf{0}\\ y\end{pmatrix}\right)^{\mathbf{r}}}{\boldsymbol{\sigma}^{\mathbf{r}}}\right) = (\mathbf{q}^T\mathbf{q}/2)y^2 + O(y^3).$$

Lemma 8.5 implies we can replace A and ψ by their leading asymptotics terms, and Problem 8.5 provides asymptotics of the resulting integral. □

Example 8.3 Continued (Asymptotics in Non-Generic Directions)

Above we reduced coefficient asymptotics of

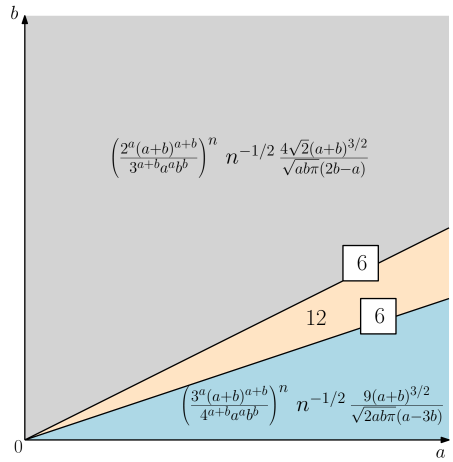


Fig. 8.11 The three asymptotic regimes for the coefficient asymptotics $[x^{an}y^{bn}]F(x,y)$ of $F(x,y) = 1\Big/\left(1-\frac{2x+y}{3}\right)\left(1-\frac{3x+y}{4}\right)$, with asymptotics in the non-generic directions filled in.

$$F(x,y) = \frac{1}{\left(1-\frac{2x+y}{3}\right)\left(1-\frac{3x+y}{4}\right)}$$

in the direction $\mathbf{r} = (2,1)$ to (8.42). Replacing the analytic functions in this integral by their leading terms gives

$$[x^{2n}y^n]F(x,y) = \frac{-6}{\pi \mathrm{i}}\int_{\mathbb{R}+\mathrm{i}\varepsilon} \frac{e^{-48ny^2}}{y}dt\left(1+O\left(\frac{1}{n}\right)\right) = 6 + O\left(\frac{1}{n}\right).$$

In the simple pole case, the fact that the leading asymptotic term in Theorem 8.3 is half what it would be in the non-generic case was noted by Pemantle and Wilson [6]. Additional information on asymptotics in non-generic directions, including results on how asymptotics transition between different asymptotic regimes as $\mathbf{r}$ varies around a non-generic direction, can be found in Baryshnikov et al. [1].

Problems

8.1 Which of the plots in Figure 8.12 contains contributing singularities that could actually arise in an analysis of a multivariate rational function whose denominator is the product of real linear factors?

8.2 Consider the Laurent expansion of the rational function

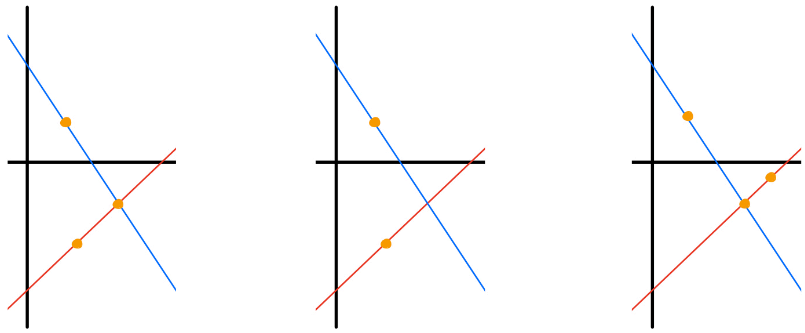


Fig. 8.12 A hyperplane arrangement with three potential arrangements of contributing singularities; not all of these arrangements can actually arise.

$$F(x,y) = \frac{1}{\left(1 - \frac{2x+y}{3}\right)\left(1 - \frac{3x+y}{4}\right)}$$

defined by

$$f_{i,j} = \frac{1}{(2\pi \mathrm{i})^2} \int_{|x|=1/5,\,|y|=3} F(x,y) \frac{dxdy}{x^{i+1}y^{j+1}}.$$

Show that this Cauchy integral can be expressed as a signed sum of Cauchy integrals over the imaginary fibers with basepoints $(\pm 1/5, \pm 3)$. If $\mathbf{r}$ is a direction with positive coordinates, express the coefficients of interest as a sum of Cauchy integrals over linking tori and imaginary fibers in unbounded components of $\mathcal{M}_{\mathbb{R}}$; determine asymptotics in the direction $\mathbf{r}$. Determine asymptotics in directions $\mathbf{r}$ with potentially negative coordinates (note that Cauchy integrals over imaginary fibers with basepoints in unbounded regions of $\mathcal{M}_{\mathbb{R}}$ may no longer be zero).

8.3 Let $F(\mathbf{z})$ be a simple rational function and fix both $\mathbf{r} \in \mathbb{R}^d_{>0}$ and a convergent Laurent expansion $F(\mathbf{z}) = \sum_{\mathbf{i}\in\mathbb{Z}^d} f_{\mathbf{i}}\mathbf{z}^{\mathbf{i}}$ in an open domain containing a point $\mathbf{x} \in \mathbb{R}^d$. Prove that the Cauchy integral expression

$$f_{n\mathbf{r}} = \frac{1}{(2\pi \mathrm{i})^d} \int_{T(\mathbf{x})} \frac{F(\mathbf{z})}{\mathbf{z}^{n\mathbf{r}+\mathbf{1}}} d\mathbf{z}$$

can be decomposed as a sum of integrals over imaginary fibers. Deduce a generalization of Theorem 8.1 for this situation. What changes if $\mathbf{r}$ has negative coordinates?

8.4 For $n > 0$ let $\mathcal{A}_n = (-n^{-2/5}, n^{2/5})$ denote a shrinking real interval centred at the origin and $\varepsilon_n = 1/\sqrt{n}$. Let $A(y)$ and $\phi(y)$ be as in Lemma 8.5.

1. Prove that for any fixed $\kappa > 0$ the integral $\int_{\mathbb{R}\setminus(-\kappa,\kappa)+\mathrm{i}\varepsilon_n} y^{-k} A(y) e^{-n\phi(y)} dy$ decays faster than any fixed power of n as $n \to \infty$.

2. Prove that $\int_{\mathbb{R}\setminus\mathcal{A}_n+\mathrm{i}\varepsilon_n} y^{-k}A(y)e^{-n\phi(y)}dy$ decays faster than any fixed power of n by bounding $|A(y)|$ and the real part of $\phi(y)$ for y near the origin using the leading terms of their power series expansions at the origin.

3. Prove that $\int_{\mathcal{A}_n+\mathrm{i}\varepsilon_n} \left|y^{-k}A(y)e^{-n\phi(y)} - ay^{-k}e^{-bny^2}\right| dy = O\left(n^{-(k+1)/2-1/5}\right)$.

4. Prove that $\int_{\mathbb{R}\setminus\mathcal{A}_n+\mathrm{i}\varepsilon_n} y^{-k}e^{-ny^2}dy$ decays faster than any fixed power of n.

5. Conclude that Lemma 8.5 holds.

8.5 For any $\kappa \in \mathbb{N}$ and $b > 0$ define the integral

$$I_\kappa(b) = \int_{\mathbb{R}+\mathrm{i}\varepsilon} \frac{e^{-bt^2}}{t^\kappa}dt.$$

By differentiating under the integral sign, show that I_κ satisfies the recursion $I_\kappa(b) = -\int I_{\kappa-2}(b)\, db$ for all $\kappa \geq 2$. After finding $I_0(b)$ and $I_1(b)$, prove that

$$I_\kappa(b) = \frac{(-\mathrm{i})^\kappa b^{(\kappa-1)/2}\pi}{\Gamma\left(\frac{\kappa+1}{2}\right)},$$

where $\Gamma(z)$ is the Euler gamma function.

8.6 Determine asymptotics of the bivariate generating function

$$F(x, y) = \frac{1}{(1-\rho_{1,1}x-\rho_{2,1}y)(1-\rho_{1,2}x-\rho_{2,2}y)}$$

for a closed multiclass queuing network with no infinite servers. How does the asymptotic behaviour depend on the weights $\rho_{i,j} > 0$? For what values of the weights are all directions $\mathbf{r} \in \mathbb{R}^2_{>0}$ generic?

8.7 Let A be an $m \times n$ matrix and B be a $n \times m$ matrix, and let I_n and I_m denote the $n \times n$ and $m \times m$ identity matrices, respectively. Prove *Sylvester's determinant identity,* which states that $\det(I_n + BA) = \det(I_m + AB)$, then show $\det(zI_n + BA) = z^{n-m}\det(zI_m + AB)$. *Hint:* If $\mathbf{0}_{n\times m}$ is the $n\times m$ zero matrix, direct calculation implies

$$\begin{pmatrix} I_m & -A \\ B & I_n \end{pmatrix}\begin{pmatrix} I_m & A \\ \mathbf{0}_{n\times m} & I_n \end{pmatrix} = \begin{pmatrix} I_m & A \\ \mathbf{0}_{n\times m} & I_n \end{pmatrix}\begin{pmatrix} I_m & -A \\ B & I_n \end{pmatrix}.$$

8.8 Let C be an $n\times n$ matrix and for any $1 \leq k \leq n$ let $\mathcal{S}_k$ denote the collection of k element subsets of $\{1, \dots, n\}$. Prove that for any $1 \leq k \leq n$ the coefficient of z^{n-k} in $\det(zI_n + C)$ equals

$$\sum_{S\in\mathcal{S}_k} \det\left(C_{S,S}\right),$$

where $C_{S,S}$ denotes the $k\times k$ submatrix of C consisting of its rows and columns with indices in S.

8.9 Let A be an $m \times n$ matrix and B be a $n \times m$ matrix, where $n \geq m$. Use the results of Problems 8.7 and 8.8 to prove the *Cauchy-Binet formula:*

$$\det(AB) = \sum_{S \in \mathcal{S}_m} \det\left(A_{S,[m]}\right) \det\left(B_{[m],S}\right),$$

where $[m] = \{1, \ldots, m\}$ and for any matrix M the notation $M_{P,Q}$ refers to the submatrix of M whose rows have indices in the set P and whose columns have indices in the set Q.

References

1. Y. Baryshnikov, S. Melczer, and R. Pemantle. Asymptotics of multivariate sequences IV: generating functions with poles on a hyperplane arrangement. In preparation, 2020.
2. Andrea Bertozzi and James McKenna. Multidimensional residues, generating functions, and their application to queueing networks. *SIAM Rev.*, 35(2):239–268, 1993.
3. Yaakov Kogan. Asymptotic expansions for large closed and loss queueing networks. *Math. Probl. Eng.*, 8(4-5):323–348, 2002.
4. Yaakov Kogan and Andrei Yakovlev. Asymptotic analysis for closed multichain queueing networks with bottlenecks. *Queueing Systems Theory Appl.*, 23(1–4):235–258, 1996.
5. Peter Orlik and Hiroaki Terao. *Arrangements of hyperplanes*, volume 300 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1992.
6. Robin Pemantle and Mark C. Wilson. Asymptotic expansions of oscillatory integrals with complex phase. In *Algorithmic probability and combinatorics*, volume 520 of *Contemp. Math.*, pages 221–240. Amer. Math. Soc., Providence, RI, 2010.

Chapter 9
Multiple Points and Beyond

Only geometry can give us the thread which will lead us through the labyrinth of the continuum's composition, the maximum and the minimum, the infinitesimal and the infinite, and no one will arrive at a solid metaphysic except he who has passed through it.
— Gottfried Wilhelm von Leibniz

In this chapter we detail more general aspects of ACSV. To begin, we combine techniques from Chapter 5 (ACSV in the smooth case) and Chapter 8 (ACSV when the singular set is a hyperplane arrangement) to study functions whose singular sets are *locally* hyperplane arrangements; asymptotics are determined through the use of multivariate residues, leading to Theorems 9.1 and 9.2. Following this we survey current aspects of ACSV, including new work making use of tools from topology, leading up to Theorem 9.3 on the asymptotic contributions of (potentially non-minimal) critical points. Combining these topological methods with the computational results for D-finite functions discussed in Chapter 2 gives a powerful approach to the connection problem for asymptotics of sequences satisfying linear recurrence relations with polynomial coefficients.

9.1 Local Geometry of Algebraic and Analytic Varieties

Before studying rational (and meromorphic) generating functions with more complicated singular varieties, we need a better understanding of the 'local' geometry that can arise. In order to talk about local geometry, we introduce the following notion.

Definition 9.1 (local rings) Given $\mathbf{w} \in \mathbb{C}^d$ the *local ring at* $\mathbf{w}$, denoted $\mathcal{O}_\mathbf{w}$, is the ring of power series which converge in some neighbourhood of $\mathbf{w}$ with the usual addition and multiplication. Equivalently, the local ring is given by complex functions $f(\mathbf{z})$ which are analytic at $\mathbf{w}$, modulo the equivalence relation $f \sim g$ if $f = g$ on some neighbourhood of $\mathbf{w}$. An element $f \in \mathcal{O}_\mathbf{w}$ is a *unit* if it does not vanish at $\mathbf{w}$ (so $1/f$ is also in $\mathcal{O}_\mathbf{w}$) and *irreducible* if $f = gh$ implies one of g or h is a unit. Two irreducible elements $f, g \in \mathcal{O}_\mathbf{w}$ are *coprime* if there does not exist a unit $u \in \mathcal{O}_\mathbf{w}$ with $f = ug$.

Hörmander [17, Thm. 6.2.2] shows that for any $\mathbf{w} \in \mathbb{C}^d$ the local ring $\mathcal{O}_\mathbf{w}$ forms a unique factorization domain, meaning any element of $\mathcal{O}_\mathbf{w}$ can be written as a product of irreducible elements and this representation is unique up to the ordering of factors and multiplication by units.

Remark 9.1 If $f(\mathbf{z}) = g(\mathbf{z})h(\mathbf{z})$ for analytic functions g and h which vanish at $\mathbf{w}$ then the product rule implies any partial derivative of f vanishes at $\mathbf{w}$. Thus, a sufficient (but not necessary) condition for an element of $\mathcal{O}_{\mathbf{w}}$ represented by the analytic function $f(\mathbf{z})$ to be irreducible is that the gradient $(\nabla f)(\mathbf{w})$ is non-zero.

Suppose we want to determine coefficient asymptotics of a rational function $F(\mathbf{z}) = G(\mathbf{z})/H(\mathbf{z})$. The most explicit results in this chapter will hold when the denominator $H(\mathbf{z})$ locally factors into irreducibles whose zero sets have nice geometry.

Definition 9.2 (factorizations and multiple points) Let $\mathbf{w} \in \mathbb{C}^d$ and $H(\mathbf{z})$ be an analytic function at $\mathbf{w}$. We call an expression of the form

$$H(\mathbf{z}) = u(\mathbf{z})H_1(\mathbf{z})^{p_1} \cdots H_s(\mathbf{z})^{p_s} \tag{9.1}$$

a *factorization of* H *in* $\mathcal{O}_{\mathbf{w}}$ if u and the H_j are analytic in some polydisk $\mathcal{D}$ centred at $\mathbf{w}$, each p_j is a positive integer, and (9.1) holds for all $\mathbf{z} \in \mathcal{D}$. By convention, when writing a factorization of the form (9.1) we always assume that $u(\mathbf{w}) \neq 0$ while each $H_j(\mathbf{w}) = 0$. If the $H_j(\mathbf{z})$ are irreducible in $\mathcal{O}_{\mathbf{w}}$ and pairwise coprime then (9.1) is called a *square-free factorization of* H *in* $\mathcal{O}_{\mathbf{w}}$. When H has a square-free factorization in $\mathcal{O}_{\mathbf{w}}$ with each exponent $p_j = 1$ then H is said to be *square-free* at $\mathbf{w}$.

We call $\mathbf{w}$ a *multiple point of* H if H has a factorization in $\mathcal{O}_{\mathbf{w}}$ of the form (9.1) where the gradient vectors $(\nabla H_1)(\mathbf{w}), \dots, (\nabla H_s)(\mathbf{w})$ are all non-zero. If, in addition, the gradient vectors $(\nabla H_1)(\mathbf{w}), \dots, (\nabla H_s)(\mathbf{w})$ are linearly-independent then $\mathbf{w}$ is called a *transverse (multiple) point*.

Remark 9.2 Recall from previous chapters that the zero set of a complex-valued function f is denoted $\mathcal{V}(f)$, that the *square-free part* of a polynomial $H(\mathbf{z}) \in \mathbb{C}[\mathbf{z}]$, denoted $H^{\mathrm{s}}(\mathbf{z})$, is the product of its distinct irreducible factors, and that the *smooth points* of $\mathcal{V}(H)$ are the points where the gradient $(\nabla H^{\mathrm{s}})(\mathbf{z})$ is non-zero. Checking definitions shows that any smooth point of $\mathcal{V}(H)$ is a transverse point.

Remark 9.3 Given an open set $U \subset \mathbb{C}^d$ and analytic function $f(\mathbf{z})$ on U, let $\mathcal{V}_U(f)$ denote the elements $\mathbf{z} \in U$ such that $f(\mathbf{z}) = 0$. If (9.1) gives a factorization of $H(\mathbf{z})$ in $\mathcal{O}_{\mathbf{w}}$ then there exists some neighbourhood U of $\mathbf{w}$ in $\mathbb{C}^d$ such that $H, H_1, \dots, H_s$ are analytic on U and

$$\mathcal{V}_U(H) = \mathcal{V}_U(H_1) \cup \cdots \cup \mathcal{V}_U(H_s).$$

Thus, if $\mathbf{w}$ is a multiple point of $H(\mathbf{z})$ then the points of $\mathcal{V}(H)$ contained in some neighbourhood of $\mathbf{w}$ form a finite union of smooth sets, each defined by the vanishing of a single analytic function. If $\mathbf{w}$ is a transverse point then the tangent spaces of these smooth sets at $\mathbf{w}$ are hyperplanes with linearly independent normal vectors.

Example 9.1 (Multiple Points in an Algebraic Set)

Suppose $H(x, y) = H_1(x, y)H_2(x, y)$ where $H_1 = 1 - x^2 + y$ and $H_2 = 1 - x^2 - y^2$, so that $\mathcal{V} = \mathcal{V}(H)$ is the union of the two irreducible algebraic varieties $\mathcal{V}_1 = \mathcal{V}(H_1)$

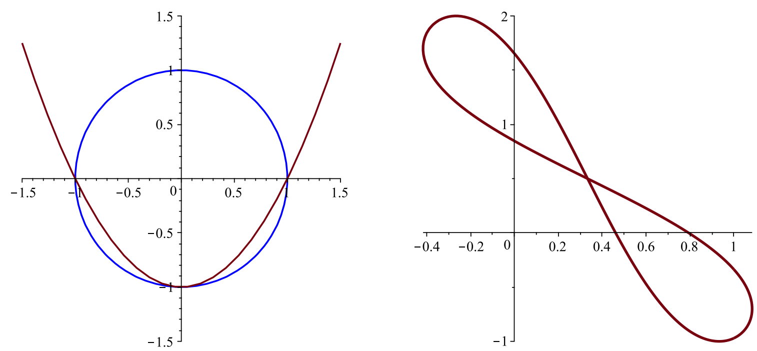


Fig. 9.1 *Left:* The real points in the zero set $\mathcal{V}((1-x^2+y)(1-x^2-y^2)) = \mathcal{V}(1-x^2+y) \cup \mathcal{V}(1-x^2-y^2)$ form the union of a circle and a parabola, which have two transverse points of intersection (on the sides of the circle) and one non-transverse point of intersection (at the bottom of the circle). The transverse points of intersection move smoothly if the parabola or circle are perturbed, but the non-transverse point can disappear or split into two points of intersection when the parabola or circle are perturbed. *Right:* The real points in the zero set of $H(x,y) = (1/3-x)^2(1/2-y)^2 + (5/6-x-y)(11/6-4x-y)$ form a figure-eight pattern.

and $\mathcal{V}_2 = \mathcal{V}(H_2)$; see the left of Figure 9.1. Any element of $\mathcal{V}$ in either $\mathcal{V}_1$ or $\mathcal{V}_2$ but not both is a smooth point, and thus a transverse point. Solving the system

$$H(x,y) = H_x(x,y) = H_y(x,y) = 0$$

gives the three points $\mathcal{V}_1 \cap \mathcal{V}_2 = \{(-1,0),(0,-1),(1,0)\}$ where $\mathcal{V}$ is non-smooth. The gradients ∇H_1 and ∇H_2 never vanish on $\mathcal{V}_1$ and $\mathcal{V}_2$, respectively, so all three points in $\mathcal{V}_1 \cap \mathcal{V}_2$ are multiple points. The gradients $(\nabla H_1)(x,y)$ and $(\nabla H_2)(x,y)$ are linearly independent when $(x,y) = (-1,0)$ and $(x,y) = (1,0)$, which are thus transverse points, but these gradients are linearly *dependent* at $(x,y) = (0,-1)$, which is not a transverse point.

Note that the transverse points vary smoothly with the coefficients of H under small perturbations. On the other hand, by slightly perturbing coefficients it is possible to remove the non-transverse multiple point or have it split into two points of intersection. This gives a geometric interpretation of transverse points: they are the multiple points arising from intersections of smooth sets which are stable under small coefficient changes.

Non-Example 9.2 (Non-Multiple Points)

Consider the polynomial $H(x,y,z) = z^2 - x^2 - y^2$. Problem 9.1 asks you to show that H is irreducible in the local ring $\mathcal{O}_{\mathbf{0}}$, but the gradient of H vanishes at the origin. The zero set $\mathcal{V}(H)$ is not smooth at the origin; its real points form the union of two cones joined to make an hourglass shape.

In practice, when $H(\mathbf{z})$ is a polynomial it often factors into irreducible polynomials whose zero sets behave nicely.

Remark 9.4 Suppose $H(\mathbf{z}) \in \mathbb{C}[\mathbf{z}]$ has a factorization $H(\mathbf{z}) = H_1(\mathbf{z})^{p_1} \cdots H_m(\mathbf{z})^{p_m}$ in $\mathbb{C}[\mathbf{z}]$, where each p_j is a positive integer. Further assume that

(1) whenever $H_j(\mathbf{w}) = 0$ for some $\mathbf{w} \in \mathbb{C}^d$ then $(\nabla H_j)(\mathbf{w})$ is not the zero vector,

(2) if $\mathbf{w} \in \mathbb{C}^d$ is a common root of some factors $H_{k_1}(\mathbf{w}) = \cdots = H_{k_s}(\mathbf{w}) = 0$ then the vectors $(\nabla H_{k_1})(\mathbf{w}), \dots, (\nabla H_{k_s})(\mathbf{w})$ are linearly independent.

Then every point in the zero set $\mathcal{V} = \mathcal{V}(H)$ is a transverse point of H. If only condition (1) holds then every point in $\mathcal{V}$ is a multiple point, but some are not transverse points. If $H_{k_1}, \dots, H_{k_s}$ are the factors of H which vanish at $\mathbf{w} \in \mathcal{V}$ then the factorization

$$H(\mathbf{z}) = \underbrace{\left(\prod_{j \notin \{k_1,\dots,k_s\}} H_j(\mathbf{z})^{p_j} \right)}_{u(\mathbf{z})} H_{k_1}(\mathbf{z})^{p_{k_1}} \cdots H_{k_s}(\mathbf{z})^{p_{k_s}}$$

gives a square-free factorization of H in $\mathcal{O}_{\mathbf{w}}$.

Definition 9.3 (transverse polynomial factorization) A factorization $H(\mathbf{z}) = H_1(\mathbf{z})^{p_1} \cdots H_m(\mathbf{z})^{p_m}$ in $\mathbb{C}[\mathbf{z}]$ satisfying properties (1) and (2) of Remark 9.4 is called a *transverse polynomial factorization* of $H(\mathbf{z})$.

Example 9.3 (A Quadrant Lattice Path Model)

If $\mathcal{A} = \{(-1,1), (0,-1), (1,1)\}$ then Theorem 4.2 in Chapter 4 implies that the generating function for the number of walks in $\mathbb{N}^2$ which begin at the origin and use the steps in $\mathcal{A}$ is the main diagonal of the rational function

$$F(x, y, t) = \frac{(1+x)(1-xy^2+x^2)}{(1-t(1+x^2+xy^2))(1-y)(1+x^2)}.$$

The denominator $H(x, y, t)$ of $F(x, y, t)$ is the product of the polynomials

$$H_1 = 1 - t(1 + x^2 + xy^2), \qquad H_2 = 1 - y, \qquad H_3 = 1 + x^2.$$

When $H_1(x, y, t) = 0$ the derivative $-\partial H_1/\partial t = 1 + x^2 + xy^2 = 1/t \neq 0$, so the gradients of H_1, H_2, and H_3 do not vanish on $\mathcal{V}(H_1), \mathcal{V}(H_2)$, and $\mathcal{V}(H_3)$, respectively. Since the gradients of these polynomials are linearly independent at common intersections of their zero sets, the decomposition $H = H_1 H_2 H_3$ is a transverse polynomial factorization of H.

