12

Combinatorial applications

The utility of any mathematical theory is ultimately determined by the breadth of problems it can solve. In this chapter we illustrate the techniques of ACSV developed in previous chapters on a large selection of combinatorial examples. These examples are arranged taxonomically in Section 12.1, helping readers identify a template for their work when trying to apply ACSV to new problems. Sections 12.2, 12.3, and 12.4 give detailed applications of our basic theory to the study of Riordan arrays, Lagrange inversion, and the transfer matrix method, respectively. Section 12.5 discusses the use of higher order asymptotics, and Section 12.6 studies algebraic generating functions by encoding them as subseries of higher-dimensional rational generating functions. Section 12.7 presents miscellaneous examples chosen to illustrate particular aspects of the theory. Combinatorics and discrete probability are closely related, and Section 12.8 applies the results of Chapter 9 to prove probabilistic limit laws for asymptotics governed by smooth points, leading to a local central limit theorem in Theorem 12.36.

12.1 Some classifications

We begin with a guide to help users of ACSV find examples similar to their intended application, with some of the examples occurring in earlier chapters and some in this chapter. We classify the examples by local geometry, by form of generating function, and by intended application. We also point to some examples where our standard hypotheses fail to hold. The website for this book contains links to Sage worksheets computing many of the examples listed here.

https://doi.org/10.1017/9781108874144.017 Published online by Cambridge University Press

Local geometry of GF	Examples
Smooth nondegenerate; explicit critical points	9.10, 9.11, 9.13, 9.14, 12.11, 12.13, 12.15, 12.27, 12.30
Smooth nondegenerate; implicit critical points	12.5, 12.6, 12.9, 12.10, 12.17, 12.18, 12.20, 12.22, 12.23, 12.24, 12.25, 12.26
Smooth nondegenerate; periodicity	9.14, 12.11, 12.13
Smooth nondegenerate; torality	12.11, 12.13, 11.11
Smooth degenerate	9.32, 9.39
Multiple transverse $n < d$	12.29
Multiple transverse $n = d$	10.28, 10.27, 10.66
Multiple arrangement $n > d$	10.35, 10.47, 12.28
Multiple not arrangement	10.69, 13.3
Cone point	11.41, 11.43, 11.45, 11.47, 11.50

Table 12.1 *Guide by local geometry (dimension d, number of local sheets n).*

Classification by geometry of contributing points

Table 12.1 collates examples arranged by local geometry. Because smoothness is a generic property, smooth singular critical points dictate asymptotics in many applications. Although the coordinates of critical points can be solved in radicals for simple examples, such as Examples 9.10 and 9.11, this is usually not the case. Thus, many examples use the algebraic techniques discussed in Chapter 8 to work with critical points implicitly. A generating function with more than one contributing point for a given direction leads to periodicity in coefficients – for instance, the rational function

$$F(x, y) = \frac{1}{1 - x^2 - y^2}$$

has four contributing singularities in the main diagonal direction, reflecting the fact that the only terms that appear in any series expansion of F are those with even exponents. We may even have a continuum of critical points, such as when $F(x, y) = 1/(1 - xy)$, which can be handled under the strong torality hypothesis discussed in Section 9.1 of Chapter 9. Although the vast majority of our results require nondegenerate critical points, an example with cubic degeneracy was studied in Example 9.39 of Chapter 9. Degeneracies of any order can be handled in two dimensions using Theorem 9.38.

The difficulty of analyzing a multiple arrangement point $\boldsymbol{w}$ depends both on the dimension of the problem and on the number of smooth sheets intersecting at $\boldsymbol{w}$. The simplest cases occur when there is a single sheet (which is the smooth case) or when the number of sheets equals the dimension (where there is a complete intersection). Arrangement points with more sheets than factors are handled through an algebraic decomposition.

Form of GF	Examples
Denominator linear in a variable	9.10, 9.11, 9.13, 9.14, 9.39, 12.5, 12.6, 12.9, 12.10, 12.11, 12.13, 12.15, 12.17
Bivariate	9.10, 9.11, 9.13, 9.14, 9.32, 9.39, 10.28, 10.27, 10.35, 10.47, 10.66, 10.69, 11.11, 12.5, 12.6, 12.9, 12.15, 12.18, 12.22, 12.23, 12.25
Trivariate	11.41, 11.43, 11.45, 11.47, 11.50, 12.20, 12.26, 12.29
Higher/arbitrary dimension	12.10, 12.17, 12.27
Repeated factors, $m < d$	12.29
Repeated factors, $m = d$	10.28, 10.27, 10.66
Repeated factors, $m > d$	10.32, 10.35, 10.47, 12.28
Meromorphic, not rational	12.24, 12.30
Algebraic	12.18, 12.20
Non-combinatorial	9.32

Table 12.2 *Guide by form of GF (dimension d, number of denominator factors m).*

Multiple points that are not arrangement points are tricky, and the general theory has not been worked out. Two sheets that are tangent behave basically like a single repeated sheet, as seen in Proposition 10.68 of Chapter 10, although more complicated singularities can arise, as in Example 10.9. Chapter 11 contains essentially all that we know for more complicated singularities, with explicit results for cone points.

Classification by form of generating function

Table 12.2 classifies our examples by the algebraic form of the generating function $F(z) = P(z)/Q(z)$. The simplest case occurs when Q is linear in one of its variables, and the (perhaps surprising) ubiquity of examples of this form is a reflection of the fact that the sequence construction on combinatorial classes (described in Section 2.2 of Chapter 2) corresponds to the quasi-inverse map $f \mapsto 1/(1 - f)$ on generating functions. The technique of Lagrange inversion can also be incorporated into this framework. Sections 12.2 and 12.3 cover applications to Riordan arrays and Lagrange inversion, respectively. Section 12.4 discusses the transfer matrix method.

Our formulae are flexible enough to work in any dimension, and even for families with arbitrary dimension as a parameter, although computations are often simpler in lower dimensions. Repeated denominator factors correspond to higher order poles, and thus change asymptotic behavior, while multiple distinct factors lead to multiple points.

Structure/application	Typical type of GF	Examples
regular languages, words, strings	Rational, smooth	12.6, 12.10, 12.17
lattice walks	Rational, smooth/multiple	12.26
trees	Algebraic or Riordan	12.9
quantum walks	Rational, toral	12.11, 12.13
tilings	Rational, cone/nasty	11.43, 11.45, 11.47
sums of independent random variables	Riordan	12.34
number triangles	Riordan	9.10, 9.11, 9.14, 12.15, 12.20, 12.24
constant coefficient linear recurrences	Rational, smooth/multiple	12.22
partitions	Infinite product	12.30

Table 12.3 *Guide by application area.*

Most of our generating functions are rational, but our asymptotic results hold more generally for meromorphic functions. Some examples, such as $F(x,y) = 1/(1 - e^x - e^y)$, require solving transcendental equations for critical points, while others, such as

$$F(x,y) = \prod_{i=1}^{\infty} \frac{1}{1 - x^i - y^i}$$

can be reduced to cases with polynomial denominators. We can also find asymptotics of algebraic generating functions by embedding them in rational series of higher dimension.

Classification by application

Many combinatorial families gives rise to multivariate generating functions with the same type of behavior, which can be analyzed together, and Table 12.3 gives a rough guide for readers seeking to quickly find a relevant application. Note that some problems fall into multiple areas, due to bijections between various combinatorial objects.

Examples where our standard hypotheses fail to hold

Most asymptotic expansions derived in this text hold for nondegenerate contributing points, require intersections of multiple denominator factors to be transverse, and occur in directions in the interior of cones where asymptotic behavior transitions smoothly with direction. Furthermore, the first-order terms

Exception	Examples
Degenerate contributing points	9.39
Non-transverse intersection	10.69
Boundary directions	10.66
Vanishing numerator	12.18, 12.20

Table 12.4 *Guide to non-generic examples.*

in our asymptotic expansions typically do not vanish. Table 12.4 collects examples where the above assumptions fail to hold.

Non-transverse intersections may be arbitrarily complicated, and we cover only a simple example. Similarly, we discuss one example with a degenerate contributing point. From the point of view of Fourier–Laplace integrals, asymptotics in directions on boundaries of the cones dictating uniform behavior are half of what they would be if the direction was interior to the cone. The first-order term in our asymptotic expansions of a sequence with generating function F vanishes when the numerator of F is zero at its contributing singularities. In this case, we can usually determine dominant asymptotic behavior by computing higher order coefficients in the expansion.

12.2 Powers, quasi-powers, and Riordan arrays

Let $v(\boldsymbol{z})$ be a power series (or polynomial) and suppose that we want to estimate the coefficient $[\boldsymbol{z}^{\boldsymbol{r}}]v(\boldsymbol{z})^k$ of a large power of v. This coefficient equals the coefficient of $\boldsymbol{z}^{\boldsymbol{r}} w^k$ in the power series expansion

$$F(\boldsymbol{z}, w) = \frac{1}{1 - wv(\boldsymbol{z})}, \tag{12.1}$$

so we can determine asymptotics using the tools of ACSV. The combinatorial constructions discussed in Section 2.2 of Chapter 2 show some ways in which generating functions of this form arise. Another common application comes from probability: if $v(\boldsymbol{z}) = \sum_{\boldsymbol{r}} a_{\boldsymbol{r}} \boldsymbol{z}^{\boldsymbol{r}}$, where $a_{\boldsymbol{r}} = \Pr(X_j = \boldsymbol{r})$ for a family $\{X_j\}$ of independent, identically distributed random variables taking values in $\mathbb{N}^d$ then $v(\boldsymbol{z})^n$ is the probability generating function for the partial sum $S_n = \sum_{j=1}^n X_j$, and hence

$$\mathbb{P}(S_n = \boldsymbol{r}) = [\boldsymbol{z}^{\boldsymbol{r}}]v(\boldsymbol{z})^n .$$

It has long been known that, under suitable hypotheses, such *large powers* lead to Gaussian behavior. An early work on multivariate analytic combinatorics [BR83] observed that this behavior is robust enough to hold not only

https://doi.org/10.1017/9781108874144.017 Published online by Cambridge University Press

for exact powers, but also for ***quasi-powers***, meaning sequences of functions $\{f_n(\boldsymbol{z})\}$ satisfying

$$f_n(\boldsymbol{z}) \sim C_n g(\boldsymbol{z}) \cdot h(\boldsymbol{z})^n \tag{12.2}$$

uniformly as $\boldsymbol{z}$ ranges over certain polydisks. Gaussian behavior of coefficients of quasi-powers is the basis for the *GF-sequence method* developed by Bender, Richmond, and collaborators in a series of papers including [Ben73; BR83; GR92; BR99]; see also the work of Hwang extending this to algebraic-logarithmic singularities [Hwa96; Hwa98a; Hwa98b]. These papers give conditions under which a multivariate generating function

$$F(z_1, \ldots, z_d, w) = \sum_{n=0}^{\infty} f_n(\boldsymbol{z}) w^n \tag{12.3}$$

is a quasi-power in the sense of (12.2). They then show that if g and h are analytic in a Δ-domain (recall Figure 3.1), if h has a unique dominant singularity where the boundary of the region intersects the positive real axis, and if the quadratic part of h is nondegenerate there, then, after a rescaling, the coefficients of $f_n(\boldsymbol{z})$ have a Gaussian limit distribution as $n \to \infty$.

Riordan arrays

An important combinatorial family of quasi-powers is the set of ***Riordan arrays*** $\{a_{nk} : n, k \geq 0\}$ whose generating functions $F(x, y) = \sum_{n,k \geq 0} a_{nk} x^n y^k$ satisfy

$$F(x, y) = \frac{\phi(x)}{1 - yv(x)} \tag{12.4}$$

for some analytic functions ϕ and v with $\phi(0) \neq 0$. Just as (12.1) represents sums of independent, identically distributed random variables when v is a probability generating function, the function (12.4) is a *delayed renewal sum* [Dur04, Section 3.4], where an initial random variable X_0 may be added that is distributed differently from the others. Riordan arrays cover an enormous number of examples arising in applications, including many lattice path problems, and are useful for simplifying sums because of the Pascal-like recurrences the terms satisfy – see [Mer+97; Spr94].

Remark 12.1. Some authors require that $v(0) = 0$ in (12.4), but we do not make this assumption.

https://doi.org/10.1017/9781108874144.017 Published online by Cambridge University Press

Define the functions

$$\mu(v; x) := \frac{xv'(x)}{v(x)} \tag{12.5}$$

$$\sigma^2(v; x) := \frac{x^2 v''(x)}{v(x)} + \mu(v; x) - \mu(v; x)^2 = x\frac{d\mu(v; x)}{dx}, \tag{12.6}$$

whose symbols are motivated by probability theory (see Section 12.8 below).

Theorem 12.2. *Let* $(v(x), \phi(x))$ *determine a Riordan array with generating function* (12.4). *Suppose that* $v(x)$ *has radius of convergence* $R \in (0, \infty]$ *and is aperiodic with nonnegative coefficients. If* ϕ *has radius of convergence at least* R *then*

(i) the function $\mu(v; x)$ *is strictly increasing for* $x \in (0, R)$*, and its range contains the interval* $J = (A, B)$*, where* $A = \mu(v; 0)$ *and* $B = \mu(v; R)$ *are defined as one-sided limits;*

(ii) there is an asymptotic expansion of the form

$$a_{rs} \approx x^{-r} v(x)^s s^{-1/2} \sum_{k=0}^{\infty} b_k(r/s) s^{-k} \tag{12.7}$$

uniformly as r/s *varies over compact subsets of* J*, where* x *is the unique positive real solution to* $\mu(v; x) = r/s$*. The leading term in this expansion is*

$$a_{rs} \sim x^{-r} v(x)^s \frac{\phi(x)}{\sqrt{2\pi s \sigma^2(v; x)}}. \tag{12.8}$$

Proof Writing $P(x, y) = \phi(x)$ and $Q(x, y) = 1 - yv(x)$, we see that $\mathcal{V}$ is smooth because Q and Q_y never simultaneously vanish. Furthermore, the smooth critical point equations simplify to show that $(x, 1/v(x))$ is critical in the direction (r, s) if and only if $\mu(v; x) = r/s$. Lemma 6.41 shows that all points of the form $(x, 1/v(x))$ for $x \in (0, R)$ are strictly minimal points. The Hessian determinant appearing in the asymptotic expansion (9.2) simplifies to $\sigma^2(v; x)$, so the result follows from an application of Theorem 9.4 after we show that μ is strictly increasing for $x \in (0, R)$. The latter result follows from (12.6). □

Remark 12.3. If v has coefficients of mixed sign, more complicated behavior can occur – see Exercise 12.8.

Remark 12.4. Riordan arrays are often specified by a recursion of the form

$$a_{n+1,k+1} = \sum_{j=0}^{\infty} c_j a_{n,k+j},$$

https://doi.org/10.1017/9781108874144.017 Published online by Cambridge University Press

where the generating function $C(t) = \sum_{j=1}^{\infty} c_j t^j$ is known explicitly but $v(x)$ is known only implicitly through the equation $v(x) = xC(v(x))$. Subtleties that arise in computations when dealing with such implicitly defined v are discussed in [Wil05].

