Strong greedoid structure of r-removed P-orderings

Dmitrii Krachun[a] Rozalina Mirgalimova[b]

Submitted: Sep 21, 2023; Accepted: Oct 13, 2024; Published: Dec 9, 2024

© The authors. Released under the CC BY-ND license (International 4.0).

Abstract

Inspired by the notion of *r-removed P-orderings* introduced in the setting of Dedekind domains by Bhargava we generalize it to the framework of arbitrary ultrametric spaces. We show that sets of maximal "r-removed perimeter" can be constructed by a greedy algorithm and form a strong greedoid. This gives a simplified proof of several theorems previously obtained by Bhargava, as well as generalises some results of Grinberg and Petrov who considered the case $r = 0$ corresponding, in turn, to simple P-orderings.

Mathematics Subject Classifications: 05B35, 13F05

1 Introduction

Motivated by questions in polynomial function theory, Bhargava [2] introduced the notion of *P-orderings* for a subset X of a Dedekind domain D. The construction is as follows. Given a prime ideal $P \subset D$, let a_0 be an arbitrary element of X and for $k = 1, 2, \dots$ choose $a_k \in X$ to minimize

$$\nu_P\left((a_k - a_0)(a_k - a_1)\dots(a_k - a_{k-1})\right),$$

where ν_P denotes the P-adic valuation on D. One of the results of [2] is the surprising fact that, despite the fact that typically the choice of each a_k is non-unique, the sequence of the resulting valuations does not depend on the specific choice of $\{a_i\}$ but only on X and P. Later, to study bases of the ring of polynomials with integer-valued divided differences, Bhargava [3] generalised this construction to *r-removed P-orderings*. For an r-removed P-ordering one again chooses a sequence $\{a_i\}$ of elements from X but now the first $r+1$ elements $a_0, \dots, a_r$ are chosen arbitrary and then each new element a_k minimizes

$$\min_{\substack{A\subset\{a_0,\dots,a_{k-1}\}\\ |A|=k-r}} \sum_{a\in A} \nu_P(a_k - a).$$

[a]Department of Mathematics, Princeton University, Princeton, US, (dk9781@princeton.edu).
[b]Faculty of Mathematics and Computer Science, St. Petersburg State University, Saint Petersburg, Russia (rozalina.mirgalimova@gmail.com).

Again, one of the results of [3] is that the resulting sequence of minimum valuations does not depend on the choice of $\{a_i\}$.

Recently, Grinberg and Petrov [5] generalised the notion of P-orderings to the context of *ultra triples*, which is a certain extension of ultrametric spaces, obtaining new proofs of several results of [2] and showing that all (prefixes of) P-orderings form a strong greedoid. A natural question which has been asked by Bhargava [1] is then

Do r-removed P-orderings also fit in the framework of [5]?

We resolve the question in the affirmative by showing that virtually all theorems of [5] can be indeed generalised to cover the case of r-removed P-orderings. This also gives a new proof of [3, Theorems 3, 4, 30], see Corollary 15 and Remark 16.

Let us now introduce the objects we work with. An *ultra triple* (E, w, d) is given by a ground set E, an arbitrary *weight function* $w : E \to \mathbb{R}$, and a *distance function* $d : \{(e, f) \in E \times E \mid e \neq f\} \to \mathbb{R}$, which is symmetric, $d(e, f) = d(f, e)$, and satisfying the ultrametric triangle inequality $d(a, b) \leqslant \max\{d(a, c), d(b, c)\}$. Note that unlike in the case of an ultrametric space we do not require distances to be non-negative.

We consider an *r-removed distance* from a point $v \in E$ to a finite subset $C \subset E$ defined by

$$\operatorname{dist}_r(C, v) := \begin{cases} \max\limits_{A \subset C, |A|=|C|-r} \sum_{x \in A} d(v, x) & |C| > r; \\ 0 & |C| \leqslant r. \end{cases}$$

In words, $\operatorname{dist}_r(C, v)$ is the maximum among sums of distances from v to the points of C except for some r of them. Given this notion of a distance from a point to a set, we then define a *greedy r-removed m-permutation* of a finite subset $C \subseteq E$ to be an ordered subset $(c_1, \ldots, c_m)$ of C which is defined greedily by choosing elements $c_1, c_2, \ldots$ one by one in such a way that for each $n \in \{1, 2, \ldots, m\}$ the element c_n maximizes

$$w(c) + \operatorname{dist}_r(\{c_1, \ldots, c_{n-1}\}, c)$$

over all possible $c \in C \setminus \{c_1, \ldots, c_{n-1}\}$, c_1 maximizes only $w(c)$. A related concept is that of an *r-removed perimeter* of a finite ordered set $A = (a_1, \ldots, a_n)$ which is defined to be

$$\operatorname{PER}_r(A) := \sum_{a \in A} w(a) + \sum_{i=1}^{n} \operatorname{dist}_r(\{a_1, a_2, \ldots, a_{i-1}\}, a_i).$$

We prove that the r-removed perimeter does not depend on the ordering of a given set, see Lemma 10, allowing us to naturally extend the notion of an r-removed perimeter to finite (non-ordered) sets. Then, in Theorem 13 we show that any greedy r-removed m-permutation of a finite set C has maximal r-removed perimeter among all subsets of C of size m, and as a corollary we show that, in fact, the sequence of numbers $w(c_j)+\operatorname{dist}_r(\{c_1, c_2, \ldots, c_{j-1}\}, c_j)$ does not depend on the choice of a greedy r-removed m-permutations. Then in Theorem 19 we show that all sets of maximal r-removed perimeter form a strong greedoid. At the end of the paper we also discuss some other notions of perimeter that fit into our framework.