Not every polynomial has a transverse polynomial factorization. In fact, it is possible for a polynomial $H(\mathbf{z})$ to be irreducible as a complex polynomial but still factor in the local ring at a point. This is one reason why it is important to discuss local rings, even if one only cares about the zero sets of polynomials.

Example 9.4 (A Lemniscate)

Historically significant examples of irreducible plane curves, which will allow for a nice illustration of our theory, are figure-eight shapes known as lemniscates[1]. For instance, the real zeroes of the polynomial

$$L(x,y) = \left(\frac{1}{3} - x\right)^2 \left(\frac{1}{2} - y\right)^2 + \left(\frac{5}{6} - x - y\right)\left(\frac{11}{6} - 4x - y\right)$$

are displayed on the right-hand side of Figure 9.1. Using the quadratic formula it is easy to see that $L(x, y)$ does not factor into linear factors. Solving the system $L = L_x = L_y = 0$ shows that $\varrho = (1/3, 1/2)$ is the only non-smooth point of $\mathcal{V}(L)$. Near ϱ the zero set $\mathcal{V}(H)$ is the union of two zero sets $\mathcal{V}(y - a_1(x)) \cup \mathcal{V}(y - a_2(x))$ where $a_1(x)$ and $a_2(x)$ are analytic at $x = 1/3$ with convergent series expansions

$$a_1(x) = 1/2 - (x - 1/3) - (1/3)(x - 1/3)^3 - (7/27)(x - 1/3)^5 + \cdots$$
$$a_2(x) = 1/2 - 4(x - 1/3) + (16/3)(x - 1/3)^3 - (128/27)(x - 1/3)^5 + \cdots .$$

Since L is quadratic in y we can determine $a_1(x)$ and $a_2(x)$ using the quadratic formula, but finding such expansions can be done for any bivariate function using the Newton polygon method discussed in Section 2.3 of Chapter 2. Because the gradients of $y - a_1(x)$ and $y - a_2(x)$ are linearly independent at ϱ, this is a transverse point. The algorithms of Chapter 7 automatically prove that ϱ is minimal.

We conclude this introductory section by giving a necessary condition for multiple points. Stating this condition requires the following concept.

Definition 9.4 (leading homogeneous terms) Given $\mathbf{w} \in \mathbb{C}^d$ and $H(\mathbf{z})$ analytic at $\mathbf{w}$, the *leading homogeneous term of H at* $\mathbf{w}$ is the sum of the lowest-degree terms appearing in the power series expansion of $H(\mathbf{w} + \mathbf{z})$ at the origin.

If $H(\mathbf{w}) \neq 0$ then the leading homogeneous term of H at $\mathbf{w}$ is the constant $H(\mathbf{w})$.

Example 9.4 Continued (Lemniscate Leading Homogeneous Term)

If $L(x, y)$ is the lemniscate from above then $L(1/3+x, 1/2+y) = 4x^2+5xy+y^2+x^2y^2$ so the leading homogeneous term of L at the non-smooth point $(1/3, 1/2)$ equals

[1] The term lemniscate was coined by Jakob Bernoulli from a Greek word $\lambda\eta\mu\nu\iota\sigma\kappa o\varsigma$ (*lemniskos*) for ribbon; see Schappacher [29] for additional historical context. Our example $L(x, y)$ is obtained from the simple polynomial $x^2y^2+(x+y)(4x+y)$ after making the (somewhat arbitrary) substitution $(x, y) \mapsto (1/3 - x, 1/2 - y)$ so that $1/L(x, y)$ has a power series expansion at the origin.

$$\ell(x, y) = 4x^2 + 5xy + y^2 = (x + y)(4x + y).$$

Lemma 9.1 *Let* $\mathbf{w} \in \mathbb{C}^d$ *and* $H(\mathbf{z})$ *be analytic at* $\mathbf{w}$ *with* $H(\mathbf{w}) = 0$. *Let* $\ell(\mathbf{z})$ *be the leading homogeneous term of* $H(\mathbf{z})$ *at* $\mathbf{w}$. *If* $\mathbf{w}$ *is a multiple point of* H *then* ℓ *factors into linear polynomials.*

Proof We know that $H(\mathbf{z})$ has a factorization in $\mathcal{O}_{\mathbf{w}}$ of the form (9.1), which implies $\ell(\mathbf{z}) = \ell_1(\mathbf{z})^{p_1} \cdots \ell_s(\mathbf{z})^{p_s}$ where ℓ_j is the leading homogeneous term of H_j at $\mathbf{w}$. When $\mathbf{w}$ is a multiple point then each ℓ_j is linear. □

Example 9.5 (Applying Lemma 9.1)

The leading homogeneous term of $H(x, y, z) = w^2 - x^2 - y^2 - z^2$ at the origin is itself, which does not factor into linear polynomials, so Lemma 9.1 implies the origin is not a multiple point of H.

With this knowledge of local geometry in hand, we now turn back to analytic combinatorics in several variables.

9.2 ACSV for Transverse Points

In this section we generalize our previous ACSV methods to cover multivariate functions whose coefficient asymptotics are determined by transverse points. For simplicity we introduce most of our constructions and results for power series expansions of rational functions, noting afterwards how they vary for general Laurent expansions of meromorphic functions.

To that end, let $F(\mathbf{z}) = G(\mathbf{z})/H(\mathbf{z})$ be the ratio of coprime polynomials $G(\mathbf{z}), H(\mathbf{z}) \in \mathbb{C}[\mathbf{z}]$ with power series expansion $F(\mathbf{z}) = \sum_{\mathbf{i} \in \mathbb{N}^d} f_{\mathbf{i}} \mathbf{z}^{\mathbf{i}}$ on the domain of convergence $\mathcal{D} \subset \mathbb{C}^d$. Recall that because G and H are coprime, the set $\mathcal{V} = \mathcal{H}(H)$ forms the singular variety of F (its set of singularities). As in previous chapters, our analysis in a direction $\mathbf{r} \in \mathbb{R}^d_*$ begins with a multivariate Cauchy integral

$$f_{n\mathbf{r}} = [\mathbf{z}^{n\mathbf{r}}]F(\mathbf{z}) = \frac{1}{(2\pi i)^d} \int_{\mathcal{T}} F(\mathbf{z}) \frac{d\mathbf{z}}{\mathbf{z}^{n\mathbf{r}+\mathbf{1}}}, \tag{9.2}$$

where $\mathcal{T}$ is any product of circles in the domain of convergence $\mathcal{D}$. Once again the growth of the Cauchy integrand in (9.2) is captured by the height function

$$h_{\mathbf{r}}(\mathbf{z}) = h(\mathbf{z}) = -\sum_{j=1}^{d} r_j \log |z_j|,$$

which gives a bound on the exponential growth

$$\limsup_{n\to\infty} |f_{n\mathbf{r}}|^{1/n} \leq e^{h_{\mathbf{r}}(\mathbf{w})}$$

for every $\mathbf{w} \in \overline{\mathcal{D}}$, and in particular for every minimal point $\mathbf{w} \in \mathcal{V} \cap \partial\mathcal{D}$ (this bound is established through the same process as (5.15) in Chapter 5). Following the approach of Chapter 5, the first step in our asymptotic analysis is to determine the singularities where this upper bound could be tight.

9.2.1 Critical Points and Stratifications

In Chapter 5 we introduced the smooth critical point equations (5.16) to characterize potential minimizers of the height function $h(\mathbf{z})$ on $\mathcal{V} \cap \partial\mathcal{D}$. Unfortunately, as noted in Chapter 8, any non-smooth point of $\mathcal{V}$ trivially satisfies the smooth critical point equations, so these equations are too weak to properly characterize non-smooth points of interest. In Chapter 8, when the singular variety $\mathcal{V}$ was assumed to form a hyperplane arrangement, we decomposed $\mathcal{V}$ into the union of smooth sets and used the geometry of $\mathcal{V}$ to create a new set of 'hyperplane' critical point equations (8.2). Now that we have a better understanding of local geometry, we can generalize the setup of Chapter 8. Recall the logarithmic gradient map $\nabla_{\log} f = (z_1 f_{z_1}, \ldots, z_d f_{z_d})$ from previous chapters, where f_{z_j} denotes the partial derivative $\partial f / \partial z_j$.

Definition 9.5 (critical points) Let $\mathbf{w} \in \mathbb{C}_*^d$ and $H(\mathbf{z})$ be analytic at $\mathbf{w}$. Fix a direction $\mathbf{r} \in \mathbb{R}_*^d$ and suppose $\mathbf{w}$ is a transverse point where $H(\mathbf{z}) = u(\mathbf{z})H_1(\mathbf{z})^{p_1} \cdots H_s(\mathbf{z})^{p_s}$ is a square-free factorization of H in $\mathcal{O}_{\mathbf{w}}$. If the number of factors $s < d$ then $\mathbf{w}$ is a *critical point in the direction* $\mathbf{r}$ if the $(s+1) \times d$ matrix

$$N_{\mathbf{w}}(H_1, \ldots, H_s) = \begin{pmatrix} -(\nabla_{\log} H_1)(\mathbf{w}) \\ \vdots \\ -(\nabla_{\log} H_s)(\mathbf{w}) \\ \mathbf{r} \end{pmatrix} \tag{9.3}$$

is rank deficient, meaning all $(s+1) \times (s+1)$ minors of $N_{\mathbf{w}}$ vanish. If $s = d$ then $\mathbf{w}$ is always called a *critical point in the direction* $\mathbf{r}$. Since $\mathbf{w}$ is a transverse point, it cannot be the case that $s > d$.

Definition 9.5 generalizes Definition 5.4 in Chapter 5, for critical points in the smooth case, and Definition 8.3 of Chapter 8, for critical points in the hyperplane arrangement setting.

Remark 9.5 Suppose $\mathbf{w}$ is a transverse point of H with square-free factorization as in Definition 9.5. Then the leading homogeneous term of $H(\mathbf{z})$ at $\mathbf{w}$ is $u(\mathbf{w})\ell_1(\mathbf{z})^{p_1} \cdots \ell_s(\mathbf{z})^{p_s}$, where $\ell_j(\mathbf{z})$ is the (linear) leading homogeneous term of $H_j(\mathbf{z})$ at $\mathbf{w}$. In particular, the gradients $(\nabla H_j)(\mathbf{w})$ can be determined implicitly from the evaluations of the partial derivatives of $H(\mathbf{z})$ at $\mathbf{z} = \mathbf{w}$ by computing a power series expansion of $H(\mathbf{z})$ at $\mathbf{z} = \mathbf{w}$ and factoring the leading homogeneous term. When $H(\mathbf{z}) \in \mathbb{Q}[\mathbf{z}]$ then all critical points have algebraic coordinates and the evaluations of the partial derivatives of H and factorization of its leading homogeneous term can be done implicitly [19].

Definition 9.5 characterizes when a *specific* point is critical. In order to determine a procedure to compute *all* critical points, we decompose the singular variety $\mathcal{V}$ into a finite collection of smooth sets such that each smooth set has consistent local geometry. The easiest case occurs when $H(\mathbf{z})$ admits a transverse polynomial factorization. The following generalizes Definition 8.1 in Chapter 8 from the hyperplane arrangement setting to this context.

Definition 9.6 (flats and strata) Suppose $H(\mathbf{z})$ admits a transverse polynomial factorization $H(\mathbf{z}) = H_1(\mathbf{z})^{p_1} \cdots H_m(\mathbf{z})^{p_m}$. For any $S = \{k_1, \ldots, k_s\} \subset \{1, \ldots, m\}$ the *flat* defined by $H_{k_1}, \ldots, H_{k_s}$ is the set $\mathcal{V}_S = \mathcal{V}_{k_1,\ldots,k_s} = \mathcal{V}(H_{k_1}, \ldots, H_{k_s})$ of their common solutions in $\mathbb{C}^d$. The *stratum* defined by S is the flat $\mathcal{V}_S$ minus any other flats it strictly contains,

$$\mathcal{S}_S = \mathcal{V}_S \setminus \bigcup_{\mathcal{V}_T \subsetneq \mathcal{V}_S} \mathcal{V}_T .$$

The *dimension* of the flat $\mathcal{V}_S$, and of the stratum $\mathcal{S}_S$, is the value of $d - |S|$.

Remark 9.6 Because we restrict to transverse polynomial factorizations there are no non-empty flats defined by subsets S of size $|S| > d$, meaning the dimension of any non-empty flat is a natural number. Any zero-dimensional flat, defined by a subset S with $|S| = d$, is a finite set. Further properties of the dimension of algebraic sets, including the general definition of dimension, can be found in Mumford [24, Ch. 1].

The singular variety $\mathcal{V}$ is the (finite and disjoint) union of the strata $\mathcal{S}_S$ as S runs over the subsets of $\{1, \ldots, m\}$. The critical points on a stratum can be characterized by one set of algebraic equations.

Definition 9.7 (transverse critical points equations) Suppose $H(\mathbf{z})$ admits a transverse polynomial factorization $H(\mathbf{z}) = H_1(\mathbf{z})^{p_1} \cdots H_m(\mathbf{z})^{p_m}$ and fix a direction $\mathbf{r} \in \mathbb{R}^d_*$. For any $S = \{k_1, \ldots, k_s\} \subset \{1, \ldots, m\}$ let $N_{\mathbf{z}}(H_{k_1}, \ldots, H_{k_s})$ denote the matrix in (9.3). The system of polynomial equalities and inequalities

$$\begin{aligned} H_{k_j}(\mathbf{z}) &= 0, & & j = 1, \ldots, s \\ \det(M) &= 0, & & M \text{ is an } (s+1) \times (s+1) \text{ minor of } N_{\mathbf{z}}(H_{k_1}, \ldots, H_{k_s}) \\ z_j &\neq 0, & & j = 1, \ldots, d \\ H_i(\mathbf{z}) &\neq 0, & & i \in \{1, \ldots, m\} \setminus \{k_1, \ldots, k_s\} \end{aligned} \tag{9.4}$$

form the *critical point equations for the stratum $\mathcal{S}_S$ in the direction* $\mathbf{r}$.

If $H(\mathbf{z})$ admits a transverse polynomial factorization then the solutions of the critical point equations over all strata of $\mathcal{V}$ give all critical points. We discuss notions of critical points beyond transverse points in Section 9.3 below.

Remark 9.7 (an optional perspective from differential geometry) Fix a stratum $\mathcal{S}$ and let $\mathcal{S}_*$ denote the submanifold of points in $\mathcal{S}$ whose coordinates are non-zero. As in Remark 5.9 of Chapter 5, we may view the height function h as a smooth map of manifolds from $\mathcal{S}_*$ to $\mathbb{R}$. Then the 'critical points' in the usual sense of differential geometry (points where the differential of h restricted to $\mathcal{S}_*$ vanishes) match our definition of critical points. See Section 9.3 below for more details.

Example 9.3 Continued (Lattice Path Critical Points)

Consider again the lattice path model in $\mathbb{N}^2$ defined by the set of steps $\mathcal{A} = \{(-1,1),(0,-1),(1,1)\}$. As noted above, the generating function of this model is the main diagonal of a rational function whose denominator has a transverse polynomial factorization with factors $H_1 = 1-t(1+x^2+xy^2)$, $H_2 = 1-y$, and $H_3 = 1+x^2$. To determine critical points in the $\mathbf{r} = (1,1,1)$ direction

- on the stratum $\mathcal{S}_1$, we build the matrix

$$N = \begin{pmatrix} -\nabla_{\log} H_1 \\ \mathbf{1} \end{pmatrix} = \begin{pmatrix} tx(y^2+2x) & 2txy^2 & t(1+x^2+xy^2) \\ 1 & 1 & 1 \end{pmatrix}$$

  and determine the points in $\mathcal{S}_1$ where H_1 and the maximal minors of N vanish. This recovers the smooth critical point equations from Chapter 5,

$$H_1 = 0, \qquad x(\partial H_1/\partial x) = y(\partial H_1/\partial y) = t(\partial H_1/\partial t),$$

  subject to the condition $(1-y)(1+x^2) \neq 0$. There are four solutions, given by $(\omega^2, \omega\sqrt{2}, 1/4)$ for $\omega \in \{\pm 1, \pm \mathrm{i}\}$, which are all smooth critical points. None of these critical points are minimal, as they have y-coordinate of modulus $\sqrt{2}$ and the denominator of $F(x,y,t)$ contains $H_2 = 1-y$ as a factor.

- on the stratum $\mathcal{S}_{1,2}$, we compute the matrix

$$N = \begin{pmatrix} -\nabla_{\log} H_1 \\ -\nabla_{\log} H_2 \\ \mathbf{1} \end{pmatrix} = \begin{pmatrix} tx(y^2+2x) & 2txy^2 & t(1+x^2+xy^2) \\ 0 & y & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

  and solve $H_1 = H_2 = \det N = 0$ subject to $1 + x^2 \neq 0$. This gives two critical points, $\boldsymbol{\sigma} = (1,1,1/3)$ and $(-1,1,1)$, of which the second is not minimal since it has larger coordinate-wise modulus than $\boldsymbol{\sigma}$.

- on the stratum $\mathcal{S}_{1,3}$, we compute the matrix

$$N = \begin{pmatrix} -\nabla_{\log} H_1 \\ -\nabla_{\log} H_3 \\ \mathbf{1} \end{pmatrix} = \begin{pmatrix} tx(y^2+2x) & 2txy^2 & t(1+x^2+xy^2) \\ -2x & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

and solve $H_1 = H_3 = \det N = 0$ subject to $1 - y \neq 0$. This system of equations has no solutions, so $\mathcal{S}_{1,3}$ contains no critical points.

- on the stratum $\mathcal{S}_{1,2,3}$, we note that both points $(\mathrm{i}, \mathrm{i}, -\mathrm{i})$ and $(-\mathrm{i}, -\mathrm{i}, \mathrm{i})$ in the stratum are critical points by definition, but neither is minimal.

- on the strata $\mathcal{S}_2, \mathcal{S}_3$, and $\mathcal{S}_{2,3}$, we note that $H_2 = 1 - y$ and $H_3 = 1 + x^2$ are independent of the variable t, so $\mathbf{r} = (1, 1, 1)$ can never be in the span of $\nabla_{\log} H_2$ and $\nabla_{\log} H_3$, and there are no critical points (alternatively one can compute matrix minors as above to obtain systems with no solutions).

Problem 9.2 below asks the reader to prove that $\boldsymbol{\sigma} = (1, 1, 1/3)$ is a minimal point, but $\boldsymbol{\sigma}$ is not strictly minimal since any point $(x, 1, t)$ with $|x| = 1$ and $|t| = 1/3$ lies in $\mathcal{V}(H)$ and has the same coordinate-wise modulus as $\boldsymbol{\sigma}$.

As in previous chapters, not all critical points contribute to dominant asymptotics. To find the critical points that will appear in our analysis we introduce the following notion, generalizing Definition 5.5 in Chapter 5 and Definition 8.8 in Chapter 8.

Definition 9.8 (contributing points) Let $\mathbf{w} \in \mathbb{C}_*^d$ be a transverse point of $H(\mathbf{z})$ and let $H(\mathbf{z}) = u(\mathbf{z}) H_1(\mathbf{z})^{p_1} \cdots H_s(\mathbf{z})^{p_s}$ be a square-free factorization of H in $\mathcal{O}_{\mathbf{w}}$. For each $1 \leq j \leq s$ let $k_j \in \{1, \ldots, d\}$ be an index such that the partial derivative $(\partial H_j / \partial z_{k_j})(\mathbf{w}) \neq 0$. Proposition 3.6 in Chapter 3 implies the vector

$$\mathbf{v}_j = \frac{(\nabla_{\log} H_j)(\mathbf{w})}{w_{k_j} (\partial H_j / \partial z_{k_j})(\mathbf{w})} = \left(\frac{w_1 (\partial H_j / \partial z_1)(\mathbf{w})}{w_{k_j} (\partial H_j / \partial z_{k_j})(\mathbf{w})}, \ldots, \frac{w_d (\partial H_j / \partial z_d)(\mathbf{w})}{w_{k_j} (\partial H_j / \partial z_{k_j})(\mathbf{w})} \right)$$

has real coordinates. The *normal cone* of H at $\mathbf{w}$ is the set

$$N(\mathbf{w}) = \left\{ \sum_{j=1}^{s} a_j \mathbf{v}_j : a_j > 0 \right\} \subset \mathbb{R}^d.$$

The point $\mathbf{w}$ is called a *contributing singularity*, or *contributing point*, for the direction $\mathbf{r} \in \mathbb{R}_*^d$ if $\mathbf{r} \in N(\mathbf{w})$. Since the gradients of the $H_j(\mathbf{z})$ at $\mathbf{z} = \mathbf{w}$ are linearly independent the vectors $\mathbf{v}_j$ are also linearly independent, thus if $\mathbf{r} = a_1 \mathbf{v}_1 + \cdots a_s \mathbf{v}_s$ then the a_j are uniquely determined.

Remark 9.8 If $\mathbf{w}$ is a contributing point for the direction $\mathbf{r}$ then $\mathbf{r}$ is a non-trivial linear combination of the logarithmic gradients $(\nabla_{\log} H_j)(\mathbf{w})$, so every contributing point is critical. Furthermore, a check of the definitions shows that for the power series expansion of rational $F(\mathbf{z})$ any smooth minimal critical point is contributing (as should be the case since Definition 9.8 generalizes Definition 5.5 of Chapter 5).

Example 9.3 Continued (Lattice Path Contributing Points)

In our running lattice path example the denominator factors $H_1 = 1 - t(1 + x^2 + xy^2)$ and $H_2 = 1 - y$ which vanish at the minimal critical point $\boldsymbol{\sigma} = (1, 1, 1/3)$ have logarithmic gradients $(\nabla_{\log} H_1)(\boldsymbol{\sigma}) = -(1, 2/3, 1)$ and $(\nabla_{\log} H_2)(\boldsymbol{\sigma}) = -(0, 1, 0)$. Thus, the normal cone at $\boldsymbol{\sigma}$ is

$$N(\boldsymbol{\sigma}) = \left\{ a\left(1, \frac{2}{3}, 1\right) + b(0, 1, 0) : a, b > 0 \right\}.$$

Since $(1, 1, 1) \in N(\boldsymbol{\sigma})$, we see that $\boldsymbol{\sigma}$ is a contributing point for the direction $\mathbf{r} = \mathbf{1}$.

Example 9.6 (Another Quadrant Lattice Path Model)

If $\mathcal{B} = \{(1, -1), (-1, -1), (0, 1)\}$ then Theorem 4.2 in Chapter 4 implies that the generating function for the number of walks in $\mathbb{N}^2$ which begin at the origin and use the steps in $\mathcal{B}$ is the main diagonal of the rational function

$$F(x, y, t) = \frac{(x - x^2y^2 - y^2)(1 + x)}{x(1 - t(x + y^2 + x^2y^2))(1 - y)}.$$

Although we develop ACSV for general Laurent expansions, here we can stick to a power series expansion by noting that the nth main diagonal coefficient of $F(x, y, t)$ is the $(n + 1)$st main diagonal coefficient of

$$E(x, y, t) = xytF(x, y, t) = \frac{(x - x^2y^2 - y^2)(1 + x)yt}{(1 - t(x + y^2 + x^2y^2))(1 - y)}. \tag{9.5}$$

The denominator $H(x, y, t)$ of $E(x, y, t)$ has the transverse polynomial factorization $H = H_1H_2$ where $H_1(x, y, t) = 1 - t(x + y^2 + x^2y^2)$ and $H_2(x, y, t) = 1 - y$. Solving the critical point equations on each stratum shows that $\mathcal{S}_1 = \mathcal{V}(H_1) \setminus \mathcal{V}(H_2)$ has four critical points in the main diagonal direction,

$$\boldsymbol{\sigma}_\omega = \left(\omega^2, \frac{\omega}{\sqrt{2}}, \frac{\omega^2}{2}\right), \qquad \omega \in \{1, -1, \mathrm{i}, -\mathrm{i}\},$$

which are finitely minimal points where the singular variety is smooth. The stratum $\mathcal{S}_2 = \mathcal{V}(H_2) \setminus \mathcal{V}(H_1)$ contains no critical points, but the stratum $\mathcal{S}_{1,2} = \mathcal{V}(H_1, H_2)$ contains two critical points: $\boldsymbol{\tau} = (1, 1, 1/3)$, which is minimal, and $(-1, 1, 1)$, which is not minimal. Note, in particular, that $\boldsymbol{\tau}$ and the $\boldsymbol{\sigma}_\omega$ are minimal critical points but they do not all have the same coordinate-wise modulus. Because we are considering a power series expansion any smooth minimal critical point is contributing, so the $\boldsymbol{\sigma}_\omega$ are all contributing points (the logarithmic gradient of H_1 at each point is a non-zero multiple of $\mathbf{1}$). Computing the logarithmic gradients of H_1 and H_2 at $\boldsymbol{\tau}$ shows

$$N(\boldsymbol{\tau}) = \left\{ a\left(1, \frac{4}{3}, 1\right) + b(0, 1, 0) : a, b > 0 \right\}$$

so $(1,1,1) \notin N(\boldsymbol{\tau})$ and $\boldsymbol{\tau}$ is *not* a contributing point for $\mathbf{r} = \mathbf{1}$.

