

Counting Lattice Points in Polytopes: The Ehrhart Theory

Ubi materia, ibi geometria.

Johannes Kepler (1571–1630)

Given the profusion of examples that gave rise to the polynomial behavior of the integer-point counting function $L_{\mathcal{P}}(t)$ for special polytopes \mathcal{P} , we now ask whether there is a general structure theorem. As the ideas unfold, the reader is invited to look back at Chapters 1 and 2 as appetizers and indeed as special cases of the theorems developed below.

3.1 Triangulations and Pointed Cones

Because most of the proofs that follow work like a charm for a simplex, we first dissect a polytope into simplices. This dissection is captured by the following definition.

A **triangulation** of a convex d -polytope \mathcal{P} is a finite collection T of d -simplices with the following properties:

- $\mathcal{P} = \bigcup_{\Delta \in T} \Delta$.
- For any $\Delta_1, \Delta_2 \in T$, $\Delta_1 \cap \Delta_2$ is a face common to both Δ_1 and Δ_2 .

We say that \mathcal{P} can be **triangulated using no new vertices** if there exists a triangulation T such that the vertices of any $\Delta \in T$ are vertices of \mathcal{P} .

Theorem 3.1 (Existence of triangulations). *Every convex polytope can be triangulated using no new vertices.*

This theorem seems intuitively obvious but is not entirely trivial to prove. We carefully work out a proof in the appendix.

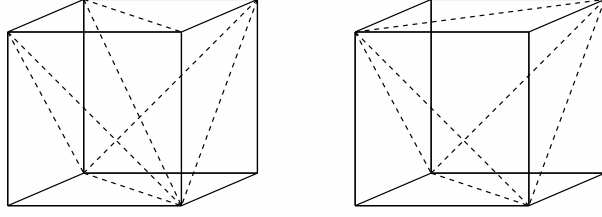


Fig. 3.1. Two (very different) triangulations of the 3-cube.

A **pointed cone** $\mathcal{K} \subseteq \mathbb{R}^d$ is a set of the form

$$\mathcal{K} = \{\mathbf{v} + \lambda_1 \mathbf{w}_1 + \lambda_2 \mathbf{w}_2 + \cdots + \lambda_m \mathbf{w}_m : \lambda_1, \lambda_2, \dots, \lambda_m \geq 0\},$$

where $\mathbf{v}, \mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m \in \mathbb{R}^d$ are such that there exists a hyperplane H for which $H \cap \mathcal{K} = \{\mathbf{v}\}$; that is, $\mathcal{K} \setminus \{\mathbf{v}\}$ lies strictly on one side of H . The vector \mathbf{v} is called the **apex** of \mathcal{K} , and the \mathbf{w}_k 's are the **generators** of \mathcal{K} . The cone is **rational** if $\mathbf{v}, \mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m \in \mathbb{Q}^d$, in which case we may choose $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m \in \mathbb{Z}^d$ by clearing denominators. The **dimension** of \mathcal{K} is the dimension of the affine space spanned by \mathcal{K} ; if \mathcal{K} is of dimension d , we call it a d -cone. The d -cone \mathcal{K} is **simplicial** if \mathcal{K} has precisely d linearly independent generators.

Just as polytopes have a description as an intersection of half-spaces, so do pointed cones: A rational pointed d -cone is the d -dimensional intersection of finitely many half-spaces of the form

$$\{\mathbf{x} \in \mathbb{R}^d : a_1 x_1 + a_2 x_2 + \cdots + a_d x_d \leq b\}$$

with integral parameters $a_1, a_2, \dots, a_d, b \in \mathbb{Z}$ such that the corresponding hyperplanes of the form

$$\{\mathbf{x} \in \mathbb{R}^d : a_1 x_1 + a_2 x_2 + \cdots + a_d x_d = b\}$$

meet in exactly one point.

Cones are important for many reasons. The most practical for us is a process called *coning over a polytope*. Given a convex polytope $\mathcal{P} \subset \mathbb{R}^d$ with vertices $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$, we lift these vertices into \mathbb{R}^{d+1} by adding a 1 as their last coordinate. So, let

$$\mathbf{w}_1 = (\mathbf{v}_1, 1), \mathbf{w}_2 = (\mathbf{v}_2, 1), \dots, \mathbf{w}_n = (\mathbf{v}_n, 1).$$

Now we define the **cone over** \mathcal{P} as

$$\text{cone}(\mathcal{P}) = \{\lambda_1 \mathbf{w}_1 + \lambda_2 \mathbf{w}_2 + \cdots + \lambda_n \mathbf{w}_n : \lambda_1, \lambda_2, \dots, \lambda_n \geq 0\} \subset \mathbb{R}^{d+1}.$$

This pointed cone has the origin as apex, and we can recover our original polytope \mathcal{P} (strictly speaking, the translated set $\{(\mathbf{x}, 1) : \mathbf{x} \in \mathcal{P}\}$) by cutting $\text{cone}(\mathcal{P})$ with the hyperplane $x_{d+1} = 1$, as shown in Figure 3.2.

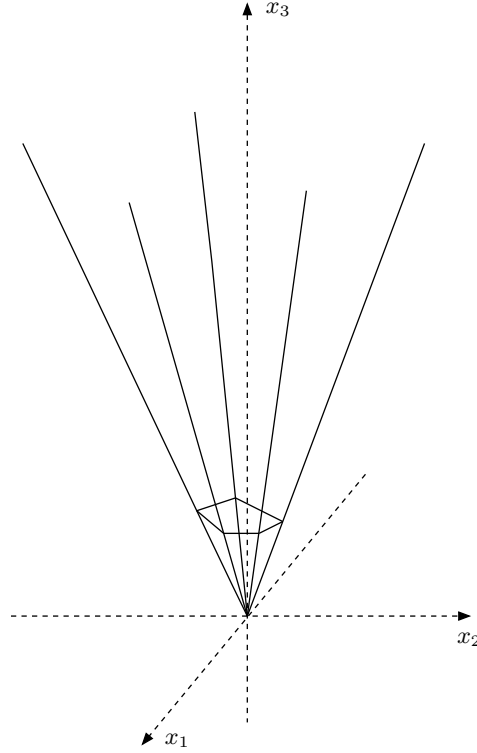


Fig. 3.2. Coning over a polytope.

By analogy with the language of polytopes, we say that the hyperplane $H = \{\mathbf{x} \in \mathbb{R}^d : \mathbf{a} \cdot \mathbf{x} = b\}$ is a **supporting hyperplane** of the pointed d -cone \mathcal{K} if \mathcal{K} lies entirely on one side of H , that is,

$$\mathcal{K} \subset \{\mathbf{x} \in \mathbb{R}^d : \mathbf{a} \cdot \mathbf{x} \leq b\} \quad \text{or} \quad \mathcal{K} \subset \{\mathbf{x} \in \mathbb{R}^d : \mathbf{a} \cdot \mathbf{x} \geq b\}.$$

A **face** of \mathcal{K} is a set of the form $\mathcal{K} \cap H$, where H is a supporting hyperplane of \mathcal{K} . The $(d-1)$ -dimensional faces are called **facets** and the 1-dimensional faces **edges** of \mathcal{K} . The apex of \mathcal{K} is its unique 0-dimensional face.

Just as polytopes can be triangulated into simplices, pointed cones can be triangulated into simplicial cones. So, a collection T of simplicial d -cones is a **triangulation** of the d -cone \mathcal{K} if it satisfies:

- $\mathcal{K} = \bigcup_{\mathcal{S} \in T} \mathcal{S}$.
- For any $\mathcal{S}_1, \mathcal{S}_2 \in T$, $\mathcal{S}_1 \cap \mathcal{S}_2$ is a face common to both \mathcal{S}_1 and \mathcal{S}_2 .

We say that \mathcal{K} can be **triangulated using no new generators** if there exists a triangulation T such that the generators of any $\mathcal{S} \in T$ are generators of \mathcal{P} .

Theorem 3.2. *Any pointed cone can be triangulated into simplicial cones using no new generators.*

Proof. This theorem is really a corollary to Theorem 3.1. Given a pointed d -cone \mathcal{K} , there exists a hyperplane H that intersects \mathcal{K} only at the apex. Now translate H “into” the cone, so that $H \cap \mathcal{K}$ consists of more than just one point. This intersection is a $(d-1)$ -polytope \mathcal{P} , whose vertices are determined by the generators of \mathcal{K} . Now triangulate \mathcal{P} using no new vertices. The cone over each simplex of the triangulation is a simplicial cone. These simplicial cones, by construction, triangulate \mathcal{K} . \square

3.2 Integer-Point Transforms for Rational Cones

We want to encode the information contained by the lattice points in a set $S \subset \mathbb{R}^d$. It turns out that the following multivariate generating function allows us to do this in an efficient way if S is a rational cone or polytope:

$$\sigma_S(\mathbf{z}) = \sigma_S(z_1, z_2, \dots, z_d) := \sum_{\mathbf{m} \in S \cap \mathbb{Z}^d} \mathbf{z}^{\mathbf{m}}.$$

The generating function σ_S simply lists all integer points in S in a special form: not as a list of vectors, but as a formal sum of monomials. We call σ_S the **integer-point transform** of S ; the function σ_S also goes by the name *moment generating function* or simply *generating function* of S . The integer-point transform σ_S opens the door to both algebraic and analytic techniques.