Example 12.5 (Packing paths in paths)**.** Došlić [Doš19] derives the bivariate generating function

$$F(x,y) = \sum_{n,k} a_{nk} x^n y^k = \frac{1 - x^m}{1 - x - x^m(1 - x^m)y}$$

for the number a_{nk} of ways to maximally pack a path of length m in a path of length n, using exactly k copies of the smaller path. This is a Riordan array with $\phi(x) = (1 - x^m)/(1 - x)$ and $v(x) = x^m\phi(x)$, and ϕ and v have infinite radii of convergence since they are both polynomials.

Taking derivatives shows that

$$\mu(v;x) = m + \frac{x + 2x^2 + \cdots + (m-1)x^{m-1}}{1 + x + x^2 + \cdots + x^{m-1}},$$

and μ increases from m to $2m - 1$ as x increases from zero to infinity, covering all directions of combinatorial interest [Doš19, Proposition 2.1]. Thus, when $\lambda = n/k$ remains in a compact sub-interval of $(m, 2m - 1)$ the Gaussian asymptotic expansion (12.7) holds.

To determine the leading term in this expansion with respect to λ, one can compute a lexicographic Gröbner basis of the ideal $\langle \mu(v;x) - \lambda, \sigma^2(v;x) - S, \phi(x) - T\rangle$ in $\mathbb{Q}(\lambda)[x, S, T]$ to eliminate x and write σ^2 and ϕ in terms of λ. For instance, when $m = 3$ we obtain the irreducible elimination polynomials

$$p(S;\lambda) = S^2 + \left(2\lambda^2 - 16\lambda + \frac{88}{3}\right)S + \lambda^4 - 16\lambda^3 + \frac{281}{3}\lambda^2 - \frac{712}{3}\lambda + 220$$
$$q(T;\lambda) = (\lambda - 5)^2T^2 + (3\lambda - 16)T + 3.$$

In this case we can use the quadratic formula to express σ^2 and ϕ explicitly in terms of λ, but for general m one must identify the correct branch of the elimination polynomials implicitly. Note that the system $p(S,\lambda) = p_S(S,\lambda) = p_\lambda(S,\lambda) = 0$ has no solutions, so the two branches given by solving $p(S,\lambda)$ for S do not meet. Because σ^2 is a continuous function of λ on the interval $[m, 2m - 1]$, and $\lim_{\lambda\to m^+} \sigma^2(v;x) = \lim_{x\to 0^+} \sigma^2(v;x) = 0$, the value of σ^2 as a function of λ is the branch of $p(S,\lambda) = 0$ passing through the point $(S,\lambda) = (0, m)$. The correct branch of $q(S,\lambda) = 0$ for ϕ with respect to λ is determined analogously. ◁

The condition in Theorem 12.2 that ϕ has at least as large a radius of convergence as v is satisfied in many, but not all, applications.

https://doi.org/10.1017/9781108874144.017 Published online by Cambridge University Press

Example 12.6 (Maximum number of distinct subsequences)**.** Flaxman, Harrow, and Sorkin [FHS04] study strings of length n over the alphabet $\{1, 2, \ldots, d\}$ that contain as many distinct (not necessarily contiguous) subsequences of length k as possible. Let a_{nk} denote the maximum number of distinct subsequences of length k that can be found in a single string of length n. Initial segments $S|_n$ of the infinite string S consisting of repeated blocks of the string $12 \cdots d$ turn out always to be maximizers, meaning $S|_n$ has exactly a_{nk} distinct subsequences of length k. The generating function for $\{a_{nk}\}$ is then computed to be

$$F(x, y) = \sum_{n,k} a_{nk} x^n y^k = \frac{1}{1 - x - xy(1 - x^d)},$$

meaning a_{nk} is a Riordan array with $\phi(x) = (1-x)^{-1}$ and $v(x) = x + x^2 + \cdots + x^d$.

Assume for non-triviality that $d \geq 2$. The singular variety $\mathcal{V}$ is the union of the line $x = 1$ and the smooth curve $y = 1/v(x)$, which meet transversely at the double point $(1, 1/d)$; see Figure 12.1 for an illustration with $d = 3$.

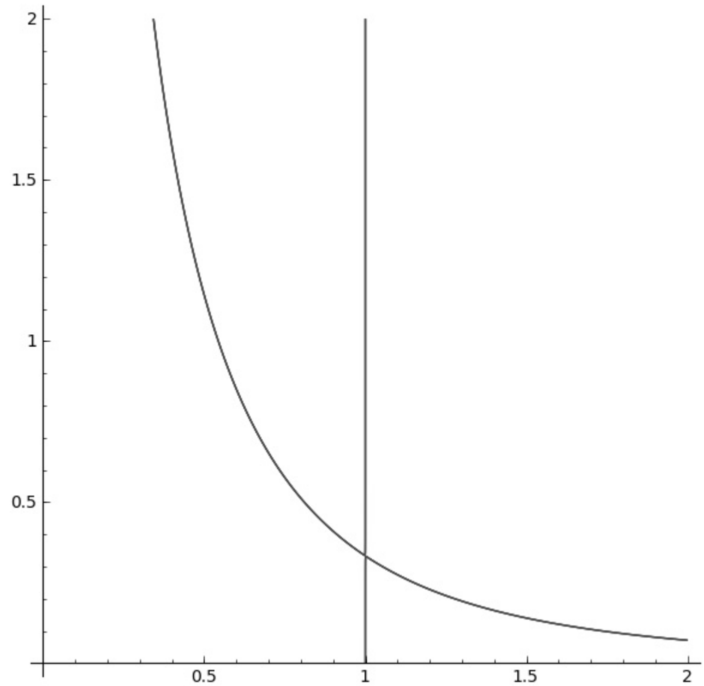

Figure 12.1 $\mathcal{V}$ in the case $d = 3$.

The radius of convergence of ϕ, namely 1, is now less than the radius of convergence of v, which is infinity. Taking derivatives shows that

$$\mu(v; x) = \frac{1}{1-x} - \frac{dx^d}{1-x^d} = \frac{1 + 2x + 3x^2 + \cdots + dx^{d-1}}{1 + x + x^2 + \cdots + x^{d-1}},$$

so that μ increases from 1 to $(d+1)/2$ as x increases from 0 to 1. The Gaussian asymptotics of (12.7) still hold, but only when $\lambda = n/k$ remains in a compact sub-interval of $(1, \frac{d+1}{2})$, and the computations to compute terms in this asymptotic expansion are analogous to those for Example 12.5 above.

Proposition 6.41 implies that there is a strictly minimal point in the positive quadrant controlling asymptotics in any direction. When $\lambda \geq (d+1)/2$ this turns out to be the non-smooth point $(1, 1/d)$, and Corollary 10.14 from Chapter 10 implies that $a_{\lambda k,k} \sim d^k$ in this case. Note that this is trivial when $\lambda \geq d$ because any prefix of the infinite string S with length at least dk will allow all possible k-subsequences to occur, meaning $a_{nk} = d^k$ in this regime. ◃

12.3 Lagrange inversion

On many families of recursively defined combinatorial classes, such as tree enumeration problems, the combinatorial constructions discussed in Chapter 2 yield generating functions satisfying functional equations of the form $f(z) = zv(f(z))$ for some function v that is analytic at the origin and does not vanish there. Although this equation can be solved exactly for f in some simple cases, the method of *Lagrange inversion* gives an invaluable tool for computing the coefficients of f, and their asymptotic behavior, directly from this functional equation.

Proposition 12.7 (*Lagrange inversion formula*)**.** *If $f(z) = zv(f(z))$ with v analytic and $v(0) \neq 0$ then*

$$[z^n]f(z) = \frac{1}{n}\left[y^{n-1}\right] v(y)^n. \tag{12.9}$$

Proof Change variables to $y = f(z)$ so that the implicit equation for f implies $z = y/v(y)$ and $dz = dy[1/v(y) - yv'(y)/v(y)^2]$. The Cauchy integral representation

$$[z^n]f(z) = \frac{1}{2\pi i}\int z^{-n-1} f(z)\, dz$$

over a circle sufficiently close to the origin becomes the integral

$$\frac{1}{2\pi i}\int \left[\left(\frac{v(y)}{y}\right)^n - \left(\frac{v(y)}{y}\right)^{n-1} v'(y)\right] dy$$

https://doi.org/10.1017/9781108874144.017 Published online by Cambridge University Press

around the origin in the y-plane. The difference between this integral and

$$\frac{1}{n}\left[y^{n-1}\right]v(y)^n = \frac{1}{2\pi i}\int \frac{1}{n}\left(\frac{v(y)}{y}\right)^n dy$$

equals

$$\frac{1}{2\pi i}\int\left[\left(\frac{v(y)}{y}\right)^{n-1} - \frac{n-1}{n}\left(\frac{v(y)}{y}\right)^n\right]dy = \frac{1}{2\pi i}\int d\left[\frac{y}{n}\left(\frac{v(y)}{y}\right)^n\right]dy = 0\,,$$

as claimed. □

To estimate the right-hand side of (12.9) via multivariate asymptotic analysis, we consider the generating function

$$\frac{1}{1-xv(y)} = \sum_{n=0}^{\infty} x^n v(y)^n$$

which generates the powers of v, so that

$$[z^n]f(z) = \frac{1}{n}\,[x^n y^n]\,\frac{y}{1-xv(y)}\,. \tag{12.10}$$

This formula holds at the level of formal power series and, if v has a non-zero radius of convergence, at the level of analytic functions.

Asymptotics for $[z^n]f(z)$ can be derived in terms of the power series coefficients of v using (12.10). For example, it follows from Theorem VI.6 of [FS09] that

$$[z^n]f(z) \sim \frac{1}{\sqrt{2\pi v''(y_0)/v(y_0)}}\,n^{-3/2}v'(y_0)^n, \tag{12.11}$$

where y_0 is the least $y > 0$ such that the tangent line to v at $(y, v(y))$ passes through the origin. In our notation y_0 is the smallest positive solution to $\mu(v; y) = 1$.

For a fixed power k, the generalization

$$[z^n]f(z)^k \sim \frac{k}{n}\frac{y_0^{k-1}}{\sqrt{2\pi n v''(y_0)/v(y_0)}}v'(y_0)^n$$

can also be obtained with univariate methods. Using multivariate methods, however, we may go further and derive bivariate asymptotics for $[z^n]f(z)^k$ as $k, n \to \infty$, holding uniformly as $\lambda = k/n$ varies over compact subsets of $(0, 1)$.

Proposition 12.8. *Let v be analytic and nonvanishing at the origin, where its power series expansion is aperiodic with nonnegative coefficients, and of order at least* 2 *at infinity. Let f be the nonnegative series satisfying $f(z) = zv(f(z))$ and define μ and σ^2 by equations (12.5) and* (12.6) *above, respectively. Let*

https://doi.org/10.1017/9781108874144.017 Published online by Cambridge University Press

$\lambda = \lambda(k,n) = k/n$ and let x_λ be the positive real solution of the equation $\mu(v;x) = 1-\lambda$. Then

$$[z^n]f(z)^k \sim v(x_\lambda)^n x_\lambda^{k-n} \frac{\lambda}{\sqrt{2\pi n \sigma^2(v;x_\lambda)}} = (1-\lambda)^{-n} v'(x_\lambda)^n \frac{\lambda x_\lambda^k}{\sqrt{2\pi n \sigma^2(v;x_\lambda)}}, \tag{12.12}$$

uniformly as λ varies over any compact subset of $(0,1)$.

Proof Exercise 12.5 below asks the reader to prove that if ψ is analytic at the origin then

$$[z^n]\psi(f(z)) = \frac{1}{n}\left[y^{n-1}\right]\psi'(y)v(y)^n, \tag{12.13}$$

a classic extension of Proposition 12.7. Assuming this result and taking $\psi(y) = y^k$, we see that

$$\begin{aligned} [z^n]f(z)^k &= \frac{k}{n}\left[y^{n-k}\right]v(y)^n \\ &= \frac{k}{n}[x^n y^{n-k}]\frac{1}{1-xv(y)}, \end{aligned} \tag{12.14}$$

representing coefficients in the powers of f as a Riordan array determined by the (known) function v. We thus apply Theorem 12.2 with $\phi \equiv 1$, after reversing the roles of x and y to obtain the Riordan array $\{a_{rs}\}$ with generating function $1/(1-yv(x))$, yielding

$$\begin{aligned} [z^n]f(z)^k &= \frac{k}{n}a_{n-k,k} \\ &\sim \lambda v(x_\lambda)^n x_\lambda^{k-n} n^{-1/2} \frac{1}{\sqrt{2\pi\sigma^2(v;x_\lambda)}}. \end{aligned}$$

The final equality in (12.12) follows from the defining equation $\mu(v;x_\lambda) = 1-\lambda$. □

Examples seen in previous chapters, including binomial coefficients and Delannoy numbers, fit into this framework, but we present a more interesting example here.

Example 12.9 (Forests of trees with restricted offspring sizes)**.** Consider the class of unlabeled plane trees with the restriction that the number of children of each node must lie in a prescribed finite subset $\Omega \subseteq \mathbb{N}$. The generating function $f(z)$ counting such trees by their number of vertices satisfies $f(z) = zv(f(z))$ where $v(z) = \sum_{j\in\Omega} y^j$ (see [FS09, Section VII.3]) and asymptotic behavior follows from Proposition 12.8. For instance, unary-binary trees are defined by

https://doi.org/10.1017/9781108874144.017 Published online by Cambridge University Press

$\Omega = \{0, 1, 2\}$, giving $v(z) = 1 + z + z^2$ and

$$\mu(v; z) = \frac{2z^2}{z^2 + z + 1}$$
$$\sigma^2(v; z) = \frac{z\left(z^2 + 4z + 1\right)}{(z^2 + z + 1)^2}.$$

When $k/n = 1/4$, so that the average tree size in a forest is 4, Proposition 12.8 implies that the number $[z^n]f(z)^k$ of forests of unary-binary trees with n nodes and k trees is asymptotically given by $C\alpha^n n^{-1/2}$ where $C \approx 0.12666642608296$ and $\alpha \approx 2.8610046903287$. ◁

12.4 Transfer matrices

The univariate transfer matrix method, discussed in Chapter 2, is easily extended to multivariate generating functions that enumerate multiplicatively weighted paths. If M is a matrix indexed by a finite set S, the ***weight*** $w(\boldsymbol{x})$ of the path $\boldsymbol{x} = (x_0, \ldots, x_n) \in S^n$ under M is $\prod_{r=1}^n M_{x_{r-1},x_r}$. The sum of weights of all paths of length n from i to j under M is given by $(M^n)_{i,j}$, yielding the generating function $F_{ij}(z) = (I - zM)^{-1}_{ij}$ enumerating weighted walks from i to j by length. Similarly, the multivariate generating function enumerating weighted walks by length while also tracking d additive integer valued functions $v_1, \ldots, v_d$ defined by their values $\boldsymbol{v}(i, j)$ on single steps (i, j) is

$$F_{\boldsymbol{v}}(\boldsymbol{y}, z)_{i,j} = (I - zM_{\boldsymbol{v}})^{-1}_{ij},$$

where $M_{\boldsymbol{v}}$ is the matrix whose (i, j)-coefficient is $M_{i,j}\boldsymbol{y}^{\boldsymbol{v}(i,j)}$. We consider two applications of this observation, message passing and quantum random walks.

