Séminaire Lotharingien de Combinatoire XX (2020)
Article #YY, 12 pp.

Proceedings of the 32nd Conference on Formal Power Series and Algebraic Combinatorics (Ramat Gan)

The Bhargava greedoid

Darij Grinberg*[1], and Fedor Petrov†[2]

[1]*Mathematics Department, Drexel University, Philadelphia, PA 19104, U.S.A.*
[2]*St. Petersburg State University and St. Petersburg Department, Steklov Mathematical Institute of Russian Academy of Sciences, St. Petersburg, Russia*

Abstract. The *greedoids* (as defined in 1981 by Korte and Lóvász) are a class of set systems similar to the matroids (but more permissive). Inspired by Bhargava's generalized factorials, we introduce a greedoid of maximum-perimeter subsets in a finite ultrametric space (actually, in a somewhat more general setting than that). We prove that it is a *Gaussian elimination greedoid* over any sufficiently large (e.g., infinite) field; this is a greedoid analogue of a representable matroid. We establish further properties of this greedoid, generalizing two facts shown by Bhargava. We also briefly discuss its relation to a greedoid appearing in phylogenetics, as well as a multiset analogue.

Keywords: greedoids, generalized factorials, phylogenetic trees, Vandermonde determinant

1 Introduction

Since it was coined in 1981 by Korte and Lóvász, the concept of a *greedoid* has seen significant developments [9] and shown up in multiple mathematical fields. It is a type of set system (i.e., a set of subsets of a given ground set E) that can be viewed as a weakening of the notion of a matroid; more precisely, the independent sets of a matroid form a greedoid. In this abstract, we shall define a greedoid stemming from Bhargava's theory of generalized factorials (in a setting significantly more general than Bhargava's) and prove that it is a *Gaussian elimination greedoid* over any sufficiently large (e.g., infinite) field (a greedoid analogue of a representable matroid). Roughly speaking, the sets that belong to this greedoid are subsets of maximum perimeter (among all subsets of their size) of an ultrametric space. (Our setup is actually more general, allowing distances in an arbitrary ordered abelian group and allowing vertex weights to influence the perimeters.) We shall contrast this greedoid – which we have dubbed the *Bhargava greedoid* – with the strong greedoid of maximum-diversity leaf sets of a phylogenetic tree found by Moulton, Semple and Steel in [10]. We shall also derive analogues of some of our results to multisets instead of sets.

The results in this abstract are proved in detail in [8] and in [7].

*darijgrinberg@gmail.com.
†fedyapetrov@gmail.com.

2 Greedoids and strong greedoids

We begin with some fundamental definitions.

Definition 2.1. *Let E be a set. A* set system *on the ground set E is defined to mean a set of subsets of E.*

Greedoids – just as matroids – can be defined as a certain class of set systems:

Definition 2.2. *Let E be a finite set. A set system $\mathcal{F}$ on the ground set E is said to be a* greedoid *if it satisfies the following three axioms:*

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

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

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

Furthermore, $\mathcal{F}$ is said to be a strong greedoid *if it satisfies the following axiom (in addition to **(i)**, **(ii)** and **(iii)**):*

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

Note that axiom **(iv)** implies axiom **(iii)**. We notice that a greedoid $\mathcal{F}$ that is subclusive (i.e., that satisfies $(A \in \mathcal{F} \text{ and } B \subseteq A) \implies (B \in \mathcal{F})$) is the same as a matroid (defined in terms of independent sets).

The fruits of the first 10 years of greedoid theory are gathered in the monograph [9]; more has been done since. Greedoids have shown up in phylogenetics [10] and chip-firing theory [5]; they have also proven to be a source of shellable simplicial complexes ([4], [9, Ch. XII]). Strong greedoids appear in [9] under the name "Gauss greedoids", albeit with a different definition (whose equivalence to our definition is proved in [6]).

In this note, we shall concern ourselves with an even more restrictive class of greedoids: the *Gaussian elimination greedoids*. Our definition of this class follows [9, §IV.2.3], except that we will be using vector families instead of matrices (which is equivalent, since any matrix can be identified with the vector family consisting of its columns):

Definition 2.3. *Let E be a finite set. Let*[1] *$m \in \mathbb{N}$ satisfy $m \geqslant |E|$. Let $\mathbb{K}$ be a field. For each $k \in \{0, 1, \ldots, m\}$, let $\pi_k : \mathbb{K}^m \to \mathbb{K}^k$ be the canonical projection that removes all but the first k coordinates of a column vector. (That is, $\pi_k\left((a_1, a_2, \ldots, a_m)^T\right) = (a_1, a_2, \ldots, a_k)^T$.)*

For each $e \in E$, let $v_e \in \mathbb{K}^m$ be a column vector. The family $(v_e)_{e \in E}$ will be called a vector family *over $\mathbb{K}$.*

[1] Here and in the following, $\mathbb{N}$ denotes the set $\{0, 1, 2, \ldots\}$.