Example 3.3. As a warm-up example, consider the 1-dimensional cone $\mathcal{K} = [0, \infty)$. Its integer-point transform is our old friend

$$\sigma_{\mathcal{K}}(z) = \sum_{m \in [0, \infty) \cap \mathbb{Z}} z^m = \sum_{m \geq 0} z^m = \frac{1}{1-z}. \quad \square$$

Example 3.4. Now we consider the two-dimensional cone

$$\mathcal{K} := \{\lambda_1(1, 1) + \lambda_2(-2, 3) : \lambda_1, \lambda_2 \geq 0\} \subset \mathbb{R}^2$$

depicted in Figure 3.3. To obtain the integer-point transform $\sigma_{\mathcal{K}}$, we tile \mathcal{K} by copies of the *fundamental parallelogram*

$$\Pi := \{\lambda_1(1, 1) + \lambda_2(-2, 3) : 0 \leq \lambda_1, \lambda_2 < 1\} \subset \mathbb{R}^2.$$

More precisely, we translate Π by nonnegative integer linear combinations of the generators $(1, 1)$ and $(-2, 3)$, and these translates will exactly cover

\mathcal{K} . How can we list the integer points in \mathcal{K} as monomials? Let's first list all vertices of the translates of Π . These are nonnegative integer combinations of the generators $(1, 1)$ and $(-2, 3)$, so we can list them using geometric series:

$$\sum_{\substack{\mathbf{m}=j(1,1)+k(-2,3) \\ j,k \geq 0}} \mathbf{z}^{\mathbf{m}} = \sum_{j \geq 0} \sum_{k \geq 0} \mathbf{z}^{j(1,1)+k(-2,3)} = \frac{1}{(1 - z_1 z_2)(1 - z_1^{-2} z_2^3)}.$$

We now use the integer points $(m, n) \in \Pi$ to generate a subset of \mathbb{Z}^2 by adding to (m, n) nonnegative linear integer combinations of the generators $(1, 1)$ and $(-2, 3)$. Namely, we let

$$\mathcal{L}_{(m,n)} := \{(m, n) + j(1, 1) + k(-2, 3) : j, k \in \mathbb{Z}_{\geq 0}\}.$$

It is immediate that $\mathcal{K} \cap \mathbb{Z}^2$ is the disjoint union of the subsets $\mathcal{L}_{(m,n)}$ as (m, n) ranges over $\Pi \cap \mathbb{Z}^2 = \{(0, 0), (0, 1), (0, 2), (-1, 2), (-1, 3)\}$. Hence

$$\begin{aligned} \sigma_{\mathcal{K}}(\mathbf{z}) &= (1 + z_2 + z_2^2 + z_1^{-1} z_2^2 + z_1^{-1} z_2^3) \sum_{\substack{\mathbf{m}=j(1,1)+k(-2,3) \\ j,k \geq 0}} \mathbf{z}^{\mathbf{m}} \\ &= \frac{1 + z_2 + z_2^2 + z_1^{-1} z_2^2 + z_1^{-1} z_2^3}{(1 - z_1 z_2)(1 - z_1^{-2} z_2^3)}. \end{aligned} \quad \square$$

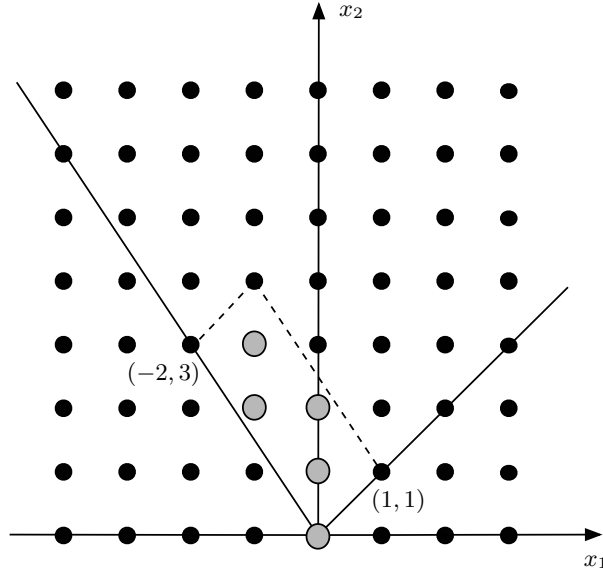


Fig. 3.3. The cone \mathcal{K} and its fundamental parallelogram.

Similar geometric series suffice to describe integer-point transforms for simplicial d -cones. The following result utilizes the geometric series in several directions simultaneously.

Theorem 3.5. *Suppose*

$$\mathcal{K} := \{\lambda_1 \mathbf{w}_1 + \lambda_2 \mathbf{w}_2 + \cdots + \lambda_d \mathbf{w}_d : \lambda_1, \lambda_2, \dots, \lambda_d \geq 0\}$$

is a simplicial d -cone, where $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_d \in \mathbb{Z}^d$. Then for $\mathbf{v} \in \mathbb{R}^d$, the integer-point transform $\sigma_{\mathbf{v}+\mathcal{K}}$ of the shifted cone $\mathbf{v}+\mathcal{K}$ is the rational function

$$\sigma_{\mathbf{v}+\mathcal{K}}(\mathbf{z}) = \frac{\sigma_{\mathbf{v}+\Pi}(\mathbf{z})}{(1 - \mathbf{z}^{\mathbf{w}_1})(1 - \mathbf{z}^{\mathbf{w}_2}) \cdots (1 - \mathbf{z}^{\mathbf{w}_d})},$$

where Π is the half-open parallelepiped

$$\Pi := \{\lambda_1 \mathbf{w}_1 + \lambda_2 \mathbf{w}_2 + \cdots + \lambda_d \mathbf{w}_d : 0 \leq \lambda_1, \lambda_2, \dots, \lambda_d < 1\}.$$

The half-open parallelepiped Π is called the **fundamental parallelepiped** of \mathcal{K} .

Proof. In $\sigma_{\mathbf{v}+\mathcal{K}}(\mathbf{z}) = \sum_{\mathbf{m} \in (\mathbf{v}+\mathcal{K}) \cap \mathbb{Z}^d} \mathbf{z}^{\mathbf{m}}$, we list each integer point \mathbf{m} in $\mathbf{v}+\mathcal{K}$ as the monomial $\mathbf{z}^{\mathbf{m}}$. Such a lattice point can, by definition, be written as

$$\mathbf{m} = \mathbf{v} + \lambda_1 \mathbf{w}_1 + \lambda_2 \mathbf{w}_2 + \cdots + \lambda_d \mathbf{w}_d$$

for some numbers $\lambda_1, \lambda_2, \dots, \lambda_d \geq 0$. Because the \mathbf{w}_k 's form a basis of \mathbb{R}^d , this representation is unique. Let's write each of the λ_k 's in terms of their integer and fractional part: $\lambda_k = \lfloor \lambda_k \rfloor + \{\lambda_k\}$. So

$$\mathbf{m} = \mathbf{v} + (\{\lambda_1\} \mathbf{w}_1 + \{\lambda_2\} \mathbf{w}_2 + \cdots + \{\lambda_d\} \mathbf{w}_d) + \lfloor \lambda_1 \rfloor \mathbf{w}_1 + \lfloor \lambda_2 \rfloor \mathbf{w}_2 + \cdots + \lfloor \lambda_d \rfloor \mathbf{w}_d,$$

and we should note that, since $0 \leq \{\lambda_k\} < 1$, the vector

$$\mathbf{p} := \mathbf{v} + \{\lambda_1\} \mathbf{w}_1 + \{\lambda_2\} \mathbf{w}_2 + \cdots + \{\lambda_d\} \mathbf{w}_d$$

is in $\mathbf{v}+\Pi$. In fact, $\mathbf{p} \in \mathbb{Z}^d$, since \mathbf{m} and $\lfloor \lambda_k \rfloor \mathbf{w}_k$ are all integer vectors. Again the representation of \mathbf{p} in terms of the \mathbf{w}_k 's is unique. In summary, we have proved that any $\mathbf{m} \in \mathbf{v} + \mathcal{K} \cap \mathbb{Z}^d$ can be uniquely written as

$$\mathbf{m} = \mathbf{p} + k_1 \mathbf{w}_1 + k_2 \mathbf{w}_2 + \cdots + k_d \mathbf{w}_d \tag{3.1}$$

for some $\mathbf{p} \in (\mathbf{v} + \Pi) \cap \mathbb{Z}^d$ and some integers $k_1, k_2, \dots, k_d \geq 0$. On the other hand, let us write the rational function on the right-hand side of the theorem as a product of series:

$$\frac{\sigma_{\mathbf{v}+\Pi}(\mathbf{z})}{(1 - \mathbf{z}^{\mathbf{w}_1}) \cdots (1 - \mathbf{z}^{\mathbf{w}_d})} = \left(\sum_{\mathbf{p} \in (\mathbf{v}+\Pi) \cap \mathbb{Z}^d} \mathbf{z}^{\mathbf{p}} \right) \left(\sum_{k_1 \geq 0} \mathbf{z}^{k_1 \mathbf{w}_1} \right) \cdots \left(\sum_{k_d \geq 0} \mathbf{z}^{k_d \mathbf{w}_d} \right).$$

If we multiply everything out, a typical exponent will look exactly like (3.1). \square

Our proof contains a crucial geometric idea. Namely, we *tile* the cone $\mathbf{v} + \mathcal{K}$ with translates of $\mathbf{v} + \Pi$ by nonnegative integral combinations of the \mathbf{w}_k 's. It is this tiling that gives rise to the nice integer-point transform in Theorem 3.5. Computationally, we therefore favor cones over polytopes due to our ability to tile a simplicial cone with copies of the fundamental domain, as above. Another reason for favoring cones over polytopes appears in Brion's theorem in Chapter 9.

Theorem 3.5 shows that the real complexity of computing the integer-point transform $\sigma_{\mathbf{v}+\mathcal{K}}$ is embedded in the location of the lattice points in the parallelepiped $\mathbf{v} + \Pi$.

By mildly strengthening the hypothesis of Theorem 3.5, we obtain a slightly easier generating function, a result we shall need in Section 3.4 and Chapter 4.

