Continuous Representation Techniques in Combinatorial Optimization

Oksana Pichugina¹, Sergey Yakovlev²
¹(Department of Applied Mathematics, Kharkiv National University of Radio Electronics, Ukraine)
²(Department of Computer Science, Kharkiv National Aerospace University, Ukraine)

Abstract: Approaches to an analytic description of vertex located discrete sets are presented. They are based on algebraic-topological features of the sets and properties of functions over them. One is a polyhedral-surface approach that represents a set as an intersection of its convex hull (a combinatorial polytope) and a circumscribed surface. Another one describes a set as an intersection of surfaces. These are applied for deriving various analytic representations of such sets as the Boolean set, binary set, permutation matrices set, the set of Boolean vectors with a given number of ones, and the cross-polytope vertex set. Among the sets' representations, polyhedral-spherical ones represent the sets by a quadratic system of constraints and tangent representations – by polynomial equations of a degree less or equal to four. These results open new opportunities for applying standard continuous optimization methods to combinatorial problems on vertex located sets, as well as for developing novel approaches, which use a specific of the sets and problems. Amongst relaxation techniques, the most promising directions are surface relaxations and algorithms that combine polyhedral and surface relaxations. In the sphere of reformulation approaches, it is important to develop new ways to convexify the original problem and then solve it to the optimum.

Keywords: Combinatorial Set, Continuous Functional Representation, Discrete Optimization, Lagrangian, The Boolean Set, The Permutation Set

I. Introduction

The traditional approach to constructing continuous relaxations of discrete problems is switching to optimization over the corresponding polyhedron. For that purpose, the problem of analytic description of the polyhedrons is solved. The problem is the main challenge in Polyhedral Combinatorics [1], [2], [3]. Among discrete sets mapped onto Euclidean space, there is a special class of vertex located sets [4] for whom there is applicable another type of relaxation – optimization on smooth convex surfaces circumscribed around them. A search of such surfaces along with finding extrema of some classes of functions is a type of problems deserving special attention.

We study issues of an analytic description of combinatorial sets mapped onto Euclidean space. Such sets, called Euclidean combinatorial sets [4], allows using geometric properties of the sets and the corresponding combinatorial polytopes in optimization over them. We use the geometric approach and describe the sets analytically as an intersection of surfaces and bodies, and the description we call a functional representation of the sets. We present classification of the sets' representations, in particular, we introduce concepts of tangent and irredundant representations. We describe two approaches to their construction and apply them to some combinatorial sets such as the Boolean set, the permutation set, the set of permutation matrices, and so forth. The common feature of the sets is, in addition to vertex locality, inscription into a hypersphere. It allows constructing a number of their functional representations: as an intersection of polytopes with surfaces (polyhedral-spherical representations); as an intersection of surfaces, in particular, tangent representations. The functional representations can be applied in optimization, namely, in reformulations of discrete problems in continuous variables, constructing continuous relaxations of the problems, regularizations of objective functions and so on. A traditional approach to the construction of continuous relaxations of discrete problems is to transfer them to optimization over the corresponding polytope [1]. In order to do this, an H-representation of the polytope is found. This, in itself, is the main challenge in Polyhedral Combinatorics [1]-[3].

Amongst discrete sets mapped onto Euclidean space (referred to as Euclidean combinatorial sets), there are a special class of vertex located sets [4] for which it is possible to apply another relaxation – optimization on smooth convex surfaces circumscribed around them. Our focus is to find such surfaces along with extrema of some classes of functions over these surfaces. For a given vertex located set $E$, an equation of the circumscribed surface with an H-representation of the polytope $P = \text{conv}E$ yields an analytic description of $E$. This allows reformulating optimization problems over $E$ in terms of continuous variables.
Continuous Representation Techniques in Combinatorial Optimization

However, other ways of continuous representations of discrete sets and, respectively, continuous reformulations of discrete problems are possible. We investigate the issues of an analytic description of vertex located sets with the aid of geometric features of the sets and properties of functions on the sets.

II. Problem Statement

Consider a discrete optimization problem over a finite set:

$$\min f(x), \quad x \in E \subseteq \mathbb{R}^n, \quad |E| < \infty.$$ (2.1)

Let $J_n = \{1, \ldots, n\}$. Likewise to any optimization problem, the feasible set of this discrete problem can be represented in the following form:

$$E = \{x \in D \subseteq \mathbb{R}^n : f_i(x) = 0, \ i \in J_m \}; f_i(x) \leq 0, \ i \in J_m \setminus J_m'. \}.$$ (2.4)

Traditionally, the condition $x \in D$ is referred to as regional, and the rest are the functional constraints [5]:

$$f_i(x) = 0, \ i \in J_m = \{1, \ldots, m'\}, \quad f_i(x) \leq 0, \ i \in J_m \setminus J_m'. \) (2.5) (2.6)

Depending on a choice of regional and functional constraints, different formulations of (2.1)-(2.3) are possible. For discrete problems, the set $D$ is usually discrete. Thus, the regional constraints describe discreteness of the feasible region, while (2.5), (2.6) are additional continuous conditions. As a consequence, firstly, combinatorial properties of $D$ are studied and underlie optimization techniques to (2.1)-(2.4). At the same time, there is a standard form (2.1),(2.5),(2.6) of an optimization problem [6], which corresponds to $D = \mathbb{R}^n$. If this form is chosen, firstly, it implies investigating properties of functions $f(x), f_i(x), i \in J_m$ over the feasible set and applying in optimization techniques.

