More Vector Spaces with Atoms of Finite Lengths

Jingjie Yang
University of Oxford

Mikołaj Bojańczyk
University of Warsaw

Bartek Klin
University of Oxford

***Abstract*—*CRITICAL: Do Not Use Symbols, Special Characters, Footnotes, or Math in Paper Title or Abstract.**

I. Introduction

II. Rado graph, sans cogs

$\mathbb{A}$ is:

- oligomorphic if, for $k = 0, 1, 2, \ldots$, $\mathbb{A}^k$ only has finitely many orbits;
- $\mathfrak{F}$-oligomorphic if, for $k = 0, 1, 2, \ldots$, $\mathrm{Lin}_{\mathfrak{F}}\, \mathbb{A}^k$ only has finitely long chains.

Of note: with Stirling numbers of the second kind and Gaussian 2-binomial coefficients, the orbit counts are given by

$$\#\mathbb{N}^k = \sum_{d=0}^{k} \left\{ {k \atop d} \right\}$$

$$\#\mathbb{Q}^k = \sum_{d=0}^{k} \left\{ {k \atop d} \right\} d!$$

$$\#\mathbb{G}^k = \sum_{d=0}^{k} \left\{ {k \atop d} \right\} 2^{\binom{d}{2}}$$

$$\#\mathbb{V}_\infty^k = \sum_{d=0}^{k} \begin{bmatrix} k \\ d \end{bmatrix}_2$$

$$\#\mathbb{W}_\infty^k = \sum_{d=0}^{k} \begin{bmatrix} k \\ d \end{bmatrix}_2 2^{\binom{d}{2}}$$

To introduce:

- *smooth approximation* by *homogeneous substructures* [2] (N.B. 'smooth approximation' from [3, Definition 4] seems to be entirely different)
- *rough approximation* of a homogeneous structure by finite substructures with few orbits (i.e., types) that cover the age of $\mathbb{A}$

A. Symplectic vector spaces

Throughout this subsection let $\mathfrak{f}$ denote a finite field.

Definition II.1. A *symplectic vector space* is an $\mathfrak{f}$-vector space $\mathbb{W}$ equipped with a bilinear form $\omega : \mathbb{W} \times \mathbb{W} \to \mathfrak{f}$ that is

- alternating: $\omega(v, v) = 0$ for all v; and
- non-degenerate: if $\omega(v, w) = 0$ for all w then $v = 0$.

Example II.2. Let $\mathbb{W}_n$ be the $\mathfrak{f}$-vector space with basis $e_1, \ldots, e_n, f_1, \ldots, f_n$. Define ω by bilinearly extending

$$\omega(e_i, f_i) = 1 = -\omega(f_i, e_i), \quad \omega(-, *) = 0 \text{ elsewhere}; \qquad (\S)$$

one may straightforwardly check that ω is alternating and non-degenerate. Moreover, noticing that $\mathbb{W}_0 \subseteq \mathbb{W}_1 \subseteq \mathbb{W}_2 \subseteq \cdots$, we obtain a countable-dimensional symplectic vector space $\mathbb{W}_\infty = \bigcup_n \mathbb{W}_n$.

We will refer to vectors satisfying (§) as a *symplectic basis* — indeed, they must be linearly independent. Such bases behave very much like the usual bases.

Proposition II.3. *Assume that $\mathbb{W}$ is a symplectic vector space that is at most countable. Then any finite symplectic basis $e_1, \ldots, e_n, f_1, \ldots, f_n$ can be extended to a symplectic basis that spans the whole $\mathbb{W}$.*

