Chapter 2
Polyhedra

2.1 Basic Definitions

Polyhedra play a central role in the theory of Linear Programming (LP), Integer Linear Programming (ILP) and combinatorial optimization, beside other branches of mathematics. The importance of polyhedra is due to the fact that they are at the same time both convex objects, and as such they share all the important properties of convex sets, and also combinatorial objects due to the rich structure of vertices, edges, faces and facets.

For a thorough treatment of all properties of polyhedra and convex sets the reader is referred to the monographs by Grünbaum (2003) and Rockafellar (2015). In this chapter we limit ourselves to present the most relevant properties with respect to the topics covered in this book.

Definition 2.1 *A polyhedron in* $\mathbb{R}^n$ *is the intersection of a finite number of closed half-spaces.*

We recall that a closed half-space is the set of points $x \in \mathbb{R}^n$ that satisfy a linear inequality $a\,x \le b$ (or equivalently $a\,x \ge b$), where a is a vector in $\mathbb{R}^n$ and b is a scalar. A hyper-plane (or simply a 'plane', as we will always refer to) is the boundary of the half-space and is the set of points $x \in \mathbb{R}^n$ that satisfy the linear equation $a\,x = b$. Therefore, a polyhedron can be also defined as the feasible set of a finite number of linear inequalities. Linear equations can be also considered in the definition because each linear equation $a\,x = b$ is equivalent to the pair of linear inequalities $a\,x \le b$ and $a\,x \ge b$.

We recall that a *convex combination* of $k \ge 1$ vectors $u^1, \dots, u^k$ is any vector v that can be expressed as $v = \sum_{i \in [k]} \lambda_i\, u^i$ with $\lambda_i \ge 0$, $i \in [k]$, and $\sum_{i \in [k]} \lambda_i = 1$, while a *conic combination* is any vector v that can be expressed as $v = \sum_{i \in [k]} \lambda_i\, u^i$ with $\lambda_i \ge 0, i \in [k]$. The combinations are said *strict* if all coefficients λ_i are positive. The *convex hull* $\operatorname{conv}(K)$ (*conic hull* $\operatorname{cone}(K)$) of a set $K \subset \mathbb{R}^n$ is the set of all points in $\mathbb{R}^n$ that can be expressed as convex (conic) combinations of any finite set of points of K.

G. Lancia and P. Serafini, *Compact Extended Linear Programming Models*,
EURO Advanced Tutorials on Operational Research,
DOI 10.1007/978-3-319-63976-5_2

A set K is convex if $K = \text{conv}(K)$ or, equivalently, if it contains all convex combinations of points in K. A *vertex* (or *extreme point*) is any point of K which cannot be represented as a strict convex combination of two points of K. A polyhedron is *pointed* if it has at least one vertex. A set C is a *cone* if, for any $u \in C$ also $\lambda\, u \in C$ for any $\lambda \geq 0$. Moreover, a set C is a *convex cone* if $C = \text{cone}(C)$.

According to the definitions, a half-space is a convex cone, and a polyhedron, being the intersection of convex sets, is convex. A polyhedron is a convex cone if all its defining inequalities are homogeneous, i.e., they can all be written in the form $a\, x \leq 0$.

We say that a polyhedron P is bounded if there exists M such that $-M \leq x_i \leq M$, $i = 1, \dots, n$, for any $x \in P$. Otherwise, the polyhedron is unbounded. A bounded polyhedron is also called a *polytope*, and it is necessarily pointed.

We remind that sometimes in the literature we may find geometrical objects in $\mathbb{R}^3$ that are called polyhedra but they are not convex. Etymologically the term polyhedron comes from the Greek words πολύζ (*polús*, many) and ἕδρα (*hedra*, seat) that simply refer to an object made up of straight faces joined together by segments. This is why we may encounter these 'strange' polyhedraIn Convex Analysis a polyhedron is always convex and is defined as above.

Given m points $u^1, \dots, u^m \in \mathbb{R}^n$ and $1 \leq k \leq m$, we may define the set

$$K = \Big\{ v \in \mathbb{R}^n : v = \sum_{i \in [m]} \lambda_i\, u^i, \quad \sum_{i=1}^{k} \lambda_i = 1, \quad \lambda_i \geq 0, \quad i \in [m] \Big\}. \tag{2.1}$$

K is therefore the convex hull of the points $u^1, \dots, u^k$ plus the conic hull of the points $u^{k+1}, \dots, u^m$. The points $u^{k+1}, \dots, u^m$ are also referred to as *rays* or *directions* because all points $w + \lambda\, u^j$, for $j = k+1, \dots, m$, are in K for any $w \in K$ and any $\lambda \geq 0$. Note that each ray is defined up to a positive multiplicative factor. For this reason when we say that two rays u^1 and u^2 are different we actually mean $\text{cone}(u^1) \neq \text{cone}(u^2)$. The following fundamental theorem holds:

Theorem 2.2 *(Minkowski-Weyl) The set K as defined in (2.1) is a polyhedron. Conversely, every polyhedron can be represented as in (2.1).*

Definition 2.1 corresponds to an 'external' representation of a polyhedron: there are planes that cut off open half-spaces, like a carving operation on $\mathbb{R}^n$. What remains inside is the polyhedron. Theorem 2.2 yields an 'internal' representation: there are points (and possibly rays) and the space between the points is filled by the convex (and possibly conic) combination operation. The remarkable fact expressed by the theorem is that the two representations are equivalent and in both cases we only need a finite information.

As a simple example consider the polyhedron in $\mathbb{R}^2$ defined by only one inequality, i.e., $x_1 + x_2 \geq 2$ which is just the closed half-plane 'above' the straight line $x_1 + x_2 = 2$. It does not look like a polyhedron as we may have it in our imagination, but it is nonetheless a polyhedron. One internal representation is given by

$$u^1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad u^2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad u^3 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad u^4 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

where $m = 4$ and $k = 1$ as in (2.1). This representation is not unique, not only because u^2, u^3 and u^4 are defined up to a multiplicative positive constant, but also because u^1 can be replaced by any point on the straight line, and u^4 can be replaced by any vector with both positive entries.

This 'unpleasant' situation arises because this polyhedron has no vertices. Vertices are very special points in a polyhedron. Since they cannot be expressed as a strict convex combination of other points of the polyhedron, they must always be included in the list of points of any internal representation.

If the polyhedron is unbounded also rays are necessary. In particular we need special rays, called *extreme rays*, that are defined as the only rays that cannot be expressed as conic combination of two different rays of the polyhedron. Extreme rays play the same role of vertices with respect to rays and indeed they can be thought of as vertices at infinity. Therefore also extreme rays must be listed within the internal representation if the polyhedron is unbounded.

At this point we may wonder whether vertices and possibly extreme rays are not only necessary but also sufficient for an internal representation. We have seen in the previous example that in general this is not true. However, under a mild assumption, the answer is affirmative and further characterizes the Minkowski-Weyl theorem.

Theorem 2.3 *If a polyhedron K does not contain straight lines then it is the convex hull of its vertices plus the conic hull of its extreme rays.*

In particular we may also say that

Theorem 2.4 *A polyhedron has at least one vertex if and only if it does not contain straight lines.*

The non existence of straight lines is almost always very easily checked, either because the polyhedra we deal with are bounded by construction or because the feasible points must be nonnegative, and so even if the polyhedron is unbounded we know that it cannot contain straight lines.

