MATHEMATICAL ENGINEERING TECHNICAL REPORTS

Discrete Convexity and Polynomial Solvability in Minimum 0-Extension Problems

Hiroshi HIRAI

METR 2012–18 October 2012

DEPARTMENT OF MATHEMATICAL INFORMATICS
GRADUATE SCHOOL OF INFORMATION SCIENCE AND TECHNOLOGY
THE UNIVERSITY OF TOKYO
BUNKYO-KU, TOKYO 113-8656, JAPAN

WWW page: http://www.keisu.t.u-tokyo.ac.jp/research/techrep/index.html

Discrete Convexity and Polynomial Solvability in Minimum 0-Extension Problems

Hiroshi HIRAI
Department of Mathematical Informatics,
Graduate School of Information Science and Technology,
University of Tokyo, Tokyo, 113-8656, Japan.
hirai@mist.i.u-tokyo.ac.jp

October, 2012

Abstract

The minimum 0-extension problem **0-Ext**$[\Gamma]$ on a graph Γ is: given a set V including the vertex set V_Γ of Γ and a nonnegative cost function c defined on the set of all pairs of V, find a 0-extension d of the path metric d_Γ of Γ with $\sum_{xy} c(xy)d(x,y)$ minimum, where a 0-extension is a metric d on V such that the restriction of d to V_Γ coincides with d_Γ and for all $x \in V$ there exists a vertex s in Γ with $d(x,s) = 0$. **0-Ext**$[\Gamma]$ includes a number of basic combinatorial optimization problems, such as minimum (s,t)-cut problem and multiway cut problem.

Karzanov proved the polynomial solvability for a certain large class of modular graphs, and raised the question: What are the graphs Γ for which **0-Ext**$[\Gamma]$ can be solved in polynomial time? He also proved that **0-Ext**$[\Gamma]$ is NP-hard if Γ is not modular or not orientable (in a certain sense).

In this paper, we prove the converse: if Γ is orientable and modular, then **0-Ext**$[\Gamma]$ can be solved in polynomial time. This completes the classification of the tractable graphs for the 0-extension problem. To prove our main result, we develop a theory of discrete convex functions on orientable modular graphs, analogous to discrete convex analysis by Murota, and utilize a recent result of Thapper and Živný on Valued-CSP.

1 Introduction

By a *(semi)metric* d on a finite set V we mean a nonnegative symmetric function on $V \times V$ satisfying $d(x,x) = 0$ for all $x \in V$ and the triangle inequalities $d(x,y) + d(y,z) \geq d(x,z)$ for all $x,y,z \in V$. An *extension* of a metric space (S,μ) is a metric space (V,d) with $V \supseteq S$ and $d(s,t) = \mu(s,t)$ for $s,t \in S$. An extension (V,d) of (S,μ) is called a 0-*extension* if for all $x \in V$ there exists $s \in S$ with $d(s,x) = 0$.

Let Γ be a simple connected undirected graph with vertex set V_Γ. Let d_Γ denote the shortest path metric on V_Γ with respect to the uniform unit edge-length of Γ. The *minimum* 0-*extension problem* **0-Ext**$[\Gamma]$ on Γ is formulated as:

0-Ext$[\Gamma]$: Given $V \supseteq V_\Gamma$ and $c : \binom{V}{2} \to \mathbf{Q}_+$,

$$\text{minimize} \sum_{xy \in \binom{V}{2}} c(xy)d(x,y) \text{ over all 0-extensions } (V,d) \text{ of } (V_\Gamma, d_\Gamma).$$

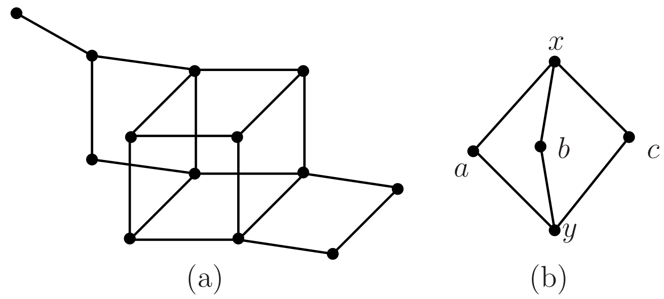

Figure 1: (a) a median graph, (b) a, b, c have two medians x, y, and (c) u, v, w have no median

Here $\binom{V}{2}$ denotes the set of all pairs of V. The minimum 0-extension problem is formulated by Karzanov [26], and is equivalent to the following classical facility location problem, known as *multifacility location problem*; see [45].

Minimize $\sum_{xy \in \binom{V}{2}} c(xy) d_\Gamma(\rho(x), \rho(y))$ over all maps $\rho : V \to V_\Gamma$ being the identity on V_Γ.

This type of problems arises in many practical situations such as computer vision, clustering, and learning theory; see [30]. Also **0-Ext**$[\Gamma]$ includes a number of basic combinatorial optimization problems. For example, take as Γ the graph K_2 consisting of a single edge st. Then **0-Ext**$[K_2]$ is the minimum (s,t)-cut problem. More generally, **0-Ext**$[K_m]$ is the multiway cut problem on m terminals. Therefore **0-ext**$[K_m]$ is solvable in polynomial time if $m = 2$ and is NP-hard if $m > 2$ [14].

This paper addresses the following problem considered by Karzanov [26, 28, 29].

What are the graphs Γ for which **0-Ext**$[\Gamma]$ *is solvable in polynomial time?*

A classical result in location theory in the 1970's is

Theorem 1.1 ([41]; also see [31])**.** *If Γ is a tree, then* **0-Ext**$[\Gamma]$ *is solvable in polynomial time.*

The tractability of graphs Γ is preserved under the Cartesian product. Therefore, cubes, grid graphs, and the Cartesian product of trees are graphs for which **0-Ext** is tractable. Chepoi [12] extended this classical result to median graphs as follows. A *median* of a triple p_1, p_2, p_3 of vertices is a vertex m satisfying $d_\Gamma(p_i, p_j) = d_\Gamma(p_i, m) + d_\Gamma(m, p_j)$ for $1 \leq i < j \leq 3$. A *median graph* is a graph in which every triple of vertices has a *unique* median. Trees and their products are median graphs. See Figure 1 for illustration of the median concept.

Theorem 1.2 ([12])**.** *If Γ is a median graph, then* **0-Ext**$[\Gamma]$ *is solvable in polynomial time.*

Karzanov [26] introduced the following LP-relaxation of **0-Ext**$[\Gamma]$.

Ext$[\Gamma]$**:** Given $V \supseteq V_\Gamma$ and $c : \binom{V}{2} \to \mathbf{Q}_+$,

$$\text{minimize} \sum_{xy \in \binom{V}{2}} c(xy) d(x, y) \text{ over all extensions } (V, d) \text{ of } (V_\Gamma, d_\Gamma).$$

This relaxation $\mathbf{Ext}[\Gamma]$ is a linear program with size polynomial in the input size. Therefore, if, for every input (V, c), $\mathbf{Ext}[\Gamma]$ has an optimal solution that is a 0-extension, then $\mathbf{0\text{-}Ext}[\Gamma]$ is solvable in polynomial time. In this case we say that $\mathbf{Ext}[\Gamma]$ *is exact.* In the same paper, Karzanov gave a combinatorial characterization of graphs Γ for which $\mathbf{Ext}[\Gamma]$ is exact. A graph Γ is called a *frame* if

(1.1) (1) Γ is bipartite,

(2) Γ has no isometric cycle of length greater than 4,

(3) Γ has an orientation o with the property that for every 4-cycle $uv, vv', v'u', u'u$, one has $u \swarrow_o v$ if and only if $u' \swarrow_o v'$.

Here an *isometric cycle* in Γ means a cycle C such that every pair of vertices in C has a shortest path for Γ in this cycle C, and $p \swarrow_o q$ means that edge pq is oriented from q to p by o.

Theorem 1.3 ([26])**.** $\mathbf{Ext}[\Gamma]$ *is exact if and only if* Γ *is a frame.*

Theorem 1.4 ([26])**.** *If* Γ *is a frame, then* $\mathbf{0\text{-}Ext}[\Gamma]$ *is solvable in polynomial time.*

It is noted that the class of frames is *not* closed under the Cartesian product, whereas the tractability of graphs is preserved under the Cartesian product. Also it should be noted that $\mathbf{Ext}[\Gamma]$ is the LP-dual to the d_Γ-weighted maximum *multiflow* problem, and $\mathbf{0\text{-}Ext}[\Gamma]$ describes a combinatorial dual problem [26, 27]; see also [18, 19, 20, 21] for further ramifications of this duality.

Karzanov [26] also proved the following hardness result. For an undirected graph Γ, an orientation with the property (1.1) (3) is said to be *admissible.* Γ is said to be *orientable* if it has an admissible orientation. Γ is said to be *modular* if every triple of vertices has a (not necessarily unique) median.

Theorem 1.5 ([26])**.** *If* Γ *is not orientable or not modular, then* $\mathbf{0\text{-}Ext}[\Gamma]$ *is NP-hard.*

In fact, a frame is precisely an orientable modular graph with the hereditary property that every isometric subgraph is modular; see [2]. A median graph is an orientable modular graph but the converse is not true. Moreover, a median graph is not necessarily a frame, and a frame is not necessarily a median graph. In [28], Karzanov proved a tractability theorem extending Theorem 1.2. He conjectured that $\mathbf{0\text{-}Ext}[\Gamma]$ is tractable for a certain *proper* subclass of orientable modular graphs including frames and median graphs. He also conjectured that $\mathbf{0\text{-}Ext}[\Gamma]$ is NP-hard for any graph Γ not in this class.

The main result of this paper is the tractability theorem for *all* orientable modular graphs, and thus disproves his second conjecture.

Theorem 1.6. *If* Γ *is orientable modular, then* $\mathbf{0\text{-}Ext}[\Gamma]$ *is solvable in polynomial time.*

Combining this result with Theorem 1.5, we obtain a complete classification of the graphs Γ for which $\mathbf{0\text{-}Ext}[\Gamma]$ is solvable in polynomial time.

Overview. In proving Theorem 1.6, we employ an axiomatic approach to optimizations on orientable modular graphs. This approach is inspired by the theory of discrete convex analysis developed by Murota and his coworkers; see [37, 39] and also [16, Chapter VII]. Discrete convex analysis is a theory of convex functions on integer lattice points $\mathbf{Z}^n$, aiming at providing a unified framework for polynomially solvable combinatorial optimization problems including network flows, matroids, and submodular functions.

The theory that we are going to develop here is, in a sense, *a theory of discrete convex functions on orientable modular graphs*, aiming at providing a unified framework for polynomially solvable 0-extension problems and related multiflow problems. We believe that our theory establishes a new link between previously unrelated fields, broadens the scope of discrete convex analysis, and opens a new perspective and new research directions.

Let us start with a simple observation to illustrate our basic idea. Consider a path P_m of length m, and consider $\mathbf{0\text{-}Ext}[P_m]$, where P_m is trivially an orientable modular graph. Then $\mathbf{0\text{-}Ext}[P_m]$ for input V, c can be regarded as an optimization problem on the integer lattice $\mathbf{Z}^n$ as follows. Let $V_{P_m} := \{1, 2, 3, \ldots, m\}$ and $V := \{1, 2, 3, \ldots, n\}$ for $n \geq m$. Then any map $\rho : V \to V_{P_m}$ is identified with point $(x_1, x_2, \ldots, x_n)$ in an integer box $[0, m]^n \cap \mathbf{Z}^n$ (by the correspondence $\rho(i) \leftrightarrow x_i$). In particular, $d_{P_m}(x_i, x_j) = |x_i - x_j|$. Therefore $\mathbf{0\text{-}Ext}[P_m]$ is equivalent to the minimization of the function

$$\sum_{ij} c_{ij}|x_i - x_j| \tag{1.2}$$

over all $(x_1, x_2, \ldots, x_n) \in [0, m]^n \cap \mathbf{Z}^n$ with $x_i = i$ for $i = 1, 2, \ldots, m$. This function is a simple instance of *L-convex functions*, one of the fundamental classes of discrete convex functions. We do not give a formal definition of L-convex functions here. The only important facts for us are the following properties of L-convex functions in optimization:

(a) A local optimality implies the global optimality.

(b) The local optimality can be checked by *submodular function minimization*.

(c) An efficient descent algorithm can be designed based on successive submodular function minimizations.

As is well-known, submodular functions can be minimized in polynomial time [17, 24, 44]. Actually the function (1.2) can be minimized by successive minimum-cut computations [31, 41], a special case of successive submodular function minimizations.

Motivated by this observation, we regard $\mathbf{0\text{-}Ext}[\Gamma]$ as a minimization of a function defined on the vertex set of a product of Γ, which is also orientable modular. We will introduce a class of functions, called *L-convex functions*, on an orientable modular graph. We show that our L-convex function satisfies analogues of (a), (b) and (c) above, and also that a multifacility location function, the objective function of $\mathbf{0\text{-}Ext}[\Gamma]$, is an L-convex function, in our sense, on the product of Γ. Theorem 1.6 is a consequence of these properties.

Let us briefly mention how to define L-convex functions, which constitutes the main body of this paper. Our definition is based on the *Lovász extension* [36], a well-known concept in submodular function theory [16], and the polyhedral complex constructions, due to Karzanov [26] and Chepoi [13], from a class of modular graphs.

Let Γ be an orientable modular graph with admissible orientation o. We call a pair (Γ, o) a *modular complex* since it turns out that (Γ, o) can be regarded as a system of modular (semi)lattices that gives rise to a simplicial complex as follows. Consider a cube subgraph B of Γ. The digraph $\vec{B}$ oriented by o coincides with the Hasse diagram of a Boolean lattice. Consider the simplicial complex $\Delta(\Gamma, o)$ whose simplices are sets of vertices forming a chain of the Boolean lattice corresponding to some cube subgraph of Γ; see Figure 2. Each (abstract) simplex is naturally regarded as a geometric simplex in the Euclidean space. $\Delta(\Gamma, o)$ is naturally regarded as a metrized simplicial complex, which we call the *geometric modular complex* associated with (Γ, o). Then any function

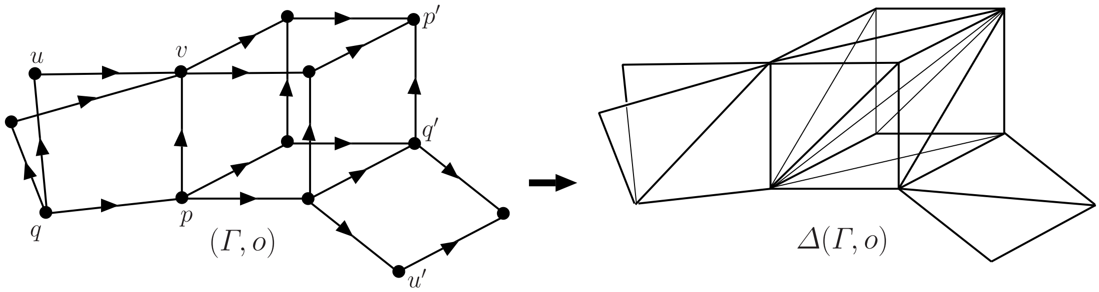

Figure 2: A construction of a geometric modular complex

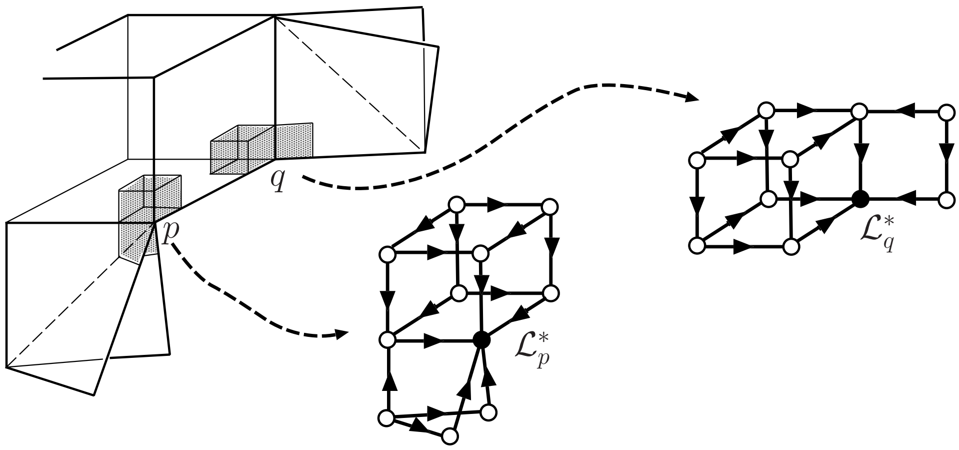

Figure 3: Neighborhood semilattices

$g : V_\Gamma \to \mathbf{R}$ is extended to $\bar{g} : \Delta(\Gamma, o) \to \mathbf{R}$ by interpolating g on each simplex linearly; this is an analogue of the Lovász extension. The geometric simplicial complex $\Delta(\Gamma, o)$ enables us to consider the *neighborhood* $\mathcal{L}_p^*$ around each vertex $p \in V_\Gamma$, as well as the local behavior of $\bar{g}$ in $\mathcal{L}_p^*$. As in Figure 3, neighborhood $\mathcal{L}_p^*$ can be described as a partially ordered set with p the unique minimal element. Then, by restricting $\bar{g}$ to $\mathcal{L}_p^*$, we obtain a function on $\mathcal{L}_p^*$ associated with each vertex p, called the *derivative* of g at p. In fact, the poset $\mathcal{L}_p^*$ is a *modular semilattice*, a semilattice analogue of a modular lattice introduced by Bandelt, van de Vel, and Verheul [5]. We first give a definition of *submodular functions on modular semilattices*, and next define an *L-convex function* as a function on V_Γ such that the derivative on each vertex is submodular.

Then our problem reduces to the minimization of a submodular function f on the product of modular semilattices $\mathcal{L}_1, \mathcal{L}_2, \ldots, \mathcal{L}_n$, where the input of the problem is $\mathcal{L}_1, \mathcal{L}_2, \ldots, \mathcal{L}_n$, and an evaluating oracle of f. We do not know whether this problem in general is tractable in the oracle model, but the submodular functions arising from **0-Ext**$[\Gamma]$ take a special form; they are *the sum of submodular functions with arity* 2, where the arity of a function f is the number of variables of f. See (1.2). This type of optimization problem with bounded arity is well-studied in the literature of *Valued-CSP* [7, 34, 43]. Recently, Thapper and Živný [46] gave a surprising condition for the *basic LP-relaxation* of Valued-CSP to exactly solve the original Valued-CSP instance. They showed that if the class of Valued-CSP (the class of input objective functions) admits a certain nice *fractional polymorphism*, then the basic LP-relaxation is exact.

We prove that the class of submodular functions on modular semilattice has such a nice fractional polymorphism. As a consequence of these two facts, the sum of submodular functions with bounded arity can be minimized in polynomial time. Therefore we can solve **0-Ext**$[\Gamma]$ in polynomial time.

We believe that our classes of functions deserve to be called submodular and L-convex. Indeed, they include not only (ordinary) submodular/L-convex functions but also other submodular/L-convex-type functions. Examples are *bisubmodular functions* [10, 40, 42] (see [16, Section 3.5]), *multimatroid rank functions* by Bouchet [8], *submodular functions on trees* by Kolmogorov [32], and *k-submodular functions* by Huber and Kolmogorov [23]. Moreover, combinatorial dual problems arising from a large class of (well-behaved) multicommodity flow problems, discussed in [18, 19, 20, 21, 25, 26, 27], fall into submodular/L-convex function minimization in our sense. This can be understood as a multiflow analogue of a fundamental fact in network flow theory: the minimum cut problem, the dual of maxflow problem, is a submodular function minimization. The detailed discussion on these topics will be given in a separate paper [22].

Organization. In Section 2, we describe basic facts on modular graphs and modular lattices. In Section 3, we prove some structural properties of orientable modular graphs. We show that an orientable modular graph is obtained by gluing the covering graphs of complemented modular lattices (Theorem 3.1). In place of the geometric modular complex, which is defined as the union of the order complexes of those lattices, we use a graph-theoretic operation on orientable modular graphs, called the 2*-subdivision operation*. This operation is adequate for our purpose, in that it keeps the orientability and the modularity, and enables us to define neighborhood semilattices $\mathcal{L}^*_p$. In Section 4, we define submodular functions on modular semilattices and L-convex functions on a modular complex, according to the idea mentioned above. We prove that our L-convex functions indeed have properties analogous to (a), (b) and (c) above, and that the sum of submodular functions with bounded arity can be minimized in polynomial time. In Section 5, we reformulate **0-Ext**$[\Gamma]$ as an optimization problem on a modular complex. We show that a multifacility location function, the objective function of **0-Ext**$[\Gamma]$, is indeed an L-convex function, and we prove Theorem 1.6. Our framework is applicable to a certain weighted version of **0-Ext**$[\Gamma]$. As a corollary, we give a generalization of Theorem 1.6 to general metrics, which completes classification of metrics μ for which the 0-extension problem on μ is polynomial time solvable (Theorem 5.12).

Notation. Let $\mathbf{Z}, \mathbf{Q}$, and $\mathbf{R}$ denote the sets of integers, rationals, and reals, respectively. Let $\mathbf{Z}_+, \mathbf{Q}_+$, and $\mathbf{R}_+$ denote the sets of nonnegative integers, nonnegative rationals, and nonnegative reals, respectively. For a graph Γ, the vertex set and the edge set are denoted by V_Γ and E_Γ, respectively. For a vertex subset X, $\Gamma[X]$ denotes the subgraph of Γ induced by X. For a nonnegative edge-length $h : E_\Gamma \to \mathbf{R}_+$, $d_{\Gamma,h}$ denotes the shortest path metric on V_Γ with respect to the edge-length h. When $h(e) = 1$ for every edge e, $d_{\Gamma,h}$ is denoted by d_Γ. A path is represented by a chain $(p_1, p_2, \ldots, p_n)$ of vertices with $p_i p_{i+1} \in E_\Gamma$.

The *Cartesian product* $\Gamma \times \Gamma'$ of graphs Γ and Γ' is the graph with vertex set $V_\Gamma \times V_{\Gamma'}$ and edge set given as: (p, p') and (q, q') are connected by an edge if and only if $p = q$ and $p'q' \in E_{\Gamma'}$ or $p' = q'$ and $pq \in E_\Gamma$.

2 Preliminaries on modular graphs and modular lattices

In this section, we summarize basic facts on modular graphs and modular lattices with emphasis on their metric aspects. Our references are [3, 5, 13, 47] for modular graphs, and the first edition of [6] for modular lattices.

2.1 Modular metric spaces and modular graphs

For a metric space (X, d), the *interval* $I(x, y)$ of $x, y \in X$ is defined as

$$I(x, y) := \{z \in X \mid d(x, z) + d(z, y) = d(x, y)\}.$$

For two subsets A, B, $d(A, B)$ denotes the minimum distance between A and B, i.e.,

$$d(A, B) = \inf_{x \in A, y \in B} d(x, y).$$

For $x_1, x_2, x_3 \in X$, an element m in $I(x_1, x_2) \cap I(x_2, x_3) \cap I(x_3, x_1)$ is called *a median* of x_1, x_2, and x_3. We note that the distance between x_i and m is given by

$$d(m, x_i) = \frac{d(x_i, x_j) + d(x_i, x_k) - d(x_j, x_k)}{2} \quad (i, j, k \text{ distinct}). \tag{2.1}$$

A metric space (X, d) is said to be *modular* if every triple of elements in X has a median. In particular, a graph Γ is modular if the shortest path metric space (V_Γ, d_Γ) is modular. We will often use the following characterization of modular graphs.

Lemma 2.1 ([5, Proposition 1.7]; see [47, Proposition 6.2.6, Chapter I])**.** *A connected graph Γ is modular if and only if*

1. *Γ is bipartite, and*
2. *for vertices p, q and neighbors p_1, p_2 of p with $d_\Gamma(p, q) = 1 + d_\Gamma(p_1, q) = 1 + d_\Gamma(p_2, q)$, there exists a common neighbor p^* of p_1, p_2 with $d_\Gamma(p, q) = 2 + d_\Gamma(p^*, q)$.*

The condition (2) is called the *quadrangle condition* [3, 13] (or the *semimodularity condition* in [5, 47]).

