Chapter 1
Introduction

What this book is about. What is a partition function?

The answer depends on who you ask. You get one (multi)set of answers if you ask physicists, and another (multi)set if you ask mathematicians (we allow multisets, in case we want to account for the popularity of each answer). In this book, we adopt a combinatorial view of partition functions. Given a family $\mathcal{F}$ of subsets of the set $\{1, \ldots, n\}$, we define the *partition function* of $\mathcal{F}$ as a polynomial in n real or complex variables $x_1, \ldots, x_n$,

$$p_{\mathcal{F}}(x_1, \ldots, x_n) = \sum_{S \in \mathcal{F}} \prod_{i \in S} x_i. \tag{1.1}$$

Under typical circumstances, it is unrealistic to try to write $p_{\mathcal{F}}$ as a sum of monomials explicitly, for at least one of the following two reasons:

(1) the family $\mathcal{F}$ is very large
or
(2) we are not really sure how large $\mathcal{F}$ is and it will take us a while to go over all subsets S of $\{1, \ldots, n\}$ and check whether $S \in \mathcal{F}$.

Typically, however, we will have no trouble checking if any particular subset S belongs to $\mathcal{F}$. A good example is provided by the family $\mathcal{H}$ of all Hamiltonian cycles in a given graph G (undirected, without loops or multiple edges) with n edges: we say that a collection S of edges forms a *Hamiltonian cycle* in G if the set of edges in S is connected and every vertex of G belongs to exactly two edges from S, see Fig. 1.1.

A graph with m vertices may contain as many as $\frac{(m-1)!}{2}$ different Hamiltonian cycles and it is believed (known, if $\mathbf{P} \neq \mathbf{NP}$) that it is computationally hard to find at least one for a graph G supplied by a clever adversary.

A. Barvinok, *Combinatorics and Complexity of Partition Functions*,
Algorithms and Combinatorics 30, DOI 10.1007/978-3-319-51829-9_1

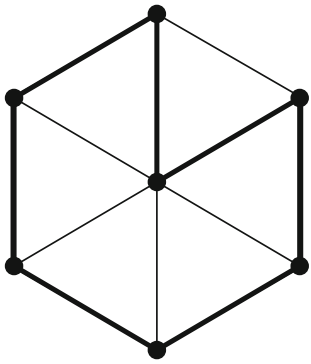

Fig. 1.1 A graph with 7 vertices, 12 edges and a Hamiltonian cycle (*thick lines*)

Sometimes we allow $\mathcal{F}$ to be a family of multisets, in which case we replace

$$\prod_{i \in S} x_i \quad \longrightarrow \quad \prod_{i \in S} x_i^{\mu_i}$$

in formula (1.1), where μ_i is the multiplicity of i in S.

Sometimes we know $p_{\mathcal{F}}$ perfectly well even if we are unable to write it explicitly as a sum of monomials due to the lack of time. For example, if $\mathcal{F} = 2^{\{1,\ldots,n\}}$ is the set of all subsets, we have

$$p_{2^{\{1,\ldots,n\}}}(x_1, \ldots, x_n) = \sum_{S \subset \{1,\ldots,n\}} \prod_{i \in S} x_i = \prod_{i=1}^{n} (1 + x_i) \tag{1.2}$$

and it is hard to argue that we can know $p_{2^{\{1,\ldots,n\}}}$ any better than by the succinct product in (1.2). Our experience teaches us, however, that the cases like (1.2) are quite rare. For some mysterious reasons they all seem to reduce eventually to some determinant enumerating perfect matchings in a planar graph, see [Ba82], [Va08] and Chap. 10 of [Ai07] for examples and recall that a *perfect matching* in a graph is a collection of edges that contains every vertex of the graph exactly once (see Fig. 4.1) and that the graph is *planar* if it can be drawn in the plane so that no two edges can possibly intersect in a point other than their common vertex (see Fig. 4.8).

