INTRODUCTION

Lattice polytopes arise naturally in algebraic geometry, analysis, combinatorics, computer science, number theory, optimization, probability and representation theory. They possess a rich structure arising from the interaction of algebraic, convex, analytic, and combinatorial properties. In this chapter, we concentrate on the theory of lattice polytopes and only sketch their numerous applications. We briefly discuss their role in optimization and polyhedral combinatorics (Section 7.1). In Section 7.2 we discuss the decision problem, the problem of finding whether a given polytope contains a lattice point. In Section 7.3 we address the counting problem, the problem of counting all lattice points in a given polytope. The asymptotic problem (Section 7.4) explores the behavior of the number of lattice points in a varying polytope (for example, if a dilation is applied to the polytope). Finally, in Section 7.5 we discuss problems with quantifiers. These problems are natural generalizations of the decision and counting problems. Whenever appropriate we address algorithmic issues. For general references in the area of computational complexity/algorithms see [AB09]. We summarize the computational complexity status of our problems in Table 7.0.1.

TABLE 7.0.1  Computational complexity of basic problems.

<table>
<thead>
<tr>
<th>PROBLEM NAME</th>
<th>BOUNDED DIMENSION</th>
<th>UNBOUNDED DIMENSION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision problem</td>
<td>polynomial</td>
<td>NP-hard</td>
</tr>
<tr>
<td>Counting problem</td>
<td>polynomial</td>
<td>#P-hard</td>
</tr>
<tr>
<td>Asymptotic problem</td>
<td>polynomial</td>
<td>#P-hard*</td>
</tr>
<tr>
<td>Problems with quantifiers</td>
<td>unknown; polynomial for $\forall \exists$ **</td>
<td>NP-hard</td>
</tr>
</tbody>
</table>

* in bounded codimension, reduces polynomially to volume computation
** with no quantifier alternation, polynomial time

7.1 INTEGRAL POLYTOPES IN POLYHEDRAL COMBINATORICS

We describe some combinatorial and computational properties of integral polytopes. General references are [GLSS88], [GW93], [Sch86], [Lag95], [DL97] and [Zie00].

GLOSSARY

$\mathbb{R}^d$: Euclidean $d$-dimensional space with scalar product $(x, y) = x_1y_1 + \ldots + x_dy_d$, where $x = (x_1, \ldots, x_d)$ and $y = (y_1, \ldots, y_d)$.

$\mathbb{Z}^d$: The subset of $\mathbb{R}^d$ consisting of the points with integral coordinates.
**Polytope:** The convex hull of finitely many points in \( \mathbb{R}^d \).

**Face of a polytope** \( P \): The intersection of \( P \) and the boundary hyperplane of a halfspace containing \( P \).

**Facet:** A face of codimension 1.

**Vertex:** A face of dimension 0; the set of vertices of \( P \) is denoted by \( \text{Vert} \, P \).

**H-description of a polytope (H-polytope):** A representation of the polytope as the set of solutions of finitely many linear inequalities.

**V-description of a polytope (V-polytope):** The representation of the polytope by the set of its vertices.

**Integral polytope:** A polytope with all of its vertices in \( \mathbb{Z}^d \).

**(0, 1)-polytope:** A polytope \( P \) such that each coordinate of every vertex of \( P \) is either 0 or 1.

An integral polytope \( P \subset \mathbb{R}^d \) can be given either by its H-description or by its V-description or (somewhat implicitly) as the convex hull of integral points in some other polytope \( Q \), so \( P = \text{conv}(Q \cap \mathbb{Z}^d) \). In most cases it is difficult to translate one description into another. The following examples illustrate some typical kinds of behavior.

**INTEGRALITY OF H-POLYTOPES**

It is an NP-hard problem to decide whether an \( H \)-polytope \( P \subset \mathbb{R}^d \) is integral. However, if the dimension \( d \) is fixed then the straightforward procedure of generating all the vertices of \( P \) and checking their integrality has polynomial time complexity. A rare case where an \( H \)-polytope \( P \) is a priori integral is known under the general name of “total unimodularity.” Let \( A \) be an \( n \times d \) integral matrix such that every minor of \( A \) is either 0 or 1 or \(-1\). Such a matrix \( A \) is called **totally unimodular**. If \( b \in \mathbb{Z}^n \) is an integral vector then the set of solutions to the system of linear inequalities \( Ax \leq b \), when bounded, is an integral polytope in \( \mathbb{R}^d \). Examples of totally unimodular matrices include matrices of vertex-edge incidences of oriented graphs and of bipartite graphs. A complete characterization of totally unimodular matrices and a polynomial time algorithm for recognizing a totally unimodular matrix is provided by a theorem of P. Seymour (see [Sch86]). A family of integral polytopes, called **transportation polytopes**, are much studied in the literature (see [EKK84] and [DLK14]). An example of a transportation polytope is provided by the set of \( m \times n \) nonnegative matrices \( x = (x_{ij}) \) whose row and column sums are given positive integers. Integral points in this polytope are called **contingency tables**; they play an important role in statistics. A particular transportation polytope, called the **Birkhoff polytope**, is the set \( B_n \) of \( n \times n \) nonnegative matrices with all row and column sums equal to 1. Alternatively, it may be described as the convex hull of the \( n! \) permutation matrices \( \pi(\sigma)_{ij} = \delta_{i\pi(i)} \) for all permutations \( \sigma \) of the set \( \{1, \ldots, n\} \).

The notion of total unimodularity has been generalized in various directions, thus leading to new classes of integral polytopes (see [Cor01] and [Sch03]).

Reflexive polytopes, that is, integral polytopes whose polar dual are also integral polytopes play an important role in mirror symmetry in algebraic geometry.
COMBINATORIALLY DEFINED V-POLYTOPES

There are several important situations where the explicit V-description of an integral polytope is too long and a shorter description is desirable although not always available. For example, a (0,1)-polytope may be given as the convex hull of the characteristic vectors

\[ \chi_S(i) = \begin{cases} 
1 & \text{if } i \in S, \\
0 & \text{otherwise}
\end{cases} \]

for some combinatorially interesting family \( S \) of subsets \( S \subset \{1, \ldots, d\} \) (see [GLS88] for various examples). One of the most famous example is the traveling salesman polytope, the convex hull \( TSP_n \) of the \((n-1)!\) permutation matrices \( \pi(\sigma) \) where \( \sigma \) is a permutation of the set \( \{1, \ldots, n\} \) consisting of precisely one cycle (cf. the Birkhoff polytope \( B_n \) above). The problem of the H-description of the traveling salesman polytope has attracted a lot of attention (see [GW93], [EKK84] and [Sch03] for some references) because of its relevance to combinatorial optimization. C.H. Papadimitriou proved that it is a co-NP-complete problem to establish whether two given vertices of \( TSP_n \) are adjacent, i.e., connected by an edge [Pap78]. L. Billera and A. Sarangarajan proved that every (0,1)-polytope can be realized as a face of \( TSP_n \) for sufficiently large \( n \) [BS90]. Thus the combinatorics of \( TSP_n \) contrasts with the combinatorics of the Birkhoff polytope \( B_n \).

Another important polytope arising in this way is the cut polytope, the famous counterexample to the Borsuk conjecture (see [DL97]). It is defined as the convex hull of the set of \( n \times n \) matrices \( x_S \), where

\[ x_S(i, j) = \begin{cases} 
1 & \text{if } |\{i, j\} \cap S| = 1 \text{ and } i \neq j, \\
0 & \text{otherwise},
\end{cases} \]

where \( S \) ranges over all subsets of the set \( \{1, \ldots, n\} \).

CONVEX HULL OF INTEGRAL POINTS

Let \( P \subset \mathbb{R}^d \) be a polytope. Then the convex hull \( P_I \) of the set \( P \cap \mathbb{Z}^d \), if nonempty, is an integral polytope. Generally, the number of facets or vertices of \( P_I \) depends not only on the number of facets or vertices of \( P \) but also on the actual numerical size of the description of \( P \) (see [CHKM92]). Furthermore, it is an NP-complete problem to check whether a given point belongs to \( P_I \), where \( P \) is given by its H-description. If, however, the dimension \( d \) is fixed then the complexity of the facial description of the polytope \( P_I \) is polynomial in the complexity of the description of \( P \). In particular, the number of vertices of \( P_I \) is bounded by a polynomial of degree \( d - 1 \) in the input size of \( P \) [CHKM92].

Integrality imposes some restrictions on the combinatorial structure of a polytope. It is known that the combinatorial type of any 2- or 3-dimensional polytope can be realized by an integral polytope. J. Richter-Gebert constructed a 4-dimensional polytope with a nonintegral (and, therefore, nonrational) combinatorial type [Ric96]. Earlier, N. Mnëv had shown that for all sufficiently large \( d \) there exist nonrational \( d \)-polytopes with \( d + 4 \) vertices [Mnë83]. The number \( N_d(V) \) of classes of integral \( d \)-polytopes having volume \( V \) and nonisomorphic with respect to
affine transformations of $\mathbb{R}^d$ preserving the integral lattice $\mathbb{Z}^d$ has logarithmic order
\[ c_1(d) V^{\frac{d-1}{d+1}} \leq \log N_d(V) \leq c_2(d) V^{\frac{d-1}{d+1}} \]
for some $c_1(d), c_2(d) > 0$ [BV92].

