THE ELLENBERG-GIJSWIJT THEOREM

Notes for our seminar

Lex Schrijver

The following is a variant of a lemma given in Terry Tao's blog of May 18, 2016:

Lemma 1. *Let L be a linear space, let S be a subspace of L, and let $e_1, \ldots, e_n \in L$ be linearly independent. Let $t \geq 2$ and define $\tau := \sum_{i=1}^{n} e_i^{\otimes t}$. If τ is a a linear combination of tensors $b_1 \otimes \cdots \otimes b_t$ with $b_1, \ldots, b_t \in L$ and $b_j \in S$ for at least one j, then $n \leq 2\dim(S)$.*

Proof. Consider the quotient space L/S, and define $f_i := e_i + S$ for $i = 1, \ldots, n$. So $d := \dim\langle f_1, \ldots, f_n\rangle \geq n - \dim(S)$ and, by the condition in the theorem, $\sum_{i=1}^{n} f_i^{\otimes t} = 0$.

Define $g_i := f_i^{\otimes t-1}$ for $i = 1, \ldots, n$. Then $\dim\langle g_1, \ldots, g_n\rangle \geq d$, since if (say) $f_1, \ldots, f_d$ are linearly independent, then also $f_1^{\otimes t-1}, \ldots, f_d^{\otimes t-1}$ are linearly independent.

Now $\sum_{i=1}^{n} g_i \otimes f_i = 0$. This implies $n \geq \dim\langle g_1, \ldots, g_n\rangle + \dim\langle f_1, \ldots, f_n\rangle \geq 2d$. To see this, let G and F be the matrices $[g_1, \ldots, g_n]$ and $[f_1, \ldots, f_n]$. Then $GF^{\mathsf{T}} = 0$, so the row space of G is orthogonal to the row space of F, hence $\operatorname{rank}(G) + \operatorname{rank}(F) \leq n$.

Concluding, $d \geq n - \dim(S)$ and $n \geq 2d$, hence $2\dim(S) \geq n$. ∎

Let q be a prime power and let $t, n \in \mathbb{Z}_+$. Let $A \subseteq \mathbb{F}_q^n$ and $\lambda_1, \ldots, \lambda_t \in \mathbb{F}_q$, with $\lambda_1 + \cdots + \lambda_t = 0$. For any real $d \geq 0$, let $m_{q,n}(d)$ be the number of monomials in $x_1, \ldots, x_n$ of degree at most d in which each x_i has degree at most $q-1$. Jordan Ellenberg and Dion Gijswijt [1] showed:

Theorem. *If for all $a_1, \ldots, a_t \in A$, $\lambda_1 a_1 + \cdots + \lambda_t a_t = 0$ implies $a_1 = \cdots = a_t$, then*

(1) $$|A| \leq 2m_{q,n}(\tfrac{1}{t}(q-1)n).$$

Proof. Let $p(x) := \prod_{i=1}^{n}(1 - x_i^{q-1})$ for $x \in \mathbb{F}_q^n$. Define the following tensor $\tau \in (\mathbb{F}_q^A)^{\otimes t}$:

(2) $$\tau := \sum_{a_1,\ldots,a_t \in A} p(\lambda_1 a_1 + \cdots + \lambda_t a_t) e_{a_1} \otimes \cdots \otimes e_{a_t} = \sum_{a \in A} e_a^{\otimes t},$$

the latter equality by the condition on A, since $p(x) = \delta_{x,0}$. (e_a denotes the indicator function $A \to \mathbb{F}_q$ of a.)

Now, for $z_1, \ldots, z_t \in \mathbb{F}_q^n$, $p(\lambda_1 z_1 + \cdots + \lambda_t z_t)$ is a sum of products $p_1(z_1) \cdots p_t(z_t)$ of polynomials with at least one p_j having degree at most $\frac{1}{t}(q-1)n$. Hence τ is a sum of tensors $b_1 \otimes \cdots \otimes b_t$ with $b_j = \sum_{a \in A} p_j(a) e_a = p_j|A$ for some polynomial p_j, with at least one p_j having degree at most $\frac{1}{t}(q-1)n$. In other words, τ is a sum of tensors $b_1 \otimes \cdots \otimes b_t$ with at least one b_j in

