A note on product-free sets in distal groups
Atticus Stonestrom
(Email: atticusstonestrom@yahoo.com)

Abstract: Recall that a subset \( X \) of a group \( G \) is ‘product-free’ if \( X^2 \cap X = \emptyset \), i.e. if \( xy \notin X \) for all \( x, y \in X \). Let \( G \) be a group interpretable in a distal structure; for example, \( G \) may be any complex algebraic group. We prove there are constants \( c > 0 \) and \( \delta \in (0,1) \) such that every finite subset \( X \subseteq G \) distinct from \( \{1\} \) contains a product-free subset of size at least \( \delta|X|^{c+1}/|X|^c \). In particular, every finite \( k \)-approximate subgroup of \( G \) distinct from \( \{1\} \) contains a product-free subset of density at least \( \delta/k^c \).

The proof is short, and follows quickly from Ruzsa calculus and an iterated application of Chernikov and Starchenko’s distal regularity lemma.

Acknowledgements: I would first like to thank Ben Green and Artem Chernikov for helpful discussion, and I would especially like to thank Jonathan Pila and Udi Hrushovski for their enormous support, encouragement, and feedback in these past months; I am very grateful.

Notation: Throughout we will use the standard terminology and notation of ‘Ruzsa calculus’; thus a ‘multiplicative set’ is simply a finite subset of a group. Given multiplicative sets \( X, Y \) in the same ambient group, we have \( XY = \{xy : x \in X, y \in Y\} \) and \( X^{-1} = \{x^{-1} : x \in X\} \) and, for any natural number \( n \), \( X^n = \{x_1 \cdots x_n : x_i \in X\} \). Also, we take log to mean the logarithm base 2.

1 Introduction
1.1 Product-free sets

Product-free sets – especially in the abelian context, where they are called ‘sum-free’ – are a rich topic of study in additive combinatorics; one of the earliest results is Erdős’ 1965 proof that every finite set of non-zero integers contains a sum-free subset of density at least 1/3. Given a multiplicative set \( X \), two natural questions of interest include the following: what is the largest density of a product-free subset of \( X \), and how many product-free subsets does \( X \) contain? There is an extensive literature on these and related topics, and we refer the reader to the survey [12] for a picture of some of this landscape in the abelian case. The starting point of our motivation here is a question raised by Babai and Sós in [1]: is there a constant \( \delta > 0 \) such that every finite non-trivial group contains a product-free subset of density at least \( \delta \)?

In [1], Gowers gave a negative answer to this question. Indeed, let \( H \) be a non-trivial finite group and \( d \) the minimal dimension of a non-trivial complex representation of \( H \). Gowers then shows that \( H \) does not contain any product-free set of size greater than \( |H|/d^{1/3} \) by a classical result of Frobenius, this is enough to conclude that, for every prime power \( q \), \( PSL_2(q) \) contains no product-free set of size greater than \( 2|PSL_2(q)|^{8/9} \).

Gowers also obtained a bound in the reverse direction, showing that \( H \) does contain a product-free set of size at least \( |H|/2000^d \). In [6], Nikolov and Pyber improved the exponential dependence on \( d \) to a polynomial one, showing the existence of an absolute constant \( \delta \in (0,1) \) such that \( H \) contains a product-free set of size at least \( \delta|H|/d \).

These results (and others – see [5], for instance) give perspective on the maximal densities of product-free subsets of finite groups. In this paper we are motivated by the natural generalization of this line of inquiry to the class of finite approximate groups. We show that, for fixed \( k \), a positive answer to the analogue of Babai and Sós’ question for \( k \)-approximate...
groups can be obtained in a certain model-theoretic context: namely, in the context of a group interpretable in a *distal structure*. Specifically, given such a group $G$, we show the existence of constants $c > 0$ and $\delta \in (0,1)$ such that every finite subset $X \subseteq G$ distinct from $\{1\}$ contains a product-free subset of density at least $\delta|X|^c/|X^2|^c$.