Let $\mathcal{G}$ be the following set of subsets of E:

$$\left\{F \subseteq E \mid \text{the family } \left(\pi_{|F|}\left(v_e\right)\right)_{e \in F} \in \left(\mathbb{K}^{|F|}\right)^F \text{ is linearly independent}\right\}.$$

Then, $\mathcal{G}$ is called the Gaussian elimination greedoid *of the vector family $(v_e)_{e \in E}$.*

Example 2.4. *Let $\mathbb{K} = \mathbb{Q}$ and $E = \{1,2,3,4,5\}$ and $m = 4$. Let $v_1, v_2, v_3, v_4, v_5 \in \mathbb{K}^5$ be the columns of the 5×5-matrix*

$$\begin{pmatrix} 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Then, the Gaussian elimination greedoid of the vector family $(v_e)_{e \in E} = (v_1, v_2, v_3, v_4, v_5)$ is

$$\{\varnothing, \{2\}, \{3\}, \{5\}, \{1,2\}, \{1,3\}, \{1,5\}, \{2,3\}, \{2,5\}, \{1,2,3\}, \{1,2,5\}, \{1,2,3,5\}\}.$$

For example, the 3-element set $\{1,2,5\}$ belongs to this greedoid because the corresponding family $(\pi_3(v_e))_{e \in \{1,2,5\}} \in \left(\mathbb{K}^3\right)^{\{1,2,5\}}$ is linearly independent (indeed, this family consists of the vectors $(0,1,0)^T$, $(1,1,2)^T$ and $(1,0,1)^T$).

Any Gaussian elimination greedoid is a strong greedoid. (This is implicit in [9], and not hard to prove using the Plücker identity; see [7, §11] for details.)

If $\mathcal{F}$ is a strong greedoid, and $k \in \mathbb{N}$, then the k-element sets in $\mathcal{F}$ (if they exist) are the bases of a matroid. If $\mathcal{F}$ is a Gaussian elimination greedoid over a field $\mathbb{K}$, then this matroid is representable over $\mathbb{K}$.

3 $\mathbb{V}$-ultra triples

We now come to our main concept: the "$\mathbb{V}$-ultra triple", a generalization of ultrametric spaces that supports vertex weights. First, an auxiliary notation is needed:

Definition 3.1. *Let E be a set. Then, $E \underline{\times} E$ shall denote the set $\{(e,f) \in E \times E \mid e \neq f\}$.*

Convention 3.2. *Fix a totally ordered abelian group $(\mathbb{V}, +, 0, \leqslant)$ (with ground set $\mathbb{V}$, group operation $+$, zero 0 and smaller-or-equal relation $\leqslant$). We shall refer to this group simply as $\mathbb{V}$. We will use the $\sum$ sign for finite sums inside this group, and we will denote the maximum of a subset S of $\mathbb{V}$ (with respect to the relation $\leqslant$) by $\max S$.*

For almost all examples we know, it suffices to set $\mathbb{V}$ to be the abelian group $\mathbb{R}$, or even $\mathbb{Z}$. Nevertheless, we shall work in full generality, as it serves to separate objects that would otherwise easily be confused.

Definition 3.3. *A* $\mathbb{V}$-ultra triple *shall mean a triple* (E, w, d) *consisting of:*

- *a set* E*, called the* ground set *of this* $\mathbb{V}$*-ultra triple;*
- *a map* $w : E \to \mathbb{V}$*;*
- *a map* $d : E \underline{\times} E \to \mathbb{V}$ *required to satisfy the following axioms:*
  - ***Symmetry:*** *We have* $d(a,b) = d(b,a)$ *for any two distinct elements* a *and* b *of* E*.*
  - ***Ultrametric triangle inequality:*** *We have* $d(a,b) \leqslant \max\{d(a,c), d(b,c)\}$ *for any three distinct elements* a*,* b *and* c *of* E*.*

If (E, w, d) *is a* $\mathbb{V}$*-ultra triple and* $e \in E$*, then the value* $w(e) \in \mathbb{V}$ *is called the* weight *of* e*.*

If (E, w, d) *is a* $\mathbb{V}$*-ultra triple and* e *and* f *are two distinct elements of* E*, then the value* $d(e, f) \in \mathbb{V}$ *is called the* distance *between* e *and* f*.*

Example 3.4. *Let* E *be a subset of* $\mathbb{Z}$*. Let* m *be any integer. Let* $\varepsilon, \alpha \in \mathbb{V}$ *satisfy* $\varepsilon \leqslant \alpha$*. Define a map* $w : E \to \mathbb{V}$ *arbitrarily. Define a map* $d : E \underline{\times} E \to \mathbb{V}$ *by*

$$d(a,b) = \begin{cases} \varepsilon, & \text{if } a \equiv b \mod m; \\ \alpha, & \text{if } a \not\equiv b \mod m \end{cases} \qquad \text{for all } (a,b) \in E \underline{\times} E.$$