Proof. Suppose that $e_1, \ldots, e_n, f_1, \ldots, f_n$ does not already span $\mathbb{W}$; take v to be a witness (that is least according to some fixed enumeration of $\mathbb{W}$ in the case it is infinite). Put

$$e_{n+1} = v - \sum_{i=1}^{n} \omega(e_i, v) f_i + \sum_{i=1}^{n} \omega(f_i, v) e_i$$

so that $\omega(e_i, e_{n+1}) = 0 = \omega(f_i, e_{n+1})$. This cannot be the zero vector lest we contradict the choice of v. By the non-degeneracy of ω, there is — rescaling if necessary — some w such that $\omega(e_{n+1}, w) = 1$. Now define

$$f_{n+1} = w - \sum_{i=1^n} \omega(e_i, w) f_i + \sum_{i=1}^{n} \omega(f_i, w) e_i$$

in a similar manner, making $e_1, \ldots, e_n, e_{n+1}, f_1, \ldots, f_n, f_{n+1}$ a symplectic basis that spans v. We go through every element of $\mathbb{W}$ by continuing this way. ☐

THROW IN THE APPENDIX:

In fact, we will also make use of the "symplectic basis and a half" variant below.

Proposition II.4. *Now assume $\mathbb{W}$ is a finite-dimensional symplectic vector space. Let*

$$\begin{aligned} &e_1, \cdots, e_n, e_{n+1}, \cdots, e_{n+k}, \\ &f_1, \cdots, f_n \end{aligned}$$

be linearly independent vectors satisfying (§). *Then we can find the missing $f_{n+1}, \cdots, f_{n+k}$ to complete the symplectic basis.*

Proof. We first need the following notion. Given a subspace $V \subseteq \mathbb{W}$, consider its orthogonal complement

$$V^\perp = \{ w \in \mathbb{W} \mid \forall v \in V : \omega(v, w) = 0 \}.$$

It is the kernel of the composite linear map

$$\begin{aligned}\mathbb{W} &\to (\mathbb{W} \xrightarrow{\text{lin.}} \mathfrak{f}) \to (V \xrightarrow{\text{lin.}} \mathfrak{f})\\ w &\mapsto \omega(-, w) \mapsto \omega(-, w)|_V.\end{aligned}$$

Note this map is surjective: the first part is injective by non-degeneracy and hence surjective for dimension reasons, and the second part is surjective since we can extend a basis of V to one of $\mathbb{W}$. Therefore

$$\dim V^{\perp} = \dim \mathbb{W} - \dim V,$$

and in particular $V^{\perp\perp}$ is precisely equal to V.

Now suppose we have found $f_{n+1}, \cdots, f_{n+i}$ already. If e_{n+i+1} were to be spanned by

$$\begin{aligned}&e_1, \cdots, e_{n+i}, \cancel{e_{n+i+1}}, e_{n+i+2}, \cdots, e_{n+k},\\ &f_1, \cdots, f_{n+i},\end{aligned}$$

it would be spanned by $e_{n+i+2}, \cdots, e_{n+k}$ alone because of (§); but this is impossible as we assumed linear independency. So

$$\begin{aligned}e_{n+i+1} &\notin \langle e_1, \cdots, e_{n+i}, e_{n+i+2}, f_1, \cdots, f_{n+i}\rangle\\ &= \langle e_1, \cdots, e_{n+i}, e_{n+i+2}, f_1, \cdots, f_{n+i}\rangle^{\perp\perp},\end{aligned}$$

i.e., some $f_{n+i+1} \in \langle e_1, \cdots, e_{n+i}, e_{n+i+2}, f_1, \cdots, f_{n+i}\rangle^{\perp}$ satisfies $\omega(e_{n+i+1}, f_{n+i+1}) = 1$. □

Given two symplectic vector spaces $\mathbb{W}$ and $\mathbb{W}'$, we call a function α between $X \subseteq \mathbb{W}$ and $X' \subseteq \mathbb{W}'$ *isometric* if $\omega(\alpha(x_1), \alpha(x_2)) = \omega(x_1, x_2)$ for all $x_1, x_2 \in X$. We can make an easy observation:

Lemma II.5. *Let $\{e_i, f_i \mid i \in I\} \subseteq \mathbb{W}$, $\{e'_j, f'_j \mid j \in J\} \subseteq \mathbb{W}'$ be two symplectic bases and let $\alpha : I \to J$ be a bijection. Then*

$$e_i \mapsto e'_{\alpha(i)}, f_i \mapsto f'_{\alpha(i)}$$

defines an isometric linear isomorphism $\langle e_i, f_i\rangle \to \langle e'_j, f'_j\rangle$.

