# ON EHRHART THEORY AND POSITIVITY

RAJIT PANDEY

Abstract. We investigate the geometric and combinatorial properties of polytope triangulations, focusing on their connections to Ehrhart positivity. The Ehrhart polynomial L(P,t) encodes the number of integer points in dilated copies of a polytope, and establishing the positivity of the coefficients of L(P,t) remains a central and challenging problem in Ehrhart theory. To address this, we use unimodular triangulations to provide insight into the behavior of the polynomials' coefficients.

## 1. Introduction

The idea of counting integer points inside shapes has a long history, tied to the evolution of modern discrete geometry. Perhaps the earliest breakthrough occurred in 1899, when Austrian mathematician Georg Pick published a simple and elegant formula [7] for the area of a polygon whose vertices lie on a grid, namely $A = I + \frac{B}{2} - 1$. Pick's Theorem showed that the area $A$ of a polygon could be found by counting the grid points inside the shape $I$ and those on its boundary $B$. Over the next several decades, his discovery sparked effort to find a similar rule for three dimensions and beyond. However, higher dimensions turned out to be much more complicated than anyone expected. In 1957, John Reeve created a family of tetrahedra (shown in Fig. 1) which had the exact same number of grid points but had wildly different volumes. Reeve's counterexample [8] demonstrated that the strict relationship between lattice points and volume found in two dimensions does not directly extend to higher dimensions, initially limiting further development in the field.

The idea behind this higher-dimensional breakdown becomes apparent when examining the geometry of the Reeve tetrahedron, denoted as $R_h$ and shown in figure 1. This three-dimensional lattice polytope is defined as the convex hull (Section 3) of its four vertices:

$$R_h = \text{conv}\,\{(0,0,0),\ (1,0,0),\ (0,1,0),\ (1,1,h)\}$$

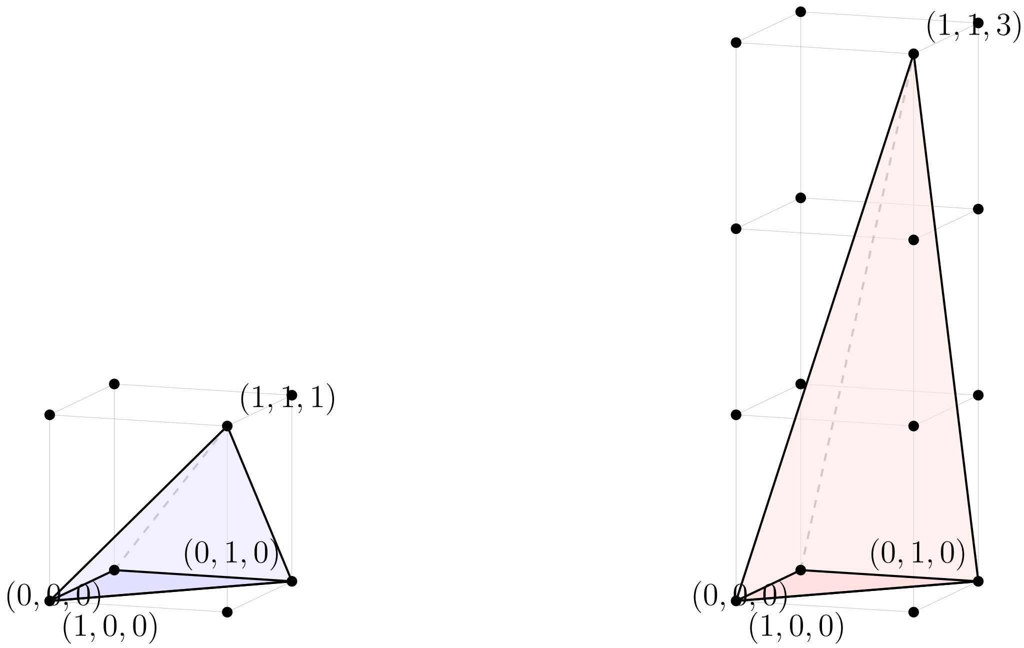
where $h \in \mathbb{Z}^+$. Evaluating the discrete point sets shows that the interior lattice count $int(R_h) \cap \mathbb{Z}^3 = \emptyset$. Similarly, the boundary contains no integer points beyond the four original vertices. Thus, the total lattice point count remains fixed at four for any value of $h$.

The Euclidean volume of $R_h$ is calculated using the determinant of the matrix formed by its edge vectors:

$$\text{Vol}(R_h) = \frac{1}{3!}\left|\det\begin{pmatrix}1 & 0 & h\\ 0 & 1 & h\\ 0 & 0 & h\end{pmatrix}\right| = \frac{h}{6}$$

The volume scales linearly with $h$ while the discrete grid interactions remain static, so any attempt to formulate an invariant higher-dimensional analogue of Pick's Theorem based on a direct linear combination of boundary and interior lattice points yields a contradiction.

*Date*: June 2026.




**(a)** Reeve Tetrahedron $R_1$ ($h = 1$)

**(b)** Reeve Tetrahedron $R_3$ ($h = 3$)

**Figure 1.** The Reeve Tetrahedron shown at different heights $h = 1$ and $h = 3$. Although the interior and boundary lattice counts remain completely fixed, the geometric volume scales proportionally, showing why 2D formulas cannot directly generalize to higher dimensions.

A significant turning point emerged in the early 1960s through the work of Eugène Ehrhart, a French high school mathematics teacher who pursued research independently. Working primarily on his own in Strasbourg, Ehrhart looked at the problem from a new angle. Instead of trying to find a fixed formula for a single polytope, he decided to look at what happens when one stretches, or *dilates*, the polytope by a whole number factor. In a series of notes presented to the French Academy of Sciences between 1962 and 1965, Ehrhart proved that the number of grid points inside these stretched polytopes grows exactly like a polynomial. This discovery, which became known as Ehrhart's Theorem, completely changed the direction of the field. It showed that what looked like a chaotic counting problem actually follows a strict algebraic pattern.

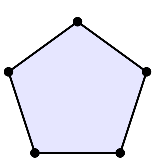
Ehrhart's discoveries [5] caught the attention of major mathematicians at the time, helping to turn a niche puzzle into a major area of study. In the 1970s, combinatorialist Richard Stanley brought Ehrhart's work into the mainstream of American mathematics. Stanley used powerful tools from commutative algebra to prove new things about the coefficients of Ehrhart polynomials, which helped connect Ehrhart's geometric ideas to algebraic ideas. Over the next few decades, this research grew into what we now call *Ehrhart theory*. Today, Ehrhart theory is a connection between geometry, combinatorics, and number theory, evolving from a simple question about grid points into a large field of modern mathematical research.

