Chapter 8

Complex Asymptotic Methods

Asymptotic methods play an essential role in the analysis of algorithms, information theory, and other problems of engineering and science. In this chapter, we primary focus on the methods of *precise* asymptotic estimates. Precise analysis of algorithms was launched in 1962 by D.E. Knuth who analyzed in his "Notes on Open Addressing" hashing tables with linear probing. We discuss a variety of (mostly analytic) asymptotic techniques such as the Euler-Maclaurin summation formula, methods of applied mathematics (i.e., WKB method and matched asymptotics), asymptotics of generating functions (polar singularities and algebraic singularities), saddle point method, methods of evaluating finite sums (e.g., Rice's method), and limiting distributions. In the application section, we enumerate words with approximate self-overlaps, discuss minimax redundancy for memoryless sources, and derive the limiting distribution of the depth in a digital search tree.

IN THE ANALYSIS OF ALGORITHMS we often aim at predicting the rates of growth of time or space complexities for large inputs n. Such analyses rarely produce exact expressions that are easy to use or provide insight into the dependence of the algorithm performance on the parameters of the problem. In fact, quite often we end up with a complicated sum, an integral, a recurrence, or even a functional equation. In such situations, one may turn to an asymptotic analysis. It allows to focus on those terms of the solution that contribute most to the answer. The extra insight from the asymptotic analysis may lead to improve the performance of an algorithm since we simply focus on precisely those parts of the algorithm that effect the leading terms of the asymptotic solution.

Asymptotic analysis can be approached from many angles. We have already used in this book "big oh" O notation. Basically, $f(n) = O(g(n))$ if $f(n)/g(n)$ is bounded from above for n sufficiently large. A more precise expression is Θ notation that assures $|f(n)/g(n)|$ is bounded from below and above for sufficiently large n. There are situations, however, when such asymptotics are still too crude and one must turn to *asymptotic expansions*.

Let us consider an example from data compression. Assume that we are to compress a binary string of length n. Knowing that the ratio of the lengths of the compressed string and the original string is $\Theta(1)$ is not too helpful! As a matter of fact, we would like to know how close this ratio is to the entropy of the source (cf. Chapter 6) in order to estimate how far away we are from the optimal compression. Thus, we need at least the first-order asymptotic expansion. But, Jacob Ziv in his *1997 Shannon Lecture* [364] presented compelling arguments for "backing off" to a certain degree from the first-order asymptotic analysis of information systems, in order to predict the behavior of real systems where we always face *finite* (and often small) lengths (of sequences, files, programs, codes, etc.) One way of overcoming these difficulties is to increase accuracy of asymptotic analysis and replace first-order analyses (e.g., a leading term of the average code length) by more complete asymptotic expansions, thereby extending their range of applicability to smaller values while providing more accurate analyses (like constructive error bounds, large deviations, local or central limit laws).

Another example provides a vast area of combinatorial optimization (cf. Section 1.5). Let us consider the problem that we have already discussed in Section 1.4, namely, the shortest common superstring problem. It is an NP-hard problem, but in Section 6.5.2 we proved that with high probability the ratio of the optimal solution to a greedy solution tends to 1 as the size of the problem $n \to \infty$. In this case, it is again important that the ratio is close to 1, not just bounded. How good is this approximation? Well, we then need more terms in the asymptotic expansions.

The analysis of algorithms that aims at precise information about algorithm performance was launched on July 27, 1963 by Donald E. Knuth in his "Notes on Open Addressing". Generally, an asymptotic expansion we have in mind is a series that represents a complicated expression in a certain limiting sense. More precisely, if x_n denotes a quantity of interest with input size n, we may look for a simple explicit function a_n (e.g., $a_n = \log n$ or $a_n = \sqrt{n}$) such that $x_n \sim a_n$ (i.e., $\lim_{n\to\infty} x_n/a_n = 1$) or we may be aiming at a very precise asymptotic expansion such as $x_n = a_n^1 + a_n^2 + \cdots + o(a_n^k)$ where for each $1 \le i \le k$ we require that $a_n^{i+1} = o(a_n^i)$. The "little oh" notation is another example of "less precise"asymptotic information and it means that $a_n^{i+1}/a_n^i \to 0$ as $n \to \infty$. Hereafter, we mostly discuss *analytic methods* of asymptotic analysis. As argued by Odlyzko [262] "analytic methods are extremely powerful and when they apply, they often yield estimates of unparalleled precision".

This chapter begins with definitions of asymptotic notation. We used some of them already in this book, but we here extend them to complex functions and introduce more rigor. Following that, we present a few tools used in asymptotic analysis. We start with elementary methods such as the Euler-Maclaurin summation formula, methods of applied mathematics (e.g., WKB and matched asymptotics) and sequences distributed modulo 1. We illustrate every method by an example taken from analysis of algorithms or information theory. Then comes our main topic, namely, how to extract asymptotic information from generating functions that are either known exactly or approximately. We first deal with the so–called small singularities of analytic functions, that is, either polar singularities or algebraic singularities. Then we turn our attention to large singularities and discuss saddle point methods. We also discuss certain finite sums (e.g., alternating sums) that occur in analysis of algorithms and are not easy to evaluate by direct methods. We represent these sums by complex integrals and use the tools of this chapter to compute them (we have already seen this approach in Section 7.6.3). Finally, we study limiting distributions through generating functions and analytic tools described in this chapter. In the application section, we discuss how to enumerate words that approximately overlap, deal with interesting problems of code redundancy, and finally derive the limiting distribution for the depth in a digital search tree.

The standard reference in asymptotic analysis is the book of De Bruijn [67]. The reader is also referred to an excellent recent survey by Odlyzko [262] on asymptotic methods. Detailed coverage of many topics consider here may be found in [136, 138, 160, 218, 219, 220, 302, 263, 355]. The reader will find many interesting methods and examples of a combinatorial nature in the forthcoming book by Flajolet and Sedgewick [118]. The book by Greene and Knuth [138] is a joy to read and highly recommended.

8.1 Introduction to Asymptotic Expansions

In this section, we introduce formally the Poincaré notion of an asymptotic expansion. This concept is quite important since it enables us to manipulate a large class of divergent series – often arising in the analysis of algorithms – in much the same way as convergent power series. Even more importantly, in the spirit of Ziv's remark mentioned in the introduction, asymptotic expansions enable to obtain numerical and quantitative results of increasing accuracy.

8.1.1 Definitions and Notations

Let $\mathcal{S}$ be a point set in the complex plane and z_0 a limit point of $\mathcal{S}$. For two complex functions $f(z)$ and $g(z)$ defined on $\mathcal{S}$ we write

$$f(z) = O(g(z)), \qquad z \to z_0 \tag{8.1}$$

to mean that there exists a constant K and a neighborhood $U(z_0)$ of z_0 such that $|f(z)| \leq K|g(z)|$ for all $z \in \mathcal{S} \cap U(z_0)$. If $|f(z)| \geq K'|g(z)|$ for a constant $K' > 0$, then we have $f(z) = \Omega(g(z))$. Furthermore, if $f(z) = O(g(z))$ and $f(z) = \Omega(g(z))$, then $f(z) = \Theta(g(z))$. We shall also use the "little oh"' notation

$$f(z) = o(g(z)), \qquad z \to z_0 \tag{8.2}$$

if for every $\varepsilon > 0$ there is neighborhood $U_\varepsilon(z_0)$ of z_0 such that $|f(z)| \leq \varepsilon|g(z)|$ for all $z \in \mathcal{S} \cap U_\varepsilon(z_0)$. Finally, if $f(z)/g(z) \to 1$ for $z \to z_0$, then

$$f(z) \sim g(z), \qquad z \to z_0. \tag{8.3}$$

Observe that (8.3) can be also written as $f(z) = (1 + o(1))g(z)$.

We mostly have either $z_0 = 0$ or $z_0 = \infty$ and restrict $\mathcal{S}$ either to a sector

$$\mathcal{S}_{\alpha,\beta} = \{z : \ 0 < |z| < \infty, \ \ \alpha < \arg(z) < \beta\},$$

in the complex plane, to the set of nonnegative integers $\mathcal{S} = \mathbb{N}$ or to the set all positive reals $\mathcal{S} = \mathbb{R}^+$. We write $\mathcal{S}_\theta = \mathcal{S}_{-\theta,\theta}$ for a symmetric sector in the complex plane. Finally, we say that the relations O or o are to hold *uniformly* in a set of parameters λ, if the constant K in the big Oh does not depend on λ in this set.

Here are some examples:

$$\begin{aligned}
T(z) - 1 + \sqrt{2(1-ez)} + \frac{2}{3}(1-ez) &= O((1-ez)^{3/2}), \quad z_0 = e^{-1},\ \mathcal{S} = \mathcal{S}_{\pi/2}, \\
\log(1 + a\sqrt{x} + bx + cx^{3/2}) - a\sqrt{x} - (b - \frac{1}{2}a^2)x &= O(x^{3/2}), \qquad z_0 = 0, \quad \mathcal{S} = \mathbb{R}^+, \\
n! &\sim n^{n+1/2}e^{-n}\sqrt{2\pi}, \quad z_0 = \infty, \quad \mathcal{S} = \mathbb{N},
\end{aligned}$$

where $T(z)$ is the tree function defined in Chapter 7 (e.g., see Example 15).

In many applied problems one must deal with divergent series that nevertheless give an excellent numerical approximation. To handle it, Poincaré introduced in 1886 asymptotic power series. For $z \in \mathcal{S}_\theta$ when $z_0 = \infty$ Poincaré defined: *A power series $\sum_{n=0}^{\infty} a_n z^{-n}$ is called an asymptotic expansion of $f(z)$ defined on $\mathcal{S}_\theta$ if for every fixed integer $N \geq 0$*

$$f(z) = \sum_{n=0}^{N} a_n z^{-n} + O(z^{-(N+1)}), \qquad z \to \infty. \tag{8.4}$$

If (8.4) holds, we write

$$f(z) \sim \sum_{n=0}^{\infty} a_n z^{-n}, \qquad z \to \infty.$$

Poincaré's asymptotic expansions are too restrictive for our applications. For example, the average height in PATRICIA trie built from n independent strings generated by unbiased memoryless source is $\log_2 n + \sqrt{2\log_2 n} + \cdots$, while the average height in a digital search tree is $\log_2 n + \sqrt{2\log_2 n} - \log_2\sqrt{2\log_2 n} + \cdots$ as $n \to \infty$ (cf. [215, 216]). These are legitimate asymptotic expressions, but they do not fall under the Poincaré definition. To encompass them, we definite below *generalized* Poincaré-type asymptotic expansions.

A sequence of functions $\phi_k(z)$ is said to be an *asymptotic sequence* as $z \to z_0$ in $\mathcal{S}$ if

$$\phi_{k+1}(z) = o(\phi_k(z)) \tag{8.5}$$

for all k. The original Poincaré definition is recovered if one sets $\phi_k(z) = 1/z^k$ and $z_0 = \infty$. The function $f(z)$ is said to have an *asymptotic expansion* of order N with respect to the asymptotic sequence $\{\phi_k(z)\}$ if there exist constants a_k such that

$$f(z) = \sum_{k=0}^{N} a_k\phi_k(z) + o(\phi_N(z)), \quad z \to z_0. \tag{8.6}$$

If the above holds for all N, then we write

$$f(z) \sim \sum_{k=0}^{\infty} a_k\phi_k(z), \qquad z \to z_0.$$

Observe that the above expansion is unique since one can determine the constants a_k by means of

$$a_k = \lim_{z \to z_0} \left(f(z) - \sum_{i=0}^{k-1} a_i \phi_i(z) \right) / \phi_k(z).$$

If the constants $a_k = 0$ for all k, then such an asymptotic expansion is called *asymptotically zero* with respect to a give asymptotic series $\{\phi_k(z)\}$. For example,

$$e^{-x} \sim 0 \cdot 1 + 0 \cdot x^{-1} + 0 \cdot x^{-2} + \cdots, \qquad x \to \infty$$

is asymptotically (exponentially) zero as $x \to \infty$ in $\mathcal{S} = \mathbb{R}^+$. If $f(z)$ is asymptotically zero, we shall write $f(z) \sim 0$. It is usual to replace such functions by zero in any stage of a proof establishing the validity of Poincaré asymptotic expansion. We shall follow this approach in this book.

Here are some examples of asymptotic expansions as $x \to \infty$ in $\mathcal{S} = \mathbb{R}^+$

$$\begin{aligned} \frac{1}{1+x} &\sim \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{x^k}, \\ \int_0^{\infty} (t+x)^{-1} e^{-t} dt &\sim \sum_{k=1}^{\infty} \frac{(-1)^{k-1}(k-1)!}{x^k}, \\ x^{-\log x} &\sim 0. \end{aligned}$$

It is worth pointing out that the series on the right in the second example above is *not* convergent, but we get better approximations (for larger value of x) to the integral by taking more terms of the asymptotic series.

Asymptotic expansions can be manipulated as convergent series, however, some caution must be exercised with differentiation. In Exercises 1-2 we ask the reader to prove some of these facts.

We also should point out that asymptotic expansions of a complex function f in a sector $\mathcal{S}_\theta$ may crucially depend on θ. For example, $\exp(-1/z) \sim 0$ in any sector contained in $\mathcal{S}_{\pi/2}$, but for purely imaginary $z = ix$ we have $|\exp(-1/z)| = 1$, so the asymptotics brake down.

8.1.2 From Taylor Expansion to Asymptotic Series

We end this section with a useful result that allows us to compute asymptotic series of a complex function $f(z)$ as $z \to \infty$ from its Taylor series at $z \to 0$. We shall follow Ford [128].

Let $f(z)$ be defined by the power series

$$f(z) = \sum_{n=0}^{\infty} a_n z^n, \qquad z \to 0 \tag{8.7}$$

with a positive radius of convergence. We aim at finding conditions under which there exist coefficients b_n such that

$$f(z) \sim \sum_{n=1}^{\infty} \frac{b_n}{z^n}, \qquad z \to \infty \tag{8.8}$$

in a sector of the complex plane. Throughout, we assume that

(A) *there exists an analytic function* $a(w)$ *in the right half plane* $\Re(w) > 0$ *such that*

$$a(w)|_{w=n} = a_n \tag{8.9}$$

where $w = x + iy$.

As the first step, we represent the function $f(z)$ as a complex integral. We have already seen this trick in Section 7.6.3. This method is surprisingly successful while dealing with finite sums and infinite series. The thrust of this approach is to evaluate the integral through the Cauchy residue theorem (cf. Section 2.3). Under certain conditions such an integral can be evaluated inside and outside the contour of integration leading to two different representations. We shall see how it works in the problem at hand (cf. Section 8.5 for a more detailed discussion).

We start with a simple observation: If $P(w)$ is analytic around a given point a and $Q(w)$ has a simple pole at a, then by Cauchy's theorem

$$\frac{1}{2\pi i} \int_{C(a)} \frac{P(w)}{Q(w)} dw = \frac{P(a)}{Q'(a)} \tag{8.10}$$

where the integration takes place around a sufficiently small closed contour $C(a)$ encircling the point $w = a$ in the positive direction. Let us now set

$$P(w) = \pi a(w)(-z)^w, \qquad Q(w) = \sin \pi w,$$

where $(-z)^w$ is understood as a single-valued function in the sector $0 < \arg(z) < 2\pi$ (which excludes the positive half of the real axis), defined as follows

$$(-z)^w = \exp\left(w \log(-z)\right) = \exp\left(w(\log|z| + i(\arg(z) - \pi))\right), \qquad 0 < \arg(z) < 2\pi. \tag{8.11}$$

We write $\overline{\mathcal{S}}_\theta = \{z : \ 0 < |z| < \infty, \ \theta < \arg(z) < 2\pi - \theta\}$ for some $0 < \theta < \pi/2$, that is, $\overline{\mathcal{S}}_\theta$ is a set in the complex plane that excludes the sector $\mathcal{S}_\theta$ around the positive half of the real axis (e.g., for $\theta = 0$ we exclude only the real positive half axis).

Let $\mathcal{C}_{n,l}$ be a contour show in Figure 8.1 that surrounds each of the following points $w = -l, -l+1, \ldots, -1, 0, 1, \ldots, 2n-1, 2n$, where l is a positive integer. Since for any integer $k \in \mathbb{Z}$

$$\frac{\pi}{\sin \pi w} = \frac{(-1)^k}{(w-k)} + O(1)$$

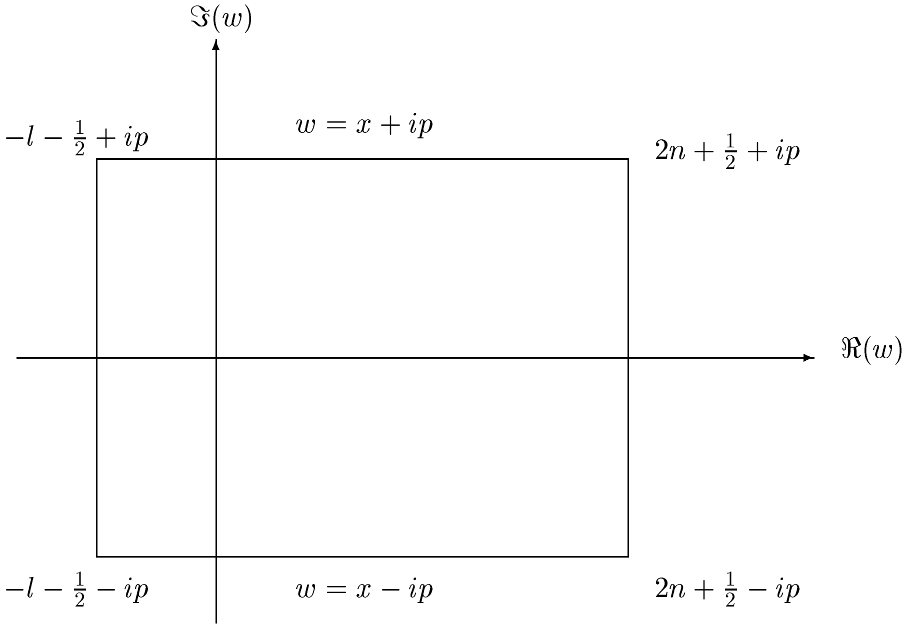


Figure 8.1: The integration contour $\mathcal{C}_{n,l}$.

the Cauchy residue theorem[1], as in (8.10), leads to

$$\frac{1}{2\pi i}\int_{\mathcal{C}_{n,l}}\frac{\pi a(w)(-z)^w}{\sin\pi w}dw=\sum_{m=1}^{l}\frac{a(-m)}{z^m}+\sum_{k=0}^{2n}a(k)z^k. \tag{8.12}$$

We now evaluate the integral in (8.12) along the four line segments of the curve $\mathcal{C}_{n,l}$ shown in Figure 8.1, namely, the lines $w=-l-\frac{1}{2}+iy$, $w=2n+\frac{1}{2}+iy$ and $w=x\pm ip$ where $-\infty<x,y<\infty$ and we allow p to be positive and large. We denote by D_p and D_{-p} the contributions of the integral along the upper and lower horizontal lines, respectively. For D_p we observe that for $w=x+ip$

$$\sin\pi w=\sinh\pi p(\sin\pi x\coth\pi p+i\cos\pi x)$$

which yields

$$D_p=\frac{(-z)^{ip}}{2i\sinh\pi p}\int_{2n+\frac{1}{2}}^{-l-\frac{1}{2}}\frac{a(x+ip)(-z)^x}{\sin\pi x\coth\pi p+i\cos\pi x}dx.$$

To assure that $D_p\to 0$ as $p\to\infty$, we adopt another assumption regarding the function $a(w)$, namely:

[1]The reader may want to go back to Section 2.3 to review some facts from complex analysis.

(B) *for sufficiently large y and any positive $\varepsilon > 0$*

$$|a(x+iy)| < Ke^{\varepsilon|y|} \tag{8.13}$$

where K is a constant that may depend only on ε.

In other words, $a(w)$ cannot grow faster than exponentially along any line parallel to the imaginary axis.

Suppose for the present that $z \in \overline{\mathcal{S}}_\theta$ is *real and negative* so that $|-z^{ip}| = 1$. Then

$$|D_p| < \frac{Ke^{\varepsilon p}}{2\sinh \pi p} \int_{-l-\frac{1}{2}}^{2n+\frac{1}{2}} (-z)^x \, dx$$

since

$$|\sin \pi x \coth \pi p + i \cos \pi x|^2 = \sin^2 \pi x \cosh^2 \pi p + \cos^2 \pi x > 1.$$

In summary,

$$\lim_{p\to\infty} D_p = 0.$$

In a similar manner we prove that

$$\lim_{p\to\infty} D_{-p} = 0.$$

Next, let us consider the integral along the vertical right line $w = 2n + \frac{1}{2} + iy$ which we denote as D_n. Since

$$\sin \pi w = \sin(\pi/2 + i\pi y) = \cos i\pi y = \cosh \pi y,$$

we arrive at (imposing $p \to \infty$)

$$D_n = \frac{(-z)^{2n+\frac{1}{2}}}{2} \int_{-\infty}^{\infty} \frac{a(2n + \frac{1}{2} + iy)(-z)^{iy}}{\cosh \pi y} dy.$$

Under hypothesis (B) we deduce that the integral above converges absolutely. *Restricting further z to the real segment* $-1 < z < 0$ we certainly have

$$\lim_{n\to\infty} D_n = 0$$

for $-1 < z < 0$. Soon we extend the validity of these estimates to $\overline{\mathcal{S}}_\theta$ with $\theta = 0$.

Finally, we deal with the left vertical line $w = -l - \frac{1}{2} + iy$ which contribution we denote as D_l. We obtain

$$D_l = (-1)^l \frac{(-z)^{-l-\frac{1}{2}}}{2} \int_{-\infty}^{\infty} \frac{a(-l - \frac{1}{2} + iy)(-z)^{iy}}{\cosh \pi y} dy.$$

Under hypothesis (B) this integral converges absolutely for $-1 < z < 0$.

In view of the above we conclude that (8.12) becomes

$$f(z) = (-1)^l \frac{(-z)^{-l-\frac{1}{2}}}{2} \int_{-\infty}^{\infty} \frac{a(-l-\frac{1}{2}+iy)(-z)^{iy}}{\cosh \pi y} dy - \sum_{m=1}^{l} \frac{a(-m)}{z^m} \tag{8.14}$$

for $n \to \infty$ as long as $-1 < z < 0$.

We now analytically continue the solution (8.14) to the sector $\overline{\mathcal{S}}_\theta$ (excluding real positive z). For this we define a subsector $\overline{\mathcal{S}}_{\theta+\delta} \subset \overline{\mathcal{S}}_\theta$ for any $\delta > 0$ in which we prove that the integral in (8.14) converges. It suffices to show that the integral converges for $y \in (-\infty, -d)$ and $y \in (d, \infty)$ for arbitrarily large d. Let $z = |z|(\cos\phi + i \sin\phi) \in \overline{\mathcal{S}}_{\theta+\delta}$. Then, under hypothesis (B) and (8.11), the above two integrals are bounded by

$$K \int_{-\infty}^{-d} \frac{e^{-(\phi-\pi+\varepsilon)y}}{\cosh \pi y} dy; \qquad K \int_{d}^{\infty} \frac{e^{-(\phi-\pi-\varepsilon)y}}{\cosh \pi y} dy$$

that converge uniformly for $z \in \overline{\mathcal{S}}_{\theta+\delta}$. Thus the integral in (8.14) defines an analytic function and this representation is valid inside $\overline{\mathcal{S}}_\theta$. But the factor $(-z)^{-l-\frac{1}{2}}$ will assure that the last omitted term in (8.14) is $o(z^{-l})$ leading to an asymptotic expansion. More precisely, it remains but to note that (8.14) is of the form

$$f(z) = -\sum_{m=1}^{l} \frac{a(-m)}{z^m} + \frac{\eta(z,l)}{z^l}$$

where $\eta(z,l) = O(1/\sqrt{z}) \to 0$ as $|z| \to \infty$.

In summary, we proved the following theorem

Theorem 8.1 *Let $f(z)$ be analytic and have the following power series representation*

$$f(z) = \sum_{n=0}^{\infty} a_n z^n, \qquad z \to 0$$

with a positive radius of convergence. Assume a_n possesses an analytic continuation $a(w)$ with $w = x + iy$ such that hypotheses (A) and (B) are satisfied, that is,

(A) there exists an analytic function $a(w)$ in the right half plane $\Re(w) > 0$ such that $a(w)|_{w=n} = a_n$

(B) for sufficiently large y and any positive $\varepsilon > 0$

$$|a(x+iy)| < K e^{\varepsilon |y|}$$

where K is a constant that may depend only on ε.

Then the function $f(z)$ defined in any sector $\overline{\mathcal{S}}_\theta$ (for arbitrary small $\theta > 0$) that excludes the positive half of the real axis can be represented by the following asymptotic series

$$f(z) \sim -\sum_{m=1}^{\infty} \frac{a(-m)}{z^m}, \qquad z \to \infty \tag{8.15}$$

in $\overline{\mathcal{S}}_\theta$.

In passing, we point out that if $a(w)$ is not analytic but has some singularities, then one must subtract this contribution from the left-hand side of (8.15).

Example 8.1 *Asymptotic Expansion of a Series*
Let us consider

$$f(z) = \sum_{n=0}^{\infty} \frac{z^n}{n+\omega}$$

where ω is a constant not equal to zero or a negative integer. In this case, we have $a(w) = 1/(w+\omega)$, which clearly satisfies hypotheses (A) and (B) everywhere except $w = -\omega$, where there is a pole of $a(w)$. Proceeding like above, we end up with an equation like (8.12) with an addition term on the right-hand side coming from the residue at $w = -\omega$. This residue depends on whether ω is a positive integer or not. In the former case, we have a double pole at $w = \omega$, in the latter just a simple pole. Let us handle first the latter situation, and we assume that $\omega \notin \mathbb{N}$. Then

$$\operatorname{Res}\left(\frac{\pi(-z)^w}{(w+\omega)\sin \pi w}, w = -\omega\right) = -\frac{\pi(-z)^{-\omega}}{\sin \pi\omega}.$$

Let us now consider the case when $\omega \in \mathbb{N}$, say $\omega = K$. At $w = -K$ we have a double pole, a singularity of $1/\sin \pi K$ and $1/(w+\omega)$. To find the residue of

$$\frac{\pi(-z)^w}{(w+K)\sin \pi w}$$

at $w = -K$, we expand the functions involved in the above expression into Taylor series around $w = -K$. Note that

$$\begin{aligned}
(w+K)\sin \pi w &= (-1)^K \pi (w+K)\left(1 + \frac{(w+K)^2}{3!} + O((w+K)^4)\right), \\
\pi(-z)^w &= \pi(-z)^{-K}\left(1 + (w+K)\log(-z) + O((w+K)^2\right).
\end{aligned}$$

Upon dividing and finding the coefficient at $(w+K)^{-1}$ we arrive at

$$\operatorname{Res}\left(\frac{\pi(-z)^w}{(w+K)\sin \pi w}, w = -K\right) = z^{-K}\log(-z).$$

In summary, we obtain

$$f(z) \sim \frac{\pi(-z)^{-\omega}}{\sin \pi\omega} - \sum_{n=1}^{\infty} \frac{1}{(\omega - n)z^n}, \qquad \omega \notin \mathbb{N}$$

and

$$f(z) \sim -\frac{\log(-z)}{z^K} - \sum_{n \neq K} \frac{1}{(K-n)z^n}, \qquad \omega = K \in \mathbb{N}$$

where the last sum is over all nonnegative integers not equal to K. □

It is to be noted, finally, that if instead of the original series $f(z) = \sum_{n \geq 0} a_n z^n$, we start with

$$f(z) = \sum_{n=0}^{\infty} a_n (-z)^n,$$

where $a(w)$ satisfies hypotheses (A) and (B), and if we replace $\pi a(w)(-z)^w$ by $\pi a(w) z^w$ defined in the sector $\mathcal{S}_\pi$, then the asymptotic expansion

$$f(z) \sim -\sum_{m=1}^{\infty} \frac{a(-m)}{(-z)^m}$$

is valid in any sector $\mathcal{S}_\theta$ that excludes the *negative* half of the real axis. The reader is asked in Exercise 3 to prove the following

$$\begin{aligned} \log(1+x) &= \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} x^k, & x \to 0, \\ &= \log x + \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} x^{-k}, & x \to \infty. \end{aligned}$$

Observe that above we have $=$ (not $\sim$) for both limits.

8.2 Basic Methods

In this section we discuss asymptotic methods that are not based on extracting singularities of generating functions. We start with the Euler-Maclaurin summation formula, followed by methods of applied mathematics such as the WKB method and matched asymptotics. Finally, we deal with sequences distributed modulo 1 in order to estimate certain sums involving fractional parts.

8.2.1 Euler-Maclaurin Summation Formula

From a first course of calculus, we know that there is a relationship between integrals and series. In fact, when the function f is monotone, then $\sum_{1\le k\le n} f(k)$ is well approximated by $\int_1^n f(x)dx$. (The reason that one prefers the integral over the sum lies in fact that such integrals are usually much easier to evaluate.) The Euler-Maclaurin summation formula makes this statement very precise for sufficiently smooth functions. It finds plenty of applications in discrete mathematics, combinatorics and analysis of algorithms.

Let f be a differentiable function defined on the interval $[m,n]$ where $m<n$ are integers. Observe that for any $m\le j<n$

$$\begin{aligned}\frac{f(j)+f(j+1)}{2} &= \int_j^{j+1}\frac{d}{dx}\left[\left((x-j-\frac{1}{2}\right)f(x)\right]dx\\ &= \int_j^{j+1} f(x)dx+\int_j^{j+1}\left(x-j-\frac{1}{2}\right)f'(x)dx \qquad (8.16)\end{aligned}$$

where the last line follows from integration by parts, and the fact that $B_1'(x-j)=\frac{d}{dx}(x-j-\frac{1}{2})=1$ and $B_1(1)=-B_1(0)=\frac{1}{2}$ where $B_1(x)$ is the first Bernoulli polynomial. Definition and properties of Bernoulli numbers and Bernoulli polynomials are reviewed in Table 8.1. Let $\langle x\rangle = x-\lfloor x\rfloor$ be the fractional part of x. If we write $B_1(x-\langle x\rangle)=x-\langle x\rangle-\frac{1}{2}=x-j-\frac{1}{2}$ for $j\le x<j+1$, and sum (8.16) over $j=m,m+1,\dots,n-1$, we immediately obtain

$$\sum_{k=m+1}^{n-1} f(k)+\frac{f(m)+f(n)}{2}=\int_m^n f(x)dx+\int_m^n B_1(x-\langle x\rangle)f'(x)dx. \qquad (8.17)$$

This is a particular case of the Euler-Maclaurin formula.

Example 8.2 *Moments of Nonnegative Random Variables*

Let M_n be a sequence of nonnegative random variables having all moments such that $\mathbf{E}[M_n^k]=o(\mathbf{E}[M_n^{k+1}])$ with respect to n. For fixed k, we calculate the $(k+1)$st moment of M_n as

$$\begin{aligned}\mathbf{E}[M_n^{k+1}] &= \sum_{m=0}^{\infty} m^{k+1}\mathsf{Pr}\{M_n=m\}=\sum_{m=0}^{\infty} m^{k+1}(\mathsf{Pr}\{M_n>m-1\}-\mathsf{Pr}\{M_n>m\})\\ &= \sum_{m=0}^{\infty}((m+1)^{k+1}-m^{k+1})\mathsf{Pr}\{M_n>m\}\\ &= (k+1)\sum_{m\ge 0} m^k\mathsf{Pr}\{M_n>m\}+O(\mathbf{E}[M_n^k])\ ,\end{aligned}$$

where the last line follows from the identity: $a^k-b^k=(a-b)(a^{k-1}+a^{k-2}b+\cdots+b^{k-1})$. Using the Euler-Maclaurin formula to shall show that

$$\sum_{m\ge 0} m^k\mathsf{Pr}\{M_n>m\}=\int_0^\infty x^k\mathsf{Pr}\{\widetilde{M}_n>x\}dx+O(\mathbf{E}[M_n^k]),$$

Table 8.1: Bernoulli Numbers and Polynomials

1. *The Bernoulli numbers* B_m for $m = 0, 1, \ldots$ are defined by the expression

$$\frac{t}{e^t - 1} = \sum_{k=0}^{\infty} \frac{B_k}{k!} t^k, \qquad |t| < 2\pi. \tag{8.18}$$

Since $\frac{t}{e^t-1} = -t + \frac{-t}{e^{-t}-1}$, hence all Bernoulli numbers with odd index $m \geq 3$ are zero, that is, $B_3 = B_5 = \cdots B_{2m+1} = \cdots = 0$.
2. Bernoulli polynomials $B_m(x)$ of order $m = 0, 1, 2, \ldots$ are defined as

$$\frac{te^{xt}}{e^t - 1} = \sum_{k=0}^{\infty} \frac{B_k(x)}{k!} t^k, \qquad |t| < 2\pi. \tag{8.19}$$

In particular, $B_0(x) = 1$, $B_1(x) = x - \frac{1}{2}$ and $B_2(x) = x^2 - x + \frac{1}{6}$.
3. *Properties* of Bernoulli polynomials directly following from the definition:

$$\begin{aligned}
B_m(1-x) &= (-1)^m B_m(x), & (8.20)\\
B_m(0) &= B_m, & (8.21)\\
B'_m(x) &= mB_{m-1}(x), \quad m \geq 1, & (8.22)\\
|B_{2n}(x)| &\leq |B_{2n}|, \quad 0 < x < 1, & (8.23)
\end{aligned}$$

The first property follows from $\frac{te^{xt}}{e^t-1} = \frac{(-t)e^{(1-x)(-t)}}{e^{-t}-1}$ while the other two are obvious. The last is a consequence of the fact that $B_{2n}(x)$ for $x \in (0,1)$ achieves maximum at $x = \frac{1}{2}$ and $B_{2n}(\frac{1}{2}) = (2^{1-2n} - 1)B_{2n}$.
4. *Asymptotics.* Consider now $n \geq 1$ and evaluate B_{2n} by Cauchy's formula

$$\frac{B_{2n}}{(2n)!} = \frac{1}{2\pi i} \oint \frac{dz}{z^{2n}(e^z - 1)}.$$

Computing the residue at $z = \pm 2\pi i k$, $k = 1, 2, \ldots$ we find

$$B_{2n} = 2(-1)^{n-1}\zeta(2n)\frac{(2n)!}{(2\pi)^{2n}} \sim 2(-1)^{n-1}\frac{(2n)!}{(2\pi)^{2n}} \tag{8.24}$$

where the asymptotics follows from $\zeta(n) = 1 + O(2^{-n})$ as $n \to \infty$.
5. *The Euler functions* $P_m(x)$ are defined as periodic functions of period 1 such that

$$\begin{aligned}
P_m(x) &= \frac{B_m(x)}{m!}, \quad 0 \leq x < 1, & (8.25)\\
P_m(0) &= P_m(1) = \frac{B_m}{m!}, \quad m \neq 1. & (8.26)
\end{aligned}$$

By (8.22) we have $P'_{m+1}(x) = P_m(x)$ for $m \geq 1$.