Lemma 2.2. *For a modular graph, every admissible orientation is acyclic*

Proof. Suppose indirectly that the statement is false. Take a vertex p belonging to a directed cycle, and take a directed cycle C containing p with $\sum_{u \in V_C} d_\Gamma(p, u)$ minimum. The length k of C is at least four (by simpleness and bipartiteness). By the definition of admissible orientation, $k = 4$ is impossible. Hence $k > 4$. Take a vertex q in C with $d_\Gamma(p, q)$ maximum. Take two neighbors q', q'' of q in C. Then $d_\Gamma(q, p) = d_\Gamma(q', p) + 1 = d_\Gamma(q'', p) + 1$ (by the maximality of q and the bipartiteness of Γ). By the quadrangle condition, there is a common neighbor q^* of q', q'' with $d_\Gamma(p, q^*) = d_\Gamma(p, q) - 2$. Here the cycle C' obtained from C by replacing q by q^* is a directed cycle, since the orientation is admissible. Then we have $\sum_{u' \in V_{C'}} d_\Gamma(p, u') < \sum_{u \in V_C} d_\Gamma(p, u)$. This is a contradiction to the minimality of C. □

2.2 Convex sets and gated sets

Let (X, d) be a metric space. A subset $Y \subseteq X$ is called *convex* if $I(p, q) \subseteq Y$ for every $p, q \in Y$. A subset $Y \subseteq X$ is called *gated* if for every $p \in X$ there is $p^* \in Y$, called a *gate* of p at Y, such that $d(p, q) = d(p, p^*) + d(p^*, q)$ holds for every $q \in Y$. One can easily see that gate p^* is uniquely determined for each p [15, p. 112]. Therefore we obtain a map $\mathrm{Pr}_Y : X \to Y$ by defining $\mathrm{Pr}_Y(p)$ to be the gate of p at Y.

Theorem 2.3 ([15]). *Let A and A' be gated subsets of (X, d) and let $B := \mathrm{Pr}_A(A')$ and $B' := \mathrm{Pr}_{A'}(A)$.*

(1) Pr_A *and* $\mathrm{Pr}_{A'}$ *induce isometries, inverse to each other, between B' and B.*

(2) *The following are equivalent for $p \in A$ and $p' \in A'$:*

  (i) $d(p, p') = d(A, A')$.

  (ii) $p = \mathrm{Pr}_A(p')$ *and* $p' = \mathrm{Pr}_A(p)$.

(3) *B and B' are gated, and* $\mathrm{Pr}_B = \mathrm{Pr}_A \circ \mathrm{Pr}_{A'}$ *and* $\mathrm{Pr}_{B'} = \mathrm{Pr}_{A'} \circ \mathrm{Pr}_A$.

As remarked in [15], every gated set is convex (see the proof of Lemma 2.4 below). The converse is not true in general, but is true for modular graphs. The following useful characterization of convex (gated) sets in a modular graph is due to Chepoi [11]. Here, for a graph Γ, a subset Y of vertices is said to be convex (resp. gated) if Y is convex (resp. gated) in (V_Γ, d_Γ).

Lemma 2.4 ([11]). *Let Γ be a modular graph. For $Y \subseteq V_\Gamma$, the following conditions are equivalent:*

(1) *Y is convex.*

(2) *Y is gated.*

(3) *$\Gamma[Y]$ is connected and $I(p, q) \subseteq Y$ holds for every $p, q \in Y$ with $d_\Gamma(p, q) = 2$.*

We give a proof for the convenience of readers as the original paper is in Russian.

Proof. d_Γ is denoted by d. (1) $\Rightarrow$ (3) is obvious.

We show (3) $\Rightarrow$ (1). Take $p, q \in Y$, and take $a \in I(p, q)$. We are going to show $a \in Y$. Since $\Gamma[Y]$ is connected, we can take a path $P = (p = p_0, p_1, \ldots, p_k = q)$ with $p_i \in Y$. Take such a path P with $\kappa_P := \sum_{i=0}^{k} d(a, p_i)$ minimum. If $d(a, p_{i-1}) + 1 = d(a, p_i) = d(a, p_{i+1}) + 1$ for some i, then, by the quadrangle condition in Lemma 2.1, there is a common neighbor p^* of p_{i-1}, p_{i+1} with $d(a, p^*) = d(a, p_i) - 2$. Since $I(p_{i-1}, p_{i+1}) \subseteq Y$ by (3), p^* belongs to Y. Then we can replace p_i by p^* in P to get another path P' connecting p, q with $\kappa_{P'} = \kappa_P - 2$; a contradiction to the minimality. Therefore there is no index j with $d(a, p_{j-1}) < d(a, p_j) > d(a, p_{j+1})$. Thus there is a unique index i with $d(a, p_i)$ minimum. Then we have $d(p, p_i) + d(p_i, a) = d(p, a)$ and $d(q, p_i) + d(p_i, a) = d(q, a)$. Adding them, we get

$$d(p, p_i) + d(p_i, q) + 2d(p_i, a) = d(p, a) + d(a, q).$$

Since $a \in I(p, q)$, we have $d(p, a) + d(a, q) = d(p, q)$. Obviously $d(p, p_i) + d(p_i, q) \geq d(p, q)$ holds. Hence $d(p_i, a) = 0$, i.e., $a = p_i \in Y$.

We show (2) $\Rightarrow$ (1). As already mentioned, any gated set is convex. Indeed, suppose that Y is gated. Take $p, q \in Y$, and take $a \in I(p, q)$. Consider the gate a^* of a in Y.

Then $d(p, a) = d(p, a^*) + d(a^*, a)$ and $d(q, a) = d(q, a^*) + d(a^*, a)$. Since $a \in I(p, q)$, we have $d(p, q) = d(p, a) + d(a, q) = d(p, a^*) + d(a^*, q) + 2d(a^*, a) \geq d(p, q) + 2d(a^*, a)$, implying $d(a^*, a) = 0$ and $a = a^* \in Y$. Thus we get (2) $\Rightarrow$ (1).

Finally we show (1) $\Rightarrow$ (2). Suppose that Y is convex. Let p be an arbitrary vertex. Let p^* be a point in Y satisfying $d(p, Y) = d(p, p^*)$. We show that p^* is a gate of p at Y. Take arbitrary $q \in Y$. Consider a median m of p, q, p^*. Then $d(p, m) = \{d(p, p^*) + d(p, q) - d(p^*, q)\}/2 = d(p, p^*) - \{d(p, p^*) + d(p^*, q) - d(p, q)\}/2 \leq d(p, p^*)$. By convexity, m belongs to Y. By definition of p^*, $d(p, p^*) = d(p, m)$ must hold. Thus $d(p, q) - d(p, p^*) - d(p^*, q) = 0$ holds for every $q \in Y$. This means that p^* is the gate of p, and therefore Y is gated. □

2.3 Modular lattices and modular semilattices

Let $\mathcal{L}$ be a partially ordered set with partial order $\preceq$. For $a, b \in \mathcal{L}$, the least common upper bound, if it exists, is denoted by $a \vee b$, and the greatest common lower bound, if it exists, is denoted by $a \wedge b$. $\mathcal{L}$ is said to be a *lattice* if both $a \vee b$ and $a \wedge b$ exist for every $a, b \in \mathcal{L}$, and said to be a *(meet-)semilattice* if $a \wedge b$ exists for every $a, b \in \mathcal{L}$. In a semilattice, if a and b have a common upper bound, then $a \vee b$ exists. Such (a, b) is said to be *bounded*. By the expression $a \vee b \in \mathcal{L}$ we mean that $a \vee b$ exists. A pair (a, b) is said to be *comparable* if $a \preceq b$ or $b \preceq a$, and *incomparable* otherwise. We say "*b covers a*" if $a \prec b$ and there is no $c \in \mathcal{L}$ with $a \prec c \prec b$, where $a \prec b$ means $a \preceq b$ and $a \neq b$. The universal upper bound and the universal lower bound, if they exist, are denoted by $\mathbf{1}$ and $\mathbf{0}$, respectively. For $a \preceq b$, the interval $\{c \in \mathcal{L} \mid a \preceq c \preceq b\}$ is denoted by $[a, b]$. A *chain* from a to b is a sequence $(a = u_0, u_1, u_2, \ldots, u_k = b)$ with $u_{i-1} \prec u_i$ for $i = 1, 2, \ldots, k$; the number k is the length of the chain. The length $r[a, b]$ of the interval $[a, b]$ is defined as the maximum length of a chain from a to b. The rank $r(a)$ of an element a is defined by $r(a) = r[\mathbf{0}, a]$.

A lattice $\mathcal{L}$ is called *modular* if $a \vee (b \wedge c) = (a \vee b) \wedge c$ for every $a, b, c \in \mathcal{L}$ with $a \preceq c$. Modular lattices are also characterized by the modular equality of the rank function.

Lemma 2.5 (see [6, Chapter III, Corollary 1])**.** *A lattice $\mathcal{L}$ is modular if and only if*

$$r(a) + r(b) = r(a \vee b) + r(a \wedge b) \quad (a, b \in \mathcal{L}).$$

The underlying undirected graph of the Hasse diagram of $\mathcal{L}$ is called the *covering graph* of $\mathcal{L}$. Modular lattice $\mathcal{L}$ is regarded as a metric space by the shortest path metric d_Γ of its covering graph Γ. It is a folklore that a lattice is modular if and only if its covering graph is modular. The modularity concept has been extended for semilattices by Bandelt, van de Vel, and Verheul [4]. A semilattice $\mathcal{L}$ is said to be *modular* if $[\mathbf{0}, p]$ is a modular lattice for every $p \in \mathcal{L}$, and $a \vee b \vee c \in \mathcal{L}$ provided $a \vee b, b \vee c, c \vee a \in \mathcal{L}$.

Theorem 2.6 ([5, Theorem 5.4])**.** *A semilattice is modular if and only if its covering graph is modular.*

Obviously the Hasse diagram of $\mathcal{L}$ is admissibly oriented.

Corollary 2.7. *The covering graph of a modular semilattice is orientable modular.*

The following metric properties of a modular semilattice play important roles in Section 4.

Lemma 2.8 ([5])**.** *Let $\mathcal{L}$ be a modular semilattice with covering graph Γ. For $p, q \in \mathcal{L}$, we have the following.*

(1) $d_\Gamma(p,q) = r[p \wedge q, p] + r[p \wedge q, q]$.

(2) $I(p,q) = \{c \in \mathcal{L} \mid c = a \vee b, a \in [p \wedge q, p], b \in [p \wedge q, q]\}$.

(3) *If* $c = a \vee b$ *for* $a \in [p \wedge q, p], b \in [p \wedge q, q]$, *then* $a = p \wedge c$ *and* $b = q \wedge c$.

Proof. The proof is given for completeness. (3). Necessarily $a \preceq p \wedge c$ and $b \preceq q \wedge c$, implying $a \vee b \preceq (p \wedge c) \vee (q \wedge c) \preceq c = a \vee b$. Hence $(p \wedge c) \vee (q \wedge c) = c$. Also $(p \wedge c) \wedge (q \wedge c) = p \wedge q$ (since $p \wedge q \preceq c$). By the modularity equality, we have $r(a) + r(b) = r(c) + r(p \wedge q) = r(p \wedge c) + r(q \wedge c)$, which implies $r[a, p \wedge c] = r[b, q \wedge c] = 0$. Thus $p \wedge c = a$ and $q \wedge c = a$ must hold.

(1). We use the induction on $d_\Gamma(p,q)$. Take a neighbor q' of q in $I(p,q)$. Then $d_\Gamma(p,q') = r[p \wedge q', p] + r[p \wedge q', q']$ by induction, and either (i) q covers q' or (ii) q' covers q. In the first case (i), we must have $p \wedge q \preceq q'$ as follows. Suppose that $p \wedge q \preceq q'$ is not in the case. Then $p \wedge q$ covers $p \wedge q'$, and the modularity equality yields $r[p \wedge q, q] = r[p \wedge q', q']$, which means that there is a (p,q)-path passing through $p \wedge q$ with the length $d_\Gamma(p,q')$. A contradiction to $d_\Gamma(p,q) = d_\Gamma(p,q') + 1$. It follows from $p \wedge q \preceq q'$ that $p \wedge q = p \wedge q'$, and the claim follows. In the second case (ii) where q' covers q, $p \wedge q'$ covers $p \wedge q$; since otherwise $p \wedge q' = p \wedge q$ which leads to a contradiction $d_\Gamma(p,q') > d_\Gamma(p,q)$. By the modularity equality $r[p \wedge q', q'] = r[p \wedge q, q]$ and the claim follows.

(2). By (1), $p \wedge q \in I(p,q)$. By the modularity equality we have LHS $\supseteq$ RHS. We show the reverse inclusion by induction on $d_\Gamma(p,q)$. Take $c \in I(p,q)$ with $p \neq c \neq q$, and take a neighbor q' of q in $I(c,q) (\subseteq I(p,q))$. Then $c \in I(p,q')$. If q covers q', then $p \wedge q = p \wedge q'$, and apply induction. Suppose that q' covers q, Then necessarily $p \wedge q'$ covers $p \wedge q$ (as above). By induction, there are $a \in [p \wedge q', p] \subseteq [p \wedge q, p]$ and $b' \in [p \wedge q', q']$ with $c = a \vee b'$. By $b' \in I(p \wedge q', q)$, the induction with (3), and $p \wedge b' = p \wedge q'$, there is $b \in [p \wedge q, q]$ with $b' = b \vee (p \wedge q')$; b' covers b. Clearly (a,b) is bounded; $a \vee b \preceq c$. By modularity, we must have $a \vee b = c$. $\square$

A lattice $\mathcal{L}$ is called *complemented* if for every $p \in \mathcal{L}$ there is $q \in \mathcal{L}$ such that $p \vee q = \mathbf{1}$ and $p \wedge q = \mathbf{0}$, and *relatively complemented* if $[a,b]$ is complemented for every $a, b \in \mathcal{L}$ with $a \preceq b$.

Theorem 2.9 (See [6, Chapter IV, Theorem 4.1])**.** *Let* $\mathcal{L}$ *be a modular lattice. The following conditions are equivalent:*

(1) $\mathcal{L}$ *is complemented.*

(2) $\mathcal{L}$ *is relatively complemented.*

(3) *Every element is the join of atoms.*

(4) $\mathbf{1}$ *is the join of atoms.*

Here an *atom* is an element of rank 1. A modular semilattice is said to be *complemented* if $[\mathbf{0}, a]$ is a complemented modular lattice for every $a \in \mathcal{L}$.

2.4 Orbits in modular graphs and modular lattices

Let Γ be a modular graph. Edges e and e' are said to be *projective* if there is a sequence $(e = e_0, e_1, e_2, \ldots, e_m = e')$ of edges such that e_i and e_{i+1} belong to a common 4-cycle and share no common vertex.

Lemma 2.10. *Let* Γ *be a modular graph. For edges* pq *and* $p'q'$, *suppose that* $d_\Gamma(p,p') = d_\Gamma(q,q')$ *and* $d_\Gamma(p,q') = d_\Gamma(p,p') + 1 = d_\Gamma(p',q)$.

(1) *pq and $p'q'$ are projective.*

(2) *In addition, if Γ has an admissible orientation o, then $p \searrow_o q$ implies $p' \searrow_o q'$.*

Proof. We use the induction on $k := d_\Gamma(p,p') = d_\Gamma(q,q')$. The case of $k = 1$ is obvious. Take a neighbor p^* of p with $d_\Gamma(p^*,p') = d_\Gamma(p,p') - 1 = k - 1$. Then $d_\Gamma(p^*,q') = k$. By the quadrangle condition for p,q,p^*,q', there is a common neighbor q^* of q,p^* with $d_\Gamma(q^*,q') = k-1$. Also $d_\Gamma(q^*,p') = k$. Obviously pq and p^*q^* are projective. and $p \searrow_o q$ implies $p^* \searrow_o q^*$. Apply the induction for p^*q^* and $p'q'$. □

In the case where Γ is the covering graph of a modular lattice $\mathcal{L}$, the projectivity relation is equivalent to the *projectivity relation on prime quotients* of $\mathcal{L}$ in the sense of [6, Chapter III, Definition 3.2]. An *orbit* is an equivalence class of the projectivity relation. The union of several orbits is called an *orbit-union*. For an orbit-union U, Γ/U is the graph obtained by contracting all edges not in U and by identifying multiple edges. The vertex in Γ/U that corresponds to $p \in V_\Gamma$ is denoted as p/U.

Lemma 2.11 ([1], also see [28]). *Let Γ be a modular graph, and U an orbit-union.*

(1) *Γ/U is a modular graph.*

(2) *For every $p,q \in V_\Gamma$, every shortest (p,q)-path P, and every (p,q)-path P', we have $|P \cap U| \leq |P' \cap U|$.*

(3) *For every $p,q \in V_\Gamma$ and every shortest (p,q)-path P, the image P/U of P is a shortest $(p/U,q/U)$-path in Γ/U.*

In particular, for any partition $\mathcal{U}$ of E_Γ into orbit-unions, we have

$$d_\Gamma(p,q) = \sum_{U \in \mathcal{U}} d_{\Gamma/U}(p/U,q/U) = \sum_{Q:\text{orbit}} d_{\Gamma/Q}(p/Q,q/Q) \quad (p,q \in V_\Gamma).$$

For an orbit-union U in a complemented modular semilattice $\mathcal{L}$ with covering graph Γ, we can define $\mathcal{L}|U \subseteq \mathcal{L}$, which will turn out to be a complementary modular subsemilattice (Lemma 2.13 (3)), as follows. The underlying set of $\mathcal{L}|U$ consists of elements q such that all edges of the covering graph of $[\mathbf{0},q]$ are contained in U. For any element $p \in \mathcal{L}$, there exists a maximal chain $P = (\mathbf{0} = p_0, p_1, \ldots, p_k, \ldots, p_m = p)$ such that $p_{i-1}p_i \in U$ if $i < k$ and $p_{i-1}p_i \notin U$ if $i \geq k$. (*Proof sketch*: If $p_{i-1}p_i \notin U$ and $p_ip_{i+1} \in U$, then we can take $p' \neq p_i$ such that $p_{i-1}p' \in U$ and $p'p_{i+1} \notin U$ since $[\mathbf{0},p]$ is relatively-complemented. Replace p_i by p' to get a new maximal chain, and repeat this process.) Furthermore, p_k is independent of the choice of P, and is denoted by $p|U$. (*Proof sketch*: Take another chain P' and $p'_{k'}$ in P'. Then $p_k \vee p'_{k'}$ must be in $\mathcal{L}|U$ and in $[\mathbf{0},p]$. Consider a maximal chain P^* of $[\mathbf{0},p]$ including $p_k \vee p'_{k'}$. Then P and P^* are both shortest paths between $\mathbf{0}$ and p, and this violates Lemma 2.11 (2) if $p_k \neq p'_{k'}$.) Note that $p|U \in \mathcal{L}|U$ for every $p \in \mathcal{L}$ and moreover $\mathcal{L}|U = \{p|U \mid p \in \mathcal{L}\}$.

For a partition $\mathcal{U}$ of E_Γ into orbit-unions, every element p is (uniquely) represented as

$$p = \bigvee_{U \in \mathcal{U}} p|U.$$

In the case where $\mathcal{L}$ is a modular lattice, each $\mathcal{L}|U$ is also a modular lattice, and this decomposition yields a lattice-isomorphism between $\mathcal{L}$ and the direct product of $\mathcal{L}|U$ over $U \in \mathcal{U}$.

Remark 2.12. For an orbit Q, $\mathcal{L}|Q$ is a complemented modular lattice having exactly one orbit. Such a lattice is a *projective space lattice*, which means a complemented modular lattice each rank-2 elements of which is the join of at least three atoms [6]. So the above-described decomposition leads to a well-known fact that every complemented modular lattice is the product of projective space lattices; take the set of all orbits as $\mathcal{U}$.

We note some convexity properties of modular semilattice. Here a subset X of semilattice $\mathcal{L}$ is said to be convex if X is convex in the covering graph Γ of $\mathcal{L}$.

Lemma 2.13. *Let $\mathcal{L}$ be a modular semilattice.*

(1) *Any convex set in $\mathcal{L}$ is a modular subsemilattice of $\mathcal{L}$.*

(2) *Suppose that $\mathcal{L}$ is a lattice. Then a subset C is convex if and only if $C = [a, b]$ for some $a, b \in \mathcal{L}$ with $a \preceq b$.*

(3) *Suppose that $\mathcal{L}$ is complemented. For any orbit-union U, $\mathcal{L}|U$ is convex, and is a complemented modular subsemilattice of $\mathcal{L}$.*

Proof. (1) follows from Lemma 2.8. The if part follows of (2) from Lemma 2.8. To see the only if part of (2), consider $a := \bigwedge_{u \in C} u$ and $b := \bigvee_{u \in C} u$. Then $C \subseteq [a, b]$. From $a, b \in C$, we can see $[a, b] \subseteq C$. (3) follows from Lemma 2.8 and Lemma 2.11 (2). □

Orbit-invariant functions and valuations. Let Γ be a modular graph. A function h on edge set E_Γ is called *orbit-invariant* if $h(e) = h(e')$ provided e and e' belong to the same orbit. For an orbit Q, let h_Q denote the value of h on Q. An orbit-invariant function h is said to be *nonnegative* if $h(e) \geq 0$ for $e \in E_\Gamma$, and is said to be *positive* if $h(e) > 0$ for $e \in E_\Gamma$. If $h(e) = 1$ for all edges e, then h is denoted by 1; in particular $d_\Gamma = d_{\Gamma,1}$. By taking the value of h of the preimage, we can define a function on the edge set of Γ/U for any orbit-union U, which is also orbit-invariant in Γ/U and is denoted by h. By Lemma 2.11 (2), the shortest path structures of (V_Γ, d_Γ) and $(V_\Gamma, d_{\Gamma,h})$ are the same in the following sense.

Lemma 2.14. *If an orbit-invariant function h is nonnegative, then* (1) *implies* (2), *where*

(1) *P is a shortest (p, q)-path with respect to* 1,

(2) *P is a shortest (p, q)-path with respect to h.*

If h is positive, then the converse also holds.

As a consequence of Lemmas 2.11 and 2.14, for any partition $\mathcal{U}$ of E_Γ into orbit-unions, we have

$$d_{\Gamma,h}(p, q) = \sum_{U \in \mathcal{U}} d_{\Gamma/U,h}(p/U, q/U) = \sum_{Q:\text{orbit}} h_Q d_{\Gamma/Q,1}(p/Q, q/Q). \tag{2.2}$$

An orbit-invariant function is a graph-theoretic analogue of a *valuation* in a modular lattice; see [6, Chapter III, 50] (but we follow the terminology in the 3rd edition of this book). Here, a *valuation* v of a modular lattice $\mathcal{L}$ is a function on $\mathcal{L}$ satisfying

$$v(p) + v(q) = v(p \wedge q) + v(p \vee q) \tag{2.3}$$

for each $p, q \in \mathcal{L}$; it is said to be *positive* if $v(p) < v(q)$ for each p, q with $p \prec q$. It is known that every positive valuation v is uniquely represented as

$$v(p) = C + d_{\Gamma,h}(\mathbf{0}, p) \quad (p \in \mathcal{L}) \tag{2.4}$$

for a constant C and a positive orbit-invariant function h on the covering graph Γ of $\mathcal{L}$. Indeed, take $C := v(\mathbf{0})$ and define h by $h(pq) := v(q) - v(p)$ if q covers p. Conversely, for every positive orbit-invariant function h on the covering graph and every constant C, the function v on $\mathcal{L}$ defined by (2.4) is a positive valuation.

We can naturally extend the valuation concept to modular semilattices. A *valuation* v of a modular semilattice $\mathcal{L}$ is a function on $\mathcal{L}$ satisfying (2.3) for every bounded pair $p, q \in \mathcal{L}$; it is said to be *positive* if $v(p) < v(q)$ for each p, q with $p \prec q$. For p, q with $p \preceq q$, let $v[p, q]$ denote $v(q) - v(p)$. Again one can easily see that v has a unique expression (2.4) for a positive orbit-invariant function h on the covering graph Γ of $\mathcal{L}$. Note that the rank function r is nothing but the positive valuation corresponding to $C = 0$ and $h = 1$.

3 Modular complex

In this section, we reveal some structural properties of orientable modular graphs. In particular, we show that an orientable modular graph is naturally regarded as a union of covering graphs of complemented modular lattices (Section 3.1). This enables us to define a simplicial complex as the union of the order complexes of these lattices, and also to define the 2-subdivision operation of orientable modular graphs (Section 3.2). So a triple (Γ, o, h) of an orientable modular graph Γ, an admissible orientation o, and a positive orbit-invariant function h is called a *modular complex.*

3.1 Boolean pairs

Let Γ be an orientable modular graph. Fix an admissible orientation o. Consider a cube subgraph B of Γ, and consider the digraph $\vec{B}$ of B induced by o. One can easily see from the definition of an admissible orientation that $\vec{B}$ is isomorphic to the Hasse diagram of a Boolean lattice. Hence $\vec{B}$ determines the greatest element and the least element of the corresponding Boolean lattice.

A pair (p, q) of vertices is called a *Boolean pair* (with respect to o) if p and q are the least element and the greatest element, respectively, of the Boolean lattice associated with some cube subgraph of Γ. By convention, (p, p) is defined to be a Boolean pair. The set of Boolean pairs is denoted by $\mathcal{B}(\Gamma, o)$. In Figure 2 in the introduction, (p, p), (q, v), (v, p') are examples of Boolean pairs.

