

A finite calculus approach to Ehrhart polynomials

Steven V. Sam

Department of Mathematics
Massachusetts Institute of Technology

`ssam@math.mit.edu`
`http://math.mit.edu/~ssam`

Kevin M. Woods

Department of Mathematics
Oberlin College

`kevin.woods@oberlin.edu`
`http://www.oberlin.edu/faculty/kwoods`

Submitted: Nov 24, 2009; Accepted: Apr 20, 2010; Published: Apr 30, 2010

Mathematics Subject Classification: 52C07

Abstract

A rational polytope is the convex hull of a finite set of points in $\mathbb{R}^d$ with rational coordinates. Given a rational polytope $\mathcal{P} \subseteq \mathbb{R}^d$, Ehrhart proved that, for $t \in \mathbb{Z}_{\geq 0}$, the function $\#(t\mathcal{P} \cap \mathbb{Z}^d)$ agrees with a quasi-polynomial $L_{\mathcal{P}}(t)$, called the Ehrhart quasi-polynomial. The Ehrhart quasi-polynomial can be regarded as a discrete version of the volume of a polytope. We use that analogy to derive a new proof of Ehrhart's theorem. This proof also allows us to quickly prove two other facts about Ehrhart quasi-polynomials: McMullen's theorem about the periodicity of the individual coefficients of the quasi-polynomial and the Ehrhart–Macdonald theorem on reciprocity.

1 Introduction.

Let us first look at an (easy) example of computing a *volume*. Let $\Delta_d \subseteq \mathbb{R}^d$ be the convex hull of the following $d + 1$ points: the origin and the standard basis vectors e_i , $1 \leq i \leq d$. Let $t\Delta_d$ be its dilation by a factor of t (for nonnegative t). A straightforward way of computing the volume of $t\Delta_d$ would be inductively in d , using the fact that the $(d - 1)$ -dimensional cross section of $t\Delta_d$ at $x_d = s$ is a dilated copy of Δ_{d-1} :

$$\text{vol}(t\Delta_d) = \int_0^t \text{vol}(s\Delta_{d-1}) \, ds,$$

and evaluating this iteratively gives us $\text{vol}(t\Delta_d) = t^d/d!$.

A generalization of volume is the *Ehrhart (quasi-)polynomial*, which we define as follows. A *polytope*, $\mathcal{P}$, is the convex hull of finitely many points in $\mathbb{R}^n$, and the dimension, $\dim(\mathcal{P})$, of the polytope is the dimension of the affine hull of $\mathcal{P}$. A *rational* (resp., *integral*) polytope is a polytope all of whose vertices are rational (resp., integral). Given a

polytope $\mathcal{P}$ and a nonnegative t , let $t\mathcal{P}$ be the dilation of $\mathcal{P}$ by a factor of t , and define the function $L_{\mathcal{P}}: \mathbb{Z}_{\geq 0} \rightarrow \mathbb{Z}_{\geq 0}$ by

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

Ehrhart proved [Ehr] that, if $\mathcal{P}$ is an integral polytope, then $L_{\mathcal{P}}(t)$ is a polynomial of degree $\dim(\mathcal{P})$. More generally, if $\mathcal{P}$ is a rational polytope of dimension $d = \dim(\mathcal{P})$, then

$$L_{\mathcal{P}}(t) = c_0(t) + c_1(t)t + \cdots + c_d(t)t^d,$$

where the c_i are periodic functions $\mathbb{Z} \rightarrow \mathbb{Q}$ (periodic meaning that there exists an s such that $c_i(t) = c_i(t+s)$ for all t). Such functions are called *quasi-polynomials*. Ehrhart quasi-polynomials can be considered as a generalization of volume, because, for full-dimensional $\mathcal{P}$, $c_d(t)$ is the constant $\text{vol}(\mathcal{P})$. That is, $L_{\mathcal{P}}(t)$ is approximately $\text{vol}(t\mathcal{P}) = \text{vol}(\mathcal{P})t^d$, with lower degree terms correcting for the fact that this is a discrete version of the volume computation.

