Introducción al análisis de Fourier de orden superior

Pablo Candela (UAM, ICMAT)

Julia Wolf (Universidad de Cambridge)

AGRA 2021 (IMPA-ICTP)

Introducción

En la combinatoria aritmética se estudian las estructuras combinatorias que pueden existir en subconjuntos de grupos abelianos, bajo suposiciones que son típicamente débiles. El siguiente teorema es un resultado central en esta área.

Teorema (Szemerédi, 1975 [4])

Para cada entero positivo k y cada $\alpha > 0$, existe $N_0 > 0$ tal que se cumple lo siguiente. Para cada entero $N > N_0$, todo subconjunto $A \subset [N]$ de cardinal mayor que αN contiene una progresión aritmética de longitud k.

Denotamos por $[N]$ el intervalo $\{1, 2, \ldots, N\} \subset \mathbb{Z}$, y escribimos "PA-$k$" para "progresión aritmética de longitud k", que es el tipo de estructura estudiada en este caso. La suposición sobre A aquí es simplemente que su *densidad* $|A|/N$ es mayor que α en $[N]$ (para N suficientemente grande).

A menudo es útil reformular este resultado fundamental desde el punto de vista del *conteo* de las PA-k incluidas en A.

Notación: dados un conjunto finito X y una función $f : X \to \mathbb{C}$, definimos *el promedio de f sobre X* por $\mathbb{E}_{x\in X} f(x) = \frac{1}{|X|} \sum_{x\in X} f(x)$. Dado $A \subset X$, denotamos por 1_A la *función indicatriz* de A.

Ejemplo: para $A \subset [N]$ tenemos $\mathbb{E}_{x\in[N]} 1_A(x) = |A|/N$, la densidad de A en $[N]$.

Un enunciado equivalente al Teorema de Szemerédi

Sea $A \subset \mathbb{Z}_N = \mathbb{Z}/N\mathbb{Z}$. Entonces el conteo de cuántas PA-k hay en A se puede expresar, de modo normalizado (i.e. dividiendo por el número total de PA-k en $\mathbb{Z}_N$), como el promedio siguiente:

$$\mathbb{E}_{x,d \in \mathbb{Z}_N} 1_A(x)\, 1_A(x+d) \,\cdots\, 1_A(x+(k-1)d).$$

<u>Notación</u>: dadas funciones cualesquiera $f_1, \dots, f_k : \mathbb{Z}_N \to \mathbb{C}$, denotamos por $S_k(f_1, \dots, f_k)$ el promedio $\mathbb{E}_{x,d \in \mathbb{Z}_N} f_1(x)\, f_2(x+d) \cdots f_k(x+(k-1)d)$. Escribimos $S_k(f)$ cuando $f_1 = \cdots = f_k = f$.

Un resultado de Varnavides [6] (ver la hoja de ejercicios) dice que el Teorema de Szemerédi es equivalente al enunciado siguiente.

Teorema

Sea $k \in \mathbb{N}$ y $\alpha > 0$. Entonces existe $c = c(\alpha, k) > 0$ tal que para cada $N \in \mathbb{N}$ y cada función $f : \mathbb{Z}_N \to [0,1]$ con $\mathbb{E}_{\mathbb{Z}_N} f \geq \alpha$, tenemos $S_k(f) \geq c$.

Como un primer paso hacia este teorema, intentemos controlar el valor de $S_k(f)$ usando propiedades analíticas o combinatorias de f.

El caso $k = 3$ y el análisis de Fourier

A partir de ahora suponemos que N es un número primo impar.

En los años 1950, Roth observó lo siguiente [3]: dado $A \subset \mathbb{Z}_N$, el promedio $S_3(1_A)$ está controlado por la *transformada de Fourier* de 1_A, de un modo que permite demostrar el teorema de Szemerédi para $k = 3$.

Elementos básicos del análisis de Fourier discreto:
Denotemos por e la función $\mathbb{R}/\mathbb{Z} \to \mathbb{C}$, $\theta \mapsto \exp(2\pi i\theta)$.

- Los *caracteres* en $\mathbb{Z}_N$ son las funciones $e_r : \mathbb{Z}_N \to \mathbb{C}$, $x \mapsto e(\frac{rx}{N})$, $(r \in \mathbb{Z}_N)$.
  Forman el *grupo dual* $\widehat{\mathbb{Z}_N} \cong \mathbb{Z}_N$, identificando $e(\frac{rx}{N})$ con su *frecuencia* r.
- Equipemos $\mathbb{C}^{\mathbb{Z}_N}$ con un producto escalar $\langle f, g\rangle = \mathbb{E}_{x\in\mathbb{Z}_N} f(x)\overline{g(x)}$.
  Los caracteres forman una *base ortonormal* de $(\mathbb{C}^{\mathbb{Z}_N}, \langle\cdot,\cdot\rangle)$.
- La *transformada de Fourier* de $f \in \mathbb{C}^{\mathbb{Z}_N}$ es $\widehat{f} : \widehat{\mathbb{Z}_N} \to \mathbb{C}$, $r \mapsto \mathbb{E}_{x\in\mathbb{Z}_N} f(x)\overline{e(\frac{rx}{N})}$.
  El número $\widehat{f}(r) = \langle f, e_r\rangle$ es el *coeficiente de Fourier* de f de *frecuencia* r.
- *Teorema de inversión*: para cada $f \in \mathbb{C}^{\mathbb{Z}_N}$, $f(x) = \sum_{r\in\mathbb{Z}_N} \widehat{f}(r)\, e(\frac{rx}{N})$.
- Definimos las normas L^p y ℓ^p en $\mathbb{C}^{\mathbb{Z}_N}$ por $\|f\|_{L^p} = (\mathbb{E}_{x\in\mathbb{Z}_N}|f(x)|^p)^{1/p}$, $\|f\|_{\ell^p} = (\sum_{x\in\mathbb{Z}_N} |f(x)|^p)^{1/p}$, $p \in [1, \infty)$. Además, $\|f\|_\infty = \sup_x |f(x)|$.

- *Teorema de Plancherel*: para cada $f, g \in \mathbb{C}^{\mathbb{Z}_N}$, $\langle f, g\rangle = \langle \widehat{f}, \widehat{g}\rangle$.
  Un caso particular es el *Teorema de Parseval*: $\forall\, f \in \mathbb{C}^{\mathbb{Z}_N}$, $\|f\|_{L^2} = \|\widehat{f}\|_{\ell^2}$.

Estos fundamentos del análisis de Fourier se extienden a todo grupo abeliano finito Z. En particular, $\widehat{\mathrm{Z}} \cong \mathrm{Z}$. Para dar un isomorfismo explícito, el carácter $\chi \in \widehat{\mathrm{Z}}$ se identifica con una frecuencia única $r \in \mathrm{Z}$, a través de una *forma bilineal* $\mathrm{Z} \times \mathrm{Z} \to \mathbb{R}/\mathbb{Z}$ (generalizando la aplicación $(r, x) \mapsto \frac{r\,x}{N}$ que usamos en $\mathbb{Z}_N$).

Resultado básico [5, §4.1]: para todo grupo abeliano finito Z existe una forma bilineal simétrica no degenerada $\mathrm{Z} \times \mathrm{Z} \to \mathbb{R}/\mathbb{Z}$.

Los únicos ejemplos explícitos que usaremos en este curso son los siguientes:

- En $\mathbb{Z}_N$, como hemos visto, $r \cdot x = \frac{r\,x}{N} \bmod 1$ (considerando $\mathbb{R}/\mathbb{Z}$ como $[0, 1)$ con la operación $+ \bmod 1$).
- Sea $\mathbb{F}^n$ un espacio vectorial sobre un cuerpo finito $\mathbb{F}$ de orden p (para cualquier primo p). Entonces para $r = (r_1, \ldots, r_n), x = (x_1, \ldots, x_n) \in \mathbb{F}^n$, definimos $r \cdot x = \frac{1}{p}(r_1 x_1 + \cdots + r_n x_n) \bmod 1$.

El caso $k = 3$ y el análisis de Fourier

La observación de Roth fue la siguiente.

La transformada de Fourier controla la deviación del conteo de PA-3

Sea $A \subset Z$ con $|Z|$ impar, $|A| = \alpha|Z|$. Entonces $|S_3(1_A) - \alpha^3| \leq \sup_{r\neq 0} |\widehat{1_A}(r)|$.

Vamos a demostrar una versión algo más general. Para ello, denotemos por f_A la *función centrada* de A, i.e. $f_A(x) = 1_A(x) - \alpha$ (“centrada" porque $\mathbb{E}_{x\in Z} f_A(x) = 0$).

Entonces $S_3(1_A) = \mathbb{E}_{x,d\in Z}\, (f_A + \alpha)(x)\ \ (f_A + \alpha)(x + d)\ \ (f_A + \alpha)(x + 2d)$.

Es fácil ver que este promedio es igual a $\alpha^3 + S_3(f_A)$.

Observamos también que $\widehat{f_A}(r) = \widehat{1_A}(r)$ para $r \neq 0$, y que $\widehat{f_A}(0) = \mathbb{E} f_A = 0$.

Con todo esto, el enunciado anterior se deduce del resultado siguiente.

Proposición (La transformada de Fourier controla S_3)

Sean $f_1, f_2, f_3 : Z \to \mathbb{C}$. Entonces $S_3(f_1, f_2, f_3) = \sum_{r\in Z} \widehat{f_1}(r)\, \widehat{f_2}(-2r)\, \widehat{f_3}(r)$. Además, si $|Z|$ es impar y $\forall j \in [3]$, $\|f_j\|_\infty \leq 1$, entonces $|S_3(f_1, f_2, f_3)| \leq \min_{j\in[3]} \|\widehat{f_j}\|_\infty$.