Corollary 3.6. *Suppose*

$$\mathcal{K} := \{\lambda_1 \mathbf{w}_1 + \lambda_2 \mathbf{w}_2 + \cdots + \lambda_d \mathbf{w}_d : \lambda_1, \lambda_2, \dots, \lambda_d \geq 0\}$$

is a simplicial d -cone, where $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_d \in \mathbb{Z}^d$, and $\mathbf{v} \in \mathbb{R}^d$, such that the boundary of $\mathbf{v} + \mathcal{K}$ contains no integer point. Then

$$\sigma_{\mathbf{v}+\mathcal{K}}(\mathbf{z}) = \frac{\sigma_{\mathbf{v}+\Pi}(\mathbf{z})}{(1 - \mathbf{z}^{\mathbf{w}_1})(1 - \mathbf{z}^{\mathbf{w}_2}) \cdots (1 - \mathbf{z}^{\mathbf{w}_d})},$$

where Π is the open parallelepiped

$$\Pi := \{\lambda_1 \mathbf{w}_1 + \lambda_2 \mathbf{w}_2 + \cdots + \lambda_d \mathbf{w}_d : 0 < \lambda_1, \lambda_2, \dots, \lambda_d < 1\}.$$

Proof. The proof of Theorem 3.5 goes through almost verbatim, except that $\mathbf{v} + \Pi$ now has no boundary lattice points, so that there is no harm in choosing Π to be open. \square

Since a general pointed cone can always be triangulated into simplicial cones, the integer-point transforms add up in an inclusion–exclusion manner (note that the intersection of simplicial cones in a triangulation is again a simplicial cone, by Exercise 3.2). Hence we have the following corollary.

Corollary 3.7. *Given any pointed cone*

$$\mathcal{K} = \{\mathbf{v} + \lambda_1 \mathbf{w}_1 + \lambda_2 \mathbf{w}_2 + \cdots + \lambda_m \mathbf{w}_m : \lambda_1, \lambda_2, \dots, \lambda_m \geq 0\}$$

with $\mathbf{v} \in \mathbb{R}^d$, $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m \in \mathbb{Z}^d$, the integer-point transform $\sigma_{\mathcal{K}}(\mathbf{z})$ evaluates to a rational function in the coordinates of \mathbf{z} . \square

Philosophizing some more, one can show that the original infinite sum $\sigma_{\mathcal{K}}(\mathbf{z})$ converges only for \mathbf{z} in a subset of \mathbb{C}^d , whereas the rational function that represents $\sigma_{\mathcal{K}}$ gives us its meromorphic continuation. Later, in Chapters 4 and 9, we make use of this continuation.

3.3 Expanding and Counting Using Ehrhart's Original Approach

Here is *the* fundamental theorem concerning the lattice-point count in an integral convex polytope.

Theorem 3.8 (Ehrhart's theorem). *If \mathcal{P} is an integral convex d -polytope, then $L_{\mathcal{P}}(t)$ is a polynomial in t of degree d .*

This result is due to Eugène Ehrhart, in whose honor we call $L_{\mathcal{P}}$ the **Ehrhart polynomial** of \mathcal{P} . Naturally, there is an extension of Ehrhart's theorem to rational polytopes, which we will discuss in Section 3.7.

Our proof of Ehrhart's theorem uses generating functions of the form $\sum_{t \geq 0} f(t) z^t$, similar in spirit to the ones discussed in the beginning of Chapter 1. If f is a polynomial, this power series takes on a special form, which we invite the reader to prove (Exercise 3.8):

Lemma 3.9. *If*

$$\sum_{t \geq 0} f(t) z^t = \frac{g(z)}{(1-z)^{d+1}},$$

then f is a polynomial of degree d if and only if g is a polynomial of degree at most d and $g(1) \neq 0$. \square

The reason we introduced generating functions of the form $\sigma_S(\mathbf{z}) = \sum_{\mathbf{m} \in S \cap \mathbb{Z}^d} \mathbf{z}^{\mathbf{m}}$ in Section 3.2 is that they are extremely handy for lattice-point problems. The connection to lattice points is evident, since we are summing over them. If we're interested in the lattice-point *count*, we simply evaluate σ_S at $\mathbf{z} = (1, 1, \dots, 1)$:

$$\sigma_S(1, 1, \dots, 1) = \sum_{\mathbf{m} \in S \cap \mathbb{Z}^d} \mathbf{1}^{\mathbf{m}} = \sum_{\mathbf{m} \in S \cap \mathbb{Z}^d} 1 = \#(S \cap \mathbb{Z}^d).$$

(Here we denote by $\mathbf{1}$ a vector all of whose components are 1.) Naturally, we should make this evaluation only if S is bounded; Theorem 3.5 already tells us that it's no fun evaluating $\sigma_{\mathcal{K}}(\mathbf{1})$ if \mathcal{K} is a cone.

But the magic of the generating function σ_S does not stop there. To literally take it to the next level, we come over a convex polytope \mathcal{P} . If $\mathcal{P} \subset \mathbb{R}^d$ has the vertices $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n \in \mathbb{Z}^d$, recall that we lift these vertices into \mathbb{R}^{d+1} , by adding a 1 as their last coordinate. So let

$$\mathbf{w}_1 = (\mathbf{v}_1, 1), \mathbf{w}_2 = (\mathbf{v}_2, 1), \dots, \mathbf{w}_n = (\mathbf{v}_n, 1).$$

Then

$$\text{cone}(\mathcal{P}) = \{\lambda_1 \mathbf{w}_1 + \lambda_2 \mathbf{w}_2 + \dots + \lambda_n \mathbf{w}_n : \lambda_1, \lambda_2, \dots, \lambda_n \geq 0\} \subset \mathbb{R}^{d+1}.$$

Recall that we can recover our original polytope \mathcal{P} by cutting $\text{cone}(\mathcal{P})$ with the hyperplane $x_{d+1} = 1$. We can recover more than just the original polytope in $\text{cone}(\mathcal{P})$: By cutting the cone with the hyperplane $x_{d+1} = 2$, we obtain a copy of \mathcal{P} dilated by a factor of 2. (The reader should meditate on why this cut is a 2-dilate of \mathcal{P} .) More generally, we can cut the cone with the hyperplane $x_{d+1} = t$ and obtain $t\mathcal{P}$, as suggested by Figure 3.4.

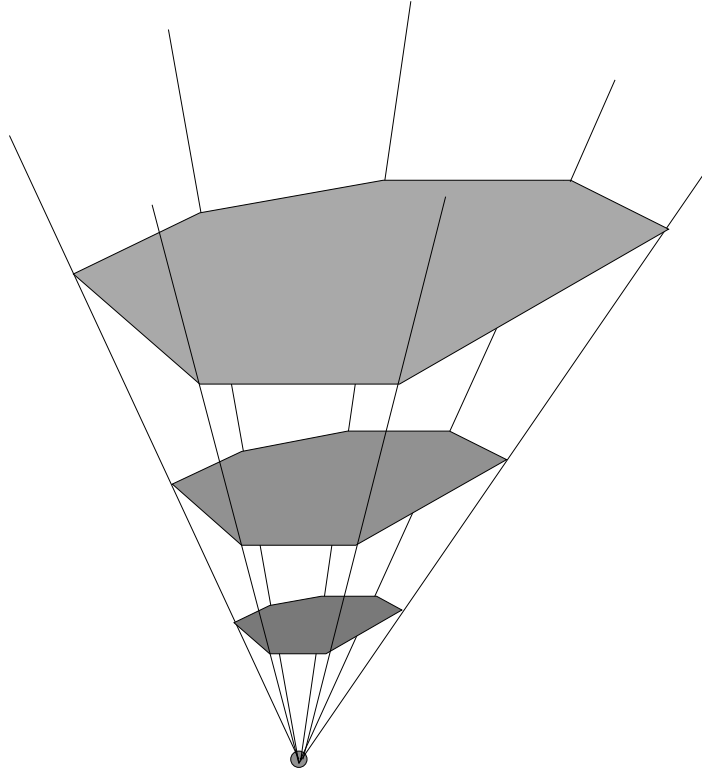


Fig. 3.4. Recovering dilates of \mathcal{P} in $\text{cone}(\mathcal{P})$.

Now let's form the integer-point transform $\sigma_{\text{cone}(\mathcal{P})}$ of $\text{cone}(\mathcal{P})$. By what we just said, we should look at different powers of z_{d+1} : there is one term (namely, 1), with z_{d+1}^0 , corresponding to the origin; the terms with z_{d+1}^1 correspond to lattice points in \mathcal{P} (listed as monomials in z_1, z_2, \dots, z_d), the terms with z_{d+1}^2 correspond to points in $2\mathcal{P}$, etc. In short,

$$\begin{aligned}
\sigma_{\text{cone}(\mathcal{P})}(z_1, z_2, \dots, z_{d+1}) &= 1 + \sigma_{\mathcal{P}}(z_1, \dots, z_d) z_{d+1} + \sigma_{2\mathcal{P}}(z_1, \dots, z_d) z_{d+1}^2 + \dots \\
&= 1 + \sum_{t \geq 1} \sigma_{t\mathcal{P}}(z_1, \dots, z_d) z_{d+1}^t.
\end{aligned}$$

Specializing further for enumeration purposes, we recall that $\sigma_{\mathcal{P}}(1, 1, \dots, 1) = \#(\mathcal{P} \cap \mathbb{Z}^d)$, and so

$$\begin{aligned}
\sigma_{\text{cone}(\mathcal{P})}(1, 1, \dots, 1, z_{d+1}) &= 1 + \sum_{t \geq 1} \sigma_{t\mathcal{P}}(1, 1, \dots, 1) z_{d+1}^t \\
&= 1 + \sum_{t \geq 1} \#(t\mathcal{P} \cap \mathbb{Z}^d) z_{d+1}^t.
\end{aligned}$$

But by definition, the enumerators on the right-hand side are just evaluations of Ehrhart's counting function, that is, $\sigma_{\text{cone}(\mathcal{P})}(1, 1, \dots, 1, z_{d+1})$ is nothing but the Ehrhart series of \mathcal{P} :

Lemma 3.10. $\sigma_{\text{cone}(\mathcal{P})}(1, 1, \dots, 1, z) = 1 + \sum_{t \geq 1} L_{\mathcal{P}}(t) z^t = \text{Ehr}_{\mathcal{P}}(z)$. \square

With all this machinery at hand, we can prove Ehrhart's theorem.

