*J. Aust. Math. Soc.* **89** (2010), 127–132
doi:10.1017/S1446788710000236

# ON A NONABELIAN BALOG–SZEMERÉDI-TYPE LEMMA

TOM SANDERS

(Received 12 December 2009; accepted 24 April 2010)

Communicated by E. A. O'Brien

**Abstract**

We show that if $G$ is a group and $A \subset G$ is a finite set with $|A^2| \leq K|A|$, then there is a symmetric neighbourhood of the identity $S$ such that $S^k \subset A^2A^{-2}$ and $|S| \geq \exp(-K^{O(k)})|A|$.

2000 *Mathematics subject classification*: primary 11B13; secondary 11P99, 20F05, 20F99, 20P99.

*Keywords and phrases*: Balog–Szemerédi–Gowers lemma, nonabelian groups, noncommutative groups, approximate group, Katz–Koester trick, induction on doubling, Freĭman's theorem.

Suppose that $G$ is a group and $A \subset G$ is a finite set with doubling $K$, that is, $|A^2| \leq K|A|$. Clearly if $A$ is a collection of free generators then $K = |A|$, but if $K$ is much smaller then it tells us that there must be quite a lot of overlap in the products $aa'$ with $a, a' \in A$. The extreme instance of this is when $K = 1$ and $A$ is necessarily a coset of a subgroup of $A$. We are interested in the extent to which some sort of structure persists when $K$ is slightly larger than 1, say $O(1)$ as $|A| \to \infty$.

If $G$ is abelian then the structure of $A$ is comprehensively described by the Green–Ruzsa–Freĭman theorem [GR07], but in the nonabelian case no analogue is known. A number of remarkable results have been established (see [BG09a, BG09b, FKP09, Hru09, Tao09] for details of these), but a clear description has not yet emerged. The interested reader may wish to consult [Gre09] for a discussion of the state of affairs.

Freĭman-type theorems for abelian groups are applied to great effect throughout additive combinatorics, and many of these applications can make do with a considerably less detailed description of the set $A$. Moreover, additive combinatorics is now beginning to explore many nonabelian questions and so naturally a Freĭman-type theorem in this setting would be very useful. This is the motivation behind our present work: we wish to trade in some of the strength of the description of $A$ in exchange for the increased generality of working in arbitrary groups. Tao proved a result in this direction in [Tao09] for which we require a short definition. A set $S$ in a (discrete) group $G$ is a *symmetric neighbourhood of the identity* if $1_G \in S$ and $S = S^{-1}$.

THEOREM 1.1 [Tao09, Proposition C.3]. *Suppose that $G$ is a group, $A \subset G$ is a finite nonempty set such that $|AA^{-1}| \leq K|A|$, and $k \in \mathbb{N}$ and $\epsilon \in (0, 1]$ are a pair of parameters. Then there is a symmetric neighbourhood of the identity $S \subset AA^{-1}$ with $|S| = \Omega_{K,k,\epsilon}(|A|)$ such that for all $l \leq k$,*

$$\mathbb{P}(a_1 \cdots a_l \in AA^{-1} \mid a_1, \ldots, a_l \in S) \geq 1 - \epsilon.$$

The proof uses the celebrated regularity lemma of Szemerédi and so the resulting bounds are of tower type.

One would like to remove the $\epsilon$-dependence in Tao's result, but this cannot be done (even in the abelian case; see [Ruz91]) if we are only prepared to accept containment in the twofold product set $AA^{-1}$. We shall prove the following $\epsilon$-free result.

THEOREM 1.2. *Suppose that $G$ is a group, $A \subset G$ is a finite nonempty set such that $|A^2| \leq K|A|$, and $k \in \mathbb{N}$ is a parameter. Then there is a symmetric neighbourhood of the identity $S$ such that*

$$S^k \subset A^2A^{-2} \quad \textit{and} \quad |S| \geq \exp(-K^{O(k)})|A|.$$

It should be remarked that in the abelian setting the result follows from Green–Ruzsa modelling and Bogolioùboff's lemma. Indeed, this essentially amounts to following the proof of the Green–Ruzsa–Freĭman theorem and stopping before the covering argument. The resulting bound has significantly better $k$ dependence, as it gives $|S| \geq k^{-K^{O(1)}}|A|$.

