Notes on "Approximate subgroups of linear groups" in *On pseudo-finite dimensions* by E. Hrushovski

Gabriel Conant

March 22, 2018

These notes cover most of Section 3 of *On pseudo-finite dimensions*, by Ehud Hrushovksi [2]. They were prepared for the reading seminar on pseudofinite structures during the Model Theory, Combinatorics, and Valued fields trimester at Institut Henri Poincaré in the spring of 2018. The notes pick up from the start of Section 3, and thus assume familiarity with Sections 1 and 2. Thanks to Martin Bays, Darío Garcia, Amador Martin-Pizarro, Pierre Simon, Frank Wagner, and Tingxiang Zou for helpful discussions and comments, and also to Charlotte Kestner for lending me her notes from a similar seminar she previously gave on this topic.

Theorem 1 below is [2, Corollary 3.2], which is a result of Breuillard-Green-Tao [1] and Pyber-Szabó [4]. Theorem 2 is [2, Theorem 3.1].

1 Setup and theorems

Theorem 1. *Let $\underline{G}$ be an algebraic group. For any $0 < \epsilon < \epsilon'$ there is some $m > 0$ satisfying the following properties. Suppose F is a field, such that $\underline{G}(F)$ is simple nonabelian, and fix a finite set $X = X^{-1} \subseteq \underline{G}(F)$. Then one of the following holds.*

(i) X is contained in a subgroup of $\underline{G}(F)$ of size at most $|X|^{1+\epsilon'}$;

(ii) there is a proper algebraic subgroup $\underline{H}(F^{alg})$ of $\underline{G}(F^{alg})$ such that $X \subseteq \underline{H}(F)$ and $\underline{H}$ has complexity at most m; or

(iii) $|X^{\cdot m}| \geq |X|^{1+\epsilon}$ (where $X^{\cdot m} = \underbrace{X \cdot \ldots \cdot X}_{m \text{ times}}$).

In particular, if X is a k-approximate subgroup, and (i) and (ii) fail, then $k \geq |X|^{\frac{\epsilon}{m-1}}$.

Remark 1. For the moreover statement about approximate subgroups, suppose $XX \subseteq XA$, where $A \subseteq \underline{G}(F)$ has size at most k. Then $|X|^{1+\epsilon} \leq |X^{\cdot m}| \leq |XA^{\cdot(m-1)}| \leq |X|k^{m-1}$, and so $k \geq |X|^{\frac{\epsilon}{m-1}}$.

Remark 2. Our formulation Theorem 1 is not completely identical to [2, Corollary 3.2] (which itself is taken from [1] and [4]). The notion of "$\underline{G}$ being an algebraic group", as well as the "complexity" of such an object, have not been precisely explained. The reader should think of $\underline{G}$ functorially as a definable object in the language of rings, possibly allowing for parameters, with definitions for an underlying set and (the graph of) a binary operation. So, given a field F, asking whether $\underline{G}(F)$ is a simple nonabelian group makes sense (although the notation $\underline{G}(F)$ suppresses the underlying choice of parameters from F). The "complexity" of $\underline{G}$ can be calculated from the definition of $\underline{G}$ (e.g. based on degrees of polynomials, etc...). Theorem 1 resembles [1, Theorem 5.5], which fixes a a field F outright, and thus is stated for concrete algebraic groups. There is also a more precise discussion of "complexity" in [1]. The main difference between Theorem 1 and [2, Corollary 3.2] is in condition (ii), where Hrushovski just says "X is not contained in $\underline{H}(F)$, for $\underline{H}$ a proper algebraic subgroup of $\underline{G}$ of bounded index". On the other hand, our formulation of condition (ii) matches the corresponding part of [1, Theorem 5.5].

We will deduce Theorem 1 from Theorem 2 below. To state this second result, we recall the setup involving coarse pseudofinite dimensions.

Setup. Consider a structure

$$M = (G, \cdot, X_*, \ldots) = \prod_{\mathcal{U}} (G_n, \cdot, X_n, \ldots)$$

where $(G_n, \cdot, X_n, \ldots)$ is an expansion of a group $(G_n, \cdot)$, and X_n is a finite subset of G_n (and so $X_* = \prod_{\mathcal{U}} X_n$ is pseudofinite). Let δ be the coarse pseudofinite dimension normalized so that $\delta(X_*) = 1$. Throughout, "type-definable" means an intersection of *countably* many definable sets, where definability is with respect to the structure M (unless otherwise stated).

Recall that a subgroup $H \leq G$ is *Zariski dense* if it is not contained in a proper $(G, \cdot)$-definable (with parameters) subgroup of G.

