# UNIVERSIDAD NACIONAL DE INGENIERÍA

## FACULTAD DE CIENCIAS

SECCIÓN DE POSGRADO Y SEGUNDA ESPECIALIZACIÓN PROFESIONAL



# EL TEOREMA DE SZEMERÉDI, CONSECUENCIAS EN LA DISTRIBUCIÓN DE NÚMEROS PRIMOS Y PERSPECTIVAS

TESIS

PARA OPTAR EL GRADO ACADÉMICO DE MAESTRO EN CIENCIAS CON MENCIÓN EN MATEMÁTICA APLICADA


ELABORADO POR

ANDRÉS VICENTE CHULLUNCUY CENTENO

ASESOR

DR. OSWALDO JOSÉ VELÁSQUEZ CASTAÑÓN



LIMA-PERÚ

2014

*Agradezco a mi familia, a mi asesor Dr. Oswaldo Velásquez por su gran ayuda, y a mis profesores y amigos del IMCA y de la UNI.*

# Índice general

**Resumen**

El teorema de Szemerédi clásico muestra la existencia de progresiones aritméticas de longitud arbitraria en subconjuntos de los números naturales que tengan densidad superior positiva. En este trabajo abordamos varias pruebas del teorema de Szemerédi a través de distintos marcos teóricos: análisis de Fourier, teoría ergódica, y normas de Gowers. Inicialmente estudiamos el teorema de Szemerédi para progresiones aritméticas de longitud tres, es decir, el teorema de Roth, por medio del análisis de Fourier en $\mathbb{Z}_N$ y argumentos de *incrementos de la densidad*, y hacemos uso de las normas de Gowers para controlar el número de progresiones aritméticas de longitud tres. Estas normas simplifican significativamente los cálculos obtenidos por uso exclusivo del análisis de Fourier en $\mathbb{Z}_N$, a costa de una ralentización del incremento de la densidad. A pesar de ello, vía las normas de Gowers se obtiene el caso general del teorema de Szemerédi para progresiones aritméticas de longitud mayor o igual que tres. Abordaremos el teorema de Szemerédi para progresiones de longitud cuatro haciendo uso de las normas de Gowers, y en seguida probaremos el caso general del teorema de Szemerédi por medio de la teoría ergódica. Finalmente, siguiendo los trabajos de Green y Tao, mostraremos que los números primos contienen progresiones aritméticas de longitud arbitraria. Dado que el conjunto de los números primos tiene densidad superior nula dos hechos son relevantes para demostrar el teorema de Green y Tao: una importante variante del teorema de Szemerédi en el ámbito de medidas pseudoaleatorias, y la existencia de una medida pseudoaleatoria que mayore cierto conjunto de números primos.

**Abstract**

The classical Szemerédi's theorem shows the existence of arbitrarily long arithmetic progressions in subsets of natural numbers that have positive upper density. In this paper we address several proofs of the Szemerédi's theorem through different theoretical frameworks: Fourier analysis, ergodic theory, and Gowers norms. Initially we study Szemerédi's theorem for arithmetic progressions of length three, known as Roth's theorem, through Fourier analysis in $\mathbb{Z}_N$ and incremental density arguments, and we make use of Gowers norms to control the number of arithmetic progressions of length three. These norms significantly simplify the calculations obtained by the exclusive use of Fourier analysis in $\mathbb{Z}_N$, while slowering the density increment. However, by the use of Gowers norms we obtain the general case of Szemerédi's theorem for arithmetic progressions of lenght greater or equal than three. We deal with Szemerédi's theorem of progressions of length four by making use of Gowers norms, and then we prove the general case of Szemerédi's theorem with ergodic theory. Finally, following the work of Green and Tao, we show that the set of prime numbers contain arbitrarily long arithmetic progressions. Because the set of primes numbers has upper density zero, two facts are relevant to demonstrate Green-Tao theorem: an important variant of Szemerédi's theorem in the context of pseudorandom measures, and the existence of a pseudorandom measure that bounds a certain set of prime numbers.

**Notaciones**

| | |
|---|---|
| $\mathbb{Z}$: | conjunto de los números enteros. |
| $\mathbb{N}$: | conjunto de los números naturales, $\mathbb{N} = \{1, 2, 3, \ldots\}$. |
| $\mathbb{R}$: | conjunto de los números reales. |
| $\mathbb{R}^+$: | conjunto de los números reales no negativos. |
| $\mathbb{C}$: | conjunto de los números complejos. |
| $\mathbb{Z}_N$: | anillo de enteros módulo $N$. |
| $\lvert G\rvert$: | cardinalidad del conjunto $G$. |
| $\widehat{f}$: | transformada de Fourier de $f$ |
| $[1, N]$: | conjunto de enteros de 1 a $N$ |
| $f * g$: | convolución de las funciones $f$ y $g$ |
| $\lvert\lvert\alpha\rvert\rvert$: | resto de $\alpha$ módulo 1 |
| $a \equiv b \pmod p$: | $a$ congruente con $b$ módulo $p$. |
| $f \ll g$: | $f = O(g)$. |
| $f = O(g)$: | existen $R > 0$ y $K > 0$ tales que si $x \geq R$ entonces $\lvert f(x)\rvert \leq Kg(x)$. |
| $f(x) \sim g(x)$: | $\lim\limits_{x\to\infty} \dfrac{f(x)}{g(x)} = 1.$ |
| $C(X)$: | conjunto de funciones continuas en $X$. |
| $C^k(X)$: | conjunto de funciones de clase $C^k$ en $X$. |
| $\mathrm{diam}(B)$: | diámetro del conjunto $B$. |
| $a\lvert b$: | el entero $a$ divide a $b$. |
| $f = o(g)$: | $\lim\limits_{x\to+\infty} \dfrac{f(x)}{g(x)} = 0.$ |
| $\mathbb{E}(f)$: | valor promedio de la función $f : \mathbb{Z}_N \to \mathbb{R}$. |

# Capítulo 1

# Introducción

Un importante problema en combinatoria aditiva es determinar si existen progresiones aritméticas de determinada longitud en un conjunto dado $A$ que posea alguna estructura aditiva, como por ejemplo en $\mathbb{Z}$ con la suma usual. En este trabajo discutiremos quizás uno de los teoremas más profundos en combinatoria aditiva, el "teorema de Szemerédi", así como una de sus principales consecuencias indirectas, la existencia de progresiones aritméticas propias de longitud arbitraria formadas solamente por números primos, que fue mostrada por Ben Green y Terence Tao en 2004. Nuestro punto de partida es el siguiente resultado.

**Teorema 1.1** (Szemerédi, 1975)**.** *Sea $A$ un subconjunto de los enteros positivos con densidad superior positiva $\delta(A) > 0$. Entonces $A$ contiene progresiones aritméticas de longitud arbitraria.*

Para precisar lo que es densidad superior damos la definición.

**Definición 1.1.** *Sea $A \subseteq \mathbb{N}$, la* densidad superior *de $A$ en $\mathbb{N}$, denotada por $\delta(A)$, se define como*

$$\delta(A) = \limsup_{N\to\infty} \frac{|A \cap [1, N]|}{N},$$

*donde $[1, N] = \{1, 2, \ldots, N\}$, y $|B|$ denota la cardinalidad del conjunto $B$.*

Este teorema fue probado originalmente por Szemerédi en 1975 a través de un sofisticado argumento combinatorio, introduciendo por primera vez el *lema de regularidad* que hoy lleva su nombre. Existen varias pruebas de este teorema, entre ellas la de Furstenberg, que usa métodos de teoría ergódica [17], y la de Tim Gowers que hace uso de la combinatoria aditiva

[18]. Podemos reformular de una manera más cuantitativa el teorema de Szemerédi a partir de la siguiente definición.

**Definición 1.2** (Constante de Erdös-Turán)**.** *Sea $A$ un subconjunto finito de los enteros positivos, y sea $k \geq 1$. Denotemos por $r_k(A)$ a la cardinalidad del mayor subconjunto de $A$ que no contiene ninguna progresión aritmética propia de longitud $k$.*

Tenemos que $r_1(A) = 0$ y $r_2(A) = 1$ para cualquier conjunto $A$ finito y no vacío. Para cada $k$ fijo, $r_k(A)$ es creciente en $A$ en el sentido de que si $A \subseteq B$ entonces $r_k(A) \leq r_k(B)$. Además, $r_k(A) \leq |A|$ para cualquier conjunto $A$ finito y no vacío.
El teorema 1.1 es equivalente al siguiente enunciado, que fue inicialmente conjeturado por Erdös y Turán.

**Teorema 1.2** (Szemerédi, $2^{da}$ versión)**.** *Sean $k \geq 1$ y $N \geq 1$ enteros. Entonces $r_k\big([1, N]\big) = o(N)$.*

Para $k = 1$ o $k = 2$ el resultado es trivial. Sin embargo, para valores de $k$ mayores que dos el resultado no se da fácilmente. El caso $k = 3$ es bien manejado por métodos de análisis de Fourier, y tenemos en este caso el famoso teorema de Roth, (cf. [3]).

**Teorema 1.3** (Roth, 1956)**.** *Tenemos que $r_3\big([1, N]\big) = o\big(N\big)$.*

Los dos últimos teoremas tienen una variedad de pruebas diferentes, las que usan técnicas enmarcadas dentro de diversas áreas de la matemática tales como: análisis armónico, teoría ergódica, teoría de grafos, teoría de hipergrafos, combinatoria aditiva y teoría de Ramsey. El propósito de este trabajo es proporcionar varias de estas pruebas. Estas pruebas sirven como modelo para el teorema de Green-Tao.
Es preciso mencionar que el análisis de Fourier lineal es una herramienta bien adaptada para detectar progresiones aritméticas de longitud tres; sin embargo, para progresiones de mayor longitud requeriremos de un análisis de Fourier de mayor orden.

Otra conjetura debida a Erdös es la siguiente.

**Conjetura 1.1** (de Erdös sobre progresiones aritméticas)**.** *Sea $A = \big\{a_1, a_2, a_3, \dots\big\}$ un subconjunto infinito de $\mathbb{N}$, tal que $\displaystyle\sum_{k=1}^{+\infty} \frac{1}{a_k} = +\infty$. Entonces, $A$ contiene infinitas progresiones aritméticas de longitud $k$ para todo $k \geq 3$ entero.*

Esta conjetura está actualmente abierta, aún no resuelta, incluso para progresiones de longitud $k = 3$. Sin embargo, un caso especial de esta conjetura, restringida al conjunto de los números primos $\mathcal{P} = \{2, 3, 5, 7, 11, \ldots\}$ ha sido probada en años recientes por Green y Tao.

**Teorema 1.4** (Green-Tao, 2004)**.** *Sean $k \geq 1$ y $N \geq 1$ enteros. Entonces tenemos que*

$$r_k\big(\mathcal{P} \cap [1, N]\big) = o\big(|\mathcal{P} \cap [1, N]|\big).$$

*En particular, los números primos contienen infinitas progresiones aritméticas propias de longitud arbitraria.*

Mostraremos en el capítulo final el teorema de Green-Tao (teorema 7.1). Más aún, es posible mostrar un resultado más general siguiendo su método.

