The minimum period of the Ehrhart quasi-polynomial of a rational polytope

Tyrrell B. McAllister and Kevin M. Woods

March 31, 2005

Abstract

If $P \subset \mathbb{R}^d$ is a rational polytope, then $i_P(n) := \#(nP \cap \mathbb{Z}^d)$ is a quasi-polynomial in n, called the Ehrhart quasi-polynomial of P. The period of $i_P(n)$ must divide $\mathcal{D}(P) = \min\{n \in \mathbb{Z}_{>0} : nP \text{ is an integral polytope}\}$. Few examples are known where the period is not exactly $\mathcal{D}(P)$. We show that for any $\mathcal{D}$, there is a 2-dimensional triangle P such that $\mathcal{D}(P) = \mathcal{D}$ but such that the period of $i_P(n)$ is 1, that is, $i_P(n)$ is a polynomial in n. We also characterize all polygons P such that $i_P(n)$ is a polynomial. In addition, we provide a counterexample to a conjecture by T. Zaslavsky about the periods of the coefficients of the Ehrhart quasi-polynomial.

1 Introduction

An *integral* (respectively, *rational*) *polytope* is a polytope whose vertices have integral (respectively, rational) coordinates. Given a rational polytope $P \subset \mathbb{R}^d$, the *denominator* of P is

$$\mathcal{D}(P) = \min\{n \in \mathbb{Z}_{>0} : nP \text{ is an integral polytope}\}.$$

Ehrhart proved ([1]) that if $P \subset \mathbb{R}^d$ is a rational polytope, then there is a quasi-polynomial function $i_P : \mathbb{Z} \mapsto \mathbb{Z}$ with period $\mathcal{D}(P)$ such that, for $n \geq 0$,

$$i_P(n) = \#\left(nP \cap \mathbb{Z}^d\right).$$

In other words, there exist polynomial functions $f_1, \ldots, f_{\mathcal{D}(P)}$ such that $i_P(n) = f_j(n)$ for $n \equiv j \pmod{\mathcal{D}(P)}$. In particular, if P is integral, then $\mathcal{D}(P) = 1$, so i_P is a polynomial function.

We call i_P the *Ehrhart quasi-polynomial of P*. This counting function satisfies several important properties:

1. The degree of each f_j is the dimension of P.

2. The coefficient of the leading term of each f_j is the volume of P, normalized with respect to the sublattice of $\mathbb{Z}^d$ which is the intersection of $\mathbb{Z}^d$ with the affine hull of P (in particular, if P is full dimensional, the coefficient is simply the Euclidean volume of P).

3. (Law of Reciprocity) For $n \geq 1$, let
$$i_P^{\circ}(n) = \#\left(interior(nP) \cap \mathbb{Z}^d\right).$$
Then $i_P^{\circ}(n) = (-1)^d i_P(-n)$.

Properties (1) and (2) were proved by Ehrhart in [1]. Property (3) was conjectured by Ehrhart and proved in full generality by I.G. MacDonald in [2]. For an excellent introduction to Ehrhart quasi-polynomials that includes proofs of all these properties, see [3].

We know that $\mathcal{D}(P)$ is *a* period of the Ehrhart quasi-polynomial of P, but what is the *minimum* period? Of course, it must divide $\mathcal{D}(P)$, and it very often equals $\mathcal{D}(P)$. Though this is not always the case, very few counterexamples were previously known. R.P. Stanley ([3], Example 4.6.27) provided an example of a polytope P with denominator $\mathcal{D}(P) = 2$ where the minimum period is 1, that is, where the Ehrhart quasi-polynomial is actually a polynomial. Stanley's example is a 3-dimensional pyramid P with vertices $(0,0,0)$, $(1,0,0)$, $(0,1,0)$, $(1,1,0)$, and $(1/2,0,1/2)$. In this case, $i_P(n) = \binom{n+3}{3}$.