The analysis of Ehrhart polynomials comes from foundational point-counting limitations to complex algebraic transformations and combinatorial positivity conjectures. The initial topic centers on the limitation of extending lattice point counting to dimensions three and higher. Here, the variable volume of the shifting Reeve tetrahedron under boundary and interior lattice point counts demonstrates why a simple extension of Pick's Theorem fails. This shows the importance of evaluating dilated shapes through the standard *Ehrhart polynomial*, whose leading, second, and constant coefficients correspond to global Euclidean volume, relative surface area, and the topological Euler characteristic. Computing these global values requires partitioning arbitrary polytopes into *simplicial complexes* using triangulations, which reveals a sign-alternating relationship between global coefficients and the local combinatorial data of the $h^*$*-vector* (see Definition 3.6) . Standard hypercubes and *permutohedra* (Definition 5.1) are used as examples for this interaction, showing how regular structures preserve uniform point counts. In general shapes, distorted grid alignments cause the components of the $h^*$-vector to drift into later indices, where a binomial transformation (Section 7) driven by *Stirling numbers of the first kind* introduces alternating negative values that pull the intermediate polynomial coefficients below zero. Such distortion is prevented in *totally unimodular* contexts; decomposing *zonotopes* (see Section 4) into *unit parallelotopes* restricts the $h^*$-vector to early indices, bounding the consecutive coefficient ratios and forcing the complex roots of the polynomial to cluster symmetrically along the vertical line where the real part equals $-1/2$.

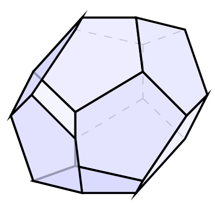
## 2. Ehrhart's Discovery

**Definition 2.1** (Polytope)**.**
A polytope is the convex hull (Section 3) of finitely many points in $d$ dimensions.

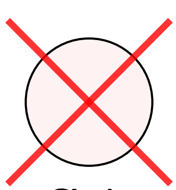
*Example.*
A pentagon is a 2-dimensional polytope, a dodecahedron is 3-dimensional, etc. *Shapes with curves, like circles or tori, are not polytopes.*



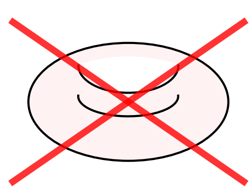
**Pentagon (2D)**



**Dodecahedron (3D)**



**Circle**



**Torus**

Ehrhart showed the function of counting the lattice points inside the polytope denoted $L(P,t)$ behaves like a polynomial where $P$ is the polytope and $t$ is the dilation factor. We find the lattice point count inside by evaluating the Ehrhart polynomial at different values of the dilation factor t.

**Definition 2.2** (Ehrhart Polynomial)**.**

For a $d$-dimensional polytope, the Ehrhart polynomial is defined as this:

$$L(P,t) = c_d t^d + c_{d-1} t^{d-1} + \cdots + c_1 t + c_0.$$

(1) The Leading Coefficient ($c_d$): This is the normalized lattice volume (Euclidean volume times $d!$) of the shape.
(2) The Second Coefficient ($c_{d-1}$): This is half of the total normalized lattice volume of the boundary, or surface area, of the shape.
(3) The Constant Term ($c_0$): This is always 1 for any solid, connected shape (it represents the Euler characteristic).

Knowing the geometric interpretations of the leading, second, and constant coefficients is useful, but explicitly calculating these coefficients for a given polytope requires deeper work. Just as a polygon in a two-dimensional plane can be cut into triangles to calculate its total area, a lattice polytope can be subdivided into fundamental units. To formalize this decomposition, let us first establish the ideas of *affine independence*, *convex hulls*, and *simplicial complexes*.

## 3. Triangulations and Simplicial Complexes

A standard mathematical strategy to analyze the Ehrhart polynomial of a complicated polytope is to decompose the shape into simpler pieces. In two dimensions, this is the familiar process of cutting a polygon into triangles. To generalize this concept to higher dimensions, we rely on simplices and simplicial complexes.

Before constructing these decompositions, let us define the boundaries of our shapes.

**Definition 3.1** (Affine Independence and Simplices)**.**
A set of points $\{v_0, v_1, \ldots, v_k\} \subset \mathbb{R}^d$ is said to be *affinely independent* if the linear displacement vectors $v_1 - v_0, v_2 - v_0, \ldots, v_k - v_0$ are linearly independent in $\mathbb{R}^d$.

**Definition 3.2** (Convex Hull)**.**
The convex hull of a finite set of points $X = \{x_1, x_2, \ldots, x_k\}$ in $\mathbb{R}^d$ is the smallest convex set containing $X$. Algebraically, it is the set of all convex combinations:

$$\mathrm{conv}(X) = \left\{ \sum_{i=1}^{k} \lambda_i x_i \;\middle|\; \sum_{i=1}^{k} \lambda_i = 1 \text{ and } \lambda_i \geq 0 \text{ for all } i \right\}.$$

**Analogy:** A way to visualize a convex hull is to imagine a rubber band (in 2D) or a balloon (in 3D) wrapping around the set of points in $\mathbb{R}^d$. It is depicted in Figure 2.

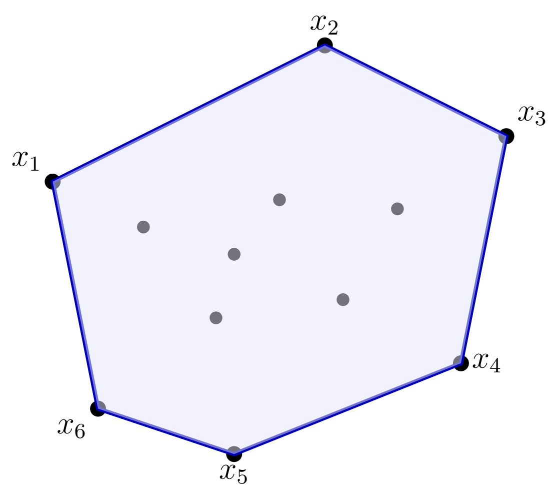



**Figure 2.** The convex hull of a finite set of points in $\mathbb{R}^2$. As described by the rubber band analogy, the boundary segments wrap around the outermost coordinates, securely enclosing all interior points within a minimal convex set.

**Definition 3.3** (Simplex)**.**
A $k$-dimensional simplex (or $k$-simplex) $\sigma$ is the convex hull of $k+1$ affinely independent points, which constitute the vertices of $\sigma$. A *face* of $\sigma$ is the convex hull of any subset of these vertices. A simplex represents the simplest possible shape one can build in any given dimension:

*Example.* Four points forming a tetrahedron would be a 3-simplex.

**Definition 3.4** (Simplicial Complex and Triangulations)**.**
A *simplicial complex* $\Delta$ in $\mathbb{R}^d$ (shown in Fig. 3) is a finite collection of simplices satisfying the following two properties:

(1) If $\sigma \in \Delta$ and $\tau$ is a face of $\sigma$, then $\tau \in \Delta$.
(2) If $\sigma_1, \sigma_2 \in \Delta$ and $\sigma_1 \cap \sigma_2 \neq \emptyset$, then $\sigma_1 \cap \sigma_2$ is a face of both $\sigma_1$ and $\sigma_2$.

A *lattice triangulation* of a lattice polytope $P$ is a simplicial complex $\Delta$ such that the union of all simplices in $\Delta$ equals $P$, and the vertices of every simplex $\sigma \in \Delta$ lie strictly in $\mathbb{Z}^d$.

