PCMI GRADUATE SUMMER SCHOOL 2025
PRELIMINARY NOTES

SARAH PELUSE

These are a preliminary version of my lecture notes, and may contain typos. The exercises in the problem sessions will have hints and intermediate steps not present in the exercises interspersed throughout the text. These are merely my lecture notes, and the exercise sheets provided during the problem sessions will contain the official exercises for this minicourse.

1. Coloring results

Arithmetic Ramsey theory is a branch of Ramsey theory that studies the emergence of arithmetic configurations in finite colorings, or large subsets of, groups or rings. The earliest results in Ramsey theory, Schur's theorem and van der Waerden's theorem, lie in this area.

For any $N \in \mathbf{N}$, we set $[N] := \{1, \ldots, N\}$.

Theorem 1 (van der Waerden's theorem, 1927)**.** *Let $r, k \in \mathbf{N}$. There exists $W = W(r,k) \in \mathbf{N}$ such that, whenever $N \geq W$, any r-coloring of $[N]$ contains a monochromatic k-term nontrivial arithmetic progression*

$$x, x+y, \ldots, x+(k-1)y \qquad (y \neq 0).$$

van der Waerden's theorem and its generalizations (e.g., the Gallai–Witt theorem and the Hales–Jewett theorem) were first proven using a technique called "color focusing", which usually produces very poor bounds for $W(r,k)$ and its analogues.

Exercise 1. Prove that Theorem 1 is equivalent to the following statement: If $\mathbf{N}$ is finitely colored, then some color class contains arithmetic progressions of all finite lengths.

Exercise 2. Construct a 2-coloring of $\mathbf{N}$ having no monochromatic infinite arithmetic progressions $\{a + qn : n\mathbf{N}\}$, $q \in \mathbf{N}$. Thus, the word "finite" in the statement in the previous exercise is necessary.

2. Density results

In 1936, Erdős and Turán conjectured that any subset of the natural numbers having positive (upper) density must contain arbitrarily long nontrivial arithmetic progressions. Thus, van der Waerden's theorem holds simply because, in any finite coloring, some color class must contain a positive proportion of the natural numbers. Roth proved this conjecture in the case of three-term arithmetic progressions in 1953, and the general case was settled by Szemerédi.

Theorem 2 (Szemerédi's theorem, 1975)**.** *Any subset $A \subset \mathbf{N}$ having positive upper density (i.e., $\limsup_N \frac{|A\cap[N]|}{N} > 0$) contains arbitrarily long nontrivial arithmetic progressions.*

Exercise 3. Prove that Szemerédi's theorem implies van der Waerden's theorem.

The above formulation of Szemerédi's theorem is equivalent to the following statement.

Theorem 3. *For all $k \in \mathbf{N}$, the following holds: If $A \subset [N]$ has no nontrivial k-term arithmetic progressions, then*

$$|A| = o_k(N).$$

We will now introduce the asymptotic notation used in these notes. We write $X_N = O(Y_N)$, $X'_N = \Omega(Y'_N)$, and $Z_N = o(W_N)$ mean, respectively, that there exist constants $C, C' > 0$ for which $|X_N| \leq CY_N$ and $|X'_N| \geq C'Y'_N$ for all $N \in \mathbf{N}$ sufficiently large and that $\lim_{N\to\infty} \frac{Z_N}{W_N} = 0$. We will frequently use Vinogradov's asymptotic notation, which is standard in analytic number theory. We write $X_N \ll Y_N$ for $X_N = O(Y_N)$, $X'_N \gg Y'_N$ for $X'_N = \Omega(Y'_N)$, and $X''_N \asymp Y''_N$ to mean that $Y''_N \ll X''_N \ll Y''_N$. We will write $O(Z_N)$, $\Omega(Z_N)$, and $o(Z_N)$ to mean quantities bounded above by CZ_N for all N sufficiently large, bounded below by $C'Z_N$ for all N sufficiently large, and whose quotient by Z_N tends to zero as N tends to ∞, respectively. When O, Ω, $\ll$, $\gg$, and $\asymp$, appear with subscripts, this means that the implied constants may depend on the parameters in the subscript, and when o appears with subscripts it means that the rate of decay to 0 may depend on the parameters in the subscript.