Table 8.2: ABEL'S PARTIAL SUMMATION FORMULA

One often has to deal with sums like $\sum_i a_i b_i$ where a_i, b_i, $1 \le i \le n$ are two real-valued sequences. A basic technique to handle such sums is *Abel's partial summation* formula. Let

$$\begin{aligned} A(k) &= \sum_{j=1}^{k} a_j, \\ \widetilde{A}(k) &= \sum_{j=n-k}^{n} a_j. \end{aligned}$$

Then

$$\sum_{j=1}^{n} a_j b_j = \sum_{k=1}^{n-1} A(k)(b_k - b_{k-1}) + A(n)b_n \tag{8.27}$$

$$= \sum_{i=0}^{n-1} \widetilde{A}(n-i-1)(b_{i+1} - b_i), \tag{8.28}$$

where $b_0 = 0$. The formulas work fine for $n = \infty$ provided the series involved converge. In this case the last term in (8.27) is not necessary. A derivation of Abel's formula is pretty straightforward and left as an exercise.

where, for the illustrative purpose, we assume that $\widetilde{M}_n$ is a continuous approximation to M_n. Indeed, after applying (8.17) with $m = 0$, $n = \infty$ and $f(x) = x^k \mathsf{Pr}\{\widetilde{M}_n > x\}$ we arrive at

$$\sum_{m \ge 0} m^k \mathsf{Pr}\{M_n > m\} = \int_0^\infty x^k \mathsf{Pr}\{\widetilde{M}_n > x\} dx + O\left(\int_0^\infty |f'(x)| dx\right).$$

But

$$\int_0^\infty \left(x^k \frac{d}{dx} \mathsf{Pr}\{\widetilde{M}_n > x\} + kx^{k-1}\mathsf{Pr}\{\widetilde{M}_n > x\}\right) dx = O(\mathbf{E}[M_n^k]).$$

In summary,

$$\mathbf{E}[M_n^{k+1}] = (k+1)\int_0^\infty x^k \mathsf{Pr}\{\widetilde{M}_n > x\} dx + O(\mathbf{E}[M_n^k])$$

for fixed k. □

We now refine (8.17) by having a closer look at the last term, that is,

$$\int_j^{j+1} B_1(x - \langle x \rangle) f'(x) dx$$

for $j = m, \ldots, n-1$. It turns out that instead of working with the Bernoulli polynomials $B_m(x - \langle x \rangle)$ it is more convenient to use Euler's function $P_m(x) = B_m(x - \langle x \rangle)/m!$ (cf. (8.25)-(8.26)) of Table 8.1) since they are periodic with period 1 and such that $P'_{m+1}(x) = P_m(x)$. Integrating by parts the above integral gives

$$\int_j^{j+1} P_1(x - \langle x \rangle) f'(x) dx = P_2(0)(f'(j+1) - f'(j)) - \int_j^{j+1} P_2(x) f''(x) dx$$

since $P_m(0) = P_m(1) = B_m/m!$ for $m > 1$ (cf. (8.26)). Summing up over $j = m, \ldots, n-1$, and plugging into (8.17) we finally obtain

$$\begin{aligned} \sum_{k=m+1}^{n-1} f(k) + \frac{f(m) + f(n)}{2} &= \int_m^n f(x)dx + \frac{B_2}{2!}(f'(n) - f'(m)) \\ &- \int_m^n B_2(x - \langle x \rangle) f''(x) dx. \end{aligned}$$

Using mathematical induction and properties of Bernoulli numbers and polynomials (in particular, (8.23) and (8.24)) we can generalize the above procedure leading to the following theorem.

Theorem 8.2 (Euler-Maclaurin) *Suppose f has continuous derivatives up to order s. Then*

$$\begin{aligned} \sum_{k=m+1}^{n} f(k) &= \int_m^n f(x)dx + \sum_{l=1}^{s} (-1)^l \frac{B_l}{l!} \left(f^{(l-1)}(n) - f^{(l-1)}(m) \right) && (8.29) \\ &+ \frac{(-1)^{s-1}}{s!} \int_m^n B_s(x - \langle x \rangle) f^{(s)}(x) dx && (8.30) \\ &= \int_m^n f(x)dx + \frac{1}{2} f(n) - \frac{1}{2} f(m) \\ &+ \sum_{\ell=1}^{S-1} \frac{B_{2\ell}}{(2\ell)!} \left(f^{(2\ell-1)}(n) - f^{(2\ell-1)}(m) \right) + O\left((2\pi)^{-2S} \right) \int_m^n |f^{(2S)}(x)| dx \end{aligned}$$

where $S = 2s$.

Proof. The first part (8.29) follows directly from our discussion above. To derive the error bound in (8.30) we observe that only even terms in the sum above contribute since $B_{2k+1} = 0$ for $k \geq 1$. We set $S = 2s$ and use (8.22) together with (8.24) to prove (8.30). ■

Example 8.3 *Asymptotic Expansion of the Euler Gamma Function*

Let us compute $\log \Gamma(z)$ for $z \in \mathbb{C} - (-\infty, 0)$. We start with the Euler representation for the gamma function, namely

$$\Gamma(z) = \lim_{n \to \infty} \prod_{k=1}^{n} \frac{k}{z + k - 1} \left(\frac{k+1}{k} \right)^{z-1},$$

which implies

$$\log\Gamma(z) = \lim_{n\to\infty}\left((z-1)\log(n+1) - \sum_{k=1}^{n}\log\frac{z+k-1}{k}\right).$$

We apply the Euler-Maclaurin formula to the sum above, that is, we assume $m = 0$ and $f(x) = \log(x+z-1) - \log x$ in (8.29). This leads to

$$\begin{aligned}\sum_{k=1}^{n}\log\frac{z+k-1}{k} &= \log z + \int_1^n (\log(x+z-1) - \log x)dx \\ &+ \sum_{j=1}^{m}\frac{B_{2j}}{2j(2j-1)}\left(\frac{1}{(n+z-1)^{2j-1}} - \frac{1}{n^{2j-1}} - \frac{1}{z^{2j-1}} + 1\right) \\ &+ \frac{1}{2}(\log(n+z-1) - \log n - \log z) \\ &+ \frac{1}{2m}\int_0^n B_{2m}(x - \langle x\rangle)\left(\frac{1}{(z+x-1)^{2m}} - \frac{1}{x^{2m}}\right)dx.\end{aligned}$$

In the above we used $B_1 = 1/2$ and $B_{2k+1} = 0$ for $k \geq 1$. After computing the first integral and taking the limit $n \to \infty$ we finally arrive at

$$\begin{aligned}\log\Gamma(z) &= \left(z - \frac{1}{2}\right)\log z - z + 1 + \sum_{j=1}^{m}\frac{B_{2j}}{2j(2j-1)}\left(\frac{1}{z^{2j-1}} - 1\right) \\ &- \frac{1}{2m}\int_0^n B_{2m}(x - \langle x\rangle)\left(\frac{1}{(z+x-1)^{2m}} - \frac{1}{x^{2m}}\right)dx.\end{aligned}$$

Further simplification can be obtained for $z \to \infty$ knowing that

$$\lim_{z\to\infty}(\log\Gamma(z) - (z - 1/2)\log z + z) = \frac{1}{2}\log 2\pi.$$

Then, one obtains Stirling's formula

$$\begin{aligned}\log\Gamma(z) &= \left(z - \frac{1}{2}\right)\log z - z + \sum_{j=1}^{m}\frac{B_{2j}}{2j(2j-1)}\frac{1}{z^{2j-1}} \\ &+ \frac{1}{2}\log 2\pi - \frac{1}{2m}\int_0^n \frac{B_{2m}(x - \langle x\rangle)}{(z+x)^{2m}}dx.\end{aligned}$$

for $z \to \infty$ in $\mathbb{C} - (-\infty, 0)$. This leads directly to a useful approximation of $n!$

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n\left(1 + \frac{1}{12n} + \frac{1}{288n^2} - \frac{139}{5140n^3} + \cdots\right) \tag{8.31}$$

as $n \to \infty$. $\square$

8.2.2 Matched Asymptotics and the WKB Method

Numerous problems in mathematics, physics, combinatorics and analysis of algorithms reduce to either functional/differential equations or recurrence equations with a small (large) parameter $\varepsilon > 0$ ($\lambda > 0$). We discuss here two methods of applied mathematics that are routinely used to solve such problems. We have in mind *matched asymptotics* and the *WKB approximation* named after the physicists Wentzel, Kramers and Brillouin (however, Olver in his book [263], rightfully, calls this method the *Liouville-Green approximation* who proposed it already in 1837). We also briefly mention another method, namely, *linearization* that finds many applications in analysis of algorithms. We must add, however, that the WKB method makes certain assumptions about the forms of the asymptotic expansions without a rigorous justification of the assumed forms.

We will not discuss the methods in depth, referring the interested reader to many excellent books such as Bender and Orszag [30], Nayfeh [258], and Olver [263]. We start with a general and brief overview of the two methods. Then, we illustrate them by deriving asymptotic distributions for the heights in tries and b-tries (cf. Section 1.1).

Let $\varepsilon > 0$ be a small number and $\lambda > 0$ a large number. Consider the following two differential equations (arising in physics)

$$\begin{aligned} \varepsilon y''(x) + (1+\varepsilon^2)y'(x) + (1-\varepsilon^2)y(x) &= 0, \qquad y(0)=\alpha, \quad y(1)=\beta, && (8.32)\\ \frac{1}{\lambda^2}y''(x) - q(x)y(x) &= 0, && (8.33) \end{aligned}$$

where α and β are two given constants. The first problem above is a simple boundary-value problem, and can be solved exactly. However, we shall try to solve it approximately to illustrate the idea of *matched asymptotic expansions.* We look for a solution in the form

$$y(x;\varepsilon) = \sum_{j=0}^{\infty} \varepsilon^j y_j(x) \tag{8.34}$$

where $y_j(x)$ are unknown functions. For the second problem above, we shall seek a (nontrivial) asymptotic solution in the following form

$$y(x;\lambda) = e^{\lambda\Phi(x)} \sum_{j=0}^{\infty} y_j(x)\frac{1}{\lambda^j} \tag{8.35}$$

where $\Phi(x)$ and $A_0(x), A_1(x), \ldots$ are unknown functions. The form (8.35) is known as the *WKB approximation.* Here's what Fedoryuk [97] has to say about such approximations: "... *It is necessary first of all to guess (and no other word will do) in what form to search for the asymptotic form. Of course, this stage – guessing the form of the asymptotic form – is not subject to any formalization. Analogy, experiments, numerical simulation, physical* *considerations, intuitions, random guesswork; these are the arsenal of means used by any research worker*".

To illustrate the methods, we discuss a concrete example. Let us first illustrate the *matching principle.* We shall follow Nayfeh [258] and start with the boundary problem (8.32) and assume that (8.34) is its solution. Looking only at the leading term $y_0(x)$ we find, after some simple algebra, that $y_0(x) = c_0 e^{-x}$, where c_0 is a constant that we should determine from the boundary condition. We either get $c_0 = \alpha$ or $c_0 = \beta e$, and usually both cannot be satisfied. Therefore, we must choose one, say the value at $x = 1$, and ignore for now the other boundary value. We then obtain a full asymptotic expansion as

$$y(x) = \beta e^{1-x} + \varepsilon\beta(1-x)e^{1-x} + \cdots \tag{8.36}$$

which works fine everywhere but near the origin where $y(0) = \beta e(1 + \varepsilon + O(\varepsilon^2)) \neq \alpha$. In order to overcome these difficulties, we consider another scale, namely $\xi = x/\varepsilon$. Keeping ξ fixed and solving the equation we obtain

$$\begin{aligned} y(x) &= \beta e(1-x) + (\alpha - \beta e)(1+x)e^{-x/\varepsilon} + \cdots & (8.37)\\ &= \beta e + (\alpha - \beta e)e^{-\xi} + \varepsilon[(\alpha - \beta e)e^{-\xi} - \beta e\xi] + O(\varepsilon^2) & (8.38) \end{aligned}$$

Observe that at the origin $y(0) = \alpha$, as desired.

In conclusion, we got two solutions (8.36) and (8.38). One is good in the interior and near the right end while the other near the left end of the interval $[0, 1]$. In order to obtain a *uniform* expansion over $[0, 1]$, we may try to blend or match these two solutions. The basic idea of *matched asymptotic expansion* is that an approximate solution to a given problem is sought not as a single expansion in terms of a single scale but as two or more separate expansions in terms of two or more scales each of which is valid in part of the domain. We chose the scales so that the overall expansion covers the whole domain and that the domains of validity of neighboring expansions *overlap.* Because the domains overlap, we can *match* or blend the neighboring expansions. If this is possible, the resulting solution is called the *matched asymptotic expansion.* The *uniform* expansion is obtained adding (8.36) and (8.38) and subtracting their "common" part, so it is not counted twice. This common part is either the expansion of (8.36) as $x \to 0$ or that of (8.38) as $\xi \to \infty$. The matching condition insures that the two are equal.

We illustrate this basic idea on an example dealing with the height distribution in a trie. The reader may want to go back to Section 1.1 to review some definition related to tries.

Example 8.4 *Limiting Distribution of the Height in a Trie*

Let H_n be the height of a trie built from n binary strings generated by independent memoryless unbiased sources, and let $h_n^k = \Pr\{H_n \leq k\}$. Observe that h_n^k satisfies the following recurrence

$$h_n^{k+1} = 2^{-n}\sum_{i=0}^{n}\binom{n}{i}h_i^k h_{n-i}^k, \qquad k \geq 0 \tag{8.39}$$

$$h_0^0 = h_1^0 = 1, \quad \text{and} \quad h_n^0 = 0, \; n > 1. \tag{8.40}$$

This follows from $H_n = \max\{H_i^{LT}, H_{n-i}^{RT}\} + 1$, where H_i^{LT} and H_{n-i}^{RT} denote, respectively, the left subtree and the right subtree of sizes i and $n-i$, which happens with probability $2^{-n}\binom{n}{i}$ for memoryless unbiased sources. To solve this recurrence, we set

$$H^k(z) = \sum_{n\geq 0} h_n^k \frac{z^n}{n!}$$

to be the exponential generating function of h_n^k. Then (8.39) implies (cf. Table 7.3 in Chapter 7)

$$H^k(z) = \left(H^0(z2^{-k})\right)^{2^k} \tag{8.41}$$

where $H^0(z) = 1 + z$. Thus, by Cauchy's formula

$$\begin{aligned} h_n^k &= \frac{n!}{2\pi i}\oint (1+z2^{-k})^{2^k} z^{-n-1} dz \\ &= \begin{cases} 0, & n > 2^k \\ \frac{(2^k)!}{2^{nk}(2^k-n)!}, & 0 \leq n \leq 2^k \end{cases} \end{aligned} \tag{8.42}$$

where the second line follows directly from the binomial formula. Applying Stirling's formula to (8.42), we prove the following.

Theorem 8.3 (Devroye 1984, Pittel 1986, Knessl & Szpankowski 1999) *The distribution of the height in a trie has the following asymptotic expansions:*

(i) Right-Tail Region: $k \to \infty$, $n = O(1)$

$$\Pr\{H_n \leq k\} = h_n^k = 1 - n(n-1)2^{-k-1} + O(2^{-2k}).$$

(ii) Central Region: $k, n \to \infty$ *with* $\xi = n2^{-k}$, $0 < \xi < 1$

$$h_n^k \sim A(\xi)e^{n\phi(\xi)},$$

where

$$\begin{aligned} \phi(\xi) &= \left(1 - \frac{1}{\xi}\right)\log(1-\xi) - 1, \\ A(\xi) &= (1-\xi)^{-1/2}. \end{aligned}$$

(iii) Left-Tail Region: $k, n \to \infty$ *with* $2^k - n = j = O(1)$

$$h_n^k \sim n^j \frac{e^{-n-j}}{j!}\sqrt{2\pi n}.$$

This shows that there are three ranges of k and n where the asymptotic form of h_n^k is different.

We now use matched asymptotics to show that these three regions cover the whole domain where either k or n are large. If we expand (i) of Theorem 8.3 for n large, we obtain $1 - h_n^k \sim n^2 2^{-k-1}$. For $\xi \to 0$ we have $A(\xi) \sim 1$ and $\phi(\xi) \sim -\xi/2$ so that the result in (ii) becomes

$$A(\xi)e^{n\phi(\xi)} \sim e^{-n\xi/2} = \exp\left(-\frac{1}{2}n^2 2^{-k}\right) \sim 1 - \frac{1}{2}n^2 2^{-k}$$

where the last approximation assumes that $n, k \to \infty$ in such a way that $n^2 2^{-k} \to 0$. Since (ii) agrees precisely with the expansion of (i) as $n \to \infty$, we say that (i) and (ii) asymptotically match. To be precise, we say they match the leading order; higher order matchings can be verified by computing higher order terms in the asymptotic series in (i) and (ii). We can also easily show that the expansion of (ii) as $\xi \to 1^-$ agrees with the expansion of (iii) as $j \to \infty$, so that (ii) and (iii) also asymptotically match. Indeed, we observe that in this case $n(1/\xi - 1) = j$ and

$$n^j \frac{e^{-n-j}}{j!}\sqrt{2\pi n} \quad \to_{j\to\infty} \quad \exp\left(n(1-\frac{1}{\xi})\log(1-\xi) - 1\right)\frac{1}{\sqrt{1-\xi}} = A(\xi)e^{n\Phi(\xi)}.$$

The matching verifications imply that, at least to leading order, there are no "gaps" in the asymptotics. In other words, one of the results in (i)-(iii) applies for any asymptotic limit which has k and/or n large. □

We now briefly discuss the WKB approximation. We shall follow Nayfeh [258] and Fedoryuk [97]. In general, it assumes that an asymptotic solution is of the form

$$y(x;\lambda) \sim e^{\lambda\Phi(x)}\left(a(x) + \frac{1}{\lambda}A_1(x)\cdots\right)$$

or equivalently

$$y(x;\lambda) \sim \exp\left(\lambda(\Phi_0(x) + \lambda^{-1}\Phi_1(x) + \cdots\right)$$

where $\Phi_i(x)$ are unknown functions that one tries to determine from the equation. Note that $\Phi_0 = \Phi$ and $\Phi_1 = \log A$. For example, consider differential equation (8.33) and assume it has the above asymptotic expansion. Plugging it into (8.33), after some algebra, one arrives at the following system of equations

$$\begin{aligned} \Phi_0'^2(x) - q(x) &= 0, \\ \Phi_0''(x) + 2\Phi_0'(x)\Phi_1'(x) &= 0, \end{aligned}$$

which for $q(x) > 0$ yields

$$\begin{aligned} \Phi_0(x) &= \pm\int_{x_0}^{x}\sqrt{q(x)}dx, \\ \Phi_1(x) &= -\log\sqrt{|\Phi_0'(x)|} = -\frac{1}{4}\log|q(x)| \end{aligned}$$

for some constant x_0. Substituting this into the postulated asymptotic solution we finally show that

$$y(x;\lambda) \sim q^{-1/4}(x) \exp\left(\pm\lambda \int_{x_0}^{x} \sqrt{q(x)}dx\right).$$

We close this section with a more sophisticated application of the WKB method, matched asymptotics, and another applied mathematics tool known as *linearization*. This is a continuation of Example 4.

Example 8.5 *Limiting Distribution of the Height in a b-Tries*

We consider an extension of tries discussed in the previous example, namely, b-tries. We adopted this example from Knessl and Szpankowski [215]. In these tries external node is allowed to store up to b strings. The parameter b is assumed to be fixed. We study the height H_n (i.e., the longest path in such a tree), and let $h_n^k = \Pr\{H_n \le k\}$. As before, h_n^k satisfies the recurrence equation (8.39) with a new initial condition, namely

$$h_n^0 = 1, \ n = 0, 1, 2, \dots, b; \qquad \text{and} \qquad h_n^0 = 0, \ n > b.$$

The exponential generating function $H^k(z)$ of h_n^k satisfies the same equation (8.41) except that now $H^0(z) = 1 + z + \cdots + z^b/b!$. In other words,

$$h_n^k = n![z^n]\left(1 + z2^{-k} + \frac{z^2}{2!}2^{-2k} + \cdots + \frac{z^b}{b!}2^{-bk}\right)^{2^k}. \tag{8.43}$$

We can use the saddle point method (cf. Section 8.4 and Exercise 7) to extract asymptotic expansion of h_n^k from the above, however, in this example we apply indirect methods of WKB, matched asymptotics, and linearization to solve the problem. It will pay off in problems where one cannot obtain explicit expressions for the generating function (e.g., heights of PATRICIA tries and digital search trees, as analyzed in [215, 216]).

We consider first the case $k, n \to \infty$ with $\xi = n2^{-k}$ fixed. (The reader finds a complete analysis in Knessl and Szpankowski [215].) We set

$$h_n^k = G(\xi; n) = G(n2^{-k}; n)$$

and note that $h_i^{k-1} = G(2i\xi/n; i)$ and $h_{n-i}^{k-1} = G(2\xi(1 - i/n); n - i)$. From the original recurrence (8.39) we obtain

$$G(\xi; n) = \left(\frac{1}{2}\right)^n \sum_{i=0}^{n} \binom{n}{i} G\left(\frac{2i}{n}\xi; i\right) G\left(2\left(1 - \frac{i}{n}\right)\xi; n - i\right).$$

We analyze above by the WKB method. Thus, we seek an asymptotic solution of the form

$$G(\xi; n) \sim e^{n\phi(\xi)}\left[A(\xi) + \frac{1}{n}A^{(1)}(\xi) + \frac{1}{n^2}A^{(2)}(\xi) + \cdots\right]$$

By symmetry, the major contribution to the sum will come from $i \approx n/2$. We also note that Stirling's formula (8.31) yields, for $i = xn$ $(0 < x < 1)$,

$$2^{-n}\binom{n}{i} = \frac{e^{nf_0(x)}}{\sqrt{2\pi n}}\frac{1}{\sqrt{x(1-x)}}\left[1 + \frac{1}{12n}\left(1 - \frac{1}{x} - \frac{1}{1-x}\right) + O(n^{-2})\right], \tag{8.44}$$

where $f_0(x) = -\log 2 - x\log x - (1-x)\log(1-x)$. For $x = 1/2 + y/\sqrt{n}$ (i.e., $i = n/2 + O(\sqrt{n})$), (8.44) simplifies to the Gaussian form

$$2^{-n}\binom{n}{n/2 + y\sqrt{n}} = \sqrt{\frac{2}{\pi n}}e^{-2y^2}\left[1 + \frac{1}{n}\left(-\frac{1}{4} + 2y^2 - \frac{4}{3}y^4\right) + O(n^{-2})\right].$$

Thus, we are led to

$$\begin{aligned} A(\xi)e^{n\phi(\xi)} \quad &\sim \quad \sum_{i=0}^{n}\sqrt{\frac{2}{\pi n}}e^{-2y^2}A\left(\frac{2i}{n}\xi\right)A\left(\frac{2(n-i)}{n}\xi\right) \\ &\times \quad \exp\left[i\phi\left(\frac{2i}{n}\xi\right) + (n-i)\phi\left(2\left(1 - \frac{i}{n}\right)\xi\right)\right]. \end{aligned}$$

For $x = i/n$, we set $\psi(x) = x\phi(2x\xi) + (1-x)\phi(2(1-x)\xi)$ and expand this function about $x = 1/2$. We have $\psi(1/2) = \phi(\xi)$, $\psi'(1/2) = 0$ and $\psi''(1/2) = 8\xi\phi'(\xi) + 4\xi^2\phi''(\xi)$. Thus the exponent in the above expression becomes

$$\exp(n\psi(x)) = \exp\left(n\phi(\xi) + y^2(4\xi\phi'(\xi) + 2\xi^2\phi''(\xi)) + O(y^3)\right).$$

Using the Euler-Maclaurin formula, substituting $y = y'/\sqrt{n}$, and approximating the sum above by an integral we obtain as $n \to \infty$

$$e^{n\phi(\xi)}A(\xi) \sim \sqrt{\frac{2}{\pi}}e^{n\phi(\xi)}A^2(\xi)\int_{-\infty}^{\infty}\exp\left(-2y^2 + y^2(4\xi\phi'(\xi) + 2\xi^2\phi''\xi)\right)dy.$$

The exponential factors $e^{n\phi}$ cancel and we have

$$1 = \sqrt{\frac{2}{\pi}}A(\xi)\sqrt{\frac{\pi}{2 - 4\xi\phi'(\xi) - 2\xi^2\phi''(\xi)}}. \tag{8.45}$$

Thus, the asymptotic series of the WKB approximation is known, up to the function $\phi(\xi)$. It does not seem to be possible to determine ϕ using only the recurrence (8.39) itself. This function is apparently very sensitive to the initial conditions. However, using matched asymptotics, we shall show how to determine $\phi(\xi)$ for $\xi \to 0$ (in a similar manner, one can find $\phi(\xi)$ as $\xi \to b$).

For this, we turn our attention to the case $n = O(1)$ and $k \to \infty$. We set $h_n^k = 1 - G_n^k$. Then G_n^k satisfies

$$G_n^{k+1} = 2\left(\frac{1}{2}\right)^n \sum_{i=0}^{n} \binom{n}{i} G_{n-i}^k - \left(\frac{1}{2}\right)^n \sum_{i=0}^{n} \binom{n}{i} G_i^k G_{n-i}^k \qquad (8.46)$$
$$= 2\left(\frac{1}{2}\right)^n \sum_{i=0}^{n-b-1} \binom{n}{i} G_{n-i}^k - \left(\frac{1}{2}\right)^n \sum_{i=b+1}^{n-b-1} \binom{n}{i} G_i^k G_{n-i}^k.$$

Here we have used the fact that $G_n^k = 0$ for $0 \le n \le b$. For $b+1 \le n \le 2b+1$, the non-linear term in (8.46) vanishes and we are left with a *linear* recurrence. We can solve (8.46) exactly by first solving the linear problem for $n \in [b+1, 2b+1]$, and using this solution to compute explicitly the nonlinear term in (8.46) for $n \in [2b+2, 3b+2]$, etc. However, the resulting expressions become complicated and we only need asymptotics of G_n^k. For ranges of k, n where h_n^k is *asymptotically close to* 1, we can replace (8.46) by the asymptotic relation (thus *linearize* the recurrence) that yields

$$G_n^{k+1} \sim 2^{1-n} \sum_{i=0}^{n-b-1} \binom{n}{i} G_{n-i}^k.$$

Solving this recurrence asymptotically one proves

$$G_n^k \sim 2^{-b} \binom{n}{b+1}, \qquad n \ge b+1.$$

Thus for $k \to \infty$ and $n = O(1)$ we have

$$\mathsf{Pr}\{H_n \le k\} = h_n^k \sim 1 - \binom{n}{b+1} 2^{-kb}.$$

As in Example 4, we now match the two expansions just derived, that is, we require

$$1 - \frac{n^{b+1}}{(b+1)!} 2^{kb} \Big|_{n\to\infty} \sim A(\xi) e^{n\phi(\xi)} \Big|_{\xi \to 0}.$$

Since $A(\xi) \to 1$ as $\xi \to 0$, we obtain

$$\phi(\xi) \sim -\frac{\xi^b}{(b+1)!}, \qquad \xi \to 0,$$

and hence

$$\mathsf{Pr}\{H_n^T \le k\} \sim A(\xi) e^{n\phi(\xi)} \sim \exp\left(-\frac{n\xi^b}{(b+1)!}\right) \qquad \xi \to 0.$$

In passing we should point out that this range is the most interesting since the probability mass is concentrated around $k = (1+1/b)\log_2 n + x$ where $x = O(1)$. In fact,

$$\begin{aligned}\Pr\{H_n^T \le (1+1/b)\log_2 n + x\} &= \Pr\{H_n^T \le \lfloor (1+1/b)\log_2 n + x \rfloor\} \\ &\sim \exp\left(-\frac{1}{(1+b)!}2^{-bx+b\langle (1+b)/b\cdot \log_2 n + x\rangle}\right),\end{aligned}$$

where $\langle x\rangle$ is the fractional part of x, that is, $\langle x\rangle = x - \lfloor x\rfloor$. Due to the erratic behavior of $\langle \log_2 n\rangle$ (cf. next section), the limiting distribution of the height H_n does not exist! The reader can find detailed analyses on this and similar topics in a series of recent papers by Knessl and Szpankowski [214, 215, 216, 217]. $\square$

8.2.3 Uniform Distribution of Sequences

In the analysis of algorithms and information theory one often deals with sequences like $\langle \log n\rangle$ and $\langle \alpha n\rangle$ where $\langle x\rangle = x - \lfloor x\rfloor$ is the fractional part of x. We just saw it in the example above. In fact, we are often interested in asymptotics of sums like

$$\sum_{k\ge 0} p_{n,k} f(\langle x_k\rangle)$$

as $n\to\infty$, where $p_{n,k}$ is a probability distribution (i.e., $\sum_k p_{n,k} = 1$) and f is a function. For example, the following sum

$$\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \left\langle \log\left(p^k(1-p)^{n-k}\right)\right\rangle$$

arises in coding theory. Our goal here is to present certain tools from theory of *uniform distribution of sequences* to extract asymptotics of such sums. We shall follow Drmota and Tichy [89].

We start with some definitions.

Definition 8.4 (P-u.d. mod 1) *A sequence $x_n \in \mathbb{R}$ with respect to the probability distribution $P = \{p_{n,k}\}_{k\ge 0}$ is said to be P-uniformly distributed modulo 1 (in short P-u.d. mod 1) if*

$$\lim_{n\to\infty} \sum_{k\ge 0} p_{n,k} I_A(\langle x_k\rangle) = \lambda(A) \qquad (8.47)$$

holds uniformly for every interval $A\subset[0,1]$, where $I_A(x_n)$ is the characteristic function of A (i.e., it equals to 1 if $x_n\in A$ and 0 otherwise) and $\lambda(A)$ is the Lebesgue measure of A.

Two special cases are of interest to us:

- *uniform distributed sequences* modulo 1 (in short u.d. mod 1) in which case $p_{n,k} = 1/n$, that is, (8.47) becomes

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n} I_A(\langle x_k \rangle) = \lambda(A) \tag{8.48}$$

- *Bernoulli distributed sequences* modulo 1 (B-u.d. mod 1) that assumes the binomial distribution for $p_{n,k}$, that is, for $p > 0$ (8.47) becomes

$$\lim_{n\to\infty} \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} I_A(\langle x_k \rangle) = \lambda(A). \tag{8.49}$$

The following result summarizes the main property of P-u.d. modulo 1 sequences. It provides the leading term of asymptotics for sums like $\sum_k p_{n,k} f(\langle x_k + y \rangle)$ where x_k is P-u.d. mod 1 and y is a shift.

Theorem 8.5 *Suppose that the sequence x_n is P-uniformly distributed modulo 1. Then for every Riemann integrable function $f : [0,1] \to \mathbb{R}$ we have*

$$\lim_{n\to\infty} \sum_{k\geq 0} p_{n,k} f(\langle x_k + y \rangle) = \int_0^1 f(t)\, dt, \tag{8.50}$$

where the convergence is uniform for all shifts $y \in \mathbb{R}$.

Proof. The proof is standard but details are quite tedious. The main idea is to notice that by definition (8.50) holds for characteristic functions $I_A(x_k)$. Next, we approximate f by a step function (i.e., a linear combination of characteristic functions) and use the definition of the Riemann integral to bound the integral from below and above proving (8.50).

Details are here. Let $\mathcal{L}$ be the linear space of all (Lebesgue) integrable functions $f : [0,1]^k \to \mathbb{C}$ satisfying

$$\lim_{n\to\infty} \sup_{y\in\mathbb{R}} \left| \sum_{k\geq 0} p_{n,k} f(\langle x_k + y \rangle) - \int_0^1 f(t)\, dt \right| = 0. \tag{8.51}$$

Of course, $\mathcal{L}$ contains all constant functions, i.e. $\mathcal{L} \neq \emptyset$. Next we prove the following closure property of $\mathcal{L}$.

> **Fact A**. Suppose $f : [0,1] \to \mathbb{R}$ is an integrable function such that for every $\varepsilon > 0$ there exist functions $g_1, g_2 \in \mathcal{L}$ with $g_1 \leq f \leq g_2$ and
>
> $$\int_0^1 (g_2(t) - g_1(t))\, dt < \varepsilon.$$
>
> Then $f \in \mathcal{L}$.

For this purpose, let $m_n(f,y)$ denote the positive linear functionals

$$m_n(f,y) := \sum_{k\ge 0} p_{n,k} f(\langle x_k + y\rangle)$$

and $m(f)$ the positive linear functional

$$m(f) := \int_0^1 f(t)\,dt.$$

By $m_n(g_1,y) \le m_n(f,y) \le m_n(g_2,y)$ and $m(g_1) \le m(f) \le m(g_2)$ we immediately get

$$\begin{array}{rcl} m(g_1) & = & \liminf_{n\to\infty} m_n(g_1,y) \quad \le \quad \liminf_{n\to\infty} m_n(f,y) \\ & \le & \limsup_{n\to\infty} m_n(f,y) \quad \le \quad \limsup_{n\to\infty} m_n(g_2,y) \\ & = & m(g_2) \end{array}$$

that implies

$$|m(f) - \liminf_{n\to\infty} m_n(f,y)| < \varepsilon$$

and

$$|m(f) - \limsup_{n\to\infty} m_n(f,y)| < \varepsilon$$

for every $\varepsilon > 0$. Thus

$$\lim_{n\to\infty} \sup_{y\in\mathbb{R}} |m_n(f,y) - m(f)| = 0,$$

and $f \in \mathcal{L}$.

Next, suppose that $f : [0,1] \to \mathbb{R}$ is a continuous function. Then by the Weierstrass approximation theorem, for every $\varepsilon > 0$ there exist two trigonometric polynomials g_1, g_2 with $g_1 \le f \le g_2$ and

$$\int_0^1 (g_2(t) - g_1(t))\,dt < \varepsilon.$$

By fact A, all continuous functions $f : [0,1] \to \mathbb{R}$ are contained in $\mathcal{L}$.

Now, it is easy to derive that every characteristic function $f = I_A$ of any interval $A \subseteq [0,1]$ is in $\mathcal{L}$, too. It is obvious, that for every $\varepsilon > 0$ there exist two continuous functions g_1, g_2 with $g_1 \le I_A \le g_2$ and

$$\int_0^1 (g_2(t) - g_1(t))\,dt < \varepsilon.$$

Again by Fact A all step functions (i.e. finite linear combinations of characteristic functions) are in $\mathcal{L}$.

Finally, if $f : [0,1] \to \mathbb{R}$ is a Riemann integrable function, then by definition for every $\varepsilon > 0$ there exist two step functions g_1, g_2 with $g_1 \le f \le g_2$ and

$$\int_0^1 (g_2(t) - g_1(t))\,dt < \varepsilon,$$

and by Fact A this completes the proof. ∎

In passing we observe that Theorem 8.5 holds if f is continuous. Furthermore, due to $\langle x_k \rangle$, we may restrict f to complex-valued continuous function with period 1.

To apply Theorem 8.5, one needs an easy criteria to verify whether a sequence x_k is P-u.d. mod 1 or not. Fortunately, such a result exists and is due to H. Weyl.

Theorem 8.6 (Weyl 1916) *A sequence x_n is P-u.d. mod 1 if and only if*

$$\lim_{n\to\infty} \sum_{k\geq 0} p_{n,k} e^{2\pi i m x_k} = 0 \tag{8.52}$$

holds for all $m \in \mathbb{Z} \setminus \{0\}$.

