

# 1 Introduction

The Minkowski-Weyl Theorem is a theorem of utmost importance in the theory of polyhedra. Ziegler refers to it as the “main theorem.” It is a theorem which is easy to intuitively understand, but is surprisingly difficult to prove. Going through the proof in detail, however, is a good exercise because it introduces various other important notions and techniques, as well as helps to mentally clarify the basic definitions relevant to the theory.

Polyhedra are “finitely generated” convex subsets of  $\mathbb{R}^n$ . “Finitely generated” here means that they can be described in a finite manner, in particular as a convex hull of either: a finite set of points and ray, or a finite intersection of halfspaces. The Minkowski-Weyl Theorem states that these two finite representations are “equivalent,” that every polyhedra can be described either way.

For a more concrete example, consider a polygon in the plane (provide some nice pictures here).

In this paper, the theorem and relevant definitions will be stated, and a proof will be given. The proof will closely follow Ziegler’s, however it is the author’s intent to make this paper as accessible as possible (who is the target audience?). Therefore, some concepts will be stated in somewhat more painstaking detail than some may find necessary, however these details will have to be addressed to implement an enumeration-algorithm that will accompany the proof.

Perhaps the most important part of the paper is the diagram in figure 1. This diagram illustrates the main steps of the proof, and will serve as a “roadmap” to the implementation.

## 2 Definitions

### 2.1 Informal Notions

Here shall be stated some of the fundamental ideas involved in the theory of polyhedra (formal definitions will be given in the next section).

**Convexity** Convexity is an essential notion in the study of polyhedra. Here, it will be considered a property of subsets of  $\mathbb{R}^n$ . Let  $A \subseteq \mathbb{R}^n$ .  $A$  is *convex* if, given any two points of  $A$ , the line segment connecting these two points is also contained in  $A$ . It formalizes a notion of being “filled-in,” and implies that (given the standard metric) the shortest path between any two points of  $A$  is included in  $A$ .

**Convex Hulls** Associated with the notion of convexity is the operation of *convex hull*. Given any subset  $B \subseteq \mathbb{R}^n$ , the *convex hull* of  $B$  is the smallest set which contains  $B$  and is convex. For two isolated points, their convex hull is the line segment with the given points as endpoints. For more complex sets, it can be thought of as an iterative procedure where the convex hull of every

two points is added to the set, repeating this operation until no more points are added. (a picture showing how the procedure must iterate).

## Minkowski Sums

**Projections, Lifts, Relaxations, Restrictions** These terms will be used informally in this paper, however when one is needed in a formal context it will be specified as necessary.

*Projections* are very simple operations, which basically cause one to “ignore” the coordinates of a set for a given index set of coordinates. For example,  $\pi : (x, y) \mapsto (x)$  is a projection from  $\mathbb{R}^2$  to  $\mathbb{R}$ . Projections are often given a subscript to indicate which dimensions are being “projected to,” or a superscript to indicate which dimensions are being “projected away.” So  $\pi$  above may be written as  $\pi_x$  or perhaps  $\pi^y$ , however in formal contexts the precise nature of the projection will be indicated; this notation is meant only to be useful for the reader.

A *lift* refers to taking a set and representing it in a higher dimension. For example, one may take the unit interval  $I = [0, 1] \subseteq \mathbb{R}^1$  and *lift* it to  $\mathbb{R}^2$  by considering  $I \times 0$ .

*Relaxations and Restrictions:* A *relaxation* of a set will typically be a way to let a set “grow,” however it will typically be in a “controlled” or “reversible” manner. These “reversals” will be referred to as *restrictions*. Take, for instance, this relaxation of the unit interval  $I = [0, 1]$ :

$$\bigcup_{t \geq 0} [0, t] \times t = \{(x, y) \mid x \geq 0\} \cap \{(x, y) \mid y - x \geq 0\}$$

From this relaxation (an “infinite triangle” of sorts) we may recover  $I$  with a respective restriction, in this case by intersecting the relaxation with  $\{(x, y) \mid y = 1\}$  and projecting it to the x axis. Again, the terms *relaxation* and *restriction* will be used informally, while specific instances will be specified as needed.

## 2.2 Formal Definitions

Here are the formal definitions required for the statement of the theorem and its proof. In all definitions, unless otherwise stated,  $I$  represents an arbitrary finite index set, and  $J$  is an arbitrary index set.