Proposición (La transformada de Fourier controla S_3)

Sean $f_1, f_2, f_3 : Z \to \mathbb{C}$. *Entonces* $S_3(f_1, f_2, f_3) = \sum_{r\in Z} \widehat{f_1}(r)\, \widehat{f_2}(-2r)\, \widehat{f_3}(r)$. *Además, si* $|\,Z\,|$ *es impar y* $\forall\, j \in [3]$, $\|f_j\|_\infty \le 1$, *entonces* $|S_3(f_1, f_2, f_3)| \le \min_{j\in[3]} \|\widehat{f_j}\|_\infty$.

Prueba. Por la fórmula de inversión, $\forall\, j \in [3]$ tenemos $f_j(x) = \sum_{r_j\in Z} \widehat{f_j}(r_j) e(r_j \cdot x)$. Sustituyendo en $S_3(f_1, f_2, f_3) = \mathbb{E}_{x,d\in Z}\, f_1(x)\, f_2(x + d)\, f_3(x + 2d)$, vemos que

$$S_3(f_1, f_2, f_3) = \mathbb{E}_{x,d} \sum_{r_1,r_2,r_3} \widehat{f_1}(r_1)\, \widehat{f_2}(r_2)\, \widehat{f_3}(r_3)\, e\big(r_1 \cdot x + r_2 \cdot (x + d) + r_3 \cdot (x + 2d)\big)$$
$$= \sum_{r_1,r_2,r_3} \widehat{f_1}(r_1)\, \widehat{f_2}(r_2)\, \widehat{f_3}(r_3) \big[\mathbb{E}_{x,d}\, e\big(x \cdot (r_1 + r_2 + r_3)\big)\, e\big(d \cdot (r_2 + 2r_3)\big)\big].$$

Utilizando la ortonormalidad de los caracteres, se ve que el promedio interior es $\big[\mathbb{E}_x e(x \cdot (r_1 + r_2 + r_3))\big]\big[\mathbb{E}_d e(d \cdot (r_2 + 2r_3))\big] = 1_{r_1+r_2+r_3=0}\, 1_{r_2=-2r_3}$ $= 1_{r_1=r_3}\, 1_{r_2=-2r_3}$. Sustituyendo en S_3, vemos que

$$S_3(f_1, f_2, f_3) = \sum_{\substack{r_1,r_2,r_3:\\ r_1=r_3,\; r_2=-2r_3}} \widehat{f_1}(r_1)\ \widehat{f_2}(r_2)\ \widehat{f_3}(r_3) = \sum_r \widehat{f_1}(r)\ \widehat{f_2}(-2r)\ \widehat{f_3}(r).$$

Para la segunda afirmación, es suficiente dar la prueba para $j = 1$. Tenemos $|S_3(f_1, f_2, f_3)| \le \sum_r |\widehat{f_1}(r)|\, |\widehat{f_2}(-2r)|\, |\widehat{f_3}(r)| \le \|\widehat{f_1}\|_\infty \sum_r |\widehat{f_2}(-2r)|\, |\widehat{f_3}(r)|$.
La desigualdad de Cauchy-Schwarz implica $\sum_r |\widehat{f_2}(-2r)|\, |\widehat{f_3}(r)| \le \|\widehat{f_2}\|_{\ell^2}\, \|\widehat{f_3}\|_{\ell^2}$.
Por el Teorema de Parseval, esta cota es igual a
$\|f_2\|_{L^2}\ \|f_3\|_{L^2} \le \|f_2\|_\infty\ \|f_3\|_\infty \le 1$. $\square$

Se deduce de la linealidad de la aplicación $f \mapsto \widehat{f}$ que $\|\widehat{f}\|_\infty$ es una *norma* en $\mathbb{C}^Z$, a menudo denotada por $\|f\|_{u^2}$. Reformulemos nuestra proposición.

Proposición (La transformada de Fourier controla S_3)

Sean $f_1, f_2, f_3 : Z \to \mathbb{C}$, $\|f_i\|_\infty \leq 1$, $|Z|$ *impar. Entonces* $|S_3(f_1, f_2, f_3)| \leq \min_{j\in[3]} \|f_j\|_{u^2}$.

Este enunciado forma la base de la prueba de Roth del caso $k = 3$ del Teorema de Szemerédi. En efecto, dado un subconjunto $A \subset \mathbb{Z}_N$ de densidad α, aplicamos la proposición con $f_1 = f_2 = f_3 = f_A$ para obtener $|\alpha^3 - S_3(1_A)| \leq \|f_A\|_{u^2}$. (De hecho, nótese que obtenemos la cota mas potente $|\alpha^3 - S_3(1_A)| \leq \alpha\|f_A\|_{u^2}$.)
Esto implica que, si $\|f_A\|_{u^2}$ es pequeño en comparación con α^2, entonces A contiene aproximadamente $\alpha^3 N^2$ PA-3.
Desde aquí, se deduce el teorema usando la dicotomía siguiente:

- O bien: $\|f_A\|_{u^2} = \sup_{r\neq 0} |\widehat{1_A}(r)|$ es suficientement pequeño y concluimos que existe una PA-3 en A (para N suficientemente grande),
- O bien: existe $r \neq 0$ tal que $|\widehat{1_A}(r)|$ es al menos una fracción de α, y entonces se emplea un argumento de *incremento de densidad*, usando e_r para localizar una PA larga en la cual A tiene densidad notablemente mayor que α.

No entraremos en los detalles de esta prueba para $k = 3$, sino que nos centraremos en generalizar esta estrategia para $k > 3$.

El caso $k = 4$: el análisis de Fourier ya no nos sirve

El resultado que quisiéramos obtener, para poder trabajar con la dicotomía anterior en el caso $k = 4$, es el siguiente análogo de la proposicíon anterior:

$\forall\, \epsilon > 0$, $\exists\, \delta > 0$ tal que si $f_1, f_2, f_3, f_4 : \mathbb{Z}_N \to \mathbb{C}$ cumplen $\|f_j\|_\infty \leq 1$ y $\|\widehat{f_j}\|_\infty \leq \delta$ para cada $j \in [4]$, entonces $|\mathbb{E}_{x,d}\, f_1(x)\, f_2(x+d)\, f_3(x+2d)\, f_4(x+3d)| \leq \epsilon$.

Lamentablemente, esta proposición hipotética es falsa.

Ejemplo: sean $f_1(x) = e(\frac{x^2}{N})$, $f_2(x) = e(\frac{-3x^2}{N})$, $f_3(x) = e(\frac{3x^2}{N})$, $f_4(x) = e(\frac{-x^2}{N})$.

Observación: $\forall\, x, d \in \mathbb{Z}_N$, $\quad x^2 - 3(x+d)^2 + 3(x+2d)^2 - (x+3d)^2 = 0$.

Obtenemos entonces

$$S_4(f_1, f_2, f_3, f_4) = \mathbb{E}_{x,d\in\mathbb{Z}_N} e\big(\tfrac{1}{N}(x^2 - 3(x+d)^2 + 3(x+2d)^2 - (x+3d)^2)\big) = 1.$$

Sin embargo, para cada j tenemos $\|\widehat{f_j}\|_\infty \leq N^{-\frac{1}{2}}$ que tiende a 0 cuando $N \to \infty$. En efecto, $\forall\, r$,

$$|\widehat{f_1}(r)|^2 = \overline{\mathbb{E}_x e(\tfrac{x^2 - r\,x}{N})}\, \mathbb{E}_y e(\tfrac{y^2 - r\,y}{N}) = \mathbb{E}_{x,h}\, e(\tfrac{-x^2 + r\,x}{N})\, e(\tfrac{(x+h)^2 - r(x+h)}{N})$$

$$= \mathbb{E}_{x,h}\, e(\tfrac{(x+h)^2 - x^2 - rh}{N}) = \mathbb{E}_{x,h}\, e(\tfrac{2hx + h^2 - rh}{N}) \leq \mathbb{E}_h |\mathbb{E}_x e(\tfrac{2hx}{N})| = \mathbb{E}_h\, 1_{\{0\}}(h) = \tfrac{1}{N}.$$

En consecuencia, para estas funciones obtenemos $\|\widehat{f_j}\|_\infty = o_{N\to\infty}(1)$ para todo $j \in [4]$, pero $S_4(f_1, f_2, f_3, f_4) = 1$ permanece alejado de 0 cuando $N \to \infty$.

El ejemplo anterior muestra que la norma $\|\widehat{f}\|_\infty$ no es suficientemente sensible para controlar los promedios de PA-k para $k > 3$. ¿Qué alternativas tenemos?

Idea: identificar una cantidad (o función) asociada con la norma $\|\widehat{f}\|_\infty$, que tenga una generalización permitiendo el control de promedios sobre progresiones más largas. Con este objetivo, Gowers consideró la norma $\|\widehat{f}\|_{\ell^4} = \big(\sum_r |\widehat{f}(r)|^4\big)^{1/4}$.

Por un lado, si $\|f\|_\infty \leq 1$, entonces las normas $\|\widehat{f}\|_{\ell^4}$ y $\|\widehat{f}\|_\infty$ son aproximadamente equivalentes. De hecho, para todo f, tenemos las siguientes desigualdades (fáciles de verificar): $\|\widehat{f}\|_\infty^4 \leq \|\widehat{f}\|_{\ell^4}^4 \leq \|\widehat{f}\|_\infty^2$.

Por otro lado, se da la fórmula siguiente, válida en todo grupo abeliano finito Z:

$$\|\widehat{f}\|_{\ell^4} = \Big(\mathbb{E}_{x,h_1,h_2 \in Z}\, f(x)\, \overline{f(x+h_1)}\, \overline{f(x+h_2)}\, f(x+h_1+h_2)\Big)^{\frac{1}{4}}.$$

Para ambos enunciados, véase la hoja de ejercicios.

Las cuaternas $(x,\, x+h_1,\, x+h_2,\, x+h_1+h_2)$ se pueden considerar como *cubos* de dimensión 2 (o paralelogramos) en Z. Es natural tratar de generalizar esta fórmula, calculando el promedio sobre *cubos de dimensión superior* en Z.

Así, llegamos a la definición de las normas de Gowers en Z.