The rest of the paper is organised as follows. In Section 2 we formally define the objects we study and introduce a device we call a *projection* which is later frequently used. Then, in Section 3 we prove that r-removed perimeter of a set is well-defined and that the construction of a greedy r-removed m-permutation gives the maximal r-removed perimeter among all subsets of size m. In Section 4 we prove that subsets of maximal r-removed perimeter form a strong greedoid. Finally, Section 5 is devoted to the discussion of a more general concept of a perimeter of a set, covered by our approach.

2 Basic definitions and constructions

We largely follow the notation used in [5], which we now briefly recall. Throughout the paper, we consider a set E as our ground set, and refer to the elements of E as points. For a non-negative integer m, an m-set means a subset A of E with $|A| = m$, and an m-permutation means an ordered set $A = (a_1, \ldots, a_m)$ formed by distinct elements of E. Analogously, if $B \subseteq E$ is a subset and m is a non-negative integer, an m-subset of B means an m-element subset of B and an m-permutation of B means an ordered set A formed by m distinct elements of B. The following definition already appeared in [5].

Definition 1. An *ultra triple* is a triple (E, w, d), where E is a set, $w : E \to \mathbb{R}$ is an arbitrary *weight function*, and $d : \{(e, f) \in E \times E \mid e \neq f\} \to \mathbb{R}$ is a *distance function*[1] satisfying

- $d(a, b) = d(b, a)$ for any two distinct $a, b \in E$.
- $d(a, b) \leqslant \max\{d(a, c),\, d(b, c)\}$ for any three distinct $a, b, c \in E$.

The inequality above is commonly known as the ultrametric triangle inequality; but unlike the distance function of an ultrametric space, we allow d to take negative values. We refer to $d(a, b)$ as the *distance* between a and b.

The following are formal definitions of the objects already mentioned in the introduction, namely, r-removed distance, r-removed perimeter, and an r-removed m-permutation.

Definition 2. Let (E, w, d) be an ultra triple, $C \subseteq E$ be a finite subset, and v be any point in $E \setminus C$. We define $\mathrm{dist}_r(C, v)$ to be the maximum among all possible sums of distances from v to some $|C| - r$ distinct points of the set C. If $|C| \leqslant r$, we set $\mathrm{dist}_r(C, v) := 0$.

Definition 3. Let (E, w, d) be an ultra triple. For a k-permutation $(a_1, a_2, \ldots, a_k)$ of a finite subset $A \subseteq E$, we define its *r-removed perimeter* by

$$\mathrm{PER}_r((a_1, a_2, \ldots, a_k)) := \sum_{a \in A} w(a) + \sum_{i=1}^{k} \mathrm{dist}_r(\{a_1, a_2, \ldots, a_{i-1}\}, a_i).$$

[1] One could also extend the domain of d to the whole of $E \times E$ by setting $d(a, a) := -\infty$ for all $a \in E$.

Definition 4. For an ultra triple (E, w, d) let $C \subseteq E$ be a finite set, and m be a non-negative integer. A *greedy r-removed m-permutation* of C is a list $(c_1, c_2, \ldots, c_m)$ of m distinct elements of C such that for each $i \in \{1, \ldots, m\}$ and each $x \in C \setminus \{c_1, c_2, \ldots, c_{i-1}\}$, we have

$$\mathrm{PER}_r((c_1, c_2, \ldots, c_i)) \geqslant \mathrm{PER}_r((c_1, c_2, \ldots, c_{i-1}, x)). \tag{1}$$

We now define a useful construction that we are going to use in the proofs, it earlier implicitly appeared in [5]. We first recall the following definition from [5].

Definition 5. Let (E, w, d) be an ultra triple, $A \subseteq E$ be a finite non-empty subset and $c \in E$ be any point. We define a subset $\mathrm{proj}_A(c)$ of A as follows:

- If $c \in A$, then $\mathrm{proj}_A(c) := \{c\}$;
- If $c \notin A$, then $\mathrm{proj}_A(c)$ is the set of all $a \in A$ that minimize the distance $d(c, a)$.

Later in the paper $\mathrm{proj}_A(c)$ is called a *projection* of c onto A. We extend this definition by the following construction which already appeared in [5, proof of Theorem 21].