Proof of Theorem 3.8. It suffices to prove the theorem for *simplices*, because we can triangulate any integral polytope into integral simplices, using no new vertices. Note that these simplices will intersect in lower-dimensional integral simplices.

By Lemma 3.9, it suffices to prove that for an integral d -simplex Δ ,

$$\text{Ehr}_{\Delta}(z) = 1 + \sum_{t \geq 1} L_{\Delta}(t) z^t = \frac{g(z)}{(1-z)^{d+1}}$$

for some polynomial g of degree at most d with $g(1) \neq 0$. In Lemma 3.10 we showed that the Ehrhart series of Δ equals $\sigma_{\text{cone}(\Delta)}(1, 1, \dots, 1, z)$, so let's study the integer-point transform attached to $\text{cone}(\Delta)$.

The simplex Δ has $d+1$ vertices $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{d+1}$, and so $\text{cone}(\Delta) \subset \mathbb{R}^{d+1}$ is simplicial, with apex the origin and generators

$$\mathbf{w}_1 = (\mathbf{v}_1, 1), \mathbf{w}_2 = (\mathbf{v}_2, 1), \dots, \mathbf{w}_{d+1} = (\mathbf{v}_{d+1}, 1) \in \mathbb{Z}^{d+1}.$$

Now we use Theorem 3.5:

$$\sigma_{\text{cone}(\Delta)}(z_1, z_2, \dots, z_{d+1}) = \frac{\sigma_{\Pi}(z_1, z_2, \dots, z_{d+1})}{(1 - \mathbf{z}^{\mathbf{w}_1})(1 - \mathbf{z}^{\mathbf{w}_2}) \dots (1 - \mathbf{z}^{\mathbf{w}_{d+1}})},$$

where $\Pi = \{\lambda_1 \mathbf{w}_1 + \lambda_2 \mathbf{w}_2 + \dots + \lambda_{d+1} \mathbf{w}_{d+1} : 0 \leq \lambda_1, \lambda_2, \dots, \lambda_{d+1} < 1\}$. This parallelepiped is bounded, so the attached generating function σ_{Π} is a Laurent polynomial in z_1, z_2, \dots, z_{d+1} .

We claim that the z_{d+1} -degree of σ_Π is at most d . In fact, since the x_{d+1} -coordinate of each \mathbf{w}_k is 1, the x_{d+1} -coordinate of a point in Π is $\lambda_1 + \lambda_2 + \cdots + \lambda_{d+1}$ for some $0 \leq \lambda_1, \lambda_2, \dots, \lambda_{d+1} < 1$. But then $\lambda_1 + \lambda_2 + \cdots + \lambda_{d+1} < d+1$, so if this sum is an integer it is at most d , which implies that the z_{d+1} -degree of $\sigma_\Pi(z_1, z_2, \dots, z_{d+1})$ is at most d . Consequently, $\sigma_\Pi(1, 1, \dots, 1, z_{d+1})$ is a polynomial in z_{d+1} of degree at most d . The evaluation $\sigma_\Pi(1, 1, 1, \dots, 1)$ of this polynomial at $z_{d+1} = 1$ is not zero, because $\sigma_\Pi(1, 1, 1, \dots, 1) = \#(\Pi \cap \mathbb{Z}^{d+1})$ and the origin is a lattice point in Π .

Finally, if we specialize $\mathbf{z}^{\mathbf{w}_k}$ to $z_1 = z_2 = \cdots = z_d = 1$, we obtain z_{d+1}^1 , so that

$$\sigma_{\text{cone}(\Delta)}(1, 1, \dots, 1, z_{d+1}) = \frac{\sigma_\Pi(1, 1, \dots, 1, z_{d+1})}{(1 - z_{d+1})^{d+1}}.$$

The left-hand side is $\text{Ehr}_\Delta(z_{d+1}) = 1 + \sum_{t \geq 1} L_\Delta(t) z_{d+1}^t$ by Lemma 3.10. \square

3.4 The Ehrhart Series of an Integral Polytope

We can actually take our proof of Ehrhart's theorem one step further by studying the polynomial $\sigma_\Pi(1, 1, \dots, 1, z_{d+1})$. As mentioned above, the coefficient of z_{d+1}^k simply counts the integer points in the parallelepiped Π cut with the hyperplane $x_{d+1} = k$. Let us record this.

Corollary 3.11. *Suppose Δ is an integral d -simplex with vertices $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{d+1}$, and let $\mathbf{w}_j = (\mathbf{v}_j, 1)$. Then*

$$\text{Ehr}_\Delta(z) = 1 + \sum_{t \geq 1} L_\Delta(t) z^t = \frac{h_d z^d + h_{d-1} z^{d-1} + \cdots + h_1 z + h_0}{(1 - z)^{d+1}},$$

where h_k equals the number of integer points in

$$\{\lambda_1 \mathbf{w}_1 + \lambda_2 \mathbf{w}_2 + \cdots + \lambda_{d+1} \mathbf{w}_{d+1} : 0 \leq \lambda_1, \lambda_2, \dots, \lambda_{d+1} < 1\}$$

with last variable equal to k . \square

This result can actually be used to compute Ehr_Δ , and therefore the Ehrhart polynomial, of an integral simplex Δ in low dimensions very quickly (a fact that the reader may discover in some of the exercises). We remark, however, that things are not as simple for arbitrary integral polytopes. Not only is triangulation a nontrivial task in general, but one would also have to deal with overcounting where simplices of a triangulation meet.

Corollary 3.11 implies that the numerator of the Ehrhart series of an integral simplex has nonnegative coefficients, since those coefficients count something. Although the latter cannot be said of the coefficients of the Ehrhart series of a general integral polytope, the nonnegativity property magically survives.

Theorem 3.12 (Stanley’s nonnegativity theorem). *Suppose \mathcal{P} is an integral convex d -polytope with Ehrhart series*

$$\text{Ehr}_{\mathcal{P}}(z) = \frac{h_d z^d + h_{d-1} z^{d-1} + \cdots + h_0}{(1-z)^{d+1}}.$$

Then $h_0, h_1, \dots, h_d \geq 0$.

Proof. Triangulate $\text{cone}(\mathcal{P}) \subset \mathbb{R}^{d+1}$ into the simplicial cones $\mathcal{K}_1, \mathcal{K}_2, \dots, \mathcal{K}_m$. Now Exercise 3.14 ensures that there exists a vector $\mathbf{v} \in \mathbb{R}^{d+1}$ such that

$$\text{cone}(\mathcal{P}) \cap \mathbb{Z}^d = (\mathbf{v} + \text{cone}(\mathcal{P})) \cap \mathbb{Z}^d$$

(that is, we neither lose nor gain any lattice points when shifting $\text{cone}(\mathcal{P})$ by \mathbf{v}) and neither the facets of $\mathbf{v} + \text{cone}(\mathcal{P})$ nor the triangulation hyperplanes contain any lattice points. This implies that every lattice point in $\mathbf{v} + \text{cone}(\mathcal{P})$ belongs to exactly one simplicial cone $\mathbf{v} + \mathcal{K}_j$:

$$\text{cone}(\mathcal{P}) \cap \mathbb{Z}^d = (\mathbf{v} + \text{cone}(\mathcal{P})) \cap \mathbb{Z}^d = \bigcup_{j=1}^m ((\mathbf{v} + \mathcal{K}_j) \cap \mathbb{Z}^d), \quad (3.2)$$

and this union is a *disjoint* union. If we translate the last identity into generating-function language, it becomes

$$\sigma_{\text{cone}(\mathcal{P})}(z_1, z_2, \dots, z_{d+1}) = \sum_{j=1}^m \sigma_{\mathbf{v} + \mathcal{K}_j}(z_1, z_2, \dots, z_{d+1}).$$

But now we recall that the Ehrhart series of \mathcal{P} is just a special evaluation of $\sigma_{\text{cone}(\mathcal{P})}$ (Lemma 3.10):

$$\text{Ehr}_{\mathcal{P}}(z) = \sigma_{\text{cone}(\mathcal{P})}(1, 1, \dots, 1, z) = \sum_{j=1}^m \sigma_{\mathbf{v} + \mathcal{K}_j}(1, 1, \dots, 1, z). \quad (3.3)$$

It suffices to show that the rational generating functions $\sigma_{\mathbf{v} + \mathcal{K}_j}(1, 1, \dots, 1, z)$ for the simplicial cones $\mathbf{v} + \mathcal{K}_j$ have nonnegative numerator. But this fact follows from evaluating the rational function in Corollary 3.6 at $(1, 1, \dots, 1, z)$. \square

This proof shows a little more: Since the origin is in precisely *one* simplicial cone on the right-hand side of (3.2), we get on the right-hand side of (3.3) precisely *one* term that contributes $1/(1-z)^{d+1}$ to $\text{Ehr}_{\mathcal{P}}$; all other terms contribute to higher powers of the numerator polynomial of $\text{Ehr}_{\mathcal{P}}$. That is, the constant term h_0 equals 1. The reader might feel that we are chasing our tail at this point, since we assumed from the very beginning that the constant term of the infinite series $\text{Ehr}_{\mathcal{P}}$ is 1, and hence h_0 has to be 1, as a quick look at the expansion of the rational function representing $\text{Ehr}_{\mathcal{P}}$ shows. The above argument shows merely that this convention is geometrically sound. Let us record this:

Lemma 3.13. *Suppose \mathcal{P} is an integral convex d -polytope with Ehrhart series*

$$\text{Ehr}_{\mathcal{P}}(z) = \frac{h_d z^d + h_{d-1} z^{d-1} + \cdots + h_0}{(1-z)^{d+1}}.$$

Then $h_0 = 1$. □

For a general integral polytope \mathcal{P} , the reader has probably already discovered how to extract the Ehrhart polynomial of \mathcal{P} from its Ehrhart series:

Lemma 3.14. *Suppose \mathcal{P} is an integral convex d -polytope with Ehrhart series*

$$\text{Ehr}_{\mathcal{P}}(z) = 1 + \sum_{t \geq 1} L_{\mathcal{P}}(t) z^t = \frac{h_d z^d + h_{d-1} z^{d-1} + \cdots + h_1 z + 1}{(1-z)^{d+1}}.$$