Proof. The proof is standard and we only sketch it here (for details see [89, 230]). First observe that $f(x) = e^{2\pi i m x}$ satisfies (8.50). To prove sufficiency, we note that by Weierstrass's *approximation theorem* every Riemann integrable function f of period 1 can be uniformly approximated by a trigonometric polynomial (i.e., a finite combination of functions of the type $e^{2\pi i m x}$). That is, for any $\varepsilon > 0$ there exists a trigonometric polynomial $\Psi(x)$ such that

$$\sup_{0\leq x\leq 1} |f(x) - \Psi(x)| \leq \varepsilon. \tag{8.53}$$

Then

$$\begin{aligned}
\left| \int_0^1 f(x)dx - \sum_{k\geq 0} p_{n,k} f(x_k) \right| &\leq \left| \int_0^1 (f(x) - \Psi(x))\, dx \right| \\
&+ \left| \int_0^1 \Psi(x)dx - \sum_{k\geq 0} p_{n,k}\Psi(x_k) \right| \\
&+ \left| \sum_{k\geq 0} p_{n,k} \left(\Psi(x_k) - f(x_k)\right) \right| \\
&\leq 3\varepsilon
\end{aligned}$$

since the first and the third term are bounded by ε due to (8.53) and the second by ε due to (8.52). This completes the proof. ∎

One of the most important and well studied sequence is $\{\alpha n\}_{n\geq 0}$ for α irrational. We discuss its properties in the example below.

Example 8.6 *Special Sequence $\langle \alpha n \rangle_{n\geq 0}$*

Let us consider $\langle \alpha n \rangle_{n\geq 0}$ for α irrational. We prove that it is u.d mod 1 and B-u.d. mod 1. For the former, observe that by Weyl's criterion for every integer $m \neq 0$

$$\left| \frac{1}{n} \sum_{k=1}^{n} e^{2\pi i k m \alpha} \right| = \frac{|e^{2\pi i m n \alpha} - 1|}{n|e^{2\pi i m \alpha} - 1|} \leq \frac{1}{n|\sin \pi m \alpha|} \to 0, \qquad n \to \infty$$

as long as α is not rational. Thus $\langle \alpha n\rangle_{n\geq 0}$ is uniformly distributed sequence modulo 1.

Consider now the B-u.d. mod 1 case. We have by Weyl's criterion (where, as usual, $q = 1-p$)

$$\lim_{n\to\infty}\sum_{k=0}^{n}\binom{n}{k}p^kq^{n-k}e^{2\pi im(k\alpha)} = \lim_{n\to\infty}\left(pe^{2\pi im\alpha}+q\right)^n = 0$$

provided α is irrational. □

Let us stay with sequences $\langle \alpha n\rangle_{n\geq 0}$ but now we assume that α is rational, say $\alpha = N/M$ such that $\gcd(N, M) = 1$. Can we say something meaningful about the sum $\sum_k p_{n,k} f(\langle \alpha k\rangle)$ in this case? To simplify our presentation we restrict our analysis to the binomial distribution $p_{n,k} = \binom{n}{k}p^kq^{n-k}$ $(q = 1-p)$. We prove the following general result.

Theorem 8.7 *Let $0 < p < 1$ be a fixed real number and suppose that $\alpha = \frac{N}{M}$ is a rational number with* $\gcd(N, M) = 1$. *Then, for every bounded function $f : [0, 1] \to \mathbb{R}$ we have for some $\rho < 1$*

$$S_n(f) = \sum_{k=0}^{n}\binom{n}{k}p^k(1-p)^{n-k}f(\langle k\alpha + y\rangle) = \frac{1}{M}\sum_{l=0}^{M-1}f\left(\frac{l}{M}+\frac{\langle My\rangle}{M}\right)+O(\rho^n)$$

uniformly for all $y \in \mathbb{R}$.

Proof. Observe that (with $p_{n,k} = \binom{n}{k}p^kq^{n-k}$)

$$\begin{aligned}
S_n(f) &= \sum_{k=0}^{n}p_{n,k}f\left(\left\langle k\frac{N}{M}+y\right\rangle\right)\\
&= \sum_{\ell=0}^{M-1}\sum_{m:\ k=\ell+mM\leq n}p_{n,k}f\left(\left\langle \ell\frac{N}{M}+N+y\right\rangle\right)\\
&= \sum_{\ell=0}^{M-1}\sum_{m:\ k=\ell+mM\leq n}p_{n,k}f\left(\left\langle \frac{\ell}{M}+y\right\rangle\right)\\
&= \sum_{\ell=0}^{M-1}f\left(\left\langle \frac{\ell}{M}+y\right\rangle\right)\sum_{m:\ k=\ell+mM\leq n}\binom{n}{k}p^kq^{n-k}.
\end{aligned}$$

To proceed, we need to evaluate the second sum. This sum takes every Mth term from the binomial distribution. In Section 7.6.1 we learned how to handle such sums through the Mth root of unity $\omega_k = e^{2\pi ik/M}$ for $k = 0, 1, \ldots, M-1$. In particular, applying (7.11) we easily obtain

$$\begin{aligned}
\sum_{m:\ k=\ell+mM\leq n}\binom{n}{k}p^kq^{n-k} &= \frac{1+(p\omega_1+q)^{n-\ell}+\cdots+(p\omega_{M-1}+q)^{n-\ell}}{M}\\
&= \frac{1}{M}+O(\rho^n), \qquad (8.54)
\end{aligned}$$

where $\rho < 1$ since $|(p\omega_r + q)| = p^2 + q^2 + 2pq\cos(2\pi r/M) < 1$ for $r \neq 0$. Thus

$$S_n(f) = \frac{1}{M}\sum_{\ell=0}^{M-1} f\left(\left\langle \frac{\ell}{M} + y \right\rangle\right) + O(\rho^n).$$

To complete the proof, we observe that the function

$$F(y) = \frac{1}{M}\sum_{l=0}^{M-1} f\left(\left\langle \frac{l}{M} + y \right\rangle\right)$$

is periodic with period $1/M$. Thus, for $0 \leq y < \frac{1}{M}$ we have the representation

$$F(y) = \frac{1}{M}\sum_{l=0}^{M-1} f\left(\frac{l}{M} + y\right).$$

For general $y \in \mathbb{R}$ we only need to use the relation $F(y) = F(\langle My\rangle/M)$ to finish the proof. (In Exercises 8–9 we ask the reader to use another method, that of Fourier series, to establish the above result.) ■

We complete this section with an interesting example arising in source coding.

Example 8.7 *Redundancy of the Shannon-Fano Code*
We consider here a very simple code, known as the Shannon-Fano code, that assigns code length $\lceil -\log_2 p^k(1-p)^{n-k})\rceil$ to a sequence x_1^n of length n occurring probability $p(k) = p^k q^{n-k}$ where k is the number of "1" in x_1^n (cf. [59]) From Shannon First Theorem 6.18 we know that the average code length cannot be smaller than the entropy of the source which in our case is $\sum_{k=0}^{n}\binom{n}{k}p^k q^{n-k}\log_2(p^k q^{n-k})$. From Chapter 6, we know that the difference between the average code length and the entropy is known as the average *redundancy* (cf. Section 8.7.2 for a more in depth discussion of redundancy issues). We denote it as $\bar{R}_n$. For the Shannon-Fano block code of length n generated by a memoryless source we have

$$\begin{aligned}
\bar{R}_n &= \sum_{k=0}^{n}\binom{n}{k}p^k q^{n-k}\left(\lceil -\log_2 p(k)\rceil + \log_2 p(k)\right) \\
&= 1 + \sum_{k=0}^{n}\binom{n}{k}p^k q^{n-k}\left(\log_2 p(k) + \lfloor -\log_2 p(k)\rfloor\right) \\
&= 1 - \sum_{k=0}^{n}\binom{n}{k}p^k q^{n-k}\langle \alpha k + \beta n\rangle \qquad (8.55)
\end{aligned}$$

where $\alpha = \log_2((1-p)/p)$ and $\beta = \log_2(1/(1-p))$. But the sum above is the one we analyzed in Theorem 8.5 and Theorem 8.7. If α is irrational by Theorem 8.5 (applied to the binomial distribution with $f(t)=t$) we obtain $\bar{R}_n = \frac{1}{2}+o(1)$. For $\alpha = N/M$, we apply Theorem 8.7 to arrive at

$$\begin{aligned}
\bar{R}_n &= 1 - \frac{1}{M}\sum_{l=0}^{M-1}\left(\frac{l}{M}+\frac{\langle M\beta n\rangle}{M}\right)+O(\rho^n) \\
&= 1-\frac{M-1}{2M}+\frac{\langle M\beta n\rangle}{M}+O(\rho^n) \\
&= \frac{1}{2}-\frac{1}{M}\left(\langle M\beta n\rangle-\frac{1}{2}\right)+O(\rho^n)
\end{aligned}$$

In summary, we prove the following.

Theorem 8.8 (Szpankowski 1999) *Consider the Shannon–Fano block code of length n binomially(n,p) distributed over a binary alphabet. Then, for $p<\frac{1}{2}$ as $n\to\infty$*

$$\bar{R}_n = \begin{cases} \frac{1}{2}+o(1) & \alpha \quad \text{irrational} \\ \\ \frac{1}{2}-\frac{1}{M}\left(\langle Mn\beta\rangle - \frac{1}{2}\right)+O(\rho^n) & \alpha=\frac{N}{M}, \quad \gcd(N,M)=1 \end{cases} \tag{8.56}$$

where $\rho<1$. □

8.3 Small Singularities of Analytic Functions

It is time to deliver what the title of this section says, namely, complex asymptotics. From now on we shall deal with the generating function

$$A(z)=\sum_{n=0}^{\infty} a_n z^n$$

and study various methods of extracting asymptotic expansions of a_n from an explicit knowledge of $A(z)$, and even more often, from partial knowledge of the generating function $A(z)$ around its singularities. Often the generating functions we work with (in combinatorics, analysis of algorithms, information theory, etc.) are *analytic functions.*[2] Therefore, by Cauchy's formula the coefficient a_n can be evaluated as

$$a_n = \frac{1}{2\pi i}\int_{\mathcal{C}} \frac{f(z)}{z^{n+1}}dz \tag{8.57}$$

[2]The reader may want to review elements of complex analysis discussed in Section 2.3. There are also excellent books on the subject; e.g., Henrici [158], Hille [159], Titchmarsh [337] and Whittaker and Watson [353].

where $\mathcal{C}$ is a closed contour that encircles the origin in the positive direction (i.e., counterclockwise). One may ask why to use Cauchy's integral formula for a_n when simpler expressions exist (e.g., $a_n = \frac{1}{n!}A^{(n)}(z)|_{z=0}$ where $A^{(n)}(z)$ is the nth derivative of $A(z)$). The point to observe is that integration preserves many nice properties of analytic function; the contour $\mathcal{C}$ can be deformed in many ways to take advantages of analytic continuations of $A(z)$; and finally this integral is usually easily computable through the Cauchy residue theorem (cf. Theorem 2.5).

We shall argue that complex methods are successful in extracting asymptotics of a_n since the *asymptotic growth of a_n is determined by the location of the singularities closest to the origin.* To see this, we first recall that a singularity of an analytic function is a point where the function ceases to be analytic. Next, we recall from Hadamard's Theorem 7.1 that the radius of convergence of the series involved in $A(z)$ is

$$R^{-1} = \limsup_{n\to\infty} |a_n|^{1/n},$$

or informally $\frac{1}{n}\log|a_n| \sim 1/R$. That is, for every $\varepsilon > 0$ there exists N such that for $n > N$ we have

$$|a_n| \le (R^{-1} + \varepsilon)^n \ ;$$

and for infinitely many n we have

$$|a_n| \ge (R^{-1} - \varepsilon)^n \ .$$

Thus, the exponential growth of a_n is determined by R^{-n}. A classical theorem (attributed to Pringsheim) says that an analytic function of finite radius of convergence must have a singularity on the boundary of its disk of convergence. Moreover, as we have already seen in Section 7.3.3, a bound on a_n is easily achievable through Cauchy's Integral Theorem 2.6 (cf. (2.21) of Section 2.3), namely,

$$|a_n| \le \frac{M(r)}{r^n} \tag{8.58}$$

where $M(r)$ is the maximum value of $|A(z)|$ for any circle $r < R$. This is the *Cauchy bound*, and we shall use it often in this chapter. In summary, locations of singularities of $A(z)$ determine asymptotic behavior of its coefficients. To obtain a more refined information about such asymptotics we must study more carefully various types of singularities, which we do next.

Finally, a last word of this introduction. One of the most important aspect of complex asymptotics is to provide the so–called *transfer theorems* under which the following central implication is valid:

$$A(z) \sim f(z) \quad \Longrightarrow \quad [z^n]A(z) \sim [z^n]f(z) \tag{8.59}$$

where $A(z)$ is an implicitly known generating function and $f(z)$ is an asymptotic expansion of $A(z)$ near a singularity. (The above is not true in general; e.g., $A(z) = 1/(1-z) \sim 2/(1-z^2) =$

$1/(1-z)+1/(1+z) =: B(z)$ at $z=1$ but $[z^n]A(z) = 1$ while $[z^n]B(z) = 1+(-1)^n$.) Following Odlyzko [262], we shall talk about *small singularity* if $A(z)$ has polynomial growth around such a singularity; otherwise it is called a *large singularity.* For example, $(1-z)^{3/2}\log(1-z)$ has a small singularity at $z_0 = 1$, while $\exp(1/(1-z))$ is said to have a large singularity at $z_0 = 1$. These are not precise concepts, but still clearly different methods are used to handle them. Therefore, we divide our discussion into two parts: In this section, we deal with methods for small singularities (i.e., poles and algebraic singularities) while in the next section we turn our attention to large singularities.

8.3.1 Polar Singularities

Here we deal with polar singularities that we define precisely below. Let $A(z)$ be an analytic function in a ring $0 < |z - z_0| < R \le \infty$. Then, the Laurent expansion applies and one finds (cf. [159, 337])

$$A(z) = \sum_{n=-M}^{-1} a_n(z-z_0)^n + \sum_{n=1}^{\infty} a_n(z-z_0)^n \tag{8.60}$$

where the coefficients a_n, $n = -M, -M+1, \ldots, 0, 1, \ldots$ are computed by the Cauchy formula (8.57). The point z_0 is called a *singularity*, and three cases are to be considered. If $a_{-n} = 0$ for all $n = 1, \ldots, M$, then z_0 is a *removable singularity* and $\lim_{z\to z_0} A(z) = a_0$. If M is finite, then z_0 is called a *pole* of order M. Finally, if $M = \infty$ then we have an *essential singularity.* We mostly deal here with poles since a function behaves very weird near essential singularities (e.g., according to the Casorati-Weierstrass theorem the function can take *any* value around such a singularity).

To continue with our definitions: a function $A(z)$ is called *meromorphic* if its only singularities are poles. A special case of meromorphic functions are *rational functions* which are the ratios of two polynomials.

We first discuss **rational functions**, that is, we assume that $A(z) = N(z)/D(z)$ where both $N(z)$ and $D(z)$ are polynomials. Hence, we can decompose $A(z)$ into a finite sum like

$$A(z) = \sum_{\rho,r} \frac{a_{\rho,r}}{(z-\rho)^r}$$

where the summation is over all poles (roots) of $A(z)$ (of $D(z)$) and r ranges over a set of integers determined by the largest multiplicity of a pole of A. To extract the coefficient a_n, it suffices to study asymptotics of $1/(z-\rho)^r$. But, in fact, we can do it *exactly*, not only asymptotically. Indeed, from Table 7.2 Entry 1, we find

$$\begin{aligned}
[z^n]\frac{1}{(z-\rho)^j} &= [z^n]\frac{(-1)^j}{\rho^j(1-z/\rho)^j} \\
&= (-1)^j\binom{n+j-1}{n}\rho^{-(n+j)} = (-1)^j\binom{n+j-1}{j-1}\rho^{-(n+j)}.
\end{aligned}$$

In summary, *if $A(z)$ is rational which is analytic at zero and has poles at points $\rho_1, \ldots, \rho_m$, then there exist polynomials $P_j(n)$ of degree equal to the order of pole ρ_j minus one such that exactly*

$$a_n := [z^n]A(z) = \sum_{j=1}^{m} P_j(n)\rho_j^{-n}. \tag{8.61}$$

We illustrate this result with a simple example which is a continuation of the pattern occurrence problem discussed in Section 7.6.2.

Example 8.8: *Frequency of a Given Pattern Occurrence*
Let H be a *given* pattern of size m, and consider a random text of length n generated by a memoryless source. In Section 7.6.2 we studied the probability generating function $T_r(z) = \sum_{n=0}^{\infty} \Pr\{O_n = r\}z^n$ of the number O_n of pattern H occurrences in the text. In Theorem 7.13 we proved (for a memoryless source) that

$$T_r(z) = \frac{z^m P(H)(D(z) + z - 1)^{r-1}}{D^{r+1}(z)}$$

where $D(z) = P(H)z^m + (1-z)A_H(z)$ and $A(z)$ is the *autocorrelation polynomial* (a polynomial of degree m). Let now $\rho < \rho_1 < \rho_2 < \cdots < \rho_{m-1}$ be m *distinct* roots of $D(z) = 0$. Then, an easy application of the above analysis, leads to

$$\begin{aligned} \Pr\{O_n(H) = r\} &= \sum_{j=1}^{r+1}(-1)^j a_j \binom{n}{j-1}\rho^{-(n+j)} + \sum_{k=1}^{m-1}\sum_{i=1}^{r+1}(-1)^i b_{ki}\binom{n}{i-1}\rho_k^{-(n+i)} \\ &= \sum_{j=1}^{r+1}(-1)^j a_j \binom{n}{j-1}\rho^{-(n+j)} + O(\rho_1^{-n}) \end{aligned} \tag{8.62}$$

where $a_{r+1} = \rho^m P(H)\,(\rho-1)^{r-1}\,(D'(\rho))^{-r-1}$ while the other coefficients a_j and b_{ki} can be computed when needed. □

Let us now turn our attention to a more general case, namely, **meromorphic functions**. We derive an asymptotic expansion for the coefficients a_n similar to the one for rational functions except that this time we only obtain asymptotic approximation not an exact formula. There are two approaches: one is called *subtracted singularities* and the other is known as the *contour integral* method.

We start with the subtracted singularities method. To make our discussion more concrete we study the following function

$$A(z) = \sum_{j=1}^{r} \frac{a_{-j}}{(z-\rho)^j} + \sum_{j=0}^{\infty} a_j(z-\rho)^j \ .$$

Then , we assume that $A(z)$ has a Laurent expansion around a pole ρ of multiplicity r. We further assume that the pole ρ is the closest to the origin, that is, the radius of convergence R is $R = |\rho|$, and there are no other poles on the circle of convergence. In other words, the second term of $A(z)$, which we denote for simplicity as $A_1(z)$, is analytic in the circle $|z| \leq |\rho|$, and its radius of convergence is $R_1 > |\rho|$. Thus, coefficients of $A_1(n)$ are bounded by $O(R_1^{-n})$ due to Cauchy's bound (8.58). Using the exact expansion of the first part of $A(z)$, which is a rational function, we arrive at

$$[z^n]A(z) = \sum_{j=1}^{r} (-1)^j a_{-j} \binom{n}{j-1} \rho^{-(n+j)} + O(R_1^{-n})$$

for $R_1 > \rho$. Observe that in the last example formula (8.62) is presented in this form.

The main thrust of the method just described is as follows: Imagine we are interested in the asymptotics for coefficients a_n of a function $A(z)$ whose circle of convergence is R. Let us also assume that we can find a simpler function, say $\bar{A}(z)$ that has the same singularities as $A(z)$ (e.g., in the example above $\bar{A}(z) = \sum_{j=1}^{r} \frac{a_{-j}}{(z-\rho)^j}$). Then, $A_1(z) = A(z) - \bar{A}(z)$ is analytic in a larger disk, of radius $R_1 > R$, say, and its coefficients are bounded by $O(R_1^{-n})$.

We now re-derive the above formula using the *contour integral* approach. Applying Cauchy's formula and extending the contour of integration to a circle of radius r that contains inside all singularities of $A(z)$, we shall deal with the following integral

$$I_n = \frac{1}{2\pi i} \int_{|z|=r} \frac{A(z)}{z^{n+1}} dz.$$

By Cauchy's residue theorem, this integral is equal to

$$I_n = \mathrm{Res}[A; z=0] + \sum_{\rho} \mathrm{Res}[A; z=\rho]$$

where the summation is over all singularities ρ contained inside the circle $|z| = r$. But by Cauchy's bound (8.58) we also know that $|I_n| = O(r^{-n})$. Thus

$$a_n = [z^n]A(z) = -\sum_{\rho} \mathrm{Res}[A; z=\rho] + O(r^{-n}).$$

This leads to the same estimate as above.

We can summarize our discussion in the following theorem.

Theorem 8.9 *Let $A(z)$ be a meromorphic function for $|z| \leq r$ with poles at $\rho_1, \rho_2, \ldots, \rho_m$ and analytic on $|z| = r$ and $z = 0$. Then there exist polynomials $P_i(n)$ such that*

$$a_n := [z^n]A(z) = \sum_{j=1}^{m} P_j(n) \rho_j^{-n} + O(r^{-n})$$

where the degree of P_j is equal to the order of the pole at ρ_j minus one.

We finish this section with an example that will combine what we have learned in the last two sections.

Example 8.9 *Enumeration of (d,k) Sequences*

For run-length coding the so called (d,k) $(d \le k)$ sequences are of prime importance. A binary string is a (d,k) sequence if it does not contain any runs of 0's shorter than d and longer than k. In other words, between any pair of 1's there must be at least d and at most k zeros. We want to enumerate all such (d,k) sequences. We first construct the ordinary generating function $W_{d,k}(z) = \sum_{w \in \mathcal{W}_{d,k}} z^{|w|}$ of all (d,k) words denoted as $\mathcal{W}_{d,k}$. We apply the symbolic methods of Chapter 7. Let $\mathcal{A}_{d,k}$ be the set of all words (strings) consisting only of 0's whose length is between d and k. The generating function $A(z)$ is clearly equal to

$$A(z) = z^d + z^{d+1} + \cdots + z^k = z^d \frac{1 - z^{k-d+1}}{1-z}.$$

We now observe that $\mathcal{W}_{d,k}$ can by symbolically written as

$$\mathcal{W}_{d,k} = \mathcal{A}_{d,k}\left(\{1\} \times \epsilon + \bar{\mathcal{A}}_{d,k} + \bar{\mathcal{A}}_{d,k} \times \bar{\mathcal{A}}_{d,k} + \cdots + \bar{\mathcal{A}}_{d,k}^{k} + \cdots\right) \tag{8.63}$$

where $\bar{\mathcal{A}}_{d,k} = \{1\} \times \mathcal{A}_{d,k}$. Above basically says that the collection of words $\mathcal{W}_{d,k}$ is a concatenation of $\{1\} \times \mathcal{A}_{d,k}$. Thus, (8.63) translates into the generating functions $W_{d,k}(z)$ as follows

$$\begin{aligned} W_{d,k}(z) &= A(z)\frac{1}{1 - zA(z)} = \frac{z^d(1 - z^{k+1-d})}{1 - z - z^{d+1} + z^{k+2}} \\ &= \frac{z^d + z^{d+1} + \cdots + z^k}{1 - z^{d+1} - z^{d+2} - \cdots - z^{k+1}}. \end{aligned} \tag{8.64}$$

Of particular interest is the so called capacity $\mathcal{C}_{d,k}$ defined as (cf. [14])

$$\mathcal{C}_{d,k} = \lim_{n\to\infty} \frac{\log[z^n]W_{d,k}(z)}{n}.$$

If ρ is the smallest root of $1 - z^{d+1} - z^{d+2} - \cdots - z^{k+1} = 0$, then clearly

$$C_{d,k} = -\log \rho.$$

But we can do much better. The function $W_{d,k}(z)$ is rational and we can compute $|\mathcal{W}_{d,k}|$ exactly. Let us consider a particular case, namely, $d = 1$ and $k = 3$. Then the denumerator in (8.64) becomes $1 - z^2 - z^3 - z^4$, and its roots are

$$\rho_{-1} = -1, \quad \rho_0 = 0.682327\ldots, \quad \rho_1 = -0.341164\ldots + i1.161541\ldots, \quad \rho_2 = \bar{\rho}_1.$$

Using the contour integral approach and computing the residues we obtain

$$[z^n]W_{1,3}(z) = \frac{\rho_0 + \rho_0^2 + \rho_0^3}{(\rho_1+1)(\rho_0-\rho_1)(\rho_0-\bar{\rho}_1)}\rho_0^{-n-1} + (-1)^{n+1}\frac{1}{(\rho_0+1)(\rho_1+1)(\bar{\rho}_1+1)} + O(r^{-n})$$

where $r \approx 0.68$. More specifically,

$$[z^n]W_{1,3}(z) = 0.594(1.465)^{n+1} + 0.189(-1)^{n+1} + O(0.68^n).$$

If we wish, we could compute the "Big Oh" term more precisely. □

8.3.2 Algebraic Singularities: Singularity Analysis

In the last subsection, we considered one type of singularity, namely, poles. For such isolated singularities, a function was well defined in its whole punctured neighborhood of, say z_0. In particular, in $0 < |z - z_0| < R$ the function $A(z)$ has its Laurent series expansion. In this subsection, we study algebraic singularities for which Laurent's expansion does not apply since around these singularities the function is not defined in the whole ring neighborhood $0 < |z - z_0| < R$ of z_0.

We first review some simple facts from complex analysis. Let us consider a particular example. Say, $w := f(z) = \sqrt{z}$. It is known that the function is multi-valued. In our case for $z = re^{i\theta}$ there are exactly two different values of w corresponding to z, namely, $w_1 = \sqrt{r}e^{i\theta/2}$ and $w_2 = \sqrt{r}e^{i\theta/2+\pi} = -w_1$. These two functions are known as *branches*. Observe the following phenomenon: when we start from a given point z and choose one of the branches, say w_1, then when traversing any closed contour we come back to w_1, *unless* we enclose the origin. If the contour encloses the origin, then w attains the value w_2 after one circle, and it comes back to w_1 after the second loop. We should conclude that the multi-valued function is well defined *except* at $z = 0$ where the function interchanges branches. Such a point is called a *branch point*. The function ceases to be analytic at this point, and hence it has a singularity (in fact, there is another explanation why $\sqrt{z}$ must have a singularity at $z_0 = 0$: its derivative does not exist in this point). We shall call it an *algebraic singularity*. To make $\sqrt{z}$ analytic, we must restrict the domain of definition to $\mathbb{C} - (-\infty, 0)$, that is, complex plane *cut* along the negative real axis (in fact, it can be cut along any ray emanating from the origin). This way we prevent z to encircle the origin and thus avoid ambiguity. As a consequence, we cannot expect to develop $\sqrt{z}$ in the Laurent series around $z = 0$, however, another expansion due to Puiseux will work, as we shall see soon. Naturally, $\sqrt{z}$ is not an exception. Functions like z^α for α real and not an integer, $\log z$, etc. belong to this category.

Let us now define more precisely a branch point and its order. The reader may review this material from any standard textbook on complex analysis (e.g., [159]). Consider an analytic function $A(z)$ that can be analytically continued along any path in the ring $0 < |z - z_0| < R$. If the function (or a branch of the function) returns to itself after its analytic continuation along any closed path, then A is a single-valued function and z_0 is a pole or an essential singularity. If, however, after a continuation along a closed path the function does not return to its original branch (as in the case of $\sqrt{z}$), then the function is multi-valued and z_0 is a branch point. To further characterize it, let $\mathcal{C}$ be a circle $|z - z_0| = r$ contained in the ring $0 < |z - z_0| < R$, and $p\mathcal{C}$ be a closed contour that encircles p times z_0 along $\mathcal{C}$. The point z_0 is called a branch point of order $p-1$, if the function returns to its original branch along $p\mathcal{C}$, but not along any closed path $m\mathcal{C}$ where $m < p$. For example, $(z-1)^{1/3}$ has a branch point at $z_0 = 1$ which is of order two; the function $\sqrt{z(z-1)}$ has two branch points at 0 and 1, each of order 1.

Although the Laurent expansion does not apply for a function around its algebraic singularity, another expansion due to Puiseux works fine (cf. [159]).

Theorem 8.10 (Puiseux 1850) *Let $A(z)$ be an analytic (multi-valued) function and z_0 its branch point of order $p-1$. Then, in the neighborhood of z_0, $A(z)$ can be represented as the following series*

$$A(z) = \sum_{k=-\infty}^{\infty} a_k (z - z_0)^{\frac{k}{p}} \tag{8.65}$$

that converges in a neighborhood of z_0.

The idea of the proof of Puiseux's expansion is quite simple. It suffices to substitute $z = z_0 + w^p$, and observe that $A(z)$ as a function of w is a singled-valued function defined in a ring $0 < |z - z_0| < R$, hence one can apply Laurent's expansion to it yielding

$$A(z_0 + w^p) = \sum_{k=-\infty}^{\infty} a_k w^k$$

which is the desired (8.65).

Puiseux's expansion suggests that for algebraic singularities, one must study asymptotics of functions like $(z - z_0)^{k/p}$. In general, in applications we often deal with functions like $(z - z_0)^{k/p} w(z)$ where $w(z)$ is either analytic in a neighborhood of z_0 or not (e.g., consider $A(z) = \frac{e^{-z-z^2/2}}{\sqrt{1-z}}$ and $\sqrt{1-z}\log\frac{1}{1-z}$). We next discuss in some depth how to obtain precise asymptotic expansions for such functions around its algebraic singularities. There are many techniques to deal with such singularities. However, in analysis of algorithms the *singularity analysis* of Flajolet and Odlyzko is the most useful, and we devote some time to it in the sequel.

Before we proceed with general discussions, we observe that because

$$[z^n]A(z) = \rho^n [z^n] A(z/\rho) \ ,$$

one only needs to study singularities at, say, $z = 1$. That's what we will do next. We shall follow Flajolet and Odlyzko [112]. The reader will also find an excellent account of

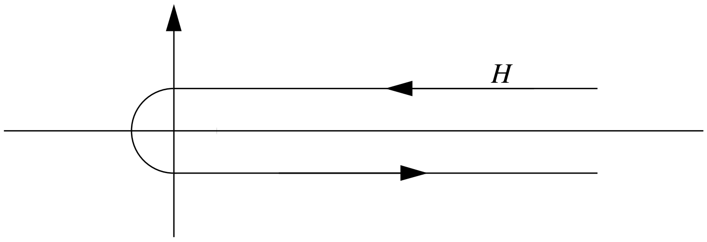


Figure 8.2: Hankel's Contour

it in Odlyzko [262] and Flajolet and Sedgewick [118]. Actually, because of these excellent references, we focus here on explaining things rather than on detailed proofs.

We start with obtaining asymptotic series for $(1-z)^{-\alpha}$ where α is real and $\alpha \notin \{0, -1, -2, \dots\}$. First of all, we observe that for $\alpha = k \in \{1, 2, \dots\}$ we have from Entry 1 of Table 7.2

$$(1-z)^{-k} = \sum_{n=0}^{\infty} \binom{n+k-1}{n} z^n.$$

It is then natural to expect that

$$\begin{aligned} [z^n](1-z)^{-\alpha} &= \binom{n+\alpha-1}{n} = \frac{\Gamma(n+\alpha)}{\Gamma(\alpha)\Gamma(n+1)} \\ &= \frac{n^{\alpha-1}}{\Gamma(\alpha)} \left(1 + \frac{\alpha(\alpha-1)}{2n} + O\left(\frac{1}{n^2}\right)\right) \end{aligned}$$

provided $\alpha \notin \{0, -1, -2, \dots\}$. We will prove it using Hankel's contour that we review in Table 8.3.

Theorem 8.12 (Flajolet and Odlyzko 1990) *Let $\alpha \notin \{0, -1, -2, \dots\}$. Then*

$$\begin{aligned} [z^n](1-z)^{-\alpha} &\sim \frac{n^{\alpha-1}}{\Gamma(\alpha)} \left(1 + \sum_{k=1}^{\infty} \frac{e_k(\alpha)}{n^k}\right) && (8.67) \\ &= \frac{n^{\alpha-1}}{\Gamma(\alpha)} \left(1 + \frac{\alpha(\alpha-1)}{2n} + \frac{\alpha(\alpha-1)(\alpha-2)(3\alpha-1)}{24n^2} + \cdots\right) \end{aligned}$$

where

$$e_k(\alpha) = \sum_{j=k}^{2k} c_{k,j}(\alpha-1)(\alpha-2)\cdots(\alpha-j)$$

and $c_{k,j}$ are coefficients in the following expansion

$$e^x(1+yt)^{-1-1/y} = \sum_{k,j} c_{k,j} x^j y^k. \qquad (8.68)$$

Table 8.3: HANKEL'S CONTOUR INTEGRAL FOR GAMMA FUNCTION

Let $\mathcal{H}$ be a contour starting at $+\infty$ in the upper half-plane, winding counterclockwise around the origin, and proceeding back towards $+\infty$ in the lower half-plane. (cf. Figure 8.2). We denote the integral along such loop as $\int_{\mathcal{H}}$ or $\int_{+\infty}^{(0)}$. The following theorem is a particular result for the Euler gamma function, but it plays an important role in the analysis of algebraic singularities.

Theorem 8.11 (Hankel's Contour Integral) *For all $s \in \mathbb{C}$*

$$\frac{1}{\Gamma(s)} = -\frac{1}{2\pi i}\int_{+\infty}^{(0)} (-w)^{-s} e^{-s} ds \tag{8.66}$$

where the function $(-w)^{-s}$ has its principal value, that is,

$$(-w)^{-s} = \exp\left(-s \ln r - si(\theta - \pi)\right)$$

with $w = re^{i\theta}$.

To see why (8.66) is true, let us specify Hankel's contour as follows (cf. Figure 8.2).

$$\mathcal{H} = \begin{cases} \mathcal{H}^+ & w = t + \delta i \quad t \geq 0 \\ \mathcal{H}^- & w = t - \delta i \quad t \geq 0 \\ \mathcal{H}^o & w = \delta e^{i\phi} \quad\ \ \phi \in [-\pi/2, \pi/2] \end{cases}$$

where $\delta \to 0$. Easy calculations show

$$\begin{aligned} \int_{\mathcal{H}^+} (-w)^{(-s)} e^{-w} ds &= e^{is\pi}\int_{\infty}^{\delta} e^{-t} t^{-s} dt, \\ \int_{\mathcal{H}^-} (-w)^{(-s)} e^{-w} ds &= e^{-is\pi}\int_{\delta}^{\infty} e^{-t} t^{-s} dt, \\ \int_{\mathcal{H}^o} (-w)^{(-s)} e^{-w} ds &= i\delta^{1-s}\int_{-\pi/2}^{\pi/2} \exp\left(-\delta^{i\phi} - si(\phi - \pi)\right) d\pi. \end{aligned}$$

The last integral converges to zero for $\delta \to 0$ as long as $\Re(1 - s) > 0$. Then, the first two integrals summing up to

$$-\frac{e^{-i\pi s} - e^{i\pi s}}{2\pi i}\int_0^{\infty} e^{-t} t^{-s} dt = \frac{\sin \pi s}{\pi}\Gamma(1 - s) = \frac{1}{\Gamma(s)}$$

where the last implication follows from the fundamental relationship (2.32) for the gamma function (cf. of Section 2.4.1). Thus, we prove (8.66) for $\Re(1-s) > 0$. By analytic continuation (since both sides of (8.66) are analytic) we can extend it to all $s \in \mathbb{C}$.
Observe that if we traverse the Hankel loop in the reverse order (i.e., from $+\infty$ in the lower plane, winding the origin clockwise and moving towards $+\infty$ in the upper plane) we must reverse the sign in (8.66).