It then follows from Proposition II.3 that, up to isometric linear isomorphisms, $\mathbb{W}_0, \mathbb{W}_1, \mathbb{W}_2, \cdots, \mathbb{W}_\infty$ are all the countable symplectic vector spaces. Whilst we may deduce that $\mathbb{W}_\infty$ is oligomorphic by appealing to Ryll-Nardzewski, we will opt for a more direct proof that also establishes smooth approximation.

Proposition II.6 (Witt Extension)**.** *Any isometric linear injection $\alpha : \langle X\rangle \subseteq \mathbb{W}_n \to \mathbb{W}_n$ can be extended to an isometric linear automorphism of $\mathbb{W}_n$ and in turn to one of $\mathbb{W}_\infty$.*

Proof. To begin with, find a basis $x_1, \cdots, x_k$ for $\langle X\rangle^{\perp} = \{w \in W \mid \forall x \in X : \omega(w, x) = 0\}$ and extend it to a basis $x_1, \cdots, x_k, x_{k+1}, \cdots, x_d$ for $\langle X\rangle$. Notice that

$$U = \langle x_{k+1}, \cdots, x_d\rangle$$

must be a symplectic subspace: as it intersects with $\langle X\rangle^{\perp}$ trivially, given any non-zero vector $u \in U$ we must have $0 \neq \omega(u, x + u') = \omega(u, u')$ for some $x \in \langle X\rangle^{\perp}$ and $u' \in U$. Hence use Proposition II.3 to find a symplectic basis $e_1, \cdots, e_n, f_1, \cdots, f_n$ for U. Observe that

$$\begin{aligned}&e_1, \cdots, e_n, x_1, \cdots, x_k,\\ &f_1, \cdots, f_n\end{aligned}$$

form a basis for $\langle X\rangle$ and satisfy (§). On the other hand,

$$\begin{aligned}&\alpha(e_1), \cdots, \alpha(e_n), \alpha(x_1), \cdots, \alpha(x_k),\\ &\alpha(f_1), \cdots, \alpha(f_n)\end{aligned}$$

form a basis for $\alpha(\langle X\rangle)$ and also satisfy (§). Therefore apply Proposition II.4 twice to find the missing $y_1, \cdots, y_k$ and $y'_1, \cdots, y'_k$ to complete the two symplectic bases — call them $\mathcal{B}$ and $\mathcal{B}'$. They are of the same size.

Now, by using Proposition II.3, extend $\mathcal{B}$ and $\mathcal{B}'$ to symplectic bases $\mathcal{C}$ and $\mathcal{C}'$ that span $\mathbb{W}_n$. These must both have size $2n$, so by Lemma II.5 we obtain an isometric linear automorphism $\beta : \mathbb{W}_n \to \mathbb{W}_n$ extending α.

To finish, notice that $\mathcal{C}, e_{n+1}, \cdots, f_{n+1}, \cdots$ as well as $\mathcal{C}', e_{n+1}, \cdots, f_{n+1}, \cdots$ form a symplectic basis spanning $\mathbb{W}_\infty$. We obtain from Lemma II.5 another time an isometric linear automorphism $\gamma : \mathbb{W}_\infty \to \mathbb{W}_\infty$ extending β that is the identity almost everywhere. □

Proposition II.7. *$\mathbb{W}^k_\infty$ has precisely $\sum_{d=0}^{k} \left[{k \atop d}\right]_q \cdot q^{\binom{d}{2}}$ orbits under isometric linear automorphisms, where $q = |\mathfrak{f}|$ and*

$$\begin{bmatrix} k \\ d \end{bmatrix}_q = \frac{(q^k - 1)(q^{k-1} - 1) \cdots (q^{k-d+1} - 1)}{(q^d - 1)(q^{d-1} - 1) \cdots (q^1 - 1)}$$

is the q-binomial coefficient.