Because there are finitely minimal smooth critical points, the non-smooth points of $E(x,y,t)$ in (9.5) are irrelevant. In particular, an asymptotic expansion for the number w_n of walks in $\mathbb{N}^2$ using the steps in $\mathcal{B}$ can be computed using Corollary 5.2 of Chapter 5. Adding the asymptotic contributions of each $\boldsymbol{\sigma}_\omega$ to the main diagonal of E, then shifting the index of n by one to return to the diagonal of F, gives

$$w_n = \frac{\left(2\sqrt{2}\right)^n}{n^2}\left(\frac{A_n}{\pi} + O\left(\frac{1}{n}\right)\right), \qquad A_n = \begin{cases} 12\sqrt{2} & : n \text{ even} \\ 32 & : n \text{ odd} \end{cases}.$$

Note that the numerator of $E(x,y,t)$ vanishes at each $\boldsymbol{\sigma}_\omega$, so the second order term in the expansion (5.27) must be calculated for each point, and that the leading constant has a periodicity coming from the different contributions of the $\boldsymbol{\sigma}_\omega$.

Minimal contributing points are important because, as in the smooth case, they are minimizers of the height function on the closure of the domain of convergence. The fastest way to see this is to recall the discussion of polynomial amoebas and the Relog map from Section 3.3.1 in Chapter 3.

Proposition 9.1 *Let* $\mathbf{w} \in \mathbb{C}_*^d$ *be a transverse point of* $H(\mathbf{z}) \in \mathbb{C}[\mathbf{z}]$. *If* $\mathbf{w}$ *is a contributing point of* H *for the direction* $\mathbf{r}$ *and* $\mathbf{w}$ *is minimal (i.e.,* $\mathbf{w} \in \partial\mathcal{D}$*) then* $\mathbf{w}$ *is a minimizer of the height function* $h_{\mathbf{r}}(\mathbf{z})$ *on* $\overline{\mathcal{D}}$.

Proof Let $H(\mathbf{z}) = u(\mathbf{z})H_1(\mathbf{z})^{p_1}\cdots H_s(\mathbf{z})^{p_s}$ be a square-free factorization of H in $\mathcal{O}_{\mathbf{w}}$ and let $B = \text{Relog}(\mathcal{D})$ be the component of amoeba$(H)^c$ corresponding to the power series expansion of $F(\mathbf{z})$. Near Relog$(\mathbf{w})$ the amoeba of H contains (at least) the zeroes of the $H_j(\mathbf{z})$ under the Relog map. Proposition 3.6 from Chapter 3 implies that the vectors $\mathbf{v}_j$ in Definition 9.8 are outward normals for support hyperplanes of B at Relog$(\mathbf{w})$ (where Remark 3.2 covers the case when the H_j are analytic at $\mathbf{w}$ but not polynomials). Thus, as illustrated in Figure 9.2, any element of $N(\mathbf{w})$ is the outward normal to a support hyperplane of B at Relog$(\mathbf{w})$. Since $h_{\mathbf{r}}$ becomes the linear function $\tilde{h}(\mathbf{x}) = -\mathbf{r}\cdot\mathbf{x}$ after taking the Relog map, this implies Relog$(\mathbf{w})$ is a minimizer of $\tilde{h}$ on the convex set $\overline{B}$, and thus $\mathbf{w}$ is a minimizer of $h_{\mathbf{r}}$ on $\overline{\mathcal{D}}$. □

Example 9.3 Continued (Lemniscate Contributing Points)

Recall the lemniscate $L(x,y) = (1/3-x)^2(1/2-y)^2 + (5/6-x-y)(11/6-4x-y)$ from above, whose zero set $\mathcal{V}(L)$ contains a single non-smooth point $\boldsymbol{\varrho} = (1/3, 1/2)$. We have seen that the leading homogeneous term of L at $\boldsymbol{\varrho}$ factors as $\ell(x,y) = (x+y)(4x+y)$ and there is a factorization $L = L_1L_2$ in $\mathcal{O}_{\boldsymbol{\varrho}}$ where

$$(\nabla_{\log}L_1)(\boldsymbol{\varrho}) = (w_1, w_2) = (1/3, 1/2), \qquad (\nabla_{\log}L_2)(\boldsymbol{\varrho}) = (4w_1, w_2) = (4/3, 1/2).$$

This gives a normal cone

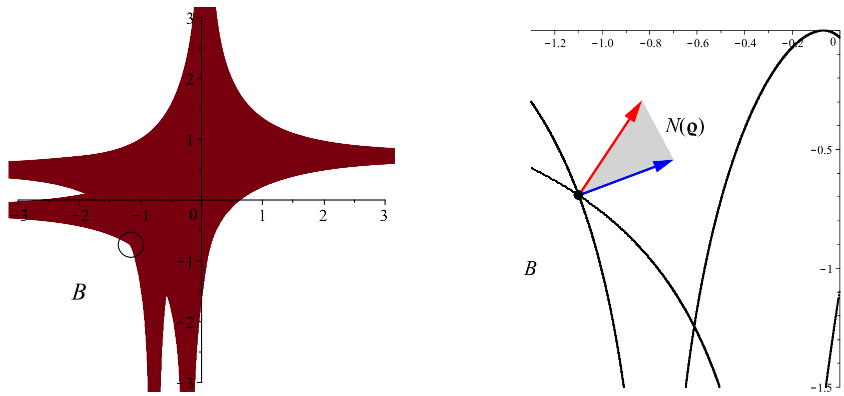


Fig. 9.2 *Left:* The amoeba of the polynomial $L(x, y)$ with the component $B \subset \text{amoeba}(L)^c$ corresponding to the power series expansion of $1/L(x, y)$ labelled. The image of the multiple point $\boldsymbol{\varrho} = (1/3, 1/2)$ under the Relog map, where ∂B has a cusp, is circled. *Right:* A portion of the contour of $L(x, y)$ with the normal cone $N(\boldsymbol{\varrho})$ displayed. The normals to the two smooth curves making up the contour near $\text{Relog}(\boldsymbol{\varrho})$, which form the boundary of $N(\boldsymbol{\varrho})$, are the logarithmic gradients of the factors of L in the local ring $\mathcal{O}_{\boldsymbol{\varrho}}$.

$$N(\boldsymbol{\varrho}) = \left\{ a\left(\frac{2}{3}, 1\right) + b\left(\frac{8}{3}, 1\right) : a, b > 0 \right\};$$

see Figure 9.2. The vector $\mathbf{r} \in \mathbb{R}^2_{>0}$ lies in $N(\boldsymbol{\varrho})$ if and only if $r_1/r_2 \in (2/3, 8/3)$, which gives the directions where $\boldsymbol{\varrho}$ is a contributing singularity. The rational function $1/L(x, y)$ also has smooth critical points, whose coordinates are degree six algebraic functions in the coordinates of $\mathbf{r} \in \mathbb{R}^d_{>0}$. There will be a minimal smooth critical point only when $\boldsymbol{\varrho}$ is not contributing (otherwise, since $\boldsymbol{\varrho}$ is strictly minimal, there would be two minimal contributing points with different coordinate-wise moduli, which cannot occur since the boundary of amoeba(L) does not contain a line segment).

Example 9.7 (Lattice Path Amoebas)

Visualizations of the amoebas of the lattice path models discussed above, together with two-dimensional projections of the relevant critical points and logarithmic gradients, are given in Figures 9.3 and 9.4.

We end this section by discussing general (non-power series) Laurent expansions. The only concept above which does not immediately apply to general Laurent expansions is Definition 9.8 for contributing points, but Proposition 3.13 from Chapter 3 allows for a natural extension.

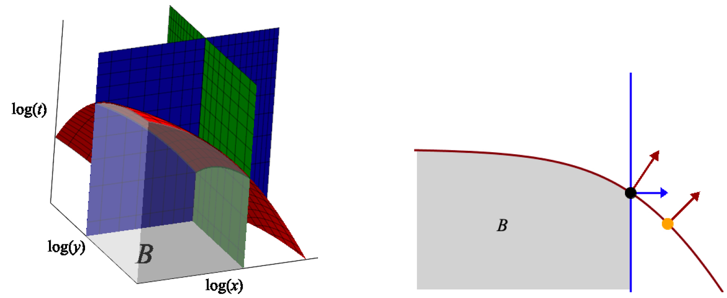


Fig. 9.3 *Left:* A portion of the amoeba contour for the denominator of the rational function corresponding to the quadrant model with step set $\mathcal{A}$, with the power series component B of the amoeba complement shaded. *Right:* The projection of this plot onto the plane $\log(x) = 1$, with the projections of the critical points and logarithmic gradients displayed. The vector $\mathbf{r} = \mathbf{1}$ is in the normal cone at the minimal critical point, which is a transverse point. The smooth critical point is not minimal, and does not lie on the boundary of B.

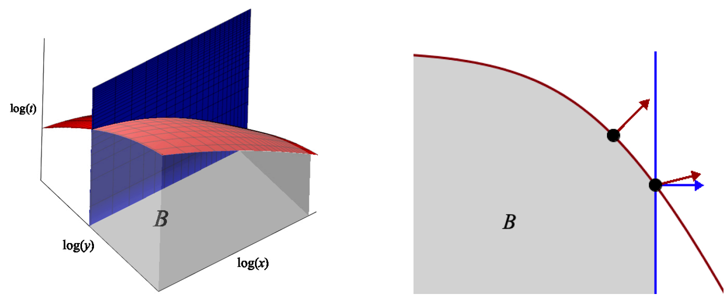


Fig. 9.4 *Left:* A portion of the amoeba contour for the denominator of the rational function corresponding to the quadrant model with step set $\mathcal{B}$, with the power series component of the amoeba complement shaded. *Right:* The projection of this plot onto the plane $\log(x) = 1$, with the projections of the critical points and logarithmic gradients displayed. The logarithmic gradient at the smooth critical point is a multiple of $\mathbf{r} = \mathbf{1}$, and this is a contributing singularity. The vector $\mathbf{r} = \mathbf{1}$ does not lie in the normal cone at the non-smooth minimal critical point.

Definition 9.9 (contributing points for Laurent expansions) Let $\mathbf{w} \in \mathbb{C}_*^d$ be a transverse point of $H(\mathbf{z})$, and let $H(\mathbf{z}) = u(\mathbf{z})H_1(\mathbf{z})^{p_1} \cdots H_s(\mathbf{z})^{p_s}$ be a square-free factorization of H in $\mathcal{O}_{\mathbf{w}}$. Consider a Laurent expansion of $F(\mathbf{z}) = G(\mathbf{z})/H(\mathbf{z})$ with domain of convergence $\mathcal{D}$ and for each $1 \leq j \leq s$ let $\mathbf{v}_j$ be a real vector which points away from $\mathrm{Relog}(\mathcal{D})$ such that $(\nabla_{\log} H_j)(\mathbf{w}) = \tau \mathbf{v}_j$ for some $\tau \in \mathbb{C}$ (the existence of such $\mathbf{v}_j$ follows from Proposition 3.13). The *normal cone* $N(\mathbf{w})$ of H at $\mathbf{w}$ with respect to this Laurent expansion is the positive span of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_s$. The point $\mathbf{w}$ is called a *contributing singularity*, or *contributing point*, for this Laurent expansion in the direction $\mathbf{r} \in \mathbb{R}_*^d$ if $\mathbf{r} \in N(\mathbf{w})$.

Remark 9.9 If $\mathcal{D}$ is the power series domain of convergence of rational $F(\mathbf{z})$ then Proposition 3.12 in Chapter 3 implies the convex set $\mathrm{Relog}(\mathcal{D})$ contains some translate of the negative orthant of $\mathbb{R}^d$. Thus, any normal vector to a supporting hyperplane of the boundary of Relog($\mathcal{D}$) which contains a positive coordinate will point away from Relog($\mathcal{D}$), verifying that Definition 9.9 is a generalization of Definition 9.8.

We are now ready to determine coefficient asymptotics.

9.2.2 Asymptotics via Residue Forms

Asymptotics for transverse points are determined by extending the approach of Chapter 8, where the singular variety is a hyperplane arrangement, to transverse points, where the singular variety locally looks like a hyperplane arrangement. Unfortunately, the most natural way to generalize our previous uses topological concepts like homology which are outside the scope of this text. We thus sketch proofs in this section and refer to Pemantle and Wilson [26] for full details.

We begin by introducing some constructions necessary to state our results. Suppose $\mathbf{w}$ is a transverse point of $H(\mathbf{z})$ such that there is a square-free factorization $H(\mathbf{z}) = u(\mathbf{z})H_1(\mathbf{z})^{p_1} \cdots H_s(\mathbf{z})^{p_s}$ in $\mathcal{O}_\mathbf{w}$, and let $S = \mathcal{V}(H_1) \cap \cdots \cap \mathcal{V}(H_s) \cap \mathcal{N}$ for any neighbourhood $\mathcal{N}$ of $\mathbf{w}$ in $\mathbb{C}^d$ on which all H_j are defined. Because the gradients of the H_j are linearly independent at $\mathbf{w}$, there exists a set $P = \{\pi_1, \dots, \pi_{d-s}\}$ of $d - s$ distinct coordinates which analytically parametrize the remaining s coordinates of S. In other words, for each $j \notin P$ there exists an analytic function $\zeta_j(z_{\pi_1}, \dots, z_{\pi_{d-s}})$ such that $\mathbf{z} \in S$ if and only if $z_j = \zeta_j(z_{\pi_1}, \dots, z_{\pi_{d-s}})$ for all $j \notin P$.

Definition 9.10 (constructions for transverse points) Under the setup of the preceding paragraph, we call P a set of *parameterizing coordinates* for $\mathcal{V}(H)$ at $\mathbf{w}$, and the set of functions $\{\zeta_j : j \notin P\}$ a *local parameterization of* $\mathcal{V}(H)$ *at* $\mathbf{w}$. The *modified log-normal matrix of* H *at* $\mathbf{w}$ is the $d \times d$ non-singular matrix

$$\Gamma_\mathbf{w} = \begin{pmatrix} (\nabla_{\log} H_1)(\mathbf{w}) \\ \vdots \\ (\nabla_{\log} H_s)(\mathbf{w}) \\ w_{\pi_1}\mathbf{e}^{(\pi_1)} \\ \vdots \\ w_{\pi(d-s)}\mathbf{e}^{(\pi_{d-s})} \end{pmatrix},$$

where $\mathbf{e}^{(j)}$ is the jth elementary basis vector. Finally, define the function

$$g(\theta_1, \dots, \theta_{d-s}) = \sum_{j \notin P} r_j \log\left[\zeta_j\left(w_{\pi_1}e^{i\theta_1}, \dots, w_{\pi_{d-s}}e^{i\theta_{d-s}}\right)\right].$$

The *parameterized Hessian matrix* of H at $\mathbf{w}$ is the $(d-s)\times(d-s)$ Hessian matrix of g at $\boldsymbol{\theta} = \mathbf{0}$, written $\mathcal{Q}_\mathbf{w}$. If $\mathcal{Q}_\mathbf{w}$ has non-zero determinant then we call $\mathbf{w}$ *nondegenerate*.

As in Chapter 8, we can be most precise about asymptotics at a transverse point where the number of factors of H in the local ring equals the number of variables.

Recall from previous chapters the notation $T(\mathbf{x})$ for the points of $\mathbb{C}^d$ with the same coordinate-wise modulus as $\mathbf{x}$. Given $\mathbf{p} \in \mathbb{N}^s$ we write $(\mathbf{p}-\mathbf{1})! = (p_1-1)!\cdots(p_s-s)!$.

Theorem 9.1 *Fix a Laurent expansion of rational $F(\mathbf{z}) = G(\mathbf{z})/H(\mathbf{z})$ with domain of convergence $\mathcal{D}$ and let $\mathbf{r} \in \mathbb{R}^d_*$. Suppose $\mathbf{x} \in \partial\mathcal{D}$ minimizes $|\mathbf{z}^{-\mathbf{r}}|$ on $\overline{\mathcal{D}}$, all minimizers of $|\mathbf{z}^{-\mathbf{r}}|$ on $\overline{\mathcal{D}}$ lie in $T(\mathbf{x})$, and all elements of $\mathcal{V} \cap T(\mathbf{x})$ are transverse points. Suppose also that any critical point $\mathbf{z}$ in $T(\mathbf{x})$ has $\mathbf{r} \notin \partial N(\mathbf{z})$, and that the set*

$$E = \{\mathbf{z} \in T(\mathbf{x}) : \mathbf{z} \text{ is a contributing point for } \mathbf{r}\} = \{\mathbf{w}\}$$

contains a single point. If H has square-free factorization $H(\mathbf{z}) = u(\mathbf{z})H_1(\mathbf{z})^{p_1}\cdots H_d(\mathbf{z})^{p_d}$ in $\mathcal{O}_\mathbf{w}$ then there exists $0 < \tau < |\mathbf{w}^{-\mathbf{r}}|$ such that

$$f_{n\mathbf{r}} = \mathbf{w}^{-n\mathbf{r}}\, n^{p_1+\cdots+p_d-d}\, \frac{G(\mathbf{w})\ \left(\mathbf{r}\Gamma_\mathbf{w}^{-1}\right)^{\mathbf{p}-\mathbf{1}}}{u(\mathbf{w})|\det\Gamma_\mathbf{w}|\,(\mathbf{p}-\mathbf{1})!} + O(\tau^n), \tag{9.6}$$

where $\Gamma_\mathbf{w}$ is the parameterized Hessian matrix from Definition 9.10.

Proof (sketch) As always, our asymptotic analysis starts with the Cauchy integral representation (9.2). Because all minimizers of $|\mathbf{z}^{-\mathbf{r}}|$ on $\overline{\mathcal{D}}$ have the same coordinate-wide modulus as $\mathbf{w}$, and $\mathbf{w}$ is the only contributing singularity in $T(\mathbf{w})$, the Cauchy domain of integration can be deformed so that the only points of asymptotic interest are arbitrarily close to $\mathbf{w}$. Near $\mathbf{w}$ the image of the singular variety $\mathcal{V}$ under the coordinate-wise logarithm map looks like a complete intersection of a simple hyperplane arrangement. Because $\mathbf{w}$ is a contributing point, it seems reasonable to think our asymptotic arguments from Section 8.3 of Chapter 8 would apply. Indeed, Proposition 10.3.6 and Theorem 10.2.6 of Pemantle and Wilson [26] use the language of relative homology to show that the coefficients $f_{n\mathbf{r}}$ can be represented by an integral over a product of circles around $S = \mathcal{V}(H_1)\cap\cdots\cap\mathcal{V}(H_d)$ (generalizing Proposition 8.3 of Chapter 8). Theorem 10.3.1 of [26] then shows how to use the theory of multivariate Leray residues to compute the expansion (9.6). Our work in Section 8.3 of Chapter 8 can be seen as an explicit development of this theory for the restricted case when H factors into real linear polynomials, so that linear algebra techniques can replace homological methods. □

Remark 9.10 As was the case for hyperplane arrangements, when the number of local factors equals the number of variables and the numerator $G(\mathbf{w}) \neq 0$ then we obtain dominant asymptotics up to an exponentially lower term.

Example 9.4 Continued (Asymptotics of the Lemniscate)

Recall again the lemniscate

$$L(x,y) = \left(\frac{1}{3}-x\right)^2\left(\frac{1}{2}-y\right)^2 + \left(\frac{5}{6}-x-y\right)\left(\frac{11}{6}-4x-y\right)$$

with strictly minimal critical point $\varrho = (1/3, 1/2)$ such that L factors in $\mathcal{O}_\varrho$ as $L = L_1 L_2$ where $(\nabla_{\log} L_1) = (1/3, 1/2)$ and $(\nabla_{\log} L_2) = (4/3, 1/2)$. The matrix

$$\Gamma_\varrho = \begin{pmatrix} 1/3 & 1/2 \\ 4/3 & 1/2 \end{pmatrix}$$

has determinant $-1/2$ and the normal cone at ϱ equals

$$N(\varrho) = \left\{ a\left(\frac{2}{3}, 1\right) + b\left(\frac{8}{3}, 1\right) : a, b > 0 \right\}.$$

Because ϱ is strictly minimal and the boundary of amoeba(L) doesn't contain a line segment, for any direction $(a, b) \in N(\varrho)$ the point ϱ is the unique minimizer of $|x^a y^b|$ on the boundary of the power series domain of convergence of $1/L(x, y)$. Thus, if $(a, b) \in N(\varrho)$ then Theorem 9.1 implies

$$[x^{an} y^{bn}] L(x, y)^{-1} = 2\left(3^a 2^b\right)^n + O(\tau^n),$$

for some $0 < \tau < 3^a 2^b$.

When asymptotics are determined by a transverse point where the number of factors of H in the local ring is less than the number of variables we still determine dominant asymptotics, but the error term obtained is polynomially small instead of exponentially small.

Theorem 9.2 *Fix a Laurent expansion of rational $F(\mathbf{z}) = G(\mathbf{z})/H(\mathbf{z})$ with domain of convergence $\mathcal{D}$ and let $\mathbf{r} \in \mathbb{R}^d_*$. Suppose $\mathbf{x} \in \partial\mathcal{D}$ minimizes $|\mathbf{z}^{-\mathbf{r}}|$ on $\overline{\mathcal{D}}$, all minimizers of $|\mathbf{z}^{-\mathbf{r}}|$ on $\overline{\mathcal{D}}$ lie in $T(\mathbf{x})$, and all elements of $\mathcal{V} \cap T(\mathbf{x})$ are transverse points. Suppose also that any critical point $\mathbf{z}$ in $T(\mathbf{x})$ has $\mathbf{r} \notin \partial N(\mathbf{z})$, and that the set*

$$E = \{\mathbf{z} \in T(\mathbf{x}) : \mathbf{z} \text{ is a contributing point for } \mathbf{r}\} = \{\mathbf{w}\}$$

contains a single point, which is nondegenerate. If H has a square-free factorization $H(\mathbf{z}) = u(\mathbf{z}) H_1(\mathbf{z})^{p_1} \cdots H_s(\mathbf{z})^{p_s}$ in $\mathcal{O}_\mathbf{w}$ then there exists $0 < \tau < |\mathbf{w}^{-\mathbf{r}}|$ such that

$$f_{n\mathbf{r}} = \mathbf{w}^{-n\mathbf{r}} n^{(s-d)/2 + p_1 + \cdots + p_s - s} \left(\frac{(2\pi)^{(s-d)/2} G(\mathbf{w}) \left(\mathbf{r}\Gamma_\mathbf{w}^{-1}\right)^{\mathbf{p}-\mathbf{1}}}{u(\mathbf{w})\sqrt{\det(r_d \mathcal{Q}_\mathbf{w})}\, |\det \Gamma_\mathbf{w}|\, (\mathbf{p}-\mathbf{1})!} + O\left(\frac{1}{n}\right)\right), \tag{9.7}$$

where $\Gamma_\mathbf{w}$ and $\mathcal{Q}_\mathbf{w}$ are defined in Definition 9.10.

Proof (sketch) As in the proof sketch of Theorem 9.1, we can manipulate the Cauchy integral representation (9.2) into an integral arbitrarily close to $\mathbf{w}$. Now, Proposition 10.3.6 and Theorem 10.2.6 of [26] show that the coefficients $f_{n\mathbf{r}}$ can be represented by an integral over the product of s circles around $S = \mathcal{V}(H_1) \cap \cdots \cap \mathcal{V}(H_d)$ with a $(d-s)$-dimensional domain of integration (similar to Proposition 8.3 of

Chapter 8). Integrating over the product of s circles leaves a $(d-s)$-dimensional integral over a subset of S whose integrand is a 'multivariate residue form' in the parametrizing variables $z_{\pi_1}, \dots, z_{\pi_{d-s}}$ of Definition 9.10 (generalizing Proposition 8.4 of Chapter 8). Theorem 10.3.4 of Pemantle and Wilson [26] shows how to massage this integral into something amenable to a saddle-point analysis, which is then asymptotically approximated. □

Remark 9.11 Although Theorem 9.2 applies when the number of factors $s = d$, in this case it is weaker than Theorem 9.1. When asymptotics are determined by minimal critical points, Theorems 9.1 and 9.2 generalize Theorem 8.2 from Chapter 8, which applies when the factors H_j are linear. Unlike Theorem 8.2, however, Theorems 9.1 and 9.2 only apply to minimal points. Dealing with non-minimal critical points is a large open problem in analytic combinatorics in several variables, discussed further in Section 9.3 below.