It is easy to see that (E, w, d) *is a* $\mathbb{V}$*-ultra triple.*

Example 3.5. *For this example, we set* $\mathbb{V} = \mathbb{Z}$*, we fix a prime number* p *and a subset* E *of* $\mathbb{Z}$*, and we define a map* $d' : E \underline{\times} E \to \mathbb{Z}$ *by setting*

$$d'(a,b) = -v_p(a-b) \qquad \text{for all } (a,b) \in E \underline{\times} E.$$

Here, for any nonzero $m \in \mathbb{Z}$*, we let* $v_p(m)$ *denote the* p-adic valuation *of* m *(that is, the largest nonnegative integer* k *such that* $p^k \mid m$*). We define the weights* $w(a) \in \mathbb{Z}$ *arbitrarily. Then,* (E, w, d) *is an* $\mathbb{V}$*-ultra triple.*

Another example of a $\mathbb{V}$-ultra triple originates in graph theory:

Example 3.6. *Let* T *be a tree (i.e., a connected finite undirected graph that has no cycles). For each edge* e *of* T*, fix some nonnegative* $\lambda(e) \in \mathbb{V}$*; this* $\lambda(e)$ *will be called the* weight *of* e*.*

For any two vertices u *and* v *of* T*, let* $\lambda(u,v)$ *denote the sum of the weights of all edges on the (unique) path from* u *to* v*. This* $\lambda(u,v)$ *generalizes the usual (graph-theoretical) distance between* u *and* v *(which is obtained when* $\lambda(e) = 1$ *for each edge* e*).*

Let E *be a subset of the vertex set of* T*. Define a map* $w : E \to \mathbb{V}$ *arbitrarily. Fix any vertex* r *of* T*, and define a map* $d : E \underline{\times} E \to \mathbb{V}$ *by setting*

$$d(x,y) = \lambda(x,y) - \lambda(x,r) - \lambda(y,r) \qquad \text{for each } (x,y) \in E \underline{\times} E.$$

Then, (E, w, d) *is a* $\mathbb{V}$*-ultra triple.*

The notion of a $\mathbb{V}$-ultra triple generalizes the notion of an ultra triple as defined in [8] (where further examples are given). More precisely, if $\mathbb{V} = \mathbb{R}$, then a $\mathbb{V}$-ultra triple is the same as what is called an "ultra triple" in [8]. All the definitions and results stated in [8] for ultra triples adapt easily to the more general setting of $\mathbb{V}$-ultra triples.

Definition 3.7. *Let (E, w, d) be a $\mathbb{V}$-ultra triple. Let A be a finite subset of E. Then, the* perimeter *of A (with respect to (E, w, d)) is defined to be*

$$\operatorname{PER}(A) := \sum_{a \in A} w(a) + \sum_{\substack{\{a,b\} \subseteq A; \\ a \neq b}} d(a, b) \in \mathbb{V}.$$

(Here, the second sum ranges over all ***unordered*** *pairs $\{a, b\}$ of distinct elements of A.)*

For example, if $A = \{p, q, r\}$ is a 3-element set, then

$$\operatorname{PER}(A) = w(p) + w(q) + w(r) + d(p, q) + d(p, r) + d(q, r).$$

Definition 3.8. *Let (E, w, d) be a $\mathbb{V}$-ultra triple. The* Bhargava greedoid *of (E, w, d) is defined to be the set*

$$\begin{aligned} &\{A \subseteq E \mid A \text{ has maximum perimeter among all } |A| \text{-element subsets of } E\} \\ &= \{A \subseteq E \mid \operatorname{PER}(A) \geqslant \operatorname{PER}(B) \text{ for all } B \subseteq E \text{ satisfying } |B| = |A|\}. \end{aligned}$$

4 The main theorem

Convention 4.1. *For this whole section, we fix a $\mathbb{V}$-ultra triple (E, w, d) such that E is finite.*

Our main result is the following:

Theorem 4.2. *Let $\mathcal{F}$ be the Bhargava greedoid of (E, w, d). Let $\mathbb{K}$ be a field of size $|\mathbb{K}| \geqslant |E|$. Then, $\mathcal{F}$ is the Gaussian elimination greedoid of a vector family over $\mathbb{K}$.*

We shall prove a stronger fact, which will however require some more terminology:

Definition 4.3. *Let $\alpha \in \mathbb{V}$. An α-*clique *of (E, w, d) will mean a subset F of E such that any two distinct elements $a, b \in F$ satisfy $d(a, b) = \alpha$.*

Definition 4.4. *A* clique *of (E, w, d) will mean a subset of E that is an α-clique for some $\alpha \in \mathbb{V}$.*

Thus, any subset of E having size $\leqslant 1$ is a clique (and an α-clique for every $\alpha \in \mathbb{V}$).