Theorem 4 (Bergelson–Leibman, 1996)**.** *Let $k, m \in \mathbf{N}$, $P_{i,j} \in \mathbf{Z}[y]$ with $P_{i,j}(0) = 0$ for all $(i,j) \in [k] \times [m]$, and $v_1, \dots, v_k \in \mathbf{Z}^d$. If $A \subset [N]^d$ contains no nontrivial configurations*

$$x, x + \sum_{i=1}^{k} P_{i,1}(y)v_i, \dots, x + \sum_{i=1}^{k} P_{i,m}(y)v_i,$$

then

$$|A| = o_{P_{1,1},\dots,P_{m,k},v_1,\dots,v_k}(N^d).$$

(1) The assumption that $P_{i,j}(0) = 0$ prevents congruence obstructions to the theorem holding. Indeed, the even numbers are a positive density subset of $\mathbf{N}$, but contain no progressions $x, x + 2y + 1$.
(2) Examples of configurations to which this result applies: the nonlinear Roth configuration $x, x+y, x+y^2$, corners $(x,y), (x,y+d), (x+d,y)$, and "sqorners" $(x,y), (x,y+d), (x+d^2,y)$.

(3) The proof of Bergelson and Leibman was via ergodic theory, and thus yields no quantitative bounds. Reasonable quantitative bounds are known in Szemerédi's original theorem, however.

Theorem 5 (Roth 1953 ($k = 3$), Gowers 1998/2001 ($k \geq 4$)). *If $A \subset [N]$ has no nontrivial k-term arithmetic progressions, then*

$$|A| \ll \frac{N}{(\log\log N)^{c_k}}.$$

(1) Reasonable bounds are also known for the size of subsets of $[N]^2$ lacking corners, thanks to work of Shkredov. All of these bounds have since been improved. It is a major open problem to prove reasonable bounds for other configurations, with the most ambitious goal being to prove a quantitative version of the Bergelson–Leibman theorem with reasonable bounds
(2) It is also of interest to do this in other settings (provided it makes sense), such as in cyclic groups, high-dimensional vector spaces over finite fields, and in nonabelian groups.

Theorem 6 (P.–Prendiville, 2019/2024). *If $A \subset [N]$ has no nontrivial nonlinear Roth configurations*

$$x, x+y, x+y^2,$$

then

$$|A| \ll \frac{N}{(\log\log N)^{c}}.$$

The upper bound can be improved to $\ll \frac{N}{(\log N)^c}$ by being a bit more careful (P.–Prendiville, 2020). The goal of these lectures is to give a proof of this theorem from scratch, introducing several key techniques in additive combinatorics along the way.

Additional notation: Whenever S is a finite nonempty set and $f : S \to \mathbf{C}$, $\mathbf{E}_{s\in S} f(s)$ means $\frac{1}{|S|}\sum_{s\in S} f(s)$ and f is said to be 1-*bounded* if $|f(s)| \leq 1$ for all $s \in S$. For any set $T \subset S$, we denote the indicator function of T by

$$1_T(t) = \begin{cases} 1 & t \in T \\ 0 & t \notin T \end{cases}.$$

For all $z \in \mathbf{C}$, we define $e(z) := e^{2\pi i z}$. For any $q \in \mathbf{N}$, we define $e_q(z) := e(z/q)$. Note that, as a function on $\mathbf{R}$, $e_q(z)$ is periodic with period q. Thus, $e_q(z)$ is also well-defined on $\mathbf{Z}/q\mathbf{Z}$.

3. Repeatedly applying the Cauchy–Schwarz inequality

We can formulate the problem of proving reasonable bounds in the Bergelson–Leibman theorem in various finite field model settings, such as by replacing $[N]$ itself with $\mathbf{F}_p$.