Definition 6. Let (E, w, d) be an ultra triple, $C = (c_1, c_2, \ldots, c_k) \subseteq E$ be a finite ordered set and A be any n-subset of E, where $n \geqslant k$. We define a k-permutation $(v_1, v_2, \ldots, v_k)$ of A recursively as follows: each v_i is an element of the projection of c_i onto $A \setminus \{v_1, v_2, \ldots, v_{i-1}\}$ for each $i = 1, 2, \ldots, k$. We denote $(v_1, v_2, \ldots, v_k)$ by $\mathrm{proj}(C \to A)$. If $\mathrm{proj}_{A\setminus\{v_1,v_2,\ldots,v_{i-1}\}}(c_i)$ contains more than one element, an arbitrary element of the set is selected.

There are three important observations about these constructions that we now make. The first proposition already appeared in [5, Lemma 13(c)] and we give its short proof for completeness.

Proposition 7. *Let (E, w, d) be an ultra triple and $A \subseteq E$ be a non-empty finite set. Then for a point $c \in E$, its projection $b \in \mathrm{proj}_A(c)$ and any $x \in A \setminus \{b\}$ we have $d(b, x) \leqslant d(c, x)$.*

Proof. If $c \in A$ then $b = c$ and we trivially have an equality. Otherwise, since $x \in A$, by the definition of the projection we have $d(c, x) \geqslant d(c, b)$ and so by the ultrametric triangle inequality we have

$$d(b, x) \leqslant \max\{d(c, b), d(c, x)\} = d(c, x).$$

□

Proposition 8. *Let (E, w, d) be an ultra triple, $C = (c_1, c_2, \ldots, c_k) \subseteq E$ be a finite ordered set and A be an n-subset of E, with $n \geqslant k$. Denote* $\mathrm{proj}(C \to A)$ *by $(v_1, v_2, \ldots, v_k)$. Then for each $j \in \{1, 2, \ldots, k\}$ one has*

$$(A \setminus \{v_1, \ldots, v_j\}) \cap \{c_1, c_2, \ldots, c_j\} = \varnothing.$$

Proof. Arguing by contradiction we assume for some $i \leqslant j \leqslant k$ that $c_i \in A \setminus \{v_1, \ldots, v_j\}$. In particular, this implies that $c_i \in A \setminus \{v_1, v_2, \ldots, v_{i-1}\}$. By definition this means that $v_i \in \mathrm{proj}_{A\setminus\{v_1,v_2,\ldots,v_{i-1}\}}(c_i) = \{c_i\}$. Hence, $v_i = c_i \in A \setminus \{v_1, \ldots, v_j\}$ which is impossible. □

Proposition 9. *Let (E, w, d) be an ultra triple, $C = (c_1, c_2, \dots, c_k) \subseteq E$ be a finite ordered set and A be an n-subset of E, with $n > k$. Then for each $v \in A \setminus \operatorname{proj}(C \to A)$*

$$\operatorname{dist}_r(\operatorname{proj}(C \to A), v) \leqslant \operatorname{dist}_r(C, v).$$

Proof. Denote $\operatorname{proj}(C \to A)$ by $(v_1, v_2, \dots, v_k)$. The statement of the proposition would follow from the inequality $d(v_i, v) \leqslant d(c_i, v)$ for each $i \in \{1, 2, \dots, k\}$. But since $v \in A \setminus \{v_1, v_2, \dots, v_{i-1}\}$ this inequality is given by Proposition 7 applied to the set $A \setminus \{v_1, v_2, \dots, v_{i-1}\}$ and points c_i, $v_i \in \operatorname{proj}_{A \setminus \{v_1, v_2, \dots, v_{i-1}\}}(c_i)$ and v. □

3 Perimeter and greedy r-removed m-permutations

We first prove that any two orderings of a given set have the same r-removed perimeter.

Lemma 10. *Any two orderings of a finite set $A \subseteq E$ have the same r-removed perimeter.*

Proof. It suffices to prove the statement for pairs of orderings which differ by one transposition. The general case is then reduced to it by consecutive transpositions.

Let us prove the statement for $(a_1, \dots, a_t, a_{t+1}, \dots, a_k)$ and $(a_1, \dots, a_{t+1}, a_t, \dots, a_k)$. Denote by C the set $\{a_1, a_2, \dots, a_{t-1}\}$. Many summands from the definition of r-removed perimeter coincide; all that remains to prove is

$$\operatorname{dist}_r(C, a_t) + \operatorname{dist}_r(C \cup \{a_t\}, a_{t+1}) = \operatorname{dist}_r(C, a_{t+1}) + \operatorname{dist}_r(C \cup \{a_{t+1}\}, a_t). \qquad (2)$$

If $t \leqslant r$, both sides are 0. Otherwise, we let $z = d(a_t, a_{t+1})$, $x_j = d(a_t, a_j)$ and $y_j = d(a_{t+1}, a_j)$, where $j = 1, \dots, t-1$. In what follows we only consider triangles of the form $a_t a_{t+1} a_j$ for some $j = 1, 2, \dots, t-1$ and use the ultrametric triangle inequality for them.