**Teorema 1.5** (Szemerédi en los primos, 2004)**.** *Sea $A$ un subconjunto cualquiera del conjunto de los números primos $\mathcal{P}$ con densidad superior positiva, esto es,*

$$\limsup_{N\to\infty} \frac{\big|A \cap [1, N]\big|}{\big|\mathcal{P} \cap [1, N]\big|} > 0.$$

*Entonces $A$ contiene infinitas progresiones aritméticas de longitud $k$ para todo $k \geq 3$ entero.*

Si cambiamos el conjunto de los números primos $\mathcal{P}$, por el conjunto de los números naturales $\mathbb{N}$, obtenemos el teorema de Szemerédi clásico, teorema 1.1, que reescribimos ahora.

**Teorema 1.6** (Szemerédi, versión clásica)**.** *Sea $A \subseteq \mathbb{N}$ con densidad superior positiva, esto es*

$$\limsup_{N\to+\infty} \frac{|A \cap [1, N]|}{N} > 0,$$

*entonces $A$ contiene infinitas progresiones aritméticas de longitud $k$ para todo entero $k \geq 3$.*

Resaltamos otras dos versiones del teorema de Szemerédi a continuación.

**Proposición 1.1** (Szemerédi, versión finitaria)**.** *Sean $N$ un entero positivo y $\mathbb{Z}_N = \mathbb{Z}/N\mathbb{Z}$. Sea $1 \geq \delta > 0$ un número real fijado, y $k \geq 3$ un entero. Entonces existe un menor natural $N_0(\delta, k) \in \mathbb{Z}^+$ con la propiedad de que, si $N \geq N_0(\delta, k)$ y $A \subseteq \mathbb{Z}_N$ con $|A| \geq \delta N$, entonces tenemos que $A$ contiene alguna progresión aritmética de longitud $k$.*

**Proposición 1.2** (Szemerédi, versión funcional)**.** *Sea $\nu_{\text{const}} : \mathbb{Z}_N \to \mathbb{R}^+$ la función constante unidad, $\nu_{\text{const}} \equiv 1$. Sean $0 < \delta \leq 1$ y $k \geq 1$ un entero fijado. Consideremos también $N \in \mathbb{Z}^+$ un parámetro entero grande y una función $f : \mathbb{Z}_N \to \mathbb{R}^+$ tal que*

$$0 \leq f(x) \leq \nu_{\text{const}}(x), \text{ para todo } x \in \mathbb{Z}_N, \tag{1.1}$$

*y*

$$\mathbb{E}(f(x)|x \in \mathbb{Z}_N) \geq \delta. \tag{1.2}$$

*Entonces tenemos que*

$$\mathbb{E}\Big(f(x)f(x+r)\cdots f\big(x+(k-1)r\big)\Big|x, r \in \mathbb{Z}_N\Big) \geq c(k,\delta) - o_{k,\delta}(1)$$

*para alguna constante* $c(k,\delta) > 0$ *que no depende ni de* $f$ *ni de* $N$*.*

Observemos en este punto que el teorema de Green-Tao no puede ser obtenido directamente de la versión clásica del teorema de Szemerédi, puesto que por el teorema del número primo, $\lim\limits_{N\to+\infty} \dfrac{\big|\mathcal{P}\cap[1,N]\big|\log N}{N} = 1$ y por lo tanto $\limsup\limits_{N\to+\infty} \dfrac{\big|\mathcal{P}\cap[1,N]\big|}{N} = 0$. Es así que requerimos de una adaptación del teorema de Szemerédi (clásico) a un contexto más general en el que cualquier subconjunto de un conjunto suficientemente pseudoaleatorio con densidad relativa positiva contenga progresiones aritméticas de longitud arbitraria. Así, la prueba del *teorema de Green-Tao* puede ser concebida, burdamente hablando en dos pasos:

(1) generalizar el teorema de Szemerédi para el contexto de *medidas pseudoaleatorias.*

(2) probar la existencia de una medida pseudoaleatoria en el conjunto de los números primos.

El primer paso lo trataremos en el capítulo 6, donde estableceremos una especie de *principio de transferencia* que parte del teorema de Szemerédi clásico a un ámbito más general, el de medidas pseudoaleatorias, y del que se concluye el teorema siguiente.

**Teorema 1.7** (Szemerédi relativo a medidas pseudoaleatorias)**.** *Sean* $k \geq 3$ *un entero y* $0 < \delta \leq 1$ *un parámetro fijado. Supongamos que* $\nu : \mathbb{Z}_N \to \mathbb{R}^+$ *es una medida* $k$*-pseudoaleatoria. Sea* $N \geq 1$ *un parámetro entero grande y* $f : \mathbb{Z}_N \to \mathbb{R}^+$*, una función cualquiera no negativa tal que*

$$0 \leq f(x) \leq \nu(x) \text{ para todo } x \in \mathbb{Z}_N \tag{1.3}$$

*y*

$$\mathbb{E}(f) \geq \delta. \tag{1.4}$$

*Entonces tenemos que*

$$\mathbb{E}\Big(f(x)f(x+r)\cdots f\big(x+(k-1)r\big)\Big|x, r \in \mathbb{Z}_N\Big) \geq c(k,\delta) - o_{k,\delta}(1) \tag{1.5}$$

*donde* $c(k,\delta) > 0$ *es la misma constante que aparece en la proposición 1.2.*

El concepto preciso de lo que es una medida pseudoaleatoria será dado en la definición 6.5. Notemos que toda vez que mostremos que $\nu_{\text{const}}$, la función idénticamente 1, constituye una medida pseudoaleatoria, el teorema 1.7 resulta una generalización de la versión funcional del teorema de Szemerédi, proposición 1.2.

El segundo paso lo abordaremos en el capítulo 7, construyendo una medida pseudoaleatoria que es soportada en el conjunto de los números primos en el sentido que tal medida se anula fuera de los primos y es positivo en una cantidad infinita de números primos. Este capítulo concluye con la proposición siguiente, que muestra la existencia de la medida pseudoaleatoria requerida.

**Proposición 1.3.** *Sean* $\epsilon_k = \dfrac{1}{2^k(k+4)!}$ *y* $N$ *un primo suficientemente grande. Entonces existe una medida* $k$*-pseudoaleatoria* $\nu : \mathbb{Z}_N \to \mathbb{R}^+$ *tal que*

$$\nu(n) \geq k^{-1}2^{-k-5}\tilde{\Lambda}(n), \text{ para todo } n \in [\epsilon_k N, 2\epsilon_k N].$$

Del teorema 1.7 y la proposición 1.3 tenemos como consecuencia el teorema de Green-Tao, como lo mostraremos en el teorema 7.1.

Dada la importancia del teorema de Szemerédi en el desarrollo de la combinatoria aditiva, y particularmente en la demostración del teorema de Green-Tao, abordaremos este importante teorema en los capítulos 3, 4 y 5. En el capítulo 3 trataremos el caso de progresiones aritméticas de longitud 3, esto es, el teorema de Roth, a través de dos métodos, primero haciendo uso del análisis de Fourier en $\mathbb{Z}_{\text{N}}$, método inicialmente empleado por Roth, (cf. [3]), y el segundo, haciendo uso de la $U^2$-norma de Gowers. En el capítulo 4 abordaremos el caso de progresiones aritméticas de longitud 4, para lo cual seguimos los trabajos de Gowers, que en 1998 y haciendo uso exclusivo de métodos analíticos, consiguió probar el resultado en mención. Son referencias importantes [1] y [2]. En el capítulo 5 se estudiará el caso de progresiones aritméticas de longitud $k$, en general, esta vez vía métodos de teoría ergódica, a través del teorema de la recurrencia múltiple de Furstenberg.

# Capítulo 2

# Preliminares

Con la finalidad de contar el número de progresiones aritméticas de longitud tres en un subconjunto $A$ de $\mathbb{Z}$, requeriremos inicialmente determinar el número de ternas $(x, y, z)$ que sean solución del problema:

$$\begin{cases} x + z \equiv 2y \mod N \\ x, y, z \in A \end{cases} \tag{2.1}$$

Siendo $\eta_0$ el número de soluciones, ternas $(x, y, z)$, de (2.1). Tenemos que

$$\eta_0 = \sum_{\substack{x,y,z\in A \\ x+z\equiv 2y \bmod N}} 1 = \sum_{x\in A}\sum_{y\in A}\sum_{z\in A} \frac{1}{N}\sum_{k=0}^{N-1} \left(e^{2\pi i/N}\right)^{-(x+z-2y)k},$$

donde aprovechamos naturalmente la relación $\sum_{k=0}^{N-1} \left(e^{2\pi i/N}\right)^{rk} = \begin{cases} N & , \text{ si } r \equiv 0 \bmod N \\ 0 & , \text{ si } r \not\equiv 0 \bmod N \end{cases}$, para cada $r \in \mathbb{Z}$. Es por ello que a continuación abordaremos el análisis de Fourier en $\mathbb{Z}_{\mathrm{N}}$. Observe que dados $r, s \in \mathbb{Z}$, si $r \equiv s \bmod N$ entonces $\left(e^{2\pi i/N}\right)^{r-s} = 1$, y por lo tanto $\left(e^{2\pi i/N}\right)^{r} = \left(e^{2\pi i/N}\right)^{s}$. De modo que, para cada $x \in \mathbb{Z}_{\mathrm{N}}$, $\sum_{k=0}^{N-1} \left(e^{2\pi i/N}\right)^{xk}$ está bien definido.

## 2.1. Transformada de Fourier discreta

Sea $N$ un entero positivo. Consideremos $\omega = e^{2\pi i/N}$.

**Definición 2.1.** *Dada una función $f : \mathbb{Z}_N \to \mathbb{C}$ y $r \in \mathbb{Z}_N$, definimos*

$$\widehat{f}(r) = \sum_{s\in\mathbb{Z}_N} f(s)\omega^{-rs}.$$

*La función $\widehat{f} : \mathbb{Z}_N \to \mathbb{C}$ es llamada la* **transformada de Fourier discreta** *de $f$.*

**Definición 2.2.** *Dadas las funciones $f, g : \mathbb{Z}_N \to \mathbb{C}$ definimos la* **convolución** *de $f$ y $g$ como*

$$(f * g)(s) = \sum_{t \in \mathbb{Z}_N} f(t)\overline{g(t-s)}.$$

**Proposición 2.1.** *Las siguientes identidades son válidas:*

$$\widehat{(f * g)}(r) = \widehat{f}(r)\overline{\widehat{g}(r)}, \tag{2.2}$$

$$\sum_{r \in \mathbb{Z}_N} \widehat{f}(r)\overline{\widehat{g}(r)} = N \sum_{s \in \mathbb{Z}_N} f(s)\overline{g(s)}, \tag{2.3}$$

$$\sum_{r \in \mathbb{Z}_N} |\widehat{f}(r)|^2 = N \sum_{s \in \mathbb{Z}_N} |f(s)|^2 \tag{2.4}$$

$$f(s) = \frac{1}{N} \sum_{r \in \mathbb{Z}_N} \widehat{f}(r)\omega^{rs}. \tag{2.5}$$

*Las identidades (2.3) y (2.4) son denominadas* **identidades de Parseval**, *siendo claramente (2.4) un caso particular de (2.3). La identidad (2.5) es llamada la* **fórmula de inversión**, *por lo que a los $\widehat{f}(r)$ se les denomina* **coeficientes de Fourier**.