**Analogy:** A way to visualize a simplicial complex is to imagine a set of LEGOs (simplices) building up a set (the polytope) such as in Figure 3. The first property described intuitively can be seen as if one has a LEGO set, they need all the pieces of it in the set to analyze it. The second property can be visualized as two LEGOs intersecting. One brick cannot intersect the inside of another, they can only meet along common faces.

**Definition 3.5** (Unimodular Simplex and Triangulation)**.**
A $d$-dimensional lattice simplex $\sigma \subset \mathbb{R}^d$ is called *unimodular* if its vertices $v_0, v_1, \ldots, v_d \in \mathbb{Z}^d$ generate the entire affine lattice over the integers. This means the absolute value of the determinant of the matrix formed by the row vectors $v_1 - v_0, v_2 - v_0, \ldots, v_d - v_0$ is equal to 1. Thus, we have that the normalized volume of a unimodular simplex is 1. A lattice triangulation $\Delta$ is a *unimodular triangulation* if every simplex $\sigma \in \Delta$ is unimodular.

(a) Valid Simplicial Complex (b) Invalid Simplicial Complex

**Figure 3.** A comparison of simplicial intersection rules. In a valid configuration, faces meet cleanly along a shared edge or vertex. In an invalid configuration, the interiors overlap improperly without sharing face boundaries.

*Example.* Consider two different triangles in $\mathbb{R}^2$:

(1) **Unimodular Simplex:** Let $\Delta_1$ be the triangle with vertices $(0,0)$, $(1,0)$, and $(10,1)$. The displacement vectors from the origin are $v_1 = (1,0)$ and $v_2 = (10,1)$. Computing the determinant of the matrix formed by these vectors gives:

$$\det \begin{pmatrix} 1 & 10 \\ 0 & 1 \end{pmatrix} = (1)(1) - (10)(0) = 1$$

Since the absolute value of the determinant is 1, $\Delta_1$ is a unimodular simplex. It has a normalized volume of 1 and contains no integer lattice points other than its three vertices.

(2) **Non-Unimodular Simplex:** Let $\Delta_2$ be the triangle with vertices $(0,0)$, $(2,0)$, and $(0,2)$. The displacement vectors are $v_1 = (2,0)$ and $v_2 = (0,2)$. The determinant calculation yields:

$$\det \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = 4$$

Because the determinant is greater than 1, $\Delta_2$ is not unimodular. Geometrically, this grid misalignment manifests as an extra lattice point $(1,1)$ in the interior of the triangle.

Among all lattice triangulations, a particularly well-behaved class consists of those whose constituent simplices are as "small" as possible relative to the grid. We formalize this idea of minimality by looking at the volume of individual simplices. When we dilate a single unimodular simplex by a positive integer factor $t$, counting its internal and boundary lattice points becomes a combinatorial problem that depends only on the dimension, rather than the specific coordinates of its vertices. By breaking a large polytope $P$ into these fundamental building blocks, we can determine the global Ehrhart polynomial $L(P,t)$ simply by understanding how these simplices glue together within the complex $\Delta$.

This decomposition process is tracked using a tool known as the $h^*$-polynomial (or $\delta$-polynomial). Instead of analyzing the coefficients $c_i$ of $L(P,t)$ directly, we rewrite the Ehrhart series—the generating function of the lattice point counts—as a rational function.

**Definition 3.6** ($h^*$-polynomial)**.**

$$\sum_{t\geq 0} L(P,t)x^t = \frac{h_0^* + h_1^* x + \cdots + h_d^* x^d}{(1-x)^{d+1}}.$$

The vector of coefficients $(h_0^*, h_1^*, \ldots, h_d^*)$ is called the $h^*$-vector of the polytope. In 1980, Richard Stanley proved that for any valid lattice polytope, these $h^*$ coefficients are always non-negative integers. If a polytope has a unimodular triangulation, these $h^*$ numbers carry a geometric meaning; each $h_i^*$ counts the number of simplices in the triangulation satisfying different conditions. However, moving from the guaranteed non-negativity of the $h^*$-vector back into the standard coefficients $c_i$ of the Ehrhart polynomial is difficult. Because the transformation between the two sets of coefficients involves alternating signs, a polytope can have positive $h^*$ numbers while still having negative coefficients in its standard Ehrhart polynomial. We can identify families of polytopes where the underlying geometry retains the positivity of the Ehrhart coefficients.

## 4. Zonotopes and Vector Configurations

While finding a clean unimodular triangulation for an arbitrary polytope can be difficult, there exists a very symmetric class of shapes known as *zonotopes*. Intuitively, a zonotope is a multi-dimensional shadow of a hypercube. We construct one by starting with a collection of vectors and taking a Minkowski sum.

**Definition 4.1** (Minkowski Sum)**.**
Let $A$ and $B$ be two subsets of $\mathbb{R}^d$. The Minkowski sum of $A$ and $B$, denoted by $A+B$, is the set formed by adding every vector in $A$ to every vector in $B$:

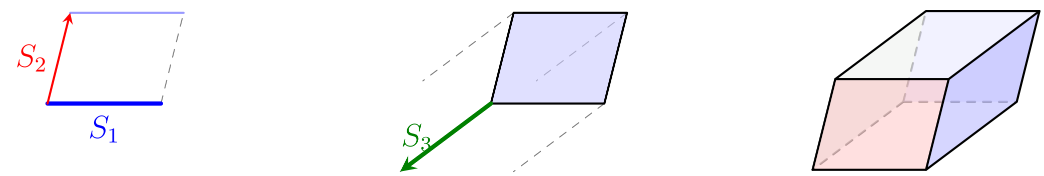
$$A + B = \{a + b \mid a \in A,\, b \in B\}.$$

**Definition 4.2** (Zonotope)**.**
Let $V = \{v_1, v_2, \ldots, v_n\}$ be a finite set of vectors in $\mathbb{R}^d$. For each vector $v_i$, let $I_i = \{\lambda v_i \mid 0 \leq \lambda \leq 1\}$ denote the line segment connecting the origin to $v_i$. The zonotope $Z(V)$ generated by $V$ is the Minkowski sum of these segments:

$$Z(V) = I_1 + I_2 + \cdots + I_n = \sum_{i=1}^{n} I_i.$$

For instance, if one takes two line segments (Fig. 4a) whose extensions intersect (essentially, the segments being non-parallel) in a plane and sweep one along the path of the other, they trace out a parallelogram (Fig. 4b). If they add a third line segment out of the plane and sweep the parallelogram along it, they form a three-dimensional parallelepiped (Fig. 4c) such as the following.

Because zonotopes are constructed from flat vector segments, they have an inherent *tiling property*. Instead of forcing a triangulation, we can naturally decompose a zonotope into a collection of smaller blocks called *parallelotopes* (see Definition 6.1). If the initial generating vectors in $V$ are chosen from the integer grid, these parallelotopes align with our lattice, giving us a controlled way to study the point counting behavior.




**1. Line + Line**

**(a)** Sweeping $S_1$ along $S_2$

**2. Face + Line**

**(b)** Sweeping face along $S_3$

**3. Parallelepiped**

**(c)** Resulting 3D Solid

**Figure 4.** Minkowski sum expansion ($S_1 + S_2 + S_3$) transitioning from 1D, to a 2D plane, to a skewed 3D Parallelepiped.