Example 12.10 (Message passing)**.** Let G be the graph on $K + L + 2$ vertices which is the union of two complete graphs of sizes $K + 1$ and $L + 1$, with a loop at every vertex and one edge between them. Paths on this graph correspond to a message or task being passed around two workgroups, with communication between the workgroups not allowed except between the bosses. If we sample uniformly among paths of length n, how much time does the message spend among the common (non-boss) members?

To analyze this problem we build a new graph, with vertices $\{v_1, v_2, v_3, v_4\}$ where v_1 represents the common Group 1 members, v_2 represents the Group 1 boss, v_3 represents the Group 2 boss, and v_4 represents the common Group 2 members. Every time the message moves to v_1 it can do so in K ways, and every move to v_4 can be done in L ways. The generating function counting

https://doi.org/10.1017/9781108874144.017 Published online by Cambridge University Press

paths by time spent among the common members of each workgroup and by total length is $(I - zA)^{-1}$, where

$$A = \begin{bmatrix} Ku & Ku & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & Lv & Lv \end{bmatrix}$$

with u tracking time among common Group 1 members, v tracking time among common Group 2 members, and z tracking the total time. The entries of $(I - zA)^{-1}$ are rational functions with common denominator $Q(Ku, Lv, z)$, where

$$Q(u, v, z) = uz^2 + uz^2v - uz - uz^4v + z^2v - 2z - zv + 1 + z^3v + uz^3 ,$$

and the coefficient of z^n in the power series expansion of any entry gives a probability distribution for the amount of time spent among the common members of each group after time n. Using Gröbner bases (or simply by computing a resultant) it can be shown that the system $Q(u, v, z) = Q_z(u, v, z) = 0$ has no solutions with $u, v > 0$. Thus, by the limit theorem discussed in Theorem 12.33 below, the times spent in Group 1 and Group 2 as a portion of the length n converge to $\boldsymbol{m} = -\nabla_{\log} Q(K, L, z_0)$, where z_0 is the minimal modulus root of $f(z) = Q(K, L, z)$.

The portion of time spent among the common members of Group 1 is given by $KQ_u(K, L, z)/(z_0 Q_z(K, L, z_0))$. Plugging in $K = L = 1$, for example, we see that $Q(1, 1, z) = 1 - 4z + 3z^2 + 2z^3 - z^4$ so $z_0 \cong 0.382$ and a proportion of approximately 0.154 of the time is spent among the common members of Group 1. If bosses and employees had equal access to communication then, by symmetry, this portion would be 1/4, so the effect of communicating only through bosses reduced the time each message spends with each non-boss by nearly 40%. This effect is more marked when the workgroups have different sizes: increasing the size of the second group to 2, we plug in $K = 1$ and $L = 2$ to find that $z_0 \cong 0.311$ and the fraction of time spent among the common members in Group 1 plummets to just under 0.038. ◁

Example 12.11 (One-dimensional quantum walk)**.** In Example 9.47 the notion of a quantum random walk, and its associated spacetime generating function F, was introduced. Letting $p(r, n)$ denote the amplitude for the random walk to be at location r at time n, we have

$$F(\boldsymbol{z}) = \sum_{r,n \geq 0} p(r, n) x^r y^n = (I - yMU)^{-1},$$

where U is a $k \times k$ unitary matrix and M is a $k \times k$ diagonal matrix whose entries

https://doi.org/10.1017/9781108874144.017 Published online by Cambridge University Press

x^a run through the k possible steps (a, b) of the walk. It is shown in [BP07] that when $k = 2$ there is no loss of generality in taking U to be the real matrix

$$U_c = \begin{bmatrix} c & \sqrt{1-c^2} \\ \sqrt{1-c^2} & -c \end{bmatrix}$$

and taking the entries of M to be 1 and x, meaning that the walk either stays where it is or moves one to the right. The universal spacetime generating function for a two-dimensional quantum walk is therefore given by

$$F_c(x, y) = \frac{P_c(x, y)}{Q_c(x, y)} = \frac{P_c(x, y)}{1 - cy + cxy - xy^2},$$

where the numerator P_c depends on initial chiralities (one of k hidden states of the walk) and plays no special role. For example, if $k = 2$ and starting and ending chiralities are both in state 2 then $P_c(x, y) = 1 - cy$. ◁

Theorem 12.12 (Spacetime asymptotics for one-dimensional quantum walk)**.** *There is a real phase function $\rho(r, s)$ such that*

$$p(r, s) = \frac{2}{\pi} \frac{\lambda\sqrt{1-c^2}}{(1-\lambda)s\sqrt{-((1-c^2) - 4\lambda + 4\lambda^2)}} \cos^2(\rho(r, s)) + O\left(s^{-3/2}\right)$$

uniformly as $\lambda = \frac{r}{s}$ varies over any compact subset of the interior of $J_c = [(1-c)/2, (1+c)/2]$. Conversely, if λ varies over a compact subset of the complement of J_c then $p(r, s) \to 0$ exponentially.

The variation of probabilities in the feasible region for $c = 1/2$ is illustrated in Figure 12.2. Qualitatively similar results hold for the other starting and ending chiralities, and for combinations of chiralities.

Proof The denominator $Q = 1 - c(1-x)y - xy^2$ is quadratic in y and linear in x, so the critical point equations can be solved explicitly in radicals. Furthermore, Q satisfies the strong torality hypothesis (Definition 9.20) so that Corollary 9.21 applies. The intersection $\mathcal{V}_1$ of $\mathcal{V}$ with the unit torus is a topological circle $x = (cy-1)/(cy-y^2)$ winding twice around the torus in the x direction and once in the y direction. The logarithmic Gauss map is a smooth map on this circle with two extreme values, $(1-c)/2$ and $(1+c)/2$, and no other critical points. Therefore, for each λ in the interior of J_c there are precisely two points (x_1, y_1) and (x_2, y_2) in $\mathcal{V}_1$ with $\nabla_{\log} Q(x_j, y_j) \parallel (\lambda, 1)$ and an application of (9.7) implies

$$p(r, s) = \sum_{j=1}^{2} \frac{1 - cy_j}{\sqrt{2\pi}} x_j^{-r} y_j^{-s} \sqrt{\frac{-y_j Q_y(x_j, y_j)}{s\,\underline{Q}(x_j, y_j)}} + O\left(s^{-3/2}\right),$$

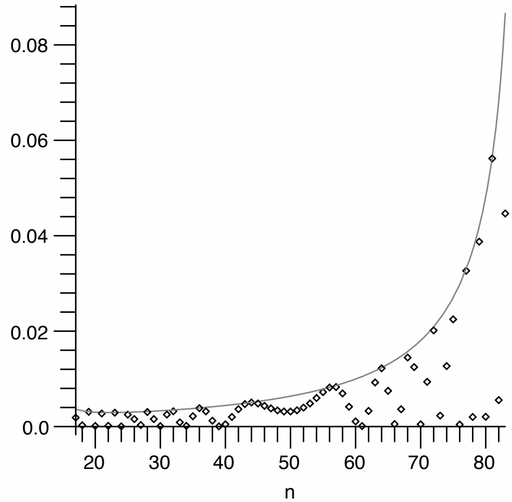

Figure 12.2 The time $n = 100$ probabilities starting and ending in state 2 when $c = 1/2$, and their upper envelope obtained by dropping the $\cos^2(\rho)$ term in Theorem 12.12.

where $\underline{Q}$ is given by (9.6). The two critical points (x_1, y_1) and (x_2, y_2) are conjugate, meaning

$$p(r,s) = 2\,\mathrm{Re}\left\{\frac{1-cy_j}{\sqrt{2\pi}} x_j^{-r} y_j^{-s} \sqrt{\frac{-y_j Q_y(x_j,y_j)}{s\,\underline{Q}(x_j,y_j)}}\right\} + O\left(s^{-3/2}\right).$$

Defining ρ to be the argument of the expression in braces, writing (x, y) for either one of the two points, and taking the square modulus, we obtain

$$p(r,s) = \frac{2}{\pi}\cos^2(\rho)\left|(1-cy)^2 \frac{-yQ_y(x,y)}{s\,\underline{Q}(x,y)}\right| + O\left(s^{-3/2}\right).$$

Using the techniques of Chapter 8, we compute that at the critical point (x, y) the quantity

$$w = (1-cy)^2 \frac{-yQ_y(x,y)}{n\,\underline{Q}(x,y)}$$

satisfies

$$\lambda^2(1-c^2) + 4\left(\frac{(1+c)}{2} - \lambda\right)\left(\lambda - \frac{(1-c)}{2}\right)(1-\lambda)^2 w^2 = 0,$$

https://doi.org/10.1017/9781108874144.017 Published online by Cambridge University Press

where λ and c are parameters. Solving, we find

$$|w| = \frac{\sqrt{1-c^2}\lambda}{(1-\lambda)s\sqrt{-((1-c^2)-4\lambda+4\lambda^2)}},$$

proving the claimed asymptotics when $\lambda \in J_c$.

If $\lambda \notin J_c$ then either $(0,0)$ does not minimize the height function in the direction $(\lambda, 1)$ on the amoeba complement component corresponding to the power series expansion of F_c, meaning $\overline{\beta} \le \beta^* < 0$, or $(0,0)$ is the minimizing point but there are no critical points on $T(1,1)$, so Theorem 11.4 implies $\overline{\beta} < 0$. In either case, $p(r,s)$ decays exponentially. □

Example 12.13 (QRWs in higher dimensions)**.** Example 9.47 of Chapter 9 shows that, in general, the spacetime generating function for a quantum walk with steps $\boldsymbol{v}^{(1)}, \ldots, \boldsymbol{v}^{(k)}$ has the form

$$F(\boldsymbol{z}) = (I - z_{d+1}MU)^{-1},$$

where $\boldsymbol{z}^\circ = (z_1, \ldots, z_d)$ are d space variables, the variable z_{d+1} tracks time, and M is the diagonal matrix whose (j,j)-entry is $(\boldsymbol{z}^\circ)^{\boldsymbol{v}^{(j)}}$. The common denominator of the F_{ij} is $Q(\boldsymbol{z}) = \det(I - z_{d+1}MU)$. As noted in Example 9.47, the feasible velocity region R is the image of the logarithmic Gauss map and the limit law for the amplitudes can be written, up to an oscillatory term, in terms of the Gaussian curvature.

Consider the two-dimensional family of walks with the nearest neighbor unit vectors $\boldsymbol{v}^{(1)}, \ldots, \boldsymbol{v}^{(4)}$ as steps and the unitary matrix

$$U = S(p) = \begin{pmatrix} \frac{\sqrt{p}}{\sqrt{2}} & \frac{\sqrt{p}}{\sqrt{2}} & \frac{\sqrt{1-p}}{\sqrt{2}} & \frac{\sqrt{1-p}}{\sqrt{2}} \\ -\frac{\sqrt{p}}{\sqrt{2}} & \frac{\sqrt{p}}{\sqrt{2}} & -\frac{\sqrt{1-p}}{\sqrt{2}} & \frac{\sqrt{1-p}}{\sqrt{2}} \\ \frac{\sqrt{1-p}}{\sqrt{2}} & -\frac{\sqrt{1-p}}{\sqrt{2}} & -\frac{\sqrt{p}}{\sqrt{2}} & \frac{\sqrt{p}}{\sqrt{2}} \\ -\frac{\sqrt{1-p}}{\sqrt{2}} & -\frac{\sqrt{1-p}}{\sqrt{2}} & \frac{\sqrt{p}}{\sqrt{2}} & \frac{\sqrt{p}}{\sqrt{2}} \end{pmatrix}.$$

Setting $\alpha = \sqrt{2p}$ to simplify notation, we have

$$\begin{aligned} Q(x,y,z) &= (x^2y^2 + y^2 - x^2 - 1 + 2xyz^2)z^2 - 2xy \\ &\quad - \alpha z(xy^2 - y - x + z^2y - z^2x + z^2xy^2 + z^2x^2y - x^2y) \end{aligned}$$

and a lexicographic Gröbner Basis for the ideal generated by Q and its partial derivatives includes the polynomial

$$z\alpha^2(\alpha^2-1)(\alpha^2-2) = 2zp(2p-1)(2p-2).$$

The only root of this polynomial with $x, y, z \neq 0$ and $0 < p < 1$ occurs when

https://doi.org/10.1017/9781108874144.017 Published online by Cambridge University Press

$p = 1/2$, and back substitution of this value implies that $-z+z^5 = z^3+2y-z = -z-z^3+2x = 0$. The first of these polynomials vanishes on the unit circle when $z \in \{\pm 1, \pm i\}$, however when $z = \pm 1$ the second polynomial only vanishes when $y = 0$, and when $z = \pm i$ the third polynomial only vanishes when $x = 0$, so Q and its partial derivatives do not simultaneously vanish on $T(\mathbf{1})$.

Thus, if Q_z vanishes at some point (a, b, c) of $\mathcal{V}_1 = \mathcal{V} \cap T(\mathbf{1})$ then (a, b, c) contributes to nonvanishing asymptotics in a direction $(r, s, 0)$ for some $(r, s) \neq (0, 0)$. This is ruled out from our knowledge of the generating function, because the velocity of QRW is at most the longest step, so Q_z does not vanish on $\mathcal{V}_1$ and the projection $(x, y, z) \to (x, y)$ is a smooth four-fold cover of the unit torus in $\mathbb{C}^2$. There are many four-fold covers of the 2-torus, but in this case some trigonometry [Bar+10, Proposition 4.6] shows that $\mathcal{V}_1$ is in fact the union of four 2-tori, each mapping diffeomorphically to the 2-torus under the logarithmic Gauss map. Figure 12.13 shows the four components for the parameter value $p = 1/2$ by graphing z as a function of x and y with the torus depicted as the unit cube with wraparound boundary conditions.

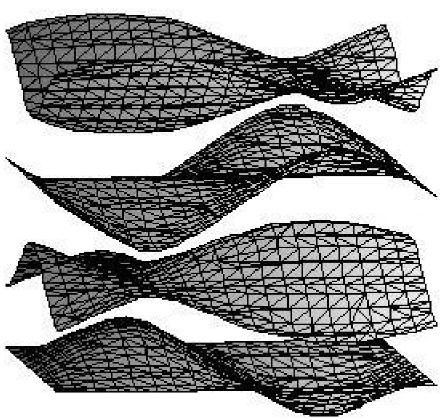

Figure 12.3 The four tori comprising $\mathcal{V}_1$ for $S(1/2)$.

Asymptotic behavior follows from Corollary 9.46 in Chapter 9.