### 7.2 DECISION PROBLEM

We consider the following general decision problem: Given a polytope $P \subset \mathbb{R}^d$ and a lattice $\Lambda \subset \mathbb{R}^d$, decide whether $P \cap \Lambda = \emptyset$ and, if the intersection is nonempty, find a point in $P \cap \Lambda$. We describe the main structural and algorithmic results for this problem. General references are [GL87], [GLS88], [GW93], [Sch86], and [Lag95].

### GLOSSARY

**Lattice:** A discrete additive subgroup $\Lambda$ of $\mathbb{R}^d$, i.e., $x - y \in \Lambda$ for any $x, y \in \Lambda$ and $\Lambda$ does not contain limit points.

**Basis of a lattice:** A set of linearly independent vectors $u_1, \ldots, u_k$ such that every vector $y \in \Lambda$ can be (uniquely) represented in the form $y = m_1 u_1 + \ldots + m_k u_k$ for some integers $m_1, \ldots, m_k$.

**Rank of a lattice:** The cardinality of any basis of the lattice. If $\Lambda \subset \mathbb{R}^d$ has rank $d$, $\Lambda$ is said to be of full rank.

**Determinant of a lattice:** For a lattice of rank $k$ the $k$-volume of the parallelepiped spanned by any basis of the lattice.

**Reciprocal lattice:** For a full rank lattice $\Lambda \subset \mathbb{R}^d$, the lattice $\Lambda^* = \{ x \in \mathbb{R}^d \mid \langle x, y \rangle \in \mathbb{Z} \text{ for all } y \in \Lambda \}$.

**Polyhedron:** An intersection of finitely many halfspaces in $\mathbb{R}^d$.

**Convex body:** A compact convex set in $\mathbb{R}^d$ with nonempty interior.

**Lattice Polytope:** For a given lattice $\Lambda$, a polytope with all of its vertices in $\Lambda$.

Applying a suitable linear transformation one can reduce the decision problem to the case in which $\Lambda = \mathbb{Z}^k$ and $P \subset \mathbb{R}^k$ is a full-dimensional polytope, $k = \text{rank } \Lambda$.

The decision problem is known to be NP-complete for $\mathcal{H}$-polytopes as well as for $\mathcal{V}$-polytopes, although some special cases admit a polynomial time algorithm. In particular, if one fixes the dimension $d$ then the decision problem becomes polynomially solvable. The main tool is provided by the so-called “flatness results.”

### FLATNESS THEOREMS

Let $P \subset \mathbb{R}^d$ be a convex body and let $l \in \mathbb{R}^d$ be a nonzero vector. The number
\[ \max \{ \langle l, x \rangle \mid x \in P \} - \min \{ \langle l, x \rangle \mid x \in P \} \]
is called the width of $P$ with respect to $l$. For a full rank lattice $\Lambda \subset \mathbb{R}^d$, the minimum width of $P$ with respect to a nonzero vector $l \in \Lambda^*$ is called the lattice width of $P$. 

The following general result is known under the unifying name of “flatness theorem.”

THEOREM 7.2.1
There is a function \( f : \mathbb{N} \to \mathbb{R} \) such that for any full rank lattice \( \Lambda \subset \mathbb{R}^d \) and any convex body \( P \subset \mathbb{R}^d \) with \( P \cap \Lambda = \emptyset \), the lattice width of \( P \) does not exceed \( f(d) \).

There are two types of results relating to the flatness theorem.

First, one may be interested in making \( f(d) \) as small as possible. One can observe that \( f(d) \geq d \): for some small \( \epsilon > 0 \), consider \( \Lambda = \mathbb{Z}^d \) and the polytope \( P \) defined by the inequalities \( x_1 + \ldots + x_d \leq d - \epsilon, \ x_i \geq \epsilon \) for \( i = 1, \ldots, d \). It is known that one can choose \( f(d) = O(d^{3/2}) \) and it is conjectured that one can choose \( f(d) \) as small as \( O(d) \). W. Banaszczyk proved that if \( P \) is centrally symmetric, then one can choose \( f(d) = O(d \log d) \), which is optimal up to a logarithmic factor. For these and related results, see \[BLPS99\]. There are results regarding the lattice width of some interesting classes of convex sets. Thus, if \( P \subset \mathbb{R}^d \) is an ellipsoid that does not contain lattice points, then the lattice width of \( P \) is \( O(d) \) \[BLPS99\]. A lattice polytope with no lattice points other than its vertices is called sometimes empty polytope. J.-M. Kantor \[Kan99\] showed that the lattice width of a \( d \)-dimensional empty simplex can grow linearly in \( d \) and then A. Sebő \[Seb99\] constructed explicit examples of \( d \)-dimensional empty simplices of width \( d - 2 \). If \( P \) is a \( 3 \)-dimensional empty polytope, then the lattice width of \( P \) is \( 1 \) (see \[Sca85\] and Section 16.6.1 of this Handbook for more on lattice width).

Second, one may be interested in the best width bound for which the corresponding vector \( l \in \Lambda^* \) can be computed in polynomial time. The best bound known is \( 2O(d) \), where \( l \) is polynomially computable even if the dimension \( d \) varies; see \[GLSS88\]. For the computational complexity of lattice problems, such as finding the shortest non-zero lattice vector or the nearest lattice vector to a given point, see \[MG02\].

ALGORITHMS FOR THE DECISION PROBLEMS

Flatness theorems allow one to reduce the dimension in the decision problem: Assuming that \( \Lambda = \mathbb{Z}^d \) and that the body \( P \) does not contain an integral point, one constructs a vector \( l \in \mathbb{Z}^d \) for which \( P \) has a small width and reduces the \( d \)-dimensional decision problem to a family of \((d - 1)\)-dimensional decision problems \( P_i = \{ x \in P \mid \langle l, x \rangle = i \} \), where \( i \) ranges between \( \min \{ \langle l, x \rangle \mid x \in P \} \) and \( \max \{ \langle l, x \rangle \mid x \in P \} \). This reduction is the main idea of polynomial time algorithms in fixed dimension. The best complexity known for the decision problem in terms of the dimension \( d \) is \( d^{O(d)} \), see \[Dad14\] for recent advances.

Constructing \( l \) efficiently relies on two major components (see \[GLSS88\]). First, a linear transformation \( T \) is computed, such that the image \( T(P) \) is “almost round,” meaning that \( T(P) \) is sandwiched between a pair of concentric balls with the ratio of their radii bounded by some small constant depending only on the dimension \( d \). At this stage, a linear programming algorithm is used. Second, a reasonably short nonzero vector \( u \) is constructed in the lattice \( \Lambda^* \) reciprocal to \( \Lambda = T(\mathbb{Z}^d) \). A basis reduction algorithm is used at this stage. Then we let \( l = (T^*)^{-1}u \).

One can streamline the process by using the generalized lattice reduction \[LS92\] tailored to a given polytope. A polynomial time algorithm based on counting lattice points in the polytope and not using the flatness argument is sketched in \[BP99\].
MINKOWSKI’S CONVEX BODY THEOREM

The following classical result, known as “Minkowski’s convex body theorem,” provides a very useful criterion.

THEOREM 7.2.2

Suppose that $B \subset \mathbb{R}^d$ is a convex body, centrally symmetric about the origin 0, and $\Lambda \subset \mathbb{R}^d$ is a lattice of full rank. If $\text{vol} \ B \geq 2^d \text{det} \ \Lambda$ then $B$ contains a nonzero point of $\Lambda$.

For the proof and various generalizations see, for example, [GL87]. An important generalization (Minkowski’s Second Theorem) concerns the existence of $i$ linearly independent lattice points in a convex body. Namely, if

$$\lambda_i = \inf \left\{ \lambda > 0 \mid \lambda B \cap \Lambda \text{ contains } i \text{ linearly independent points} \right\}$$

is the “$i$th successive minimum,” then $\lambda_1 \ldots \lambda_d \leq (2^d \text{det} \ \Lambda) / (\text{vol} B)$.

If $B$ is a symmetric convex body such that $\text{vol} \ B = 2^d \text{det} \ \Lambda$ but $B$ does not contain a nonzero lattice point in its interior, then $B$ is called extremal. Every extremal body is necessarily a polytope. Moreover, this polytope contains at most $2(2^d - 1)$ facets, and therefore, for every dimension $d$, there exist only finitely many combinatorially different extremal polytopes. The contracted polytope $P = \{x/2 \mid x \in B\}$ has the property that its lattice translates $P + x$, $x \in \Lambda$, tile the space $\mathbb{R}^d$. Such a tiling polytope is called a parallelohedron. Similarly, for every dimension $d$ there exist only finitely many combinatorially different parallelohedra. Parallelohedra can be characterized intrinsically: a polytope is a parallelohedron if and only if it is centrally symmetric, every facet of it is centrally symmetric, and every class of parallel ridges ($(d-2)$-dimensional faces) consists of four or six ridges. If $q : \mathbb{R}^d \to \mathbb{R}$ is a positive definite quadratic form, then the Dirichlet-Voronoi cell $P_q = \{x \mid q(x) \leq q(x - \lambda) \text{ for any } \lambda \in \Lambda\}$ is a parallelohedron.