Let us return to our polytope $t\Delta_d \subseteq \mathbb{R}^d$ and compute its Ehrhart polynomial. For this example, our inductive approach to computing volume works out well when translated to the discrete problem. When $d = 1$, $L_{\Delta_1}(t) = t + 1$. We see that

$$L_{\Delta_d}(t) = \sum_{s=0}^t L_{\Delta_{d-1}}(s),$$

which we can prove by induction on d and t is

$$\frac{(t+1)(t+2) \cdots (t+d)}{d!}.$$

This calculation works out so well because expressions like the falling factorial,

$$t^{\underline{d}} := t(t-1)(t-2) \cdots (t-d+1),$$

sum well. This is a well-known fact from finite calculus [GKP, Chapter 2], and just as the polynomials t^d form the perfect basis of $\mathbb{R}[t]$ (as a vector space over $\mathbb{R}$) for integrating because of the power rule, the polynomials $t^{\underline{d}}$ form the perfect basis for summing, since

$$\sum_{i=0}^t i^{\underline{d}} = \frac{1}{d+1} (t+1)^{\underline{d+1}} \quad (1.1)$$

(this fact can be proved quickly, by induction on t).

In Section 2, we prove that this method of computing $L_{\mathcal{P}}(t)$ works for any simplex (and hence for any polytope by triangulation). This provides a new proof of Ehrhart's theorem that uses more minimal (but less powerful) tools than other traditional proofs, such as proofs via generating functions [Ehr, Sta] or via valuations [McM]. Unlike these other proofs, proving it for integral polytopes requires the full power of the proof for rational

polytopes. To prove it, we'll need a nice basis for the vector space of quasi-polynomials of period s , which we shall present in Section 2.

This inductive computation of $L_{\mathcal{P}}(t)$ has two more desirable outcomes: new and basic proofs of McMullen's theorem about periods of the individual coefficients, $c_i(t)$, of the quasi-polynomial and of Ehrhart–Macdonald reciprocity. We describe both of these results below.

McMullen's theorem [McM, Theorem 4], is as follows. The i -index of a rational polytope $\mathcal{P}$ is the smallest number s_i such that, for each i -dimensional face F of $\mathcal{P}$, the affine hull of $s_i F$ contains integer points. For this definition, we include $\mathcal{P}$ as a d -dimensional face of itself. Note that if $i \geq j$, then we must have $s_i | s_j$: any i -dimensional face, F , contains j -dimensional faces, and so the affine hull of $s_j F$ contains integer points, though it may not be the smallest dilate to do so.

Theorem 1.2 (McMullen's theorem). *Given a rational polytope $\mathcal{P} \subseteq \mathbb{R}^n$, let $d = \dim(\mathcal{P})$, and let*

$$L_{\mathcal{P}}(t) = \#(t\mathcal{P} \cap \mathbb{Z}^n) = c_0(t) + c_1(t)t + \cdots + c_d(t)t^d$$

be the Ehrhart quasi-polynomial. Given i , with $0 \leq i \leq d$, let s_i be the i -index of $\mathcal{P}$. Then s_i is a period of $c_i(t)$.

For example, let $\mathcal{D}$ be the smallest positive integer such that $\mathcal{D}\mathcal{P}$ is integral. Then s_i divides $\mathcal{D}$, for all i , and so $\mathcal{D}$ is a period of each $c_i(t)$. If $\mathcal{P}$ is an integral polytope, then $\mathcal{D} = 1$, and we recover that $L_{\mathcal{P}}(t)$ is actually a *polynomial*. As another example, if $\mathcal{P}$ is full-dimensional then the affine span of $\mathcal{P}$ is all of $\mathbb{R}^d$, and therefore $c_d(t)$ has period 1. As mentioned, $c_d(t)$ is the constant which is the volume of $\mathcal{P}$. McMullen's theorem is proven in Section 2, concurrently with Ehrhart's theorem.

Now we describe *Ehrhart–Macdonald reciprocity*. Since the function $L_{\mathcal{P}}(t)$ agrees with a quasi-polynomial $p(t)$ for all positive integers, a natural question to ask is if $p(t)$ has any meaning when t is a negative integer, and indeed it does. Given a polytope $\mathcal{P}$, let $\mathcal{P}^\circ$ be the relative interior of $\mathcal{P}$, that is, the interior when considering $\mathcal{P}$ as a subset of its affine hull. The number of integer points in $t\mathcal{P}^\circ$ is similarly counted by $L_{\mathcal{P}^\circ}(t)$.

Theorem 1.3 (Ehrhart–Macdonald reciprocity). *Let $\mathcal{P}$ be a rational polytope. Then*

$$L_{\mathcal{P}^\circ}(t) = (-1)^{\dim(\mathcal{P})} L_{\mathcal{P}}(-t).$$

This statement was conjectured by Ehrhart, and he proved it in many special cases. The general case was proven by Macdonald in [Mac]. This will be proven in Section 3, using the following idea, which could be called a reciprocity theorem for finite calculus.