Theorem 12.14. *For each $\hat{\mathbf{r}}$ in the image of the logarithmic Gauss map on $\mathcal{V}_1$*

https://doi.org/10.1017/9781108874144.017 Published online by Cambridge University Press

let **W** *be the set of four preimages of* $\boldsymbol{r}$ *in* $\mathcal{V}_1$ *for the* $S(p)$ *walk with* $0 < p < 1$ *and fixed states i and j. Then, as* $\boldsymbol{r} \to \infty$ *with* $\hat{\boldsymbol{r}}$ *in a compact subset of the feasible velocity region R, the amplitude* $p(\boldsymbol{r})$ *satisfies*

$$p(\boldsymbol{r}) = (-1)^{\delta} \frac{1}{2\pi|\boldsymbol{r}|} \sum_{\boldsymbol{z} \in \mathbf{W}} \boldsymbol{z}^{-\boldsymbol{r}} \frac{P(\boldsymbol{z})}{\| \nabla_{\log} Q(\boldsymbol{z}) \|_2^2} \frac{1}{\sqrt{|\mathcal{K}(\boldsymbol{z})|}} e^{-i\pi\tau(\boldsymbol{z})/4} + O\left(|\boldsymbol{r}|^{-3/2}\right), \tag{12.15}$$

where $\delta = 1$ *if* $\nabla_{\log} Q$ *is a negative multiple of* $\hat{\boldsymbol{r}}$ *(to account for the absolute value in Corollary 9.46) and zero otherwise.*

◁

12.5 Higher order asymptotics

Our results give effective methods for computing asymptotic expansions of multivariate generating functions. Although the first term in such an expansion generically dictates behavior of the sequence under consideration, there are several reasons for wanting to compute higher order terms. Most obviously, computing more terms in such expansions gives better approximations.

Example 12.15. Consider the $(4, 3)$-diagonal of the Delannoy numbers with generating function $F(x, y) = 1/(1 - x - y - xy)$. The critical points in this direction are $(-2, -3)$ and $(1/2, 1/3)$, both of which are smooth, and the point $\boldsymbol{w} = (1/2, 1/3)$ is strictly minimal. Corollary 5.17 implies

$$[x^{4n} y^{3n}] F(x, y) = 432^n \left(a_1 n^{-1/2} + a_2 n^{-3/2} + O(n^{-5/2})\right)$$

as $n \to \infty$, where

$$a_1 = \frac{\sqrt{30}}{10\sqrt{\pi}} \approx 0.3090193616$$

$$a_2 = -\frac{\sqrt{30}}{288\sqrt{\pi}} \approx -0.01072983895\,.$$

Comparing this approximation with the actual Delannoy numbers for small n gives (after scaling out the exponential growth) the results in Table 12.5. The error in the 1-term approximation clearly decays as $1/n$ while the 2-term error decays as $1/n^2$ – note the extreme accuracy even for $n = 1$. The question of the optimal order to which to truncate an asymptotic series in a given application goes beyond our scope here.

https://doi.org/10.1017/9781108874144.017 Published online by Cambridge University Press

n	1	2	4	8	16
Exact	0.2986	0.2148	0.1532	0.1088	0.07709
1-term	0.3090	0.2185	0.1545	0.1093	0.07726
Rel. error	-0.03486	-0.01742	-0.008698	-0.004345	-0.002171
2-term	0.2983	0.2147	0.1532	0.1088	0.07709
Rel. error	0.001077	0.0002450	0.00005820	0.00001417	0.000003496

Table 12.5 *Approximations to scaled Delannoy numbers* $432^n F_{4n,3n}$.

Higher order terms in an asymptotic expansion for any direction $\hat{\boldsymbol{r}}$ with positive coordinates can be computed symbolically with the coordinates of $\hat{\boldsymbol{r}}$ as parameters. ◁

Although the first terms in our asymptotic expansions are typically non-zero, and thus determine dominant asymptotic behavior, they can vanish. In this case, higher order terms need to be computed in order to determine how the sequence behaves.

Example 12.16. The rational functions

$$F_1(x, y) = \frac{1}{1 - x - y}, \quad F_2(x, y) = \frac{y(1 - 2y)}{1 - x - y}, \text{ and } F_3(x, y) = \frac{x - y}{1 - x - y}$$

all admit $(x, y) = (1/2, 1/2)$ as a minimal contributing point in the main diagonal direction. Applying the results of smooth ACSV gives an asymptotic expansion (9.4) of the coefficients of F_1 whose leading term is non-zero. The leading term in the asymptotic expansion of the coefficients of F_2 vanishes, as the numerator of F_2 vanishes at the contributing point, but the second order term in this expansion is non-zero (in fact, the main diagonal of F_2 is the generating function of the Catalan numbers – see Example 12.18 below). On the other hand, it can be shown that the main diagonal of F_3 is identically zero, so all coefficients in the asymptotic expansion must vanish. It is not obvious how to detect automatically the vanishing of all terms in the asymptotic expansion for a general rational function. ◁

Another reason why we may need higher order terms is because of cancellation of terms when we combine asymptotic expansions of related functions.

Example 12.17. Consider the $(d + 1)$-variate function

$$W(x_1, \ldots, x_d, y) = \frac{A(\boldsymbol{x})}{1 - yB(\boldsymbol{x})},$$

https://doi.org/10.1017/9781108874144.017 Published online by Cambridge University Press

where

$$A(\boldsymbol{x}) = \frac{1}{1 - \sum_{j=1}^{d} \frac{x_j}{1+x_j}}$$

$$B(\boldsymbol{x}) = 1 - (1 - e_1(\boldsymbol{x}))A(\boldsymbol{x})$$

for the elementary symmetric function $e_1(\boldsymbol{x}) = \sum_{i=1}^{d} x_j$. The combinatorial constructions discussed in Chapter 2 imply that W enumerates words over the alphabet $\{1, \ldots, d\}$, where x_j marks occurrences of the letter j and y marks occurrences of ***snaps***, which are nonoverlapping pairs of duplicate letters. The factor $A(\boldsymbol{x})$ counts snapless words over X, which are the Smirnov words described in Example 2.11. If ψ denotes the random variable counting snaps among words with n occurrences of each letter then the expected value and variance of ψ satisfy

$$\mathbb{E}(\psi) = \frac{[\boldsymbol{x}^{n\mathbf{1}}]W_y(\boldsymbol{x}, 1)}{[\boldsymbol{x}^{n\mathbf{1}}]W(\boldsymbol{x}, 1)} = \frac{[\boldsymbol{x}^{n\mathbf{1}}]A(\boldsymbol{x})^{-1}B(\boldsymbol{x})(1 - e_1(\boldsymbol{x}))^{-2}}{[\boldsymbol{x}^{n\mathbf{1}}](1 - e_1(\boldsymbol{x}))^{-1}}$$

and $\mathbb{V}(\psi) = \mathbb{E}(\psi^2) - \mathbb{E}(\psi)^2$ with

$$\begin{aligned}\mathbb{E}(\psi^2) &= \frac{[x^{n\mathbf{1}}]\left(W_{yy}(\boldsymbol{x}, 1) + W_y(\boldsymbol{x}, 1)\right)}{[\boldsymbol{x}^{n\mathbf{1}}]W(\boldsymbol{x}, 1)} \\ &= \frac{[\boldsymbol{x}^{n\mathbf{1}}]A(\boldsymbol{x})^{-2}B(\boldsymbol{x})(B(\boldsymbol{x}) + 1)(1 - e_1(\boldsymbol{x}))^{-3}}{[\boldsymbol{x}^{n\mathbf{1}}](1 - e_1(\boldsymbol{x}))^{-1}}.\end{aligned}$$

Each of the above coefficient extractions applies to a rational function whose denominator is a power of $Q(\boldsymbol{x}) = 1 - e_1(\boldsymbol{x})$. By Lemma 6.41, asymptotics can be determined by applying the smooth ACSV results of Chapter 9 to the strictly minimal critical point $(1/d, \ldots, 1/d)$ of each function, giving

$$\begin{aligned}\mathbb{E}(\psi) &= \frac{\frac{3\sqrt{3}}{8\pi} - \frac{61\sqrt{3}}{192\pi}n^{-1} + O(n^{-2})}{\frac{\sqrt{3}}{2\pi}n^{-1} - \frac{\sqrt{3}}{9\pi}n^{-2} + O(n^{-3})} \\ &= (3/4)n - 15/32 + O(n^{-1})\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}(\psi^2) &= \frac{\frac{9\sqrt{3}}{32\pi}n - \frac{35\sqrt{3}}{128\pi} + O(n^{-1})}{\frac{\sqrt{3}}{2\pi}n^{-1} - \frac{\sqrt{3}}{9\pi}n^{-2} + O(n^{-3})} \\ &= (9/16)n^2 - (27/64)n + O(1)\end{aligned}$$

so that $\mathbb{V}(\psi) = (9/32)n + O(1)$. Note that determining the leading term of $\mathbb{V}(\psi)$ requires the second-order terms in the expansions of $\mathbb{E}(\psi)$ and $\mathbb{E}(\psi^2)$.

Comparing these approximations when $d = 3$ with the actual values for small n, we obtain the following table.

n	1	2	4	8
$\mathbb{E}(\psi)$	0	1.00000	2.50909	5.52056
$(3/4)n$	0.75	1.5	3	6
$(3/4)n - 15/32$	0.28125	1.03125	2.53125	5.53125
One-term relative error	Undefined	0.50000	0.19565	0.086846
Two-term relative error	Undefined	0.031250	0.0088315	0.0019363
$\mathcal{V}(\psi)$	0	0.8	1.20057	2.31961
$(9/32)n$	0.28125	0.5625	1.125	2.25
Relative error	Undefined	0.29688	0.062942	0.030013

◁

12.6 Algebraic generating functions

Algebraic generating functions in one variable can be analyzed asymptotically by the transfer theorems of Section 3.4 in Chapter 3, and often exactly by Lagrange inversion as described in Section 12.3 above. Of course, the situation in several variables is more complicated. For the purposes of asymptotics, we can sometimes ignore the fact that the generating function is algebraic, because the contributing points determining asymptotics might occur in regions where the generating function is meromorphic (this happens in several Riordan array examples). The more difficult cases occur when asymptotics are determined by an algebraic singularity. Assuming that we have an analytic branch at the origin, we aim to compute asymptotics of the coefficients of its power series expansion as done for rational and meromorphic functions in previous chapters.

Section 2.4 gives an approach to asymptotics by embedding algebraic generating functions as subseries of higher-dimensional rational functions. Unfortunately, there are several difficulties with this approach. First, the results of Chapters 9–11 cannot capture asymptotics of all algebraic generating functions: for any $\beta \in \mathbb{Z}$ there exists a univariate sequence a_n with an algebraic generating function such that $a_n \sim C\alpha^n n^\beta$, but our results for nondegenerate critical points only give asymptotics of this form with $\beta = k/2$ for $k \in \mathbb{Z}$. This asymptotic behavior also shows that the leading terms in our asymptotic expansions often vanish when applied to rational functions constructed by such lifting procedures (Exercise 12.11 asks you to prove that this vanishing always occurs for an embedding given by the method presented in Proposition 2.36).

Algebraic generating functions are often well-behaved as they are built from nice combinatorial constructions, but the rational functions obtained from a lifting procedure are built from algebraic machinery and are thus often worse

https://doi.org/10.1017/9781108874144.017 Published online by Cambridge University Press

behaved. For instance, the lifted functions usually have negative power series coefficients, making the determination of minimal points harder to verify, and may have critical points at infinity. Finally, there is no guarantee that the rational generating function will have a contributing singularity of a type that we can deal with using current technology.

Despite these qualifications, the rational embedding approach does work in many situations.

Example 12.18. Consider the shifted Catalan number GF $f(x) = \sum_n a_n x^n = (1 - \sqrt{1-4x})/2$ with minimal polynomial $P(x,y) = y^2 - y + x$. Proposition 2.34 implies that $f(x)$ is the main diagonal of the bivariate rational function $F(y,z) = y(1-2y)/(1-x-y)$. The singular variety $\mathcal{V}$ of F is globally smooth and the critical point equations yield the single solution $(1/2, 1/2)$ which is strictly minimal. The first term in the expansion (9.4) vanishes, as expected, while the second term does not vanish and implies $a_n \sim 4^{n-1}/\sqrt{\pi n^3}$. ◁

Remark 12.19. For bivariate algebraic singularities it is also possible to determine asymptotics through a transfer theorem [Gre18].

Example 12.20. As shown in Example 2.37, the generating function for the Narayana numbers a_{rs} is the subseries of

$$F(w,x,y) = \frac{w(1-2w-wx(1-y))}{1-w-xy-wx(1-y)}$$

consisting of terms whose powers of w and x are equal. Since F can be written

$$F(w,x,y) = \frac{(1-w)^{-1}}{1-xy-\frac{wx}{1-w}},$$

its power series expansion is combinatorial. To determine asymptotics of $a_{\alpha n,\beta n}$ as $n \to \infty$ with $\alpha > \beta > 0$, we compute asymptotics of the power series coefficients of F in the direction (α,α,β). The critical point equations yield the unique solution

$$(w,x,y) = \left(\frac{\beta}{\alpha}, \frac{(\alpha-\beta)^2}{\alpha\beta}, \frac{\beta^2}{(\alpha-\beta)^2}\right),$$

which can be shown to be minimal as F is combinatorial. An asymptotic expansion can thus be computed using Theorem 9.5 in Chapter 9. In particular, the exponential growth of $a_{\alpha n,\beta n}$ is $\left(\frac{\alpha^\alpha}{\beta^\beta(\alpha-\beta)^{\alpha-\beta}}\right)^2$, with the maximum exponential growth of 4 achieved when $\alpha/\beta = 2$. ◁

One positive feature of the embedding approach is that there are many possible embeddings of a given algebraic generating function, so if a rational func-

https://doi.org/10.1017/9781108874144.017 Published online by Cambridge University Press

tion obtained by one embedding procedure is too difficult to analyze, it may be possible to find one better suited to ACSV.

Example 12.21. The combinatorial univariate algebraic generating function $\sqrt{1-x}-1 = \sum_n a_n x^n$ is the main diagonal of

$$\frac{2\,(Yx-1)(Y+1)Y}{Y^2x+2\,Yx-Y+x-2}.$$

It is easily checked that among affine points of the singular variety, there are no non-smooth points and no critical points for the main diagonal direction, so asymptotics are determined by a critical point at infinity. The generating function $2xY/(2-x-Y)$ has the same main diagonal but is much easier to analyze, giving $a_n \sim (\pi n)^{-1/2}$, as we expect from Newton's binomial theorem and Stirling's approximation. ◁

See Section 13.5 for further discussion on ACSV and algebraic generating functions.

12.7 Additional worked examples

We now run through several other examples.