The problem of deciding whether a centrally symmetric polyhedron $P$ contains a nonzero point from a given lattice $\Lambda$ is known to be NP-complete even in the case of the standard cube $P = \{(x_1, \ldots, x_d) \mid -1 \leq x_i \leq 1\}$. For fixed dimension $d$ there exists a polynomial time algorithm since the problem obviously reduces to the decision problem (one can add the extra inequality $x_1 + \ldots + x_d \geq 1$).

VOLUME BOUNDS

An integral simplex in $\mathbb{R}^d$ containing no integral points other than its vertices has volume $1/2$ if $d = 2$ but already for $d = 3$ can have an arbitrarily large volume (the smallest possible volume of such a simplex is $1/d!$). On the other hand, D. Hensley proved if an integral polytope $P$ contains precisely $k > 0$ integral points then its volume is bounded by a function of $k$ and $d$. J.C. Lagarias and G.M. Ziegler proved that $\text{vol} P \leq k(7(k + 1))^{d+1}$, see [Lag95] and also [Pik01] for some sharpening.

G. Averkov, J. Krümpelmann and B. Nill found the maximum volume of an integral simplex that contains exactly one integer point in its interior [AKN15], thus proving a conjecture of D. Hensley. Namely, let $s_1, \ldots, s_d$ be the Sylvester...
sequence, defined recursively by

\[ s_1 = 2 \quad \text{and} \quad s_i = 1 + \prod_{j=1}^{i-1} s_j \quad \text{for} \quad i \geq 2. \]

The volume of an integral simplex in \( \mathbb{R}^d \) with precisely one interior point does not exceed \( 2(s_d - 1)^2/d! \) and the bound is attained for the simplex with vertices at \( 0, s_1 e_1, \ldots, s_d - 1 e_d, \) and \( 2(s_d - 1)e_d, \) where \( e_1, \ldots, e_d \) is the standard basis of \( \mathbb{R}^d \).

### 7.3 COUNTING PROBLEM

We consider the following problem: Given a polytope \( P \subset \mathbb{R}^d \), compute exactly or approximately the number of integral points \( |P \cap \mathbb{Z}^d| \) in \( P \).

For counting in general convex bodies see [CHKM92]. For applications in the combinatorics of generating functions, see [Sta86]. For applications in representation theory, see [BZ88], [CDW12] and [PP17]. For applications in statistical physics (computing permanents) and statistics (counting contingency tables), see [JS97]. For applications in social sciences, see [GL11]. For general information see surveys [GW93] and [BP99] and books [BR07] and [Bar08].

### GLOSSARY

**Rational polyhedron:** The set

\[ P = \{ x \in \mathbb{R}^d \mid \langle a_i, x \rangle \leq \beta_i, \ i = 1, \ldots, m \}, \]

where \( a_i \in \mathbb{Z}^d \) and \( \beta_i \in \mathbb{Z} \) for \( i = 1, \ldots, m \). Generally, for a **rational polyhedron with respect to lattice** \( \Lambda \subset \mathbb{R}^d \) of full rank, we have \( a_i \in \Lambda^* \) and \( \beta_i \in \mathbb{Z} \) for \( i = 1, \ldots, m \).

**Polyhedral cone:** A set \( K \subset \mathbb{R}^d \) of the form \( K = \{ \sum_{i=1}^{k} \lambda_i u_i \mid \lambda_i \geq 0, \ i = 1, \ldots, k \} \) for some vectors \( u_1, \ldots, u_k \in \mathbb{R}^d \). The vectors \( u_1, \ldots, u_k \) are called **generators** of \( K \).

**Rational cone:** A polyhedral cone having a set of generators belonging to \( \mathbb{Z}^d \).

Generally, a **rational cone with respect to lattice** \( \Lambda \) is a cone generated by vectors from \( \Lambda \). A rational cone is a rational polyhedron.

**Simple cone:** A polyhedral cone generated by linearly independent vectors.

**Cone of feasible directions at a point:** The cone

\[ K_v = \{ x \mid v + \epsilon x \in P \text{ for all sufficiently small } \epsilon > 0 \} \]

for a point \( v \) of a polytope \( P \). If \( v \) is a vertex, then the cone \( K_v \) is generated by the vectors \( u_i = v_i - v, \) where \( [v_i, v] \) is an edge of \( P \).

**Unimodular cone:** A rational simple cone \( K \subset \mathbb{R}^d \) generated by a basis of \( \mathbb{Z}^d \).

Generally, a **unimodular cone with respect to lattice** \( \Lambda \subset \mathbb{R}^d \) is a cone generated by a basis of \( \Lambda \).

**Simple polytope:** A polytope \( P \) such that the cone \( K_v \) of feasible directions is simple for every vertex \( v \) of \( P \).
**Totally unimodular polytope:** An integral polytope $P$ such that the cone $K_v$ of feasible directions is unimodular for every vertex $v$ of $P$.

**GENERAL INFORMATION**

The counting problem is known to be $\#P$-hard even for an integral $\mathcal{H}$- or $\mathcal{V}$-polytope. However, if the dimension $d$ is fixed, one can solve the counting problem in polynomial time (see [BP99] and [Bar08]).

**SOME EXPLICIT FORMULAS IN LOW DIMENSIONS**

The classical Pick formula expresses the number of integral points in a convex integral polygon $P \subset \mathbb{R}^2$ in terms of its area and the number of integral points on the boundary $\partial P$:

$$|P \cap \mathbb{Z}^2| = \text{area}(P) + \frac{1}{2} \cdot |\partial P \cap \mathbb{Z}^2| + 1$$

(see, for example, [Mor93b], [GW93] and [BR07]). This formula almost immediately gives rise to a polynomial time algorithm for counting integral points in integral polygons.

An important explicit formula for the number of integral points in a lattice tetrahedron of a special kind was proven by L. Mordell, see [BR07]. Let $a,b,c$ be pairwise coprime positive integers and $\Delta(a,b,c) \subset \mathbb{R}^3$ be the tetrahedron with vertices $(0,0,0)$, $(a,0,0)$, $(0,b,0)$, and $(0,0,c)$. Then

$$|\Delta(a,b,c) \cap \mathbb{Z}^3| = \frac{abc}{6} + \frac{ab + ac + bc + a + b + c}{4} + \frac{1}{12}\left(\frac{ac}{b} + \frac{bc}{a} + \frac{ab}{c} + \frac{1}{abc}\right) - s(bc,a) - s(ac,b) - s(ab,c) + 2.$$  \hspace{1cm} (7.3.1)

Here

$$s(p,q) = \sum_{i=1}^{q} \left( \left\lfloor \frac{i}{q} \right\rfloor \right) \left( \left\lfloor \frac{pi}{q} \right\rfloor \right), \hspace{1cm} \text{where} \hspace{1cm} \left\lfloor x \right\rfloor = x - 0.5(\lfloor x \rfloor + \lceil x \rceil),$$

is the Dedekind sum. A similar formula was found in dimension 4. The reciprocity relation $s(p,q) + s(q,p) = (p/q + q/p + 1/pq - 3)/12$ allows one to compute the Dedekind sum $s(p,q)$ in polynomial time. A version of formula (7.3.1) was used by M. Dyer to construct polynomial time algorithms for the counting problem in dimensions 3 and 4. Formula (7.3.1) was generalized to an arbitrary tetrahedron by J. Pommersheim (see [BP99] and [BR07]).

Computationally efficient formulas for the number of lattice points are known for some particular polytopes, most notably zonotopes. Given integral points $v_1, \ldots, v_n$ in $\mathbb{R}^d$, a zonotope spanned by $v_1, \ldots, v_n$ is the polytope

$$P = \left\{ \lambda_1v_1 + \ldots + \lambda_nv_n \mid 0 \leq \lambda_i \leq 1 \text{ for } i = 1,\ldots,n \right\}.$$  

For each subset $S \subset \{v_1, \ldots, v_n\}$ of linearly independent points, let $a_S$ be the index of the sublattice generated by $S$ in the lattice $\mathbb{Z}^d \cap \text{span}(S)$, where $a_{\emptyset} = 1$. Then

$$|P \cap \mathbb{Z}^d| = \sum_S a_S$$  \hspace{1cm} (see Chapter 4, Problem 31 of [Sta86]).
A powerful tool for solving the counting problem exactly is provided by \textit{exponential sums}.

Let $P \subset \mathbb{R}^d$ be a polytope and $c \in \mathbb{R}^d$ be a vector. We consider the exponential sum
\[
\sum_{x \in P \cap \mathbb{Z}^d} \exp\{\langle c, x \rangle\}.
\]
If $c = 0$ we get the number of integral points in $P$. The reason for introducing the parameter $c$ is that for a “generic” $c$ exponential sums reveal some nontrivial algebraic properties that remain invisible when $c = 0$. To describe these properties we need to consider exponential sums over rational polyhedra and, in particular, over cones.