If the standard form of the discrete optimization problem (2.1)-(2.3) was found, then continuous methods could be applied. For instance, if the hard constraints (2.5), (2.6) contain only the equality constraints (2.5) ($m = m'$), then the method of Lagrange multipliers [7] can be used. In the case of presence of inequality constraints (2.6) ($m' < m$), first-order necessary conditions for a solution such as the Karush-Kuhn-Tucker conditions (KKT) can be applied [7].

An issue is to find the standard form of discrete problems (referred to as Problem 1). If this form is not uniquely defined, which form should be chosen? We will refer to the problem as Problem 2. Another issue is that Problems 1, 2 require a specific solution for each type of discrete sets. The goal of our study is to find a solution of Problem 1 for several classes of combinatorial sets. Different solutions of the Problem will be presented along with their comparison, which is, in fact, an initial step to the solution of Problem 2.

We restrict our study to discrete sets coinciding with a vertex set of their convex hull:

$$E = \text{vert conv}(E).$$ (2.7)

called vertex located sets [4] and consider Problems 1.2 on them.

III. Related Work

In [13],[16],[24]-[26], functional representations of discrete sets have been considered and their classification was introduced. Here, we outline some terminologies necessary for presenting this material.

Definition. A functional representation of a set $E$ is its representation in the form (2.5), (2.6).

Thus, Problem 1 is a problem of a search of a functional representation of $E$.

Denote sets of points in $\mathbb{R}^n$ given by (2.5), (2.6) as $\{S_i\}_{i \in J_m}$:

$$S_i = \{x \in \mathbb{R}^n : f_i(x) = 0\}, i \in J_m'.$$ (3.1)
Continuous Representation Techniques in Combinatorial Optimization

\[ S_i = \{ x \in \mathbb{R}^n : f_i(x) \leq 0 \}, \ i \in J_m \setminus J_{m'} . \]  

Then (2.5), (2.6) can be rewritten as:

\[ E = \bigcap_{i \in J_m} S_i . \]  

Depending on a type of functions in (2.5), (2.6) the functional representations can be linear and non-linear, continuous and differentiable, smooth, convex (on a convex set) and so on.

The functional \( E \)-representation is called:

- **strict** if \( m' = m \), otherwise **non-strict**;
- **irredundant** if eliminating of any constraint from (2.5), (2.6) leads to a violation of (3.3):

\[ \forall j \in J_m \setminus \bigcap_{i \in J_m, i \neq j} S_i \neq E , \]

otherwise, it is **redundant**.

The number \( m, m', m - m' \) is the **order** of a functional representation, its strict and non-strict parts, respectively.

Concerning the number of functions in the strict part (2.5) of an irredundant representation of a discrete set, the representation is called:

- **intersecting** if \( m' = n \) (this implies that \( E \) is formed as an intersection of \( n \) surfaces);
  - **tangent** if \( m' = 2 \) and \( E \) is a subset of touching points of \( S_1, S_2 \) that satisfies (2.6);
- **one-component** if \( m = 1 \);
- **otherwise** mixed.

A solution of Problem 1 is a way to reformulate an original discrete problem in terms of continuous variables. In other words, it is a way of its **continuous reformulations**. These approaches can be found in literature [10],[11]. Indeed, methods of combinatorial optimization are roughly divisible into discrete and continuous [11]. Continuous algorithms, in turn, are classified into continuous relaxations and reformulations [10].

Let us consider a class of the unconstrained binary problems (UBP). In our notations, it is (2.1), (2.2) on \( E = B_n \). Moreover, this problem belongs to the class (2.1), (2.7) since points of \( B_n \) are vertices of the unit hypercube \( PB_n = [0,1]^n \).

Outline some continuous approaches to (UBP): a typical relaxation is **semi-definite** [27] and one way of continuous reformulation is to lift into a higher dimensional space by adding new variables. For instance, instead of the discrete constraint:

\[ x \in B_n , \]  

the following lifting technique can be used [10]:

\[ x_i + y_i = 1, x_i \cdot y_i = 0, x_i, y_i \geq 0, i, j \in J_n . \]

If \( f(x) \) is quadratic, \( f(x) = \sum_i a_{ij} x_i y_j + \sum_i b_{ij} x_i \), the change of variables \( y_{ij} = x_i x_j \) allows transforming (UBP) into linear in an extended space by \( C^2 \)-variables [28].

Also, an alternating analytic representation of \( B_n \) is known and widely applied in optimization [10], [11],[13]:

\[ f_i(x) = x_i (x_i - 1) = 0, i \in J_n . \]  

In our classification, (3.5) is a strict, continuous, convex, quadratic, irredundant, intersecting functional representation of \( B_n \). The problem (2.1), (3.5) is an equivalent reformulation of (UBP) in the original space and a Problem 1 solution.

Now, continuous optimization approaches, such as the method of Lagrangian multipliers, penalty methods, their combinations such as the augmented Lagrangian method, smoothing techniques [7], [11] and so forth become applicable to (UBP).

For instance, an equivalent reformulation of (2.1), (3.5) in a higher dimensional space is to minimize the Lagrangian:

\[ \Phi(x, \lambda) = f(x) + \sum_i \lambda_i (x_i^2 - x_i) \rightarrow \min, \lambda = (\lambda_i)_{i \in J_n} . \]  

DOI: 10.9790/5728-1302051225  www.iosrjournals.org  14 | Page
The problem (3.6) can be solved explicitly for simple \( f(x) \), in other cases, numerical schemes should be applied [10].

The penalty method with a quadratic penalty function [13],[24] utilizes the representation (3.5) in the following way:

\[
\Phi'(x, \mu) = f(x) + \mu \sum_{i} (x_i^2 - x_i) \to \min.
\]  

(3.7)

The problems (2.1) and (3.7) are equivalent only on \( E \). However, if the penalty method is applied to the problem (3.7) for an increasing sequence of \( \mu \), it approaches a local minimizer of (UBP) with any prescribed accuracy.