Example 12.22 (A constant coefficient bivariate recurrence)**.** Consider the constant coefficient linear recurrence

$$f(m,n) = f(m-1,n-1) + f(m-1,n-2) + 2f(m-2,n-1) \qquad (m,n \geq 2)$$

with boundary conditions $f(m,n) = 0$ if $m < 0$ or $n < 0$ and $f(m,n) = 1$ when $0 \leq m \leq 1$ or $0 \leq n \leq 1$. Grau Ribas [Gra] asked for the limit of $f(n+1,n+1)/f(n,n)$ which, based on numerical computations such as $f(100,100)/f(99,99) = 2.70265\ldots$, appears to exist and be close to e.

Introducing the generating function $F(x,y) = \sum_{m,n} f(m,n)x^m y^n$, the defining recurrence of f implies

$$F(x,y) = \frac{1}{Q(x,y)}$$

with $Q(x,y) = 1 - xy - xy^2 - 2x^2y$. Gröbner basis computations verify that $\mathcal{V}$ is smooth, and the critical points in the main diagonal direction are the points $(x, 2x)$ where x is a root of $8x^3 + 2x^2 - 1$. If $w = 0.85\ldots$ denotes the unique real root of $8x^3 + 2x^2 - 1$ then Lemma 6.41 implies that $(w, 2w)$ is minimal, and it is the only critical point on the torus $T(w, 2w)$. The limit

$$\frac{f(n+1,n+1)}{f(n,n)} \to \frac{1}{2w^2}$$

https://doi.org/10.1017/9781108874144.017 Published online by Cambridge University Press

is algebraic, and therefore not equal to e. Reducing $R = 1/(2w^2)$ modulo the minimal polynomial for w implies R is the unique positive root of the polynomial $8 - R + 2R^2 - R^3$. ◁

Example 12.23 (Horizontally Convex Polyominoes)**.** The generating function

$$F(x,y) = \sum_{r,s\geq 0} a_{rs}x^r y^s = \frac{xy(1-x)^3}{(1-x)^4 - xy(1-x-x^2+x^3+x^2y)} \tag{12.16}$$

enumerates ***horizontally convex polyominoes*** (HCPs) [Pól69; Odl95; Wil06; Sta97] by total size r and number of rows s. Letting Q denote the denominator of F in (12.16), we know from Example 8.4 that $\mathcal{V}$ is smooth except at the point $(1,0)$. By Proposition 6.38, there is a part of the graph of Q in the first quadrant consisting of minimal points, which are the points shown in Figure 12.4 with $x \leq 1$.

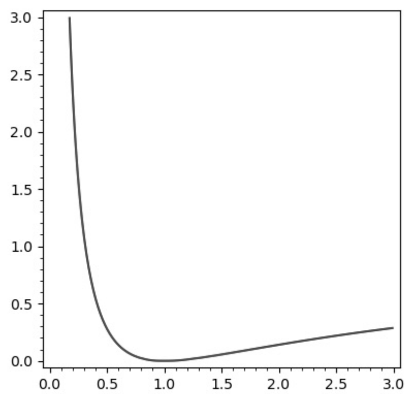

Figure 12.4 Minimal points of $\mathcal{V}$ in the positive real quadrant.

The only combinatorially interesting directions occur in $\Xi = \{(\hat{r},\hat{s}) : 0 < \hat{s} < 1/2\}$, because an HCP cannot have more rows than its size. As $\hat{\mathbf{r}}$ varies over Ξ from the horizontal to the diagonal, the unique contributing point in the direction $\hat{\mathbf{r}}$ moves along this graph from $(1,0)$ to $(0,\infty)$. The numerator $P = xy(1-x)^3$ of F in (12.16) is nonvanishing on this component and, using Gröbner bases, we find that the quantity $\underline{Q}$ in the asymptotic expansion (9.7) does not vanish at any of these contributing points.

Since for each direction there are only finitely many critical points, and no others lie on the same torus as the one we have identified, it follows from Theorem 9.12 that asymptotics for a_{rs} are uniform as s/r varies over a compact subset of the interval $(0,1)$ and have the form $a_{rs} \sim Cx^{-r}y^{-s}s^{-1/2}$. Algebraic

https://doi.org/10.1017/9781108874144.017 Published online by Cambridge University Press

methods may then be used to determine x, y, and C as explicit functions of $\lambda = s/r$, giving asymptotics for the number of HCPs that are uniform as long as s/r remains in a compact subinterval of $(0, 1)$. For instance, when $\lambda = 1/2$ the smooth critical points (x, y) satisfy $3x^2 + 18x - 5$ and $75y^2 - 288y + 256$, with the contributing singularity occurring at

$$(x_0, y_0) = \left(\sqrt{\frac{32}{3}} - 3, \frac{48 - \sqrt{512}}{25} \right) \approx (0.265986, 1.397442),$$

and a floating point computation gives $a_{n,n/2} \sim (0.2373\ldots)(3.1803\ldots)^n n^{-1/2}$. Note that the exponential growth of $3.1803\ldots$ in this case is only slightly less than the exponential growth $3.2055\ldots$ for all HCPs – a reflection of the fact that the exponential growth varies quadratically around its maximum.

We can also derive limit laws. For instance, let $q(x) = Q(x, 1) = 1-5x+7x^2-4x^3$ and let $a \approx 0.3120$ be the root of q with minimum modulus. Theorem 12.33 below, with the roles of x and y switched, proves that the number of rows k of a uniformly chosen HCP of total size n satisfies $k/n \to m$ in probability, where

$$m = \frac{yQ_y(a, y)}{aQ_x(a, y)} = \frac{5 - 9a + 11a^2}{4(5 - 14a + 12a^2)} = \frac{1}{2.207\ldots}.$$

In other words, the average row size converges to a little over 2.2.

◁

Example 12.24 (Symmetric Eulerian numbers)**.** The symmetric ***Eulerian numbers*** $A(r, s)$ count the number of permutations of the set $[r+s+1] = \{1, 2, \ldots, r+s+1\}$ with precisely r descents, and admit the exponential generating function

$$F(x, y) = \frac{e^x - e^y}{xe^y - ye^x} = \sum_{r,s \geq 0} \frac{A(r, s)}{r!\, s!}, \tag{12.17}$$

see [Com74, page 246] and [GJ04, p. 2.4.21]. To represent the numerator and denominator of F as analytic functions with no common divisor we factor out a term $x-y$ from each, writing $F(x, y) = P(x, y)/Q(x, y)$ for $P = (e^x - e^y)/(x-y)$ and $Q = (xe^y - ye^x)/(x - y)$. We know by its combinatorial definition that the power series coefficients of F are nonnegative, and the power series expansion of Q is aperiodic, so the results of Section 6.4 imply that the minimal points of F have positive coordinates. The graph of Q in this quadrant of $\mathbb{R}^2$ is monotone decreasing with x and asymptotic to both axes; see Figure 12.5.

It can be verified by a direct computation that the singular variety of F is smooth, and the quantity $\underline{Q}$ in the asymptotic expansion (9.7) does not vanish at positive real points – in fact, $\underline{Q}(x, y)$ reaches its minimum value of $e^3/12$ at the point $(1, 1)$. The logarithmic gradient of Q at any point (x, y) in the first

https://doi.org/10.1017/9781108874144.017 Published online by Cambridge University Press

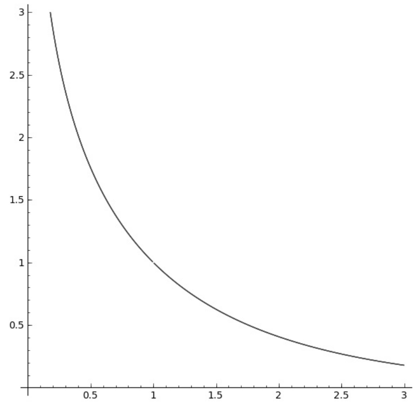

Figure 12.5 The zero set of $(xe^y - ye^x)/(x - y)$ in the first quadrant.

quadrant is a non-zero multiple of $(\alpha, 1 - \alpha)$, where

$$\alpha = \frac{xQ_x(x, y)}{xQ_x(x, y) + yQ_y(x, y)},$$

which simplifies to

$$\alpha = \frac{1 - x}{y - x}$$

for points on $\mathcal{V}$, except when $(x, y) = (1, 1)$ and $\alpha = 1/2$. Thus, Theorem 9.5 implies that asymptotics in the direction $(\alpha, 1 - \alpha)$ can be computed by solving the system

$$\alpha = \frac{1 - x}{y - x}, \qquad xe^y = ye^x$$

for positive real values (x, y) and using the expansion (9.7). The result, recalling that F is an exponential generating function, is that $A(r, s)$ is asymptotically estimated by

$$A(r, s) \sim C_\alpha (r + s)^{-1/2} \gamma^{r+s} r! s!$$

where $\gamma = x^{-\alpha} y^{\alpha - 1}$ and C_α is a messy constant determined by (9.6) and (9.7). ◃

Example 12.25 (Number of successes in a coin-flipping game)**.** Consider a single player game with biased coins, so that heads appears with probability $p = 2/3$ for the first n flips and $p = 1/3$ thereafter. The player is told to get r heads and s tails, and is allowed to choose n. On average, how many choices of $n \leq r + s$ will be winning choices?

Decomposing the probability that n is a winning choice for the player by how many heads are rolled in the first n throws shows that the probability is

$$\sum_{a+b=n} \binom{n}{a} (2/3)^a (1/3)^b \binom{r+s-n}{r-a} (1/3)^{r-a} (2/3)^{s-b} .$$

If a_{rs} is the sum of the winning probability over all n, then the array $\{a_{rs}\}$ is the convolution of the arrays $\binom{r+s}{r}(2/3)^r(1/3)^s$ and $\binom{r+s}{r}(1/3)^r(2/3)^s$, so the generating function $F(x,y) = \sum_{rs} a_{rs} x^r y^s$ is the product

$$F(x,y) = \frac{P(x,y)}{Q(x,y)} = \frac{1}{\left(1 - \frac{1}{3}x - \frac{2}{3}y\right)\left(1 - \frac{2}{3}x - \frac{1}{3}y\right)} .$$

The complete intersection asymptotic results of Chapter 10 imply that $a_{rs} = 3$ plus an error term which is exponentially small as $r, s \to \infty$ provided that $r/(r+s)$ stays in a compact subinterval of $(1/3, 2/3)$. A purely combinatorial analysis of the sum may be carried out to yield the leading term 3, but says nothing about the error terms. ◃

Example 12.26 (Lattice paths constrained to a quadrant)**.** Consider the class of lattice paths on $\mathbb{Z}^2$ that start at the origin and must remain in the nonnegative quadrant $\mathbb{N}^2$ while using the set $\{(1,0),(1,-1),(-1,0),(-1,1)\}$ of allowable steps. The *kernel method* [Mel21, Chapter 4] allows one to prove that the generating function enumerating such walks by number of steps is the main diagonal of

$$F(x,y,t) = \frac{(x+1)(x^{-2} - y^{-1})(x-y)(x+y)}{1 - xyt(x + xy^{-1} + yx^{-1} + x^{-1})}.$$

This is a rational function with smooth singular variety, and minimal critical points $\boldsymbol{p}_1 = (1,1,1/4)$ and $\boldsymbol{p}_2 = (-1,1,1/4)$. As the numerator of F has a zero of order 2 at $\boldsymbol{p}_1$ and a zero of order 3 at $\boldsymbol{p}_2$, only $\boldsymbol{p}_1$ contributes to dominant asymptotics. Computation with a computer algebra system shows that the counting sequence for the number of walks on these steps satisfies

$$s_n = \frac{4^n}{n^2} \cdot \frac{8}{\pi} + O\left(\frac{4^n}{n^3}\right).$$

ACSV provides a powerful framework to determine asymptotics of such lattice path models [Mel21, Chapters 6 and 9]. ◃

Example 12.27 (alignments)**.** Fix a positive integer d. The $(d, \boldsymbol{n})$*-alignments* are $d \times m$ binary matrices for some m, such that no columns are identically zero and the ith row sum is n_i. These structures have relevance to bioinformatics [RT05; Wat95], and the generating function enumerating such alignments

https://doi.org/10.1017/9781108874144.017 Published online by Cambridge University Press

by multi-index $\boldsymbol{n}$ is

$$F(\boldsymbol{z}) = \frac{1}{2 - \prod_{i=1}^{d}(1+z_i)} = \sum_{\boldsymbol{n}\in\mathbb{N}^d} a_{\boldsymbol{n}} \boldsymbol{z}^{\boldsymbol{n}}.$$

In the main diagonal direction, the symmetry of F and the combinatorial aperiodic nature of the problem combine with Lemma 6.41 to imply that there is a single, strictly minimal, contributing point $(z, \ldots, z)$ in the positive orthant, where $z = 2^{1/d} - 1$. Theorem 9.5 then yields

$$a_{n,\ldots,n} \sim \frac{2^{(1-d^2)/d}}{(2^{1/d}-1)\sqrt{d\pi^{d-1}}}(2^{1/d}-1)^{-kn} n^{-\frac{d-1}{2}},$$

recovering a result of [Gri+90]. ◁

Example 12.28 (integer solutions to linear equations)**.** Let A be a $d \times m$ matrix of nonnegative integers and, for $\boldsymbol{r} \in \mathbb{N}^d$, let $a_{\boldsymbol{r}}$ denote the number of nonnegative integer solutions to $A\boldsymbol{x} = \boldsymbol{r}$. The generating function for the array $\{a_{\boldsymbol{r}}\}$ is given by

$$F(\boldsymbol{z}) = \sum_{\boldsymbol{r}\in\mathbb{N}^d} a_{\boldsymbol{r}} \boldsymbol{z}^{\boldsymbol{r}} = \prod_{j=1}^{m} \frac{1}{1 - \boldsymbol{z}^{Ae_j^T}},$$

where e_j is the jth elementary vector. This enumeration problem is discussed at length in [DS03] (see also [Sta97, Section 4.6]) which uses the running example

$$A = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix}$$

with generating function

$$F(\boldsymbol{z}) = \frac{1}{Q_1 Q_2 Q_3 Q_4 Q_5} = \frac{1}{(1-x)(1-y)(1-z)(1-xy)(1-xz)}.$$

The divisors of F are all binomials of the form $1 - \boldsymbol{z}^{\boldsymbol{\alpha}}$, and their zero sets all intersect at $(1,1,1)$. The logarithmic gradients of the divisors at $(1,1,1)$ are the columns of A.

