The Green-Tao Theorem: An Exposition

by *David Conlon, Jacob Fox and Yufei Zhao*

presented by Martin Balko

The celebrated theorem of Green and Tao states that the prime numbers contain arbitrarily long arithmetic progressions. Incorporating discoveries that have been made since the original paper, the authors give an exposition of the proof with substantial simplifications in almost every aspect.

Theorem 1 (Green-Tao, 2004)**.** *The prime numbers contain arbitrarily long arithmetic progressions (APs).*

Proof Strategy of Green-Tao Theorem:

- An *upper density* of a set $A \subseteq \mathbb{N}$ is $\limsup\limits_{N\to\infty} |A \cap [N]|/N$. A *relative upper density* of a A in S is $\limsup\limits_{N\to\infty} |A \cap [N]|/|S \cap [N]|$.

  Theorem 2 (Szemerédi, 1975)**.** *Every subset of* $\mathbb{N}$ *with positive upper density contains arbitrarily long arithmetic progressions.*

- Use of *transference principle*: 1) show that if $S \subseteq \mathbb{N}$ satisfies certain pseudorandomness conditions, then every subset of S of positive relative upper density contains long APs (Relative Szemerédi Theorem). 2) construct a superset of primes that satisfies the conditions.

1)	Dense Model Theorem (Theorem 7) Counting Lemma (Theorem 9)	Relative Szemerédi Theorem (Theorem 6)	Green-Tao Theorem (Theorem 1)
2)		Constructing the Majorant (Proposition 10)	

Roth's Theorem via Graph Theory:

- For $A \subseteq \mathbb{Z}_N$ let G_A be a tripartite graph whose vertex sets are $X = Y = Z = \mathbb{Z}_N$ and edges E, where $(x, y) \in X \times Y$ is in E iff $2x + y \in A$, $(x, z) \in X \times Z$ is in E iff $x - z \in A$, and $(y, z) \in Y \times Z$ is in E iff $-y - 2z \in A$.

  Theorem 3 (Roth, 1952)**.** *If* $A \subseteq \mathbb{Z}_N$ *is 3-AP-free, then* $|A| = o(N)$.

Relative Roth Theorem:

- We write $\mathbb{E}_{x_1\in X_1,\dots,x_k\in X_k}$ as a shorthand for $|X_1|^{-1}\cdots|X_k|^{-1}\sum_{x_1\in X_1}\cdots\sum_{x_k\in X_k}$.

  Theorem 4 (Roth's Theorem, weighted version)**.** *For every* $\delta > 0$, *there exists* $c = c(\delta) > 0$ *such that every* $f\colon \mathbb{Z}_N \to [0, 1]$ *with* $\mathbb{E}f \geq \delta$ *satisfies* $\mathbb{E}_{x,d\in\mathbb{Z}_N}[f(x)f(x + d)f(x + 2d)] \geq c - o_\delta(1)$.

- Functional setting: replace the set S by a function (*majorizing measure*) $\nu\colon \mathbb{Z}_N \to [0, \infty)$ satisfying $\mathbb{E}\nu = 1 + o(1)$. Replace $A \subseteq S$ by a function $f\colon \mathbb{Z}_N \to [0, \infty)$ majorized by ν. Modify the graph G_S to a weighted graph G_ν with edge weights given by functions $\nu_{XY}, \nu_{XZ}, \nu_{YZ}$.

- Main motivating example: take $\nu(x) = \frac{N}{|S|}1_S(x)$ and $f(x) = \nu(x)1_A(x)$.

- A weighted tripartite graph ν with vertex sets X,Y, and Z satisfies the *3-linear forms condition* if $\mathbb{E}_{x,x'\in X,y,y'\in Y,z,z'\in Z}[\nu(y, z)\nu(y', z)\nu(y, z')\nu(y', z')\nu(x, z)\nu(x', z)\nu(x, z')\nu(x', z')\nu(x, y)\nu(x', y)\nu(x, y')\cdot \nu(x', y')] = 1 + o(1)$ and this also holds when one or more of the twelve ν factors are erased.

- A function $\nu\colon \mathbb{Z}_N \to [0,\infty)$ satisfies the *3-linear forms condition* if $\mathbb{E}_{x,x'\in X, y,y'\in Y, z,z'\in Z}[\nu(-y-2z)\nu(-y'-2z)\nu(-y-2z')\nu(-y'-2z')\nu(x-z)\nu(x'-z)\nu(x-z')\nu(x'-z')\nu(2x+y)\nu(2x'+y)\nu(2x+y')\nu(2x'+y')] = 1+o(1)$ and this also holds when one or more of the twelve ν factors are erased.

  Theorem 5 (Relative Roth)**.** *Suppose $\nu\colon \mathbb{Z}_N \to [0,\infty)$ satisfies the 3-linear forms condition. For every $\delta > 0$, there exists $c = c(\delta) > 0$ such that every $f\colon \mathbb{Z}_N \to [0,\infty)$ with $0 \le f \le \nu$ and $\mathbb{E}f \ge \delta$ satisfies $\mathbb{E}_{x,d\in\mathbb{Z}_N}[f(x)f(x+d)f(x+2d)] \ge c - o_\delta(1)$.*

Relative Szemerédi Theorem:

- Instead of constructing a weighted graph, we now construct a weighted $(k-1)$-uniform hypergraph corresponding to k-APs.