We colour triangles with two sides strictly greater than z in red, in which case $x_j = y_j$ by the ultrametric inequality. Triangles coloured in red correspond to some largest distances from points a_t and a_{t+1} to the set C which coincide. In any other triangle we must have $x_i = z \geqslant y_i$ or $y_i = z \geqslant x_i$.

If the number of red triangles is at least $t - r$, then $\operatorname{dist}_r(C, a_t) = \operatorname{dist}_r(C, a_{t+1})$, $\operatorname{dist}_r(C \cup \{a_t\}, a_{t+1}) = \operatorname{dist}_r(C \cup \{a_{t+1}\}, a_t)$ and (2) is true.

If there are less than $t - r$ red triangles, then $\operatorname{dist}_r(C \cup \{a_t\}, a_{t+1}) = \operatorname{dist}_r(C, a_{t+1}) + z$ and $\operatorname{dist}_r(C \cup \{a_{t+1}\}, a_t) = \operatorname{dist}_r(C, a_t) + z$. By substituting these expressions into (2), we again get an equality. □

In light of this lemma we have the following definition.

Definition 11. For a finite subset $A \subseteq E$, we define its *r-removed perimeter* $\operatorname{PER}_r(A)$ to be the common r-removed perimeter of all orderings of A.

Remark 12. For the case $r = 0$, the r-removed perimeter is the sum of the distances between all unordered pairs of points plus the sum of the weight function of all points. This case was considered in [5].

Theorem 13. *Let (E, w, d) be an ultra triple and $C \subseteq E$ be a finite subset, and m and r be non-negative integers. Let $(c_1, c_2, \ldots, c_m)$ be any greedy r-removed m-permutation of C. Then, for each $k \in \{0, 1, \ldots, m\}$, the set $\{c_1, c_2, \ldots, c_k\}$ has maximum r-removed perimeter among all k-subsets of C.*

Proof. Given a greedy r-removed m-permutation $(c_1, c_2, \ldots, c_m)$, we want to prove that for any k-subset $A \subseteq C$, $\mathrm{PER}_r(A) \leqslant \mathrm{PER}_r(\{c_1, c_2, \ldots, c_k\})$. We induct on k.

For $k = 0$ both perimeters are 0, and so the inequality is trivially true. For the induction step from $k-1$ to k, let $(v_1, v_2, \ldots, v_k) := \mathrm{proj}((c_1, c_2, \ldots, c_k) \to A)$, which is an ordering of A. Then by Proposition 8 and the fact that $v_k \in A \setminus \{v_1, \ldots, v_{k-1}\}$, we have:

$$v_k \notin \{c_1, c_2, \ldots, c_{k-1}\}.$$

By induction hypothesis we know that

$$\mathrm{PER}_r(\{v_1, v_2, \ldots, v_{k-1}\}) \leqslant \mathrm{PER}_r(\{c_1, c_2, \ldots, c_{k-1}\}),$$

and so to complete the induction step it suffices to show that

$$\mathrm{PER}_r(A) - \mathrm{PER}_r(\{v_1, v_2, \ldots, v_{k-1}\}) \leqslant \mathrm{PER}_r(\{c_1, c_2, \ldots, c_k\}) - \mathrm{PER}_r(\{c_1, c_2, \ldots, c_{k-1}\}),$$

which, implicitly using Lemma 10, can be equivalently written as

$$w(v_k) + \mathrm{dist}_r(\{v_1, v_2, \ldots, v_{k-1}\}, v_k) \leqslant w(c_k) + \mathrm{dist}_r(\{c_1, c_2, \ldots, c_{k-1}\}, c_k). \tag{3}$$

We now turn to proving (3). Since $v_k \in A \setminus \{c_1, c_2, \ldots, c_{k-1}\} \subseteq C \setminus \{c_1, c_2, \ldots, c_{k-1}\}$ (recall that $A \subseteq C$), we have $\mathrm{PER}_r\{c_1, c_2, \ldots, c_{k-1}, v_k\} \leqslant \mathrm{PER}_r\{c_1, c_2, \ldots, c_k\}$ by the definition of a greedy r-removed m-permutation. Subtracting $\mathrm{PER}_r(\{c_1, c_2, \ldots, c_{k-1}\})$ from both sides we arrive at

$$w(v_k) + \mathrm{dist}_r(\{c_1, c_2, \ldots, c_{k-1}\}, v_k) \leqslant w(c_k) + \mathrm{dist}_r(\{c_1, c_2, \ldots, c_{k-1}\}, c_k).$$

And so to deduce (3) it remains to show that

$$\mathrm{dist}_r(\{v_1, v_2, \ldots, v_{k-1}\}, v_k) \leqslant \mathrm{dist}_r(\{c_1, c_2, \ldots, c_{k-1}\}, v_k).$$

Which is nothing else but the statement of Proposition 9 for the point $v = v_k$ and $(v_1, v_2, \ldots, v_{k-1}) = \mathrm{proj}((c_1, c_2, \ldots, c_{k-1}) \to A)$. ☐

Remark 14. It follows from the proof that if the equality

$$\mathrm{PER}_r(\{v_1, v_2, \ldots, v_k\}) = \mathrm{PER}_r(\{c_1, c_2, \ldots, c_k\})$$

holds for $(v_1, v_2, \ldots, v_k) := \mathrm{proj}(\{c_1, c_2, \ldots, c_k\} \to A)$, then for each $j < k$ one also has an equality $\mathrm{PER}_r(\{v_1, v_2, \ldots, v_j\}) = \mathrm{PER}_r(\{c_1, c_2, \ldots, c_j\})$.