Every triple of columns of A except for $(1,2,4)$ and $(1,3,5)$ forms a linearly independent set. The circuits of the matroid defined by A are therefore these triples and the only quadruple $(2,3,4,5)$ not containing either triple, so the broken circuits are $(1,2), (1,3)$, and $(2,3,4)$, and the bases containing no broken circuit are $(1,4,5), (2,3,5), (2,4,5)$, and $(3,4,5)$. By Theorem 10.33, a

https://doi.org/10.1017/9781108874144.017 Published online by Cambridge University Press

basis for rational functions with simple poles on $\mathcal{V}_{Q_1}, \ldots, \mathcal{V}_{Q_5}$ is given by

$$\left\{ \frac{1}{Q_1Q_4Q_5}, \frac{1}{Q_2Q_3Q_5}, \frac{1}{Q_2Q_4Q_5}, \frac{1}{Q_3Q_4Q_5} \right\}.$$

To reduce F to the sum of terms whose support is in $\mathtt{BC}(F)$ we use relations expressing Q_i in terms of $\{Q_j : j \in C \setminus \{i\}\}$, where i is the greatest element in a circuit. A little scratch work uncovers these relations for the respective circuits $(1, 2, 4), (1, 3, 5)$, and $(2, 3, 4, 5)$, giving

$$Q_4 = Q_1 + Q_2 - Q_1Q_2 \tag{12.18a}$$

$$Q_5 = Q_1 + Q_3 - Q_1Q_3 \tag{12.18b}$$

$$Q_5 = -y^{-1}Q_2 + y^{-1}Q_3 + xQ_4 . \tag{12.18c}$$

The first relation (12.18a) divided by QQ_4 yields

$$\frac{1}{Q} = \frac{1}{Q_2Q_3Q_4^2Q_5} + \frac{1}{Q_1Q_3Q_4^2Q_5} - \frac{1}{Q_3Q_4^2Q_5} . \tag{12.19}$$

We are finished manipulating the third term of (12.19), as its support $\{3, 4, 5\}$ is in $\mathtt{BC}(F)$. The second term of (12.19), after an application of (12.18b), becomes

$$\frac{1}{Q_3Q_4^2Q_5^2} + \frac{1}{Q_1Q_4^2Q_5^2} - \frac{1}{Q_4^2Q_5^2} .$$

The first term of (12.19), after an application of (12.18c), yields

$$\frac{-1/y}{Q_3Q_4^2Q_5^2} + \frac{1/y}{Q_2Q_4^2Q_5^2} + \frac{x}{Q_2Q_3Q_4Q_5^2}$$

and, using (12.18c) once again on the last of these three terms, replaces that term by

$$\frac{-x/y}{Q_3Q_4Q_5^3} + \frac{x/y}{Q_2Q_4Q_5^3} + \frac{x^2}{Q_2Q_3Q_5^3} .$$

Putting this all together gives the decomposition

$$\begin{aligned} F = &-\frac{1}{Q_3Q_4^2Q_5} + \frac{1}{Q_3Q_4^2Q_5^2} + \frac{1}{Q_1Q_4^2Q_5^2} - \frac{1}{Q_4^2Q_5^2} \\ &+ \frac{-1/y}{Q_3Q_4^2Q_5^2} + \frac{1/y}{Q_2Q_4^2Q_5^2} \\ &+ \frac{-x/y}{Q_3Q_4Q_5^3} + \frac{x/y}{Q_2Q_4Q_5^3} + \frac{x^2}{Q_2Q_3Q_5^3} , \end{aligned}$$

and the techniques of Chapter 10 can now be applied to each term to determine asymptotic behavior. ◁

https://doi.org/10.1017/9781108874144.017 Published online by Cambridge University Press

Example 12.29 (Serial Dictatorship)**.** The trivariate sequence

$$a_{njr} = \begin{cases} \dfrac{\binom{r-s}{t-s}}{\binom{r}{t-1}} & \text{if } 1 \le s \le t \le r \\ 0 & \text{otherwise} \end{cases}$$

arises in the study of the Serial Dictatorship algorithm for allocating indivisible goods (it gives the probability, under IID uniform strict preferences of r agents over r items, that the sth agent receives its tth most preferred item).

The generating function for the numerators $b_{rst} = \binom{r-s}{t-s}$ is

$$F(x,y,z) = \sum_{r,s,t \ge 0} b_{rst} x^r y^s z^t = \frac{xyz}{(1-x-xz)(1-xyz)}.$$

Let $Q_1(x,y,z) = 1 - z - xz$ and $Q_2(x,y,z) = 1 - xyz$. Because Q_1 is independent of y, points on the stratum where Q_1 vanishes but Q_2 is non-zero can only be critical points for directions (r,s,t) with $s = 0$. Similarly, the stratum where Q_1 is non-zero but Q_2 vanishes contains only critical points in the main diagonal direction (which results in a trivial sequence). Thus, all interesting directions correspond to points where both factors vanish, which is the curve parametrized by $(x, 1/(1-x), (1-x)/x)$ for $x \neq 0$.

A point on this curve is a contributing point for asymptotics in the direction (r,s,t) if and only if the direction lies in the lognormal cone at the point. This means that $(r,s,t) = \lambda(1,0,1-x) + \mu(1,1,1)$ for some $\lambda, \mu \ge 0$. Solving this system gives the unique contributing point

$$\left(\frac{r-t}{r-s}, \frac{r-s}{t-s}, \frac{t-s}{r-t}\right)$$

for any such direction, when $r > t > s > 0$. Each such point is minimal by Corollary 6.39. Theorem 10.38 applies and yields the dominant asymptotic behavior. We obtain the first-order approximation for the numerator

$$b_{rst} \sim \frac{(r-s)^{r-s+\frac{1}{2}}}{\sqrt{2\pi}(r-t)^{r-t+\frac{1}{2}}(t-s)^{t-s+\frac{1}{2}}}.$$

The relative error is less than 2% even for $(r,s,t) = (20,5,10)$. Note that the approximation is uniform in the union of the cones spanned by $(1,1,1)$ and $(1,0,1-x)$ as x ranges over any compact subset of $(0,1)$. The cones cover the entire range of interest $r > t > s > 0$. ◁

Example 12.30 (Infinite products: Quivers and Littlewood–Richardson coefficients)**.** Consider the infinite product generating function

$$F(x,y) = \prod_{i=1}^{\infty} \left(1 - x^i - y^i\right)^{-1}$$

https://doi.org/10.1017/9781108874144.017 Published online by Cambridge University Press

arising both in the study of chiral operators in four-dimensional quiver gauge theories [RWZ20] and in the enumeration of *Littlewood–Richardson coefficients* $c^{\lambda}_{\mu\nu}$, since

$$F(x, y) = \sum_{\lambda,\mu,\nu} \left(c^{\lambda}_{\mu\nu}\right)^2 x^{|\mu|} y^{|\nu|}$$

(see Harris and Willenbring [HW14] for a derivation and more on Littlewood–Richardson coefficients). Although F is meromorphic and not rational, its analysis is surprisingly simple. Write

$$F(x, y) = \frac{P(x, y)}{1 - x - y},$$

where $P(x, y) = \prod_{i=2}^{\infty} \left(1 - x^i - y^i\right)^{-1}$. It is easy to compute that for a direction (r, s) with positive coordinates the unique critical point on $\mathcal{V}_{1-x-y}$ occurs at $\boldsymbol{w} = \left(\frac{r}{r+s}, \frac{s}{r+s}\right)$. Because the coordinates of $\boldsymbol{w}$ lie in $(0, 1)$, the polynomial $1 - x^i - y^i$ does not vanish at $(x, y) = \boldsymbol{w}$ when $i > 1$ and $\boldsymbol{w}$ is a smooth strictly minimal singularity of F where P does not vanish. Theorem 9.5 and Equation (9.7) thus imply that

$$a_{r,s} \sim \frac{P\left(\frac{r}{r+s}, \frac{s}{r+s}\right)}{\sqrt{2\pi}} \frac{(r + s)^{r+s+\frac{1}{2}}}{r^r s^s \sqrt{rs}}.$$

◁

12.8 Limit laws from probability theory

We now show how to use the techniques of ACSV to derive limit laws. Although probabilistic interpretations and limit laws can hold for series with negative coefficients, to simplify our presentation we restrict to the most common case of a series F with nonnegative coefficients. Suppose the array $\{a_{\boldsymbol{r}}\}$ encodes some combinatorial structure, and that the size of an object is given by a map $\gamma(\boldsymbol{r})$ into the integers – most commonly $\gamma(\boldsymbol{r}) = r_d$ or $\gamma(\boldsymbol{r}) = |\boldsymbol{r}|$.

To discuss typical behavior, we define a ***grand measure*** μ on $\mathbb{Z}^d$ as a sum of point mass measures

$$\mu = \sum_{\boldsymbol{r}\in\mathbb{Z}^d} a_{\boldsymbol{r}} \delta_{\boldsymbol{r}},$$

where $\delta_{\boldsymbol{r}}(\boldsymbol{s}) = 1$ if $\boldsymbol{s} = \boldsymbol{r}$ and 0 otherwise, and normalize its slices

$$\mu_k = \frac{\sum_{\gamma(\boldsymbol{r})=k} a_{\boldsymbol{r}} \delta_{\boldsymbol{r}}}{\sum_{\gamma(\boldsymbol{r})=k} a_{\boldsymbol{r}}}$$

https://doi.org/10.1017/9781108874144.017 Published online by Cambridge University Press

on the objects of size k to be probability measures. One can ask for limit laws with varying degrees of subtlety. A *weak law* tells us that μ_k is concentrated on a region of diameter $o(k)$. More precisely, a weak law with limit $\boldsymbol{m}$ holds for the sequence $\{\mu_k\}$ if, for any $\varepsilon > 0$,

$$\mu_k \left\{ \boldsymbol{r} \,:\, \left| \frac{\boldsymbol{r}}{k} - \boldsymbol{m} \right| > \varepsilon \right\} \to 0 \tag{12.20}$$

as $k \to \infty$. More delicate is a *central limit theorem*, which states that the distribution of μ_k in a region of diameter $O(k^{1/2})$ around $k\boldsymbol{m}$ has Gaussian behavior:

$$\mu_k \left\{ \boldsymbol{r} : \frac{\boldsymbol{r} - k\boldsymbol{m}}{k^{1/2}} \in A \right\} \to \Phi(A) \tag{12.21}$$

as $k \to \infty$, where A is any "nice" region and Φ is a multivariate normal distribution.

Exercise 12.1. When a central limit theorem holds, what (if anything) can you conclude from (12.21) about the probabilities $\mu_k\{\lfloor k\boldsymbol{m} \rfloor\}$, where the greatest integer function is applied coordinatewise? What about the probabilities

$$\mu_k \left\{ \boldsymbol{r} : \frac{|\boldsymbol{r} - k\boldsymbol{m}|}{k} > c \right\}$$

for fixed $c > 0$ as $k \to \infty$?

Even better is a *local central limit theorem* (LCLT) which estimates μ_k at individual points. A LCLT tells us that $\mu_k(\boldsymbol{r}) \sim \mathfrak{n}(\boldsymbol{r})$, where $\mathfrak{n}$ is the density of a multivariate normal distribution. In the case where the size parameter is r_d, for example, the normal density looks like

$$\mathfrak{n}(\boldsymbol{r}) = (2\pi\, r_d)^{(1-d)/2} \sqrt{\det N} \exp \left[\frac{-1}{2} (\boldsymbol{r}^\circ - r_d \boldsymbol{v})^T N \,(\boldsymbol{r}^\circ - r_d \boldsymbol{v}) \right], \tag{12.22}$$

where $\boldsymbol{r}^\circ$ denotes $(r_1, \ldots, r_{d-1})$, the vector $\boldsymbol{v}$ is the mean of the limit Gaussian distribution, and the matrix N is the inverse covariance matrix. As we will see, so-called *Ornstein–Zernike* behavior $a_{\boldsymbol{r}} \sim C(\hat{\boldsymbol{r}})|\boldsymbol{r}|^{(1-d)/2} \boldsymbol{z}^{-\boldsymbol{r}}$ leads to Gaussian estimates for individual probabilities if the coordinates of $\boldsymbol{z}$ are nonnegative real numbers. Ornstein–Zernike asymptotics are precisely the conclusion of our asymptotic estimates in the case where there is a single smooth point and the Hessian matrix is nondegenerate. Thus, all we require for a weak law and LCLT is that asymptotics are governed by a single smooth point with nondegenerate Hessian matrix. The remainder of this section is devoted to the statement and proof of a weak law and local central limit theorem, holding under smoothness and nondegeneracy assumptions.

Exercise 12.2. Suppose $a_{\boldsymbol{r}} = 0$ unless $\sum_{j=1}^{d} r_j$ is even.

https://doi.org/10.1017/9781108874144.017 Published online by Cambridge University Press

(a) How would you expect (12.22) to be altered?
(b) How would you expect this to be reflected in the generating function?

Weak laws

Unless otherwise specified, our results in this section apply to the power series expansion of a generating function $F(\boldsymbol{z}) = P(\boldsymbol{z})/Q(\boldsymbol{z})$ with nonnegative coefficients, where Q is a polynomial containing all variables $z_1, \ldots, z_d$. Confining our statements to power series rather than Laurent series simplifies matters by ensuring finite support of the cross-sectional measures μ_k.

Theorem 12.31 (weak law for diagonal slices)**.** *Let $q(z) = Q(z, \ldots, z)$ and let $p > 0$ be the smallest value such that $q(p) = 0$. If $z = p$ is a strictly minimal simple zero of $q(z)$ and $P(p, \ldots, p) \neq 0$ then the sequence $\{\mu_k\}$ of probability measures defined above satisfies a weak law* (12.20) *with limit $\boldsymbol{m}$, where*

$$\boldsymbol{m} = -\nabla_{\log} Q(p, \ldots, p).$$

Proof The domain of convergence of the power series defined by F corresponds to the component B of the complement of $\texttt{amoeba}(Q)^c$ that contains all points of the form $-(N, \ldots, N)$ for all sufficiently large $N > 0$. In fact, because B is convex the set of points $(x, \ldots, x)$ for $x \in \mathbb{R}$ must intersect ∂B (or else B would be all of $\mathbb{R}^d$). Proposition 6.38 then implies that $q(e^x) = Q(e^x, \ldots, e^x) = 0$ has a real solution, so the polynomial q has positive roots and $p > 0$ is well defined.

Due to nonnegativity of the coefficients, every point of ∂B is a zero of $Q \circ \exp$. Write $\boldsymbol{p} = (p, \ldots, p)$ and define $\log \boldsymbol{p}$ coordinate-wise. If $\nabla\, Q(\boldsymbol{p}) = \boldsymbol{0}$ then $q(z)$ would have a zero of order at least two at $z = p$, so our assumptions imply $\nabla\, Q(\boldsymbol{p}) \neq \boldsymbol{0}$, hence $\nabla(Q \circ \exp)(\log \boldsymbol{p}) \neq \boldsymbol{0}$ and the zero set of $Q \circ \exp$ in a real neighborhood of $\log \boldsymbol{p}$ is a smooth hypersurface normal to $\boldsymbol{m}$.

Letting Z denote the real points of $\mathcal{V}$, we claim that the logarithm maps a neighborhood of $\boldsymbol{p}$ in Z into ∂B. Denoting the homeomorphism type of a $(d-1)$-ball with a distinguished interior point (N, x), the intersection I of $\overline{B}$ with a closed ball around $\log \boldsymbol{p}$ is compact and, via projection from the center of a sphere in the interior, has homeomorphism type (N, x). By the implicit function theorem, a neighborhood of $\boldsymbol{p}$ in Z also has type (N, x). Any homeomorphism from one pair of type (N, x) to another covers a neighborhood of the distinguished point. Therefore, the logarithm maps some neighborhood of $\boldsymbol{p}$ in Z into a neighborhood of $\log \boldsymbol{p}$ in ∂B, and we see that Z coincides locally with ∂B.