Furthermore, for any β

$$[z^n](1-z)^{-\alpha}\left(\frac{1}{z}\log\frac{1}{1-z}\right)^{\beta} \sim \frac{n^{\alpha-1}}{\Gamma(\alpha)}(\log n)^{\beta}\left(1+\sum_{k=1}^{\infty}\frac{C_k(\alpha,\beta)}{(\log n)^k}\right) \tag{8.69}$$

where

$$C_k(\alpha,\beta)=\binom{\beta}{k}\Gamma(\alpha)\frac{d^k}{ds^k}\frac{1}{\Gamma(s)}\Big|_{s=\alpha}.$$

Proof. We give a detailed proof of (8.67) and only a sketch of (8.69). By Cauchy's formula

$$[z^n](1-z)^{-\alpha}=\frac{1}{2\pi i}\int_{\mathcal{C}}(1-z)^{-\alpha}\frac{dz}{z^{n+1}}$$

where $\mathcal{C}$ is a small contour encircling the origin (e.g., $\mathcal{C}=\{z: \ |z|=\frac{1}{2}\}$). We deform it in such a way that the new contour, say $\mathcal{C}_1$ consists of a large circle of radius R with a notch that comes back as close to $z=1$ as possible (cf. Figure 8.3). By Cauchy's bound the integral on the large circle can be bounded by $O(R^{-n-\alpha})$, and we are left with estimating the integral along the notch which resembles the Hankel contour except that it is traversed in the opposite direction (cf. Table 8.3). We denote it by $\bar{\mathcal{H}}$ and specify as follows

$$\bar{\mathcal{H}}=\begin{cases}\mathcal{H}^- & w=t-\frac{i}{n} & t\ge 0\\ \mathcal{H}^+ & w=t+\frac{i}{n} & t\ge 0\\ \mathcal{H}^o & w=1-\frac{e^{i\phi}}{n} & \phi\in[-\pi/2,\pi/2]\end{cases}$$

Then, after a change of variable $z=1+t/n$ we arrive at

$$[z^n](1-z)^{-\alpha}=\frac{n^{\alpha-1}}{2\pi i}\int_{\bar{\mathcal{H}}}(-t)^{-\alpha}\left(1+\frac{t}{n}\right)^{-n-1}dt+O(R^{-n-\alpha})$$

We now proceed formally, omitting $O(R^{-n-\alpha})$, and later justify it. Using (8.68) we have

$$e^t\left(1+\frac{t}{n}\right)^{-1-n}=\sum_{k,j}c_{k,j}t^jn^{-k} \tag{8.70}$$

which yields

$$\begin{aligned}[z^n](1-z)^{-\alpha} &= \frac{n^{\alpha-1}}{2\pi i}\int_{\bar{\mathcal{H}}}(-t)^{-\alpha}e^{-t}\sum_{k,j}c_{k,j}t^jn^{-k}dt\\ &= n^{\alpha-1}\sum_{k,j}c_{k,j}n^{-k}\frac{1}{2\pi i}\int_{\bar{\mathcal{H}}}(-t)^{-\alpha+j}e^{-t}dt\\ &\overset{(8.66)}{=} n^{\alpha-1}\sum_{k,j}c_{k,j}n^{-k}\frac{1}{\Gamma(\alpha-j)}\\ &= \frac{n^{\alpha-1}}{\Gamma(\alpha)}\sum_{k,j}c_{k,j}\frac{(\alpha-1)(\alpha-2)\cdots(\alpha-j)}{n^k}\end{aligned}$$

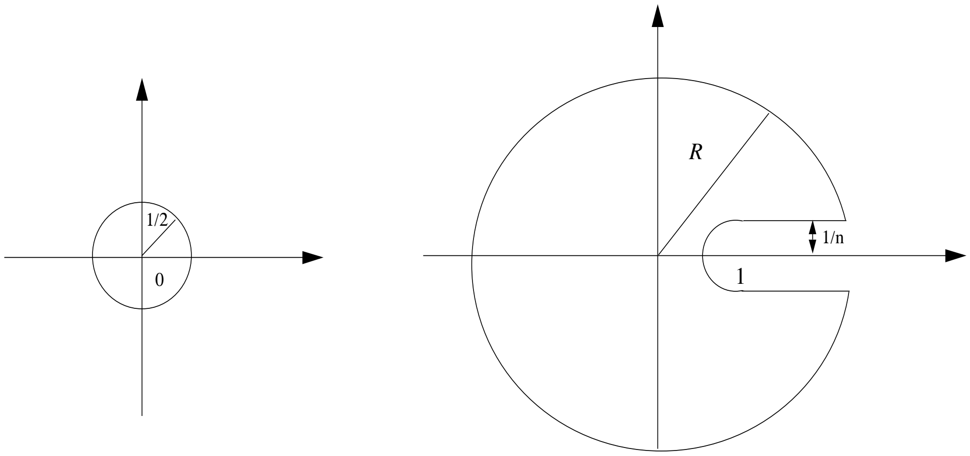


Figure 8.3: Contours Used in Theorem 8.12.

which formally establishes our result. To justify it rigorously, we must prove convergence of above series in order to interchange the integral and the sum.

But, actually, this is not difficult at all. All we need to do it to split the integral according to, say $|t| \le \log^2 n$ and $|t| \ge \log^2 n$. This is necessary in order to justify (8.70), that is, to make t small. The part corresponding to $|t| \ge \log^2 n$ is negligible since

$$\left(1+\frac{t}{n}\right)^{-n} = O(\exp(-\log^2 n)).$$

On the remaining part of the contour, that is, for $|t| \le \log^2 n$, the term t/n is small and (8.70) is applicable. This completes the proof of (8.67).

The proof of (8.69) is similar. We only sketch it. The basic expansion is

$$\begin{aligned} z^{-n-1}(1-z)^{-\alpha}\left(\frac{1}{z}\log\frac{1}{1-z}\right)^{\beta} z^{-n-1}\Big|_{z=1+t/n} &\sim e^{-t}\left(\frac{n}{-t}\right)^{\alpha}\log^{\beta}\left(\frac{-n}{t}\right) \\ = \frac{e^{-t}(-t)^{\alpha}}{n^{\alpha}}\log^{\beta} n\left(1-\frac{\log(-t)}{\log n}\right)^{\beta} \\ = \frac{e^{-t}(-t)^{\alpha}}{n^{\alpha}}\log^{\beta} n\left(1-\beta\frac{\log(-t)}{\log n}\frac{\beta(\beta-1)}{2}\left(\frac{\log(-t)}{\log n}\right)^{2}+\cdots\right). \end{aligned}$$

By considering $|t| \le \log^2 n$ and $|t| \ge \log^2 n$, we can justify the expansion. The last observation is that we end up with the Hankel integrals of the form

$$\frac{1}{2\pi i}\int_{\bar{\mathcal{H}}}(-t)^{-(\alpha-j)}e^{-t}\log^k(-t)dt$$

which reduces to the derivatives of $1/\Gamma(\alpha-j)$. This completes the sketch of the proof for (8.69). For details the reader is referred to [112]). ■

The most important aspect of the singularity theory comes next: In the analysis of algorithms, analytic combinatorics, information theory, and many engineering problems, we rarely have an explicit expression for the generating function $A(z)$. We are lucky if we get an expansion of $A(z)$ around its singularities. For example, we are often left with $A(z) \sim (1-z)^{-\alpha} = (1-z)^{-\alpha} + o((1-z)^{-\alpha})$. In order to evaluate a_n as $n \to \infty$ we need a "transfer theorem" that will allow us to pass to coefficients of $o((1-z)^{-\alpha})$ under the "small oh" notation. These transfer theorems are jewels of Flajolet and Odlyzko singularity analysis [112], and we discuss them next. However, in order not to clutter too much our expositions, we consider a particular example adopted from Odlyzko [262].

Let $A(z) \sim \sqrt{1-z} = \sqrt{1-z} + f(z)$ where $f(z) = o(\sqrt{1-z})$. From Theorem 8.12 we conclude that

$$[z^n](1-z)^{\frac{1}{2}} = -\frac{1}{\sqrt{\pi}n^{3/2}}\left(\frac{1}{2} + \frac{3}{16n} + O(n^{-2})\right).$$

Hence, we would like to prove that

$$[z^n]f(z) = [z^n]o(\sqrt{1-z}) = o(n^{-3/2}) \tag{8.71}$$

as $n \to \infty$. The analyticity of $f(z)$ in a region around $z_0 = 1$ is crucial for obtaining any further information. Let us consider a few different scenarios:

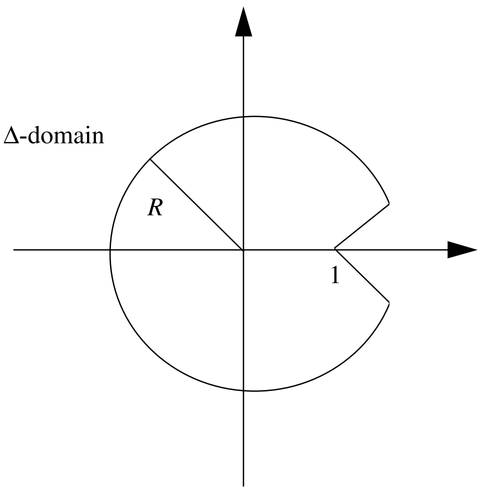
- Assume $f(z)$ is analytic in a larger disk, say, of radius $1+\delta$ for some $\delta > 0$. Then by Cauchy's bound we immediately conclude that $[z^n]f(z) = O((1+\delta')^{-n})$ for $0 < \delta' < \delta$, a conclusion that is much stronger than the desired (8.71). Furthermore, this assumption will not allow us to not handle expression like $\sqrt{1-z}\log\frac{1}{1-z}$ that often arises in analysis of algorithms.

- If we only know that $f(z) = o(\sqrt{1-z})$ in $|z| \leq 1$, then we may only conclude that $[z^n]f(z) = O(1)$. This is true since $|f(z)| \leq C$ in $|z| < 1$, where C is a constant, and for all $r < 1$ and $n \geq 0$ we know that by Cauchy's bound $[z^n]f(z) \leq Cr^{-n}$. But this only suggests that $f(z) = O(1)$, a result that is far from what is required. However, we if would know some smoothness condition for the function f on the circle $|z| = 1$, then some further progress can be achieved. This is due to Darboux and we briefly discuss it at the end of this section.

We now consider the third scenario analyzed by Flajolet and Odlyzko [112]. Let us define a domain around $z_0 = 1$ as follows

$$\Delta(R, \phi) = \{z : \ |z| < R, \ z \neq 1, \ |\arg(z-1)| > \phi\} \tag{8.72}$$

for some $R > 1$ and $0 < \phi < \pi/2$. The domain Δ is an extended disk around $z = 1$ with a circular part rooted at $z = 1$ deleted, as shown in Figure 8.4. In our case, we assume that



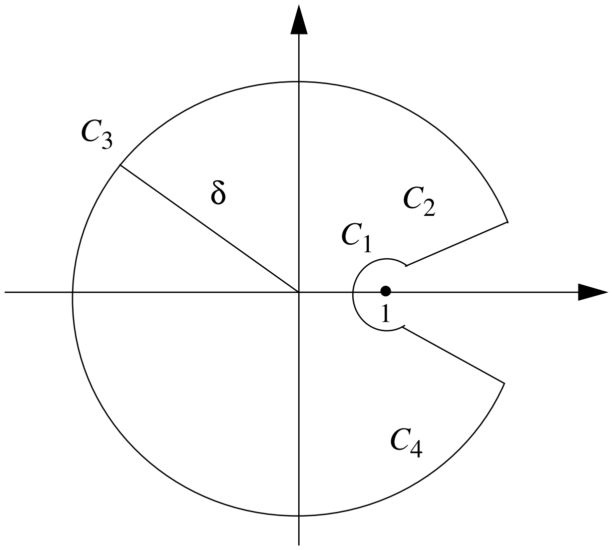


Figure 8.4: The Δ domain and the contour used in the singularity analysis.

$f(z)$ can be analytically continued to $\Delta(R, \pi/4)$. We apply Cauchy's formula to extract the coefficients of $f(z)$, and as the contour $\mathcal{C}$ we choose (cf. Figure 8.4)

$$\begin{aligned}
\mathcal{C}_1 &= \{z: \ |z-1| = 1/n, \ |\arg(z-1) \geq \pi/4\}, \\
\mathcal{C}_2 &= \{z: \ z = 1 + re^{i\pi/4}, \ 1/n \leq r \leq \delta\}, \\
\mathcal{C}_3 &= \{z: \ |z| = |1+\delta e^{i\pi/4}|, \ |\arg(z-1) \geq \pi/4\}, \\
\mathcal{C}_4 &= \{z: \ z = 1 + re^{-i\pi/4}, \ 1/n \leq r \leq \delta\}
\end{aligned}$$

where $0 < \delta < 1/2$. We evaluate the following four integrals for $j = 1, 2, 3, 4$

$$f_j = \frac{1}{2\pi i} \int_{\mathcal{C}_j} f(z) \frac{dz}{z^{n+1}}.$$

The easiest to handle is f_3, since $f(z)$ is bounded on $\mathcal{C}_3$ (outer large circle). By Cauchy's bound

$$f_3 = O((1+\delta)^{-n}).$$

On $\mathcal{C}_1$ (inner small circle), we have $|f(z)| = o(n^{-1/2})$ and the length of the circle is $2\pi/n$, so that

$$f_1 = o(n^{-3/2})$$

as $n \to \infty$. Finally, we are left with $\mathcal{C}_2$ and $\mathcal{C}_4$ which can be handled in a similar manner. Let us consider $\mathcal{C}_2$. Observe that for $z = 1 + re^{i\pi/4}$

$$|z^{-n}| = |1 + r\cos\pi/4 + ir\sin\pi/4|^{-n}$$

$$\leq \quad (1+r)^{-n/2} \leq \exp(-nr/10)$$

for $0 \leq r < 1$. Since $f(z) = o(\sqrt{1-z})$, for any $\varepsilon > 0$ we have

$$|f(1+re^{i\pi/4})| \leq \varepsilon\sqrt{r}.$$

This yields, after the change $nr = z$,

$$|f_2| \leq \varepsilon \int_0^\delta \sqrt{r}\exp(-nr/10)dr \leq \varepsilon n^{-3/2} \int_0^\delta \sqrt{z}\exp(-z/10)dz = o(n^{-3/2})$$

as desired. In a similar manner we show that $f_4 = o(n^{-3/2})$. Thus we prove that if $f(z)$ is expendable to the Δ-domain, then $f_n := [z^n]f(z) = o(n^{-3/2})$. We should observe that the critical derivation was the one on the rectilinear curves $\mathcal{C}_2$ and $\mathcal{C}_4$.

The method just described extends to "Big Oh" notation as long as one can analytically continue a function to the $\Delta(R,\phi)$ domain. For example, let $f(z) = O((1-z)^{-\alpha})$. Considering the contour $\mathcal{C}$ as above, and bounding the integral on $\mathcal{C}_1$ and $\mathcal{C}_3$ through Cauchy's bound, we are left with estimating

$$|f_2| \leq \frac{1}{2\pi n}\int_1^\infty K\left(\frac{t}{n}\right)^{-\alpha} |1+e^{i\phi}t/n|^{-n-1}dt$$

where K is the constant from the "Big Oh" notation, that is, such that $|f(z)| \leq K(1-z)^{-\alpha}$. Since $|1+e^{i\phi}t/n| \geq 1 + \Re(e^{i\phi}t/n) = 1 + n^{-1}t\cos\phi$ we finally obtain

$$|f_2| \leq \frac{K}{2\pi} J_n n^{\alpha-1},$$

where $J_n = \int_1^\infty t^{-\alpha}(1+t\frac{1}{n}\cos\phi)^{-n}dt$ us bounded for $0 < \phi < \pi/2$. We just established one of the transfer theorem of Flajolet and Odlyzko.

In summary, the singularity analysis of Flajolet and Odlyzko is based on the following transfer theorems. The proof follows the footsteps of the above derivations, and details can be found in [112].

Theorem 8.13 (Flajolet and Odlyzko 1990) *Let $A(z)$ be Δ-analytical where $\Delta(R,\phi)$ is defined (8.72). If $A(z)$ satisfies in $\Delta(R,\pi)$ either*

$$A(z) = O\left((1-z)^{-\alpha}\log^\beta(1-z)^{-1}\right)$$

or

$$A(z) = o\left((1-z)^{-\alpha}\log^\beta(1-z)^{-1}\right),$$

then either

$$[z^n] = O\left(n^{\alpha-1}\log^\beta n\right)$$

or

$$[z^n] = o\left(n^{\alpha-1}\log^\beta n\right),$$

respectively.

A classical example of singularity analysis is the Flajolet and Odlyzko [110] analysis of the height of binary trees. However, we finish this subsection with a simpler application that quite well illustrates the theory.

Example 8.10: *Certain Sums from Coding Theory*
In coding theory the following sum is of interest (cf. [211])

$$S_n = \sum_{i=0}^{n} \binom{n}{i} \left(\frac{i}{n}\right)^i \left(1 - \frac{i}{n}\right)^{n-i}$$

Let $s_n = n^n S_n$ and $S(z)$ be the exponential generating function of s_n. In addition, let

$$B(z) = \frac{1}{1 - T(z)}$$

where $T(z)$ is the "tree function" defined in Example 15 (cf. (7.30) and (7.31)). Since $[z^n]T(z) = n^{n-1}/n!$, hence $[z^n]B(z) = n^n/n!$. Thus by (7.16) of Table 7.3 we find that

$$S(z) = (B(z))^2.$$

From Example 15 we also know that $T(z)$ is singular at $z = e^{-1}$. Using MAPLE we find (see also [58] for a more detailed analysis)

$$T(z) - 1 = \sqrt{2(1 - ez)} + \frac{2}{3}(1 - ez) + \frac{11\sqrt{2}}{36}(1 - ez)^{3/2} + \frac{43}{135}(1 - ez)^2 + O((1 - ez)^{5/2}).$$

in a Δ domain around $z = e^{-1}$. Indeed, expanding $z = ye^{-y}$ in a power series around $z = e^{-1}$ we have

$$z = e^{-1} - \frac{e^{-1}}{2}(y-1)^2 + \frac{e^{-1}}{3}(y-1)^3 + O((y-1)^4$$

and after solving it for y we obtain the above expansion for $T(z)$. Now let $h(z) = \sqrt{1 - ez}$. Then

$$\begin{aligned} S(z) &= \frac{1}{2h(z)\left(1 + \frac{\sqrt{2}}{3}\sqrt{h(z)} + \frac{11}{36}h(z) + O(h^{3/2}(z)\right)^2} \\ &= \frac{1}{2(1-ez)} + \frac{\sqrt{2}}{3\sqrt{(1-ez)}} + \frac{1}{36} + \frac{\sqrt{2}}{540}\sqrt{1-ez} + O(1-ez)\ . \end{aligned}$$

An application of the singularity analysis (i.e., Theorem 8.12 and Theorem 8.13) yields

$$S_n = \sqrt{\frac{n\pi}{2}} + \frac{2}{3} + \frac{\sqrt{2\pi}}{24}\frac{1}{\sqrt{n}} - \frac{4}{135}\frac{1}{n} + O(1/n^{3/2})$$

as shown in [326], from where this example is taken. We shall discuss another example in the application Section 8.7.2. □

What if $A(z)$ cannot be analytically continued to a Δ domain? We still have some options. One is real analysis, another is Tauberian theorem. However, the most successful seems to be the Darboux method that we briefly discuss next.

We start with a quick review of the *Riemann-Lebesgue lemma* that is often useful, as an intermediate result, in establishing asymptotics.

Lemma 8.14 *Let $f(t)$ be continuous in $(0, \infty)$. Then*

$$\lim_{x\to\infty} \int_0^\infty f(t)e^{ixt}dt = 0 \tag{8.73}$$

provided the integral converges uniformly at 0 and ∞ for all sufficiently large x.

The proof of the Riemann-Lebesgue lemma can be found in [263, 355]. We just observe that it follows from a few simple facts. First of all, by the hypothesis, there are positive a and b such that

$$\left|\int_0^a f(t)e^{ixt}dt\right| \le \frac{\varepsilon}{3}, \qquad \left|\int_b^\infty f(t)e^{ixt}dt\right| \le \frac{\varepsilon}{3}$$

for any $\varepsilon > 0$. In addition, since $f(t)$ is continuous, it is bounded in $[a, b]$. Approximating f by its value at a partition $a = t_0 < t_1 < \ldots < t_n = b$ (e.g., $|f(t) - f(t_j)| \le \varepsilon/(6(b-a))$) we arrive at

$$\int_a^b f(t)e^{ixt}dt = \sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}} |f(t) - f(t_j)|e^{ixt}dt + \sum_{j=0}^{n-1} f(t_j)\int_{t_j}^{t_{j+1}} e^{itx}dt.$$

But the first sum is bounded by, say $\varepsilon/6$, while the second is bounded by $2n\max\{f\}/x < \varepsilon/6$, hence can be made as small as desired for large x, which completes the sketch of the proof.

Now, we present the Darboux method. In this case, we only assume that $A(z)$ is smooth enough on the disk of convergence.

Theorem 8.15 (Darboux 1878) *Assume that $A(z)$ is continuous in the closed disk $|z| \le 1$, and in addition k times continuously differentiable on $|z| = 1$. Then*

$$[z^n]A(z) = o(n^{-k}) \tag{8.74}$$

Proof. By the continuity assumption on $|z| = 1$ we have

$$\begin{aligned} a_n = [z^n]A(z) &= \frac{1}{2\pi i}\int_{|z|=1} A(z)\frac{dz}{z^{n+1}} \\ &\overset{z=e^{i\theta}}{=} \frac{1}{2\pi}\int_0^{2\pi} A(e^{i\theta})e^{-ni\theta}d\theta. \end{aligned}$$

By the Riemann-Lebesgue lemma we must conclude that $a_n = o(1)$ as $n \to \infty$. This proves the result for $k = 0$. To get a stronger result we integrate the above by parts k times to see that

$$[z^n]A(z) = \frac{1}{2\pi(in)^k} \int_0^{2\pi} A^{(k)}(e^{i\theta})e^{-ni\theta} d\theta$$

where $A^{(k)}$ is the kth derivative of A. Again, the proof follows from the Riemann-Lebesgue lemma. ■

As an example, consider $A(z) = v(z)/\sqrt{1-z}$ where $v(z)$ is analytic at $z = 1$. Then

$$A(z) = \frac{v(1)}{\sqrt{1-z}} - v'(1)\sqrt{1-z} + O((1-z)^{3/2}).$$

Since $(1-z)^{3/2}$ is once continuously differentiable on $|z| \leq 1$, we can apply Darboux's Theorem with $k = 1$ to show that $[z^n]O((1-z)^{3/2}) = o(1/n)$.

Finally, we should mention that so far we only dealt with a single singularity. Multiple singularities can be handled in a similar fashion, however, some complications arise. The reader is referred to Flajolet and Sedgewick [118] for details.

8.4 Large Singularities: Saddle Point Method

This section presents asymptotic methods for generating functions whose dominant singularities are large. We recall that, roughly speaking, a singularity is large if the growth of the function near such a singularity is rapid, say, larger than any polynomial. We shall see that methods of previous section that work fine for small singularities are useless for large singularities. To focus, we concentrate on asymptotics of the Cauchy integral and hence coefficients of some generating functions. In short, we shall investigate integrals like

$$I(\lambda) = \int_{\mathcal{C}} f(z)e^{-\lambda h(z)} dz \tag{8.75}$$

where $f(z)$ and $h(z)$ are analytic functions, and λ is a large positive parameter (often we assume that $\lambda = n$ with n being a natural number and the coefficients of f and h are nonnegative).

We start with the Laplace method which assumes that the path $\mathcal{C}$ in (8.75) is an interval $[a, b]$, and $f(z)$ and $h(z)$ are real-valued functions. Then, we discuss the *saddle point method* (also known as the *steepest descent method*) which by far is the most useful and popular technique to extract asymptotic information of rapidly growing functions. It is one of the most complicated method, therefore, given the purpose and limitations, we do not present a full discussion of it. The reader is referred to [67, 263, 355] for a complete and insightful presentations. We, however, describe here a very useful (and largely forgotten) method of

Van der Waerden [349], and well known method of Hayman [155]. We complete this section with an application of the saddle point method to large deviations. The reader is also referred to excellent survey by Odlyzko [262] and the book by Flajolet and Sedgewick [118].

8.4.1 Laplace Method

Let us consider a special case of (8.75), namely

$$I(x) = \int_a^b f(t)e^{-xh(t)}dt \tag{8.76}$$

where $-\infty \le a < b \le \infty$ and x is a large positive parameter. Furthermore, let the minimum of $h(x)$ be attained at the point $a < t_0 < b$. Assume that $f(t_0) \neq 0$ and $h''(t_0) > 0$. Observe then that as $x \to \infty$

$$f(t)\exp\left((-x(h(t) - h(t_0))\right) \to f(t_0)\delta(t - t_0)$$

where $\delta(t)$ is the Dirac-delta function taking the value ∞ at $t = 0$ and zero otherwise. In view of this, we can approximate $I(x)$ by

$$\begin{aligned} I(x) &\sim \int_a^b f(t_0)\exp\left(-x(h(t_0) + \frac{1}{2}(t-t_0)^2h''(t_0)\right) dt \\ &\sim f(t_0)e^{-xh(t_0)} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}x(t-t_0)^2h''(t_0)\right) dt \\ &= f(t_0)e^{-xh(t_0)}\sqrt{\frac{2\pi}{xh''(t_0)}}. \end{aligned}$$

In above we used twice the fact that the function under the integral is concentrated around t_0. While above is only a heuristic, we will make it rigorous soon. We should add that the last line of the above derivation follows from the Gauss integral. The following lemma is often used in this book.

Lemma 8.16 (Gauss Integral) *The following identities hold:*

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} x^k e^{-tx^2}dx = \begin{cases} 0 & k = 1, 3, 5, \dots \\ \frac{t^{-1/2-k/2}k!}{(k/2)!2^{k+1/2}} & k = 2, 4, 6, \dots \end{cases} \tag{8.77}$$

and for fixed k

$$\int_\theta^\infty x^k e^{-\alpha x^2}dx = O\left(e^{-\alpha\theta^2}\right) \tag{8.78}$$

for any $\alpha > 0$.

The first identity above is nothing but moments of the Gaussian distribution. For the latter observe that

$$\int_\theta^\infty x^k e^{-\alpha x^2} dx = e^{-\alpha\theta^2} \int_\theta^\infty x^k e^{-\alpha(x^2-\theta^2)} dx = O\left(e^{-\alpha\theta^2}\right)$$

since the integral converges.

On the other hand, if the minimum of $h(x)$ is achieved at the end of the interval and $f(a) \neq 0$, say at $t = a$, then we obtain

$$\begin{aligned} I(x) &\sim \int_a^b f(a) \exp\left(-x(h(a) + (t-a)h'(a)\right) dt \\ &\sim f(a)e^{-xh(a)} \int_a^b \exp\left(-x(t-a)h'(a)\right) dt \\ &= \frac{f(a)e^{-xh(a)}}{xh'(a)}. \end{aligned} \tag{8.79}$$

Finally, before we provide a rigorous proof of the Laplace's approach, let us say a word about the *principle of stationary phase* which applies to integrals of the form

$$F(x) = \int_a^b f(t) e^{ixh(t)} dt$$

where x is assumed to be large and positive. The underlying principle, due to Lord Kelvin, is the assertion that the major asymptotic contribution comes from points where the *phase of* $h(t)$ *is stationary*, that is, where $h'(t)$ vanishes. We should add that oscillations near such points slow down. Heuristically, if $t_0 \in (a, b)$ is such that $h'(t_0) = 0$, then

$$\begin{aligned} F(x) &\sim \int_{t_0-\varepsilon}^{t_0+\varepsilon} f(t) e^{ixh(t)} dt, \qquad \varepsilon > 0, \\ &\sim f(t_0) e^{ixh(t_0)} \int_{-\infty}^{\infty} e^{ixh''(t_0)(t-t_0)^2/2} dt \\ &= f(t_0) \sqrt{\frac{2\pi}{x|h''(t_0)|}} e^{i(xh(t_0) + \frac{\pi}{4}\mathrm{sgn}\ h''(t_0))} \end{aligned}$$

where $\mathrm{sgn}(x) = 1$ for $x > 0$ and $\mathrm{sgn}(x) = -1$ for $x < 0$. The last equation above follows from the Fresnel integral (cf. [2])

$$\int_{-\infty}^{\infty} e^{\frac{1}{2}it^2} dt = \sqrt{2\pi} e^{i\pi/4}.$$

As we mention before it is not difficult to make above analyses completely rigorous. We now focus on the Laplace method and prove the following theorem.

Theorem 8.17 (Laplace's Method) *Let*

$$I(x) = \int_a^b f(t) e^{-xh(t)} dt$$

where $-\infty \le a < b \le \infty$ *and* x *is a large positive parameter. Assume that*

(i) $h(t)$ *has only one minimum inside* (a, b) *at point* $t = t_0$*, that is,* $h'(t_0) = 0$ *and* $h''(t_0) > 0$.

(ii) $h(t)$ *and* $f(t)$ *are continuously differentiable in a neighborhood of* t_0 *with* $f(t_0) \ne 0$ *and*

$$h(t) = h(t_0) + \frac{1}{2} h''(t_0)(t - t_0)^2 + O((t - t_0)^3);$$

(iii) The integral $I(x)$ *exists for sufficiently large* x*, say for* $x \ge \xi$.

Then, the following holds

$$\int_a^b f(t) e^{-xh(t)} dt = f(t_0) \sqrt{\frac{2\pi}{xh''(t_0)}} e^{-xh(t_0)} \left(1 + O\left(\frac{1}{\sqrt{x}}\right)\right) \tag{8.80}$$

for $x \to \infty$.

Proof. We follow Lauwerier [231]. Without loss of generating we take $t_0 = 0$ (hence $a < 0$ and $b > 0$) and $h(t_0) = 0$ (otherwise substitute $t' = t - t_0$). We first assume $f(t) \equiv 1$ to spell out succinctly the main idea. From (ii) we conclude that there exist constants c and δ such that

$$|h(t) - \frac{1}{2} h''(0) t^2| \le c|t|^3, \qquad |t| < \delta. \tag{8.81}$$

We split the integral into two parts, namely

$$I(x) = \int_{-\delta}^{\delta} e^{-xh(t)} dt + \int_{t \notin [-\delta, \delta]} e^{-xh(t)} dt.$$

The latter integral can be evaluated as follows: By (i), $\mu(\delta) = \inf_{|t| > \delta} h(t) > 0$ and hence

$$\begin{aligned} \left| \int_{t \notin [-\delta, \delta]} e^{-xh(t)} dt \right| &\le \int_{t \notin [-\delta, \delta]} e^{-\xi h(t) - (x - \xi) h(t)} dt \\ &\le e^{-(x - \xi)\mu(\delta)} \int_{t \notin [-\delta, \delta]} e^{-\xi h(t)} dt = O(e^{-x\beta}) \end{aligned}$$

for some $\beta > 0$, where the last assertion follows from (iii). Let $\omega = h''(t_0)$. Now, we deal with the first integral

$$\begin{aligned} I_1(x) &= \int_{-\delta}^{\delta} e^{-xh(t)} dt \\ &= \int_{-\delta}^{\delta} e^{-\frac{1}{2}x\omega t^2} dt + \int_{-\delta}^{\delta} e^{-\frac{1}{2}x\omega t^2} \left[\exp\left(-x(h(t) - \frac{1}{2}\omega t^2)\right) - 1\right] dt \end{aligned}$$

From Lemma 8.16 we estimate the first integral as

$$\int_{-\delta}^{\delta} e^{-\frac{1}{2}x\omega t^2} dt = \sqrt{\frac{2\pi}{h''(0)x}} + O(e^{-x\delta^2}).$$

With the help of

$$|e^t - 1| < |t|e^{|t|}$$

and using (8.81) several times, we evaluate the second integral in $I_1(x)$ as

$$\begin{aligned} \left|\int_{-\delta}^{\delta} e^{-\frac{1}{2}x\omega t^2} \left[\exp\left(-x(h(t) - \frac{1}{2}\omega t^2)\right) - 1\right] dt\right| &\le 2cx \int_0^{\delta} t^3 \exp\left(-\left(\frac{1}{2}\omega x t^2 - cxt^3\right)\right) dt \\ &\le 2cx \int_0^{\delta} t^3 \exp\left(-xt^2\left(\frac{1}{2}\omega - \delta c\right)\right) dt \\ &\le 2cx \int_0^{\delta} t^3 e^{-xbt^2} dt \le \frac{2c}{b^2 x} \int_0^{\infty} u^3 e^{-u^2} du \\ &= O(x^{-1}) \end{aligned}$$

where $b = \frac{1}{2}\omega - \delta c > 0$ for sufficiently small δ. Combining everything we obtain (8.80) for $f(t) \equiv 1$. To extend to general f, we just follow the steps of the above proof using $f(t) = f(t_0) + O(t - t_0)$. This completes the proof. ■

Laplace's method has many applications in the analysis of algorithms. We illustrate it on the evaluation of a certain integral arising the the analysis of b-tries with large b.

Example 8.11 *Height in b-Tries for Large b*

In Example 5 we studied the height of a b-tries. In particular, the height distribution $h_n^k = \Pr\{H_n \le k\}$ is given by (8.43) where the following truncated exponential is involved

$$1 + z2^{-k} + \cdots + \frac{z^b 2^{-kb}}{b!} = e^{z2^{-k}} \int_{z2^{-k}}^{\infty} e^{-w} \frac{w^b}{b!} dw = e^{z2^{-k}} \left[1 - \int_0^{z2^{-k}} e^{-w} \frac{w^b}{b!} dw\right].$$

It proves useful to have the asymptotic behavior of the above integral, and this is summarized in the lemma below.