Example 9.4 Continued (Lattice Path Asymptotics)

Consider again the lattice path model in $\mathbb{N}^2$ defined by the step set $\mathcal{A} = \{(-1,1),(0,-1),(1,1)\}$, whose generating function is the main diagonal of

$$F(x,y,t) = \frac{(1+x)(1-xy^2+x^2)}{(1-t(1+x^2+xy^2))(1-y)(1+x^2)}.$$

Above we have seen that F admits a minimal contributing point $\boldsymbol{\sigma} = (1,1,1/3)$ in the direction $\mathbf{r} = \mathbf{1}$, which is thus a minimizer of $|xyt|^{-1}$ on $\overline{\mathcal{D}}$. Problem 9.3 asks you to prove that all minimizers of $|xyt|^{-1}$ have the same coordinate-wise modulus as $\boldsymbol{\sigma}$ (this also follows from our general argument in Chapter 10). We have already calculated the logarithmic gradients

$$(\nabla_{\log} H_2)(\boldsymbol{\sigma}) = (-1,-2/3,-1) \quad \text{and} \quad (\nabla_{\log} H_2)(\boldsymbol{\sigma}) = (0,-1,0),$$

and noted that the only critical point of $T(\boldsymbol{\sigma})$ is $\boldsymbol{\sigma}$ itself. Furthermore, on the flat $\mathcal{V}_{1,2} = \mathcal{V}(1-y, 1-t(1+x^2+xy^2))$ containing $\boldsymbol{\sigma}$ we can parametrize y and t by their x-coordinates as $y = 1$ and $t = 1/(1+x+x^2)$, giving

$$g(\theta) = \log(1) + \log\left(\frac{1}{1+e^{i\theta}+e^{2i\theta}}\right) = \log\left(\frac{1}{1+e^{i\theta}+e^{2i\theta}}\right)$$

and $Q = g''(0) = 2/3$. Since

$$\Gamma_{\boldsymbol{\sigma}} = \begin{pmatrix} & (\nabla_{\log} H_1)(\boldsymbol{\sigma}) & \\ & (\nabla_{\log} H_2)(\boldsymbol{\sigma}) & \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} -1 & -2/3 & -1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$

Theorem 9.2 implies the number w_n of walks of length n on the steps $\mathcal{A}$ which start at the origin and stay in $\mathbb{N}^2$ satisfies

$$w_n = 3^n n^{-1/2} \frac{\sqrt{3}}{2\sqrt{\pi}} \left(1 + O\left(\frac{1}{n}\right)\right).$$

As usual, if there are a finite number of points determining asymptotics then we can add the contributions from each of the points.

Corollary 9.1 *If the set E in Theorems 9.1 and 9.2 contains a finite set of points, all of which satisfy the conditions of Theorems 9.1 or 9.2, then one can sum the contributions of each element of E given by* (9.6) *and* (9.7) *to find asymptotics of $f_{n\mathbf{r}}$.*

We apply Theorem 9.2 and Corollary 9.1 to a variety of lattice path enumeration problems in Chapter 10.

9.3 A Geometric Approach to ACSV

We end this chapter with a discussion of the most general aspects of ACSV, using geometric constructions to give a high-level view of the theory and clarifying some earlier remarks by putting our previous results in context. To discuss this approach we first need to introduce some additional concepts from differential geometry. The definitions presented here are tailored to our setting.

Definition 9.11 (smooth and complex manifolds) Let $\mathcal{M} \subset \mathbb{C}^k$ for some positive integer k. A *chart of dimension s* on $\mathcal{M}$ is a pair (U, ϕ) consisting of an open subset $U \subset \mathcal{M}$ and a homeomorphism[2] ϕ from U onto an open subset of $\mathbb{R}^s$. An *atlas of dimension s* for $\mathcal{M}$ is a collection of charts $\{(U_\alpha, \phi_\alpha)\}$ of dimension s such that every $\mathbf{p} \in \mathcal{M}$ is contained in at least one U_α. If $\mathcal{M}$ admits an atlas of dimension s such that for any charts (U, ϕ) and (V, ψ) the maps

$$\phi \circ \psi^{-1} : \psi(U \cap V) \to \phi(U \cap V) \quad \text{and} \quad \psi \circ \phi^{-1} : \phi(U \cap V) \to \psi(U \cap V)$$

between subsets of $\mathbb{R}^s$ are smooth (all coordinates have derivatives of all orders) then $\mathcal{M}$ is a *smooth manifold of dimension s*. By replacing $\mathbb{R}^s$ by $\mathbb{C}^s$ in the above constructions we can analogously define *complex charts* and *complex atlases*, and $\mathcal{M}$ is a *complex manifold of dimension s* if it admits a complex atlas of dimension s such that for any charts (U, ϕ) and (V, ψ) the maps $\phi \circ \psi^{-1}$ and $\psi \circ \phi^{-1}$ are analytic.

Whenever we talk about charts in a manifold we always mean charts with respect to an atlas giving the manifold structure under consideration.

Example 9.8 (Smooth Varieties are Complex Manifolds)

Fix a polynomial $H(\mathbf{z})$ and let $\mathcal{V} = \mathcal{V}(H)$. If $\mathbf{w} \in \mathcal{V}$ and $(\partial H/\partial z_d)(\mathbf{w}) \neq 0$ then the implicit function theorem (Proposition 3.1 in Chapter 3) implies the existence of

---

[2] Recall that a homeomorphism is a continuous bijective map whose inverse is also continuous.

- a neighbourhood U of $\mathbf{w}$ in $\mathcal{V}$,
- a neighbourhood $\mathcal{N}$ of $\widehat{\mathbf{w}}$ in $\mathbb{C}^{d-1}$,
- an analytic function $g\colon \mathcal{N} \to \mathbb{C}$,

such that $U = \{(\hat{\mathbf{z}}, g(\hat{\mathbf{z}})) : \hat{\mathbf{z}} \in \mathcal{N}\}$. In other words, if $\pi\colon U \to \mathbb{C}^{d-1}$ denotes the projection $\pi(\mathbf{z}) = \hat{\mathbf{z}}$ of a point in U to its first $d-1$ coordinates then $\pi^{-1}\colon \mathcal{N} \to \mathbb{C}^d$ is the analytic function g, so (U, π) forms a $(d-1)$-dimensional complex chart of $\mathcal{V}$.

Parameterizing other variables as necessary, if for each $\mathbf{w} \in \mathcal{V}$ some partial derivative of H does not vanish at $\mathbf{w}$ then this procedure constructs a complex atlas of $\mathcal{V}$, showing that $\mathcal{V}$ is a complex manifold of dimension $d-1$.

Remark 9.12 By writing $z_j = x_j + iy_j$ to identify a complex variable z_j with the real variables x_j and y_j a complex manifold $\mathcal{M}$ of dimension s can be viewed as a smooth manifold of dimension $2s$.

Our next definitions help us understand the behaviour of functions on a manifold.

Definition 9.12 (mappings, differentials, and stationary points) If $\mathcal{M}$ is a smooth manifold of dimension s then a function $f\colon \mathcal{M} \to \mathbb{R}$ is called *smooth* at $\mathbf{p} \in \mathcal{M}$ if there is a chart (U, ϕ) with $\mathbf{p} \in U$ such that the map $f \circ \phi^{-1}\colon \mathbb{R}^s \to \mathbb{R}$ is smooth on $\phi(U)$. The *differential* of f at $\mathbf{p}$ is the vector $d_{\mathbf{p}} f \in \mathbb{R}^s$ whose jth entry equals $(\partial g/\partial x_j)(\mathbf{p})$ for $g(x_1, \ldots, x_s) = (f \circ \phi^{-1})(x_1, \ldots, x_s)$.

Similarly, if $\mathcal{M}$ is a complex manifold of dimension s then $f\colon \mathcal{M} \to \mathbb{C}$ is *analytic* at $\mathbf{p} \in \mathcal{M}$ if there is a chart (U, ϕ) with $\mathbf{p} \in U$ such that $f \circ \phi^{-1}\colon \mathbb{C}^s \to \mathbb{C}$ is analytic on $\phi(U)$, and the *differential* of f at $\mathbf{p}$ is the vector $d_{\mathbf{p}} f \in \mathbb{C}^s$ whose jth entry equals $(\partial g/\partial z_j)(\mathbf{p})$ for $g(z_1, \ldots, z_s) = (f \circ \phi^{-1})(z_1, \ldots, z_s)$.

In either case, $\mathbf{p} \in \mathcal{M}$ is called a *stationary point* of f if $d_{\mathbf{p}} f = \mathbf{0}$. Note that stationary points are usually called *critical points*, but we do not use this terminology to prevent confusion with the critical points of ACSV. Whether $\mathbf{p}$ is a stationary point does not depend on the chart used to calculate the differential of f.

Example 9.9 (Stationary Points)

Let $H(x, y) = 1 - x - y$ and $\mathcal{V}_* = \mathcal{V}(H) \cap \mathbb{C}_*^d$ be the roots of H with non-zero coordinates. An atlas of $\mathcal{V}_*$ is given by a single chart $(\mathcal{V}_*, \pi)$ where $\pi(x, y) = x$ is the projection map with inverse $\pi^{-1}(x) = (x, 1 - x)$. With respect to this chart the function $\psi(x, y) = \log(x) + \log(y)$ is represented by

$$g(x) = (\psi \circ \pi^{-1})(x) = \log(x) + \log(1 - x).$$

Since $g'(x) = 0$ has the unique solution $x = 1/2$, the only stationary point of ψ on $\mathcal{V}_*$ is $(1/2, 1/2)$.

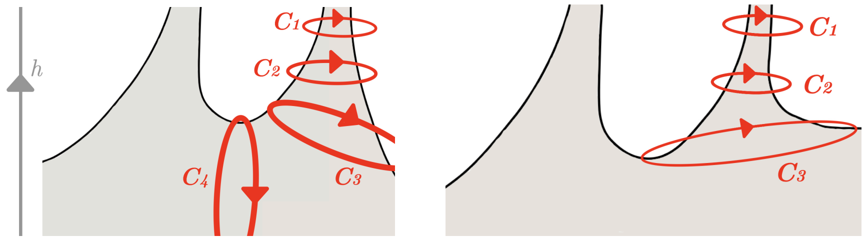


Fig. 9.5 *Left:* We can push down a cycle of integration at high height to a cycle arbitrarily close to a stationary point, except for points of lower height (the curve C_1 gets pushed down to C_2, then C_3, then C_4). *Right:* A problem arises when the gradient flow keeps decreasing but an asymptote forces it to stay at bounded height.

9.3.1 A Gradient Flow Interpretation for Analytic Combinatorics

Before returning to ACSV let us reflect on our first analytic argument for asymptotics, back in Section 2.1.1 of Chapter 2, where we studied alternating permutations by analyzing the coefficients of the meromorphic function $f(z) = \tan(z)$. The general outline of that approach, reflected in Figure 2.2 of Chapter 2, can be summarized as

1. start with the univariate Cauchy integral formula for coefficients,
2. push the domain of integration away from the origin without crossing singularities until it consists of points where the Cauchy integrand is exponentially small, *except* for points arbitrarily close to the dominant singularities $z = \pm\pi/2$,
3. compute asymptotic contributions of the dominant singularities using residues.

Our goal is to generalize this thinking (as much as possible) to the multivariate setting. To that end, fix a Laurent expansion of rational $F(\mathbf{z}) = G(\mathbf{z})/H(\mathbf{z})$ with domain of convergence $\mathcal{D}$ and let $\mathbf{r} \in \mathbb{R}^d_*$. Again we have the multivariate Cauchy integral representation

$$f_{n\mathbf{r}} = \frac{1}{(2\pi i)^d} \int_{\mathcal{T}} \frac{F(\mathbf{z})}{\mathbf{z}^{n\mathbf{r}+\mathbf{1}}} d\mathbf{z}$$

where $\mathcal{T}$ is a polytorus in $\mathcal{D}$. In the univariate case we pushed the domain of integration away from the origin because that made the Cauchy integrand small. Now the exponential growth of the Cauchy integrand is captured by the familiar height function $h(\mathbf{z}) = -\sum_{j=1}^{d} r_j \log|z_j|$, so it makes sense to deform the domain of integration $\mathcal{T}$ (without crossing the singular variety $\mathcal{V}$) so that the maximum of the height function on the domain of integration is minimized. The immediate problem in the multivariate setting is that there are *too many* ways to deform the domain of integration: it is not clear how the domain should be deformed in each of its d coordinates, and it is possible to leave the domain of convergence $\mathcal{D}$ without crossing a singularity.

Thankfully, we are searching for special points on the singular variety. Starting with a domain of integration at some (usually high[3]) height, we want to push the domain of integration to points of lower height until it locally gets stuck; see the left side of Figure 9.5. For any $r \in \mathbb{R}$, let $\mathcal{V}_r$ denote the points in $\mathcal{V}$ with height at least r. The points where the domain of integration can get stuck arise from a change in the topology of $\mathcal{V}_r$, and a study of such points is a classic geometric problem. In particular, Morse theory [22] studies changes in the topology of super-level sets for compact real manifolds sorted by a real height function. Stratified Morse theory [16] extends this to sets (like algebraic varieties) which are not manifolds but can be partitioned into a finite collection of manifolds which 'fit together nicely'. Similar results for complex varieties are also studied in Picard-Lefschetz theory [32].

These high-level theories tell us to partition the singular variety $\mathcal{V}$ into manifolds and study the stationary points of the height function h on each of the manifolds. Unfortunately, because h depends on the moduli of the coordinates it is only analytic in trivial cases. There are two ways to work around this difficulty,

1. If $\mathcal{M}$ is a complex manifold of dimension r then, as in Remark 9.12, we can view $\mathcal{M} \subset \mathbb{C}^d$ as a smooth manifold $\mathcal{W} \subset \mathbb{R}^{2d}$ of real dimension $2r$. Setting $z_j = x_j + iy_j$ and multiplying by two gives a modified height function $\eta(\mathbf{x}, \mathbf{y}) = 2h(\mathbf{x} + i\mathbf{y}) = -\sum_{j=1}^{d} r_j \log(x_j^2 + y_j^2)$ which *is* a smooth mapping, and we can look for stationary points of η on the smooth manifold $\mathcal{W}$.
2. Because h is the real part of the analytic function $\psi(\mathbf{z}) = -\sum_{j=1}^{d} r_j \log z_j$, the Cauchy-Riemann equations [17, Sec. 2.3] imply that the stationary points of the smooth map $\eta\colon \mathcal{W} \to \mathbb{R}$ are the stationary points of $\psi\colon \mathcal{M} \to \mathbb{C}$. One can thus determine interesting topological information using the analytic function ψ.

Standard algebro-geometric constructions can be used to partition an arbitrary algebraic set $\mathcal{V}$ into a finite collection of manifolds (see, for instance, Mumford [24, Ch. 1A]). The techniques of stratified Morse theory, however, require a *Whitney stratification*, which is a partition of $\mathcal{V}$ into manifolds with an extra condition ensuring that for any manifold $\mathcal{M}$ in the partition the local picture of $\mathcal{V}$ near all points of $\mathcal{M}$ is consistent. A full discussion of Whitney stratifications is beyond the scope of this text, but modern treatments can be found in Mather [20] and Goresky and MacPherson [16, Ch. 1.2]. Of particular interest to us is the fact that there exist algorithms [23, 28] which take an algebraic set $\mathcal{V}$ defined by explicit polynomial equations and return a finite sequence of algebraic sets

$$\varnothing = \mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots \subset \mathcal{F}_r = \mathcal{V},$$

each defined by explicit polynomial equations, such that the connected components of the differences $\mathcal{F}_j \setminus \mathcal{F}_{j-1}$ form a Whitney stratification of $\mathcal{V}$. This sequence of algebraic sets, known as a *canonical Whitney filtration of* $\mathcal{V}$, allows us to give a general definition of critical points.

[3] For instance, when dealing with power series expansions we can start with any product of circles sufficiently close to the origin, which can thus contain points of arbitrarily high height.

Definition 9.13 (critical points) Let $F(\mathbf{z})$ be a rational function with singular variety $\mathcal{V}$ and fix $\mathbf{r} \in \mathbb{R}^d_*$. Suppose $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots \subset \mathcal{F}_r$ is a canonical Whitney filtration of $\mathcal{V}$. The *critical points of F in the direction* $\mathbf{r}$ is the set of stationary points of the function $\psi(\mathbf{z}) = -\sum_{j=1}^d r_j \log z_j$ restricted to the differences $\mathcal{F}_j \setminus \mathcal{F}_{j-1}$. A critical point $\mathbf{w}$ in the stratum $\mathcal{S}$ is *non-degenerate* if there exists a chart (U, ϕ) of $\mathcal{S}$ with $\mathbf{w} \in U$ and $\phi(\mathbf{w}) = \mathbf{0}$ such that the Hessian of $\psi \circ \phi^{-1}$ at the origin is non-singular.

Given polynomials generating the elements of a canonical Whitney filtration, vanishing of the differential of ψ can be encoded as the vanishing of minors of a matrix with explicit polynomial entries. Thus, when the denominator of F is a polynomial with rational coefficients then all critical points have algebraic coordinates.

Remark 9.13 Definition 9.13 extends our previous definitions for critical points (in the smooth, hyperplane, and transverse multiple point settings). As proven in Section 5.3.4 of Chapter 5, generically the singular variety $\mathcal{V}$ is smooth, so a Whitney stratification for $\mathcal{V}$ is just $\mathcal{V}$ itself. When $\mathcal{V}$ is smooth then the critical points of Definition 9.13 are the solutions of the smooth critical point equations $H^s(\mathbf{z}) = 0$ and $r_j z_1 H^s_{z_1}(\mathbf{z}) - r_1 z_j H^s_{z_j}(\mathbf{z}) = 0$ for $2 \leq j \leq d$ from (5.16) in Chapter 5. If $\mathcal{V}$ is not smooth but has a transverse polynomial factorization, the next most common situation in practice, then critical points can be calculated using (9.4) above.

Roughly, stratified Morse theory tells us that the critical points of $F(\mathbf{z})$ represent the only obstructions to pushing down the Cauchy domain of integration to lower height. Such results are obtained via *gradient flows*, which use the differential of the height function on the Whitney strata to determine how to deform a domain of integration arbitrarily close to $\mathcal{V}$ to points of lower height. At stationary points the differential is zero, and the flow gets stuck.

Aside from stationary points there is one other situation in which the flow doesn't work as intended: when there is an 'asymptote' of $\mathcal{V}$ such that points on $\mathcal{V}$ go to infinity with decreasing height, but the height of these points is bounded from below[4] (see the right side of Figure 9.5). To work around this issue, Baryshnikov et al. [6] introduce a notion of *stationary points at infinity*, along with an algorithm to determine when there are no such points. Certifying that no stationary points at infinity exist then allows one to apply the usual techniques of Morse theory to reduce the Cauchy integral for coefficients to a sum of saddle-point integrals which can be asymptotically approximated. A characterization of stationary points at infinity was a long-time sticking point in the use of Morse-like topological constructions in ACSV contexts.

Recall the discussion on polynomial ideals and algebraic varieties from Section 7.3 in Chapter 7. Our discussion of stationary points at infinity uses the following construction.

---

[4] In fact, not all asymptotes of this form are a problem, only those that influence gradient flows. Indeed, sequences of points going out to infinity while staying at finite height appear generically in at least three dimensions, so if the existence of any asymptote was a problem then the methods developed using these arguments would not work on generic examples.

Definition 9.14 (saturation of ideals) Let $I, J \subset \mathbb{C}[\mathbf{z}]$ be two polynomial ideals. The *quotient* of I by J is the ideal $(I : J) = \{f \in \mathbb{C}[\mathbf{z}] : fg \in I \text{ for all } g \in J\}$. If $(I : J^n)$ denotes the result of dividing I by J repeatedly n times then the sequence $(I : J) \subset (I : J^2) \subset (I : J^3) \subset \cdots$ eventually stabilizes[5] to an ideal $(I : J^\infty)$ known as the *saturation of I by J*.

Remark 9.14 Geometrically, the zero set $\mathcal{V}(I : J^\infty)$ is the smallest algebraic set containing $\mathcal{V}(I) \setminus \mathcal{V}(J)$. There are algorithms for computing ideal saturations using Gröbner bases [10, Table 6.6], and such algorithms have been implemented in several[6] algebra systems.

Although Baryshnikov et al. give more general results, to be explicit we focus on the generic case when $F(\mathbf{z}) = G(\mathbf{z})/H(\mathbf{z})$ has a smooth singular variety. In order to avoid introducing even more geometric language, we do not fully define stratified points at infinity. Instead, we give the effective test of Baryshnikov et al. [6, Prop. 3.2] which characterizes their absence. Recall from Chapter 5 that the square-free factorization of a polynomial is the product of its irreducible factors and that the *homogenization* of a polynomial $f(\mathbf{z}) \in \mathbb{C}[\mathbf{z}]$ of degree k is the polynomial $z_0^k f(z_1/z_0, \ldots, z_d/z_0) \in \mathbb{C}[z_0, \mathbf{z}]$.

Definition 9.15 (absence of stratified points at infinity) Let $F(\mathbf{z})$ be the ratio of coprime polynomials $G(\mathbf{z})$ and $H(\mathbf{z})$ with smooth singular variety $\mathcal{V} = \mathcal{V}(H)$, and fix $\mathbf{r} \in \mathbb{R}_*^d$. Let $\tilde{H}(z_0, \mathbf{z})$ be the homogenization of the square-free part of H, and let I be the ideal of $\mathbb{C}[z_0, \mathbf{z}, \mathbf{y}]$ generated by the polynomials

$$\tilde{H}(z_0, \mathbf{z}) \quad \text{and} \quad y_j z_1 \tilde{H}_{z_1}(z_0, \mathbf{z}) - y_1 z_j \tilde{H}_{z_j}(z_0, \mathbf{z}) \quad \text{for} \quad 2 \le j \le d.$$

Let C be the saturation of I by the ideal (z_0) and $C_\mathbf{r}$ be the result of substituting $\mathbf{y} = \mathbf{r}$ and $z_0 = 0$ in C. We say F *has no stratified points at infinity in the direction* $\mathbf{r}$ if the only element of $\mathcal{V}(C_\mathbf{r})$ is $\mathbf{z} = \mathbf{0}$.

The rough idea behind Definition 9.15 is the following. By homogenizing H we can move from $\mathbb{C}^d$ to complex projective space, so that 'points at infinity' can be captured algebraically by points where $z_0 = 0$. The polynomials in the ideal I generate the smooth critical points equations on $\mathcal{V}$ when the direction $\mathbf{y}$ is a parameter, and saturation by z_0 removes extra 'points at infinity' which are not limit points of critical points in $\mathbb{C}^d$. The elements of the saturation with $\mathbf{y} = \mathbf{r}$ and $z_0 = 0$ are thus limit points of smooth critical points in directions approaching $\mathbf{r}$ which do not lie in $\mathbb{C}^d$. If any solution with $z_0 = 0$ also has $\mathbf{z} = \mathbf{0}$ then there are no solutions in projective space (because $\tilde{H}$ is homogeneous the point $(z_0, \mathbf{z}) = (0, \mathbf{0})$ will always be a solution).

[5] If $I_1 \subset I_2 \subset \cdots$ is an infinite ascending chain of ideals then the union of all I_j is also an ideal. The Hilbert basis theorem states that this ideal has a finite set of generators, each of which must lie in I_j with j sufficiently large, so the inclusions eventually become identities. The condition that any infinite ascending chain of ideals stabilizes characterizes a *Noetherian* ring.

[6] For instance, saturation of ideals is implemented by the Saturate command of the PolynomialIdeals package in Maple.

Example 9.10 (A Stratified Point at Infinity)

Consider the rational function $F(x, y) = 1/H(x, y)$ with square-free denominator $H(x, y) = 2 + y - x(1 + y)^2$. Since the system

$$H(x, y) = H_x(x, y) = H_y(x, y) = 0$$

has no solutions, the singular variety $\mathcal{V}$ is smooth. Furthermore, the system

$$H(x, y) = xH_x(x, y) - yH_y(x, y) = 0$$

has no solutions, so F has no critical points in the main diagonal direction. Homogenizing H gives $\tilde{H}(z, x, y) = z^3 H(x/z, y/z) = 2z^3 + z^2 y - x(z + y)^2$. Forming the ideal I, using the PolynomialIdeals package of Maple to saturate by (z), then substituting in the main diagonal direction and $z = 0$ gives $C_{\mathbf{1}} = (x)$. Since $\mathcal{V}(C_{\mathbf{1}}) \subset \mathbb{C}^2$ contains non-zero solutions, such as $(0, 1)$, there is a stratified point at infinity. Figure 5.5 of Chapter 5 shows the amoeba of H. The existence of a stratified point at infinity is reflected by a limit direction of amoeba(H) which is normal to $\mathbf{r} = \mathbf{1}$. Further details are given in the worksheet corresponding to this example.

Example 9.11 (No Stratified Points at Infinity)

Consider the rational function $F(x, y) = 1/H(x, y)$ with $H(x, y) = 1 - x - y$. Then $\tilde{H}(z, x, y) = z - x - y$ and for $\mathbf{r} = (r, s)$ we compute $C_{\mathbf{r}} = (y(r + s), x + y)$. There are no stratified points at infinity if $r + s \neq 0$, in which case $C_{\mathbf{r}} = (x, y)$. There are stratified points at infinity in directions which are a multiple of $\mathbf{r} = (1, -1)$.