Thus, for any $\boldsymbol{r}$ not parallel to $\boldsymbol{m}$ the maximum value of $\boldsymbol{r} \cdot \boldsymbol{x}$ over $\boldsymbol{x} \in B$

https://doi.org/10.1017/9781108874144.017 Published online by Cambridge University Press

is strictly greater than $\boldsymbol{r} \cdot \log \boldsymbol{p}$, so $a_{\boldsymbol{r}} = O((p+\varepsilon)^{-|\boldsymbol{r}|})$ for some $\varepsilon > 0$ whose choice is uniform as $\hat{\boldsymbol{r}}$ varies over any neighborhood not containing $\hat{\boldsymbol{m}}$. The generating function $\sum_{k \geq 0} C_k z^k$ of the sequence $C_k = \sum_{\gamma(\boldsymbol{r})=k} a_{\boldsymbol{r}}$ is $F(z, \ldots, z)$, which has a strictly minimal simple pole at $z = p$. Proposition 3.1 then implies that $C_k \sim cp^{-k}$ for some constant c. In particular, $\sum_{\boldsymbol{r} \in \mathcal{R}} a_{\boldsymbol{r}} = o(C_k)$ as $k \to \infty$ where $\mathcal{R}$ consists of all indices whose directions are in a compact set not containing $\hat{\boldsymbol{m}}$, so a weak law holds with limit $\boldsymbol{m}$. □

Example 12.32 (multinomial distribution)**.** Let $c_1, \ldots, c_d > 0$ sum to 1 and let $F(\boldsymbol{z}) = 1/(1 - \mathbf{c} \cdot \boldsymbol{z})$ be the generating function for the multinomial distribution

$$a_{\boldsymbol{r}} = \binom{|\boldsymbol{r}|}{r_1, \ldots, r_d} c_1^{r_1} \cdots c_d^{r_d}$$

with parameters $\mathbf{c}$. The denominator of F is $Q(\boldsymbol{z}) = 1 - \mathbf{c} \cdot \boldsymbol{z}$. It follows that $Q(z, \ldots, z) = 1 - z$ regardless of $\mathbf{c}$ and the hypotheses of Theorem 12.31 are satisfied with $p = 1$. Thus, there is a weak law with limit $\boldsymbol{m} = -\nabla_{\log} Q(p, \ldots, p) = \mathbf{c}$, recovering the well-known weak law for repeated rolls of a die with weights $\mathbf{c}$. ◁

Often combinatorial classes are enumerated by a size parameter that is not the sum $|\boldsymbol{r}|$ of the indices but is just one of the indices, say r_d. A similar weak law holds in this case, except that the probability measures $\{\mu_k\}$ no longer have finite support and an added hypothesis is required. Adding this hypothesis allows us to work with Laurent series, instead of only power series.

Theorem 12.33 (weak law for coordinate slices)**.** *Let $F(\boldsymbol{z}) = P(\boldsymbol{z})/Q(\boldsymbol{z}) = \sum_{\boldsymbol{r}} a_{\boldsymbol{r}} \boldsymbol{z}^{\boldsymbol{r}}$ be a Laurent series with nonnegative coefficients converging on a component B of* $\texttt{amoeba}(Q)^c$. *Suppose there is some $p > 0$ such that $Q(\mathbf{1}, p) = 0$ and $(\mathbf{0}, \log x) \in B$ if and only if $0 < x < p$. Then $C_k = \sum_{r_d = k} a_{\boldsymbol{r}}$ is finite for all k and if $z = p$ is also a strictly minimal simple zero of $q(z) = Q(\mathbf{1}, z)$ and $P(\mathbf{1}, p) \neq 0$ then the sequence $\{\mu_k\}$ of probability measures defined by*

$$\mu_k = \frac{1}{C_k} \sum_{r_d = k} a_{\boldsymbol{r}} \delta_{\boldsymbol{r}}$$

satisfies a weak law with limit

$$\boldsymbol{m} = -\nabla_{\log} Q(\mathbf{1}, p).$$

Proof Arguing as in the proof of Theorem 12.31, we see again that $\exp(\boldsymbol{u}) \in \mathcal{V}$ for every $\boldsymbol{u} \in \partial B$ and that $\nabla(Q \circ \exp)$ is nonvanishing at $(\mathbf{0}, \log p)$. Because $(\mathbf{1}, \log x) \in B$ for $0 < x < p$, we have convergence of the sum

$$F(\mathbf{1}, z_d) = \sum_{r_d = k} a_{\boldsymbol{r}} z_d^{r_d} = \sum_{k \geq 0} C_k z_d^k$$

https://doi.org/10.1017/9781108874144.017 Published online by Cambridge University Press

whenever $0 < z_d < p$, so C_k takes on finite values and its univariate generating function has radius of convergence p. Just as in the proof of Theorem 12.31, our hypothesis that $q(z)$ has a simple strictly minimal zero at $z = p$ implies $C_k \sim cp^{-k}$ for some constant c, and the total weight of μ_k is $o(p^{-k})$ on sets for which $\hat{\boldsymbol{r}}$ is bounded away from $\hat{\boldsymbol{m}}$, since any hyperplane through $(\mathbf{0}, \log p)$ other than the hyperplane normal to the dth coordinate plane intersects the interior of B. $\square$

Example 12.34 (IID sums)**.** Let μ be a probability measure on a finite subset $E \subseteq \mathbb{Z}^{d-1}$. The spacetime generating function F for convolutions of μ is given by

$$F(\boldsymbol{z}) = \sum_{k \geq 0} \sum_{\boldsymbol{r} \in \mathbb{Z}^{d-1}} \mu^{(k)}(\boldsymbol{r}) \boldsymbol{z}^{(\boldsymbol{r},k)} = \frac{1}{1 - z_d \phi_\mu(z_1, \ldots, z_{d-1})},$$

where ϕ_μ is the $(d-1)$-variable generating function for μ. Then $\mathbf{1}$ is a simple pole of Q and is strictly minimal as long as ϕ is aperiodic. Directly, $\nabla_{\log} Q(\mathbf{1}) = (\boldsymbol{m}, 1)$, where $\boldsymbol{m}$ is the mean vector of μ. Theorem 12.33 then recovers the weak law of large numbers for sums of IID samples from μ. $\triangleleft$

Remark. More generally, we may allow μ to be any measure on $\mathbb{Z}^{d-1}$ whose moment generating function is finite everywhere. This takes us out of the theory of amoebas of polynomials. However, all the facts that are required concerning logarithmic domains of convergence still hold. Because we have not developed the theory of analytic amoebas, we do not include a statement or proof of this result. The greatest generality for weak laws via this type of argument is achieved by weakening the hypothesis to finiteness of the moment generating function in a neighborhood of the origin.

Exercise 12.3. Find a weak law for the binomial coefficients $a_{rs} = \binom{r+s}{s}$ with generating function $1/(1-x-y)$ under the size $\gamma(r,s) = r + 2s$. Rather than using Stirling's formula, maximize $m^{-1} \log a_{rs}$ over $r + 2s = m$ using that the exponential growth of a_{rs} is $x^{-r}y^{-s}$ where $(x, y) = (r/(r+s), s/(r+s))$ is the critical point in the direction $\boldsymbol{r} = (r, s)$ computed in Example 9.10.

Central limits

We now derive a local central limit theorem for the profile $\{a_{\boldsymbol{r}} : r_d = k\}$ as $k \to \infty$. Similar limit theorems for profiles such as $\{a_{\boldsymbol{r}} : |\boldsymbol{r}| = k\}$ also hold, but the arguments are similar and we find it simplest to stick to the case where the size parameter is the last coordinate, mirroring classical limit theorems for the spacetime generating function of a stochastic process on $\mathbb{Z}^{d-1}$. We do, however, weaken our hypotheses to allow F to be a (non-rational) meromorphic function in a suitable domain.

Before giving a limit theorem we need a lemma describing Ornstein–Zernike behavior in this setting. We use the notation $\boldsymbol{e}_j$ for the jth elementary unit vector and recall the notation $\mathbf{T}_e(\boldsymbol{x})$ for the set of complex vectors $\boldsymbol{z}$ with $(\log|z_1|, \ldots, \log|z_d|) = \boldsymbol{x}$.

Lemma 12.35. *Let $F(\boldsymbol{z}) = \sum_{\boldsymbol{r}\in\mathbb{Z}^{d-1}\times\mathbb{N}} a_{\boldsymbol{r}} \boldsymbol{z}^{\boldsymbol{r}}$ be a d-variate series with logarithmic domain of convergence of convergence $B \subseteq \mathbb{R}^d$. Suppose that $\overline{B}$ intersects the negative e_d-axis in the ray $(-\infty, t]$ for some real number t, that F is meromorphic on a neighborhood of the torus $\mathbf{T}_e(\mathbf{0}, t)$ with $F(\boldsymbol{z}) = P(\boldsymbol{z})/Q(\boldsymbol{z})$ for analytic functions P and Q in a neighborhood of $\boldsymbol{w} = (\mathbf{1}, e^t)$, and that $\boldsymbol{w}$ is a strictly minimal pole of F where $Q_{z_d}(\boldsymbol{w}) \neq 0$. Then the logarithmic pole variety* $\log \mathcal{V}$ *of F is a smooth complex analytic hypersurface in a neighborhood of* $\log \boldsymbol{w} = (\mathbf{0}, t)$.

Let $\boldsymbol{m}$ denote the vector $(\nabla_{\log} Q)(\mathbf{1}, e^t)$ scaled so that $m_d = 1$, and let g be the function parametrizing $(\boldsymbol{x}^\circ, x_d) \in \log \mathcal{V}$ by $x_d = g(\boldsymbol{x}^\circ)$ near $\log \boldsymbol{w}$. *If the Hessian matrix $\mathcal{H}$ for g is nonsingular at the origin then as $\boldsymbol{r}$ varies over a neighborhood of $\boldsymbol{m}$ in $S = \mathbb{R}^{d-1} \times \{1\}$ the point $\boldsymbol{w}(\boldsymbol{r}) \in \mathcal{V}$ near $\boldsymbol{w}$ with $(\nabla_{\log} Q)(\boldsymbol{w}(\boldsymbol{r})) = \boldsymbol{r}$ varies smoothly. In this case there is an* Ornstein–Zernike estimate

$$a_{\boldsymbol{r}} \sim \frac{(2\pi\,|r_d|)^{(1-d)/2}}{\sqrt{\operatorname{sgn}(r_d)\det\mathcal{H}}} \cdot \frac{-\operatorname{sgn}(r_d) P(\boldsymbol{w}(\boldsymbol{r}))}{z_d Q_{z_d}(\boldsymbol{w}(\boldsymbol{r}))} \cdot \exp(-\boldsymbol{r}\cdot\boldsymbol{w}(\boldsymbol{r}))\,. \tag{12.23}$$

Proof The first conclusion, that $\log \mathcal{V}$ is a smooth complex analytic hypersurface near $\log \boldsymbol{w}$, follows from the implicit function theorem, since $\nabla(Q \circ \exp)$ is nonvanishing at $\log \boldsymbol{w}$ under our assumptions. Proposition 9.44 implies that nonsingularity of the Jacobian of the Gauss map is equivalent to nondegeneracy of the critical point $\boldsymbol{w}$, and to nonvanishing of the Gaussian curvature of $\mathcal{V}$. Thus, the inverse of the map $\boldsymbol{z} \mapsto (\nabla_{\log} Q)(\boldsymbol{z})$ is smooth near $\boldsymbol{w}$ and $\boldsymbol{w}(\boldsymbol{r})$ varies smoothly with $\boldsymbol{r}$. Theorem 9.4 and Remark 9.19 from Chapter 9 then yield (12.23). □

Theorem 12.36 (LCLT)**.** *Let F be a d-variate generating function satisfying the hypotheses of Lemma 12.35 and let $M = \mathcal{H}(\mathbf{0})$ denote the Hessian matrix of g at the origin. If M is nonsingular then there is a constant c such that*

$$e^{tk}\, a_{\boldsymbol{r}^\circ,k} \sim c\ \mathfrak{n}_k(\boldsymbol{r}^\circ)$$

as $k \to \infty$ with $|\boldsymbol{r}^\circ - k\boldsymbol{m}| = o(k^{2/3})$, where

$$\mathfrak{n}_k(\boldsymbol{r}^\circ) = \frac{(2\pi k)^{(1-d)/2}}{\sqrt{\det M}} \exp\left[-\frac{1}{2k}(\boldsymbol{r}^\circ - k\boldsymbol{m})^T M^{-1}(\boldsymbol{r}^\circ - k\boldsymbol{m})\right] \tag{12.24}$$

denotes the $(d-1)$*-variate* ***normal density*** *with mean* $k\boldsymbol{m}$ *and covariance* M. *It follows that*

$$\sup_{\boldsymbol{r}:r_d=k} k^{(d-1)/2}\left|e^{tk}\,a_{\boldsymbol{r}} - c\,\mathfrak{n}(\boldsymbol{r}^\circ)\right| \to 0$$

as $k \to \infty$.

Remark. We can compute the constant c by comparing (12.23) and (12.24), although a LCLT traditionally does not require knowledge of the normalizing constant.

Proof Let $\boldsymbol{x}(\boldsymbol{r}) = \log \boldsymbol{w}(\boldsymbol{r})$. Comparing (12.24) to (12.23) in Lemma 12.35, it is sufficient to prove that the rate function $\beta(\boldsymbol{r}) = -\boldsymbol{r}\cdot\boldsymbol{x}(\boldsymbol{r})$ satisfies

$$\beta(k\boldsymbol{m}+\boldsymbol{y}) = -\frac{1}{2k}\boldsymbol{y}^T M^{-1}\boldsymbol{y} + C_k + o(1)$$

as $k \to \infty$ with $|\boldsymbol{y}| = o(k^{2/3})$. The rate function is homogeneous of degree one, so this is the same as

$$\beta(\boldsymbol{m}+\boldsymbol{y}) = -\frac{1}{2}\boldsymbol{y}^T M^{-1}\boldsymbol{y} + C'_k + o(k^{-1}), \tag{12.25}$$

where we have scaled $\boldsymbol{y}$ by $1/k$ so it is now restricted to be $o(k^{-1/3})$.

