A semi-algebraic version of Zarankiewicz's problem

by *Jacob Fox, János Pach, Adam Sheffer, Andrew Suk, and Joshua Zahl*

presented by Martin Balko

We discuss semi-algebraic graphs and hypergraphs and show that some of the most important results in extremal combinatorics can be substantially strengthened when restricted to semi-algebraic hypergraphs. In particular, we discuss such a strengthening of the Kövári-Sós-Turán theorem.

- A hypergraph $H = (P, E)$ is called *r-partite* if it is r-uniform and P is partitioned into r parts, $P = P_1 \cup \cdots \cup P_r$, such that every edge has precisely one vertex in each part.
- An r-partite hypergraph $H = (P_1 \cup \cdots \cup P_r, E)$ is called *semi-algebraic* in $(\mathbb{R}^{d_1}, \ldots, \mathbb{R}^{d_r})$, $d = \sum_{i=1}^{r} d_i$, if there are polynomials $f_1, \ldots, f_t \in \mathbb{R}[x_1, \ldots, x_d]$ and a boolean function $\Phi(X_1, \ldots, X_t)$ such that for every $(p_1, \ldots, p_r) \in P_1 \times \cdots P_r \subset \mathbb{R}^d$, we have $(p_1, \ldots, p_r) \in E \Leftrightarrow \Phi(f_1(p_1, \ldots, p_r) \geq 0; \ldots; f_t(p_1, \ldots, p_r) \geq 0) = 1$.
- If our r-uniform hypergraph $H = (P, E)$ is a priori not r-partite, we fix an enumeration $p_1, p_2, \ldots$ of the elements of $P \subset \mathbb{R}^d$, and say that H is *semi-algebraic* if for every $1 \leq i_1 < \cdots < i_r \leq n$, $(p_{i_1}, \ldots, p_{i_r}) \in E \Leftrightarrow \Phi(f_1(p_{i_1}, \ldots, p_{i_r}) \geq 0; \ldots; f_t(p_{i_1}, \ldots, p_{i_r}) \geq 0) = 1$, where Φ is a boolean function and $f_1, \ldots, f_t$ are polynomials satisfying the same properties as above.
- We say that the E has *description complexity* at most t if E can be described with at most t polynomial equations and inequalities, and each of them has degree at most t.

Ramsey's Theorem. *The Ramsey number* $\mathrm{R}_k(n)$ *of the complete* k*-uniform hypergraph on* n *vertices satisfies* $\mathrm{twr}_{k-1}(cn^2) \leq \mathrm{R}_k(n) < \mathrm{twr}_k(c'n)$ *where the tower function* $\mathrm{twr}_k(x)$ *is defined by* $\mathrm{twr}_1(x) = x$ *and* $\mathrm{twr}_i(x) = 2^{\mathrm{twr}_{i-1}(x)}$ *for* $i \geq 2$.

Semi-algebraic setting: *Let* $\mathrm{R}_k^{d,t}(n)$ *be the minimum* N *such that every semi-algebraic* k*-uniform hypergraph* $H = (P, E)$ *of description complexity* t *contains* $P' \subseteq P$ *of size* n *such that* $\binom{P'}{k} \subseteq E$ *or* $\binom{P'}{k} \cap E = \emptyset$. *For* $k \geq 2$ *and* $d, t \geq 1$, $\mathrm{R}_k^{d,t}(n) \leq \mathrm{twr}_{k-1}(n^{c_1})$ *where* $c_1 = c_1(d, k, t)$.

Szemerédi's Regularity Lemma. *For every* $\varepsilon > 0$ *there is* $K = K(\varepsilon)$ *such that every graph has an equitable vertex partition into at most* K *parts such that all but at most an* ε *fraction of the pairs are* ε*-regular.*

Semi-algebraic setting: *For any positive integers* r, d, t, D *there exists a constant* $c = c(r, d, t, D) > 0$ *with the following property. Let* $0 < \varepsilon < 1/2$ *and* $H = (P, E)$ *be an* r*-uniform semi-algebraic hypergraph in* $\mathbb{R}^d$ *with complexity* (t, D). *Then* P *has an equitable partition* $P = P_1 \cup \cdots \cup P_K$ *into at most* $K \leq (1/\varepsilon)^c$ *parts such that all but an* ε*-fraction of the* r*-tuples of parts are homogeneous in the sense that either* $P_{i_1} \times \cdots \times P_{i_r} \subseteq E$ *or* $P_{i_1} \times \cdots \times P_{i_r} \cap E = \emptyset$.