Suppose $f(s)$ is a quasi-polynomial, and suppose we are examining $F(t) = \sum_{i=0}^t f(i)$. We will show in Section 2 that there is a quasi-polynomial $p(t)$ such that $F(t) = p(t)$, for nonnegative integers t . How about for negative integers? Certainly we can evaluate p at a negative integer, $-t$, but we need to define what

$$F(-t) = \sum_{i=0}^{-t} f(i)$$

should mean. Assuming that we want the summation rule

$$\sum_{i=0}^a f(i) + \sum_{i=a+1}^b f(i) = \sum_{i=0}^b f(i)$$

to hold for all integers a and b , we must have that

$$\sum_{i=0}^{-t} f(i) + \sum_{i=-t+1}^0 f(i) = \sum_{i=0}^0 f(i),$$

which means we should define

$$F(-t) = \sum_{i=0}^{-t} f(i) := - \sum_{i=-t+1}^{-1} f(i).$$

Fortunately, when we plug in negative values, we still have $F(-t) = p(-t)$. This is the content of the following lemma, which we prove in Section 3.

Lemma 1.4 (Reciprocity for finite calculus). *Suppose that $f(i)$ is a quasi-polynomial in i . For nonnegative integers t define the function*

$$F(t) = \sum_{i=0}^t f(i).$$

Then there is a quasi-polynomial $p(t)$ such that $F(t) = p(t)$ for all nonnegative integers t , and furthermore,

$$p(-t) = - \sum_{i=-t+1}^{-1} f(i)$$

for all $t > 0$.

2 Ehrhart's theorem and McMullen's theorem.

As mentioned in the introduction, “discrete integration” of polynomials is made easy by using the basis t^d of $\mathbb{R}[t]$. We will use the following generalization, which tells us how to discretely integrate quasi-polynomials.

Lemma 2.1. *Let $f(t) = c_0(t) + c_1(t)t + \cdots + c_d(t)t^d$ be a quasi-polynomial of degree d , where $c_i(t)$ is a periodic function of period s_i , for each i . Define $F: \mathbb{Z}_{\geq 0} \rightarrow \mathbb{Q}$ by*

$$F(t) = \sum_{i=0}^{\lfloor \frac{at}{b} \rfloor} f(i),$$

where $a, b \in \mathbb{Z}$ and $\lfloor \cdot \rfloor$ is the greatest integer function. Let $S_i = \frac{s_i b}{\gcd(s_i, a)}$. Then $F(t) = C_0(t) + C_1(t)t + \cdots + C_{d+1}(t)t^{d+1}$ is a quasi-polynomial of degree $d+1$. Furthermore, a period of $C_i(t)$ is $\text{lcm}\{S_d, S_{d-1}, \dots, S_i\}$, for $0 \leq i \leq d$, and C_{d+1} has period 1.

Before we prove this lemma, let's look at an example. Suppose that

$$f(t) = \begin{cases} t/2 & \text{if } t \text{ even} \\ 0 & \text{if } t \text{ odd} \end{cases},$$

and we would like to evaluate the sum

$$F(t) = \sum_{i=0}^{\lfloor 3t/2 \rfloor} f(i).$$

We have that $s_1 = 2$ and $s_0 = 1$, which give us $S_1 = 4$ and $S_0 = 2$. The lemma tells us that the period of the t^2 coefficient of $F(t)$ should be 1, the period of the t^1 coefficient should be $S_1 = 4$, and the period of the t^0 coefficient should be $\text{lcm}\{S_1, S_0\} = 4$. Indeed,

$$F(t) = \sum_{i=0}^{\lfloor 3t/2 \rfloor} f(i) = \sum_{j=0}^{\lfloor 3t/4 \rfloor} j = \frac{1}{2} (\lfloor 3t/4 \rfloor + 1)^2 = \begin{cases} \frac{9}{32}t^2 + \frac{3}{8}t & \text{if } t \equiv 0 \pmod{4} \\ \frac{9}{32}t^2 - \frac{3}{16}t - \frac{3}{32} & \text{if } t \equiv 1 \pmod{4} \\ \frac{9}{32}t^2 - \frac{1}{8} & \text{if } t \equiv 2 \pmod{4} \\ \frac{9}{32}t^2 + \frac{3}{16}t - \frac{3}{32} & \text{if } t \equiv 3 \pmod{4} \end{cases}.$$

Notice that this shows why taking the lcm of $S_d, \dots, S_i$ is necessary: $f(t)$ has periodicity only in the t^1 coefficient, but affects the period of both the t^1 and t^0 coefficients of $F(t)$.

