On a certain generalization of the Balog-Szemerédi-Gowers theorem

Evan Borenstein and Ernie Croot

December 30, 2010

The Balog-Szemerédi-Gowers theorem has a rich history, and is a very useful tool in additive combinatorics. It began with a paper by Balog and Szemerédi [2], and then was refined by Gowers [3] to the following basic result (actually, Gowers proved somewhat more than we state here):

Theorem 1 *There exists an absolute constant $\kappa > 0$ such that the following holds for all finite subsets X and Y of size $n > n_0$ of an abelian group: Suppose that there are at least Cn^3, $C > 0$, solutions to $x_1 + y_1 = x_2 + y_2$, $x_i \in X$ and $y_i \in Y$. Then X contains a subset X', of size at least $C^{\kappa}n$, such that*

$$|X' + X'| \ \leq \ C^{-\kappa}n.$$

Sudakov, Szemerédi and Vu [6] proved a refinement of this theorem (Balog [1] independently obtained a similar result). Before we state it, we need to introduce some notation: Suppose G is a graph connecting two vertex sets A and B, which we think of as two subsets of some additive group. We then use the notation $A \overset{G}{+} B$ to denote the set of all sums $a + b$, where (a, b) is an edge of the graph G. Now we can state the theorem of Sudakov, Szemerédi and Vu:

Theorem 2 *Let n, C, K be positive numbers, and let A and B be two sets of n integers. Suppose that there is a bipartite graph $G(A, B, E)$ with at least n^2/K edges and $|A \overset{G}{+} B| \leq Cn$. Then one can find a subset $A' \subset A$ and a subset $B' \subset B$ such that $|A'| \geq n/16K^2$, $|B'| \geq n/4K$ and $|A' + B'| \leq 2^{12}C^3K^5n$.*

Remark. It is not difficult to show that this theorem, along with some lemmas and theorems of Ruzsa (the Ruzsa triangle inequality [7], and the Ruzsa-Plünnecke Theorem [4]), implies that we may take $\kappa < 20$ in Theorem 1.

In the same paper, Sudakov, Szemerédi and Vu [6, Theorem 4.3] proved the following powerful hypergraph version of the Balog-Szemerédi-Gowers Theorem:

Theorem 3 *For any positive integer k, there are polynomials $f_k(x,y)$ and $g_k(x,y)$ with degrees and coefficients depending only on k, such that the following holds. Let n, C, K be positive numbers. If $A_1, ..., A_k$ are sets of n positive integers, $H(A_1, ..., A_k, E)$ is the k-partite, k-uniform hypergraph with at least n^k/K edges, and $|\overset{H}{\oplus}_{1\le i\le k} A_i| \le Cn$, then one can find subsets $A_i' \subset A_i$ such that*

- $|A_i'| \ge n/f_k(C,K)$ *for all* $1 \le i \le k$;
- $|A_1' + \cdots + A_k'| \le g_k(C,K)n$.

The notation $\overset{H}{\oplus}$ means that the sum is restricted to the hypergraph H.

Beautiful and useful as it is, it would be nice if one had some control on the degrees of these polynomials f and g. And, it would be good to be able to control the rate of growth of sums $A_1' + \cdots + A_\ell'$, where ℓ is much smaller than k – it would be good to be able to bound the size of this sum from above by

$$C^{1+\varepsilon} K^{d_k} n, \tag{1}$$

where d_k depends only on k. Perhaps such a bound can be developed by modifying the proof of Sudakov, Szemerédi and Vu (though we were not able to see how to do this); however, in the present paper, we take a different tack, and produce an alternate proof of a related hypergraph Balog-Szmeredi-Gowers theorem, where an upper bound such as (1) will be implicit, though only for the case where $A_1 = \cdots = A_k$. In our proof, we will use some of the same standard tricks as Sudakov, Szemerédi and Vu do in their proof.

The notation we use to describe this theorem, and its proof, will be somewhat different from that used by Sudakov, Szemerédi and Vu. Furthermore, we will not attempt here to give the most general formulation of the theorem.