Absence of stratified points at infinity implies that the gradient flow approach can decompose the Cauchy integral into a sum of saddle-point integrals near stationary points. The integrands of these saddle-point integrals are expressed by residues. If $\mathcal{V}$ is smooth and $\mathbf{w} \in \mathcal{V}$ we write

$$\operatorname*{Res}_{\mathbf{w}, \mathcal{V}} \left(\frac{F(\mathbf{z})}{\mathbf{z}^{n\mathbf{r}+\mathbf{1}}} d\mathbf{z} \right) = \left(\operatorname*{Res}_{z_k = g(\mathbf{z}_{\hat{k}})} \frac{F(\mathbf{z})}{\mathbf{z}^{n\mathbf{r}+\mathbf{1}}} \right) d\mathbf{z}_{\hat{k}}$$

for a coordinate z_k and analytic function $g(\mathbf{z}_{\hat{k}})$ such that $\mathbf{z}$ lies in some neighbourhood of $\mathbf{w}$ in $\mathcal{V}$ if and only if $z_k = g(\mathbf{z}_{\hat{k}})$. When we use this notation it is implicit that the stated formula is valid for any possible pair of coordinate z_k and analytic function g. Lemma 5.2 in Chapter 5 gives an explicit expression for the residue.

We are now ready to state the main result of Baryshnikov et al., simplified under our hypothesis of smoothness.

Theorem 9.3 (Baryshnikov et al. [6, Thm. 2.16]) *Let $F(\mathbf{z})$ be the ratio of coprime polynomials $G(\mathbf{z})$ and $H(\mathbf{z})$ with smooth singular variety $\mathcal{V} = \mathcal{V}(H)$. Suppose there are no stationary points at infinity in the direction $\mathbf{r} \in \mathbb{R}^d_*$ and that the critical points* *of F in the direction* $\mathbf{r}$ *are all nondegenerate and form the finite set* $E = \{\mathbf{w}_1, \ldots, \mathbf{w}_r\}$. *Then each* $\mathbf{w}_j$ *can be paired with an integer* κ_j *such that*

$$f_{n\mathbf{r}} = \sum_{j \in \chi} \frac{\kappa_j}{(2\pi i)^{d-1}} \int_{\mathcal{N}_j} \operatorname*{Res}_{\mathbf{w}_j, \mathcal{V}} \left(\frac{F(\mathbf{z})}{\mathbf{z}^{n\mathbf{r}+\mathbf{1}}} d\mathbf{z} \right) + O(\tau^n), \tag{9.8}$$

where χ *contains the elements* $\mathbf{w}_j \in E$ *which have highest height among those with* κ_j *non-zero, each* $\mathcal{N}_j$ *is a domain making the summands of* (9.8) *saddle-point integrals (meaning after a coordinate change they satisfy the hypotheses of Proposition 5.3 in Chapter 5), and* $0 < \tau < |\mathbf{w}^{\mathbf{r}}|$ *for any* $\mathbf{w} \in \chi$.

Theorem 9.3 expresses $f_{n\mathbf{r}}$ as a linear combination of saddle-point integrals, which can be asymptotically approximated using Proposition 5.3 of Chapter 5. The integers κ_j in the theorem arise from expressing the original Cauchy cycle of integration $\mathcal{T}$ in a basis for 'relative singular homology of a pair consisting of subsets of $\mathbb{C}_*^d \setminus \mathcal{V}$ and points of lower height'; they can be very difficult to determine directly using geometric arguments, and we do not give a full definition[7].

Remark 9.15 If one of the critical points $\mathbf{w}_j$ in Theorem 9.3 is minimal, we have seen in Chapter 5 that the corresponding integer coefficient κ_j equals one if $\mathbf{w}_j$ is contributing and zero otherwise. The power of Theorem 9.3 is that it tells us something about the asymptotic contributions of *non-minimal* critical points, an extremely difficult task outside of very specialized circumstances (since it is hard to deform the Cauchy domain of integration outside the domain of convergence). Other than minimal points, DeVries et al. [13] show how to determine the non-zero coefficients of highest height for bivariate functions with a smooth singular variety. There are also extensions of Theorem 9.3 to non-smooth singular varieties, and we have seen in Chapter 8 how to compute the coefficients κ_j when $\mathcal{V}$ is a hyperplane arrangement (in this situation each κ_j is ± 1 or 0, depending on the orthant of the critical point $\mathbf{w}_j$ and whether or not it is contributing).

An important corollary of Theorem 9.3 is the following, which bounds the exponential growth of $\mathbf{r}$-diagonals by the maximum of a finite algebraic set.

Corollary 9.2 *Under the assumptions of Theorem 9.3, let* ρ *be the maximum height of a critical point of* F *in the direction* $\mathbf{r}$. *Then* $\limsup_{n\to\infty} |f_{n\mathbf{r}}|^{1/n} \le e^{\rho}$.

Example 9.12 (Quantum Random Walks)

Discrete quantum random walks have been studied as a computational primitive for quantum computation (see, for instance, Ambainis et al. [1]). Bressler and Pemantle [12] analyze the behaviour of one-dimensional quantum walks which either stay stationary or move one step to the right at each time step, and always evolve an underlying quantum state, using bivariate generating functions of the form

[7] In fact, if $\eta = \max\{h(\mathbf{w}_j) : \kappa_j \neq 0\}$ then only the coefficients κ_j with $h(\mathbf{w}_j) \geq \eta$ are uniquely defined. These non-zero coefficients of high height are the only ones which appear in (9.8).

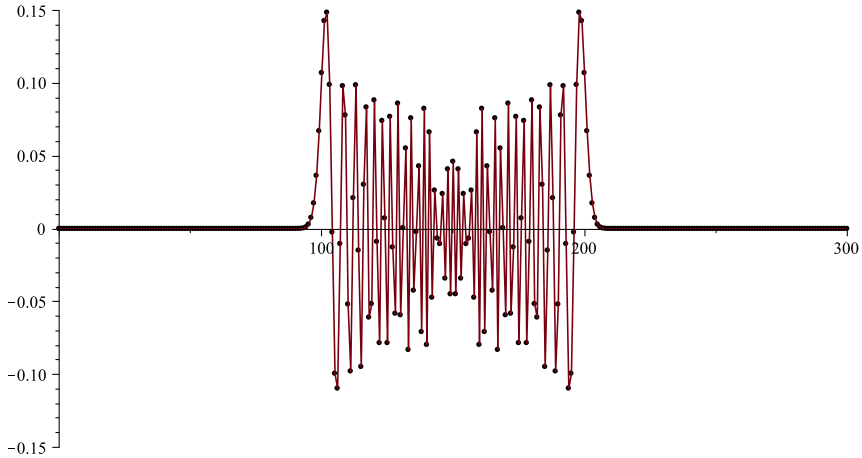


Fig. 9.6 Series coefficients of the polynomial $f(x) = [y^{300}]F(x, y)$ where $F(x, y) = (1 - cy + cxy - xy^2)^{-1}$ for $c = 1/3$: the coefficients are represented by dots, with consecutive terms connected by a line for visibility. Note the exponential drop outside of the interval [100, 200].

$$F(x, y) = \frac{G(x, y)}{1 - cy + cxy - xy^2} = \sum_{i,n \geq 0} f_{i,n} x^i y^n,$$

where $c \in (0, 1)$ is a parameter encoding the transition probabilities of the walk model, $G(x, y)$ is a polynomial encoding the initial fixed state of the model, and $|f_{i,n}|^2$ is the probability that a walk ends at $x = i$ after n steps. Unlike 'classical' walk models, where the endpoint probabilities for walks of large length usually approach a normal distribution, quantum walks experience an 'interference pattern' which makes the endpoint distribution far from normal (see Figure 9.6). Because the y variable tracks the length of a walk, and a walk takes at most one step at each time period, we are interested in asymptotic behaviour of the coefficients $[x^{\lambda n} y^n]F(x, y)$ for $0 < \lambda < 1$, given by the $\mathbf{r} = (\lambda, 1)$ diagonal. Of particular interest is the fact that the walk probability decays exponentially for λ outside $J = [(1 - c)/2, (1 + c)/2]$.

A quick computation shows that for any $c \in (0, 1)$ there are no stratified points at infinity. For any fixed values of c and $\lambda \notin J$ a computer algebra system can easily solve the smooth critical point equations $H = xH_x - \lambda y H_y = 0$ and verify that all critical points have negative height. Corollary 9.2 then implies exponential decay of the $\mathbf{r} = (\lambda, 1)$ diagonal by examining this finite set of algebraic points (without the need to determine minimality or any other properties of the singular variety). Detailed studies of quantum walks using analytic combinatorics in several variables can be found in Bressler and Pemantle [12] and Baryshnikov et al. [7].

Although it is usually not possible to determine the integers κ_j directly using geometric arguments, or even say if they are non-zero, knowing that these coefficients are integers can be very useful. Combined with the techniques on numeric analytic continuation for D-finite functions from Section 2.4 of Chapter 2, which allows one to determine asymptotics up to numeric constants with rigorous error bounds, this can be used to create a powerful tool for attacking the connection problem for sequences satisfying linear recurrence relations. We illustrate the possibilities of this combination on an example.

9.3.2 Attacking the Connection Problem through ACSV and Numeric Analytic Continuation

While resolving certain conjectures about the eventual positivity of sequences encoded by multivariate rational functions, Baryshnikov et al. [8] consider[8] the main power series diagonal coefficients of

$$F(w,x,y,z) = \frac{1}{H(w,x,y,z)} = \frac{1}{1-(w+x+y+z)+27wxyz}.$$

The system

$$H(w,x,y,z) = H_w(x,y,z) = H_x(w,x,y,z) = H_y(w,x,y,z) = H_z(w,x,y,z) = 0$$

has a unique solution at $(w,x,y,z) = \boldsymbol{\sigma} = (1/3, 1/3, 1/3, 1/3)$, which is the only non-smooth point of the singular variety $\mathcal{V} = \mathcal{V}(H)$. One thus obtains a Whitney stratification of the singular variety $\mathcal{V}$ by taking $\mathcal{S}_1 = \{\boldsymbol{\sigma}\}$ and $\mathcal{S}_2 = \mathcal{V} \setminus \{\boldsymbol{\sigma}\}$ as strata. Because $\boldsymbol{\sigma}$ lies in a stratum with a single element, it is a critical point (the height function is trivially constant on the stratum, so its differential is zero). The critical points in $\mathcal{S}_2$ are obtained by solving the smooth critical point equations

$$H = wH_w - xH_x = wH_w - yH_y = wH_w - zH_z = 0,$$

giving two smooth critical points $\boldsymbol{\tau}_+ = (\omega_+, \omega_+, \omega_+, \omega_+)$ and $\boldsymbol{\tau}_- = (\omega_-, \omega_-, \omega_-, \omega_-)$ for $\omega_\pm = (-1 \pm i\sqrt{2})/3$. Because $\boldsymbol{\sigma}$ has smaller coordinate-wise modulus than $\boldsymbol{\tau}_\pm$, the points $\boldsymbol{\tau}_\pm$ are not minimal. Applying either the algorithms of Chapter 7, or the famous Grace-Walsh-Szegő theorem [11, Thm. 1.1] on the distributions of zeroes of symmetric polynomials which are linear in each variable (see Problem 9.7 below), shows that $\boldsymbol{\sigma}$ is minimal. Thus, there is a single minimal critical point $\boldsymbol{\sigma}$, which is

[8] Continuing work of Gillis et al. [15] and Straub and Zudilin [30], Baryshnikov et al. study values of the real parameter C such that the main diagonal of $F(\mathbf{z}) = (1-(z_1+\cdots+z_d)+Cz_1\cdots z_d)^{-1}$ contains only a finite number of negative values. The most interesting case is $C = C_* = (d-1)^{d-1}$, which is the threshold for different behaviour (when $C < C_*$ there are only a finite number of negative values, and when $C > C_*$ there are an infinite number of both positive and negative values). When d is even and at least four, and $C = C_*$, the topology of the singular variety is perfectly aligned so that asymptotics are determined by non-minimal critical points; the rational function here is the smallest such example. See also the discussion in Section 3.4.5 of Chapter 3 for some context on these positivity problems.

not smooth, and two smooth critical points $\boldsymbol{\tau}_\pm$, which are not minimal. Computing the ideal C_1 in Definition 9.15 proves there are no stratified points at infinity.

To better understand the minimal critical point $\boldsymbol{\sigma}$ we examine the series expansion

$$H\left(w + \frac{1}{3}, x + \frac{1}{3}, y + \frac{1}{3}, z + \frac{1}{3}\right) = 3(wx + wy + wz + xy + xz + yz) \ + \text{ higher terms},$$

of H at $\boldsymbol{\sigma}$. Because the leading term is an irreducible quadratic, Lemma 9.1 implies $\boldsymbol{\sigma}$ is not a transverse multiple point. In fact, the invertible change of variables[9]

$$\begin{aligned} w &\mapsto a + \sqrt{6}b + \sqrt{2}c + d + 1/3 & \qquad x &\mapsto a - 3d + 1/3 \\ y &\mapsto a - 2\sqrt{2}c + d + 1/3 & \qquad z &\mapsto a - \sqrt{6}b + \sqrt{2}c + d + 1/3 \end{aligned}$$

transforms the leading term into $18(a^2 - b^2 - c^2 - d^2)$. A singularity whose leading term can put into the form $z_1^2 - z_2^2 - \cdots - z_d^2$ is called a *cone point*, since the real solutions of such a polynomial form two d-dimensional cones meeting at a point.

Baryshnikov and Pemantle [9] show that minimal critical cone points typically determine asymptotic behaviour of rational diagonals, *except* when they occur in even dimensions greater than four (which is precisely our situation). In even dimensions greater than four, a peculiar *lacuna* phenomenon[10] appears, suggesting that cone points may not contribute to coefficient asymptotics. Indeed, Baryshnikov et al. [5, Thm. 2.3] use an in-depth topological argument to show that for our rational function (and others of a similar form) the Cauchy domain of integration can be pushed past the minimal critical cone point $\boldsymbol{\sigma}$, which thus does not affect asymptotics. In particular, our example is a pathological case where asymptotics are determined by non-minimal critical points. We find asymptotics by expressing dominant coefficient behaviour in two ways: as a linear combination with unknown integer coefficients using the geometric approach to ACSV from this chapter, and as a linear combination with complex number coefficients that can be rigorously approximated using the computational tools for D-finite functions from Chapter 2.

Expression #1 from ACSV

Because there are no stationary points at infinity and the Cauchy domain of integration can be pushed past the critical point $\boldsymbol{\sigma}$, which has highest height among the three critical points in our example, asymptotic coefficient behaviour will be determined by the smooth (non-minimal) critical points $\boldsymbol{\tau}_\pm$ of height $h(\boldsymbol{\tau}_\pm) = \log 9$. Theorem 2.16 of Baryshnikov et al. [6], which is an extension of Theorem 9.3 that allows non-smooth points such as $\boldsymbol{\sigma}$, implies the asymptotic contributions of $\boldsymbol{\tau}_\pm$ are integer multiples of what they would be if these points were minimal critical points. In other words, there exist integers κ_+ and κ_- such that the power series coefficients

[9] This change of variables is found by diagonalizing the matrix M such that the leading term of H at $\boldsymbol{\sigma}$ is $\mathbf{v}M\mathbf{v}^T$ for $\mathbf{v} = \begin{pmatrix} w & x & y & z \end{pmatrix}$; see the worksheet corresponding to this example for details.

[10] See, for instance, Petrowsky [27] and Atiyah et al. [4].

of $F(w, x, y, z)$ satisfy

$$f_{n,n,n,n} = \kappa_- \Phi_{\boldsymbol{\tau}_-}(n) + \kappa_+ \Phi_{\boldsymbol{\tau}_+}(n) + O\left(\rho^n\right) \tag{9.9}$$

for some $0 < \rho < 9$, where $\Phi_{\boldsymbol{\tau}_\pm}(n)$ are the asymptotic expansions

$$\Phi_{\boldsymbol{\tau}_-}(n) = \frac{\left(-7+4\mathrm{i}\sqrt{2}\right)^n}{n^{3/2}\pi^{3/2}}\left(-\frac{\left(\sqrt{2}-5\mathrm{i}\right)\sqrt{-2\mathrm{i}\sqrt{2}-8}}{24} + O\left(\frac{1}{n}\right)\right)$$

$$\Phi_{\boldsymbol{\tau}_+}(n) = \frac{\left(-7-4\mathrm{i}\sqrt{2}\right)^n}{n^{3/2}\pi^{3/2}}\left(-\frac{\left(\sqrt{2}+5\mathrm{i}\right)\sqrt{2\mathrm{i}\sqrt{2}-8}}{24} + O\left(\frac{1}{n}\right)\right)$$

given by the right-hand side of (5.27) in Chapter 5 assuming $\boldsymbol{\tau}_\pm$ are minimal. As mentioned above, it is not immediately clear how to determine the integers $\kappa_\pm$ using geometric arguments.

Expression #2 from Numeric Analytic Continuation

Recall from Chapter 3 that the diagonal of any rational function satisfies a linear differential equation with polynomial coefficients. In fact, the creative telescoping techniques discussed in Section 3.2.1 of Chapter 3 allow us to compute that the diagonal $f(t) = (\Delta F)(t)$ of our rational function satisfies

$$\begin{aligned} t^2(81t^2+14t+1)f^{(3)}(t) + 3t(162t^2+21t+1)f^{(2)}(t) \\ + (21t+1)(27t+1)f'(t) + 3(27t+1)f(t) = 0. \end{aligned} \tag{9.10}$$

The analysis of D-finite functions in Section 2.4 of Chapter 2 implies that the only singularities of $f(t)$ occur at the roots $t = \omega_\pm^4$ of the leading coefficient factor $81t^2+14t+1 = 0$. Following the techniques discussed in Section 2.4.1 of Chapter 2, we can compute a basis of solutions to (9.10) for $t \in \{0, \omega_\pm^4\}$ and use numeric analytic continuation to determine singular expansions of $f(t)$ at $t = \omega_\pm^4$. Performing these calculations with the Sage package of Mezzarobba [21] allows us, in a few seconds of computation time, to determine an asymptotic expansion

$$\begin{aligned} f_{n,n,n,n} = {} & \frac{\left(-7+4\mathrm{i}\sqrt{2}\right)^n}{n^{3/2}}\left((0.3066\ldots) + (0.146\ldots)\mathrm{i} + O\left(\frac{1}{n}\right)\right) \\ & + \frac{\left(-7-4\mathrm{i}\sqrt{2}\right)^n}{n^{3/2}}\left((0.3066\ldots) - (0.146\ldots)\mathrm{i} + O\left(\frac{1}{n}\right)\right) \end{aligned} \tag{9.11}$$

whose coefficients are rigorously certified to over a thousand digits. Without additional information it is impossible to exactly determine the coefficients in this expansion, even though they can be approximated to arbitrary accuracy.

Determining Asymptotics Exactly

Equations (9.9) and (9.11), one with unknown integers and the other with approximated coefficients from a larger field, are perfect complements, and combining the expressions allows us to approximate $\kappa_1, \kappa_2 = 2.999\ldots$ to over a thousand decimal digits in a few seconds. Even knowing these integers to a single decimal digit is enough to rigorously conclude that $\kappa_1 = \kappa_2 = 3$ and thus exactly determine asymptotics of the diagonal sequence under consideration. One can think of the κ_j as a sort of 'winding number' of the starting Cauchy domain of integration $\mathcal{T}$. Working in four-dimensional complex space (which is extremely difficult to picture) one can deform $\mathcal{T}$ to points of height smaller than $h(\boldsymbol{\sigma})$, eventually moving to neighbourhoods of $\boldsymbol{\tau}_\pm$ and points of even lower height, but this process 'wraps' $\mathcal{T}$ three times around the smooth critical points, giving a multiplicity. Our combination of geometric and computer algebra based techniques allows us to determine the multiplicity. A direct geometric argument for this multiplicity is current unknown.

9.3.3 The State of Analytic Combinatorics in Several Variables

The methods of this chapter suggest a general strategy for analyzing Laurent coefficients of a function $F(\mathbf{z})$ with singular variety $\mathcal{V}$,

1. Determine a Whitney stratification of $\mathcal{V}$ and, for a fixed direction, compute the critical points of the height function on each strata using Definition 9.13.
2. Verify there are no stationary points at infinity using Definition 9.15 (or a generalization which takes into account non-smooth points).
3. Express, up to negligible error, the Cauchy integral for coefficients as an integer sum of integrals over neighbourhoods of critical points (as in Theorem 9.3).
4. Asymptotically approximate each of the integrals appearing in this expression.

As detailed in Baryshnikov et al. [6], Steps 1 and 2 are now essentially solved. Step 3 is a topological problem, and is typically the most difficult task of an analysis. Aside from the few special cases mentioned above, such as minimal contributing points, no general algorithm for Step 3 is known. Step 4 depends heavily on the local geometry of $\mathcal{V}$ at each critical point. For a critical point which is a non-degenerate transverse multiple point, the corresponding integrals are saddle-point integrals which can be asymptotically approximated. Asymptotics of integrals corresponding to degenerate transverse multiple points can in theory be computed [31], although such examples are not prominent in the analytic combinatorics literature.

Using results of Atiyah et al. [4] on inverse Fourier transforms of homogeneous hyperbolic functions, Baryshnikov and Pemantle [9] and Baryshnikov et al. [5] determine asymptotics for certain types of cone points. Explicit results for singular varieties with more exotic singular behaviour are harder to find, although much of the necessary background needed to attack such problems can be found in [2, 3].

In order to limit the necessary background, our presentation of this advanced approach to analytic combinatorics in several variables did not dig deeply into the underlying theory. A full discussion on such topics can be found in Baryshnikov et al. [6] and the textbook of Pemantle and Wilson [26].

Problems

9.1 Prove that if $A(x, y, z)$ and $B(x, y, z)$ are power series whose product is $z^2 - x^2 - y^2$ then one of A or B is invertible as a power series.

9.2 Prove that $\boldsymbol{\sigma} = (1, 1, 1/3)$ is a minimal point of

$$F(x, y, t) = \frac{(1 + x)(1 - xy^2 + x^2)}{(1 - t(1 + x^2 + xy^2))(1 - y)(1 + x^2)}.$$

Hint: Although F is not combinatorial a short argument allows one to apply Proposition 5.4 from Chapter 5. Alternatively, one can use the algorithms of Chapter 7.

9.3 Using Proposition 4.5 of Chapter 4 and Proposition 5.4 from Chapter 5, prove that all minimizers of $|xyt|^{-1}$ on the power series domain of convergence of

$$F(x, y, t) = \frac{(1 + x)(1 - xy^2 + x^2)}{(1 - t(1 + x^2 + xy^2))(1 - y)(1 + x^2)}$$

have the same coordinate-wise modulus as $\boldsymbol{\sigma} = (1, 1, 1/3)$.

9.4 The number of three-dimensional 'singular vector tuples of generic tensors' (as discussed in [25] and Problem 5.5 of Chapter 5) has multivariate generating function

$$F(x, y, z) = \frac{xyz}{(1 - x)(1 - y)(1 - z)(1 - x - y - z)(1 - xy - xz - yz)}.$$

Using algebraic relations between the denominator factors, decompose F as a sum of rational functions such that the denominator of each summand has at most three distinct linear factors. Use Theorems 9.1 and 9.2 to determine asymptotics along different directions $\mathbf{r} \in \mathbb{R}^d_{>0}$. For which directions do these theorems not apply?

9.5 Let $F(x, y) = \frac{1}{(1-x^2-y)(1-x-y^2)}$. Determine the directions to which Theorems 9.1 and 9.2 apply and determine asymptotics of the power series expansion of F in these directions.

9.6 In statistical signal processing, the design of multivariate autoregressive filters requires estimating series coefficients of certain *spectral density functions*. In this paradigm, Geronimo et al. [14] study the Laurent series coefficients of

$$F(x, y) = \frac{xy}{\left(1 - \frac{x+y}{C}\right)\left(xy - \frac{x+y}{C}\right)}$$

for a real parameter $C > 2$, expanded in the unique domain of convergence containing the point $(1, 1)$. Use Theorems 9.1 and 9.2 to find asymptotics in different directions, as a function of C.

9.7 Let $F(\mathbf{z}) = (1 - (z_1 + \cdots + z_d) + Cz_1 \cdots z_d)^{-1}$ and let $C_* = (d-1)^{d-1}$. Prove the singular variety $\mathcal{V}$ is smooth if and only if $C \neq C_*$. If $C < C_*$ show that there is a single smooth minimal critical point for the main diagonal direction, which has positive coordinates, and determine asymptotics of the main diagonal of F. You may use without proof the Grace-Walsh-Szegő theorem [11, Thm. 1.1]: if $f(\mathbf{z})$ is unchanged by permutations of the variables, f is linear in each variable individually, and $f(\mathbf{w}) = 0$ where each $|w_j| < r$ then there exists $|\omega| < r$ with $f(\omega, \omega, \ldots, \omega) = 0$.