Lemma 8.18 *We let*

$$I = I(A,b) = \frac{1}{b!}\int_0^A e^{b\log w - w}dw = \frac{e^{-b}b^{b+1}}{b!}\int_0^{A/b} e^{b(\log u - u+1)}du.$$

Then the asymptotic expansions of I are as follows:
(i) $b, A \to \infty$, $\alpha = b/A > 1$

$$I = e^{-A}\frac{A^b}{b!}\left[\frac{1}{b/A-1} - \frac{b}{A^2}\frac{1}{(b/A-1)^3} + O(A^{-2})\right].$$

(ii) $b, A \to \infty$, $b/A < 1$

$$I = 1 - e^{-A}\frac{A^b}{b!}\left[\frac{1}{1-b/A} - \frac{b}{A^2}\frac{1}{(1-b/A)^3} + O(A^{-2})\right].$$

(iii) $b, A \to \infty$, $A - b = \sqrt{b}B$, $B = O(1)$

$$\begin{aligned} I &= \frac{1}{\sqrt{2\pi}}\left(\int_{-\infty}^B e^{-x^2/2}dx - \frac{1}{3\sqrt{b}}(B^2+2)e^{-B^2/2}\right.\\ &+ \left.\frac{1}{b}\left(-\frac{B^5}{18} - \frac{B^3}{36} - \frac{B}{12}\right)e^{-B^2/2} + O(b^{-3/2})\right). \end{aligned}$$

Proof. To establish our result we note that I is a Laplace-type integral (8.80). Setting $h(u) = \log u - u + 1$ we see that h is maximal at $u = 1$. For $A/b < 1$ we have $h'(u) > 0$ for $0 < u \le A/b$ and thus the major contribution to the integral comes from the upper endpoint (more precisely, from $u = A/b - O(b^{-1})$). Then (8.79) yields part (i) of the Lemma 8.18. If $A/b > 1$ we write $\int_0^{A/b}(\cdots) = \int_0^\infty(\cdots) - \int_{A/b}^\infty(\cdots)$, evaluate the first integral exactly and use Laplace's method on the second integral. Now $h'(u) < 0$ for $u \ge A/b$ and the major contribution to the second integral is from the lower endpoint. Obtaining the leading two terms leads to (ii) in the Lemma 8.18.

To derive part (iii), we scale $A - b = \sqrt{b}B$ to see that the main contribution will come from $u - 1 = O(b^{-1/2})$. We thus set $u = 1 + x/\sqrt{b}$ and obtain

$$\begin{aligned} I &= \frac{e^{-b}b^{b+1}}{b!}\int_{-\sqrt{b}}^B \exp\left(b\left[\log\left(1+\frac{x}{\sqrt{b}}\right) - \frac{x}{\sqrt{b}}\right]\right)\frac{dx}{\sqrt{b}} \\ &= \frac{b^b\sqrt{b}e^{-b}}{b!}\int_{-\infty}^B e^{-x^2/2}\left[1 + \frac{x^3}{3\sqrt{b}} + \frac{1}{b}\left(-\frac{x^4}{4} + \frac{x^6}{18}\right) + O(b^{-3/2})\right]dx. \end{aligned} \tag{8.82}$$

Evaluating explicitly above integrals and using Stirling's formula in the form $b! = \sqrt{2\pi b}b^b e^{-b}(1 + (12b)^{-1} + O(b^{-2}))$, we obtain part (iii) of Lemma 8.18. This example is adopted from Knessl and Szpankowski [217]. $\square$

8.4.2 Method of Steepest Descent

We shall analyze here again (8.75), that is,

$$I(n) = \int_{\mathcal{C}} f(z)e^{-nh(z)}dz \tag{8.83}$$

for n large, where $f(z)$ and $h(z)$ are *analytic* functions and $\mathcal{C}$ is a path in complex plane. As always, we can deform $\mathcal{C}$ into another contour $\mathcal{C}'$ as long as the region between $\mathcal{C}$ and $\mathcal{C}'$ does not contain singularities of f or h. The idea is to find such a path $\mathcal{C}'$ that transforms the complex integral $I(n)$ into the real case considered in the previous section, that is, to reduce $I(n)$ to the one manageable by the Laplace method. In general, this is not an easy task and requires quite a bit of experience. Given the length of this chapter, we restrict ourselves to a short account on the general method explaining the underlying principle, and later focus only on the Cauchy's coefficient formula that is easier to handle than the general case.

Let us start with a general case. We assume that $f(z) \equiv 1$, and $h(z) = u(x,y) + iv(x,y)$ with $z = x + iy$. Our goal is to find such a path $\mathcal{C}'$ that

(i) the imaginary part of $f(z)$ on $\mathcal{C}'$ is constant.

This condition allows us to reduce the complex integral to a the real case that we know how to handle. Indeed, assuming $z(\tau)$ is a parametric description of the path $\mathcal{C}'$, and z_0 is a fixed point on the path (hence, $v(x,y) = \text{constant} = \Im f(z_0)$, the integral $I(n)$ becomes

$$\begin{aligned} I(n) &= e^{-nh(z_0)} \int_a^b \exp\left[-n(h(z(\tau)) - h(z_0))\right] \frac{dz(\tau)}{d\tau} d\tau \\ &= e^{-nh(z_0)} \int_a^b \exp\left[-n\Re(h(z(\tau) - h(z_0))\right] \frac{dz(\tau)}{d\tau} d\tau \end{aligned}$$

for some $a \le \tau \le b$. Clearly, the latter integral can be treated by the Laplace method. In particular, we expect the main contribution to come from a small region around a point z_0 that minimizes the exponent, that is, $u_x(x_0,y_0) = u_y(x_0,y_0) = 0$. But, by the Cauchy-Riemann equations (i.e., $u_x = v_y$ and $u_y = -v_x$; cf. [158]) at point z_0 also $f'(z_0) = u_x(x_0,y_0) - iu_y(x_0,y_0) = 0$. The point z_0 such that $f'(z_0) = 0$ is called the *saddle point* of $f(z)$. In view of this, we add one more condition to (i), namely,

(ii) the path $\mathcal{C}'$ passes through the saddle point(s) $f'(z_0) = 0$,

which reduces (8.83) to the Laplace integral. This constitutes the backbone of the saddle point method.

We now work out more details concerning the curve $\mathcal{C}'$, which is always the hardest part of the method. Let us try to understand condition (i) and its relationship to condition (ii). The Laplace method will work only if on the curve $\mathcal{C}'$ near the saddle point the value of f

drops significantly. But this is the case, and here why. Let us consider $u(x, y)$ as a function of (x, y). Notice that $|e^{-nh(z)}| = e^{-nu(x,y)}$ so the absolute value of the integral $I(n)$ depends on the behavior of $u(x, y)$ along $\mathcal{C}'$. Let now $x(t)$, $y(t)$ $u(t)$ be a parametric description of a curve on the surface $u(x, y)$. It is known that (details can be found in Wong [355])

$$\frac{du}{dt} = u_x \cos\theta + u_y \sin\theta$$

for some θ. To measure the steepness on the surface $u(x, y)$ along the path $u(t)$, one computes the cosine of the angle α between the path and the u-axis. It can be shown that the derivative of this cosine is zero (i.e., maximized or minimized) when

$$-u_x \sin\theta + u_y \cos\theta = 0.$$

But then, again by the Cauchy-Riemann equation, this implies

$$\frac{dv}{ds} = v_x \cos\theta + v_y \sin\theta = 0$$

on the path. Thus, the steepest curve (where the derivative of $\cos\alpha$ is zero) is also where the imaginary part of $h(z)$ is constant, that is, $\Im(f(z)) = v(s)$ =constant. In other words, the path $\mathcal{C}'$ postulated in (i) is the steepest curve on $u(x, y)$. Simultaneously, the curve $\mathcal{C}'$ on $u(x, y)$ reaches the maximum/minimum value at the saddle point z_0 where $u_x = u_y = 0$. Thus, "luckily" the the second condition (ii) is also satisfied, and one expects most of the contribution to the integral coming from a small region around the saddle point z_0.

Example 8.12 *Steepest Descent Method*
Let

$$I(n) = \int_{-\infty}^{\infty} e^{-n(z^2-2zi)} f(z)dz.$$

The saddle point is at $z_0 = i$ and $\Im(z^2 - 2zi) = 2x(y-1)$ =constant represents the steepest descent curves, where $z = x + iy$. In particular, $y = 1$ is the steepest descent curve that passes through the saddle point. Accordingly, we deform the path $(-\infty, \infty)$ to the line $\Im z = 1$ reducing the integral to

$$I(n) = e^{-n} \int_{-\infty}^{\infty} e^{-nt^2} f(t+i)dt$$

provided $f(z)$ is well-behaved (e.g., no singularities in the strip $0 \le \Im z \le 1$). Then by Theorem 8.17 we directly find that

$$I(n) \sim e^{-n} f(i) \sqrt{\frac{\pi}{n}}.$$

If, however, $f(z)$ has a singularity in $0 \le \Im z \le 1$, we must add this contribution. For example, assuming that

$$f(z) = \frac{1}{z-p} + \cdots$$

where $p \neq i$, we must add $2\pi i \exp\left(-n(p^2 - 2pi)\right)$. The situation is even more interesting when $p = i$ (i.e., the pole coincides with the saddle point). We shall discuss this case later in this chapter (in particular, in Example 22). □

Hereafter, we concentrate on the Cauchy integral and study asymptotics of

$$a_n := [z^n]A(z) = \frac{1}{2\pi i}\oint \frac{A(z)}{z^{n+1}}dz = \frac{1}{2\pi i}\oint e^{h(z)}dz \tag{8.84}$$

where

$$h(z) = \log A(z) - (n+1)\log z,$$

and $A(z)$ is a generating function with positive coefficients, and of rapid growth. Since in this case, $|A(z)| \leq A(|z|)$ we have

$$\max_{|z|=r} |A(z)| = A(r) \tag{8.85}$$

for some $r < R$ where it is assumed that A is analytic for $|z| < R \leq \infty$ (and often $R = \infty$). Clearly, the saddle point z_0 of (8.84) is located at $h'(z) = 0$ or equivalently satisfying

$$\frac{A'(z_0)}{A(z_0)} = \frac{n+1}{z_0}. \tag{8.86}$$

In view of (8.85), we may also compute the saddle point by observing that it must minimizes $r^{-n}A(r)$.

We summarize what we learned so far in the next example where we also work out all details of the method that we later generalize to establish the so called Hayman's result.

Example 8.13 *Asymptotic of* $\exp(z^\alpha)$

Let us assume that $A(z) = \exp(z^\alpha)$ with $\alpha > 0$ (we later further restrict α). In particular, for $\alpha = 1$ we have $[z^n]A(z) = 1/n!$, so we can verify our analysis. Since $h(z) = z^\alpha - (n+1)\log z$, the saddle point is at

$$z_0 = \left(\frac{n+1}{\alpha}\right)^{1/\alpha} \sim \left(\frac{n}{\alpha}\right)^{1/\alpha}.$$

Thus, writing $n_1 = n/\alpha$ we have

$$\begin{aligned}
[z^n]e^{z^\alpha} &= \frac{1}{2\pi i}\int_{|z|=n_1} \frac{e^{z^\alpha}}{z^{n+1}}dz \\
&= \frac{n_1^{-n_1}}{2\pi}\int_{-\pi}^{\pi} \exp\left(n_1 e^{i\alpha\theta} - ni\theta\right) d\theta
\end{aligned}$$

where we substituted $z = z_0 e^{i\theta}$. We now choose such $\theta_0 \in (0, \pi)$ that for $-\theta_0 \le \theta \le \theta_0$ the following expansion is valid

$$e^{i\alpha\theta} = 1 + i\alpha\theta - \frac{1}{2}\alpha^2\theta^2 + O(\theta^3), \tag{8.87}$$

For above to hold, we need to assure that $\theta = o(1)$, that is, we need further restrictions of θ_0. For such θ_0, we split the integral into three parts such that

$$\begin{aligned}
I_1 &= \frac{n_1^{-n_1}}{2\pi}\int_{-\theta_0}^{\theta_0} \exp\left(n_1 - \frac{1}{2}n\alpha\theta^2 + O(n\theta^3)\right) d\theta \\
I_2 &= \frac{n_1^{-n_1}}{2\pi}\int_{\theta_0}^{\pi} \exp\left(n_1 e^{i\alpha\theta} - ni\theta\right) d\theta \\
I_3 &= \frac{n_1^{-n_1}}{2\pi}\int_{-\pi}^{-\theta_0} \exp\left(n_1 e^{i\alpha\theta} - ni\theta\right) d\theta,
\end{aligned}$$

where certainly I_2 and I_3 are of the same order, so we only estimate I_2.

We first evaluate I_1. Our goal is to reduce it to the Gauss integral, so we shall require that

$$n\theta_0^2 \to \infty, \tag{8.88}$$
$$n\theta_0^3 \to 0, \tag{8.89}$$

and, of course, $\theta_0 \to 0$ to assure validity of (8.87). A good choice is $\theta_0 = n^{-2/5}$. Then $\exp\left(n_1 - n\theta^2/2 + O(n\theta^3)\right) = (1 + O(n^{-1/5}))\exp\left(n_1 - n\theta^2/2\right)$ and this yields

$$\begin{aligned}
I_1 &= \left(1 + O(n^{-1/5})\right)\left(\frac{n}{\alpha}\right)^{-n/\alpha} e^{n/\alpha}\left(\int_{-\infty}^{\infty} e^{-\frac{1}{2}\alpha n\theta^2} d\theta - 2\int_{\theta_0}^{\infty} e^{-\frac{1}{2}\alpha n\theta^2} d\theta\right) \\
&= \left(1 + O(n^{1/5})\right)\left(\frac{n}{\alpha}\right)^{-n/\alpha} e^{n/\alpha}\left(\sqrt{\frac{2\pi}{\alpha n}} - O(e^{-\alpha n^{1/5}/2})\right) \\
&= \left(1 + O(n^{-1/5})\right)\left(\frac{n}{\alpha}\right)^{-n/\alpha} e^{n/\alpha}\sqrt{\frac{2\pi}{\alpha n}}
\end{aligned}$$

Now, we estimate $I_2 = I_3$. For this, we observe $|\exp(z^\alpha)| = \exp(\Re(z^\alpha))$, hence

$$|I_2| \le \frac{n_1^{-n_1}}{2\pi}\int_{\theta_0}^{\pi} e^{n_1 \cos\alpha\theta} d\theta.$$

We want to show that I_2 contributes negligible. Since $\theta_0 \le |\theta| \le \pi$ we have

$$\cos\alpha\theta \le \cos\alpha\theta_0 = 1 - \alpha^2\theta_0^2/2 + O(\theta_0^4)$$

provided $\alpha\theta \le 2\pi - \theta_0$ since otherwise the inequality above would not hold. To assure this condition to hold, we now restrict $\alpha \le 1$ (later we shall discuss the case $\alpha > 1$). Then we can proceed with our calculation to obtain

$$|I_2| \le \frac{1}{2\pi} \le \left(\frac{n}{\alpha}\right)^{-n/\alpha} e^{n/\alpha} e^{-\alpha n^{1/5}/3} = O(I_1 e^{-\alpha n^{1/5}/3})$$

since $n\alpha^{-1}\cos\alpha\theta \le n\alpha^{-1} - \alpha n^{1/5}/2 + O(n^{-3/5})$. This finally proves that

$$[z^n]e^{z^\alpha} = \left(1 + O(n^{-1/5})\right)\left(\frac{n}{\alpha}\right)^{-n/\alpha} e^{n/\alpha}\sqrt{\frac{2\pi}{\alpha n}}$$

for $\alpha \le 1$. When $\alpha = 1$ we obtain Stirling's approximation, as expected. □

The example above sheds some light on complications that may arise when the saddle point method is used but we do not use the steepest descent contour through the saddle. The reader should notice that above we were unable to show that I_2 is small for $\alpha > 1$. A question arises whether we can overcome this difficulty by applying a more sophisticated saddle point approximation or there is a "fatal error". Actually, the former is true. To see this, let $\alpha = 2$. Then

$$e^{z^2} = \sum_{n=0}^{\infty} \frac{z^{2n}}{n!}$$

so that

$$[z^n]e^{z^2} = \begin{cases} \frac{1}{k!} & n = 2k \\ 0 & \text{otherwise} \end{cases}$$

or simply

$$[z^n]e^{z^2} = \frac{1+(-1)^n}{2}\frac{1}{(n/2)!}.$$

In other words, the coefficients $a_n = [z^n]e^{z^2}$ oscillate. Why then we could not handle it by the saddle point method? The culprit lies in the *multiplicity* of the saddle points. For $\alpha = 2$ we have two saddle points at $z_0 = (n/\alpha)^{1/\alpha}$ and at $-z_0$. Multiple saddle points, almost always lead to some periodicity in the asymptotics coefficients, and one must be more careful with the selection of the steepest descent contour. We shall not discuss it in this book, and the reader is referred to de Bruijn [67] or Wong [355] for a detailed discussion.

Let us now generalize Example 13. Consider

$$\begin{aligned} a_n &= \frac{1}{2\pi i}\int_{|z|=r} \frac{A(z)}{z^{n+1}}dz = \frac{1}{2\pi r^n}\int_{-\pi}^{\pi} A(re^{i\theta})e^{-in\theta}d\theta \\ &= \frac{1}{2\pi r^n}\int_{-\pi}^{\pi} \exp\left(\log A(re^{i\theta}) - in\theta\right)d\theta. \end{aligned} \tag{8.90}$$

Formally,

$$\log A(re^{i\theta}) = \log A(r) + \sum_{k=1}^{\infty} \alpha_k(r) \frac{(i\theta)^k}{k!} \tag{8.91}$$

where

$$\begin{aligned} \alpha_1(r) &:= a(r) = r\frac{d}{dr}\log A(r) = r\frac{A'(r)}{A(r)}, & (8.92)\\ \alpha_2(r) &:= b(r) = \left(r\frac{d}{dr}\right)^2 \log A(r) = ra'(r), & (8.93)\\ \alpha_k(r) &:= k\alpha_{k-1}(r). & (8.94) \end{aligned}$$

Let us continue our formal derivation and assume that the whole contribution to a_n concentrates near one saddle point, say r_0, determined by

$$\frac{d}{d\theta}\left(\log A(re^{i\theta}) - in\theta\right) = 0,$$

that is,

$$a(r_0) = n. \tag{8.95}$$

Then, after an application of the Gauss integral (8.77), we obtain formally the **saddle point approximation**

$$a_n \sim \frac{A(r_0)}{2\pi r_0^n} \int_{-\theta_0}^{\theta_0} e^{-b(r_0)\theta^2/2} d\theta \sim \frac{A(r_0)}{r_0^n\sqrt{2\pi b(r_0)}} \tag{8.96}$$

provided θ_0 is chosen appropriately. The main ingredients of the saddle point approximation are summarized in Table 8.4. We shall justify them below.

However, one must be very careful with a blind application of this heuristic. We have already seen when the saddle point method breaks down due to the impossibility of establishing smallness of the integral for $\theta \in (-\pi, -\theta_0) \cup (\theta_0, \pi)$. Now we discuss an example when the main contribution from $\theta \in (-\theta_0, \theta_0)$ does not give the right asymptotic approximation and the above heuristic again breaks down.

Example 8.14 *Invalid Application of the Saddle Point Method*

Let us consider $[z^n](1-z)^{-\alpha}$ for α not being a negative integer. We apply "blindly" the saddle point approximation. We find that $r_0 = n/(n+\alpha)$ and $b(r_0) = n(n+\alpha)/\alpha$. Using (8.96) we should obtain

$$[z^n](1-z)^{-\alpha} \sim n^{\alpha-1}\left(\frac{e}{\alpha}\right)^{\alpha}\sqrt{\frac{\alpha}{2\pi}}$$

which is not equal to $n^{\alpha-1}/\Gamma(\alpha)$ as proved in Theorem 8.12. For example, for $\alpha = 1$ we have $e/\sqrt{2\pi} = 1.0844\ldots \neq 1 = [z^n](1-z)^{-1}$. □

Now we are ready to establish rigorously the saddle point approximation (8.96) and generalize Example 13. After Hayman [155] (cf. also [154]), we define *admissible functions* $A(r)$ that satisfy the following three conditions:

Table 8.4: SUMMARY OF THE SADDLE POINT APPROXIMATION

Input: A function $g(z)$ analytic in $|z| < R$ ($0 < R < +\infty$) with nonnegative Taylor coefficients and "fast growth" as $z \to R^-$. Let $h(z) := \log g(z) - (n+1)\log z$.
Output: The asymptotic formula (8.105) for $g_n := [z^n]g(z)$ derived from the Cauchy coefficient integral

$$g_n = \frac{1}{2i\pi}\int_\gamma g(z)\,\frac{dz}{z^{n+1}} = \frac{1}{2i\pi}\int_\gamma e^{h(z)}\,dz \tag{8.97}$$

where γ is a loop around $z = 0$.
(SP1). SADDLE POINT CONTOUR. *Assume that $g'(z)/g(z) \to +\infty$ as $z \to R^-$.* Let $r = r(n)$ be the unique positive root of the saddle point equation

$$h'(r) = 0 \qquad \text{or} \qquad \frac{rg'(r)}{g(r)} = n+1, \tag{8.98}$$

so that $r \to R$ as $n \to \infty$. The integral (8.97) is evaluated on $\gamma = \{z \mid |z| = r\}$.
(SP2). BASIC SPLIT. *Assume that $h'''(r)^{1/3}h''(r)^{-1/2} \to 0$.* Define $\delta = \delta(n)$ called the "range" of the saddle point by

$$\delta = \left|h'''(r)^{-1/6}h''(r)^{-1/4}\right|, \tag{8.99}$$

so that

$$\begin{aligned} \delta &\to 0, & (8.100)\\ h''(r)\delta^2 &\to \infty, & (8.101)\\ h'''(r)\delta^3 &\to 0. & (8.102) \end{aligned}$$

Split $\gamma = \gamma_0 \cup \gamma_1$, where

$$\gamma_0 = \{z \in \gamma \mid |\arg(z)| \le \delta\}, \quad \gamma_1 = \{z \in \gamma \mid |\arg(z)| \ge \delta\}.$$

(SP3) ELIMINATION OF TAILS. *Assume that $|g(re^{i\theta})| \le |g(re^{i\delta})|$ on γ_1.* Then, the tail integral satisfies the trivial bound,

$$\left|\int_{\gamma_1} e^{h(z)}\,dz\right| = O\left(|e^{-h(re^{i\delta})}|\right). \tag{8.103}$$

(SP4) LOCAL APPROXIMATION. *Assume that $h(re^{i\theta}) - h(r) - \frac{1}{2}r^2\theta^2h''(r) = O(|h'''(r)\delta^3|)$ on γ_0.* Then, the central integral is asymptotic to a complete Gaussian integral, and

$$\frac{1}{2i\pi}\int_{\gamma_0} e^{h(z)}\,dz = \frac{g(r)r^{-n}}{\sqrt{2\pi h''(r)}}\left(1 + O(|h'''(r)\delta^3|)\right). \tag{8.104}$$

(SP5) COLLECTION. Assumptions $(SP1), (SP2), (SP3), (SP4)$, imply the estimate:

$$[z^n]g(z) = \frac{g(r)r^{-n}}{\sqrt{2\pi h''(r)}}\left(1 + O(|h'''(r)\delta^3|)\right) \sim \frac{g(r)r^{-n}}{\sqrt{2\pi h''(r)}}. \tag{8.105}$$

(H1) there exists a finite number R_0 such that $A(r) > 0$ for $R_0 < r < \infty$;

(H2) $\lim_{r\to\infty} b(r) = +\infty$ where $b(r)$ is defined in (8.93);

(H3) for some $0 < \delta(r) < \pi$ we have with $a(r)$ defined in (8.92)

$$A(re^{i\theta}) \sim A(r) \exp\left(i\theta a(r) - \frac{1}{2}\theta^2 b(r)\right), \quad r \to \infty \tag{8.106}$$

uniformly in $|\theta| \le \delta(r)$, and

$$A(re^{i\theta}) = o\left(\frac{A(r)}{\sqrt{b(r)}}\right), \quad r \to \infty \tag{8.107}$$

uniformly in $\delta(r) \le |\theta| \le \pi$.

We observe that by (8.91) condition (H3) implies

$$\sum_{k=3}^{\infty} \alpha_k(r) \frac{(i\theta)^k}{k!} \to 0$$

as $r \to \infty$ for $\delta(r) < \pi$. By (H2) we also have

$$\delta^2(r) b(r) \to \infty$$

as $r \to \infty$. To evaluate $[z^n]A(z)$ for A satisfying (H1)-(H3) we use Cauchy's integral (8.90) split at $\delta = \delta(r)$, that is, $a_n r^n = I_1 + I_2$ where

$$\begin{aligned} I_1 &= \frac{1}{2\pi}\int_{-\delta}^{\delta} A(re^{i\theta})e^{-in\theta}d\theta \\ I_2 &= \frac{1}{2\pi}\int_{-\delta}^{2\pi-\delta} A(re^{i\theta})e^{-in\theta}d\theta. \end{aligned}$$

By (H3) we have $I_2 = o(A(r)/\sqrt{b(r)})$ as $r \to \infty$. The first integral becomes under another application of (H3)

$$\begin{aligned} I_1 &= \frac{A(r)}{2\pi}\int_{-\delta}^{\delta} \exp\left(i(a(r)-n)\theta - b(r)\theta^2/2\right) d\theta + o(f(r)/\sqrt{b(r)}) \\ &\overset{y=\theta\sqrt{b(r)/2}}{=} \frac{A(r)}{\pi\sqrt{2b(r)}} \exp\left(-\frac{(a(r)-n)^2}{2b(r)}\right) \int_{-\delta}^{\delta} \exp\left[-(y-icy)^2\right] d\theta \\ &\sim \frac{A(r)}{\sqrt{2\pi b(r)}} \exp\left(-\frac{(a(r)-n)^2}{2b(r)}\right) \end{aligned}$$

where $c = (a(r)-n)\sqrt{2/b(r)}$. We formulate it as follows.

Theorem 8.19 (Hayman 1956) (i) *For admissible functions* $A(z) = \sum_{n\geq 0} a_n z^n$ *satisfying conditions (H1)-(H3) we have uniformly as* $n \to \infty$

$$a_n = \frac{A(r)}{r^n\sqrt{2\pi b(r)}}\left(\exp\left(-\frac{(a(r)-n)^2}{2b(r)}\right) + o(1)\right) \tag{8.108}$$

as $r \to \infty$.

(ii) *If* $A(z)$ *and* $B(z)$ *are admissible then* $e^{A(z)}$ *and* $A(z)B(z)$ *are admissible. If* $P(z)$ *is a polynomial and* $A(z) + P(z)$ *is admissible, then* $A(z)P(z)$ *and* $P(A(z))$ *are admissible, too. Finally,* $e^{P(z)}$ *is admissible if* $[z^n]e^{P(z)} \in \mathbb{R}^+$.

In passing, we observe that to get the saddle point approximation from Hayman's theorem one defines r_n as a unique solution of $a(r_n) = n$. Such a unique solution exists since by (H2) $b(r)ra'(r) \to +\infty$, hence $a'(r) > 0$ and $a(r)$ must be an increasing function in some interval. This suffices to prove the assertion.

Example 8.15 *Asymptotics of the Bell Numbers*
In Example 10 we defined Bell numbers b_n satisfying

$$\exp(e^z - 1) = \sum_{n=0}^{\infty} b_n \frac{z^n}{n!}.$$

By Theorem 8.19(ii) $\exp(e^z - 1)$ is admissible with $a(r) = re^r$ and $b(r) = (r^2 + r)e^r$. By Theorem 8.19(i) we have

$$b_n \sim \frac{n!\exp(e^{r_n} - 1)}{r_n^n\sqrt{2\pi(r_n+1)r_n\exp(r_n)}}$$

where r_n is defined by $r_n \exp(r_n) = n$. □

Finally, we discuss one systematic approach to the saddle point method due to **van der Waerden** [349] that is particularly useful when the saddle point coincides with a pole. We present this method in a rather formal way, but it is not difficult to make it rigorous, as Hayman's method, by stating necessary restrictive conditions on the admissible functions. When discussing this method we again return to the general integral representation (8.83). We shall follow here the presentation of Lauwerier [231].

Consider again the integral

$$I(n) = \int_{\mathcal{C}} f(z)e^{-nh(z)}dz$$

where $\mathcal{C}$ is a contour that starts and ends at infinity. We define the van der Waerden transformation as

$$w = h(z) \tag{8.109}$$

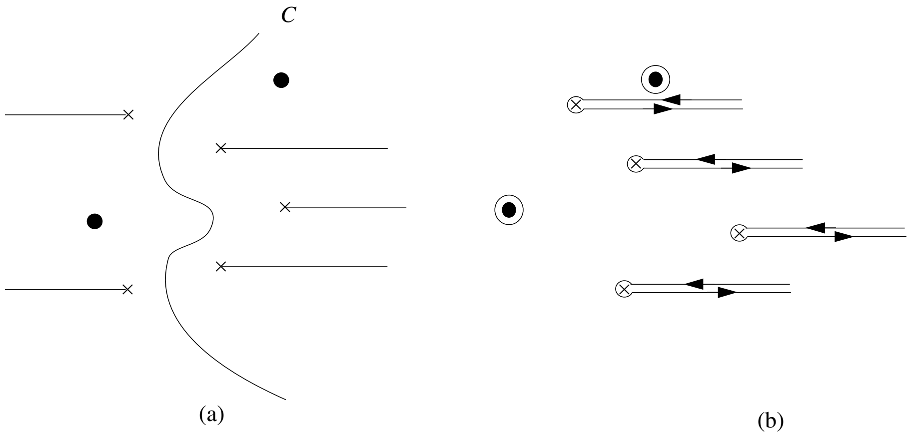


Figure 8.5: A contour in the van der Waerden method.

where $w = u+iv$. The importance of this transformation lies in changing the steepest descent lines into the horizontal lines in the (u, v) plane. Furthermore, this transformation is singular at the saddle points of $h(z)$. We assume now that:

(W1) $f(z)$ and $h(z)$ are algebraic functions of z;

(W2) the function

$$P(w) = f(z(w))\frac{dz}{dw}$$

does not grow faster than an exponential function, that is, $|P(w)| \leq Be^{c|w|}$ for some positive constants B and c.

Under this assumptions, any saddle point $h'(z) = 0$ becomes a branch point in the complex plane. Then also the integral becomes

$$I(n) = \int_{C'} e^{-nw} f(z(w))\frac{dz}{dw} dw \tag{8.110}$$

where C' is a contour that meanders between branch points and poles of the w plane, as illustrated in Figure 8.5(a). We should observe that apart from branch points and poles due to the factor dz/dw there might be some singularities coming from $f(z(w))$.

To obtain the saddle point approximation we deform the contour C' into another one C'' that will minimize $|e^{-nw}| = e^{-nu}$. This means to shift C' to the right as much as possible. It should be clear that the best contour satisfying this condition is one that consists of loops surrounding the branch points and poles, coming from infinity in straight lines from the right and going back to infinity to the right, as illustrated in Figure 8.5(b).

Let us discuss a concrete situation and assume that $w = 0$ is a single branch point of order one. A local substitution $w = s^2$ will transform the integrand function into

$$F(s) := f(z(w))\frac{dz}{dw} = f(z(\sqrt{s}))\frac{1}{2s}\frac{dz}{ds}$$

so that it has a pole of the first order at $s = 0$. This implies that $F(s)$ admits an expansion

$$F(s) = a_{-1}s^{-1} + a_0 + a_1 s + \cdots$$

which yields

$$I(n) = \int_{-\infty}^{\infty} e^{-ns^2}(a_{-1} + a_0 s + a_1 s^2 + \cdots = \sqrt{\frac{\pi}{n}}\left(2a_{-1} + \frac{a_1}{n} + \cdots\right).$$

More interestingly, if the saddle point (i.e., branch point at $w = s^2$) coincides with the pole at $s = 0$, then $F(s)$ becomes

$$F(s) = a_{-2}s^{-2} + a_{-1}s^{-1} + a_0 + a_1 s + \cdots$$

which is easily computable through the following formula (cf. [2])

$$\frac{1}{2\pi i}\int_{-\infty}^{\infty}\frac{e^{ax^2}}{x - ib}dx = e^{ab^2}(1 - \Phi(b\sqrt{2a}) \tag{8.111}$$

provided $a > 0$, where

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-t^2/2}dt$$

is the standard normal distribution.

Example 8.16 *Van der Waerden's Method*

Consider the integral

$$I(n) = \int_{-\infty}^{\infty}\frac{e^{-n(z^2-2zi)}}{(z^2+1)}dz.$$

The saddle point is at $z = i$ which coincides with the pole of $1/(z^2 + 1)$. According to van der Waerden's approach we consider the transformation

$$z^2 - 2zi = 1 + s^2$$

which simply becomes $z = i + s$. Then

$$\frac{1}{z^2+1} = \frac{1}{2is} + \frac{1}{4} + \frac{is}{8} + O(s^2)$$

and

$$\begin{aligned} I(n) &= e^{-n}\int_{-\infty}^{\infty} e^{-ns^2}\left(\frac{1}{2is}+\frac{1}{4}+\frac{is}{8}+O(s^2)\right)ds \\ &= e^{-n}\left(\frac{\pi}{2}+\frac{1}{4}\sqrt{\frac{\pi}{n}}+O(n^{-3/2})\right) \end{aligned}$$

where the last line follows from (8.77) and (8.111). □

In the next section (cf. Example 22) we present another example that better illustrate the power of the van der Waerden method.

8.4.3 Local Central Limit Theorem and Large Deviations Revisited

The saddle point method finds many applications in probability theory. In particular, it can be used to derive a local central limit theorem and a local large deviations result. The reader may want to review and compare Chapter 5 before reading this section. Following Greene and Knuth [138] we give here a short presentation (cf. also Section 8.6). Further readings on this topic are [31, 32, 65].

Let us consider $S_n = X_1 + \cdots + X_n$ where X_i are i.i.d. discrete random variables with the probability generating function $G(z) = \mathbf{E}[z^{X_1}] = \sum_{k\ge 0}\mathsf{Pr}\{X_1 = k\}z^k$. Let $\mu = \mathbf{E}[X_1]$ and $\sigma^2 = \mathbf{Var}[X_1]$. Clearly, the probability generating function of S_n is $\mathbf{E}[z^{S_n}] = G^n(z)$. Then

$$\mathsf{Pr}\{S_n = k\} := [z^k]\mathbf{E}[z^{S_n}] = \frac{1}{2\pi i}\oint \frac{G^n(z)}{z^{k+1}}dz.$$

Of particular interests are two cases, namely: (i) $k = n\mu + r$ with $r = o(\sqrt{n})$ (central limit regime), and (ii) $k = n(\mu+\Delta) = na$ where $\Delta > 0$ or $a > \mu$ (large deviations regime).

We adopt the following three assumptions:

(K1) The generating function $G(z)$ is defined in $|z| < 1+\delta$, so that

$$G(e^t) = \exp\left(\mu t + \frac{\sigma^2 t^2}{2} + \frac{\kappa_3 t^3}{3!} + \frac{\kappa_4 t^4}{4!} + \cdots\right) \tag{8.112}$$

where κ_j is the jth cumulant of X_1 (cf. Section 5.4).

(K2) $G(0) \neq 0$.

(K3) X_1 is not periodic, that is, the greatest common divisor of all k with $\mathsf{Pr}\{X_1 = k\} \neq 0$ is one.