### 1.2 Small doubling

Recall that the ‘doubling’ and ‘tripling’ constants of a multiplicative set $X$ are the ratios $|X^2|/|X|$ and $|X^3|/|X|$, respectively. There is of course an absolutely vast literature on the structure of sets with small doubling and tripling, of which the author is far from knowledgeable, and we thus refer the reader to the survey [2] for context. In this note we need only two very basic facts; the first is a consequence of Ruzsa’s ‘triangle inequality’:

**Fact 1.1.** If $X$ is a multiplicative set with $|X^2| \leq k|X|$, then $|XX^{-1}| \leq k^2|X|$.

The second is a consequence of a result of Petridis; see Theorem 1.5 in [7]:

**Fact 1.2.** If $X$ is a multiplicative set with $|X^2| \leq k|X|$, then there is a subset $Y \subseteq X$ such that $|Y| \geq |X|/k$ and $|Y^3| \leq k^3|Y|$.

### 1.3 Distal structures

The model-theoretic context for this paper is that of distality, introduced by Simon in [8] to identify those NIP theories that are in some sense ‘purely unstable’; two key examples of distal structures are the fields $\mathbb{R}$ and $\mathbb{Q}_p$. Here we will in fact not even need the definition of a distal structure, and we refer the reader either to chapter 9 of the book [9] for a thorough introduction to the notion, or to the introduction of [3] for a quick overview.

In this note we are concerned with groups $G$ that are definable in some distal structure. One non-trivial fact, which follows from Theorem 2.28 of [8], is that these are precisely the groups interpretable in a distal structure. As a concrete example, the reader may throughout consider $G$ to be any complex algebraic group.

The main tool we employ comes from the paper [3], wherein Chernikov and Starchenko showed that relations definable in distal structures enjoy strong combinatorial regularity properties. To state their theorem, we first recall the notion of a ‘homogeneous’ tuple of sets in an $n$-partite hypergraph; let $R \subseteq X_1 \times \cdots \times X_n$ be an $n$-ary relation on sets $X_i$.

**Definition 1.3.** A tuple $(U_1, \ldots, U_n)$ of subsets $U_i \subseteq X_i$ is ‘$R$-homogeneous’ if $U_1 \times \cdots \times U_n$ is either a subset of $R$ or disjoint from $R$.

**Definition 1.4.** $R$ has the ‘strong Erdős-Hajnal property’ if there exists a constant $\delta \in (0,1)$ with the following property: for any finite subsets $W_i \subseteq X_i$, there are subsets $U_i \subseteq W_i$, with each $U_i$ of size at least $\delta|W_i|$, such that $(U_1, \ldots, U_n)$ is $R$-homogeneous.

The relevant result from [3] is then the following:

**Fact 1.5.** Every relation definable in a distal structure has the strong Erdős-Hajnal property.

\*Once again, the paper actually shows a much more general result, applicable to arbitrary generically stable Keisler measures; see Corollary 4.5 for the precise statement.
2 Main Result

We now have all the necessary prerequisites. Our result holds in any group in which the 6-ary relation \( R(x, y, z) \equiv x_1x_2x_3 = y_1y_2y_3 \) has the strong Erdős-Hajnal property; by Fact 1.5, and the fact that any relation interpretable in a distal structure is definable in one, it applies in particular to any group interpretable in a distal structure. For the rest of this section thus fix a group \( G \) in which \( R \) has the strong Erdős-Hajnal property.

**Lemma 2.1.** There is a constant \( c_0 > 0 \) such that, for any \( \alpha \in (0, 1) \), there is some \( \varepsilon \in (0, 1) \) with the following property: every finite subset \( Y \subseteq G \) of size at least \( 8/\alpha \) contains subsets \( U, V, W \subseteq Y \), all of size at least \( \varepsilon |Y|^{c_0+1}/|Y^3|^{c_0} \), such that \( |UVW| \leq \alpha |Y| \).