## 5. Ehrhart Positivity

We now come back to the mystery highlighted in the introduction: *Ehrhart positivity*. We say a lattice polytope $P$ is Ehrhart positive if every single coefficient $c_i$ in its Ehrhart polynomial $L(P,t)$ is strictly greater than zero ($c_i > 0$ for all $0 \leq i \leq d$). As established by geometry, the constant term ($c_0 = 1$), the surface area component ($c_{d-1}$), and the volume term ($c_d$) are safe from being negative. The true danger lies in the intermediate, hidden coefficients like $c_1$ or $c_2$. In general dimensions, if a polytope is squeezed too thin or has a skewed shape, these middle coefficients can turn negative.

*Example* (The Unit Hypercube).
The most fundamental example of this phenomenon is the standard $d$-dimensional unit hypercube $Q_d = [0,1]^d$. This polytope is a basic zonotope generated by the standard coordinate axis vectors $\{e_1, e_2, \ldots, e_d\}$. Dilating this hypercube by an integer factor $t$ gives a larger hypercube of side length $t$. The number of grid points contained within it is trivial to compute:

$$L(Q_d, t) = (t+1)^d = \sum_{i=0}^{d} \binom{d}{i} t^i.$$

Since the binomial coefficients $\binom{d}{i}$ are always positive integers, the hypercube gives an example of Ehrhart positivity. As one may expect, a hypercube can be neatly sliced into $d!$ identical unimodular simplices.

Beyond simple hypercubes, this positivity extends to more intricate objects. For example, consider the family of *permutohedra*—polytopes whose vertices are generated by shuffling the coordinates of a given starting vector. Permutohedra are famous examples of zonotopes that naturally split into unimodular pieces. The permutohedron was first systematically studied in higher dimensions by Dutch geometer Pieter Hendrik Schoute in 1911, who analyzed it as an omnitruncated simplex whose vertices are coordinates of permutations. The modern term "permutoèdre" was later coined in 1963 by French mathematicians Georges Th. Guilbaud and Pierre Rosenstiehl in a paper studying the geometry of voting preferences.

**Definition 5.1** (Permutohedra)**.**
The standard $(d-1)$-dimensional permutohedron $\Pi_{d-1} \subset \mathbb{R}^d$ is defined as the convex hull of all vectors obtained by permuting the coordinates of the vector $(1, 2, \ldots, d)$. This polytope can be represented as a zonotope generated by the set of all root vectors of the form $e_i - e_j$ for $1 \leq i < j \leq d$, where $e_i$ denotes the standard basis vector.

*Example.* Let us construct the 2-dimensional permutohedron $\Pi_2$ in $\mathbb{R}^3$ which is a regular hexagon (Fig. 5) by permuting the coordinates of the vector $(1, 2, 3)$. Shuffling these coordinates gives $3! = 6$ distinct vertices:

$$v_1 = (1, 2, 3), \quad v_2 = (1, 3, 2), \quad v_3 = (2, 1, 3),$$

$$v_4 = (2, 3, 1), \quad v_5 = (3, 1, 2), \quad v_6 = (3, 2, 1)$$

These vertices form a regular hexagon residing on the plane $x_1 + x_2 + x_3 = 6$. To see how the edges correspond to standard root vectors, examine the edge connecting $v_1 = (1, 2, 3)$ and $v_2 = (1, 3, 2)$:

$$v_2 - v_1 = (1 - 1,\ 3 - 2,\ 2 - 3) = (0, 1, -1) = e_2 - e_3$$

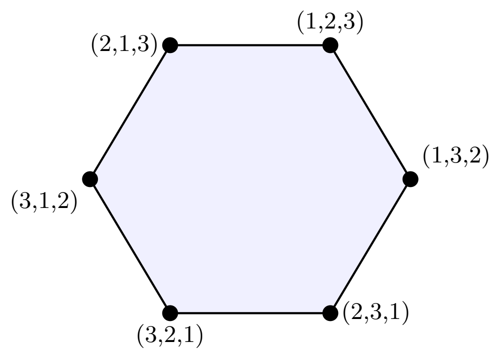



**Figure 5.** The 2D hexagonal permutohedron embedded in 3D space, generated by permuting the coordinates of the vector (1,2,3).

Because the generating configuration consists of a fully unimodular system of vectors, every linearly independent subset of these generators corresponds to a lattice parallelotope with a normalized volume equal to 1. Consequently, the natural parallelotope decomposition of the permutohedron gives components that behave well relative to the integer lattice. This ensures that the intermediate coefficients $c_1, c_2, \ldots, c_{d-2}$ are nonnegative, establishing Ehrhart positivity of the permutohedron across all dimensions.

The property shared by hypercubes and permutohedra suggests a broader question: whether all lattice zonotopes whose generating vectors form a totally unimodular matrix necessarily have Ehrhart positivity. We must determine the relation between the algebraic data of the matrix $V$ and the coefficients of the Ehrhart polynomial $L(Z(V), t)$. When the generating matrix $V$ is unimodular, every linearly independent subset of $V$ forms a basis for the sublattice it spans. This simplifies the problem because the normalized volume of every minimal parallelotope in the decomposition of $Z(V)$ is 1.

## 6. Parallelotope Decompositions and Totally Unimodular Configurations

To analyze the bigger relationship between the generating vector configuration $V$ and the coefficients of $L(Z(V), t)$, we must formalize the tiling properties of zonotopes. Because a zonotope $Z(V)$ is constructed via the Minkowski sum of line segments, it possesses an inherent property that allows it to be broken down into simpler blocks without resorting to a standard simplicial triangulation.

6.1. **Parallelotope Decompositions.** Instead of subdividing the geometric space into simplices, a zonotope can be cleanly partitioned into a collection of smaller, perfectly fitting regions called parallelotopes.

**Definition 6.1.** Parallelotopes:
Let $V = \{v_1, v_2, \dots, v_n\} \subset \mathbb{Z}^d$ be the set of generating lattice vectors. For any subset of $d$ linearly independent vectors $S = \{v_{i_1}, v_{i_2}, \dots, v_{i_d}\} \subseteq V$, we define the corresponding $d$-dimensional parallelotope $\square(S)$ by scaling each vector between 0 and 1:

$$\square(S) = \left\{ \sum_{j=1}^{d} \lambda_j v_{i_j} \;\middle|\; 0 \leq \lambda_j \leq 1 \right\}.$$

Each independent subset $S$ spans a local coordinate block whose volume is determined by the absolute value of the determinant of the matrix formed by these vectors.

**Theorem 6.1** (McMullen's Theorem on Zonotopal Tilings)**.**
*The global zonotope $Z(V)$ can be expressed as a disjoint union of the interiors of shifted copies of these parallelotopes:*

$$Z(V) = \coprod_{S \subseteq V,\, |S|=d} (c(S) + \mathit{int}(\square(S)))$$

*Here, $c(S)$ represents a suitable shifting vector chosen from the lattice $\mathbb{Z}^d$ to ensure that the pieces tile without overlapping gaps.*

This decomposition simplifies the lattice point counting problem: By counting the integer points contained within each individual parallelotope piece, the global Ehrhart polynomial $L(Z(V), t)$ can be computed as the sum of the polynomials of its constituent components. Crucially, when the vertices of these component parallelotopes do not align with the integer grid, fractional spaces occur. These fractional spaces algebraically distort the intermediate coefficients of the final polynomial, often creating the conditions under which negative terms arise. To guarantee that the parallelotope decomposition aligns with the underlying grid, we must restrict the generating configuration $V$. This brings us to the concept of total unimodularity, which provides the condition needed to preserve positivity.