Under these assumptions, we observe that the Cauchy integral above can be evaluated by the saddle point method with the saddle point at $z = 1 - O(1/n) \sim 1$, hence

$$\Pr\{S_n = \mu n + r\} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{G^n(e^{it})}{e^{it(\mu n+r)}}dt.$$

This integral can be evaluated by the Laplace method after noting that under (K3) for $\delta < |t| < \pi$ there exists $\alpha < 1$ such that $|G(e^{it})| < \alpha$, and

$$\frac{G^n(e^{it})}{e^{it(\mu n+r)}} = \exp\left(irt - \frac{\sigma^2 nt^2}{2} + \frac{\kappa_3 nt^3}{3!} + \frac{\kappa_4 nt^4}{4!} + \cdots\right).$$

If we choose $\delta^2 n \to \infty$ (say $\delta = n^{1/2+\varepsilon}$), then the Laplace method yields, after a careful evaluation,

$$\Pr\{S_n = \mu n + r\} = \frac{1}{\sigma\sqrt{2\pi n}}\exp\left(\frac{-r^2}{2\sigma^2 n}\right)\left(1 - \frac{\kappa_3}{2\sigma^4}\left(\frac{r}{n}\right) + \frac{\kappa_3}{6\sigma^6}\left(\frac{r^3}{n^2}\right)\right) + O(n^{-3/2}). \quad (8.113)$$

We mentioned this result in Section 5.4.

This is the *local central limit* law, from which we can obtained the central limit law after summing over r. Unfortunately, like all central limit results, the formula suffers its weakness since the range of its applicability is limited to $r = o(\sqrt{n})$ due to the fact that the error term is only a polynomial of $n^{-1/2}$. But there is a remedy: to expand the range of its validity one should shift the distribution to a new location. This is called the method of *shifted mean.* In Section 5.4 we used a variant of it called the *exponential change of measure* to obtain large deviations results. We shall show below how to recover a stronger large deviations result for discrete distributions by the shift of the mean method.

The idea is to shift the mean of S_n to a new value such that (8.113) is valid again. Thus, let us define a new random variable $\widetilde{X}$ whose generating function is

$$\widetilde{G}(z) = \frac{G(z\alpha)}{G(\alpha)}$$

where α is a constant that is to be determined. Observe that

$$\mathbf{E}[\widetilde{X}] = \alpha\frac{G'(\alpha)}{G(\alpha)}, \quad (8.114)$$

$$\mathbf{Var}[\widetilde{X}] = \alpha^2\frac{G''(\alpha)}{G(\alpha)} + \alpha\frac{G'(\alpha)}{G(\alpha)} - \alpha^2\left(\frac{G'(\alpha)}{G(\alpha)}\right)^2. \quad (8.115)$$

If one needs large deviations result around $m = na = n(\mu + \Delta)$ where $\Delta > 0$, then clearly (8.113) cannot be applied directly, but a proper choice of α can help. Let us select α such that the new $\widetilde{S}_n = \widetilde{X}_1 + \cdots + \widetilde{X}_n$ has mean $m = n(\mu + \Delta)$. This results in setting α to be a solution of

$$\frac{\alpha G'(\alpha)}{G(\alpha)} = \frac{m}{n} = a = \mu + \Delta \ . \tag{8.116}$$

A change of variable yields

$$[z^m]G^n(z) = \frac{G^n(\alpha)}{\alpha^m}[z^m]\left(\frac{G(\alpha z)}{G(\alpha)}\right)^n . \tag{8.117}$$

Now we can apply (8.113) (with $r = 0$) to $[z^m]\left(\frac{G(\alpha z)}{G(\alpha)}\right)^n$ since $m = an$ and $\mathbf{E}[\widetilde{X}_1] = a = \mu + \Delta$ so that $\mathbf{E}[\widetilde{S}_n] = an$. We re-write above as follows

$$\mathsf{Pr}\{S_n = na\} = \frac{1}{\widetilde{\sigma}\sqrt{2\pi n}} e^{-nI(a)} \left(1 + O(1/n)\right) \tag{8.118}$$

where $\widetilde{\sigma}^2 = \mathbf{Var}[\widetilde{X}_1]$ given by (8.115) and $I(a) = a \log \alpha - \log G(\alpha)$. This is a stronger version of Theorem 5.18 if one sets $\lambda_a = \log \alpha$.

Example 8.17: *Large Deviations for the Binomial Distribution*

This is taken from Greene and Knuth [138] (a more sophisticated example is presented in Section 8.6). Let S_n be binomially distributed with parameter 1/2, that is, $G^n(z) = ((1+z)/2)^n$. We want to estimate the probability $\mathsf{Pr}\{S_n = n/3\} = [z^{n/3}](z/2 + 1/2)^n$ which is far away from its mean $\mathbf{E}[S_n] = n/2$. Shift of mean yields such α that

$$\frac{\alpha G'(\alpha)}{G(\alpha)} = \frac{\alpha}{1+\alpha} = \frac{1}{3} ,$$

i.e., $\alpha = 1/2$. Using (8.113) we obtain

$$[z^{n/3}]\left(\frac{2}{3} + \frac{1}{3}z\right)^n = \frac{3}{2\sqrt{\pi n}}\left(1 - \frac{7}{24n}\right) + O(n^{-5/2})$$

and an application (8.118) gives

$$[z^{n/3}](z/2 + 1/2)^n = \left(\frac{3 \cdot 2^{1/3}}{4}\right)^n \frac{3}{2\sqrt{\pi n}}\left(1 - \frac{7}{24n} + O(n^{-2})\right)$$

for $n \to \infty$. $\square$

8.5 Finite Sums as Complex Integrals

In Section 8.1.2 we saw how a series may be represented by a complex integral. Such a representation is often quite useful since the integral, more than often, can easily be evaluated by the residue theorem. Here is another representative example of this kind

$$\sum_{k=m}^{n} (-1)^k f_k = \frac{1}{2\pi i} \int_{m-\frac{1}{2}-i\infty}^{m-\frac{1}{2}+i\infty} f(s) \frac{\pi}{\sin \pi s} ds$$

which is true under suitable conditions on $f(s)$ (e.g., polynomial growth), where $f(s)$ is an analytic continuation of f_k, that is, $f(s)_{s=k} = f_k$. The above is a direct consequence of the residue theorem and the fact that $\text{Res}[\pi/\sin(\pi s),\ s = k] = (-1)^k$, provided $f(s)$ does not have any other singularity right to the line of the integration.

In this section, we discuss an important finite sum that often arises in the analysis of algorithms (cf. Section 7.6.1). Namely, we deal with the following alternating sum

$$S_n[f] = \sum_{k=m}^{n} (-1)^k \binom{n}{k} f_k$$

where f_k is a known, but otherwise, general sequence. Due to the alternating sign this sum is not easy to evaluate by traditional methods. In fact, we shall show that there are some fluctuations involved in its asymptotics that can be treated by complex asymptotic methods. We will not dwell too much on this topic since there is an excellent survey on such sums by Flajolet and Sedgewick [117]. The reader is refer to it (and other references such as [220, 320] for more examples).

Theorem 8.20 (Rice's Formula) *Let $f(s)$ be an analytical continuation of $f(k) = f_k$ that contains the half line $[m, \infty)$. Then,*

$$S_n[f] := \sum_{k=m}^{n} (-1)^k \binom{n}{k} f_k = \frac{(-1)^n}{2\pi i} \int_{\mathcal{C}} f(s) \frac{n!}{s(s-1)\cdots(s-n)} ds \tag{8.119}$$

where $\mathcal{C}$ is a positively oriented curve that encircles $[m, n]$ and does not include any of the integers $0, 1, \ldots, m-1$.

Proof. It is a direct application of the residue calculus after noting that the poles of the integrand are at $k = m, m+1, \ldots, n$ and

$$\text{Res}\left[\frac{n!}{s(s-1)\cdots(s-n)} f(s);\ s = k\right] = (-1)^{n-k} \binom{n}{k} f(k).$$

This completes the proof. ■

There is another representation of $S_n[f]$ that is sometimes easier to handle computationally. (It is called by Knuth [220] the gamma method.) In this case, however, we must restrict f to have a *polynomial growth*, that is, $f(s) = O(|s|^k)$ for some k as $s \to \infty$. In fact, when evaluating the integral in the Rice method we usually must impose some growth condition on f over large circles.

Theorem 8.21 (Knuth 1973, Szpankowski 1988) *Let $f(z)$ be of polynomial growth at infinity and analytical left to the vertical line $(\frac{1}{2} - m - i\infty, \frac{1}{2} - m + i\infty)$. Then*

$$\begin{aligned} S_n[f] &= \frac{1}{2\pi i}\int_{\frac{1}{2}-m-i\infty}^{\frac{1}{2}-m+i\infty} f(-z)B(n+1,z)dz && (8.120) \\ &= \frac{1}{2\pi i}\int_{\frac{1}{2}-m-i\infty}^{\frac{1}{2}-m+i\infty} f(-z)n^{-z}\Gamma(z)\left(1 - \frac{z(z+1)}{2n}\right. && (8.121) \\ &+ \left.\frac{z(1+z)}{24n^2}(3(1+z)^2 + z - 1) + O(n^{-3})\right) dz, \qquad \Re(z) > 0 \end{aligned}$$

where $B(x,y) = \Gamma(x)\Gamma(y)/\Gamma(x+y)$ is the beta function.

Proof. We again use the residue theorem, the fact that $\mathrm{Res}[\Gamma(z),\ z = k] = (-1)^k/k!$ and $\mathrm{Res}[B(n+1,z),\ z = k] = (-1)^k\binom{n}{k}$. However, some care is needed. Let us consider a large rectangle $R_{\alpha,M}$ left to the line of the integration, with corners at four points $(\frac{1}{2} - M \pm i\alpha$, $c \pm i)$, $\alpha > 0$, M is a positive integer and $c = \frac{1}{2} - m$. By Cauchy's theorem the integral in (8.120) is equal to the sum of residues in $R_{\alpha,M}$ minus the values of the integral on the bottom, top and left lines of the rectangle. We must prove now that the integral along left, top and bottom lines are small. This is not difficult due to the behavior of the gamma function for imaginary values (cf. (2.33)) and the polynomial growth of f. We consider the integral along the top line which we denote as I_T. Using the estimate (2.34) for the gamma function we obtain

$$|I_T| = O\left(e^{-\pi\alpha/2}\int_{-\infty}^{c} (\alpha/n)^x |f(-x - i\alpha)| dx\right) \to 0$$

which tends to zero for fixed n and $\alpha \to \infty$. Let now I_L be the integral along the left line. We need a different estimate here. Using successively $\Gamma(z+1) = z\Gamma(z)$ we first observe that

$$\Gamma(\frac{1}{2} - M + iy) = \Gamma(\frac{1}{2} + iy)/O((M-1)!).$$

Then

$$|I_L| = O\left(\frac{n^M}{(M-1)!}\int_{-\infty}^{\infty} |\Gamma(1/2 + iy)||f(M - 1/2 - iy)| dy\right)$$

which also tends to zero as $M \to \infty$ due to the exponential smallness of $\Gamma(z)$ along the imaginary line and the polynomial growth of $f(z)$ so that the integral above converge. This completes the proof. ■

Example 8.18: *Asymptotics of an Alternating Sum*
In Section 7.6.1 we encountered alternating sums of the following type

$$S_{n+r} = \sum_{k=2}^{n+r} (-1)^k \binom{n+r}{k} \binom{k}{r} \frac{1}{1 - p^{-k} - q^{-k}} \tag{8.122}$$

where $p+q=1$. We now use Theorem 8.21 to obtain asymptotics of S_{n+r} as n becomes large and r is fixed. We first slightly generalize Rice's method by observing that for any sequence f_k

$$S_{n+r}[f] = \sum_{k=2}^{n+r} (-1)^k \binom{n+r}{k} \binom{k}{r} f_k = (-1)^r \binom{n+r}{r} \sum_{k=[2-r]^+}^{n} (-1)^k \binom{n}{k} f(k+r).$$

Thus by (8.121) the sum (8.122) becomes

$$S_{n+r} = \left(1 + O(n^{-1})\right) \frac{(-1)^r}{r!} \frac{1}{2\pi i} \int_{\frac{1}{2}-[2-r]^+ - i\infty}^{\frac{1}{2}-[2-r]^+ + i\infty} n^{r-z} \Gamma(z) \frac{1}{1 - p^{r-z} - q^{r-z}} dz + e_n$$

$x^+ = \max\{0, x\}$, where e_n is an error term that we discuss later. Naturally, we evaluate the integral by the residue theorem. However, first we observe that the function under the integral has infinitely many poles that are roots of

$$1 = p^{r-z} + q^{r-z} \ .$$

This equation appears many times in this book, and we need to study the locations of these roots in order to find asymptotic expansions of S_{n+r} and similar expressions. The reader is asked in Exercise 18 to provide a full proof of the following lemma.

Lemma 8.22 (Jacquet 1989, Schachinger 1993) *Let z_k for $k = \mathbb{Z} = 0, \pm1, \pm2, \ldots$ be solutions of*

$$p^{-z+r} + q^{-z+r} = 1 \tag{8.123}$$

where $p+q=1$ and z is complex.
(i) *For all $k \in \mathbb{Z} = \{0, \pm1, \pm2, \ldots\}$*

$$-1 + r \le \Re(z_k) \le \sigma_0 + r \tag{8.124}$$

where σ_0 is a positive solution of $1 + q^{-s} = p^{-s}$. Furthermore,

$$\frac{(2k-1)\pi}{\log p} \le \Im(z_k) \le \frac{(2k+1)\pi}{\log p} \ .$$

(ii) *If* $\Re(z_k) = -1 + r$ *and* $\Im(z_k) \neq 0$, *then* $\log p / \log q$ *must be rational. More precisely, if* $\frac{\log p}{\log q} = \frac{s}{t}$ *where* $\gcd(s,t) = 1$ *for* $s, t \in \mathbb{Z}$, *then*

$$z_k = -1 + r + \frac{2ks\pi i}{\log p} \tag{8.125}$$

for all $k \in \mathbb{Z}$.

Proof. We give only a sketch of the proof for $r = 1$. Let us assume that $z = -1 + it$ is a solution of $p^{-z} + q^{-z} = 1$. Then

$$\begin{aligned} |p^{-z} + q^{-z}| &= 1 = p + q = |p^{1-it}| + |q^{1-it}| \\ &= |p^{-z}| + |q^{-z}|. \end{aligned}$$

Now we know that the first and the last expression are equal and both equal to 1. But for complex numbers u and v the equality $|u + v| = |u| + |v|$ holds if and only if $u = \lambda v$, with some real $\lambda \geq 0$. Therefore, p^{-z} and q^{-z} have the same argument in the complex plane. Furthermore $p^{-z} + q^{-z}$, which is again a complex number with this argument, is equal to 1. Thus, the argument in question must be zero. This means that p^{-z} as well as q^{-z} are nonnegative real numbers as soon as z is a solution. Their imaginary parts are zero, which yields

$$\sin(t \log p) = \sin(t \log q) = 0.$$

Therefore $t \log p$ as well as $t \log q$ are an integer multiple of π, from which it follows immediately, that the ratio of $\log p$ and $\log q$ has to be rational, if t is different from zero. Also, if this ratio is not in $\mathbb{Q}$, then there does not exist a solution with real part -1 and imaginary part different from zero. ■

Based on the above lemma, we conclude that either $\Re(z_k) > r - 1$ for $\log p / \log q \notin \mathbb{Q}$ or $z_k = r - 1 + 2\pi iks / \log p$ when $\log p / \log q = s/t \in \mathbb{Q}$. In any case, the line $\Re(z_k)$ lies to the *right* of the line of the integration $(\frac{1}{2} - [2-r]^+ - i\infty, \frac{1}{2} - [2-r]^+ + i\infty)$. We now apply the closing-the-box method discussed in Section 7.6.3. We consider a big rectangle with left side being the line of integration, the right side positioned at $\Re(z) = M$ (where M is a large number), and bottom and top side positioned at $\Im(z) = \pm A$, say. We further observe that the right side contributes only $O(n^{r-M})$ due to the factor n^{r-M} in the integral. Both the bottom and top sides contributes negligible, too, since the gamma function decays exponentially fast with the increase of imaginary part when $A \to \infty$. This fact was already used in the proof of Theorem 8.21. In summary, the integral is equal to a circular integral (around the rectangle) plus a negligible part $O(n^{r-M})$. But, then by Cauchy's residue theorem the latter integral is equal to minus the sum of all residues at z_k, that is,

$$S_{n+r} = -\frac{(-1)^r}{r!} \sum_{k=-\infty}^{\infty} \operatorname{Res}\left(\frac{n^{r-z}\Gamma(z)}{1 - p^{r-z} - q^{r-z}}, \; z = z_k \right) + O(n^{r-M}) \; . \tag{8.126}$$

Table 8.5: Upper Bound on $|P_r(n)|$ as a function of r.

r	$\lvert P_r(n)\rvert$
-2	$.1725\ 10^{-6}$
-1	$.1573\ 10^{-5}$
0	$.1426\ 10^{-4}$
1	$.1300\ 10^{-3}$
2	$.1207\ 10^{-2}$
3	$.1153\ 10^{-1}$
4	$.1142\ 10^{-0}$
5	1.1823
6	12.8529
7	147.2071
8	1779.8207

The minus sign above indicates that the contour is encircled in the clockwise direction. When $\log p/\log q \notin \mathbb{Q}$, then $\Re(z_k) > r-1$, hence $S_{n+r} = o(n)$. Hereafter, we only consider $\log p/\log = s/t \in \mathbb{Q}$. In this case, the main contribution to the asymptotics comes from $z_0 = r-1$, which is a double pole. We shall use the following expansions for $w = z - z_0$

$$\begin{aligned} n^{r-z} &= n(1 - w\ln n + O(w^2)) \ , \\ (1 - p^{r-z} - q^{r-z})^{-1} &= -w^{-1}h^{-1} + \frac{1}{2}h_2 h^{-2} + O(w), \\ \Gamma(z) &= (-1)^{r+1}(w^{-1} - \gamma + \delta_{r,0}) + O(w) \qquad r = 0,1 \end{aligned}$$

where $h = -p\ln p - q\ln q$, $h_2 = p\ln^2 p + q\ln^2 q$, and $\gamma = 0.577215\ldots$ is the Euler constant. Considering, in addition, the residues coming from z_k for $k \neq 0$ we finally arrive at

$$S_{n+r} = \begin{cases} \frac{1}{h}n(\ln n + \gamma - \delta_{r,0} + \frac{1}{2}h_2) + \frac{(-1)^r}{r!}nP_r(n) + e_n & r = 0,1 \\ \\ n\frac{(-1)^r}{r(r-1)h} + (-1)^r nP_r(n) + e_n & r \geq 2 \end{cases}$$

provided $\log p/\log = s/t \in \mathbb{Q}$. In the above $P_r(n)$ is a contribution coming from z_k for $k \neq 0$, and one computes

$$\begin{aligned} P_r(n) &= \frac{1}{h}\sum_{k\in\mathbb{Z}\setminus\{0\}} \Gamma(r - 1 + 2\pi ks/\log p)\exp(2\pi iks\log_p n) \\ &= \frac{1}{h}\sum_{k\neq 0}\Gamma(r - 1 + 2\pi ks/\log p)\exp(2\pi iks\log_p n) \end{aligned}$$

which is a periodic function of $\log_p n$ with period 1, mean 0 and small amplitude for small r. (Throughout, we write $\sum_{k\neq 0} := \sum_{k\in\mathbb{Z}\setminus\{0\}}$.) For example, when $p = q = 1/2$ we have

$$P_r(n) = \frac{1}{\ln 2}\sum_{k\neq 0}\Gamma(r - 1 + 2\pi ik/\log 2)\exp(-2\pi ik\log_2 n).$$

The absolute value of $P_r(n)$ is shown in Table 8.5 for some r. Observe that the amplitude of $P_r(n)$ is very small for small r but rapidly increases even for moderate r.

Concerning the error term e_n. It can be represented as

$$e_n = O\left(\frac{1}{2n}\frac{1}{2\pi i}\frac{(-1)^n}{r!}\int_{\frac{1}{2}-[2-r]^+-i\infty}^{\frac{1}{2}-[2-r]^++i\infty} n^{r-z}(z(z+1)\Gamma(z)\frac{1}{1-p^{r-z}-q^{r-z}}dz\right)$$

hence the same procedure as above may be used to evaluate the integral. This proves that $e_n = O(1)$. □

Alternating sums arise in many computations, and we often must use various asymptotic methods to evaluate the integral involved. The reader is referred to Flajolet and Sedgewick [117] and exercises at the end of this chapter for more examples. For example, if $F(z)$ denotes the ordinary generating function of f_k and $S(z)$ the ordinary generating functions of $S_n[f]$, then one obtains, as in Example 11 of Chapter 7

$$S(z) = \frac{1}{1-z}F\left(-\frac{z}{1-z}\right).$$

Thus, if $F(z)$ has algebraic singularities, so $S(z)$ does, and one can use the singularity analysis to derive asymptotics of S_n.

Furthermore, this approach works fine also for the so called *binomial sums* defined as

$$\bar{S}_n[f] = \sum_{k=0}^{n}\binom{n}{k}p^k q^{n-k} f_k, \quad q = 1-p.$$

Again, the ordinary generating function $\bar{S}(z)$ of $\bar{S}_n[f]$ can be expressed in terms of the generating function $F(z)$ of f_k as

$$\bar{S}(z) = \frac{1}{1-qz}F\left(\frac{pz}{1-qz}\right).$$

Binomial sums were studied by Flajolet [105] using singularity analysis. We shall return to them in the last chapter of this book where they are solved by the depoissonization tool.

We complete this section with the analysis of an interesting alternating sum that requires an application of the saddle point method in the Rice formula. Moreover, in this case the saddle point coincides with a pole.

Example 8.19 *Digital Search Tree*

In the analysis of digital search trees Louchard [239] considered the following sum

$$F_n(k) = \sum_{i=0}^{k} \binom{k}{i} p^i q^{k-i} (1 - p^i q^{k-i})^n \tag{8.127}$$

where $k = O(\log n)$ and $p + q = 1$. This is related to the depth in a digital search tree, and therefore of particular interest is the case when $k = \frac{1}{h}\log n + x\sqrt{\log n}$ and $x = O(1)$. As before, $h = -p\log p - q\log q$ is the entropy. First of all, observe that

$$\begin{aligned} F_n(k) &= \sum_{i\geq 0} \binom{k}{i} p^i q^{k-i} \sum_{\ell\geq 0} \binom{n}{\ell} (-1)^{n-\ell} p^{\ell i} q^{\ell(k-i)} \\ &= \sum_{\ell\geq 0} \binom{n}{\ell} (-1)^{\ell} \sum_{i\geq 0} \binom{k}{i} p^{i(\ell+1)} q^{(k-i)(\ell+1)} \\ &= \sum_{\ell\geq 0} \binom{n}{\ell} (-1)^{\ell} (p^{\ell+1} + q^{\ell+1})^k. \end{aligned}$$

The last sum is an alternating sum and can be treated by the Rice method. Using Theorem 8.21 we arrive at the following

$$F_n(k) = \frac{1}{2\pi i}\,(1 + O(1/n)) \int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty} n^{-s}\Gamma(s)(p^{1-s} + q^{1-s})^k ds. \tag{8.128}$$

As in Louchard [239], we claim the following.

Theorem 8.23 (Louchard 1983) *Let $h_2 = p\log^2 p + q\log^2 q$ and $\sigma^2 = (h_2 - h^2)/h^3$ with $p \neq q$. For $k = \frac{1}{h}\log n + y\sigma\sqrt{\log n}$ we have as $n \to \infty$*

$$F_n(k) \sim \Phi(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-t^2/2} dt \tag{8.129}$$

provided $y = O(1)$.

Proof. We apply the saddle point method, and estimate the integral

$$I(n) = \frac{1}{2\pi i} \int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty} n^{-s}\Gamma(s)(p^{1-s} + q^{1-s})^k ds$$

for $k = c_1 \log n + x\sqrt{\log n}$ where $c_1 = 1/h$ and $x = y\sigma$. Clearly,

$$I(n) = \frac{1}{2\pi i} \int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty} \Gamma(s) \exp\left(-(s\log n - k\log(p^{1-s} + q^{1-s}))\right) ds\ .$$

Observe that

$$s\log n - k\log(p^{1-s}+q^{1-s}) = -\frac{(s-s_0)^2}{2}k(h_2-h^2) + \frac{(\log n - kh)^2}{2k(h_2-h^2)} + O(ks^3)$$

where

$$s_0 = \frac{\log n - kh}{k(h_2-h^2)} = \frac{-y}{h\sigma\sqrt{\log n}}.$$

Let now $s = s_0 + it$. Since s_0 is also close to a pole of $\Gamma(s)$ we must also use

$$\Gamma(s) \sim \frac{1}{s_0+it} = \frac{s_0}{s_0^2+t^2} - \frac{it}{s_0^2+t^2}.$$

Substituting $u = t^2$ and $x = y\sigma$, and putting everything together, we finally obtain

$$\begin{aligned} I(n) &\sim \frac{1}{\pi i}e^{-y^2/2}\int_0^\infty e^{t^2h^2\sigma^2\log n/2}\frac{s_0}{s_0^2+t^2}(1+O(x/\sqrt{\log n}))dt \\ &= (1+O(x/\sqrt{\log n}))\frac{s_0}{\pi i}e^{-y^2/2}\int_0^\infty e^{-u\beta}\frac{1}{\sqrt{u}(s_0^2+u)}du \end{aligned}$$

where $\beta = h^2\sigma^2\log n/2$. But it is known that (cf. [2])

$$\frac{1}{\pi i}\int_0^\infty e^{-u\beta}\frac{1}{\sqrt{u}(s_0^2+u)}du = \frac{1}{2s_0}e^{\beta s_0^2}\mathrm{Erfc}(s_0\sqrt{\beta}) \ ,$$

where

$$\mathrm{Erfc}(x) = \frac{2}{\sqrt{\pi}}\int_x^\infty e^{-t^2}dt \ .$$

Then

$$I(n) \sim \frac{1}{2}\mathrm{Erfc}(-y/\sqrt{2}) = 1 - \Phi(-y) = \Phi(y) \ ,$$

where Φ is the standard normal distribution. The reader is asked in Exercise 17 to use the van der Waerden method to prove this result. □

8.6 Limiting Distributions

In this section we apply the asymptotic methods learned so far to derive limiting distributions of a sequence of random variables from its sequence of generating functions.

8.6.1 Discrete Limit Laws Through PGF

From now on, we shall deal with a sequence of discrete random variables X_n and the corresponding sequence of generating function $G_n(z) = \mathbf{E}[z^{X_n}]$. A natural question to ask is whether one can infer from the limit of $G_n(z)$ $(n \to \infty)$, if it exists, whether the coefficients $p_{n,k} = [z^k]G_n(z)$ (for any fixed k) tend to a limit, say $p_k = \lim_{n\to\infty} \Pr\{X_n = k\}$ such that $\sum_{k\geq 0} p_k = 1$. If the latter holds, then X_n is said to satisfy a *discrete limit law.* A *continuity theorem* for the probability generating functions (PGF) answers the above question in the positive. Interestingly enough, convergence to a discrete law, if exists, is always *uniform*, as we shall see below.

Let $G_n(z) = \mathbf{E}[z^{X_n}]$ be a sequence of generating functions analytic in a region Ω that has at least one accumulation point. Observe that $G_n(z)$ are uniformly bounded since $|G_n(z)| \leq 1$ for $z \in \Omega \subset \{z : \ |z| < 1\}$. Vitali's theorem (cf. [337] pp. 168) then implies: $G_n(z)$ *tends uniformly to a limit inside* Ω, *and therefore the limit is an analytic function.*

Theorem 8.24 (Continuity of PGF) *Let* $G(z) = \sum_{k\geq 0} p_k z^k$ *together with the sequence* $G_n(z)$ *of generating functions be defined on a set* Ω *having at least one accumulation point. If*

$$\lim_{n\to\infty} G_n(z) = G(z) \tag{8.130}$$

pointwise for $z \in \Omega$, *then for every* k *we have*

$$\lim_{n\to\infty} [z^k]G_n(z) = p_k. \tag{8.131}$$

Proof. Following Flajolet and Sedgewick [118], we apply Cauchy's theorem enforced by Vitali's theorem to obtain

$$\begin{aligned} p_k &= \frac{1}{2\pi i}\oint \frac{G(z)}{z^{k+1}}dz \\ &= \lim_{n\to\infty} \frac{1}{2\pi i}\oint \frac{G_n(z)}{z^{k+1}}dz \\ &= \lim_{n\to\infty} [z^k]G_n(z), \end{aligned}$$

where the second, crucial, line is justified by the Vitali theorem. ■

Example 8.20 *Poisson Theorem*

Let X_n have the *binomial*(n,p) distribution, that is, $\Pr\{X_n = k\} = \binom{n}{k}p^k q^{n-k}$ where $q = 1 - p$. We shall assume that p depends on n such that $\lambda = \lim_{n\to\infty} np$ is a constant. Observe that

$$G_n(z) = (1 - p + pz)^n,$$

hence

$$\lim_{n\to\infty} G_n(z) = \lim_{n\to\infty} \left(1 + \frac{\lambda(z-1)}{n}\right)^n = e^{\lambda(z-1)}.$$

Since $e^{\lambda(z-1)}$ is PGF for the Poisson distribution, by Continuity Theorem 8.24 we proved Poisson's theorem which asserts that the binomial distribution converges to the Poisson when $np \to \lambda$. □

In passing, we should point out that one must be careful with limiting distributions of discrete random variables since often such a distribution might not exist even if an *asymptotic formula* for the "limiting" distribution can be derived. This problem is notorious for the so called *extreme distributions* that we already touched in Section 3.3.2. To see this, let us consider a sequence of i.i.d. discrete random variables, say, $X_1, \ldots, X_n$. Define

$$M_n = \max\{X_1, \ldots, X_n\}.$$

The question we ask is whether one can find sequences a_n and b_n such that

$$\frac{M_n - a_n}{b_n}$$

converges to a non-degenerate random variable. Anderson [11] proved that this is not the case as long as the probability mass $p_k = \Pr\{X_i = k\}$ is such that

$$\frac{p_k}{\sum_{j=k}^{\infty} p_j} \tag{8.132}$$

does not converge to 0 or ∞ as $k \to \infty$. We illustrate this point on an example.

Example 8.21 *Extreme Distribution for i.i.d. Geometric Distributions*

Let $\Pr\{X_i = k\} = p^k(1-p)$ for each $1 \le i \le n$ have *geometric*(p) distribution. Then (8.132) converges to $1-p > 0$, hence a limiting distribution cannot exit. But, an asymptotic formula for $\Pr\{M_n < \log_{1/p} n + x\}$ can be expressed for fixed x and large n as follows

$$\Pr\{M_n < \log_{1/p} n + x\} \sim \exp\left(-p^{x - \langle \log_{1/p} n \rangle}\right)$$

where $\langle t \rangle = t - \lfloor t \rfloor$. Indeed, let $\Pr\{X_i \ge k\} = 1 - F(k) = p^k$. Then

$$\begin{aligned}
\Pr\{M_n < k\} &= (F(k))^n = (1 - (1 - F(k))^n \\
&= \exp(n \log(1 - (1 - F(k))) \\
&\sim \exp(-n(1 - F(k)) \qquad 1 - F(k) = p^k \to 0 \quad \text{as} \quad k \to \infty.
\end{aligned}$$

Thus, for x real and $k = \lfloor \log_{1/p} n + x \rfloor$ we have

$$\begin{aligned}
\Pr\{M_n < \lfloor \log_{1/p} n + x \rfloor\} &= \Pr\{M_n < \log_{1/p} n + x\} \\
&\sim \exp\left(-np^{\lfloor \log_{1/p} n + x \rfloor}\right) \\
&= \exp\left(-p^{x - \langle \log_{1/p} n + x \rangle}\right) .
\end{aligned}$$

Observe now that by Weyl's criterion $\langle \log n \rangle$ is not uniformly distributed modulo 1 (cf. Theorem 8.6) and the limiting distribution of M_n does not exist. □

8.6.2 Continuous Limit Laws By Integral Transforms

In this book, we treat mostly discrete structures, and hence deal mostly with discrete random variables. However, when considering limit laws, a discrete random variable after a normalization may converge to a continuous random variable. The most famous example is the central limit law which asserts that the sequence $(X_n - \mathbf{E}[X_n])/\sqrt{\mathbf{Var}[X_n]}$ convergences in distribution to the standard normal distribution. We have already seen it in Section 8.4.3.

Convergence *in distribution* was already discussed in Section 2.2.2. To recall, we say that X_n converges in distribution to X, and we write $X_n \xrightarrow{d} X$, if $F_n(x) \to F(x)$ for each point of continuity of $F(x)$, where $F_n(x)$ and $F(x)$ are distribution functions of X_n and X, respectively. If X is a continuous random variable, then instead of PGF we should deal with the *moment generating function* MGF (also known as the Laplace transform but strictly speaking it is the double-sided standard Laplace transform with s replaced by $-s$)

$$M(s) := \mathbf{E}[e^{sX}] = \int_{-\infty}^{\infty} e^{sx} dF(x);$$

or the *characteristic function* CF (which is really the Fourier transform)

$$\phi(t) := \mathbf{E}[e^{itX}] = \int_{-\infty}^{\infty} e^{itx} dF(x).$$

Clearly, the characteristic function is always defined for any distribution, while the moment generating function might not exist for heavily tailed distributions. In fact, we proved in Section 5.4 that the existence of MGF near the origin implies an exponential tail.

We finish these introductory comments with a few remarks. First of all, the standard normal distribution has the following MGF and CF;

$$M(s) = e^{s^2/2}, \qquad \phi(t) = e^{-t^2/2}.$$

Secondly, we recall that the rth moment of X, if it exists, can be computed as follows

$$\mathbf{E}[X^r] = \frac{d^r}{ds^r} M(s)\bigg|_{s=0} = (-1)^r \frac{d^r}{dt^r}\phi(t)\bigg|_{t=0}.$$

Thirdly, for $\mu = \mathbf{E}[X]$ and $\sigma^2 = \mathbf{Var}[X] > 0$ we denote by

$$X^* = \frac{X - \mu}{\sigma}$$

the normalized random variable. Then

$$\begin{aligned} \phi_{X^*}(t) &= e^{-i\mu t/\sigma}\phi_X(t/\sigma), \\ M_{X^*}(s) &= e^{-s\mu/\sigma} M_X(s/\sigma). \end{aligned}$$

With this in mind, we can move to more interesting considerations. As in the case of discrete random variables, we would like to infer convergence in distribution of X_n from convergence of MGF or CF. This is fortunately possible and known as Levy's continuity theorem. We state it below, and the reader is referred to Feller [99] or Durrett [93] for a proof.

Theorem 8.25 (Levy's Continuity Theorem) *Let X_n and X be random variables with characteristic functions $\phi_n(t)$ and $\phi(t)$, and moment generating functions $M_n(s)$, $M(s)$ (if exist), respectively.*
CHARACTERISTIC FUNCTION. *If X has a continuous distribution, then a necessary and sufficient condition for $X_n \xrightarrow{d} X$ is*

$$\lim_{n\to\infty} \phi_n(t) = \phi(t)$$

pointwise for all real t.
LAPLACE TRANSFORM. *If $M_n(t)$ and $M(t)$ exist in a common interval $[-t_0, t_0]$, and the following holds pointwise for all real $t \in [-t_0, t_0]$*

$$\lim_{n\to\infty} M_n(t) = M(t),$$

then $X_n \xrightarrow{d} X$.

In combinatorics and analysis of algorithms, the following two corollaries are very useful. They take advantage of the fact that often generating functions of discrete structures are analytic for all complex z in a neighborhood of zero. We know that analytic functions have all derivatives, hence the underlying random variables should possess all moments.