We say that *period collapse* occurs when the minimum period is strictly less than the denominator of the polytope. We say that P has *full period* if the minimum period equals the denominator of the polytope. Stanley's example raises some natural questions. In what dimensions can period collapse occur? Can period collapse occur for P such that $\mathcal{D}(P) > 2$? What values may the minimum period be when it is not $\mathcal{D}(P)$? This note answers all of these questions.

In Section 2, we provide (Theorem 2.2) an infinite class of 2-dimensional triangles such that, for any $\mathcal{D}$, there is a triangle P in this class with denominator $\mathcal{D}$, but such that $i_P(n)$ is actually a polynomial. In fact, for any $d \geq 2$ and for any $\mathcal{D}$ and s with $s|\mathcal{D}$, there is a d-dimensional polytope with denominator $\mathcal{D}$ but with minimum period s. Such period collapse cannot occur in dimension 1, however: rational 1-dimensional polytopes always have full period (Theorem 2.1). Finally, in Section 3 (Theorem 3.1), we give a geometric characterization of all polygons P whose quasi-polynomials are actually polynomials. We also provide several examples, one of which settles a conjecture of Zaslavsky that we detail now.

Another way to consider the period of a quasi-polynomial is to examine the periods of its coefficients. Suppose P is a d-dimensional polytope and, for all j,
$$f_j(n) = c_{jd}n^d + c_{j,d-1}n^{d-1} + \cdots + c_{j1}n + c_{j0}.$$
Then we say that s_k, the *period of the kth coefficient*, is the minimum period of the sequence
$$c_{1k}, c_{2k}, c_{3k}, \ldots.$$

The minimal period of P is then the least common multiple of $s_0, s_1, \ldots, s_d$. T. Zaslavsky conjectured (unpublished) that the periods of the coefficients are decreasing, *i.e.*, $s_k \leq s_{k-1}$ for $1 \leq k \leq d$. In this paper, we provide a counterexample (Example 3.3) which is a 2-dimensional triangle.

2 Period Collapse

First, we prove that period collapse cannot happen in dimension 1.

Theorem 2.1. *The quasi-polynomials of rational 1-dimensional polytopes always have full period.*

Proof. In this case, P is simply a segment $[\frac{p}{q}, \frac{r}{s}]$ (where the integers p, q, r, and s are chosen so that the fractions are fully reduced). Write $\mathcal{D} = \mathcal{D}(P) = \operatorname{lcm}(s, q)$.

On the one hand, we clearly have that

$$i_P(n) = \left\lfloor n\frac{r}{s} \right\rfloor - \left\lceil n\frac{p}{q} \right\rceil + 1. \tag{1}$$

On the other hand, there exist $\mathcal{D}$ polynomials $f_1(n), \ldots, f_{\mathcal{D}}(n)$ such that $i_P(n) = f_j(n)$, for $n \equiv j \pmod{\mathcal{D}}$. The claim is that i_P has period $\mathcal{D}$. To show this, it suffices to show that the constant term of $f_j(n)$ is 1 if and only if $j = \mathcal{D}$.

Since P is one-dimensional, we have that, for each $j \in \{1, 2, \ldots, \mathcal{D}\}$, the polynomial $f_j(n)$ is linear, and therefore it is determined by its values at $n = j$ and $n = j + \mathcal{D}$. Interpolating using (1) yields

$$f_j(n) = \left(\frac{r}{s} - \frac{p}{q}\right) n + 1 - \left(\left\lceil j\frac{p}{q} \right\rceil - j\frac{p}{q}\right) - \left(j\frac{r}{s} - \left\lfloor j\frac{r}{s} \right\rfloor\right).$$

The constant term is 1 if and only if q and s both divide j, which happens if and only if $j = \mathcal{D}$. □

While in dimension 1, nothing (with respect to period collapse) is possible, in dimension 2 and higher, anything is possible, as the following theorem demonstrates.