**Definition 6.2** (Totally Unimodular Matrix)**.**
A matrix $M$ with integer entries is said to be *totally unimodular* if the determinant of every square submatrix of $M$ is equal to 0, 1, or $-1$.

In the context of Ehrhart theory, we construct a matrix whose columns are the generating vectors of our configuration $V$. When the generating matrix is totally unimodular, any subset of $d$ linearly independent vectors forms a basis for the entire integer sublattice it intersects. Then, the absolute value of the determinant of any such basis matrix is equal to 1. This uniform behavior has direct consequences for our shape: It implies that every minimal parallelotope $\square(S)$ in the decomposition acts as a fundamental unit of volume relative to the lattice $\mathbb{Z}^d$. Because every building block has a normalized volume of 1, no fractional shifts or grid misalignments can occur when the shape is dilated by an integer factor $t$. This puts constraints on the $h^*$-vector, preventing the alternating signs of the transformation function from driving the intermediate standard coefficients $c_i$ below zero.

## 7. The $h^*$-Vector and Its Applications

In this section, we explain how total unimodularity guarantees that the coefficients $c_i$ of the Ehrhart polynomial $L(P,t) = c_d t^d + \cdots + c_1 t + c_0$ are positive, using the connection between these coefficients and the polynomial's generating function.

Instead of working with the polynomial $L(P,t)$ by itself, it is useful to collect all the values of $L(P,t)$ for $t = 0, 1, 2, \ldots$ into an infinite series using a formal variable $x$. This series is called the Ehrhart series of the polytope $P$. By a theorem of Stanley, this infinite sum can always be simplified into a rational function where the denominator is $(1-x)^{d+1}$ and the numerator is a polynomial of degree at most $d$:

$$\sum_{t \geq 0} L(P,t)x^t = \frac{h_0^* + h_1^* x + h_2^* x^2 + \cdots + h_d^* x^d}{(1-x)^{d+1}}$$

**Definition 7.1** (Stirling Numbers of the First Kind)**.**
The *unsigned Stirling numbers of the first kind*, denoted by $\begin{bmatrix} n \\ k \end{bmatrix}$ or $c(n,k)$, count the number of permutations of $n$ elements with exactly $k$ disjoint cycles. They are defined as the coefficients of the rising factorial:

$$x^{\overline{n}} = x(x+1)\cdots(x+n-1) = \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix} x^k$$

Alternatively, the *signed Stirling numbers of the first kind*, denoted by $s(n,k)$, are the coefficients of the falling factorial:

$$x^{\underline{n}} = x(x-1)\cdots(x-n+1) = \sum_{k=0}^{n} s(n,k)x^k$$

where $s(n,k) = (-1)^{n-k} \begin{bmatrix} n \\ k \end{bmatrix}$.

The coefficients of the numerator polynomial, $(h_0^*, h_1^*, \ldots, h_d^*)$, form the $h^*$-vector of $P$. To find a formula for $L(P,t)$ in terms of these $h_j^*$ values, we can rewrite the fraction by expanding the denominator as an infinite series using the binomial theorem for negative exponents:

$$\frac{1}{(1-x)^{d+1}} = \sum_{k \geq 0} \binom{k+d}{d} x^k$$

Substituting this series back into the generating function equation allows us to match the coefficients for each power of $x^t$. This gives an expression for the point-counting function as a weighted sum of binomial coefficients:

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

To isolate the standard coefficients $c_i$ from this expression, we expand the binomial coefficient $\binom{t+d-j}{d}$ into a standard polynomial form. By definition, this binomial coefficient is a product of $d$ linear terms divided by $d!$:

$$\binom{t+d-j}{d} = \frac{(t+d-j)(t+d-j-1)\cdots(t-j+1)}{d!}$$

We can expand this product into a sum of powers of $t$ by using the signed Stirling numbers of the first kind. The Stirling number, denoted $\left[{d \atop k}\right]$, is a fixed integer that describes how the coefficients behave when expanding products of consecutive linear terms. Applying this definition and rearranging the terms by each power $t^i$ yields:

$$\binom{t+d-j}{d} = \frac{1}{d!}\sum_{i=0}^{d}\left(\sum_{k=i}^{d}\left[{d \atop k}\right]\binom{k}{i}(d-j)^{k-i}\right)t^i$$

Substituting this expanded form back into our formula for $L(P,t)$ and collecting the terms for each power $t^i$ gives the formula for each standard coefficient $c_i$:

$$c_i = \frac{1}{d!}\sum_{j=0}^{d} h_j^*\left(\sum_{k=i}^{d}\left[{d \atop k}\right]\binom{k}{i}(d-j)^{k-i}\right)$$

The $h^*$-vector gives a nonnegative combinatorial representation of the Ehrhart polynomial, but nonnegativity of $h_j^*$ does not automatically imply positivity of the coefficients in the standard power basis. We can now see how total unimodularity prevents negative coefficients from happening. When the generating configuration $V$ of a zonotope $Z(V)$ is totally unimodular, the tiling theorem decomposes the shape into parallelotopes where every single piece has a volume of exactly 1. For a parallelotope with volume 1, the counting function simplifies well because it contains no interior lattice points during dilation. This means its local $h^*$-vector is concentrated entirely at the beginning, where $h_0^* = 1$ and all other terms $h_j^* = 0$ for $j \geq 1$.

When these local vectors are summed together to find the $h^*$-vector of the complete zonotope, the resulting global $h^*$-vector is heavily weighted toward the early indices where $j$ is small. Because the indices $j$ stay small, the term $(d-j)$ remains strictly positive throughout the expansion. This prevents the negative signs from occurring in the inner sum, mathematically ensuring that the standard coefficients $c_i$ remain strictly greater than zero for all dimensions $0 \leq i \leq d$.

## 8. The Roots in the Complex Plane

The proof that a totally unimodular configuration guarantees positive coefficients $c_i > 0$ has a direct effect on where the Ehrhart polynomial's zeroes are. For a general lattice polytope, the positions of its roots in the complex plane can appear chaotic. However, when total unimodularity restricts the coefficients to positive real numbers, the roots are forced into some conditions. We can see this first on the real number line. Since every coefficient $c_i$ of the Ehrhart polynomial is strictly positive, it is impossible for the polynomial to have any non-negative real roots. If we plug in any real number $t \geq 0$, every single term $c_i t^i$ is non-negative, and the constant term $c_0 = 1$ ensures the total sum is at least 1. Therefore, the polynomial can never equal zero for any positive scale factor. This means all real roots of a totally unimodular zonotope must be strictly negative.