Corollary 8.26 (Goncharov) *Let X_n be a sequence of discrete random variables with generating function $G_n(z) = \mathbf{E}[z^{X_n}]$, mean $\mu_n = \mathbf{E}[X_n]$ and variance $\mathbf{Var}[X_n] = \sigma_n^2 > 0$. If*

$$\lim_{n\to\infty} e^{-\tau\mu_n/\sigma_n} G_n(e^{\tau/\sigma_n}) = e^{\tau^2/2} \tag{8.133}$$

for all $\tau = it$ and $-\infty < t < \infty$, then $X_n \xrightarrow{d} N(0,1)$ where $N(0,1)$ stands for the random variable having the standard normal distribution.
If (8.133) holds for $\tau \in [-\theta, \theta]$, $\theta > 0$, then there exists $\alpha > 0$ such that

$$\Pr\left\{\frac{X_n - \mu_n}{\sigma_n} > k\right\} \le e^{-\alpha k} \tag{8.134}$$

uniformly for all n and k.

Proof. The first part is just a simple consequence of the Levy continuity theorem. The large deviations part follows from Markov's inequality applied to $X_n^* = (X_n - \mu_n)/\sigma_n$. Indeed, since $\mathbf{E}[e^{tX_n^*}] < C$ for $t \in [-\theta, \theta]$, where C is a constant, we obtain for any $k > 0$

$$\Pr\{X_n^* > k\} \leq e^{-tk}\mathbf{E}[e^{tX_n^*}] \leq Ce^{-k\theta},$$

as needed. ■

Corollary 8.27 *Let X_n and X be respectively a sequence of real random variables and a real random variable such that $M_n(t) = \mathbf{E}[e^{tX_n}]$ and $M(t) = \mathbf{E}[e^{tX}]$ are their moment generating functions defined in a real neighborhood of $t = 0$. Suppose that $\lim_{n\to\infty} M_n(t) = M(t)$ for t belonging to such a real neighborhood of 0. Then, X_n converges to X both in distribution and in moments (i.e., $\mathbf{E}[X_n^r] \to \mathbf{E}[X^r]$ for any $r > 0$).*

Proof: The convergence in distribution follows directly from Levy's theorem. We concentrate on proving the second part. Let us first extend the definition of $M_n(t)$ to a complex neighborhood of $t = 0$. Observe that in such a neighborhood

$$|e^{tX_n}| \leq e^{\Re(t)X_n} \ .$$

Therefore, $M_n(t)$ exists and is bounded in a complex neighborhood of 0. Clearly, $M_n(t)$ and $M(t)$ are analytic functions. We know that $|M_n(t)|$ is uniformly bounded in this neighborhood by, say, a number $A > 0$. We redefine this neighborhood by removing all point with distance smaller than $\varepsilon > 0$ of the boundary of the former neighborhood, where ε is arbitrarily small. Due to Cauchy's estimate the derivatives $|M_n'(t)|$ are also uniformly bounded by $A\varepsilon^{-1}$, and therefore $M_n(t)$ are bounded and uniformly continuous. By *Ascoli's Theorem*, from every sequence of $M_n(t)$ we can extract a convergent subsequence. This limit function can only be the unique analytic continuation of $M(t)$, thus $M_n(t)$ converges to $M(t)$ in this complex of 0. The convergence in moments follows since convergence of analytic functions implies the convergence of their derivatives. ■

Finally, we should add one more comment. When deriving generating functions of discrete structures, we can often obtain an error term. To translate such an analytic error term into a distribution error term, we often use the Cauchy estimate. But, in some situations the following estimate of Berry and Essén is very useful (cf. below the proof of Theorem 8.29). The proof can be found in Durrett [93] or Feller [99]. We write $\| f \|_\infty := \sup_x |f(x)|$.

Lemma 8.28 (Berry-Essén Inequality) *Let F and G be distribution functions with characteristic functions $\phi_F(t)$ and $\phi_G(t)$. Assume that G has a bounded derivative. Then*

$$\| F - G \|_\infty \leq \frac{1}{\pi}\int_{-T}^{T}\left|\frac{\phi_F(t) - \phi_G(t)}{t}\right| dt + \frac{24}{\pi}\frac{\| G' \|_\infty}{T} \tag{8.135}$$

for any $T > 0$.

These analytic tools allow to derive many convergence results in probability theory, analytic combinatorics and analysis of algorithms. For example, consider the *central limit theorem*. It concerns the sum $S_n = X_1 + \cdots + X_n$ of i.i.d. random variables with mean μ and variance σ^2. Consider the normalized random variable

$$S_n^* = \frac{1}{\sigma\sqrt{n}} \sum_{i=1}^{n} \bar{X}_i$$

where $\bar{X}_i = X_i \mu$. The CF of S_n^* is

$$\phi_{S_n^*}(t) = \phi_{\bar{X}}^n\left(\frac{t}{\sigma\sqrt{n}}\right).$$

But Taylor's expansion of $\phi_X(\tau)$ around $\tau = 0$ gives

$$\phi_{\bar{X}}^n\left(\frac{t}{\sigma\sqrt{n}}\right) = \left(1 - \frac{t^2}{2n} + o(t^2/2n)\right)^n \to e^{-t^2/2}$$

which proves the central limit theorem.

In the analysis of algorithms, we often deal with weakly dependent random variables $X_1, \ldots, X_n$, Y such that

$$Z_n = Y + X_1 + \cdots + X_n$$

satisfies the following asymptotic formula for the generating function of Z_n

$$G_{Z_n}(z) = G_Y(z)[G_X(z)]^{\beta_n}\left(1 + O(\kappa_n^{-1})\right)$$

where $\beta_n, \kappa_n \to \infty$ as $n \to \infty$ and $G_X(z)$ is the probability generating function of X_1 Can we infer the central limit law from this asymptotic expression? The answer is *yes* as demonstrated by Bender [31], Bender and Richmond [32], Hwang [164], and Flajolet and Soria [114]. We follow below Hwang's presentation.

Theorem 8.29 (Hwang 1994) *Assume that the moment generating functions $M_n(s) = \mathbf{E}[e^{sX_n}]$ of a sequence of random variables X_n are analytic in the disc $|s| < \rho$ for some $\rho > 0$, and satisfy there the expansion*

$$M_n(s) = e^{\beta_n U(s) + V(s)}\left(1 + O\left(\frac{1}{\kappa_n}\right)\right) \tag{8.136}$$

for $\beta_n, \kappa_n \to \infty$ as $n \to \infty$, and $U(s), V(s)$ are analytic in $|s| < \rho$. Assume also that $U''(0) \neq 0$. Then

$$\begin{aligned} \mathbf{E}[X_n] &= \beta_n U'(0) + V'(0) + O(\kappa_n^{-1}), &(8.137)\\ \mathbf{Var}[X_n] &= \beta_n U''(0) + V''(0) + O(\kappa_n^{-1}), &(8.138) \end{aligned}$$

and for any fixed x

$$\Pr\left\{\frac{X_n-\beta_n U'(0)}{\sqrt{\beta_n U''(0)}}\le x\right\}=\Phi(x)+O\left(\frac{1}{\kappa_n}+\frac{1}{\sqrt{\beta_n}}\right) \tag{8.139}$$

where $\Phi(x)$ is the distribution function of the standard normal distribution.

Proof. The convergence in distribution and in moments follow from Theorem 8.25 and Corollary 8.27, so we need only to derive the rate of convergence. We shall use the Berry-Essén estimate with $T_n=\sqrt{\beta_n}$ to prove that

$$\| F_n-\Phi\|_\infty\le\frac{1}{\pi}\int_{T_n}^{T_n}\left|\frac{M_n^*(it)-e^{-t^2/2}}{t}\right|dt+\frac{24}{\pi}\frac{c}{T_n}=O\left(\kappa_n^{-1}+\beta_n^{-1/2}\right) \tag{8.140}$$

where c is a constant, and $M_n^*(s)$ stands for the MGF of $X_n^*=(X_n-\beta_n U'(0))/\sqrt{\beta_n U''(0)}$. Indeed, setting $\sigma_n=\sqrt{\beta_n U''(0)}$, and using (8.136), we obtain

$$\begin{aligned}
M_n^*(it) &= \exp\left(-\frac{\beta_n U'(0)}{\sigma_n}it+U\left(\frac{it}{\sigma_n}\right)\beta_n+V\left(\frac{it}{\sigma_n}\right)\right)\left(1+O(\kappa_n^{-1})\right)\\
&= \exp\left(-\frac{t^2}{2}+O\left(\frac{|t|+|t|^3}{\sigma_n}\right)\right)\left(1+O(\kappa_n^{-1})\right) \qquad |t|\le\sqrt{\beta_n}.
\end{aligned}$$

In the above, we used the following Taylor expansions valid for $|t|\le\sqrt{\beta_n}$

$$\begin{aligned}
U\left(\frac{it}{\sigma_n}\right) &= \frac{it}{\sigma_n}U'(0)-\frac{t^2}{2\sigma_n}U''(0)+O\left(\frac{|t|^3}{\sigma_n}\right),\\
V\left(\frac{it}{\sigma_n}\right) &= O\left(\frac{|t|}{\sigma_n}\right).
\end{aligned}$$

Since $|e^w-1|\le|w|e^{|w|}$, we find

$$\begin{aligned}
|M_n^*(it)-e^{-t^2/2}| &= O\left(\left(\frac{|t|+|t|^3}{\sigma_n}+\frac{1}{\kappa_n}\right)\exp\left(-\frac{t}{2}+O\left(\frac{|t|+|t|^3}{\sigma_n}+\frac{1}{\kappa_n}\right)\right)\right)\\
&= O\left(\left(\frac{|t|+|t|^3}{\sigma_n}+\frac{1}{\kappa_n}\right)e^{-t^2/4}\right) \qquad |t|\le\beta_n.
\end{aligned}$$

After simple algebra, this leads to (8.140) and the proof is complete. ∎

Example 8.22 *Pattern Occurrences Revisited*

We shall continue the discussion of the problem from Section 7.6.2 where we analyzed the number of occurrences O_n of a given pattern H in a random text. In particular, we extend Theorem 7.14 and prove the following: *Let $r=\mathbf{E}[O_n]+x\sqrt{\mathbf{Var}[O_n]}$ for $x=O(1)$, then*

$$\Pr\{O_n(H)=r\}=\frac{1}{\sqrt{2\pi c_1 n}}e^{-\frac{1}{2}x^2}\left(1+O\left(\frac{1}{\sqrt{n}}\right)\right)\ , \tag{8.141}$$

where c_1 is defined in (7.95) of Theorem 7.14. Let $\mu_n = \mathbf{E}[O_n(H)] = (n-m+1)P(H)$ and $\sigma_n^2 = \mathbf{Var}[O_n(H)] \sim c_1 n$. By Cauchy's theorem

$$T_n(u) = \frac{1}{2\pi i}\oint \frac{T(z,u)}{z^{n+1}}dz = \frac{1}{2\pi i}\oint \frac{uP(H)}{D_H^2(z)(1-uM_H(z))z^{n+1-m}}dz ,$$

where the integration is along a circle around the origin. To evaluate this integral we enlarge the circle of integration to a bigger one, say $R > 1$, such that the bigger circle contains the dominating pole of the integrand function. By Cauchy's estimate we conclude that the integral over the bigger circle is $O(R^{-n})$. We need now to evaluate the residues inside the circle. Let us now substitute $u = e^t$ and $z = e^\rho$ (so that $T_n(e^t)$ becomes the moment generating function of O_n). Then, the poles of the integrand are the roots of the equation

$$1 - e^t M_H(e^\rho) = 0.$$

This equation implicitly defines, in some neighborhood of $t = 0$, a unique differentiable function $\rho(t)$, satisfying $\rho(0) = 0$. Notably, all other roots ρ satisfy $\inf|\rho| = \rho' > 0$. Then, the residue theorem with $e^{\rho'} > R > e^\rho > 1$ leads to

$$T_n(e^t) = C(t)e^{-(n+1-m)\rho(t)} + O(R^{-n}) \tag{8.142}$$

where

$$C(t) = \frac{P(H)}{D_H^2(\rho(t))M_H'(\rho(t))}.$$

But (8.142) satisfies the assumption of Theorem 8.29, which implies the desired result.

Finally, we observe that (8.142) can be used to obtain the large deviations result. Indeed, we conclude that

$$\lim_{n\to\infty} \frac{\log T_n(e^t)}{n} = -\rho(t) .$$

Thus, directly from the Gärtner-Ellis Theorem 5.20 we obtain for $a > P(H)$

$$\lim_{n\to\infty} \frac{\log \Pr\{O_n > na\}}{n} = -I(a) ,$$

where

$$I(a) = a\omega_a + \rho(\omega_a)$$

with ω_a being a solution of

$$-\rho'(\omega_a) = a.$$

In passing we observe that this can be also derived directly by the shift of mean method. In Exercise 21 we propose a stronger version of the large deviations result. $\square$

8.7 Applications

It is time again to apply what we have learned so far. Since this chapter is very long, we will be rather sketchy. We start with a combinatorial problem on words where we find a set of words that approximately self-overlap. This problem combines methodologies of the last two chapters. Then we plunge into singularity analysis to provide an asymptotic expansion for the minimax redundancy for memoryless sources. Finally, in our most sophisticated application we derive the limiting distribution of the depth in a digital search tree.

8.7.1 Variations on Approximate Self-Overlapping Words

We are interested here in the structure of a word $w_k := w_1^k \in \mathcal{A}^k$ of length k such that when shifted by, say s, the shifted word is within a given Hamming distance from the original (unshifted word). More precisely, let the set of all words of length k be denoted as $\mathcal{W}_k := \mathcal{A}^k$. A prefix of length $q \leq k$ of w_k is written $\overline{w}_k(q)$ or simple $\overline{w}_k$ if there is no confusion. The distance between words is understood as the relative Hamming distance, that is, $d_n(x_1^n, \widetilde{x}_1^n) = n^{-1}\sum_{i=1}^n d_1(x_i, \widetilde{x}_i)$ where $d_1(x, \widetilde{x}) = 0$ for $x = \widetilde{x}$ and 1 otherwise $(x, \widetilde{x} \in \mathcal{A})$. We also write $M(x_1^n, \widetilde{x}_1^n) = nd_n(x_1^n, \widetilde{x}_1^n)$ for the number of mismatches between x_1^n and $\widetilde{x}_1^n$. We now fix $D > 0$. Consider a word $w_{k+s} = w_1^{k+s}$ of length $k+s$, and shift it by $s \leq k$. The shifted word of length k is w_s^{k+s}. We are interested in the set $\mathcal{W}_{k,s}(D)$ of all words w_{k+s} such that

$$d(w_1^k, w_s^{k+s}) \leq D \ . \tag{8.143}$$

This problem is well understood for the "faithful" (lossless) overlapping case, that is, when $D = 0$. Then, for $m = \lfloor k/s \rfloor$ (cf. Lemma 4.10 and [237, 324, 325])

$$\begin{aligned} \mathcal{W}_{k,s}(0) &= \{w_s \in \mathcal{W}_s : \ w_{k+s} = w_s^{(m+1)}\overline{w}_s \\ &= \bigcup_{w_s \in \mathcal{W}_s} \{w_s^{(m+1)}\overline{w}_s\} \end{aligned}$$

where $\overline{w}_s$ is a prefix of length $q = k - m \cdot s$, and $w_s^{(m)}$ is a concatenation of m words w_s.

We construct now all words w_{k+s} that belongs to $\mathcal{W}_{k,s}(D)$. First, let us define an integer ℓ such that $\ell/k \leq D < (\ell+1)/k$. Also, we write $k = s \cdot m + q$ where $0 \leq q < s$. Take now $0 \leq l \leq \ell$, and partition the integer l into $m+1$ integer terms as follows:

$$l = a_1 + a_2 + \cdots + a_m + \widetilde{a}_{m+1} \qquad 0 \leq a_i \leq s \quad \text{for} \quad 1 \leq i \leq m \tag{8.144}$$

and $0 \leq \widetilde{a}_{m+1} \leq q$. Let the set of all such partitions be denoted as $\mathcal{P}_{k,s}(l)$. We now define recursively m sets $\mathcal{W}_s(a_i)$ for $i \leq m$. We set $\mathcal{W}_s(a_0) := \mathcal{W}_s$ where $a_0 = 0$, and then

$$\mathcal{W}_s(a_k) = \{v_s \in \mathcal{W}_s : \ M(w_s, v_s) = a_k \quad \text{for} \quad w_s \in \mathcal{W}_s(a_{k-1})\} \ ,$$

together with

$$\overline{\mathcal{W}}_q(\tilde{a}_{m+1}) = \{v_q \in \mathcal{W}_q : \ M(\overline{w}_s(q), v_q) = \tilde{a}_{m+1} \quad \text{for} \quad w_s \in \mathcal{W}_s(a_m)\} \ .$$

Thus

$$\mathcal{W}_{k,s}(D) = \bigcup_{l=0}^{l} \{W_{k,s}(l)\}$$

such that

$$\mathcal{W}_{k,s}(l) = \bigcup_{w_s^0 \in \mathcal{W}_s} \bigcup_{\mathcal{P}_{s,k}(l)} \bigcup_{w_s^1 \in \mathcal{W}_s(a_1)1} \cdots \bigcup_{w_s^m \in \mathcal{W}_s(a_m)} \bigcup_{\overline{w}_s^{m+1} \in \overline{\mathcal{W}}_s(\tilde{a}_{m+1})} w_s^0 w_s^1 \cdots w_s^m \overline{w}_s^{m+1} \qquad (8.145)$$

where $w_s^0 w_s^1 \cdots w_s^m \overline{w}_s^{m+1}$ is the concatenation of words w_s^0 and $\ldots$ $\overline{w}_s^{m+1}$. Summing over all ℓ we obtain $\mathcal{W}_{k,s}(D)$.

In some applications the cardinality of $\mathcal{W}_{k,s}(D)$ and $\mathcal{P}_{s,k}(l)$ are of interest. Let $G(z)$ be the generating function of the cardinality $|\mathcal{P}_{s,k}(l)|$ of $\mathcal{P}_{s,k}(l)$. By (8.144), we immediately obtain

$$\begin{aligned} G(z) &= (1 + x + x^2 + \cdots + x^s)^m (1 + x + x^2 + \cdots + x^q) && (8.146) \\ &= \frac{(1 - x^{s+1})^m (1 - x^{q+1})}{(1 - x)^{m+1}} , && (8.147) \end{aligned}$$

where $m = \lfloor k/s \rfloor$ and $q = k - ms$. Let now $e_l = |\mathcal{P}_{s,k}(l)| = [z^l]G(z)$. As in Comtet [56] (cf. Ex. 16 page 77) we introduce the *polynomial coefficients* $\binom{n,q}{k}$ defined as

$$(1 + x + \cdots + x^{q-1})^n = \sum_{k=0}^{\infty} \binom{n,q}{k} x^k \qquad (8.148)$$

for integer q. Note that $\binom{n,2}{k} = \binom{n}{k}$. Observe that

$$\begin{aligned} e_l = |\mathcal{P}_{s,k}(l)| &= \sum_{j=0}^{q} \binom{m, s+1}{l-j} \\ &= \sum_{j=0}^{q} \sum_{(s+1)i+t=l-j} (-1)^i \binom{m}{i} \binom{m+t}{m} \end{aligned}$$

where $m = \lfloor k/s \rfloor$ and $q = k - ms$.

The enumeration of $\mathcal{W}_{k,s}$ is even easier since

$$|\mathcal{W}_{k,s}(l)| = 2^s \sum_{a_1+a_2+\cdots+a_m+a_{m+1}=l} \binom{s}{a_1} \cdots \binom{s}{a_m} \binom{q}{a_{m+1}} = 2^s \binom{k}{l} \ .$$

The above follows directly from $(1+x)^s(1+x)^s\cdots(1+x)^s(1+x)^q = (1+x)^{ms+q} = (1+x)^k$.

The next interesting question is how to get asymptotics for e_l and $\binom{n,q}{k}$. We prove the following result. Let $g(z) = (\frac{1}{q} + \frac{z}{q} + \cdots + \frac{z^{q-1}}{q})$ be the probability generating function so that

$$[z^k]q^n g^n(z) = \binom{n,q}{k}.$$

From the Cauchy formula we have

$$\binom{n,q}{k} = \frac{q^n}{2\pi i}\oint \frac{g(z)^n}{z^{k+1}}dz$$

where the path of integration encloses the origin. Judging from the binomial coefficients (i.e., $q=2$) we should expect different asymptotics for various values of k (e.g., bounded k, k around the mean $n\mu = n(q-1)/2$, and $k = \alpha n$ where $\alpha \neq (q-1)/2$). This is summarized in the lemma below.

Lemma 8.30 *For any q and large n the following holds.*

(i) *If $k = n(q-1)/2 + r$ where $r = o(\sqrt{n})$, then*

$$\binom{n,q}{k} \sim \frac{q^n}{\sigma\sqrt{2\pi n}}\exp\left(-\frac{r^2}{2n\sigma^2}\right) \tag{8.149}$$

where $\sigma^2 = (q^2-1)/12$. In particular (cf. Comtet [56] [Ex. 16, p.77]*),*

$$\sup_k \binom{n,q}{k} = \binom{n,q}{n(q-1)/2} \sim q^n\sqrt{\frac{6}{(q^2-1)\pi n}}\,. \tag{8.150}$$

(ii) *If $k = \alpha n$ where $\alpha \neq (q-1)/2$, then*

$$\binom{n,q}{k} \sim \frac{g(\beta)^n}{\beta^{\alpha n}}\frac{1}{\sigma_\alpha\sqrt{2\pi n}} \tag{8.151}$$

where β is a solution of $\beta g'(\beta) = \alpha g(\beta)$ and $\sigma_\alpha^2 = \beta^2 g''(\beta)/g(\beta) + \alpha - \alpha^2$.

(iii) *If $k = O(1)$, then*

$$\binom{n,q}{k} \sim \frac{n^k}{k!} \tag{8.152}$$

Proof. Part (i) follows directly from the saddle point method in exactly the same manner as we did in Section 8.4.3 so we omit details. Setting $r=0$ in (8.149) we obtain (8.150).

Comtet [56] suggested also another derivation of it. Namely, note that after the substitution $z = e^{ix}$, the Cauchy formula yields

$$\binom{n,q}{k} = \frac{1}{\pi}\int_{-\pi/2}^{\pi/2}\left(\frac{\sin(qx)}{\sin(x)}\right)^n \cos(x(n(q-1)-2k))dx \ .$$

Observe that for $k = n(q-1)/2$ the cosine function achieves its maximum, hence by a simple application of Laplace's method we obtain again (8.150).

Part (ii) follows from (i) and the method of shifted mean as discussed in Section 8.4.3. More precisely, we use

$$[z^{\alpha n}](g(z))^n = \frac{g(\beta)^n}{\beta^{\alpha n}}\left(\frac{g(\beta z)}{g(\beta)}\right)^n$$

and find β to be $\beta g_1'(\beta) = \alpha g_1(\beta)$.

Part (iii) can be proved as follows. From the Cauchy integral we have after, substituting $z = w/n$

$$\begin{aligned}\binom{n,q}{k} &= \frac{1}{2\pi i}\oint \frac{G(z)^n}{z^{k+1}}dz \\ &= \frac{1}{2\pi i}\oint \frac{(1+w/n+\cdots+(w/n)^{q-1})^n}{w^{k+1}}n^k dw \to \frac{n^k}{2\pi i}\oint \frac{e^w}{w^{k+1}} = \frac{n^k}{k!} \ .\end{aligned}$$

This completes the proof. ∎

8.7.2 Redundancy Rate for Memoryless Sources

We briefly discussed redundancy of source coding in Section 6.4, Example 16 of Chapter 7, and Example 8.7 of this chapter. Here, we show how the analytic tools presented in this chapter can be used to evaluate the redundancy. We start with a brief introduction, but the reader may want also to review Chapter 6. A code $\mathscr{C}_n : \mathcal{A}^n \to \{0,1\}^*$ is defined as a mapping from the set $\mathcal{A}^n$ of all sequences of length n over the finite alphabet $\mathcal{A}$ to the set $\{0,1\}^*$ of all binary sequences. Given a probabilistic source model, we let $P(x_1^n)$ be the probability of the message x_1^n while for a given code $\mathscr{C}_n$ we denote by $L(\mathscr{C}_n, x_1^n)$ the code length for x_1^n.

Shannon's Theorem 6.18 asserts that the entropy $H_n(P) = -\sum_{x_1^n} P(x_1^n)\log_2 P(x_1^n)$ is the lower bound on the expected code length. Hence, $-\log_2 P(x_1^n)$ can be viewed as the "ideal" code length. The next natural question is to ask by how much the length $L(\mathscr{C}_n, x_1^n)$ of a code differs from the ideal code length, either for individual sequences or on average. The *pointwise redundancy* $R_n(\mathscr{C}_n, P; x_1^n)$ and the *average redundancy* $\bar{R}_n(\mathscr{C}_n, P)$ are defined as

$$\begin{aligned}R_n(\mathscr{C}_n, P; x_1^n) &= L(\mathscr{C}_n, x_1^n) + \log_2 P(x_1^n) \\ \bar{R}_n(\mathscr{C}_n, P) &= \mathbf{E}_P[R_n(\mathscr{C}_n, P; X_1^n)] = \mathbf{E}[L(\mathscr{C}_n)] - H_n(P).\end{aligned}$$

Another natural measure of code performance is the *maximal* redundancy defined as

$$R_n^*(\mathscr{C}_n, P) = \max_{x_1^n}\{R_n(\mathscr{C}_n, P; x_1^n)\}.$$

Observe that while the pointwise redundancy can be negative, maximal and average redundancies cannot, by Shannon's Theorem 6.18 and Kraft's inequality (cf. Theorem 6.17).

In practice, the source probabilities are unknown, hence there is the desire to design codes for a whole class of source models $\mathcal{S}$. When the source is known, the redundancy can be as low as 1 bit, as demonstrated by Shannon-Fano codes (cf. Example 7). Therefore, for unknown probabilities, the redundancy rate can be also viewed as the penalty paid for estimating the underlying probability measure. More precisely, *universal codes* are those for which the redundancy is $o(n)$ for all $P \in \mathcal{S}$. The (asymptotic) *redundancy-rate problem* consists in determining for a class $\mathcal{S}$ of source models the rate of growth of the minimax quantities

$$R_n^*(\mathcal{S}) = \min_{\mathscr{C}_n} \max_{P\in\mathcal{S}}\{R_n^*(\mathscr{C}_n, P)\}$$

as $n \to \infty$. This minimax redundancy be viewed as the additional "price" on top of entropy incurred (at least) by any code in order to be able to cope with *all* sources.

In passing we would like to add that the redundancy rate problem is typical of a situation where second-order asymptotics play a crucial role since the leading term of $\mathbf{E}[L(\mathscr{C}_n)]$ is known to be nh, where h is the entropy rate. This problem is an ideal candidate for "analytic information theory" that applies analytic tools to information theory.

Hereafter, we aim at establishing a precise asymptotic expansion of $R^*(\mathcal{M})$ for memoryless sources $\mathcal{M}$. However, first we review Shtarkov's result [308] concerning the redundancy

$$R_n^*(\mathcal{S}) = \min_{\mathscr{C}_n} \sup_{P\in\mathcal{S}} \max_{x_1^n}\{L(\mathscr{C}_n, x_1^n) + \log_2 P(x_1^n)\}$$

for general class of sources $\mathcal{S}$, where the supremum is taken over all distributions P. Shtarkov proposed the following bound

$$\log_2\left(\sum_{x_1^n}\sup_P P(x_1^n)\right) \le R_n^*(\mathcal{S}) \le \log_2\left(\sum_{x_1^n}\sup_P P(x_1^n)\right) + 1. \tag{8.153}$$

To prove it, Shtarkov first introduced a new probability distribution, namely,

$$q(x_1^n) := \frac{\sup_P P(x_1^n)}{\sum_{x_1^n}\sup_P P(x_1^n)}.$$

By Kraft's inequality (cf. Theorem 6.17) there exists $\widetilde{x}_1^n$ such that

$$-L(\mathscr{C}_n, \widetilde{x}_1^n) \le \log_2 q(\widetilde{x}_1^n).$$

The above must be true since otherwise the Kraft inequality would be violated. The lower bound follows. For the upper bound, Shtarkov proposed the Shannon-Fano code $\widetilde{\mathscr{C}}_n$ for the distribution $q(x_1^n)$ defined above whose length is

$$L(\widetilde{\mathscr{C}}_n, x_1^n) = \left\lceil \log_2 \left(\sum_{x_1^n} \sup_P P(x_1^n) \right) - \log_2 \left(\sup_P P(x_1^n) \right) \right\rceil .$$

This gives the desired upper bound.

Now, we concentrate on studying

$$R_n := \log_2 \left(\sum_{x_1^n} \sup_P P(x_1^n) \right) .$$

for a class of memoryless sources over a m-ary alphabet $\mathcal{A}(m)$. Observe that R_n is within distance one from the minimax redundancy R_n^*. We shall prove that

$$R_n = \log D_n(m)$$

where $D_n(m)$ satisfies

$$D_n(m) = \sum_{i=1}^{m} \binom{m}{i} D_n^*(i) \tag{8.154}$$

with $D_0^*(1) = 0$, $D_n^*(1) = 1$ for $n \geq 1$, and for $i > 1$ we have

$$D_n^*(i) = \sum_{k=1}^{n} \binom{n}{k} \left(\frac{k}{n}\right)^k \left(1 - \frac{k}{n}\right)^{n-k} D_{n-k}^*(i-1) \ . \tag{8.155}$$

Indeed, we first observe that we can partition all sequences into m sets such that the ith set is built from sequences composed of symbols from the alphabet $\mathcal{A}(i)$ consisting of any i symbols from $\mathcal{A}(m)$. This implies (8.154) since

$$D_n(m) \quad = \quad \sum_{i=1}^{m} \binom{m}{i} \sum_{x^n \in \mathcal{A}(i)} \sup P(x^n) = \sum_{i=1}^{m} \binom{m}{i} D_n^*(i)$$

where $D_n^*(i)$ is defined above. To derive the recurrence of $D_n^*(i)$ we argue as follows: Consider an alphabet $\mathcal{A}(i-1)$ and assume that these $i-1$ symbols of $\mathcal{A}(i-1)$ occur on $n-k$ positions of x_1^n that contributes $D_{n-k}^*(i-1)$ to the sum. On the remaining k positions we place the ith symbol with the maximum probability which for a memoryless source is

$$\sup P(x_1^n) = \left(\frac{k}{n}\right)^k \left(1 - \frac{k}{n}\right)^{n-k} .$$

This establishes the desired recurrence (8.155).

We now prove an asymptotic expansion of $D_n^*(m)$ and $D_n(m)$.

Theorem 8.31 (Szpankowski 1998) *For fixed $m \geq 1$ the quantity $D_n^*(m)$ attains the following asymptotics*

$$\begin{aligned} D_n^*(m) &= \frac{\sqrt{\pi}}{\Gamma(\frac{m}{2})}\left(\frac{n}{2}\right)^{\frac{m}{2}-\frac{1}{2}} - \frac{\sqrt{\pi}}{\Gamma(\frac{m}{2}-\frac{1}{2})}\left(\frac{2m}{3}\right)\left(\frac{n}{2}\right)^{\frac{m}{2}-1} \\ &+ \frac{\sqrt{\pi}}{\Gamma(\frac{m}{2})}\left(\frac{n}{2}\right)^{\frac{m}{2}-\frac{3}{2}}\left(\frac{3+m(m-2)(8m-5)}{72}\right) \\ &+ O(n^{\frac{m}{2}-2}). \end{aligned} \tag{8.156}$$

Furthermore, for $m \geq 2$

$$\begin{aligned} R_n &= \log D_n(m) = \frac{m-1}{2}\log\left(\frac{n}{2}\right) + \log\left(\frac{\sqrt{\pi}}{\Gamma(\frac{m}{2})}\right) \\ &+ \frac{\Gamma(\frac{m}{2})m}{3\Gamma(\frac{m}{2}-\frac{1}{2})}\cdot\frac{\sqrt{2}}{\sqrt{n}} \\ &+ \left(\frac{3+m(m-2)(2m+1)}{36} - \frac{\Gamma^2(\frac{m}{2})m^2}{9\Gamma^2(\frac{m}{2}-\frac{1}{2})}\right)\cdot\frac{1}{n} \\ &+ O\left(\frac{1}{n^{3/2}}\right) \end{aligned} \tag{8.157}$$

for large n.

Proof. Observe that the recurrence (8.155) is the same as the one analyzed in Example 16 of Chapter 7. In particular, we prove there that the tree-like generating function

$$D_m^*(z) = \sum_{k=0}^{\infty} \frac{k^k}{k!} z^k D_k^*(m)$$

satisfies (cf. (7.37))

$$D_m^*(z) = (B(z)-1)^m$$

where $B(z)$ is related to the tree-function $T(z)$ and defined in (7.34). Observe that $B(z) = 1/(1-T(z))$ and $[z^n]B(z) = n^n/n!$. Furthermore, by (7.38)

$$D_m(z) = B^m(z) - 1.$$

We now use the singularity analysis to obtain asymptotics of $D_n^*(m)$ and $\log_2 D_n(m)$. In Example 10 we derive asymptotic expansion of $T(z)-1$ around $z = e^{-1}$ from which we directly obtain

$$B(z) = \frac{1}{\sqrt{2(1-ez)}} + \frac{1}{3} - \frac{\sqrt{2}}{24}\sqrt{(1-ez)} + \frac{4}{135}(1-ez) - \frac{23\sqrt{2}}{1728}(1-ez)^{3/2} + O((1-ez)^2)\,. \tag{8.158}$$

Then applications of Theorem 8.12 and Theorem 8.13 yield

$$\begin{aligned} e^{-n}[z^n](B(z)-1)^m &= \frac{n^{\frac{m}{2}-1}}{2^{\frac{m}{2}}\Gamma(\frac{m}{2})} - \frac{n^{\frac{m}{2}-\frac{3}{2}}}{2^{\frac{m}{2}-\frac{1}{2}}}\left(\frac{2m}{3\Gamma(\frac{m}{2}-\frac{1}{2})}\right) \\ &+ \frac{n^{\frac{m}{2}-2}}{2^{\frac{m}{2}}}\left(\frac{m(m-2)(8m-5)}{36\Gamma(\frac{m}{2})}\right) + O(n^{\frac{m}{2}-\frac{5}{2}}) \end{aligned}$$

as needed to establish the first part of the theorem. The second part follows immediately from

$$\begin{aligned} e^{-n}[z^n]B^m(z) &= \frac{n^{\frac{m}{2}-1}}{2^{\frac{m}{2}}\Gamma(\frac{m}{2})} + \frac{n^{\frac{m}{2}-\frac{3}{2}}}{2^{\frac{m}{2}-\frac{1}{2}}}\left(\frac{m}{3\Gamma(\frac{m}{2}-\frac{1}{2})}\right) \\ &+ \frac{n^{\frac{m}{2}-2}}{2^{\frac{m}{2}}}\left(\frac{m(m-2)(2m+1)}{36\Gamma(\frac{m}{2})}\right) + O(n^{\frac{m}{2}-\frac{5}{2}}) \ . \end{aligned}$$

and

$$\log(1+a\sqrt{x}+bx+cx^{3/2}) = a\sqrt{x} + (b-\frac{1}{2}a^2)x + O(x^{3/2})$$

as $x \to 0$. This completes the proof. ∎

8.7.3 Limiting Distribution of the Depth in Digital Search Trees

We shall derive here the limiting distribution of the typical depth D_m in a digital search tree (cf. Section 1.1). We assume that m independent strings are generated by a binary memoryless source with p being the probability of emitting "0" and $q = 1-p$. Recall that the typical depth is defined as follows. Let $D_m(i)$ be the depth of the ith node in a digital tree. Actually, observe that $D_m(i) = D_i(i)$ for $m \geq i$. Clearly, for various $i \leq m$ distributions of $D_m(i)$ are different, and therefore it makes sense to define the *typical depth* D_m as

$$\Pr\{D_m < x\} = \frac{1}{m}\sum_{i=1}^{m}\Pr\{D_m(i) < x\} \ .$$

Let B_m^k denote the number of internal nodes at level k in a digital tree. To derive a recurrence on D_m we introduce the so called *average profile* $\bar{B}_m^k := \mathbf{E}[B_m^k]$ to be the average number of internal nodes at level k in a digital tree. The following obvious relationship between the depth D_m and the average profile $\bar{B}_m^k$ holds

$$\Pr\{D_m = k\} = \frac{\bar{B}_m^k}{m}. \tag{8.159}$$

This follows from the definition of D_m and the definition of $\bar{B}_m^k$.