Las normas de uniformidad en un grupo abeliano finito Z

Para cada entero $d > 1$, existe una tal *norma de uniformidad* de orden d sobre $\mathbb{C}^{\mathsf{Z}}$, denotada por $\|\cdot\|_{U^d}$.

La norma U^2 es el promedio sobre cubos de dimensión 2, como acabamos de ver:

$$\|f\|_{U^2} = \Big(\mathbb{E}_{x,h_1,h_2\in \mathsf{Z}}\, f(x)\, \overline{f(x+h_1)}\, \overline{f(x+h_2)}\, f(x+h_1+h_2) \Big)^{\frac{1}{4}}.$$

La norma U^3 se obtiene calculando el promedio sobre cubos de dimensión 3:

$$\|f\|_{U^3} = \Big(\mathbb{E}_{x,h_1,h_2,h_3\in \mathsf{Z}}\, f(x)\, \overline{f(x+h_1)}\, \overline{f(x+h_2)}\, f(x+h_1+h_2) \\ \cdot \overline{f(x+h_3)}\, f(x+h_1+h_3)\, f(x+h_2+h_3)\, \overline{f(x+h_1+h_2+h_3)} \Big)^{\frac{1}{8}}.$$

La regla general ya se ve claramente. Definamos pues estas normas formalmente.

Definición (Normas de uniformidad en Z)

Para cada $d \geq 2$, la norma U^d se define, para $f \in \mathbb{C}^{\mathsf{Z}}$, por

$$\|f\|_{U^d} = \Big(\mathbb{E}_{x,h_1,h_2,\dots,h_d\in \mathsf{Z}} \prod_{v\in\{0,1\}^d} \mathcal{C}^{|v|} f(x+v_1h_1+\cdots+v_dh_d) \Big)^{\frac{1}{2^d}},$$

donde $\mathcal{C}f := \overline{f}$ es el operador de conjugación compleja, y $|v| := v_1+\cdots+v_d$.

Las normas de uniformidad en un grupo abeliano finito Z

Se puede dar también la siguiente **definición inductiva** de las normas U^d:

- Definimos la *seminorma* $\|f\|_{U^1} := |\mathbb{E}_{x\in \mathsf{Z}} f(x)|$.
- Denotamos por T el operador de traslación en $\mathbb{C}^{\mathsf{Z}}$, definido por $T^h f(x) := f(x+h)$ para todo $h \in \mathsf{Z}$. Entonces para todo $d \geq 2$, definimos

$$\|f\|_{U^d}^{2^d} = \mathbb{E}_{h\in\mathsf{Z}} \|f\, T^h\overline{f}\|_{U^{d-1}}^{2^{d-1}}. \qquad (*)$$

Por ejemplo, $\|f\|_{U^3}^8 = \mathbb{E}_{x,h_1,h_2,h}\, f(x)\overline{f(x+h)} \quad \overline{f(x+h_1)}\, f(x+h_1+h)$

$$\overline{f(x+h_2)}\, f(x+h_2+h) \quad f(x+h_1+h_2)\,\overline{f(x+h_1+h_2+h)}\Big)$$

$$= \mathbb{E}_h\Big(\mathbb{E}_{x,h_1,h_2}\, f\, T^h\overline{f}\,(x) \quad \overline{f}\, T^h f\,(x+h_1) \quad \overline{f}\, T^h f\,(x+h_2) \quad f\, T^h\overline{f}\,(x+h_1+h_2)\Big)$$

$$= \mathbb{E}_h \|f\, T^h\overline{f}\|_{U^2}^4 .$$

Observamos que $\|f\|_{U^1}^2 = \mathbb{E}_h\mathbb{E}_x f(x)\overline{f(x+h)}$. Aplicando Cauchy-Schwarz en h, obtenemos $\|f\|_{U^1}^2 \leq (\mathbb{E}_h|\mathbb{E}_x f(x)\overline{f(x+h)}|^2)^{1/2} = \|f\|_{U^2}^2$, i.e. $\|f\|_{U^1} \leq \|f\|_{U^2}$.
Usando Cauchy-Schwarz de modo similar en (*), se ve por inducción lo siguiente.

Las normas U^d forman una secuencia creciente: $\quad \forall\, d \geq 2, \quad \|f\|_{U^d} \leq \|f\|_{U^{d+1}}$.

Las normas de uniformidad en un grupo abeliano finito Z

A continuación establecemos otras propiedades fundamentales de las normas U^d. Comencemos definiendo, para cada $d \geq 2$, una operación análoga al producto interno $\langle f, g \rangle = \mathbb{E}_x f(x)\overline{g(x)}$.

Definición (Producto U^d en $\mathbb{C}^{\mathrm{Z}}$)

Sea $d \geq 2$ y sea $(f_v)_{v \in \{0,1\}^d}$ una colección de 2^d funciones $\mathrm{Z} \to \mathbb{C}$.
El *producto* U^d de estas funciones se denota por $\langle (f_v)_v \rangle_{U^d}$ y se define por

$$\langle (f_v)_v \rangle_{U^d} = \mathbb{E}_{x, h_1, h_2, \dots, h_d \in \mathrm{Z}} \prod_{v \in \{0,1\}^d} \mathcal{C}^{|v|} f_v(x + v_1 h_1 + \cdots + v_d h_d).$$

Por ejemplo, para $d = 2$, el producto U^2 de las funciones $f_{00}, f_{01}, f_{10}, f_{11} \in \mathbb{C}^{\mathrm{Z}}$ es

$$\langle f_{00}, f_{10}, f_{01}, f_{11} \rangle_{U^2} = \mathbb{E}_{x, h_1, h_2} \; f_{00}(x) \; \overline{f_{10}(x + h_1)} \; \overline{f_{01}(x + h_2)} \; f_{11}(x + h_1 + h_2).$$

Observación: si todas las f_v son iguales a una misma función f, entonces

$$\langle (f_v)_v \rangle_{U^d} = \|f\|_{U^d}^{2^d}.$$

Recordemos que la desigualdad triangular para $\| \cdot \|_{L^2}$ se deduce de la desigualdad de Cauchy-Schwarz. De modo similar, para demostrar la desigualdad triangular para $\| \cdot \|_{U^d}$, utilizaremos un resultado análogo a la desigualdad de Cauchy-Schwarz para el producto U^d.

Las normas de uniformidad en un grupo abeliano finito Z

Proposición (La desigualdad de Gowers-Cauchy-Schwarz)

Sea $d \geq 2$ *y* $(f_v)_{v \in \{0,1\}^d}$ *una colección de* 2^d *funciones* $\mathbb{Z} \to \mathbb{C}$. *Entonces*

$$|\langle (f_v)_{v \in \{0,1\}^d} \rangle_{U^d}| \leq \prod_{v \in \{0,1\}^d} \|f_v\|_{U^d}.$$

Para ver la idea principal de la prueba, es suficiente tratar el caso $d = 2$.
Sustituyendo $y = x + h_2$ en

$|\langle f_{00}, f_{10}, f_{01}, f_{11} \rangle_{U^2}| = |\mathbb{E}_{x,h_1,h_2}\, f_{00}(x)\, \overline{f_{10}(x+h_1)}\, \overline{f_{01}(x+h_2)}\, f_{11}(x+h_1+h_2)|$,

obtenemos $|\mathbb{E}_{h_1} \left(\mathbb{E}_x f_{00}(x)\, \overline{f_{10}(x+h_1)}\right) \overline{\left(\mathbb{E}_y f_{01}(y)\, \overline{f_{11}(y+h_1)}\right)}|$.

Aplicando Cauchy-Schwarz en h_1, vemos que esta expresión está acotada por la raíz de $\mathbb{E}_{h_1} \left|\mathbb{E}_x f_{00}(x)\, \overline{f_{10}(x+h_1)}\right|^2 \mathbb{E}_{h_1} \left|\mathbb{E}_y f_{01}(y)\, \overline{f_{11}(y+h_1)}\right|^2$. Esto es igual a

$\langle f_{00}, f_{10}, f_{00}, f_{10} \rangle_{U^2} \quad \langle f_{01}, f_{11}, f_{01}, f_{11} \rangle_{U^2}$. Repitiendo este argumento

en cada uno de los dos factores, el primero se acota por

$\langle f_{00}, f_{00}, f_{00}, f_{00} \rangle_{U^2}^{1/2} \quad \langle f_{10}, f_{10}, f_{10}, f_{10} \rangle_{U^2}^{1/2} = \|f_{00}\|_{U^2}^2\, \|f_{10}\|_{U^2}^2$,

y del mismo modo el segundo factor se acota por $\|f_{01}\|_{U^2}^2\, \|f_{11}\|_{U^2}^2$.

Concluimos que $|\langle f_{00}, f_{10}, f_{01}, f_{11} \rangle_{U^2}| \leq \|f_{00}\|_{U^2}\, \|f_{10}\|_{U^2}\, \|f_{01}\|_{U^2}\, \|f_{11}\|_{U^2}$.

Ahora estamos en condiciones para demostrar que la función $\|\cdot\|_{U^d}$ es una norma.

Proposición

Para cada $d \geq 2$, $\|\cdot\|_{U^d}$ es una norma en $\mathbb{C}^{\mathrm{Z}}$.

Prueba. Sabemos ya que $\|f\|_{U^d} \geq \|f\|_{U^2}$ y $\|f\|_{U^2} \geq \|\widehat{f}\|_\infty$. Por tanto está claro que $\|\cdot\|_{U^d}$ es no-negativa y no-degenerada. Para comprobar la desigualdad triangular, usamos

$$|\langle (f_v)_{v\in\{0,1\}^d}\rangle_{U^d}| \leq \prod_{v\in\{0,1\}^d} \|f_v\|_{U^d}. \qquad (*)$$

Sean $f, g : \mathrm{Z} \to \mathbb{C}$, y consideremos $\|f+g\|_{U^d}^{2^d} = \langle (f+g)_{v\in\{0,1\}^d}\rangle_{U^d}$. Este producto U^d es lineal en cada una de las coordenadas de $(f+g)_v$. Desarrollamos entonces $\|f+g\|_{U^d}^{2^d}$ en una suma de 2^d términos, cada uno de la forma $\langle (f_v)_v\rangle_{U^d}$ con $f_v = f$ o bien $f_v = g$. Aplicando $(*)$ a cada uno de estos términos, obtenemos

$$\|f+g\|_{U^d}^{2^d} \leq \sum_{r=0}^{2^d} \binom{2^d}{r} \|f\|_{U^d}^r \|g\|_{U^d}^{2^d-r} = (\|f\|_{U^d} + \|g\|_{U^d})^{2^d}.$$ □

Así pues, las normas U^d con $d > 2$ generalizan de manera natural la norma U^2.