Also, penalty methods may be combined with other techniques. For example, in [11] the quadratic penalty function (3.7) and \( PB_n \)-constrains (boxing constraints) are used to smooth and convexify an objective function.

As a result, (UBP) is reduced to a series of convex subproblems.

One more (UBP) relaxation [10],[13],[24] based on (3.5) is the following:

\[
\Phi^*(x, \lambda) = f(x) + \lambda \sum_{i} (x_i^2 - x_i) = f(x) + \lambda \left( \sum_i \left( x_i - \frac{1}{2} \right)^2 - \frac{n}{4} \right) \to \min.
\]  

(3.8)

It is seen that in (3.8), (UBP) is reduced to optimization over a hypersphere circumscribed around \( PB_n \).

An interesting property of (3.8) is that an appropriate choice of \( \lambda \) allows making \( \Phi^*(x, \lambda) \) either convex or concave. Namely, if \( f(x) \in C^2(R^n) \) then there are \( \Lambda_{\min}, \Lambda_{\max}, \Lambda_{\min} \leq \Lambda_{\max} \), such that the problem (3.8) is convex \( \forall \lambda \geq \Lambda_{\max} \) and concave \( \forall \lambda \leq \Lambda_{\min} \) [7],[10]. This allows usage of convex and concave methods to (3.8). Moreover, the regularization term \( \lambda \left( \sum_i (x_i - 0.5)^2 - 0.25n \right) \) can be added to (3.6), (3.7) as well.

Then, for fixed \( \lambda, \mu \), a choice of \( \lambda \) allows making \( \Phi(), \Phi'() \) convex or concave and again to apply continuous methods to (UBQ).

Note that this technique allows also forming equivalent reformulations of (UBP). For instance, if \( \lambda \) is chosen such that \( \Phi^*(x, \lambda) \) is concave then adding the boxing constraints to (3.8) transforms the problem into equivalent to (UBP).

**IV. Approaches to Problem**

Construction of functional representations can be done in a different way resulting different solutions of Problem 1 [12-16]. Outline three of them applicable to vertex located sets (2.7).

**Approach 1 (a polyhedral-surface approach)** includes two stages. Firstly, an \( H \)-representation

\[
P = \{ x : Ax = b, A'x \leq b', A \in R^{m-n}^{\times n}, A' \in R^{m-n}^{\times m} \},
\]  

(4.1)

of the polytope \( P \) is found.

Then a surface \( S \) circumscribed around \( E \) is derived:

\[
S = \{ x : f_i(x) = 0 \} : E \subset S, f_i(x) \text{ is strongly convex}.
\]  

(4.2)

Strong convexity of \( f_i(x) \) ensures that \( E \) is formed as an intersection of the polytope (4.1) and surface (4.2):

\[
E = P \cap S.
\]  

(4.3)

Hence, the linear system of \( P \) with the non-linear equation of \( S \) gives a non-linear, convex, non-strict functional representation of \( E \) (called a polyhedral-surface \( E \)-representation [25]) and a solution of Problem 1.

**Remark 1.** Note a particular case of polyhedral-surface representations when \( S \) is a hypersphere.

Let \( E \) be inscribed in a hypersphere:

\[
E \subset S_r(x^0),
\]  

(4.4)

where

\[
S_r(x^0) = \{ x \in R^n : (x - x^0)^2 = r^2 \}
\]  

(4.5)
is a hypersphere of the radius \( r \) centered at \( x^0 \in \mathbb{R}^n \). Then the set (4.4) is called sphere located [8]. Also, the representation (4.3) is the polyhedral-spherical representation (PSR) of \( E \) [16],[25] if it contains the equation (4.5) of a circumscribed sphere:

\[
(PSR) \quad E = P \cap S(x^0).
\]

(PSR) is important since many combinatorial sets are sphere located and H-representations of their polytopes are known [11],[2],[4],[14],[15],[17]. Also, properties of spheres allow developing specific approaches to (2.1)-(2.3),(4.4) such as the polyhedral-spherical method (PSM) [13] of combinatorial optimization over sphere located sets.

Note, the representation (4.1)-(4.3) is irredundant if and only if (4.1) is an irredundant system of \( P \). This implies that eliminating of any number of restrictions from the \( E \)-representation leads to a continuous relaxation of the original problem.

For instance, relaxing the condition \( x \in S \) yields the standard polyhedral relaxation of (2.1),(2.7) \( f(x) \rightarrow \min_p \). A relaxation of the condition \( x \in P \) yields another continuous relaxation of the problem called a surface relaxation that is an optimization problem over a convex surface \( S : f(x) \rightarrow \min \).

The combination of both these relaxations is a core of a (PSM)-generalization called the polyhedral-surface method to (2.1),(2.7) [25].

Next two approaches are based on exploring properties of non-linear functions over a set \( E \), deriving those of them that take constant values over the set, and construction of \( E \)-representations from these functions.

**Approach 2.** Similarly to Approach 1, two main stages can be extracted: 1) constructing a system of equations, a solution of which is a discrete set \( E' \supseteq E \); 2) if necessary, cut off points of \( E' \backslash E \) and adding the cuts to the constraints found in the first stage.

Note that in \( n \)-dimensional space a discrete set is formed by an intersection of at most \( n \) surfaces. Thus the first stage may be reduced to consideration of systems of order \( m' = n \). When stage 2 has been completed, a redundancy of the resulting functional representation may be investigated, and the order \( m' \) of its strict part can be decreased.

This approach is rather complicated as it requires solving non-linear systems analytically. However, it is the only known way to derive a strict representation of such sets as \( E_{ab}(G) \) [26].