- A function $\nu\colon \mathbb{Z}_N \in [0,\infty)$ satisfies the *k-linear forms condition* if

$$\mathbb{E}_{x_1^{(0)},x_1^{(1)},\ldots,x_k^{(0)},x_k^{(1)}}\left[\prod_{j=1}^{k}\prod_{\omega\in\{0,1\}^{[k]\setminus\{j\}}}\nu\left(\sum_{i=1}^{k}(j-i)x_i^{(\omega_i)}\right)^{n_{j,\omega}}\right] = 1+o(1)$$

  for any choice of exponents $n_{j,\omega} \in \{0,1\}$.

  Theorem 6 (Relative Szemerédi)**.** *Suppose $\nu\colon \mathbb{Z}_N \to [0,\infty)$ satisfies the k-linear forms condition. For every $k \ge 3$ and $\delta > 0$, there exists $c = c(\delta) > 0$ such that every $f\colon \mathbb{Z}_N \to [0,\infty)$ with $0 \le f \le \nu$ and $\mathbb{E}f \ge \delta$ satisfies $\mathbb{E}_{x,d\in\mathbb{Z}_N}[f(x)f(x+d)f(x+2d)\cdots f(x+(k-1)d)] \ge c - o_{k,\delta}(1)$.*

Dense Model Theorem:

- Informally: it is possible to approximate an unbounded (or sparse) function f by a bounded (or dense) function $\tilde{f}$.

- A *cut norm* of an edge-weighted r-uniform hypergraph $g\colon X_1 \times \cdots \times X_r \to \mathbb{R}$ is

$$\|g\|_\square := \sup |\mathbb{E}_{x_1\in X_1,\ldots,x_r\in X_r}[g(x_1,\ldots,x_r)1_{A_1}(x_{-1})\cdots 1_{A_r}(x_{-r})]|,$$

  where the supremum is taken over all choices of subsets $A_i \subseteq X_{-i} := \prod_{j\in[r]\setminus\{i\}} X_j$, $i \in [r]$, and we write $x_{-i} := (x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_r) \in X_{-i}$. For a function $f\colon \mathbb{Z}_N \to \mathbb{R}$ we define $\|f\|_{\square,r} := \sup |\mathbb{E}[f(x_1+\cdots+x_r)1_{A_1}(x_{-1})\cdots 1_{A_r}(x_{-r})]|$, where the supremum is taken over all $A_1,\ldots,A_r \subseteq \mathbb{Z}_N^{r-1}$.

  Theorem 7 (Dense Model Theorem)**.** *For every $\varepsilon > 0$, there exists an $\varepsilon' > 0$ such that the following holds. Suppose $\nu\colon \mathbb{Z}_N \to [0,\infty)$ satisfies $\|\nu - 1\|_{\square,r} \le \varepsilon'$. Then, for every $f\colon \mathbb{Z}_N \to [0,\infty)$ with $f \le \nu$, there exists a function $\tilde{f}\colon \mathbb{Z}_N \to [0,1]$ such that $\|f - \tilde{f}\|_{\square,r} \le \varepsilon$.*

Counting Lemma:

- Informally: if two weighted graphs are close in cut norm, then they have similar triangle densities.

  Proposition 8 (Dense triangle counting lemma)**.** *Let g and $\tilde{g}$ be weighted tripartite graphs on $X \cup Y \cup Z$ with weights in $[0,1]$. If $\|g - \tilde{g}\|_\square$, then $|\mathbb{E}_{x\in X,y\in Y,z\in Z}[g(x,y)g(x,z)g(y,z) - \tilde{g}(x,y)\tilde{g}(x,z)\tilde{g}(y,z)]| \le 3\varepsilon$.*

  Theorem 9 (Relative triangle counting lemma)**.** *Let ν, g, $\tilde{g}$ be weighted tripartite graphs on $X \cup Y \cup Z$. Assume that ν satisfies the 3-linear forms condition, $0 \le g \le \nu$, and $0 \le \tilde{g} \le 1$. If $\|g - \tilde{g}\|_\square$, then $|\mathbb{E}_{x\in X,y\in Y,z\in Z}[g(x,y)g(x,z)g(y,z) - \tilde{g}(x,y)\tilde{g}(x,z)\tilde{g}(y,z)]| = o(1)$.*

Constructing the Majorant:

- Let $w = w(N)$ be any function that tends to infinity slowly with N and let $W = \prod_{p\le w} p$. We define the modified von Mangoldt function by $\tilde{\Lambda}(n) := \frac{\phi(W)}{W}\log(Wn+1)$ when $Wn+1$ is prime and 0 otherwise.

  Proposition 10. *For every $k \ge 3$, there exists $\delta_k > 0$ such that for every sufficiently large N there exists a function $\nu\colon \mathbb{Z}_N \to [0,\infty)$ satisfying the k-linear forms condition and $\nu(n) \ge \delta_k\tilde{\Lambda}(n)$ for all $N/2 \le n < N$.*