**Proof.** Let \( \delta \in (0, 1) \) be a constant witnessing the strong Erdős-Hajnal property for \( R \); ie any finite subsets \( U_1, U_2, U_3, V_1, V_2, V_3 \subseteq G \) contain respective subsets \( U'_1, U'_2, U'_3, V'_1, V'_2, V'_3 \), all of density at least \( \delta \), such that \( (U'_1, \ldots, V'_3) \) is \( R \)-homogeneous. I claim we may take \( c_0 = \log(1/\delta) \); to see this, fix \( \alpha \in (0, 1) \) and let \( \varepsilon = \delta^{c_0} \).

First note that if \( U_1, U_2, U_3, V_1, V_2, V_3 \) are subsets of \( G \) of size at least 2, then we cannot have \( u_{i_1}u_{i_2}u_{i_3} = v_{i_1}v_{i_2}v_{i_3} \) for all \( u_i \in U_i \) and \( v_i \in V_i \). Thus if \( U_1, \ldots, V_3 \) is \( R \)-homogeneous, then in fact \( U_1U_2U_3 \) and \( V_1V_2V_3 \) are disjoint, and so in particular one of \( U_1U_2U_3 \) and \( V_1V_2V_3 \) has size at most \( |W_1W_2W_3|/2 \) for any sets \( W_i \supseteq U_i \cup V_i \). The strong Erdős-Hajnal property hence tells us that any finite subsets \( W_1, W_2, W_3 \subseteq G \), all of size at least 2, contain respective subsets \( U_1, U_2, U_3 \) with \( |U_i| \geq \delta |W_i| \) for each \( i \) and with \( |U_1U_2U_3| \leq |W_1W_2W_3|/2 \).

Now we simply iterate this fact; fix a finite subset \( Y \subseteq G \) with \( |Y| \geq 8/\alpha \) and let \( U_1 = V_1 = W_1 = Y \). Applying the previous remark, we inductively define sets \( U_n, V_n, W_n \) such that (i) \( U_{n+1}, V_{n+1}, W_{n+1} \) are contained in \( U_n, V_n, W_n \), respectively, (ii) \( U_{n+1}, V_{n+1}, W_{n+1} \) have size at least \( \delta |U_n|, \delta |V_n|, \delta |W_n| \), respectively, and (iii) \( |U_{n+1}V_{n+1}W_{n+1}| \leq |U_nV_nW_n|/2 \); the \((n+1)\)-th tuple of sets in this sequence can be constructed as long as all of \( U_n, V_n, W_n \) have size at least 2, and hence as long as \( \delta^{-1}|Y| \geq 2 \).

Now let \( n = \lfloor \log(|Y^3|/\alpha|Y|) \rfloor + 1 \). Then

\[
\delta^{-n} \geq \delta^{\log(|Y^3|/\alpha|Y|)} = (\alpha|Y|/|Y^3|)^{\log(1/\delta)} = (\alpha|Y|/|Y^3|)^{c_0},
\]

whence \( \delta^{-n}|Y| \geq \varepsilon |Y|^{c_0+1}/|Y^3|^{c_0} \), and so it suffices to find \( U, V, W \subseteq Y \) of size at least \( \delta^{-n}|Y| \) and with \( |UVW| \leq \alpha |Y| \). If \( \delta^{-n}|Y| \geq 2 \), then the sets \( U_n, V_n, W_n \) exist and are of appropriate size, and further satisfy \( |U_nV_nW_n| \leq |Y^3|/2^{n-1} \leq \alpha |Y| \), as needed. If instead \( 2 > \delta^{-n}|Y| \), then in particular \( 2 > \delta^{-n+1}|Y| \), and so we may take \( U, V, W \) to be any subsets of \( Y \) of size at least 2; these will satisfy \( |UVW| \leq 8 \leq \alpha |Y| \), again as needed. □

**Corollary 2.2.** There are constants \( c_1 > 0 \) and \( \varepsilon_1 \in (0, 1) \) such that every finite subset \( Y \subseteq G \) of size at least 16 contains a subset \( Z \subseteq Y \) of size at least \( \varepsilon_1 |Y|^{c_1+1}/|Y^3|^{c_1} \), such that \( |Z^{-1}Z| \leq |Y|/2 \).