Recall that any admissible orientation is acyclic (Lemma 2.2). Let $\preceq_o$ be the transitive closure of relation $\swarrow_o$ on V_Γ. Then V_Γ is regarded as a partially ordered set by this relation. For any Boolean pair (p, q), necessarily $p \preceq_o q$ holds. We define the relation $\sqsubseteq_o$ as: $p \sqsubseteq_o q$ if (p, q) is a Boolean pair. This relation $\sqsubseteq_o$ coarsens $\preceq_o$, and is not transitive in general. For a vertex p, define subset $\mathcal{L}_p(\Gamma, o)$ by

$$\mathcal{L}_p(\Gamma, o) := \{q \in V_\Gamma \mid p \sqsubseteq_o q\} = \bigcup_{p \sqsubseteq_o q} [p, q]. \tag{3.1}$$

$\mathcal{L}_p(\Gamma, o)$ is also denoted simply by $\mathcal{L}_p$. The main result in this section is the following.

Theorem 3.1. *Let Γ be an orientable modular graph with an admissible orientation o.*

(1) *For every Boolean pair (p, q), interval $[p, q]$ is a complemented modular lattice, and is convex in Γ.*

(2) *For every vertex p, $\mathcal{L}_p$ is a complemented modular semilattice, and is convex in Γ.*

(1) says that an orientable modular graph Γ can be regarded as the union of several complemented modular lattices, and (2) says that each vertex is associated with a local semilattice structure.

The rest of this subsection is devoted to proving Theorem 3.1. We denote d_Γ by d, and denote $\swarrow_o$, $\preceq_o$, and $\sqsubseteq_o$ by $\swarrow$, $\preceq$, and $\sqsubseteq$, respectively. First we show a Jordan-Dedekind-type chain condition for Γ (Lemma 3.2). Second we show that if $p \preceq q$ then $[p, q]$ is a modular lattice (Lemma 3.3 and Proposition 3.4), which immediately proves Theorem 3.1 (1). Third we prove a criterion when the transitivity $p \sqsubseteq q \swarrow q' \Rightarrow p \sqsubseteq q'$ holds (Lemma 3.6). Then we prove Theorem 3.1 (2).

The first lemma says that Γ satisfies a Jordan-Dedekind-type condition. A path $(p_0, p_1, p_2, \dots, p_k)$ is said to be *ascending* if $p_i \swarrow p_{i+1}$ for $i = 0, \dots, k-1$.

Lemma 3.2. *For $p, q \in V_\Gamma$ with $p \preceq q$, a (p, q)-path P is shortest if and only if P is an ascending path from p to q. In particular, $I(p, q) = [p, q]$, and any maximal chain in $[p, q]$ has the same length.*

Proof. Suppose $p \preceq q$. Then an ascending path $P = (p = p_0, p_1, p_2, \dots, p_k = q)$ exists. We use the induction on length k; the statement for $k = 1$ is obvious.

(If part). Suppose for contradiction that $d(p, q) < k$. By induction and bipartiteness, we have $d(p, p_k) = d(p, p_{k-1}) - 1$ and $d(p, p_{k-1}) = k - 1$. By the quadrangle condition (Lemma 2.1 (2)) for p_{k-1}, p_k, p_{k-2}, p, there is a common neighbor q^* of p_k, p_{k-2} with $d(p, q^*) = d(p, p_{k-1}) - 2 = k - 3$. Consider the 4-cycle of $p_{k-1}, p_k, q^*, p_{k-2}$. By orientability $p_{k-2} \swarrow q^*$ must hold. Hence we obtain an ascending path $(p = p_0, p_1, \dots, p_{k-2}, q^*)$ of length $k - 1$ with $d(p, q^*) = k - 3$. A contradiction.

(Only if part). Take any shortest path $Q = (p = q_0, q_1, q_2, \dots, q_{k'} = q)$ between p and q. By the if part, necessarily $k' = k$. We may assume that $p_{k-1} \neq q_{k-1}$ (by induction). By the quadrangle condition for q, p_{k-1}, q_{k-1}, p, there is a common neighbor q^* of p_{k-1}, q_{k-1} with $d(q^*, p) = d(q, p) - 2$. This means that (p_{k-1}, q^*) can be extended to a shortest path $P^* = (p = q^*_0, q^*_1, \dots, q^*_{k-3}, q^*, p_{k-1})$ between p and p_{k-1}. By induction, P^* is ascending. In particular, $q^* \swarrow p_{k-1} \swarrow q$. Consider 4-cycle of p_{k-1}, q, q_{k-1}, q^*. By orientability, we have $q^* \swarrow q_{k-1} \swarrow q$. Replacing q^* by q_{k-1} in P^*, we get an ascending path from p to q_{k-1}. By induction $Q \setminus q^*$ is ascending. Hence Q is also ascending. $\square$

Lemma 3.3. *If $p \preceq a$ and $p \preceq b$, then there uniquely exists a median m of p, a, b, which coincides with $a \wedge b$. Similarly, if $a \preceq q$ and $b \preceq q$, then there uniquely exists a median m of q, a, b, which coincides with $a \vee b$.*

Proof. It suffices to prove the former statement. Suppose that a, b, p have two distinct medians c, c'. Take a median m of c, c', p. Let $k := d(c, m) = d(c', m) > 0$. We can take an ascending path $(m = m_0, m_1, \dots, m_k = c)$ from m to c, and also can take a neighbor m' of m with $d(m, c') = 1 + d(m', c')$; necessarily $m \swarrow m'$. By the quadrangle condition for m, m_1, m', a, there is a common neighbor m'_1 of m_1, m' such that $d(a, m'_1) = d(a, m) - 2$. Also by the quadrangle condition for m, m_1, m', b there is a common neighbor m''_1 of m_1, m' such that $d(b, m''_1) = d(b, m) - 2$. By $m_1 \searrow m \swarrow m'$ and the orientability, we have $m_1 \swarrow m'_1 \searrow m'$ and $m_1 \swarrow m''_1 \searrow m'$. Hence $m'_1 = m''_1$ must hold. Similarly, by the quadrangle condition for m_1, m_2, m'_1, a and for m_1, m_2, m'_1, b, we can find a common neighbor m'_2 of m_2, m'_1 such that $d(m_2, a) = d(m'_2, a) + 1$ and $d(m_2, b) = d(m'_2, b) + 1$. Necessarily $m_2 \preceq m'_2 \preceq a, b$. Repeat this process to get a neighbor m'_k of $m_k(= c)$ such that $d(c,a) = d(m'_k, a) + 1$ and $d(c,b) = d(m'_k, b) + 1$. This implies that $d(a,b) \leq d(a, m'_k) + d(m'_k, b) = d(a,c) + d(c,b) - 2 = d(a,b) - 2$; a contradiction.

We next show $m = a \wedge b$. Indeed, take an arbitrary $p' \in [p,q]$ with $a \succeq p' \preceq b$. Consider a median m' of a, b, p'. Since there is an ascending path from p to m' using p', m' is also a median of a, b, p, and $m' = m$ by the uniqueness. Hence $p' \preceq m$. □

Proposition 3.4. *If $p \preceq q$, then $[p,q]$ is a modular lattice, and is convex in Γ.*

Proof. We first show the convexity by verifying (3) in Lemma 2.4. Clearly the subgraph of Γ induced by $[p,q]$ is connected. Take $a, b \in [p,q]$ with $d(a,b) = 2$. We show that $I(a,b) \subseteq [p,q]$. From Lemma 3.2, this is obvious when $a \preceq b$ or $b \preceq a$. Thus we may assume $a \not\preceq b$ and $b \not\preceq a$. Consider $a \wedge b$ and $a \vee b$ (the existence of $a \wedge b$ and $a \vee b$ is guaranteed by Lemma 3.3). By $d(a,b) = 2$ and the orientability we have $a \swarrow a \vee b \searrow b \searrow a \wedge b \swarrow a$. In particular, a and b cannot have a common neighbor different from $a \vee b$ and $a \wedge b$ (by orientability). This means $I(a,b) = \{a, b, a \vee b, a \wedge b\} \subseteq [p,q]$, as required.

By Lemma 3.2, the rank function r of poset $[p,q]$ is given as

$$r(a) = d(p,a) \quad (a \in [p,q]).$$

By Lemma 3.3, $a \wedge b$ and $a \vee b$ are medians of p, a, b and of q, a, b respectively, and hence we have

$$\begin{aligned} & r(a \wedge b) + r(a \vee b) = d(a \wedge b, p) + d(a \vee b, p) \\ = \ & \{d(a,p) + d(b,p) - d(a,b)\}/2 + d(p,q) - \{d(a,q) + d(b,q) - d(a,b)\}/2 \\ = \ & \{d(a,p) + d(b,p)\}/2 + \{d(p,q) - d(a,q) + d(p,q) - d(b,q)\}/2 \\ = \ & r(a) + r(b), \end{aligned}$$

where we use (2.1) and $d(p,c) + d(c,q) = d(p,q)$ $(c \in [p,q])$. By Lemma 2.5, $[p,q]$ is a modular lattice. □

Proof of Theorem 3.1 (1). $[p,q]$ contains a Boolean lattice of rank $d(p,q)$. This means that the greatest element q is the join of atoms. Hence, by Theorem 2.9, $[p,q]$ is a complemented modular lattice. □

Since every interval of a complemented modular lattice is also a complemented modular lattice, we have the following.

Lemma 3.5. *If $p \sqsubseteq q$ and $u \in I(p,q)(= [p,q])$, then $p \sqsubseteq u \sqsubseteq q$.*

We next proceed to the proof of Theorem 3.1 (2). The following important lemma is used also in the next section.

Lemma 3.6. *If $p \sqsubseteq q \swarrow q'$, and there exists a neighbor p' of p with $p' \not\in [p,q]$ and $d(p,q) = d(p',q')$, then $p \sqsubseteq q'$.*

Proof. By Lemma 3.2 and the assumption, we have $d(p,q') = d(p,q) + d(q,q') = d(p,q) + 1 = d(p',q') + 1 = d(p,p') + d(p',q')$. Therefore, by Lemma 3.2 again, there is an ascending path from p to q' passing through p'. Necessarily $p' \in [p,q']$ holds. Since $p' \not\in [p,q]$, we have $q \vee p' = q'$. Here $[p,q]$ is complemented modular by Theorem 3.1 (1), and hence q is the join of atoms in $[p,q]$. Consequently q' is the join of atoms in $[p,q']$. Therefore $[p,q']$ is complemented modular, and includes a Boolean lattice, implying $p \sqsubseteq q'$. □

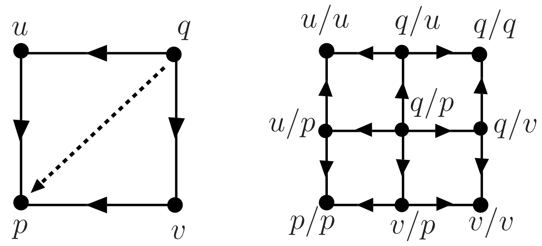

Figure 4: Orientations o and o^*

Proof of Theorem 3.1 (2). The statement that $\mathcal{L}_p$ is a semilattice immediately follows from Lemma 3.3.

Next we show the convexity. In view of Lemma 2.4, take $a, b \in \mathcal{L}_p$ with $d(a,b) = 2$, and take any common neighbor c of a, b. We show $c \in \mathcal{L}_p$. This is obvious if $a \swarrow c \swarrow b$ or $b \swarrow c \swarrow a$. Also, if $a \searrow c \swarrow b$, then $c = a \wedge b \in \mathcal{L}_p$.

So suppose that $a \swarrow c \searrow b$. Since there are two ascending paths from p to c, one including a and the other including b, we have $d(p,a) = d(p,b)$. Let $c^* := a \wedge b$, which is a common neighbor of a, b with $a \searrow c^* \swarrow b$ and $d(p, c^*) = d(p,a) - 1 = d(p,b) - 1$. Since $[p,b]$ is complemented, we can take an atom p' (neighbor of p) in $[p,b]$ such that $p' \vee c^* = b$. Then $d(p,a) = d(p',c)$, and also $p' \not\in [p,a]$ (otherwise $b = c^* \vee p' \preceq a$; a contradiction). Thus Lemma 3.6 implies $p \sqsubseteq c$, i.e., $c \in \mathcal{L}_p$.

The subgraph of Γ induced by any convex set is again a modular graph. Therefore the covering graph of $\mathcal{L}_p$ is modular. By Theorem 2.6, $\mathcal{L}_p$ is a modular semilattice. In particular, each $[p,q]$ for each $q \in \mathcal{L}_p$ is a complemented modular lattice, and $\mathcal{L}_p$ is a complemented modular semilattice. $\square$

3.2 Modular complex

Let Γ be an orientable modular graph with an admissible orientation o and a positive orbit-invariant function h. We call triple (Γ, o, h) a *modular complex*. As already seen in Theorem 3.1, (Γ, o, h) is a system of modular semilattices, and, moreover, gives rise to a geometric simplicial complex $\Delta(\Gamma, o, h)$ in the following way. For each Boolean pair (p,q) and each ascending path $(p = p_0, p_1, p_2, \ldots, p_k = q)$ from p to q, fill a k-dimensional simplex $\{x \in \mathbf{R}^k \mid 0 \leq x_j - x_{j-1} \leq h(p_{j-1}p_j)\ (j = 1, 2, \ldots, k)\}$, as in Figure 2 in the introduction. Then we obtain a geometric simplicial complex, denoted by $\Delta(\Gamma, o, h)$ (*geometric modular complex*). This generalizes the construction of the *folder complex* associated with a frame [13, 26].

Actually, we do not use this complex $\Delta(\Gamma, o, h)$ in the sequel, although our argument is based on this geometric view. Instead of dealing with $\Delta(\Gamma, o, h)$, we use a graph-theoretic operation, the 2-subdivision Γ^2 of Γ, which comes from the subdivision of $\Delta(\Gamma, o, h)$. The 2-subdivision Γ^2 is constructed as follows, where a Boolean pair $(p,q) \in \mathcal{B}(\Gamma, o)$ is denoted by q/p.

The 2-*subdivision* Γ^2 of Γ is a simple undirected graph on the set $\mathcal{B}(\Gamma, o)$ of all Boolean pairs with edges given as: q/p and q'/p' are adjacent if and only if $p = p'$ and $qq' \in E_\Gamma$ or $q = q'$ and $pp' \in E_\Gamma$. The orientation o^* for Γ^2 is given as: $q/p \swarrow_{o^*} q'/p'$ if $p = p'$ and $q \swarrow_o q'$ or if $q = q'$ and $p' \swarrow_o p$. See Figure 4. An edge joining q/p and q'/p (resp. q/p and q/p') is denoted by qq'/p (resp. q/pp'). A function $h/2$ on E_{Γ^2} is defined as $(h/2)(qq'/p) := h(qq')/2$ and $(h/2)(q/pp') := h(pp')/2$.

The main statement of this section is that this operation keeps the modularity and

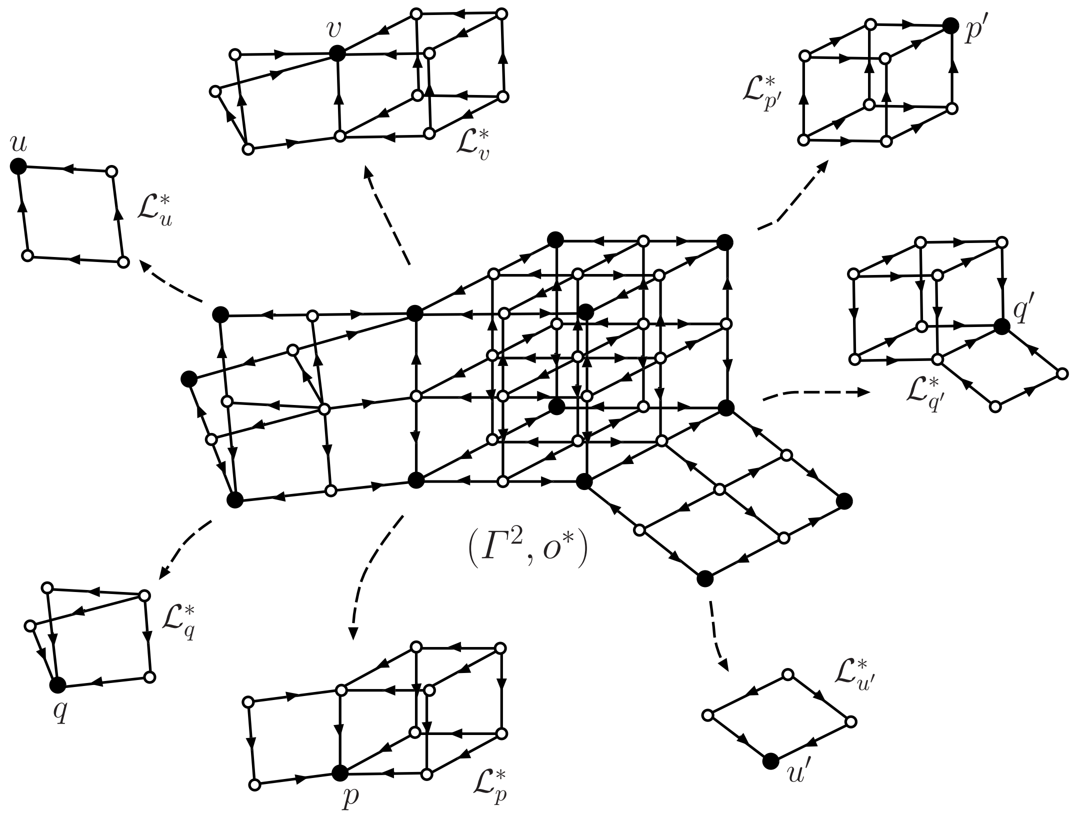

Figure 5: Construction of (Γ^2, o^*) and neighborhood semilattices

the orientability. The proof of this theorem is given in Section 3.3.

Theorem 3.7. *For an orientable modular graph Γ with an admissible orientation o and an orbit-invariant function h, the 2-subdivision Γ^2 is orientable modular, the orientation o^* is admissible, and $h/2$ is orbit-invariant.*

Hence $(\Gamma^2, o^*, h/2)$ is also a modular complex, which is called the 2-*subdivision* of (Γ, o, h). Figure 5 illustrates the 2-subdivision of (Γ, o) in Figure 2. The 2-subdivision enables us to define a neighborhood concept around vertices in Γ as follows. By embedding $p \mapsto p/p$, we can regard $V_\Gamma \subseteq V_{\Gamma^2}$. The admissible orientation o^* is oriented so that the vertices in V_Γ are all sinks. Recall the definition (3.1) of $\mathcal{L}_p = \mathcal{L}_p(\Gamma, o)$. For each vertex $p \in V_\Gamma$, define the *neighborhood semilattice* $\mathcal{L}^*_p := \mathcal{L}_p(\Gamma^2, o^*)$, which is a complementary modular semilattice with the universal lowest element p (Theorem 3.1 (2)). See Figure 5. Neighborhood semilattice $\mathcal{L}^*_p$ has much information of the local property of p than that of $\mathcal{L}_p$.

Valuation of local semilattices. A positive orbit-invariant function gives positive valuations to local semilattices. For each vertex p, modular semilattices $\mathcal{L}_p$ and $\mathcal{L}^*_p$ have positive valuations v_p and v^*_p defined by

$$(3.2) \qquad \begin{aligned} v_p(q) &:= d_{\Gamma,h}(q,p) \quad (q \in \mathcal{L}_p), \\ v^*_p(v/u) &:= d_{\Gamma^2,h/2}(v/u, p/p) \quad (v/u \in \mathcal{L}^*_p), \end{aligned}$$

respectively. In the sequel, $\mathcal{L}_p$ and $\mathcal{L}_p^*$ are supposed to be endowed with valuations v_p and v_p^*, respectively.

Isometric embedding of $(V_\Gamma, d_{\Gamma,h})$ into $(V_{\Gamma^2}, d_{\Gamma^2,h/2})$. To show that $(V_\Gamma, d_{\Gamma,h})$ is an isometric subspace of $(V_{\Gamma^2}, d_{\Gamma^2,h/2})$, we give a fundamental metric relation between Γ and Γ^2.

Proposition 3.8. *We have*

$$d_{\Gamma^2,h/2}(q/p, q'/p') = \frac{d_{\Gamma,h}(p,p') + d_{\Gamma,h}(q,q')}{2} \quad (q/p, q'/p' \in \mathcal{B}(\Gamma, o) = V_{\Gamma^2}). \tag{3.3}$$

In particular, with the embedding $p \mapsto p/p$, $(V_\Gamma, d_{\Gamma,h})$ is an isometric subspace of $(V_{\Gamma^2}, d_{\Gamma^2,h/2})$.

Proof. Take a path $P = (q/p = q_0/p_0, q_1/p_1, \ldots, q_k/p_k = q'/p')$ between q/p and q'/p'. The length of P (with respect to h) is equal to $\sum_{i=0}^{k-1}(d_{\Gamma,h}(q_i, q_{i+1}) + d_{\Gamma,h}(p_i, p_{i+1}))/2 \geq (d_{\Gamma,h}(q,q') + d_{\Gamma,h}(p,p'))/2$; hence LHS $\geq$ RHS holds.

In Lemma 3.18 in Section 3.3, we prove (3.3) for the case $h = 1$. Suppose that (3.3) is true for $h = 1$, and that P is shortest with respect to $h/2$. By Lemma 2.14, this is also shortest with respect to uniform edge-length 1/2. Necessarily the paths obtained from $(q = q_0, q_1, q_2, \ldots, q_k = q')$ and $(p = p_0, p_1, p_2, \ldots, p_k = p')$ (by identifying repetitions) are both shortest in Γ with respect to uniform edge-length 1. Again, by Lemma 2.14, they are shortest relative to h, and have the lengths $d_{\Gamma,h}(p,p')$ and $d_{\Gamma,h}(q,q')$, respectively. Thus (3.3) holds. □

The 2-subdivision operation keeps the convexity in the following sense. For a vertex set X in Γ, X^2 denotes the set of vertices $q/p \in \Gamma^2$ with $p, q \in X$.

Lemma 3.9. *For a convex set X in Γ, X^2 is convex in Γ^2.*

Proof. For $q/p, q'/p' \in X^2$, take $v/u \in I(q/p, q'/p')$. By Proposition 3.8, we have $u \in I(p,p')$ and $v \in I(q,q')$. By convexity of X, we have $u, v \in X$, i.e., $v/u \in X^2$. □

Relation between the orientations o and o^*. The relation between the orientations o and o^* is given as follows.

Lemma 3.10. (1) *For $q/p, v/u \in V_{\Gamma^2}$, $q/p \sqsubseteq_{o^*} v/u$ if and only if $q \sqsubseteq_o v$ and $u \sqsubseteq_o p$.*

(2) *For $p \in V_\Gamma$ and $v/u \in V_{\Gamma^2}$, v/u belongs to $\mathcal{L}_p^*$ if and only if $u \sqsubseteq_o p \sqsubseteq_o v$.*

Proof. (2) is a special case of (1). We show (1). From definition of o^*, $q/p \preceq_{o^*} v/u$ if and only if $q \preceq_{o^*} v$ and $u \preceq_{o^*} p$. Therefore, in proving the if part and the only if part, we can consider intervals $[q/p, v/u]$, $[q,v]$, and $[u,p]$. Take $b/a \in [q/p, v/u] = I(q/p, v/u)$. By Proposition 3.8, we have $b \in I(q,v) = [q,v]$ and $a \in I(u,p) = [u,p]$. This gives an injective map $b/a \mapsto (b,a)$ from $[q/p, v/u]$ to $[q,v] \times [u,p]$. This map preserves the partial order in the sense that $b/a \preceq_{o^*} b'/a' \Leftrightarrow a \preceq_o a'$ and $b' \preceq_o b$.

(Only if part). Suppose that $[q/p, v/u]$ contains a Boolean lattice of rank $d_{\Gamma^2}(q/p, v/u)$. Necessarily it is isomorphic to the product of a Boolean lattice of rank $d_\Gamma(q,v)$ in $[q,v]$ and a Boolean lattice of rank $d_\Gamma(u,p)$ in $[u,p]$. Thus $q \sqsubseteq_o v$ and $u \sqsubseteq_o p$.