Remark. To anticipate the next subsection, we note a similarity with the Rado graph: in $\mathbb{G}^k$ there are $\sum_{d=0}^{k} \binom{k}{d} \cdot 2^{\binom{d}{2}}$ orbits — we may impose any edge relation on d vertices.

Proof. To $(v_1, \cdots, v_k) \in \mathbb{W}^k_\infty$ we associate a *type*, which comprises the following data:

- pivot indices $I \subseteq \{1, \cdots, k\}$ containing every i such that v_i is not spanned by $v_1, \cdots, v_{i-1}$ — so we inductively ensure that
$$\{v_{i'} \mid i' \in I, i' \leq i\}$$
is a basis for $\langle v_1, \cdots, v_i\rangle$;
- for each $j \notin I$, an assignment $\Lambda_j : \{i \in I \mid i < j\} \to \mathfrak{f}$ such that $v_j = \sum_{i \in I, i < j} \Lambda_j(i) v_i$;
- a map $\Omega : \binom{I}{2} \to \mathfrak{f}$ defined by $\Omega(\{i' < i\}) = \omega(v_{i'}, v_i)$.

If $\pi : \mathbb{W}_\infty \to \mathbb{W}_\infty$ is an isometric linear automorphism, then $(v_1, \cdots, v_k)$ and $(\pi(v_1), \cdots, \pi(v_k))$ evidently share the same type. Conversely, if $(w_1, \cdots, w_k)$ has the type of $(v_i, \cdots, v_k)$, then

$$\begin{aligned}\alpha : \langle v_i \mid i \in I\rangle &\to \langle w_i \mid i \in I\rangle \subseteq \mathbb{W}_n\\ v_i &\mapsto w_i\end{aligned}$$

gives an isometric linear injection for some large enough n. Observe that α must send $v_j \mapsto w_j$ for $j \notin I$ too, and that it may be extended to an isometric linear automorphism π

of $\mathbb{W}_\infty$ by Propsoition II.6. Furthermore we can find some $(v_1, \cdots, v_k)$ that realises any given type $(I, \{\Lambda_j\}_j, \Omega)$: it suffices to put

$$v_i = e_i + \sum_{i' \in I, i' < i} \Omega(i', i) f_{i'}$$

for $i \in I$ and $v_j = \sum_{i \in I, i<j} \in \Lambda_j(i) v_i$ for $j \notin I$. Therefore the number of types is precisely the number of orbits in $\mathbb{W}_\infty^k$.

Finally, we do some combinatorics. Fix $0 \leq d \leq k$ and count the number of types with $|I| = d$. There are $q^{\binom{d}{2}}$ choices for Ω and say $\#_{k,d}$ choices for the Λ_j's; the two can be chosen separately. In total, this gives

$$\sum_{d=0}^{k} q^{\binom{d}{2}} \cdot \#_{k,d}$$

types for vectors in $\mathbb{W}_\infty^d$. So focus on $\#_{k,d}$, the number of *linear types* — i.e., $(I, \{\Lambda_j\}_j)$, ignoring Ω — in $\mathbb{W}_\infty^k$. (Incidentally $\sum_{d=0}^{k} \#_{k,d}$ is the number of orbits in $\mathbb{W}_\infty^k$ or, more generally, any countable-dimensional $\mathfrak{f}$-vector space under linear automorphisms.) On the small values we easily check that

$$\#_{0,0} = 1 = \left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right]_q,$$
$$\#_{1,0} = 1 = \left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]_q, \qquad \#_{1,1} = 1 = \left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]_q.$$