Theorem 7 (Bourgain–Chang, 2017)**.** *If $A \subset \mathbf{F}_p$ has no nontrivial nonlinear Roth configurations, then*

$$|A| \ll p^{14/15}.$$

(1) The theorem they prove is actually stronger, and guarantees a nontrivial copy of $x, x+y, x+y^2$ in $A \times B \times C$ for any three subsets $A, B, C \subset \mathbf{F}_p$ that are large enough.
(2) The exponent has since been improved, with the current best upper bound $|A| \ll p^{7/8}$, due to Kavrut and Wu.
(3) Bourgain and Chang proved their theorem via a Fourier-analytic argument using the explicit evaluation of quadratic Gauss sums

$$\sum_{x \in \mathbf{F}_p} e_p(ax^2 + bx + c)$$

that doesn't generalize to other (genuinely different) progressions or the integer setting.

One of the most important skills in the subfield of additive combinatorics that proves Szemerédi-type theorems is the ability to apply the Cauchy–Schwarz inequality. Repeatedly applying the Cauchy–Schwarz inequality followed by a judicious change of variables is surprisingly very powerful. We will illustrate the power of repeated applications of the Cauchy–Schwarz inequality by giving an alternative proof of a result of Shkredov concerning another three-term configuration mixing addition and multiplication:

Theorem 8 (Shkredov, 2010)**.** *Let $A, B, C \subset \mathbf{F}_p$. If $|A||B||C| \gg p^{5/2}$, then there exist $x, y \in \mathbf{F}_p$ such that $x \in A$, $x+y \in B$, and $xy \in C$.*

(1) Shkredov's original proof used Fourier analysis. Morally, one should be able to avoid its use, since $x, x+y, xy$ is a "complexity zero" configuration.
(2) We will actually obtain a slightly worse exponent of $11/4$, but by being more careful in argument below (keeping track of L^2-norms of various functions that arise out of applications of the Cauchy–Schwarz inequality), one can recover the exponent of $\frac{5}{2}$.

For any $f, g, h : \mathbf{F}_p \to \mathbf{C}$, define

$$\Lambda(f, g, h) := \mathbf{E}_{x,y \in \mathbf{F}_p} f(x) g(x+y) h(xy).$$

Thus, for any $A, B, C \subset \mathbf{F}_p$, we have

$$\Lambda(1_A, 1_B, 1_C) = \frac{\#\{(x,y) \in \mathbf{F}_p^2 : (x, x+y, xy) \in A \times B \times C\}}{p^2}.$$

Observe that $\Lambda(f, g, h)$ is a trilinear function of f, g, and h.

Theorem 9. *Let $f, g, h : \mathbf{F}_p \to \mathbf{C}$ be 1-bounded. Then, $\Lambda(f, g, h)$ equals*

$$\left(\mathbf{E}_{x \in \mathbf{F}_p} f(x)\right) \left(\mathbf{E}_{x \in \mathbf{F}_p} g(x)\right) \left(\mathbf{E}_{x \in \mathbf{F}_p} h(x)\right) + O\left(\frac{1}{p^{1/4}}\right).$$

Exercise 4. Prove that Theorem 9 implies the following statement: Let $A, B, C \subset \mathbf{F}_p$. If $|A||B||C| \gg p^{11/4}$, then there exist $x, y \in \mathbf{F}_p$ such that $x \in A$, $x + y \in B$, and $xy \in C$.

Our first reduction is to the case where g has mean $\mathbf{E}g$ zero.