We shall work initially with the average profile, and we define the generating function $B_m(u) = \sum_{k=0}^{\infty} \bar{B}_m^k u^k$ which satisfies the following recurrence

$$B_{m+1}(u) = 1 + u \sum_{j=0}^{m} \binom{m}{j} p^j q^{m-j} (B_j(u) + B_{m-j}(u)) \qquad (8.160)$$

with $B_0(u) = 0$. This recurrence arises naturally in our setting by considering the left and the right subtrees of the root, and noting that j strings will go to the left subtree with the probability $\binom{m}{j} p^j q^{m-j}$. A general recurrence of this type was discussed in Section 7.6.1 from which we conclude that

$$B_m(u) = m - (1-u) \sum_{k=2}^{m} (-1)^k \binom{m}{k} Q_{k-2}(u) \qquad (8.161)$$

where

$$Q_k(u) = \prod_{j=2}^{k+1} (1 - up^j - uq^j) \quad , \quad Q_0(u) = 1 \ . \qquad (8.162)$$

Actually, the derivation of (8.161) is not too complicated, so we provide a sketch of the proof. Let us start with multiplying both sides of (8.161) by $z^m/m!$ to get $B_z'(z,u) = e^z + uB(pz,u)e^{qz} + uB(qz,u)e^{pz}$ where $B(z,u) = \sum_{m=0}^{\infty} B_m(u) \frac{z^m}{m!}$, and $B_z'(z,u)$ is the derivative of $B(z,u)$ with respect to z. We now multiply this functional equation by e^{-z} and introduce $\widetilde{B}(z,u) = B(z,u)e^{-z}$. This leads to a new functional equation, namely $\widetilde{B}'(z,u) + \widetilde{B}(z,u) = 1 + u(\widetilde{B}(zp,u) + \widetilde{B}(zq,u))$. Comparing now the coefficients at z^m one immediately obtains $\widetilde{B}_{m+1}(u) = \delta_{m,0} - \widetilde{B}_m(u)(1 - up^m - uq^m)$ where $\delta_{0,m}$ is the Kronecker symbol. To prove (8.161) it only suffices to note that $B_m(u) = \sum_{k=0}^{m} \binom{m}{k} \widetilde{B}_k(u)$.

Using again the idea from Section 7.6.1, Example 19 and the Rice method we shall sketch a proof of the following result. In Exercise 22 the reader is asked to provide details.

Theorem 8.32 (Kirschenhofer and Prodinger 1988; Szpankowski 1991) *The average* $\mathbf{E}[D_m]$ *of the depth attains the following asymptotics as* $m \to \infty$

$$\mathbf{E}[D_m] = \frac{1}{h}\left(\log m + \gamma - 1 + \frac{h_2}{2h} + \theta + \delta(m)\right) + O\left(\frac{\log m}{m}\right) \qquad (8.163)$$

where h *is the entropy,* $h_2 = p\log^2 p + q\log^2 q$, $\gamma = 0.577\ldots$ *is the Euler constant, and*

$$\theta = -\sum_{k=1}^{\infty} \frac{p^{k+1}\log p + q^{k+1}\log q}{1 - p^{k+1} - q^{k+1}} \ .$$

The function $\delta(x)$ *is a fluctuating function with a small amplitude when* $\log p/\log q$ *is rational, and* $\delta(x) \equiv 0$ *for* $\log p/\log q$ *irrational. More precisely, for* $\log p/\log q = r/t$ *where* r, t *are* *integers,*

$$\delta_1(x) = \sum_{\substack{\ell=-\infty \\ \ell\neq 0}}^{\infty} \frac{\Gamma(s_0^\ell)Q(-2)}{Q(s_0^\ell - 1)} \exp\left(-\frac{2\pi i \ell r}{\log p}\log x\right) , \tag{8.164}$$

where $s_0^\ell = -1 + 2\pi i\ell r/\log p$.

(ii) *The variance of* D_m *for large* m *satisfies*

$$\mathbf{Var}[D_m] = \frac{h_2 - h^2}{h^3}\log m + A + \Delta(m) + O(\log^2 m/m) \tag{8.165}$$

where A *is a constant (cf. Exercise 22 and* $\Delta(x)$ *is a fluctuating function with a small amplitude. In the symmetric case, the coefficient at* $\log m$ *becomes zero, and then*

$$\mathbf{Var}[D_m] = \frac{1}{12} + \frac{1}{\log^2 2}\cdot\frac{\pi^2}{6} - \alpha - \beta + \Delta(\log_2 m) - [\delta_1^2]_0 + O(\log^2 m/m) \tag{8.166}$$

where

$$\alpha = \sum_{j=1}^{\infty}\frac{1}{2^j - 1} \qquad , \qquad \beta = \sum_{j=1}^{\infty}\frac{1}{(2^j-1)^2},$$

the function $\Delta(x)$ *is continuous with period* 1 *and mean zero, and* $|[\delta_1^2]_0| < 10^{-10}$ *is the mean of* $\delta_1^2(x)$.

Sketch of Proof. It suffices to derive asymptotics of the average path length $l_m := m\mathbf{E}[D_m]$ and the second factorial moment of the path length $l_m^2 := m\mathbf{E}[D_m(D_m-1)]$. But $l_m = B_m'(1)$ and $l_m^2 = B_m''(1)$. After some algebra, we arrive at the following two recurrences for all $m \geq 0$

$$l_{m+1} = m + \sum_{k=0}^{m}\binom{m}{k}p^k q^{m-k}(l_k + l_{m-k}), \tag{8.167}$$

$$l_{m+1}^2 = 2(l_{m+1} - m) + \binom{m}{k}p^k q^{m-k}(l_k^2 + l_{m-k}^2) \tag{8.168}$$

with $l_0 = l_0^2 = 0$. The recurrence (8.167) falls directly under the general recurrence (7.56) of Section 7.6.1. Its solution is

$$l_m = \sum_{k=2}^{m}(-1)^k\binom{m}{k}Q_{k-2} \tag{8.169}$$

where $Q_k = Q_k(1) = \prod_{j=2}^{k+1}(1 - p^j - q^j)$. This is an alternating sum with the kernel Q_k of a polynomial growth, so Rice's method applies. We only need analytic continuation of Q_k but this is easy. Indeed, define

$$Q(z) = \frac{P(0)}{P(z)}$$

where $P(z) = \prod_{j=2}^{\infty}(1 - p^{z+j} - q^{z+j})$. Then $Q_k = Q(z)|_{z=k}$ and by Rice's method or Theorem 8.21

$$l_m = \frac{1}{2\pi i}\int_{-\frac{3}{2}-i\infty}^{-\frac{3}{2}+i\infty} \Gamma(z) n^{-z} Q(-z-2)dz + O(\log m).$$

The integral can be evaluated in the same manner as in Example 19, so we only sketch it. Considering singularities right to the line of the integration, we must take into account singularities of the gamma function at $z_0 = 0$ and $z_{-1} = -1$ together with all roots of

$$p^{j-2-z} + q^{j-2-z} = 1$$

that we denote as $z_{k,j}$. The main contribution comes from $z_{-1} = z_{0,2} = -1$. Using the expansions

$$\begin{aligned}
\Gamma(z) &= \frac{1}{z+1} + \gamma - 1 + O(z+1), \\
n^{-z} &= n - n\log n(z+1) + O((z+1)^2), \\
\frac{1}{1-p^{-z}-q^{-z}} &= -\frac{1}{h}\frac{1}{z+1} + \frac{h_2}{2h^2} + O(z+1), \\
Q(-z-2) &= -\frac{1}{h}\frac{1}{z+1} + \frac{\theta}{h} + \frac{h_2}{2h^2} + (z+1)\frac{\theta h_2}{2h^2} + O((z+1)^2)
\end{aligned}$$

and extracting the coefficient at $(z+1)^{-1}$ we establish the leading term of the asymptotics without $\delta_1(m)$. To obtain $\delta_1(m)$ we must consider the roots at $z_{k,j}$ other than $z_{-1} = -1$. Using Lemma 8.22 we finally prove (8.163).

The asymptotics of l_m^2 can be obtained in a similar manner but with more involved algebra. To use (7.69) of Section 7.6.1 we need to find the binomial inverse to l_m. But solution (8.169) implies that $\hat{l}_m = Q_{m-2}$. With this in mind we directly obtain

$$l_m^2 = -2\sum_{k=2}^{m}\binom{m}{k}(-1)^k Q_{k-2}T_{k-2}$$

where

$$T_n = \sum_{i=2}^{n+1}\frac{p^i + q^i}{1-p^i-q^i}.$$

To apply Rice's method we need analytic continuation of T_n. A "mechanical derivation" works fine here after observing that

$$T_{n+1} = T_n + \frac{p^{n+2}+q^{n+2}}{1-p^{n+2}-q^{n+2}}.$$

Replacing n by z and solving the recurrence with respect to z we find

$$T(z) = \theta - \sum_{i=2}^{\infty} \frac{p^{z+i} + q^{z+i}}{1 - p^{z+i} - q^{z+i}}$$

where θ is defined in the statement of the theorem. The above derivation can be easily made rigorous by observing that all series involved do converge for $\Re(z) > -2$. ∎

We are now ready to derive the limiting distribution of D_m for $p \neq q$. We prove the following result.

Theorem 8.33 (Louchard and Szpankowski 1995) *In the asymmetric case when $p \neq q$, the limiting distribution of D_m is normal, that is,*

$$\frac{D_m - \mathbf{E}[D_m]}{\sqrt{\mathbf{Var}[D_m]}} \to N(0,1) \tag{8.170}$$

where $\mathbf{E}[D_m]$ and $\mathbf{Var}[D_m]$ are given by (8.163) and (8.165), respectively, and the moments of D_m converge to the appropriate moments of the normal distribution. More generally, for any complex ϑ

$$e^{-\vartheta c_1 \log m} E(e^{\vartheta D_m/\sqrt{c_2 \log m}}) = e^{\frac{\vartheta^2}{2}} \left(1 + O(\frac{1}{\sqrt{\log m}})\right) \tag{8.171}$$

where $c_1 = 1/h$ and $c_2 = (h_2 - h^2)/h^3$.

Proof. We shall work with the probability generating function $D_m(u)$ for the depth which is equal to $B_m(u)/m$, that is,

$$D_m(u) = 1 - \frac{1-u}{m} \sum_{k=2}^{m} (-1)^k \binom{m}{k} Q_{k-2}(u) \ . \tag{8.172}$$

Let $\mu_m = \mathbf{E}[D_m]$ and $\sigma_m^2 = \mathbf{Var}[D_m]$. Theorem 8.32 implies $\mu_m \sim c_1 \log m$ and $\sigma_m^2 \sim c_2 \log m$ where $c_1 = 1/h$ and $c_2 = (h_2 - h^2)/h^3$. We use Goncharov's Theorem 8.26 to establish the normal distribution of D_m by showing that

$$\lim_{m\to\infty} e^{-\vartheta \mu_m/\sigma_m} D_m(e^{\vartheta/\sigma_m}) = e^{\vartheta^2/2} \tag{8.173}$$

where ϑ is a complex number. By Corollary 8.27, this will establish both the limiting distribution and the convergence in moments for D_m.

We now derive asymptotics for the probability generating function $D_m(u)$ around $u = 1$. We assume $u = e^v$, and due to $\sigma_m = O(\sqrt{\log m})$, we define $v = \vartheta/\sigma_m \to 0$. Note that $1 - D_m(u)$ given in (8.172) has the form of an alternating sum. We now apply the Rice method, more precisely Theorem 8.21. To do so, however, we need an analytical continuation of $Q_k(u)$ defined in (8.162). Denote it as $Q(u,s)$, and observe that

$$Q(u,s) = \frac{P(u,0)}{P(u,s)} = \frac{Q_\infty(u)}{P(u,s)} \tag{8.174}$$

where $P(u,s) = \prod_{j=2}^{\infty}(1 - up^{s+j} - uq^{s+j})$. Using Theorem 8.21 we arrive at

$$1 - D_m(u) = \frac{1-u}{m2\pi i}\int_{-3/2-i\infty}^{-3/2+i\infty} \Gamma(s)m^{-s}Q(u,-s-2)ds + e_m \ , \tag{8.175}$$

where $e_m = O(1/m^2)\int_{-3/2-i\infty}^{-3/2+i\infty}\Gamma(s)m^{-s}sQ(u,-s-2)ds = O(1/m)$, as we shall see soon. As usual, we evaluate the integral in (8.175) by the residue theorem. We compute residues *right* to the line of integration in (8.175). The gamma function has its singularities at $s_{-1} = -1$ and $s_0 = 0$, and in addition we have infinite number of zero $s_{k,j}(v)$ ($j = 2, 3, \ldots$, $k = 0 \pm 1, \pm 2, \ldots$) of $P(e^v, -s-2)$ of the denominator of $Q(e^v, -s-2)$ where we substituted $u = e^v$. More precisely, $s_{k,j}(v)$ are zeros of

$$p^{-s-2+j} + q^{-s-2+j} = e^{-v} \ . \tag{8.176}$$

The dominating contribution to the asymptotics comes from $s_{0,j}(v)$ and s_1. Indeed, the contributions of the first two singularities at s_{-1} and s_0 are respectively $(1-u)Q(u,-1) = 1$ and $(1-u)Q(u,-2)/m = O(1/m)$. We now concentrate on the contribution coming from $s_{0,j}(v)$. In this case, one can solve equation (8.176) to derive

$$s_{0,j}(v) = j - 3 - \frac{v}{h} - \frac{1}{2}\left(\frac{1}{h} - \frac{h_2}{h^3}\right)v^2 + O(v^3) \tag{8.177}$$

for integer $j \geq 2$ and $v \to 0$. We also note that $\Im(s_{k,j}(v)) \neq 0$ for $k \neq 0$.

Let now $R_k^j(v)$ denote the residue of $(1 - e^v p^{-s_{k,j}-2+j} + e^v q^{-s_{k,j}-2+j})^{-1}$ at $s_{k,j}(v)$, and let $g(s) = \Gamma(s)Q(u,-s-1)$. In the sequel, we use the following expansion

$$\begin{aligned} Q(u,-s-2) &= \frac{1}{1-u(p^{-s}+q^{-s})} \cdot \frac{Q_\infty(u)}{P(u,-s-1)} \\ &= -\frac{w^{-1}}{h} - \frac{\theta}{h} + \frac{h_2}{2h^2} + w\frac{\theta h_2}{2h^2} + O(w^2) \end{aligned}$$

where $w = s - s_{0,j}(v)$. Then

$$\begin{aligned} -D_m(e^v) &= R_0^2(v)g(s_{0,2}(v))(1-e^v)m^{-1}m^{-s_{0,2}(v)} + \sum_{j=3}^{\infty} R_0^j(v)g(s_{0,j}(v))(1-e^v)m^{-1}m^{-s_{0,j}(v)} \\ &+ \sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty}\sum_{j=2}^{\infty} R_k^j(v)g(s_{k,j}(v))(1-e^v)m^{-1}m^{-s_{k,j}(v)} + O(1) \ . \end{aligned} \tag{8.178}$$

We consider the above three terms separately:

(a) $j = 2$ and $k = 0$
Set $v = \vartheta/\sigma_m = \vartheta/\sqrt{c_2 \log m}$ with $\Re(\vartheta) > 0$. Then by (8.177)

$$m^{-s_{0,2}(v)} = m \exp\left(\frac{\vartheta}{h}\sqrt{\frac{\log m}{c_2}} + \frac{\vartheta^2}{2}\right) .$$

In addition, the following holds: $R_0^2(v) = -1/h + O(v)$, and $g(s_{0,2}(v)) = -h/v + O(1)$, and finally $1 - e^{-v} = v + O(1)$. Therefore, we obtain

$$e^{-\vartheta\mu_m/\sigma_m} R_0^2(v) g(s_{0,2}(v))(1 - e^{-v}) m^{-s_{0,2}(v)-1} \to -e^{\vartheta^2/2} \qquad (8.179)$$

(b) $j \geq 3$ and $k = 0$
In this case we can repeat the analysis from case (a) to get

$$e^{-\vartheta\mu_m/\sigma_m} R_0^2(v) g(s_{0,2}(v))(1 - e^{-v}) m^{-s_{0,2}(v)-1} \to O(m^{2-j} e^{\vartheta^2/2}) , \qquad (8.180)$$

so this term is of order magnitude smaller than the first term in (8.178).

(c) $k \neq 0$
Fix $j = 2$. Observe that

$$\sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} R_k^2(v) g(s_{k,2}(v))(1 - e^{v}) m^{-1} m^{-s_{k,2}(v)} = O(vm^{-1-\Re(s_{k,2}(v))}) .$$

By the same argument as in the Lemma 8.22 we conclude that $\Re(s_{k,j}) \geq j - 3 + O(v)$. If $\Re(s_{k,j}) > j - 3 + O(v)$, then the term $O(vm^{-1-\Re(s_{k,2}(v))})$ is negligible compared to $O(m^{2-j})$. Otherwise, that is, for $s_{k,j}$ such that $\Re(s_{k,j}) = j - 3 + O(v)$ we observe that the error term is $O(vm^{j-2})$. Actually, we can do better using Jacquet and Régnier [172] observation that (the reader is asked to prove it in Exercise 23; for a generalization see Jacquet and Szpankowski [175])

$$\Re(s_{k,2}(v)) \geq s_{0,2}(\Re(v))$$

so the above sum becomes

$$\begin{aligned} \sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} R_k^2(v) g(s_{k,2}(v))(1 - e^{v}) m^{-1} m^{-s_{k,2}(v)} &= m^{-1-\Re(s_{0,2}(v))} O(vm^{\Re(s_{0,2}(v)) - s_{0,2}(\Re(v))}) \\ &= m^{-1-\Re(s_{0,2}(v))} O(vm^{-\beta v^2}) \end{aligned}$$

for some β. Finally, consider general $j \geq 3$. As in the case (b), we note that $m^{-s_{k,j}(v)}$ contributes $O(m^{2-j})$, so this term is negligible.

Putting everything together, we note that as $v = O(1/\sqrt{m}) \to 0$ for $m \to \infty$

$$e^{-\vartheta\mu_m/\sigma_m} D_m(e^{\vartheta/\sigma_m}) = e^{\vartheta^2/2}(1 + O(vm^{-\beta v^2}) + O(1/m)) \to e^{\vartheta^2/2} \qquad (8.181)$$

that proves the theorem. ■

8.8 Extensions and Exercises

8.1 Prove that if $f \sim \sum a_k \phi_k$ and $g \sim \sum b_k \phi_k$, then $\alpha f + \beta g \sim \sum(\alpha a_k + \beta b_k)\phi_k$.

8.2 Consider asymptotic expansions with respect to $\{x^k\}$ in $\mathbb{R}^+$ and $z_0 = 0$. Prove the following theorems.

Theorem 8.34 *Let $f(x) \sim \sum_{k=0}^{\infty} a_k x^k$ as $x \to 0$.*

(i) *If $g(z) \sim \sum_{k=0}^{\infty} b_k x^k$ as $x \to 0$, then $f(x)g(x) \sim \sum_{n=0}^{\infty} c_n x^n$ where $c_n = \sum_{k=0}^{n} a_k b_{n-k}$.*

(ii) *We have*

$$\int_0^x f(t)dt \sim \sum_{k=0}^{\infty} \frac{a_k}{k+1} x^{k+1}.$$

(iii) *If $f(x)$ has a continuous derivative and if $f'(x)$ possesses an asymptotic expansion, then*

$$f'(x) \sim \sum_{k=1}^{\infty} k a_k x^{k-1}$$

as $x \to 0$.

(iv) *For any power series $\sum_{k=0}^{\infty} a_k x^k$ there exists a function $f(x)$ with this as its asymptotic expansion.*

8.3 Using similar arguments to those presented in Section 8.1.2 show that

$$\begin{aligned} \log(1+x) &= \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} x^k, & x \to 0, \\ &= \log x + \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} x^{-k}, & x \to \infty. \end{aligned}$$

8.4 Prove Abel's summation formulas (8.27) and (8.28).

8.5 Consider again recurrence (8.39) for the height in a b-tries. In this exercise we will analyze it for

$$k = \frac{b+1}{b} \log_2 n + \beta$$

on the β-scale, where $\beta = O(1)$. We set

$$h_n^k = F(\beta; n) = F(k - (1 + 1/b) \log_2 n; n)$$

so that (8.39) becomes

$$F(\beta+1;n) = \left(\frac{1}{2}\right)^n \sum_{i=0}^{n} \binom{n}{i} \quad \cdot \quad F\left(\beta - \left(1+\frac{1}{b}\right)\log_2\left(\frac{i}{n}\right); i\right)$$
$$\cdot \quad F\left(\beta - \left(1+\frac{1}{b}\right)\log_2\left(1-\frac{i}{n}\right);\ n-i\right).$$

Postulating that $F(\beta;n)$ assumes an expansion of the form

$$F(\beta;n) = F_0(\beta) + \frac{1}{n}F_1(\beta) + O(n^{-2}).$$

prove that F_0 satisfies the following non-linear functional equation

$$F_0(\beta+1) = [F_0(\beta+1+1/b)]^2.$$

Next prove that a general solution of above is

$$F_0(\beta) = \exp(-c2^{-b\beta}) = \exp(-ce^{-b\beta\log 2}).$$

Finally, using matched asymptotics show that

$$c = \frac{1}{(b+1)!}.$$

8.6 Knessl and Szpankowski analyzed in [215] the height of PATRICIA trie. Among others, the following recurrence arises

$$(2-2^n)L_2(n) + 2nL_2(n-1) = 0, \qquad n \geq 3.$$

with $L_2(2) = 1$. Show that it has the following solution

$$L_2(n) = n!2^{-n^2/2}2^{n/2}\prod_{m=3}^{n}\left(\frac{1}{1-2^{1-m}}\right), \qquad n \geq 2.$$

Similarly, show that recurrence

$$(2-2^n)L_3(n) + 2nL_3(n-1) + 2\binom{n}{2}L_2(n-2) = 0, \qquad n \geq 5$$

with $L_3(3) = 1$ and $L_3(4) = 6$, leads to

$$L_3(n) = n!2^{-n^2/2}2^{3n/2}\left(\frac{1}{4} - \frac{n2^{-n}}{2} - \frac{1}{4}2^{-n}\right)\prod_{m=3}^{n}\left(\frac{1}{1-2^{1-m}}\right), \qquad n \geq 4.$$

In general, the following recurrence for $n \geq 2j-3$

$$(2-2^n)L_{j-1}(n) + 2\sum_{i=1}^{j-2}\binom{n}{i}L_{j-i}(n-i) = 0, \qquad n \geq 2j-3$$

has the following asymptotic solution

$$\begin{aligned} L_j(n) &\sim \rho_0 K_j n! 2^{-n^2/2} 2^{(j-3/2)n}, \quad j \text{ fixed}, \\ \rho_0 &= \prod_{\ell=2}^{\infty} (1-2^{-\ell})^{-1}. \end{aligned}$$

where K_j is a constant.

8.7 ⚠ Using the saddle point method and (8.43) derived in Example 5 prove the following theorem.

Theorem 8.35 (Knessl and Szpankowski 1999) *The distribution of the height of b-tries has the following asymptotic expansions for fixed b:*

(i) Right-Tail Region*:* $k \to \infty$, $n = O(1)$*:*

$$\Pr\{H_n^T \le k\} = h_n^k \sim 1 - \frac{n!}{(b+1)!(n-b-1)!} 2^{-kb}.$$

(ii) Central Regime*:* $k, n \to \infty$ *with* $\xi = n2^{-k}$, $0 < \xi < b$*:*

$$h_n^k \sim A(\xi; b) e^{n\phi(\xi;b)},$$

where

$$\begin{aligned} \phi(\xi; b) &= -1 - \log \omega_0 + \frac{1}{\xi}\left(b \log(\omega_0 \xi) - \log b! - \log\left(1 - \frac{1}{\omega_0}\right)\right), \\ A(\xi; b) &= \frac{1}{\sqrt{1 + (\omega_0 - 1)(\xi - b)}}. \end{aligned}$$

In the above, $\omega_0 = \omega_0(\xi; b)$ *is the solution to*

$$1 - \frac{1}{\omega_0} = \frac{(\omega_0 \xi)^b}{b!\left(1 + \omega_0 \xi + \frac{\omega_0^2 \xi^2}{2!} + \cdots + \frac{\omega_0^b \xi^b}{b!}\right)}.$$

(iii) Left-Tail Region*:* $k, n \to \infty$ *with* $j = b2^k - n$

$$h_n^k \sim \sqrt{2\pi n} \frac{n^j}{j!} b^n \exp\left(-(n+j)\left(1 + b^{-1} \log b!\right)\right)$$

where $j = O(1)$.

8.8 Using the following Fourier Series

$$\langle x \rangle = \frac{1}{2} - \sum_{m \in \mathbb{Z} \setminus \{0\}} c_m e^{2\pi i m x}, \qquad c_m = -\frac{i}{2\pi m},$$

prove that

$$\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \left\langle k\frac{N}{M} + y \right\rangle = \frac{1}{2} - \frac{1}{M}\left(\frac{1}{2} - \langle yM \rangle\right) + O(\rho^n)$$

where $\rho < 1$.

8.9 Using the Fourier series of $2^{-\langle x \rangle}$ prove the following

$$\frac{1}{M}\sum_{l=0}^{M-1} 2^{-\langle l/M+y \rangle} = \frac{1}{2M(1-2^{-1/M})} 2^{-\langle yM \rangle /M}$$

for any real y.

8.10 Show that for any irrational α

$$\lim_{n\to\infty} \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} 2^{-\langle \alpha k + \beta n \rangle} = \frac{1}{2\log 2}$$

for any real β.

8.11 Prove the following expansions

$$\begin{aligned}
(1-z)^{3/2} &= \frac{1}{\sqrt{\pi n^5}}\left(\frac{3}{4} + \frac{45}{32n} + O\left(\frac{1}{n^2}\right)\right), \\
\sqrt{1-z}\log(1-z)^{-1} &= -\frac{1}{\sqrt{\pi n^3}}\left(\frac{1}{2}\log n + \frac{\gamma + 2\log 2 - 2}{2} + O\left(\frac{\log n}{n}\right)\right), \\
\log^2(1-z)^{-1} &= \frac{1}{n}\left(2\log n + 2\gamma + +O\left(\frac{1}{n}\right)\right), \\
(1-z)^{-1/2} &= \frac{1}{\sqrt{\pi n}}\left(+\frac{8}{n} + O\left(\frac{1}{n^2}\right)\right), \\
(1-z)^{-3/2} &= \frac{n}{\sqrt{\pi}}\left(2 + \frac{3}{4n} + O\left(\frac{1}{n^2}\right)\right), \\
(1-z)^{-3/2}\log(1-z)^{-1} &= \frac{n}{\sqrt{\pi}}\left(2\log n + 2\gamma + 4\log 2 - 2 + O\left(\frac{\log n}{n}\right)\right).
\end{aligned}$$

8.12 Analyze asymptotically

$$A(z) = \frac{e^z}{\sqrt{1-z^2}}.$$

Notice that there are two algebraic singularities at $z = 1$ and $z = -1$. Prove that

$$[z^n]A(z) \sim \frac{1}{\sqrt{2\pi n}}\left(e + (-1)^n e^{-1}\right).$$

Find a full asymptotic expansion.

8.13 ⚠ Let for some $k \geq 0$

$$S_{n,k} = \sum_{i=1}^{n} \binom{n-k}{i} \left(\frac{i}{n}\right)^i \left(1 - \frac{i}{n}\right)^{n-i}.$$

(i) Prove that for fixed k

$$S_{n,k} = \frac{1}{2^{2k}} \binom{2k}{k} \sqrt{\frac{n\pi}{2}} - \frac{2}{3} + O\left(\frac{1}{\sqrt{n}}\right)$$

for large n (Szpankowski 1995).

(ii) Find a *uniform* (with respect to k) asymptotic expansion of $S_{n,k}$.

8.14 ⚠ Consider the following multidimensional extension of the Laplace method. Let $\mathbf{s} = (s_1, \ldots, s_n)$ and for large positive x

$$I(x) = \int_S f(\mathbf{s}) e^{-xh(\mathbf{s})} d\mathbf{s}.$$

Let $\mathbf{s}_0$ be a unique point where $h(\mathbf{s})$ attains minimum. We further assume that second partial derivatives of h exist and are continuous and the Hessian matrix $\mathsf{H} = \left(\frac{\partial^2 h(\mathbf{s}_0)}{\partial s_i \partial s_j}\right)$ is positive definite. Prove the following (cf. [158]).

Theorem 8.36 *Under the above hypothesis (see also hypothesis of Theorem 8.17)*

$$I(x) \sim \frac{f(\mathbf{s}_0)}{\sqrt{\det \mathsf{H}}} e^{-xh(\mathbf{s}_0)} \left(\frac{2\pi}{x}\right)^{n/2}$$

as $x \to \infty$.

8.15 ⚠ The *involution numbers* I_n has the following exponential generating functions

$$A(z) = e^{z+z^2/2}.$$

Using Hayman's method prove that

$$I_n = n! \frac{e^{-1/4}}{\sqrt{2\pi n}} n^{-n/2} e^{n/2+\sqrt{n}}$$

(originally due to Moser and Wyman).

8.16 Using the van der Waerden method prove that

$$I(n) = \int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty} \exp(-n(z^2 + 2/z)) dz$$

(the line of integration is $\Im z = \frac{1}{2}$) yields

$$I(n) \sim \sqrt{\frac{\pi}{3n}} \exp\left(n\left(\frac{3}{2} + i\frac{3\sqrt{3}}{2}\right)\right).$$

8.17 Provide a rigorous proof of the result from Example 22 through an application of the van der Waerden method.

8.18 Prove all the claims of the Jacquet-Schachinger Lemma 8.22.

8.19 ⚠ Let

$$S_n(\lambda) = \sum_{k=1}^{n} (-1)^n \binom{n}{k} k^{-\lambda}$$

where $\lambda \notin \mathbb{Z}$. Prove the Flajolet and Sedgewick [117] result, namely

$$S_n(\lambda) = -(\log n)^\lambda \sum_{j=0}^{\infty} (-1)^j \frac{\Gamma^{(j)}(1)}{j!\Gamma(1+\lambda-j)} \frac{1}{\log^j n}$$

where $\Gamma^{(j)}(1)$ is the jth derivative of the gamma function at $z = 1$.

8.20 ⚠ (Flajolet and Sedgewick 1995) Let

$$Y_n = \sum_{k=1}^{n} (-1)^n \binom{n}{k} \log k.$$

Prove that

$$Y_n = \log\log n + \gamma + \frac{\gamma}{\log n} - \frac{\pi^2 + 6\gamma^2}{12\log^2 n} + O(\log^{-3} n).$$

8.21 Consider the exact pattern matching problem of Section 7.6.2. Prove the following stronger version of the large deviations. Let $r = (1+\delta)\mathbf{E}[O_n]$ and $a = (1+\delta)P(H)$ with $\delta \neq 0$. For complex t, define $\rho(t)$ to be the root of

$$1 - e^t M_H(e^\rho) = 0 \ ,$$

while ω_a and σ_a are defined as

$$\begin{aligned} -\rho'(\omega_a) &= a, \\ -\rho''(\omega_a) &= \sigma_a^2. \end{aligned}$$

Then

$$\Pr\{O_n(H) = (1+\delta)\mathbf{E}[O_n]\} = \frac{1}{\sigma_a\sqrt{2\pi(n-m+1)}} e^{-(n-m+1)I(a)} \left(1 + O\left(\frac{1}{n}\right)\right)$$

where $I(a) = a\omega_a + \rho(\omega_a)$.

8.22 (Szpankowski 1991) Using Rice's method or Theorem 8.21 prove Theorem 8.32. In particular, establish a formula on the constant A in (8.165).

8.23 (Jacquet and Régnier) Let $s_{j,k}$ ($j = 1, 2, \ldots$ and $k = 0, \pm 1, \ldots$) be a solution of (8.176). Prove that

$$\Re(s_{k,2}(v)) \geq s_{0,2}(\Re(v))$$

8.24 (Flajolet 1999) Using singularity analysis and the method of Section 8.5 prove that

$$\sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} \frac{1}{4^k} \binom{2k}{k} = \frac{1}{\sqrt{pn}} \left(1 - \frac{3p-2}{8pn} + O\left(\frac{1}{n^3}\right)\right),$$

$$\sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} H_k = p \log(pn) + \gamma + \frac{1}{2n} + O\left(\frac{1}{n^2}\right),$$

where H_k is the Harmonic number.

8.25 Prove the following generalization of Theorem 8.24: *If for some $r > 0$, the PGF's $G_n(z)$ and $G(z)$ of discrete distributions are analytic in $|z| < r$ and continuous of $|z| = r$, then*

$$|p_{n,k} - p_k| \leq r^{-k} \sup_{|z|=r} |G_n(z) - G(z)|$$

where $p_{n,k} = [z^k]G_n(z)$ and $p_k = [z^k]G(z)$.

8.26 Find the limiting distribution of the profile $B_m(k)$ (defined in Section 8.7.3) for $k = \frac{1}{h} \log m + O(\sqrt{\log m})$, where $B_m(k)$ is the number of nodes on level k in digital search tree built from m strings generated by a biased memoryless source.

8.27 Using techniques of this chapter, prove **Watson's Lemma**: *If $f(t)$ is a real or complex function of positive real variable t, and*

$$f(t) \sim t^{\lambda-1} \sum_{n=0}^{\infty} a_n t^n \qquad \Re(\lambda) > 0,$$

and the integral (i.e., Laplace transform)

$$F(z) = \int_0^{\infty} f(t) e^{-zt} dt$$

is convergent for sufficiently large $\Re(z)$, then

$$F(z) \sim \sum_{n=0}^{\infty} \Gamma(n+\lambda) \frac{a_n}{z^{n+\lambda}}, \qquad z \to \infty$$

in the sector $|\arg(z)| < \frac{1}{2}\pi$ (cf. [263, 335, 353]).