*Demostración.* Mostremos (2.2). Calculamos

$$\begin{aligned}
\widehat{(f * g)}(r) &= \sum_{s \in \mathbb{Z}_N} (f * g)(s)\omega^{-rs} = \sum_{s,t \in \mathbb{Z}_N} f(t)\overline{g(t-s)}\omega^{-rs} \\
&= \sum_{s,t \in \mathbb{Z}_N} f(t)\overline{g(t-s)}\omega^{-rt}\omega^{r(t-s)} = \sum_{s,t \in \mathbb{Z}_N} f(t)\omega^{-rt}\overline{g(t-s)\omega^{-r(t-s)}} \\
&= \sum_{t \in \mathbb{Z}_N} f(t)\omega^{-rt} \sum_{s \in \mathbb{Z}_N} \overline{g(t-s)\omega^{-r(t-s)}} = \sum_{t \in \mathbb{Z}_N} f(t)\omega^{-rt}\overline{\sum_{s \in \mathbb{Z}_N} g(t-s)\omega^{-r(t-s)}} \\
&= \sum_{t \in \mathbb{Z}_N} f(t)\omega^{-rt}\overline{\sum_{p \in \mathbb{Z}_N} g(p)\omega^{-rp}} = \widehat{f}(r)\overline{\widehat{g}(r)}.
\end{aligned}$$

Mostremos (2.3). Calculamos

$$\begin{aligned}
\sum_{r \in \mathbb{Z}_N} \widehat{f}(r)\overline{\widehat{g}(r)} &= \sum_{r \in \mathbb{Z}_N} \widehat{(f * g)}(r) \\
&= \sum_{r \in \mathbb{Z}_N} \sum_{s \in \mathbb{Z}_N} (f * g)(s)\omega^{-rs} \\
&= \sum_{s \in \mathbb{Z}_N} (f * g)(s) \sum_{r \in \mathbb{Z}_N} \omega^{-rs}.
\end{aligned} \tag{2.6}$$

Para $s \in \mathbb{Z}_N \setminus \{0\}$ tenemos

$$\sum_{r \in \mathbb{Z}_N} \omega^{-rs} = \sum_{r=0}^{N-1} \omega^{-rs} = \frac{1 - (\omega^{-s})^N}{1 - \omega^{-s}} = 0,$$

donde $\omega = e^{2\pi i/N}$ es diferente de 1, por lo tanto

$$\sum_{r\in\mathbb{Z}_N} \omega^{-rs} = \begin{cases} N & \text{, si } s = 0 \\ 0 & \text{, si } s \in \mathbb{Z}_N \setminus \{0\}. \end{cases}$$

De (2.6), tenemos que

$$\sum_{r\in\mathbb{Z}_N} \widehat{f}(r)\overline{\widehat{g}(r)} = (f * g)(0)N = N \sum_{t\in\mathbb{Z}_N} f(t)\overline{g(t)}.$$

Mostremos (2.5). Sea $s \in \mathbb{Z}_N$, entonces

$$\begin{aligned} \sum_{r\in\mathbb{Z}_N} \widehat{f}(r)\omega^{rs} &= \sum_{r\in\mathbb{Z}_N}\sum_{t\in\mathbb{Z}_N} f(t)\omega^{-rt}\omega^{rs} \\ &= \sum_{t\in\mathbb{Z}_N} f(t) \sum_{r\in\mathbb{Z}_N} \omega^{-r(t-s)} \\ &= f(s)N, \end{aligned}$$

por lo tanto

$$f(s) = \frac{1}{N} \sum_{r\in\mathbb{Z}_N} \widehat{f}(r)\omega^{rs}.$$

□

**Lema 2.1.** *Sean $f, g : \mathbb{Z}_N \to \mathbb{C}$. Entonces*

$$\sum_{r\in\mathbb{Z}_N} \left|\widehat{f}(r)\right|^2 \left|\overline{\widehat{g}(r)}\right|^2 = N \sum_{t\in\mathbb{Z}_N} \left| \sum_{s\in\mathbb{Z}_N} f(s)\overline{g(s-t)} \right|^2. \tag{2.7}$$

*Demostración.* De (2.2) y (2.4) tenemos que

$$\begin{aligned} \sum_{r\in\mathbb{Z}_N} \left|\widehat{f}(r)\right|^2 \left|\overline{\widehat{g}(r)}\right|^2 &= \sum_{r\in\mathbb{Z}_N} \left|\widehat{f * g}(r)\right|^2 \\ &= N \sum_{t\in\mathbb{Z}_N} \left|(f * g)(t)\right|^2 \\ &= N \sum_{t\in\mathbb{Z}_N} \left| \sum_{s\in\mathbb{Z}_N} f(s)\overline{g(s-t)} \right|^2. \end{aligned}$$

□

Tomando $f = g$ en la identidad (2.7), tenemos

$$\begin{aligned} \sum_{r\in\mathbb{Z}_N} \left|\widehat{f}(r)\right|^4 &= N \sum_{t\in\mathbb{Z}_N} \left| \sum_{m\in\mathbb{Z}_N} f(m)\overline{f(m-t)} \right|^2 \\ &= N \sum_{t\in\mathbb{Z}_N} \sum_{m\in\mathbb{Z}_N} f(m)\overline{f(m-t)} \overline{\sum_{n\in\mathbb{Z}_N} f(n)\overline{f(n-t)}} \end{aligned}$$

$$
\begin{aligned}
&= N \sum_{t,m,n\in\mathbb{Z}_N} f(m)\overline{f(m-t)f(n)}f(n-t)\\
&= N \sum_{m\in\mathbb{Z}_N} f(m) \sum_{t,n\in\mathbb{Z}_N} \overline{f(m-t)f(n)}f(n-t)\\
&= N \sum_{m\in\mathbb{Z}_N} f(m) \sum_{t,n\in\mathbb{Z}_N} \overline{f(m+t)f(n)}f(n+t)\\
&= N \sum_{m\in\mathbb{Z}_N} f(m) \sum_{t,n\in\mathbb{Z}_N} \overline{f(m+t)f(m+n)}f(m+n+t)\\
&= N \sum_{m,n,t\in\mathbb{Z}_N} f(m)\overline{f(m+t)f(m+n)}f(m+n+t). \qquad (2.8)
\end{aligned}
$$

La igualdad (2.8) está relacionada con la $U^2$-norma de Gowers que será dada en la definición 2.5.

## 2.2. La $U^d$-norma de Gowers para funciones reales.

En esta sección desarrollaremos un tipo de norma sobre el conjunto de funciones $f$ de $\mathbb{Z}_N$ en $\mathbb{R}$, que permitirá contar el número de progresiones aritméticas existentes dentro de un conjunto $A$ en $\mathbb{Z}_N$, como se muestra en el lema 3.5 para el caso de progresiones aritméticas de longitud tres, y en el lema 4.7 para progresiones de longitud cuatro. En el primer caso tenemos que para funciones $f_1, f_2, f_3 : \mathbb{Z}_N \to \mathbb{R}$,

$$\left|\sum_{a,d} f_1(a)f_2(a+d)f_3(a+2d)\right| \leq N^2 \min_{i=1,2,3} ||f_i||_{U^2}.$$

donde $||f_i||_{U^2}$ denota la $U^2$-norma de Gowers de $f_i$. Así, si $A \subseteq \mathbb{Z}_N$ y $|A| = \delta N$, entonces

$$
\begin{aligned}
\sum_{a,d} 1_A(a)1_A(a+d)1_A(a+2d) &= \sum_{a,d} 1_A(a)1_A(a+d)\delta + \sum_{a,d} 1_A(a)1_A(a+d)\big(1_A(a+2d)-\delta\big)\\
&= \delta|A|^2 + \sum_{a,d} 1_A(a)1_A(a+d)\big(1_A(a+2d)-\delta\big)\\
&= \delta|A|^2 - \Big|\sum_{a,d} 1_A(a)1_A(a+d)\big(1_A(a+2d)-\delta\big)\Big|\\
&\geq \delta^3N^2 - N^2||1_A-\delta||_{U^2},
\end{aligned}
$$

Por lo tanto existirán progresiones aritméticas de longitud tres si la $U^2$-norma de Gowers de la función $1_A - \delta$ es pequeña.

**Definición 2.3.** *Sea $A$ un conjunto finito no vacío, que usualmente consideraremos como $\mathbb{Z}_N$, y sea una función $f : A \to \mathbb{R}$. El* **valor promedio de** *$f$, al cual denotaremos con $\mathbb{E}(f)$,*

*es definido por*

$$\mathbb{E}(f) = \frac{1}{|A|}\sum_{x\in A} f(x),$$

*donde* $|A|$ *denota la cardinalidad de* $A$*. De manera más general, si* $P(x)$ *es una afirmación que concierne a elementos de* $A$ *y que es cierta para al menos un elemento de* $A$*, entonces definimos el* **valor promedio de** $f$ **sujeto a la condición** $P(x)$*, como*

$$\mathbb{E}(f(x)|P(x)) = \frac{\sum_{x\in A, P(x)} f(x)}{|\{x \in A : P(x)\}|}.$$

**Definición 2.4.** *Sea* $d \geq 0$ *un entero que llamaremos* **dimensión***, en la práctica tomaremos* $d = k-1$*, donde* $k$ *será la longitud de las progresiones aritméticas en consideración. Sea* $\{0,1\}^d$ *el cubo discreto* $d$*-dimensional, consistente de* $d$*-uplas* $\omega = (\omega_1, \ldots, \omega_d)$ *donde* $\omega_j \in \{0,1\}$ *para* $j = 1, \ldots, d$*. Si* $h = (h_1, \ldots, h_d) \in \mathbb{Z}_N^d$*, definimos* $\omega \cdot h = \omega_1 h_1 + \cdots + \omega_d h_d$*. Si* $(f_\omega)_{\omega\in\{0,1\}^d}$ *es una* $\{0,1\}^d$*-upla de funciones en* $L^\infty(\mathbb{Z}_N)$*, definimos el* **producto interno de Gowers** $d$**-dimensional** $\langle (f_\omega)_{\omega\in\{0,1\}^d}\rangle_{U^d}$ *por la fórmula*

$$\langle (f_\omega)_{\omega\in\{0,1\}^d}\rangle_{U^d} = \mathbb{E}\left(\prod_{\omega\in\{0,1\}^d} f_\omega(x+\omega\cdot h)\middle| x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^d\right). \tag{2.9}$$

En lo que sigue nos referiremos a $\{x+\omega\cdot h : \omega \in \{0,1\}^d\}$ como un cubo $d$-dimendional.