Proof of 2.1. Given d , s , and j , define the periodic function

$$\chi_{s,j}(t) = \begin{cases} 1 & \text{if } t \equiv j \pmod{s} \\ 0 & \text{otherwise} \end{cases}$$

and the quasi-polynomial

$$g_{d,s,j}(t) = \chi_{s,j}(t) \prod_{k=0}^{d-1} \left(\frac{t-j}{s} - k \right).$$

For instance, in the preceding example, we had $f(t) = g_{1,2,0}(t)$. For $t \equiv j \pmod{s}$, substituting $t = ms + j$ gives us $g_{d,s,j}(ms + j) = m^d$. This implies that, for a given d and s , the set of $g_{d',s,j}$ such that $0 \leq d' \leq d$ and $0 \leq j < s$ forms a basis (and, as we will see, a nice basis!) for the set of all quasi-polynomials of degree at most d with period s . Writing our function $f(t)$ as a linear combination of such quasi-polynomials (for various d , s , and j), it suffices to prove that

$$G_{d,s,j}(t) = \sum_{i=0}^{\lfloor \frac{at}{b} \rfloor} g_{d,s,j}(i)$$

is a quasi-polynomial of degree $d+1$ and period $S = \frac{sb}{\text{gcd}(s,a)}$ whose leading term has period 1 coefficient.

We have that, for any $k \in \mathbb{Z}_+$,

$$\begin{aligned} \sum_{i=0}^k g_{d,s,j}(i) &= \sum_{m=0}^{\lfloor \frac{k-j}{s} \rfloor} g_{d,s,j}(ms+j) \\ &= \sum_{m=0}^{\lfloor \frac{k-j}{s} \rfloor} m^d \\ &= \frac{1}{d+1} \left(\left\lfloor \frac{k-j}{s} \right\rfloor + 1 \right)^{\frac{d+1}{2}}, \end{aligned}$$

where the last line follows from (1.1), and so

$$G_{d,s,j}(t) = \frac{1}{d+1} \left(\left\lfloor \frac{\lfloor \frac{at}{b} \rfloor - j}{s} \right\rfloor + 1 \right)^{\frac{d+1}{2}}.$$

One can check that this is a quasi-polynomial of period $S = \frac{sb}{\gcd(s,a)}$ whose leading coefficient has period 1 by substituting $t = mS + k$:

$$\begin{aligned} G_{d,s,j} \left(m \frac{sb}{\gcd(s,a)} + k \right) &= \frac{1}{d+1} \left(\left\lfloor \frac{\frac{ams}{\gcd(s,a)} + \lfloor \frac{ak}{b} \rfloor - j}{s} \right\rfloor + 1 \right)^{\frac{d+1}{2}} \\ &= \frac{1}{d+1} \left(\frac{am}{\gcd(s,a)} + \left\lfloor \frac{\lfloor \frac{ak}{b} \rfloor - j}{s} \right\rfloor + 1 \right)^{\frac{d+1}{2}}, \end{aligned}$$

a polynomial in m whose leading coefficient does not depend on k . The lemma follows. $\square$

Proof of Ehrhart's Theorem and of 1.2. We prove this by induction on d . As the base case, consider $d = 0$. Then $\mathcal{P}$ is a point in $\mathbb{Q}^n$. Let $\mathcal{D}$ be the smallest positive integer such that $\mathcal{D}\mathcal{P}$ is an integer point. Then we see that

$$L_{\mathcal{P}}(t) = c_0(t), \text{ where } c_0(t) = \begin{cases} 1 & \text{if } \mathcal{D} \mid t \\ 0 & \text{otherwise} \end{cases}.$$

The base case follows. Now we assume that the theorem is true for all $d' < d$. We divide the proof into a number of steps:

1: Without loss of generality, we may assume that $\mathcal{P}$ is full-dimensional, that is, $\dim(\mathcal{P}) = n$.

Let s' be the smallest positive integer such that the affine hull of $s'\mathcal{P}$ contains integer points. Then we must have that s' divides each s_i . Let V be the affine hull of $s'\mathcal{P}$. There is an affine transformation $T: V \rightarrow \mathbb{R}^{\dim(\mathcal{P})}$ that maps $V \cap \mathbb{Z}^n$ bijectively onto $\mathbb{Z}^{\dim(\mathcal{P})}$.