(If part). It suffices to show that the map above is surjective, and hence bijective. Take $(b,a) \in [q,v] \times [u,p]$. Then $u \preceq_o a \preceq_o p \sqsubseteq_o q \preceq_o b \preceq_o v$. By $u \sqsubseteq_o v$ and Lemma 3.5, (a,b) is a Boolean pair, and $b/a \in I(q/p, v/u) = [q/p, v/u]$. □

Remark 3.11. The proof of Lemma 3.10 shows the fact that $[p/p, v/u]$ in $\mathcal{L}_p^*$ is isomorphic to the product of $[p, v]$ in $\mathcal{L}_p(\Gamma, o)$ and $[p, u]$ in $\mathcal{L}_p(\Gamma, o^{-1})$, where o^{-1} denotes the reverse orientation of o. Moreover, consider the set $\mathcal{L}_p^{*+}$ of element of the form q/p in $\mathcal{L}_p^*$, and the set $\mathcal{L}_p^{*-}$ of element of the form p/q in $\mathcal{L}_p^*$. By Proposition 3.8 and Lemma 2.13, both $\mathcal{L}_p^{*+}$ and $\mathcal{L}_p^{*-}$ are convex, and modular subsemilattices of $\mathcal{L}_p^*$. By $q/p \mapsto q$, $\mathcal{L}_p^{*+}$ is isomorphic to $\mathcal{L}_p(\Gamma, o)$, and, by $p/q \mapsto q$, $\mathcal{L}_p^{*-}$ is isomorphic to $\mathcal{L}_p(\Gamma, o^{-1})$. Also one can see that for every $v/u \in \mathcal{L}_p^*$ we have $v/p \in \mathcal{L}_p^{*+}$, $p/u \in \mathcal{L}_p^{*-}$, and $v/u = (p/u) \vee (v/p)$ (but $\mathcal{L}_p^*$ is not the product of $\mathcal{L}_p^{*+}$ and $\mathcal{L}_p^{*-}$ in general).

Product of modular complexes. Suppose that we are given two modular complexes (Γ, o, h) and (Γ', o', h'). Then the Cartesian product $\Gamma \times \Gamma'$ is also modular. Furthermore, define the orientation $o \times o'$ of $\Gamma \times \Gamma'$ as: $(p, p') \swarrow_{o \times o'} (p, q')$ if $p' \swarrow_{o'} q'$ and $(p, p') \swarrow_{o \times o'} (q, p')$ if $p \swarrow_o q$. Then $o \times o'$ is an admissible orientation. Similarly define $h \times h'$ by $h \times h'((p, p')(q, p')) := h(pq)$ and $h \times h'((p, p')(p, q')) := h'(p'q')$, which is orbit-invariant in $\Gamma \times \Gamma'$. Thus we obtain a new modular complex $(\Gamma \times \Gamma', o \times o', h \times h')$, which is called the *product* of (Γ, o, h) and (Γ', o', h').

Lemma 3.12. *$(p, p') \sqsubseteq_{o \times o'} (q, q')$ if and only if $p \sqsubseteq_o p'$ and $q \sqsubseteq_{o'} q'$.*

Proof. By Lemma 3.5, $(p, p') \sqsubseteq_{o \times o'} (q, q')$ implies $(p, p') \sqsubseteq_{o \times o'} (q, p') \sqsubseteq_{o \times o'} (q, q')$. This in turn implies $p \sqsubseteq_o q$ and $p' \sqsubseteq_{o'} q'$. The converse follows from the observation that $[(p, p'), (q, q')]$ is isomorphic to the product $[p, q] \times [p', q']$. □

In particular the correspondence $\mathcal{B}(\Gamma \times \Gamma', o \times o') \ni (q, q')/(p, p') \mapsto (q/p, q'/p') \in \mathcal{B}(\Gamma, o) \times \mathcal{B}(\Gamma', o')$ is bijective, and we can regard

$$\mathcal{B}(\Gamma \times \Gamma', o \times o') = \mathcal{B}(\Gamma, o) \times \mathcal{B}(\Gamma', o').$$

Under this correspondence, the product operation and the 2-subdivision operation commute in the following sense.

Lemma 3.13. (1) $\mathcal{L}_{(p,p')}(\Gamma \times \Gamma', o \times o') = \mathcal{L}_p(\Gamma, o) \times \mathcal{L}_{p'}(\Gamma', o')$.

(2) $(\Gamma \times \Gamma')^2 = \Gamma^2 \times \Gamma'^2$.

(3) $\mathcal{L}^*_{(p,p')}(\Gamma \times \Gamma') = \mathcal{L}^*_p(\Gamma) \times \mathcal{L}^*_{p'}(\Gamma')$.

Proof. (1) follows from the previous lemma. (2) follows from the fact that $(q, q')/(p, p')$ and $(v, v')/(u, u')$ have an edge in $(\Gamma \times \Gamma')^2$ if and only if $d_\Gamma(q, v) + d_{\Gamma'}(q', v') + d_\Gamma(p, u) + d_{\Gamma'}(p', u') = 1$, which is equivalent to the condition that $(q/p, q'/p')$ and $(v/u, v'/u')$ have an edge in $\Gamma^2 \times \Gamma'^2$. (3) follows from (2). □

We end this subsection with some remarks.

Remark 3.14. The 2-subdivision Γ^2 is independent of the choice of an admissible orientation of Γ. Indeed, for two admissible orientations o, o' of Γ, if $p \sqsubseteq_o q$ then there uniquely exist p', q' with $p' \sqsubseteq_{o'} q'$ and $I[p, q] = I[p', q']$. This gives a bijection between $\mathcal{B}(\Gamma, o)$ and $\mathcal{B}(\Gamma, o')$, which is in fact a graph-theoretic isomorphism between the 2-subdivisions by o and o'. Furthermore o^* is also independent of an admissible orientation of Γ. So (Γ^2, o^*) is determined by Γ only. This is a fundamental fact, but we do not use this fact here, and omit the proof. It should be emphasized that the Lovász extension of $g : V_\Gamma \to \mathbf{R}_+$ depends on an admissible orientation.

Remark 3.15. We can define the l_1-length metric d_{l_1} on $\Delta(\Gamma, o, h)$, since each simplex in $\Delta(\Gamma, o, h)$ has an isometry to a simplex in the l_1-space,; see [9, Chapter I.7] for a more precise construction of such metric simplicial complexes. Then one can see from Theorem 3.7 that *metric space* $(\Delta(\Gamma, o, h), d_{l_1})$ *is modular.* This fact also justifies the term "modular complex." Indeed, consider the successive 2-subdivision Γ^{2^k}. Then we can regard $(V_{\Gamma^{2^k}}, d_{\Gamma^{2^k}, h/2^k})$ as an isometric subspace of $(\Delta(\Gamma, o, h), d_{l_1})$. For every triple $a, b, c \in \Delta(\Gamma, o, h)$, there are convergent sequences $\{a_k\}, \{b_k\}, \{c_k\}$ such that $a_k, b_k, c_k \in V_{\Gamma^{2^k}}$ and $\lim_{k \to +\infty} a_k = a$, $\lim_{k \to +\infty} b_k = b$, and $\lim_{k \to +\infty} c_k = c$. Then we can take a median $m_k \in V_{\Gamma^{2^k}}$ of a_k, b_k, c_k for each k (with the help of the axiom of choice). Hence we obtain a sequence $\{m_k\}_k$ in $\Delta(\Gamma, o, h)$. Since Γ is a finite graph and $\Delta(\Gamma, o, h)$ is compact, we can take a convergent subsequence of $\{m_k\}$, which converges to a median of a, b, c. Hence $(\Delta(\Gamma, o, h), d_{l_1})$ is modular.

Remark 3.16. Geometric modular complex itself seems to be an interesting geometric object, although we introduced this object in the study of 0-extension problems. For example, in the case where Γ is a median graph, $\Delta(\Gamma, o, h)$ is a simplicial subdivision of the *median complex* of Γ [47]; also see [13]. Such connections as well as other metric aspects, e.g., CAT(0) property under the l_2-metrization, will be studied in a future paper.

3.3 Proof of Theorem 3.7

An arbitrary 4-cycle in Γ^2 is represented as $(q/p, q'/p, q'/p', q/p')$ for some $pp', qq' \in E_\Gamma$. This immediately implies that o^* is an admissible orientation and $h/2$ is orbit-invariant.

To show that Γ^2 is modular, we are going to verify that Γ^2 actually satisfies the two conditions of Lemma 2.1. If q/p and q'/p' are joined by an edge, then $d_\Gamma(p, q)$ and $d_\Gamma(p', q')$ have different parity. This implies:

Lemma 3.17. *Γ^2 is bipartite.*

To show the second property in Lemma 2.1, take two Boolean pairs q/p and q'/p'.

Lemma 3.18. $d_{\Gamma,1/2}(q/p, q'/p') = (d_\Gamma(p, p') + d_\Gamma(q, q'))/2$.

Proof. We have seen LHS $\geq$ RHS in the first half of the proof of Proposition 3.8. We show LHS $=$ RHS by the induction on $d_\Gamma(p, p') + d_\Gamma(q, q')$. Here $p \sqsubseteq q$. So we can consider $[p, q]$, which is a convex set, and hence a gated set (Lemma 2.4). Take the gate $p^* := \mathrm{Pr}_{[p,q]}(p')$ of p' at $[p, q]$. Then $d_\Gamma(p, p') = d_\Gamma(p, p^*) + d_\Gamma(p^*, p')$. Suppose $p^* \neq p$. Then $p \prec p^* \preceq q$. Take a neighbor u of p with $p \swarrow u \preceq p^*$. By Lemma 3.5, we have $u \sqsubseteq p^*$, and $d_\Gamma(u, p') = d_\Gamma(p, p') - 1$. By induction, $d_{\Gamma,1/2}(q/u, q'/p') = (d_\Gamma(u, p') + d_\Gamma(q, q'))/2$. Since q/p is adjacent to q/u in Γ^2, we have $d_{\Gamma,1/2}(q/p, q'/p') \leq (1 + d_\Gamma(u, p') + d_\Gamma(q, q'))/2 = (d_\Gamma(p, p') + d_\Gamma(q, q'))/2$. Then the equality holds (by LHS $\geq$ RHS).

It suffices to consider the case where $p = \mathrm{Pr}_{[p,q]}(p')$, $q = \mathrm{Pr}_{[p,q]}(q')$, $p' = \mathrm{Pr}_{[p',q']}(p)$, and $q' = \mathrm{Pr}_{[p',q']}(q)$. Then $d_\Gamma(p, q') = d_\Gamma(p, p') + d_\Gamma(p', q') = d_\Gamma(p, q) + d_\Gamma(q, q')$, and $d_\Gamma(q, p') = d_\Gamma(q, q') + d_\Gamma(q', p') = d_\Gamma(q, p) + d_\Gamma(p, p')$. This implies $d_\Gamma(p', q') = d_\Gamma(p, q)$, and $d_\Gamma(p, p') = d_\Gamma(q, q')$. Take a neighbor q^* of q in $I(q, q')$, and take a median p^* of p, p', q^*. Then $d_\Gamma(p, q) = d_\Gamma(p^*, q^*)$, and p^* does not belong to $[p, q]$. Suppose $q \swarrow q^*$. By Lemma 3.6, we have $p \sqsubseteq q^*$. Therefore q/p is adjacent to q^*/p in Γ^2 with $d_\Gamma(q, q') = 1 + d_\Gamma(q, q^*)$. Apply the induction to $(q^*/p, q'/p')$, as above. Similarly, suppose $q \searrow q^*$. Then $p \searrow p^*$, and hence $p^* \sqsubseteq q$. Therefore q/p is adjacent to q/p^* in Γ^2 with $d_\Gamma(p, p') = 1 + d_\Gamma(p, p^*)$, and apply the induction to $(q/p^*, q'/p')$. □

By using Lemma 3.18, we complete the proof of Theorem 3.7 by verifying the quadrangle condition (Lemma 2.1 (2)). We use the same notation d for d_Γ and $d_{\Gamma^2,1/2}$ (since they can be distinguished by the context). In the notation above, suppose further that we are given two neighbors q_1/p_1 and q_2/p_2 of q/p with $d(q/p, q'/p') = 1/2 + d(q_1/p_1, q'/p') = 1/2 + d(q_2/p_2, q'/p')$. We are going to show the existence of a common neighbor q^*/p^* of $q_1/p_1, q_2/p_2$ with $d(q/p, q'/p') = 1 + d(q^*/p^*, q'/p')$.

It suffices to consider the following three cases:

(i) $p_1 = p = p_2$.

(ii) $p_1 = p$, $q_1 \swarrow q$, and $q_2 = q$.

(iii) $p_1 \swarrow p = p_2$ and $q_1 = q \swarrow q_2$.

Case (i). By Lemma 3.18, we have $d(q, q') = 1 + d(q_i, q')$ for $i = 1, 2$. By Lemma 2.1 (2), there is a common neighbor q^* of q_1, q_2 with $d(q, q') = 2 + d(q^*, q')$. Here $p \sqsubseteq q_i$ $(i = 1, 2)$, and hence $q_i \in \mathcal{L}_p$. By the convexity of $\mathcal{L}_p$, we have $q^* \in \mathcal{L}_p$, implying $p \sqsubseteq q^*$. Again, by Lemma 3.18, we have $d(q/p, q'/p') = 1 + d(q^*/p, q'/p')$, as required.

Case (ii). We show $p_2 \sqsubseteq q_1$, which implies that q_1/p_2 is a required common neighbor (by Lemma 3.18). If $p_2 \swarrow p$, then $p_2 \sqsubseteq q$ and $p_2 \swarrow p \preceq q$ imply $p_2 \sqsubseteq q_1$ (Lemma 3.5). Suppose $p \swarrow p_2$. Consider $\mathrm{Pr}_{[p,q]}([p', q'])$, which is a convex set (by Theorem 2.3). By Lemma 2.13, we have $\mathrm{Pr}_{[p,q]}([p', q']) = [a, b]$ for $a, b \in [p, q]$ with $a \preceq b$. Similarly $\mathrm{Pr}_{[p',q']}([p, q]) = [a', b']$ for $a', b' \in [p', q']$ with $a' \preceq b'$. By Theorem 2.3 and Lemma 2.10, $[a, b]$ is isomorphic to $[a', b']$ with $a = \mathrm{Pr}_{[p,q]}(a')$ and $b = \mathrm{Pr}_{[p,q]}(b')$. Necessarily $d(p, p') = d(p, a) + d(a, a') + d(a', p')$. In particular, $a = \mathrm{Pr}_{[p,q]}(p')$. Hence $d(p, p') = d(p, a) + d(a, p')$ and $d(p_2, p') = d(p_2, a) + d(a, p')$. By $d(p, p') = d(p_2, p') + 1$, we have $d(p, a) = d(p_2, a) + 1$. Thus $p_2 \in I(a, p) = [a, p]$ (Lemma 3.2), implying $p_2 \preceq a$. Similarly $b \preceq q_1$. Thus $p \preceq p_2 \preceq a \preceq b \preceq q_1 \preceq q$ and $p \sqsubseteq q$ imply $p_2 \sqsubseteq q_1$ (Lemma 3.5), as required.

Case (iii). We show $p_1 \sqsubseteq q_2$. Then q_2/p_1 is a required common neighbor, as above. First we claim:

(3.4) $\quad [p_1, q] \cup [p, q_2] \neq [p_1, q_2]$, and hence $[p_1, q] \cup [p, q_2]$ is not convex.

Proof. Suppose indirectly $[p_1, q] \cup [p, q_2] = [p_1, q_2]$. As above, consider $\mathrm{Pr}_{[p_1,q_2]}([p', q'])$, which is represented by $[a, b]$. Then $a = \mathrm{Pr}_{[p_1,q_2]}(p')$ and $b' = \mathrm{Pr}_{[p_1,q_2]}(q')$. Moreover $a \wedge p = p_1$. Otherwise $p \in [p_1, a]$, implying $d(p_1, p') = d(p_1, a) + d(a, p') = 1 + d(p, a) + d(a, p') = 1 + d(p, p')$; a contradiction to $d(p_1, p') = d(p, p') - 1$. Similarly $b \vee q = q_2$. In particular, $a \notin [p, q_2]$ and $b \notin [p_1, q]$. By Theorem 2.3, $[a, b]$ is a bijective image of some interval of $[p', q']$, which is complemented modular. Therefore $[a, b]$ is also complemented modular. However this is impossible. Indeed, consider an arbitrary atom g of $[a, b]$. Then $g \preceq q$ holds since $g \not\preceq q$ implies $a = g \wedge q$, $g \in [p, q_2]$, and hence $p \preceq a$; a contradiction to $a \notin [p, q_2]$. So the join of all atoms belongs to $[p_1, q]$, and is not equal to $b \notin [p_1, q]$; a contradiction. □

By Lemma 2.4, there are $u, v \in [p_1, q] \cup [p, q_2]$ and a common neighbor w of u, v with $w \notin [p_1, q] \cup [p, q_2]$. By convexity of $[p_1, q]$ and of $[p, q_2]$ we may assume that $u \in [p_1, q] \setminus [p, q_2]$ and $v \in [p, q_2] \setminus [p_1, q]$. Then $u \swarrow w \swarrow v$, $v \swarrow w \swarrow u$, or $w \in \{u \wedge v, u \vee v\}$. The second case is impossible since $v \preceq u \preceq q$ implies $v \in [p_1, q]$. The third case is also impossible since $u \wedge v \in [p_1, q]$ and $u \vee v \in [p, q_2]$. Thus we have $u \swarrow w \swarrow v$. Then $d(w, p) = 1 + d(v, p)$; otherwise $d(v, p) = 1 + d(w, p)$ implying $w \in I(p, v) = [p, v] \subseteq [p, q_1]$ and contradicting $w \notin [p_1, q] \cup [p, q_2]$. Similarly $d(w, p) = 1 + d(u, p)$. By the quadrangle condition for w, u, v, p, there exists a common neighbor x of u, v with

$d(w,p) = 2 + d(x,p)$. Necessarily $u \swarrow x \swarrow v$ and $x \in [p,q]$. Take an atom g in $[p,q_2]$ with $x \vee g = v$. Then $d(p,w) = 1 + d(p,v) = 1 + d(p,g) + 1 = d(w,g) + 1$. Similarly we have $d(p,w) = 1 + d(u,p) = 1 + d(u,p_1) + 1 = d(w,p_1) + 1$, where the second equality follows from $u \wedge p = p_1$ (by $u \not\in [p,q_2]$), By the quadrangle condition for p, p_1, g, w, there is a common neighbor h of p_1, g with $d(w,p) = 2 + d(w,h)$. Necessarily $p_1 \swarrow h \swarrow g$, and $h \not\in [p_1, q]$ (otherwise $w \in I(h,u) \subseteq [p_1,q]$; a contradiction). Also there is an ascending path from h to q_2 (passing through h, w, v, q_2 in order) of length equal to $d(p_1,q)$, i.e., $d(h,q_2) = d(p_1,q)$. Hence, by Lemma 3.6, we have $p_1 \sqsubseteq q_2$, as required. Now the proof of Theorem 3.7 is completed. $\Box$.

4 Discrete convex functions on modular complexes

In this section, following the idea outlined in the introduction, we introduce submodular functions on a modular semilattice and L-convex functions on a modular complex. The main results in this section are :

(i) a sum of submodular functions on the product of modular semilattices can be efficiently minimized provided the arity of each summand is bounded (Theorem 4.13).

(ii) L-convex functions admit a local optimality criterion for global optimality, and checking the local optimality reduces to submodular function minimization on a modular semilattice (Theorem 4.18).

In Section 4.1, we give definitions of submodular and L-convex functions. In Sections 4.2 and 4.3 we study some technical properties of these functions. In Section 4.4, we prove (i) above by utilizing a recent result of Thapper and Živný [46] on Valued-CSP. In Section 4.5, we discuss (ii) above.

4.1 Submodular functions and L-convex functions

We first give a definition of submodular functions on a modular semilattice. Next we introduce L-convex functions on a modular complex (Γ, o, h) as functions which are *locally* submodular.

Submodular functions. Let $\mathcal{L}$ be a modular semilattice with a positive valuation v. Recall that a pair (p,q) is said to be bounded if $p \vee q \in \mathcal{L}$. A pair (p,q) is said to be *antipodal* if, for every bounded pair (p',q') in $[p \wedge q, p] \times [p \wedge q, q]$, we have

$$v[p',p]v[q',q] \geq v[p \wedge q, p']v[p \wedge q, q']. \tag{4.1}$$

A geometric meaning of this concept is the following. With elements $p \wedge q, p, p', q, q', p' \vee q'$, associate, respectively, points $(0,0), (v[p\wedge q, p], 0), (v[p \wedge q, p'], 0), (0, v[p \wedge q, q]), (0, v[p \wedge q, q']), (v[p \wedge q, p'], v[p \wedge q, q'])$ in $\mathbf{R}^2$. See Figure 6. Then a pair $(p', q') \in [p \wedge q, p] \times [p \wedge q, q]$ satisfies the inequality (4.1) if and only if $p' \vee q'$ is lower than the line segment connecting the points associated with p and q in $\mathbf{R}^2$.

Let $f : \mathcal{L} \to \mathbf{R}$ be a function on $\mathcal{L}$. For a bounded pair $p, q \in \mathcal{L}$, the *submodularity inequality* is:

$$f(p) + f(q) \geq f(p \wedge q) + f(p \vee q).$$

For a pair of elements $p, q \in \mathcal{L}$, the $\wedge$-*convexity inequality* (*meet-convexity inequality*) with respect to v is:

$$v[p \wedge q, q]f(p) + v[p \wedge q, p]f(q) \geq (v[p \wedge q, p] + v[p \wedge q, q])f(p \wedge q).$$

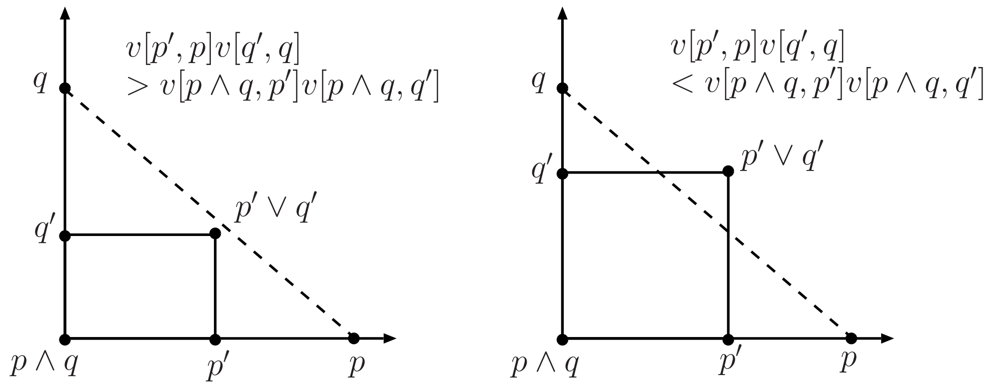

Figure 6: Antipodality

One may get an intuition of this condition by associating $p, p\wedge q, q$ with points $-v[p\wedge q, p]$, 0, $v[p \wedge q, q]$, respectively, in $\mathbf{R}$.

A function $f : \mathcal{L} \to \mathbf{R}$ is called *submodular* (relative to v) if f satisfies the submodularity inequality for every bounded pair and the $\wedge$-convexity inequality for every antipodal pair. Every incomparable bounded pair (p, q) is not antipodal since by taking (p, q) as (p', q') in (4.1) we have $v[p, p]v[q, q] = 0 < v[p \wedge q, p]v[p \wedge q, q]$. Therefore, in the case of a (modular) lattice, there is no nontrivial antipodal pair, and our submodular functions coincide with submodular functions on lattices in the ordinary sense. Note also that valuation v is not referred to in the submodularity inequality.

L-convex functions. Next we define the concept of an L-convex function $g : V_\Gamma \to \mathbf{R}$ for a modular complex (Γ, o, h). Consider the 2-subdivision $(\Gamma^2, o^*, h/2)$ of (Γ, o, h). Define $\bar{g} : V_{\Gamma^2} \to \mathbf{R}$ by

$$\bar{g}(q/p) := \frac{g(p) + g(q)}{2} \quad (q/p \in \mathcal{B}(\Gamma, o) = V_{\Gamma^2}). \tag{4.2}$$

This is the restriction of the Lovász extension of g; see the introduction for the Lovász extension. Recall from Section 3.2 the notion of the neighborhood semilattice $\mathcal{L}^*_p = \mathcal{L}_p(\Gamma^2, o^*)$, which is also a (complemented) modular semilattice. The *derivative* $\mathrm{d}^*_p g$ of g at p is a function on $\mathcal{L}^*_p$ defined as

$$\mathrm{d}^*_p g(q/q') := \bar{g}(q/q') - g(p) \quad (q/q' \in \mathcal{L}^*_p).$$

A function g on V_Γ is said to be *L-convex* on (Γ, o, h) if, for every $p \in V_\Gamma$, the derivative $\mathrm{d}^*_p g$ is submodular on $\mathcal{L}^*_p$ relative to the valuation v^*_p defined in (3.2).