Definition 4.5. *Let* $\operatorname{mcs}(E, w, d)$ *denote the maximum size of a clique of (E, w, d).*

Clearly, $\operatorname{mcs}(E, w, d) \leqslant |E|$. Thus, the following theorem generalizes Theorem 4.2:

Theorem 4.6. *Let $\mathcal{F}$ be the Bhargava greedoid of (E, w, d). Let $\mathbb{K}$ be a field of size $|\mathbb{K}| \geqslant$* $\operatorname{mcs}(E, w, d)$. *Then, $\mathcal{F}$ is the Gaussian elimination greedoid of a vector family over $\mathbb{K}$.*

5 Greedy permutations

Convention 5.1. *For this whole section, we fix a* $\mathbb{V}$*-ultra triple* (E, w, d).

The notion of a greedoid owes its name to the phenomenon that (usually) when the sets solving an optimization problem form a greedoid, then they can be constructed efficiently by a greedy algorithm (see [9] for some instances). The Bhargava greedoid is no exception. Given a $\mathbb{V}$-ultra triple (E, w, d) and a $k \in \mathbb{N}$, the greedy algorithm for constructing a k-element subset of E of maximum perimeter works as one would expect: Start with the empty set, and add elements to it one by one, each time choosing an element that maximizes the increment it contributes to the perimeter. This leads to sequences we call *greedy permutations*:

Definition 5.2. *Let* C *be a subset of* E*, and let* $m \in \mathbb{N}$*. A* greedy m-permutation *of* C *is a list* $(c_1, c_2, \ldots, c_m)$ *of* m *distinct elements of* C *such that for each* $i \in \{1, 2, \ldots, m\}$ *and each* $x \in C \setminus \{c_1, c_2, \ldots, c_{i-1}\}$*, we have* $\operatorname{PER}\{c_1, c_2, \ldots, c_i\} \geqslant \operatorname{PER}\{c_1, c_2, \ldots, c_{i-1}, x\}$.

Example 5.3. *Let* $m = 2$ *and* $E = \{1, 2, 3, 4, 5\}$*. Let* $\mathbb{V}$*,* ε*,* α *and* d *be as in Example 3.4. Define* $w : E \to \mathbb{V}$ *such that* $w(a) = 0$ *for all* $a \in E$.

Then, a pair (i, j) *of elements of* E *is a greedy* 2*-permutation of* E *if and only if* $i \not\equiv j \mod 2$.

The pair $(1, 3)$ *is a greedy* 2*-permutation of* $\{1, 3, 5\}$*, but not of* E.

Also, $(1, 2, 3, 4, 5)$ *is a greedy* 5*-permutation of* E*, but* $(1, 2, 3, 5, 4)$ *is not.*

If C is a finite subset of E, then a greedy m-permutation of C exists for each $m \in \{0, 1, \ldots, |C|\}$. Moreover, the greedy m-permutations generate all maximum-perimeter subsets of E, in the sense that the following theorems ([8, §4]) hold:

Theorem 5.4. *Let* C *be any subset of* E*, and let* m *be a nonnegative integer.*

Let $(c_1, c_2, \ldots, c_m)$ *be any greedy* m*-permutation of* C.

Then, for each $k \in \{0, 1, \ldots, m\}$*, the set* $\{c_1, c_2, \ldots, c_k\}$ *has maximum perimeter among all* k*-element subsets of* C *(and thus belongs to the Bhargava greedoid of* (E, w, d)*).*

Theorem 5.5. *Let* C *be any finite subset of* E*, and let* m *be a nonnegative integer such that* $|C| \geqslant m$*. Let* $k \in \{0, 1, \ldots, m\}$.

Let A *be a* k*-element subset of* C *having maximum perimeter (among the* k*-element subsets of* C*). Then, there exists a greedy* m*-permutation* $(v_1, v_2, \ldots, v_m)$ *of* C *such that* $A = \{v_1, v_2, \ldots, v_k\}$.

6 The connection to Bhargava's p-orderings

Greedy m-permutations generalize the notion of P-orderings introduced by Manjul Bhargava in [2, Section 2] (which, in turn, has the p-orderings in [3, Section 4] and [1, Section

2] as particular cases). Bhargava already observed the possibility of such a generalization in the paragraph after the proof of Lemma 2 in [2], but never elaborated further on it.

In this abstract, we merely outline the simplest case, in which the underlying Dedekind ring is $\mathbb{Z}$. We refer to [8, §9] for the general setting.

Let $\mathbb{V}$, p, E, $v_p(m)$ and d' be as in Example 3.5. Define a map $w : E \to \mathbb{Z}$ by setting $w(e) = 0$ for each $e \in E$.