Pero ¿para qué sirven estas normas en combinatoria aritmética?

Propósito de las normas de uniformidad

Las normas U^d nos permiten controlar promedios sobre estructuras más complejas que las progresiones aritméticas de longitud 3. Un ejemplo central en el contexto del teorema de Szemerédi es el control del promedio $S_k(f_1, \ldots, f_k) = \mathbb{E}_{x,d} f_1(x)\, f_2(x+d) \cdots f_k(x+(k-1)d)$ por la norma U^{k-1}.

Teorema (El teorema de Von Neumann generalizado)

Sea $k \geq 3$, y supongamos que $\mathrm{mcd}(|\mathrm{Z}|, (k-1)!) = 1$. Sean $f_1, ..., f_k : \mathrm{Z} \to \mathbb{C}$ con $\|f_i\|_\infty \leq 1$ para cada $i \in [k]$. Entonces $|S_k(f_1, \ldots, f_k)| \leq \min_{i\in[k]} \|f_i\|_{U^{k-1}}$.

Consecuencia combinatoria:
Para todo conjunto $A \subset \mathrm{Z}$ con $|A| = \alpha|\mathrm{Z}|$, se tiene $|S_k(1_A) - \alpha^k| \leq 2^k \|f_A\|_{U^{k-1}}$.

La demostración del teorema se realiza por inducción sobre k.

Ya hemos visto el caso $k = 3$: hemos demostrado que si $|\mathrm{Z}| > 2$ es impar, entonces $|S_3(f_1, f_2, f_3)| \leq \min_{i\in[3]} \|\widehat{f_i}\|_\infty$. Además, $\|\widehat{f_i}\|_\infty \leq \|\widehat{f_i}\|_{\ell^4} = \|f_i\|_{U^2}$. Por tanto $|S_3(f_1, f_2, f_3)| \leq \min_{i\in[3]} \|f_i\|_{U^2}$.

Teorema (El teorema de Von Neumann generalizado)

Sea $k \geq 3$*, y supongamos que* $\mathrm{mcd}(|\mathrm{Z}|, (k-1)!) = 1$*. Sean* $f_1, ..., f_k : \mathrm{Z} \to \mathbb{C}$ *con* $\|f_i\|_\infty \leq 1$ *para cada* $i \in [k]$*. Entonces* $|S_k(f_1, \ldots, f_k)| \leq \min_{i \in [k]} \|f_i\|_{U^{k-1}}$*.*

Prueba. $k = 3$ ✓. Para captar la idea del paso inductivo, consideremos el caso $k = 4$. Realizando un cambio de variables, vemos que es suficiente demostrar que $|S_4(f_1, ..., f_4)| \leq \min_{2 \leq i \leq 4} \|f_i\|_{U^3}$.

Aplicando la desigualdad de Cauchy-Schwarz a
$\left|\mathbb{E}_{x,d}\, f_1(x)\, f_2(x+d)\, f_3(x+2d)\, f_4(x+3d)\right|^2$ en x,

obtenemos la cota superior $\|f_1\|_{L_2}^2\, \mathbb{E}_x \left|\mathbb{E}_d\, f_2(x+d)\, f_3(x+2d)\, f_4(x+3d)\right|^2$.

Este promedio es igual a

$$\mathbb{E}_{x,d,d'}\; f_2(x+d)\, \overline{f_2(x+d')} \quad f_3(x+2d)\, \overline{f_3(x+2d')} \quad f_4(x+3d)\, \overline{f_4(x+3d')}.$$

Efectuando el cambio de variables $y = x + d$, $h = d' - d$, este promedio se convierte en

$$\mathbb{E}_{y,d,h}\;\; f_2(y)\, \overline{f_2(y+h)} \quad f_3(y+d)\, \overline{f_3(y+d+2h)} \quad f_4(y+2d)\, \overline{f_4(y+2d+3h)}.$$

Para cada $i \in [3]$ y cada $h \in \mathrm{Z}$, definimos $g_{i,h}(y) = f_{i+1}(y)\, \overline{f_{i+1}(y+ih)}$.

Entonces este último promedio es

$$\mathbb{E}_h\;\; \mathbb{E}_{y,d}\; g_{1,h}(y) \quad g_{2,h}(y+d) \quad g_{3,h}(y+2d) = \mathbb{E}_h\; S_3(g_{1,h}, g_{2,h}, g_{3,h}).$$

Teorema (El teorema de Von Neumann generalizado)

Sea $k \geq 3$*, y supongamos que* $\mathrm{mcd}(|\,Z\,|, (k-1)!) = 1$*. Sean* $f_1, ..., f_k : Z \to \mathbb{C}$ *con* $\|f_i\|_\infty \leq 1$ *para cada* $i \in [k]$*. Entonces* $|S_k(f_1, \ldots, f_k)| \leq \min_{i\in[k]} \|f_i\|_{U^{k-1}}$*.*

Continuación del caso $k = 4$: hasta ahora hemos establecido que $|S_4(f_1, f_2, f_3, f_4)|^2 \leq \mathbb{E}_h \ |S_3(g_{1,h}, g_{2,h}, g_{3,h})|$, habiendo definido, para cada $i \in [3]$ y $h \in Z$, la función $g_{i,h}(x) = f_{i+1}(x)\overline{f_{i+1}(x+ih)}$. Por la hipótesis inductiva (i.e. el caso $k = 3$) tenemos $\mathbb{E}_h \ |S_3(g_{1,h}, g_{2,h}, g_{3,h})| \leq \mathbb{E}_h \min_{i\in[3]} \|g_{i,h}\|_{U^2}$.
Es fácil ver que esto es como mucho $\min_{i\in[3]} \mathbb{E}_h \|g_{i,h}\|_{U^2}$, lo que, a su vez, está acotado por $\min_{i\in[3]} \left(\mathbb{E}_h \|g_{i,h}\|_{U^2}^4\right)^{1/4}$.
Puesto que $|\,Z\,|$ es coprimo con $(k-1)! = 6$, la aplicación $h \mapsto i\,h$ es una permutación de Z para cada $i \in [3]$. Por tanto, deducimos que
$$\min_{i\in[3]} \left(\mathbb{E}_h \|g_{i,h}\|_{U^2}^4\right)^{1/4} = \min_{i\in[3]} \left(\mathbb{E}_h \|f_{i+1}\, T^h \overline{f_{i+1}}\|_{U^2}^4\right)^{1/4}.$$
Recordando la definición inductiva de las normas U^d, vemos que esto es igual a $\min_{i\in[3]} \|f_{i+1}\|_{U^3}^2 = \min_{i\in\{2,3,4\}} \|f_i\|_{U^3}^2$. Esto completa la prueba del caso $k = 4$.
El paso inductivo para $k > 4$ es parecido. $\square$

Resumen

- La transformada de Fourier es provechosa en la combinatoria aritmética porque controla promedios tales como $S_3(f_1, f_2, f_3)$.
- La transformada de Fourier no es suficientemente sensible para controlar promedios sobre estructuras más complejas, como $S_k(f_1, \ldots, f_k)$ para $k > 3$.
- La norma U^2 está estrechamente relacionada con la transformada de Fourier, y se generaliza de manera natural a las normas U^d, $d > 2$.
- La norma U^{k-1} facilita el control de promedios de la forma $S_k(f_1, \ldots, f_k)$, $k > 3$ (teorema de Von Neumann generalizado).
- Para la norma U^2, tenemos $\|f\|_{U^2} \leq \|\widehat{f}\|_\infty^{1/2}$ para todo $f \in \mathbb{C}^Z$ con $|f| \leq 1$. De ahí, si $\|f\|_{U^2} \geq \eta > 0$, deducimos que existe un carácter $e_r \in \widehat{Z}$ tal que $|\langle f, e_r \rangle| \geq \eta^2$.

Para que las normas U^d sean de utilidad similar a la de la transformada de Fourier (en particular, para que posibiliten una prueba del teorema de Szemerédi), tenemos que conseguir una conclusión de utilidad similar a la del último punto anterior, a partir de la suposición que $f \in \mathbb{C}^Z$ satisface $\|f\|_{U^d} \geq \eta > 0$.

Esperaríamos poder decir que en este caso f tiene un producto interno no negligible con una función adecuadamente estructurada (análoga a un carácter). Un resultado de este tipo se conoce como un **teorema inverso para la norma U^d**.

Hacia un teorema inverso para la norma U^3

Objetivo: identificar una colección $\mathcal{F}_d$ de funciones $Z \to \mathbb{C}$ con las propiedades siguientes. Primero, las funciones en $\mathcal{F}_d$ deben ser suficientemente estructuradas (como los caracteres). Segundo, la colección debe ser suficientemente amplia para que si $f \in \mathbb{C}^Z$ con $|f| \leq 1$ cumple $\|f\|_{U^{d+1}} \geq \eta > 0$, entonces exista una función $Q \in \mathcal{F}_d$ tal que $|\langle f, Q\rangle| \geq c$, con $c = c(\eta) > 0$ *independiente de* $|Z|$.

Para $d = 1$, hemos visto que es suficiente la colección $\widehat{Z}$ de caracteres e_r, $r \in Z$, i.e. $\mathcal{F}_1 = \widehat{Z}$. ¿Qué sería una colección $\mathcal{F}_2$ adecuada para $\|\cdot\|_{U^3}$?