Remark 4.1. To see "discrete convexity" in this definition, consider the case where Γ is a path with the unit orbit-invariant function. Suppose that $V_\Gamma = \{1, 2, \dots, m\}$ and $E_\Gamma = \{ij \mid j = i + 1 \ (i = 1, 2, \dots, m - 1)\}$. Every orientation o is admissible. Then the 2-subdivision Γ^2 is the subdivision of Γ (in the ordinary sense). The new vertex between i and $i + 1$ is denoted by $i + 1/2$. Let g be a function on $V_\Gamma = \{1, 2, \dots, m\}$. Then $\bar{g} : V_{\Gamma^2} \to \mathbf{R}$ is given as

$$\bar{g}(j) = \begin{cases} g(j) & \text{if } j \in \{1, 2, \dots, m\}, \\ (g(i) + g(i + 1))/2 & \text{if } j = i + 1/2 \text{ for } i \in \{1, 2, \dots, m - 1\}. \end{cases}$$

The neighborhood semilattice $\mathcal{L}_i^*$ at $i = 1, 2, \ldots, m-1$ is given by $\mathcal{L}_i^* = \{i, i+1/2, i-1/2\}$ with ordering $i - 1/2 \succ_{o^*} i \prec_{o^*} i + 1/2$; if $i = 0, m$, then $\mathcal{L}_i^*$ is a 2-element lattice of the lowest element i. The derivative $\mathrm{d}_i^* g$ is given by

$$\begin{aligned} \mathrm{d}_i^* g(i) &= 0, \\ \mathrm{d}_i^* g(i-1/2) &= (g(i-1) - g(i))/2, \\ \mathrm{d}_i^* g(i+1/2) &= (g(i+1) - g(i))/2. \end{aligned}$$

Here $\mathcal{L}_i^*$ has no incomparable bounded pair and has only one incomparable antipodal pair $(i-1/2, i+1/2)$. Hence $\mathrm{d}_i^* g$ is submodular if and only if $\mathrm{d}_i^* g(i+1/2) + \mathrm{d}_i^* g(i-1/2) \geq 2\mathrm{d}_i^* g(i)$, that is

$$g(i-1) + g(i+1) \geq 2g(i). \tag{4.3}$$

Consequently, g is L-convex if and only if (4.3) holds for $i = 2, 3, \ldots, m-1$. This is nothing but a 1-dimensional convexity condition.

Consider, more generally, the product Γ of several paths. As mentioned in the introduction, the product of paths is naturally identified with a box subset B of integer lattice $\mathbf{Z}^n$. In contrast with the 1-dimensional case, there are many admissible orientations that yields different classes of discrete convex functions. Among them, consider the admissible orientation o defined as: $x \swarrow_o y \Leftrightarrow x \leq y$ for $x, y \in B(\subseteq \mathbf{Z}^n)$ with $xy \in E_\Gamma$. Then, in fact, L-convex functions on $(\Gamma, o, 1)$ coincide with $L^\natural$*-convex functions* on B in discrete convex analysis [39, Chapter 7]. This relation may not be obvious at first glance. We will give, in a future paper [22], detailed discussions on this relation as well as links to other L-convex/submodular-type functions mentioned in the introduction.

4.2 (p, q)-envelope

Here we introduce the concept of (p, q)-envelope of a modular semilattice, which plays a important role in the proof of Theorems 4.13 and 4.18.

Let $\mathcal{L}$ be a modular semilattice with a positive valuation v. Take a pair (p, q) of elements in $\mathcal{L}$. Recall from Lemma 2.8 that $u \in I(p, q)$ if and only if $u = a \vee b$ for $a \in [p \wedge q, p]$ and $b \in [p \wedge q, q]$, and such (a, b) is uniquely determined by $a = u \wedge p$ and $b = u \wedge q$. Then we can define map $\varphi = \varphi^{p,q} : I(p, q) \to \mathbf{R}_+^2$ by

$$\varphi(u) := (v[p \wedge q, p \wedge u], v[p \wedge q, q \wedge u]) \quad (u \in I(p, q)). \tag{4.4}$$

Note that $\varphi^{p,q}$ and $\varphi^{q,p}$ are different. Then $\varphi(I(p, q))$ is a finite set of points in a box $[0, v[p \wedge q, p]] \times [0, v[p \wedge q, q]]$ (by modularity of v), and always includes $\varphi(p \wedge q) = (0, 0)$, $\varphi(p) = (v[p \wedge q, p], 0)$, and $\varphi(q) = (0, v[p \wedge q, q])$.

Our interest lies in the convex hull $\mathrm{Conv}(\varphi(I(p, q)))$ of $\varphi(I(p, q))$. From the definition, we can easily see that

(4.5) (1) (p, q) is bounded $\Leftrightarrow \mathrm{Conv}(\varphi(I(p, q))) = \mathrm{Conv}(0, \varphi(p), \varphi(q), \varphi(p \vee q))$.

(2) (p, q) is antipodal $\Leftrightarrow \mathrm{Conv}(\varphi(I(p, q))) = \mathrm{Conv}(0, \varphi(p), \varphi(q))$.

The (p, q)-*envelope* $\mathcal{E}^{p,q}$ is the set of elements $u \in I(p, q) \setminus \{p \wedge q\}$ such that $\varphi(u)$ is an extreme point of $\mathrm{Conv}(\varphi(I(p, q)))$. We will see in Lemma 4.3 that map φ is injective on $\mathcal{E}^{p,q}$, and hence is a bijection between $\mathcal{E}^{p,q}$ and the set of extreme points of $\mathrm{Conv}(\varphi(I(p, q)))$ other than $(0, 0)$. By convention, let $\mathcal{E}^{p,q} := \{p \wedge q\}$ if $p = q$. In particular, by (4.5), (p, q) is bounded if and only if $\mathcal{E}^{p,q} = \{p, p \vee q, q\}$, and (p, q) is antipodal if and only if $\mathcal{E}^{p,q} = \{p, q\}$.

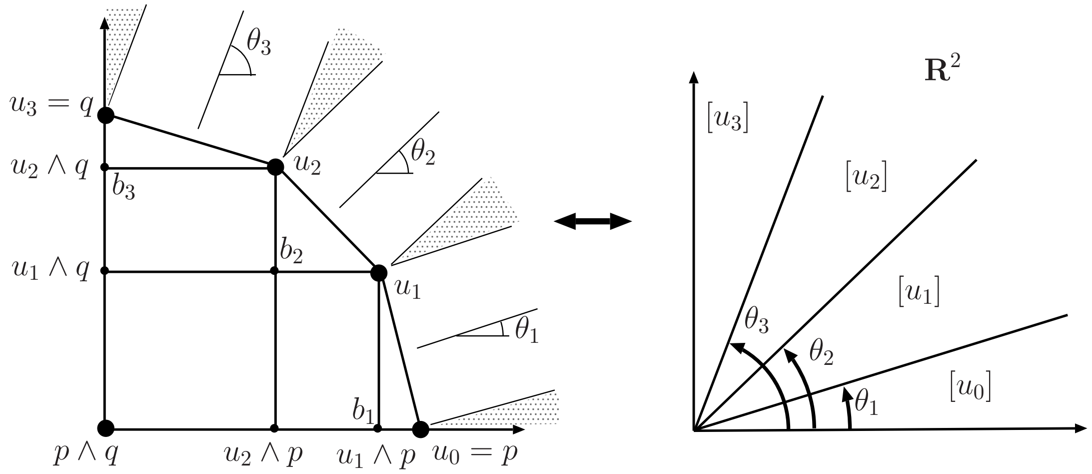

Figure 7: (p, q)-envelope $\mathcal{E}^{p,q}$ and normal cone decomposition

For $u \in \mathcal{E}^{p,q}$, let $[u](= [u]^{p,q})$ denote the set of nonnegative vectors $w \in \mathbf{R}^2_+$ with $\langle w, \varphi(u)\rangle = \max_{u' \in \mathcal{E}} \langle w, \varphi(u')\rangle$, where $\langle , \rangle$ is the standard inner product. Then $[u]$ forms a closed convex cone in $\mathbf{R}^2_+$; $[u]$ is nothing but the intersection of $\mathbf{R}^2_+$ and the normal cone at $\varphi(u)$ of $\mathrm{Conv}(\varphi(I(p, q)))$. See Figure 7. Every closed convex cone C in $\mathbf{R}^2_+$ is uniquely represented as

$$C = \{(x, y) \in \mathbf{R}^2_+ \mid -x \sin \alpha + y \cos \alpha \leq 0,\ -x \sin \beta + y \cos \beta \geq 0\}$$

for some $0 \leq \beta \leq \alpha \leq \pi/2$. Define $\nu(C)$ by

$$\nu(C) := \frac{\sin \alpha}{\cos \alpha + \sin \alpha} - \frac{\sin \beta}{\cos \beta + \sin \beta} = \frac{\sin(\alpha - \beta)}{(\cos \alpha + \sin \alpha)(\cos \beta + \sin \beta)} (\geq 0). \tag{4.6}$$

The main result of this subsection is the following.

Theorem 4.2. *Let $\mathcal{L}$ be a modular semilattice with a positive valuation v and let f be a submodular function on $\mathcal{L}$. Then we have*

$$f(p) + f(q) \geq f(p \wedge q) + \sum_{u \in \mathcal{E}^{p,q}} \nu([u]) f(u) \quad ((p, q) \in \mathcal{L} \times \mathcal{L}). \tag{4.7}$$

The inequality (4.7) generalizes and unifies the submodular inequality and the $\wedge$-convexity inequality. Indeed, if (p, q) is bounded, then $\mathcal{E}^{p,q} = \{p, p \vee q, q\}$, $\nu([p]) = \nu([q]) = 0$ and $\nu([p \vee q]) = 1$; (4.7) coincides with the submodularity inequality for (p, q). If (p, q) is antipodal, then $\mathcal{E}^{p,q} = \{p, q\}$, $\nu([p]) = v[p \wedge q, p]/(v[p \wedge q, p] + v[p \wedge q, q])$, and $\nu([q]) = v[p \wedge q, q]/(v[p \wedge q, p] + v[p \wedge q, q])$; (4.7) coincides with the $\wedge$-convexity inequality for (p, q). Therefore we can employ (4.7) as the definition of submodular functions.

Proof of Theorem 4.2. Take a pair (p, q) of elements in $\mathcal{L}$ and consider (p, q)-envelope $\mathcal{E}^{p,q}$. We may assume $v(p \wedge q) = 0$ (for notational simplicity). If $u \in I(p, q)$, then $u \wedge p, u \wedge q \in I(p, q)$, and $[0, v(u \wedge p)] \times [0, v(u \wedge q)] \subseteq \mathrm{Conv}(\varphi(I(p, q)))$. Therefore, for $a, b \in \mathcal{E}^{p,q}$, $v(a \wedge p) \leq v(b \wedge p)$ implies $v(a \wedge q) \geq v(b \wedge q)$, and $v(a \wedge p) \geq v(b \wedge p)$ implies $v(a \wedge q) \leq v(b \wedge q)$.

Lemma 4.3. *For $a, b \in \mathcal{E}^{p,q}$, if $v(a \wedge p) \leq v(b \wedge p)$ and $v(a \wedge q) \geq v(b \wedge q)$, then $a \wedge p \preceq b \wedge p$ and $a \wedge q \succeq b \wedge q$. In particular, $\varphi(a) = \varphi(b)$ implies $a = b$.*

Proof. For $u \in I(p,q)$, we denote $u \wedge p$ and $u \wedge q$ by u_p and u_q, respectively. Then all pairs (a_p, b_p), $(a_p, a_q \wedge b_q)$, and $(b_p, a_q \wedge b_q)$ from among the triple $(a_p, b_p, a_q \wedge b_q)$ are bounded. Hence, by definition of modular semilattices, their join $\eta = (a_p \vee b_p) \vee (a_q \wedge b_q)$ exists and belongs to $I(p,q)$. Similarly the join $\xi = (a_q \vee b_q) \vee (a_p \wedge b_p)$ of the triple $(a_q, b_q, a_p \wedge b_p)$ exists and belongs to $I(p,q)$. Then $(\eta_p, \eta_q) = (a_p \vee b_p, a_q \wedge b_q)$ and $(\xi_p, \xi_q) = (a_p \wedge b_p, a_q \vee b_q)$ (by Lemma 2.8 (3)). Since $a_p, b_p \in [\xi_p, \eta_p]$ and $a_q, b_q \in [\eta_q, \xi_q]$, both $\varphi(a)$ and $\varphi(b)$ belong to the box $[\varphi(\xi_p), \varphi(\eta_p)] \times [\varphi(\eta_q), \varphi(\xi_q)]$. By modularity (2.3) of v, we have

$$\begin{aligned} \varphi(\eta) + \varphi(\xi) &= (v(a_p \vee b_p) + v(a_p \wedge b_p), v(a_q \wedge b_q) + v(a_q \vee b_q)) \\ &= (v(a_p) + v(b_p), v(a_q) + v(b_q)) = \varphi(a) + \varphi(b). \end{aligned}$$

This implies $\varphi(a) = \varphi(\eta)$ and $\varphi(b) = \varphi(\xi)$, since $\varphi(a)$ and $\varphi(b)$ are extreme points of $\mathrm{Conv}(\varphi(I(p,q)))$. Since $\eta_p \succeq a_p$, $\eta_q \preceq a_q$, $\xi_p \preceq b_p$, and $\xi_q \succeq b_q$, we must have $(\eta_p, \eta_q) = (a_p, a_q)$ and $(\xi_p, \xi_q) = (b_p, b_q)$, implying $\eta = a$ and $\xi = b$ (by Lemma 2.8 (3)). Hence $a_p = a_p \vee b_p$ and $a_q = a_q \wedge b_q$, implying $a_p \succeq b_p$ and $a_q \preceq b_q$. □

Therefore we can arrange $\mathcal{E}^{p,q}$ into a sequence $(p = u_0, u_1, \ldots, u_k = q)$ such that

$$u_i \wedge p \succeq u_j \wedge p, \quad u_i \wedge q \preceq u_j \wedge q \quad (i \leq j). \tag{4.8}$$

Lemma 4.4. *For $i \leq j$, we have the following.*

(1) $u_i \wedge u_j = (u_j \wedge p) \vee (u_i \wedge q)$.

(2) $u_i = (u_i \wedge u_j) \vee (u_i \wedge p)$ *and* $u_j = (u_i \wedge u_j) \vee (u_j \wedge q)$.

(3) $d_\Gamma(p,q) = d_\Gamma(p,u_i) + d_\Gamma(u_i,u_j) + d_\Gamma(u_j,q)$, *and hence* $I(u_i,u_j) \subseteq I(p,q)$, *where* Γ *is the covering graph of* $\mathcal{L}$.

Proof. (1). By Lemma 2.8 and (4.8), we have $u_j = (u_j \wedge p) \vee (u_j \wedge q) = ((u_j \wedge p) \vee (u_i \wedge q)) \vee (u_j \wedge q)$. Here $(u_j \wedge p) \vee (u_i \wedge q) \in [u_i \wedge q, u_i]$ (by $u_j \wedge p \preceq u_i \wedge p \preceq u_i$). By applying Lemma 2.8 to (q, u_i), we get $(u_j \wedge p) \vee (u_i \wedge q) = u_j \wedge u_i$.

(2). By (1), $(u_i \wedge u_j) \vee (u_i \wedge p) = (u_j \wedge p) \vee (u_i \wedge q) \vee (u_i \wedge p) = (u_i \wedge p) \vee (u_i \wedge q) = u_i$. Similar for the second equality.

(3). By Lemma 2.8, $u_i \wedge u_j \in I(u_i,u_j)$, By (1), $u_i \wedge u_j \in I(p,q)$. By (2), $u_i \in I(p, u_i \wedge u_j)$ and $u_j \in I(u_i \wedge u_j, q)$. Thus we have $d_\Gamma(p,u_i) + d_\Gamma(u_i,u_j) + d_\Gamma(u_j,q) = d_\Gamma(p,u_i) + d_\Gamma(u_i, u_i \wedge u_j) + d_\Gamma(u_i \wedge u_j, u_j) + d_\Gamma(u_j,q) = d_\Gamma(p, u_i \wedge u_j) + d_\Gamma(u_i \wedge u_j, q) = d_\Gamma(p,q)$. □

Lemma 4.5. *For $c \in I(u_i,u_j)$ $(i < j)$, we have $\varphi^{p,q}(c) = \varphi^{u_i,u_j}(c) + \varphi^{p,q}(u_i \wedge u_j)$.*

Proof. Since $u_i \wedge u_j \preceq c$ we have $u_i \wedge u_j \wedge p \preceq c \wedge p$ and $u_i \wedge u_j \wedge q \preceq c \wedge q$. Hence

$$\begin{aligned} v(c \wedge p) &= v(u_i \wedge u_j \wedge p) + v[u_i \wedge u_j \wedge p, c \wedge p], \\ v(c \wedge q) &= v(u_i \wedge u_j \wedge q) + v[u_i \wedge u_j \wedge q, c \wedge q]. \end{aligned}$$

So we show $v[u_i \wedge u_j \wedge p, c \wedge p] = v[u_i \wedge u_j, c \wedge u_i]$ (and $v[u_i \wedge u_j \wedge q, c \wedge q] = v[u_i \wedge u_j, c \wedge u_j]$). By modularity (2.3), it suffices to show (i) $(c \wedge p) \vee (u_i \wedge u_j) = u_i \wedge c$ and (ii) $(c \wedge p) \wedge (u_i \wedge u_j) = u_i \wedge u_j \wedge p$. The second equation (ii) is easy; $(c \wedge p) \wedge (u_j \wedge u_i) = (c \wedge u_i) \wedge (u_i \wedge u_j) \wedge p = u_i \wedge u_j \wedge p$ (by $u_i \wedge u_j \preceq c \wedge u_i$). We show the first equation (i). From $c \in I(u_i,u_j) \subseteq I(u_j,p)$ (Lemma 4.4 (3)), we have $u_i \in I(c,p)$, which implies $c \wedge u_i \in [c \wedge p, c]$ by Lemma 2.8. Then we have $c \wedge u_i \wedge p = c \wedge p$. Also from $u_i \wedge q = u_i \wedge u_j \wedge q \preceq c \wedge u_i \wedge q \preceq u_i \wedge q$ we have $u_i \wedge c \wedge q = u_i \wedge q$. Therefore we obtain $(c \wedge p) \vee (u_j \wedge u_i) = (c \wedge u_i \wedge p) \vee (u_j \wedge p) \vee (u_i \wedge q) = (c \wedge u_i \wedge p) \vee (c \wedge u_i \wedge q) = c \wedge u_i$. □

Lemma 4.6. *For $i = 0, 1, \ldots, k-1$, (u_i, u_{i+1}) is antipodal.*

Proof. The line segment $[\varphi(u_i), \varphi(u_{i+1})]$ is an edge of $\mathrm{Conv}(\varphi(I(p,q)))$. By (4.5) and Lemma 4.5, if (u_i, u_{i+1}) is not antipodal, then there is $c \in I(u_i, u_{i+1}) \subseteq I(p,q)$ such that $\varphi(c)$ goes beyond $[\varphi(u_i), \varphi(u_{i+1})]$; a contradiction to $\varphi(c) \in \mathrm{Conv}(\varphi(I(p,q)))$. □

Let $b_i := u_{i-1} \wedge u_i$ for $i = 1, 2, \ldots, k$ (see Figure 7). Then we have

$$b_i \wedge (u_i \wedge q) = u_{i-1} \wedge q, \quad b_i \vee (u_i \wedge q) = u_i.$$

The first equality follows from (4.8), and the second equality follows from Lemma 4.4 (2). Let f be a submodular function on $\mathcal{L}$. Therefore by submodularity inequalities we have

$$f(b_i) + f(u_i \wedge q) \geq f(u_{i-1} \wedge q) + f(u_i) \quad (i = 1, 2, \ldots, k). \tag{4.9}$$

By adding (4.9) for $i = 1, 2, \ldots, k$ and $f(p) = f(u_0)$ (recall $(p,q) = (u_0, u_k)$), we get

$$f(p) + f(b_1) + f(b_2) + \cdots + f(b_k) + f(q) \geq f(p \wedge q) + f(u_0) + f(u_1) + \cdots + f(u_k). \tag{4.10}$$

For $i = 1, 2, \ldots, k$, let θ_i be the angle of the line normal to the line segment connecting $\varphi(u_{i-1})$ and $\varphi(u_i)$. Namely $\theta_i = \arctan v[b_i, u_{i-1}]/v[b_i, u_i]$. See Figure 7. Here (u_{i-1}, u_i) is antipodal by Lemma 4.6. Then the $\wedge$-convexity inequality for antipodal pair (u_{i-1}, u_i) with $b_i = u_{i-1} \wedge u_i$ is rewritten as

$$f(u_{i-1}) \geq f(b_i) + \frac{\sin\theta_i}{\sin\theta_i + \cos\theta_i} f(u_{i-1}) - \frac{\sin\theta_i}{\sin\theta_i + \cos\theta_i} f(u_i) \quad (i = 1, 2, \ldots, k). \tag{4.11}$$

Substituting (4.11) to (4.10), we get

$$\begin{aligned} f(p) + f(q) &\geq f(p \wedge q) + \sum_{i=0}^{k} \left(\frac{\sin\theta_{i+1}}{\sin\theta_{i+1} + \cos\theta_{i+1}} - \frac{\sin\theta_i}{\sin\theta_i + \cos\theta_i} \right) f(u_i) \\ &= f(p \wedge q) + \sum_{u \in \mathcal{E}^{p,q}} \nu([u]) f(u), \end{aligned}$$

where $\theta_0 := 0$ and $\theta_k := \pi/2$. □

4.3 Restrictions and products

In this subsection, we study the behavior of L-convexity/submodularity under restrictions and products.

Convex restriction. Let $\mathcal{M}$ be a convex set of a modular semilattice $\mathcal{L}$, which is necessarily a modular semilattice (Lemma 2.13). The restriction of a positive valuation v of $\mathcal{L}$ to $\mathcal{M}$ gives a positive valuation of $\mathcal{M}$. Let (Γ, o, h) be a modular complex. For a convex set X in Γ, the induced subgraph $\Gamma[X]$ is also modular. The restrictions of o and h to $\Gamma[X]$ are admissible and orbit-invariant in $\Gamma[X]$, respectively. Hence we get a modular complex $(\Gamma[X], o, h)$.

Lemma 4.7. (1) *If $f : \mathcal{L} \to \mathbf{R}$ is submodular on $\mathcal{L}$, then f is submodular on $\mathcal{M}$ for any convex set $\mathcal{M} \subseteq \mathcal{L}$ (regarded as a modular semilattice).*

(2) *If $g : V_\Gamma \to \mathbf{R}$ is L-convex on (Γ, o, h), then g is L-convex on $(\Gamma[X], o, h)$ for any convex set $X \subseteq V_\Gamma$.*

Proof. (1). By convexity of $\mathcal{M}$, we have $I(p,q) \subseteq \mathcal{M}$. Hence a pair (p,q) is antipodal in $\mathcal{M}$ if and only if it is antipodal in $\mathcal{L}$. (2) follows from (1) and Lemma 3.9. □

Recall that each $\mathcal{L}_p$ is endowed with a positive valuation v_p defined in (3.2). Then the restriction of an L-convex function to $\mathcal{L}_p$ is submodular.

Lemma 4.8. *For every vertex $p \in V_\Gamma$, L-convex function g is submodular on $\mathcal{L}_p$.*

Proof. Consider the set $\mathcal{L}_p^{*+}$ of elements in $\mathcal{L}_p^*$ of the form q/p. $\mathcal{L}_p^{*+}$ is convex in $\mathcal{L}_p^*$ and is a modular subsemilattice of $\mathcal{L}_p^*$; see Remark 3.11. Therefore $\mathrm{d}_p^* g$ is submodular on $\mathcal{L}_p^{*+}$ by Lemma 4.7 (1). Here $\mathcal{L}_p^{*+}$ is isomorphic to $\mathcal{L}_p$ by $q/p \mapsto q$. By using relation $g(q) = g(p) + 2\mathrm{d}_p^* g(q/p)$ $(q \in \mathcal{L}_p)$, we see the submodularity of g on $\mathcal{L}_p$. □

Product and sections. Let $\mathcal{L}$ be the product of modular semilattices $\mathcal{L}_i$ indexed by $i \in I$, where $I = \{1, 2, \ldots, n\}$. For $p \in \mathcal{L}$, the i-th component of p is denoted by p_i, and p is represented as $p = (p_1, p_2, \ldots, p_n) \in \mathcal{L}_1 \times \mathcal{L}_2 \times \cdots \times \mathcal{L}_n$. Suppose that each $\mathcal{L}_i$ has a positive valuation v_i. Valuation v of $\mathcal{L}$ is given by $\sum_{i \in I} v_i$.

For $p, q \in \mathcal{L}$, $I(p,q)$ is the product of $I(p_i, q_i)$ over $i = 1, 2, \ldots, n$. Consequently, we have

$$\varphi^{p,q}(u) = \sum_{i \in I} \varphi^{p_i,q_i}(u_i) \quad (u \in I(p,q)).$$

Hence we obtain the following decomposition property of $\mathrm{Conv}(\varphi^{p,q}(I(p,q)))$.