The next approach allows forming an irredundant representation directly.

**Approach 3** is representing \( E \) as a set of touching points of two surfaces \( S_1, S_2 \):

\[
E = S_1 \cap S_2, \quad (4.6)
\]

where

\[
S_i = \{ x \in \mathbb{R}^n : f_i(x) = 0 \}, \quad f_i(x) \text{ are differentiable, } i = 1,2. \quad (4.7)
\]

Let us assume that the equations (4.7) of surfaces circumscribed around \( E \) are known. Then there are two ways to derive the representation (4.6).

The first one is to solve the following problems:

\[
f_1(x) \rightarrow \min_{S_1}, \quad f_2(x) \rightarrow \min_{S_1}. \quad (4.8)
\]

If a solution of one of them is attained at \( E \), the tangent representation (4.6) has been found.

Another way is to find all extrema of one of the functions, \( f_1(x) \) or \( f_2(x) \), over another surface, e.g.

\[
f_2(x) \rightarrow \text{extr}_{S_1}. \quad (4.9)
\]

**V. Notations**

Let \( E \in \mathbb{R}^n \). Introduce some notations and terminologies:

- \( x \) is a \( n \)-vector, \( X \) is a \( n \times n \)-matrix: \( x = [x_i] \in \mathbb{R}^n \), \( X = [x_{ij}] \in \mathbb{R}^{nm} \);

- \( a \) is a \( n \)-vector of \( a \)-s, \( A = [a] \in \mathbb{R}^n \), \( A = [a] \in \mathbb{R}^{nm} \), where \( a \in \mathbb{R} \);

- \( E^m \) is the Cartesian power of \( E \), \( E^m = \bigotimes_{i=1}^{m} E \);

- \( \forall \alpha \in \mathbb{R} \quad \alpha E = \{ \alpha x : x \in E \} \);
• $G$ is a $n$-element multiset, which the underlying set $\{e_j\}_{j \in J_n}$ consists of $k$ different elements with
multiplicities $n_j$, $j \in J_k$, such that $\sum_j n_j = n$:

$$G = \{e_j^n\}_{j \in J_n}.$$  \hspace{1cm} (5.1)

We will consider the following vertex located combinatorial sets induced by a multiset (5.1) [2],[4],[14],[15],[17],[20]:

1. $E_{nk}(G)$ – is the general permutation set induced by (5.1).

2. $B_n = \{0,1\}^n$, $B'_n = \{-1,1\}^n$ are sets of Boolean/binary $n$-vectors (the Boolean set and binary set), respectively. For them, the induced multisets are $G = \{0^n, 1^n\}$, $G = \{-1^n, 1^n\}$, correspondingly;

3. $B_n(m)$ is a set of Boolean $n$-vectors that sum to $m$:

$$B_n(m) = \{x \in B_n : x^T 1 = m\}, m \in J_n^0.$$  

Interestingly that $B_n(m)$ is simultaneously a subset of $B_n$ and the general permutation set:

$$B_n(m) \subset B_n, B_n(m) = E_{n2}(G) : G = \{0^{n-m}, 1^m\}.$$  

In particular, $B_n(1)$ is a set of 0–1-permutations with one unit:

$$B_n(1) = E_{n2}(G),$$  \hspace{1cm} (5.2)

where

$$G = \{0^{n-1}, 1\;}.$$  \hspace{1cm} (5.3)

4. a set $CE_n(b)$ of $n$-vectors induced by $n-1$ zeros, $b>0$ and $-b$:

$$CE_n(b) = \pm b \cdot B_n(1) = E_{n2}(b \cdot G) \cup E_{n2}(-b \cdot G)$$  

where $G$ is defined by (5.3);

5. a set $\Pi_n$ of $n$-permutation matrices:

$$\Pi_n = \{X \in \mathbb{R}^{en} : X1 = X^T 1 = 1, x_{ij} \in \{0,1\}, i, j \in J_n\}.$$  \hspace{1cm} (5.4)

Notice, that if a matrix $X \in \Pi_n$, then its vector-columns $x_j = \{x_{ij}\}_{i \in J_n} (j \in J_n)$ and vector-columns $x'_j = \{x_{ij}\}_{i \in J_n} (j \in J_n)$ are elements of the set (5.2):

$$x_j, x'_j \in B_n(1), i, j \in J_n.$$  \hspace{1cm} (5.5)

We map the matrix $X$ into $\mathbb{R}^{e^2}$: $X \rightarrow x$ where $x = \bigotimes_{i \in J_n} x_i$ and safe double indices for its coordinates.

Similarly, a $n^2$-vector corresponds to $X' = X^T$: $X' \rightarrow x'$ where $x'^T = \bigotimes_{i \in J_n} x'_j$. The map of $\Pi_n$ into $\mathbb{R}^{e^2}$ denote $\Pi'_n$. There is a one-to-one correspondence between the pairs $X$ and $x$. Therefore, instead of matrices, we can consider their images as points in the Euclidean space $\mathbb{R}^{e^2}$. Sets of objects of arbitrary nature that allow mapping into the Euclidean space, alongside with the images, are called the Euclidean combinatorial sets [4]. In this manner, to the image $\Pi'_n$ of the whole set (5.4) we will refer as called the Euclidean combinatorial set of permutation matrices. It can be represented in two ways:

$$\Pi'_n = \{x \in (B_n(1))^{en} : X1 = 1\};$$

$$\Pi'_n = \{x' \in (B_n(1))^{n^2} : X1 = 1\}.$$  \hspace{1cm} (5.6)