We assume that $(G, \cdot)$ is a (definably) simple nonabelian group of finite Morley rank. We also make a further assumption, denoted $\diamond$, about the geometry of centralizers, which holds when $(G, \cdot)$ is a simple nonabelian algebraic group over an algebraically closed field (see, e.g., [1, Lemma 5.1] and references cited there). The precise formulation of $\diamond$ is postponed until Section 2.

Theorem 2. *Suppose $\Gamma \leq G$ is Zariski dense and type-definable, with $0 < \delta(\Gamma) < \infty$. Then there is a definable $S \leq G$ such that $\Gamma \leq S$ and $\delta(S) = \delta(\Gamma)$.*

1.1 Proof of Theorem 1 from Theorem 2

Fix $0 < \epsilon < \epsilon'$ and suppose there is no such $m > 0$, i.e., for any $m > 0$ there is a field F_m, a finite set $X_m = X_m^{-1} \subseteq \underline{G}(F_m)$ such that $\underline{G}(F_m)$ is simple nonabelian and:

(1) X_m is not contained in a subgroup of $\underline{G}(F_m)$ of size at most $|X|^{1+\epsilon'}$;

(2) for any proper algebraic subgroup $\underline{H}(F^{alg})$ of $\underline{G}(F^{alg})$, if $\underline{H}$ has complexity at most m then X is not contained in $\underline{H}(F)$; and

(3) $|X_m^{\cdot m}| < |X_m|^{1+\epsilon}$.

For m a power of 2, let $M_m = (\underline{G}(F_m^{alg}), \cdot, F_m, X_m^{\cdot m}, X_m^{\cdot m/2}, X_m^{\cdot m/4}, \ldots, X_m)$. Let

$$M = (G, \cdot, F, Y_0, Y_1, Y_2, \ldots) = \prod_{\mathcal{U}} M_m,$$

for some nonprincipal ultrafilter $\mathcal{U}$. In particular, $G = \underline{G}(F^{alg})$, $F = \prod_{\mathcal{U}} F_m$, and $Y_n = \prod_{\mathcal{U}} X_m^{\cdot m/2^n} \subseteq \underline{G}(F)$ for all $n \in \mathbb{N}$. Note that $Y_{n+1}Y_{n+1} = Y_n$, and so $\Gamma := \bigcap_{n\in\mathbb{N}} Y_n$ is a type-definable subgroup of G. Let δ be the coarse pseudofinite dimension normalized so that $\delta(Y_0) = 1$. So we are in the setup above, with $X_* = Y_0$.

By assumption (3), we have $|X_m^{\cdot m/2^n}|^{1+\epsilon} \geq |X_m|^{1+\epsilon} > |X_m^{\cdot m}|$ for all sufficiently large m. So $(1 + \epsilon)\delta(Y_n) \geq \delta(Y_0) = 1$. In particular, it follows that $0 < \frac{1}{1+\epsilon} \leq \delta(\Gamma) \leq 1 < \infty$.

Next, we claim that Γ is Zariski dense in G. Otherwise, there a $(G,\cdot)$-definable proper subgroup $\underline{H}(F^{alg})$ of G such that $\Gamma \subseteq \underline{H}(F)$. By compactness, there is some $n \in \mathbb{N}$ such that $Y_n \subseteq \underline{H}(F) = \prod_{\mathcal{U}} \underline{H}(F_m)$. So there is some m, larger than the complexity of $\underline{H}$, such that $X_m \subseteq X_m^{\cdot m/2^n} \subseteq \underline{H}(F_m)$, which contradicts assumption (2).

Finally, by Theorem 2 there is a definable $S \leq G$ such that $\Gamma \leq S$ and $\delta(S) = \delta(\Gamma)$. By compactness, there is $n \in \mathbb{N}$ such that $Y_n \subseteq S$. So for $\mathcal{U}$-many m, $X_m \subseteq X_m^{\cdot m/2^n} \subseteq S_m \cap G(F_m)$ and $S_m \cap \underline{G}(F_m)$ is a finite subgroup of $\underline{G}(F_m)$. By assumptions (1) and (3), we have

$$|X_m^{\cdot m}|^{\frac{1+\epsilon'}{1+\epsilon}} < |X_m|^{1+\epsilon'} < |S_m|$$

for $\mathcal{U}$-many m. Therefore

$$\frac{1+\epsilon'}{1+\epsilon}\delta(Y_0) \leq \delta(S) = \delta(\Gamma) \leq \delta(Y_0),$$

which is a contradiction since $\epsilon < \epsilon'$.