Lemma 1. *Let $f, g, h : \mathbf{F}_p \to \mathbf{C}$ be 1-bounded. Then, $\Lambda(f, g, h)$ equals*

$$\left(\mathbf{E}_{x\in\mathbf{F}_p} f(x)\right)\left(\mathbf{E}_{x\in\mathbf{F}_p} g(x)\right)\left(\mathbf{E}_{x\in\mathbf{F}_p} h(x)\right) + \Lambda\left(f, g - \mathbf{E}g, h\right) + O\left(\frac{1}{p}\right).$$

Proof. Writing $g = g - \mathbf{E}g + \mathbf{E}g$, by trilinearity we have

$$\Lambda(f, g, h) = \Lambda(f, g - \mathbf{E}g, h) + \mathbf{E}g \cdot \Lambda(f, 1, h).$$

We have

$$\Lambda(f, 1, h) = \mathbf{E}_{x,y\in\mathbf{F}_p} f(x)h(xy) = \mathbf{E}_{x\in\mathbf{F}_p^\times}\mathbf{E}_{y\in\mathbf{F}_p} f(x)h(xy) + O\left(\frac{1}{p}\right)$$

since f and h are 1-bounded and $|\mathbf{F}_p^\times| = p - 1$. More generally, observe that restricting an average over $\mathbf{F}_p$ to one over $\mathbf{F}_p^\times$ or extending an average from $\mathbf{F}_p^\times$ to one over $\mathbf{F}_p$ introduces an error of magnitude $\ll p^{-1}$. Making the change of variables $y \mapsto \frac{y}{x}$ yields

$$\begin{aligned}\mathbf{E}_{x\in\mathbf{F}_p^\times}\mathbf{E}_{y\in\mathbf{F}_p} f(x)h(xy) &= \mathbf{E}_{x\in\mathbf{F}_p^\times}\mathbf{E}_{y\in\mathbf{F}_p} f(x)h(y)\\ &= \frac{p}{p-1}\left(\mathbf{E}f - \frac{f(0)}{p}\right)\mathbf{E}h\\ &= \mathbf{E}f \cdot \mathbf{E}h + O\left(\frac{1}{p}\right),\end{aligned}$$

again using that f and h are 1-bounded and $|\mathbf{F}_p^\times| = p - 1$. Plugging this back in above and using that g is 1-bounded completes the proof. □

It thus suffices to prove the following key upper bound.

Proposition 1. *Let $f, g, h : \mathbf{F}_p \to \mathbf{C}$ be 1-bounded with $\mathbf{E}g = 0$. Then,*

$$|\Lambda(f, g, h)| \ll \frac{1}{p^{1/4}}.$$

Proof. Arguing as in the proof of the previous lemma, we can write

$$\begin{aligned}\Lambda(f, g, h) &= \mathbf{E}_{x,y\in\mathbf{F}_p^\times} f(x)g(x+y)h(xy) + O\left(\frac{1}{p}\right)\\ &= \mathbf{E}_{x\in\mathbf{F}_p^\times} h(x)\mathbf{E}_{y\in\mathbf{F}_p^\times} f\left(\frac{x}{y}\right) g\left(\frac{x}{y} + y\right) + O\left(\frac{1}{p}\right)\end{aligned}$$

by making the change of variables $x \mapsto \frac{x}{y}$, and apply the Cauchy–Schwarz inequality to obtain that $|\Lambda(f, g, h)|^2$ is bounded above by

$$\mathbf{E}_{x\in\mathbf{F}_p^\times}\left|\mathbf{E}_{y\in\mathbf{F}_p^\times} f\left(\frac{x}{y}\right) g\left(\frac{x}{y} + y\right)\right|^2,$$

which, expanding the square to double the y variable, equals

$$\mathbf{E}_{x,y,z\in\mathbf{F}_p^\times} f\left(\frac{x}{y}\right)\overline{f}\left(\frac{x}{z}\right)g\left(\frac{x}{y}+y\right)\overline{g}\left(\frac{x}{z}+z\right), \tag{1}$$

plus an error that's $\ll p^{-1}$. Making the change of variables $x\mapsto xz$ and then setting $a=\frac{z}{y}$ in (1) yields

$$\mathbf{E}_{x,y,a\in\mathbf{F}_p^\times} f(xa)\overline{f}(x)g(xa+y)\overline{g}(x+ay).$$