*Ejemplo.* Cuando $d = 2$, tenemos

$$\langle (f_{00}, f_{10}, f_{01}, f_{11})\rangle_{U^2} = \mathbb{E}\Big(f_{00}(x)f_{10}(x+h_1)f_{01}(x+h_2)f_{11}(x+h_1+h_2)\Big| x, h_1, h_2 \in \mathbb{Z}_N\Big).$$

Supongamos que la colección $\{f_\omega\}_{\omega\in\{0,1\}^d}$ no dependa del dígito final $\omega_d$ de $\omega$, esto es, $f_{\omega',0} = f_{\omega',1}$ para todo $\omega' \in \{0,1\}^{d-1}$. Entonces podemos reescribir (2.9) como

$$\begin{aligned}
\langle (f_\omega)_{\omega\in\{0,1\}^d}\rangle_{U^d} &= \mathbb{E}\left(\prod_{\omega\in\{0,1\}^d} f_\omega(x+\omega\cdot h)\middle| x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^d\right)\\
&= \mathbb{E}\left(\prod_{\omega'\in\{0,1\}^{d-1}} f_{\omega',0}(x+\omega'\cdot h')f_{\omega',1}(x+h_d+\omega'\cdot h')\right.\\
&\qquad\qquad \left.\middle| x \in \mathbb{Z}_N, h' \in \mathbb{Z}_N^{d-1}, h_d \in \mathbb{Z}_N\right),
\end{aligned}$$

donde $\omega' = (\omega_1, \ldots, \omega_{d-1})$ y $h' = (h_1, \ldots, h_{d-1})$. Siendo $f_{\omega'} = f_{\omega',0} = f_{\omega',1}$, tenemos que

$$\langle (f_\omega)_{\omega\in\{0,1\}^d}\rangle_{U^d} = \mathbb{E}\left(\prod_{\omega'\in\{0,1\}^{d-1}} f_{\omega'}(x+\omega'\cdot h') \prod_{\omega'\in\{0,1\}^{d-1}} f_{\omega'}(x+h_d+\omega'\cdot h')\right.$$
$$\left.\middle| x \in \mathbb{Z}_N, h' \in \mathbb{Z}_N^{d-1}, h_d \in \mathbb{Z}_N\right)$$

$$= \mathbb{E}\Bigg( \prod_{\omega' \in \{0,1\}^{d-1}} f_{\omega'}(x + \omega' \cdot h') \mathbb{E}\Big( \prod_{\omega' \in \{0,1\}^{d-1}} f_{\omega'}(x + h_d + \omega' \cdot h') \Big| h_d \in \mathbb{Z}_N \Big) \Bigg| x \in \mathbb{Z}_N, h' \in \mathbb{Z}_N^{d-1} \Bigg)$$

$$= \mathbb{E}\Bigg( \mathbb{E}\Big( \prod_{\omega' \in \{0,1\}^{d-1}} f_{\omega'}(x + \omega' \cdot h') \mathbb{E}\Big( \prod_{\omega' \in \{0,1\}^{d-1}} f_{\omega'}(x + h_d + \omega' \cdot h') \Big| h_d \in \mathbb{Z}_N \Big) \Big| x \in \mathbb{Z}_N \Big) \Bigg| h' \in \mathbb{Z}_N^{d-1} \Bigg)$$

$$= \mathbb{E}\Bigg( \mathbb{E}\Big( \prod_{\omega' \in \{0,1\}^{d-1}} f_{\omega'}(x + \omega' \cdot h') \mathbb{E}\Big( \prod_{\omega' \in \{0,1\}^{d-1}} f_{\omega'}(y + \omega' \cdot h') \Big| y \in \mathbb{Z}_N \Big) \Big| x \in \mathbb{Z}_N \Big) \Bigg| h' \in \mathbb{Z}_N^{d-1} \Bigg)$$

$$= \mathbb{E}\Bigg( \mathbb{E}\Big( \prod_{\omega' \in \{0,1\}^{d-1}} f_{\omega'}(y + \omega' \cdot h') \Big| y \in \mathbb{Z}_N \Big) \mathbb{E}\Big( \prod_{\omega' \in \{0,1\}^{d-1}} f_{\omega'}(x + \omega' \cdot h') \Big| x \in \mathbb{Z}_N \Big) \Bigg| h' \in \mathbb{Z}_N^{d-1} \Bigg)$$

$$= \mathbb{E}\Bigg( \Big| \mathbb{E}\Big( \prod_{\omega' \in \{0,1\}^{d-1}} f_{\omega'}(x + \omega' \cdot h') \Big| x \in \mathbb{Z}_N \Big) \Big|^2 \Bigg| h' \in \mathbb{Z}_N^{d-1} \Bigg). \tag{2.10}$$

En particular tenemos que $\big\langle (f_\omega)_{\omega \in \{0,1\}^d} \big\rangle_{U^d} \geq 0$ cuando la colección $\{f_\omega\}_{\omega \in \{0,1\}^d}$ es independiente de $\omega_d$. Así, hemos probado que

$$\big\langle (f)_{\omega \in \{0,1\}^d} \big\rangle_{U^d} \geq 0 \text{ para cada función } f : \mathbb{Z}_N \to \mathbb{R} \tag{2.11}$$

cuando $d \geq 1$, donde $f_\omega = f$ para cada $\omega \in \{0,1\}^d$. Podemos ahora definir la *$U^d$-norma de Gowers*.

**Definición 2.5.** *Sean $f : \mathbb{Z}_N \to \mathbb{R}$ y $d \geq 1$. Definimos la* **$U^d$-norma de Gowers** *de $f$, que denotaremos por $||f||_{U^d}$, por la expresión*

$$||f||_{U^d} = \big\langle (f)_{\omega \in \{0,1\}^d} \big\rangle_{U^d}^{1/2^d} = \mathbb{E}\Bigg( \prod_{\omega \in \{0,1\}^d} f(x + \omega \cdot h) \Bigg| x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^d \Bigg)^{1/2^d}. \tag{2.12}$$

Notemos que en la definición 2.5, en los casos particulares de $U^2$ y $U^3$ normas de Gowers para una función real $f : \mathbb{Z}_N \to \mathbb{R}$, se pueden expresar de la siguiente manera:

$$||f||_{U^2} = \Bigg( \frac{1}{N^3} \sum_{a,b,c} f(a) f(a+b) f(a+c) f(a+b+c) \Bigg)^{1/4}, \tag{2.13}$$

y

$$||f||_{U^3} = \Big(\frac{1}{N^4}\sum_{a,b,c,d} f(a)f(a+b)f(a+c)f(a+d)f(a+b+c)f(a+b+d)\times \\ \times f(a+c+d)f(a+b+c+d)\Big)^{1/8},$$

estas expresiones se usarán para *contar* progresiones aritméticas de longitud 3 y 4, respectivamente.

Cuando la colección $\{f_\omega\}_{\omega\in\{0,1\}^d}$ depende de $\omega_d$, tenemos que

$$\begin{aligned}
\big\langle (f_\omega)_{\omega\in\{0,1\}^d}\big\rangle_{U^d} &= \mathbb{E}\Big(\prod_{\omega'\in\{0,1\}^{d-1}} f_{\omega',0}(x+\omega'\cdot h')\prod_{\omega'\in\{0,1\}^{d-1}} f_{\omega',1}(x+h_d+\omega'\cdot h') \\
&\qquad \Big| x\in\mathbb{Z}_N, h'\in\mathbb{Z}_N^{d-1}, h_d\in\mathbb{Z}_N\Big) \\
&= \mathbb{E}\Big(\prod_{\omega'\in\{0,1\}^{d-1}} f_{\omega',0}(x+\omega'\cdot h')\mathbb{E}\Big(\prod_{\omega'\in\{0,1\}^{d-1}} f_{\omega',1}(x+h_d+\omega'\cdot h')\Big|h_d\in\mathbb{Z}_N\Big) \\
&\qquad \Big| x\in\mathbb{Z}_N, h'\in\mathbb{Z}_N^{d-1}\Big) \\
&= \mathbb{E}\Big(\prod_{\omega'\in\{0,1\}^{d-1}} f_{\omega',0}(x+\omega'\cdot h')\mathbb{E}\Big(\prod_{\omega'\in\{0,1\}^{d-1}} f_{\omega',1}(y+\omega'\cdot h')\Big|y\in\mathbb{Z}_N\Big) \\
&\qquad \Big| x\in\mathbb{Z}_N, h'\in\mathbb{Z}_N^{d-1}\Big) \\
&= \mathbb{E}\Big(\mathbb{E}\Big(\prod_{\omega'\in\{0,1\}^{d-1}} f_{\omega',0}(x+\omega'\cdot h')\mathbb{E}\Big(\prod_{\omega'\in\{0,1\}^{d-1}} f_{\omega',1}(y+\omega'\cdot h')\Big|y\in\mathbb{Z}_N\Big)\Big|x\in\mathbb{Z}_N\Big) \\
&\qquad \Big| h'\in\mathbb{Z}_N^{d-1}\Big) \\
&= \mathbb{E}\Big(\mathbb{E}\Big(\prod_{\omega'\in\{0,1\}^{d-1}} f_{\omega',0}(x+\omega'\cdot h')\Big|x\in\mathbb{Z}_N\Big)\mathbb{E}\Big(\prod_{\omega'\in\{0,1\}^{d-1}} f_{\omega',1}(y+\omega'\cdot h')\Big|y\in\mathbb{Z}_N\Big) \\
&\qquad \Big| h'\in\mathbb{Z}_N^{d-1}\Big).
\end{aligned}$$

De la desigualdad de Cauchy-Schwarz en las $h'$ variables, tenemos que

$$\begin{aligned}
\big|\big\langle (f_\omega)_{\omega\in\{0,1\}^d}\big\rangle_{U^d}\big| &\le \mathbb{E}\Big(\Big|\mathbb{E}\Big(\prod_{\omega'\in\{0,1\}^{d-1}} f_{\omega',0}(x+\omega'\cdot h')\Big|x\in\mathbb{Z}_N\Big)\Big|^2\Big|h'\in\mathbb{Z}_N^{d-1}\Big)^{1/2}\times \\
&\times \mathbb{E}\Big(\Big|\mathbb{E}\Big(\prod_{\omega'\in\{0,1\}^{d-1}} f_{\omega',1}(y+\omega'\cdot h')\Big|y\in\mathbb{Z}_N\Big)\Big|^2\Big|h'\in\mathbb{N}_N^{d-1}\Big)^{1/2}. \qquad (2.14)
\end{aligned}$$

Así, como las colecciones $\{f_{\omega',0}\}_{\omega\in\{0,1\}^d}$ y $\{f_{\omega',1}\}_{\omega\in\{0,1\}^d}$ no dependen de $\omega_d$, donde $\omega'$ representa a las primeras $d-1$ coordenadas de $\omega$, esto es $\omega=(\omega',\omega_d)$; de (2.10) y (2.14) tenemos que

$$\left|\left\langle (f_\omega)_{\omega\in\{0,1\}^d}\right\rangle_{U^d}\right| \leq \left\langle (f_{\omega',0})_{\omega\in\{0,1\}^d}\right\rangle_{U^d}^{1/2}\left\langle (f_{\omega',1})_{\omega\in\{0,1\}^d}\right\rangle_{U^d}^{1/2}.$$