¿Podemos tomar como $\mathcal{F}_2$ otra vez la colección de caracteres? No. Para la norma U^3, esta familia no es suficientemente amplia para dar lugar a un teorema inverso.

Ejemplo: consideremos de nuevo $f(x) = e(x^2/N)$ sobre $\mathbb{Z}_N$. Tenemos $\|f\|_{U^3} = 1$, ya que

$$x^2 - (x+h_1)^2 - (x+h_2)^2 + (x+h_1+h_2)^2 \\ -(x+h_3)^2 + (x+h_1+h_3)^2 + (x+h_2+h_3)^2 - (x+h_1+h_2+h_3)^2 \equiv 0.$$

Sin embargo, tenemos $\sup_r |\langle f, e_r\rangle| \leq \frac{1}{\sqrt{N}}$ (como vimos anteriormente), lo cual tiende a 0 cuando $N \to \infty$.

El ejemplo anterior indica que es necesario añadir más funciones para obtener una familia $\mathcal{F}_2$ adecuada. ¿Pero qué funciones?

Las normas U^d desde el punto de vista del "cálculo discreto":
Sea Z un grupo abeliano finito. Para todo $h \in Z$ y $f \in \mathbb{C}^Z$, denotamos por $\Delta_h f$ la **derivada multiplicativa (discreta)** de f con diferencia h, definida por

$$\Delta_h f(x) = f(x+h)\overline{f(x)} = T^h f\, \overline{f}.$$

Es posible expresar $\|f\|_{U^d}$ en términos de derivadas multiplicativas de grado d:

$$\|f\|_{U^d}^{2^d} = \mathbb{E}_{h_1,\ldots,h_d,x\in Z}\, \Delta_{h_d}\cdots\Delta_{h_1} f(x).$$

En particular, supongamos que $f(x) = e(\phi(x))$ para cierta aplicación $\phi : Z \to \mathbb{R}/\mathbb{Z}$. Para $h \in Z$, sea $h\nabla$ el **operador diferencia**, con diferencia h, definido por $(h\nabla)\phi(x) = \phi(x+h) - \phi(x)$. Entonces el promedio anterior se convierte en

$$\mathbb{E}_{h_1,\ldots,h_d,x\in Z}\; e\Big((h_d\nabla)(h_{d-1}\nabla)\cdots(h_1\nabla)\phi(x)\Big).$$

Desde este punto de vista, los caracteres se pueden considerar *puntos extremos* para la norma U^2, en el sentido siguiente.

Hacia un teorema inverso para la norma U^3

Proposición (Los caracteres son máximos de la norma U^2)

Sea $f : Z \to \mathbb{C}$ *con* $\|f\|_\infty \leq 1$. *Entonces* $\|f\|_{U^2} = 1 \Leftrightarrow f(x) = e(\phi(x) + \theta)$, *donde* $\phi : Z \to \mathbb{R}/\mathbb{Z}$ *es un homomorfismo y* $\theta \in \mathbb{R}/\mathbb{Z}$ *es una constante.*

Prueba. Si $\|f\|_\infty \leq 1$ y $\|f\|_{U^2} = 1$, entonces f debe ser de módulo 1, así que $f(x) = e(\psi(x))$ para cierta función $\psi : Z \to \mathbb{R}/\mathbb{Z}$. Entonces la ecuación $\|f\|_{U^2} = 1$ es equivalente a la validez de la siguiente ecuación para todo $x, h_1, h_2 \in Z$:

$$(h_2 \nabla)(h_1 \nabla)\psi(x) = \psi(x + h_1 + h_2) - \psi(x + h_1) - \psi(x + h_2) + \psi(x) = 0.$$

A su vez, esto es equivalente a que ψ sea de la forma $\phi(x) + \psi(0)$ para cierto homomorfismo $\phi : Z \to \mathbb{R}/\mathbb{Z}$. □

Llamaremos a cualquier función $Z \to \mathbb{R}/\mathbb{Z}$ una **función de fase** sobre Z.

El resultado precedente se puede considerar como un caso extremo del teorema inverso para la norma U^2. Existe un enunciado análogo para la norma U^3, en el cual las funciones lineales $x \mapsto \phi(x) + \theta$ son reemplazadas por funciones de fase *cuadráticas*.

Nota: de ahora en adelante suponemos que $|Z|$ es impar y mayor que 3, y fijamos una forma bilineal símetrica no degenerada $Z \times Z \to \mathbb{R}/\mathbb{Z}$, $(x, y) \mapsto x \cdot y$.

Hacia un teorema inverso para la norma U^3

Un homomorfismo $M : \mathrm{Z} \to \mathrm{Z}$ es **autoadjunto** (o *simétrico*) si $Mx \cdot y = My \cdot x$ para todo $x, y \in \mathrm{Z}$.

Proposición (Las fases cuadráticas son máximos de la norma U^3)

Sea $f : \mathrm{Z} \to \mathbb{C}$ con $\|f\|_\infty \leq 1$. Entonces $\|f\|_{U^3} = 1$ si y solo si existen un homomorfismo autoadjunto $M : \mathrm{Z} \to \mathrm{Z}$, un $\xi \in \mathrm{Z}$ y un $\theta \in \mathbb{R}/\mathbb{Z}$ tales que, para todo $x \in \mathrm{Z}$, $f(x) = e\big(Mx \cdot x + \xi \cdot x + \theta\big)$. (*)

Véase la hoja de ejercicios.

Este resultado puede motivar la reflexión siguiente: a lo mejor, en lugar de las *fases lineales* $e(\phi(x) + \theta)$, la familia de *fases cuadráticas* de la forma (*) nos basta para un teorema inverso para la norma U^3. Desgraciadamente, para ciertos grupos esta familia aún no es suficientemente amplia para servir como familia $\mathcal{F}_2$.

Para ver esto, debemos considerar funciones de fase definidas **localmente**, i.e. sobre un subconjunto estricto de Z.

Usamos la notación siguiente: para $x, h_1, h_2, \ldots, h_d \in \mathrm{Z}$, escribimos $\mathrm{C}_{x,h_1,\ldots,h_d}$ para $(x + v \cdot (h_1, \ldots, h_d))_{v \in \{0,1\}^d}$, el cubo de dimensión d en $\mathrm{Z}^{\{0,1\}^d}$.

Por ejemplo $\mathrm{C}_{x,h_1,h_2} = (x,\ x + h_1,\ x + h_2,\ x + h_1 + h_2)$.

Hacia un teorema inverso para la norma U^3

Definición (Funciones de fase localmente lineales/cuadráticas)

Sea X un subconjunto de un grupo abeliano. Una función $\phi : X \to \mathbb{R}/\mathbb{Z}$ se dice **lineal sobre** X si
$(h_2 \nabla)(h_1 \nabla)\phi(x) = \phi(x + h_1 + h_2) - \phi(x + h_2) - \phi(x + h_1) + \phi(x) = 0$
para todo $C_{x,h_1,h_2} \in X^{\{0,1\}^2}$. La función ϕ se dice **cuadrática sobre** X si $(h_3 \nabla)(h_2 \nabla)(h_1 \nabla)\phi(x) = 0$ para todo $C_{x,h_1,h_2,h_3} \in X^{\{0,1\}^3}$.

Ahora podemos ver que las fases cuadráticas $Mx \cdot x + \xi \cdot x + \theta$ definidas *globalmente* no bastan para un teorema inverso para la norma U^3.

Ejemplo (Gowers, 2001): Sea $K = \lfloor\sqrt{N}\rfloor$, y sea $P \subset \mathbb{Z}_N$ la progresión aritmética de dimensión 2 definida por $P = \{x_1 + Kx_2 : -K/10 \leq x_1, x_2 \leq K/10\}$.

Sea $f(x) = e\big((x_1^2 + x_2^2)/N\big)\, 1_P(x)$. Se verifican las afirmaciones siguientes:

- $\phi(x) = (x_1^2 + x_2^2)/N$ es cuadrático sobre P.
- Tenemos $\|f\|_{U^3} \gg 1$ cuando $N \to \infty$, esencialmente porque las únicas contribuciones a $\|f\|_{U^3}$ se deben a cubos de dimensión 3 en $P^{\{0,1\}^3}$, y cada una de estas contribuciones vale 1, ya que ϕ es cuadrático sobre P.
- Toda fase cuadrática *global* sobre $\mathbb{Z}_N$ es de la forma $\psi(x) = \frac{r\,x^2}{N} + \frac{s\,x}{N} + \theta$, y $|\langle f, e(\psi)\rangle|^2 \leq \mathbb{E}_{h\in\mathbb{Z}_N}|\mathbb{E}_{x\in\mathbb{Z}_N}\Delta_h f(x)\, e(-2r\,h\,x/N)| = o(1)_{N\to\infty}$.

Hacia un teorema inverso para la norma U^3

Así pues, a la hora de definir $\mathcal{F}_2$ en $\mathbb{Z}_N$, debemos tener en cuenta fases cuadráticas definidas localmente sobre *progresiones aritméticas multidimensionales*. Resulta que con estas fases más generales, obtenemos por fin una colección $\mathcal{F}_2$ suficientemente amplia para un teorema inverso para la norma U^3 en $\mathbb{Z}_N$.

En el marco de este curso, será más conveniente concentrarnos en una versión del teorema inverso que es técnicamente menos exigente, a saber, la versión que se da en espacios vectoriales $\mathbb{F}^n$ sobre un cuerpo finito $\mathbb{F}$. (En un tal espacio, las PA multidimensionales se reemplazan por *subespacios afines*, más fáciles de manejar.)

Para enunciar el teorema, necesitamos la noción siguiente, que es válida en todo grupo abeliano finito Z.

Dado $X \subset \mathrm{Z}$, el **sesgo local polinomial de grado** d de $f : X \to \mathbb{C}$ es la cantidad $\|f\|_{u^d(X)} = \sup_{\phi: X \to \mathbb{R}/\mathbb{Z}} |\langle f, e(\phi)\rangle|$, donde el supremo se toma sobre todos los **polinomios** $\phi : X \to \mathbb{R}/\mathbb{Z}$ **de grado** $d-1$ en X, i.e. ϕ es tal que para todo cubo $\mathrm{C}_{x,h_1,\dots,h_d}$ de dimensión d en $X^{\{0,1\}^d}$, tenemos $(h_d \nabla) \cdots (h_1 \nabla) \phi(x) = 0$.