We recall that by *linear combination* of the vectors $u^1, \dots, u^k$ we mean any vector v that can be expressed as $v = \sum_{i \in [k]} \lambda_i \, u^i$ with no restriction on the scalars λ_i, and by *affine combination* any vector v that can be expressed as $v = \sum_{i \in [k]} \lambda_i \, u^i$ with $\sum_{i \in [k]} \lambda_i = 1$. The *linear hull* $\text{lin}(K)$ (*affine hull* $\text{aff}(K)$) of a set $K \subset \mathbb{R}^n$ is the set of all points in $\mathbb{R}^n$ that can be expressed as linear (affine) combination of any finite set of points of K. Compare with the convex and conic combinations previously defined.

Since an affine set is a translated subspace (this subspace is unique for a given affine set), the dimension of an affine set is defined as the dimension of its generating subspace. Moreover, the dimension of a polyhedron (or, in general, of a convex set) is defined as the dimension of its affine hull. When the dimension of a convex set is the same as the space where it is defined then we say that the set is *full dimensional* or that it is a *convex body*.

If a polyhedron is not full dimensional it is more convenient to view the polyhedron as embedded in the affine manifold given by its affine combination. In particular, the topology is the one induced by the affine manifold, i.e., all neighborhoods are the intersection of the neighborhoods in $\mathbb{R}^n$ with the affine manifold, and this is called the *relative topology* of the polyhedron. Therefore when we speak of *boundary* of a polyhedron we mean boundary in the relative topology (otherwise the whole polyhedron would be boundary and the definition would become useless).

The relative boundary of a polyhedron consists of one or more polyhedra, called *facets* that have dimension one less than the dimension of the polyhedron. Given an external representation of a polyhedron, we say that an inequality is *facet-defining* if there is a facet such that all points of the facet satisfy the inequality as an equation. In order to identify facet-defining inequalities the following theorem is useful.

Theorem 2.5 *Let a polyhedron be defined by the inequalities $a^j\, x \le b_j$, $j \in [m]$, and possibly by equations. If there exists a point $\hat{x}$, necessarily on the relative boundary of the polyhedron, and an inequality, say $a^1\, x \le b_1$, such that $a^1\, \hat{x} = b_1$, while $a^j\, \hat{x} < b_j$, for all $j \neq 1$, then the inequality $a^1\, x \le b_1$ is facet-defining.*

Since the facets are polyhedra themselves, their relative boundaries have in turn facets, that, with respect to the original polyhedron, are simply called *faces*. Going down we arrive at faces of dimension one, that we call *edges*, and faces of dimension zero, that are the vertices. Necessarily, edges join exactly two vertices and vertices joined by an edge are called *adjacent*.

We have stated at the beginning that vertices and faces may exhibit a rich combinatorial structure. A very interesting example is given by the permutahedron, a polyhedron whose vertices are in one-to-one correspondence with the permutations of $[n]$ and whose facets are in a one-to-one correspondence with the proper subsets of $[n]$. We will provide an extensive analysis of the permutahedron in Chap. 7.

One of the most interesting questions is understanding the vertex structure from the given inequality set, or vice versa. The former case happens when a polyhedron is described by a set of linear inequalities and we have to find some vertices, as it happens in linear programming. In other cases, like in combinatorial optimization, a set of points is defined that corresponds to the set of solutions of some combinatorial structure and we want to find a description of the convex hull of these points, i.e., the linear inequalities that define this convex hull.

This task may be inherently intractable because there may be exponentially many vertices with respect to the given set of inequalities and conversely exponentially many facets with respect to a given set of vertices. As a relevant example of the first type consider the *n-dimensional cube*, also called *hyper-cub*, which we simply refer to as cube, that is defined by the $2\,n$ inequalities

$$0 \le x_j \le 1 \qquad j \in [n],$$

that give rise to 2^n vertices (all 0-1 vectors in $\mathbb{R}^n$). It is not difficult to see that the cube has

$$\binom{n}{k}\, 2^{n-k}$$

faces of dimension k.

In the reverse direction, there can be polyhedra with a small number of vertices whose external representation requires many inequalities, like the *orthoplex* (also called *cross-polytope*, or *hyperoctahedron*, or *cocube*), whose vertices are the $2\,n$ points $(0, 0, \ldots, \pm 1, \ldots, 0, 0)$. Their convex hull is described by the 2^n inequalities

$$\sum_{j \in [n]} \pm x_j \leq 1.$$

The number of faces of dimension k of the orthoplex is given by

$$\binom{n}{k+1}\, 2^{k+1}.$$

2.2 Convex Hulls of Infinitely Many Points

If S is an infinite set of points, then $\text{conv}(S)$ is not necessarily a polyhedron. However, if S exhibits special properties it may happen that $\text{conv}(S)$ is a polyhedron. For the problems considered in this monograph the case $S \subset \mathbb{Z}^n$ is relevant, i.e., when S consists of points with integral coordinates.

This is an issue that may be easily overlooked. A typical case in integer linear programming is when $S \subset \{0, 1\}^n$. Then, S is clearly finite and the Minkowski-Weyl theorem implies that $\text{conv}(S)$ is a polyhedron. This fact directs the search for optimal solutions to finding the facets of $\text{conv}(S)$, at least near the optimum. Since integer linear programming is not computationally different than binary linear programming, one is inclined to think that also the convex hull of the (infinite) feasible points of an ILP instance is a polyhedron and therefore it makes sense to find the external representation of $\text{conv}(S)$. However, some caution is necessary. Consider in $\mathbb{R}^2$ the set S

$$S = \left\{x \in \mathbb{Z}^2 : x_2 \leq \sqrt{2}\, x_1, \quad x_2 \geq 1\right\}.$$

It can be proved that $\text{conv}(S)$ has the extreme ray $(1, 0)$ and infinitely many vertices of coordinates $(k, \lfloor\sqrt{2}\, k\rfloor)$ for all k such that $2\,k^2 - 1$ is a square (sequence A001653 in OEIS (2017) that goes as $1, 5, 29, 169, 985, 5741, 33461, \ldots$). Therefore it is not a polyhedron.

However, a positive result is given by the following important theorem due to Meyer (1974).

Theorem 2.6 *Let* $S = \left\{x \in \mathbb{Z}^n : A\,x \leq b\right\}$. *If* A *and* b *have rational entries, then* $\textit{conv}(S)$ *is a polyhedron.*

2.3 The Slack Matrix

The following important concept has been introduced by Yannakakis (1991). Let a polyhedron P be defined by the set of inequalities $A\,x \leq b$ (where possible non negativity constraints are embedded in the matrix A). We assume that all inequalities are non redundant, i.e., they are facet-defining. Let V be the matrix whose columns are all vertices of P. Let $\mathbf{1}$ be a vector of all ones of size the number of vertices.

Definition 2.7 *The slack matrix of P is the nonnegative matrix*

$$S = b\,\mathbf{1}^T - A\,V.$$

In other words the entry s_{ij}, associated to the ith facet and to the jth vertex, measures the 'distance' of the vertex j from the facet i, given by $b_i - \sum_k a_{ik}\,v^j_k$. Defining the slack matrix, as the slack matrix of the polyhedron P, is somehow misleading, because the slack matrix is based on how the polyhedron is represented through the matrix A and the vector b. It is clear that an inequality can be multiplied by a positive constant and the polyhedron is not altered while the slack matrix is altered by a constant factor along a row. However, this would not alter the structural properties of the slack matrix, so that we may, with abuse of terminology, speak of the slack matrix of a polyhedron.

For instance the slack matrix of the cube is the $2n \times 2^n$ matrix

$$\begin{pmatrix} V \\ \bar{V} \end{pmatrix}$$

where V is the matrix of all vertices of the cube, i.e., all possible 0-1 vectors and $\bar{V}$ is the same matrix with all entries flipped.