Although in Sect. 4.3 of the book we describe the classical Kasteleyn's construction expressing the partition function of perfect matchings in a planar graph as a determinant (more precisely, as a Pfaffian), the focus of the book is different. Since the efficient exact computation of $p_{\mathcal{F}}$ in most interesting situations is believed to be impossible (unless the computational complexity hierarchy collapses, that is, unless $\mathbf{P} = \#\mathbf{P}$), we are interested in situations when $p_{\mathcal{F}}$ can be efficiently approximated. By *efficiently approximated* we understand that we can compute $p_{\mathcal{F}}$ approximately for all $x = (x_1, \ldots, x_n)$ in some sufficiently interesting domain, but not only. We also approximate $p_{\mathcal{F}}$ by some "nice function", whose behavior we understand reasonably well. We concentrate mostly on the following three approaches.

Scaling. It may happen that there is a sufficiently rich group of transformations, for example of the type $x_i \longmapsto \lambda_i x_i$ for some λ_i, which change the value of the polynomial $p_{\mathcal{F}}(x_1, \ldots, x_n)$ in some obvious way and such that after factoring that group out, we are left with a function that varies little. This is the case for the permanent (Sect. 3.5), hafnian (Sect. 4.2) and their higher-dimensional extensions (Sects. 4.4 and 4.5). A closely related approach expresses $p_{\mathcal{F}}$ as the coefficient of a monomial $y_1^{\alpha_1} \cdots y_N^{\alpha_N}$ in some explicit polynomial $P(y_1, \ldots, y_N)$ and obtains an estimate of $p_{\mathcal{F}}$ via solution of a convex optimization problem of minimizing $y_1^{-\alpha_1} \cdots y_N^{-\alpha_N} P(y_1, \ldots, y_N)$ for $y_1, \ldots, y_N > 0$. We apply this approach to estimate partition functions of flows (Chap. 8).

Correlation decay. We choose a variable (or a small set of variables), say x_n, and define $p_{\mathcal{F}_n}$ as the sum of the monomials of $p_{\mathcal{F}}$ containing x_n. It may happen that there is some metric on the set $\{x_1, \ldots, x_n\}$ of variables such that the ratio $p_{\mathcal{F}_n}(x_1, \ldots, x_n) / p_{\mathcal{F}}(x_1, \ldots, x_n)$ does not depend much on the variables x_i that are sufficiently far away from x_n in that metric. This allows us to fix values of those remote variables to our convenience and quickly approximate the ratio. We then recover $p_{\mathcal{F}}$ by iterating this procedure and telescoping. As a result, we approximate $\ln p_{\mathcal{F}}(x_1, \ldots, x_n)$ by a sum of functions, each of which depends on a small number of coordinates. We apply this method to the matching polynomial (Sect. 5.2) and to the independence polynomial of a graph (Sects. 6.3 and 6.4).

Interpolation. Suppose that the polynomial $p_{\mathcal{F}}$ has no zeros in a domain $\Omega \subset \mathbb{C}^n$. It turns out that $\ln p_{\mathcal{F}}$ is well approximated in a slightly smaller domain $\Omega' \subset \Omega$ by a low degree Taylor polynomial, sometimes after a change of coordinates (Sect. 2.2). We demonstrate this approach for the permanent (Sects. 3.6 and 3.7) and hafnian (Sect. 4.1), their higher-dimensional extensions (Sect. 4.4), for the matching polynomial (Sect. 5.1) and the independence polynomial of a graph (Sect. 6.1), and for the graph homomorphism partition function (Chap. 7). In our opinion, this is the most general approach.

The correlation decay approach appears to be closely related to a probabilistic approach, known as the Markov Chain Monte Carlo method. Assuming that $x_1 > 0, \ldots, x_n > 0$, we consider the family $\mathcal{F}$ as a finite probability space, with

$$\mathbf{Pr}(S) = \left(\prod_{i \in S} x_i \right) / p_{\mathcal{F}}(x_1, \ldots, x_n) \quad \text{for} \quad S \in \mathcal{F}. \tag{1.3}$$