9.8 In Section 9.3.2 we found asymptotics of a four variable rational function whose coefficient behaviour was determined by two non-minimal critical points. Consider the next smallest example in this family, the six variable rational function

$$F(u, v, w, x, y, z) = \frac{1}{1 - (u + v + w + x + y + z) + 3125uvwxyz}.$$

Show that there is a single non-smooth point, when all variables equal $1/5$, which is a cone point, and show that there are four (non-minimal) smooth critical points for the main diagonal. Assuming the results of Baryshnikov et al. [5], which imply the Cauchy integral can be deformed past the cone point, mirror the arguments of Section 9.3.2 to determine asymptotics of the main diagonal of F. You may also assume the Grace-Walsh-Szegő theorem from Problem 9.7. An annihilating differential equation for the diagonal can be computed using the MAGMA package of Lairez [18], or found on the textbook website. Note that, unlike the four-dimensional case, not all smooth critical points determine dominant asymptotics.

References

1. Andris Ambainis, Eric Bach, Ashwin Nayak, Ashvin Vishwanath, and John Watrous. One-dimensional quantum walks. In *Proceedings of the Thirty-Third Annual ACM Symposium on Theory of Computing*, pages 37–49. ACM, New York, 2001.
2. V. I. Arnold, S. M. Gusein-Zade, and A. N. Varchenko. *Singularities of differentiable maps. Volume 1*. Modern Birkhäuser Classics. Birkhäuser/Springer, New York, 2012. Classification of critical points, caustics and wave fronts, Translated from the Russian by Ian Porteous based on a previous translation by Mark Reynolds, Reprint of the 1985 edition.
3. V. I. Arnold, S. M. Gusein-Zade, and A. N. Varchenko. *Singularities of differentiable maps. Volume 2*. Modern Birkhäuser Classics. Birkhäuser/Springer, New York, 2012. Monodromy

and asymptotics of integrals, Translated from the Russian by Hugh Porteous and revised by the authors and James Montaldi, Reprint of the 1988 translation.
4. M. F. Atiyah, R. Bott, and L. Gårding. Lacunas for hyperbolic differential operators with constant coefficients. I. *Acta Math.*, 124:109–189, 1970.
5. Y. Baryshnikov, S. Melczer, and R. Pemantle. Asymptotics of multivariate sequences in the presence of a lacuna. *arXiv preprint, arXiv:1905.04187*, 2019. Submitted.
6. Y. Baryshnikov, S. Melczer, and R. Pemantle. Stationary points at infinity for analytic combinatorics. *arXiv preprint, arXiv:1905.05250*, 2020. Submitted.
7. Yuliy Baryshnikov, Wil Brady, Andrew Bressler, and Robin Pemantle. Two-dimensional quantum random walk. *J. Stat. Phys.*, 142(1):78–107, 2011.
8. Yuliy Baryshnikov, Stephen Melczer, Robin Pemantle, and Armin Straub. Diagonal asymptotics for symmetric rational functions via ACSV. In *29th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms*, volume 110 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages Art. No. 12, 15. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2018.
9. Yuliy Baryshnikov and Robin Pemantle. Asymptotics of multivariate sequences, part III: Quadratic points. *Adv. Math.*, 228(6):3127–3206, 2011.
10. Thomas Becker and Volker Weispfenning. *Gröbner bases*, volume 141 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1993. A computational approach to commutative algebra, In cooperation with Heinz Kredel.
11. Julius Borcea and Petter Brändén. The Lee-Yang and Pólya-Schur programs. II. Theory of stable polynomials and applications. *Comm. Pure Appl. Math.*, 62(12):1595–1631, 2009.
12. A. Bressler and R. Pemantle. Quantum random walks in one dimension via generating functions. In *2007 Conference on Analysis of Algorithms, AofA 07*, Discrete Math. Theor. Comput. Sci. Proc., AM, pages 403–414. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2007.
13. Timothy DeVries, Joris van der Hoeven, and Robin Pemantle. Automatic asymptotics for coefficients of smooth, bivariate rational functions. *Online J. Anal. Comb.*, (6):24, 2011.
14. Jeffrey S. Geronimo, Hugo J. Woerdeman, and Chung Y. Wong. Spectral density functions of bivariable stable polynomials. *Submitted*, 2020.
15. J. Gillis, B. Reznick, and D. Zeilberger. On elementary methods in positivity theory. *SIAM J. Math. Anal.*, 14(2):396–398, 1983.
16. Mark Goresky and Robert MacPherson. *Stratified Morse theory*, volume 14 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1988.
17. Lars Hörmander. *An introduction to complex analysis in several variables*, volume 7 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam, third edition, 1990.
18. Pierre Lairez. Computing periods of rational integrals. *Math. Comp.*, 85(300):1719–1752, 2016.
19. Arjen K. Lenstra. Factoring multivariate polynomials over algebraic number fields. *SIAM J. Comput.*, 16(3):591–598, 1987.
20. John Mather. Notes on topological stability. *Bull. Amer. Math. Soc. (N.S.)*, 49(4):475–506, 2012.
21. Marc Mezzarobba. Rigorous multiple-precision evaluation of D-finite functions in SageMath. Technical Report 1607.01967, arXiv, 2016.
22. J. Milnor. *Morse theory*. Based on lecture notes by M. Spivak and R. Wells. Annals of Mathematics Studies, No. 51. Princeton University Press, Princeton, N.J., 1963.
23. T. Mostowski and E. Rannou. Complexity of the computation of the canonical Whitney stratification of an algebraic set in $\mathbf{C}^n$. In *Applied algebra, algebraic algorithms and error-correcting codes (New Orleans, LA, 1991)*, volume 539 of *Lecture Notes in Comput. Sci.*, pages 281–291. Springer, Berlin, 1991.
24. David Mumford. *Algebraic geometry. I*. Springer-Verlag, Berlin-New York, 1976. Complex projective varieties, Grundlehren der Mathematischen Wissenschaften, No. 221.
25. Jay Pantone. The asymptotic number of simple singular vector tuples of a cubical tensor. *Online Journal of Analytic Combinatorics*, 12, 2017.

26. Robin Pemantle and Mark C. Wilson. *Analytic combinatorics in several variables*, volume 140 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2013.
27. I. Petrowsky. On the diffusion of waves and the lacunas for hyperbolic equations. *Rec. Math. [Mat. Sbornik] N. S.*, 17(59):289–370, 1945.
28. E. Rannou. The complexity of stratification computation. *Discrete Comput. Geom.*, 19(1):47–78, 1998.
29. Norbert Schappacher. Some milestones of lemniscatomy. In *Algebraic geometry (Ankara, 1995)*, volume 193 of *Lecture Notes in Pure and Appl. Math.*, pages 257–290. Dekker, New York, 1997.
30. Armin Straub and Wadim Zudilin. Positivity of rational functions and their diagonals. *J. Approx. Theory*, 195:57–69, 2015.
31. A. N. Varčenko. Newton polyhedra and estimates of oscillatory integrals. *Funkcional. Anal. i Priložen.*, 10(3):13–38, 1976.
32. V. A. Vassiliev. *Applied Picard-Lefschetz theory*, volume 97 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2002.

Chapter 10
Application: Lattice Paths, Revisited

We have continually to make our choice among different courses of action open to us, and upon the discretion with which we make it, in little matters and in great, depends our prosperity and our happiness.
— William A. Whitworth

Thus all actions have one or more of these seven causes: chance, nature, compulsion, habit, reasoning, passion, desire.
– Plato

In our final chapter we return to lattice path enumeration, combining the generating function expressions from Chapter 4 with the asymptotic results of Chapter 9. Once again, we see that lattice path models with similar combinatorial properties have generating functions encoded by multivariate diagonals with similar analytic properties, and can thus be treated uniformly.

First, Section 10.1 generalizes the results of Chapter 6 from lattice path models whose step sets are symmetric over ever axis to models whose step sets are symmetric over all but one axis[1]. This work, originally carried out in Melczer and Wilson [9], solved previous conjectures on the asymptotics of two-dimensional lattice path models restricted to a quadrant which had resisted purely univariate techniques. Our presentation of this material follows [9]. Section 10.2 then discusses further work from the lattice path literature where diagonal representations and analytic combinatorics in several variables have appeared. A collection of exercises from this literature is presented for the reader to hone their ACSV skills.

Although we reintroduce terminology as needed, the reader who hasn't yet gone through Chapter 4 is encouraged to do so. In particular, recall that for any dimension $d \in \mathbb{N}$ a lattice path model of dimension d is defined by a non-empty finite set of steps $\mathcal{S} \subset \mathbb{Z}^d$, a restricting region $\mathcal{R} \subset \mathbb{R}^d$, a starting point $\mathbf{p} \in \mathcal{R}$, and a terminal set $\mathcal{T} \subset \mathcal{R}$. The model defined by these parameters consists of all finite tuples $(\mathbf{s}_1, \dots, \mathbf{s}_r) \in \mathcal{S}^r$, called *walks* or *paths*, such that $\mathbf{p} + \mathbf{s}_1 + \cdots + \mathbf{s}_r \in \mathcal{T}$ and $\mathbf{p}+\mathbf{s}_1+\cdots+\mathbf{s}_k \in \mathcal{R}$ for all $1 \le k \le r$. We often associate a weight $w_\mathbf{s} > 0$ to each step $\mathbf{s} \in \mathcal{S}$ and enumerate walks by weight, where the weight of a walk $(\mathbf{s}_1, \dots, \mathbf{s}_r)$ is the product $w_{\mathbf{s}_1} \cdots w_{\mathbf{s}_r}$ of the weights of its steps. An unweighted model, where one simply counts walks by the number of steps they contain, is the same as the weighted model where each step has weight one. Unless otherwise stated the terminal set $\mathcal{T}$ of a model is taken to be the entire restricting region $\mathcal{R}$.

[1] Recall from Proposition 4.10 in Chapter 4 that in any dimension $d \ge 2$ there exists a lattice path model restricted to $\mathbb{N}^d$ whose set of steps is symmetric over all but *two* axes which cannot be written as the diagonal of a rational function.

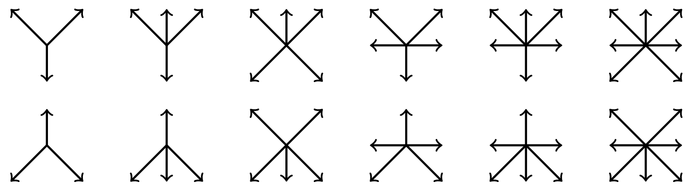


Fig. 10.1 The mostly symmetric step sets of dimension two, up to isomorphism. The step sets in the first row have positive drift, while those in the second row have negative drift.

10.1 Mostly Symmetric Models in an Orthant

We begin by studying walks defined by a weighted step set $\mathcal{S} \subset \{\pm 1, 0\}^d$, where each $\mathbf{s} \in \mathcal{S}$ is given a weight $w_\mathbf{s} > 0$, such that

- Walks on the step set can move forwards and backwards in each direction,

$$\text{For all } j = 1, \dots, d \text{ there exists a step } \mathbf{s} \in \mathcal{S} \text{ with } s_j = 1$$

$$\text{For all } j = 1, \dots, d \text{ there exists a step } \mathbf{s} \in \mathcal{S} \text{ with } s_j = -1$$

- Walks are restricted to $\mathcal{R} = \mathbb{N}^d$,
- The step set and weighting are symmetric over every axis except for one.

Recall from Chapter 4 that a lattice path model satisfying these conditions is called *mostly symmetric*, while a lattice path model defined by a step set that is symmetric over every axis is called *highly symmetric*. In Chapter 6 we gave explicit asymptotic results for highly symmetric models using the analytic methods of Chapter 5 and the fact that the generating function of a highly symmetric model can be represented by the diagonal of a rational function with a smooth singular variety. Now that we have developed analytic combinatorics in several variables for non-smooth points in Chapter 9, we will be able to derive asymptotics for mostly symmetric models.

Example 10.1 (Mostly Symmetric Models in the Quadrant)

Recall from Section 4.1.4 of Chapter 4 that many two-dimensional lattice path models in the quadrant $\mathbb{N}^2$ are isomorphic to each other (if their defining step sets differ by a reflection over the line $y = x$) or are isomorphic to half-space models (if one of the bounding axes is never interacting with). Figure 10.1 displays all mostly symmetric step sets of dimension two, up to reflection over the line $y = x$, after step sets corresponding to half-space models are removed.

As in previous chapters, for a variable z we adopt the notation $\overline{z} = 1/z$, which is common in the lattice path literature (we never use an overline to denote complex

Class	Exponential Growth	Order	Geometry	Covered By
highly symmetric	$S(\mathbf{1})$	$n^{-d/2}$	smooth	Theorem 6.1 in Ch. 5
positive drift	$S(\mathbf{1})$	$n^{1/2-d/2}$	non-smooth	Theorem 10.1
negative drift	$< S(\mathbf{1})$	$n^{-1-d/2}$	smooth	Theorem 10.2

Table 10.1 Summary of our asymptotic results, including whether or not the exponential growth is the same as the number of unrestricted walks, and whether the contributing singularities in our asymptotic analysis are smooth or non-smooth points.

conjugation). Without loss of generality we may assume the axis of non-symmetry for a mostly symmetric model corresponds to its final coordinate, meaning the *(weighted) characteristic polynomial*

$$S(\mathbf{z}) = \sum_{\mathbf{s}\in\mathcal{S}} w_{\mathbf{s}}\mathbf{z}^{\mathbf{s}}$$

satisfies $S(z_1, \dots, z_{j-1}, \overline{z}_j, z_{j+1}, \dots, z_d) = S(\mathbf{z})$ for all $1 \leq j \leq d-1$. Because the step sets under consideration are subsets of $\{\pm 1, 0\}^d$, this implies we can write

$$S(\mathbf{z}) = \overline{z}_d A(\hat{\mathbf{z}}) + Q(\hat{\mathbf{z}}) + z_d B(\hat{\mathbf{z}}) \tag{10.1}$$

for Laurent polynomials $A(\hat{\mathbf{z}}), Q(\hat{\mathbf{z}})$, and $B(\hat{\mathbf{z}})$ symmetric in $\hat{\mathbf{z}} = (z_1, \dots, z_{d-1})$.

A mostly symmetric step set only has one unrestricted coordinate, so the properties of the step set in this coordinate are crucial to its asymptotic behaviour. In particular, we will see that models share a similar asymptotic template depending on whether more steps move towards or away from the boundary axes. This concept is captured by the following definition.

Definition 10.1 (drift for mostly symmetric models) Given a mostly symmetric model with characteristic polynomial $S(\mathbf{z})$ satisfying (10.1), the *drift* of the model is the real number $\delta = B(\mathbf{1}) - A(\mathbf{1})$, which is the total weight of steps with positive dth coordinate minus the total weight of steps with negative dth coordinate. The model has *negative, zero, or positive drift* depending on whether $\delta < 0$, $\delta = 0$, or $\delta > 0$.

Although we will make use of a uniform diagonal expression for all mostly symmetric models, the local geometry of the singular variety at the contributing singularities (in particular, whether the contributing singularities are smooth or non-smooth transverse multiple points) will depend on the drift of the model. Table 10.1 outlines the different cases that we cover. As can be expected, when the total weight of a step set is fixed then negative drift models, where walks tend towards the boundary axes, grow the slowest (in fact, exponentially slower than models on the same set of steps that are not restricted to an orthant) while positive drift models grow the fastest. Our analytic calculations will explicitly show how the drift comes into play when determining asymptotics.

We now describe our various results in more detail, before proving them.

Asymptotics of Positive Drift Models

For each $1 \leq k \leq d-1$ let $b_k = \sum_{\mathbf{s}\in\mathcal{S}, i_k=1} w_\mathbf{s}$ be the total weight of the steps moving forwards in the kth coordinate (which by symmetry is also the total weight of steps moving backwards in the kth coordinate).

Theorem 10.1 (Positive Drift Asymptotics) *Let $\mathcal{S} \subset \{-1,0,1\}^d$ be a weighted set of steps that is symmetric over all but one axis and moves forwards and backwards in each coordinate. If the model defined by $\mathcal{S}$ has positive drift then the number c_n of walks taking n steps in $\mathcal{S}$, beginning at the origin, and never leaving $\mathbb{N}^d$ satisfies*

$$c_n = S(\mathbf{1})^n n^{1/2-d/2}\left[\left(1-\frac{A(\mathbf{1})}{B(\mathbf{1})}\right)\left(\frac{S(\mathbf{1})}{\pi}\right)^{\frac{d-1}{2}}\frac{1}{\sqrt{b_1\cdots b_{d-1}}}\right]\left(1+O\left(n^{-1}\right)\right). \quad (10.2)$$

As with our results in Chapter 6 on highly symmetric models, Theorem 10.1 can be instantly applied to any given model, and families of models in varying dimension.

Example 10.2 (A Family of Positive Drift Models)

For any $d \in \mathbb{N}$ consider the (unweighted) model with characteristic polynomial

$$S(\mathbf{z}) = \overline{z}_d + z_d\prod_{j<d}(z_j+\overline{z}_j),$$

so that $\mathcal{S}$ has a single step with negative dth coordinate. Then Theorem 10.1 implies

$$c_n = \left(1+2^{d-1}\right)^n n^{1/2-d/2}\left[\frac{2^{d-1}-1}{(2^d\pi)^{(d-1)/2}}\right]\left(1+O\left(n^{-1}\right)\right).$$

This lattice path model can be viewed a queuing system with d servers where at every discrete time step either the final server processes a job, or a job is added onto the queue for the final server and the first $d-1$ servers each simultaneously receive or process one additional job (when there is at least one job to be processed).

Theorem 10.1 is proven in Section 10.1.2.

Asymptotics of Negative Drift Models

Let $\rho = \sqrt{A(\mathbf{1})/B(\mathbf{1})}$ and for each $1 \leq k \leq d-1$ let $b_k(\mathbf{z}_{\hat{k}}) = [z_k]S(\mathbf{z})$. We will soon see that unlike the positive drift case, when there is always a unique contributing singularity, under certain circumstances asymptotics of negative drift models will be determined by two contributing singularities. With that in mind, we define

$$C_\rho = \frac{S(\mathbf{1},\rho)\,\rho}{2\,\pi^{d/2}\,A(\mathbf{1})(1-1/\rho)^2}\sqrt{\frac{S(\mathbf{1},\rho)^d}{\rho\,b_1(\mathbf{1},\rho)\cdots b_{d-1}(\mathbf{1},\rho)B(\mathbf{1})}}$$

and let $C_{-\rho}$ be the constant obtained by replacing ρ by $-\rho$ in C_ρ. The term under the square-root in C_ρ and $C_{-\rho}$ will always be positive when it appears in our formulas.

Theorem 10.2 (Negative Drift Asymptotics) *Let $\mathcal{S} \subset \{-1,0,1\}^d$ be a weighted set of steps that is symmetric over all but one axis and moves forwards and backwards in each coordinate. If the model defined by $\mathcal{S}$ has negative drift and $Q(\hat{\mathbf{z}}) \neq 0$ (i.e., if there are steps in $\mathcal{S}$ whose dth coordinate is zero) then the number c_n of walks taking n steps in $\mathcal{S}$, beginning at the origin, and never leaving the orthant $\mathbb{N}^d$ satisfies*

$$c_n = S(\mathbf{1},\rho)^n n^{-1-d/2} C_\rho \left(1 + O\left(n^{-1}\right)\right).$$

If the model defined by $\mathcal{S}$ has negative drift and $Q(\hat{\mathbf{z}})$ is identically zero then

$$c_n = \left[S(\mathbf{1},\rho)^n n^{-1-d/2} C_\rho + S(\mathbf{1},-\rho)^n n^{-1-d/2} C_{-\rho}\right]\left(1 + O\left(n^{-1}\right)\right).$$

Example 10.3 (A Family of Negative Drift Models)

For any $d \in \mathbb{N}$ consider the (unweighted) model with characteristic polynomial

$$S(\mathbf{z}) = \overline{z}_d \prod_{j<d} (z_j + \overline{z}_j) + z_d,$$

whose set of steps is the negation of our last example. A quick calculation shows

$$\rho = 2^{(d-1)/2}, \quad C_\rho = \frac{2^{2d-3/2}}{\pi^{d/2}\left(2^{(d-1)/2}-1\right)^2}, \quad \text{and} \quad C_{-\rho} = \frac{2^{2d-3/2}}{\pi^{d/2}\left(2^{(d-1)/2}+1\right)^2},$$

so Theorem 10.2 implies

$$c_n = \left(2^{(d+1)/2}\right)^n n^{-1-d/2}\left[\frac{2^{2d-3/2}\kappa_n}{\pi^{d/2}\left(2^{d-1}-1\right)^2}\right]\left(1 + O\left(n^{-1}\right)\right),$$

where κ_n equals $2^d + 2$ when n is even and $2^{(d+3)/2}$ when n is odd. This model can be viewed a queuing system with d servers where at every discrete time step either the final server adds a job to its queue, or the final server processes a job from its queue and the first $d-1$ servers each simultaneously receive or process one additional job.

Theorem 10.2 is proven in Section 10.1.3.

Zero Drift Models

The zero drift models break down into two sub-cases, one easy to handle and the other more difficult. First, if $A(\hat{\mathbf{z}}) = B(\hat{\mathbf{z}})$ then the model with characteristic polynomial $S(\mathbf{z})$ is highly symmetric and asymptotics for the number of walks are determined in Chapter 6. For unweighted models in dimension two, the motivating case for much of our asymptotic work, this is the only situation that can occur. Unfortunately, for unweighted models in dimension three and greater, and weighted models in dimension two, it is possible for a step set to have zero drift and not be highly symmetric[2]. The diagonal representations in Chapter 4 for generating functions of mostly symmetric models have particularly nasty analytic behaviour for non-highly symmetric models with zero drift (similar to our discussion of asymptotics along non-generic directions in Chapter 8). Because of this additional difficulty we do not cover the zero drift models in detail.

10.1.1 Diagonal Expressions and Contributing Points

Our asymptotic results follow from the diagonal expressions derived in Chapter 4. First, we note that Proposition 4.8 of Chapter 4 states that the generating function enumerating walks in a mostly symmetric lattice path model by length satisfies

$$C(t) = \Delta\left(\frac{(1+z_1)\cdots(1+z_{d-1})\left(B(\hat{\mathbf{z}}) - z_d^2 A(\hat{\mathbf{z}})\right)}{(1-z_d)B(\hat{\mathbf{z}})(1-tz_1\cdots z_d\overline{S}(\mathbf{z}))}\right), \tag{10.3}$$

where the rational function is expanded in the ring $\mathcal{R} = \mathbb{Q}((\mathbf{z}))[[t]]$, the diagonal operator Δ was introduced in Definition 3.23 of Chapter 3, and

$$\overline{S}(\mathbf{z}) = S(z_1, \dots, z_{d-1}, \overline{z}_d).$$

Equation (10.3) was used in Chapter 9 to derive asymptotics for two mostly symmetric lattice path models, one of which was the running example for that chapter.

Unfortunately, the presence of the factor $B(\hat{\mathbf{z}})$ in the denominator of (10.3) can make the contributing singularities that determine diagonal asymptotics non-minimal. Thus, to simplify our asymptotic arguments we use the alternative diagonal expression given by Proposition 4.9 of Chapter 4. Proposition 4.9 states that the generating function enumerating walks in a mostly symmetric lattice path model by length is the main power series diagonal of $F(\mathbf{z}, t) = G(\mathbf{z}, t)/H(\mathbf{z}, t)$, where

[2] For example, the mostly symmetric three-dimensional model with $A(x, y) = x + \overline{x} + y + \overline{y}$, $B(x, y) = x/y + y/x + xy + 1/(xy)$, and $Q(x, y) = 0$ has zero drift but is not highly symmetric.

$$\begin{aligned} G(\mathbf{z},t) &= (1+z_1)\cdots(1+z_{d-1})\,(1-tz_1\cdots z_d\,(Q(\hat{\mathbf{z}})+2z_dA(\hat{\mathbf{z}}))) \\ H(\mathbf{z},t) &= \left(1-tz_1\cdots z_d\overline{S}(\mathbf{z})\right)(1-tz_1\cdots z_d\,(Q(\hat{\mathbf{z}})+z_dA(\hat{\mathbf{z}})))\,(1-z_d). \end{aligned} \tag{10.4}$$

Note that under our assumptions G and H are coprime as polynomials, so the singularities of F form the algebraic set $\mathcal{V} = \mathcal{V}(H)$. Following the setup of Chapter 9, we define the factors

$$\begin{aligned} H_1 &= 1 - tz_1\cdots z_d\overline{S}(\mathbf{z}) \\ H_2 &= 1 - tz_1\cdots z_d\,(Q(\hat{\mathbf{z}}) + z_dA(\hat{\mathbf{z}})) \\ H_3 &= 1 - z_d \end{aligned}$$

and the flats $\mathcal{V}_1 = \mathcal{V}(H_1), \mathcal{V}_2 = \mathcal{V}(H_2)$, and $\mathcal{V}_3 = \mathcal{V}(H_3)$.