Theorem 4 *For every $0 < \varepsilon < 1/2$ and $c > 1$, there exists $\delta > 0$, such that the following holds for all k sufficiently large, and all sufficiently large finite subsets A of an additive abelian group: suppose that*

$$S \subseteq A \times A \times \cdots \times A = A^k,$$

and let

$$\Sigma(S) := \{a_1 + \cdots + a_k : (a_1, ..., a_k) \in S\}.$$

If

$$|S| \geq |A|^{k-\delta}, \text{ and } |\Sigma(S)| < |A|^c,$$

then there exists

$$A' \subseteq A, \ |A'| \geq |A|^{1-\varepsilon},$$

such that

$$|\ell A'| = |A' + \cdots + A'| \leq |A'|^{c(1+\varepsilon\ell)}.$$

To appreciate the strength of the conclusion here, note that the exponent $c(1 + \varepsilon\ell)$ appearing in the last displayed equation is not much larger than the exponent c (at least for ε small enough relative to ℓ) appearing in the assumption $|\Sigma(S)| < |A|^c$. Furthermore, this is about the best we could hope to prove, apart from that term $\varepsilon\ell$ in the exponent $c(1 + \varepsilon\ell)$ because of the following example: take A to be a Sidon subset of $\{1, 2, ..., n\}$ (recall that a Sidon set is one having no solutions $a + b = c + d$ except trivial ones where $\{a, b\} = \{c, d\}$) having size $\sim \sqrt{n}$, which is known to exist from the work of Singer [5]. Then, if we just set $S = A^k$ we will have that

$$|A|^2 \ll_k |\Sigma(S)| \ll_k |A|^2,$$

which means that S satisfies the hypotheses of the above theorem with $c \sim 2$ (the larger n is relative to k, the closer to 2 we can take c to be). Now, regardless of what subset $A' \subseteq A$ we choose, it too must be a Sidon set, and therefore will satisfy $|A' + A'| \gg |A'|^2$, and more generally,

$$|\ell A'| \gg |A'|^2.$$

We can likewise develop extremal examples showing that Theorem 4 is sharp for higher values of c by using generalized Sidon sets that avoid non-trivial solutions to $x_1 + \cdots + x_k = y_1 + \cdots + y_k$, where $k \geq 3$ is some fixed integer.

1 Proof of Theorem 4

1.1 Notation and basic assumptions

It will be advantageous to describe the proof in terms of strings. So, the set $S \subseteq A^k$ will be thought of as a collection of strings $x_1x_2\cdots x_k$ (where $x_i \in A$) of length k.

Often, we split these strings up into substrings; for example, the string $x = x_1\cdots x_k$ can be written as a product of a "left substring ℓ of length $k/2$" (assume k is even) and a "right substring r of length $k/2$". So, $x = \ell r$.

We may assume that $k = 2^n$, since if this is not the case, then we let k' be the largest power of 2 of size at most k, and proceed as follows: Given a string $x_1\cdots x_k$ in S, we write it as a product $\ell_x r_x$, where

$$\ell_x := x_1\cdots x_{k'} \text{ and } r_x := x_{k'+1}\cdots x_k.$$

Now, for some string y we will have that $r_x = y$ for at least $|S|/|A|^{k-k'}$ choices for $x \in S$. Letting S' denote the set of all strings ℓ_x with $r_x = y$, we will have

$$|S'| \geq |A|^{k'-\delta},$$

and clearly

$$|\Sigma(S')| \leq |\Sigma(\{\ell_x y : x \in S'\})| \leq |\Sigma(S)| < |A|^c.$$

So, we could just assume that our k had this value k' all along (remember, we get to choose k to be as large as needed to get the desired conclusion).

1.2 Lengths of iterations and the choice of δ and k

Our proof will be highly iterative, and will produce a sequence of sets

$$S_0 := S,\ S_1,\ S_2,\ ...,\ \text{each } S_m \subseteq A^{k_m},$$

until one is found that has certain nice properties.

We will think of this process in terms of 'replacing' the set $S_m \subseteq A^{k_m}$ with a set $S_{m+1} \subseteq A^{k_{m+1}}$ that satisfies 'better' inequalities, specifically

$$|S_{m+1}| \geq |A|^{k_{m+1}-\delta_{m+1}}, \text{ and } |A|^{1-\delta_{m+1}} \leq |\Sigma(S_{m+1})| \leq |\Sigma(S_m)|^{1-\varepsilon/400c},$$

where each $\delta_i \leq 5^i\delta$. The lower bound on $|\Sigma(S_{m+1})|$ comes from the fact that each element of $\Sigma(S_{m+1})$ can correspond to at most $|A|^{k_{m+1}-1}$ strings of S_{m+1}, along with the lower bound on $|S_{m+1}|$.