**EXPONENTIAL SUMS OVER RATIONAL POLYHEDRA**

Let $K \subset \mathbb{R}^d$ be a rational cone without lines generated by vectors $u_1, \ldots, u_k$ in $\mathbb{Z}^d$. Then the series $\sum_{x \in K \cap \mathbb{Z}^d} \exp\{\langle c, x \rangle\}$ converges for any $c$ such that $\langle c, u_i \rangle < 0$ for all $i = 1, \ldots, k$ and defines a meromorphic function of $c$, which we denote by $f_K(c)$. In particular, if $K$ is unimodular then
\[
f_K(c) = \prod_{i=1}^{k} \frac{1}{1 - \exp\{\langle c, u_i \rangle\}},
\]

since the corresponding sum is just the multiple geometric series. Generally speaking, the farther a given cone is from being unimodular, the more complicated the formula for $f_K(c)$ will be.

These results are known in many different forms (see, for example, [Sta86, Section 4.6]). Furthermore, the function $f_K(c)$ can be extended to a finitely additive measure, defined on rational polyhedra in $\mathbb{R}^d$ and taking its values in the space of meromorphic functions in $d$ variables, so that the measure of a rational polyhedron with a line is equal to 0. To state the result precisely, let us associate with every set $A \in \mathbb{R}^d$ its \textbf{indicator function} $[A] : \mathbb{R}^d \rightarrow \mathbb{R}$, given by
\[
[A](x) = \begin{cases} 
1 & \text{if } x \in A, \\
0 & \text{otherwise.}
\end{cases}
\]

The following result was proved by A.G. Khovanskii and A. Pukhlikov and, independently, by J. Lawrence, see [Bar08] for an exposition.

**THEOREM 7.3.1** \textit{Lawrence-Khovanskii-Pukhlikov Theorem}

There exists a map that associates, to every rational polyhedron $P \subset \mathbb{R}^d$, a meromorphic function $f_P(c)$, $c \in \mathbb{C}^d$, with the following properties:

The correspondence $P \mapsto f_P$ preserves linear dependencies among indicator functions of rational polyhedra:
\[
\sum_{i=1}^{m} \alpha_i [P_i] = 0 \quad \text{implies} \quad \sum_{i=1}^{m} \alpha_i f_{P_i}(c) = 0
\]
for rational polyhedra $P_i$ and integers $\alpha_i$.

If $P$ does not contain lines, then

$$f_P(c) = \sum_{x \in P \cap \mathbb{Z}^d} \exp\{\langle c, x \rangle\}$$

for all $c$ such that the series converges absolutely.

If $P$ contains a line then $f_P(c) \equiv 0$.

If $P + m$ is a translation of $P$ by an integral vector $m$ then

$$f_{P+m}(c) = \exp\{\langle c, m \rangle\} f_P(c).$$

For example, suppose that $d = 1$ and let us choose $P_+ = [0, +\infty)$, $P_- = (-\infty, 0]$, $P_0 = \{0\}$, and $P = (-\infty, +\infty)$. Then

$$f_{P_+}(c) = \sum_{x=0}^{+\infty} \exp\{cx\} = \frac{1}{1 - \exp\{c\}} \quad \text{and} \quad f_{P_-}(c) = \sum_{x=0}^{-\infty} \exp\{cx\} = \frac{1}{1 - \exp\{-c\}}.$$

Moreover, $f_{P_0} = 1$ and $f_P = 0$ since $P$ contains a line. We see that $|P| = |P_+| + |P_-| - |P_0|$ and that $f_P = f_{P_+} + f_{P_-} - f_{P_0}$.

Let $P \subset \mathbb{R}^d$ be a rational polytope and let $v \in P$ be its vertex. Let us consider the translation $v + K_v$ of the cone $K_v$ of feasible directions at $v$. The following crucial result was proved by M. Brion, see [BR07] and [Bar08].

THEOREM 7.3.2 Brion’s Theorem

Let $P \subset \mathbb{R}^d$ be a rational polytope. Then

$$\sum_{x \in P \cap \mathbb{Z}^d} \exp\{\langle c, x \rangle\} = \sum_{v \in \text{Vert } P} f_{v + K_v}(c).$$

If the polytope is integral, we have $f_{v + K_v}(c) = \exp\{\langle c, v \rangle\} f_{K_v}(c).$ We note that if $K$ is a unimodular cone and $v$ is a rational vector then $f_{K+v} = \exp\{\langle c, w \rangle\} f_K(c)$, where $w \in \mathbb{Z}^d$ is a certain “rounding” of $v$ with respect to $K$. Namely, assume that $K$ is the conic hull of some integral vectors $u_1, \ldots, u_d$ that constitute a basis of $\mathbb{Z}^d$. Let $u^*_1, \ldots, u^*_d$ be the biorthogonal basis such that $\langle u^*_i, u_j \rangle = \delta_{ij}$. Then $w = \sum_{i=1}^d \left\lfloor \langle v, u^*_i \rangle \right\rfloor u_i$.

Essentially, Theorem 7.3.2 can be deduced from Theorem 7.3.1 by noticing that the indicator function of every (rational) polyhedron $P$ can be written as the sum of the indicator functions $|v + K_v|$ modulo indicator functions of (rational) polyhedra with lines, see [BP99] and [Bar08].

Brion’s formula allows one to reduce the counting of integral points in polytopes to the counting of points in polyhedral cones, a much easier problem. Below we discuss two instances where the application of exponential sums and Brion’s identities leads to an efficient computational solution of the counting problem.

COUNTING IN FIXED DIMENSION

The following result was obtained by A. Barvinok (see [BP99] and [Bar08] for an exposition).
THEOREM 7.3.3
Let us fix the dimension \(d\). Then there exists a polynomial time algorithm that, for any given rational polytope \(P \subset \mathbb{R}^d\), computes the number \(|P \cap \mathbb{Z}^d|\) of integral points in \(P\).

THE IDEA OF THE ALGORITHM
We assume that the polytope is given by its \(V\)-description. Let us choose a “generic” \(c \in \mathbb{Q}^d\). We can compute the number \(|P \cap \mathbb{Z}^d|\) as the limit of the exponential sum
\[
\lim_{t \to 0} \sum_{x \in P \cap \mathbb{Z}^d} \exp\{\langle tc, x \rangle\},
\]
where \(t\) is a real parameter. Using Brion’s Theorem 7.3.2, we reduce the problem to the computation of the constant term in the Laurent expansion of the meromorphic function \(f_v(t) = f_{v+K_v}(tc)\), where \(v\) is a vertex of \(P\) and \(K_v\) is the cone of feasible directions at \(v\). If \(K_v\) is a unimodular cone, we have an explicit formula for \(f_{v+K_v}(c)\) (see above) and thus can easily compute the desired term. However, for \(d > 1\) the cone \(K_v\) does not have to be unimodular. It turns out, nevertheless, that for any given rational cone \(K\) one can construct in polynomial time a decomposition
\[
[K] = \sum_{i \in I} \epsilon_i [K_i] \quad \text{where} \quad \epsilon_i \in \{-1, 1\},
\]
of the “inclusion-exclusion” type, where the cones \(K_i\) are unimodular (see below). Thus one can get an explicit expression
\[
f_{v+K_v}(c) = \sum_{i \in I} \epsilon_i \cdot f_{v+K_i}(c)
\]
and then compute the constant term of the Laurent expansion of \(f_v(t)\). The complexity of the algorithm in terms of the dimension \(d\) is \(dO(d)\). The algorithm have been implemented in packages \textsc{LattE} by J.A. De Loera et al. \cite{DLH04} and \textsc{barvinok} by S. Verdoolaege et al. \cite{VSB07}.

COUNTING IN TOTALLY UNIMODULAR POLYTOPES
One can efficiently count the number of integral points in a totally unimodular polytope given by its vertex description even in varying dimension.

THEOREM 7.3.4 \cite{BP99}
There exists an algorithm that, for any \(d\) and any given integral vertices \(v_1, \ldots, v_m \in \mathbb{Z}^d\) such that the polytope \(P = \text{conv}\{v_1, \ldots, v_m\}\) is totally unimodular, computes the number of integral points of \(P\) in time linear in the number \(m\) of vertices.

Moreover, the same result holds for rational polytopes with unimodular cones of feasible directions at the vertices. The algorithm uses Brion’s formulas (Theorem 7.3.2) and the explicit formula above for the exponential sum over a unimodular cone.
EXAMPLE: COUNTING CONTINGENCY TABLES

Suppose $A$ is an $n \times d$ totally unimodular matrix (see Section 7.1). Let us choose $b \in \mathbb{Z}^n$ such that the set $P_b$ of solutions to the system $Ax \leq b$ of linear inequalities is a simple polytope. Then $P_b$ is totally unimodular.

For example, if we know all the vertices of a simple transportation polytope $P$, we can compute the number of integral points of $P$ in time linear in the number of vertices of $P$.

One can construct an efficient algorithm for counting integral points in a polytope that is somewhat “close” to totally unimodular and for which the explicit formulas for $f_{K_c}(c)$ are therefore not too long.

One particular application is counting contingency tables (see Section 7.1), see [DLH+04].