Bhargava defines a *P-ordering* of E to be a sequence $(a_0, a_1, a_2, \ldots)$ of elements of E constructed recursively as follows: For each $k \in \mathbb{N}$, we define a_k (assuming that $a_0, a_1, \ldots, a_{k-1}$ are already defined) to be an element of E minimizing the quantity

$$v_p((a_k - a_0)(a_k - a_1) \cdots (a_k - a_{k-1}))$$

(where we set $v_p(0) = +\infty$ and consider $+\infty$ to be larger than any integer). In order not to be distracted by the infiniteness of such sequences, we instead consider a finite analogue of P-orderings, which we call *(P, m)-orderings* (for some $m \in \mathbb{N}$); these are defined just as P-orderings, except for being m-tuples instead of infinite sequences.

Now, it is easy to see that if C is a subset of E such that $|C| \geqslant m$, then a (P, m)-ordering of C is the same as a greedy m-permutation of C (with respect to the $\mathbb{V}$-ultra triple (E, w, d')). Thus, [2, Theorem 1 and Lemma 2] can easily be derived from the above Theorems 5.4 and 5.5.

The "P-orderings of order h" defined in [1, Section 2.2] can also be regarded as a particular case of greedy m-permutations; we only need to modify the map d'.

7 Greedy subsequences

Greedy m-permutations of a subset of a $\mathbb{V}$-ultra triple consist of distinct elements by definition. One might wonder if there is an analogue that allows equal entries. Of course, this requires having distances of the form $d(a, a)$, which are not supported by the notion of a $\mathbb{V}$-ultra triple; thus, we need a somewhat stronger notion:

Definition 7.1. *A* full $\mathbb{V}$-ultra triple *shall mean a triple (E, w, d) consisting of:*

- *a set E, called the* ground set *of this full $\mathbb{V}$-ultra triple;*
- *a map $w : E \to \mathbb{V}$;*
- *a map $d : E \times E \to \mathbb{V}$ required to satisfy the following axioms:*
  - ***Symmetry:*** *We have $d(a, b) = d(b, a)$ for any two elements a and b of E.*
  - ***Ultrametric triangle inequality:*** *We have $d(a, b) \leqslant \max\{d(a, c), d(b, c)\}$ for any three elements a, b and c of E.*

If (E, w, d) is a $\mathbb{V}$-ultra triple such that E is finite, then it is easy to see that d can be extended to a map $\overline{d} : E \times E \to \mathbb{V}$ such that $\left(E, w, \overline{d}\right)$ is a full $\mathbb{V}$-ultra triple. For infinite E, such an extension may or may not exist.

Convention 7.2. *For the rest of this section, we fix a full $\mathbb{V}$-ultra triple (E, w, d).*

We shall not consider multisets, but instead simply work with tuples: The *perimeter* of any m-tuple $\mathbf{a} = (a_1, a_2, \ldots, a_m) \in E^m$ is defined as

$$\operatorname{PER}(\mathbf{a}) := \sum_{k=1}^{m} w(a_k) + \sum_{1 \leqslant i < j \leqslant m} d(a_i, a_j).$$

In analogy to Definition 5.2, we now define:

Definition 7.3. *Let C be a subset of E, and let $m \in \mathbb{N}$. A* greedy m-subsequence *of C is an m-tuple $(c_1, c_2, \ldots, c_m) \in C^m$ such that for each $i \in \{1, 2, \ldots, m\}$ and each $x \in C$, we have $\operatorname{PER}(c_1, c_2, \ldots, c_i) \geqslant \operatorname{PER}(c_1, c_2, \ldots, c_{i-1}, x)$.*

Thus, a greedy m-subsequence may revisit elements of C multiple times (but will generally try to avoid doing so as long as it can, since the ultrametric triangle inequality entails $d(a, a) \leqslant d(a, c)$ for any $a, c \in E$). We now have the following analogues of Theorems 5.4 and 5.5 (see [8, §8]):

Theorem 7.4. *Let C be any subset of E, and let m be a nonnegative integer.*

Let $(c_1, c_2, \ldots, c_m)$ be any greedy m-subsequence of C.

Then, for each $k \in \{0, 1, \ldots, m\}$, the k-subsequence $(c_1, c_2, \ldots, c_k)$ has maximum perimeter among all k-subsequences of C.

Theorem 7.5. *Let C be any finite nonempty subset of E, and let m be a nonnegative integer. Let $k \in \{0, 1, \ldots, m\}$.*

Let $\mathbf{a}$ be any k-subsequence of C with maximum perimeter. Then, there exists a greedy m-subsequence $(c_1, c_2, \ldots, c_m)$ of C such that $\mathbf{a}$ is a permutation of the k-tuple $(c_1, c_2, \ldots, c_k)$.