Corollary 15. *Let $C \subseteq E$ be a set, m and r be non-negative integers, $j \in \{1, 2, \ldots, m\}$. If $(c_1, c_2, \ldots, c_m)$ is a greedy r-removed m-permutation of C, then the number*

$$w(c_j) + \mathrm{dist}_r(\{c_1, c_2, \ldots, c_{j-1}\}, c_j)$$

does not depend on the choice of this greedy r-removed m-permutation but only depends on C, r and j.

Proof. By Theorem 13, for each $k \leqslant m$ the set $\{c_1, c_2, \ldots, c_k\}$ has maximum r-removed perimeter among all k-subsets of C, which implies that $\mathrm{PER}_r(\{c_1, c_2, \ldots, c_k\})$ does not depend on the choice of the greedy r-removed m-permutation of C. It remains to note that

$$w(c_j) + \mathrm{dist}_r(\{c_1, c_2, \ldots, c_{j-1}\}, c_j) = \mathrm{PER}_r(\{c_1, c_2, \ldots, c_j\}) - \mathrm{PER}_r(\{c_1, c_2, \ldots, c_{j-1}\}).$$

□

Remark 16. As a special case of this corollary we obtain the results of [3, Theorems 3, 4, 30]. Indeed, for a Dedekind domain D, a prime ideal $P \subset D$, and $h \in \mathbb{Z}_{\geqslant 0}$, the distance function $d_{P,h}(a, b) := -\max(h, \nu_P(a - b))$ satisfies the ultrametric triangle inequality and so the result follows from Corollary 15 applied to an ultra triple $(S, w \equiv 0, d_{P,h})$.

We now prove the converse of Theorem 13, namely, that any set of maximal r-removed perimeter is a prefix of some greedy r-removed m-permutation.

Theorem 17. *Let (E, w, d) be an ultra triple, $C \subseteq E$ be a finite set, and m be a non-negative integer such that $|C| \geqslant m$. Let $k \in \{0, 1, \ldots, m\}$ and A be a k-subset of C having maximum r-removed perimeter (among all k-subsets of C). Then, there exists a greedy r-removed m-permutation of C for which A is a prefix of this permutation.*

Proof. Choose an arbitrary greedy r-removed m-permutation $(c_1, c_2, \ldots, c_m)$ of C by starting with any point and continuing the sequence greedily choosing elements from the remaining points. By Theorem 13, the set $(c_1, c_2, \ldots, c_k)$ has maximum perimeter among all k-subsets of C. Hence, $\mathrm{PER}_r(A) = \mathrm{PER}_r(\{c_1, c_2, \ldots, c_k\})$ since the set A also has maximum r-removed perimeter among them.

Let $(v_1, v_2, \ldots, v_k) := \mathrm{proj}((c_1, c_2, \ldots, c_k) \to A)$. What we want to prove is that there exists a greedy r-removed m-permutation of C which starts from $(v_1, v_2, \ldots, v_k)$, which is equivalent to checking that for each $p \leqslant k$ the point v_p maximizes

$$w(x) + \mathrm{dist}_r(\{v_1, \ldots, v_{p-1}\}, x)$$

over all $x \in C \setminus \{v_1, \ldots, v_{p-1}\}$.

As mentioned in Remark 14, the fact that

$$\mathrm{PER}_r(\{v_1, v_2, \ldots, v_k\}) = \mathrm{PER}(A) = \mathrm{PER}_r(\{c_1, c_2, \ldots, c_k\})$$

implies that for each $j \leqslant k$ we have $\mathrm{PER}_r(\{v_1, v_2, \ldots, v_j\}) = \mathrm{PER}_r(\{c_1, c_2, \ldots, c_j\})$. In particular, this holds for $j = p$. Now, arguing by contradiction we assume that there exists $x \in C \setminus \{v_1, \ldots, v_{p-1}\}$ such that

$$w(x) + \mathrm{dist}_r(\{v_1, \ldots, v_{p-1}\}, x) > w(v_p) + \mathrm{dist}_r(\{v_1, \ldots, v_{p-1}\}, v_p).$$

This would mean that

$$\mathrm{PER}_r(\{v_1, \ldots, v_{p-1}, x\}) > \mathrm{PER}_r(\{v_1, \ldots, v_{p-1}, v_p\}) = \mathrm{PER}_r(\{c_1, \ldots, c_{p-1}, c_p\}),$$

contradicting the fact that $\{c_1, \ldots, c_{p-1}, c_p\}$ has the largest r-removed perimeter among all subsets of C of size p. □

4 Strong greedoid of maximum perimeter sets

In [5] it was shown that sets maximizing the perimeter (i.e. r-removed perimeter with $r = 0$) form a strong greedoid. In this section we generalize this statement to all $r \geqslant 0$. We start by recalling the relevant definitions from the theory of greedoids.