**VI. $B_n, B_n(1), CE_n(1), \Pi_n$ FUNCTIONAL REPRESENTATIONS**

In this section, Problem 1 is solved for the following combinatorial sets:

$$E \in \{B_n, B_n(1), CE_n(1), \Pi_n\}.$$  \hspace{1cm} (6.1)

First, in Sect. VI.1 Approach 1 to the construction of functional representations is applied to the sets (6.1). Then, in Sect. VI.2 tangent representations of three combinatorial sets are constructed with the aid of Approach 3. Finally, in Sect. VI.3 some tangent and other strict representations of the sets (6.1) are presented along with their
6.1 THE SETS’ POLYHEDRAL-SPHERICAL REPRESENTATIONS

It turns out, a common feature of the combinatorial sets (6.1) is a sphere locality [4],[13],[25],[26]. Also, an H-representation of the corresponding polytopes are known [2],[10],[17],[20]. Thus, Approach 1 can be applied to their analytic representation and yields polyhedral-spherical representations of these sets. To specify each of them, list the polytopes H-representations in the form (4.1) and parameters of hyperspheres (4.5).

Some of the sets (6.1) are inscribed into families of hyperspheres. Among them, we will indicate a sphere $S_r(0)$ centered at the origin and a minimum circumscribed sphere $S_{\min}$.

So, for the sets (6.1), the polytope $P$ and sphere $S_r(x^0)$ in (PSR) are defined by:

1. $E = B_n$:

$$P = PB_n = \{ x \in R^n : 0 \leq x \leq 1 \}, \quad S_r(x^0) = S_{\min} = \frac{1}{\sqrt{n}} \cdot \frac{1}{2}.$$ 

2. $E = B_n(1)$:

$$P = \text{conv}(B_n(1)) = PB_n(1) = \text{a unit } n - 1 \text{-simplex} \quad PB_n(1) = \{ x \in R^n : x \geq 0, x^T 1 = 1 \}.$$ 

3. $E = CE_n(1)$: $P = \text{conv}(CE_n(1)) = CP_n(1)$ is the cross-polytope [20]:

$$CP_n(1) = \{ x \in R^n : |x^T 1| \leq 1 \},$$

where $|x| = (|x_j|)_{j \in J_n}$, $S_r(x^0) = S_r(0) = S_{\min}$.

4. $E = \Pi_n$ – is the Euclidean combinatorial set of permutation matrices. A convex hull of the set of permutation matrices $D_n = \text{conv}(\Pi_n)$ is the Birkhoff polytope [17]:

$$D_n = \{ X \in R^{m \times n} : X \succeq (0^n), X 1 = X^T 1 = 1 \},$$

hence

$$D_n = \text{conv}(\Pi_n) = \{ x \in R^n : x_{ij} \geq 0, i, j \in J_n; \sum_{i=1}^n x_{ij} = 1, j \in J_n; \sum_{j=1}^n x_{ij} = 1, i \in J_n \}.$$ (6.5)

By (5.6), (6.2), the set $\Pi_n$ is inscribed into a family of hyperspheres, $\Pi_n \subset S_{r(a^n)}(a^n)$, which can be derived from (6.2):

$$S_{r(a^n)}(a^n), r'(a) = \sqrt{n \cdot r(a)} = \sqrt{n^2 \cdot a^2 - 2a \cdot n + n}.$$ (6.6)

The family (6.6) includes the following hyperspheres of our interest – $S_{\min}(0^n)$ and $S_{\min} = S_{\sqrt{m}((1/n)^n)}$.

The above H-representations are irredundant, except for $D_n$. (6.5) becomes irredundant after eliminating anyone constraint from:

$$X 1 = X^T 1 = 1.$$ (6.7)

Complementing each of the H-representation by an equation of $S_r(x^0)$, we get irredundant representations of the sets (6.1). Moreover, for $B_n(1)$ and $\Pi_n$ families of such representations are obtained.

Example 1. For $n = 3$, a graphical illustration of the sets (6.1), their polyhedrons, and (PSR) is given in
Figs. 1, 2:
- the hypercube $PB_3$ is in Fig. 1.a; the cross-polytope $CP_3(1)$ – is in Fig. 1.b; the 2-simplex $PB_3(1)$ – is in Fig. 1.c);
- (PSR) of: $B_3$ is shown in Fig. 2.a); $CP_3(1)$ – is in Fig. 2.b); $B_3(1)$ – is in Fig. 2.c). It includes a sphere $S_1(0)$.

6.2 TANGENT REPRESENTATIONS OF $B'_n$, $A_n(\sqrt{n})$, $B_n(1)$

We search for a tangent representation using the second way presented in Sect. IV and rewriting (4.9) in the form:

$$f(x) \to \text{extr}_{x \in S},$$

where the surface $S = \{x: f_i(x) = 0\}$. $f(x) = f_2(x)$.

For minimization and maximization version of (6.8), denote its solutions:

$$z^\text{min} = \min_S f(x), \quad X^\text{min} = \text{Argmin}_S f(x); \quad z^\text{max} = \max_S f(x), \quad X^\text{max} = \text{Argmax}_S f(x).$$

Now, we formulate main results of the paper.