Let $\mathcal{P}' = T(s'\mathcal{P})$. Then $\mathcal{P}'$ is a full-dimensional polytope. If we can prove the theorem for $\mathcal{P}'$, it will follow for $\mathcal{P}$, because

$$L_{\mathcal{P}}(t) = \begin{cases} L_{\mathcal{P}'}\left(\frac{t}{s'}\right) & \text{if } s' \text{ divides } t \\ 0 & \text{otherwise} \end{cases}.$$

2: Without loss of generality, we may assume that $\mathcal{P} = \text{conv}\{0, \mathcal{Q}\}$, where $\mathcal{Q}$ is a $(d-1)$ -dimensional rational polytope.

Assume that we have a general rational polytope $\mathcal{P}$, with $\dim(\mathcal{P}) = d$. Without loss of generality, translate it by an integer vector so that it does not contain the origin. We simply write $L_{\mathcal{P}}(t)$ as sums and differences of polytopes of the form $\text{conv}\{0, \mathcal{Q}\}$ (including lower dimensional $\mathcal{Q}$), using inclusion-exclusion to make sure that the intersections of faces are counted properly. The exact form of this decomposition is not important for this proof, but it will be important in Section 3, so we will present it now. We examine two types of faces of $\mathcal{P}$:

- The collection $\mathcal{F}_v$ of faces F of $\mathcal{P}$ that are “visible”: a facet (that is, a $(d-1)$ -dimensional face) is said to be visible if, for all $a \in F$ and all λ with $0 < \lambda < 1$, we have $\lambda a \notin \mathcal{P}$, and a lower dimensional face is visible if every facet that it is contained in is visible.
- The collection $\mathcal{F}_h$ of faces F of $\mathcal{P}$ that are “hidden”: a facet is “hidden” if it is not visible, and a lower dimensional face is hidden if every facet that it is contained in is hidden.

Some lower dimensional faces may be neither visible nor hidden. For a face F of $\mathcal{P}$, let $\mathcal{P}_F = \text{conv}(0, F)$. Then, using inclusion-exclusion,

$$L_{\mathcal{P}}(t) = \sum_{F \in \mathcal{F}_h} (-1)^{d-1-\dim(F)} L_{\mathcal{P}_F}(t) - \sum_{F \in \mathcal{F}_v} (-1)^{d-1-\dim(F)} (L_{\mathcal{P}_F}(t) - L_F(t)). \quad (2.2)$$

An example of this decomposition for a polygon is given in Figure 2.3. So as not to interrupt the flow of the proof, we offer a proof of the correctness of (2.2) at the end of the section.

For a given face F of $\mathcal{P}$, the i -dimensional faces F' of $\mathcal{P}_F$ are either faces of $\mathcal{P}$ or contain the origin. In either case, the affine hull of $s_i F'$ contains integer points, so they meet the conditions of the theorem. The theorem is true for the third piece of the sum, $\sum_{F \in \mathcal{F}_v} (-1)^{d-\dim(F)} L_F(t)$, by the induction hypothesis, since these are faces of smaller dimension than $\mathcal{P}$.

3: Without loss of generality, we may assume that $\mathcal{P} = \text{conv}\{0, \mathcal{Q}\}$, where $\mathcal{Q}$ is a $(d-1)$ -dimensional rational polytope lying in the hyperplane $x_d = q$, where $q \in \mathbb{Q}_{>0}$.

Perform a unimodular transformation such that this is true.

4: We prove the theorem for such a $\mathcal{P}$.

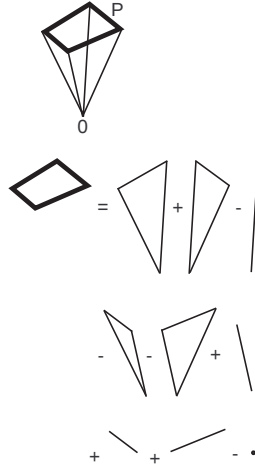


Figure 2.3: Decomposition of a 2-dimensional polytope.