Definition 18. A collection $\mathcal{F} \subseteq 2^E$ of subsets of a finite set E is called a *greedoid* (on the ground set E) if it satisfies the following three axioms:

(i) $\varnothing \in \mathcal{F}$.

(ii) If $A \in \mathcal{F}$ satisfies $|A| > 0$, then there exists $a \in A$ such that $A \setminus \{a\} \in \mathcal{F}$.

(iii) If $A, B \in \mathcal{F}$ satisfy $|A| = |B| + 1$, then there exists $a \in A \setminus B$ such that $B \cup \{a\} \in \mathcal{F}$.

A greedoid $\mathcal{F}$ on a ground set E is called a *strong greedoid* (also known as "Gauss greedoid") if it additionally satisfies the following axiom:

(iv) If $A, B \in \mathcal{F}$ satisfy $|A| = |B| + 1$, then there exists $a \in A \setminus B$ such that $B \cup \{a\} \in \mathcal{F}$ and $A \setminus \{a\} \in \mathcal{F}$.

There are several equivalent definitions of a greedoid in the literature, ours is taken from [6, Section IV.1]. Specifically, our axioms (i) and (iii) align with conditions (1.4) and (1.6) in [6, Section IV.1], while axioms (i) and (ii) establish $(E, \mathcal{F})$ as an accessible set system. The definition of a strong greedoid can be found in [4].

Now we assume that the set E is finite. The following theorem shows that sets with maximal r-removed perimeter form a strong greedoid.

Theorem 19. *Let (E, w, d) be an ultra triple on a finite ground set and $\mathcal{F}_r$ denote the collection of subsets $A \subseteq E$ that have maximum r-removed perimeter among all $|A|$-sets:*

$$\mathcal{F}_r := \{A \subseteq E \mid \mathrm{PER}_r(A) \geqslant \mathrm{PER}_r(B) \text{ for all } B \subseteq E \text{ satisfying } |B| = |A|\}.$$

Then $\mathcal{F}_r$ is a strong greedoid on the ground set E.

We start by proving the following lemma.

Lemma 20. *Let A and B be two subsets of E such that $|A| = |B| + 1$. Then, there exists $u \in A \setminus B$ satisfying*

$$\mathrm{PER}_r(A \setminus \{u\}) + \mathrm{PER}_r(B \cup \{u\}) \geqslant \mathrm{PER}_r(A) + \mathrm{PER}_r(B). \tag{4}$$

Proof. Let $k = |B|$ and so $|A| = k + 1$. With a slight abuse of notation we denote by B an arbitrary ordering of B, which we fix from now on. Define $(v_1, v_2, \dots, v_k) := \mathrm{proj}(B \to A)$ and let u be the unique element of $A \setminus \{v_1, v_2, \dots, v_k\}$. By Proposition 8 we have $u \notin B$. We now want to prove (4) for this choice of u. Subtracting $\mathrm{PER}_r(A \setminus \{u\}) + \mathrm{PER}_r(B) + w(u)$ from both sides we arrive at an equivalent inequality

$$\mathrm{dist}_r(B, u) \geqslant \mathrm{dist}_r(A \setminus \{u\}, u),$$

which is simply the result of Proposition 9 applied to u and $A \setminus \{u\} = \mathrm{proj}(B \to A)$. ☐

Proof of Theorem 19. First note that property (i) is trivial, and (iii) immediately follows from (iv). Furthermore, since E is finite, for any $s \leqslant |E|$ there exists $B \in \mathcal{F}_r$ with $|B| = s$, and so by choosing arbitrary $B \in \mathcal{F}_r$ with $|B| = |A| - 1$ we can deduce (ii) from (iv).

To prove (iv) we use Lemma 20 to construct $u \in A \setminus B$ satisfying (4). Since $A \in \mathcal{F}_r$ we must have $\mathrm{PER}_r(B \cup \{u\}) \leqslant \mathrm{PER}_r(A)$. Similarly, $B \in \mathcal{F}_r$ implies $\mathrm{PER}_r(A \setminus \{u\}) \leqslant \mathrm{PER}_r(B)$. Together with (4) these two inequalities immediately imply that

$$\mathrm{PER}_r(A \setminus \{u\}) = \mathrm{PER}_r(B), \qquad \mathrm{PER}_r(B \cup \{u\}) = \mathrm{PER}_r(A),$$

which means that both $A \setminus \{u\}$ and $B \cup \{u\}$ are in $\mathcal{F}_r$. This shows that $\mathcal{F}_r$ is a strong greedoid. ☐

5 Other perimeters

In this section we consider more general notions of perimeter $\mathrm{PER}(A)$ which fit into our framework. Lemma 24 provides a large class of perimeters for which theorems of previous sections still hold true, Lemma 25 then shows that for the ultra triple $(\mathbb{Z}, w, -\nu_p)$ the perimeters we construct are the only ones.