A general concept associated to any $m \times n$ matrix A is the *rank* of the matrix, rank$A$ (also called *linear rank*). One possible way to define the rank is by finding a factorization of A into the product of an $m \times q$ matrix B and a $q \times n$ matrix C. The minimum number q such that a factorization exists is the rank of A.

An important concept associated to a nonnegative $m \times n$ matrix A is the *nonnegative rank* of the matrix. The minimum number q such that a factorization of A into the product of an $m \times q$ nonnegative matrix B and a $q \times n$ nonnegative matrix C exists is called the nonnegative rank of A and is denoted $\text{rank}_+ A$.

The problem of finding a nonnegative factorization is called *Nonnegative Matrix Factorization* (NMF). NMF is very important in various areas of applied mathematics, like quantum mechanics, probability theory, polyhedral combinatorics, communication complexity, demography, chemiometrics, machine learning, data mining (see for instance Gregory and Pullman 1983; Cohen and Rothblum 1993). In particular in data analysis the NMF corresponds to extracting features from a dataset. This explains why the two matrix factors are often denoted as F (features) and W (weights).

We have the following relations

$$\operatorname{rank} A \le \operatorname{rank}_+ A \le \min\{m, n\}.$$

The first inequality is obvious. As for the second inequality we note that if $m \le n$, by taking $B = I$ (or any permutation matrix of order m), we obtain a nonnegative factorization with $q = \min\{m, n\}$. Similarly if $n \le m$, by taking $C = I$ (or any permutation matrix of order n), we obtain a nonnegative factorization with $q = \min\{m, n\}$. Whenever we have $\operatorname{rank}_+ A = \min\{m, n\}$ we speak of a trivial nonnegative rank factorization.

As we shall see later (Sect. 6.8) the nonnegative rank of the slack matrix plays an important role relating the facets of a polyhedron and its projections.

2.4 Projections of Polyhedra

In this book we are interested in polyhedra and projections of polyhedra, and it is useful to briefly recall some facts associated to projections. Originally, a projection is a geometrical operation for sending points of the three dimensional space onto a plane along some straight lines precisely defined. This operation can be abstracted to general vector spaces. If we require the operation to be linear we have the following definition.

Definition 2.8 *A projection linear operator $\mathscr{P} : R^n \to R^n$ is any linear operator that satisfies the idempotency property $\mathscr{P}^2 = \mathscr{P}$.*

Let $\mathscr{N}$ be the null space of $\mathscr{P}$ and $\mathscr{R}$ its range. Equivalently, $\mathscr{P}$ is a projection operator if $\mathscr{P}$ restricted to $\mathscr{R}$ is the identity. By the idempotency property $x - \mathscr{P}x \in \mathscr{N}$.

Clearly a particular projection operator is identified by its null space and its range. Therefore the way a projection operator can be represented depends on how the range and the null space are represented. We say that a projection is orthogonal if the null space and the range are orthogonal subspaces. In this case, we only need the information either for the range or for the null space.

As a first case, let us assume that we know a basis $r^1, \ldots, r^k$ for $\mathscr{N}$ and a basis $r^{k+1}, \ldots, r^n$ for $\mathscr{R}$. Let R be the $n \times n$ matrix whose columns are the n vectors $r^1, \ldots, r^n$. By taking $r^1, \ldots, r^n$ as a new basis for R^n the relationship between the old coordinates x and the new coordinates ξ of a generic point is given by

$$x = R\,\xi.$$

Hence we have for any linear operator $A : R^n \to R^n$ (where η are the new coordinates of a point y)

$$y = Ax \quad \Longrightarrow \quad R\,\eta = AR\,\xi \quad \Longrightarrow \quad \eta = R^{-1}AR\,\xi$$

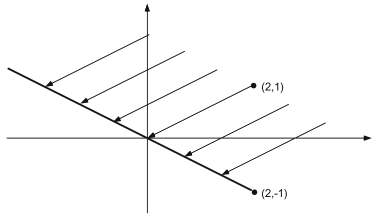

Fig. 2.1 A non orthogonal projection in $\mathbb{R}^2$

and the operator A is represented in the new system as $R^{-1} A\, R$. Any vector represented in the new system has the first k coordinates spanning $\mathcal{N}$ and the remaining $n-k$ coordinates spanning $\mathcal{R}$. Therefore if $\mathcal{P}$ is a projection operator

$$R^{-1} \mathcal{P}\, R = J_k$$

where J_k is a diagonal matrix with the first k entries equal to 0 and the remaining entries equal to 1. Therefore

$$\mathcal{P} = R\, J_k\, R^{-1}. \tag{2.2}$$

As a simple example consider the non orthogonal projection in Fig. 2.1 where

$$r^1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, r^2 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, R = \begin{pmatrix} 2 & 2 \\ 1 & -1 \end{pmatrix}, J_k = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

$$R^{-1} = \frac{1}{4} \begin{pmatrix} 1 & 2 \\ 1 & -2 \end{pmatrix}, \mathcal{P} = \frac{1}{4} \begin{pmatrix} 2 & -4 \\ -1 & 2 \end{pmatrix}.$$

As a second case, let us now assume that the null space $\mathcal{N}$ is defined as

$$\mathcal{N} = \{x : K\, x = 0\}$$

for some $(n-k) \times n$ matrix K with full rank and we do not have available a basis for $\mathcal{N}$. In other words, we explicitly know the orthogonal vectors to the null space but not the basis of the null space. Let R_1 be the $n \times (n-k)$ matrix with columns $r^{k+1}, \ldots, r^n$, a basis of $\mathcal{R}$ which we assume to be available. Any vector in $\mathcal{R}$ can be written as $R_1\, y$ for some vector y. Note that $K\, R_1$ is non singular because for any $y \neq 0$ we must have $K\, R_1 y \neq 0$ (the range and the null space are two complementary subspaces). Since $x - \mathcal{P}\, x \in \mathcal{N}$ we have $K\,(x - \mathcal{P}\, x) = 0$. Since $\mathcal{P}\, x = R_1\, y$ for some vector y, we may write $K\, x - K\, R_1\, y = 0$ from which

$$y = (K\, R_1)^{-1}\, K\, x$$

and therefore

$$\mathscr{P} = R_1\,(K\,R_1)^{-1}K.$$

In the example

$$K = \begin{pmatrix} 1 & -2 \end{pmatrix},\ R_1 = \begin{pmatrix} 2 \\ -1 \end{pmatrix} \Longrightarrow \mathscr{P} = \begin{pmatrix} 2 \\ -1 \end{pmatrix} \frac{1}{4} \begin{pmatrix} 1 & -2 \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 2 & -4 \\ -1 & 2 \end{pmatrix}.$$

As a third case suppose now that $\mathscr{R}$ is defined via equations, i.e.,

$$\mathscr{R} = \{x : H\,x = 0\}$$

for some $k \times n$ matrix H with full rank and that we have available a representation of $\mathscr{N}$ via a basis $r^1, \ldots, r^k$. Let R_0 be the $n \times k$ matrix whose columns are $r^1, \ldots, r^k$. Note that $H\,\mathscr{P} = 0$. Then we have $x - \mathscr{P}\,x = R_0\,y$ for some vector y and, by applying H, $H\,x - H\,\mathscr{P}\,x = H\,R_0\,y$, i.e., $H\,x = H\,R_0\,y$ and we may write

$$y = (H\,R_0)^{-1}H\,x.$$

Therefore from $\mathscr{P}\,x = x - R_0\,y$ we may write

$$\mathscr{P} = I - R_0\,(H\,R_0)^{-1}H. \tag{2.3}$$

In the example

$$H = \begin{pmatrix} 1 & 2 \end{pmatrix},\quad R_0 = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad \Longrightarrow$$

$$\mathscr{P} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 2 \\ 1 \end{pmatrix} \frac{1}{4} \begin{pmatrix} 1 & 2 \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 2 & -4 \\ -1 & 2 \end{pmatrix}.$$