**Convex Combination** Let  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ ,  $\lambda \in [0, 1]$ . Then

$$\lambda \mathbf{x} + (1 - \lambda) \mathbf{y}$$

is called a *convex combination* of  $\mathbf{x}$  and  $\mathbf{y}$ . More generally,  $\mathbf{x}_i \in \mathbb{R}^n$ ,  $\lambda_i \in [0, 1]$ , and  $\sum_i \lambda_i = 1$ . Then

$$\sum_i \lambda_i \mathbf{x}_i$$

is called a *convex combination* of  $\mathbf{x}_i$ .

**Convex** Let  $A \subseteq \mathbb{R}^n$ .  $A$  is by definition *convex* if every convex combination of its members is again a member. In symbols:

$$\mathbf{x}_i \in A \Rightarrow \sum_i \lambda_i \mathbf{x}_i \in A$$

Note: Let  $(\forall j \in J) A_j \subseteq \mathbb{R}^n$  be convex, and  $(\forall i \in I) \mathbf{x}_i \in \bigcap_{j \in J} A_j$ . Then

$$(\forall j \in J) \sum_i \lambda_i \mathbf{x}_i \in A_j \Rightarrow \sum_i \lambda_i \mathbf{x}_i \in \bigcap_{j \in J} A_j$$

This implies that  $\bigcap_{i \in I} A_i$  is also convex. In other words, the property of being convex is closed under the operation of intersection.

**Halfspace** Let  $\mathbf{y} \in \mathbb{R}^n$ ,  $c \in \mathbb{R}$ . Then the set

$$\{\mathbf{x} \mid \langle \mathbf{x}, \mathbf{y} \rangle \leq c\}$$

is called a *halfspace*. In particular,  $H_k$  shall denote the halfspace given by

$$\{\mathbf{x} \mid \langle \mathbf{x}, \mathbf{e}_k \rangle \leq 0\} = \{\mathbf{x} \mid x_k \leq 0\}$$

Note: Suppose that  $\langle \mathbf{x}, \mathbf{y} \rangle \leq c$ ,  $\langle \mathbf{z}, \mathbf{y} \rangle \leq c$ , and  $\lambda \in [0, 1]$ . Then

$$\langle \lambda \mathbf{x} + (1 - \lambda) \mathbf{z}, \mathbf{y} \rangle = \lambda \langle \mathbf{x}, \mathbf{y} \rangle + (1 - \lambda) \langle \mathbf{z}, \mathbf{y} \rangle \leq \lambda \cdot c + (1 - \lambda) \cdot c = c$$

This means that halfspaces are *convex*.

**H-polyhedron** An *H-polyhedron* is a finite intersection of halfspaces. By the remark given under the definition of **convexity**, *H-polyhedra are convex*. Here is a useful way of “writing down” an H-polyhedron:

$$\{\mathbf{x} \in \mathbb{R}^n \mid \forall i : \langle \mathbf{x}, \mathbf{a}_i \rangle \leq b_i\}$$

If all the  $\mathbf{a}_i$  are gathered into a matrix  $A$  (where the  $\mathbf{a}_i$  become the rows of  $A$ ), this is sometimes also written as:

$$\{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} \leq \mathbf{b}\}$$

**Convex Hull** Let  $A \subseteq \mathbb{R}^n$ .  $B$  is by definition the *convex-hull* of  $A$  if

$$B = \text{conv}(A) = \{\mathbf{x} \mid \exists \mathbf{a}_i \in A, \exists \lambda_i \in [0, 1] : \sum_i \lambda_i = 1, \mathbf{x} = \sum_i \lambda_i \mathbf{a}_i\}$$

Note: Another useful, equivalent definition of *convex hull* can be given as:

$$B' = \text{conv}'(A) = \bigcap \{C \mid A \subseteq C, C \text{ is convex}\}$$

This more clearly illustrates the idea of “smallest convex set containing  $A$ .” To see that these definitions are equivalent, first note that  $B$  is convex, and  $A \subseteq B$ , so  $B' \subseteq B$  (since  $B$  will be included in the intersection). Then note that if  $C$  is convex and  $A \subseteq C$ , then any convex combination of members of  $A$  must (by definition) be in  $C$ , so it will be in every such  $C$  and therefore in their intersection, so  $B' \subseteq B$ .