Similarmente ocurre si reemplazamos el rol de la última coordenada de $\omega$, $\omega_d$, por cualquiera de las otras coordenadas $\omega_i$, $1\leq i\leq d$. Así

$$\left|\left\langle (f_{\omega',0})_{\omega\in\{0,1\}^d}\right\rangle_{U^d}\right| \leq \left\langle (f_{\omega'',0,0})_{\omega\in\{0,1\}^d}\right\rangle_{U^d}^{1/2}\left\langle (f_{\omega'',1,0})_{\omega\in\{0,1\}^d}\right\rangle_{U^d}^{1/2},$$

y

$$\left|\left\langle (f_{\omega',1})_{\omega\in\{0,1\}^d}\right\rangle_{U^d}\right| \leq \left\langle (f_{\omega'',0,1})_{\omega\in\{0,1\}^d}\right\rangle_{U^d}^{1/2}\left\langle (f_{\omega'',1,1})_{\omega\in\{0,1\}^d}\right\rangle_{U^d}^{1/2},$$

donde $\omega''\in\{0,1\}^{d-2}$ representa a las primeras $d-2$ coordenadas de $\omega$. Continuando de esta manera obtenemos la *desigualdad de Gowers-Cauchy-Schwarz*, esto es,

$$\begin{aligned}\left|\left\langle (f_\omega)_{\omega\in\{0,1\}^d}\right\rangle_{U^d}\right| &\leq \prod_{\omega\in\{0,1\}^d}\left\langle (f_\omega)_{\eta\in\{0,1\}^d}\right\rangle_{U^d}^{1/2^d}\\ &= \prod_{\omega\in\{0,1\}^d}||f_\omega||_{U^d}. \end{aligned} \tag{2.15}$$

De la multilinealidad del producto interno de Gowers, y de la fórmula binomial, obtenemos la desigualdad

$$\left|\left\langle (f+g)_{\omega\in\{0,1\}^d}\right\rangle_{U^d}\right| \leq \left(||f||_{U^d}+||g||_{U^d}\right)^{2^d}.$$

Por consiguiente obtenemos la *desigualdad triangular de Gowers*

$$\begin{aligned}\left|\left\langle (f+g)_{\omega\in\{0,1\}^d}\right\rangle_{U^d}\right|^{1/2^d} &\leq ||f||_{U^d}+||g||_{U^d}\\ ||f+g||_{U^d} &\leq ||f||_{U^d}+||g||_{U^d}.\end{aligned}$$

**Ejemplo 2.1.** *Si $d=2$ tenemos para $f,g:\mathbb{Z}_N\to\mathbb{R}$*

$$\begin{aligned}&\left\langle (f+g)_{\omega\in\{0,1\}^2}\right\rangle_{U^2}\\ &\quad=\left\langle \big((f+g)_{00},(f+g)_{01},(f+g)_{10},(f+g)_{11}\big)\right\rangle_{U^2}\\ &\quad=\left\langle \big((g)_{00},(g)_{01},(g)_{10},(g)_{11}\big)\right\rangle_{U^2}\\ &\quad+\left\langle \big((f)_{00},(g)_{01},(g)_{10},(g)_{11}\big)\right\rangle_{U^2}+\cdots+\left\langle \big((g)_{00},(g)_{01},(g)_{10},(f)_{11}\big)\right\rangle_{U^2}\\ &\quad+\left\langle \big((f)_{00},(f)_{01},(g)_{10},(g)_{11}\big)\right\rangle_{U^2}+\cdots+\left\langle \big((g)_{00},(g)_{01},(f)_{10},(f)_{11}\big)\right\rangle_{U^2}\\ &\quad+\left\langle \big((f)_{00},(f)_{01},(f)_{10},(g)_{11}\big)\right\rangle_{U^2}+\cdots+\left\langle \big((g)_{00},(f)_{01},(f)_{10},(f)_{11}\big)\right\rangle_{U^2}\\ &\quad+\left\langle \big((f)_{00},(f)_{01},(f)_{10},(f)_{11}\big)\right\rangle_{U^2}\end{aligned}$$

$$\leq \binom{4}{0}||g||_{U^2}^4 + \binom{4}{1}||f||_{U^2}||g||_{U^2}^3 + \binom{4}{2}||f||_{U^2}^2||g||_{U^2}^2 + \binom{4}{3}||f||_{U^2}^3||g||_{U^2} + \binom{4}{4}||f||_{U^2}^4$$
$$= \Big(||f||_{U^2} + ||g||_{U^2}\Big)^4.$$

**Ejemplo 2.2.** *Cuando $d = 2$, tenemos*

$$||f||_{U^2} = \mathbb{E}\bigg(f(x)f(x+h_1)f(x+h_2)f(x+h_1+h_2)\Big|x, h_1, h_2 \in \mathbb{Z}_N\bigg)^{1/4},$$

*y la desigualdad de Gowers-Cauchy-Schwarz será*

$$\big|\mathbb{E}\big(f_{00}(x)f_{10}(x+h_1)f_{01}(x+h_2)f_{11}(x+h_1+h_2)\big|x, h_1, h_2 \in \mathbb{Z}_N\big)\big|$$
$$\leq ||f_{00}||_{U^2}||f_{10}||_{U^2}||f_{01}||_{U^2}||f_{11}||_{U^2}.$$

De (2.8), en el caso que $f$ sea una función real, tenemos que

$$\sum_{r\in\mathbb{Z}_N} \big|\widehat{f}(r)\big|^4 = N \sum_{m,n,t\in\mathbb{Z}_N} f(m)f(m+t)f(m+n)f(m+n+t),$$

y por lo tanto, de (2.13), tenemos que

$$||f||_{U^2} = \frac{1}{N}\bigg(\sum_{\xi\in\mathbb{Z}_N} |\widehat{f}(\xi)|^4\bigg)^{1/4}.$$

Así, si $||f||_{U^2} = 0$ entonces

$$\widehat{f}(\xi) = 0, \text{ para todo } \xi \in \mathbb{Z}_N.$$

Como $f(s) = \dfrac{1}{N}\displaystyle\sum_{\xi\in\mathbb{Z}_N} \widehat{f}(\xi)\omega^{s\xi}$ para todo $s \in \mathbb{Z}_N$, donde $\omega = e^{2\pi i/N}$, deducimos que $f = 0$, esto es, $f$ es la función idénticamente nula. Recíprocamente, si $f = 0$, de la definición de $U^2$-norma de Gowers, tenemos que $||f||_{U^2} = 0$. Por lo tanto la $U^2$-norma de Gowers es una norma auténtica.

Retornamos al estudio de la $U^d$-norma de Gowers. Como

$$||\nu_{\text{const}}||_{U^d} = ||1||_{U^d} = 1, \tag{2.16}$$

donde $\nu_{\text{const}}(x) = 1$ para cada $x \in \mathbb{Z}_N$, vemos de (2.15) que

$$\big|\big\langle(f_\omega)_{\omega\in\{0,1\}^d}\big\rangle_{U^d}\big| \leq ||f||_{U^d}^{2^{d-1}},$$

donde $f_\omega = 1$ cuando $\omega_d = 1$, y $f_\omega = f$ cuando $\omega_d = 0$. Pero el lado izquierdo de la desigualdad anterior es

$$\begin{aligned}
\left\langle (f_\omega)_{\omega\in\{0,1\}^d} \right\rangle_{U^d} &= \mathbb{E}\left( \prod_{\omega\in\{0,1\}^d} f_\omega(x+\omega\cdot h) \middle| x\in\mathbb{Z}_N, h\in\mathbb{Z}_N^d \right) \\
&= \mathbb{E}\Bigg( \prod_{\omega'\in\{0,1\}^{d-1}} f_{\omega',0}(x+\omega'\cdot h') \prod_{\omega'\in\{0,1\}^{d-1}} f_{\omega',1}(x+h_d+\omega'\cdot h') \\
&\qquad \Bigg| x\in\mathbb{Z}_N, h'\in\mathbb{Z}_N^{d-1}, h_d\in\mathbb{Z}_N \Bigg) \\
&= \mathbb{E}\left( \prod_{\omega'\in\{0,1\}^{d-1}} f(x+\omega'\cdot h') \middle| x\in\mathbb{Z}_N, h'\in\mathbb{Z}_N^{d-1}, h_d\in\mathbb{Z}_N \right) \\
&= \mathbb{E}\left( \prod_{\omega'\in\{0,1\}^{d-1}} f(x+\omega'\cdot h') \middle| x\in\mathbb{Z}_N, h'\in\mathbb{Z}_N^{d-1} \right) \\
&= ||f||_{U^{d-1}}^{2^{d-1}}.
\end{aligned}$$

Por lo tanto

$$||f||_{U^{d-1}}^{2^{d-1}} \leq ||f||_{U^d}^{2^{d-1}},$$

y entonces

$$||f||_{U^{d-1}} \leq ||f||_{U^d} \tag{2.17}$$

para todo $d \geq 2$. Puesto que la $U^2$-norma de Gowers es una norma genuina, tenemos que las $U^d$-normas de Gowers son también normas auténticas. Si $||f||_{U^d} = 0$, $d \geq 2$, de la expresión (2.17), $||f||_{U^2} = 0$ y por lo tanto $f = 0$. Por otro lado, la $U^1$-norma no es una norma. En efecto, dado $f : \mathbb{Z}_N \to \mathbb{R}$ se tiene

$$\begin{aligned}
||f||_{U^1} &= \left\langle (f)_{\omega\in\{0,1\}} \right\rangle_{U^1}^{1/2} \\
&= \mathbb{E}\left( \prod_{\omega\in\{0,1\}} f(x+\omega\cdot h) \middle| x\in\mathbb{Z}_N, h\in\mathbb{Z}_N \right)^{1/2} = \left( \frac{1}{N^2} \sum_{x\in\mathbb{Z}_N, h\in\mathbb{Z}_N} f(x)f(x+h) \right)^{1/2} \\
&= \left( \frac{1}{N^2} \sum_{x\in\mathbb{Z}_N} f(x) \sum_{h\in\mathbb{Z}_N} f(x+h) \right)^{1/2} = \left( \frac{1}{N} \sum_{x\in\mathbb{Z}_N} f(x)\mathbb{E}(f) \right)^{1/2} = |\mathbb{E}(f)|.
\end{aligned}$$

Luego, que $||f||_{U^1} = 0$, esto es, $|\mathbb{E}(f)| = 0$ no implica que $f = 0$.