Theorem 2.2. *Given $d \geq 2$, and given $\mathcal{D}$ and s such that $s|\mathcal{D}$, there exists a d-dimensional polytope with denominator $\mathcal{D}$ whose Ehrhart quasi-polynomial has minimum period s.*

Proof. We first prove the theorem in the case where $d = 2$ and $s = 1$; that is, we exhibit a polygon with denominator $\mathcal{D}$ for which $i_P(n)$ is actually a polynomial in n. Given $\mathcal{D} \geq 2$, let P be the triangle with vertices $(0, 0), (1, \frac{\mathcal{D}-1}{\mathcal{D}})$, and $(\mathcal{D}, 0)$ (see Figure 1). We will prove that

$$i_P(n) = \frac{\mathcal{D}-1}{2}n^2 + \frac{\mathcal{D}+1}{2}n + 1.$$

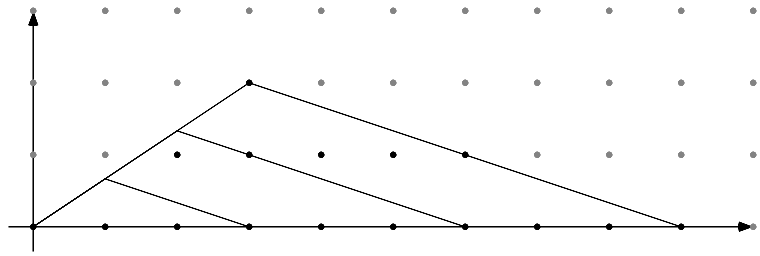

Figure 1: The first three dilations of P when $\mathcal{D} = 3$

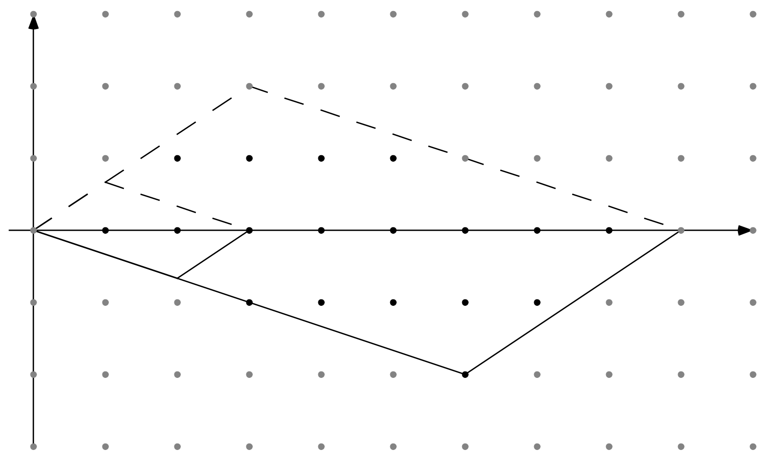

Figure 2: Q and $3Q$ when $\mathcal{D} = 3$

First we will calculate $i_Q(n)$, where Q is the half-open parallelogram with vertices $(0,0), (1, \frac{\mathcal{D}-1}{\mathcal{D}}), (\mathcal{D}, 0)$, and $(\mathcal{D}-1, -\frac{\mathcal{D}-1}{\mathcal{D}})$ and with top two edges open. That is, to construct Q, take the closed parallelogram with these vertices and remove the line segments $\Big[(0,0), (1, \frac{\mathcal{D}-1}{\mathcal{D}})\Big]$ and $\Big[(1, \frac{\mathcal{D}-1}{\mathcal{D}}), (\mathcal{D}, 0)\Big]$ (see Figure 2). Q has the nice property that, for $n \in \mathbb{N}$, nQ can be tiled by translates of Q with no overlap. It is clear that Q contains exactly $\mathcal{D} - 1$ lattice points (the lattice points $(1,0), (2,0), \ldots, (\mathcal{D}-1, 0)$). To tile nQ, however, we must use translates of Q that are not lattice translates, so it is not immediately clear how many lattice points these translates contain. In fact, they all contain $\mathcal{D} - 1$ points, as we shall show.