Suppose that we can sample a random set $S \in \mathcal{F}$ in accordance with the probability distribution (1.3). Then we can measure the frequency of how often a random S contains a particular element of the ground set, say n, and hence we can estimate the ratio $p_{\mathcal{F}_n}(x_1, \ldots, x_n) / p_{\mathcal{F}}(x_1, \ldots, x_n)$, which is also the goal of the correlation decay method. To sample a random $S \in \mathcal{F}$, we perform a random walk on $\mathcal{F}$ by starting with some particular S and, at each step, trying to modify $S \longmapsto \widehat{S}$ by a random move of the type $\widehat{S} := (S \setminus I) \cup J$ for some small sets $I, J \subset \{1, \ldots, n\}$ performed with probability proportional to

$$\frac{\mathbf{Pr}\,(\widehat{S})}{\mathbf{Pr}\,(S)} = \left(\prod_{j\in J} x_j\right)\left(\prod_{i\in I} x_i^{-1}\right).$$

It stands to reason that if the ratios of the type $p_{\mathcal{F}_n}(x_1,\ldots,x_n)\,/\,p_{\mathcal{F}}(x_1,\ldots,x_n)$ depend effectively only on a small set of variables, then we can expect the resulting walk to mix rapidly, that is, we should hit more or less random S after performing a moderate number of moves.

The Markov Chain Monte Carlo method resulted in a number of remarkable successes, most notably in a randomized polynomial time approximation algorithm for the permanent of a non-negative matrix [J+04]. However, we do not discuss it in this book. First, there are excellent books such as [Je03] describing the method in detail and second, we are interested in analytic properties of partition functions that make them amenable to computation (approximation). Granted, the fact that randomized algorithms are often very efficient must be telling us something important about analytic properties of the functions they approximate, but at the moment we hesitate to say what exactly.

Why this is interesting. Why do we care to approximate $p_{\mathcal{F}}$ in (1.1)?

For one thing, it gives us some information about complicated combinatorial families. As an example, let us consider the family $\mathcal{H}$ of all Hamiltonian cycles in a complete graph K_m (undirected, without loops or multiple edges) with m vertices $1,\ldots,m$. Hence to every edge (i,j) of K_m we assign a variable x_{ij}, to every Hamiltonian cycle in K_m we assign a monomial that is the product of the variables x_{ij} on the edges of the cycle, and we define $p_{\mathcal{H}}$ by summing up all monomials attached to the Hamiltonian cycles in K_m. If we let $x_{ij}=1$ for all edges (i,j) then the value of $p_{\mathcal{H}}$ is just the number of Hamiltonian cycles in K_m, which is $(m-1)!/2$. If we assign $x_{ij}=1$ for some edges of K_m and $x_{ij}=0$ for all other edges of K_m, then the value of $p_{\mathcal{H}}$ is the number of Hamiltonian cycles in the graph G consisting of the edges selected by the condition $x_{ij}=1$ (generally, it is computationally hard even to tell $p_{\mathcal{H}}$ from 0).

Looking at the problem of counting Hamiltonian cycles through the prism of the partition function $p_{\mathcal{H}}$ allows us to interpolate between a trivial problem (counting Hamiltonian cycles in the complete graph) and an impossible one (counting Hamiltonian cycles in an arbitrary graph) and find some middle ground. Given a graph G with vertices $1,\ldots,m$, let us fix a small $\epsilon>0$ (think $\epsilon=10^{-10}$) and let us define

$$x_{ij} = \begin{cases} 1 & \text{if } (i,j) \text{ is an edge of } G \\ \epsilon & \text{otherwise.} \end{cases}$$

In this case, $p_{\mathcal{H}}$ still enumerates Hamiltonian cycles in the complete graph K_m, but it does so deliberately. It counts every Hamiltonian cycle in G with weight 1, while every Hamiltonian cycle in K_m that contains r non-edges of G is counted with weight ϵ^r. In Sect. 3.8, we show that it is quite easy to approximate $p_{\mathcal{H}}$ within a factor of $m^{O(\ln m)}$, where the implicit constant in the "O" notation depends on ϵ. This gives us

some idea about Hamiltonian cycles in G: for example, we can separate graphs G with many Hamiltonian cycles (the value of $p_{\mathcal{H}}$ is large) from graphs G that do not acquire a single Hamiltonian cycle unless sufficiently many new edges are added to G (the value of $p_{\mathcal{H}}$ is small).

