Monochromatic Corners on the Integer Lattice

Frank Xiao

UGA REU 2010

Theorem (Graham and Solymosi) *Given any integer $r > 0$, if the lattice points in the $N \times N$ grid are arbitrarily r-colored, and $N > 2^{2^{3r}}$, then there exist at least $\delta(r)N^3$ monochromatic "corners", i.e. triples of points $(x, y), (x + d, y), (x, y + d)$ for some $d > 0$, where $\delta(r) = (3r)^{-2^{r+2}}$.*

Proof:

The proof by Graham and Solymosi [1] proceeds in stages. In each stage we identify a new line in the $N \times N$ grid that contains at least some constant proportion of N^2 monochromatic corners. The goal is to then show, first, that we can guarantee finding such a line in each stage, and second, that the number of these lines that we can find is at least some constant proportion of N.

Stage 1

Suppose that the $N \times N$ integer lattice (that is the set $\{(x, y)|1 \leq x, y \leq N\}$) is arbitrarily r-colored. Let L_0 denote the line $x + y = N + 1$, and let T_0 be the bottom left triangle region of all points lying on or below L_0, so that $|T_0| = \binom{N+1}{2}$.

By the pigeonhole principle, some color, call it c_1, must occur at least $\frac{1}{r}\binom{N+1}{2} > \frac{1}{2r}N^2$ times in T_0. Again by pigeonholing, we get that some line, call it L_1, of the form $x + y = m$ (for $2 \leq m \leq N + 1$) must contain at least $\frac{1}{2r}N$ points of color c_1 on it. Let S_1 be the set of points lying on L_1 that are colored c_1, and let $s_1 = |S_1|$. Now define T_1, the "lower vertex region" of S_1, to be the set of vertices that form the bottom left vertex of the corners formed by all pairs of points in S_1, i.e.,

$$T_1 = \{(x, y) : \exists s, t \ni (x, t), (s, y) \in S_1, s > x\}$$

We say L_1 is *good* if T_1 has at least $\alpha(r)N^2$ color c_1 points in it, where α is a constant (to be determined later) that depends only on r. Note that if L_1 is good, then we are guaranteed the existence of at least αN^2 monochromatic corners, since each point in T_1 that's colored c_1 is the bottom left vertex of a c_1-colored corner with the two other vertices lying in S_1 that "spawned" it.

If L_1 is good, then we move on to Stage 2.

If L_1 is not good, then we know that T_1 has at least $\binom{s_1}{2} - \alpha N^2$ points not colored c_1. Again by applying the pigeonhole principle, we then know that some color that's **not** c_1, call it c_2, occurs at least $\frac{1}{r-1}\left[\binom{s_1}{2} - \alpha N^2\right]$ times. Some line of slope -1, call it L_2, lying parallel and below to L_1 must then contain at least $\frac{1}{(r-1)N}[\binom{s_1}{2} - \alpha N^2]$ points colored c_2. Let S_2 be the set of points lying on L_2 that are colored c_2, and let T_2 be the lower vertex region of S_2, so $|T_2| = \binom{s_2}{2}$.

We note here that $T_2 \subset T_1$. To verify this, let $p = (a, b)$ be any point in T_2. Then there exists an integer d_1 such that the pair of points $(a, b + d_1)$ and $(a + d_1, b)$ lie in S_2 - these are the points in S_2 that "spawned" p. But $S_2 \subseteq T_1$, so each of these two points must have been spawned by a pair of points in S_1. In particular, for the point $(a, b + d_1)$, there must exist an integer d_2 such that the points $(a, b + d_1 + d_2)$ and $(a + d_2, b + d_1)$ lie in $S_1 \subseteq L_1$. But since L_1 and L_2 both are lines of slope

-1, that same d_2 must work for the other point $(a+d_1, b)$, so that $(a+d_1, b+d_2)$ and $(a+d_1+d_2, b)$ lie in S_1. But then if we let $d = d_1 + d_2$, we have that (a, b) is the bottom left vertex of a corner formed with the points $(a, b+d)$ and $(a+d, b)$, both of which we have shown to lie in S_1, so that $p \in T_1$. Since p was arbitrary, $T_2 \subseteq T_1$.

We now have two possibilities, either L_2 is good, that is there exist at least αN^2 points in T_2 that are colored c_2, in which case we move on to Stage 2, or L_2 is not good.