Instead of the r-removed distance $\mathrm{dist}_r(C, v)$ we could start from some other notion of a distance from a point to a set, call it $\mathrm{dist}(C, v)$, and define $\mathrm{PER}((a_1, \dots, a_n))$ of an ordered set $A := (a_1, \dots, a_n)$ by setting $\mathrm{PER}(A) := \sum_{a \in A} w(a) + \sum_{i=1}^{n} \mathrm{dist}(\{a_1, a_2, \dots, a_{i-1}\}, a_i)$. Tracking the proofs of Theorems 13, 17 and 19 we see that the only two properties of the dist function that we use are given by the following

Definition 21. We say that a real-valued function dist satisfies property **S** if for any set $\{c_1, c_2, \dots, c_n\} =: C \subseteq E$ and distinct $x, y \in E \setminus C$ one has

(S1) $\mathrm{dist}(C, x) + \mathrm{dist}(C \cup \{x\}, y) = \mathrm{dist}(C, y) + \mathrm{dist}(C \cup \{y\}, x)$;

(S2) If $d(c_i, x) \leqslant d(c_i, y)$ for each $i \in \{1, 2, \ldots, n\}$, then

$$\mathrm{dist}(\{c_1, c_2, \ldots, c_n\}, x) \leqslant \mathrm{dist}(\{c_1, c_2, \ldots, c_n\}, y).$$

Remark 22. Indeed, property **S1** is used in Lemma 10 to prove that the perimeter of a set is well-defined, and is, in fact, equivalent to this lemma. Property **S2** is first used in the proof of Proposition 9, and through it, indirectly in Theorems 13, 17, 19, and Lemma 20.

We now give a large family of distances satisfying property **S**.

Lemma 23. *Let (E, w, d) be an ultra triple and $f = (f_j)_{j=1}^{\infty}$ be a sequence of functions $f_j : \mathbb{R} \to \mathbb{R}$. For a set $\{c_1, \ldots, c_n\} = C \subseteq E$ and $x \in E \setminus C$, let*

$$\mathrm{dist}_f(C, x) := \sum_{j=1}^{n} f_j(d_j),$$

where $(d_1, d_2, \ldots, d_n)$ is the ordered set of values $d(c_1, x), \ldots, d(c_n, x)$ arranged in non-decreasing order. Then dist_f *satisfies property* **S1**.

Proof. We will prove that

$$\mathrm{dist}_f(C, x) + \mathrm{dist}_f(C \cup \{x\}, y) = \mathrm{dist}_f(C, y) + \mathrm{dist}_f(C \cup \{y\}, x) \tag{5}$$

for fixed C and $x, y \in E \setminus C$, with $|C| = n$ and $x \neq y$. Essentially, we can focus on the set of points $E = C \cup \{x, y\}$, $|E| = n + 2$. Since (5) depends only on the values of f_j at finitely many points, we can restrict each f_j to the finite set of distances that appear in it.

Then, since the property (5) is linear in $f = (f_j)_{j=1}^{n+1}$, it suffices to check it for dist_f with

$$f_j := \begin{cases} 0 & j \leqslant r; \\ g & j > r. \end{cases}$$

Where $g : \mathbb{R} \to \mathbb{R}$ is some fixed non-decreasing function. Indeed, any sequence $f = (f_j)_{j=1}^{n+1}$ of functions defined on a finite set can be written as a linear combination of sequences of the form $(0, 0, \ldots, 0, g, g, \ldots, g)$, and any such function g can be written as a difference of two non-decreasing functions on this finite set.

But for this specific choice of f, the distance dist_f is nothing else but the r-removed distance for the ultra triple (E, w, d_g) in which the distance function d_g is given by[2] $d_g(a, b) := g(d(a, b))$ and so the equality **S1** is given by (2) from the proof of Lemma 10. ☐

Lemma 24. *In the setting of Lemma 23, if one additionally requires that each function f_j is non-decreasing, then* dist_f *satisfies property* **S**.

[2] One easily sees that d_g satisfies the ultrametric inequality for any non-decreasing g.

Proof. Property **S1** is checked in Lemma 23. To check **S2** it suffices to note that if one n-tuple is entry-wise smaller than another, then the same remains true after each of the n-tuples is sorted from the smallest value to the largest, and then use the fact that each f_j is non-decreasing. $\square$

For sufficiently large spaces and under certain natural conditions on the distance function dist from a point to a set we manage to prove the reverse of Lemma 24. To avoid stating technical conditions we prove the result for the space $(\mathbb{Z}, w, -\nu_p)$ in which distance between points $a, b \in \mathbb{Z}$ is given by $-\nu_p(a-b)$, where ν_p stands for the p-adic valuation and we further assume that $p > 2$.