**Proof.** Let \( c_0 \) be given by Lemma 2.1, and let \( \varepsilon_0 \in (0, 1) \) witness the lemma for the case \( \alpha = 1/2 \); I claim we may take \( c_1 = 3c_0 \) and \( \varepsilon_1 = 4\varepsilon_0^3 \). To see this, fix a finite subset \( Y \subseteq G \) of size at least 16, and for notational convenience let \( k = |Y^3|/|Y| \). By Lemma 2.1 we can find subsets \( U, V, W \subseteq Y \), each of size at least \( \varepsilon_0 |Y|/k^{c_0} \), such that \( |UVW| \leq |Y|/2 \).

Now, for each \( g, h \in G \), let \( Z_{g,h} = U^{-1}g\cap V\cap hW^{-1} \). The map \((g, h, z) \mapsto (gz^{-1}, z, z^{-1}h)\) gives a bijection from \( \{(g, h, z) : g, h \in G, z \in Z_{g,h}\} \) to \( U \times V \times W \), with inverse given by \((u, v, w) \mapsto (uv, wv, wv)\), whence \( \sum_{g,h \in G} Z_{g,h} = |U| |V| |W| \). On the other hand, \( Z_{g,h} \neq \emptyset \) only if \( g \in UV \) and \( h \in VW \). The size of these sets is bounded above by \( |UVW| \leq |Y|/2 \),
so we have $|U||V||W| \leq (|Y|^2/4)\sup_{g,h \in G} |Z_{g,h}|$, and there thus exist some $g, h \in G$ such that

$$|Z_{g,h}| \geq 4|U||V||W|/|Y|^2 \geq 4\varepsilon_0^3 |Y|/k^{3c_0} = \varepsilon_1 |Y|^{c_1+1}/|Y|^3 c_1.$$ 

On the other hand, $Z_{g,h}^{-1}Z_{g,h}Z_{h,g}^{-1} \subseteq g^{-1}UVW^{-1}$, whence $|Z_{g,h}^{-1}Z_{g,h}Z_{h,g}^{-1}| \leq |UVW| \leq |Y|/2$. So taking $Z = Z_{g,h}$ gives the desired result. 

**Theorem 2.3.** There are constants $c_2 > 0$ and $\varepsilon_2 \in (0, 1)$ such that every finite subset $X \subseteq G$ distinct from $\{1\}$ contains a product-free subset of size at least $\varepsilon_2 |X|^{c_2 + 1}/|X|^{c_2}$.

**Proof.** Let $c_1$ and $\varepsilon_1$ be given by Corollary 2.2; I claim we may take $c_2 = 3c_1 + 4$ and $\varepsilon_2 = \min\{\varepsilon_1/2, 1/16\}$. Thus fix a finite subset $X \subseteq G$ distinct from $\{1\}$, and for notational convenience let $k = |X^2/||X|$. If $|X| < 16k$, then $\varepsilon_2 |X|/k^{c_2} < 16k/16k^{c_2} < 1$, so we may pick any $z \in X \setminus \{1\}$ and then $\{z\}$ will be a product-free subset of $X$ of appropriate size. Hence we may assume $|X| \geq 16k$. By Fact 1.2, there is a subset $Y \subseteq X$ with $|Y| \geq |X|/k$, hence $|Y| \geq 16$, and $|Y^3| \leq k^3|Y|$. By Corollary 2.2, there is then a subset $Z \subseteq Y$ with $|Z| \geq \varepsilon_1 |Y|/k^{3c_1}$ and $|Z^{-1}ZZ^{-1}| \leq |Y|/2$.