If L_2 is not good, then we know that T_2 contains at least $\binom{s_2}{2} - 2\alpha N^2$ points not colored either c_1 or c_2 (since T_2 lies entirely in T_1, the worst case would be if all the points colored c_1 in T_1 were also in T_2). Applying the pigeonhole principle in the same manner as before, we know that some color, call it c_3, occurs at least $\frac{1}{r-2}[\binom{s_2}{2} - 2\alpha N^2]$ times in T_2. Some line, call it L_3, parallel and below to L_2, contains at least $\frac{1}{(r-2)N}[\binom{s_2}{2} - 2\alpha N^2]$ points colored c_3. Let S_3 be the set of points lying on L_3 that are colored c_3, and let T_3 be the lower vertex region of S_3, so $|T_3| = \binom{s_3}{2}$. A similar argument to the one in the previous step shows that $T_3 \subseteq T_2$.

We again now have two possibilities, either L_3 is good, in which case we move on to Stage 2, or L_3 is not good, in which case we move on to the next step.

At the kth step of this process (assuming we haven't found a good line in the previous $k-1$ steps), we will have two possibilities. Either L_k is good, or is not good. If L_k is good, we then move on to Stage 2. If L_k is not good, then we know that T_k contains at least $\binom{s_k}{2} - k\alpha N^2$ points not colored $c_1, c_2, \ldots, c_k$. We can apply the same pigeonholing argument to get that some color not equal to c_1 through c_k, call it c_{k+1}, occurs at least $\frac{1}{r-k}[\binom{s_k}{2} - k\alpha N^2]$ times. Then some line parallel and below to L_k, call it L_{k+1}, contains at least $\frac{1}{(r-k)N}[\binom{s_k}{2} - k\alpha N^2]$ points colored c_{k+1}. Let S_{k+1} be the set of points lying on L_{k+1} that are colored c_{k+1}, and let T_{k+1} be the lower vertex region of S_{k+1}. Note that $T_{k+1} \subseteq T_k \subseteq \ldots \subseteq T_1$.

Suppose we reach the rth step of this process, and so have not found a good line in any of the previous $r-1$ steps. Then we have two possibilities. If L_r is good, then we move on to Stage 2. If L_r is not good, then we know that T_r has at least $M_1 = \binom{s_r}{2} - r\alpha N^2$ points not colored $c_1, c_2, \ldots, c_r$. However, provided we choose α and N appropriately, i.e. α small enough and N large enough, we can then bound M_1 below such that $M_1 > 0$. But this would imply the existence of at least 1 point in T_r which has the property that it avoid all the colors $c_1, c_2, \ldots c_r$. But then we would have run out of colors for that point, which is impossible. Hence, somewhere in Stage 1, we must have found a good line, which we denote by L_1^*.

Stage 2

We start this stage by again restricting our attention to T_0, this time knowing the existence of a good line L_1^*. Some color, call it $c_1^{(2)}$, occurs in at least $\frac{1}{r}[\binom{N+1}{2} - N]$ points that do **not** lie on L_1^*. Hence, some line that is **not** L_1^*, call it $L_1^{(2)}$, must contain at least $\frac{1}{rN}[\binom{N+1}{2} - N]$ points colored $c_1^{(2)}$. Let $S_1^{(2)}$ be the set of points lying on $L_1^{(2)}$ that are colored $c_1^{(2)}$. Let $T_1^{(2)}$ be the lower vertex region of $S_1^{(2)}$.

We have two possibilities. If $L_1^{(2)}$ is good (that is $T_1^{(2)}$ contains at least αN^2 points colored $c_1^{(2)}$), then we go on to Stage 3.

If $L_1^{(2)}$ is not good, then we know that $T_1^{(2)}$ contains at least $\binom{s_1^{(2)}}{2} - \alpha N^2 - N$ points which do not lie on L_1^* and are not colored $c_1^{(2)}$. Some color that's not $c_1^{(2)}$, call it $c_2^{(2)}$, occurs at least $\frac{1}{r-1}[\binom{s_1^{(2)}}{2} - \alpha N^2 - N]$ times in $T_1^{(2)}$, so that some line that is **not** L_1^*, call it $L_2^{(2)}$, that is parallel and below to $L_1^{(2)}$, contains at least $\frac{1}{(r-1)N}[\binom{s_1^{(2)}}{2} - \alpha N^2 - N]$ points colored $c_2^{(2)}$. Let $S_2^{(2)}$ be the set of points lying on $L_2^{(2)}$ that are colored $c_2^{(2)}$. Let $T_2^{(2)}$ be the lower vertex region of $S_2^{(2)}$

We have two possibilities. If $L_2^{(2)}$ is good, then we move on to Stage 3. If $L_2^{(2)}$ is not good, then we know that $T_2^{(2)}$ contains at least $\binom{s_2^{(2)}}{2} - 2\alpha N^2 - N$ points which do not lie on L_1^* and are not colored $c_1^{(2)}$ or $c_2^{(2)}$. We can repeat this process to recursively define $c_k^{(2)}$, $L_k^{(2)}$, $S_k^{(2)}$, and $T_k^{(2)}$.