2 Theorem 2

Let M and δ be as in the setup above. Let $g = \mathrm{RM}(G)$ (and recall that Morley rank is definable in our setting). Given $a \in G$, define

$$T_a = C_G(a) \text{ and } T_a^r = \{b \in T_a : T_a = T_b\}.$$

Let $t = \mathrm{RM}(T_a)$ for some (any) generic $a \in G$. Define

$$R = \{a \in G : \mathrm{RM}(N_G(T_a)) = \mathrm{RM}(T_a) = t \text{ and } \mathrm{RM}(T_a \backslash T_a^r) < t\}.$$

Note that $R = \bigcup_{a \in R} T_a^r$. Moreover, $t < g$ since otherwise $[G : T_a] < \infty$, and so $G = T_a$ since G is connected, contradicting $Z(G) = \{1\}$.

Assumption $\diamond$: R is generic (i.e., has Morley rank g).

The following is the generalized Larsen-Pink inequality from [3] (see also [2, Section 2.15]).

Fact 1. [Larsen-Pink] *Suppose $\Gamma \leq G$ is Zariski dense and type-definable, with $0 < \delta(\Gamma) < \infty$. If $Z \subseteq G$ is $(G,\cdot)$-definable then $\delta(\Gamma \cap Z) \leq \frac{\mathrm{RM}(Z)}{g}\delta(\Gamma)$. Moreover, if $Z = T_a$ for some $a \in \Gamma$, then equality holds.*

Proof of "moreover". Fix $a \in \Gamma$, and, for $X \subseteq G$, let $a^X = \{g^{-1}ag : g \in X\}$. Set $\beta = \frac{g}{\delta(\Gamma)}$. Then

$$\mathrm{RM}(T_a) + \mathrm{RM}(a^G) = \mathrm{RM}(G) = \beta\delta(\Gamma) = \beta\delta(\Gamma \cap T_a) + \beta\delta(a^\Gamma).$$

By Larsen-Pink, $\beta\delta(\Gamma \cap T_a) \leq \mathrm{RM}(T_a)$ and $\beta\delta(a^\Gamma) \leq \beta\delta(\Gamma \cap a^G) \leq \mathrm{RM}(a^G)$, and so we must have $\beta\delta(\Gamma \cap T_a) = \mathrm{RM}(T_a)$. □

2.1 Proof of Theorem 2

Let $\Gamma \leq G$ be type-definable and Zariski dense, with $0 < \delta(\Gamma) < \infty$. Let $\mathcal{Y} = \{T_a : a \in \Gamma \cap R\}$ and set

$$S = N_G(\mathcal{Y}) = \{s \in G : s^{-1}\mathcal{Y}s = \mathcal{Y}\}.$$

We are going to show that S is the desired subgroup.

Claim 2.1.

(*a*) *$\delta(\Gamma \backslash R) < \delta(\Gamma)$.*

(*b*) *For any $a \in \Gamma \cap R$, $\delta(\Gamma \cap (T_a \backslash R)) < \delta(\Gamma \cap T_a)$.*

Proof. Part (*a*). We have

$$\delta(\Gamma \backslash R) = \delta(\Gamma \cap (G \backslash R)) \leq \frac{g-1}{g}\delta(\Gamma) < \delta(\Gamma),$$

where the first inequality follows from Larsen-Pink, and since $\mathrm{RM}(G \backslash R) \leq g - 1$ by $\diamond$.

Part (*b*). Fix $a \in \Gamma \cap R$, then

$$\delta(\Gamma \cap (T_a \backslash R)) \leq \delta(\Gamma \cap (T_a \backslash T_a^r)) \leq \frac{t-1}{g}\delta(\Gamma) = \frac{t-1}{t}\delta(\Gamma \cap T_a) < \delta(\Gamma \cap T_a),$$

where the second inequality follows from Larsen-Pink and since $a \in R$, and the equality follows from Larsen-Pink for T_a. □

Fix $X \subseteq G$ definable such that $\Gamma \subseteq X$ and $\delta(XX) < \delta(\Gamma) + \frac{\delta(\Gamma)}{2g}$.