One of the main applications of Theorem 1.2 is to produce pairs of sets that are 'almost invariant'. Indeed, if $|A^3| = O(|A|)$ then one can apply the theorem to get a large set $S$ such that

$$A \subset S^kA \subset A^2A^{-2}A.$$

By the nonabelian Plünnecke–Ruzsa inequalities of Tao [Tao08], we have that $|A^2A^{-2}A| = O(|A|)$ and hence by the pigeon-hole principle there is some $l \leq k - 1$ such that

$$|SS^lA| \leq (1 + O(1/k))|A| \leq (1 + O(1/k))|S^lA|.$$

Writing $A' := S^lA$, we see that the pair $(S, A')$ is almost invariant in the sense that $SA' \approx A'$ with the accuracy of approximation increasing as $k$ increases.

Exactly this argument is given as a 'cheat' argument for the proof of [Tao09, Proposition 5.1] where Tao applies [Tao09, Proposition C.3] and first sketches a proof assuming $\epsilon = 0$. In view of the above that 'cheat' is now sufficient. (In fact this entails a very slight weakening of the conclusion, but the resulting proposition is still more than sufficient for its intended use.) A similar pigeon-holing argument, but this time on multiple scales, is also used in [San09] on the way to proving a weak nonabelian Freĭman-type theorem for so-called multiplicative pairs.

We turn now to the proof of Theorem 1.2 which uses symmetry sets, popularized in the abelian setting by the book [TV06]. Suppose that $G$ is a group. Recall that the

https://doi.org/10.1017/S1446788710000236 Published online by Cambridge University Press

convolution of two functions $f, g \in \ell^1(G)$ is defined by

$$f * g(x) = \sum_{y \in G} f(y)g(y^{-1}x),$$

so that if $A, B \subset G$ then

$$\operatorname{supp} 1_A * 1_B = AB \quad \text{and} \quad 1_A * 1_B(x) = |A \cap xB^{-1}|.$$

Given $\eta \in (0, 1]$, the *symmetry set of $A$ at threshold $\eta$* is

$$\operatorname{Sym}_\eta(A) := \{x \in G : 1_A * 1_{A^{-1}}(x) \geq \eta|A|\}.$$

It is immediate that $\operatorname{Sym}_\eta(A)$ is a symmetric neighbourhood of the identity contained in $AA^{-1}$, and that we have the nesting property

$$\operatorname{Sym}_\eta(A) \subset \operatorname{Sym}_{\eta'}(A) \quad \text{whenever } \eta \geq \eta'.$$

A straightforward pigeon-hole argument shows that they also enjoy the following useful submultiplicativity property:

$$\operatorname{Sym}_{1-\epsilon}(A) \cdot \operatorname{Sym}_{1-\epsilon'} \subset \operatorname{Sym}_{1-(\epsilon+\epsilon')}(A)$$

for all $\epsilon, \epsilon' \in [0, 1)$ with $\epsilon + \epsilon' < 1$. See [TV06, Lemma 2.33] for the abelian details, which are exactly the same.

Our main result provides a plentiful supply of large symmetry sets with threshold close to 1.

PROPOSITION 1.3. *Suppose that $G$ is a group, $A$ is a nonempty subset of $G$ with $|A^2| \leq K|A|$, and $\epsilon \in (0, 1]$ is a parameter. Then there is a nonempty set $A' \subset A$ such that*

$$|\operatorname{Sym}_{1-\epsilon}(A'A)| \geq \exp(-K^{O(1/\log(1/(1-\epsilon)))}\log K)|A|.$$

One perhaps expects $\epsilon$ to be close to 0, where $1/\log(1/(1-\epsilon)) = O(\epsilon^{-1})$ is a strong estimate and would simplify the expression above. However, Tao has pointed out that the result already has content for $\epsilon = 1 - K^{-\eta}$ and this has been used in the abelian setting in [San10].

With this in hand the proof of our main theorem is immediate.