It suffices to prove this for $Q_t = Q - (0, \frac{t}{\mathcal{D}})$, where $t = 0, 1, \ldots, \mathcal{D}-1$, because all of the translates of Q that we need to tile nQ are lattice translates of one of these Q_t. The only horizontal lines $y = a$, with a integral, that possibly intersect Q_t are $y = 0$ and $y = -1$, and they intersect Q_t with x-coordinates in the intervals $(\frac{t}{\mathcal{D}-1}, \mathcal{D} - t)$ and $[\mathcal{D} - t, \mathcal{D} - 1 + \frac{t-1}{\mathcal{D}-1}]$, respectively. These intervals contain $\mathcal{D}-t-1$ and t integral points, respectively, so in all, Q_t contains $\mathcal{D}-1$ integer points. Therefore, we must have that

$$i_Q(n) = (\mathcal{D}-1)n^2.$$

Let $\overline{Q}$ be the closure of Q. To calculate $i_{\bar{Q}}(n)$, we must add to $i_Q(n)$ the number of integer points in $n\overline{Q} \setminus nQ$, which is $n+1$ (one can check that the number of lattice points on the interval $\Big[(0,0),(n, n\frac{\mathcal{D}-1}{\mathcal{D}})\Big)$ is $\lfloor \frac{n-1}{\mathcal{D}} \rfloor + 1$ and the number of lattice points on the interval $\Big[(n, n\frac{\mathcal{D}-1}{\mathcal{D}}),(0, n\mathcal{D})\Big]$ is $n - \lfloor \frac{n-1}{\mathcal{D}} \rfloor$, so there are $n+1$ in all). So

$$i_{\bar{Q}}(n) = (\mathcal{D}-1)n^2 + n + 1.$$

$n\overline{Q}$ is the union (not disjoint) of 2 copies of nP (one rotated by a half-turn), each with the same number of lattice points. The overlap of these two copies of nP is the line segment $\Big[(0,0),(0,\mathcal{D}n)\Big]$, which contains $\mathcal{D}n+1$ integer points. Therefore

$$i_P(n) = \frac{1}{2}\Big(i_{\bar{Q}}(n) + (\mathcal{D}n+1)\Big) = \frac{\mathcal{D}-1}{2}n^2 + \frac{\mathcal{D}+1}{2}n + 1,$$

as desired.

Now suppose d is 2, but s is not necessarily 1. Let P' be the pentagon with vertices $(0,0),(1,\frac{\mathcal{D}-1}{\mathcal{D}}),(\mathcal{D},0),(\mathcal{D},-\frac{1}{s})$, and $(0,-\frac{1}{s})$. If P is the triangle defined as before, then $nP' \setminus nP$ contains $\lfloor \frac{n}{s} \rfloor \cdot (\mathcal{D}n+1)$ lattice points, and so

$$i_{P'}(n) = i_P(n) + \left\lfloor \frac{n}{s} \right\rfloor \cdot (\mathcal{D}n+1),$$

which has minimum period s.

Now suppose d is greater than 2. Let P' be the pentagon defined as before, and let $P'' = P' \times [0,1]^{d-2}$, a polytope of dimension d. Then

$$i_{P''}(n) = (n+1)^{d-2} i_{P'}(n),$$

which also has minimum period s. □

3 The 2-dimensional Case

We have seen (in Theorem 2.2) an infinite class of rational polygons P in dimension 2 such that $i_P(n)$ is a polynomial. Can we characterize such polygons? We know that, for all *integer* polygons P, $i_P(n)$ is a polynomial. One property that an integer polygon P has is that it and its dilates satisfy Pick's theorem, i.e., if we let $\partial_P(n) = \#\Big(boundary(nP) \cap \mathbb{Z}^d\Big)$, then

$$\begin{aligned} i_P(n) &= \text{Area}(nP) + \frac{1}{2}\partial_P(n) + 1 \\ &= n^2 \text{Area}(P) + \frac{1}{2}\partial_P(n) + 1. \end{aligned}$$

Another property that an integer polygon, P, and its dilates satisfy is that the number of points on their boundary is linear, i.e.,

$$\partial_P(n) = n\partial_P(1).$$

In fact, these two properties are exactly what we need to guarantee that a rational polygon's Ehrhart quasi-polynomial is actually a polynomial.