Claim 2.2. *For any $a \in R$, $\delta(X \cap T_a) < \delta(\Gamma \cap T_a) + \frac{\delta(\Gamma)}{2g}$.*

Proof. Fix $a \in R$. Define $f \colon (X \cap T_a) \times \Gamma \to X\Gamma$ such that $f(t, \gamma) = t\gamma$. For any $(t, \gamma), (u, \eta) \in \mathrm{dom}(f)$, if $f(t, \gamma) = f(u, \eta)$ then $u^{-1}t = \eta\gamma^{-1} \in \Gamma \cap T_a$. So, for a fixed $c \in \mathrm{Im}(f)$, projection onto the first coordinate gives a well-defined injective map from $f^{-1}(c)$ to $t(\Gamma \cap T_a)$, where t is such that $f(t, \gamma) = c$ for some γ. So $\delta(f^{-1}(c)) \leq \delta(\Gamma \cap T_a)$ for any $c \in \mathrm{Im}(f)$. Altogether,

$$\delta(X \cap T_a) + \delta(\Gamma) = \delta((X \cap T_a) \times X) \leq \delta(\Gamma \cap T_a) + \delta(X\Gamma).$$

Therefore

$$\delta(X \cap T_a) - \delta(\Gamma \cap T_a) \leq \delta(X\Gamma) - \delta(\Gamma) \leq \delta(XX) - \delta(\Gamma) < \frac{\delta(\Gamma)}{2g},$$

where the last inequality follows by choice of X. □

Let $D = \{a \in R : T_a = T_{a'} \text{ for some } a' \in \Gamma \cap R\}$.

Claim 2.3. *D is definable.*

Proof. Given $a \in G$, we show that the following are equivalent:

(*i*) $a \in D$.

(*ii*) $a \in R$ and $\delta(X \cap T_a) > \frac{t-\frac{1}{2}}{g}\delta(\Gamma)$.

(*iii*) $a \in R$ and $\delta(X \cap T_a) \geq \frac{t}{g}\delta(\Gamma)$.

This suffices to prove the claim since, by continuity of δ, (*ii*) is a co-type-definable condition and (*iii*) is a type-definable condition. (Note that D is also type-definable by type-definability of Γ.)

(*i*) ⇒ (*iii*). Assume $a \in D$. Then $a \in R$ and there is $a' \in \Gamma \cap R$ such that $T_{a'} = T_a$. By Larsen-Pink for $T_{a'}$,

$$\delta(X \cap T_a) = \delta(X \cap T_{a'}) \geq \delta(\Gamma \cap T_{a'}) = \frac{t}{g}\delta(\Gamma).$$

$(iii) \Rightarrow (ii)$. Trivial.

$(ii) \Rightarrow (i)$. Assume $a \in R$ and $\delta(X \cap T_a) > \frac{t-\frac{1}{2}}{g}\delta(\Gamma)$. Then

$$\delta(\Gamma \cap T_a) > \delta(X \cap T_a) - \frac{\frac{1}{2}\delta(\Gamma)}{g} > \frac{t-1}{g}\delta(\Gamma) \geq \delta(\Gamma \cap (T_a \backslash T_a^r)),$$

where the first inequality is by Claim 2.2, and the third inequality is by Larsen-Pink and since $a \in R$. So there is some $a' \in \Gamma \cap T_a^r$. Then $a' \in \Gamma \cap R$ and $T_a = T_{a'}$, and so $a \in D$. $\square$

Note that $\mathcal{Y} = \{T_a : a \in D\}$ and so, by Claim 2.3, S is definable. It is also easy to see that $x^{-1}\mathcal{Y}x = \mathcal{Y}$ for all $x \in \Gamma$, and so $\Gamma \leq S$.

In the following, we identify $\mathcal{Y}$ with $(D/\sim) \subseteq M^{\mathrm{eq}}$, where $a \sim a'$ if and only if $T_a = T_{a'}$ (in particular, given $a \in D$, $T_a \in \mathcal{Y}$ is identified with $a/\sim$).

Claim 2.4. *$\delta(\mathcal{Y}) = (1 - \frac{t}{g})\delta(\Gamma)$.*

Proof. For any $a \in \Gamma \cap R$,

$$\frac{t}{g}\delta(\Gamma) = \delta(\Gamma \cap T_a) = \max\{\delta(\Gamma \cap T_a \cap R), \delta(\Gamma \cap (T_a \backslash R))\} = \delta(\Gamma \cap T_a \cap R),$$

where the first equality is by Larsen-Pink for T_a, and the third equality is by Claim 2.1(b). Now we have

$$\delta(\Gamma) = \max\{\delta(\Gamma \cap R), \delta(\Gamma \backslash R)\} = \delta(\Gamma \cap R) = \frac{t}{g}\delta(\Gamma) + \delta(\mathcal{Y}),$$

where the second equality is by Claim 2.1(a), and the third equality follows from the fact that $a \mapsto T_a$ is a surjective definable map from $\Gamma \cap R$ to $\mathcal{Y}$, all of whose fibers have dimension $\frac{t}{g}\delta(\Gamma)$ (note that each fiber is of the form $\Gamma \cap T_a \cap R$ for some $a \in \Gamma \cap R$). So $\delta(\mathcal{Y}) = (1 - \frac{t}{g})\delta(\Gamma)$. $\square$