8 Proof outline for Theorem 4.6

Let us briefly survey how Theorem 4.6 is proved; for details, we refer to [7, §§4–9]. The proofs of the combinatorial results can be found in [8].

The proof relies on constructing a rather wide but manageable class of $\mathbb{V}$-ultra triples and proving that their Bhargava greedoids are Gaussian elimination greedoids, and subsequently showing that every finite $\mathbb{V}$-ultra triple is isomorphic to one in this class. We call these the *valadic* $\mathbb{V}$-ultra triples.

To define them, we proceed as follows. We fix a field $\mathbb{K}$. Let $\mathbb{L}$ denote the group algebra $\mathbb{K}[\mathbb{V}]$ of the group $\mathbb{V}$ over $\mathbb{K}$. This is a free $\mathbb{K}$-module with basis $(t_\alpha)_{\alpha\in\mathbb{V}}$; it becomes a commutative $\mathbb{K}$-algebra with unity t_0 and with multiplication given by $t_\alpha t_\beta = t_{\alpha+\beta}$ for all $\alpha, \beta \in \mathbb{V}$. Let $\mathbb{V}_{\geqslant 0} = \{\alpha \in \mathbb{V} \mid \alpha \geqslant 0\}$; this is a submonoid of the group $\mathbb{V}$. Let $\mathbb{L}_+$ be the span of the t_α with $\alpha \in \mathbb{V}_{\geqslant 0}$; this is a $\mathbb{K}$-subalgebra of $\mathbb{L}$.

Example 8.1. *If $\mathbb{V} = \mathbb{Z}$ (with the usual addition and total order), then $\mathbb{V}_{\geqslant 0} = \mathbb{N}$. In this case, the group algebra $\mathbb{L}$ is the Laurent polynomial ring $\mathbb{K}\left[X, X^{-1}\right]$ in a single indeterminate X over $\mathbb{K}$ (indeed, each t_α plays the role of X^α), and its subalgebra $\mathbb{L}_+$ is the polynomial ring $\mathbb{K}[X]$.*

If $a \in \mathbb{L}$ is nonzero, then the *order* of a is defined to be the smallest $\beta \in \mathbb{V}$ such that the coefficient of t_β in a is nonzero. This order is an element of $\mathbb{V}$, and is denoted by $\operatorname{ord} a$. For example, $\operatorname{ord}(t_2 - t_3 + 5t_6) = 2$ (if $\mathbb{V} = \mathbb{Z}$).

The following properties of orders are straightforward to check:

Lemma 8.2. ***(a)*** *A nonzero element $a \in \mathbb{L}$ belongs to $\mathbb{L}_+$ if and only if its order $\operatorname{ord} a$ is nonnegative (i.e., we have $\operatorname{ord} a \geqslant 0$).*

(b) *We have $\operatorname{ord}(-a) = \operatorname{ord} a$ for any nonzero $a \in \mathbb{L}$.*

(c) *Let a and b be two nonzero elements of $\mathbb{L}$. Then, ab is nonzero and satisfies $\operatorname{ord}(ab) = \operatorname{ord} a + \operatorname{ord} b$. (In particular, $\mathbb{L}$ is an integral domain.)*

(d) *Let a and b be two nonzero elements of $\mathbb{L}$ such that $a + b$ is nonzero. Then, $\operatorname{ord}(a+b) \geqslant \min\{\operatorname{ord} a, \operatorname{ord} b\}$.*

We can now assign a $\mathbb{V}$-ultra triple to each subset of $\mathbb{L}$:

Definition 8.3. *Let E be a subset of $\mathbb{L}$. Define a map $d : E \underline{\times} E \to \mathbb{V}$ by setting*

$$d(a,b) = -\operatorname{ord}(a-b) \qquad \text{for all } (a,b) \in E \underline{\times} E.$$

Then, (E, w, d) is a $\mathbb{V}$-ultra triple whenever $w : E \to \mathbb{V}$ is a map (this follows easily from Lemma 8.2). Such a $\mathbb{V}$-ultra triple (E, w, d) will be called valadic.

Now, the first half of the proof of Theorem 4.6 is showing the theorem for valadic $\mathbb{V}$-ultra triples:

Theorem 8.4. *Let E be a finite subset of $\mathbb{L}$. Define d as in Definition 8.3. Let $w : E \to \mathbb{V}$ be a map. Then, the Bhargava greedoid of the $\mathbb{V}$-ultra triple (E, w, d) is the Gaussian elimination greedoid of a vector family over $\mathbb{K}$.*

The main tool in the proof of this theorem is a slightly generalized Vandermonde determinant:

Lemma 8.5. *Let R be a commutative ring. For each $j \in \{1, 2, \ldots, m\}$, let $f_j \in R[X]$ be a monic polynomial of degree $j-1$. Let $u_1, u_2, \ldots, u_m$ be m elements of R. Then,*

$$\det\left(\left(f_j(u_i)\right)_{1\leqslant i\leqslant m,\ 1\leqslant j\leqslant m}\right) = \prod_{m\geqslant i>j\geqslant 1} (u_i - u_j).$$

Here, the notation $(b_{i,j})_{1 \leqslant i \leqslant p,\ 1 \leqslant j \leqslant q}$ denotes the $p \times q$-matrix whose (i,j)-th entry is $b_{i,j}$ for all i and j.