## 2.3. La topología débil estrella

Sea $\big(\mathcal{N}, ||\cdot||\big)$ un espacio vectorial normado sobre $\mathbb{F} = \mathbb{R}$ o $\mathbb{F} = \mathbb{C}$. Definimos el espacio dual de $\mathcal{N}$ como

$$\mathcal{N}^* = \Big\{ f : \mathcal{N} \to \mathbb{F};\ f \text{ es lineal y continua} \Big\}. \tag{2.18}$$

Este es un espacio normado con la norma $||f||_{\mathcal{N}^*} = \sup\big\{|f(z)|;\ ||z||_{\mathcal{N}} \leq 1\big\}$ para $f \in \mathcal{N}^*$ ($||f||_{\mathcal{N}^*} < +\infty$ pues $f$ es continua), y es un espacio de Banach pues $\mathbb{F}$ es completo.
Sea $\mathcal{N}^{**} = (\mathcal{N}^*)^*$, llamado segundo dual o bidual de $\mathcal{N}$. Para cada $\xi \in \mathcal{N}$ asociamos la funcional de evaluación en el punto $\xi$, $\widehat{\xi} \in \mathcal{N}^{**}$, definida por

$$\widehat{\xi}(f) = f(\xi), \text{ para cada } f \in \mathcal{N}^*. \tag{2.19}$$

**Proposición 2.2.** *La aplicación $\widehat{\ } : \mathcal{N} \to \mathcal{N}^{**}$ definida en (2.19) es una isometría lineal.*

*Demostración.* La linealidad de $\widehat{\ }$ es clara. El hecho que $\widehat{\xi} \in \mathcal{N}^{**}$ para $\xi \in \mathcal{N}$ se sigue de

$$||\widehat{\xi}||_{\mathcal{N}^{**}} = \sup_{f \in \mathcal{N}^*,\, ||f||_{\mathcal{N}^*} \leq 1} |\widehat{\xi}(f)| = \sup_{f \in \mathcal{N}^*,\, ||f||_{\mathcal{N}^*} \leq 1} |f(\xi)| \leq \sup_{f \in \mathcal{N}^*,\, ||f||_{\mathcal{N}^*} \leq 1} ||f||_{\mathcal{N}^*} ||\xi||,$$

y por tanto

$$||\widehat{\xi}||_{\mathcal{N}^{**}} \leq ||\xi||_{\mathcal{N}}. \tag{2.20}$$

Si $\xi = 0$ tenemos que $||\widehat{\xi}||_{\mathcal{N}^{**}} = ||\xi||_{\mathcal{N}}$. Sea $\mathcal{B} = \langle \xi \rangle$ el subespacio generado por $\xi \neq 0$. Definamos $\tilde{g} : \mathcal{B} \to \mathbb{F}$ tal que $\tilde{g}(t\xi) = t||\xi||_{\mathcal{N}}$. Tenemos que

$$\tilde{g}(x) \leq p(x), \text{ para todo } x \in \mathcal{B},$$

donde $p : \mathcal{N} \to \mathbb{R}$ es definida por $p(x) = ||x||$, para cada $x \in \mathcal{N}$. La función $p$ es sublineal (esto es, $p(x+y) \leq p(x) + p(y)$ y $p(tx) = tp(x)$, para todo $x, y \in \mathcal{N}$ y todo $t > 0$).
Por el teorema de Hanhn-Banach, existe $g : \mathcal{N} \to \mathbb{F}$, extensión de $\tilde{g}$, esto es

$$g(x) = \tilde{g}(x), \text{ para cada } x \in \mathcal{B}, \tag{2.21}$$

tal que

$$g(x) \leq p(x) = ||x||, \text{ para todo } x \in \mathcal{N}. \tag{2.22}$$

Así, $g(\xi) = \tilde{g}(\xi) = ||\xi||$. Por lo tanto

$$||g||_{\mathcal{N}^*} \geq |g(\xi/||\xi||)| = |g(\xi)|/||\xi|| = 1,$$

y de (2.22), $||g||_{\mathcal{N}^*} \leq 1$. Por consiguiente $||g||_{\mathcal{N}^*} = 1$.

Así, tenemos

$$||\widehat{\xi}||_{\mathcal{N}^{**}} = \sup_{f \in \mathcal{N}^*, ||f||_{\mathcal{N}^*} \leq 1} |\widehat{\xi}(f)| \geq |\widehat{\xi}(g)| = |g(\xi)| = ||\xi||. \tag{2.23}$$

De (2.20) y (2.23) concluimos que $||\widehat{\xi}||_{\mathcal{N}^{**}} = ||\xi||_{\mathcal{N}}$. □

**Definición 2.6.** *Una sucesión* $(\xi_n)_n \subseteq \mathcal{N}$ converge débilmente *a* $\xi \in \mathcal{N}$ *si*

$$\lim_{n\to\infty} f(\xi_n) = f(\xi) \text{ para todo } f \in \mathcal{N}^*,$$

*lo que denotaremos por* $\xi_n \xrightarrow{w} \xi$ *o* $\xi_n \rightharpoonup \xi$.

Si $(\xi_n)_n$ converge a $\xi$ fuertemente, esto es, en el sentido usual de la norma $||\cdot||_{\mathcal{N}}$, entonces

$$\xi_n \rightharpoonup \xi,$$

pues, por la continuidad de los elementos de $\mathcal{N}^*$ tenemos que

$$|f(\xi_n) - f(\xi)| = |f(\xi_n - \xi)| \leq ||f||_{\mathcal{N}^*} ||\xi_n - \xi||_{\mathcal{N}}.$$

**Proposición 2.3.** *Sea* $(\xi_n)_n \subset \mathcal{N}$ *una sucesión tal que* $\xi_n \rightharpoonup \xi$. *Entonces el límite* $\xi$ *es único y* $(\xi_n)_n$ *es acotada.*

*Demostración.* Supongamos que $\xi_n \rightharpoonup \xi$ y $\xi_n \rightharpoonup \eta$. Entonces para todo $f \in \mathcal{N}^*$ se tiene que

$$f(\xi - \eta) = f(\xi) - f(\eta) = \lim_{n\to\infty} \big(f(\xi_n) - f(\xi_n)\big) = 0. \tag{2.24}$$

Supongamos que $\xi \neq \eta$, sea $\theta = \xi - \eta$, $\theta \neq 0$. Consideremos $p : \mathcal{N} \to \mathbb{R}$, dada por $p(x) = ||x||$, el subespacio $Z = \langle\theta\rangle$ y la funcional lineal $\tilde{g} : Z \to \mathbb{R}$ definida por $\tilde{g}(t\theta) = t||\theta||$ para todo $t \in \mathbb{R}$. Por el teorema de Banach-Steinhauss existe $g : \mathcal{N} \to \mathbb{R}$ tal que

$$g(x) = \tilde{g}(x) \text{ para todo } x \in Z,$$

y

$$g(x) \leq p(x) = ||x|| \text{ para todo } x \in \mathcal{N}.$$

Como antes, deducimos que $||g|| = 1$ y $g(\theta) = ||\theta|| > 0$, lo que contradiría (2.24). Por lo tanto $\xi = \eta$.

Ahora mostremos que $(\xi_n)_n$ es acotada. De la proposición 2.2 tenemos que $\widehat{\xi}_n \in \mathcal{N}^{**}$ y

$||\widehat{\xi}_n|| = ||\xi_n||$. Para cada $f \in \mathcal{N}^*$, $\widehat{\xi}_n(f) = f(\xi_n)$ para todo $n \in \mathbb{N}$, y $\big(f(\xi_n)\big)_n$ es convergente, luego acotada. Así

$$\sup_{n\in\mathbb{N}} |\widehat{\xi}_n(f)| < +\infty \text{ para cada } f \in \mathcal{N}^*. \tag{2.25}$$

Como $\mathcal{N}^*$ es un espacio de Banach, de (2.25) y del teorema de acotación uniforme

$$\sup_{n\in\mathbb{N}} ||\xi_n||_{\mathcal{N}} = \sup_{n\in\mathbb{N}} ||\widehat{\xi}_n||_{\mathcal{N}^{**}} < +\infty.$$

□

Recordemos que la convergencia débil en un espacio normado $\mathcal{N}$ se define a través de su dual $\mathcal{N}^*$. De manera análoga se puede definir la convergencia débil en $\mathcal{N}^*$ a través de $\mathcal{N}^{**}$. Como $\widehat{\mathcal{N}} \subset \mathcal{N}^{**}$, podemos introducir una noción de convergencia en $\mathcal{N}^*$, aún más débil usando $\widehat{\mathcal{N}}$ en vez de $\mathcal{N}^{**}$, que llamaremos convergencia débil estrella.

**Definición 2.7.** *Dado un espacio normado $\mathcal{N}$, se dice que la sucesión $(f_n)_n \subset \mathcal{N}^*$* **converge débilmente estrella** *a $f \in \mathcal{N}^*$, si*

$$\lim_{n\to\infty} \widehat{\xi}(f_n) = \widehat{\xi}(f) \text{ para todo } \widehat{\xi} \in \widehat{\mathcal{N}}.$$

*En este caso denotaremos $f_n \xrightarrow{w^*} f$.*

**Proposición 2.4.** *Sea $(f_n)_n \subset \mathcal{N}^*$ una sucesión tal que $f_n \xrightarrow{w^*} f$. Entonces el límite $f$ es único. Además, $\widehat{\mathcal{N}}$ separa puntos de $\mathcal{N}^*$ y, si $\mathcal{N}$ es un espacio de Banach, entonces la sucesión $\big(||f_n||\big)_n$ es acotada.*

*Demostración.* Dados $f, g \in \mathcal{N}^*$. Supongamos que para todo $\xi \in \mathcal{N}$ se tenga que $\widehat{\xi}(f) = \widehat{\xi}(g)$, entonces

$$f(\xi) = g(\xi) \text{ para todo } \xi \in \mathcal{N}, \text{ esto es, } f = g. \tag{2.26}$$

Esto muestra la unicidad del límite.

De (2.26), si $f \neq g$ entonces existe $\xi \in \mathcal{N}$ tal que $\widehat{\xi}(f) \neq \widehat{\xi}(g)$.

Como $f_n \xrightarrow{w^*} f$ tenemos que para todo $\xi \in \mathcal{N}$, la sucesión $\big(\widehat{\xi}(f_n)\big)_n = \big(f_n(\xi)\big)_n$ es convergente y por lo tanto acotada, esto es

$$\sup_{n\in\mathbb{N}} |f_n(\xi)| < +\infty \text{ para todo } \xi \in \mathcal{N}.$$

Así, si $\mathcal{N}$ es un espacio de Banach, por el teorema de la acotación uniforme

$$\sup_{n\in\mathbb{N}} ||f_n|| < +\infty.$$

□

**Definición 2.8.** *Sea $\mathcal{N}$ un espacio normado con norma $||\cdot|| = ||\cdot||_{\mathcal{N}}$. La* **topología fuerte** *en $\mathcal{N}$ es la topología métrica inducida por la norma $||\cdot||$. Una base de esta topología está dada por las bolas*

$$B_{\mathcal{N}}(\xi, r) = \Big\{\eta \in \mathcal{N};\, ||\eta - \xi|| < r\Big\},$$

*con centro en $\xi \in \mathcal{N}$ y radio $r > 0$.*

En lo que sigue $\mathbb{F}$ denotará al cuerpo $\mathbb{R}$ o $\mathbb{C}$ sobre el cual se define el espacio normado.