When we look at the complex roots, the front-heavy nature of the $h^*$-vector introduces a pattern. Because total unimodularity concentrates the $h^*$-vector at the lower indices, the standard coefficients $c_i$ grow and decay at a predictable rate as the degree increases. The distribution of a polynomial's roots is determined by the relative ratios of its consecutive coefficients. Because the ratios are bounded by the unit-volume geometry, the complex roots cannot wander arbitrarily far into space. Instead, they are forced to form a tight, symmetric cluster around the vertical line $\mathrm{Re}(t) = -1/2$ for reflexive polytopes. This clustering implies that the roots of a totally unimodular zonotope offer a a bound for the shape's dimensionality and structure.

This gives us a method to detect when a shape violates the lattice grid. If a polytope cannot be tiled because its building blocks have a volume greater than 1, the resulting shift in its $h^*$-vector alters the coefficient ratios. If an intermediate coefficient drops below zero, the alternating signs break the barrier on the positive real axis. The appearance of a positive real root or a scattered, unstable cloud of complex roots acts as a "mathematical signature", immediately signaling that the shape contains grid misalignments and fractional spaces violating total unimodularity.

## 9. Order Polytopes of Finite Posets

Another way to see how combinatorial rules shape geometric polytopes is by looking at partially ordered sets, called *posets*. A poset is a set of elements where some items are ranked or ordered relative to one another, while other pairs might be completely incomparable. Following a construction by Richard Stanley [2], we can translate these rankings directly into an *order polytope*. To build an order polytope for a poset with $d$ elements, we assign one variable (or coordinate axis) to each element. If our poset has elements labeled $1, 2, \ldots, d$, we look at points $(x_1, x_2, \ldots, x_d)$ in $d$-dimensional space. We then restrict this space using two rules: every coordinate must sit between 0 and 1, and if element $i$ is smaller than or equal to element $j$ in our poset, its coordinate must be less than or equal to $j$'s coordinate $(x_i \leq x_j)$. The resulting shape is a bounded, flat-faced geometric object. Because the vertices of this polytope turn out to be made entirely of 0s and 1s, it forms an integral lattice polytope.

When we dilate this order polytope by a whole number factor $t$, we are looking for all the lattice points inside the stretched shape $t\mathcal{O}(P)$. Geometrically, this means our coordinates must now fall between 0 and $t$, while still obeying the same matching inequalities $(x_i \leq x_j)$ dictated by our poset. In combinatorics, assigning a whole number to each element of a poset in a way that respects the ordering is called an order-preserving map. The number of ways to do this is a known formula called the order polynomial. Because counting these integer points is exactly what an Ehrhart polynomial does, the Ehrhart polynomial of an order polytope is equivalent to this order polynomial.

To see this visually, let us look at a simple example with just two elements, $\{1, 2\}$, where element 1 is smaller than element 2. The order polytope lives in a two-dimensional plane with axes $x_1$ and $x_2$. Our rules state that $0 \leq x_1 \leq 1$, $0 \leq x_2 \leq 1$, and $x_1 \leq x_2$. If we plot these boundaries on a grid, they slice out a regular right triangle with vertices at $(0, 0)$, $(0, 1)$, and $(1, 1)$. If we dilate this triangle by an integer factor $t$, the new boundaries become $0 \leq x_1 \leq x_2 \leq t$. Counting the lattice points inside this stretched triangle is a matter of listing pairs of integers where the second number is at least as big as the first. For a dilation of $t = 1$, we find 3 points: $(0, 0)$, $(0, 1)$, and $(1, 1)$. For a dilation of $t = 2$, we find 6 points: $(0, 0)$, $(0, 1)$, $(0, 2)$, $(1, 1)$, $(1, 2)$, and $(2, 2)$. Following this pattern, the total number of points for any dilation $t$ can be counted using a basic combinations-with-repetition formula:

$$L(\mathcal{O}(P_2), t) = \frac{1}{2}t^2 + \frac{3}{2}t + 1. \tag{9.1}$$

Just like our early examples of hypercubes, every single coefficient in this polynomial ($c_2 = \frac{1}{2}$, $c_1 = \frac{3}{2}$, and $c_0 = 1$) is strictly positive.

The reason order polytopes are important for Ehrhart positivity in lower dimensions is due to how their $h^*$-vectors behave. The coefficients in the $h^*$-vector count the number of ways to arrange the poset into a strict line, called a *linear extension*, based on how many times the sequence steps downward. For smaller posets with 13 elements or fewer, the balanced distribution of positive values is not affected too significantly by algebraic grid shifts caused by the Stirling numbers, ensuring that the final Ehrhart polynomial has positive coefficients.

However, when we move to spaces with 14 or more dimensions, this breaks down, as proved by Stanley. For complicated branching posets, the geometry of the order polytope becomes skewed, causing the values in the $h^*$-vector to compile at one end. When these counts are passed through the sign-alternating formulas used to find the standard coefficients, the negative terms exceed the positive volume, causing coefficients to drop below zero.

## 10. The Geometry of Magic Squares

To see how Ehrhart theory can unexpectedly solve classical puzzles, we can look toward a topic many mathematicians are familiar with: magic squares. A magic square is a grid of numbers where every single row, column, and main diagonal adds up to the exact same "magic" constant. Arranging these numbers can be solved as a logic puzzle, but we can turn the rules of a magic square into geometry, and then use lattice points to count how many puzzles exist.

Let us look at a simple example using a $3 \times 3$ grid. For the purpose of simplicity, we will look at semi-magic squares, where we only require the rows and columns to add up to our constant, ignoring the diagonals for a moment. We can treat each of the nine squares in the grid as a coordinate axis in a nine-dimensional space: $(x_1, x_2, \ldots, x_9)$. To build our shape, we set up two sets of rules. First, because no slot can have a negative value, every coordinate must be greater than or equal to zero. Second, we write out our row and column rules as equations. For example, if the first row consists of the first three slots, we write $x_1 + x_2 + x_3 = 1$. If we do this for every row and column, these flat equations slice through our nine-dimensional space and carve out a bounded shape called the Birkhoff polytope. Stretching this shape by a whole-number factor $t$ is where more "magic" happens. Dilating the Birkhoff polytope by $t$ changes our equations so that every row and column must now add up to $t$. Counting the lattice points inside this stretched shape means finding every possible arrangement of non-negative integers that forms a valid semi-magic square with a magic constant of $t$. Since counting the lattice points is what the Ehrhart polynomial does, the polynomial for this shape acts as a formula that tells us the number of magic squares that exist for any target sum. For our $3 \times 3$ grid, counting these points across different dilations reveals a pattern. If the target sum is $t = 0$, there is only 1 solution: a grid filled entirely with zeros. If the target sum is $t = 1$, it turns out there are 6 solutions. If we calculate the total number of points for any target sum $t$, the resulting Ehrhart polynomial expands to:

$$L(P,t) = \frac{3}{8}t^4 + \frac{3}{4}t^3 + \frac{9}{8}t^2 + \frac{3}{4}t + 1. \tag{10.1}$$

When we look closely at this polynomial, its coefficients ($c_4 = \frac{3}{8}$, $c_3 = \frac{3}{4}$, $c_2 = \frac{9}{8}$, $c_1 = \frac{3}{4}$, and $c_0 = 1$) are all positive. Even though this shape lives in a high-dimensional space and is made from flat slices rather than solid blocks, its balance prevents any negative numbers from showing up.This connection shows that Ehrhart theory is not only about measuring distances or volumes. It is also a tool that allows us to take a discrete counting puzzle and stretch it out into a continuous geometric space to find an answer.