If we get to the rth step, then either $L_r^{(2)}$ is good, in which case we move on to Stage 3, or $L_r^{(2)}$ is not good, in which case we know that T_r contains at least $M_2 = \binom{s_r^{(2)}}{2} - r\alpha N^2 - N$ points which do not lie on L_1^* and are not colored $c_1^{(2)}$ through $c_r^{(2)}$. Provided we choose α and N appropriately such that we can bound $M_2 > 0$, we'd then know that there was at least one point in T_r that avoided all the colors, which is impossible. Hence, we can show that we must have found a good line in Stage 2.

Stage g

We can repeat this process, going through M stages and finding a *good* line in each stage, provided that α and N are chosen such that in the Mth stage, we have that

$$M_g = \binom{s_r^{(M)}}{2} - r\alpha N^2 - (g-1)N > 0$$

Here, the subtraction of $(g-1)N$ points is to make sure that we do not double count good lines, and so we subtract out a number greater than the maximum possible total number of points in the good lines already found.

The number of stages we go through, i.e., the number of good lines we find, we want to show is at least some constant proportion of N, say $g = \gamma(r)N$, where γ is a constant that depends only on r.

Computing α and γ

In general, in the gth phase, we want the $s_i^{(g)}$'s to satisfy the inequality, for $1 \leq i \leq r$

$$s_{i+1}^{(g)} \geq \frac{1}{((r-i)N} \left[\binom{s_i^{(g)}}{2} - i\alpha N^2 - gN \right]$$

(If we show that $s_{r+1}^{(g)} > 0$, then this would imply that $M_g > 0$, which would guarantee our good line).

For convenience, we drop the g in the superscript. Let $g = \gamma N$, and $s_i = \sigma_i N$. Then modifying our inequality slightly, we get, for $1 \leq i \leq r$,

$$\begin{aligned} s_{i+1} &\geq \frac{1}{((r-i)N} \left[\binom{s_i}{2} - i\alpha N^2 - gN \right] \\ &\geq \frac{1}{rN} \left[\binom{s_i}{2} - i\alpha N^2 - \gamma N^2 \right] \\ &\geq \frac{1}{rN} \left[\frac{s_i^2 - s_i}{2} - r\alpha N^2 - \gamma N^2 \right] \\ &> \frac{1}{rN} \left[\frac{s_i^2}{2} - (r+1)\alpha N^2 - \gamma N^2 \right] \end{aligned} \tag{1}$$

provided that $s_i < 2\alpha N^2$.

Let $\alpha(r) = \gamma(r) = (3r)^{-2^{r+1}}$. Let $\omega = (r+1)\alpha + \gamma$, so that $\omega = (r+2)(3r)^{-2^{r+1}}$. Then we can rewrite (1) in terms of σ_i's as

$$\sigma_{i+1} > \frac{1}{2r}\left(\sigma_i^2 - 2\omega\right) \tag{2}$$

provided that $\sigma_i < 2\alpha N$.

We can use induction (on i) to show that when $r \geq 2$ (in the case $r = 1$, we have a monochromatic $N \times N$ grid, where the result we want clearly holds), $\sigma_i > (3r)^{-(2^i-1)}$ for $1 \leq i \leq r+1$. For the base case $i = 1$, we have that $\sigma_i > \frac{1}{3r}$, which is true. Now for the inductive step, assume that it holds for k, then from (2), we have

$$\begin{aligned}
\sigma_{k+1} &> \frac{1}{2r}\left(\sigma_i^2 - 2\omega\right) \\
&> \frac{1}{2r}\left(\left(\frac{1}{(3r)^{2^i-1}}\right)^2 - 2\left(\frac{r+2}{(3r)^{2^{r+1}}}\right)\right) \\
&= \frac{3}{2}\left(\frac{1}{(3r)^{2^{i+1}-1}}\right) - \frac{1}{2r}\left(\frac{2(r+2)}{(3r)^{2^{r+1}}}\right) \\
&= \frac{1}{(3r)^{2^{i+1}-1}} + \frac{1}{2}\left(\frac{1}{(3r)^{2^{i+1}-1}} - \frac{\frac{2(r+2)}{r}}{(3r)^{2^{r+1}}}\right) \\
&> \frac{1}{(3r)^{2^{i+1}-1}}
\end{aligned}$$