Theorem 3.1. *Let $P \subset \mathbb{Z}^2$ be a rational polygon, let A be the area of P, and let $\mathcal{D}$ be the denominator of P. Then the following are equivalent:*

1. *$i_P(n)$ is a polynomial in n;*
2. *$i_P(n) = An^2 + \frac{1}{2}\partial_P(1)n + 1$;*
3. *For all $n \in \mathbb{N}$,*
   - *(a) nP obeys Pick's theorem, i.e., $i_P(n) = An^2 + \frac{1}{2}\partial_P(n) + 1$, and*
   - *(b) $\partial_P(n) = n\partial_P(1)$; and*
4. *For $n = 1, 2, \ldots, \mathcal{D}$, 3a and 3b hold.*

Proof. We will prove that $1 \Rightarrow 2 \Rightarrow 3 \Rightarrow 4 \Rightarrow 2 \Rightarrow 1$. Two of these steps, $3 \Rightarrow 4$ and $2 \Rightarrow 1$, are trivial. To prove the remaining implications, we will repeatedly use the law of reciprocity for Ehrhart quasi-polynomials, which was stated in the introduction.

$1 \Rightarrow 2$. If 1 holds, then $i_P(n) = An^2 + bn + c$ for some b and c. Since $i_P(0) = 1$, we know that $c = 1$. By the reciprocity law, we know that

$$i_P^\circ(n) = A(-n)^2 + b(-n) + c,$$

and so

$$\partial_P(1) = i_P(1) - i_P^\circ(1) = 2b.$$

Therefore $i_P(n) = An^2 + \frac{1}{2}\partial_P(1)n + 1$, as desired.

$2 \Rightarrow 3$. If 2 holds, then, again using reciprocity, for all $n \in \mathbb{N}$,

$$i_P^\circ(n) = An^2 - \frac{1}{2}\partial_P(1)n + 1,$$

and so

$$\partial_P(n) = i_P(n) - i^\circ_P(n) = \partial_P(1)n,$$

and so 3b holds. Then

$$\begin{aligned} i_P(n) &= An^2 + \frac{1}{2}\partial_P(1)n + 1 \\ &= An^2 + \frac{1}{2}\partial_P(n) + 1, \end{aligned}$$

and so 3a holds.

$4 \Rightarrow 2$. If 4 holds, then let

$$f_j(n) = An^2 + b_j n + c_j,$$

for $j = 1, 2, \ldots, \mathcal{D}$, be the polynomials such that $i_P(n) = f_j(n)$ for $n \equiv j \pmod{\mathcal{D}}$. Given j with $1 \le j \le \mathcal{D}$, we again use reciprocity, and we have

$$\begin{aligned} j\partial_P(1) &= \partial_P(j) \\ &= f_j(j) - f_{\mathcal{D}-j}(-j) \\ &= (b_j + b_{\mathcal{D}-j}) \cdot j + (c_j - c_{\mathcal{D}-j}) \end{aligned} \tag{2}$$

and

$$\begin{aligned} (\mathcal{D} - j)\partial_P(1) &= \partial_P(\mathcal{D} - j) \\ &= f_{\mathcal{D}-j}(\mathcal{D} - j) - f_j(j - \mathcal{D}) \\ &= (b_j + b_{\mathcal{D}-j}) \cdot (\mathcal{D} - j) + (c_{\mathcal{D}-j} - c_j) \end{aligned} \tag{3}$$

Multiplying Equation (2) by $\mathcal{D} - j$ and Equation (3) by j and subtracting,

$$0 = \mathcal{D} \cdot (c_j - c_{\mathcal{D}-j}),$$

and so

$$c_j = c_{\mathcal{D}-j}. \tag{4}$$

Adding Equations (2) and (3),

$$\mathcal{D} \cdot \partial_P(1) = \mathcal{D} \cdot (b_j + b_{\mathcal{D}-j}),$$

and so

$$b_j + b_{\mathcal{D}-j} = \partial_P(1). \tag{5}$$

Using the facts that Pick's theorem holds and that $j\partial_P(1) = \partial_P(j)$, we have