We now will show that the number of such iterations we can take will be bounded from above in terms of ε and c for $\delta > 0$ sufficiently small: first note that $|\Sigma(S_0)| = |\Sigma(S)| < |A|^c$, by hypothesis. Next, observe that for $\delta > 0$ small enough in terms of m we have the lower bound $|\Sigma(S_m)| \geq |A|^{1-\delta_m} > |A|^{1/2}$. And we have a companion upper bound

$$\begin{aligned} |\Sigma(S_{m-1})|^{1-\varepsilon/400c} \;\leq\; |\Sigma(S_{m-2})|^{(1-\varepsilon/400c)^2} \;\leq\; \cdots \;&\leq\; |\Sigma(S_0)|^{(1-\varepsilon/400c)^m} \\ &<\; |A|^{c(1-\varepsilon/400c)^m}. \end{aligned}$$

Putting the upper and lower bounds together we have that for $\delta > 0$ small enough in terms of ε and c, the iteration process can continue only so long as $1/2 < c(1-\varepsilon/400c)^m$, which gives the following upper bound on the number of iterations:

$$m \;<\; \log(1/2c)/\log(1-\varepsilon/400c).$$

Of course, there is also the issue of whether we "run out of dimensions" before performing this many iterations, because at each step the $k_{m+1} \leq k_m$; in fact, we do not run out, since $k_0 = k$ and since at each step $k_{m+1} \geq k_m/2$, meaning that so long as the initial value of k is large enough in terms of c and ε we will certainly have enough dimensions to play with to run though $\lfloor \log(1/2c)/\log(1-\varepsilon/400c) \rfloor$ iterations.

Since our theorem is a qualitative result, in that it does not even attempt to explain how δ or k depends on ε and c, there is no need to be more precise about just how small to take δ or how large to take k, in order for our iteration process to terminate and prove our theorem.

We will now describe the iterative process, but before we do, we initialize parameters as follows:

$$S_0 \;:=\; S,\; k_0 \;:=\; k,\; \delta_0 \;:=\; \delta, \text{ and set } m \;:=\; 0.$$

1.3 The iteration part of the argument

Given a string x of length $k_m/2$, we let $R_m(x)$ denote the set of all strings y of length $k_m/2$ such that

$$xy \;\in\; S_m.$$

We analogously define $L_m(y)$ to be those strings x such that $xy \in S_m$.

We will now select an x, and therefore $R_m(x)$, very carefully, so that it satisfies certain useful properties: We begin with the inequality

$$\sum_x |R_m(x)| \;=\; |S_m| \;\geq\; |A|^{k_m - \delta_m}.$$

We now apply the following lemma.

Lemma 1 *Suppose that V is a set of n elements, and suppose that*

$$U_1, U_2, ..., U_r \;\subseteq\; V$$

satisfy

$$\sum_{i=1}^{r} |U_i| \;\geq\; rn^{1-\delta}.$$

Then, there exists $1 \leq j \leq r$ such that

$$\sum_{1 \leq i \leq r} |U_i \cap U_j| \;\geq\; rn^{1-2\delta}.$$

Proof of the lemma. Let $r(v)$ denote the number of sets U_i that contain the element $v \in V$. One easily sees that

$$\sum_{v \in V} r(v)^2 \;=\; \sum_{1 \leq i,j \leq r} |U_i \cap U_j|,$$

and

$$\sum_{v \in V} r(v) \;=\; \sum_{i=1}^{r} |U_i|.$$

So, the Cauchy-Schwarz inequality tells us that

$$\sum_{1 \leq i,j \leq r} |U_i \cap U_j| \;\geq\; \left(\sum_{i=1}^{r} |U_i| \right)^2 |V|^{-1} \;\geq\; r^2 n^{1-2\delta}.$$

Picking out any value j making the sum over i on the corresponding terms on the left-hand-side maximal, we see that

$$\sum_{i=1}^{r} |U_i \cap U_j| \;\geq\; rn^{1-2\delta},$$

as claimed. ■

From this lemma we easily deduce that there exists x such that

$$\sum_y |R_m(x) \cap R_m(y)| \ \geq \ |A|^{k_m - 2\delta_m}.$$

Next, we let

$$S_{m+1} \ := \ \{yz \in S_m \ : \ z \in R_m(x)\}, \tag{2}$$

and we observe that

$$|S_{m+1}| \ = \ \sum_y |R_m(x) \cap R_m(y)| \ \geq \ |A|^{k_m - 2\delta_m};$$

so, S_{m+1} is not too much smaller than S_m.