Proof of Theorem 8.4 (outline). Let $m = |E|$. Let $(c_1, c_2, \ldots, c_m)$ be a greedy m-permutation of E. For each $k \in \{1, 2, \ldots, m\}$, define a $\rho_k \in \mathbb{V}$ by

$$\rho_k = w(c_k) + \sum_{i=1}^{k-1} d(c_i, c_k) = \operatorname{PER}\{c_1, c_2, \ldots, c_k\} - \operatorname{PER}\{c_1, c_2, \ldots, c_{k-1}\}.$$

Consider the polynomial ring $\mathbb{L}[X]$. For each $j \in \{1, 2, \ldots, m\}$, we define a monic polynomial $f_j \in \mathbb{L}[X]$ by $f_j = (X - c_1)(X - c_2) \cdots (X - c_{j-1})$. It is easy to see (using the definition of a greedy m-permutation) that

$$t_{\rho_j - w(e)} f_j(e) \in \mathbb{L}_+ \qquad \text{for every } e \in E \text{ and } j \in \{1, 2, \ldots, m\}.$$

This allows us to define an $a(e,j) \in \mathbb{L}_+$ by $a(e,j) = t_{\rho_j - w(e)} f_j(e)$ for every $e \in E$ and $j \in \{1, 2, \ldots, m\}$. Let $\pi : \mathbb{L}_+ \to \mathbb{K}$ be the $\mathbb{K}$-algebra homomorphism that sends each $x \in \mathbb{L}_+$ to the coefficient of t_0 in x. For each $e \in E$, define a column vector $v_e \in \mathbb{K}^m$ by

$$v_e = (\pi(a(e,1)), \pi(a(e,2)), \ldots, \pi(a(e,m)))^T.$$

Using Lemma 8.5, we can see that $\mathcal{G}$ is the Gaussian elimination greedoid of the vector family $(v_e)_{e \in E}$. □

The second ingredient in the proof of Theorem 4.6 is a theorem stating that if (E, w, d) is a $\mathbb{V}$-ultra triple with finite ground set E, and if $\mathbb{K}$ is a field of size $|\mathbb{K}| \geqslant \operatorname{mcs}(E, w, d)$, then (E, w, d) is isomorphic[2] to a valadic $\mathbb{V}$-ultra triple. This is proved by strong induction on $|E|$, where the induction step involves identifying a hierachical structure in $\mathbb{V}$-ultra triples: If (E, w, d) is a $\mathbb{V}$-ultra triple with $1 < |E| < \infty$, then we can define the *diameter* of (E, w, d) to be $\alpha = \max(d(E \underline{\times} E))$. If we now pick a maximum-size α-clique $\{e_1, e_2, \ldots, e_m\}$ of (E, w, d), then the balls $B^\circ_\alpha(e_1), B^\circ_\alpha(e_2), \ldots, B^\circ_\alpha(e_m)$ (where $B^\circ_\alpha(f) := \{g \in E \mid f = g \text{ or } d(f,g) < \alpha\}$ is an open ball as defined in metric space theory) are disjoint and constitute a set partition of E, and we can restrict our $\mathbb{V}$-ultra triple to any of these balls to obtain a smaller $\mathbb{V}$-ultra triple with diameter smaller than α. By the induction hypothesis, each of these smaller $\mathbb{V}$-ultra triples is isomorphic to a valadic one; by appropriately shifting the ground sets of the latter valadic $\mathbb{V}$-ultra triples and taking their union, we then obtain a valadic $\mathbb{V}$-ultra triple isomorphic to (E, w, d).

[2] Isomorphism between $\mathbb{V}$-ultra triples is defined in the natural way (i.e., there must be a bijection that forms commutative triangles with w and d).

9 A phylogenetic analogue

In Example 3.6, we have constructed a $\mathbb{V}$-ultra triple (E, w, d) for any tree T, any vertex r of T, any weight function $\lambda : \{\text{edges of } T\} \to \mathbb{V}$ and any map $w : E \to \mathbb{V}$. A particularly prominent instance of this kind of data is when T is an *edge-weighted phylogenetic tree* (one of the basic concepts in mathematical biology – see, e.g., [11]), the vertex r is its root, while E is the set of its leaves. In this case, we can regard the perimeter $\text{PER}(A)$ of a subset A of E to be a measure of "broadness" of A. The sets in the Bhargava greedoid of (E, w, d) thus are the optimal solutions to a "Noah's ark" type problem: how do we choose a maximally representative (i.e., maximally broad) small sample from the taxa (i.e., the leaves of our phylogenetic tree)? (The weights $w(e)$ can come useful if some taxa are more desirable than others.)