Then

$$L_{\mathcal{P}}(t) = \binom{t+d}{d} + h_1 \binom{t+d-1}{d} + \cdots + h_{d-1} \binom{t+1}{d} + h_d \binom{t}{d}.$$

Proof. Expand into a binomial series:

$$\begin{aligned} \text{Ehr}_{\mathcal{P}}(z) &= \frac{h_d z^d + h_{d-1} z^{d-1} + \cdots + h_1 z + 1}{(1-z)^{d+1}} \\ &= (h_d z^d + h_{d-1} z^{d-1} + \cdots + h_1 z + 1) \sum_{t \geq 0} \binom{t+d}{d} z^t \\ &= h_d \sum_{t \geq 0} \binom{t+d}{d} z^{t+d} + h_{d-1} \sum_{t \geq 0} \binom{t+d}{d} z^{t+d-1} + \cdots \\ &\quad + h_1 \sum_{t \geq 0} \binom{t+d}{d} z^{t+1} + \sum_{t \geq 0} \binom{t+d}{d} z^t \\ &= h_d \sum_{t \geq d} \binom{t}{d} z^t + h_{d-1} \sum_{t \geq d-1} \binom{t+1}{d} z^t + \cdots \\ &\quad + h_1 \sum_{t \geq 1} \binom{t+d-1}{d} z^t + \sum_{t \geq 0} \binom{t+d}{d} z^t. \end{aligned}$$

In all infinite sums on the right-hand side, we can start the index t with 0 without changing the sums, by the definition (2.1) of the binomial coefficient. Hence

$$\begin{aligned} \text{Ehr}_{\mathcal{P}}(z) &= \sum_{t \geq 0} \left(h_d \binom{t}{d} + h_{d-1} \binom{t+1}{d} + \cdots + h_1 \binom{t+d-1}{d} + \binom{t+d}{d} \right) z^t. \end{aligned}$$

□

The representation of the polynomial $L_{\mathcal{P}}(t)$ in terms of the coefficients of $\text{Ehr}_{\mathcal{P}}$ can be interpreted as the Ehrhart polynomial expressed in the basis $\binom{t}{d}, \binom{t+1}{d}, \dots, \binom{t+d}{d}$ (see Exercise 3.9). This representation is very useful, as the following results show.

Corollary 3.15. *If \mathcal{P} is an integral convex d -polytope, then the constant term of the Ehrhart polynomial $L_{\mathcal{P}}$ is 1.*

Proof. Use the expansion of Lemma 3.14. The constant term is

$$L_{\mathcal{P}}(0) = \binom{d}{d} + h_1 \binom{d-1}{d} + \cdots + h_{d-1} \binom{1}{d} + h_d \binom{0}{d} = \binom{d}{d} = 1. \quad \square$$

This proof is exciting, because it marks the first instance where we extend the domain of an Ehrhart polynomial beyond the positive integers, for which the lattice-point enumerator was initially defined. More precisely, Ehrhart's theorem (Theorem 3.8) implies that the counting function

$$L_{\mathcal{P}}(t) = \#(t\mathcal{P} \cap \mathbb{Z}^d),$$

originally defined for positive integers t , can be extended to all real or even complex arguments t (as a polynomial). A natural question arises: are there nice *interpretations* of $L_{\mathcal{P}}(t)$ for arguments t that are not positive integers? Corollary 3.15 gives such an interpretation for $t = 0$. In Chapter 4, we will give interpretations of $L_{\mathcal{P}}(t)$ for *negative* integers t .

Corollary 3.16. *Suppose \mathcal{P} is an integral convex d -polytope with Ehrhart series*

$$\text{Ehr}_{\mathcal{P}}(z) = \frac{h_d z^d + h_{d-1} z^{d-1} + \cdots + h_1 z + 1}{(1-z)^{d+1}}.$$

Then $h_1 = L_{\mathcal{P}}(1) - d - 1 = \#(\mathcal{P} \cap \mathbb{Z}^d) - d - 1$.

Proof. Use the expansion of Lemma 3.14 with $t = 1$:

$$L_{\mathcal{P}}(1) = \binom{d+1}{d} + h_1 \binom{d}{d} + \cdots + h_{d-1} \binom{2}{d} + h_d \binom{1}{d} = d + 1 + h_1. \quad \square$$

The proof of Corollary 3.16 suggests that there are also formulas for h_2, h_3, \dots in terms of the evaluations $L_{\mathcal{P}}(1), L_{\mathcal{P}}(2), \dots$, and we invite the reader to find them (Exercise 3.10).

A final corollary to Theorem 3.12 and Lemma 3.14 states how large the denominators of the Ehrhart coefficients can be:

Corollary 3.17. *Suppose \mathcal{P} is an integral polytope with Ehrhart polynomial $L_{\mathcal{P}}(t) = c_d t^d + c_{d-1} t^{d-1} + \cdots + c_1 t + 1$. Then all coefficients satisfy $d! c_k \in \mathbb{Z}$.*

Proof. By Theorem 3.12 and Lemma 3.14,

$$L_{\mathcal{P}}(t) = \binom{t+d}{d} + h_1 \binom{t+d-1}{d} + \cdots + h_{d-1} \binom{t+1}{d} + h_d \binom{t}{d},$$

where the h_k 's are integers. Hence multiplying out this expression yields a polynomial in t whose coefficients can be written as rational numbers with denominator $d!$. \square

We finish this section with a general result that gives relations between negative integer roots of a polynomial and its generating function. This theorem will become handy in Chapter 4, in which we find an interpretation for the evaluation of an Ehrhart polynomial at negative integers.

Theorem 3.18. *Suppose p is a degree- d polynomial with the rational generating function*

$$\sum_{t \geq 0} p(t) z^t = \frac{h_d z^d + h_{d-1} z^{d-1} + \cdots + h_1 z + h_0}{(1-z)^{d+1}}.$$

Then $h_d = h_{d-1} = \cdots = h_{k+1} = 0$ and $h_k \neq 0$ if and only if $p(-1) = p(-2) = \cdots = p(-(d-k)) = 0$ and $p(-(d-k+1)) \neq 0$.

Proof. Suppose $h_d = h_{d-1} = \cdots = h_{k+1} = 0$ and $h_k \neq 0$. Then the proof of Lemma 3.14 gives

$$p(t) = h_0 \binom{t+d}{d} + \cdots + h_{k-1} \binom{t+d-k+1}{d} + h_k \binom{t+d-k}{d}.$$

All the binomial coefficients are zero for $t = -1, -2, \dots, -d+k$, so those are roots of p . On the other hand, all binomial coefficients but the last one are zero for $t = -d+k-1$, and since $h_k \neq 0$, $-d+k-1$ is not a root of p .

Conversely, suppose $p(-1) = p(-2) = \cdots = p(-(d-k)) = 0$ and $p(-(d-k+1)) \neq 0$. The first root -1 of p gives

$$0 = p(-1) = h_0 \binom{d-1}{d} + h_1 \binom{d-2}{d} + \cdots + h_{d-1} \binom{0}{d} + h_d \binom{-1}{d} = h_d \binom{-1}{d},$$

so we must have $h_d = 0$. The next root -2 forces $h_{d-1} = 0$, and so on, up to the root $-d+k$, which forces $h_{k+1} = 0$. It remains to show that $h_k \neq 0$. But if h_k were zero then, by a similar line of reasoning as in the first part of the proof, $p(-d+k-1) = 0$, a contradiction. \square

3.5 From the Discrete to the Continuous Volume of a Polytope

Given a geometric object $S \subset \mathbb{R}^d$, its **volume**, defined by the integral $\text{vol } S := \int_S dx$, is one of the fundamental data of S . By the definition of the integral, say

in the Riemannian sense, we can think of computing $\text{vol } S$ by approximating S with d -dimensional boxes that get smaller and smaller. To be precise, if we take the boxes with side length $1/t$ then they each have volume $1/t^d$. We might further think of the boxes as filling out the space between grid points in the lattice $(\frac{1}{t}\mathbb{Z})^d$. This means that volume computation can be approximated by counting boxes, or equivalently, lattice points in $(\frac{1}{t}\mathbb{Z})^d$:

$$\text{vol } S = \lim_{t \rightarrow \infty} \frac{1}{t^d} \cdot \# \left(S \cap \left(\frac{1}{t}\mathbb{Z} \right)^d \right).$$

It is a short step to counting integer points in dilates of S , because

$$\# \left(S \cap \left(\frac{1}{t}\mathbb{Z} \right)^d \right) = \# (tS \cap \mathbb{Z}^d).$$

Let us summarize:

Lemma 3.19. *Suppose $S \subset \mathbb{R}^d$ is d -dimensional. Then*

$$\text{vol } S = \lim_{t \rightarrow \infty} \frac{1}{t^d} \cdot \# (tS \cap \mathbb{Z}^d). \quad \square$$

We emphasize here that S is d -dimensional, because otherwise (since S could be lower-dimensional although living in d -space), by our current definition $\text{vol } S = 0$. (We will extend our volume definition in Chapter 5 to give nonzero *relative* volume to objects that are not full-dimensional.)

Part of the magic of Ehrhart's theorem lies in the fact that for an integral d -polytope \mathcal{P} , we do not have to take a limit to compute $\text{vol } \mathcal{P}$; we need to compute “only” the $d+1$ coefficients of a polynomial.

Corollary 3.20. *Suppose $\mathcal{P} \subset \mathbb{R}^d$ is a integral convex d -polytope with Ehrhart polynomial $c_d t^d + c_{d-1} t^{d-1} + \cdots + c_1 t + 1$. Then $c_d = \text{vol } \mathcal{P}$.*

Proof. By Lemma 3.19,

$$\text{vol } \mathcal{P} = \lim_{t \rightarrow \infty} \frac{c_d t^d + c_{d-1} t^{d-1} + \cdots + c_1 t + 1}{t^d} = c_d. \quad \square$$

On the one hand, this should not come as a surprise, because counting integer points in some object should grow roughly like the volume of the object as we make it bigger and bigger. On the other hand, the fact that we can compute the volume as one term of a polynomial should be very surprising: the polynomial is a counting function and as such is something *discrete*, yet by computing it (and its leading term), we derive some *continuous* data. Even more, we can—at least theoretically—compute this continuous datum (the volume) of the object by calculating a few values of the polynomial and then interpolating; this can be described as a completely discrete operation!