(3) $$S := \{f|A \mid f \in \mathbb{F}_q[x_1, \ldots, x_n], \deg(f) \leq \tfrac{1}{t}(q-1)n\}.$$

So, by the Lemma, $|A| \leq 2\dim(S) \leq 2m_{q,n}(\frac{1}{t}(q-1)n)$. ∎

Proposition. *For all q, α with $\alpha < \frac{1}{2}(q-1)$: $\lim_{n\to\infty} m_{q,n}(\alpha n)^{1/n} < q$.*

Proof. Define $f(x) := \sum_i x^{i-\alpha}$ (here and below i ranges over $0,\ldots,q-1$). As $\lim_{x\downarrow 0} f(x) = \infty$ and $f'(1) = \sum_i (i-\alpha) > 0$, there exists $r \in (0,1)$ that attains the minimum of f on $(0,1]$. So $f'(r) = 0$, hence $\sum_i (i-\alpha)r^i = 0$.

Define $s := \sum_i r^i$ and $z_i := s^{-1}r^i$ for $i = 0,\ldots,q-1$. So $\sum_i z_i = 1$ and $\sum_i iz_i = \alpha$. Now let $y = (y_0,\ldots,y_{q-1}) \in \mathbb{R}^q_+$ with $y_i \geq 0$ for each i, $\sum_i y_i = 1$, and $\sum_i iy_i \leq \alpha$. Consider the functions $h(x) := \sum_i x_i \log x_i$ for $x = (x_0,\ldots,x_{q-1})$ and $g(\lambda) := h(z + \lambda(y-z))$ for $\lambda \geq 0$. Then

$$\text{(4)}\qquad \frac{dg(\lambda)}{d\lambda}\bigg\rfloor_{\lambda=0} = \sum_i \frac{dh(x)}{dx_i}\bigg\rfloor_{x=z}(y_i - z_i) = \sum_i (1+\log z_i)(y_i - z_i) = \sum_i (1 - \log s + i\log r)(y_i - z_i) = \log r \sum_i i(y_i - z_i) \geq 0$$

(as $\sum_i y_i = 1 = \sum_i z_i$, $\sum_i iy_i \leq \alpha = \sum_i iz_i$, and $\log r < 0$). Hence, since h is convex, $g(1) \geq g(0)$, so $h(y) \geq h(z)$; that is, $\prod_i y_i^{y_i} \geq \prod_i z_i^{z_i}$.

Now, for any monomial in $x_1,\ldots,x_n$, consider the number n_i of variables x_j of degree i (for $i = 0,\ldots,q-1$). So $\sum_i n_i = n$ and the degree is $\sum_i in_i$. As the number of $(n_0,\ldots,n_{q-1}) \in \mathbb{Z}^q_+$ with $\sum_i n_i = n$ is at most n^q, we have, with Stirling,

$$\text{(5)}\qquad \lim_{n\to\infty} m_{q,n}(\alpha n)^{1/n} = \lim_{n\to\infty}\Big(\sum_{\substack{n_0,\ldots,n_{q-1}\in\mathbb{Z}_+\\ \sum_i n_i = n\\ \sum_i in_i \leq \alpha n}} \binom{n}{n_0,\ldots,n_{q-1}}\Big)^{1/n} = \lim_{n\to\infty} \sup_{\substack{n_0,\ldots,n_{q-1}\in\mathbb{Z}_+\\ \sum_i n_i = n\\ \sum_i in_i \leq \alpha n}} \binom{n}{n_0,\ldots,n_{q-1}}^{1/n} = \sup_{\substack{y_0,\ldots,y_{q-1}\geq 0\\ \sum_i y_i = 1\\ \sum_i iy_i \leq \alpha}} \prod_i y_i^{-y_i} = \prod_i z_i^{-z_i} = \prod_i (s^{-1}r^i)^{-z_i} = s^{\sum_i z_i} r^{-\sum_i iz_i} = sr^{-\alpha} = \sum_i r^{i-\alpha} = f(r) < f(1) = q. \qquad \blacksquare$$

References

[1] J.S. Ellenberg, D. Gijswijt, On large subsets of $\mathbb{F}_q^n$ with no three-term arithmetic progression, *Annals of Mathematics* 185 (2017) 339–343.