We have $\mathcal{P} = \text{conv}\{0, \mathcal{Q}\}$, where $\mathcal{Q}$ is a $(d-1)$ -dimensional rational polytope lying in the hyperplane $x_d = \frac{a}{b}$, where $a, b \in \mathbb{Z}_{>0}$ and $\gcd(a, b) = 1$. Since faces of $\mathcal{Q}$ are also faces of $\mathcal{P}$, it follows that the affine hull of $s_i F$, where F is an i -dimensional face of $\mathcal{Q}$, contains integer points. Let $\bar{\mathcal{Q}} = \frac{b}{a}\mathcal{Q}$, lying in the hyperplane $x_d = 1$. Then the affine hull of $s_i \frac{a}{b}\bar{F}$, where $\bar{F}$ is an i -dimensional face of $\bar{\mathcal{Q}}$, contains integer points. We have that $t\mathcal{P} \cap \mathbb{Z}^d$ is the disjoint union

$$\bigcup_{i=0}^{\lfloor \frac{ta}{b} \rfloor} i\bar{\mathcal{Q}} \cap \mathbb{Z}^d,$$

and so

$$L_{\mathcal{P}}(t) = \sum_{i=0}^{\lfloor \frac{ta}{b} \rfloor} L_{\bar{\mathcal{Q}}}(i).$$

By Lemma 2.1, this is a quasi-polynomial of degree d . Furthermore, the S_i in the statement of Lemma 2.1 are given by

$$S_i = \frac{(s_i \frac{a}{b})b}{\gcd(s_i \frac{a}{b}, a)} = \frac{as_i}{a} = s_i.$$

Since $s_d | s_{d-1} | \cdots | s_0$, $s_i = \text{lcm}(s_d, s_{d-1}, \dots, s_i)$, and the coefficients of $L_{\mathcal{P}}(t)$ have the desired periods. The theorem follows. $\square$

More can be said about the period of $c_{d-1}(t)$. In this case, s_{d-1} is not only a period but is guaranteed to be the *minimum* period. A proof of this fact along with a study of maximal period behavior is given in [BSW] and relies only on McMullen's theorem and Ehrhart–Macdonald reciprocity, which we prove in the next section.

We also remark that, following the constant term through the induction, we get for free another well-known fact: that the constant term of the Ehrhart polynomial of a polytope is 1. More precisely, the constant term of the Ehrhart polynomial for a polytopal complex (open or closed) is equal to its Euler characteristic.

We close this section with a proof of (2.2).

Proof of (2.2). One can prove that this inclusion-exclusion is correct combinatorially, but the quickest proof to understand is topological. Let $\mathcal{C} = \bigcup_{F \in \mathcal{F}_h} F$ and $\mathcal{C}' = \bigcup_{F \in \mathcal{F}_v} F$. We only need to prove that the first sum in (2.2) counts each $x \in \text{conv}\{0, \mathcal{C}\}$ exactly once, and that the second sum counts each $x \in \text{conv}\{0, \mathcal{C}'\} \setminus \mathcal{C}'$ exactly once. Let's examine the first sum. It suffices to prove that each $x \in \mathcal{C}$ is counted correctly, as each $\lambda x \in \text{conv}\{0, \mathcal{C}\}$ is counted identically to x .

Assume for the moment that x lies in the interior of $\mathcal{C}$, considered as a $(d-1)$ -dimensional CW complex. Let B be the intersection of $\mathcal{C}$ with the closure of a sufficiently small ball around x (small enough so that B only intersects faces F that contain x). B inherits a CW complex structure from $\mathcal{C}$. In particular, there is a one-to-one correspondence between $F \in \mathcal{F}_h$ that contain x and cells of B that are not contained in the boundary ∂B . Therefore, in the first sum of (2.2), the number of times the point x is counted is

$$n(x) = (-1)^{d-1} \sum_{F \ni x} (-1)^{\dim(F)} = (-1)^{d-1} (\chi(B) - \chi(\partial B)),$$

where χ is the Euler characteristic (the alternating sum of the number of cells of each dimension). Since B is contractible and ∂B is homeomorphic to a $(d-2)$ -sphere, $\chi(B) = 1$ and $\chi(\partial B) = 1 + (-1)^{d-2}$. Hence

$$n(x) = (-1)^{d-1} [1 - (1 + (-1)^{d-2})] = 1,$$

so x is properly counted in the sum.

If x is not in the interior of $\mathcal{C}$, notice that x is counted exactly the same as any “nearby” point that is in the interior: the key is that faces on the boundary of $\mathcal{C}$ are not defined to be hidden faces in $\mathcal{F}_h$, because they are also contained in visible facets. Therefore these x are also counted correctly. A similar argument shows that the second sum properly counts each $x \in \text{conv}\{0, \mathcal{C}'\} \setminus \mathcal{C}'$. $\square$

3 Reciprocity.

In this section, we prove Theorem 1.3. First we prove Lemma 1.4, which gives a reciprocity theorem for finite calculus.