## 11. Cross-Polytopes and Diagonal Grid Cuts

When we evaluated hypercubes, the lattice point counting was straightforward because the boundaries of the shape ran parallel to our coordinate axes. To see how changing the orientation of a shape affects our counting formulas, we can examine a family of shapes known as *cross-polytopes*. In two dimensions, a cross-polytope is a square rotated 45°, forming a diamond (Fig. 6 left). In three dimensions, it forms a regular octahedron (Fig. 6 right). This family allows us to study shapes defined by diagonal boundaries that cut across the grid lines.

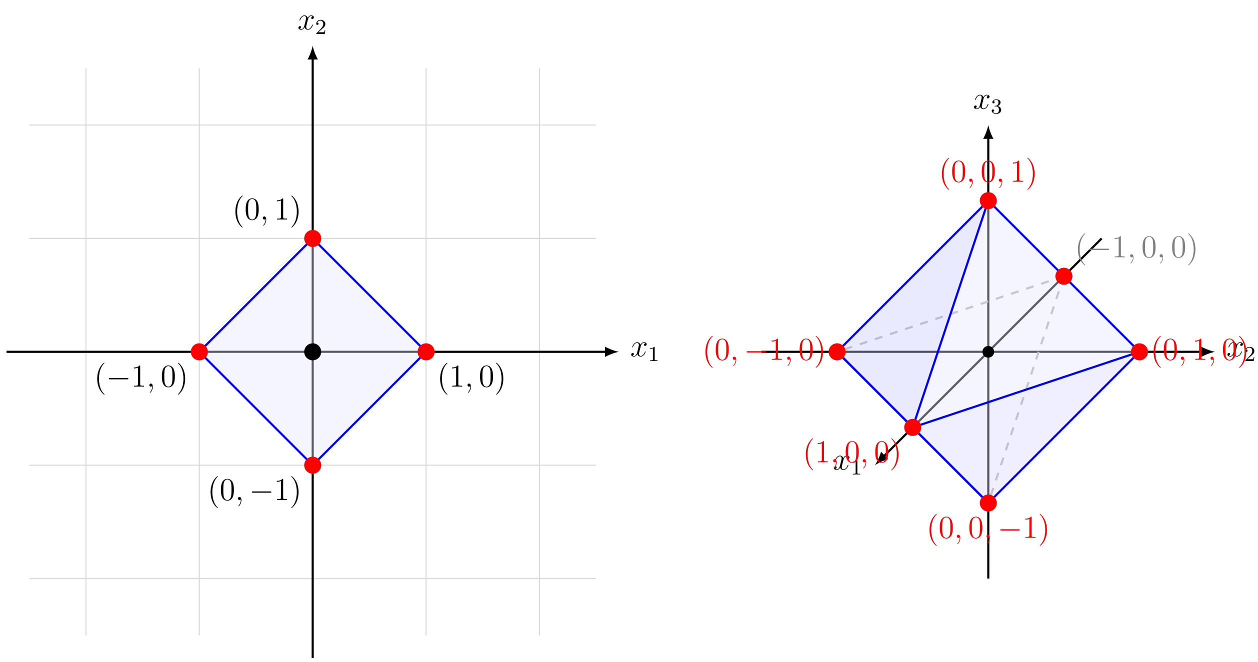



**Figure 6.** The 2D cross-polytope $\Diamond_2$ (left) and 3D cross-polytope $\Diamond_3$ (right).

We can define a standard $d$-dimensional cross-polytope $\Diamond_d$ using a single inequality based on absolute values. The shape consists of all points $(x_1, x_2, \ldots, x_d)$ in space such that:

$$|x_1| + |x_2| + \cdots + |x_d| \le 1. \tag{11.1}$$

The sharp corners (vertices) of this shape sit exactly at the coordinate points $(1, 0, \ldots, 0)$, $(-1, 0, \ldots, 0)$, and their rearrangements across the other axes. Because these corners rest entirely on integer grid intersections, $\Diamond_d$ is an integral lattice polytope.

When we dilate the cross-polytope by a whole-number factor $t$, the boundary condition expands to follow the same rule:

$$|x_1| + |x_2| + \cdots + |x_d| \le t. \tag{11.2}$$

Counting the integer points inside this stretched diamond becomes a sorting problem. For any fixed total sum $k$ (where $0 \le k \le t$), we have to choose which coordinates will be non-zero, assign positive or negative signs to those active slots, and distribute the remaining value among them. Because each non-zero coordinate has two options (positive or negative), the formulas naturally involve powers of 2 alongside standard combination choices.

For a two-dimensional diamond $\Diamond_2$, counting these points across different dilations gives a clean quadratic pattern. When $t = 1$, we find 5 points: the origin $(0, 0)$ and the four sharp corners $(1, 0)$, $(-1, 0)$, $(0, 1)$, and $(0, -1)$. When $t = 2$, the boundary expands to include points like $(1, 1)$ and $(2, 0)$, giving a total of 13 points. If we compute the total number of points for any dilation factor $t$, the resulting Ehrhart polynomial expands to:

$$L(\Diamond_2, t) = 2t^2 + 2t + 1. \tag{11.3}$$

Evaluating the coefficients of this polynomial gives us $c_2 = 2$, $c_1 = 2$, and $c_0 = 1$. Just like the standard hypercube, every single coefficient is positive. In fact, because the underlying combinatorial structure of a cross-polytope relies on adding together strictly positive combinations of choices, its Ehrhart polynomial is guaranteed to have positive coefficients in all dimensions. Though a cross-polytope tilts across the grid lines, its reflectional symmetry across the coordinate axes keeps the lattice point counts well balanced, showing that diagonal cuts do not inherently disrupt positivity.

## 12. Beyond Total Unimodularity: Matroids and Non-Unimodular Zonotopes

The previous section establishes that the standard coefficients $c_i$ of an Ehrhart polynomial are strictly positive whenever the generating configuration $V$ of a zonotope $Z(V)$ is totally unimodular. This uniform behavior depends on the restriction that every parallelotope in the decomposition has a normalized volume of 1. A natural follow-up question is whether total unimodularity is a necessary condition for Ehrhart positivity, or simply a sufficient one.

We can look to an answer by introducing a vector configuration that violates total unimodularity. Suppose a single square submatrix yields a determinant of 2 instead of $1, 0$, or $-1$. In our tiling theorem, this shift disrupts the lattice grid alignment, forcing at least one parallelotope $\square(S)$ to have a normalized volume greater than 1.