We now let

$$\delta_{m+1} \ := \ 2\delta_m, \text{ and } k_{m+1} \ := \ k_m,$$

and observe that S_{m+1} satisfies

$$|S_{m+1}| \ \geq \ |A|^{k_{m+1} - \delta_{m+1}},$$

and we in addition have that every element of S_{m+1} can be expressed as yz, where $z \in R_m(x) = R_{m+1}(x)$.

Now suppose that there is a string y of length $k_{m+1}/2$ such that

$$|R_{m+1}(y)| \ \geq \ |A|^{k_{m+1}/2 - 2\delta_{m+1}},$$

and such that

$$|\Sigma(R_{m+1}(y))| \ \leq \ |\Sigma(S_{m+1})|^{1-\varepsilon/400c}.$$

If this occurs, then we let

$$S_{m+2} \ := \ R_{m+1}(y), \ k_{m+2} \ := \ k_{m+1}/2, \ \delta_{m+2} \ := \ 2\delta_{m+1}.$$

In addition, we set

$$m \ \leftarrow \ m+2,$$

and then we start back at the very beginning of subsection 1.3.

1.4 The sets H' and H''

When we come out of the iteration loops from the previous subsection, we finish with a set S_m having a number of highly useful properties, among them:

- $|S_m| \ \geq \ |A|^{k_m - \delta_m}$;

- For a particular string x of length $k_m/2$, for all y we have $R_m(y) \subseteq R_m(x)$; and,

- If we let H denote those strings h of length $k_m/2$ such that

$$|R_m(h)| \ \geq \ |A|^{k_m/2 - 2\delta_m},$$

then for every such h we will have that

$$|\Sigma(S_m)|^{1-\varepsilon/400c} \ < \ |\Sigma(R_m(h))| \ \leq \ |\Sigma(R_m(x))| \ \leq \ |\Sigma(S_m)|. \tag{3}$$

One can easily show, using the lower bound for $|S_m|$, that for $|A|$ sufficiently large,

$$|H| \ > \ |A|^{k_m/2 - 2\delta_m}.$$

Since

$$\sum_{z \in R_m(x)} |\{h \in H \ : \ hz \in S_m\}| \ \geq \ |H| \cdot |A|^{k_m/2 - 2\delta_m},$$

we deduce that there exists $z \in R_m(x)$ such that there are at least

$$|H| \cdot |A|^{-2\delta_m} \ \geq \ |A|^{k_m/2 - 4\delta_m}$$

vectors $h \in H$ satisfying

$$hz \ \in \ S_m. \tag{4}$$

Fix one of these z, and let

$$H' \ \subseteq \ H$$

denote all those $h \in H$ such that (4) holds. Note that

$$|H'| \ \geq \ |A|^{k_m/2 - 4\delta_m}. \tag{5}$$

Next, let $H'' \subseteq H'$ denote those $h \in H'$ such that there are at least

$$|H'| \cdot |\Sigma(H')|^{-1}/2 \tag{6}$$

other $h' \in H'$ satisfying $\Sigma(h') = \Sigma(h)$. We have that

$$|H' \setminus H''| < |\Sigma(H')|(|H'| \cdot |\Sigma(H')|^{-1}/2) = |H'|/2$$

So,

$$|H''| > |H'|/2 \geq |A|^{k_m/2-5\delta_m}, \tag{7}$$

for $|A|$ sufficiently large.

We also note that

$$|\Sigma(H'')| \leq |\Sigma(H')| = |\Sigma(\{hz : h \in H'\})| \leq |\Sigma(S_m)|.$$

This is one of the places where it was essential to have that $z \in R_m(h)$ for all $h \in H'$.

Now suppose that, in fact,

$$|\Sigma(H'')| \leq |\Sigma(S_m)|^{1-\varepsilon/400c}. \tag{8}$$

Then set

$$S_{m+1} := H'', \; k_{m+1} := k_m/2, \; \delta_{m+1} := 5\delta_m,$$

update m to

$$m \leftarrow m+1$$

and repeat the iteration process again, starting in subsection 1.3.