**Theorem 1.** The solution of (6.8) with

$$f(x) = \sum_{i=1}^n x_i^4$$

over a hypersphere

$$S = S_{\sqrt{n}}(0)$$

is

$$z^\text{min} = n, \quad X^\text{min} = B'_n;$$

$$z^\text{max} = n^2, \quad X^\text{max} = CE_n(\sqrt{n}).$$

**Proof.** The problem (6.8)-(6.10) is equivalent to the following:
Continuous Representation Techniques in Combinatorial Optimization

\[ H(x, \lambda) = f(x) + \lambda \left( n - \sum_{i=1}^{n} x_i^2 \right) = \sum_{i=1}^{n} x_i^4 + \lambda \left( n - \sum_{i=1}^{n} x_i^2 \right) \rightarrow \text{extr} . \]  

(6.13)

Setting the gradient \( \nabla_{x, \lambda} H(x, \lambda) = 0 \) we get the following system for determining stationary points of \( H(x, \lambda) \):

\[
\frac{\partial H}{\partial x_i} = 4 \cdot x_i^3 - \lambda \cdot 2x_i = 0, \quad i \in J_n ;
\]  

(6.14)

\[
\frac{\partial H}{\partial \lambda} = n - \sum_{i=1}^{n} x_i^2 = 0.
\]  

(6.15)

Denote a stationary point as \( (x^*, \lambda^*) \). Then (6.13) gives:

\[
\frac{\partial H}{\partial x_i} = 0 \implies x^*_i \in \left\{ 0, \pm \sqrt{\frac{\lambda^*}{2}} \right\}, \quad i \in J_n .
\]  

(6.16)

Let \( I \) be a set of indices of non-zero \( x^*_i \)-coordinates and \( m \) be a number of such indices:

\[
I \subseteq J_n : i \in I \iff x^*_i \neq 0, \quad m = |I| .
\]  

(6.17)

Then, in our case, \( I = \{ i : x^*_i = \pm \sqrt{\frac{\lambda^*}{2}} \} \) and \( I \neq \emptyset \) that follows from (6.15).

From (6.15), \( \sum_{i=1}^{n} x_i^{*2} = \sum_{i \in I} x_i^{*2} + \sum_{i \notin I} x_i^{*2} = m \frac{\lambda^*}{2} = n \) wherefrom

\[
\lambda^* = \frac{2n}{m} .
\]  

(6.18)

Moreover, by (6.16),

\[
x_i^* = \begin{cases} \pm \sqrt{\frac{\lambda^*}{2}} = \pm \sqrt{\frac{n}{m}}, & i \in I, \\ 0, & \text{otherwise}. \end{cases}
\]  

(6.19)

Substitution of (6.18), (6.19) into the Lagrangian yields:

\[
H(x^*, \lambda^*) = \sum_{i \in I} x_i^{*4} + \lambda^* \cdot \left( n - \sum_{i \notin I} x_i^{*2} \right) = m \left( \frac{n}{m} \right)^2 + \lambda^* \cdot \left( n - m \cdot \frac{n}{m} \right) = \frac{n^2}{m} .
\]

Thus, (6.13) and, respectively, the original problem (6.8) are equivalent to the following one:

So, the problems \( (x^*, \lambda^*) \) and, respectively, (6.8), are equivalent to the following:

\[
h^*(m) = \frac{n^2}{m \cdot \text{me} J_n} \rightarrow \text{extr} .
\]  

(6.20)

Thus, the original problem was reduced to optimization of the function \( h^* \) of one variable, which is decreasing and attains the extrema at endpoints of \( \left\{ 1, n \right\} \). Namely, by (6.18) and (6.19),

- the minimum of (6.20) occurs at \( m = n \) and, hence \( I = J_n \):

\[
\lambda_{\min} = 2, \quad \forall i \in J_n, \quad x_{i,\min} = \pm 1, \quad h_{\min} = \min_{m \in J_n} h^*(m) = h^*(n) = n ;
\]  

(6.21)

- its maximum occurs at \( m = 1 \) implying that \(|I| = 1 \) and \( \exists i^0 \in J_n : I = \{ i^0 \} \). Correspondingly,

\[
\lambda_{\max} = 2n, \quad \forall i \in J_n, \quad x_{i,\max} = \begin{cases} \pm \sqrt{n}, \quad i = i^0, \\ 0, \quad \text{otherwise}; \end{cases} \quad h_{\max} = \max_{m \in J_n} h^*(m) = h(1) = n^2 .
\]  

(6.22)
Due to an equivalence of (6.8)-(6.10) and (6.20): a) from (6.21) we have $X^\min = B_n'$, $z^\min = h^\min = n$, thus (6.11); b) from (6.22), $X^\max = CE_n(\sqrt{n})$, $z^\max = h^\max = n^2$, hence (6.12) holds.

**Corollary 1.** The sets $B_n'$, $CE_n(\sqrt{n})$ have the following convex tangent representations in the form (2.5):

$$
B_n' : f_1(x) = \sum_{i \in J_n} x_i^2 - n = 0, \quad f_2(x) = \sum_{i \in J_n} x_i^4 - n = 0 ;
$$

(6.23)

$$
CE_n(\sqrt{n}) : f_1(x) = \sum_{i \in J_n} x_i^2 - n = 0, \quad f_2(x) = \sum_{i \in J_n} x_i^4 - n^2 = 0 .
$$

(6.24)

**Theorem 2.** The solution of (6.8) with

$$
f(x) = \sum_{i=1}^{n} x_i^3
$$

over a hypersphere

$$
S = \{ x : \sum_{i=1}^{n} x_i^2 = 1 \}
$$

(6.26)

is $z^\min = -1, X^\min = -B_n(1)$;

$z^\max = 1, X^\max = B_n(1)$.