CONNECTIONS WITH TORIC VARIETIES

It has been known since the 1970s that the number of integral points in an integral polytope is related to some algebro-geometric invariants of the associated toric variety (see [Oda88]). Naturally, for smooth toric varieties (they correspond to totally unimodular polytopes) computation is much easier. Various formulas for the number of integral points in polytopes were first obtained for totally unimodular polytopes and then, by the use of resolution of singularities, generalized to arbitrary integral polytopes (see, for example, [BP99]). Resolution of singularities of toric varieties reduces to dissection of a polyhedral cone into unimodular cones. However, as one can see, it is impossible to subdivide a rational cone into polynomially (in the input) many unimodular cones even in dimension $d = 2$. For example (see Figure 8.3.1), the plane cone $K$ generated by the points $(1, 0)$ and $(1, n)$ cannot be subdivided into fewer than $2n - 1$ unimodular cones, whereas a polynomial time subdivision would give a polynomial in $\log n$ cones. On the other hand, if we allow a signed linear combination of the inclusion-exclusion type, then one can easily represent this cone as a combination of 3 unimodular cones: $[K] = [K_1] + [K_2] + [K_3]$, where $K_1$ is generated by the basis $(1, 0)$ and $(0, 1)$, $K_2$ is generated by $(0, 1)$ and $(1, n)$, and $K_3$ is generated by $(1, n)$.

Moreover, modulo rational cones with lines (cf. Theorem 7.3.1), we need to use only two unimodular cones: $[K] = [K_3] + [K_4]$ modulo rational cones with lines, where $K_3$ is the cone generated by $(1, n)$ and $(0, -1)$ and $K_4$ is the cone generated by $(0, 1)$ and $(1, 0)$. Consequently, from Theorem 7.3.1,

$$f_K(c) = \frac{1}{(1 - e^{c_1} + e^{c_2})(1 - e^{-c_2})} + \frac{1}{(1 - e^{c_1})(1 - e^{c_2})} \quad \text{for} \quad c = (c_1, c_2).$$

As we have mentioned above, once we allow “signed” combinations, any rational polyhedral cone can be decomposed into unimodular cones in polynomial time, provided the dimension is fixed. Moreover, if we allow decompositions modulo rational cones with lines, the algorithm can be sped up further; roughly from $2^{O(d^2)}$ to $2^{O(d \ln d)}$ (see [BP99] and [Bar08]).
Chapter 7: Lattice points and lattice polytopes

Connections with valuations

The number of integral points \( \nu(P) = |P \cap \mathbb{Z}^d| \) in an integral polytope \( P \subset \mathbb{R}^d \) is a valuation, that is, it preserves linear relations among indicator functions of polytopes; and it is lattice-translation-invariant, i.e., \( \nu(P + l) = \nu(P) \) for any \( l \in \mathbb{Z}^d \).

General properties of valuations and the related notion of the “polytope algebra” have been intensively studied (see, for example, 

\[ \text{[McM93]} \] and 

\[ \text{[Mor93a]} \]). Various identities discovered in this area might prove useful in dealing with particular counting problems (see 

\[ \text{[BP99]} \] and 

\[ \text{[Bar08]} \]). For example, if the transportation polytope \( P_b \) is not simple, one can apply the following recipe. First, triangulating the normal cone at the vertex, we represent it as a combination of unimodular cones (we discard lower-dimensional cones). Then, passing to the dual cones, we get the desired representation of the cone of feasible directions (we discard cones with lines).

Exponential sums can be considered as a way to extend the counting valuation (the number of integer points in a polytope) from rational polytopes to rational polyhedra, possibly unbounded. Similarly, one can extend volume (which is obviously a valuation) from polytopes to possibly unbounded polyhedra by using exponential integrals

\[
\int_P \exp \{ \langle c, x \rangle \} \, dx
\]

and prove the corresponding version of Theorem 7.3.1, see 

\[ \text{[Bar08]} \]. In 

\[ \text{[Bar06]} \] a family of intermediate valuations interpolating between discrete (integer points) and continuous (volumes) valuations was introduced: for a given \( k \)-dimensional lattice subspace \( L \subset \mathbb{R}^d \) (subspace spanned by lattice points), with a polytope \( P \subset \mathbb{R}^d \), we associate the sum of \( k \)-dimensional volumes of sections \( P \cap A \), where \( A \) ranges over different lattice translates of \( L \). The corresponding theory of intermediate exponential valuations was developed by V. Baldoni, N. Berline, J.A. De Loera, M. Köppe, M. Vergne, see 

\[ \text{[BBD+12]} \], 

\[ \text{[BBK+13]} \] and references therein. These valuations turn out to be useful for efficient asymptotic counting of integer points in polytopes of varying dimension, see Section 7.4 below.

Analytic Methods

In 

\[ \text{[BH10]} \], the following approach to approximate counting was suggested. Suppose that the polytope \( P \subset \mathbb{R}^d \) is the intersection of an affine subspace of \( \mathbb{R}^d \) defined by the system of linear equations \( Ax = b \), where \( A \) is a \( k \times d \) integer matrix of rank \( k < d \) and the non-negative orthant \( \mathbb{R}^d_+ \). Suppose further that the relative interior

---

of \( P \) is non-empty. Then the strictly concave function \( g : \mathbb{R}_+^d \rightarrow \mathbb{R} \),

\[
g(x) = \sum_{i=1}^d \left( (x_i + 1) \ln (x_i + 1) - x_i \ln x_i \right) \quad \text{for} \quad x = (x_1, \ldots, x_d)
\]

attains its maximum on \( P \) at a (necessarily unique) point \( z = (z_1, \ldots, z_n) \), which can be easily computed by interior point methods. Suppose that \( \xi_1, \ldots, \xi_d \) are independent geometric random variables with expectations \( z_1, \ldots, z_d \) respectively. In other words, each random variable \( \xi_i \) accepts non-negative integer values and

\[
P(\xi_i = k) = p_i q_i^k \quad \text{for} \quad p_i = \frac{1}{1 + z_i}, \quad q_i = \frac{z_i}{1 + z_i} \quad \text{and} \quad k = 0, 1, 2, \ldots.
\]

Let \( \xi = (\xi_1, \ldots, \xi_d) \) be the corresponding random vector. It is shown in [BH10] that

\[
P(\xi = m) = e^{-g(z)} \quad \text{for all} \quad m \in P \cap \mathbb{Z}^d
\]

and hence one can express the number of integer points in \( P \) as

\[
|P \cap \mathbb{Z}^d| = e^{g(z)} P(A\xi = b). \tag{7.3.2}
\]

In fact, the random vector \( \xi \) has the largest entropy among all random vectors supported on the set \( \mathbb{Z}^d \) of non-negative integer vectors and with expectation in the affine subspace \( Ax = b \) while \( g(x) \) is the entropy of the vector of independent geometric random variables with expectation \( x \). The random \( k \)-dimensional vector \( A\xi \) is a linear combination of the \( d \) columns of matrix \( A \) with independent random coefficients and \( \mathbf{E}(A\xi) = Ax = b \). If \( k \ll d \), one can hope that in the spirit of the Local Central Limit Theorem, the distribution of \( A\xi \) is close to a multivariate Gaussian distribution, so the right-hand side of (7.3.2) can be estimated via the covariance matrix of \( A\xi \). Computing the covariance matrix, we obtain a heuristic formula

\[
|P \cap \mathbb{Z}^d| \approx \frac{e^{g(z)} (\det \Lambda)}{(2\pi)^{k/2} \sqrt{\det (BB^T)}}, \tag{7.3.3}
\]

where \( \Lambda \subset \mathbb{Z}^k \) is the lattice spanned by the columns of matrix \( A \) and \( B \) is the matrix obtained from \( A \) by multiplying the \( i \)-th column of \( A \) onto \( \sqrt{z_i^2 + z_i} \) for \( i = 1, \ldots, d \). It is shown in [BH10] that for many classes of polytopes \( P \) formula (7.3.3) indeed provides an asymptotically tight approximation for the number of integer points. J.A. De Loera reported encouraging results of some numerical experiments. For example, the exact number of \( 4 \times 4 \) non-negative integer matrices with row sums 220, 215, 93 and 64 and column sums 108, 286, 71 and 127 is 1,225,914,276,768,514, while (7.3.3) gives a 6\% error (we approximate the sum of 16 independent 7-dimensional vectors by the Gaussian distribution). The exact number of \( 3 \times 3 \times 3 \) arrays of non-negative integers with sums \([31, 22, 87], [50, 13, 77]\) and \([42, 87, 11]\) along the coordinate hyperplane “slices” is 8,846,838,772,161,591, while (7.3.3) gives a 0.19\% error (we approximate the sum of 27 independent 7-dimensional vectors by the Gaussian distribution). It is shown in [Ben14] that (7.3.3) is asymptotically exact for multiway contingency tables (3- and higher-dimensional arrays of non-negative integers with prescribed sums along the coordinate hyperplane slices). In [BH12], it is shown that in the case of classical contingency tables (non-negative integer matrices with prescribed row and column
sums) to obtain an asymptotically exact formula, one needs to introduce a correction to (7.3.3) (the so-called Edgeworth correction), which is an explicit, efficiently computable factor based on the third and fourth moments of the random vector $A\xi$.