We finish this section by showing how to retrieve the continuous volume of an integer polytope from its Ehrhart series.

Corollary 3.21. *Suppose $\mathcal{P} \subset \mathbb{R}^d$ is an integral convex d -polytope, and*

$$\text{Ehr}_{\mathcal{P}}(z) = \frac{h_d z^d + h_{d-1} z^{d-1} + \cdots + h_1 z + 1}{(1-z)^{d+1}}.$$

Then $\text{vol } \mathcal{P} = \frac{1}{d!} (h_d + h_{d-1} + \cdots + h_1 + 1)$.

Proof. Use the expansion of Lemma 3.14. The leading coefficient is

$$\frac{1}{d!} (h_d + h_{d-1} + \cdots + h_1 + 1). \quad \square$$

3.6 Interpolation

We now use the polynomial behavior of the discrete volume $L_{\mathcal{P}}$ of an integral polytope \mathcal{P} to compute the continuous volume $\text{vol } \mathcal{P}$ and the discrete volume $L_{\mathcal{P}}$ from finite data.

Two points uniquely determine a line. There exists a unique quadratic passing through any three given points. In general, a degree- d polynomial p is determined by $d+1$ points $(x, p(x)) \in \mathbb{R}^2$. Namely, evaluating $p(x) = c_d x^d + c_{d-1} x^{d-1} + \cdots + c_0$ at distinct inputs x_1, x_2, \dots, x_{d+1} gives

$$\begin{pmatrix} p(x_1) \\ p(x_2) \\ \vdots \\ p(x_{d+1}) \end{pmatrix} = \mathbf{V} \begin{pmatrix} c_d \\ c_{d-1} \\ \vdots \\ c_0 \end{pmatrix}, \quad (3.4)$$

where

$$\mathbf{V} = \begin{pmatrix} x_1^d & x_1^{d-1} & \cdots & x_1 & 1 \\ x_2^d & x_2^{d-1} & \cdots & x_2 & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ x_{d+1}^d & x_{d+1}^{d-1} & \cdots & x_{d+1} & 1 \end{pmatrix},$$

so that

$$\begin{pmatrix} c_d \\ c_{d-1} \\ \vdots \\ c_0 \end{pmatrix} = \mathbf{V}^{-1} \begin{pmatrix} p(x_1) \\ p(x_2) \\ \vdots \\ p(x_{d+1}) \end{pmatrix}. \quad (3.5)$$

(Exercise 3.16 makes sure that \mathbf{V} is invertible.) The identity (3.5) gives the famous *Lagrange interpolation formula*.

This gives us an efficient way to compute $L_{\mathcal{P}}$, at least when $\dim \mathcal{P}$ is not too large. The continuous volume of \mathcal{P} will follow instantly, since it is the leading coefficient c_d of $L_{\mathcal{P}}$. In the case of an Ehrhart polynomial $L_{\mathcal{P}}$, we know that $L_{\mathcal{P}}(0) = 1$, so that (3.4) simplifies to

$$\begin{pmatrix} L_{\mathcal{P}}(x_1) - 1 \\ L_{\mathcal{P}}(x_2) - 1 \\ \vdots \\ L_{\mathcal{P}}(x_d) - 1 \end{pmatrix} = \begin{pmatrix} x_1^d & x_1^{d-1} & \cdots & x_1 \\ x_2^d & x_2^{d-1} & \cdots & x_2 \\ \vdots & \vdots & & \vdots \\ x_d^d & x_d^{d-1} & \cdots & x_d \end{pmatrix} \begin{pmatrix} c_d \\ c_{d-1} \\ \vdots \\ c_1 \end{pmatrix}.$$

Example 3.22 (Reeve's tetrahedron). Let \mathcal{T}_h be the tetrahedron with vertices $(0, 0, 0)$, $(1, 0, 0)$, $(0, 1, 0)$, and $(1, 1, h)$, where h is a positive integer (see Figure 3.5).

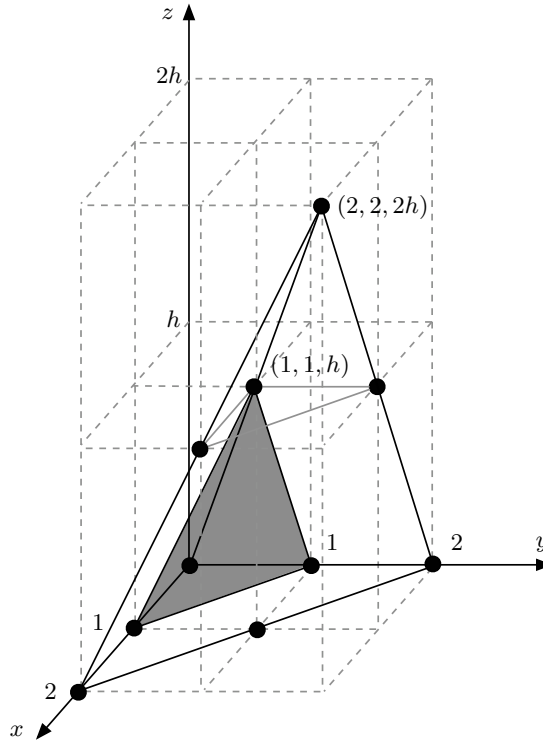


Fig. 3.5. Reeve's tetrahedron \mathcal{T}_h (and $2\mathcal{T}_h$).

To interpolate the Ehrhart polynomial $L_{\mathcal{T}_h}(t)$ from its values at various points, we use Figure 3.5 to deduce the following:

$$\begin{aligned} 4 &= L_{\mathcal{T}_h}(1) = \text{vol}(\mathcal{T}_h) + c_2 + c_1 + 1, \\ h + 9 &= L_{\mathcal{T}_h}(2) = \text{vol}(\mathcal{T}_h) \cdot 2^3 + c_2 \cdot 2^2 + c_1 \cdot 2 + 1. \end{aligned}$$

Using the volume formula for a pyramid, we know that

$$\text{vol}(\mathcal{T}_h) = \frac{1}{3}(\text{base area})(\text{height}) = \frac{h}{6}.$$

Thus $h + 1 = h + 2c_2 - 1$, which gives us $c_2 = 1$ and $c_1 = 2 - \frac{h}{6}$. Therefore

$$L_{\mathcal{T}_h}(t) = \frac{h}{6}t^3 + t^2 + \left(2 - \frac{h}{6}\right)t + 1. \quad \square$$

3.7 Rational Polytopes and Ehrhart Quasipolynomials

We do not have to change much to study lattice-point enumeration for *rational* polytopes, and most of this section will consist of exercises for the reader. The structural result paralleling Theorem 3.8 is as follows.

Theorem 3.23 (Ehrhart’s theorem for rational polytopes). *If \mathcal{P} is a rational convex d -polytope, then $L_{\mathcal{P}}(t)$ is a quasipolynomial in t of degree d . Its period divides the least common multiple of the denominators of the coordinates of the vertices of \mathcal{P} .*

We will call the least common multiple of the denominators of the coordinates of the vertices of \mathcal{P} the **denominator** of \mathcal{P} . Theorem 3.23, also due to Ehrhart, extends Theorem 3.8, because the denominator of an integral polytope \mathcal{P} is one. Exercises 3.21 and 3.22 show that the word “divides” in Theorem 3.23 is far from being replaceable by “equals.”

We start the path toward a proof of Theorem 3.23 by stating the analogue of Lemma 3.9 for quasipolynomials (see Exercise 3.19):

Lemma 3.24. *If*

$$\sum_{t \geq 0} f(t) z^t = \frac{g(z)}{h(z)},$$

then f is a quasipolynomial of degree d with period dividing p if and only if g and h are polynomials such that $\deg(g) < \deg(h)$, all roots of h are p^{th} roots of unity of multiplicity at most $d + 1$, and there is a root of multiplicity equal to $d + 1$ (all of this assuming that g/h has been reduced to lowest terms). \square

Our goal is now evident: we will prove that if \mathcal{P} is a rational convex d -polytope with denominator p , then

$$\text{Ehr}_{\mathcal{P}}(z) = 1 + \sum_{t \geq 1} L_{\mathcal{P}}(t) z^t = \frac{g(z)}{(1 - z^p)^{d+1}},$$

for some polynomial g of degree less than $p(d + 1)$. As in Section 3.3, we will have to prove this only for the case of a rational *simplex*. So suppose the d -simplex Δ has vertices $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{d+1} \in \mathbb{Q}^d$, and the denominator of Δ is p . Again we will cone over Δ : let

$$\mathbf{w}_1 = (\mathbf{v}_1, 1), \mathbf{w}_2 = (\mathbf{v}_2, 1), \dots, \mathbf{w}_{d+1} = (\mathbf{v}_{d+1}, 1);$$

then

$$\text{cone}(\Delta) = \{\lambda_1 \mathbf{w}_1 + \lambda_2 \mathbf{w}_2 + \dots + \lambda_{d+1} \mathbf{w}_{d+1} : \lambda_1, \lambda_2, \dots, \lambda_{d+1} \geq 0\} \subset \mathbb{R}^{d+1}.$$

To be able to use Theorem 3.5, we first have to ensure that we have a description of $\text{cone}(\Delta)$ with integral generators. But since the denominator of Δ is p , we can replace each generator \mathbf{w}_k by $p\mathbf{w}_k \in \mathbb{Z}^{d+1}$, and we're ready to apply Theorem 3.5. From this point, the proof of Theorem 3.23 proceeds exactly like that of Theorem 3.8, and we invite the reader to finish it up (Exercise 3.20).