PROOF OF THEOREM 1.2. We apply Proposition 1.3 with parameter $\epsilon := 1/(k+1)$ to get a nonempty set $A' \subset A$ such that

$$|\operatorname{Sym}_{1-\epsilon}(A'A)| \geq \exp(-K^{O(k)})|A|.$$

However, by the submultiplicativity property of symmetry sets,

$$\operatorname{Sym}_{1-\epsilon}(A'A)^k \subset \operatorname{Sym}_{1-k/(k+1)}(A'A) \subset A'A(A'A)^{-1} \subset A^2A^{-2}.$$

The result follows on setting $S := \operatorname{Sym}_{1-\epsilon}(A'A)$. □

The proof of the proposition involves iterating the following lemma.

https://doi.org/10.1017/S1446788710000236 Published online by Cambridge University Press

LEMMA 1.4. *Suppose that $G$ is a group, $A \subset G$ is nonempty and finite, $A' \subset A$ has $|A'| \geq c|A|$ and $|A'A| \leq K|A|$, and $\epsilon \in (0, 1]$ is a parameter. Then at least one of the following is true:*

(i) *there is a subset $A'' \subset A' \subset A$ such that*

$$|A''| \geq c^4|A|/2K \quad \text{and} \quad |A''A| \leq K(1-\epsilon)|A|;$$

(ii) *we have the bound*

$$|\mathrm{Sym}_{1-\epsilon}(A'A)| \geq c^3|A|/2K.$$

PROOF. Since $A' \subset A$ we have that $|A'A'| \leq |A'A|$ and

$$\sum_{x \in G} 1_A * 1_{A'}(x)^2 \geq \sum_{x \in G} 1_{A'} * 1_{A'}(x)^2.$$

Now, the Cauchy–Schwarz inequality can be used to bound the right-hand side:

$$\sum_{x \in G} 1_{A'} * 1_{A'}(x)^2 \geq \frac{1}{|A'^2|}\left(\sum_{x \in G} 1_{A'} * 1_{A'}(x)\right)^2.$$

However,

$$\sum_{x \in G} 1_{A'} * 1_{A'}(x) = |A' \times A'| = |A'|^2$$

and so

$$\sum_{x \in G} 1_{A'} * 1_{A'}(x)^2 \geq |A'|^4/|A'A|.$$

On the other hand, for arbitrary sets $B, C, D, E \subset G$,

$$\langle 1_B * 1_C, 1_D * 1_E \rangle_{\ell^2(G)} = |\{(b, c, d, e) \in B \times C \times D \times E : bc = de\}|,$$

and $bc = de$ if and only if $d^{-1}b = ec^{-1}$, whence

$$\langle 1_A * 1_{A'}, 1_A * 1_{A'} \rangle_{\ell^2(G)} = \langle 1_{A^{-1}} * 1_A, 1_{A'} * 1_{A'^{-1}} \rangle_{\ell^2(G)}.$$

For $t \in G$ write $A'_t := A' \cap (tA')$ and define

$$L := \{t \in G : |A'_t| \geq |A'|^4/(2|A'A||A|^2)\}.$$

It is easy to check that

$$|L||A||A'| + \frac{|A'|^4}{2|A'A||A|^2} \cdot |A|^2 \geq \sum_{x \in G} 1_A * 1_{A'}(x)^2,$$

from which it follows that

$$|L| \geq |A'|^3/2|A'A||A| \geq c^3|A|/2K$$

since $|A'A| \leq K|A|$ and $|A'| \geq c|A|$.

https://doi.org/10.1017/S1446788710000236 Published online by Cambridge University Press

Now, if there is some $t \in L$ such that $|A'_t A| \le (1-\epsilon)|A'A|$, then we terminate in the first case of the lemma with $A'' = A'_t$: simply note that $A'' = A'_t \subset A' \subset A \subset G$,

$$|A''| = |A'_t| \ge \frac{|A'|^4}{2|A'A||A|^2} \ge \frac{c^4}{2K}|A|,$$

since $|A'A| \le K|A|$ and $|A'| \ge c|A|$, and

$$|A''A| \le (1-\epsilon)|A'A| \le K(1-\epsilon)|A|.$$

In light of this we may assume that there is no such $t \in L$, that is,

$$|A'_t A| \ge (1-\epsilon)|A'A| \quad \forall t \in L.$$

However, $A'_t A = (A' \cap tA')A \subset (A'A) \cap t(A'A)$, whence

$$1_{A'A} * 1_{(A'A)^{-1}}(t) \ge (1-\epsilon)|A'A| \quad \forall t \in L,$$

and we are in the second case in view of the lower bound on the size of $L$. □

PROOF OF PROPOSITION 1.3. We apply Lemma 1.4 iteratively to get a sequence of nonempty sets $(A'_i)_i$ satisfying

$$A'_{i+1} \subset A, \quad |A'_{i+1}| \ge |A|/(2K)^{(4^i-1)/3} \quad \text{and} \quad |A'_i A| \le (1-\epsilon)^i K|A|.$$