As a fourth final case, suppose that both $\mathscr{N}$ and $\mathscr{R}$ are defined via equations, i.e.,

$$\mathscr{N} = \{x : K\,x = 0\}, \qquad \mathscr{R} = \{x : H\,x = 0\}.$$

Then we have

$$K\,\mathscr{P} = K, \qquad H\,\mathscr{P} = 0.$$

The first relation is due to $x - \mathscr{P}\,x \in \mathscr{N}$ for all x, and the second relation is due to $\mathscr{P}\,x \in \mathscr{R}$ for all x. The previous relations can be written as

$$\begin{pmatrix} K \\ H \end{pmatrix} \mathscr{P} = \begin{pmatrix} K \\ 0 \end{pmatrix}$$

and since the rows of K and H are linearly independent, the matrix is invertible and we have

$$\mathscr{P} = \begin{pmatrix} K \\ H \end{pmatrix}^{-1} \begin{pmatrix} K \\ 0 \end{pmatrix}.$$

In the example

$$\begin{pmatrix} K \\ H \end{pmatrix}^{-1} \begin{pmatrix} K \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & -2 \\ 1 & 2 \end{pmatrix}^{-1} \begin{pmatrix} 1 & -2 \\ 0 & 0 \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 2 & 2 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 0 & 0 \end{pmatrix}.$$

For orthogonal projections the formulas are simpler. Indeed if $\mathscr{R}$ is represented as

$$\mathscr{R} = \{x : H\,x = 0\}$$

for some $k \times n$ matrix H with full rank, then the rows of H are a basis for $\mathscr{N}$ and therefore (2.3) becomes

$$\mathscr{P} = I - H^T (H\,H^T)^{-1} H. \tag{2.4}$$

These projection operators can be also used to find the projection of a polyhedron if we have an internal representation of the polyhedron as the convex hull of vertices and the conic hull of estreme rays.

If the polyhedron Q is given by a set of inequalities, i.e.,

$$Q = \left\{x \in \mathbb{R}^n : A\,x \le d\right\}$$

finding its projection $\overline{Q}$ is more complex. Let us start by studying the simpler projection in which $\mathscr{R}$ is the subspace generated by $\mathbf{e}_1, \dots, \mathbf{e}_{n-k}$, and the projection is orthogonal. Let

$$Q = \left\{(x, y) \in \mathbb{R}^n : T\,x + R\,y \le d\right\} \tag{2.5}$$

with T a $m \times (n-k)$-matrix and R a $m \times k$ matrix. We want to find the external representation of the polyhedron $\overline{Q} = \mathscr{P}\,Q$ on the subspace $\{(x, y) : y = 0\}$. Let, for each $x \in \mathbb{R}^{n-k}$,

$$Q(x) = \left\{y \in \mathbb{R}^k : R\,y \le d - T\,x\right\}. \tag{2.6}$$

The polyhedron $\overline{Q}$ can be also expressed as

$$\overline{Q} = \left\{x \in \mathbb{R}^{n-k} : Q(x) \ne \emptyset\right\}.$$

Now we introduce one of the most fundamental tools in linear algebra, polyhedral theory, linear programming, combinatorial optimization, etc. This is the celebrated Farkas' lemma that dates back to 1894 (Farkas 1894, 1902) and can be considered in some sense the ancestor of linear programming. The lemma defines two geometrical sets and states that one is empty if and only if the other one is not empty. There are several variants of the lemma according to little variations about the definitions of the two sets. The theorem can be proved from the separation theorem of convex sets. We do not report here the proof.

Theorem 2.9 (Farkas' lemma) *Given an* $m \times n$*-matrix* Q *and an* m*-vector* q*, the feasible set of*

$$Q\,u \le q, \qquad u \ge 0 \tag{2.7}$$

is non empty if and only if the feasible set of

$$v\,Q \ge 0, \qquad v\,q < 0, \qquad v \ge 0 \tag{2.8}$$

is empty.

The Farkas' lemma can be rephrased in polyhedral terms by saying that the polyhedron

$$\left\{u \in \mathbb{R}^n : Q\,u \le q,\; u \ge 0\right\}$$

is not empty if and only $v\,q \ge 0$ for all points in the polyhedron

$$\left\{v \in \mathbb{R}^m : v\,Q \ge 0,\; v \ge 0\right\}.$$

In a variant of the lemma we may say that the polyhedron

$$\left\{u \in \mathbb{R}^n : Q\,u \le q\right\}$$

is not empty if and only $v\,q \ge 0$ for all points in the polyhedron

$$\left\{v \in \mathbb{R}^m : v\,Q = 0,\; v \ge 0\right\}.$$

According to this variant of the Farkas' lemma we have that $Q(x)$ in (2.6) is not empty if and only if for all u such that $u \ge 0$ and $u\,R = 0$ one has $u\,(d - T\,x) \ge 0$. Since the set $C := \{u \ge 0 : u\,R = 0\}$ is a polyhedral cone, the condition $u\,(d - T\,x) \ge 0$ can be expressed only with respect to the generators of C, and, if C is pointed, just to the extreme rays of C. Let $u^1, \dots, u^q$ be these extreme rays. Therefore $Q(x) \ne \emptyset$ if and only if $u^i\,(d - T\,x) \ge 0$, $i = 1, \dots, q$, so that

$$\overline{Q} = \left\{x \in \mathbb{R}^k : u^i\,(d - T\,x) \ge 0,\; i = 1, \dots, q\right\}.$$

Denoting by U the matrix whose rows are the extreme rays u^i, the polyhedron $\overline{Q}$ is given by

$$U\,T\,x \le U\,d.$$

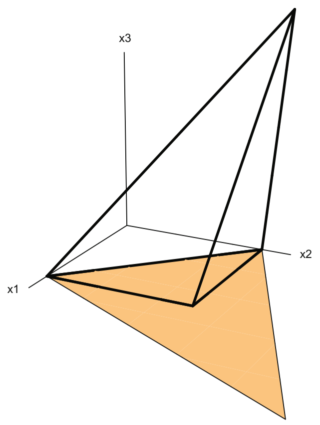

Fig. 2.2 A polyhedron in $\mathbb{R}^3$ and its projection

As a small example consider the polyhedron in $\mathbb{R}^3$ defined by

$$\begin{pmatrix} -2 & -2 & 3 \\ 2 & 0 & -1 \\ 0 & 2 & -1 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \le \begin{pmatrix} -2 \\ 2 \\ 2 \\ 0 \end{pmatrix}$$

and shown in Fig. 2.2. We want to project the polyhedron onto the subspace generated by $\mathbf{e}_1$ and $\mathbf{e}_2$ (see in the figure also the projected polyhedron, actually a polygon). Hence

$$T = \begin{pmatrix} -2 & -2 \\ 2 & 0 \\ 0 & 2 \\ 0 & 0 \end{pmatrix}, \qquad R = \begin{pmatrix} 3 \\ -1 \\ -1 \\ -1 \end{pmatrix}, \qquad d = \begin{pmatrix} -2 \\ 2 \\ 2 \\ 0 \end{pmatrix}$$

and

$$C = \left\{ u \in \mathbb{R}^4 : 3\,u_1 - u_2 - u_3 - u_4 = 0,\ u_1 \ge 0,\ u_2 \ge 0,\ u_3 \ge 0,\ u_4 \ge 0 \right\}.$$