Remark 10.1 One may be tempted to prove that $H = H_1H_2H_3$ is a transverse polynomial factorization of H, in the sense of Definition 9.3 from Chapter 9, but this is not always the case. For instance, in the three-dimensional model with characteristic polynomial $S(x,y,z) = \overline{z} + z(x + \overline{x} + y + \overline{y})$ the point $\mathbf{w} = (1,-1,1,-1)$ satisfies $H_1(\mathbf{w}) = H_2(\mathbf{w}) = 0$ and $(\nabla H_1)(\mathbf{w}) = (\nabla H_2)(\mathbf{w})$. In this example $\mathbf{w}$ is not even a transverse multiple point! Such pathological points, if they exist, have lower height than the singularities determining asymptotics and thus will not affect our methods.

Our first goal is to find the finite set of contributing singularities where a local analysis of F gives diagonal asymptotics. Because we are considering the main diagonal direction, we begin by characterizing the minimal points of F (the elements of $\mathcal{V}$ coordinate-wise closest to the origin) that minimize the product $|z_1\cdots z_dt|^{-1}$. The general case will follow from a study of the singularities with positive coordinates.

Proposition 10.1 *Let $\mathcal{S} \subset \{-1,0,1\}^d$ be a weighted set of steps that is symmetric over all but one axis and moves forwards and backwards in each coordinate. Then $|z_1\cdots z_dt|^{-1}$ is minimized at a unique minimal singularity of $F(\mathbf{z})$ with positive coordinates, equal to*

$$\mathbf{p}_1 = \left(1,1,\ldots,1,\sqrt{\frac{B(\mathbf{1})}{A(\mathbf{1})}},\frac{\sqrt{A(\mathbf{1})/B(\mathbf{1})}}{2\sqrt{A(\mathbf{1})B(\mathbf{1})}+Q(\mathbf{1})}\right) \quad \textit{when the drift is negative, and}$$

$$\mathbf{p}_2 = \left(1,1,\ldots,1,\frac{1}{S(\mathbf{1})}\right) \quad \textit{otherwise.}$$

Proof Because $|z_1\cdots z_dt|^{-1}$ decreases as $(\mathbf{z},t)$ moves away from the origin, any minimizer among the minimal points of F must be an element of $\mathcal{V}_1$ or $\mathcal{V}_2$. If

$$\begin{aligned} P_1(\mathbf{z}) &= z_1\cdots z_d\,(\overline{z}_dB(\hat{\mathbf{z}}) + Q(\hat{\mathbf{z}}) + z_dA(\hat{\mathbf{z}})) = z_1\cdots z_d\overline{S}(\mathbf{z}) \\ P_2(\mathbf{z}) &= z_1\cdots z_d\,(Q(\hat{\mathbf{z}}) + z_dA(\hat{\mathbf{z}})) \end{aligned}$$

then the points in $\mathcal{V}_1$ and $\mathcal{V}_2$ are characterized by $t = 1/P_1(\mathbf{z})$ and $t = 1/P_2(\mathbf{z})$, respectively. Since P_1 and P_2 are polynomials with positive coordinates, and the terms appearing in P_2 are a strict subset of the terms appearing in P_1, we see that $\mathcal{V}_2$ contains no minimal points with positive coordinates (the value of t on $\mathcal{V}_1$ is always less than the value of t on $\mathcal{V}_2$ when $\mathbf{z}$ with positive coordinates is fixed). Furthermore, increasing any coordinate of $\mathbf{z} \in \mathbb{R}^d_{>0}$ decreases the value of $t = 1/P_1(\mathbf{z})$, so the minimal points of $\mathcal{V}_1$ form the set

$$M = \left\{ \left(\mathbf{z}, P_1(\mathbf{z})^{-1} \right) : \mathbf{z} \in \mathbb{R}^d_{>0} \text{ and } z_d \leq 1 \right\},$$

where the inequality on z_d accounts for the factor $H_3 = 1 - z_d$.

In other words, we want to minimize $|z_1 \cdots z_d t|^{-1}$ for $(\mathbf{z}, t) \in M$, which is equivalent to minimizing $|\overline{z}_1 \cdots \overline{z}_d P_1(\mathbf{z})| = \overline{S}(\mathbf{z})$ over points with positive coordinates where $z_d \leq 1$. Ignoring the constraint on z_d, Proposition 4.5 of Chapter 4 implies that $\overline{S}(\mathbf{z})$ has a unique minimizer with positive coordinates, given by the solution of

$$\overline{S}_{z_1}(\mathbf{z}) = \cdots = \overline{S}_{z_d}(\mathbf{z}) = 0$$

with positive coordinates (as usual subscripted variables refer to partial derivatives). Simplifying this system to

$$\left(1 - z_1^2\right) \left[z_1^{-1}\right] \overline{S}(\mathbf{z}) = \cdots = \left(1 - z_{d-1}^2\right) \left[z_{d-1}^{-1}\right] \overline{S}(\mathbf{z}) = B(\hat{\mathbf{z}}) - z_d^2 A(\hat{\mathbf{z}}) = 0,$$

and noting that $\overline{S}$ has non-negative coefficients, we see that the desired minimizer of $\overline{S}(\mathbf{z})$ is $(\mathbf{1}, \sqrt{B(\mathbf{1})/A(\mathbf{1})})$. If the drift of the model is positive or zero then $B(\mathbf{1}) \leq A(\mathbf{1})$ and we have shown that $\mathbf{p}_1$ is the unique minimizer of $|z_1 \cdots z_d t|^{-1}$ among the minimal singularities of F with positive coordinates.

If the drift of the model is negative then $B(\mathbf{1}) > A(\mathbf{1})$, so $\mathbf{p}_1$ is not a minimal point. In this case the desired minimizer of $\overline{S}(\mathbf{z})$ must occur when $z_d = 1$ and $\hat{\mathbf{z}}$ minimizes $\overline{S}(\hat{\mathbf{z}}, 1)$ among points with positive coordinates. Once again invoking Proposition 4.5 of Chapter 4, this time on $\overline{S}(\hat{\mathbf{z}}, 1)$, shows that the minimizer occurs when $\hat{\mathbf{z}} = \mathbf{1}$. Thus, in the negative drift case $\mathbf{p}_2$ is the unique minimizer of $|z_1 \cdots z_d t|^{-1}$ among the minimal singularities of F with positive coordinates. □

Remark 10.2 The non-smooth point $\mathbf{p}_2$ is always a minimal point, but it is only a minimizer of $|z_1 \cdots z_d t|^{-1}$ among the minimal points when the drift is negative so that the smooth point $\mathbf{p}_1$ is non-minimal. In the zero drift case the points $\mathbf{p}_1$ and $\mathbf{p}_2$ coincide. Recall also Figures 9.3 and 9.4 in Chapter 9: although these figures illustrate the singular sets of a different diagonal encoding of lattice path generating functions, the pictures are similar to our situation.

One nice thing about the diagonal representation under consideration is that the power series expansion of $1/H(\mathbf{z}, t)$ is *combinatorial*, meaning the series coefficients are all non-negative. Lemma 5.7 in Chapter 5 states that the minimal singularities of a combinatorial series are characterized by the minimal singularities with positive real coordinates. Thus, we can extend Proposition 10.1 to obtain an explicit classification of the singularities that dictate lattice path asymptotics. The formal definitions of critical points and contributing singularities can be found in Chapter 5 for smooth points and Chapter 9 for transverse multiple points; intuitively, contributing singularities can simply be viewed as those where the local behaviour of F determines asymptotics of its diagonal coefficient sequence. We always discuss critical and contributing points with respect to the main diagonal direction $\mathbf{r} = \mathbf{1}$.

Theorem 10.3 (mostly symmetric contributing points) *Consider the lattice path model defined by a weighted set of steps $\mathcal{S} \subset \{-1, 0, 1\}^d$ that is symmetric over all but one axis and moves forwards and backwards in each coordinate. Let $\mathcal{D}$ denote the power series domain of convergence of $F(\mathbf{z})$ and let $\mathbf{p}_1$ and $\mathbf{p}_2$ denote the points from Proposition 10.1.*

- *If the drift of the model is negative then every point of $\mathcal{V} \cap T(\mathbf{p}_1)$ is a minimal smooth point and this set contains at most 2^{d+1} points, given by $(\widehat{\mathbf{w}}, w_d, t)$ where*

$$\widehat{\mathbf{w}} \in \{\pm 1\}^{d-1}, \qquad w_d = \nu \sqrt{\frac{B(\widehat{\mathbf{w}})}{A(\widehat{\mathbf{w}})}}, \qquad t = \frac{1}{w_d S\left(\widehat{\mathbf{w}}, \overline{w}_d\right)},$$

$$|w_d| = \sqrt{\frac{B(\mathbf{1})}{A(\mathbf{1})}}, \quad \textit{and} \quad |t| = \frac{\sqrt{A(\mathbf{1})}}{\sqrt{B(\mathbf{1})}\, S\left(\mathbf{1}, \sqrt{A(\mathbf{1})/B(\mathbf{1})}\right)},$$

  with ν a fourth root of unity[3]. If w_d is imaginary at such a point then $\widehat{\mathbf{w}}$ has at least one coordinate equal to -1. Each of these points is a smooth critical point. When the drift of the model is negative then these points are contributing singularities which form the minimizers of $|z_1 \cdots z_d t|^{-1}$ on $\overline{\mathcal{D}}$.

- *Every point of $\mathcal{V} \cap T(\mathbf{p}_2)$, which contains the elements of $\mathcal{V}$ with the same coordinate-wise modulus as $\mathbf{p}_2$, is a transverse multiple point. The set $\mathcal{V} \cap T(\mathbf{p}_2)$ contains at most 2^{d-1} critical points of F, given by the points $(\widehat{\mathbf{w}}, 1, t)$ where*

$$\widehat{\mathbf{w}} \in \{\pm 1\}^{d-1}, \quad t = \frac{1}{w_1 \cdots w_{d-1} S\left(\widehat{\mathbf{w}}, 1\right)}, \quad \textit{and} \quad |t| = \frac{1}{S(\mathbf{1}, 1)}.$$

  When the drift of the model is positive then each of these points is a contributing singularity, and these points are the minimizers of $|z_1 \cdots z_d t|^{-1}$ on $\overline{\mathcal{D}}$.

- *If the drift of the model is zero then $\mathbf{p}_1 = \mathbf{p}_2$ and every point of $\mathcal{V} \cap T(\mathbf{p}_1)$ is a transverse multiple point. Under this assumption of zero drift the set $\mathcal{V} \cap T(\mathbf{p}_1)$ contains at most 2^d elements, given by the points $(\widehat{\mathbf{w}}, w_d, t)$ where*

$$(\widehat{\mathbf{w}}, w_d) \in \{\pm 1\}^d, \qquad t = \frac{1}{w_1 \cdots w_d S\left(\widehat{\mathbf{w}}, w_d\right)}, \textit{ and } \quad |t| = \frac{1}{S(\mathbf{1}, 1)},$$

  and these points are the minimizers of $|z_1 \cdots z_d t|^{-1}$ on $\overline{\mathcal{D}}$.

[3] In order to satisfy the condition on $|t|$ it must be that $B(\widehat{\mathbf{w}})/A(\widehat{\mathbf{w}}) > 0$, so the square root can be taken unambiguously.

Proof Suppose first that the drift is negative. Since $1/H(\mathbf{z}, t)$ is combinatorial, Corollaries 5.4 and 5.5 from Chapter 5 imply the minimizers of $|z_1 \cdots z_d t|^{-1}$ on $\overline{\mathcal{D}}$ will be the points in $\mathcal{V} \cap T(\mathbf{p}_1)$, and all these points are smooth critical points (and thus also contributing points). If $(\mathbf{w}, s) \in \mathcal{V}_1$ then s is uniquely determined by $\mathbf{w}$ and if $(\mathbf{w}, s)$ has the same coordinate-wise modulus as $\mathbf{p}_1$ then $\overline{S}(\mathbf{w}) = \overline{S}(\widehat{\mathbf{p}_1})$. Because $z_1 \cdots z_d \overline{S}(\mathbf{z})$ is a polynomial with nonnegative coefficients, and every coordinate of $\widehat{\mathbf{p}_1}$ is positive, this equality implies every monomial of $\overline{S}(\mathbf{z})$ has the same argument at $\mathbf{z} = \mathbf{w}$. Symmetry over the first $d-1$ axes then implies $w_1, \ldots, w_{d-1}$ are real, and thus lie in $\{\pm 1\}$ as they have modulus one. Our condition that $\mathcal{S}$ moves forwards and backwards in each coordinate then implies $\overline{S}(\mathbf{z})$ evaluated at $\mathbf{z} = \mathbf{w}$ contains two monomials of the form $\pm w_d$ and $\pm w_d^{-1}$, so w_d is a fourth root of unity with modulus $\sqrt{B(\mathbf{1})/A(\mathbf{1})}$. The value of w_d is non-real only if one of the coefficients of the terms w_d and w_d^{-1} is negative, meaning $w_k = -1$ for at least one $k \neq d$.

The proof of Proposition 10.1 implies no point of $T(\mathbf{p}_2)$ lies in $\mathcal{V}_2$, and the gradients of H_1 and H_3 are non-zero and linearly-independent on $\mathcal{V}$, so every point of $\mathcal{V} \cap T(\mathbf{p}_2)$ is a transverse multiple point. An argument analogous to the negative drift case shows that any point in $\mathcal{V} \cap T(\mathbf{p}_2)$ satisfies $\widehat{\mathbf{w}} \in \{\pm 1\}^{d-1}$. Suppose now that $(\mathbf{w}, t_{\mathbf{w}})$ is one of the points in the statement of the theorem for the positive drift case. If $1 \leq k \leq d-1$ then

$$w_k(\partial H_1/\partial z_k)(\mathbf{w}) = -t_{\mathbf{w}} w_1 \cdots w_d \overline{S}(\mathbf{w}) - w_k(t_{\mathbf{w}} w_1 \cdots w_d)(\partial \overline{S}/\partial z_k)(\mathbf{w}) = -1,$$

since $t z_1 \cdots z_d \overline{S}(\mathbf{z}) = 1$ on $\mathcal{V}_1$ and $(\partial \overline{S}/\partial z_k)(\mathbf{w}) = 0$ when $w_k \in \{\pm 1\}$. Similarly,

$$w_d(\partial H_1/\partial z_d)(\mathbf{w}) = -1 - \frac{A(\mathbf{w}) - B(\mathbf{w})}{\overline{S}(\mathbf{w})} = -1 - \frac{A(\mathbf{1}) - B(\mathbf{1})}{\overline{S}(\mathbf{1})} = -\frac{2A(\mathbf{1}) + Q(\mathbf{1})}{A(\mathbf{1}) + Q(\mathbf{1}) + B(\mathbf{1})}$$

since $\overline{S}(\mathbf{z}) = z_d A(\hat{\mathbf{z}}) + Q(\hat{\mathbf{z}}) + \overline{z}_d B(\hat{\mathbf{z}})$ and our hypotheses imply $A(\mathbf{w})$, $B(\mathbf{w})$, and $Q(\mathbf{w})$ differ by positive real factors. Thus, when there is a positive drift

$$-(\nabla_{\log} H_1)(\mathbf{w}, t_{\mathbf{w}}) = (1, \ldots, 1, r, 1) \qquad \text{and} \qquad -(\nabla_{\log} H_3)(\mathbf{w}, t_{\mathbf{w}}) = (0, \ldots, 0, 1, 0)$$

for some $0 < r < 1$, and $(\mathbf{w}, t_{\mathbf{w}})$ is a contributing point. □

Having characterized the contributing singularities, we are now ready to determine asymptotics. In order to complete the necessary calculations we require some additional notation.

Definition 10.2 For any $1 \leq j \leq d$ let ∂_j denote the partial differential operator $\partial/\partial\theta_j$. Furthermore, for any $1 \leq k \leq d-1$ define B_k and Q_k as the unique Laurent polynomials such that

$$\overline{S}(\mathbf{z}) = (z_k + \overline{z}_k) B_k(\mathbf{z}_{\widehat{k}}) + Q_k(\mathbf{z}_{\widehat{k}}), \tag{10.5}$$

and let A_k', B_k', A_k'', B_k'' be the unique Laurent polynomials such that

$$A(\hat{\mathbf{z}}) = (z_k + \overline{z}_k)A'_k(\mathbf{z}_{\widehat{k}}) + A''_j(\mathbf{z}_{\widehat{k}})$$
$$B(\hat{\mathbf{z}}) = (z_k + \overline{z}_k)B'_k(\mathbf{z}_{\widehat{k}}) + B''_j(\mathbf{z}_{\widehat{k}}).$$

To simplify notation we occasionally write $B_k(\mathbf{p})$ for $B_k(\mathbf{p}_{\hat{k}})$, understanding that the kth entry of $\mathbf{p}$ should be removed, and similarly for A'_k, B'_k, A''_k, B''_k.

10.1.2 Asymptotics for Positive Drift Models

Asymptotics for the number of walks in a positive drift model follow directly from Theorem 9.2 and Corollary 9.1 in Chapter 9, which describe the asymptotic contributions of transverse multiple points. In particular, the numerator of our diagonal expression vanishes at all contributing points except for $\mathbf{p}_2 = (\mathbf{1}, 1/S(\mathbf{1}))$, which is thus the only contributing singularity affecting dominant asymptotics.

Proof (of Theorem 10.1) First, we note that

$$\frac{G(\mathbf{p}_2)}{H_2(\mathbf{p}_2)} = 2^{d-1}\frac{B(\mathbf{1}) - A(\mathbf{1})}{S(\mathbf{1})}\,\frac{S(\mathbf{1})}{B(\mathbf{1})} = 2^{d-1}\left(1 - \frac{A(\mathbf{1})}{B(\mathbf{1})}\right).$$

Next, the calculations in the proof of Theorem 10.3 imply that the matrix $\Gamma_{\mathbf{p}_2}$ in the asymptotic expansion (9.7) given by Theorem 9.2 equals

$$\Gamma_{\mathbf{p}_2} = \begin{pmatrix} -1 & -1 & -1 & \cdots & -1 & -r & -1 \\ 0 & 0 & 0 & \cdots & 0 & -1 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 & 0 \end{pmatrix}$$

for a real number $0 < r < 1$. Finally, to calculate the matrix $\mathcal{Q}_{\mathbf{p}_2}$ appearing in (9.7) we parametrize $\mathcal{V}_1 \cap \mathcal{V}_3$ by $z_1, \ldots, z_{d-1}$ using $z_d = 1$ and $t = \left(z_1 \cdots z_{d-1}\overline{S}(\hat{\mathbf{z}}, 1)\right)^{-1}$. This implies $\mathcal{Q}_{\mathbf{p}_2}$ is the Hessian of the function

$$g(\widehat{\boldsymbol{\theta}}) = \log\left(\frac{1}{e^{\mathrm{i}(\theta_1+\cdots+\theta_{d-1})}\overline{S}\left(e^{\mathrm{i}\widehat{\boldsymbol{\theta}}}\right)}, 1\right) + \log 1 = -\log \overline{S}\left(e^{\mathrm{i}\widehat{\boldsymbol{\theta}}}, 1\right) - \mathrm{i}(\theta_1 + \cdots + \theta_{d-1})$$

at the origin. We claim

$$Q_{\mathbf{p}_2} = \begin{pmatrix} \frac{2B_1(\mathbf{1})}{S(\mathbf{1})} & 0 & 0 & \cdots & 0 \\ 0 & \frac{2B_2(\mathbf{1})}{S(\mathbf{1})} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \ddots & \frac{2B_{d-2}(\mathbf{1})}{S(\mathbf{1})} & 0 \\ 0 & 0 & \cdots & 0 & \frac{2B_{d-1}(\mathbf{1})}{S(\mathbf{1})} \end{pmatrix},$$

so that substitution into the asymptotic expansion (9.7) of Theorem 9.2 gives Theorem 10.1. To prove this claim let $\tilde{S}(\boldsymbol{\theta}) = \overline{S}(\mathbf{p}_2 e^{\mathrm{i}\boldsymbol{\theta}})$ and for $1 \leq k \leq d-1$ define

$$\tilde{B}(\boldsymbol{\theta}_{\hat{k}}) = B_k\left(\mathbf{p}_{\hat{k}} e^{\mathrm{i}\boldsymbol{\theta}_{\hat{k}}}\right) \quad \text{and} \quad \tilde{Q}(\boldsymbol{\theta}_{\hat{k}}) = Q_k\left(\mathbf{p}_{\hat{k}} e^{\mathrm{i}\boldsymbol{\theta}_{\hat{k}}}\right).$$

Writing $\tilde{S}(\boldsymbol{\theta}) = (p_k e^{\mathrm{i}\theta_k} + \overline{p_k} e^{-\mathrm{i}\theta_k})\tilde{B}(\boldsymbol{\theta}_{\hat{k}}) + \tilde{Q}(\boldsymbol{\theta}_{\hat{k}})$ shows

$$(\partial_k \tilde{S})(\boldsymbol{\theta}) = \mathrm{i}(p_k e^{\mathrm{i}\theta_k} - \overline{p_k} e^{-\mathrm{i}\theta_k})\tilde{B}(\boldsymbol{\theta}_{\hat{k}})$$

and, since $p_k = \overline{p_k}$ under our assumptions, it follows that the partial derivatives of $\tilde{S}$ satisfy $(\partial_k \tilde{S})(\mathbf{0}) = 0$, $(\partial_k^2 \tilde{S})(\mathbf{0}) = -2p_k B_k(\mathbf{p}_{\hat{k}})$, and $(\partial_j \partial_k \tilde{S})(\mathbf{0}) = 0$ whenever $j \neq k$. Writing

$$\tilde{S}(\boldsymbol{\theta}) = p_d e^{i\theta_d} A\left(\hat{\mathbf{p}} e^{\mathrm{i}\widehat{\boldsymbol{\theta}}}\right) + Q\left(\hat{\mathbf{p}} e^{\mathrm{i}\widehat{\boldsymbol{\theta}}}\right) + \overline{p_d} e^{-i\theta_d} B\left(\hat{\mathbf{p}} e^{\mathrm{i}\widehat{\boldsymbol{\theta}}}\right),$$

partial derivatives involving θ_d follow from a similar argument. The chain rule then expresses the partial derivatives of g in terms of the partial derivatives of $\tilde{S}$, giving the claimed value of $Q_{\mathbf{p}_2}$ and completing the argument. □

10.1.3 Asymptotics for Negative Drift Models

Asymptotics for the number of walks in a negative drift model are determined by smooth minimal contributing singularities using Theorem 5.2 and Corollary 5.2 from Chapter 5. Smoothness simplifies the analysis, but the leading term in the expansion of Theorem 5.2 vanishes at all contributing points, meaning higher order constants must be determined.

With this in mind, let W be the set of contributing points of $\mathcal{V} \cap T(\mathbf{p}_1)$ for a negative drift model, described in Theorem 10.3 above. For any $\mathbf{p} \in W$ we define the functions $\psi_{\mathbf{p}}(\boldsymbol{\theta}) = -\log \overline{S}(p_1 e^{\mathrm{i}\theta_1}, \ldots, p_d e^{\mathrm{i}\theta_d}) + \log \overline{S}(\mathbf{p}) + \mathrm{i}(\theta_1 + \cdots + \theta_d)$ and

$$P_{\mathbf{p}}(\boldsymbol{\theta}) = (1+z_1)\cdots(1+z_{d-1})\left(1 - z_d^2 \frac{A(\hat{\mathbf{z}})}{B(\hat{\mathbf{z}})}\right)(1-z_d)^{-1}\bigg|_{z_k = p_k e^{\mathrm{i}\theta_k},\ 1 \leq k \leq d}.$$

A similar calculus computation to the positive drift case shows that the Hessian of the function $\psi_{\mathbf{p}}$ at the origin is the non-singular matrix

$$\mathcal{H}_{\mathbf{p}} = \begin{pmatrix} \frac{2p_1B_1(\mathbf{p})}{\overline{S}(\mathbf{p})} & 0 & 0 & \cdots & 0 \\ 0 & \frac{2p_2B_2(\mathbf{p})}{\overline{S}(\mathbf{p})} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \ddots & \frac{2p_{d-1}B_{d-1}(\mathbf{p})}{\overline{S}(\mathbf{p})} & 0 \\ 0 & 0 & \cdots & 0 & \frac{2B(\hat{\mathbf{p}})}{p_d\overline{S}(\mathbf{p})} \end{pmatrix}.$$

Finally, for any $N \in \mathbb{N}$ define the somewhat cumbersome quantity

$$\Phi_N^{(\mathbf{p})}(n) = S(\mathbf{p})^n\, n^{-d/2}\,(2\pi)^{-d/2}(\det \mathcal{H}_{\mathbf{p}})^{-1/2}\sum_{k=0}^{N-1} C_k^{(\mathbf{p})} n^{-k}, \tag{10.6}$$

where

$$C_k^{(\mathbf{p})} = \sum_{\ell=0}^{2k} \frac{\mathcal{E}^{k+l}(P_{\mathbf{p}}(\boldsymbol{\theta})\,\widetilde{\psi_{\mathbf{p}}}(\boldsymbol{\theta})^{\ell})}{(-1)^k 2^{k+\ell} l!(k+\ell)!}\Bigg|_{\boldsymbol{\theta}=\mathbf{0}},$$

the function

$$\widetilde{\psi_{\mathbf{p}}}(\boldsymbol{\theta}) = \psi_{\mathbf{p}}(\boldsymbol{\theta}) - (1/2)\boldsymbol{\theta}^T \cdot \mathcal{H} \cdot \boldsymbol{\theta},$$

and $\mathcal{E}$ is the differential operator

$$\mathcal{E} = -\frac{\overline{S}(\mathbf{p})}{2}\left(\sum_{1\le k<d} \frac{1}{p_k B_k(\mathbf{p}_{\hat{k}})}\partial_k^2 + \frac{p_d}{B(\hat{\mathbf{p}})}\partial_d^2\right).$$

Corollary 5.2 from Chapter 5 states that for any $N \in \mathbb{N}$ the number of walks in the orthant model defined by the characteristic polynomial $S(\mathbf{z})$ satisfies

$$c_n = \sum_{\mathbf{p}\in W} \Phi_N^{(\mathbf{p})}(n) + O(n^{-N}).$$

To establish Theorem 10.2 we need to determine the dominant non-zero terms in such an expansion. This follows from some straightforward, but rather tedious, calculus computations. For any fixed dimension these computations can be performed by a computer algebra system.