**Definición 2.9.** *Sea $\mathcal{N}$ un espacio normado con norma $||\cdot|| = ||\cdot||_{\mathcal{N}}$. La* **topología débil en** $\mathcal{N}$ *es la topología $\tau(\mathcal{N}, \mathcal{N}^*)$ generada por las funcionales lineales en $\mathcal{N}^*$, es decir, es la topología menos fina en $\mathcal{N}$ en la que todos los elementos de $\mathcal{N}^*$ son continuos. Una subbase (abierta) de $\tau(\mathcal{N}, \mathcal{N}^*)$ es la colección*

$$V(\xi; f; r) = f^{-1}B_{\mathbb{F}}\big(f(\xi); r\big) = \{\eta \in \mathcal{N} :\, |f(\xi) - f(\eta)| < r\},$$

*donde $\xi \in \mathcal{N}$, $r > 0$ y $f \in \mathcal{N}^*$.*

**Definición 2.10.** *Sea $\mathcal{N}$ un espacio normado con norma $||\cdot|| = ||\cdot||_{\mathcal{N}}$. La* **topología débil estrella** *en $\mathcal{N}^*$ es la topología $\tau(\mathcal{N}^*, \widehat{\mathcal{N}})$ generada por las funcionales lineales en $\widehat{\mathcal{N}}$, es decir, es la topología menos fina en $\mathcal{N}^*$ en la que todos los elementos de $\widehat{\mathcal{N}}$ sean continuos. Una subbase (abierta) de $\tau(\mathcal{N}^*, \widehat{\mathcal{N}})$ es la colección*

$$V^*(f; \xi; r) = \widehat{\xi}^{-1}B_{\mathbb{F}}\big(\widehat{\xi}(f); r\big) = \{g \in \mathcal{N}^* :\, |\widehat{\xi}(f) - \widehat{\xi}(g)| < r\}, \tag{2.27}$$

*donde $f \in \mathcal{N}^*$, $\xi \in \mathcal{N}$ y $r > 0$.*

Un elemento típico de la base generada por la subbase dada de la topología débil estrella $\tau(\mathcal{N}^*, \widehat{\mathcal{N}})$ es

$$V^*(f; \xi_1, \ldots, \xi_n; r) = \Big\{g \in \mathcal{N}^* :\, \max_{1 \le j \le n} \big|\widehat{\xi}_j(f) - \widehat{\xi}_j(g)\big| < r\Big\}.$$

Un elemento típico de la base generada por la subbase dada de la topología débil $\tau(\mathcal{N}, \mathcal{N}^*)$ es

$$V(\xi; f_1, \ldots, f_n; r) = \Big\{\eta \in \mathcal{N} :\, \max_{1 \le j \le n} \big|f_j(\xi) - f_j(\eta)\big| < r\Big\},$$

Uno de los motivos para introducir la topología débil estrella es el hecho que, si $\dim \mathcal{N} = +\infty$ la bola cerrada $\overline{B}_{\mathcal{N}^*}(0; 1)$ no es compacta en la topología usual de $\mathcal{N}^*$ (como espacio normado con norma $||h||_{\mathcal{N}^*} = \sup_{x \in \mathcal{N},\, ||x|| \le 1} |h(x)|$). Sin embargo, tenemos lo siguiente (cf. [20]).

**Teorema 2.1** (Alaoglu)**.** *Sea $\mathcal{N}$ un espacio normado con norma $||\cdot|| = ||\cdot||_{\mathcal{N}}$. Entonces la bola cerrada*

$$B^* = \overline{B}_{\mathcal{N}^*}(0;1) = \{f \in \mathcal{N}^*;\ ||f|| \leq 1\}$$

*es un espacio topológico de Hausdorff compacto en la topología débil estrella.*

*Demostración.* Dados $f, g \in \mathcal{N}^*$ tales que $f \neq g$, entonces existe $\xi \in \mathcal{N}$ tal que

$$0 < \delta = |f(\xi) - g(\xi)| = |\widehat{\xi}(f) - \widehat{\xi}(g)|.$$

Así, $V^*\big(f;\xi;\delta/3\big)$ y $V^*\big(g;\xi;\delta/3\big)$ son vecindades abiertas, no vacías y disjuntas de $f$ y $g$, respectivamente.

A cada $\xi \in \mathcal{N}$ asociemos $K_\xi = \{z \in \mathbb{F};\ |z| \leq ||\xi||\}$, $K_\xi$ es compacto en $\mathbb{F}$, y por el teorema de Tychonov, el producto cartesiano

$$K = \prod_{\xi \in \mathcal{N}} K_\xi = \Big\{f : \mathcal{N} \to \mathbb{F};\ f \text{ es función },|f(\xi)| \leq ||\xi|| \text{ para cada } \xi \in \mathcal{N}\Big\}$$

de todos los $K_\xi$ es compacto en la topología producto. Así, la bola unitaria $B^*$ es un subconjunto de $K$. Más aún,

$$B^* = K \cap \mathcal{N}^*. \tag{2.28}$$

Para $f \in B^*$ consideremos $V^*(f;\xi;r)$ como en (2.27), y

$$U(f;\xi;r) = \{g \in K;\ |f(\xi) - g(\xi)| < r\} = p_\xi^{-1}\big(B_{\mathbb{F}}(f(\xi);r)\big),$$

donde $p_\xi : K \to K_\xi$ es la proyección $\xi$-ésima (o sea, $p_\xi(f) = f(\xi)$). La familia formada por los $U(f;\xi;r)$ donde $\xi \in \mathcal{N}$ y $r > 0$ es una subbase local de vecindades de $f \in B^*$ en la topología producto.

Como $B^* = K \cap \mathcal{N}^*$, tenemos que

$$\begin{aligned} V^*(f;\xi;r) \cap B^* &= \{g \in \mathcal{N}^*;\ |\widehat{\xi}(g) - \widehat{\xi}(f)| < r\} \cap B^* \\ &= \{g \in K;\ |\widehat{\xi}(g) - \widehat{\xi}(f)| < r\} \cap \mathcal{N}^* \\ &= U(f;\xi;r) \cap B^*. \end{aligned}$$

Así, la topología débil* de $\mathcal{N}^*$ y la topología producto de $K$ coinciden en $B^*$. Como $B^*$ es un subconjunto del compacto $K$, considerando $K$ provisto de la topología producto, solamente resta mostrar que $B^*$ es un subconjunto cerrado de $K$ (y por ende compacto en la topología producto de $K$ en $B^*$).

Sea $g$ un elemento en la cerradura de $B^*$ en $K$. Por la definición de $K$ tenemos que $|g(\xi)| \leq ||\xi||$. Luego, de (2.28) resta mostrar que $g$ es lineal. Dados $\xi, \eta \in \mathcal{N}$ y $\epsilon > 0$, existe

$$h \in B^* \cap U(g; \xi; \epsilon/3) \cap U(g; \eta; \epsilon/3) \cap U(g; \xi + \eta; \epsilon/3).$$

De la linealidad de $h$ se sigue que

$$\begin{aligned} |g(\xi+\eta) - g(\xi) - g(\eta)| &= |g(\xi+\eta) - h(\xi+\eta) + h(\xi+\eta) - g(\xi) - g(\eta)| \\ &= |g(\xi+\eta) - h(\xi+\eta) + h(\xi) - g(\xi) + h(\eta) - g(\eta)| \\ &\leq |g(\xi+\eta) - h(\xi+\eta)| + |h(\xi) - g(\xi)| + |h(\eta) - g(\eta)| < \epsilon, \end{aligned}$$

por lo tanto $g(\xi + \eta) = g(\xi) + g(\eta)$, puesto que $\epsilon$ fue arbitrario.
Análogamente, dados $\xi \in \mathcal{N}$ y $\lambda \in \mathbb{F} \setminus \{0\}$ y $\epsilon > 0$, existe

$$h_1 \in B^* \cap U(g; \xi; \epsilon/2|\lambda|) \cap U(g; \lambda\xi; \epsilon/2).$$

Por la linealidad de $h_1$,

$$\begin{aligned} |g(\lambda\xi) - \lambda g(\xi)| &= |g(\lambda\xi) - h_1(\lambda\xi) + h_1(\lambda\xi) - \lambda g(\xi)| \\ &= |g(\lambda\xi) - h_1(\lambda\xi) + \lambda h_1(\xi) - \lambda g(\xi)| \\ &\leq |g(\lambda\xi) - h_1(\lambda\xi)| + |\lambda h_1(\xi) - \lambda g(\xi)| < \epsilon, \end{aligned}$$

por tanto $g(\lambda\xi) = \lambda g(\xi)$, dado que $\epsilon$ fue arbitrario. Cuando $\lambda = 0$ es claro que $g(\lambda\xi) = \lambda g(\xi)$ puesto que $g \in K$. Así $g \in B^*$, y por lo tanto $B^*$ es cerrado en $K$, considerando $K$ provisto de la topología producto. □

Como una aplicación del teorema de Alaoglu tenemos el siguiente resultado.

**Teorema 2.2** (Krylov-Bogolioubov)**.** *Sea $\Omega$ un espacio métrico compacto. Entonces para toda aplicación continua $A : \Omega \to \Omega$ existe una medida de probabilidad (boreliana) $\nu$ tal que $A$ es $\nu$-invariante.*

Para la prueba, requerimos de las siguientes proposiciones.

**Proposición 2.5.** *Sea $(X, \tau)$ un espacio topológico compacto. Si existe una sucesión de funciones continuas $\left(f_n : X \to \mathbb{F}\right)_{n\in\mathbb{N}}$ que separa puntos de $X$, entonces $(X, \tau)$ es metrizable, esto es, existe una métrica $d$ en $X$ tal que $\tau$ es la colección de conjuntos abiertos en el espacio métrico $X$.*

*Demostración.* Podemos suponer que $||f_n||_\infty > 0$, para todo $n \in \mathbb{N}$. Definamos $d : X \times X \to \mathbb{R}$ por

$$d(x,t) = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{|f_n(x) - f_n(t)|}{2||f_n||_\infty}, \text{ donde } x, t \in X.$$

Notemos que la serie anterior converge uniformemente en $X \times X$, puesto que dados $x, t \in X$,

$$0 \leq \frac{1}{2^n} \frac{|f_n(x) - f_n(t)|}{2||f_n||_\infty} \leq \frac{1}{2^n} \frac{|f_n(x)| + |f_n(t)|}{2||f_n||_\infty} \leq \frac{1}{2^n} \text{ para cada } n \in \mathbb{N}.$$

Mostremos que $d$ es una métrica en $X$. De hecho, $d(x,t) \geq 0$ para todo $x, t \in X$. Además como $(f_n)$ separa puntos de $X$, dados $x, t \in X$ con $x \neq t$, tenemos que existe $n \in \mathbb{N}$ tal que $f_n(x) \neq f_n(t)$, y por lo tanto $d(x,t) > 0$. Claramente, $d(x,t) = d(t,x)$, para todo $x, t \in X$. Dados $x, y, t \in X$, $|f_n(x) - f_n(t)| \leq |f_n(x) - f_n(y)| + |f_n(y) - f_n(t)|$, de donde

$$d(x,t) \leq d(x,y) + d(y,t).$$

Para todo $n \in \mathbb{N}$ definimos la función continua $g_n : X \times X \to \mathbb{F}$ por

$$g_n(x,t) = \frac{1}{2^n} \frac{|f_n(x) - f_n(t)|}{2||f_n||_\infty}.$$

Tenemos que $\sum_{n=1}^{\infty} g_n$ converge uniformemente a $d$ en $X \times X$. Por ende, $d$ es una función continua. En particular, para todo $x \in X$ la función $d_x : (X, \tau) \to [0, +\infty)$ dada por

$$d_x(t) = d(x,t), \text{ para todo } t \in X,$$

es continua. Así, las bolas $B(x;r) = d_x^{-1}([0,r))$ con centro $x \in X$ y radio $r > 0$ son abiertas en $(X, \tau)$, Por lo tanto, la topología $\tau_d$ generada por esta métrica en $X$ es tal que $\tau_d \subseteq \tau$. Recíprocamente veamos que $\tau \subseteq \tau_d$. Sea $A$ un abierto en $(X, \tau)$, entonces $F = A^c$ es cerrado en $(X, \tau)$ y, siendo $(X, \tau)$ compacto, tenemos que también $F$ es compacto en $(X, \tau)$. Como $\tau_d \subseteq \tau$, todo cubrimiento de $F$ por abiertos de $\tau_d$ es también un cubrimiento por abiertos de $\tau$; por tanto posee un subcubrimiento abierto finito de $F$, y así $F$ es compacto en $(X, \tau_d)$. Como todo compacto en un espacio métrico es cerrado, se sigue que $F$ es cerrado en $(X, \tau_d)$, y por lo tanto $A$ es abierto en $(X, \tau_d)$. □

**Proposición 2.6.** *Sea $\mathcal{N}$ un espacio normado. Si $\mathcal{N}$ es separable y $S \subset \mathcal{N}^*$ es compacto en la topología débil estrella, entonces $S$ es metrizable en la topología débil estrella.*

*Demostración.* Como $\mathcal{N}$ es separable, existe una sucesión $(\xi_n)_{n \in \mathbb{N}}$ densa en $\mathcal{N}$. Por la definición de topología débil estrella en $\mathcal{N}^*$ tenemos que cada $\widehat{\xi_n} : \mathcal{N}^* \to \mathbb{F}$ es continua.