Lemma 25. *Consider an ultra triple $(\mathbb{Z}, w, -\nu_p)$ with arbitrary weight function w and $p > 2$. Assume that the value of the distance function* $\operatorname{dist}(C, x)$ *from a point x to a set C depends only on the multiset of distances from x to the points of C and that* dist *satisfies property* **S**. *Then* $\operatorname{dist} \equiv \operatorname{dist}_f$ *for some sequence of non-decreasing functions $f = (f_n)_{n=1}^{\infty}$.*

Proof. By assumption there exists a sequence of symmetric functions $g_n : \mathbb{Z}^n_{\leqslant 0} \to \mathbb{R}$ indexed by $n \geqslant 1$, such that for any point x and any set $\{c_1, \ldots, c_n\}$ not containing x we have

$$\operatorname{dist}(x, \{c_1, \ldots, c_n\}) = g_n(d(c_1, x), \ldots, d(c_n, x)).$$

We want to prove the existence of a sequence of non-decreasing functions $(f_n)_{n=1}^{\infty}$ such that for any non-positive integers $d_1 \leqslant d_2 \leqslant \ldots \leqslant d_n$

$$g_n(d_1, d_2, \ldots d_n) = \sum_{j=1}^{n} f_j(d_j). \tag{6}$$

We prove the existence of functions f_j by induction on j, and for the base case we set $f_1 := g_1$ which is non-decreasing by the second condition of property **S**.

Now assume that $f_1, \ldots, f_{m-1}$ are already defined in such a way that (6) is satisfied for all $n < m$ and we want to define f_m. For each $d \in \mathbb{Z}_{\leqslant 0}$ we set

$$f_m(d) := g_m(d, d, \ldots, d) - g_{m-1}(d, d, \ldots, d),$$

where we have m arguments equal to d in the first case and $m-1$ arguments equal to d in the second.

First, we check that (6) is satisfied for $n = m$. For this, given non-positive integers $d_1 \leqslant d_2 \leqslant \ldots \leqslant d_n$ we consider two points $x, y \in \mathbb{Z}$ with $d(x, y) = d_n$ and a set of points $C := \{c_1, \ldots, c_{n-1}\}$ such that for each $j = 1, \ldots, n-1$ we have $d_j := d(c_j, x)$ and $d(c_j, y) = d_n$. The existence of such a set follows from the property of $\mathbb{Z}$ with the p-adic distance (where $p > 2$) which guarantees that for any two points $a, b \in \mathbb{Z}$ and any $\ell \leqslant d(a, b)$ there exists $c \in \mathbb{Z}$ such that $d(a, c) = \ell$ and $d(b, c) = d(a, b)$. Using **S1** we write

$$g(d_1, \ldots, d_n) = \operatorname{dist}(C \cup \{y\}, x) = \operatorname{dist}(C, x) + \left(\operatorname{dist}(C \cup \{x\}, y) - \operatorname{dist}(C, y)\right),$$

and it remains to observe that the difference in parentheses is equal to $f_n(d_n)$ by the definition of f_n and

$$\operatorname{dist}(C, x) = \sum_{j=1}^{n-1} f_j(d_j)$$

by induction hypothesis.

Second, we show that f_m is non-decreasing. Let $\ell_1, \ell_2 \in \mathbb{Z}_{\leqslant 0}$ satisfy $\ell_1 \leqslant \ell_2$. By (6) we have

$$f_m(\ell_2) - f_m(\ell_1) = g_m(\ell_1, \ldots, \ell_1, \ell_2) - g_m(\ell_1, \ldots, \ell_1, \ell_1)$$

and so $f_m(\ell_2) \geqslant f_m(\ell_1)$ directly follows from **S2**. □

Acknowledgements

We would like to thank Fedor Petrov for his guidance in this project and Darij Grinberg for carefully proofreading an earlier version of the paper and providing helpful feedback, especially on Section 5. This work is supported by RSF grant 22-11-00131.

References

[1] M. Bhargava. private communication, 2022.

[2] M. Bhargava. P-orderings and polynomial functions on arbitrary subsets of Dedekind rings. *Journal für die reine und angewandte Mathematik*, 490:101–128, 1997.

[3] M. Bhargava. On p-orderings, rings of integer-valued polynomials, and ultrametric analysis. *Journal of the American Mathematical Society*, 22(4):963–993, 2009.

[4] V. Bryant and I. Sharpe. Gaussian, Strong and Transversal Greedoids. *European Journal of Combinatorics*, 20(4):259–262, 1999.

[5] D. Grinberg and F. Petrov. A greedoid and a matroid inspired by Bhargava's p-orderings. *The Electronic Journal of Combinatorics*, 28(3), 2021.

[6] B. Korte, R. Schrader, and L. Lovász. Greedoids. *Algorithms and Combinatorics*, vol. 4. Springer-Verlag Berlin Heidelberg, 1991.