Nótese que el teorema inverso para U^2 es equivalente a la siguiente desigualdad: para todo $f : \mathrm{Z} \to \mathbb{C}$ con $|f| \leq 1$, se tiene $\|f\|_{U^2(\mathrm{Z})} \leq \|f\|_{u^2(\mathrm{Z})}^{1/2} = \|\widehat{f}\|_{\ell^\infty}^{1/2}$.

El teorema inverso para la norma U^3

Teorema inverso para U^3 en espacios vectoriales finitos

Sea $\mathbb{F}$ un cuerpo finito de orden primo e impar. Sea $f : \mathbb{F}^n \to \mathbb{C}$ con $|f| \leq 1$ tal que $\|f\|_{U^3} \geq \eta > 0$. Entonces existe un subespacio $W \leq \mathbb{F}^n$ con $\dim(W) \geq n - O(\eta^{-O(1)})$, tal que $\mathbb{E}_{y \in \mathbb{F}^n} \|f\|_{u^3(y+W)} \gg \eta^{O(1)}$.
En particular, existe $y \in \mathbb{F}^n$ tal que $\|f\|_{u^3(y+W)} = \Omega(\eta^{O(1)})$.

Comentarios:

- Este teorema fue demostrado por Green y Tao [2], en base a ideas de Gowers.
- Tiene aplicaciones combinatorias. En particular, como veremos más adelante, el teorema permite demostrar el teorema de Szemerédi para $k = 4$ en $\mathbb{F}^n$, usando un argumento de incremento de densidad.
- Existe un resultado análogo en $\mathbb{Z}_N$, pero es técnicamente bastante más complicado. Además, la generalización a U^d para $d > 3$ en términos de fases polinomiales localmente definidas no es obvia (Gowers demostró el teorema de Szemerédi usando las normas U^d en $\mathbb{Z}_N$, pero usando un enunciado más débil que el teorema inverso para U^d).

Prueba del teorema inverso para la norma U^3 en $\mathbb{F}^n$

Paso 1. Si $f : \mathbb{F}^n \to \mathbb{C}$ con $|f| \leq 1$ tiene norma U^3 grande, entonces el argumento complejo de la función $h \mapsto \Delta_h(f)$ es una función "bastante lineal", en el sentido que se parece a una función ξ cuyo grafo contiene muchas cuaternas aditivas.

Proposición

Sea $f : \mathbb{F}^n \to \mathbb{C}$ con $|f| \leq 1$ y $\|f\|_{U^3} \geq \eta$. Entonces existe $H \subset \mathbb{F}^n$ con densidad $|H|/|\mathbb{F}^n| \geq \eta^8/2$, y $\xi : H \to \mathbb{F}^n$ tal que $|\widehat{\Delta_h f}(\xi(h))| \geq \eta^4/\sqrt{2}$ para todo $h \in H$, y tal que el grafo Γ de ξ contiene al menos $\frac{\eta^{64}}{2^8}|\mathbb{F}^n|^3$ cuaternas aditivas (i.e. cuaternas (a_1, a_2, a_3, a_4) tales que $a_1 + a_2 = a_3 + a_4$).

Prueba. Recuerden la definición inductiva de U^3, $\|f\|_{U^3}^8 = \mathbb{E}_h \|\Delta_h f\|_{U^2}^4$. Entonces la suposición $\|f\|_{U^3} \geq \eta$ es equivalente a $\mathbb{E}_h \|\Delta_h f\|_{U^2}^4 \geq \eta^8$.

Por el teorema inverso para la norma U^2, tenemos $\mathbb{E}_h \|\widehat{\Delta_h f}\|_{\ell^\infty}^2 \geq \eta^8$.

Definimos el conjunto $H := \{h \in \mathbb{F}^n : \|\widehat{\Delta_h f}\|_{\ell^\infty}^2 \geq \eta^8/2\}$. Ya que $|f| \leq 1$, tenemos $\|\widehat{\Delta_h f}\|_{\ell^\infty}^2 \leq 1$, así que $\eta^8 \leq \mathbb{E}_h \|\widehat{\Delta_h f}\|_{\ell^\infty}^2 \leq |H|/|\mathbb{F}^n| + \eta^8/2$. Deducimos la cota inferior $|H|/|\mathbb{F}^n| \geq \eta^8/2$.

Por definición de $\|\widehat{\Delta_h f}\|_{\ell^\infty}$ y de H, existe una función $\xi : H \to \mathbb{F}^n$ que cumple $|\widehat{\Delta_h f}(\xi(h))|^2 = \left|\mathbb{E}_x \Delta_h f(x) e(-\xi(h) \cdot x)\right|^2 \geq \eta^8/2$ para todo $h \in H$.

Ahora mostramos que el grafo de ξ contiene muchas cuaternas aditivas, aplicando la desigualdad de Cauchy-Schwarz. Primero, observando la cota $|H|/|\mathbb{F}^n| \geq \eta^8/2$ para la densidad de H, y la condición $|\widehat{\Delta_h f}(\xi(h))|^2 \geq \eta^8/2$ para todo $h \in H$, obtenemos $\eta^{16}/4 \leq \mathbb{E}_h \left| \mathbb{E}_x T^h f(x) \overline{f(x)} e(-\xi(h) \cdot x) \right|^2 1_H(h)$

$$= \left| \mathbb{E}_{h,x,k} \overline{T^h f(x)} f(x) \overline{T^k f(x)} T^k(T^h f)(x) e(-\xi(h) \cdot k) 1_H(h) \right|.$$

Aplicando la desigualdad triangular en x, k (y $|f(x)\overline{T^k f(x)}| \leq 1$), deducimos que $\mathbb{E}_{x,k} \left| \mathbb{E}_h T^k(T^h f)(x) \overline{T^h f(x)} e(-\xi(h) \cdot k) 1_H(h) \right| \geq \eta^{16}/4$.

Aplicando ahora la desigualdad de Cauchy-Schwarz en x, k, obtenemos

$$\mathbb{E}_{x,k,h,\ell} T^k(T^{h+\ell} f)(x) \overline{T^{h+\ell} f(x)}\, e(-\xi(h+\ell) \cdot k)\, 1_H(h+\ell) \cdot \overline{T^k(T^h f)(x)} T^h f(x)\, e(\xi(h) \cdot k)\, 1_H(h) \geq \eta^{32}/16.$$

Multiplicando las dos funciones exponenciales, obtenemos $e\big((\ell \nabla)\xi(h) \cdot k\big)$.

Además, la función $T^k(T^{h+\ell} f)(x) \overline{T^{h+\ell} f(x) T^k(T^h f)(x)} T^h f(x)$ es igual a $g_{k,\ell}(y) := T^k(T^\ell f)(y) \overline{T^\ell f(y) T^k f(y)} f(y)$ con el cambio de variables $y = x + h$.

Deducimos que $\eta^{32}/16 \leq |\mathbb{E}_{y,k,h,\ell}\, g_{k,\ell}(y)\, e\big((\ell \nabla)\xi(h) \cdot k\big)\, 1_H(h+\ell) 1_H(h)|$

$\leq \mathbb{E}_{k,\ell} |\mathbb{E}_h\, e\big((\ell \nabla)\xi(h) \cdot k\big)\, 1_H(h+\ell) 1_H(h)|$.

Usando Cauchy-Schwarz en k, ℓ, y un cambio de variables similar, obtenemos

$\mathbb{E}_{h,k,\ell_1,\ell_2}\, e\big((\ell_2 \nabla)(\ell_1 \nabla)\xi(h) \cdot k\big) 1_H(h+\ell_1+\ell_2) 1_H(h+\ell_2) 1_H(h+\ell_1) 1_H(h) \geq \frac{\eta^{64}}{2^8}$.

¡Hemos terminado! ¿Por qué?

Consideremos el promedio sobre k como función en h, ℓ_1, ℓ_2:
$(h, \ell_1, \ell_2) \mapsto \mathbb{E}_k\, e\big((\ell_2\nabla)(\ell_1\nabla)\xi(h)\cdot k\big) 1_H(h+\ell_1+\ell_2) 1_H(h+\ell_2) 1_H(h+\ell_1) 1_H(h)$.
Esto es igual a 1 cuando $\xi(h+\ell_1+\ell_2) - \xi(h+\ell_2) - \xi(h+\ell_1) + \xi(h) = 0$ y $h+\ell_1+\ell_2, \quad h+\ell_2, \quad h+\ell_1, \quad h \in H$, y es igual a 0 en cualquier otro caso. $\square$

Paso 2. Sobre un subespacio de alta dimensión, la función ξ se puede remplazar por una aplicación (perfectamente) lineal.
Para demostrar esto, vamos a aplicar los tres resultados siguientes, que son herramientas muy conocidas y útiles en la combinatoria aritmética.

Teorema de Balog-Szemerédi-Gowers (BSG) (cotas debidas a Schoen, 2013). Existe $c > 0$ tal que, si $A \subset Z$ contiene al menos $\delta|A|^3$ cuaternas aditivas, entonces existe $B \subset A$ con $|B| \geq c\,\delta|A|$ y $|B - B| \leq c^{-1}\delta^{-4}|B|$.

Teorema de Chang en $\mathbb{F}^n$. Sea $\mathbb{F}$ un cuerpo de característica p. Si $A \subset \mathbb{F}^m$ contiene $\delta|A|^3$ cuaternas aditivas, entonces existe un subespacio $V \leq \mathbb{F}^m$ con $V \subseteq 2A - 2A$ y $|V| \geq p^{-O(\delta^{-O(1)})}|A|$.