Lemma 4.9. $\mathrm{Conv}(\varphi^{p,q}(I(p,q))) = \sum_{i=1}^n \mathrm{Conv}(\varphi^{p_i,q_i}(I(p_i,q_i)))$, *where the sum means the Minkowski sum.*

For $J \subseteq I$ and $q_l \in \mathcal{L}_l$ for $l \in I \setminus J$, the set of elements $(p_i)_{i \in I}$ in $\mathcal{L}$ with $p_l = q_l$ for $l \in I \setminus J$ is called the *section* of $\mathcal{L}$ with respect to J and $(q_l)_{l \in I \setminus J}$, and is particularly called a *k-section* when $|J| = k$. Let (Γ, o, h) be the product of modular complexes (Γ_i, o_i, h_i) indexed by $i \in I$. We can define sections of V_Γ analogously. Since any section is convex, this operation yields a modular complex induced by a section, which is called a section of (Γ, o, h). By Lemma 4.7 and the fact that every section is convex, we have the following.

Lemma 4.10. (1) *If f is submodular on $\mathcal{L}$, then f is submodular on every section of $\mathcal{L}$.*

(2) *If g is L-convex on (Γ, o, h), then g is L-convex on every section of (Γ, o, h).*

A bounded pair (p,q) in $\mathcal{L}$ is said to be *2-bounded* if $p \vee q$ covers p and q (in which case both p and q cover $p \wedge q$). Note that every 2-bounded pair necessarily belongs to a 2-section. The following criterion is useful to check the submodularity.

Proposition 4.11. *f is submodular on $\mathcal{L}$ if and only if*

(1) *f satisfies the submodularity inequality for every 2-bounded pair, and*

(2) *f satisfies the $\wedge$-convexity inequality for every antipodal pair belonging to a 1-section.*

Proof. The only if part is obvious. We prove the if part. We first show that every submodularity inequality is implied by submodularity inequalities for 2-bounded pairs. For a bounded pair (p,q), take maximal chains $(p \wedge q = p_0, p_1, \ldots, p_k = p)$ and $(p \wedge q = q_0, q_1, \ldots, q_l = q)$. Let $a_{i,j} := p_i \vee q_j$. Then $f(p) + f(q) - f(p \wedge q) - f(p \vee q) =$

$\sum_{i,j}(f(a_{i+1,j}) + f(a_{i,j+1}) - f(a_{i+1,j+1}) - f(a_{i,j})) \geq 0$. Here we use the fact seen from modularity that $(a_{i+1,j}, a_{i,j+1})$ is a 2-bounded pair with $a_{i+1,j+1} = a_{i+1,j} \vee a_{i,j+1}$ and $a_{i,j} = a_{i+1,j} \wedge a_{i,j+1}$.

Next we show the $\wedge$-convexity inequality. We may consider the case $\mathcal{L} = \mathcal{L}_1 \times \mathcal{L}_2$. Take an (incomparable) antipodal pair $(p, q) = ((p_1, p_2), (q_1, q_2))$ in $\mathcal{L}$. Then $\mathcal{E}^{p,q} = \{p, q\}$. By (4.5), $\mathrm{Conv}(\varphi^{p,q}(I(p, q)))$ is a triangle Δ. By Lemma 4.9, $\Delta = \mathrm{Conv}(\varphi^{p_1,q_1}(I(p_1, q_1))) + \mathrm{Conv}(\varphi^{p_2,q_2}(I(p_2, q_2)))$. Thus $\mathrm{Conv}(\varphi^{p_i,q_i}(I(p_i, q_i)))$ is a triangle congruent to Δ for $i = 1, 2$. Hence (p_i, q_i) is antipodal in $\mathcal{L}_i$ with $\nu([p]) = \nu([p_1]) = \nu([p_2])$ and $\nu([q]) = \nu([q_1]) = \nu([q_2])$. Therefore both $((q_1, q_2), (q_1, p_2))$ and $((q_1, p_2), (p_1, p_2))$ are antipodal pairs belonging 1-sections of $\mathcal{L}$. Hence we have

$$
\begin{aligned}
(1 - \nu([q_2]))f(q_1, q_2) + (1 - \nu([p_2]))f(q_1, p_2) &\geq f(q_1, p_2 \wedge q_2),\\
(1 - \nu([q_1]))f(q_1, p_2) + (1 - \nu([p_1]))f(p_1, p_2) &\geq f(p_1 \wedge q_1, p_2).
\end{aligned}
$$

Also, by submodularity inequality (shown above), we have

$$
f(q_1, p_2 \wedge q_2) + f(p_1 \wedge q_1, p_2) \geq f(p_1 \wedge q_1, p_2 \wedge q_2) + f(q_1, p_2).
$$

From the three inequalities, we obtain

$$
\begin{aligned}
&(1 - \nu([q]))f(q_1, q_2) + (2 - \nu([p]) - \nu([q]))f(q_1, p_2) + (1 - \nu([p]))f(p_1, p_2)\\
&\quad \geq f(p_1 \wedge q_1, p_2 \wedge q_2) + f(q_1, p_2).
\end{aligned}
$$

By using $\nu([p]) + \nu([q]) = 1$, we get the $\wedge$-convexity inequality for (p, q). □

In particular, f is submodular if and only if f is submodular on every 2-section, and g is L-convex if and only if g is L-convex on every 2-section.

We next discuss extension properties of L-convex/submodular functions. For $J = \{i_1, i_2, \ldots, i_K\} \subseteq I$, consider the product $\mathcal{L}_J$ of $\mathcal{L}_j$ over $j \in J$. A function f_J on $\mathcal{L}_J$ can be naturally extended to a function f on $\mathcal{L}$ by

(4.12) $\quad f(p) := f_J(p_{i_1}, p_{i_2}, \ldots, p_{i_K}) \quad (p = (p_1, p_2, \ldots, p_n) \in \mathcal{L} = \mathcal{L}_1 \times \mathcal{L}_2 \times \ldots \times \mathcal{L}_n).$

Similarly, for $J \subseteq I$, consider the product (Γ_J, o_J, h_J) of (Γ_j, o_j, h_j) over $j \in J$. As above, a function $g_J : V_{\Gamma_J} \to \mathbf{R}$ is extended to $g : V_\Gamma \to \mathbf{R}$.

Lemma 4.12. (1) *If f_J is submodular on $\mathcal{L}_J$ then f is submodular on $\mathcal{L}$.*

(2) *If g_J is L-convex on (Γ_J, o_J, h_J) then g is L-convex on (Γ, o, h).*

Proof. (1). It is easy to see that if f_J satisfies the submodularity inequality on $\mathcal{L}_J$ then so does f on $\mathcal{L}$. The $\wedge$-convexity inequalities follows from Proposition 4.11. (2) follows from (1). □

4.4 Minimizing a sum of submodular functions with bounded arity: an approach from Valued-CSP

We consider the problem of minimizing submodular function f on the product $\mathcal{L} = \mathcal{L}_1 \times \mathcal{L}_2 \times \cdots \times \mathcal{L}_n$ of modular semilattices, where the input of the problem is $\mathcal{L}_1, \mathcal{L}_2, \ldots, \mathcal{L}_n$ and an evaluating oracle of f. In the case where each $\mathcal{L}_i$ is a lattice of rank 1, this problem is the submodular set function minimization in the ordinary sense, and can be solved by polynomial time [17, 24, 44]. However, we do not know whether this problem in general is polynomial time solvable or not. One notable result in this direction, due to Kuivinen [35], is that if each $\mathcal{L}_i$ is a complemented modular lattice with rank 2 (a *diamond lattice*), then this problem has a *good characterization.*

Here we restrict ourselves to the case where f has a special representation of our interest. We say that $f : \mathcal{L} \to \mathbf{R}$ has *arity* K if f is represented as

$$f(\rho) := f'(\rho_{i_1}, \rho_{i_2}, \ldots, \rho_{i_K}) \quad (\rho = (\rho_1, \rho_2, \ldots, \rho_n) \in \mathcal{L} = \mathcal{L}_1 \times \mathcal{L}_2 \times \cdots \times \mathcal{L}_n),$$

for some function f' on $\mathcal{L}_{i_1} \times \mathcal{L}_{i_2} \times \ldots \times \mathcal{L}_{i_K}$. The problem we consider is minimizing a sum of submodular functions with arity K, where the input function f is given as a form

$$f = \sum \{ f^{i_1, i_2, \ldots, i_K} \mid 1 \leq i_1 < i_2 < \cdots < i_K \leq n \} \tag{4.13}$$

with some submodular functions $f^{i_1,i_2,\ldots,i_K}$ on $\mathcal{L}_{i_1} \times \mathcal{L}_{i_2} \times \ldots \times \mathcal{L}_{i_K}$, which are regarded as functions on $\mathcal{L}_1 \times \mathcal{L}_2 \times \cdots \times \mathcal{L}_n$ as (4.12). By Lemma 4.12, $f^{i_1,i_2,\ldots,i_K}$ is submodular on $\mathcal{L}$, and hence f is submodular. Such f can be minimized in polynomial time, provided K is fixed.

Theorem 4.13. *Let $\mathcal{L}$ be the product of modular semilattices $\mathcal{L}_1, \mathcal{L}_2, \ldots, \mathcal{L}_n$. A sum of submodular functions with arity K can be minimized in time polynomial in n^K and L^K, where $L = \max_i \{|\mathcal{L}_i|\}$.*

We prove Theorem 4.13 as a (nontrivial) consequence of a surprising result of Thapper and Živný [46] obtained in the context of Valued-CSP. To describe their result, let us consider a more general setting. Let $V = \{1, 2, \ldots, n\}$ and $D_1, D_2, \ldots, D_n$ be finite sets, and put $D = D_1 \times D_2 \times \cdots \times D_n$. As above, we regard a function on $D_{i_1} \times D_{i_2} \times \cdots \times D_{i_K}$ as a function on D. For a class of functions $\mathcal{F}$ on D, consider the following general problem:

$$\text{Given a function } f \in \mathcal{F}, \text{ minimize } f(\rho) \text{ over } \rho \in D. \tag{4.14}$$

Suppose that f is given as an expression

$$f = \sum \{ f^{i_1, i_2, \ldots, i_K} \mid 1 \leq i_1 < i_2 < \cdots < i_K \leq n \}$$

with $f^{i_1,i_2,\ldots,i_K} : D_{i_1} \times D_{i_2} \times \cdots \times D_{i_K} \to \mathbf{R}$ belonging to $\mathcal{F}$, where $f^{i_1,i_2,\ldots,i_K}$ are regarded as functions on D as in (4.12). The set of all K-element subsets of V is denoted by $\binom{V}{K}$. For $\mathcal{U} = \{i_1, i_2, \ldots, i_K\} \in \binom{V}{K}$, we denote $D_{i_1} \times D_{i_2} \times \cdots \times D_{i_K}$ by $D_{\mathcal{U}}$; so $f(\rho) = \sum_{\mathcal{U} \in \binom{V}{K}} f^{\mathcal{U}}(\rho|_{\mathcal{U}})$, where $\rho|_{\mathcal{U}}$ is the restriction of ρ to $\mathcal{U}$.

Consider the following linear program:

$$\begin{aligned} \text{Min.} \quad & \sum_{\mathcal{U} \in \binom{V}{K}} \sum_{p \in D_{\mathcal{U}}} \lambda_p^{\mathcal{U}} f^{\mathcal{U}}(p), \\ \text{s.t.} \quad & \sum_{p \in D_{\mathcal{U}} : p_i = a} \lambda_p^{\mathcal{U}} = \lambda_a^i \quad (i \in V, a \in \mathcal{L}_i, \mathcal{U} \in \tbinom{V}{K} : \mathcal{U} \ni i), \\ & \sum_{a \in \mathcal{L}_i} \lambda_a^i = 1, \quad (i \in V), \\ & \lambda_a^i \geq 0 \quad (i \in V, a \in D_i), \\ & \lambda_p^{\mathcal{U}} \geq 0 \quad (\mathcal{U} \in \tbinom{V}{K}, p \in D_{\mathcal{U}}). \end{aligned} \tag{4.15}$$

Here the input data is $f^{\mathcal{U}}(p)$ $(\mathcal{U} \in \binom{V}{K}, p \in D_{\mathcal{U}})$, and the (decision) variables are $\lambda_p^{\mathcal{U}}$ $(\mathcal{U} \in \binom{V}{K}, p \in D_{\mathcal{U}})$ and λ_a^i $(i \in V, a \in D_i)$. Hence the size of this LP is bounded by polynomial on n^K and L^K, where $L = \max_i |D_i|$. This is a relaxation of (4.14). Indeed, for $\rho \in D$, define λ_a^i as 1 if $a = \rho_i$, and zero otherwise. Define $\lambda_p^{i_1,i_2,\ldots,i_K}$ as 1 if $p = (\rho_{i_1}, \rho_{i_2}, \ldots, \rho_{i_K})$, and zero otherwise. Then $\lambda_p^{\mathcal{U}}$ and λ_a^i are feasible, and the objective function of (4.15) is equal to $f(\rho) = \sum f^{i_1,i_2,\ldots,i_K}(\rho_{i_1}, \rho_{i_2}, \ldots, \rho_{i_K})$. The optimal values of (4.14) and of (4.15) will be denoted by τ_f and τ_f^*, respectively. Obviously $\tau_f^* \leq \tau_f$.

Thapper and Živný [46] gave a powerful criterion for a class $\mathcal{F}$ of functions to have the property that $\tau_f^* = \tau_f$ for all $f \in \mathcal{F}$. A (binary) *operation* on D_i is a function from $D_i \times D_i$ to D_i. A (separable) *operation* g on D is a function from $D \times D$ to D such that for some operations g_i on D_i $(i = 1, 2, \ldots, n)$, $g(p, q) = (g_1(p_1, q_1), g_2(p_2, q_2), \ldots, g_n(p_n, q_n))$ for $p = (p_1, p_2, \ldots, p_n) \in D$ and $q = (q_1, q_2, \ldots, q_n) \in D$. The set of all operations on D is denoted by $\mathcal{O}$. Here $\mathcal{O}$ is a (very large) finite set. Consider a formal sum of operations with real coefficients, which we call a *fractional operation*. A fractional operation is identified with a function $\omega : \mathcal{O} \to \mathbf{R}$, which we represent as $\omega = \sum_{g \in \mathcal{O}} \omega(g) g$. A *fractional polymorphism* ω for $\mathcal{F}$ is a fractional operation on D such that

$$\sum_{g \in \mathcal{O}} \omega(g) = 1, \quad \omega(g) \geq 0 \ (g \in \mathcal{O}),$$
$$\frac{1}{2} f(p) + \frac{1}{2} f(q) \geq \sum_{g \in \mathcal{O}} \omega(g) f(g(p, q)) \quad (f \in \mathcal{F}, (p, q) \in D \times D).$$

The support of ω is the set of operations g with $\omega(g) > 0$.

Theorem 4.14 (Special case of [46, Theorem 5.1])**.** *If there exists a fractional polymorphism ω for $\mathcal{F}$ such that the support of ω includes a semilattice operation, then $\tau_f = \tau_f^*$ for every $f \in \mathcal{F}$.*

Here a *semilattice operation* is an operation g satisfying $g(a, a) = a$, $g(a, b) = g(b, a)$, and $g(g(a, b), c) = g(a, g(b, c))$ for $a, b, c \in D$.

Remark 4.15. In the setting in [46], D_i is the same set $\hat{D}$ for all i. Our problem reduces to this case by taking the disjoint union of D_i as $\hat{D}$, and by setting $f^{i_1,i_2,\ldots,i_K}(\rho) := +\infty$ if $(\rho_{i_1}, \rho_{i_2}, \ldots, \rho_{i_K}) \notin D_{i_1} \times D_{i_2} \times \cdots \times D_{i_K}$. Without such a reduction, their proof also works for our setting in a straightforward way.

Our goal is to prove that the class of submodular functions admits such a nice fractional polymorphism.

Theorem 4.16. *Suppose that each D_i is a modular semilattice and $\mathcal{F}$ is the set of submodular functions on D. Then there exists a fractional polymorphism ω for $\mathcal{F}$ such that the support of ω includes semilattice operation $\wedge$.*

We first derive Theorem 4.13 from Theorems 4.14 and 4.16 and next give the proof of Theorem 4.16.

Proof of Theorem 4.13. Put $D_i = \mathcal{L}_i$ and $D = \mathcal{L}$. Thanks to Theorems 4.14 and 4.16, we can evaluate $\tau_f (= \tau_f^*)$ in time polynomial in n^K and D^K, by solving (4.15). For $i \in V$ and $a \in D_i$, consider (4.14) with an additional constraint $\rho_i = a$. This problem is the minimization of f over a section, which is also the minimization of a sum of submodular functions of arity (at most) K. Again we can evaluate the optimal value $\tau_f^{i,a}$ of this problem. If $\tau_f = \tau_f^{i,a}$, then there exists an optimal solution ρ with $\rho_i = a$. Obviously $\tau_f = \tau_f^{i,a}$ holds for some a. Therefore we can fix ρ_i by evaluating $\tau_f^{i,a}$ for all $a \in D_i$. After n fixing steps, we can get an optimal solution. Thus we can obtain an optimal solution of (4.14) in time polynomial in n^K and L^K.

Constructing a fractional polymorphism: Proof of Theorem 4.16. We are going to construct a fractional polymorphism based on the inequality (4.7) in Theorem 4.2. Since the space of operations is huge, we need a systematic approach. The basic idea is to construct a cone-decomposition $\mathcal{C}$ of $\mathbf{R}^2$ and to associate each cone C in $\mathcal{C}$ with an operation g_C, and to let $\omega = \sum_{C \in \mathcal{C}} \nu(C) g_C$.

We start with the notion of a cone-decomposition. A closed convex cone in $\mathbf{R}^2_+$ is called a 2-*cone* if it is 2-dimensional, i.e., if it has an interior point. For a 2-cone C, a *cone-decomposition* of C is a finite set $\mathcal{C}$ of 2-cones such that every pair of cones in $\mathcal{C}$ has no common interior point, and the union of the 2-cones in $\mathcal{C}$ is equal to C. Recall the definition (4.6) of ν. We easily see the following valuation property.

(4.16) (1) $\nu(C)$ is positive for any 2-cone C.

(2) $\nu(\mathbf{R}^2_+) = 1$.

(3) For a cone-decomposition $\mathcal{C}$ of C, we have $\nu(C) = \sum_{F \in \mathcal{C}} \nu(F)$.

Let $\mathcal{C}$ and $\mathcal{C}'$ be cone-decompositions of the same 2-cone C. We say that $\mathcal{C}$ is a *refinement* of $\mathcal{C}'$ if for each 2-cone F in $\mathcal{C}$ there is a (unique) 2-cone F' in $\mathcal{C}'$ with $F \subseteq F'$. The *common refinement* $\mathcal{C} \wedge \mathcal{C}'$ is defined as

$$\mathcal{C} \wedge \mathcal{C}' := \{F \cap F' \mid F \in \mathcal{C}, F' \in \mathcal{C}', F \cap F' \text{ has an interior point}\}.$$

Clearly $\mathcal{C} \wedge \mathcal{C}'$ is also a cone-decomposition of C.

Then we can associate a modular semilattice with cone-decompositions as follows. Let $\mathcal{L}$ be a modular semilattice. For $(p,q) \in \mathcal{L} \times \mathcal{L}$, define the set of 2-cones $\mathcal{N}^{p,q}(\mathcal{L})$ by

$$\mathcal{N}^{p,q}(\mathcal{L}) := \{[u]^{p,q} \mid u \in \mathcal{E}^{p,q}, [u]^{p,q} \text{ is a 2-cone}\}.$$

Recall the definition of cone $[u]^{p,q}$ in Section 4.2; the ordering of p,q is important here. This is a cone decomposition of $\mathbf{R}^2_+$; this is nothing but the normal cone decomposition of $\mathrm{Conv}(\varphi^{p,q}(I(p,q)))$ (see Figure 7). In the case where $\mathcal{L} = \mathcal{L}_1 \times \mathcal{L}_2 \times \cdots \times \mathcal{L}_n$, by Lemma 4.9 $\mathrm{Conv}(\varphi^{p,q}(I(p,q)))$ is the Minkowski sum of $\mathrm{Conv}(\varphi^{p_i,q_i}(I(p_i,q_i)))$. By a well-known property of the Minkowski sum, we obtain

$$\mathcal{N}^{p,q}(\mathcal{L}) = \bigwedge_{i=1}^{n} \mathcal{N}^{p_i,q_i}(\mathcal{L}_i) \quad ((p,q) \in \mathcal{L} \times \mathcal{L}). \tag{4.17}$$

Next we explain a method to construct a fractional operation. Let $\mathcal{L}$ be a modular semilattice. Let $\mathcal{C}(\mathcal{L})$ be the cone-decomposition in $\mathbf{R}^2_+$ defined by

$$\mathcal{C}(\mathcal{L}) := \bigwedge_{(p,q) \in \mathcal{L} \times \mathcal{L}} \mathcal{N}^{p,q}(\mathcal{L}). \tag{4.18}$$

Suppose $\mathcal{L} = \mathcal{L}_1 \times \mathcal{L}_2 \times \cdots \times \mathcal{L}_n$. By (4.17), we have

$$\mathcal{C}(\mathcal{L}) = \bigwedge_{(p,q) \in \mathcal{L} \times \mathcal{L}} \bigwedge_{i=1}^{n} \mathcal{N}^{p_i,q_i}(\mathcal{L}_i) = \bigwedge_{i=1}^{n} \bigwedge_{(p_i,q_i) \in \mathcal{L}_i \times \mathcal{L}_i} \mathcal{N}^{p_i,q_i}(\mathcal{L}_i) = \bigwedge_{i=1}^{n} \mathcal{C}(\mathcal{L}_i). \tag{4.19}$$

For each 2-cone C in $\mathcal{C}(\mathcal{L})$, we can define an operation $g_C = ((g_C)_1, (g_C)_2, \ldots, (g_C)_n)$ on $\mathcal{L}$ by

$$(g_C)_i(u,v) := w \quad \text{where } [w]^{u,v} \in \mathcal{N}^{u,v}(\mathcal{L}_i) \text{ with } C \subseteq [w]^{u,v} \quad (1 \leq i \leq n; u,v \in \mathcal{L}_i)$$

Here $(g_C)_i$ is well-defined; for each 2-cone C in $\mathcal{C}(\mathcal{L})$ there uniquely exists a 2-cone in $\mathcal{N}^{u,v}(\mathcal{L}_i)$ including C (thanks to (4.19)). Now we arrive at the goal.

Theorem 4.17. *Let $\mathcal{L}$ be the product of modular semilattices $\mathcal{L}_1, \mathcal{L}_2, \ldots, \mathcal{L}_n$. For every submodular function f on $\mathcal{L}$, we have*

$$f(p) + f(q) \geq f(p \wedge q) + \sum_{C \in \mathcal{C}(\mathcal{L})} \nu(C) f(g_C(p,q)) \quad ((p,q) \in \mathcal{L} \times \mathcal{L}).$$

In particular, $(1/2) \wedge + \sum_{C \in \mathcal{C}(\mathcal{L})} (1/2) \nu(C) g_C$ is a fractional polymorphism for the set of submodular functions on $\mathcal{L}$.

Proof. By Theorem 4.2, we have

$$f(p) + f(q) \geq f(p \wedge q) + \sum_{u \in \mathcal{E}^{p,q}} \nu([u]) f(u).$$

By definition of $\mathcal{C}(\mathcal{L})$, $\mathcal{C}(\mathcal{L})$ is a refinement of $\mathcal{N}^{p,q}(\mathcal{L})$. Then each $[u]$ is the union of 2-cones C in $\mathcal{C}(\mathcal{L})$ with $C \subseteq [u]$. By definition of g_C and (4.19), $C \subseteq [u]$ implies $u = g_C(p,q)$. Hence by (4.16) we get

$$\begin{aligned} \sum_{C \in \mathcal{C}(\mathcal{L})} \nu(C) f(g_C(p,q)) &= \sum_{[u] \in \mathcal{N}^{p,q}(\mathcal{L})} \sum_{C \in \mathcal{C}(\mathcal{L}): C \subseteq [u]} \nu(C) f(g_C(p,q)) \\ &= \sum_{u \in \mathcal{E}^{p,q}} f(u) \sum_{C \in \mathcal{C}(\mathcal{L}): C \subseteq [u]} \nu(C) \\ &= \sum_{u \in \mathcal{E}^{p,q}} \nu([u]) f(u). \end{aligned}$$

□

4.5 L-optimality criterion and steepest descent algorithm

Here we consider minimization of L-convex functions on a modular complex (Γ, o, h). There is an optimality criterion that extends the L-optimality criterion in discrete convex analysis; see [39, Theorem 7.14]. Here $\mathcal{L}_p^-$ denotes semilattice $\mathcal{L}_p(\Gamma, o^{-1})$ for the reverse orientation o^{-1} of o. Accordingly, $\mathcal{L}_p^+$ denotes $\mathcal{L}_p(\Gamma, o)$. Recall that h gives positive valuations on $\mathcal{L}_p^+$ and on $\mathcal{L}_p^-$ according to (2.4).