First $A'_0 := A$. Suppose that we are at stage $i$ of the iteration and apply Lemma 1.4 to the pair $(A'_i, A)$. If we are in the first case of the lemma then we get a set $A'_{i+1} \subset A$ with

$$|A'_{i+1}| \ge (1/(2K)^{(4^i-1)/3})^4|A|/2K = |A|/(2K)^{(4^{i+1}-1)/3}$$

and

$$|A'_{i+1}A| \le (1-\epsilon)|A'_i A| \le (1-\epsilon)^{i+1}K|A|.$$

The sequence $(A'_i)_i$ has the desired properties and in light of the last one the iteration certainly terminates at some stage $i_0$ with $i_0 \le \lceil \log K / -\log(1-\epsilon) \rceil$ since $A'_i$ is nonempty so $|A'_i A| \ge |A|$.

When the iteration terminates we put $A' := A'_i$ and since we are in the second case of Lemma 1.4 we have the desired result. □

It is worth making a number of remarks. First, a lower bound for $|A'|$ may also be read out of the proof, although in applications it is not clear how useful this information is. The driving observation in the proof of Lemma 1.4 is that

$$(A' \cap tA')A \subset (A'A) \cap (tA'A),$$

so if the left-hand side is close to $|A'A|$ in size then $t \in \mathrm{Sym}_{1-o(1)}(A'A)$. This rather cute idea comes from the work of Katz and Koester [KK08], who use it in abelian groups to show that if a set has doubling $K$ then there is a correlating set with larger additive energy than the trivial Cauchy–Schwarz lower bound.

https://doi.org/10.1017/S1446788710000236 Published online by Cambridge University Press

Finally, at about the same time as this paper was produced, Croot and Sisask [CS09] developed a different method for analysing sumsets, which turns out to also work for sets of small doubling in nonabelian groups. Their argument gives a better bound in Theorem 1.2 showing that one may take $|S| \geq \exp(-O(k^2 K \log K))|A|$.

## Acknowledgements

The author would like to thank Ben Green for encouraging the writing of this paper, Ben Green and Terry Tao for useful discussions around this topic and the anonymous referee for many useful suggestions.

## References

[BG09a] E. Breuillard and B. J. Green, 'Approximate groups, II: The solvable linear case', arXiv:0907.0927, 2009.
[BG09b] E. Breuillard and B. J. Green, 'Approximate subgroups, I: The torsion-free nilpotent case', arXiv:0906.3598, 2009.
[CS09] E. S. Croot and O. Sisask, 'A probabilistic technique for finding almost-periods of convolutions', arXiv:1003.2978, 2010.
[FKP09] D. Fischer, N. H. Katz and I. Peng, 'On Freĭman's theorem in nilpotent groups', arXiv:math/0901.1409, 2009.
[Gre09] B. J. Green, 'Approximate groups and their applications: work of Bourgain, Gamburd, Helfgott and Sarnak', arXiv:0911.3354, 2009.
[GR07] B. J. Green and I. Z. Ruzsa, 'Freĭman's theorem in an arbitrary abelian group', *J. Lond. Math. Soc.* (2) **75**(1) (2007), 163–175.
[Hru09] E. Hrushovski, 'Stable group theory and approximate subgroups', arXiv:0909.2190, 2009.
[KK08] N. H. Katz and P. Koester, 'On additive doubling and energy', arxiv:0802.4371, 2008.
[Ruz91] I. Z. Ruzsa, 'Arithmetic progressions in sumsets', *Acta Arith.* **60**(2) (1991), 191–202.
[San09] T. Sanders, 'Indicator functions in the Fourier–Eymard algebra', arXiv:0912.0308, 2009.
[San10] T. Sanders, 'Structure in sets with logarithmic doubling', arXiv:1002.1552, 2010.
[Tao08] T. C. Tao, 'Product set estimates for non-commutative groups', *Combinatorica* **28**(5) (2008), 547–594.
[Tao09] T. C. Tao, 'Freĭman's theorem for solvable groups', arXiv:0906.3535, 2009.
[TV06] T. C. Tao and V. H. Vu, *Additive Combinatorics*, Cambridge Studies in Advanced Mathematics, 105 (Cambridge University Press, Cambridge, 2006).

TOM SANDERS, Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, UK
e-mail: t.sanders@dpmms.cam.ac.uk