(6.27)

(6.28)

**Proof.** Prove the theorem in similar way to Theorem 1. The Lagrangian of (6.8), (6.9), (6.26) is

$$
H(x, \lambda) = f(x) + \lambda \left( 1 - \sum_{i=1}^{n} x_i^2 \right) = \sum_{i=1}^{n} x_i^3 + \lambda \left( \sum_{i=1}^{n} 1 - x_i^2 \right) \rightarrow \text{extr} .
$$

(6.29)

Setting the gradient of (6.29) equal to $0$, we obtain:

$$
\frac{\partial H}{\partial x_i} = 3x_i^2 - \lambda \cdot 2x_i = 0, \quad i \in J_n ;
$$

(6.30)

$$
\frac{\partial H}{\partial \lambda} = 1 - \sum_{i=1}^{n} x_i^2 = 0 .
$$

(6.31)

From (6.30):

$$
\frac{\partial H}{\partial x_i} \bigg|_{x_i = x_i^*, \lambda} = x_i^* \cdot (3x_i^* - 2 \cdot \lambda^*) = 0 \Leftrightarrow x_i^* \in \left\{ 0, \frac{2}{3} \lambda^* \right\}, \quad i \in J_n .
$$

(6.32)

Respectively, from (6.31) and (6.32) $\sum_{i=1}^{n} x_i^2 = \sum_{i \in J_n} x_i^2 = m \left( \frac{2}{3} \lambda^* \right)^2 = m \cdot \frac{4}{9} \lambda^* = 1$. As a consequence,

$m > 0$ and

$$
\lambda^* = \pm \frac{3}{2\sqrt{m}} .
$$

(6.33)

According to (6.17), (6.32), and (6.33),

$$
x_i^* = \begin{cases} 
\pm \frac{2}{3} \lambda^* = \pm \frac{1}{\sqrt{m}}, & i \in I, \\
0, & \text{otherwise}.
\end{cases}
$$

(6.34)

The set $I$ is divisible into to subsets $I^+, I^-$ of indices of positive and negative $x_i^*$-coordinates:

$I^+ = \{ i : x_i^* = \frac{1}{\sqrt{m}} \}$, $I^- = \{ i : x_i^* = -\frac{1}{\sqrt{m}} \}$. Also, denote $|I^+| = m'$, then $|I^-| = m - m'$.

Now, (6.34) can be rewritten in the form:
Continuous Representation Techniques in Combinatorial Optimization

\[ x_i^* = \begin{cases} 
\frac{1}{\sqrt{m}}, & i \in I^+, \\
-\frac{1}{\sqrt{m}}, & i \in I^-, \\
0, & \text{otherwise}. 
\end{cases} \tag{6.35} \]

Substituting (6.35) into (6.29) we have:

\[ H(x^*, \lambda^*) = \sum_{i \in I^+} x_i^3 + \sum_{i \in I^-} \left( -\frac{1}{\sqrt{m}} \right)^3 = \frac{2m' - m}{m^{1.5}}. \tag{6.36} \]

It is seen, the problem of optimizing the function (6.36) has been reduced to optimization of a function \( h^*(m, m') = \frac{2m' - m}{m^{1.5}} \) of two variables. Within \( 0 \leq m' \leq m \leq n \), it attains:

- the maximum at \( (m, m') = (1,1) \): \( h^{\max} = h^*(1,1) = 1 \);
- the minimum at \( (m, m') = (1,0) \): \( h^{\min} = h^*(1,0) = -1 \).

Respectively, the extrema of \( f(x) - \lambda^* = h^\max - h^*(1,1) = 1, h^{\min} = h^*(1,0) = -1 \) occur at \( X^{\max}, X^{\min} \):

\[ \forall x^{\max} \in X^{\max} \exists i_0 \in J_n^+ : x_i^{\max} = \begin{cases} 1, i = i_0, \\
0, & \text{otherwise}, \end{cases} \Rightarrow X^{\max} = B_n(1); \]

\[ \forall x^{\min} \in X^{\min} \exists i_0 \in J_n^- : x_i^{\min} = \begin{cases} -1, i = i_0, \\
0, & \text{otherwise}, \end{cases} \Rightarrow X^{\min} = -B_n(1). \]

Hence, (6.27) and (6.28) hold.

**Corollary 2.** The set \( B_n(1) \) has a tangent representation:

\[ B_n(1) : f_1(x) = \sum_{i \in J_n^+} x_i^2 - 1 = 0, \quad f_2(x) = \sum_{i \in J_n^-} x_i^3 - 1 = 0. \tag{6.37} \]

### 6.3 Strict Representations of \( B_n \), \( B_n(1) \), \( CE_n(1) \), \( \Pi_n \)

In this section, the results of Sect. VI.2 are used for the construction of functional representations of the sets (6.1). New functional representations are formed from above by shifting, scaling and them applying them to specific sets of variables:

1. \( B_n^+ \):

   a. **Representation 1.1** – making in (6.23) the change of variables \( x_i \rightarrow x_i + \frac{1}{2}, i \in J_n^+ \), we obtain a tangent representation of \( B_n^+ \):

      \[ B_n^+ : f_1(x) = \sum_{i \in J_n} \left( x_i - \frac{1}{2} \right)^2 - \frac{n}{4} = 0, \quad f_2(x) = \sum_{i \in J_n} \left( x_i - \frac{1}{2} \right)^4 - \frac{n}{16} = 0; \tag{6.38} \]

These representations of the Boolean sets for dimensions \( n = 2, 3 \) are illustrated in Fig.3, e.g., in Fig.3 a) there is the representation (6.23) for \( B_n^+ \); for \( B_n^- \) there are Representation 1.1 and Representation 1.2 in Figs.3 b), 3 c).
Continuous Representation Techniques in Combinatorial Optimization

Fig. 3. $B'_2$: the representation (6.23), $B'_3$: Representations 1.1 and 1.2

(b) $B'_2$: Representations 1.1
(c) $B'_3$: Representations 1.2

Example 2. The representation (6.23) of $B'_2$ is shown in Fig. 3.a). The two above representations of the Boolean set for $n = 3$ – Representation 1.1 and Representation 1.2 – are illustrated in Figs. 3.b),c), respectively.