**Conical Hull** Let  $\cup_{i \in I} \mathbf{a}_i = A \subseteq \mathbb{R}^n$ , then  $B$  is by definition the *conical hull* of  $A$  if:

$$B = \text{cone}(A) = \{\mathbf{x} \in \mathbb{R}^n \mid \exists t_i \in \mathbb{R}^n \mid t_i \geq 0, \mathbf{x} = \sum_{i \in I} t_i \mathbf{a}_i\}$$

Note: A conical hull is convex. Let  $K$ ,  $(\forall i \in I) J_i$  be finite index sets (i.e.  $J_i$  is a collection of finite index sets, indexed by the finite set  $I$ ),  $(\forall i \in I)(j \in J_i) : t_{i,j} \geq 0$ ,  $\mathbf{a}_{i,j} \in A$ , and, as usual,  $\lambda_i \in [0, 1]$ ,  $\sum_i \lambda_i = 1$ . Then

$$\sum_i \lambda_i \left( \sum_{j \in J_i} t_{i,j} \mathbf{a}_{i,j} \right) = \sum_{i,j \in J_i} \lambda_i \cdot t_{i,j} \mathbf{a}_{i,j}$$

is a “conical combination” of vectors in  $A$  (because  $\lambda_i \cdot t_{i,j} \geq 0$ ), so every convex combination from  $\text{cone}(A)$  is again in  $\text{cone}(A)$ .

**Minkowski Sum** Let  $P, Q \subseteq \mathbb{R}^n$ . Then the *Minkowski Sum* of  $P$  and  $Q$  is given by:

$$P \oplus Q = \{p + q \mid p \in P, q \in Q\}$$

Note: If both  $P$  and  $Q$  are convex, then so is  $P \oplus Q$ . Indeed, let  $p_i \in P$ ,  $q_i \in Q$ , then

$$\sum_i \lambda_i (p_i + q_i) = \sum_i \lambda_i p_i + \sum_i \lambda_i q_i = \bar{p} + \bar{q}$$

Where  $\bar{p} \in P$  and  $\bar{q} \in Q$  (because they are convex), so  $\bar{p} + \bar{q} \in P \oplus Q$ .

**V-Polyhedron** A *V-Polyhedron* is a subset  $\mathcal{P}$  of  $\mathbb{R}^n$  given by

$$\mathcal{P} = \text{cone}(\mathcal{U}) \oplus \text{conv}(\mathcal{V})$$

where  $\mathcal{U}$  and  $\mathcal{V}$  are both finite (possibly empty) subsets of  $\mathbb{R}^n$ . By the remark given under the definition of Minkowski Sums, *V-Polyhedra are convex*.

We are now prepared to state the Minkowski-Weyl Theorem.

**Minkowski-Weyl Theorem** Every V-Polyhedron is an H-Polyhedron, and every H-Polyhedron is a V-Polyhedron. This permits the use of the term “Polyhedron” unambiguously.

Note: I’ve read that “definitions should be hard and theorems should be easy.” Perhaps there is a bit of this involved in the Minkowski-Weyl Theorem, but an unfortunate side effect is that the significance of a theorem may not be immediate. The Minkowski-Weyl Theorem is a “representation theorem,” and is a consequence of a type of duality that exists between the V and H –type definitions of polyhedra.

For a quick demonstration of why a dual representation is important for these types of sets, consider the following two tasks:

- Given a point  $\mathbf{x}$ , determine if it is a member of the set
- Produce an arbitrary member of the set

These two fundamental tasks tell completely different stories given one representation as opposed to the other. For an H-Polyhedron, determining membership is straightforward and relatively inexpensive: for each halfspace constraint, check to see that it is fulfilled. For a V-Polyhedron, you would need to either find a combination that produces the point, or provide a “certificate” that demonstrates that it is not in the set. On the other hand, given an H-Polyhedron, producing a member of the set is not straightforward at all. Given a list of half-spaces, it isn’t even obvious (or necessarily cheap to determine) that their intersection is not empty! A V-Polyhedron, on the other hand, is given by members of the set, so the task of providing a member (or determining if the set is empty or not) is already done for you. Observe that these tasks are facilitated by the qualifiers used to describe the sets (the qualifier  $\forall$  is used for the H-Polyhedron, while  $\exists$  is used for the V-Polyhedron).

It should be noted that guaranteeing the existence of two representations does not promise an efficient method of creating one from the other. The proof of the Minkowski-Weyl Theorem indicates an enumeration algorithm, and an implementation will be provided. This enumeration can be inefficient, however.