The number of $m \times n$ non-negative integer matrices with equal row sums $s$ and equal column sums $t$ (and hence the total sum $N = ms = tn$ of entries) was earlier computed by E.R. Canfield and B.D. McKay [CM10]. Asymptotically, this number is

$\begin{align*}
\frac{(n+s-1)^m (m+t-1)^n}{(mn+N-1)} \exp \left\{ \frac{1}{2} \right\},
\end{align*}$

provided $n$ and $m$ grow roughly proportionately, see [CM10] for details. In [Sha10], non-trivial upper bounds for the number of integer points in polytopes are obtained via anti-concentration inequalities. In particular, by bounding the right-hand side of (7.3.2), the following simple and useful bound is obtained in [Sha10]:

$$|P \cap \mathbb{Z}^d| \leq e^{g(z)} \min_{j_1, \ldots, j_k} \prod_{i=1}^k \frac{1}{1 + z_{j_i}},$$

where the minimum is taken over all collections $j_1, \ldots, j_k$ of linearly independent columns of matrix $A$.

The following simple observation often leads to practically efficient (although theoretically exponential time) algorithms. Suppose that the polytope $P$ is defined as above as the intersection of the non-negative orthant in $\mathbb{R}^d$ with a $k$-dimensional affine subspace defined by the system $Ax = b$, where $A = (a_{ij})$ is a $k \times d$ integer matrix. Let $z_1, \ldots, z_k$ be (complex) variables and let

$$f_A(z_1, \ldots, z_k) = \prod_{j=1}^d \prod_{m=0}^{+\infty} \frac{z_1^{a_{1j}m} \cdots z_k^{a_{kj}m}}{1 - z_1^{a_{1j}} \cdots z_k^{a_{kj}}},$$

Thus $|P \cap \mathbb{Z}^d|$ is equal to the coefficient of $z_1^{b_1} \cdots z_k^{b_k}$ in the expansion of $f_A(z_1, \ldots, z_k)$ in a neighborhood of $z_1 = \ldots = z_k = 0$. This coefficient may be extracted by numerical differentiation, or by (repeated) application of the residue formula. M. Beck and D. Pixton [BP03] report results on numerical computation for the problem of counting contingency tables using repeated application of the residue formula.

As discussed in [BV97], various identities relating functions $f_A$ mirror corresponding identities among indicator functions of rational polyhedra. In particular, decompositions of $f_A$ into “simple fractions” correspond to decompositions of $P$ into simple cones.

Quite a few useful inequalities for the number of lattice points can be found in [GW93], [Lag95], and [GL87]. Blichfeldt’s inequality states that

$$|B \cap \Lambda| \leq \frac{d!}{\det \Lambda} \text{vol } B + d,$$

where $B$ is a convex body containing at least $d + 1$ affinely independent lattice points. Davenport’s inequality implies that

$$|B \cap \mathbb{Z}^d| \leq \sum_{i=0}^d \binom{d}{i} V_i(B),$$
where the $V_i$ are the intrinsic volumes. A conjectured stronger inequality, $|B \cap \mathbb{Z}^d| \leq V_0(K) + \ldots + V_d(K)$, was shown to be false in dimensions $d \geq 207$, although it is correct for $d = 2, 3$. Furthermore, H. Hadwiger proved that

$$|B \cap \mathbb{Z}^d| \geq \sum_{i=0}^{d} (-1)^{d-i}V_i(B),$$

provided $B \subset \mathbb{R}^d$ is a convex body having a nonempty interior (see \cite{Lag95}).

For applications of harmonic analysis, see \cite{BR07}.

**PROBABILISTIC METHODS**

Often, we need the number of integral points only approximately. Probabilistic methods based on Monte-Carlo methods have turned out to be quite successful. The main idea can be described as follows (see \cite{JS97}). Suppose we want to approximate the cardinality of a finite set $X$ (for example, $X$ may be the set of lattice points in a polytope). Suppose, further, that we can present a “filtration” $X_0 \subset X_1 \subset \ldots \subset X_n = X$, where $|X_0| = 1$ (in general, we require $|X_0|$ to be small) and $|X_{i+1}|/|X_i| \leq 2$ (in general, we require the ratio $|X_{i+1}|/|X_i|$ to be reasonably small). Finally, suppose that we have an efficient procedure for sampling an element $x \in X_i$ uniformly at random (in practice, we settle for “almost uniform” sampling). Given an $\epsilon > 0$ and a $\delta > 0$, with probability at least $1 - \delta$ one can estimate the ratio $|X_{i+1}|/|X_i|$, within a relative error $\epsilon/n$, by sampling $O(n\epsilon^{-2} \ln \delta^{-1})$ points at random from $X_{i+1}$ and counting how many times the points end up in $X_i$. Then, by “telescoping,” with probability at least $(1 - \delta)^n$, we estimate

$$|X| = |X_n| = \frac{|X_n|}{|X_{n-1}|} \ldots \frac{|X_{i+1}|}{|X_i|} \ldots \frac{|X_2|}{|X_1|}$$

within relative error $\epsilon$.

The bottleneck of the method is the ability to sample a point $x \in X_i$ uniformly at random. To achieve that, a Markov chain on $X_i$ is designed, which converges fast (“mixes rapidly”) to the uniform distribution. Usually, there are some natural candidates for such Markov chains and the main difficulty is to establish whether they indeed mix rapidly.

Counting various combinatorial structures can be interpreted as counting vertices in a certain $(0, 1)$-polytope. For example, computing the number of perfect matchings in a given bipartite graph on $n + n$ vertices, or, equivalently, computing the permanent of a given $n \times n$ matrix of 0’s and 1’s, can be viewed as counting the number of vertices in a particular face of the Birkhoff polytope $B_n$. M. Jerrum, A. Sinclair, and E. Vigoda \cite{JSV04} have constructed a polynomial-time probabilistic algorithm to approximate the permanent of any given nonnegative matrix. B. Morris and A. Sinclair \cite{MS99} have presented a polynomial-time probabilistic algorithm to compute the number of $(0, 1)$-vectors $(x_1, \ldots, x_n)$ satisfying the inequality $a_1x_1 + \ldots + a_nx_n \leq b_n$, where $a_i$ and $b$ are given positive integers, see also \cite{Dye03} for a deterministic algorithm and \cite{CDR10} for further applications of dynamic programming in derandomization.

R. Kannan and S. Vempala proved \cite{KV99} that if a polytope $P \subset \mathbb{R}^d$ with $m$ facets contains a ball of radius $d\sqrt{\ln m}$ then the number of integer points in $P$ is well approximated by the volume of $P$ and, moreover, sampling a random
point from the uniform (or almost uniform) distribution \( P \cap \mathbb{Z}^d \) can be achieved by sampling a random point from the uniform (or almost uniform) distribution in \( P \) (which is an easier problem, see [JS97] and [Vem10]), rounding the obtained point to an integer point and accepting it if the resulting integer point lies in \( P \). This leads to a polynomial time algorithm for sampling and counting of integer points in such polytopes \( P \).

### 7.4 ASYMPTOTIC PROBLEMS

If \( P \subset \mathbb{R}^d \) is an integral polytope then the number of integral points in the dilated polytope \( nP = \{ nx \mid x \in P \} \) for a natural number \( n \) is a polynomial in \( n \), known as the Ehrhart polynomial. We review several results concerning the Ehrhart polynomial and its generalizations.

#### GLOSSARY

**Cone of feasible directions at a face of a polytope:** The cone \( K_F \) of feasible directions at any point in the relative interior of the face \( F \subset P \).

**Tangent cone at a face of a polytope:** The translation \( x + K_F \) of the cone of feasible directions \( K_F \) by any point \( x \) in the face \( F \subset P \).

**Apex of a cone:** The largest linear subspace contained in the cone.

**Dual cone:** The cone \( K^* = \{ x \in \mathbb{R}^d \mid \langle x, y \rangle \leq 0 \text{ for all } y \in K \} \), where \( K \subset \mathbb{R}^d \) is a given cone.

**\( \text{vol}_k \):** The normalized \( k \)-volume of a \( k \)-dimensional rational polytope \( P \subset \mathbb{R}^d \) computed as follows. Let \( L \subset \mathbb{R}^d \) be the \( k \)-dimensional linear subspace parallel to the affine span of \( P \). Then \( \text{vol}_k(P) \) is the Euclidean \( k \)-dimensional volume of \( P \) in the affine span of \( P \) divided by the determinant of the lattice \( \Lambda = \mathbb{Z}^d \cap L \).

**Lattice subspace:** A subspace spanned by lattice points.

#### EHRHART POLYNOMIALS

The following fundamental result was suggested by Ehrhart.

**THEOREM 7.4.1**

Let \( P \subset \mathbb{R}^d \) be an integral polytope. For a natural number \( n \) we denote by \( nP = \{ nx \mid x \in P \} \) the \( n \)-fold dilation of \( P \). Then the number of integral points in \( nP \) is a polynomial in \( n \):

\[
|nP \cap \mathbb{Z}^d| = E_P(n) \quad \text{for some polynomial } \ E_P(x) = \sum_{i=0}^{d} e_i(P) \cdot x^i.
\]

Moreover, for positive integers \( n \) the value of \( (-1)^{\deg E_P} E_P(-n) \) is equal to the number of integral points in the relative interior of the polytope \( nP \) (the “reciprocity law”).
The polynomial $E_P$ is called the **Ehrhart polynomial** and its coefficients $e_i(P)$ are called **Ehrhart coefficients**. For various proofs of Theorem 7.4.1 see, for example, [Sta83], [Sta86], [BR07] and [Bar08]. The existence of the Ehrhart polynomials and the reciprocity law can be derived from the single fact that the number of integral points in a polytope is a lattice-translation-invariant valuation (see [McM93] and Section 7.3 above).