The hyperplane with normal $\boldsymbol{r}$ going through $\boldsymbol{x}(\boldsymbol{r})$ is a support hyperplane to B, so $\boldsymbol{x}(\boldsymbol{r})$ is a minimizing point for $-\boldsymbol{r}\cdot\boldsymbol{x}$ on $\overline{B}$. When $\boldsymbol{r}\in S$ we may write $\boldsymbol{r} = (\boldsymbol{r}^\circ, 1)$, and we write $\boldsymbol{x} = (\boldsymbol{x}^\circ, g(\boldsymbol{x}^\circ))$ for points $\boldsymbol{x}$ in a neighborhood of $\boldsymbol{x}(\boldsymbol{m})$ in $\log \mathcal{V}\cap\partial B$. The function g is concave, because locally the logarithmic domain of convergence $\overline{B}$ is described by $\{(\boldsymbol{x}, u) : u \le g(\boldsymbol{x})\}$ and logarithmic domains of convergence are convex. Thus,

$$\beta(\boldsymbol{r}^\circ, 1) = \inf_{\boldsymbol{x}^\circ\in\mathbb{R}^{d-1}}\{-g(\boldsymbol{x}^\circ) - \boldsymbol{r}^\circ\cdot\boldsymbol{x}^\circ\}$$

which is the negative of the convex dual of the convex function $-g$. As discussed in Chapter 6, the convex dual of the quadratic form $\boldsymbol{x}\mapsto\boldsymbol{x}^T A\boldsymbol{x}$ represented by a positive definite matrix A is represented in the dual basis by the inverse matrix $\boldsymbol{r}\mapsto\boldsymbol{r}^T A^{-1}\boldsymbol{r}$. The quadratic Taylor expansion of the convex dual at a point $\boldsymbol{r}$ is determined by the quadratic Taylor expansion of the function (assuming this is nondegenerate) at the point where the minimum occurs.

The minimizing point for $\boldsymbol{r} = \boldsymbol{m}$ is at the origin and the quadratic term in the expansion of $-g$ at the origin is the matrix $-M$ representing the Hessian of $-g$ at the origin. Therefore the Taylor expansion of β about $\boldsymbol{m}$ on S is given by

$$\beta(\boldsymbol{m}+\boldsymbol{y}) = \beta(\boldsymbol{m}) - \frac{1}{2}\boldsymbol{y}^T M^{-1}\boldsymbol{y} + O(|\boldsymbol{y}|^3)\,.$$

The condition $|\boldsymbol{y}| = o(k^{-1/3})$ is exactly what is needed for $O(|\boldsymbol{y}|^3)$ to be $o(k^{-1})$

https://doi.org/10.1017/9781108874144.017 Published online by Cambridge University Press

and taking C'_k to be $\beta(\boldsymbol{m})$ (not depending on k after all) establishes (12.25) and the first conclusion.

Pick ν with $1/2 < \nu < 2/3$. When $|\boldsymbol{r} - k\boldsymbol{m}| \le k^\nu$, the first conclusion implies that

$$|e^{tk} a_{\boldsymbol{r}} - c\,\mathfrak{n}(\boldsymbol{r})| = o(\mathfrak{n}(\boldsymbol{r})) = o(k^{(1-d)/2}).$$

It remains to establish the second conclusion when $|\boldsymbol{r} - k\boldsymbol{m}| \ge k^\nu$, which we do by showing that both terms being compared are small separately. The term $\mathfrak{n}(\boldsymbol{r})$ is in fact bounded above by $\exp(-ck^{\nu-1/2})$ for some $c > 0$. On the other hand, when $\boldsymbol{r} \in S$ the quantity $\boldsymbol{r} \cdot \boldsymbol{x}(\boldsymbol{r})$ differs from its value at $\boldsymbol{r} = \boldsymbol{m}$ by at least a constant multiple of $|\boldsymbol{r} - \boldsymbol{m}|^2$. In general, when $r_d = k$ the value of $\boldsymbol{r} \cdot \boldsymbol{x}(\boldsymbol{r})$ differs from its value at $\boldsymbol{r} = k\boldsymbol{m}$ by at least a constant multiple of $k^{-1}|\boldsymbol{r} - k\boldsymbol{m}|^2$. When $|\boldsymbol{r} - k\boldsymbol{m}| \ge k^\nu$ this is of order at least $k^{2\nu-1}$, which is a positive power of k. Plugging this into (12.23) shows that $a_{\boldsymbol{r}}$ is also at most $\exp(-ck^{2\nu-1})$, completing the proof. □

Classical LCLT

We end by deriving the classical LCLT for sums of independent lattice random variables whose moment generating functions are everywhere finite using Theorem 12.36. Let μ be an aperiodic probability distribution on $\mathbb{Z}^d$, let $\mu^{(k)}$ denote the k-fold convolution of μ, and let

$$F(\boldsymbol{z}, z_{d+1}) = \sum_{(\boldsymbol{r},k)\in\mathbb{Z}^d\times\mathbb{N}} \mu^{(k)}(\boldsymbol{r})\boldsymbol{z}^{\boldsymbol{r}} z_{d+1}^k$$

denote the spacetime generating function for the random walk with increments distributed as μ. The d-dimensional probability generating function ϕ for μ is defined by

$$\phi(\boldsymbol{z}) = \sum_{\boldsymbol{r}\in\mathbb{Z}^d} \mu(\boldsymbol{r})\boldsymbol{z}^{\boldsymbol{r}}$$

and finiteness of the moment generating function implies ϕ is an entire function. The spacetime generating function F is related to the moment generating function ϕ for the distribution μ via

$$F(\boldsymbol{z}) = \frac{1}{1 - z_{d+1}\,\phi(\boldsymbol{z})}.$$

The singular variety of F is globally defined by $z_{d+1} = 1/\phi(\boldsymbol{z})$ so the logarithmic singular variety is the graph of the function

$$g(\boldsymbol{x}) = -\log\phi(\exp(\boldsymbol{x})).$$

https://doi.org/10.1017/9781108874144.017 Published online by Cambridge University Press

Nonnegativity of series coefficients implies that $(\boldsymbol{x}, t)$ is in the interior of the domain of convergence when $t < g(\boldsymbol{x})$, while no point $(\boldsymbol{x}, g(\boldsymbol{x}))$ lies in the logarithmic domain because it is on $\log \mathcal{V}$. When $\boldsymbol{z}$ is on the torus $\mathbf{T}_e(\boldsymbol{x}, g(\boldsymbol{x}))$, aperiodicity implies that $\phi(\boldsymbol{z}) \neq 0$ unless $\boldsymbol{z}$ is real. Therefore, each point of $\mathcal{V}_{\mathbb{R}} = \mathcal{V} \cap \mathbb{R}^d$ is the only point of $\mathcal{V}$ on its torus, and such a minimal point is strictly minimal.

As μ is a probability measure, $g(\mathbf{0}) = \log(1) = 0$. The chain rule further implies

$$\frac{\partial}{\partial x_j} g(\mathbf{0}) = -\frac{\partial}{\partial x_j}(\log \circ \phi \circ \exp)(\mathbf{0}) = -\frac{e^{\mathbf{0}} \phi_{z_j}(\exp(\mathbf{0}))}{\phi(\exp(\mathbf{0}))}$$
$$= -\phi_{z_j}(\mathbf{1}),$$

and this partial derivative evaluates to $\sum_{\boldsymbol{r} \in \mathbb{Z}^d} r_j \mu(\boldsymbol{r})$, so

$$\boldsymbol{m} = -\nabla(\phi \circ \exp)(\mathbf{0}) = \sum_{\boldsymbol{r} \in \mathbb{Z}^d} \boldsymbol{r} \mu(\boldsymbol{r})$$

is the mean of the distribution μ. Differentiating again, if $i \neq j$ we find that

$$\frac{\partial}{\partial x_i x_j} g(\mathbf{0}) = -\left[\frac{e^{x_i} e^{x_j} \phi_{z_i z_j} \circ \exp}{\phi \circ \exp} - \frac{e^{x_i} e^{x_j} (\phi_{z_i} \circ \exp)(\phi_{z_j} \circ \exp)}{(\phi \circ \exp)^2}\right](\mathbf{0})$$
$$= \phi_{z_i z_j}(\mathbf{1}) - m_i m_j$$

so that the (i, j) entry of the Hessian of g at the origin is indeed the covariance of the i and j coordinates under μ. A similar computation works for $i = j$ and establishes that the Hessian matrix of g at the origin is the covariance matrix for μ. Applying Theorem 12.36, we see that $\mu^{(k)}(\boldsymbol{r})$ is asymptotically equal to $c\, \mathfrak{n}_k(\boldsymbol{r})$. There is no need to compute c because we know $\sum_{\boldsymbol{r}^\circ} \mu^{(k)}(\boldsymbol{r}^\circ) = 1$. Thus $c = 1$ and we recover the classical LCLT.

Theorem 12.37. *If μ is an irreducible aperiodic probability measure on $\mathbb{Z}^n$ with moment generating function ϕ everywhere finite, then*

$$\mu^{(k)}(\boldsymbol{r}) \sim \mathfrak{n}_k(\boldsymbol{r})$$

as $k \to \infty$ with $|\boldsymbol{r} - k\boldsymbol{m}| = o(k^{2/3})$, where $\mathfrak{n}_k$ is defined by (12.24) *with $d = n+1$, the vector $\boldsymbol{m}$ equal to the mean of μ, and M equal to the covariance matrix of μ. It follows that*

$$\sup_{\boldsymbol{r} \in \mathbb{Z}^d} k^{d/2} \left|a_{\boldsymbol{r}} - \mathfrak{n}(\boldsymbol{r})\right| \to 0$$

as $k \to \infty$. □

https://doi.org/10.1017/9781108874144.017 Published online by Cambridge University Press

Example 12.38. Recalling nonnegative Riordan arrays from Section 12.2, observe that setting $x = 1$ gives

$$\mu(v; 1) = \frac{v'(1)}{v(1)}$$

and

$$\sigma^2(v; 1) = \frac{v(1)v''(1) - v'(1)^2 + v(1)v'(1)}{v(1)^2} .$$

Thus, when the hypotheses of Theorem 12.33 are satisfied, a WLLN holds with mean $\boldsymbol{m} = \mu(v; 1)$. Here $\mu(v; 1)$ is simply the mean of the renormalized distribution on the nonnegative integers with probability generating function v, and $\sigma^2(v; 1)$ is the variance of the renormalized distribution. The quadratic form in the exponent of (12.24) in Theorem 12.36 is given by $(s-\mu(v; 1)r)^2/(2k\sigma^2(v; 1))$, meaning a local central limit theorem holds with variance $\sigma^2(v; 1)$. ◁

Exercise 12.4. Large deviation theory is used to provide bounds on $\mu^{(k)}\{|\boldsymbol{r} - k\boldsymbol{m}| > Ck\}$, which are exponentially small in k. What bounds on this kind of event follow from Theorem 12.37?

Notes

The material in Sections 12.2 and 12.3 is largely taken from [PW08, Section 4.3], as is the message passing example in Section 12.4. The idea for Exercise 12.6 comes from [Nob10], and Exercise 12.8 comes from [PW02, Example 3.4]. Exercise 12.9 is suggested by a line of work on rook walks, see for example [KZ11]. The results of Exercise 12.10 are generalized in [Wil15] to diagonal asymptotics of products of combinatorial classes. Elementary derivations of some of the limit theorems presented here are given in Melczer [Mel21, Section 5.3.3].

Riordan arrays have been widely studied. In addition to enumerating a great number of combinatorial classes, Riordan arrays also behave in an interesting way under matrix multiplication (note that the condition $v(0) = 0$ implies $a_{nk} = 0$ for $k < n$, and, by triangularity of the infinite array, that multiplication in the Riordan group is well defined). Surveys of the Riordan group and its combinatorial applications may be found in [Spr94; Sha+91].

There are many multivariate generalizations of the Lagrange Inversion Formula, but we know of none that are useful for our purposes. Proposition 12.8 was given in [Wil05]. The asymptotic behavior of univariate QRWs is derived in several papers, of which [CIR03] is perhaps the most complete. Some of our presentation comes from [BP07], and our second QRW example comes from [Bar+10, Section 4.1].

It is an interesting question to pick how many terms to compute in an asymptotic expansion when the goal is to numerically approximate a fixed coefficient. The books [PK01; Par11] give a good introduction, and applications to integrals arising from coefficient extraction are treated in [DH02].

Additional exercises

Exercise 12.5. (general Lagrange inversion) Use the change of variables described in the proof of Proposition 12.7 and the exact differential $d\left[\frac{\psi(y)}{n}\left(\frac{v(y)}{y}\right)^n\right]$ to prove the more general Lagrange inversion formula (12.13).

Exercise 12.6. Compute dominant asymptotics for $\mu(n,n)$, where $\mu(m,n) = \sum_{k=0}^{n}(-1)^k\binom{n}{k}\binom{2m}{k}$. *Hint:* Replace -1 by z, multiply by $x^m y^n$ and sum over k, m, n to obtain a trivariate generating function.

Exercise 12.7. Using the results in Section 12.2, compute asymptotics for the generalized Dyck paths described in Section 2.3.

Exercise 12.8. Let $F(x,y) = 1/(3 - 3x + x^2 - y)$ be the generating function of a generalized Riordan array $\{a_{rs}\}$. Compute asymptotics for directions (r, s) when $r/s > 1$, then do the same thing when $0 < r/s < 1$. What happens when $r = s$?

Exercise 12.9. Let $a_{\boldsymbol{r}}$ count the number of ways in which a chess rook can move from the origin to $\boldsymbol{r}$ by moves that increase one coordinate and do not decrease any other. The methods of Chapter 2 yield the generating function

$$F(\boldsymbol{z}) = \sum_{\boldsymbol{r}\in\mathbb{N}^d} a_{\boldsymbol{r}} \boldsymbol{z}^{\boldsymbol{r}} = \frac{1}{1 - \sum_{i=1}^{d} \frac{z_i}{1-z_i}} = \frac{\prod_{i=1}^{d}(1 - z_i)}{\sum_{j=0}^{d}(-1)^j (j+1) e_j(\boldsymbol{z})},$$

where e_j is the jth elementary symmetric polynomial.

a) Use Theorem 2.32 in Chapter 2 to find the generating function of the main diagonal of F when $d = 2$. What happens when you try this for $d = 3$?
b) Compute the first-order asymptotic approximation to $a_{\boldsymbol{r}}$ for $d = 3$.
c) Use a computer algebra system, or write your own program, to compute $a_{\boldsymbol{r}}$ exactly, for values of d up to 10. Compare with the first-order asymptotic when $\boldsymbol{r} = (100, \ldots, 100)$.

https://doi.org/10.1017/9781108874144.017 Published online by Cambridge University Press

d) Compute the next term in the expansion and determine how much better is the accuracy of the 2-term asymptotic approximation when $d = 3$ and $\boldsymbol{r} = (100, 100, 100)$.

Exercise 12.10. Derive the bivariate generating function $\sum_{r,s\geq 0} a_{rs} x^r y^s$ for the number of ordered pairs of ordered sequences of integers with parts in a fixed set $A \subset \mathbb{N}$, the first summing to r and the second to s, each having the same number of parts. Compute the asymptotics of the coefficients on the main diagonal. Compare your results and methods with those in [BH12].

Exercise 12.11. Prove that embedding an algebraic generating function A into a rational one R using the method of Lemma 2.36 always makes the numerator of R vanish at the contributing points of R.

https://doi.org/10.1017/9781108874144.017 Published online by Cambridge University Press