## 2.3 Equivalence of Polyhedra

The notion of equivalence is worth a small discussion. Consider the two polyhedra  $I = [0, 1]$  and  $I \times \{0\}$  (i.e. the unit interval in  $\mathbb{R}^1$  and the unit interval in  $\mathbb{R}^2$ ). It would be unfortunate if the theory of polyhedra didn’t consider these two entities as *equivalent* in some manner. The basic operation of equivalence in the theory is the *affine transformation*. Then polyhedra are considered equivalent if there is an affine transformation from one to the other, which is bijective *when restricted to the second polyhedron*. In our setting, it is enough to consider a subset of the affine transformations called *projections*. For example, one may write a projection  $\pi_x : (x, y) \mapsto (x)$ . Then we have  $\pi_x(I \times \{0\}) = I$ . Note that this projection, when restricted to  $I$  (and, more generally, the  $x$ -axis) is bijective. Therefore, given our notion of equivalence, we can conclude that  $I$  and  $I \times \{0\}$  are equivalent. As an anti-example, consider  $I \times I$  (i.e. the unit square in  $\mathbb{R}^2$ ). While  $\pi_x(I \times I) = I$ , this mapping is not bijective when restricted to  $I$ . Observe that  $(0, 0) \mapsto (0)$ , and  $(0, 1) \mapsto (0)$ , so the projection fails to be injective on  $I$ . To indicate why this restricted definition of equivalence may not be suitable for a more robust theory, note that the two line segments  $[0, 1]$  and  $[0, 2]$  are not equivalent (given our definition), even though they are geometrically congruent. Considering the more general affine transformations resolves this problem, but is not necessary here (mention projective transformations and give a reference for further reading).

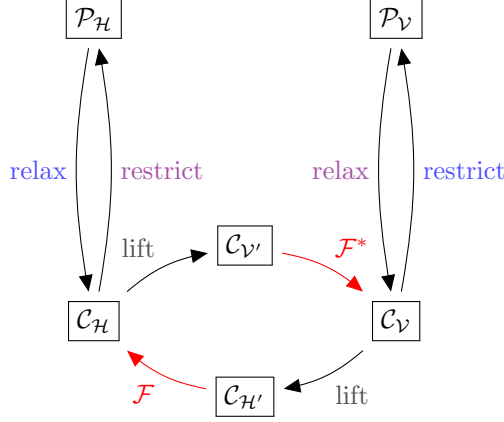


Figure 1: Diagram of the proof

### 3 Proof of the Minkowski-Weyl Theorem

#### 3.1 Overview

The proof shall proceed by first considering an H-Polyhedron  $\mathcal{P}_H$ , and constructing from it a V-Polyhedron  $\mathcal{P}_V$  which represents the same set of points, then starting with a V-Polyhedron  $\mathcal{P}_V$  and constructing an H-Polyhedron  $\mathcal{P}_H$ . The high-level steps are almost identical, and is illustrated in Figure 1.

First,  $\mathcal{P}_H$  will be relaxed to a cone  $\mathcal{C}_H$ . While technically unnecessary, in the steps that follow it is convenient to have a cone, as opposed to a more general polyhedron.  $\mathcal{C}_H$  is then immediately lifted to a higher dimension so that it can be directly represented as a V-Cone  $\mathcal{C}_V$  (have we defined V-Cone?). It is at this point that the “real work” must be done (which is why this arrow is colored red). This step requires intersecting the V-Cone with a number of hyperplanes, and requires a process which is dual to the so-called “Fourier-Motzkin Elimination.” At this point, a simple observation will allow us to restrict  $\mathcal{C}_V$  to get us to  $\mathcal{P}_V$ , completing this direction of the proof. Note that this restriction is the opposite of the relaxation which occurs at the beginning of this direction of the proof, as the colors of the diagram reflect.

For the other direction, the exact same steps are taken, however the relaxing and lifting that occur are designed to convert the V-Polyhedron into a V-Cone and H-Cone, instead of the other way around<sup>1</sup>. Also, the full fledged “Fourier-Motzkin Elimination” shall take place, this time because we need to project an H-Cone down a number of dimensions.

The transformation which go from  $V \rightarrow H$  and  $H \rightarrow V$  representations in the proof seem like a bit of trickery, but in earnest these come from the field of Linear Programming. For more information on these types of transformations, and a more systematic treatment of them, see (that book by Matusek...), for instance.

Finally, before we get started, I'd like to mention that Ziegler provides the questions:

- “Is a polyhedron intersected with a hyperplane also a polyhedron?”
- “Is the projection of a polyhedron also a polyhedron?”

as examples of the utility of the two different representations. The answer to the first question is clear for H-Polyhedra, this is merely adding another constraint to the system, while it is not so clear for a V-Polyhedra (it seems as though you must somehow “solve” something to prove this...) Similarly, the second question is clear for V-Polyhedra, you simply take the vectors that you already have and forget about one of the coordinates, while for H-Polyhedra it again seems like some system needs to somehow be solved in order to prove this statement. Before you go and try to prove it yourself, know that the “solving” of these problems is essentially Fourier-Motzkin Elimination, and is the bulk of the work in the proof that follows.

### **3.2 H-Polyhedra $\rightarrow$ V-Polyhedra**

### **3.3 V-Polyhedra $\rightarrow$ H-Polyhedra**

## **4 Program**

## **5 Discussion**