Theorem 4.18 (L-optimality criterion)**.** *Let $g : V_\Gamma \to \mathbf{R}$ be an L-convex function on a modular complex (Γ, o, h). For a vertex $p \in V_\Gamma$, the following conditions are equivalent:*

(1) *$g(p) \leq g(q)$ holds for every $q \in V_\Gamma$.*

(2) *$g(p) \leq g(q)$ holds for every $q \in V_\Gamma$ with $p \sqsubseteq_o q$ or $q \sqsubseteq_o p$. That is*

$$g(p) = \min\{g(q) \mid q \in \mathcal{L}_p^+\} = \min\{g(q) \mid q \in \mathcal{L}_p^-\}.$$

The proof of this theorem is given at the end of this section. The condition (2) implies that g can be minimized by tracing the 1-skeleton graph of the geometric modular complex $\Delta(\Gamma, o)$. Recall Lemma 4.8. The checking of the condition (2) reduces to the submodular function minimization on a modular semilattice, analogous to the case of discrete convex analysis [39, Section 10.3].

Next consider the problem of minimizing an L-convex function g on the product (Γ, o, h) of modular complexes (Γ_i, o_i, h_i) for $i = 1, 2, \ldots, n$. Again we say nothing about the complexity under the oracle model, and hence we impose an arity condition. We say that $g : V_\Gamma \to \mathbf{R}$ has *arity* K if g is represented as $g(\rho) = g'(\rho_{i_1}, \rho_{i_2}, \ldots, \rho_{i_K})$ for some $\{i_1, i_2, \ldots, i_K\}$ and a function $g' : V_{\Gamma_{i_1} \times \Gamma_{i_2} \times \cdots \times \Gamma_{i_K}} \to \mathbf{R}$. So suppose that the input function g is given as an expression

$$g = \sum \{g^{i_1, i_2, \ldots, i_K} \mid 1 \leq i_1 < i_2 < \cdots < i_K \leq n\}$$

for some $0 < K \leq n$ and some L-convex functions $g^{i_1,i_2,\ldots,i_K}$ on the product of (Γ_i, o_i, h_i) over $i \in \{i_1, i_2, \ldots, i_K\}$, where $g^{i_1,i_2,\ldots,i_K}$ is regarded as a function on V_Γ as in (4.12). Then $g^{i_1,i_2,\ldots,i_K}$ is L-convex on (Γ, o, h) (Lemma 4.12 (2)), and so is g. The restriction of g to $\mathcal{L}_p(\Gamma, o) = \mathcal{L}_{p_1}(\Gamma_1, o_1) \times \mathcal{L}_{p_2}(\Gamma_2, o_2) \times \cdots \times \mathcal{L}_{p_n}(\Gamma_n, o_n)$ is a sum of submodular functions with arity K. Hence, by Theorem 4.13, we get the following.

Theorem 4.19. *Let (Γ, o, h) be the product of modular complexes (Γ_i, o_i, h_i) $(i = 1, 2, \ldots, n)$. The optimality of a sum of L-convex functions with arity K can be checked in time polynomial in n^K and N^K, where $N = \max_{i=1,2,\ldots,n}\{|V_{\Gamma_i}|\}$.*

Steepest descent algorithm. Theorem 4.18, Lemma 4.8, and Theorem 4.13 naturally lead us to a descent algorithm for L-convex functions on modular complexes, analogous to the steepest descent algorithm for L-convex function minimization in discrete convex analysis.

Starting from an arbitrary point p, each descent step is to find, for $\sigma \in \{-, +\}$, an optimal solution q^σ of the problem:

(4.20) Minimize $g(q)$ over $q \in \mathcal{L}_p^\sigma$.

As mentioned already, this is a submodular function minimization. If $g(p) = g(q^+) = g(q^-)$, then p is optimal. Otherwise, take $\sigma \in \{-, +\}$ with $g(q^\sigma) = \min\{g(q^-), g(q^+)\}(< g(p))$, let $p := q^\sigma$ (*steepest direction*), and repeat the descent step. After a finite number of descent steps, we can obtain an optimal solution (a minimizer of g).

In the case where f is an L-convex function on a box subset B of $\mathbf{Z}^n$, as already mentioned, the problem (4.20) is a submodular function minimization (in the ordinary sense) and can be solved in polynomial time under the oracle model. Moreover Murota [38] proved that, by appropriate choices of steepest directions, the number of the descent steps is bounded by l_1-diameter of B; later Kolmogorov and Shioura [33] improved this bound. We, however, do not know whether a similar upper bound exists for L-convex function minimizations on the product of general modular complexes. This issue will be studied in the next paper [22].

Proof of L-optimality criterion (Theorem 4.18). Let (Γ, o, h) be a modular complex and let g be a function on V_Γ. Let g^* and g_* denote $\max_p g(p)$ and $\min_p g(p)$, respectively. Let $\bar{\Gamma}$ denote the graph obtained from Γ by joining each Boolean pair (p, q) (with $d_\Gamma(p, q) \geq 2$). Namely $\bar{\Gamma}$ is the 1-skeleton graph of geometric modular complex $\Delta(\Gamma, o)$. For $\alpha \in [g_*, g^*]$, the *level-set subgraph* $\bar{\Gamma}(g, \alpha)$ is the subgraph of $\bar{\Gamma}$ induced by the set of vertices p with $g(p) \leq \alpha$. The following connectivity property of $\bar{\Gamma}(g, \alpha)$ rephrases the L-optimality criterion (Theorem 4.18).

Theorem 4.20. *Suppose that g is an L-convex function on (Γ, o, h). For every $\alpha \in [g_*, g^*]$,*

(1) *$\bar{\Gamma}(g, \alpha)$ is connected, and*

(2) *if $\alpha > g_*$ and $g^{-1}(\alpha) \neq \emptyset$, then each vertex in $g^{-1}(\alpha)$ is adjacent to some vertex of $\bar{\Gamma}(g, \alpha) \setminus g^{-1}(\alpha)$ in $\bar{\Gamma}(g, \alpha)$.*

In a crucial step of the proof, we use the following general property of submodular functions on a modular semilattice.

Lemma 4.21. *Let f be a submodular function on a modular semilattice $\mathcal{L}$. For $p, q \in \mathcal{L}$, if $f(p) \leq 0$ and $f(q) < 0$, there exists a sequence $(p = p_0, p_1, p_2, \ldots, p_m = q)$ such that $f(p_i) < 0$ for $i = 1, 2, \ldots, m-1$, and $p_i \preceq p_{i+1}$ or $p_{i+1} \preceq p_i$ for $i = 0, 1, 2, \ldots, m-1$.*

Proof. We may assume that p and q are incomparable; (p, q)-envelope $\mathcal{E}^{p,q}$ is a polygon. Consider inequality (4.7) in Theorem 4.2:

$$(1 - \nu([p]))f(p) + (1 - \nu([q]))f(q) \geq f(p \wedge q) + \sum_{u \in \mathcal{E}^{p,q} \setminus \{p,q\}} \nu([u])f(u).$$

Since $0 \leq \nu([p]) < 1$ and $0 \leq \nu([q]) < 1$, LHS is negative, and hence RHS is negative. If $f(p \wedge q) < 0$, then $(p, p \wedge q, q)$ is a required sequence. Otherwise there exists $u \in \mathcal{E}^{p,q} \setminus \{p, q\}$ with $f(u) < 0$ (since ν is nonnegative-valued). By an inductive argument (on distance between p and q), there are $(p, p_1, p_2, \ldots, p_k = u)$ and $(u, q_1, q_2, \ldots, q_{k'} = q)$ with $f(p_i) < 0$ and $f(q_j) < 0$. Concatenating them, we obtain a required sequence. □

Proof of Theorem 4.20. (1). Suppose (indirectly) that $\bar{\Gamma}(g, \alpha')$ is disconnected for some α'. Clearly $\bar{\Gamma}(g, g^*)(= \bar{\Gamma})$ is connected. Also, for a sufficiently small $\epsilon > 0$, $\bar{\Gamma}(g, \alpha - \epsilon) = \bar{\Gamma}(g, \alpha) \setminus g^{-1}(\alpha)$. This implies that there exists $\alpha^* \in [g_*, g^*]$ such that $\bar{\Gamma}(g, \alpha^*)$ is connected, and $\bar{\Gamma}(g, \alpha^*) \setminus g^{-1}(\alpha^*)$ is disconnected. Then there exists a pair of vertices p, p' belonging to different components in $\bar{\Gamma}(g, \alpha^*) \setminus g^{-1}(\alpha^*)$; in particular $g(p) < \alpha^*$ and $g(q') < \alpha^*$. Take such a pair (p, p') with $k := d_{\bar{\Gamma}(g,\alpha^*)}(p, p')$ minimum. There exists a path $(p = p_0, p_1, \ldots, p_k = p')$ in $\bar{\Gamma}(g, \alpha^*)$ with $g(p_i) = \alpha^*$ for $i = 1, 2, \ldots, k-1$.

We first show $k = 2$. Consider $\mathcal{L}^*_{p_1}$ and the derivative $\mathrm{d}^*_{p_1} g$. We may assume that $p_1 \sqsubseteq_o p_2$. Let $u := p/p_1$ if $p_1 \sqsubseteq_o p$ and $u := p_1/p$ if $p \sqsubseteq_o p_1$. Then $\mathrm{d}^*_{p_1} g(u) < 0$ and $\mathrm{d}^*_{p_1} g(p_2/p_1) \leq 0$. Therefore, by Lemma 4.21, there exists a comparable sequence $(u = u_0, u_1, \ldots, u_{m-1}, u_m = p_2/p_1)$ in $\mathcal{L}^*_{p_1}$ such that $\mathrm{d}^*_{p_1} g(u_i) < 0$ for $i = 0, 1, 2, \ldots, m-1$. Consider u_{m-1}. Then (i) $p_2/p_1 \sqsubseteq_{o^*} u_{m-1}$ or (ii) $u_{m-1} \sqsubseteq_{o^*} p_2/p_1$. Consider case (i): $p_2/p_1 \sqsubseteq_{o^*} u_{m-1}$. By Lemma 3.10, we have $u_{m-1} = q/q'$ for some $q, q' \in V_\Gamma$ with $q' \sqsubseteq_o q$ and $q' \sqsubseteq_o p_1 \sqsubseteq_o p_2 \sqsubseteq_o q$. By Lemma 3.5, we have $q' \sqsubseteq_o p_i \sqsubseteq_o q$ for $i = 1, 2$. Thus both q and q' are adjacent to each of p_1 and p_2 (in $\bar{\Gamma}$). By $\mathrm{d}^*_{p_1} g(u_{m-1}) < 0$, we have $g(q) < \alpha^*$ or $g(q') < \alpha^*$. Say $g(q) < \alpha^*$; q is adjacent to p_1 and p_2 in $\bar{\Gamma}[g, \alpha^*]$. If q and p' belongs different components in $\bar{\Gamma} \setminus g^{-1}(\alpha^*)$, then path $(q, p_2, p_3, \ldots, p_k = p')$ violates the minimality assumption. This means that q and p' belong to the same component, which is different from the component that p belongs to. Thus we could have chosen path (p, p_1, q) of length 2, This implies that $k = 2$ and $g(p_2) < \alpha^*$. Consider case (ii): $u_{m-1} \sqsubseteq_{o^*} p_2/p_1$. By Lemma 3.10, $u_{m-1} = q/p_1$ for some $q \in V_\Gamma$ with $p_1 \sqsubseteq_o q \sqsubseteq_o p_2$. Also, we have $g(q) < \alpha^*$, and q is adjacent to each of p_1 and p_2. As above, by the minimality, we must have $k = 2$ and $g(p_2) < \alpha^*$.

Suppose that u_i is represented by $u_i = q_i/q'_i$ for $q_i, q'_i \in V_\Gamma$ with $q'_i \sqsubseteq_o q_i$ ($i = 0, 1, 2, \ldots, m$). By Lemmas 3.10 (1) and 3.5, both q_i and q'_i are adjacent to each of q_{i+1} and q'_{i+1} in $\bar{\Gamma}$. Also, by $\mathrm{d}^*_p g(q_i/q'_i) < 0$, at least one of $g(q_i)$ and $g(q'_i)$ is less than α^*. This means that there is a path in $\bar{\Gamma}(g, \alpha^*) \setminus g^{-1}(\alpha^*)$ connecting p and p'. This is a contradiction to the initial assumption that p and p' belong to distinct components in $\bar{\Gamma}(g, \alpha^*) \setminus g^{-1}(\alpha^*)$.

(2). Take $p \in g^{-1}(\alpha)$. By (1), there is a pair of $q \in V_\Gamma$ and a path $(q = p_0, p_1, p_2, \ldots, p_k = p)$ in $\bar{\Gamma}$ such that $g(q) < \alpha$ and $g(p_i) = \alpha$ for $i = 1, 2, \ldots, k$. Take such a pair with the minimum length k. We show $k = 1$. Suppose that $k \geq 2$. As above, by considering $\mathrm{d}^*_{p_1} g$ on $\mathcal{L}^*_{p_1}$, we can find a neighbor q' of p_2 with $f(q') < \alpha$. This is a contradiction to the minimality of k. Hence $k = 1$, as required. □

5 Minimum 0-extension problems

In this section, we study, from the viewpoint developed in the previous sections, the minimum 0-extension problem **0-Ext**$[\Gamma]$ on an orientable modular graph Γ. In Section 5.1, we show that **0-Ext**$[\Gamma]$ can be formulated as an L-convex function minimization on a modular complex. In Section 5.2, we present a powerful optimality criterion (Theorem 5.4) for **0-Ext**$[\Gamma]$ by specializing the L-optimality criterion (Theorem 4.18). In Section 5.3 we prove the main theorem (Theorem 1.6) of this paper. In Section 5.4 we consider the minimum 0-extension problem for metrics, not necessarily graph metrics, and extend Theorem 1.6 to metrics.

5.1 L-convexity of multifacility location functions

We introduce a weighted version of **0-Ext**$[\Gamma]$. Let Γ be an orientable modular graph with an orbit-invariant function h. Fix an admissible orientation o of Γ. We are given a finite set V with $V_\Gamma \subseteq V$ and a nonnegative cost function $c : \binom{V}{2} \to \mathbf{Q}_+$. A *location* is a map $\rho : V \to V_\Gamma$ satisfying $\rho(s) = s$ for all $s \in V_\Gamma$. The *cost* of location ρ is defined as $(c \cdot d_{\Gamma,h})(\rho) := \sum_{xy \in \binom{V}{2}} c(xy) d_{\Gamma,h}(\rho(x), \rho(y))$. The weighted version of **0-Ext**$[\Gamma]$ is as follows:

Multifac$[\Gamma, h; V, c]$**:** Minimize $(c \cdot d_{\Gamma,h})(\rho)$ over all locations ρ,

where the unweighted version corresponds to $h = 1$.

We next reformulate **Multifac**$[\Gamma, h; V, c]$ as a discrete optimization problem on a modular complex. Let $X := V \setminus V_\Gamma$; a location is identified with a map $X \to V_\Gamma$. For $x \in X$, let (Γ_x, o_x, h_x) be a copy of the modular complex (Γ, o, h). Let (Γ_X, o_X, h_X) be the product of (Γ_x, o_x, h_x) over $x \in X$. Then any location ρ is identified with a point $\rho = (\rho(x))_{x \in X}$ in V_{Γ_X}, and hence $(c \cdot d_{\Gamma,h})$ is a function on V_{Γ_X}. Accordingly **Multifac**$[\Gamma, h; V, c]$ is reformulated as follows:

$$\text{Minimize } (c \cdot d_{\Gamma,h})(\rho) \text{ over all } \rho \in V_{\Gamma_X}.$$

The following theorem says that this problem is an L-convex function minimization.

Theorem 5.1. *$(c \cdot d_{\Gamma,h})$ is an L-convex function on modular complex (Γ_X, o_X, h_X).*

Since $(c \cdot d_{\Gamma,h})$ is a nonnegative combination of metric functions $d_{\Gamma,h} : V_\Gamma \times V_\Gamma \to \mathbf{R}$ (regarded as $V_{\Gamma_X} \to \mathbf{R}$) and the set of L-convex functions is closed under nonnegative combinations and restrictions/extensions (Lemmas 4.10 and 4.12). Theorem 5.1 follows from the following.

Theorem 5.2. *Metric function $d_{\Gamma,h}$ is L-convex on $(\Gamma \times \Gamma, o \times o, h \times h)$.*

Proof. Consider the 2-subdivision $(\Gamma \times \Gamma)^2$, which is isomorphic to $\Gamma^2 \times \Gamma^2$ by correspondence $(q, q')/(p, p') \leftrightarrow (q/p, q'/p')$ (Lemma 3.13). Consider $\overline{d_{\Gamma,h}} : V_{(\Gamma \times \Gamma)^2} \to \mathbf{R}$. Then we have

$$\overline{d_{\Gamma,h}}((q, q')/(p, p')) = \frac{d_{\Gamma,h}(p, p') + d_{\Gamma,h}(q, q')}{2} = d_{\Gamma^2,h/2}(q/p, q'/p'),$$

where the first equality is the definition (4.2) and the second follows from Proposition 3.8. Hence it suffices to show that $d_{\Gamma^2,h/2} : V_{\Gamma^2} \times V_{\Gamma^2} \to \mathbf{R}$ is submodular on $\mathcal{L}_{(a/a,b/b)}$ for every $(a, b) \in V_\Gamma \times V_\Gamma$. Therefore, by taking $(\Gamma^2, h/2)$ as (Γ, h), this follows from the next lemma. □

Lemma 5.3. *Metric function $d_{\Gamma,h}$ is submodular on $\mathcal{L}_a \times \mathcal{L}_b$ for every $a, b \in V_\Gamma$.*

Proof. By Proposition 4.11, it suffices to show the following, where we denote $d_{\Gamma,h}$ by d, and denote the valuation on $\mathcal{L}_a$ (defined in (3.2)) by v.

(1) For every $u \in \mathcal{L}_b$ and every antipodal pair (p, q) in $\mathcal{L}_a$, we have

$$v[p \wedge q, q]d(p, u) + v[p \wedge q, p]d(q, u) \geq (v[p \wedge q, p] + v[p \wedge q, q])d(p \wedge q, u).$$

(2) For every $u \in \mathcal{L}_b$ and every 2-bounded pair (p, q) in $\mathcal{L}_a$, we have

$$d(p, u) + d(q, u) \geq d(p \wedge q, u) + d(p \vee q, u).$$

(2′) For every $p, q \in \mathcal{L}_a$ with $p \swarrow q$ and every $p', q' \in \mathcal{L}_b$ with $p' \swarrow q'$, we have

$$d(q, p') + d(p, q') \geq d(q, q') + d(p, p').$$

Note that (2) and (2′) correspond to the submodularity condition for 2-bounded pairs.

(1). We may assume that $p \wedge q = a$ (by considering $\mathcal{L}_{p\wedge q}$). Take a median m of p, q, u. Then $m \in I(p, q)$. By Lemma 2.8, there are $p' \in [a, p]$ and $q' \in [a, q]$ with $m = p' \vee q'$. Let $D := d(m, u)$. Then we have $d(p, u) = v[p', p] + v(q') + D$, $d(q, u) = v[q', q] + v(p') + D$, and $d(a, u) = v(p') + v(q') + D$. Hence we get

$$\begin{aligned}
&v(q)d(p, u) + v(p)d(q, u) - (v(p) + v(q))d(a, u) \\
&= \{v(q') + v[q', q]\}\{v[p', p] + v(q') + D\} + \{v(p') + v[p', p]\}\{v[q', q] + v(p') + D\} \\
&\quad - \{v(p') + v(q') + v[p', p] + v[q', q]\}\{v(p') + v(q') + D\} \\
&= 2v[q', q]v[p', p] - 2v(p')v(q').
\end{aligned}$$

This must be nonnegative since (p, q) is antipodal.

(2). Recall the notion of gated sets (Section 2.2); $[p \wedge q, p \vee q]$ is convex, and is gated (Lemmas 2.4 and 2.13). Let $m := \mathrm{Pr}_{[p\wedge q,p\vee q]}(u)$, and $D := d(m, u)$. Then we have $d(x, u) = d(x, m) + D$ for $x \in \{p, q, p\wedge q, p\vee q\}$. There are three cases: (i) $m \in \{p, q\}$, (ii) $m \in \{p\wedge q, p\vee q\}$, and (iii) $m \notin \{p, q, p\wedge q, p\vee q\}$. Note that $(p, p\wedge q, q, p\vee q)$ forms a 4-cycle since (p, q) is 2-bounded. Let $\alpha := v[p, p\vee q] = v[p\wedge q, q]$ and $\beta := v[p\wedge q, p] = v[q, p\vee q]$. Consider the case (i). Then $\{d(p \wedge q, m), d(p \vee q, m)\} = \{\alpha, \beta\}$ and $\{d(p, m), d(q, m)\} = \{0, \alpha + \beta\}$. Hence $d(p, u) + d(q, u) - d(p \wedge q, u) - d(p \vee q, u) = 0$. Consider the case (ii). Then $\{d(p, m), d(q, m)\} = \{\alpha, \beta\}$, and $\{d(p \wedge q, m), d(p \vee q, m)\} = \{0, \alpha + \beta\}$. Hence $d(p, u) + d(q, u) - d(p \wedge q, u) - d(p \vee q, u) = 0$. Consider the case (iii). Then m is a common neighbor of $p\wedge q, p\vee q$ different from p, q. Hence all edges in $[p\wedge q, p\vee q]$ belong to the same orbit. Thus $\alpha = \beta$, $d(p, m) = d(q, m) = 2\alpha$, $d(p \wedge q, m) = d(p \vee q, m) = \alpha$, and $d(p, u) + d(q, u) - d(p \wedge q, u) - d(p \vee q, u) = 2\alpha > 0$.

(2′). Consider $\mathrm{Pr}_{\{p',q'\}}(\{p, q\})$. Let $D := d(\{p, q\}, \{p', q'\})$. There are two cases: (i) $|\mathrm{Pr}_{\{p',q'\}}(\{p, q\})| = 1$ and (ii) $\{p', q'\} = \mathrm{Pr}_{\{p',q'\}}(\{p, q\})$. Consider the case (i). For u, v, u', v' with $\{u, v\} = \{p, q\}$ and $\{u', v'\} = \{p', q'\}$, we have $d(v, u') = D$, $d(u, u') = D + h(uv)$, $d(v, v') = D + h(u'v')$, and $d(u, v') = D + h(uv) + h(u'v')$. Thus $d(u, u') + d(v, v') = d(u, v') + d(v, u')$, and the equality holds in (2′). Consider the case (ii). By Theorem 2.3 and Lemma 2.10, we have $p' = \mathrm{Pr}_{\{p',q'\}}(p)$, $q' = \mathrm{Pr}_{\{p',q'\}}(q)$, $d(p, p') = d(q, q') = D$, and that pq and $p'q'$ must belong to the same orbit Q. Then $d(p, q') = d(q, p') = D + h_Q$. Therefore (2′) holds. □

5.2 Optimality criterion and orbit-additivity

Let ρ be a location. A location ρ' is said to be a *forward neighbor* of ρ if $\rho(x) \sqsubseteq_o \rho'(x)$ for all $x \in X$, and is said to be a *backward neighbor* of ρ if $\rho'(x) \sqsubseteq_o \rho(x)$ for all $x \in X$. A forward or backward neighbor is simply called a *neighbor*. This terminology is due to [19]. Regard ρ as a vertex in Γ_X. By the definition, the set of forward (resp. backward) neighbors of ρ is the product of $\mathcal{L}^+_{\rho(x)}(\Gamma_x, o_x)$ (resp. $\mathcal{L}^-_{\rho(x)}(\Gamma_x, o_x)$) over $x \in X$, which coincides with $\mathcal{L}^+_\rho = \mathcal{L}^+_\rho(\Gamma_X, o_X)$ (resp. $\mathcal{L}^-_\rho = \mathcal{L}^-_\rho(\Gamma_X, o_X)$) by Lemma 3.13.

We need a sharper neighbor concept. Let Q be an orbit. A location ρ' is called a *forward Q-neighbor* of ρ if, for all $x \in X$, $\rho(x) \sqsubseteq_o \rho'(x)$ and every ascending path from $\rho(x)$ to $\rho'(x)$ belongs to Q. Analogously, a location ρ' is called a *backward Q-neighbor* of ρ if, for all $x \in X$, $\rho'(x) \sqsubseteq_o \rho(x)$ and every ascending path from $\rho'(x)$ to $\rho(x)$ belongs to Q. A forward or backward Q-neighbor is simply called a *Q-neighbor*. For an orbit Q in Γ, the set of forward (resp. backward) Q-neighbors of ρ is denoted by $\mathcal{L}^+_{\rho,Q}$ (resp. $\mathcal{L}^-_{\rho,Q}$).

The main result in this section is the following optimality criterion, which has been shown for some special cases of orientable modular graphs: trees by Kolen [31, Chapter 3], median graphs by Chepoi [12, p.11–12], and frames by Hirai [19, Section 4.1].

Theorem 5.4. *Let Γ be an orientable modular graph with an admissible orientation o and a positive orbit-invariant function h. For a location ρ the following conditions are equivalent:*

(1) *ρ is optimal to* **Multifac**$[\Gamma, h; V, c]$.