Applying the Cauchy–Schwarz inequality again to double the y variable gives that $|\Lambda(f,g,h)|^4$ is bounded above by

$$\mathbf{E}_{x,a\in\mathbf{F}_p^\times}\left|\mathbf{E}_{y\in\mathbf{F}_p^\times} g\left(x+ya\right)\overline{g}\left(xa+y\right)\right|^2,$$

which, expanding the square, equals

$$\mathbf{E}_{x,a,y,z\in\mathbf{F}_p^\times} g\left(x+ya\right)\overline{g}\left(xa+y\right)\overline{g}\left(x+za\right)g\left(xa+z\right), \tag{2}$$

plus an error that's $\ll p^{-1}$. Since g is 1-bounded, we can extend the average above to one over all of $\mathbf{F}_p^4$ to get that (2) equals

$$\mathbf{E}_{x,a,y,z\in\mathbf{F}_p} g\left(x+ya\right)\overline{g}\left(xa+y\right)\overline{g}\left(x+za\right)g\left(xa+z\right), \tag{3}$$

up to an error that $\ll p^{-1}$.

Now, as (x,a,y,z) ranges over $\mathbf{F}_p^4$, the quadruple

$$(x+ya, xa+y, x+za, xa+z)$$

is essentially equidistributed over $\mathbf{F}_p^4$. Indeed, set

$$\mu(\mathbf{c})=\frac{\#\{(x,a,y,z)\in\mathbf{F}_p^4 : (x+ya,xa+y,x+za,xa+z)=\mathbf{c}\}}{p^4}$$

for all $\mathbf{c}\in\mathbf{F}_p^4$, so that (3) equals

$$\sum_{\mathbf{c}\in\mathbf{F}_p^4} g(c_1)\overline{g}(c_2)\overline{g}(c_3)g(c_4)\mu(\mathbf{c}),$$

and let $\tilde{\mu}$ denote the uniform measure on $\mathbf{F}_p^4$.

Exercise 5. Prove that

$$\sum_{\mathbf{c}\in\mathbf{F}_p^4}|\mu(\mathbf{c})-\tilde{\mu}(\mathbf{c})|\ll\frac{1}{p}.$$

Writing $\mu(\mathbf{c})=\mu(\mathbf{c})-\tilde{\mu}(\mathbf{c})+\tilde{\mu}(\mathbf{c})$, it thus follows that (3) equals

$$\begin{aligned}&\sum_{\mathbf{c}\in\mathbf{F}_p^4} g(c_1)\overline{g}(c_2)\overline{g}(c_3)g(c_4)(\mu-\tilde{\mu})(\mathbf{c})+\sum_{\mathbf{c}\in\mathbf{F}_p^4} g(c_1)\overline{g}(c_2)\overline{g}(c_3)g(c_4)\tilde{\mu}(\mathbf{c})\\ &=\sum_{\mathbf{c}\in\mathbf{F}_p^4} g(c_1)\overline{g}(c_2)\overline{g}(c_3)g(c_4)(\mu-\tilde{\mu})(\mathbf{c})+\left(\mathbf{E}g\right)^4.\end{aligned}$$

Thus, since $\mathbf{E}g = 0$ and g is 1-bounded,

$$\left|\sum_{\mathbf{c}\in\mathbf{F}_p^4} g(c_1)\overline{g}(c_2)\overline{g}(c_3)g(c_4)\mu(\mathbf{c})\right| \leq \sum_{\mathbf{c}\in\mathbf{F}_p^4} |(\mu-\tilde{\mu})(\mathbf{c})| \ll \frac{1}{p}.$$

Putting everything together, we deduce that $|\Lambda(f,g,h)|^4 \ll p^{-1}$, and so the conclusion of the lemma now follows. □

Exercise 6. Be more careful and obtain the exponent $\frac{5}{2}$ in Shkredov's result.

Exercise 7. Give a similar "Cauchy–Schwarz" proof of the Bourgain–Chang theorem.

If you manage to find a good enough solution to this last exercise, then you may be able to address an open problem of Manners, which is to (essentially) give a Fourier-free proof of reasonable bounds in the nonlinear Roth theorem.