Dilating this shape by an integer factor $t$ introduces fractional spaces relative to the integer grid within the non-unimodular blocks. This distorts the local $h^*$-vector. The weight of the $h^*$-vector shifts away from the initial index ($h_0^* = 1$) toward the later indices, inflating terms like $h_{d-1}^*$ or $h_d^*$. Substituting these values into our formula for the standard coefficients,

$$c_i = \frac{1}{d!} \sum_{j=0}^{d} h_j^* \left( \sum_{k=i}^{d} \begin{bmatrix} d \\ k \end{bmatrix} \binom{k}{i} (d-j)^{k-i} \right)$$

reveals the consequences of this shift. When $h_j^*$ is large for indices where $j > d$, the base term $(d - j)$ becomes negative. Raised to odd powers, these terms put negative values into the summation. Because total unimodularity no longer suppresses these later $h_j^*$ terms, the internal negative signs can overwhelm the early positive terms. This often causes intermediate coefficients to drop below zero. Stepping outside the unimodular case consequently causes negative coefficients.

### 12.1. **Matroids.**

Arbitrary violations of total unimodularity frequently break Ehrhart positivity, yet a large class of objects retain positive coefficients despite lacking a fully unimodular structure. Formalizing these shapes involves turning our understanding of linear independence into the idea of *matroids*.

**Definition 12.1** (Matroid)**.**
A matroid $M = (E, \mathcal{I})$ is a combinatorial structure consisting of a finite set of elements $E$ (the ground set) and a collection $\mathcal{I}$ of subsets of $E$ (the independent sets) satisfying three conditions:

(1) The empty set is independent ($\emptyset \in \mathcal{I}$).
(2) If a set is independent, all of its subsets are also independent (if $A \in \mathcal{I}$ and $B \subset A$, then $B \in \mathcal{I}$).
(3) If two independent sets have different sizes, the smaller set can be expanded by taking an element from the larger set while remaining independent (if $A, B \in \mathcal{I}$ and $|A| > |B|$, there exists an element $x \in A \setminus B$ such that $B \cup \{x\} \in \mathcal{I}$).

The maximal independent sets in a matroid are called its *bases*.

A matroid serves as a generalized template for linear independence. When the ground set $E$ consists of the columns of a matrix, and $\mathcal{I}$ contains the subsets of columns that are linearly independent in the traditional sense, the structure forms a *representable matroid.* However, matroids also model the independence found in the edges of a graph, where independent sets are collections of edges containing no cycles.

12.2. **Matroid Polytopes and the Generalized Positivity Conjecture.** Much like a configuration of vectors constructs a geometric zonotope via a Minkowski sum, the bases of a matroid define a unique lattice polytope.

**Definition 12.2** (Matroid Polytope)**.**
Let $M$ be a matroid on the ground set $E = \{1, 2, \dots, n\}$ with a collection of bases $\mathcal{B}$. For each basis $B \in \mathcal{B}$, let $e_B \in \mathbb{R}^n$ be the *characteristic vector* where the $i$-th coordinate is 1 if $i \in B$ and 0 if $i \notin B$. The matroid polytope $P_M$ is the convex hull of these characteristic vectors:

$$P_M = \operatorname{conv}\{e_B \in \mathbb{R}^n \mid B \in \mathcal{B}\}$$

Matroid polytopes are an active area of study of Ehrhart theory. Because they are constructed from standard binary vectors, their vertices are always bound to the integer lattice. Yet, unlike the hypercubes and permutohedra explored in Section 5, the underlying combinatorial structures of general matroids do not guarantee a totally unimodular matrix representation. As a result, their local parallelotope or simplicial decompositions often produce blocks with volumes strictly greater than 1. Despite these grid misalignments and the resulting shifts in their $h^*$-vectors, evidence shows that the intermediate coefficients of their Ehrhart polynomials remain positive. This observation leads to a major open problem in modern discrete geometry, originally formulated by De Loera, Haws, and Köppe:

**Conjecture 12.1** (The Generalized Ehrhart Positivity Conjecture)**.**
*Let $M$ be any matroid. Then its corresponding matroid polytope $P_M$ is Ehrhart positive. That is, every coefficient $c_i$ of the polynomial $L(P_M, t)$ is strictly greater than zero.*

The mystery of this conjecture lies in the contrast it presents with our previous results. For zonotopes, Ehrhart positivity is tied to the local geometry of unit volumes and total unimodularity. For matroid polytopes, total unimodularity fails, meaning the algebraic cancellation of negative signs in the Stirling expansions cannot be prevented by local volume constraints alone. Instead, a global combinatorial structure must be ensuring that even when fractional spaces distort the $h^*$-vector, they do so in a restricted manner that preserves overall positivity across all the dimensions. Despite some intuition supporting the conjecture, the hypothesis was ultimately disproven in a breakthrough result by Luis Ferroni [6]. For nearly fifteen years, the conjecture stood as a major open problem in discrete geometry, verified for extensive families such as uniform matroids and hypersimplices. This validated the idea that a global combinatorial mechanism protected the coefficients from negativity. However, Ferroni demonstrated that this breaks down in sufficiently high dimensions by providing explicit counterexamples within the class of sparse paving matroids.

Ferroni proved that for every rank $k \geq 3$, there exists a sparse paving matroid whose base polytope possesses at least one negative coefficient in its Ehrhart polynomial. The failure of the conjecture shows the limits of relying on combinatorial constraints to guarantee positivity. When these asymmetric matroid structures are dilated, the distribution of internal lattice points shifts. This yields an $h^*$-vector that eventually follows the sign-alternating Stirling numbers of the first kind during the algebraic transformation, pulling the intermediate coefficients below zero. The disproof of De Loera, Haws, and Köppe's conjecture has reoriented the path of modern research in the field. Rather than seeking a universal proof of positivity for all matroids, researchers have shifted toward mapping the boundary where Ehrhart positivity holds. This has led to an active frontier in combinatorics, including recent successful proofs of Ehrhart positivity for large classes of matroids, including lattice path matroids, Schubert matroids, and positroids.

To conclude, Ehrhart theory shows how discrete mathematics and geometry are fundamentally connected. Though the questions around coefficient positivity prove that higher-dimensional shapes are complex, finding the exact boundaries of that positivity remains a great open problem in the field.

## 13. Acknowledgements

The author would like to thank Dr. Simon Rubinstein-Salzedo for giving the opportunity to write this paper and Elsa Frankel for advice and support throughout this project. The author would also like to express gratitude to peers and classmates whose helpful discussions and suggestions contributed to the refinement of this paper.

## References

1. Federico Ardilo, Mariel Supina, and Andrés R. Vindas-Meléndez, *The equivariant ehrhart theory of the permutahedron.*
2. Matthias Beck, *Stanley's major contributions to ehrhart theory.*
3. Matthias Beck and Sinai Robins, *Computing the continuous discretely*, Springer, 2007.
4. Federico Castillo and Fu Liu, *Ehrhart positivity for generalized permutohedra.*
5. Eugène Ehrhart, *Sur les polyèdres homothétiques bordés à n dimensions*, Comptes Rendus de l'Académie des Sciences **254** (1962), 988–990.
6. Luis Ferroni, *Matroids are not ehrhart positive.*
7. Georg Pick, *Geometrisches zur Zahlenlehre*, Sitzungsberichte des Deutschen Naturwissenschaftlich-Medizinischen Vereins für Böhmen "Lotos" in Prag **19** (1899), 311–319.
8. J. E. Reeve, *On the volume of lattice polyhedra*, Proceedings of the London Mathematical Society. Third Series **7** (1957), 378–395.