Desigualdad de Plünnecke-Ruzsa. Sean A, B subconjuntos finitos no vacíos de un grupo abeliano tales que $|A + B| \leq C|A|$, para alguna constante $C \geq 1$. Entonces, para enteros $k, \ell \geq 0$ cualesquiera, tenemos $|kB - \ell B| \leq C^{k+\ell}|A|$.

Recuerden: el grafo $\Gamma \subset \mathbb{F}^n \times \mathbb{F}^n$ de ξ contiene al menos $\frac{\eta^{64}}{2^8}|\Gamma|^3$ cuaternas aditivas.

Lema. Existen un subespacio $V \subset \mathbb{F}^n \times \mathbb{F}^n$ con $\dim(V) \geq n - O(\eta^{-O(1)})$ y una pareja $(h_1, \xi_1) \in \mathbb{F}^n \times \mathbb{F}^n$ tales que $|\Gamma \cap (V + (h_1, \xi_1))| \gg \eta^{O(1)}|V|$.

Prueba. Aplicando BSG con $\delta = \frac{\eta^{64}}{2^8}$, existe $B \subset \Gamma$ con $|B - B| \ll \eta^{-O(1)}|B|$. En particular, B contiene $\gg \eta^{O(1)}|B|^3$ cuaternas aditivas (véanse los ejercicios).

El teorema de Chang nos aporta un subespacio $V \subset 2B - 2B$ de codimensión $O(\eta^{-O(1)})$. Ahora la desigualdad de Plünnecke-Ruzsa implica que $|B + V| \leq |B + 2B - 2B| \ll \eta^{-O(1)}|B|$. Entonces $|\Gamma \cap (B + V)| \geq |\Gamma \cap B| = |B| \gg \eta^{O(1)}|B + V|$.

Consideremos la partición de $B + V$ en $|B + V|/|V|$ clases laterales de V. Por el principio del palomar, deducimos que existe una clase lateral $V + (h_1, \xi_1)$ tal que $|\Gamma \cap (V + (h_1, \xi_1))| \gg \eta^{O(1)}|V|$, como queríamos demostrar. □

Podemos ahora demostrar el resultado principal de este paso.

Teorema. Sea $f : \mathbb{F}^n \to \mathbb{C}$ con $|f| \leq 1$ y $\|f\|_{U^3} \geq \eta > 0$. Entonces existen un subespacio $W_1 \leq \mathbb{F}^n$ con $\dim(W_1) \geq n - O(\eta^{-O(1)})$, $x_0 \in \mathbb{F}^n$, una aplicación lineal $M : W_1 \to \mathbb{F}^n$, y un $\xi_0 \in \mathbb{F}^n$, tales que $\mathbb{E}_{h \in W_1} |\widehat{\Delta_{x_0+h} f}\,(M(h) + \xi_0)| \gg \eta^{O(1)}$.

Prueba. Sea V el subespacio obtenido en el lema anterior, y escribamos $\Gamma_1 = \Gamma \cap (V + (h_1, \xi_1))$. Sea $V_0 = V \cap (\{0\} \times \mathbb{F}^n)$. Puesto que Γ es el grafo de una función, todas las sumas en $\Gamma_1 + V_0$ son distintas, así que $|\Gamma_1 + V_0| = |\Gamma_1||V_0|$. También tenemos $|\Gamma_1 + V_0| \leq |V|$, y por el lema anterior $|\Gamma_1| \gg \eta^{O(1)}|V|$.
Por tanto $|V_0| \ll \eta^{-O(1)}$.

Se deduce del álgebra lineal (véase la hoja de ejercicios) que existen un subespacio $W_1 \leq \mathbb{F}^n$ y una aplicación lineal $M : W_1 \to \mathbb{F}^n$ tales que el subespacio $V_1 := \{(h, M(h)) : h \in W_1\} \leq \mathbb{F}^n \times \mathbb{F}^n$ cumple $V_0 + V_1 = V$
(en particular, $\dim(W_1) = \dim(V_1) = \dim(V) - \dim(V_0) \geq n - O(\eta^{-O(1)})$).

Tomemos la partición de V en trasladados de V_1. La cota inferior para $|\Gamma_1|$ y el principio del palomar implican que para cierta pareja (x_0, ξ_0) tenemos $|\Gamma \cap \big(V_1 + (x_0, \xi_0)\big)| \gg \eta^{O(1)}|V_1|$. Por definición de Γ, esto es equivalente a

$$\mathbb{P}_{h \in W_1}\Big(h + x_0 \in H,\ \xi(h + x_0) = M(h) + \xi_0\Big) \gg \eta^{O(1)}.$$

Recuerden: por definición de ξ, tenemos $\widehat{\Delta_{h+x_0} f}(\xi(h + x_0)) \geq \frac{\eta^4}{2^{1/2}}$ para todo $h + x_0 \in H$.

Combinando todo, obtenemos $\mathbb{E}_{h \in W_1}|\widehat{\Delta_{h+x_0} f}\,(M(h) + \xi_0)|$

$$\geq\ \mathbb{P}_{h \in W_1}\Big(h + x_0 \in H,\ \xi(h + x_0) = M(h) + \xi_0\Big)\ \frac{\eta^4}{2^{1/2}}\ \gg\ \eta^{O(1)}. \qquad \square$$

Paso 3 (Argumento de simetría). Sea W el subespacio en el cual M es simétrico: $W = \{h \in W_1 : M(x) \cdot h = M(h) \cdot x \text{ for all } x \in W_1\} \leq W_1$.

Lema. Tenemos $|W| \gg \eta^{O(1)}|W_1|$, lo que implica $\dim(W) \geq n - O(\eta^{-O(1)})$.

Prueba. Comencemos por el resultado del paso 2, es decir la cota $\eta^{O(1)} \ll \mathbb{E}_{h\in W_1}|\widehat{\Delta_{x_0+h}f}\,(M(h)+\xi_0)|$.
Denotando por $\mathbf{b}$ una función cualquiera cuyo valor absoluto está acotado por 1, este promedio se puede expresar como $\mathbb{E}_{h\in W_1}\,\mathbf{b}(h)\,\mathbb{E}_{x\in\mathbb{F}^n}\,T^{h+x_0}f(x)\overline{f(x)}e(-(M(h)+\xi_0)\cdot x)$. Esta última expresión, a su vez, se puede escribir como $\mathbb{E}_{h\in W_1}\mathbb{E}_{x\in\mathbb{F}^n}\,\mathbf{b}(h)\,\mathbf{b}(x)\,\mathbf{b}(x+h)\,e(-M(h)\cdot x) \gg \eta^{O(1)}$.

Dividiendo $\mathbb{F}^n$ en clases laterales de W_1, el principio del palomar implica que existe $x_1 \in \mathbb{F}^n$ tal que $\mathbb{E}_{x,h\in W_1}\,\mathbf{b}(h)\,\mathbf{b}(x+x_1)\,\mathbf{b}(x+x_1+h)\,e(-M(h)\cdot(x+x_1)) \gg \eta^{O(1)}$.

Ahora que hemos fijado x_1, lo podemos ignorar usando la notación $\mathbf{b}$. Así, la desigualdad final se convierte en $\eta^{O(1)} \ll \mathbb{E}_{x,h\in W_1}\,\mathbf{b}(h)\,\mathbf{b}(x)\,\mathbf{b}(x+h)\,e(-M(h)\cdot x)$ $\leq \mathbb{E}_{h\in W_1}|\mathbb{E}_{x\in W_1}\,\mathbf{b}(x)\,\mathbf{b}(x+h)\,e(-M(h)\cdot x)|$, donde $\mathbf{b}(h)$ ha sido eliminado por la desigualdad triangular.

Aplicando Cauchy-Schwarz en $h \in W_1$, deducimos que $\eta^{O(1)} \ll \mathbb{E}_{x,y,h\in W_1}\,\mathbf{b}(x)\,\mathbf{b}(y)\,\mathbf{b}(x+h)\,\mathbf{b}(y+h)\,e(-M(h)\cdot(y-x))$.

El objetivo es colocar términos en x, y de tal manera que se pueda cuantificar la potencial asimetría de M. Con este fin, empleamos un cambio de variable $h = z - x - y$, lo que resulta en la desigualdad $\eta^{O(1)} \ll \mathbb{E}_{x,y,z \in W_1} \mathbf{b}(x,z)\, \mathbf{b}(y,z)\, e(-M(z-x-y) \cdot (y-x))$.

Observamos que $e(-M(z-x-y) \cdot (y-x)) = e(M(x) \cdot y - M(y) \cdot x)\, e(-M(z) \cdot y + M(y) \cdot y)\, e(M(z) \cdot x - M(x) \cdot x)$. Esto nos permite absorber más términos en las funciones $\mathbf{b}$: la desigualdad anterior se puede escribir $\eta^{O(1)} \ll \mathbb{E}_{x,y,z \in W_1} \mathbf{b}(x,z)\, \mathbf{b}(y,z)\, e(M(x) \cdot y - M(y) \cdot x)$.

Ahora quisiéramos eliminar las funciones acotadas $\mathbf{b}$. Con este fin, empleamos primero el principio del palomar relativo a la variable z, deduciendo que $\eta^{O(1)} \ll \mathbb{E}_{x,y \in W_1} \mathbf{b}(x)\, \mathbf{b}(y)\, e(M(x) \cdot y - M(y) \cdot x)$.
Aplicando Cauchy-Schwarz en x (parecido a la aplicacíon precedente), se deduce que $\eta^{O(1)} \ll \mathbb{E}_{x,y,h \in W_1} \mathbf{b}(y)\, \mathbf{b}(y+h)\, e(M(x) \cdot h - M(h) \cdot x)$.
Finalmente, aplicando la desigualdad triangular en y, h, eliminamos $\mathbf{b}(y)\mathbf{b}(y+h)$, obteniendo que $\eta^{O(1)} \ll \mathbb{E}_{h \in W_1} |\mathbb{E}_{x \in W_1} e(M(x) \cdot h - M(h) \cdot x)|$.