$$\begin{aligned} Aj^2 + \frac{1}{2}\partial_P(1) \cdot j + 1 &= Aj^2 + \frac{1}{2}\partial_P(j) + 1 \\ &= f_j(j) \\ &= Aj^2 + b_j \cdot j + c_j, \end{aligned}$$

and so

$$\frac{1}{2}\partial_P(1) \cdot j + 1 = b_j \cdot j + c_j. \tag{6}$$

Similarly,

$$\frac{1}{2}\partial_P(1) \cdot (\mathcal{D} - j) + 1 = b_{\mathcal{D}-j} \cdot (\mathcal{D} - j) + c_{\mathcal{D}-j}. \tag{7}$$

Multiplying Equation (6) by $\mathcal{D} - j$ and Equation (7) by j and adding together (and then using Equations (4) and (5)),

$$\begin{aligned} \partial_P(1) \cdot j \cdot (\mathcal{D} - j) + \mathcal{D} &= (b_j + b_{\mathcal{D}-j}) \cdot j \cdot (\mathcal{D} - j) + (\mathcal{D} - j) \cdot c_j + j \cdot c_{\mathcal{D}-j} \\ &= \partial_P(1) \cdot j \cdot (\mathcal{D} - j) + \mathcal{D} \cdot c_j, \end{aligned}$$

and so $c_j = 1$. Substituting $c_j = 1$ into Equation (6), we see that $b_j = \frac{1}{2}\partial_P(1)$. Therefore, for all $n \in \mathbb{N}$,

$$i_P(n) = An^2 + \frac{1}{2}\partial_P(1)n + 1,$$

as desired. □

Example 3.2. *P is the triangle with vertices $(0, 0), (\mathcal{D}, 0)$, and $(1, \frac{\mathcal{D}-1}{\mathcal{D}})$, for some $\mathcal{D} \in \mathbb{N}$.*

This is the example from Theorem 2.2 with denominator $\mathcal{D}$ for which the Ehrhart quasi-polynomial is a polynomial. One can check that conditions 3a and 3b are met.

Example 3.3. *P is the triangle with vertices $(-\frac{1}{2}, -\frac{1}{2}), (\frac{1}{2}, -\frac{1}{2})$, and $(0, \frac{3}{2})$.*

One can check that nP, for $n \in \mathbb{N}$, satisfies 3a (Pick's theorem), but not 3b. Indeed, we have

$$i_P(n) = \begin{cases} n^2 + 1, & \text{if } n \text{ is odd} \\ n^2 + n + 1, & \text{if } n \text{ is even,} \end{cases}$$

which is not a polynomial. This example disproves a conjecture of T. Zaslavsky that the period of the coefficient of n^k in the quasi-polynomial increases as k decreases (in the example, the coefficients of n^2 and n^0 have period 1, but the coefficient of n^1 has period 2). A similar counterexample has been found independently by D. Einstein.

Example 3.4. *P is the triangle with vertices $(0, 0), (1, 0)$, and $(0, \frac{1}{2})$.*

In this example, nP, for $n \in \mathbb{N}$ satisfies 3b, but not 3a. We have

$$i_P(n) = \begin{cases} \frac{1}{4}n^2 + n + \frac{3}{4}, & \text{if } n \text{ is odd,} \\ \frac{1}{4}n^2 + n + 1, & \text{if } n \text{ is even.} \end{cases}$$

Acknowledgements

We would like to thank Matthias Beck and Jesus De Loera for helpful conversations. Special thanks to David Einstein for a simplification of the example in Theorem 2.2.

References

[1] EHRHART, E., *Polynmes arithmtiques et mthode des polydres en combinatoire*. International Series of Numerical Mathematics, Vol. 35. Birkhuser Verlag, Basel-Stuttgart, 1977.

[2] MACDONALD, I. G., Polynomials associated with finite cell complexes, *J. London Math. Soc.* **4** (1971), 181-192.

[3] STANLEY, R. P., *Enumerative Combinatorics, Volume I*, Cambridge University Press, 1997.