The extreme rays of C have only two components different from zero and so they are

$$U := \begin{pmatrix} u^1 \\ u^2 \\ u^3 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 0 & 0 \\ 1 & 0 & 3 & 0 \\ 1 & 0 & 0 & 3 \end{pmatrix}$$

so that

$$U\,T = \begin{pmatrix} 4 & -2 \\ -2 & 4 \\ -2 & -2 \end{pmatrix} \qquad U\,d = \begin{pmatrix} 4 \\ 4 \\ -2 \end{pmatrix}$$

and $\overline{Q} \subset \mathbb{R}^2$ is described by

$$\begin{aligned} 2\,x_1 - \ x_2 &\leq 2 \\ -x_1 + 2x_2 &\leq 2 \\ x_1 + x_2 &\geq 1 \end{aligned}$$

A slightly more complicated situation arises when we want to project a polyhedron Q onto a subspace $\mathscr{R}$ defined by

$$\mathscr{R} = \{x : H\,x = 0\}$$

where H is a $k \times n$ matrix (not necessarily of full rank). We note that $x \in \overline{Q}$ if and only if there exists a linear combination of the rows of H, i.e., $H^T\,y$, such that $x + H^T\,y \in Q$. Hence $x \in \overline{Q}$ if and only if

$$A\,(x + H^T\,y) \leq d, \quad H\,x = 0$$

i.e.,

$$\begin{aligned} A\,x + A\,H^T\,y &\leq d \\ H\,x \qquad\qquad &= 0. \end{aligned}$$

In this way we have a new polyhedron in a higher dimensional space of dimension $(n + k)$ that we have to project back to $\mathbb{R}^n$, i.e., to the subspace spanned by the first n axes. We apply the previous results with the only warning that we have also a set of equations (not all inequalities) and this amounts to having a corresponding free variable in the Farkas' lemma. Hence, with respect to the previous notation

$$T = \begin{pmatrix} A \\ H \end{pmatrix}, \qquad R = \begin{pmatrix} A\,H^T \\ 0 \end{pmatrix}.$$

We have to find the generators of C

$$C = \left\{(u, v) : u\,A\,H^T + v\,0 = 0,\ u \geq 0\right\}.$$

The cone C (note that C is not pointed because v is free) can be represented by the extreme rays of

$$u\,A\,H^T = 0, \quad u \geq 0, \tag{2.9}$$

plus a linear combination of the unit vectors of the axes $v^1, \ldots, v^k$. Let u^j be one of the extreme rays of (2.9). By the previous results, an inequality describing $\overline{Q}$ is given by

$$u^j A + v\,H\,x \leq u^j d.$$

Note that, by putting $u = 0$ we obtain the constraint $v\,H\,x \le 0$, from which $H\,x = 0$ since v can have any sign. Therefore, denoting by U the matrix whose rows are the extreme rays of (2.9), the inequalities defining $\overline{Q}$ are

$$U\,A\,x \le U\,d, \qquad H\,x = 0.$$

We reconsider the previous example, this time by projecting onto the subspace (see Fig. 2.3)

$$\left\{x \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 0\right\}$$

so that

$$H = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}$$

and

$$A\,H^T = \begin{pmatrix} -2 & -2 & 3 \\ 2 & 0 & -1 \\ 0 & 2 & -1 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 1 \\ -1 \end{pmatrix}.$$

The extreme rays are the rows of

$$U = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

and therefore

$$U\,A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} -2 & -2 & 3 \\ 2 & 0 & -1 \\ 0 & 2 & -1 \\ 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & -2 & 2 \\ -2 & 0 & 2 \\ 2 & 0 & -2 \\ 0 & 2 & -2 \end{pmatrix}, \quad U\,d = \begin{pmatrix} 0 \\ 0 \\ 2 \\ 2 \end{pmatrix}.$$

Hence $\overline{Q}$ is described by

$$\begin{aligned} -2\,x_2 + 2\,x_3 &\le 0 \\ -2\,x_1 \qquad + 2\,x_3 &\le 0 \\ 2\,x_1 \qquad - 2\,x_3 &\le 2 \\ 2\,x_2 - 2\,x_3 &\le 2 \\ x_1 + \quad x_2 + \quad x_3 &= 0 \end{aligned}$$

i.e.,

$$\left\{x \in \mathbb{R}^3 : 0 \le x_1 - x_3 \le 1,\ 0 \le x_2 - x_3 \le 1,\ x_1 + x_2 + x_3 = 0\right\}.$$

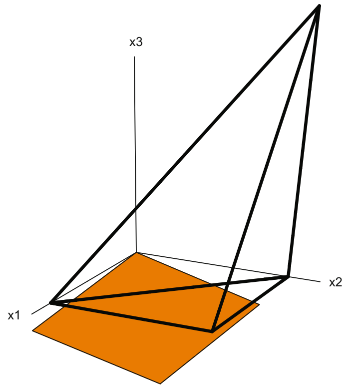

Fig. 2.3 Projection of the polyhedron onto the subspace $x_1 + x_2 + x_3 = 0$

The vertices of Q are

$$\hat{x}^1 = \begin{pmatrix} 2 & 2 & 2 \end{pmatrix}, \quad \hat{x}^2 = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}, \quad \hat{x}^3 = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}, \quad \hat{x}^4 = \begin{pmatrix} 1 & 1 & 0 \end{pmatrix},$$

and the vertices of $\overline{Q}$ are

$$\bar{x}^1 = \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}, \bar{x}^2 = \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \end{pmatrix}, \bar{x}^3 = \begin{pmatrix} -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \end{pmatrix}, \bar{x}^4 = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} \end{pmatrix}.$$

The projection matrix $\mathscr{P}$ is given, according to (2.4)

$$\mathscr{P} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 1 & 1 & 1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}$$

and we may verify that $\mathscr{P}\,\hat{x}^j = \bar{x}^j$.

Enumerating the extreme rays of a cone can be a time-consuming task. However, if we split the procedure into several steps, one variable at a time, we obtain the so-called *Fourier elimination scheme*, that is easily implementable. Assume that R in (2.5) has one column (and T has $n-1$ columns). From R we form three index sets, namely

$$I^0 = \{i : R_i = 0\}, \qquad I^+ = \{i : R_i > 0\}, \qquad I^- = \{i : R_i < 0\}.$$

Now we have to find the extreme rays of the cone

$$u \geq 0, \qquad u\,R = 0.$$

If $R_h = 0$ for some $h \in I^0$ then one extreme ray is given by $u_h = 1$, $u_i = 0$, $i \neq h$. The other extreme rays are obtained by picking $h \in I^+$ and $k \in I^-$ and defining $u_h = -R_k$ and $u_k = R_h$, $u_i = 0$, $i \neq h, k$. This leads to the following new inequalities in the subspace of $x_1, \ldots, x_{n-1}$

$$\begin{aligned} \sum_{j=1}^{n-1} (R_h\, T_{kj} - R_k\, T_{hj})\, x_j \leq R_h\, d_k - R_k\, d_h, \qquad & h \in I^+, k \in I^- \\ \sum_{j=1}^{n-1} T_{hj}\, x_j \leq d_h \qquad & h \in I^0 \end{aligned}$$

The number of generated inequalities can grow very quickly. But this drawback is inherent in the projection process itself. Some inequalities may be redundant and we should be able to identify these inequalities. This is also a time-consuming task. Just think that if we project a polyhedron onto $\mathbb{R}^1$, only two inequalities are eventually needed but the elimination scheme generates many inequalities.