Although the proofs of Theorem 3.23 and Theorem 3.8 are almost identical, the arithmetic structure of Ehrhart quasipolynomials is much more subtle and less well known than that of Ehrhart polynomials.

3.8 Reflections on the Coin-Exchange Problem and the Gallery of Chapter 2

At this point, we encourage the reader to look back at the first two chapters in light of the basic Ehrhart-theory results. Popoviciu's theorem (Theorem 1.5) and its higher-dimensional analogue give a special set of Ehrhart quasipolynomials. On the other hand, in Chapter 2 we encountered many integral polytopes. Ehrhart's theorem (Theorem 3.8) explains why their lattice-point enumeration functions were all polynomials.

Notes

1. Triangulations of polytopes and manifolds are an active source of research with many interesting open problems; see, e.g., [68].

2. Eugène Ehrhart laid the foundation for the central theme of this book in the 1960s, starting with the proof of Theorem 3.8 in 1962 [78]. The proof we give here follows Ehrhart's original lines of thought. An interesting fact is that he did his most beautiful work as a teacher at a *lycée* in Strasbourg (France), receiving his doctorate at age 60 on the urging of some colleagues.

3. Given any d linearly independent vectors $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_d \in \mathbb{R}^d$, the **lattice** generated by them is the set of all integer linear combinations of $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_d$. Alternatively, one can define a lattice as a discrete subgroup of \mathbb{R}^d , and these two notions can be shown to be equivalent. One might wonder whether replacing the lattice \mathbb{Z}^d by an arbitrary lattice \mathcal{L} throughout the statements of the theorems—requiring now that the vertices of a polytope be in \mathcal{L} —gives us any different results. The fact that the theorems of this chapter

remain the same follows from the observation that any lattice can be mapped to \mathbb{Z}^d by an invertible linear transformation.

4. Richard Stanley developed much of the theory of Ehrhart (quasi-)polynomials, initially from a commutative-algebra point of view. Theorem 3.12 is due to him [169]. The proof we give here appeared in [30]. An extension of Theorem 3.12 was found by Takayuki Hibi; he proved that if $h_d > 0$ then $h_k \geq h_1$ for all $1 \leq k \leq d-1$ (using the notation of Theorem 3.12) [97].

5. The tetrahedron \mathcal{T}_h of Example 3.22 was used by John Reeve to show that Pick's theorem does not hold in \mathbb{R}^3 (see Exercise 3.18) [153]. Incidentally, the formula for $L_{\mathcal{T}_h}$ also proves that the coefficients of an Ehrhart polynomial (of a closed polytope) are not always positive.

6. There are several interesting questions (some of which are still open) regarding the periods of Ehrhart quasipolynomials. Some particularly nice examples about what can happen with periods were given by Tyrrell McAllister and Kevin Woods [125].

7. Most of the results remain true if we replace “convex polytope” by “polytopal complex,” that is, a finite union of polytopes. One important exception is Corollary 3.15: the constant term of an “Ehrhart polynomial” of an integral polytopal complex C is the *Euler characteristic* of C .

8. The reader might wonder why we do not discuss polytopes with *irrational* vertices. The answer is simple: nobody has yet found a theory that would parallel the results in this chapter, even in dimension two. One notable exception is [11], in which irrational extensions of Brion's theorem are given; we will study the rational case of Brion's theorem in Chapter 9. On the other hand, Ehrhart theory has been extended to functions other than strict lattice-point counting; one instance is described in Chapter 11.

Exercises

3.1. To any permutation $\pi \in S_d$ on d elements, we associate the simplex

$$\Delta_\pi := \text{conv} \{ \mathbf{0}, \mathbf{e}_{\pi(1)}, \mathbf{e}_{\pi(1)} + \mathbf{e}_{\pi(2)}, \dots, \mathbf{e}_{\pi(1)} + \mathbf{e}_{\pi(2)} + \dots + \mathbf{e}_{\pi(d)} \},$$

where $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_d$ denote the unit vectors in \mathbb{R}^d .

- Prove that $\{\Delta_\pi : \pi \in S_d\}$ is a triangulation of the unit d -cube $[0, 1]^d$.
- Prove that all Δ_π are congruent to each other, that is, each one can be obtained from any other by reflections, translations, and rotations.
- Show that for all $\pi \in S_d$, $L_{\Delta_\pi}(t) = \binom{d+t}{d}$.

3.2. ♣ Suppose T is a triangulation of a pointed cone. Prove that the intersection of two simplicial cones in T is again a simplicial cone.

3.3. Find the generating function $\sigma_{\mathcal{K}}(\mathbf{z})$ for the following cones:

- (a) $\mathcal{K} = \{\lambda_1(0, 1) + \lambda_2(1, 0) : \lambda_1, \lambda_2 \geq 0\}$;
- (b) $\mathcal{K} = \{\lambda_1(0, 1) + \lambda_2(1, 1) : \lambda_1, \lambda_2 \geq 0\}$;
- (c) $\mathcal{K} = \{(3, 4) + \lambda_1(0, 1) + \lambda_2(2, 1) : \lambda_1, \lambda_2 \geq 0\}$.

3.4. ♣ Let $S \subseteq \mathbb{R}^m$ and $T \subseteq \mathbb{R}^n$. Show that $\sigma_{S \times T}(z_1, z_2, \dots, z_{m+n}) = \sigma_S(z_1, z_2, \dots, z_m) \sigma_T(z_{m+1}, z_{m+2}, \dots, z_{m+n})$.

3.5. ♣ Let \mathcal{K} be a rational d -cone, and let $\mathbf{m} \in \mathbb{Z}^d$. Show that $\sigma_{\mathbf{m}+\mathcal{K}}(\mathbf{z}) = \mathbf{z}^{\mathbf{m}} \sigma_{\mathcal{K}}(\mathbf{z})$.

3.6. ♣ For a set $S \subset \mathbb{R}^d$, let $-S := \{-x : x \in S\}$. Prove that

$$\sigma_{-S}(z_1, z_2, \dots, z_d) = \sigma_S\left(\frac{1}{z_1}, \frac{1}{z_2}, \dots, \frac{1}{z_d}\right).$$

3.7. Given a pointed cone $\mathcal{K} \subset \mathbb{R}^d$ with apex at the origin, let $S := \mathcal{K} \cap \mathbb{Z}^d$. Show that if $\mathbf{x}, \mathbf{y} \in S$ then $\mathbf{x} + \mathbf{y} \in S$. (In algebraic terms, S is a *semi-group*, since $\mathbf{0} \in S$ and associativity of the addition in S follows trivially from associativity in \mathbb{R}^d .)

3.8. ♣ Prove Lemma 3.9: If

$$\sum_{t \geq 0} f(t) z^t = \frac{g(z)}{(1-z)^{d+1}},$$

then f is a polynomial of degree d if and only if g is a polynomial of degree at most d and $g(1) \neq 0$.

3.9. Prove that $\binom{x+n}{n}, \binom{x+n-1}{n}, \dots, \binom{x}{n}$ is a basis for the vector space Pol_n of polynomials (in the variable x) of degree less than or equal to n .

3.10. For a polynomial $p(t) = c_d t^d + c_{d-1} t^{d-1} + \dots + c_0$, let $H_p(z)$ be defined by

$$\sum_{t \geq 0} p(t) z^t = \frac{H_p(z)}{(1-z)^{d+1}}.$$

Consider the map $\phi_d : \text{Pol}_d \rightarrow \text{Pol}_d$, $p \mapsto H_p$.

- (a) Show that ϕ_d is a linear transformation.
- (b) Compute the matrix describing ϕ_d for $d = 0, 1, 2, \dots$
- (c) Deduce formulas for h_2, h_3, \dots , similar to the one in Corollary 3.16.

3.11. Compute the Ehrhart polynomials and the Ehrhart series of the simplices with the following vertices:

- (a) $(0, 0, 0)$, $(1, 0, 0)$, $(0, 2, 0)$, and $(0, 0, 3)$;
 (b) $(0, 0, 0, 0)$, $(1, 0, 0, 0)$, $(0, 2, 0, 0)$, $(0, 0, 3, 0)$, and $(0, 0, 0, 4)$.

3.12. Define the **hypersimplex** $\Delta(d, k)$ as the convex hull of

$$\{\mathbf{e}_{j_1} + \mathbf{e}_{j_2} + \cdots + \mathbf{e}_{j_k} : 1 \leq j_1 < j_2 < \cdots < j_k \leq d\},$$

where $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_d$ are the standard basis vectors in \mathbb{R}^d . For example, $\Delta(d, 1) = \Delta(d, d-1)$ are regular $(d-1)$ -simplices. Compute the Ehrhart polynomial and the Ehrhart series of $\Delta(d, k)$.

3.13. ♣ Suppose H is the hyperplane given by

$$H = \{\mathbf{x} \in \mathbb{R}^d : a_1x_1 + a_2x_2 + \cdots + a_dx_d = 0\}$$

for some $a_1, a_2, \dots, a_d \in \mathbb{Z}$, which we may assume to have no common factor. Prove that there exists $\mathbf{v} \in \mathbb{Z}^d$ such that $\bigcup_{n \in \mathbb{Z}} ((n\mathbf{v} + H) \cap \mathbb{Z}^d) = \mathbb{Z}^d$. (This implies, in particular, that the points in $\mathbb{Z}^d \setminus H$ are all at least some minimal distance away from H ; this minimal distance is essentially given by the dot product of \mathbf{v} with (a_1, a_2, \dots, a_d) .)

3.14. ♣ A hyperplane H is **rational** if it can be written in the form

$$H = \{\mathbf{x} \in \mathbb{R}^d : a_1x_1 + a_2x_2 + \cdots + a_dx_d = b\}$$

for some $a_1, a_2, \dots, a_d, b \in \mathbb{Z}$. A **hyperplane arrangement** in \mathbb{R}^d is a finite set of hyperplanes in \mathbb{R}^d . Prove that a rational hyperplane arrangement \mathcal{H} can be translated so that no hyperplane in \mathcal{H} contains any integer points.