Zarankiewicz's Problem. *What is the maximum number of edges in a* $K_{k,k}$*-free bipartite graph* $G = (P, Q, E)$ *with* $|P| = m$ *and* $|Q| = n$?

Kövári-Sós-Turán Theorem: *Every bipartite graph* $G = (P, Q, E)$, $|P| = m$, $|Q| = n$, *which does not contain* $K_{k,k}$ *satisfies* $|E(G)| < c_k(mn^{1-1/k} + n)$ *where* c_k *depends on* k.

Semi-algebraic setting: *Let* $G = (P, Q, E)$ *be a semi-algebraic bipartite graph in* $(\mathbb{R}^{d_1}, \mathbb{R}^{d_2})$ *such that* E *has description complexity at most* t, $|P| = m$, *and* $|Q| = n$. *If* G *is* $K_{k,k}$*-free, then*

$$|E(G)| \leq c_1 \left((mn)^{2/3} + m + n \right) \quad \text{for } d_1 = d_2 = 2,$$

and more generally,

$$|E(G)| \leq c_3 \left(m^{\frac{d_2(d_1-1)}{d_1 d_2 - 1} + \varepsilon} n^{\frac{d_1(d_2-1)}{d_1 d_2 - 1}} + m + n \right) \quad \text{for all } d_1, d_2.$$

Here, ε *is an arbitrary small constant and* $c_1 = c_1(t, k)$ *and* $c_3 = c_3(d_1, d_2, t, k, \varepsilon)$.

Proof of the semi-algebraic version of the Kövári-Sós-Turán Theorem:

- For a bipartite graph $G = (P, Q, E)$, let $\mathcal{F} = \{N_G(q) \subseteq P : q \in Q\}$ be a set system with ground set P and let the *dual* of $(P, \mathcal{F})$ be the set system $(\mathcal{F}, \mathcal{F}^*)$ where $\mathcal{F}^* = \{\{A \in \mathcal{F} : p \in A\} : p \in P\}$.

- The *VC dimension* of $(P, \mathcal{F})$ is the largest integer d_0 for which there exists a d_0-element set $S \subseteq P$ such that for every $B \subseteq S$, one can find a member $A \in \mathcal{F}$ with $A \cap S = B$.

- The *primal shatter function* of $(P, \mathcal{F})$ is defined as $\pi_{\mathcal{F}}(z) = \max_{P' \subseteq P, |P'| = z} |\{A \cap P' : A \in \mathcal{F}\}|$.

  Theorem 1. *Let $G = (P, Q, E)$ be a bipartite graph with $|P| = m$ and $|Q| = n$, such that the set system $\mathcal{F}_1 = \{N(q) : q \in Q\}$ satisfies $\pi_{\mathcal{F}_1}(z) \leq cz^d$ for all z. Then if G is $K_{k,k}$-free, we have $|E(G)| \leq c_1(mn^{1-1/d} + n)$, where $c_1 = c_1(c, d, k)$.*

  Theorem 2 (Milnor-Thom)**.** *Let $f_1, \ldots, f_\ell$ be d-variate real polynomials of degree at most t. The number of cells in the arrangement of their zero-sets $V_1, \ldots, V_\ell \subseteq \mathbb{R}^d$ is at most $\left(\frac{50t\ell}{d}\right)^d$ for $\ell \geq d \geq 2$.*

  Corollary 3. *Let $G = (P, Q, E)$ be a bipartite semi-algebraic graph in $(\mathbb{R}^{d_1}, \mathbb{R}^{d_2})$ with $|P| = m$ and $|Q| = n$, such that E has complexity at most t. If G is $K_{k,k}$-free, then $|E(G)| \leq c'(mn^{1-1/d_2} + n)$ where $c' = c'(d_1, d_2, t, k)$.*

- The *distance* between two sets $A_1, A_2 \in \mathcal{F}$ is $|A_1 \Delta A_2| = |(A_1 \cup A_2) \setminus (A_1 \cap A_2)|$. The *unit distance graph* $UD(\mathcal{F})$ is the graph with vertex set $\mathcal{F}$, and its edges are pairs of sets (A_1, A_2) that have distance one.