We stress the fact that the number of inequalities of the projected polyhedron can be higher with respect to the original polyhedron. It is indeed this property that is exploited when, going in the reverse direction, we solve a problem in a higher dimension because there are much less inequalities. Consider for instance the 4-dimensional unit cube that is rotated so that one diagonal is on the $\mathbf{e}_4$ axis and we want to project it on the 3-dimensional subspace spanned by $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$. The cube in the new coordinate system is described by the following eight inequalities

$$\begin{aligned} 0 \leq \quad x_1 + x_2 + x_3 + x_4 \leq 2, \qquad & 0 \leq \quad x_1 - x_2 - x_3 + x_4 \leq 2, \\ 0 \leq -x_1 + x_2 - x_3 + x_4 \leq 2, \qquad & 0 \leq -x_1 - x_2 + x_3 + x_4 \leq 2. \end{aligned}$$

These inequalities have been obtained by defining new orthonormal axes $r^1, \ldots, r^4$ that form as columns the matrix

$$R = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \\ -1 & 1 & -1 & 1 \\ -1 & -1 & 1 & 1 \end{pmatrix} \tag{2.10}$$

and then the original inequality matrix has been multiplied by R.

By the Fourier method one obtains 16 inequalities. Four of them are redundant (their coefficients are all zero) and the projection (shown in Fig. 2.4) is described by the following 12 facet-defining inequalities

$$\begin{array}{rrr} x_2 + x_3 \leq 1, & x_1 + x_2 \leq 1, & x_1 + x_3 \leq 1, \\ x_2 - x_3 \leq 1, & x_1 - x_2 \leq 1, & x_1 - x_3 \leq 1, \\ -x_2 - x_3 \leq 1, & -x_1 - x_2 \leq 1, & -x_1 - x_3 \leq 1, \\ -x_2 + x_3 \leq 1, & -x_1 + x_2 \leq 1, & -x_1 + x_3 \leq 1. \end{array} \tag{2.11}$$

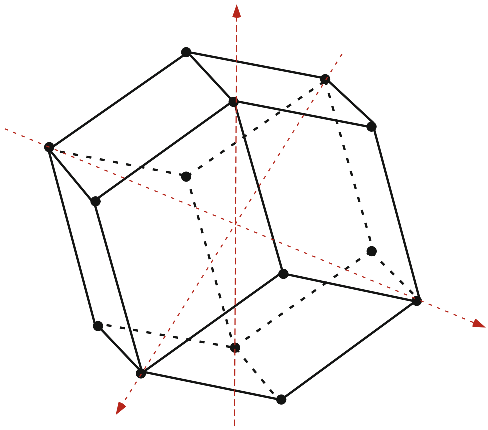

Fig. 2.4 A projection in $\mathbb{R}^3$ of the four dimensional cube

The projected polyhedron has the following 14 vertices

$$(0,0,-1),\ \frac{1}{2}(-1,1,-1)\ ,\frac{1}{2}(1,1,-1),\ \frac{1}{2}(1,-1,-1),\ \frac{1}{2}(-1,-1,-1),$$

$$(-1,0,0),\ (0,-1,0),\ (1,0,0),\ (0,1,0),$$

$$\frac{1}{2}(-1,-1,1),\ \frac{1}{2}(-1,1,1),\ \frac{1}{2}(1,-1,1),\ \frac{1}{2}(1,1,1),\ (0,0,1).$$

All facets are equal diamonds with angles equal to either $\arccos(1/3) \simeq 70.53°$ or $\arccos(-1/3) \simeq 109.47°$.

Let us do again the same computation, this time however by projecting the cube in the usual formulation (i.e., $0 \le x_i \le 1$, $i \in [n]$) onto the subspace $\{x : x_1 + x_2 + x_3 + x_4 = 0\}$. Hence according to (2.9) we have to find the extreme rays of $u\,A\,H^T$, i.e., of

$$u \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = u_1 - u_2 + u_3 - u_4 + u_5 - u_6 + u_7 - u_8 = 0.$$

There are 16 extreme rays consisting of 0-1 vectors with exactly two 1's, one on an odd-index entry and the other one on an even-index entry. Hence the 12 inequalities that are obtained from $U\,A \leq U\,d$ (as a simple calculation shows) are

$$x_h - x_k \leq 1, \qquad h, k = 1, \ldots, 4, \; h \neq k. \tag{2.12}$$

To understand the relation between (2.12) and (2.11) we have to find a new coordinate system for $\mathbb{R}^4$ such that the fourth axis is orthogonal to the projection range and the other three axes span the range. For instance we may use the same matrix R in (2.10) to establish the relation $x = R\,\xi$ between the old x and the new ξ coordinates. If we multiply the matrix underlying the inequalities in (2.12) by R we obtain (an easy exercise) exactly (2.11) (with the only difference that ξ replaces x)

It is possible to establish a lower bound on the number of inequalities needed for the higher dimensional polyhedron. More specifically, let $Q \subset \mathbb{R}^{n+m}$ and $P \subset \mathbb{R}^n$ be two polyhedra such that $\mathscr{P}Q = P$. Moreover, we know the number $v(P)$ of vertices of P. Then the following theorem holds (Goemans 2015).

Theorem 2.10 *The number $t(Q)$ of facets of Q satisfies $t(Q) \geq \log_2 v(P)$.*

Proof It is easy to show that each face of P is the projection of some face of Q. This implies that the number $f(P)$ of faces of P is at most the number $f(Q)$ of faces of Q, i.e., $f(P) \leq f(Q)$. This implies $v(P) \leq f(P) \leq f(Q)$. Now we note that each face of a polyhedron is the intersection of some subset of its facets. This implies that $f(Q)$ is upper bounded by the number of subsets of facets, i.e., $f(Q) \leq 2^{t(Q)}$. Putting together the inequalities the thesis follows. ∎

Since the lower bound given by the theorem grows very slowly, a polynomial external description of Q is not excluded even if the number of vertices is exponential. For instance the permutahedron has $v(P) = n!$ vertices, so that the bound is $\log_2 n! = O(n \, \log n)$ and there exists a polyhedron with this bound as it will be shown in Sect. 7.4. We shall provide another lower bound in Sect. 6.8 based on the slack matrix of P.

2.5 Union of Polyhedra Defined by Inequalities

While the intersection of convex sets is always a convex set, the union of convex sets is in general not convex. However, disjunctive sets are almost pervasive in combinatorial optimization and it is natural to model problems by considering the union of polyhedra. As long as we have linear objective functions it is equivalent to optimize over a set or over its convex hull. Therefore, the question we would like to answer is: given m polyhedra P^i, $i \in [m]$, how to find the external representation of

$$\overline{\text{conv}} \bigcup_{i \in [m]} P^i$$

where $\overline{\text{conv}}$ is the closure of the convex hull. In order to obtain a polyhedron it is necessary to take the closure if one of the polyhedra is unbounded. Consider the simple example in $\mathbb{R}^2$ with $P^1 = \{x : x_2 = 0, x_1 \geq 0\}$ and $P^2 = \{(0, 1)\}$. The convex hull of $P^1 \cup P^2$ is the non closed strip $\{x : x_1 \geq 0, 0 \leq x_2 < 1\}$ plus the point (0, 1).