If $P$ is a rational polytope, one can define $e_k(P) = n^{-k}e_k(nP)$, where $n$ is a positive integer such that $nP$ is an integral polytope. For an integral polytope $P \subseteq \mathbb{R}^d$, one has $|P \cap \mathbb{Z}^d| = e_0(P) + e_1(P) + \ldots + e_d(P)$. (This formula is no longer true, however, if $P$ is a general rational polytope.) The Ehrhart coefficients constitute a basis of all additive functions (valuations) $\nu$ on rational polytopes that are invariant under unimodular transformations (see [McM93] and [GW93]).

**GENERAL PROPERTIES**

It is known that $e_0(P) = 1$, $e_d(P) = \text{vol}_d(P)$, and $e_{d-1}(P) = \frac{1}{2} \sum_F \text{vol}_{d-1} F$, where the sum is taken over all the facets of $P$. Thus, computation of the two highest coefficients reduces to computation of the volume. In fact, the computation of any fixed number of the highest Ehrhart coefficients of an $\mathcal{H}$-polytope reduces in polynomial time to the computation of the volumes of faces; see [BP99], [Bar08] and also below.

**EXISTENCE OF LOCAL FORMULAS**

The Ehrhart coefficients can be decomposed into a sum of “local” summands. The following theorem was proven by P. McMullen (see [McM93], [Mor93a], and [BP99]).

**THEOREM 7.4.2**

For any natural numbers $k$ and $d$ there exists a real valued function $\mu_{k,d}$, defined on the set of all rational polyhedral cones $K \subseteq \mathbb{R}^d$, such that for every rational full-dimensional polytope $P \subseteq \mathbb{R}^d$ we have

$$e_k(P) = \sum_F \mu_{k,d}(K_F) \cdot \text{vol}_k F,$$

where the sum is taken over all $k$-dimensional faces $F$ of $P$ and $K_F$ is the cone of feasible directions of $P$ at the face $F$. Moreover, one can choose $\mu_{k,d}$ to be an additive measure on polyhedral cones.

Different explicit and also computationally efficient constructions of $\mu_{d,k}$ were described by R. Morelli [Mor93b], J. Pommersheim and H. Thomas [PT04] and by N. Berline and M. Vergne [BV07].

In general, suppose $V$ is a finite-dimensional real space and let $\Lambda \subset V$ be a lattice that spans $V$. Then we consider lattice polytopes in $V$ and in every lattice subspace $L \subset V$ we define volume so that $\det(\Lambda \cap L) = 1$. Theorem 7.4.2 holds in this generality, though the function $\mu_{k,d}$ that satisfies the conditions of Theorem 7.4.2 is not unique. To make a canonical choice one has to introduce some additional structure, such as an inner product as in [Mor93b] and [BV07] or
fix a flag of subspaces in $V$ as in [PT04]. Essentially, one needs to be able to choose canonically the complement to a given subspace.

For some specific values of $k$ and $d$ convenient choices of $\mu_{k,d}$ has long been known.

**EXAMPLE**

For a cone $K \subset \mathbb{R}^d$, let $\gamma(K)$ be the spherical measure of $K$ normalized in such a way that $\gamma(\mathbb{R}^d) = 1$. Thus $\gamma(K) = 0.5$ if $K$ is a halfspace. One can choose $\mu_{d,d} = \mu_{d-1,d} = \gamma$ because of the formulas for $e_d(P)$ and $e_{d-1}(P)$ (see above).

On the other hand, one can choose $\mu_{0,d}(K) = \gamma(K^*)$, where $K^*$ is the dual cone, since it is known that $e_0(P) = 1$. We note that if $\mu(K)$ is an additive measure on polyhedral cones then $\nu(K) = \mu(K^*)$ is also an additive measure on polyhedral cones, see, for example, [Bar08]. Moreover, for integral zonotopes (see Section 7.3), one can always choose $\mu_{k,d}(K_F) = \gamma(K_F^*)$ [BP99]. If $F$ is a $k$-dimensional face of $P$ then $K_F^*$ is a $(d-k)$-dimensional face of $K_F^*$.

**BERLINE–VERGNE FORMULAS**

We describe an elegant and computationally efficient choice of $\mu_{k,d}$ suggested by N. Berline and M. Vergne in [BV07], see also [Bar08] of an exposition.

Let $V$ be a real vector space endowed with scalar product $\langle \cdot, \cdot \rangle$, let $\Lambda \subset V$ be a lattice that spans $V$ and let $K \subset V$ be a pointed rational polyhedral cone with non-empty interior. Let $d = \dim V$. Our immediate goal is to construct a meromorphic function $\psi(K;c) : V \to \mathbb{C}$, which we define recursively for $d = 0, 1, \ldots$. If $\dim V = 0$, we define $\psi(K;c) = 1$. Suppose that $d > 0$ and let $L \subset V$ be a lattice subspace. We introduce the volume form $dx_L$ in $L$ so that $\det(\Lambda \cap L) = 1$. Let $V/L$ be the orthogonal complement of $L$, let $\Lambda/L \subset V/L$ be the orthogonal projection of $\Lambda$ onto $V/L$, which is necessarily a lattice there, and let $K/L \subset V/L$ be the orthogonal projection of $K$, which is necessarily a rational cone with non-empty interior. We define $\psi(K;c)$ so that the following identity holds:

$$
\sum_{m \in K \cap \Lambda} e^{\langle c, m \rangle} = \sum_L \psi(K/L; c) \int_{K \cap L} e^{\langle c, x \rangle} dx_L, \tag{7.4.1}
$$

where the sum is taken over all subspaces $L$ spanned by the faces of $K$. We note that unless $L = \{0\}$, we have $\dim V/L < d$ and hence $\psi(K/L; c)$ has been already defined. Since $K$ is a pointed cone, there is a non-empty open set $U \subset V$ for which all the integrals in (7.4.1) converge absolutely, which allows us to define $\psi(K;c)$. Note that the identity (7.4.1) can be understood in terms of exponential valuations (see Section 7.3); in which case one can formally consider the sum over all lattice subspaces $L$: indeed, if $\dim(K \cap L) < \dim L$ then the corresponding integral is 0 and if $L$ intersects the interior of $K$ then $K/L$ contains a line, in which case $\psi(K/L; c) \equiv 0$. Thus only subspaces $L$ spanned by faces of $K$ contribute non-zero terms in (7.4.1). It turns out that $K \mapsto \psi(K;c)$ extends to a valuation with values in the ring of meromorphic functions, that $\psi(K;c) = 0$ provided $K$ contains a line or has empty interior and that $\psi(K;c)$ is regular at $c = 0$. We then define $\mu_{k,d}(K_F)$ in Theorem 7.4.2 as $\psi(K;0)$, where $K = K_F/L_F$ and $L_F$ is the apex of $K_F$. 
Using the technique of exponential sums, one can show that \( \psi(K; c) \) is computable in polynomial time if \( \dim K \) is fixed. Consequently, computation of any fixed number of the highest Ehrhart coefficients reduces in polynomial time to computation of the volumes of faces for an \( \mathcal{H} \)-polytope (see \[HP99\] and \[Bar08\]).

**The \( h^* \)-Vector**

General properties of generating functions (see \[Sta86\]) imply that for every integral \( d \)-dimensional polytope \( P \) there exist integers \( h_k^0(P), \ldots, h_k^d(P) \) such that

\[
\sum_{n=0}^{\infty} E_P(n)x^n = \frac{h_0^k(P) + h_1^k(P)x + \ldots + h_d^k(P)x^d}{(1-x)^{d+1}}.
\]

The \((d+1)\)-vector \( h^*(P) = (h_0^k(P), \ldots, h_d^k(P)) \) is called the \( h^*-vector \) of \( P \). It is clear that \( h^*(P) \) is a (vector-valued) valuation on the set of integral polytopes and that \( h^*(P) \) is invariant under unimodular transformations of \( \mathbb{Z}^d \). Moreover, the functions \( h_k^*(P) \) constitute a basis of all valuations on integral polytopes that are invariant under unimodular transformations. Unlike the Ehrhart coefficients \( e_k(P) \), the numbers \( h_k^*(P) \) are not homogeneous. However, \( h_k^*(P) \) are monotone (and, therefore, nonnegative): if \( Q \subset P \) are two integral polytopes then \( h_k^*(P) \geq h_k^*(Q) \) \( \[Sta93\] \).

The largest \( k \) such that \( h_k^*(P) \neq 0 \) is called the degree of \( P \). Equivalently, \( k \) is the smallest non-negative integer such that the dilated polytope \((d-k)P\) has no interior lattice points. C. Haase, B. Nill, and S. Payne proved a decomposition theorem for polytopes of a fixed degree \[HNP09\], and, as a corollary, established that the volume of \( P \) is bounded from above by a function of its degree \( k \) and the value of \( h_k^*(P) \), independently of the dimension, thus confirming a conjecture of V. Batyrev.