¡Hemos terminado! En efecto, observen que $\mathbb{E}_{x \in W_1} e(M(x) \cdot h - M(h) \cdot x) = 1_W(h)$ (ya que $x \mapsto e(M(x) \cdot h - M(h) \cdot x)$ es un carácter en W_1). Así la última desigualdad implica que $\eta^{O(1)} \ll \mathbb{E}_{h \in W_1} 1_W(h) = |W|/|W_1|$, y esto es lo que afirma el lema. $\square$

Ahora pasamos a la última etapa de la prueba.

Paso final. Comencemos de nuevo con la conclusión del paso 2, es decir $\eta^{O(1)} \ll \mathbb{E}_{h\in W_1} |\widehat{\Delta_{x_0+h} f}\,(M(h)+\xi_0)|$. Dividimos W_1 en clases laterales de W. Por el principio del palomar, existe $x_1 \in W_1$ tal que

$$\eta^{O(1)} \ll \mathbb{E}_{h\in W} |\,\widehat{\Delta_{x_0+x_1+h} f}\,(M(h+x_1)+\xi_0)|$$
$$= \mathbb{E}_{h\in W}\, \mathbf{b}(h)\, \mathbb{E}_{x\in\mathbb{F}^n}\, T^{x_0+x_1+h} f(x)\, \overline{f(x)}\, e(-(M(h+x_1)+\xi_0)\cdot x).$$

Usando la notación $\mathbf{b}$ como antes, simplificamos aún más esta expresión: noten que para todo r, x, $e(r\cdot x) = e(r\cdot(x+h))\overline{e(-r\cdot h)}$, luego la desigualdad anterior se escribe $\eta^{O(1)} \ll \mathbb{E}_{h\in W, x\in\mathbb{F}^n}\, \mathbf{b}(h)\, \mathbf{b}(h+x) \overline{f(x)}\, e(-M(h)\cdot x)$.

Dividiendo $\mathbb{F}^n$ en clases laterales $y+W$, $y\in\mathbb{F}^n$, vemos que

$$\eta^{O(1)} \ll \mathbb{E}_{y\in\mathbb{F}^n}\mathbb{E}_{h,x\in W}\, \mathbf{b}(h)\, \mathbf{b}(h+x+y)\, \overline{f(x+y)}\, e(-M(h)\cdot(x+y))$$
$$= \mathbb{E}_{y\in\mathbb{F}^n}\mathbb{E}_{h,x\in W}\, \mathbf{b}(h,y)\, \mathbf{b}(h+x,y)\, \overline{f(x+y)}\, e(-M(h)\cdot x).$$

Por la simetría $M(h)\cdot x = M(x)\cdot h$ para todo $x, h\in W$, y notando que p es impar, obtenemos $e(-M(h)\cdot x) = e\big(-\frac{1}{2}M(x+h)\cdot(x+h)\big)\, e\big(\frac{1}{2}M(x)\cdot x\big)\, e\big(\frac{1}{2}M(h)\cdot h\big)$.

Sustituyendo esto en la última desigualdad, y recolocando factores, obtenemos $\eta^{O(1)} \ll \mathbb{E}_{y\in\mathbb{F}^n}\mathbb{E}_{h,x\in W}\, \mathbf{b}(h,y)\, \mathbf{b}(h+x,y)\, \overline{f(x+y)}\, e\big(\frac{1}{2}M(x)\cdot x\big)$.

Para cada elemento y, escribimos $f_y(x) = \overline{f(x+y)}\, e\big(\frac{1}{2}M(x)\cdot x\big)$. Deducimos (véase la hoja de ejercicios) que $\eta^{O(1)} \ll \mathbb{E}_{y\in\mathbb{F}^n} \|f_y\|_{u^2(W)} \le \mathbb{E}_{y\in\mathbb{F}^n} \|f\|_{u^3(y+W)}$. $\square$

Incremento de densidad para $k = 4$ en $\mathbb{F}^n$

Supongamos que $A \subset \mathbb{F}^n$ con $|A|/|\mathbb{F}^n| = \alpha$ no contiene ninguna PA-4 no trivial. Entonces, por el teorema de Von Neuman generalizado, la función centrada f_A cumple $\|f_A\|_{U^3} \geq \eta(\alpha) > 0$. Ahora el teorema de Szemerédi para $k = 4$ se deduce del resultado siguiente.

Proposición. Sea $f : \mathbb{F}^n \to [-1, 1]$ con $\mathbb{E}_{\mathbb{F}^n} f = 0$ y $\|f\|_{U^3} \geq \eta > 0$. Entonces existen un subespacio $V \leq \mathbb{F}^n$ con $\dim V \geq n/2 - O(\eta^{-O(1)})$, y un elemento $x_0 \in \mathbb{F}^n$, tales que $\mathbb{E}_{x \in x_0 + V} f(x) \gg \eta^{O(1)}$.

Prueba. Sea $W \leq \mathbb{F}^n$ dado por el teorema inverso. Tenemos $\mathbb{E}_{y \in \mathbb{F}^n} \mathbb{E}_{x \in y+W} f(x) = \mathbb{E}_{\mathbb{F}^n} f = 0$. Notamos que $\max(t, 0) = (|t| + t)/2$, $\forall\, t \in \mathbb{R}$. Deducimos que $\mathbb{E}_{y \in \mathbb{F}^n} |\mathbb{E}_{x \in y+W} f(x)| = 2\mathbb{E}_{y \in \mathbb{F}^n} \max(\mathbb{E}_{y+W} f, 0)$. Podemos suponer que $\mathbb{E}_{y+W} f \leq \eta^C / C$ para una constante $C > 0$ de cualquier tamaño fijo que queramos (porque en caso contrario hemos terminado ya la prueba). Se deduce que $\mathbb{E}_{y \in \mathbb{F}^n} |\mathbb{E}_{x \in y+W} f(x)| \leq 2\eta^C / C$. Eligiendo C suficientemente grande, y aplicando el teorema inverso, deducimos que
$\mathbb{E}_{y \in \mathbb{F}^n} \big(\|f\|_{u^3(y+W)} - 2|\mathbb{E}_{x \in y+W} f(x)|\big) \gg \eta^{O(1)}$.
Entonces existe $y \in \mathbb{F}^n$ tal que $\|f\|_{u^3(y+W)} \geq 2|\mathbb{E}_{x \in y+W} f(x)| + \Omega(\eta^{O(1)})$. Sin pérdida de generalidad supongamos que $y = 0$. Por definición de u^3, existen una aplicación lineal autoadjunta $M : W \to W$ y $\xi \in W$ tales que

$$|\mathbb{E}_{x \in W} f(x) e(-M(x) \cdot x - \xi \cdot x)| \geq 2|\mathbb{E}_W(f)| + \Omega(\eta^{O(1)}).$$

Recuerden: tenemos $|\mathbb{E}_{x\in W} f(x) e(-M(x)\cdot x - \xi\cdot x)| \geq 2|\mathbb{E}_W f| + \Omega(\eta^{O(1)})$.

Formemos la partición de W en conjuntos de nivel

$$S_j = \left\{x \in W : M(x)\cdot x + \xi\cdot x = \tfrac{j}{p}\right\},\ j \in [p].$$

Por la desigualdad triangular, $\sum_{j=1}^{p} |\mathbb{E}_{x\in W} 1_{S_j}(x) f(x)| \geq 2|\mathbb{E}_W f| + \Omega(\eta^{O(1)})$.

Utilizamos otra vez la identidad $\max(t,0) = \frac{1}{2}(|t| + t)$, deduciendo así que

$\sum_{j=1}^{p} \max(\mathbb{E}_{x\in W} 1_{S_j}(x) f(x),\, 0) \geq \frac{1}{2}\sum_{j=1}^{p} |\mathbb{E}_{x\in W} 1_{S_j}(x) f(x)| + \frac{1}{2}\mathbb{E}_W f$
$\geq |\mathbb{E}_W f| + \frac{1}{2}\mathbb{E}_W f + \Omega(\eta^{O(1)}) \geq \Omega(\eta^{O(1)}) = \Omega(\eta^{O(1)}) \sum_{j=1}^{p} \frac{|S_j|}{|W|}$.

Por el principio del palomar, existe j tal que $\mathbb{E}_{x\in W} 1_{S_j}(x) f(x) \gg \eta^{O(1)} \frac{|S_j|}{|W|}$.

Formamos ahora una partición de W en subespacios afines adecuados.
Sea $U \leq W$ de cardinal máximo tal que $Mx \cdot y = 0$ para todo $x, y \in U$.
Tenemos $\dim(U) \geq \frac{1}{2}\dim(W) - \frac{3}{2} \geq \frac{n}{2} - O(\eta^{-O(1)})$ (véanse los ejercicios).

La partición en cuestión divide W en clases laterales de U. Por el principio del palomar, existe $x_1 + U$ tal que $\mathbb{E}_{x\in x_1+U} 1_{S_j}(x) f(x) > \Omega(\eta^{O(1)})\, \mathbb{P}_{x_1+U}(S_j)$.

Sea $V = S_j \cap (x_1 + U)$. Entonces $\mathbb{E}_{x\in V} f(x) \gg \eta^{O(1)}$, en particular $V \neq \emptyset$.
Observamos que V es un subespacio afín (¿por qué?), y que además
$\dim V \geq \dim(U) - 1 \geq \frac{n}{2} - O(\eta^{-O(1)})$. $\square$

References

[1] W. T. Gowers, *A new proof of Szemerédi's theorem*, GAFA **11** (2001), 465–588.

[2] B. Green, T. Tao, *An inverse theorem for the Gowers U^3-norm*, Proc. Edinb. Math. Soc. (1) **51** (2008) 73–153.

[3] K. F. Roth, *On certain sets of integers*, J. London Math. Soc. **28** (1953), 104–109.

[4] E. Szemerédi, *On sets of integers containing no k elements in arithmetic progressions*, Acta Arith. **27** (1975), 199–245.

[5] T. Tao, V. Vu, *Additive combinatorics*, Cambridge University Press, 2006.

[6] P. Varnavides, *On certain sets of positive density*, J. London Math. Soc. **34** (1959), 358–360.