It turns out that a similar problem has been considered in the biology literature ([10] and references therein). The main difference is that biologists define the *phylogenetic diversity* of a subset A of E to be the sum of the weights of all edges on the minimal subtree connecting all leaves in A. This is not the same as $\text{PER}(A)$, and indeed behaves somewhat differently; for example, for a star-shaped tree, the phylogenetic diversity grows linearly in the size of A, while the perimeter grows quadratically. Nevertheless, the two concepts are sufficiently similar that the maximum-diversity subsets (in the biologists' sense) also form a strong greedoid; this was proved by Moulton, Semple and Steel in 2006 [10, Theorem 3.2]. (The two greedoids are generally different, and we don't know whether the latter is a Gaussian elimination greedoid as well.)

Let us finally remark that the assignment of $\mathbb{V}$-ultra triples to edge-weighted phylogenetic trees is not a one-way road. There is a construction that (to some extent) inverts it: If (E, w, d) is a $\mathbb{V}$-ultra triple with finite ground set E, then the iterative disassembling procedure we sketched at the end of Section 8 (which begins by splitting E into a disjoint union of balls, and then proceeds recursively by applying the same construction to these balls) produces a tree which, if appropriately edge-weighted, encodes the $\mathbb{V}$-ultra triple (E, w, d). If we apply the construction from Example 3.6 to it, we recover an isomorphic copy of (E, w, d). The two transformations (triple to tree and backwards) are not literally inverse, and the precise connection is yet to be explored.

10 Further questions

Several questions are left to be answered:

1. Is the bound $|\mathbb{K}| \geqslant \text{mcs}(E, w, d)$ in Theorem 4.6 optimal? (See [7, §4] for details.)

2. Does some multiset analogue of Bhargava greedoids underlie Theorems 7.4 and 7.5?

3. Is the Moulton–Semple–Steel greedoid [10] a Gaussian elimination greedoid?

4. Both the Bhargava greedoid and the Moulton–Semple–Steel greedoid are obtained by picking the subsets of E maximizing a statistic (perimeter or phylogenetic diversity). Is there a more general class of statistics defined on subsets of E whose maximizers form a greedoid?

Acknowledgements

The first author thanks the Mathematisches Forschungsinstitut Oberwolfach for its hospitality. He also thanks Alexander Postnikov for enlightening conversations.

References

[1] M. Bhargava. "On P-orderings, rings of integer-valued polynomials, and ultrametric analysis". In: *J. Amer. Math. Soc.* 22 (4 2009), pp. 963–993.

[2] M. Bhargava. "P-orderings and polynomial functions on arbitrary subsets of Dedekind rings". In: *J. reine. angew. Math.* 490 (1997), pp. 101–127.

[3] M. Bhargava. "The Factorial Function and Generalizations". In: *The American Mathematical Monthly* 107 (9 2000), pp. 783–799. DOI: 10.2307/2695734.

[4] A. Björner, B. Korte, and L. Lovász. "Homotopy properties of greedoids". In: *Adv. in Appl. Math.* 6 (4 1985), pp. 447–494. DOI: 10.1016/0196-8858(85)90021-1.

[5] A. Björner, L. Lovász, and P.W. Shor. "Chip-firing Games on Graphs". In: *European Journal of Combinatorics* 12 (4 1991), pp. 283–291. DOI: 10.1016/S0195-6698(13)80111-4. URL: https://doi.org/10.1016/S0195-6698(13)80111-4.

[6] V. Bryant and I. Sharpe. "Gaussian, Strong and Transversal Greedoids". In: *European Journal of Combinatorics* 20.4 (1999), pp. 259–262. DOI: 10.1006/eujc.1998.0287.

[7] D. Grinberg. *The Bhargava greedoid as a Gaussian elimination greedoid*. 2019. URL: http://www.cip.ifi.lmu.de/~grinberg/algebra/greedrepv2.pdf.

[8] D. Grinberg and F. Petrov. *A greedoid and a matroid inspired by Bhargava's p-orderings*. 2019. arXiv: 1909.01965 [math.CO].

[9] B. Korte, L. Lovász, and R. Schrader. *Greedoids*. Algorithms and Combinatorics #4. Springer, 1991. DOI: 10.1007/978-3-642-58191-5.

[10] V. Moulton, C. Semple, and M. Steel. "Optimizing phylogenetic diversity under constraints". In: *Journal of Theoretical Biology* 246 (2007), pp. 186–194. DOI: 10.1016/j.jtbi.2006.12.021.

[11] T. Warnow. *Computational Phylogenetics*. Cambridge University Press, 2018.