3.15. The conclusion of the previous exercise can be strengthened: Prove that a rational hyperplane arrangement \mathcal{H} can be translated such that no hyperplane in \mathcal{H} contains any *rational* points.

3.16. Show that, given distinct numbers x_1, x_2, \dots, x_{d+1} , the matrix

$$\mathbf{V} = \begin{pmatrix} x_1^d & x_1^{d-1} & \cdots & x_1 & 1 \\ x_2^d & x_2^{d-1} & \cdots & x_2 & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ x_{d+1}^d & x_{d+1}^{d-1} & \cdots & x_{d+1} & 1 \end{pmatrix}$$

is not singular. (\mathbf{V} is known as the *Vandermonde matrix*.)

3.17. Let \mathcal{P} be an integral d -polytope. Show that

$$\text{vol } \mathcal{P} = \frac{1}{d!} \left((-1)^d + \sum_{k=1}^d \binom{d}{k} (-1)^{d-k} L_{\mathcal{P}}(k) \right).$$

3.18. As in Example 3.22, let \mathcal{T}_n be the tetrahedron with vertices $(0, 0, 0)$, $(1, 0, 0)$, $(0, 1, 0)$, and $(1, 1, n)$, where n is a positive integer. Show that the volume of \mathcal{T}_n is unbounded as $n \rightarrow \infty$, yet for all n , \mathcal{T}_n has no interior and four boundary lattice points. This example proves that Pick's theorem does not hold for a three-dimensional integral polytope \mathcal{P} , in the sense that there is no linear relationship among $\text{vol } \mathcal{P}$, $L_{\mathcal{P}}(1)$, and $L_{\mathcal{P}^\circ}(1)$.

3.19. ♣ Prove Lemma 3.24: If

$$\sum_{t \geq 0} f(t) z^t = \frac{g(z)}{h(z)},$$

then f is a quasipolynomial of degree d with period dividing p if and only if g and h are polynomials such that $\deg(g) < \deg(h)$, all roots of h are p^{th} roots of unity of multiplicity at most $d+1$, and there is a root of multiplicity equal to $d+1$ (all of this assuming that g/h has been reduced to lowest terms).

3.20. ♣ Provide the details for the proof of Theorem 3.23: If \mathcal{P} is a rational convex d -polytope, then $L_{\mathcal{P}}(t)$ is a quasipolynomial in t of degree d . Its period divides the least common multiple of the coordinates of the vertices of \mathcal{P} .

3.21. Let \mathcal{T} be the rational triangle with vertices $(0, 0)$, $(1, \frac{p-1}{p})$, and $(p, 0)$, where p is a fixed integer ≥ 2 . Show that $L_{\mathcal{T}}(t) = \frac{p-1}{2}t^2 + \frac{p+1}{2}t + 1$; in particular, $L_{\mathcal{T}}$ is a *polynomial*.

3.22. Prove that for any $d \geq 2$ and any $p \geq 1$, there exists a d -polytope \mathcal{P} whose Ehrhart quasipolynomial is a *polynomial* (i.e., it has period 1), yet \mathcal{P} has a vertex with denominator p .

3.23. Prove that the period of the Ehrhart quasipolynomial of a 1-dimensional polytope is *always* equal to the lcm of the denominators of its vertices.

3.24. Let \mathcal{T} be the triangle with vertices $(-\frac{1}{2}, -\frac{1}{2})$, $(\frac{1}{2}, -\frac{1}{2})$, and $(0, \frac{3}{2})$. Show that $L_{\mathcal{T}}(t) = t^2 + c(t)t + 1$, where

$$c(t) = \begin{cases} 1 & \text{if } t \text{ is even,} \\ 0 & \text{if } t \text{ is odd.} \end{cases}$$

(This shows that the periods of the “coefficients” of an Ehrhart quasipolynomial do not necessarily increase with decreasing power.) Find the Ehrhart series of \mathcal{T} .

3.25. Prove the following extension of Theorem 3.12: Suppose \mathcal{P} is a rational d -polytope with denominator p . Then

$$\text{Ehr}_{\mathcal{P}}(z) = \frac{f(z)}{(1 - z^p)^{d+1}},$$

where f is a polynomial with nonnegative integral coefficients.

3.26. Find and prove a statement that extends Lemma 3.14 to Ehrhart quasipolynomials.

3.27. Prove the following extension of Corollary 3.15 to rational polytopes. Namely, the Ehrhart quasipolynomial $L_{\mathcal{P}}$ of the rational convex polytope $\mathcal{P} \subset \mathbb{R}^d$ satisfies $L_{\mathcal{P}}(0) = 1$.

3.28. Prove the following analogue of Corollary 3.17 for rational polytopes: Suppose \mathcal{P} is a rational polytope with Ehrhart quasipolynomial $L_{\mathcal{P}}(t) = c_d(t)t^d + c_{d-1}(t)t^{d-1} + \cdots + c_1(t)t + c_0(t)$. Then for all $t \in \mathbb{Z}$ and $0 \leq k \leq d$, we have $d! c_k(t) \in \mathbb{Z}$.

3.29. ♣ Prove that Corollary 3.20 also holds for rational polytopes: Suppose $\mathcal{P} \subset \mathbb{R}^d$ is a rational convex d -polytope with Ehrhart quasipolynomial $c_d(t)t^d + c_{d-1}(t)t^{d-1} + \cdots + c_0(t)$. Then $c_d(t)$ equals the volume of \mathcal{P} ; in particular, $c_d(t)$ is constant.

3.30. Suppose f and g are quasipolynomials. Prove that the **convolution**

$$F(t) := \sum_{s=0}^t f(s) g(t-s)$$

is also a quasipolynomial. What can you say about the degree and the period of F , given the degrees and periods of f and g ?

3.31. Given two positive, relatively prime integers a and b , let

$$f(t) := \begin{cases} 1 & \text{if } a|t, \\ 0 & \text{otherwise,} \end{cases} \quad \text{and} \quad g(t) := \begin{cases} 1 & \text{if } b|t, \\ 0 & \text{otherwise.} \end{cases}$$

Form the convolution of f and g . What function is it?

3.32. Suppose $\mathcal{P} \subset \mathbb{R}^m$ and $\mathcal{Q} \subset \mathbb{R}^n$ are rational polytopes. Prove that the convolution of $L_{\mathcal{P}}$ and $L_{\mathcal{Q}}$ equals the Ehrhart quasipolynomial of the polytope given by the convex hull of $\mathcal{P} \times \{\mathbf{0}_n\} \times \{0\}$ and $\{\mathbf{0}_m\} \times \mathcal{Q} \times \{1\}$. Here $\mathbf{0}_d$ denotes the origin in \mathbb{R}^d .

3.33. We define the **unimodular group** $\mathrm{SL}_d(\mathbb{Z})$ as the set of all $d \times d$ matrices with integer entries and determinant ± 1 .

- (a) Show that $\mathrm{SL}_d(\mathbb{Z})$ acts on the integer lattice \mathbb{Z}^d as a one-to-one, onto map. That is, fix any $\mathbf{A} \in \mathrm{SL}_d(\mathbb{Z})$. Then \mathbf{A} maps \mathbb{Z}^d to itself in a bijective fashion.
- (b) For any open simplex $\Delta^\circ \subset \mathbb{R}^d$ and any $\mathbf{A} \in \mathrm{SL}_d(\mathbb{Z})$, consider the image of Δ° under \mathbf{A} , defined by $\mathbf{A}(\Delta^\circ) := \{\mathbf{A}\mathbf{x} : \mathbf{x} \in \Delta^\circ\}$. Show that

$$\#\{\Delta^\circ \cap \mathbb{Z}^d\} = \#\{\mathbf{A}(\Delta^\circ) \cap \mathbb{Z}^d\}.$$

- (c) Let \mathcal{P} be an integral polytope, and let $\mathcal{Q} := \mathbf{A}(\mathcal{P})$, where $\mathbf{A} \in \mathrm{SL}_d(\mathbb{Z})$, so that \mathcal{P} and \mathcal{Q} are unimodular images of each other. Show that $L_{\mathcal{P}}(t) = L_{\mathcal{Q}}(t)$. (*Hint:* Write \mathcal{P} as the disjoint union of open simplices.)

3.34. Search on the Internet for the program **LattE**: Lattice-Point Enumeration [65, 114]. You can download it for free. Experiment.

Open Problems

3.35. How many triangulations are there for a given polytope?

3.36. What is the minimal number of simplices needed to triangulate the unit d -cube? (These numbers are known for $d \leq 7$.)

3.37. Classify the polynomials of a fixed degree d that are Ehrhart polynomials. This is completely done for $d = 2$ [159] and partially known for $d = 3$ and 4 [24, Section 3].

3.38. Study the roots of Ehrhart polynomials of integral polytopes in a fixed dimension [24, 37, 41, 93]. Study the roots of the numerators of Ehrhart series.

3.39. Come up with an efficient algorithm that computes the period of an Ehrhart quasipolynomial. (See [187], in which Woods describes an efficient algorithm that checks whether a given integer is a period of an Ehrhart quasipolynomial.)

3.40. Suppose \mathcal{P} and \mathcal{Q} are integer polytopes with the same Ehrhart polynomial, that is, $L_{\mathcal{P}}(t) = L_{\mathcal{Q}}(t)$. What additional conditions on \mathcal{P} and \mathcal{Q} do we need to ensure that integer translates of \mathcal{P} and \mathcal{Q} are unimodular images of each other? That is, when is $\mathcal{Q} = \mathbf{A}(\mathcal{P}) + \mathbf{m}$ for some $\mathbf{A} \in \mathrm{SL}_d(\mathbb{Z})$ and $\mathbf{m} \in \mathbb{Z}^d$?

3.41. Find an “Ehrhart theory” for irrational polytopes.



<http://www.springer.com/978-0-387-29139-0>

Computing the Continuous Discretely
Integer-point Enumeration in Polyhedra

Beck, M.; Robins, S.

2007, XVIII, 227 p., Hardcover

ISBN: 978-0-387-29139-0