  Lemma 4 (Haussler)**.** *If $\mathcal{F}$ is a set system of VC-dimension d_0 on a ground set P, then the unit distance graph $UD(\mathcal{F})$ has at most $d_0|\mathcal{F}|$ edges.*

- We say that the set system $\mathcal{F}$ is (k, δ)-separated if among any k sets $A_1, \ldots, A_k \in \mathcal{F}$ we have

$$|(A_1 \cup \cdots \cup A_k) \setminus (A_1 \cap \cdots \cap A_k)| \geq \delta.$$

  Lemma 5 (Packing lemma)**.** *Let $\mathcal{F}$ be a set system on a ground set P such that $|P| = m$ and $\pi_{\mathcal{F}}(z) \leq cz^d$ for all z. If $\mathcal{F}$ is (k, δ)-separated, then $\mathcal{F} \leq c'(m/\delta)^d$ where $c' = c'(c, d, k)$.*

- Let $\mathcal{F}_1 = \{N_G(q) : q \in Q\}$ and $\mathcal{F}_2 = \{N_G(p) : p \in P\}$. Given a set of k points $\{q_1, \ldots, q_k\} \subseteq Q$, we say that a set $B \in \mathcal{F}_2$ *crosses* $\{q_1, \ldots, q_k\}$ if $\{q_1, \ldots, q_k\} \cap B \neq \emptyset$ and $\{q_1, \ldots, q_k\} \not\subseteq B$.

  Observation 6. *There exists k points $q_1, \ldots, q_k \in Q$ such that at most $2c'm/n^{1/d}$ sets from $\mathcal{F}_2$ cross $\{q_1, \ldots, q_k\}$, where c' is defined as in Lemma 5.*

Applications:

- Incidences with algebraic varieties in $\mathbb{R}^d$

  Theorem 7. *Let P be a set of m points and let $\mathcal{V}$ be a set of n constant-degree algebraic varieties, both in $\mathbb{R}^d$, such that the incidence graph of $P \times \mathcal{V}$ does not contain a copy of $K_{s,t}$ (here we think of $s, t,$ and d as being fixed constants, and m and n are large). Then for every $\varepsilon > 0$, we have*

$$I(P, \mathcal{V}) = O\left(m^{\frac{(d-1)s}{ds-1} + \varepsilon} n^{\frac{d(s-1)}{ds-1}} + m + n\right).$$

- Unit distances in $\mathbb{R}^d$

  Theorem 8. *Let P be a set of n points in $\mathbb{R}^d$, so that every $(d-3)$-dimensional sphere contains fewer than k points (for some constant k). Then, for any $\varepsilon > 0$, the number of unit distances spanned by P is $O(n^{2d/(d+1)+\varepsilon})$.*

- A variant for semi-algebraic hypergraphs

  Corollary 9. *Let $G = (P_1, \ldots, P_r, E)$ be an r-partite semi-algebraic hypergraph in $(\mathbb{R}^{d_1}, \ldots, \mathbb{R}^{d_r})$, such that E has description complexity at most t. For any subset $S \subseteq \{1, 2, \ldots, r\}$, we set $m = m(S) = \prod_{i \in S} |P_i|$, $n = n(S) = \prod_{i \notin S} |P_i|$, $D_1 = D_1(S) = |S|$, and $D_2 = D_2(S) = r - |S|$. If G is $K_{k,\ldots,k}$-free, then*

$$|E(G)| \leq \min_{\emptyset \neq S \subset \{1,2,\ldots,r\}} \left\{ c_3 \left(m^{\frac{D_2(D_1-1)}{D_1 D_2 - 1} + \varepsilon} n^{\frac{D_1(D_2-1)}{D_1 D_2 - 1}} + m + n \right) \right\}.$$

  Here, ε is an arbitrarily small constant and $c_3 = c_3(d_1, \ldots, d_r, t, k, \varepsilon)$.