Claim 2.5. *$\delta(S) < \infty$.*

Proof. Let $Z = \bigcap \mathcal{Y}$. Then $Z \subseteq C_G(\Gamma \cap R)$. Fix $b \in \Gamma$. By Claim 2.1(a),

$$\delta(\Gamma \backslash b(\Gamma \cap R)) = \delta(\Gamma \backslash (\Gamma \cap R)) < \delta(\Gamma).$$

It follows that $b(\Gamma \cap R) \cap (\Gamma \cap R) \neq \emptyset$, since otherwise

$$\delta(\Gamma) = \max\{\delta(\Gamma \backslash b(\Gamma \cap R)), \delta(\Gamma \backslash (\Gamma \cap R))\} < \delta(\Gamma).$$

As $\Gamma \cap R$ is symmetric, we have shown $\Gamma = (\Gamma \cap R) \cdot (\Gamma \cap R)$, and so $Z \subseteq C_G(\Gamma)$. Since Γ is Zariski dense, we have $C_G(\Gamma) = Z(G)$, since otherwise there is some $a \in C_G(\Gamma) \backslash Z(G)$, and so Γ is contained in T_a, which is proper and $(G, \cdot)$-definable. So $Z \subseteq Z(G) = \{1\}$.

By compactness, we may fix $T_1, \ldots, T_k \in \mathcal{Y}$ such that $T_1 \cap \ldots \cap T_k = Z = \{1\}$. Let $N_i = N_G(T_i)$. Let $\sigma \colon S \to \mathcal{Y}^k$ be given by $\sigma(s) = (s^{-1}T_1 s, \ldots, s^{-1}T_k s)$. For any $s \in S$,

$$\{x \in S : \sigma(x) = \sigma(s)\} = s(N_1 \cap \ldots \cap N_k).$$

Moreover,

$$|N_1 \cap \ldots \cap N_k| = [N_1 \cap \ldots \cap N_k : T_1 \cap \ldots \cap T_k] \leq \prod_{i=1}^{k} [N_i : T_i] < \infty,$$

where the last inequality follows from the fact that if $a \in R$ then $\mathrm{RM}(N_G(T_a)) = \mathrm{RM}(T_a)$ and so $[N_G(T_a)) : T_a] < \infty$. So the fibers of σ are uniformly finite, and thus $\delta(S) \leq \delta(\mathcal{Y}^k) < \infty$ by Claim 2.4. $\square$

Claim 2.6. $\delta(S) \leq \delta(\Gamma)$.

Proof. Fix $a \in \Gamma \cap R$. Let $T = T_a$ and $N = N_G(T_a)$. We have a well-defined injective map from $S/(S \cap N)$ to $\mathcal{Y}$, which sends $s(S \cap N)$ to $s^{-1}Ts$. So

$$\delta(S) - \delta(S \cap N) = \delta(S/S \cap N) \leq \delta(\mathcal{Y}).$$

Next, $[S \cap N : S \cap T] \leq [N : T] < \infty$ since $a \in R$, and so $\delta(S \cap N) = \delta(S \cap T)$. Altogether,

$$\left(1 - \frac{t}{g}\right) \delta(S) = \delta(S) - \delta(S \cap T) \leq \delta(\mathcal{Y}) = \left(1 - \frac{t}{g}\right) \delta(\Gamma),$$

where the first equality is by Larsen-Pink for T (applied with respect to the definable subgroup S, using Claim 2.5), and the second equality is by Claim 2.4. Since $t < g$, this yields $\delta(\Gamma) \leq \delta(S)$. $\square$

Altogether, we have shown that S has the desired properties.

References

[1] Emmanuel Breuillard, Ben Green, and Terence Tao, *Approximate subgroups of linear groups*, Geom. Funct. Anal. **21** (2011), no. 4, 774–819.

[2] Ehud Hrushovski, *On pseudo-finite dimensions*, Notre Dame J. Form. Log. **54** (2013), no. 3-4, 463–495.

[3] Ehud Hrushovski and Frank Wagner, *Counting and dimensions*, Model theory with applications to algebra and analysis. Vol. 2, London Math. Soc. Lecture Note Ser., vol. 350, Cambridge Univ. Press, Cambridge, 2008, pp. 161–176.

[4] László Pyber and Endre Szabó, *Growth in finite simple groups of Lie type*, J. Amer. Math. Soc. **29** (2016), no. 1, 95–146.