2. $B_n(1)$: in addition to the tangent representation (6.37) (referred to as Representation 2.1), two mixed strict representations are derivable by using the fact that $B_n(1)$ lies in the plane $x^T 1 = 1$ and the hypersphere (6.3). So, adding their equations to Representation 1.1, we obtain:

a. **Representation 2.2** of $B_n(1)$: Representation 1.1 plus

$$f_j(x) = \sum_{i \in J} x_i - 1 = 0; \quad (6.39)$$

b. **Representation 2.3** of $B_n(1)$: Representation 1.1 plus

$$f_j(x) = \sum_{i \in J} \left(x_i - \frac{1}{n}\right)^2 - 1 + \frac{1}{n} = 0 \quad (6.40)$$

Example 3. An illustration of Representation 2.1 for $B_2(1)$, $B_3(1)$ is given in Figs. 4.a), 4.b). $B_3(1)$ – Representation 2.3 is shown in Fig. 4.c).

3. $CE_n(1)$: **Representation 3** is a tangent representation of $CE_n(1)$ obtained from (6.24) by the change of variables $x_i \rightarrow \frac{x_i}{\sqrt{n}}, \ i \in J_n$:

$$CE_n(1): f_1(x) = \sum_{i \in J} x_i^2 - 1 = 0, \ f_2(x) = \sum_{i \in J} x_i^4 - 1 = 0. \quad (6.41)$$
Example 4. Representation 3 is demonstrated: for $CE_2(1)$ in Fig.5.a), for $CE_3(1)$ in Fig.5.b). Based on Representation 3 and the fact that $B_n(1) = B_n \cap CE_n(1)$, new $B_n(1)$-representation can be formed: a) Representations 1.1 and 3; b) Representations 3 plus (6.39); c) Representations 3 plus (6.40).

Fig. 5. Representation 3 ($n = 2, 3$)

4. $\Pi_n$:

Representation 4.1 is formed by the linear system (6.7) and a generalization of Representation 1.1 from a Boolean $n$-vector $\mathbf{x}$ into a Boolean $n^2$-vector:

$$f_j(x) = \sum_{i,j \in J_n} \left( x_{ij} - \frac{1}{2} \right)^2 - \frac{n^2}{4} = 0; \quad f_j(x) = \sum_{i,j \in J_n} \left( x_{ij} - \frac{1}{2} \right)^4 - \frac{n^2}{16} = 0.$$  \hspace{1cm} (6.42)

a. Representation 4.1 is of order $2n+2$, its components are linear and convex quadratic and biquadratic polynomials. So, it is a mixed, nonlinear, polynomial, convex, and smooth analytic representation of $\Pi_n$, which is redundant due to the redundancy of (6.7). Similarly to the $D_n$-system of constraints (see (6.5)), eliminating any constraint from (6.7) converts Representation 4.1 into irredundant.

b. Representation 4.2 is formed from Representation 4.1 by replacing the linear part (6.7) with quadratic. Namely, we use the property (5.5) wherefrom there follows an inscription of $x_i, x'_i, i, j \in J_n$, into a hypersphere (6.3) and get the representation (6.42) plus:

$$f_{2x/j}(x'_j) = \sum_{i \in J_n} (x_{ij} - 1/n)^2 - 1 + 1/n = 0, \quad j \in J_n;$$

$$f_{2x/n+j}(x_i) = \sum_{j \in J_n} (x_{ij} - 1/n)^2 - 1 + 1/n = 0, \quad i \in J_n.$$  

Similarly to Representation 4.1, the order of Representation 4.2 is $2n+2$. Its components are quadratic and biquadratic only. Thus, it is a mixed, polynomial, biquadratic, convex, and smooth $\Pi_n$-representation. Its redundancy needs an additional investigation.

c. Representation 4.3 is also based on (5.5) and applying Representation 3 to the rows and columns of a matrix $X$ associated with a vector $\mathbf{x}$:

$$f_j(x_j') = \sum_{i \in J_n} x_{ij}^2 - 1 = 0, \quad f_{nx/j}(x'_j) = \sum_{i \in J_n} x_{ij}^3 - 1 = 0, \quad j \in J_n;$$

$$f_{2x/n+i}(x_i) = \sum_{j \in J_n} x_{ij}^2 - 1 = 0, \quad f_{3x/n+i}(x_i) = \sum_{i \in J_n} x_{ij}^3 - 1 = 0, \quad i \in J_n.$$  

Representation 4.1 is also mixed, of order $4n$, cubic, convex in $R_n^{2n}$, and smooth.

Comparing the strict $\Pi_n$-representations, we observe that the first two are of less order and higher degree, while the last has more order and fewer degree. A comparison of all these strict representations with (PSR) given in Sect. VI.1 demonstrates that the first ones have much less components and higher degree of polynomials. The order of the strict representations decreases as the highest degree of polynomials in these representations increases. At the same time, the last ones are...
Continuous Representation Techniques in Combinatorial Optimization

polynomial representations of the minimal degree. It is due to linear representations of discrete sets if they are one point ones do not exist. The next simplest case is quadratic.

VII. Conclusion

In the paper, approaches to an analytic description of the Euclidean combinatorial sets were studied. For the sets, such as $B_n$, $B_n^r(1)$, $B_n'$, $CE_n(1)$, and the Euclidean combinatorial set $\Pi_n$ of permutation matrices, a number of strict and non-strict, tangent and mixed functional representations were derived. The results allow to develop new continuous algorithms of optimization over these sets that are based on new continuous reformulations and relaxations of the combinatorial problems, which apply the derived continuous representations of the sets.

References