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 $(p_1, \ldots, p_r) \in P_1 \times \cdots \times P_r \subseteq \mathbb{R}^d$, we have $(p_1, \ldots, p_r) \in E \Leftrightarrow \Phi(f_1(p_1, \ldots, p_r), \ldots, f_t(p_1, \ldots, p_r)) = 0$.

- 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 $P \subseteq \mathbb{R}^d$, and say that $H$ is semi-algebraic if for every $1 \leq i_1 < \cdots < i_r \leq n$, $(p_{i_1}, \ldots, p_{i_r}) \in E \Rightarrow \Phi(f_1(p_{i_1}, \ldots, p_{i_r}), \ldots, f_t(p_{i_1}, \ldots, p_{i_r})) = 0 = 1$, where $\Phi$ 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 $R_k(n)$ of the complete $k$-uniform hypergraph on $n$ vertices satisfies $t \omega_{k-1}(cn^2) \leq R_k(n) < t \omega_k(cn)$ where the tower function $t \omega_k(x)$ is defined by $t \omega_1(x) = x$ and $t \omega_i(x) = 2^{t \omega_{i-1}(x)}$ for $i \geq 2$.

Semi-algebraic setting: Let $R_k^{d_1}(n)$ 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 $(P')^c \subseteq E$ or $(P')^c \cap E = \emptyset$.

For $k \geq 2$ and $d, t \geq 1$, $R_k^{d_1}(n) \leq t \omega_{k-1}((cn)^{\varepsilon})$ 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 $\varepsilon$ fraction of the pairs are $\varepsilon$-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 $K \leq (1/\varepsilon)^{c \rho}$ parts such that all but an $\varepsilon$-fraction of the $r$-tuples of parts are homogeneous in the sense that either $P_1 \times \cdots \times P_r \subseteq E$ or $P_1 \times \cdots \times P_r \cap E = \emptyset$.

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 $|Q| = n$?

Kővári-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

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

and more generally,

$$|E(G)| \leq c_3 \left(m^{d_2(d_1^2-1)/d_4} + n^{d_2(d_1^2-1)/d_4} + m + n\right) \text{ for all } d_1, d_2.$$

Here, $\varepsilon$ 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 Kővári-Sós-Turán Theorem:

- For a bipartite graph $G = (P, Q, E)$, let $F = \{N_G(q) \subseteq P : q \in Q\}$ be a set system with ground set $P$ and let the dual of $(P, F)$ be the set system $(F, F^*)$ where $F^* = \{\{A \in F : p \in A\} : p \in P\}$.
- The VC dimension of $(P, F)$ 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 $A \in F$ with $A \cap B = S$.
- The primal shatter function of $(P, F)$ is defined as $\pi_F(z) = \max_{P \subseteq P, |P| = z} \left| \{A \cap P : A \in F\} \right|$.

**Theorem 1.** Let $G = (P, Q, E)$ be a bipartite graph with $|P| = m$ and $|Q| = n$, such that the set system $F_1 = \{N(q) : q \in Q\}$ satisfies $\pi_{F_1}(z) \leq cz^d$ for all $z$. Then if $G$ is $K_{k,k}$-free, we have $|E(G)| \leq c_1(n^{1-1/d} + n)$, where $c_1 = c_1(c, d, k)$.

**Theorem 2** (Milnor-Thom). Let $f_1, \ldots, f_d$ be $d$-variate real polynomials of degree at most $t$. The number of cells in the arrangement of their zero-sets $V_1, \ldots, V_d \subseteq \mathbb{R}^d$ is at most $\left(\frac{9rt^d}{d}\right)^d$ for $t \geq d \geq 2$.

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

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

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

- We say that the set system $F$ is $(k, \delta)$-separated if among any $k$ sets $A_1, \ldots, A_k \in F$ we have
  $$|\left((A_1 \cup \cdots \cup A_k) \setminus (A_1 \cap \cdots \cap A_k)\right)| \geq \delta.$$

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

- Let $F_1 = \{N_G(q) : q \in Q\}$ and $F_2 = \{N_G(p) : p \in P\}$. Given a set of $k$ points $\{q_1, \ldots, q_k\} \subseteq Q$, we say that a set $B \in F_2$ 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 $F_2$ 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 $V$ be a set of $n$ constant-degree algebraic varieties, both in $\mathbb{R}^d$, such that the incidence graph of $P \times V$ 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 any $\varepsilon > 0$, we have
  $$I(P, V) = O\left(m^{\frac{(d+1)}{d-1}+\varepsilon} n^{\frac{4d-4}{4d-2}} + m + n\right).$$

- 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 $m = \min_{S \subseteq \{1, 2, \ldots, r\}} |P_S|$, $n = n(S) = \prod_{i \in S} |P_i|$, $D_1 = D_1(S) = |S|$, and $D_2 = D_2(S) = r - |S|$. If $G$ is $K_{k,k}$-free, then
  $$|E(G)| \leq \min_{\emptyset \neq S \subseteq \{1, 2, \ldots, r\}} \left\{ c_3 \left( m^{\frac{D_2(D_1-1)}{D_2(D_2-1)}} + \varepsilon \frac{D_2(D_2-1)}{n^{\frac{(D_2-2)}{2}}} + m + n \right) \right\}.$$

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