Proof of Lemma 1.4. By Lemma 2.1, there is a quasi-polynomial $p(t)$ such that $F(t) = p(t)$ for nonnegative integers t . Let n be a fixed positive integer. Define

$$C_n = \sum_{i=-n}^{-1} f(i),$$

and for integers $t \geq -n$ define

$$F_n(t) = -C_n + \sum_{i=-n}^t f(i).$$

Using Lemma 2.1 (and reindexing as necessary), we see that there is a quasi-polynomial $p_n(t)$ such that $F_n(t) = p_n(t)$ for all integers $t \geq -n$. But then we see that, for nonnegative integers t ,

$$p_n(t) = -C_n + \sum_{i=-n}^t f(i) = \sum_{i=0}^t f(i) = F(t) = p(t).$$

Because $p_n(t)$ and $p(t)$ agree for all nonnegative t , they must be identical as quasi-polynomials, and in particular

$$p(-n) = -C_n + \sum_{i=-n}^{-n} f(i) = - \sum_{i=-n+1}^{-1} f(i),$$

as desired. □

Proof of Theorem 1.3. Again, we induct on the dimension d of the polytope. The inductive step will consist of two parts. First, assume $\mathcal{P}$ is a d -dimensional polytope that is the convex hull of the origin and $\mathcal{Q}$, where $\mathcal{Q}$ is a $(d-1)$ -dimensional polytope. We shall first prove reciprocity for these types of polytopes. Second, having reciprocity for pyramids, we use the explicit inclusion-exclusion formula for the indicator functions of the “exterior point triangulation” given by (2.2) to show that reciprocity holds in general.

Let $\mathcal{Q}$ be a $(d-1)$ -dimensional polytope in $\mathbb{R}^d$ contained in the hyperplane $x_d = \frac{a}{b}$ for nonzero, relatively prime integers a and b , and let $\mathcal{P}$ be the pyramid $\text{conv}\{0, \mathcal{Q}\}$. As before, define $\bar{\mathcal{Q}} = \frac{b}{a}\mathcal{Q}$, lying in the hyperplane $x_d = 1$. Let $f(i)$ give the number of lattice points in $i\bar{\mathcal{Q}}$ and $f^\circ(i)$ give the number of lattice points in $i\bar{\mathcal{Q}}^\circ$. By induction, we can assume that $f^\circ(i) = (-1)^{d-1}f(-i)$. So

$$F(t) = \sum_{i=0}^{\lfloor \frac{ta}{b} \rfloor} f(i)$$

is the number of lattice points in $\mathcal{P}$, and

$$F^\circ(t) = \sum_{i=1}^{\lceil \frac{ta}{b} \rceil - 1} f^\circ(i)$$

is the number of lattice points in $\mathcal{P}^\circ$. Let $p(t)$ be the quasi-polynomial which corresponds to $F(t)$. By Lemma 1.4, $p(t)$ agrees with $F(t)$ for both positive and negative integers.

Let $t' = \lceil \frac{ta}{b} \rceil$. Then

$$\begin{aligned}
(-1)^d p(-t) &= (-1)^d F(-t) \\
&= (-1)^d \sum_{i=0}^{\lfloor \frac{-ta}{b} \rfloor} f(i) \\
&= (-1)^d \sum_{i=0}^{-t'} f(i) \\
&= (-1)^{d+1} \sum_{i=-t'+1}^{-1} [(-1)^{d-1} f^\circ(-i)] \\
&= \sum_{i=1}^{t'-1} f^\circ(i) = F^\circ(t).
\end{aligned}$$

We now consider the case for general rational polytopes $\mathcal{P}$. As in part 2 of the proof of Ehrhart's theorem, we write $\mathcal{P}$ as a sum and difference of polytopes of the form $\text{conv}\{0, \mathcal{Q}\}$ and lower dimensional polytopes. By equation (2.2) and the inductive hypothesis,

$$\begin{aligned}
L_{\mathcal{P}}(-t) &= \sum_{F \in \mathcal{F}_h} (-1)^{d-1-\dim(F)} L_{\mathcal{P}_F}(-t) - \sum_{F \in \mathcal{F}_v} (-1)^{d-1-\dim(F)} (L_{\mathcal{P}_F}(-t) - L_F(-t)) \\
&= (-1)^d \left[\sum_{F \in \mathcal{F}_h} L_{\mathcal{P}_F^\circ}(t) - \sum_{F \in \mathcal{F}_v} (L_{\mathcal{P}_F^\circ}(t) - L_{F^\circ}(t)) \right],
\end{aligned}$$

and it is easy to see that the right hand side counts $(-1)^d$ times the number of integer points in $t\mathcal{P}^\circ$, which finishes the proof by induction. $\square$