For each $g \in G$, let $Y_g = gZ \cap Y$. The map $(g, y) \mapsto (g^{-1}y)$ gives a bijection from $\{(g, y) : g \in G, y \in Y_g\}$ to $Y \times Z$, whence $\sum_{g \in G} |Y_g| = |Y||Z|$. Moreover, $|Y_g| \leq |gZ| = |Z|$ for all $g \in G$, and $|Z^{-1}ZZ^{-1}| \leq |Y|/2$, so $\sum_{g \in Z^{-1}ZZ^{-1}} |Y_g| < |Y||Z|/2$; combining these bounds gives $\sum_{g \in Z^{-1}ZZ^{-1}} |Y_g| \geq |Y||Z|/2$.

On the other hand, $Y_g \neq \emptyset$ only if $g \in YZ^{-1}$. Since $Y, Z \subseteq X$ and $|X^2| \leq k|X|$, by Fact 1.1 we have $|YZ^{-1}| \leq |XZ^{-1}| \leq k^2|X| \leq k^3|Y|$. So in particular

$$|Y||Z|/2 \leq k^3|Y| \sup_{g \in Z^{-1}ZZ^{-1}} |Y_g|,$$

and there thus exists some $g \in G \setminus Z^{-1}ZZ^{-1}$ such that $|Y_g| \geq |Z|/2k^3$.

By definition, $Y_g \subseteq Y \subseteq X$. We further have $|Y_g| \geq \varepsilon_1 |Y|/2k^{3c_1+4}$ since $|Z| \geq \varepsilon_1 |Y|/k^{3c_1}$, and therefore

$$|Y_g| \geq \varepsilon_1 |X|/2k^{3c_1+4} \geq |\varepsilon_2 |X|/k^{c_2} = \varepsilon_2 |X|^{c_2+1}/|X|^{c_2}$$

since $|Y| \geq |X|/k$. So we need only show that $Y_g$ is product-free and we will be done; for this, since $Y_g \subseteq gZ$, it suffices to show that $gZ$ is product-free. Thus suppose otherwise; then there are $z_1, z_2, z_3 \in Z$ such that $g^{-1}z_1g^{-1}z_2 = g^{-1}z_3$. Rearranging this gives $g = z_1^{-1}z_2z_2^{-1}$, contradicting that $g \notin Z^{-1}ZZ^{-1}$, and so we are done.

This concludes the theorem. There are a number of natural extensions to consider; for example, does the analogue of Erdős’ result on sum-free subsets of $Z$ hold in $G$?

**Question 2.4.** Is there a constant $\delta > 0$ such that every finite subset $X \subseteq G$ distinct from $\{1\}$ contains a product-free subset of size at least $\delta |X|$? How about in the case $G = \text{GL}_n(\mathbb{C})$?

For completeness we include some remarks on this question in the next section. One can also ask more model-theoretic questions on definable product-free subsets of $G$; in the interest of keeping this note short and self-contained, we will not discuss these here, but hope to address them in future work.
3 Remarks on Question 2.4

In this section we will briefly discuss Question 2.4; these are standard observations but perhaps worth noting nonetheless. Let \( X \subseteq G \) be a multiplicative set distinct from \( \{1\} \); we wish to find a product-free subset of \( X \) of fixed positive density. On the one hand, if there are ‘very few’ pairs \( (x,y) \in X \times X \) with \( xy \in X \), say at most \( k|X| \)-many for some fixed \( k \geq 1 \), then one can check that \( X \) contains a product-free subset of density \( 1/O(k) \); this is true without any hypotheses on \( G \) whatsoever.