On the other hand, if (8) does not hold, then we will have that

$$|\Sigma(S_m)|^{1-\varepsilon/400c} < |\Sigma(H'')| \leq |\Sigma(H')| \leq |\Sigma(S_m)|. \tag{9}$$

1.5 The final leg of the proof

First, we will produce a lower bound on the number of quadruples

$$\sigma_1, \sigma_2 \in \Sigma(H''), \text{ and } \sigma_3, \sigma_4 \in \Sigma(R_m(x)), \tag{10}$$

satisfying

$$\sigma_1 + \sigma_3 = \sigma_2 + \sigma_4. \tag{11}$$

To do this, we begin by noting that the number of quadruples we aim to count is at least

$$Q_0 := \sum_{s \in \Sigma(S_m)} |\{\sigma \in \Sigma(H''),\ \sigma' \in \Sigma(R_m(x)) \ :\ \sigma + \sigma' = s\}|^2;$$

and the number of quadruples $\sigma_1, \sigma_2, \sigma_3, \sigma_4$ satisfying $\sigma_1 + \sigma_3, \sigma_2 + \sigma_4 \in \Sigma(S_m)$, without regard to whether they satisfy (11), is

$$\begin{aligned} Q_1 \quad &:= \quad \left(\sum_{s \in \Sigma(S_m)} |\{\sigma \in \Sigma(H''),\ \sigma' \in \Sigma(R_m(x)) \ :\ \sigma + \sigma' = s\}| \right)^2 \\ &\geq \quad |\Sigma(H'')|^2 \min_{h \in H''} |\Sigma(R_m(h))|^2. \end{aligned}$$

From the Cauchy-Schwarz inequality we have that

$$Q_0 \ \geq \ Q_1/|\Sigma(S_m)| \ \geq \ |\Sigma(H'')|^2 \min_{h \in H''} |\Sigma(R_m(h))|^2/|\Sigma(S_m)|.$$

To bound this from below, we apply (3) and (9), and deduce

$$Q_0 \ \geq \ |\Sigma(S_m)|^{3-\varepsilon/100c}.$$

Now we set

$$X_0 \ := \ \Sigma(H''), \text{ and } Y_0 \ := \ \Sigma(R_m(x)).$$

We cannot directly apply Theorem 1 to X_0 and Y_0, because $|X_0|$ and $|Y_0|$ may not be equal; however, they nearly are, since both are bounded from below by $|\Sigma(S_m)|^{1-\varepsilon/400c}$ and from above by $|\Sigma(S_m)|$. We now describe how to work around this issue: let us suppose without loss of generality that $|X_0| \leq |Y_0|$. Partition Y_0 into $q \leq |\Sigma(S_m)|^{\varepsilon/400c}$ disjoint subsets (any way you wish) $Y_{0,1}, ..., Y_{0,q}$ where all but the last one have size $|X_0|$, while the last set will have size at most $|X_0|$. For a pair of sets U, V letting $E(U,V)$ denote the number of quadruples $(u, u', v, v') \in U^2 \times V^2$ satisfying $u + v = u' + v'$, we find that

$$\sum_{i=1}^{q} E(X_0, Y_{0,i}) \ = \ E(X_0, Y_0) \ \geq \ Q_0 \ \geq \ |\Sigma(S_m)|^{3-\varepsilon/100c}.$$

Pick out the set $Y_{0,i}$ in this sum that maximizes $E(X_0, Y_{0,i})$, and note that

$$E(X_0, Y_{0,i}) \ \geq \ q^{-1}|\Sigma(S_m)|^{3-\varepsilon/100c} \ \geq \ |\Sigma(S_m)|^{3-\varepsilon/80c}.$$

We then set $X = X_0$, and if it so happens that $i < q$ then we set $Y = Y_{0,i}$; otherwise, if $i = q$, we set Y be the union of $Y_{0,i}$ with any collection of $|X_0| - |Y_{0,i}|$ elements from $Y_{0,1} \cup Y_{0,2} \cup \cdots \cup Y_{0,q-1}$, so as to make $|Y| = |X_0| = |X|$. Note that in either case we will arrive at a pair of sets X, Y satisfying

$$|\Sigma(S_m)|^{1-\varepsilon/400c} \ \leq \ |X| \ = \ |Y| \ \leq \ |\Sigma(S_m)|, \text{ and } E(X,Y) \ \geq \ |\Sigma(S_m)|^{3-\varepsilon/80c}, \tag{12}$$

where the first inequalities follow from (3) and (9).