Let us first consider the simple case of the union of two polyhedra defined by

$$P^1 := \left\{x \in \mathbb{R}^n : A^1 x \leq b^1\right\} \qquad P^2 := \left\{x \in \mathbb{R}^n : A^2 x \leq b^2\right\}$$

that we assume non empty. We want to find valid inequalities for the polyhedron

$$P := \overline{\text{conv}}(P^1, P^2).$$

Just note that $\overline{\text{conv}}(P^1, P^2)$ can be also expressed as $\text{conv}(P^1, P^2) + C^1 + C^2$ where the cones C^i are defined by $C^i := \left\{x : A^i x \leq 0\right\}$. By definition of convex hull, for each point $x \in \text{conv}(P^1, P^2)$ there exist points $y^1 \in P^1$, $y^2 \in P^2$, and coefficients $\alpha_1 \geq 0, \alpha_2 \geq 0, \alpha_1 + \alpha_2 = 1$ such that

$$x = \alpha_1 y^1 + \alpha_2 y^2$$

and P is the closure of the canonical projection over $\mathbb{R}^n$ of the set in $\mathbb{R}^{3n+2}$ given by the values $(x, y^1, y^2, \alpha_1, \alpha_2)$ feasible for the system

$$\begin{aligned} x &= \alpha_1 y^1 + \alpha_2 y^2 \\ A^1 y^1 &\leq b^1 \\ A^2 y^2 &\leq b^2 \\ \alpha_1 + \alpha_2 &= 1 \\ \alpha_1 &\geq 0 \\ \alpha_2 &\geq 0 \end{aligned} \tag{2.13}$$

The system (2.13) is clearly non linear. However, it may be easily made linear by defining $x^i := \alpha_i y^i$. Then (2.13) is equivalent to

$$\begin{array}{ccccccl} & & & \alpha_1 + & \alpha_2 & = & 1 \\ x - & x^1 - & x^2 & & & = & 0 \\ & A^1 x^1 & & - \; b^1 \alpha_1 & & \leq & 0 \\ & & A^2 x^2 & & - \; b^2 \alpha_2 & \leq & 0 \\ & & & \alpha_1 & & \geq & 0 \\ & & & & \alpha_2 & \geq & 0 \end{array} \tag{2.14}$$

Let $\tilde{P}$ be the polyhedron in $\mathbb{R}^{3n+2}$ defined by (2.14). Denote by $\mathscr{P}_x$ the canonical projection over the subspace of the x variables. Then we have:

Theorem 2.11 $P = \mathscr{P}_x \tilde{P}$.

Proof If $\alpha_i > 0$, to each solution in (2.14) there corresponds a solution of (2.13) and vice versa. If $\alpha_1 = 0$ and $\alpha_2 = 1$, then every x feasible in (2.14) is given by the sum of an element of P^2 and an element of the cone C^1 and so $x \in P$. The case $\alpha_1 = 1$ and $\alpha_2 = 0$ is symmetrical. ∎

Now we have to determine the inequalities defining P. To this aim we apply the previous results. First let us rewrite (2.14) as

$$\begin{pmatrix} 0 \\ I \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} x + \begin{pmatrix} & & 1 & 1 \\ -I & -I & & \\ A^1 & & -b^1 & \\ & A^2 & & -b^2 \\ & & -1 & \\ & & & -1 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ \alpha_1 \\ \alpha_2 \end{pmatrix} \begin{matrix} = \\ = \\ \le \\ \le \\ \le \\ \le \end{matrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{2.15}$$

Hence P is defined by the inequalities $u\,x \le u_0$ for all those values u and u_0 that correspond to the generators of the polyhedral cone

$$\tilde{C} := \left\{(u_0, u, v^1, v^2) : u = v^i\,A^i,\ u_0 \ge v^i\,b^i,\ v^i \ge 0,\ i = 1, 2\right\} \tag{2.16}$$

or, more exactly, to the values (u_0, u) that are generators of the cone C, projection of $\tilde{C}$ onto the subspace of the variables (u_0, u).

In order to find the generators of C we may use the Fourier elimination scheme. However, if we are not interested in the full description of P but we need only some facets, typically the ones that are near an optimal vertex, it is better to find directly those facets. To this aim we may solve an LP problem after having normalized the rays of C. We illustrate two possible normalizations. The first one is theoretically more appealing because it does not introduce redundant inequalities but it requires the a priori knowledge of an interior point of P (and consequently the assumption that P is a convex body). The second one is a simple $\|\cdot\|_\infty$ bound and as such it is always applicable but it may generate redundant inequalities.

As far as the first normalization is concerned, let x^0 be an interior point of P. This means that for every $(u_0, u) \in C$ one has $u\,x^0 < u_0$. Let us now operate a space translation so that x^0 becomes the origin and let $\xi := x - x^0$. Hence $u\,(\xi + x^0) \le u_0$ for every $\xi \in P - x^0$. Therefore $u\,\xi \le u_0 - u\,x^0$. Since $u_0 - u\,x^0 > 0$ one has

$$\frac{1}{u_0 - u\,x^0}\,u\,\xi \le 1, \qquad \xi \in P - x^0, \qquad (u, u_0) \in C.$$

The set $\left\{u' : u'\,\xi \le 1,\ \xi \in P - x^0\right\}$ is by definition the polar set of $P - x^0$. Since the vertices of the polar set define the facets of $P - x^0$, it is enough to set $u_0 - u\,x^0 = 1$ and maximize $u\,c$ for an arbitrary vector c. In conclusion, a facet of P can be found by solving the following LP problem

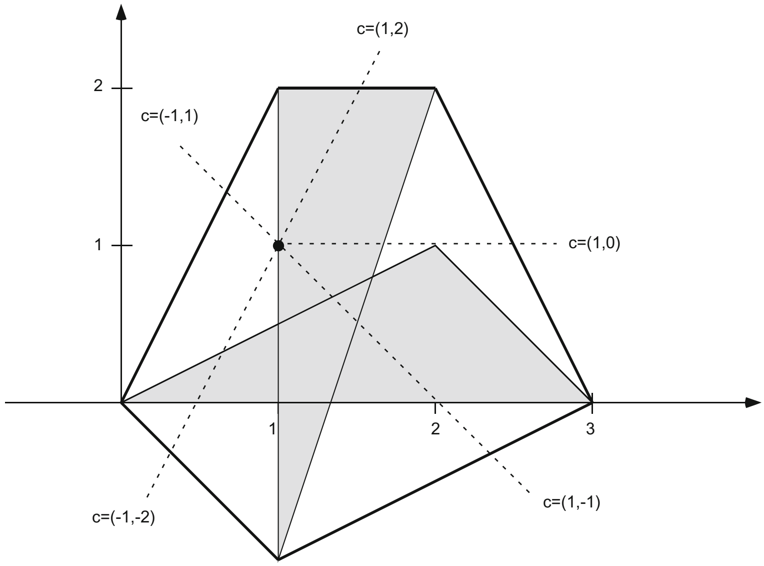

Fig. 2.5 Two polyhedra and their convex hull

$$
\begin{array}{llllll}
\max & & u\,c & & & \\
 & & u & -v^1 A^1 & & = 0 \\
 & & u & & -v^2 A^2 & = 0 \\
 & -u_0 & & +v^1 b^1 & & \le 0 \\
 & -u_0 & & & +v^2 b^2 & \le 0 \\
 & u_0 & -u\,x^0 & & & = 1 \\
 & & & v^i \ge 0 \quad i = 1,2 & &
\end{array}
\tag{2.17}
$$

In particular, the facet that is output by (2.17) is the one that has non empty intersection with the half line $x^0 + \alpha\, c$, $\alpha \ge 0$. This suggests the following choice for the vector c: suppose we have a point $\hat{x}$ lying outside P and we want to find a cutting inequality, i.e., an inequality that is both valid for P and makes $\hat{x}$ infeasible. This can be obtained by maximizing $u\,(\hat{x} - x^0)$. Due to the normalization choice $u_0 - u\,x^0 = 1$, maximizing $u\,(\hat{x} - x^0)$ is equivalent to maximizing $u\,\hat{x} - u^0$.