Two particular topics discussed in this book are

(1) connections between the computational complexity of partition functions and their complex zeros

and

(2) connections between computational complexity and "phase transition" in physics.

In statistical physics, one deals with the probability space $\mathcal{F}$ defined by (1.3) (sets $S \in \mathcal{F}$ are called "configurations"), where $x_i = e^{\beta_i/t}$ for some constants $\beta_i > 0$ and a real parameter t, interpreted as temperature. As the ground set $\{1, \ldots, n\}$ and the set $\mathcal{F}$ of configurations grow in some regular way, one can consider two related, though not identical notions of phase transition. The first notion has to do with a complex zero of $p_{\mathcal{F}}$, as a function of t, approaching the positive real axis at some "critical temperature" $t_c > 0$. This implies the loss of smoothness or even continuity for various physically meaningful quantities, expressed in terms of $\ln p_{\mathcal{F}}$ and its derivatives [YL52]. The second notion of phase transition has to do with the appearance or disappearance of "long-range correlations". Typically, at a high temperature t (that is, when x_i are close to 1), there is no long-range correlation: the probability that S contains a given element i of the ground set is not affected by whether S contains another element j, far away from i in some natural metric. As the temperature t falls (and hence x_i grow), such a dependence may appear. These two notions of phase transition are related though apparently not identical, see [DS87] and [Ci87], we discuss this when we talk about the Ising model in Sect. 7.4 .

The correlation decay approach emphasizing (2) was introduced by Bandyopadhyay and Gamarnik [BG08] and independently by Weitz [We06] and is generally well-known in the computational community, while (1) is relatively less articulated but appears to be no less interesting. Curiously, while the first type of phase transition is associated with complex zeros of the partition function approaching the positive real axis, as far as our ability to approximate is concerned, a priori this does not represent an insurmountable obstacle. What hinders our ability to compute are the complex zeros "blocking" the reference point in the vicinity of which $p_{\mathcal{F}}$ looks easy, such as the point $x_{ij} = 1$ for the partition function $p_{\mathcal{H}}$ of Hamiltonian cycles, see also our discussion in Sect. 2.2. The ways of statistical physics and those of computational complexity diverge at this point, which is probably explained by the fact that the temperature in the physical world is necessarily a real number, while for computational purposes we can manipulate with a complex temperature just as easily.

We stick to the language of combinatorics but the objects and phenomena discussed in this book have also their names in physics. Thus the "matching polynomial" of Chap. 5 corresponds to the "monomer-dimer model", the "graph homomorphism partition function" in Chap. 7 corresponds to a "spin system", while the cut partition

function of Sect. 7.4 corresponds to a "ferromagnetic spin system". Some of our results, such as in Sects. 3.6, 3.7, 3.8, 4.2, 4.4, 7.1 and 7.2 correspond to the "mean field theory" approach, while some others, such as in Chaps. 5 and 6 correspond to the "hard core" model. For still others, such as in Sects. 3.4, 3.5 and Chap. 8, we were unable to think of an appropriate physics name (though "renormalization" may work for those in Sects. 3.4 an 3.5). We talk about physical implications of results in Sect. 7.4 while discussing the Ising model, which connects several directions explored this book: zeros of partition functions, phase transition, correlation decay, graph homomorphisms and enumeration of perfect matchings.

Finally, this book may be interesting because it contains an exposition of quite recent breakthroughs (available before, to the best of our knowledge, only as preprints, journal or conference proceedings papers). These include the Gurvits approach connecting certain combinatorial quantities with stable polynomials (Sects. 3.3 and 8.1), Csikvári and Lelarge approach to the Bethe-approximation of the permanent (Sects. 5.3 and 5.4) and Weitz correlation decay method for the independence polynomial (Sect. 6.4).

Prerequisites, contents, notation, and assorted remarks. We use some concepts of combinatorics, but only very basic, such as graphs and hypergraphs. All other terms, also very basic, such as matchings, perfect matchings and colorings are explained in the text. We also employ some computational complexity concepts. As we are interested in establishing that some functions can be efficiently computed (approximated), and not in proving that some functions are hard to approximate, we use only some very basic complexity concepts, such as polynomial time algorithm, etc. The book [PS98] will supply more than enough prerequisites in combinatorics and computational complexity (but see also more recent and comprehensive [AB09] and [Go08]). We also require modest amounts of linear algebra, real and complex analysis. This book should be accessible to an advanced undergraduate.