Proof (of Theorem 10.2) Since $P_{\mathbf{p}}(\mathbf{0}) = 0$ for all $\mathbf{p} \in W$, it follows that $C_0^{(\mathbf{p})} = 0$ for all $\mathbf{p} \in W$. We thus determine which of the constants $C_1^{(\mathbf{p})}$ in the expansions (10.6) are non-zero. For readability we suppress $\mathbf{p}$ in P and $\widetilde{\psi}$. Because P vanishes at the origin, and $\widetilde{\psi}$ vanishes to third order at the origin, the sum for $C_1^{(\mathbf{p})}$ simplifies to

$$C_1^{(\mathbf{p})} = -\frac{1}{2}\left(\mathcal{E}(P)(\mathbf{0}) + \frac{\mathcal{E}^2(P\widetilde{\psi})(\mathbf{0})}{4}\right).$$

First, we evaluate $\mathcal{E}(P)(\mathbf{0})$. Let

$$X = \prod_{1 \le j < d} (1 + p_j e^{i\theta_j}), \quad Y = 1 - p_d^2 e^{2i\theta_d} \frac{A\left(\hat{\mathbf{p}} e^{i\widehat{\boldsymbol{\theta}}}\right)}{B\left(\hat{\mathbf{p}} e^{i\widehat{\boldsymbol{\theta}}}\right)}, \text{ and } Z = \frac{1}{1 - p_d e^{i\theta_d}}$$

so that $P(\boldsymbol{\theta}) = XYZ$, and

$$\mathcal{E}(P)(\mathbf{0}) = \frac{-\overline{S}(\mathbf{p})}{2} \left(\sum_{1 \le k < d} \frac{1}{p_k B_k(\mathbf{p}_{\hat{k}})} (\partial_k^2 XYZ)(\mathbf{0}) + \frac{p_d}{B(\hat{\mathbf{p}})} (\partial_d^2 XYZ)(\mathbf{0}) \right).$$

For $1 \le k < d$ straightforward calculus computations show that Y and its first order partial derivative with respect to θ_k vanish at the origin, so that

$$\begin{aligned}
(\partial_k^2 XYZ)(\mathbf{0}) &= X(\mathbf{0})\, (\partial_k^2 Y)(\mathbf{0})\, Z(\mathbf{0}) \\
&= -p_d^2 \frac{\prod_{1 \le j < d}(1 + p_j)}{1 - p_d} \frac{\left(-2A_k'(\mathbf{p})B(\hat{\mathbf{p}}) + 2A(\hat{\mathbf{p}})B_k'(\mathbf{p})\right)}{B(\hat{\mathbf{p}})^2} \\
&= 2 \frac{B(\hat{\mathbf{p}})}{A(\hat{\mathbf{p}})} \frac{\prod_{1 \le j < d}(1 + p_j)}{1 - p_d} \left[\frac{A_k'(\mathbf{p})B(\hat{\mathbf{p}}) - A(\hat{\mathbf{p}})B_k'(\mathbf{p})}{B(\hat{\mathbf{p}})^2} \right] \\
&= \frac{2 \prod_{1 \le j < d}(1 + p_j)}{1 - p_d} \left[\frac{A_k'(\mathbf{p})}{A(\hat{\mathbf{p}})} - \frac{B_k'(\mathbf{p})}{B(\hat{\mathbf{p}})} \right].
\end{aligned}$$

Similarly, because X is independent of θ_d and $Y(\mathbf{0}) = 0$,

$$(\partial_d^2 XYZ)(\mathbf{0}) = X(\mathbf{0})\, (\partial_d^2 Y)(\mathbf{0})\, Z(\mathbf{0}) + 2X(\mathbf{0})\, (\partial_d Y)(\mathbf{0})\, (\partial_d Z)(\mathbf{0}) = \frac{4 \prod_{j<d}(1 + p_j)}{(1 - p_d)^2}.$$

Combining these two expressions then yields

$$\mathcal{E}(P)(\mathbf{0}) = \frac{-\overline{S}(\mathbf{p}) \prod_{1 \le j < d}(1 + p_j)}{2(1 - p_d)} \left[\sum_{j=1}^{d-1} \frac{2}{p_j B_j(\mathbf{p})} \left(\frac{A_j'(\mathbf{p})}{A(\hat{\mathbf{p}})} - \frac{B_j'(\mathbf{p})}{B(\hat{\mathbf{p}})} \right) + \frac{4 p_d}{B(\hat{\mathbf{p}})(1 - p_d)} \right].$$

To determine $C_1^{(\mathbf{p})}$ we must also calculate $\mathcal{E}^2(P\widetilde{\psi})(\mathbf{0})$. The operator $\mathcal{E}^2$ is the sum of operators of the form $\partial_j^2 \partial_k^2$ for $1 \le j \le k \le d$. Differentiation shows $(\partial_j^3 \widetilde{\psi})(\mathbf{0}) = 0$ so, since P vanishes at the origin, the non-zero terms of $\mathcal{E}^2(P\widetilde{\psi})(\mathbf{0})$ are those corresponding to the evaluations $(\partial_d P)(\mathbf{0})(\partial_d \partial_j^2 \widetilde{\psi})(\mathbf{0})$ for $1 \le j < d$. More differentiation gives

$$(\partial_d P)(\mathbf{0}) = X(\mathbf{0})(\partial_d Y)(\mathbf{0}) Z(\mathbf{0}) = -2i \frac{\prod_{j<d}(1 + p_j)}{1 - p_d}$$

and

$$(\partial_d \partial_j^2 \widetilde{\psi})(\mathbf{0}) = -\frac{(\partial_d \partial_j^2 \overline{S})(\mathbf{p})}{\overline{S}(\mathbf{p})} = \frac{2\mathrm{i} p_j \left(p_d A_j'(\mathbf{p}) - p_d^{-1} B_j'(\mathbf{p})\right)}{\overline{S}(\mathbf{p})}.$$

Expanding $\mathcal{E}^2$ then gives

$$\mathcal{E}^2(P\,\widetilde{\psi})(\mathbf{0}) = \frac{4\overline{S}(\mathbf{p}) \prod_{j<d}(1+p_j)}{p_d(1-p_d)A(\hat{\mathbf{p}})} \sum_{i=1}^{d-1} \frac{p_d A_j'(\mathbf{p}) - p_d^{-1} B_j'(\mathbf{p})}{B_j(\mathbf{p})}.$$

Finally, pulling everything together and simplifying using $p_d^2 = B(\hat{\mathbf{p}})/A(\hat{\mathbf{p}})$ gives

$$C_1^{(\mathbf{p})} = \frac{\overline{S}(\mathbf{p}) \prod_{j<d}(1+p_j)}{1-p_d} \left[\frac{1}{A(\hat{\mathbf{p}})p_d(1-p_d)} + \sum_{j=1}^{d-1} \frac{1-p_j}{2p_j B_j(\mathbf{p})} \left(\frac{A_j'(\mathbf{p})}{A(\hat{\mathbf{p}})} - \frac{B_j'(\mathbf{p})}{B(\hat{\mathbf{p}})} \right) \right].$$

We now see that the only values of $\mathbf{p}$ for which $C_1^{(\mathbf{p})}$ is non-zero are those where the first $d-1$ coordinates equal one. The final coordinate p_d can equal $\pm\sqrt{B(\mathbf{1})/A(\mathbf{1})}$, but Theorem 10.3 implies $S(\mathbf{1}, p_d) = S(\mathbf{1})$ at any contributing point. If $Q(\hat{\mathbf{z}})$ is identically zero then $\overline{S}(\hat{\mathbf{z}}, -z_d) = -\overline{S}(\hat{\mathbf{z}}, z_d)$ and both the positive and negative values of p_d appear in contributing points. If $Q(\hat{\mathbf{z}})$ is not identically zero then only $\mathbf{p}_2 = (\mathbf{1}, \sqrt{B(\mathbf{1})/A(\mathbf{1})})$ is a minimal contributing singularity. Summing the corresponding asymptotic contributions and simplifying the resulting expression then gives Theorem 10.2. □

Remark 10.3 As described in Chapter 4, several approaches to walks in orthants were originally motivated by conjectures of Bostan and Kauers [2] on asymptotics of two-dimensional walks restricted to a quadrant (see Table 4.1 from Chapter 4). Four of the twenty-three models arising in this context are highly-symmetric, while another twelve are mostly symmetric (six having positive drift and six having negative drift). Asymptotics of the highly symmetric models follow immediately from Theorem 6.1 in Chapter 6, while asymptotics of the mostly symmetric models are given by Theorems 10.1 and 10.2. Six of the remaining seven models have algebraic generating functions, so that asymptotics can be determined using univariate techniques [4, 3]. The final two-dimensional quadrant models with short steps are discussed in Problem 10.3.

The number of walks ending on one or more of the boundary axes of the first orthant is also of combinatorial interest. Problem 10.1 asks you to determine a diagonal expression for walks returning to these boundary hyperplanes. Problem 10.2 asks you to give asymptotics for the two-dimensional mostly symmetric models with short steps returning to either or both axes.

10.2 Lattice Path Problems to Test Your Skills

We end by listing two other areas of lattice path enumeration where diagonal expressions naturally occur. Asymptotic behaviour for a selection models that appear in these contexts is explored in Problems 10.4 to 10.7. There are countless opportunities to apply analytic combinatorics in several variables to lattice path enumeration problems and we hope the reader, after making it through this text, is inspired to take their newly developed skills and discover additional results in the area.

Central Weightings

Let $\mathcal{S} \subset \mathbb{Z}^d$ be a finite step set not contained in a half-space of $\mathbb{Z}^d$. A weighting which assigns a positive weight $w_{\mathbf{s}} > 0$ to each $\mathbf{s} \in \mathcal{S}$ is called *central* if the weight of any path in $\mathbb{N}^d$ using the steps of $\mathcal{S}$ depends only on the length, start, and end points of the path. Courtiel et al. [5] and Melczer [8, Ch. 11] investigate, classify, and deduce asymptotics for centrally weighted models. Central weightings are interesting from the point of view of analytic combinatorics in several variables because uniform rational diagonal expressions can often be deduced for all central weightings of a fixed step set, treating the weights as parameters. As the weights vary, the contributing singularities and geometry of the singular variety at these points change, resulting in interesting shifts between different asymptotic templates. Problems 10.4 and 10.5 ask you to determine asymptotic behaviour for an important family of central weightings.

Walks with Longer Steps

Bostan et al. [1] generalize aspects of the kernel method discussed in Chapter 4 to walks with steps whose coordinates have modulus larger than one. Although this extension requires a more careful consideration of algebraic functions than the kernel method for short step models, it still results in rational diagonal expressions for several lattice path models restricted to the quadrant. Problems 10.6 and 10.7 ask you to determine asymptotics for two such models.

Problems

10.1 Let $V \subset \{1, \dots, d\}$. If $\mathcal{S} \subset \{\pm 1, 0\}^d$ is a mostly symmetric step set that takes a step forwards and backwards in each coordinate, prove that the generating function enumerating walks on the steps in $\mathcal{S}$ which start at the origin and stay in $\mathbb{N}^d$ is the main power series diagonal of the rational function

$$E(\mathbf{z}, t) = \frac{G(\mathbf{z}, t)}{H(\mathbf{z}, t)} \prod_{j \in V} (1 - z_j),$$

where G and H are the functions in (10.4).

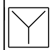
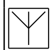
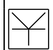
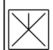
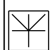
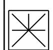
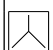
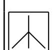
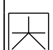
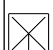
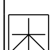
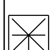
$\mathcal{S}$	Return to x-axis	Return to y-axis	Return to origin
	$\frac{3\sqrt{3}}{4\sqrt{\pi}}\frac{3^n}{n^{3/2}}$	$\delta_n\frac{4\sqrt{2}}{\pi}\frac{(2\sqrt{2})^n}{n^2}$	$\epsilon_n\frac{16\sqrt{2}}{\pi}\frac{(2\sqrt{2})^n}{n^3}$
	$\frac{8}{3\sqrt{\pi}}\frac{4^n}{n^{3/2}}$	$\delta_n\frac{4\sqrt{3}}{\pi}\frac{(2\sqrt{3})^n}{n^2}$	$\delta_n\frac{12\sqrt{3}}{\pi}\frac{(2\sqrt{3})^n}{n^3}$
	$\frac{5\sqrt{10}}{16\sqrt{\pi}}\frac{5^n}{n^{3/2}}$	$\frac{\sqrt{2}(1+\sqrt{2})^{3/2}}{\pi}\frac{(2+2\sqrt{2})^n}{n^2}$	$\frac{2(1+\sqrt{2})^{3/2}}{\pi}\frac{(2+2\sqrt{2})^n}{n^3}$
	$\frac{5\sqrt{10}}{24\sqrt{\pi}}\frac{5^n}{n^{3/2}}$	$\delta_n\frac{4\sqrt{30}}{5\pi}\frac{(2\sqrt{6})^n}{n^2}$	$\delta_n\frac{24\sqrt{30}}{25\pi}\frac{(2\sqrt{6})^n}{n^3}$
	$\frac{\sqrt{3}}{\sqrt{\pi}}\frac{6^n}{n^{3/2}}$	$\frac{2\sqrt{3}(1+\sqrt{3})^{3/2}}{3\pi}\frac{(2+2\sqrt{3})^n}{n^2}$	$\frac{2(1+\sqrt{3})^{3/2}}{\pi}\frac{(2+2\sqrt{3})^n}{n^3}$
	$\frac{7\sqrt{21}}{54\sqrt{\pi}}\frac{7^n}{n^{3/2}}$	$\frac{A}{285\pi}\frac{(2+2\sqrt{6})^n}{n^2}$	$\frac{2B}{1805\pi}\frac{(2+2\sqrt{6})^n}{n^3}$
	$\gamma_n\frac{1}{9\pi}\frac{(2\sqrt{2})^n}{n^3}$	$\delta_n\frac{4\sqrt{2}}{\pi}\frac{(2\sqrt{2})^n}{n^2}$	$\epsilon_n\frac{16\sqrt{2}}{\pi}\frac{(2\sqrt{2})^n}{n^3}$
	$\left(\delta_n\frac{36\sqrt{3}}{\pi}+\delta_{n-1}\frac{54}{\pi}\right)\frac{(2\sqrt{3})^n}{n^3}$	$\delta_n\frac{4\sqrt{3}}{\pi}\frac{(2\sqrt{3})^n}{n^2}$	$\delta_n\frac{12\sqrt{3}}{\pi}\frac{(2\sqrt{3})^n}{n^3}$
	$\frac{4(1+\sqrt{2})^{7/2}}{\pi}\frac{(2+2\sqrt{2})^n}{n^3}$	$\frac{\sqrt{2}(1+\sqrt{2})^{3/2}}{\pi}\frac{(2+2\sqrt{2})^n}{n^2}$	$\frac{2(1+\sqrt{2})^{3/2}}{\pi}\frac{(2+2\sqrt{2})^n}{n^3}$
	$\left(\delta_n\frac{72\sqrt{30}}{5\pi}+\delta_{n-1}\frac{864\sqrt{5}}{25\pi}\right)\frac{(2\sqrt{6})^n}{n^3}$	$\delta_n\frac{4\sqrt{30}}{5\pi}\frac{(2\sqrt{6})^n}{n^2}$	$\delta_n\frac{24\sqrt{30}}{25\pi}\frac{(2\sqrt{6})^n}{n^3}$
	$\frac{3(1+\sqrt{3})^{7/2}}{2\pi}\frac{(2+2\sqrt{3})^n}{n^3}$	$\frac{2\sqrt{3}(1+\sqrt{3})^{3/2}}{3\pi}\frac{(2+2\sqrt{3})^n}{n^2}$	$\frac{2(1+\sqrt{3})^{3/2}}{\pi}\frac{(2+2\sqrt{3})^n}{n^3}$
	$\frac{6C}{1805\pi}\frac{(2+2\sqrt{6})^n}{n^3}$	$\frac{A}{285\pi}\frac{(2+2\sqrt{6})^n}{n^2}$	$\frac{2B}{1805\pi}\frac{(2+2\sqrt{6})^n}{n^3}$

Table 10.2 Asymptotics of quadrant walks with mostly symmetric step sets which end on the x-axis, the y-axis, and the origin, respectively.

10.2 Use the diagonal expression in Problem 10.1 to prove the asymptotics for two-dimensional models listed in Table 10.2, where the terms

$$\delta_n = \begin{cases} 1 & : n \equiv 0 \mod 2 \\ 0 & : \text{otherwise} \end{cases} \qquad \sigma_n = \begin{cases} 1 & : n \equiv 0 \mod 3 \\ 0 & : \text{otherwise} \end{cases} \qquad \epsilon_n = \begin{cases} 1 & : n \equiv 0 \mod 4 \\ 0 & : \text{otherwise} \end{cases}$$

and

$$\gamma_n = \begin{cases} 448\sqrt{2} & : n \equiv 0 \mod 4 \\ 640 & : n \equiv 1 \mod 4 \\ 416\sqrt{2} & : n \equiv 2 \mod 4 \\ 512 & : n \equiv 3 \mod 4 \end{cases}$$

account for periodicities in the lattice path models, and A, B, and C are given by

$$A = (156+41\sqrt{6})\sqrt{23-3\sqrt{6}}, \quad B = (583+138\sqrt{6})\sqrt{23-3\sqrt{6}}, \quad C = (4571+1856\sqrt{6})\sqrt{23-3\sqrt{6}}.$$

Hint: Most of the analysis from Theorem 10.3 can be reused. Under what conditions are the contributing singularities for the diagonal expression encoding walks returning to one or both axes the same as the contributing singularities for the diagonal expression encoding walks ending anywhere in $\mathbb{N}^2$?

10.3 Quadrant models on the steps $\mathcal{S} = \{(1,0),(-1,0),(-1,1),(1,-1)\}$ are named after Gouyou-Beauchamps, who found [6] a hypergeometric formula for the number of walks returning to an axis and proved [7] that walks in the model with various restrictions are in bijection with several other combinatorial objects (for instance, pairs of non-intersecting Dyck prefixes and Young tableaux of height at most four). Let c_n enumerate the number of lattice walks of length n taking steps in $\mathcal{S}$, starting at the origin, and staying in the first quadrant. Theorem 4.1 from Chapter 4 implies that the generating function of c_{n+2} is the main power series diagonal of

$$F(x,y,t) = \frac{yt^2(x+1)(y-x^2)(x-y)(x+y)}{1-xyt(x+x\overline{y}+y\overline{x}+\overline{x})}.$$

Find dominant asymptotics of c_n.

10.4 The characteristic polynomial for a central weighting of the Gouyou-Beauchamps step set $\mathcal{S} = \{(1,0),(-1,0),(-1,1),(1,-1)\}$ can be parametrized (up to a constant multiple) as $S_{a,b}(x,y) = \frac{1}{ax} + ax + \frac{ax}{by} + \frac{by}{ax}$ with $a,b>0$. Using the kernel method of Chapter 10, or other means, prove that the generating function enumerating the number of weighted walks on $\mathcal{S}$ which start at the origin and stay in the quadrant $\mathbb{N}^2$ satisfies

$$\sum_{n\geq 0} c_{n+2}t^n = \Delta\left(\frac{yt^2(y-b)(a-x)(a+x)(a^2y-bx^2)(ay-bx)(ay+bx)}{a^4b^3(1-txyS_{a,b}(\overline{x},\overline{y}))(1-x)(1-y)}\right).$$

10.5 Use the techniques of analytic combinatorics in several variables to determine asymptotics for the rational diagonal of Problem 10.4 as a function of the positive parameters a and b. As a and b vary, characterize the values where asymptotics undergo a sharp transition.

10.6 Let $\mathcal{S} = \{(1,0),(-1,0),(0,-1),(-2,1)\}$. The generating function for the number c_n of walks on $\mathcal{S}$ which use the steps in $\mathcal{S}$, start at the origin, and stay in the quadrant $\mathbb{N}^2$ satisfies

$$\sum_{n\geq 0} c_{n+2}t^n = \Delta\left(\frac{yt^2(x^2+1)(x^2+2xy-1)(2x^3+x^2y-y)(x^2-y^2)}{(1-x)(1-y)(1-t(x^3+x^2y+xy^2+y))}\right).$$

Use the techniques of analytic combinatorics in several variables to prove

$$c_n = \frac{(2\sqrt{3})^n}{\pi n^4}\left(C_n + O\left(\frac{1}{n}\right)\right),$$

where

$$C_n = \begin{cases} 5616\sqrt{3} & : n \text{ even} \\ 9720 & : n \text{ odd} \end{cases}$$

10.7 Let $\mathcal{S} = \{(0,1),(1,-1),(-1,-1),(-2,1)\}$. The generating function for the number c_n of walks on $\mathcal{S}$ which use the steps in $\mathcal{S}$, start at the origin, and stay in the quadrant $\mathbb{N}^2$ satisfies

$$\sum_{n\geq 0} c_{n+2}t^n = \Delta\left(\frac{yt^2(2xy^2+x^2-1)(x-y^2)(x^2y^2+2x^3-y^2)}{(1-x)(1-y)(1-t(x^2y^2+x^3+y^2+x))}\right).$$

Use the techniques of analytic combinatorics in several variables to prove

$$c_n = \frac{\left(8\ 3^{-3/4}\right)^n}{\pi n^4}\left(C_n + O\left(\frac{1}{n}\right)\right),$$

where

$$C_n = \begin{cases} 5120\sqrt{3} & : n \equiv 0 \mod 4 \\ 6656\ 3^{1/4} & : n \equiv 1 \mod 4 \\ 26624/3 & : n \equiv 2 \mod 4 \\ 3840\ 3^{3/4} & : n \equiv 3 \mod 4 \end{cases}$$

References

1. Alin Bostan, Mireille Bousquet-Mélou, and Stephen Melczer. Counting walks with large steps in an orthant. *Accepted to Journal of the European Mathematical Society*, 2020.
2. Alin Bostan and Manuel Kauers. Automatic classification of restricted lattice walks. In *21st International Conference on Formal Power Series and Algebraic Combinatorics (FPSAC 2009)*, Discrete Math. Theor. Comput. Sci. Proc., AK, pages 201–215. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2009.
3. Alin Bostan and Manuel Kauers. The complete generating function for Gessel walks is algebraic. *Proc. Amer. Math. Soc.*, 138(9):3063–3078, 2010. With an appendix by Mark van Hoeij.
4. Mireille Bousquet-Mélou and Marni Mishna. Walks with small steps in the quarter plane. In *Algorithmic probability and combinatorics*, volume 520 of *Contemp. Math.*, pages 1–39. Amer. Math. Soc., Providence, RI, 2010.
5. J. Courtiel, S. Melczer, M. Mishna, and K. Raschel. Weighted lattice walks and universality classes. *Journal of Combinatorial Theory. Series A*, 152:255–302, 2017.
6. D. Gouyou-Beauchamps. Chemins sous-diagonaux et tableaux de Young. In *Combinatoire énumérative (Montréal, 1985, Québec)*, volume 1234 of *Lecture Notes in Math.*, pages 112–125. Springer, Berlin, 1986.
7. D. Gouyou-Beauchamps. Standard Young tableaux of height 4 and 5. *European J. Combin.*, 10(1):69–82, 1989.
8. Stephen Melczer. *Analytic Combinatorics in Several Variables: Effective Asymptotics and Lattice Path Enumeration*. PhD thesis, University of Waterloo and ENS Lyon, 2017. Thesis (PhD).
9. Stephen Melczer and Mark C. Wilson. Higher Dimensional Lattice Walks: Connecting Combinatorial and Analytic Behavior. *SIAM J. Discrete Math.*, 33(4):2140–2174, 2019.

Index

M

N

O

P

R

Author Index

H

I

J

K