Following the comment after Theorem 2, we have that there exists $\Sigma := X' \subseteq X \subseteq \Sigma(H'')$ satisfying $|\Sigma| \ \geq \ |X|^{1-\varepsilon/2c}$, such that

$$|\Sigma + \Sigma| \ \leq \ |\Sigma|^{1+\varepsilon/2c}. \tag{13}$$

Let H''' denote the set of all $h \ \in \ H''$, such that $\Sigma(h) \ \in \ \Sigma$. By (5), (6), and (12), we have that

$$\begin{aligned} |H'''| &\geq |\Sigma|(|H'| \cdot |\Sigma(H')|^{-1}/2) \\ &\geq |X|^{1-\varepsilon/2c}|H'| \cdot |\Sigma(S_m)|^{-1}/2 \\ &\geq |X|^{1-\varepsilon/2c}|X|^{-1/(1-\varepsilon/400c)}|H'|/2 \\ &\geq |X|^{-\varepsilon/c}|H'| \\ &\geq |A|^{k_m/2-4\delta_m-\varepsilon}, \end{aligned}$$

for $|A|$ sufficiently large.

By simple averaging, there is some vector $w \ \in \ A^{k_m/2-1}$, such that there are at least $|A|^{1-4\delta_m-\varepsilon}$ vectors $h \in H'''$ whose last $k_m/2-1$ coordinates are the vector w. The upshot of this is that if we let

$$A' \ := \ \{a \in A \ : \ aw \in H'''\},$$

then

$$|A'| \ \geq \ |A|^{1-4\delta_m-\varepsilon}, \tag{14}$$

and

$$A' + A' + 2\Sigma(w) \ \subseteq \ \Sigma(H''') + \Sigma(H''') \ = \ \Sigma + \Sigma.$$

Now we apply a weak form of the Ruzsa-Plünnecke Theorem [4], given as follows:

Theorem 5 *Suppose that X is some finite subset of an additive abelian group, such that*

$$|X+X| \ \leq \ C|X|.$$

Then, we have that

$$|kX| \ = \ |X+X+\cdots+X| \ \leq \ C^k|X|.$$

Using

$$X \ := \ \Sigma, \text{ and } C \ := \ |\Sigma|^{\varepsilon/2c},$$

we deduce that for ℓ even,

$$|\ell A'| \ \leq \ |\ell\Sigma| \ \leq \ |\Sigma|^{1+\varepsilon\ell/2c} \ \leq \ |A|^{c+\varepsilon\ell} \ \leq \ |A'|^{(c+\varepsilon\ell)/(1-4\delta_m-\varepsilon)}$$

By selecting $\delta > 0$ small enough (and therefore $\delta_m > 0$ small enough), relative to $\varepsilon > 0$, we can ensure that for $\varepsilon < 1/2$,

$$|\ell A'| \ \leq \ |A'|^{c(1+2\varepsilon\ell)}.$$

Of course, when $1/2 \leq \varepsilon < 1$ the inequality is trivial, as $c > 1$. Clearly, on rescaling ε appropriately, our theorem is proved.

2 Acknowledgements

We would like to thank Antal Balog, Harald Helfgott, Jozsef Solymosi, Terry Tao, and Van Vu for helpful comments.

References

[1] A. Balog, *Many additive quadruples*, CRM Proceedings and Lecture Notes in Additive Combinatorics, **43** (2007).

[2] A. Balog and E. Szemerédi, *A statistical theorem of set addition*, Combinatorica **14** (1994), 263-268.

[3] W. T. Gowers, *A new proof of Szemerédi's Theorem for progressions of length four*, Geom. Funct. Anal. **8** (1998), 529-551.

[4] I. Ruzsa, *Arithmetic progressions and the number of sums*, Periodica Math. Hung. **33** (1992), 105-111.

[5] J. Singer, *A theorem in finite projective geometry and some applications to number theory*, Trans. Amer. Math. Soc. **43** (1938), 377385.

[6] B. Sudakov, E. Szemerédi, and V. Vu, *On a question of Erdős and Moser*, Duke Math. Jour. **129** (2005), 129-155.

[7] T. Tao and V. Vu, *Additive Combinatorics*, Cambridge Univ. Press, 2006.