In Chap. 2, we develop our toolbox. First, we discuss various topics in convexity: convex and concave functions, entropy and Bethe-entropy, Gauss-Lucas theorem on the zeros of the derivative of a complex polynomial, the capacity of real polynomials and the Prékopa-Leindler inequality. Then we present one of our main tools, interpolation, which allows us to approximate the logarithm of a multivariate polynomial p by a low degree polynomial in a domain, given that there are no complex zeros of p in a slightly larger domain. We discuss interlacing polynomials, $\mathbb{H}$-stable polynomials (polynomials with no roots in the open upper half-plane of $\mathbb{C}$) and $\mathbb{D}$-stable polynomials (polynomials with no roots in the closed unit disc in $\mathbb{C}$).

Then we begin our study of partition functions in earnest.

In Chap. 3, we start slowly with the permanent, as it is very easy to define and it has a surprisingly rich structure. All this makes the permanent a very natural candidate to try our toolbox on.

In Chap. 4, we consider extensions of the permanent to non-bipartite graphs (hafnians) and hypergraphs (multi-dimensional permanents). We also consider the mixed discriminant, which is a generalization of the permanent and of the determinant

simultaneously. We observe that some properties of the permanent can be extended to those more general objects, while some other cannot.

In Chap. 5, we consider the matching polynomial of a graph, a relative of the permanent and hafnian. Here we introduce the correlation decay method, which, as Bayati, Gamarnik, Katz, Nair and Tetali showed [B+07], looks particularly elegant and simple in the case of the matching polynomial. It turns out to be very useful too and provides some additional insight into the permanent.

In Chap. 6, we discuss the independence polynomial of a graph. We prove Dobrushin's bound on the complex roots and also present the correlation decay approach at its most technical. We discuss an open question due to Sokal [S01b], which, if answered affirmatively, would allow us to bridge the gap between different degrees of approximability afforded by the interpolation and by correlation decay approaches.

In Chap. 7, we present combinatorial partition functions at their most general. Here we rely entirely on our interpolation technique, although some of the results can be obtained by the correlation decay approach [LY13]. We also prove the Circle Theorem of Lee and Yang and discuss the Ising model in some detail.

In Chap. 8, we consider partition functions associated with multisets. We study the partition functions of 0-1 and non-negative integer flows, which present yet another extension of permanents. Permanents also supply our main technical tool.

Sections, theorems, lemmas, and formulas are numbered separately inside each chapter. Figures are numbered consecutively in each chapter. For example, Fig. 4.3 is the third figure in Chap. 4.

We use $\Re$ to denote the real part of a complex number and $\Im$ to denote the imaginary part of a complex number, so that $\Re z = a$ and $\Im z = b$ for $z = a + ib$. We denote by $|X|$ the cardinality of a finite set X.

Finally, the product of complex numbers from an empty set is always 1.

Acknowledgements I am grateful to David Gamarnik, Leonid Gurvits, Gil Kalai, Guus Regts, Alex Samorodnitsky, Prasad Tetali, Jan Vondrak and Shira Zerbib for helpful conversations, advice and encouragement. I benefited a lot from the program "Counting Complexity and Phase Transition" and I thank the Simons Institute for the Theory of Computing for hospitality. I am grateful to many anonymous referees of the first draft of the book for their remarks. I thank Han Wu and Max Kontorovich for their work on the REU project on the complex roots and approximation of permanents in Summer 2016, where they implemented and experimented with the Taylor polynomial interpolation method for permanents [KW16] and also pointed out to [Wa03] and connections to the Szegő curve (see Sect. 5.5).

This work is partially supported by NSF Grant DMS 1361541.

http://www.springer.com/978-3-319-51828-2

Combinatorics and Complexity of Partition Functions

Barvinok, A.

2016, VI, 303 p. 51 illus., 42 illus. in color., Hardcover
ISBN: 978-3-319-51828-2