Given a linear type in $\mathbb{W}_\infty^k$ with $|I| = d$, we either have $1 \in I$ or $I \subseteq \{2, \cdots, k\}$. In the first case, the linear type is specified by one of the $\#_{k-1,d-1}$ linear types in $\mathbb{W}_\infty^{k-1}$ together with how v_1 is involved in the span of the $(k-1)-(d-1)$ non-pivot vectors. In the second case, the linear type is simply one of the $\#_{k-1,d}$ linear types in $\mathbb{W}_\infty^{k-1}$. Thus

$$\begin{aligned} \#_{k,d} &= q^{k-d} \cdot \#_{k-1,d-1} + \#_{k-1,d} \\ &= q^{k-d} \cdot \begin{bmatrix} k-1 \\ d-1 \end{bmatrix}_q + \begin{bmatrix} k-1 \\ d \end{bmatrix}_q = \begin{bmatrix} k \\ d \end{bmatrix}_q. \qquad \square \end{aligned}$$

Theorem II.8. *The symplectic vector space* $\mathbb{W}_\infty$ *is smoothly approximated by* $\mathbb{W}_0 \subseteq \mathbb{W}_1 \subseteq \mathbb{W}_2 \subseteq \cdots$.

Corollary II.9. *Provided* $\mathfrak{F}$ *is of characteristic* 0*, the symplectic* $\mathfrak{f}$*-vector space* $\mathbb{W}_\infty$ *is* $\mathfrak{F}$*-oligomorphic.*

B. Symplectic graphs

For this subsection let $\mathfrak{f}$ be the two-element field.

Definition II.10. For $n = 0, 1, 2, \cdots$, the *symplectic graph* $\widetilde{\mathbb{W}}_n$ has vertices $\mathbb{W}_n$ and edges

$$v_1 \sim v_2 \iff \omega(v_1, v_2) = 1.$$

This is indeed an undirected graph: as ω is alternating, we have $\omega(v_1, v_2) = -\omega(v_2, v_1) = \omega(v_2, v_1)$.

Proposition II.11. $\mathrm{Aut}(\widetilde{\mathbb{W}}_n) = \mathrm{Aut}(\mathbb{W}_n)$.

Proof. Clearly any isometric linear automorphism of $\mathbb{W}_n$ is a graph automorphism of $\widetilde{\mathbb{W}}_n$. Conversely, any $f \in \widetilde{\mathbb{W}}_n$ is evidently isometric. To show that f is linear, take $\lambda_1, \lambda_2 \in \mathfrak{f}$ and $v_1, v_2 \in \mathbb{W}$. We calculate:

$$\begin{aligned} &\omega\Big(f(\sum_i \lambda_i v_i) - \sum_i \lambda_i f(v_i), f(w)\Big) \\ &= \omega\Big(f(\sum_i \lambda_i v_i), f(w)\Big) - \sum_i \lambda_i \omega\big(f(v_i), f(w)\big) \\ &= \omega\Big(\sum_i \lambda_i v_i, w\Big) - \sum_i \lambda_i \omega(v_i, w) \\ &= \omega(0, w) = 0 \end{aligned}$$

for all $f(w) \in f(\mathbb{W}_n) = \mathbb{W}_n$; since ω is non-degenerate, we conclude that $f(\sum_i \lambda_i v_i) = \sum_i \lambda_i f(v_i)$. $\square$

So the number of orbits in $\widetilde{\mathbb{W}}_n^k$ is precisely equal to the number of orbits in $\mathbb{W}_n^k$ — in particular, it is bounded above by $\sum_{d=0}^{k} \left[\begin{smallmatrix} k \\ d \end{smallmatrix}\right]_2 \cdot 2^{\binom{d}{2}}$ independently of n by Proposition II.7. It remains to show $\widetilde{\mathbb{W}}_0 \subseteq \widetilde{\mathbb{W}}_1 \subseteq \widetilde{\mathbb{W}}_2 \subseteq \cdots$ embeds all finite graphs:

Proposition II.12 ([1, Theorem 8.11.2])**.** *Every graph on at most* $2n$ *vertices embed into* $\widetilde{\mathbb{W}}_n$.