(2) *ρ is optimal to* **Multifac**$[\Gamma, 1; V, c]$.

(3) *For every neighbor ρ' of ρ, we have $(c \cdot d_{\Gamma,h})(\rho) \leq (c \cdot d_{\Gamma,h})(\rho')$. That is*

$$(c \cdot d_{\Gamma,h})(\rho) = \min\{(c \cdot d_{\Gamma,h})(\rho') \mid \rho' \in \mathcal{L}^+_\rho\} = \min\{(c \cdot d_{\Gamma,h})(\rho') \mid \rho' \in \mathcal{L}^-_\rho\}.$$

(4) *For every neighbor ρ' of ρ, we have $(c \cdot d_{\Gamma,1})(\rho) \leq (c \cdot d_{\Gamma,1})(\rho')$. That is*

$$(c \cdot d_{\Gamma,1})(\rho) = \min\{(c \cdot d_{\Gamma,1})(\rho') \mid \rho' \in \mathcal{L}^+_\rho\} = \min\{(c \cdot d_{\Gamma,1})(\rho') \mid \rho' \in \mathcal{L}^-_\rho\}.$$

(5) *For every orbit Q and every Q-neighbor ρ' of ρ, we have $(c \cdot d_{\Gamma,1})(\rho) \leq (c \cdot d_{\Gamma,1})(\rho')$. That is, for every orbit Q, we have*

$$(c \cdot d_{\Gamma,1})(\rho) = \min\{(c \cdot d_{\Gamma,1})(\rho') \mid \rho' \in \mathcal{L}^+_{\rho,Q}\} = \min\{(c \cdot d_{\Gamma,1})(\rho') \mid \rho' \in \mathcal{L}^-_{\rho,Q}\}.$$

Before the proof, we explain consequences of Theorem 5.4. The first consequence is that in solving **Multifac**$[\Gamma, h; V, c]$, we may replace h with the unit function, even when h is nonnegative.

Theorem 5.5. *For every nonnegative orbit-invariant function h, every optimal location in* **Multifac**$[\Gamma, 1; V, c]$ *is optimal to* **Multifac**$[\Gamma, h; V, c]$

Proof. Let ρ be an optimal location for **Multifac**$[\Gamma, 1; V, c]$. Take an arbitrary positive $\epsilon > 0$. Consider the positive orbit invariant function $h + \epsilon 1$. By Theorem 5.4, ρ is optimal to **Multifac**$[\Gamma, h + \epsilon 1; V, c]$ Hence, for an arbitrary location ρ', we have

$$(c \cdot d_{\Gamma,h})(\rho') + \epsilon(c \cdot d_{\Gamma,1})(\rho') = (c \cdot d_{\Gamma,h+\epsilon 1})(\rho') \geq (c \cdot d_{\Gamma,h+\epsilon 1})(\rho) = (c \cdot d_{\Gamma,h})(\rho) + \epsilon(c \cdot d_{\Gamma,1})(\rho).$$

Since $\epsilon > 0$ was arbitrary, we have $(c \cdot d_{\Gamma,h})(\rho') \geq (c \cdot d_{\Gamma,h})(\rho)$. □

The second consequence is the decomposition property of **Multifac**$[\Gamma, h; V, c]$. For an orbit Q, consider the following problem on Γ/Q:

(5.1) Minimize $(c \cdot d_{\Gamma/Q,h})(\rho)$ over all $\rho : V \to V_{\Gamma/Q}$ with $\rho(s) = s/Q$ for $s \in V_\Gamma$.

The optimal value of (5.1) is denoted by $\tau_Q(\Gamma, h; V, c)$, whereas the optimal value of the original problem **Multifac**$[\Gamma, h; V, c]$ is denoted by $\tau(\Gamma, h; V, c)$. Then we have

$$\tau(\Gamma, h; V, c) \geq \sum_{Q:\text{orbit}} h_Q \tau_Q(\Gamma, 1; V, c). \tag{5.2}$$

Indeed, for any optimal location ρ in **Multifac**$[\Gamma, h; V, c]$, ρ/Q is feasible to (5.1), and by (2.2) we have

$$\tau(\Gamma, h; V, c) = (c \cdot d_{\Gamma,h})(\rho) = \sum_{Q:\text{orbit}} h_Q (c \cdot d_{\Gamma,1})(\rho/Q) \geq \sum_{Q:\text{orbit}} h_Q \tau_Q(\Gamma, 1; V, c). \tag{5.3}$$

Note that problems **Multifac**$[\Gamma, h; V, c]$ and (5.1) can be considered for a possibly nonorientable modular graph, and the inequality relation (5.2) still holds; see [28]. A modular graph Γ is said to be *orbit-additive* if (5.2) holds in equality. Karzanov [28, Section 6] conjectured that every orientable modular graph is orbit-additive. We can solve this conjecture affirmatively.

Theorem 5.6. *Every orientable modular graph is orbit-additive.*

Proof. Take an optimal solution ρ in **Multifac**$[\Gamma, 1; V, c]$. By Theorem 5.5, ρ is also optimal to **Multifac**$[\Gamma, 1_Q; V, c]$ for every orbit Q, where 1_Q is the orbit-invariant function taking 1 on Q and 0 on $E_\Gamma \setminus Q$. Here **Multifac**$[\Gamma, 1_Q; V, c]$ is equivalent to (5.1). Hence the inequality in (5.3) holds in equality. □

Remark 5.7. According to Theorem 5.1 and Lemma 4.8, the condition (3) (or (4)) in Theorem 5.4 can be checked by submodular function minimizations on the modular semilattices formed by forward and backward neighbors. Furthermore, one can see that the condition (5) can also be checked submodular function minimizations on the modular semilattices formed by forward and backward Q-neighbors for each Q. Sometimes checking (5) is easier than checking (3) (or (4)). For example, consider the case where Γ is a median graph. Then one can see that (5) can be checked by minimum-cut computations.

Remark 5.8. If problem (5.1) is solvable in (strongly) polynomial time for each orbit, then by Theorem 5.6 we can evaluate τ in (strongly) polynomial time, and hence **0-Ext**$[\Gamma]$ is solvable in (strongly) polynomial time by the fixing technique as in the proof of Theorem 4.13. As was suggested by Karzanov [28], this approach is applicable to the case where each orbit graph of Γ is a frame. Then (5.1) is a 0-extension problem on a frame, is solvable in strongly polynomial time, and hence **0-Ext**$[\Gamma]$ is solvable in strongly polynomial time; the (strong polytime) tractability of this class of orientable modular graphs was conjectured by [28]. It should be noted that our proof of the main theorem gives only a *weakly* polynomial time algorithm.

Proof of Theorem 5.4. (1) ⇔ (3) and (2) ⇔ (4) follow from Theorems 4.18 and 5.1. (4) ⇒ (5) is obvious.

(3) $\Rightarrow$ (5). Suppose that ρ' is a Q-neighbor of ρ. Then $\rho'(x)/R = \rho(x)/R$ for orbit R different from Q, and $d_{\Gamma/Q,h} = h_Q d_{\Gamma/Q,1}$. By (2.2) we have

$$\begin{aligned} 0 &\leq (c \cdot d_{\Gamma,h})(\rho') - (c \cdot d_{\Gamma,h})(\rho) = \sum_{R:\text{orbit}} (c \cdot d_{\Gamma/R,h})(\rho'/R) - (c \cdot d_{\Gamma/R,h})(\rho/R) \\ &= (c \cdot d_{\Gamma/Q,h})(\rho'/Q) - (c \cdot d_{\Gamma/Q,h})(\rho/Q) \\ &= h_Q \left((c \cdot d_{\Gamma/Q,1})(\rho'/Q) - (c \cdot d_{\Gamma/Q,1})(\rho/Q)\right) \\ &= h_Q \sum_{R:\text{orbit}} (c \cdot d_{\Gamma/R,1})(\rho'/R) - (c \cdot d_{\Gamma/R,1})(\rho/R) \\ &= h_Q \left((c \cdot d_{\Gamma,1})(\rho') - (c \cdot d_{\Gamma,1})(\rho)\right). \end{aligned}$$

(5) $\Rightarrow$ (3) and (5) $\Rightarrow$ (4). Let ρ' be a forward neighbor of ρ. For each orbit Q, let $Q^+_{\rho(x)}$ be the set of edges $e \in Q$ with both ends belonging to $\mathcal{L}^+_{\rho(x)}$. Then $Q^+_{\rho(x)}$ is an orbit-union in $\mathcal{L}^+_{\rho(x)}$. Recall notions in Section 2.4. We can define a forward Q-neighbor $\rho'|Q$ of ρ by

$$(\rho'|Q)(x) := \rho'(x)|Q^+_{\rho(x)} \quad (x \in X).$$

In the case where ρ' is a backward neighbor of ρ, we define a backward Q-neighbor $\rho'|Q$ analogously, by considering $\mathcal{L}^-_{\rho(x)}$. Then (5) $\Rightarrow$ (3) and (5) $\Rightarrow$ (4) follow from the following decomposition property.

Lemma 5.9. *For a location ρ and a neighbor ρ' of ρ, we have*

$$\begin{aligned} d_{\Gamma,h} \circ \rho' - d_{\Gamma,h} \circ \rho &= \sum_{Q:\text{orbit}} h_Q \left(d_{\Gamma/Q,1} \circ (\rho'/Q) - d_{\Gamma/Q,1} \circ (\rho/Q)\right) \\ &= \sum_{Q:\text{orbit}} h_Q \left(d_{\Gamma,1} \circ (\rho'|Q) - d_{\Gamma,1} \circ \rho\right). \end{aligned} \tag{5.4}$$

Proof. The first equality follows from (2.2). For two orbits Q, R, we observe from the definitions that

$$(\rho'|Q)/R = \begin{cases} \rho'/Q & \text{if } Q = R, \\ \rho/R & \text{otherwise.} \end{cases}$$

Indeed, if $Q \neq R$, then $\rho'(x)|Q$ and $\rho(x)$ are joined by $E_\Gamma \setminus R$, and hence $(\rho'(x)|Q)/R = \rho(x)/R$. If $Q = R$, then $\rho'(x)$ and $\rho'(x)|Q$ are joined by $E_\Gamma \setminus Q$, and hence $\rho'(x)/Q = (\rho(x)'|Q)/Q$. Therefore we have

$$\begin{aligned} d_{\Gamma,1} \circ (\rho'|Q) - d_{\Gamma,1} \circ \rho &= \sum_{R:\text{orbit}} d_{\Gamma/R,1} \circ (\rho'|Q/R) - d_{\Gamma/R,1} \circ (\rho/R) \\ &= d_{\Gamma/Q,1} \circ (\rho'/Q) - d_{\Gamma/Q,1} \circ (\rho/Q). \end{aligned}$$

This formula implies the second equality of (5.4). □

5.3 Proof of the main theorem (Theorem 1.6)

In this section, we complete the proof of the main theorem (Theorem 1.6) stating that **0-Ext**$[\Gamma]$ for every orientable modular graph Γ can be solved in polynomial time. Now we know that **0-Ext**$[\Gamma]$ is a problem of minimizing the sum of L-convex functions of arity 2. Hence, for every cost $c : \binom{V}{2} \to \mathbf{Q}_+$, every location ρ, and sign $\sigma \in \{-,+\}$, $(c \cdot d_{\Gamma,1})$ is the sum of arity-2 submodular functions on $\mathcal{L}^\sigma_\rho$. By Theorem 4.13, we can minimize $(c \cdot d_{\Gamma,1})$ over $\mathcal{L}^\sigma_\rho$ in polynomial time. Therefore we can assume that we have a *descent oracle*, an oracle that returns an optimal solution of this (local) problem.

By the steepest descent algorithm, we can obtain a global optimal solution. As mentioned already, we do not know whether the number of descent steps is polynomially bounded. Fortunately, in the case of multifacility location functions, a cost-scaling approach gives a *weakly* polynomial bound on the number of descent steps. Now the main theorem (Theorem 1.6) follows from the following.

Proposition 5.10. *Suppose that c is integer-valued.* $\mathbf{Multifac}[\Gamma, 1; V, c]$ *can be solved with $O(|V|^2 \operatorname{diam} \Gamma \log C)$ calls of the descent oracle, where $C := \max\{c(xy) \mid xy \in \binom{V}{2}\}$ and $\operatorname{diam} \Gamma$ denotes the diameter of Γ.*

Proof. Let $\lfloor c/2 \rfloor : \binom{V}{2} \to \mathbf{Z}_+$ be defined by $\lfloor c/2 \rfloor(xy) := \lfloor c(xy)/2 \rfloor$. We show:

(5.5) For an optimal location ρ in $\mathbf{Multifac}[\Gamma, 1; V, \lfloor c/2 \rfloor]$, we have

$$(c \cdot d_{\Gamma,1})(\rho) - \tau(\Gamma, 1; V, c) \leq |V|^2 \operatorname{diam} \Gamma.$$

If this is true, then the number of the descent steps from an initial starting point ρ is bounded by $|V|^2 \operatorname{diam} \Gamma$. Consequently, by recursive scaling, we obtain an optimal solution for $\mathbf{Multifac}[\Gamma, 1; V, c]$ in $O(|V|^2 \operatorname{diam} \Gamma \log C)$ descent steps.

To show (5.5), take an optimal location ρ in $\mathbf{Multifac}[\Gamma, 1; V, \lfloor c/2 \rfloor]$, and let $\epsilon := 2\lfloor c/2 \rfloor - c$. Then $\epsilon(xy) \in \{0, -1\}$. Take an optimal location ρ^* in $\mathbf{Multifac}[\Gamma, 1; V, c]$. Then $\tau(\Gamma, 1; V, c) = (c \cdot d_{\Gamma,1})(\rho^*)$. Thus we have

$$\begin{aligned} (c \cdot d_{\Gamma,1})(\rho) - (c \cdot d_{\Gamma,1})(\rho^*) \ &= \ ((2\lfloor c/2 \rfloor \cdot d_{\Gamma,1})(\rho) - (2\lfloor c/2 \rfloor \cdot d_{\Gamma,1})(\rho^*)) \\ &\quad + \sum_{xy} \epsilon(xy) \left(d_{\Gamma,1}(\rho^*(x), \rho^*(y)) - d_{\Gamma,1}(\rho(x), \rho(y)) \right). \end{aligned}$$

Here the first term on the right hand side is at most zero since ρ is an optimal location in $\mathbf{Multifac}[\Gamma, 1; V, 2\lfloor c/2 \rfloor]$ and the second term is at most $|V|^2 \operatorname{diam} \Gamma$. □

5.4 Minimum 0-extension problems for metrics

Let μ be a metric on a finite set S (not necessarily a graph metric). We can naturally consider the minimum 0-extension problem $\mathbf{0\text{-}Ext}[\mu]$ for a general μ formulated as: *Given a set $V \supseteq S$ and $c : \binom{V}{2} \to \mathbf{Q}_+$, find a 0-extension (V, d) of (S, μ) with $\sum_{xy} c(xy)d(x, y)$ minimum.* Metric μ is said to be *modular* if (S, μ) is a modular metric space (see Section 2). Let H_μ be the graph on the vertex set S with edge set E_{H_μ} given as: $xy \in E_{H_\mu} \Leftrightarrow$ there is no $z \in S \setminus \{x, y\}$ with $\mu(x, z) + \mu(z, y) = \mu(x, y)$. H_μ is called the *support graph* of μ. Karzanov [29] extended the hardness result (Theorem 1.5) to the following.

Theorem 5.11 ([29])**.** *If μ is not modular or H_μ is not orientable, then* $\mathbf{0\text{-}Ext}[\mu]$ *is NP-hard.*

We can also consider LP-relaxation $\mathbf{Ext}[\mu]$ obtained by relaxing 0-extensions into extensions in $\mathbf{0\text{-}Ext}[\mu]$. Extending Theorem 1.3, Bandelt, Chepoi, and Karzanov [4] proved that $\mathbf{Ext}[\mu]$ is exact if and only if μ is modular and H_μ is frame.

Our framework covers $\mathbf{0\text{-}Ext}[\mu]$ for a metric μ such that μ is modular and H_μ is orientable. Indeed μ induces the edge-length $\bar{\mu}$ on H_μ by $\bar{\mu}(pq) = \mu(p, q)$ $(pq \in E_{H_\mu})$. From the definition of the support graph H_μ, we have $\mu = d_{H_\mu, \bar{\mu}}$. Moreover, it was shown in [2] (see [28, Section 2]) that

(5.6) if μ is modular, then H_μ is a modular graph, and $\bar{\mu}$ is orbit-invariant.

Hence we can apply the argument in Section 5 to **Multifac**$[H_\mu, \bar{\mu}; V, c]$ to obtain results for **0-Ext**$[\mu]$. By Theorems 1.6 and 5.4, we obtain the converse of Theorem 5.11, which completes the classification of those metrics for which **0-Ext**$[\mu]$ is tractable.

Theorem 5.12. *If μ is modular and H_μ is orientable, then* **0-Ext**$[\mu]$ *is solvable in polynomial time.*

Acknowledgments

We thank Kazuo Murota for careful reading and numerous helpful comments, Kei Kimura for discussion on Valued-CSP, Satoru Iwata for communicating the paper [35] of Kuivinen, Akiyoshi Shioura for the paper [32] of Kolmogorov, and Satoru Fujishige for the paper [23] of Huber-Kolmogorov. This research is partially supported by the Aihara Project, the FIRST program from JSPS, by Global COE Program "The research and training center for new development in mathematics" from MEXT, and by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

[1] H.-J. Bandelt, Networks with Condorcet solutions, *European Journal of Operational Research* **20** (1985), 314–326.

[2] H.-J. Bandelt, Hereditary modular graphs, *Combinatorica* **8** (1988), 149–157.

[3] H.-J. Bandelt and V. Chepoi, Metric graph theory and geometry: a survey, in: J.E. Goodman, J. Pach, and R. Pollack eds., *Surveys on discrete and computational geometry: Twenty Years Later*, 49–86, American Mathematical Society, Providence, 2008.

[4] H.-J. Bandelt, V. Chepoi, and A. V. Karzanov, A characterization of minimizable metrics in the multifacility location problem, *European Journal of Combinatorics* **21** (2000), 715–725.

[5] H.-J. Bandelt, M. van de Vel, and E. Verheul, Modular interval spaces, *Mathematische Nachrichten* **163** (1993) 177–201.

[6] G. Birkhoff, *Lattice Theory*, American Mathematical Society, New York, 1940; 3rd edn., American Mathematical Society, Providence, RI, 1967.

[7] S. Bistarelli, U. Montanari, and F. Rossi, Semiring-based constraint satisfaction and optimization, *Journal of the ACM* **44** (1997), 201–236.

[8] A. Bouchet, Multimatroids. I. Coverings by independent sets, *SIAM Journal on Discrete Mathematics* **10** (1997), 626–646.

[9] M. R. Bridson and A. Haefliger, *Metric Spaces of Non-Positive Curvature*, Springer-Verlag, Berlin, 1999.

[10] R. Chandrasekaran and S. N. Kabadi, Pseudomatroids, *Discrete Mathematics* **71** (1988), 205–217.

[11] V. Chepoi, Classification of graphs by means of metric triangles, *Metody Diskretnogo Analiza* **49** (1989), 75–93 (in Russian).

[12] V. Chepoi, A multifacility location problem on median spaces, *Discrete Applied Mathematics* **64** (1996) 1–29.

[13] V. Chepoi, Graphs of some CAT(0) complexes, *Advances in Applied Mathematics* **24** (2000), 125–179.

[14] E. Dahlhaus, D. S. Johnson, C. H. Papadimitriou, P. D. Seymour, and M. Yannakakis, The complexity of multiterminal cuts, *SIAM Journal on Computing* **23** (1994), 864–894.

[15] A. W. M. Dress and R. Scharlau, Gated sets in metric spaces, *Aequationes Mathematicae* **34** (1987), 112–120.

[16] S. Fujishige, *Submodular Functions and Optimization, 2nd Edition*, Elsevier, Amsterdam, 2005.

[17] M. Grötschel, L. Lovász, and A. Schrijver, *Geometric Algorithms and Combinatorial Optimization*, Springer-Verlag, Berlin, 1988.

[18] H. Hirai, Tight spans of distances and the dual fractionality of undirected multiflow problems, *Journal of Combinatorial Theory, Series B* **99** (2009), 843–868.

[19] H. Hirai, Folder complexes and multiflow combinatorial dualities, *SIAM Journal on Discrete Mathematics* **25** (2011), 1119–1143.

[20] H. Hirai, The maximum multiflow problems with bounded fractionality, RIMS-preprint (2009).

[21] H. Hirai, Half-integrality of node-capacitated multiflows and tree-shaped facility locations on trees, *Mathematical Programming, Series A*, to appear.

[22] H. Hirai, L-convex functions on modular complexes, in preparation.

[23] A. Huber and V. Kolmogorov, Towards minimizing k-submodular functions, in: *Proceedings of the 2nd International Symposium on Combinatorial Optimization (ISCO'12)*, LNCS 7422, Springer, Berlin, 2012, pp. 451–462.

[24] S. Iwata, L. Fleischer, and S. Fujishige, A combinatorial strongly polynomial algorithm for minimizing submodular functions, *Journal of the ACM* **48** (2001), 761–777.

[25] A. V. Karzanov, Polyhedra related to undirected multicommodity flows, *Linear Algebra and its Applications* **114/115** (1989), 293–328.

[26] A. V. Karzanov, Minimum 0-extensions of graph metrics, *European Journal of Combinatorics* **19** (1998), 71–101.

[27] A. V. Karzanov, Metrics with finite sets of primitive extensions, *Annals of Combinatorics* **2** (1998), 211–241.

[28] A. V. Karzanov, One more well-solved case of the multifacility location problem, *Discrete Optimization* **1** (2004), 51–66.

[29] A. V. Karzanov, Hard cases of the multifacility location problem, *Discrete Applied Mathematics* **143** (2004), 368–373.

[30] J. Kleinberg and É. Tardos, Approximation algorithms for classification problems with pairwise relationships: metric labeling and Markov random fields, *Journal of the ACM* **49** (2002), 616–639.

[31] A. W. J. Kolen, *Tree Network and Planar Rectilinear Location Theory*, CWI Tract 25, Center for Mathematics and Computer Science, Amsterdam, 1986.

[32] V. Kolmogorov, Submodularity on a tree: Unifying $L^\natural$-convex and bisubmodular functions, in: *Proceedings of the 36th International Symposium on Mathematical Foundations of Computer Science (MFCS'11)*, LNCS 6907, Springer, Berlin, 2011, pp. 400–411

[33] V. Kolmogorov and A. Shioura, New algorithms for convex cost tension problem with application to computer vision, *Discrete Optimization* **6** (2009), 378–393.

[34] V. Kolmogorov and S. Živný, The complexity of conservative valued CSPs, in: *Proceedings of the 23rd ACM-SIAM Symposium on Discrete Algorithms (SODA'12)*, 2012, pp. 750–759.

[35] F. Kuivinen, On the complexity of submodular function minimisation on diamonds, *Discrete Optimization* **8** (2011), 459–477.

[36] L. Lovász, Submodular functions and convexity, in: A. Bachem, M. Grötschel, and B. Korte, eds., *Mathematical Programming—The State of the Art*, Springer-Verlag, Berlin, 1983, 235–257.

[37] K. Murota, Discrete convex analysis, *Mathematical Programming* **83** (1998), 313–371.

[38] K. Murota, Algorithms in discrete convex analysis, *IEICE Transactions on Systems and Information*, **E83-D** (2000), 344–352.

[39] K. Murota, *Discrete Convex Analysis*, SIAM, Philadelphia, 2003.

[40] M. Nakamura, A characterization of greedy sets: Universal polymatroids (I), *Scientific Papers of College of Arts and Science, The University of Tokyo* **38** (1988), 155–167.

[41] J. C. Picard and D. H. Ratliff, A cut approach to the rectilinear distance facility location problem, *Operations Research* **26** (1978), 422–433.

[42] L. Qi, Directed submodularity, ditroids and directed submodular flows, *Mathematical Programming* **42** (1988), 579–599.

[43] T. Schiex, H. Fargier, and G. Verfaillie, Valued constraint satisfaction problems: hard and easy problems, In *Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI'95)*, 1995.

[44] A. Schrijver, A combinatorial algorithm minimizing submodular functions in strongly polynomial time, *Journal of Combinatorial Theory, Series B* **80** (2000), 346–355.

[45] B. C. Tansel, R. L. Francis, and T. J. Lowe, Location on networks I, II, *Management Science* **29** (1983), 498–511.

[46] J. Thapper and S. Živný, The power of linear programming for valued CSPs, in: *Proceedings of the 53rd Annual IEEE Symposium on Foundations of Computer Science (FOCS'12)*, preprint available at `arXiv:1204.1079`.

[47] M. L. J. van de Vel, *Theory of Convex Structures*, North-Holland, Amsterdam, 1993.