4 Discussion.

One might hope that this proof of Ehrhart's Theorem yields an efficient algorithm to compute Ehrhart polynomials inductively. To make this work, one must be able to efficiently compute a simple expression for sums like

$$\sum_{s=0}^t \left\lfloor \frac{2s+3}{4} \right\rfloor \cdot \left\lfloor \frac{3s+2}{5} \right\rfloor$$

(the summands are called *step-polynomials* in [VW]). The only known way to compute such sums seems to be to first convert to a generating function using methods from [VW] and then manipulate the generating function using the Barvinok Algorithm (which computes the generating function of the integer points in a polyhedron) and other related techniques, see [BP]. However, computing the sum in this way would be ill-advised, because the generating function techniques can compute the Ehrhart polynomial directly.

Put another way, an elementary algorithm that, given a summation of a step-polynomial computes the sum as a new step-polynomial, would be interesting, because it would provide an alternative algorithm to Barvinok for answering questions about integer points in polytopes.

A second insight that this proof of Ehrhart's Theorem provides is the importance of picking nice bases in which to write Ehrhart polynomials and quasi-polynomials. Perhaps, rather than the standard basis for polynomials, t^d for $d \geq 0$, a basis such as

$$L_{\Delta_d}(t) = \frac{(t+1)(t+2)\cdots(t+d)}{d!} \text{ for } d \geq 0$$

(which sums nicely) would be enlightening. A similar basis has already been studied: given d , $\binom{t+d-j}{d}$ for $0 \leq j \leq d$ is a basis for polynomials of degree at most d . If we write

$$L_{\mathcal{P}}(t) = \sum_{j=0}^d h_j \binom{t+d-j}{d},$$

then the associated Hilbert series has the form

$$\sum_{s=0}^{\infty} L_{\mathcal{P}}(s)t^s = \frac{h_0 + h_1t + \cdots + h_dt^d}{(1-t)^{d+1}}.$$

See, for example, Section 3.4 of [BR] for a discussion of this, including a proof of the fact that the h_j are nonnegative. This basis has recently been used [Bra] to study the roots of the Ehrhart polynomial, inspiring further study [Pfe] of roots of polynomials whose coefficients are nonnegative in arbitrary bases.

Acknowledgements

The authors would like to thank Matthias Beck and Timothy Chow for helpful discussions, and the anonymous referee for helping improve the exposition.

References

- [BP] Alexander Barvinok and Jamie Pommersheim, An algorithmic theory of lattice points in polyhedra, *New Perspectives in Algebraic Combinatorics (Berkeley, CA, 1996–97)*, volume 38 of *Math. Sci. Res. Inst. Publ.*, Cambridge Univ. Press, Cambridge (1999), 91–147.
- [BR] Matthias Beck and Sinai Robins, *Computing the Continuous Discretely*, Springer (2007).
- [BSW] Matthias Beck, Steven V. Sam, and Kevin M. Woods, Maximal periods of (Ehrhart) quasi-polynomials, *J. Combinatorial Theory, Ser. A* **115** (2008), 517–525.

- [Bra] Benjamin Braun, Norm bounds for Ehrhart polynomial roots, *Discrete and Computational Geometry* **39** (2008), 191–193.
- [Ehr] Eugène Ehrhart, Sur les polyèdres rationnels homothétiques à n dimensions, *C. R. Acad. Sci. Paris* **254** (1962), 616–618.
- [GKP] Ronald Graham, Donald Knuth, and Oren Patashnik, *Concrete Mathematics*, Second edition, Addison-Wesley, 1994.
- [Mac] I. G. Macdonald, Polynomials associated with finite cell-complexes, *J. London Math. Soc.* (2) **4** (1971), 181–192.
- [McM] Peter McMullen, Lattice invariant valuations on rational polytopes, *Arch. Math. (Basel)* **31** (1978/79), no. 5, 509–516.
- [Pfe] Julian Pfeifle, Gale duality bounds for roots of polynomials with nonnegative coefficients, *J. Combinatorial Theory, Ser. A* **117** (2010), 248–271.
- [Sta] Richard P. Stanley, *Enumerative Combinatorics, Vol. I*, Cambridge Studies in Advanced Mathematics **49**, Cambridge University Press, 1997.
- [VW] Sven Verdoolaege and Kevin Woods, Counting with rational generating functions, *J. Symbolic Computation* **43** (2008), no. 2, 75–91