Proof. Let G be a graph on at most $2n$ vertices. The conclusion is trivial when $n = 0$. Also, if G contains no edges, we can choose any $2n$ of the 2^n vectors in $\langle e_1, \ldots, e_n \rangle \subseteq \widetilde{\mathbb{W}}_n$.

So suppose $n \geq 1$ and G has an edge $s \sim t$. Let $G_{s,t}$ be the graph on vertices $G \setminus \{s, t\}$ with edges which we will specify later. By induction, some embedding $f : G_{s,t} \to \widetilde{\mathbb{W}}_{n-1}$ exists. Define $f' : G \to \widetilde{\mathbb{W}}_n$ by

$$\begin{aligned} x \in G_{s,t} &\mapsto f(x) - [\![x \sim s]\!] f_n + [\![x \sim t]\!] e_n \\ s &\mapsto e_n \\ t &\mapsto f_n \end{aligned}$$

where $[\![\phi]\!]$ is 1 if ϕ holds and 0 otherwise. Then we have $\omega(f'(x), f'(s)) = [\![x \sim s]\!]$ and $\omega(f'(x), f'(t)) = [\![x \sim t]\!]$ as desired, on one hand. On the other,

$$\begin{aligned} \omega(f'(x_1), f'(x_2)) = [\![x_1 \sim x_2]\!] &+ [\![x_1 \sim s]\!][\![x_2 \sim t]\!] \\ &+ [\![x_1 \sim t]\!][\![x_2 \sim s]\!] \end{aligned}$$

tells us how we should define the edge relation in $G_{s,t}$ for f' to be an embedding of graphs. $\square$

Theorem II.13. *The Rado graph is roughly approximated by* $\widetilde{\mathbb{W}}_0 \subseteq \widetilde{\mathbb{W}}_1 \subseteq \widetilde{\mathbb{W}}_2 \subseteq \cdots$.

Corollary II.14. *Provided* $\mathfrak{F}$ *is of characteristic* 0*, the Rado graph is* $\mathfrak{F}$*-oligomorphic.*

III. Rado graph, with cogs

In this section we work with the following setting:

- $\mathcal{L}_0$ is a (possibly infinite) relational language;
- $\mathcal{C}_0$ is a monotone, free amalgamation class of $\mathcal{L}_0$-structures where each $R \in \mathcal{L}_0$ is interpreted irreflexively;
- $\mathcal{L}$ consists of $\mathcal{L}_0$ together with a new binary symbol $<$;
- $\mathcal{C}$ consists of $\mathcal{L}$-structures obtained from $\mathcal{C}_0$ by expanding with all possible linear orderings;

- $\mathbb{A}_0$ and $\mathbb{A}$ are the respective Fraïssé limits of $\mathcal{C}_0$ and $\mathcal{C}$;
- and O is an S-orbit in $\mathbb{A}^k_<$, where $S \subseteq \mathbb{A}$ is finite.

Definition III.1. A *cog* in O consists of atoms

$$a_1 < b_1 < a_2 < b_2 < \cdots < a_d < b_d$$

Proposition III.2. *"Cogs arise everywhere"*

Acknowledgements

Hrushovski
Evans

References

[1] Chris Godsil and Gordon Royle. *Algebraic graph theory*. 1st ed. Graduate Texts in Mathematics. Springer, 2001. ISBN: 978-0-387-95241-3.

[2] W. M. Kantor, Martin W. Liebeck, and H. D. Macpherson. "$\aleph_0$-Categorical Structures Smoothly Approximated by Finite Substructures". In: *Proceedings of the London Mathematical Society* s3-59.3 (1989), pp. 439–463. DOI: https://doi.org/10.1112/plms/s3-59.3.439.

[3] Antoine Mottet and Michael Pinsker. "Smooth approximations: An algebraic approach to CSPs over finitely bounded homogeneous structures". In: *J. ACM* 71.5 (Oct. 2024). DOI: 10.1145/3689207.