If we take $N > (3r)^{2^{r+1}}$, then $2\alpha N > 2$, so that $\sigma_i < 2\alpha N$ and we are justified in using (2) in showing that $\sigma_i > (3r)^{-(2^i-1)}$. But then we have

$$\begin{aligned}
s_{r+1} &= \sigma_{r+1} N \\
&> \frac{1}{(3r)^{2^{r+1}-1}} \times (3r)^{2^{r+1}} \\
&= 3r \\
&> 1
\end{aligned}$$

which means that we must have found some good line in Stage g. Since we are guaranteed the existence of at least $g = \gamma N$ good lines, each of which corresponds to αN^2 monochromatic corners, we get that the number of monochromatic corners in the $N \times N$ grid is at least $\alpha\gamma N^3$, provided that $N > (3r)^{r^{2+1}}$. We can use induction to show that $2^{2^{3r}} > (3r)^{2^{r+1}}$, so that the theorem then follows.

An Alternative Method

Here we give another way of showing that the number of corners is at least some constant proportion of N^3, by adapting a method used by Varnavides [2] to prove that given a subset $A \subseteq [1, n]$ with density δ, the number of solutions to $x + y = 2z$ for $x, y, z \in A$ is at least some constant proportion of N^2, provided N is big enough. We use the following result proven by Graham and Solymosi [1]:

Lemma *Given any integer $r > 0$, if the lattice points in the $N \times N$ grid are arbitrarily r-colored, and $N > (2r)^{2^r}$, then there exists at least one monochromatic corner.*

Again look at the $N \times N$ grid, and suppose that it's arbitrarily r-colored. By the lemma, if we take $k = (2r)^{2^r}$, then in any $k \times k$ subgrid of our original $N \times N$ grid, there exists at least one monochromatic corner.

To get a lower bound on the total number of monochromatic corners in our original grid, we can count how many $k \times k$ subgrids there are in it. Let d denote the step size of our $k \times k$ subgrid. When $d = 1$, then there are a total of $(N-k+1)^2$ such subgrids, provided N is greater than or equal to k. If $d = 2$, then there are a total of $(N-[1+2(k-1)]+1)^2$ such subgrids, provided that $N \geq 2k$. The maximum size that d can be, so that a $k \times k$ subgrid with step size d can exist in our original grid, is $d = \lfloor \frac{N-1}{k-1} \rfloor$.

Our total number of $k \times k$ subgrids, call it K, is then

$$\begin{aligned} K &= (N-k+1)^2 + (N-2(k-1))^2 + \ldots + (N - \lfloor (N-1)/(k-1) \rfloor (k-1))^2 \\ &> \left\lfloor \frac{N-1}{k-1} \right\rfloor \left(N^2 - 2 \left\lfloor \frac{N-1}{k-1} \right\rfloor (k-1) \right) \end{aligned}$$

Each one of these subgrids contributes at least one monochromatic corner, however, we have to adjust for overcounting. Let $Q = (a,b), (a, b+d'), (a+d', b)$ be any monochromatic corner in our $N \times N$ grid. Then the number of $k \times k$ subgrids with step size $d = d'$ that contain Q is at most $(k-1)^2$. Any smaller value of d, which necessarily has to be a divisor of d', is bounded below with $(k-1)d > d'$ (otherwise the monochromatic corner couldn't possibly fit inside the $k \times k$ subgrid.

Hence, the maximum overcount factor is $k(k-1)^2$, and so the tot-al number of monochromatic corners C is bounded below by

$$C > \frac{1}{k(k-1)^2} \left\lfloor \frac{N-1}{k-1} \right\rfloor \left(N^2 - 2 \left\lfloor \frac{N-1}{k-1} \right\rfloor (k-1) \right)$$

The proportion here is approximately $\frac{1}{k^4}$, which in terms of r is $(2r)^{-2^{(r+2)}}$, which compared to the proportion from earlier $(3r)^{-2^{(r+2)}}$ has the same exponential form, but a different constant. The ratio of the original constant to the new one here is $(2/3)^{2^{r+2}}$, which goes to zero as r goes to infinity.

References

[1] R. Graham and J. Solymosi, Monochromatic equilateral right triangles on the integer grid. (2005).

[2] P. Varnavides, On certain sets of positive density, *J. London Math. Soc.* (1959).