On the other hand, if ‘most’ of the pairs \( (x,y) \in X \times X \) satisfy \( xy \in X \), then we can couple Theorem 2.3 with the Balog-Szemerédi-Gowers theorem to obtain the desired set. Specifically, if there are at least \( |X|^2/k \) such pairs, then \( X \) contains a product-free subset of density \( 1/O(kO(1)) \); let us quickly sketch the proof of this.

The first relevant observation is that every dense subset of a ‘coset’ of an approximate group contains a dense product-free subset. More precisely, suppose that \( W \subseteq G \) is a finite \( k \)-approximate group, that \( u \in G \) is arbitrary, and that \( V \) is a subset of \( uW \) distinct from \( \{1\} \) and of size \( |W|/O(kO(1)) \). We have two cases; if \( u \notin W^{-1}WW^{-1} \), then \( uW \) and hence \( V \) as well are already product-free, so there is nothing to show. If instead \( u \) does lie in \( W^{-1}WW^{-1} = W^3 \), then \( (uW)^2 \) is contained in \( W^8 \), whence \( uW \) has doubling at most \( k^7 \). Since \( |uW|/|V| = O(kO(1)) \), thus \( V \) has doubling \( O(kO(1)) \), and now by Theorem 2.3 \( V \) contains a product-free subset of density \( 1/O(kO(1)) \), as needed.

We use two tools to apply this observation to more general sets; fix \( k \geq 1 \), and suppose there are at least \( |X|^2/k \) pairs \( (x,y) \in X \times X \) with \( xy \in X \). By the Balog-Szemerédi-Gowers theorem, there are then subsets \( Y,Z \subseteq X \), each of size \( |X|/O(kO(1)) \), such that \( |YZ| = O(kO(1))|X| \). By Theorem 4.6 in Tao’s paper [11], there is then an \( O(kO(1)) \)-approximate subgroup \( W \subseteq G \) of size \( O(kO(1))|X| \), and a finite set \( U \subseteq G \) of size \( O(kO(1)) \), such that \( Y \subseteq UW \). In particular, \( Y \) must have intersection of size \( |X|/O(kO(1)) \) with some coset \( uW \); now one applies the previous paragraph to obtain a product-free subset of \( Y \) of size \( |X|/O(kO(1)) \), giving the desired result.

So, we are able to handle the two edge cases, where either ‘very many’ or ‘very few’ pairs of elements of \( X \) have product landing in \( X \). However, it is entirely unclear to us how to handle the more general question, and we thus close our remarks here.
References

[1] László Babai, Vera Sós. Sidon sets in groups and induced subgraphs of Cayley graphs. European Journal of Combinatorics, 1985.

[2] Emmanuel Breuillard. A brief introduction to approximate groups. Thin Groups and Superstrong Approximation, 2013.

[3] Artem Chernikov and Sergei Starchenko. Regularity lemma for distal structures. Journal of the European Mathematical Society, 2018.

[4] Timothy Gowers. Quasirandom groups. Combinatorics, Probability, and Computing, 2008.

[5] Kiran Kedlaya. Product-free subsets of groups, then and now. Contemporary Mathematics, 2009.

[6] Nikolay Nikolov, László Pyber. Product decompositions of quasirandom groups and a Jordan type theorem. Journal of the European Mathematical Society, 2011.

[7] Giorgis Petridis. New proofs of Plünnecke-type estimates for product sets in groups. Combinatorica, 2012.

[8] Pierre Simon. Distal and non-distal NIP theories. Annals of Pure and Applied Logic, 2013.

[9] Pierre Simon. A guide to NIP theories. Cambridge University Press, 2015.

[10] Pierre Simon. A note on “Regularity lemma for distal structures”. Proceedings of the American Mathematical Society, 2016.

[11] Terence Tao. Product set estimates for non-commutative groups. Combinatorica, 2008.

[12] Terence Tao, Van Vu. Sum-free sets in groups: a survey. Journal of Combinatorics, 2017.