In principle, there is a combinatorial way to calculate \( h^*(P) \). Namely, let \( \Delta \) be a triangulation of \( P \) such that every \( d \)-dimensional simplex of \( \Delta \) is integral and has volume \( 1/d! \) (see Section \[72\]). Let \( f_k(\Delta) \) be the number of \( k \)-dimensional faces of the triangulation \( \Delta \). Then

\[
h_k^*(P) = \sum_{i=0}^{k} (-1)^{k-i} \binom{d-i}{d-k} f_{i-1}(\Delta),
\]

where we let \( f_{-1}(\Delta) = 1 \). Such a triangulation may not exist for the polytope \( P \) but it exists for \( mP \), where \( m \) is a sufficiently large integer \[KKMS73\]. Generally, this triangulation \( \Delta \) would be too big, but for some special polytopes with nice structure (for example, for the so-called poset polytopes) it may provide a reasonable way to compute \( h^*(P) \) and hence the Ehrhart polynomial \( E_P \).

Since the number of integral points in a polytope is a valuation, we get the following result proved by P. McMullen (see \[McM93\]).

**Theorem 7.4.3**

Let \( P_1, \ldots, P_m \) be integral polytopes in \( \mathbb{R}^d \). For an \( m \)-tuple of natural numbers \( \mathbf{n} = (n_1, \ldots, n_m) \), let us define the polytope

\[
P(\mathbf{n}) = \{n_1x_1 + \ldots + n_mx_m \mid x_1 \in P_1, \ldots, x_m \in P_m\}
\]
(using “+” for Minkowski addition one can also write \( P(n) = n_1P_1 + \ldots + n_mP_m \).
Then there exists a polynomial \( p(x_1, \ldots, x_m) \) of degree at most \( d \) such that
\[
|P(n) \cap \mathbb{Z}^d| = p(n_1, \ldots, n_m).
\]

More generally, the existence of local formulas for the Ehrhart coefficients implies that the number of integral points in an integral polytope \( P_h = \{ x \in \mathbb{R}^d \mid Ax \leq b + h \} \) is a polynomial in \( h \) provided \( P_h \) is an integral polytope combinatorially isomorphic to the integral polytope \( P_0 \), see, for example, [Bar08]. In other words, if we move the facets of an integral polytope parallel to themselves so that it remains integral and has the same facial structure, then the number of integral points varies polynomially.

### INTEGRAL POINTS IN RATIONAL POLYTOPES

If \( P \) is a rational (not necessarily integral) polytope then \( |nP \cap \mathbb{Z}^d| \) is not a polynomial but a quasipolynomial (a function of \( n \) whose value cycles through the values of a finite list of polynomials). The following result was independently proven by P. McMullen and R. Stanley (see [McM93] and [Sta86]).

**Theorem 7.4.4**

Let \( P \subset \mathbb{R}^d \) be a rational polytope. For every \( r \), \( 0 \leq r \leq d \), let \( \text{ind}_r \) be the smallest natural number \( k \) such that all \( r \)-dimensional faces of \( kP \) are integral polytopes. Then, for every \( n \in \mathbb{N} \),
\[
|nP \cap \mathbb{Z}^d| = \sum_{r=0}^{d} e_r(P, n(\text{mod} \, \text{ind}_r)) \cdot n^r
\]
for suitable rational numbers \( e_r(P, k), \ 0 \leq k < \text{ind}_r \).

P. McMullen also obtained a generalization of the “reciprocity law” (see [Sta86] and [McM93]).

Let us fix an \( n \times d \) integer matrix \( A \) such that the set \( P_b = \{ x \mid Ax \leq b \} \), \( b \in \mathbb{Z}^n \), if nonempty, is a rational polytope. Let \( B \subset \mathbb{Z}^n \) be a set of right-hand-side vectors \( b \) such that the combinatorial structure of \( P_b \) is the same for all \( b \in B \). In [BP99] it is shown that as long as the dimension \( d \) is fixed, one can find a polynomially computable formula \( F(b) \) for the number \( |P_b \cap \mathbb{Z}^d| \), where \( F \) is a polynomial of degree \( d \) in integer parts of linear functions of \( b \). It is based on Brion’s Theorem (Theorem 7.3.2) and the “rounding” of rational translations of unimodular cones.

Theorem 7.4.2 and Berline-Vergne formulas extend to rational polytopes. In Theorem 7.4.2, measures \( \mu_{k,d} \) depend on the translation class of the tangent cone \( x + K_F \) modulo integer translations. Similarly, Berline-Vergne functions \( \psi(x + K; c) \) are defined for translations of rational cones and also invariant under lattice translations. Consequently, the computation of \( e_r(P, k) \) reduces in polynomial time to the computation of the volume of faces of \( P \) as long as the codimension \( d - r \) is fixed. A different approach to computing the coefficients \( e_r(P; k) \) in fixed codimension is via intermediate valuations, see [Bar06] and [BBD+12].

Interestingly, for a “typical” (and, therefore, nonrational) polytope \( P \) the difference \( |tP \cap \mathbb{Z}^d| - t^d \text{vol } P \) has order \( O((\ln t)^{d-1+\epsilon}) \) as \( t \to +\infty \) [Skr98].

---

7.5 PROBLEMS WITH QUANTIFIERS

A natural generalization of the decision problem (see Section 7.2) is a problem with quantifiers. We describe some known results and formulate open questions for this class of problems.

FROBENIUS PROBLEM

The most famous problem from this class is the Frobenius problem:

Given \( k \) coprime positive integers \( a_1, \ldots, a_k \), find the largest integer \( m \) that cannot be represented as a non-negative integer combination \( a_1n_1 + \ldots + a_kn_k \).

The problem is known to be NP-hard in general, but a polynomial time algorithm is known for fixed \( k \) [Kan92].

PROBLEM WITH QUANTIFIERS

A general problem with quantifiers can be formulated as follows. Suppose that \( P \) is a Boolean combination of convex polyhedra: we start with some polyhedra \( P_1, \ldots, P_k \subset \mathbb{R}^d \) given by their facet descriptions and construct \( P \) by using the set-theoretical operations of union, intersection, and complement. We want to find out if the formula

\[
\exists x_1 \forall x_2 \exists x_3 \ldots \forall x_m : (x_1, \ldots, x_m) \in P \tag{7.5.1}
\]

is true. Here \( x_i \) is an integral vector from \( \mathbb{Z}^{d_i} \), and, naturally, \( d_1 + \ldots + d_m = d \), \( d_i \geq 0 \). The parameters that characterize the size of (7.5.1) can be divided into two classes. The first class consists of the parameters characterizing the combinatorial size of the formula. These are the dimension \( d \), the number \( m - 1 \) of quantifier alternations, the number of linear inequalities and Boolean operations that define the polyhedral set \( P \). The parameters from the other class characterize the numerical size of the formula. Those are the bit sizes of the numbers involved in the inequalities that define \( P \).

The following fundamental question remains open.

PROBLEM 7.5.1

Let us fix all the combinatorial parameters of the formula (8.5.1). Does there exist a polynomial time algorithm that checks whether this formula is true?

Naturally, “polynomial time” means that the running time of the algorithm is bounded by a polynomial in the numerical size of the formula. The answer to this question is unknown. A polynomial time algorithm is known if the formula contains not more than 1 quantifier alternation, i.e., if \( m \leq 2 \) [Kan90].

Let \( P \subset \mathbb{R}^n \) be a rational polytope, let \( pr : \mathbb{R}^n \rightarrow \mathbb{R}^d \) be the projection on the first \( d \) coordinates, and let \( S = pr (P \cap \mathbb{Z}^n) \) be the projection of the set of integer points in \( P \). For any fixed \( n \), A. Barvinok and K. Woods [BAW93] constructed a polynomial time algorithm that, given \( P \), computes the exponential sum over \( S \) in...
the form
\[
\sum_{m \in S} e^{(c,m)} = \sum_{i \in I} \alpha_i \left(1 - e^{(c,b_{i1})}\right) \cdots \left(1 - e^{(c,b_{ik})}\right),
\]
where \(\alpha_i \in \mathbb{Q}, a_i \in \mathbb{Z}^d\) and \(b_{ij} \in \mathbb{Z}^d \setminus \{0\}\). Besides, \(k\) depends only on \(n\) and \(d\). Such sets \(S\) can be defined by formulas with no quantifier alternations, see also \[Woo14\] for some related developments. As a corollary, for any fixed \(d\), we obtain a polynomial time algorithm that, for any given coprime positive integer \(s a_1, \ldots, a_d\), computes the number of non-negative integers \(m\) that are not non-negative integer combinations of \(a_1, \ldots, a_d\).

7.6 SOURCES AND RELATED MATERIAL

RELATED CHAPTERS

Chapter 3: Tilings
Chapter 15: Basic properties of convex polytopes
Chapter 16: Subdivisions and triangulations of polytopes
Chapter 36: Computational convexity

REFERENCES


208 A. Barvinok