As an illustrative example consider the two triangles shown in Fig. 2.5. They are defined by the data:

$$
A^1 = \begin{pmatrix} 0 & -1 \\ -1 & 2 \\ 1 & 1 \end{pmatrix},\ b^1 = \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix}, \qquad A^2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \\ 3 & -1 \end{pmatrix}\ b^2 = \begin{pmatrix} 2 \\ -1 \\ 4 \end{pmatrix}.
$$

If we apply the Fourier elimination scheme, we obtain many inequalities among which the five facet-defining inequalities that are listed below. Alternatively, we assume to know the interior point $x^0 = (1, 1)$ and we solve several times (2.17) for different values of c. For instance we may obtain the following five inequalities that are facet-defining

c	u	u^0	inequality
$(1, 0)$	$(2/3, 1/3)$	2	$2\,x_1 + x_2 \le 6$
$(1, -1)$	$(1/4, -1/)$	3/4	$x_1 - 2\,x_2 \le 3$
$(-1, -2)$	$(-1/2, -1/2)$	0	$-x_1 - x_2 \le 0$
$(-1, 1)$	$(-2, 1)$	0	$-2\,x_1 + x_2 \le 0$
$(1, 2)$	$(0, 1)$	2	$x_2 \le 2$

If we impose the normalization $\|u\|_\infty \le 1$ we obtain the following LP problem

$$\begin{array}{lrrrrl}
\max & & u\,c & & & \\
 & & u & -\,v^1 A^1 & & = 0 \\
 & & u & & -\,v^2 A^2 & = 0 \\
 & -\,u_0 & & +\,v^1 b^1 & & \le 0 \\
 & -\,u_0 & & & +\,v^2 b^2 & \le 0 \\
 & & u & & & \le 1 \\
 & & -\,u & & & \le 1 \\
 & & & v^i \ge 0 & i = 1, 2 &
\end{array} \tag{2.18}$$

If we solve (2.18) for the same previous values of c we obtain the following inequalities that are valid although not necessarily facet-defining

c	u	u^0	inequality
$(1, 0)$	$(1, -1)$	3	$x_1 - x_2 \le 3$
$(1, -1)$	$(1, -1)$	3	$x_1 - x_2 \le 3$
$(-1, -2)$	$(-1, -1)$	0	$-x_1 - x_2 \le 0$
$(-1, 1)$	$(-1, 1)$	1	$-x_1 + x_2 \le 1$
$(1, 2)$	$(1, 1)$	4	$x_1 + x_2 \le 4$

2.6 Union of Polyhedra Defined by Vertices and Extreme Rays

If the polyhedra that have to be merged as a convex hull of the union are defined by their vertices and extreme rays, the situation is somehow simpler. Obviously this is possible only if the number of vertices and extreme rays is computationally tractable. The operation we have to carry out is very simple: we just put together all vertices and all extreme rays and take the convex combination of all vertices and the conic combination of all extreme rays. This operation introduces new variables, that is, the coefficients of the convex and the conic combinations. However, we have linear expressions that define a new polyhedron in a higher dimensional space and we have to project this polyhedron onto the space of the original variables.

Hence let us assume we have m polyhedra in $\mathbb{R}^n$ whose vertices and extreme rays are

$$v^h(i),\ h = 1, \dots p_i, \qquad r^k(i),\ k = 1, \dots, q_i, \qquad i \in [m]$$

Then we have

$$\begin{aligned} \sum_{i\in[m]} \sum_{h=1}^{p_i} \alpha_{ih}\, v^h(i) + \sum_{i\in[m]} \sum_{k=1}^{q_i} \mu_{ik}\, r^k(i) &= x \\ \sum_{i\in[m]} \sum_{h=1}^{p_i} \alpha_{ih} &= 1 \\ \alpha_{ih} \geq 0 \qquad \mu_{ik} &\geq 0, \end{aligned} \tag{2.19}$$

Now we may use the Fourier elimination scheme to obtain a representation in terms of inequalities in the space of the x variables.

As a very simple example, let us consider the example of p. 25 that we rewrite here for easy reference. There are two polyhedra in $\mathbb{R}^2$. P^1 is the half line $\{x : x_2 = 0, x_1 \geq 0\}$ that has just one vertex $v(1) = (0, 0)$ and one extreme ray $r(1) = (1, 0)$. P^2 is just the point $(0, 1)$ so that we have the unique vertex $v(2) = (0, 1)$. Hence the polyhedron in $\mathbb{R}^5$ that we have to project onto $\mathbb{R}^2$ is defined by (note that the equations appearing in (2.19) have been converted into pairs of inequalities in order to comply with (2.5))

$$\begin{pmatrix} 1 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 & 0 \\ -1 & 0 & 0 & 0 & 1 \\ 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & -1 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \alpha_1 \\ \alpha_2 \\ \mu \end{pmatrix} \leq \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

We think it may be useful for the reader to follow in detail the computation required by the Fourier elimination scheme. We start from the last column and eliminate one variable at a time going backward. The first step produces $6 + 2 \cdot 1 = 8$ inequalities, of which the only facet-defining are the following

$$\begin{pmatrix} 0 & 1 & 0 & -1 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \alpha_1 \\ \alpha_2 \end{pmatrix} \leq \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

The second step produces $2 + 3 \cdot 2 = 8$ inequalities, of which the only facet-defining are the following 5

$$\begin{pmatrix} 0 & 0 & -1 \\ -1 & 0 & 0 \\ 0 & -1 & -1 \\ 0 & -1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \alpha_1 \end{pmatrix} \leq \begin{pmatrix} 0 \\ 0 \\ -1 \\ 0 \\ 1 \end{pmatrix}.$$

The final step produces $2 + 2 \cdot 1 = 4$ inequalities, but one has null coefficients, so that this is the final result

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \leq \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

As a second slightly more complex example let us reconsider the example of the Fig. 2.5. In this case we have the data (there are no extreme rays)

$$v^1(1) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad v^2(1) = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad v^3(1) = \begin{pmatrix} 3 \\ 0 \end{pmatrix}$$

$$v^1(2) = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad v^2(2) = \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \quad v^3(2) = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

and, by applying (2.19), the polyhedron in $\mathbb{R}^8$ we have to project onto $\mathbb{R}^2$ is defined by the inequalities

$$\begin{pmatrix} 1 & 0 & 0 & -2 & -3 & -1 & -2 & -1 \\ 0 & 1 & 0 & -1 & 0 & -2 & -2 & 1 \\ -1 & 0 & 0 & 2 & 3 & 1 & 2 & 1 \\ 0 & -1 & 0 & 1 & 0 & 2 & 2 & -1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & -1 & -1 & -1 & -1 & -1 & -1 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \alpha_{11} \\ \alpha_{12} \\ \alpha_{13} \\ \alpha_{21} \\ \alpha_{22} \\ \alpha_{23} \end{pmatrix} \leq \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

We leave as an exercise for the reader to carry out the Fourier elimination scheme. Starting backwards from the last column, we finally get

$$\begin{pmatrix} 12 & -24 \\ -288 & -288 \\ 0 & 4 \\ -96 & 48 \\ 24 & 12 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \leq \begin{pmatrix} 36 \\ 0 \\ 8 \\ 0 \\ 72 \end{pmatrix}.$$

These can be clearly simplified as

$$\begin{pmatrix} 1 & -2 \\ -1 & -1 \\ 0 & 1 \\ -2 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \leq \begin{pmatrix} 3 \\ 0 \\ 2 \\ 0 \\ 6 \end{pmatrix}.$$

http://www.springer.com/978-3-319-63975-8

Compact Extended Linear Programming Models
Lancia, G.; Serafini, P.
2018, IX, 208 p. 37 illus., Hardcover
ISBN: 978-3-319-63975-8