Además, la sucesión $(\widehat{\xi_n})_n$ separa puntos de $\mathcal{N}^*$. En efecto, si para todo $n \in \mathbb{N}$ se tiene que $\widehat{\xi_n}(f) = \widehat{\xi_n}(g)$, donde $f, g \in \mathcal{N}^*$; tenemos que $f(\xi_n) = g(\xi_n)$ para todo $n \in \mathbb{N}$. Por tanto, $f = g$, pues ambas funciones continuas, $f$ y $g$, coinciden en un conjunto denso en $\mathcal{N}$. Así, $(\widehat{\xi_n})_n$ es una sucesión de funciones continuas en $\mathcal{N}^*$ ($\mathcal{N}^*$ con ta topología débil estrella), que separa puntos de $\mathcal{N}^*$. En particular $(\widehat{\xi_n})_n$ separa puntos de $S$. De la proposición 2.5 se sigue que $S$ es metrizable en la topología débil estrella. □

**Proposición 2.7.** *Sea $\Omega$ un espacio métrico compacto. Entonces $C(\Omega)$, el espacio de las funciones continuas $f : \Omega \to \mathbb{F}$, es un espacio métrico separable.*

**Corolario 2.1.** *Sea $\Omega$ un espacio métrico compacto, y $\mathcal{B} = C(\Omega)$ el espacio de funciones continuas sobre $\Omega$, $f : \Omega \to \mathbb{F}$. Entonces para todo $r > 0$, la bola cerrada de radio $r$, $\overline{B}_{\mathcal{B}^*}(0, r)$, es secuencialmente compacta en la topología débil estrella de $\mathcal{B}^*$.*

*Demostración.* Del teorema de Alaoglu (teorema 2.1) tenemos que $\overline{B}_{\mathcal{B}^*}(0; r)$ es compacto en la topología débil estrella de $\mathcal{B}^*$, y de la proposición 2.7, $\mathcal{B}$ es separable. Luego, de la proposición 2.6 tenemos que $\overline{B}_{\mathcal{B}^*}(0; r)$ es metrizable en la topología débil estrella, y por ende, $\overline{B}_{\mathcal{B}^*}(0; r)$ es secuencialmente compacto. □

*Prueba del teorema de Krylov-Bogolioubov 2.2.* Sea $A : \Omega \to \Omega$ una aplicación continua y $\mathcal{B} = C(\Omega)$. Por el teorema 5.6 (Riesz-Markov) $\mathcal{B}$ es identificado con $M(\Omega)$, a través de la isometría $G : M(\Omega) \to \mathcal{B}^*$, dada por $G(\mu) = G_\mu$, para cada $\mu \in M(\Omega)$, y

$$G_\mu(\psi) = \int_\Omega \psi d\mu \text{ para todo } \psi \in \mathcal{B}.$$

Para $\widehat{\psi} \in \mathcal{B}^{**}$ ($\psi \in \mathcal{B}$)

$$\widehat{\psi}(G_\mu) = G_\mu(\psi) = \int_\Omega \psi d\mu, \text{ para todo } \mu \in M(\Omega).$$

Para $\xi \in \Omega$, definamos $\tilde{\delta}_\xi \in \mathcal{B}^*$ por

$$\tilde{\delta}_\xi(\psi) = \psi(\xi), \text{ para todo } \psi \in \mathcal{B}. \tag{2.29}$$

Como $|\tilde{\delta}_\xi(\psi)| = |\psi(\xi)| \leq \sup_{\alpha \in \Omega} |\psi(\alpha)| = ||\psi||$, entonces $\tilde{\delta}_\xi \in \mathcal{B}^*$, y de hecho

$$||\tilde{\delta}_\xi|| \leq 1, \text{ para todo } \xi \in \Omega. \tag{2.30}$$

Gracias a la isometría $G : M(\Omega) \to \mathcal{B}^*$, para cada $\xi \in \Omega$ existe una medida $\delta_\xi \in M(\Omega)$ tal que $G(\delta_\xi) = \tilde{\delta}_\xi$. Por lo tanto $\int_\Omega \psi d\delta_\xi = G_{\delta_\xi}(\psi) = \tilde{\delta}_\xi(\psi) = \psi(\xi)$, para todo $\psi \in \mathcal{B}$.

Fijemos $\xi \in \Omega$, y para cada $n \in \mathbb{N}$ consideremos

$$\tilde{\mu}_n = \frac{1}{n}\sum_{j=0}^{n-1} \tilde{\delta}_{A^j(\xi)} \in \mathcal{B}^*.$$

De la desigualdad (2.30) tenemos que

$$||\tilde{\mu}_n||_{\mathcal{B}^*} \leq \frac{1}{n}\sum_{j=0}^{n-1} ||\tilde{\delta}_{A^j(\xi)}||_{\mathcal{B}^*} \leq \frac{1}{n}\sum_{j=0}^{n-1} 1 = 1.$$

Así, la sucesión de funcionales $(\tilde{\mu}_n)_n$ en $\mathcal{B}^*$ está contenida en la bola cerrada $B^* = \overline{B}_{\mathcal{B}^*}(0;1)$. Del corolario 2.1 tenemos que $\overline{B}_{\mathcal{B}^*}(0;r)$ es secuencialmente compacto en la topología débil estrella de $\mathcal{B}^*$. Por lo tanto existe $\tilde{\nu} \in \mathcal{B}^*$, (y por la isometría $G$ existe también $\nu \in M(\Omega)$ con $G(\nu) = \tilde{\nu}$), tal que una subsucesión $(\tilde{\mu}_{k_n})_n$ de $(\tilde{\mu}_n)_n$ converge a $\tilde{\nu}$ (en la topología débil estrella). Así, para todo $\psi \in \mathcal{B}$ se tiene que

$$\begin{aligned}\int_\Omega \psi \, d\nu = [G(\nu)](\psi) = \tilde{\nu}(\psi) = \widehat{\psi}(\tilde{\nu}) &= \lim_{n\to+\infty} \widehat{\psi}(\tilde{\mu}_{k_n}) \\ &= \lim_{n\to+\infty} \tilde{\mu}_{k_n}(\psi) = \lim_{n\to+\infty} \frac{1}{k_n}\sum_{j=0}^{k_n-1} \tilde{\delta}_{A^j(\xi)}(\psi) \\ &= \lim_{n\to+\infty} \frac{1}{k_n}\sum_{j=0}^{k_n-1} \psi(A^j(\xi)), \qquad (2.31)\end{aligned}$$

y

$$\begin{aligned}\int_\Omega (\psi\circ A) \, d\nu = [G(\nu)](\psi\circ A) = \tilde{\nu}(\psi\circ A) = \widehat{\psi\circ A}(\tilde{\nu}) &= \lim_{n\to+\infty} \widehat{\psi\circ A}(\tilde{\mu}_{k_n}) \\ &= \lim_{n\to+\infty} \tilde{\mu}_{k_n}(\psi\circ A) = \lim_{n\to+\infty} \frac{1}{k_n}\sum_{j=0}^{k_n-1} \tilde{\delta}_{A^j(\xi)}(\psi\circ A) \\ &= \lim_{n\to+\infty} \frac{1}{k_n}\sum_{j=0}^{k_n-1} \psi(A^{j+1}(\xi)). \qquad (2.32)\end{aligned}$$

Por lo tanto de (2.31) y (2.32), tenemos que

$$\begin{aligned}\int_\Omega (\psi\circ A) \, d\nu - \int_\Omega \psi \, d\nu &= \lim_{n\to+\infty} \frac{1}{k_n}\sum_{j=0}^{k_n-1} \Big(\psi\circ A^{j+1}(\xi) - \psi\circ A^j(\xi)\Big) \\ &= \lim_{n\to+\infty} \frac{1}{k_n}.\Big(\psi\circ A^{k_n}(\xi) - \psi(\xi)\Big)\end{aligned}$$

Siendo $\psi$ acotada (puesto que es continua en un compacto) se tiene que

$$\lim_{n\to+\infty} \frac{1}{k_n}\Big(\psi\circ A^{k_n}(\xi) - \psi(\xi)\Big) = 0,$$

y por lo tanto

$$\int_\Omega \psi \circ A \, d\nu = \int_\Omega \psi \, d\nu \text{ para todo } \psi \in \mathcal{B} = C(\Omega).$$

Para $\psi = 1$, en (2.31), tenemos

$$\nu(\Omega) = \int_\Omega 1 \, d\nu = \lim_{k_n \to +\infty} \frac{1}{k_n} \sum_{j=0}^{k_n - 1} 1(A^j(\xi)) = 1.$$

□

# Capítulo 3

# Dos perspectivas del teorema de Roth

Probaremos el teorema de Roth, es decir, el teorema de Szemerédi para progresiones de longitud tres. El teorema de Szemerédi clásico muestra la existencia de progresiones aritméticas de longitud $k$ ($k \geq 3$), en subconjuntos de $[1, N]$ con densidad positiva, para $N$ suficientemente grande. Abordaremos el caso $k = 3$ por métodos analíticos, a través de la transformada de Fourier discreta en $\mathbb{Z}_N$. Estos métodos, empleados inicialmente por Roth en [3], produjeron mejores cotas sobre las constantes $N_0(\delta, 3)$ y $r_3(A)$ comparadas con aquellas obtenidas mediante procedimientos combinatorios, donde $\delta$ está relacionada con la densidad del subconjunto de $[1, N]$, y $k$ con la longitud de la progresión aritmética; vea la proposición 6.1.

## 3.1. El teorema de Roth, método analítico

En esta sección mostraremos el teorema de Szemerédi en su versión más sencilla, esto es, para progresiones aritméticas de longitud tres.

**Teorema 3.1** (Roth)**.** *Existe alguna constante positiva $C \in \mathbb{R}$ tal que para todo $\delta \in ]0, 1[$, si $N \geq \exp\exp(C\delta^{-1})$, entonces cualquier conjunto $A \subseteq \{0, 1, \ldots, N-1\}$ con $|A| = \delta N$, contiene necesariamente alguna progresión aritmética no trivial, cuya diferencia común es diferente de cero, de longitud tres.*

Como veremos en (3.6), la definición siguiente será conveniente en la prueba del teorema

de Roth y se dará naturalmente.

**Definición 3.1.** *Sean $\delta > 0$ y $A \subseteq \mathbb{Z}_N$ tal que $|A| = \delta N$. Definimos la* **función balanceada** de $A$ *como $f_A : \mathbb{Z}_N \to [-1, 1]$, donde*

$$f_A(s) = \begin{cases} 1 - \delta & , \text{ si } s \in A, \\ -\delta & , \text{ si } s \notin A. \end{cases}$$

*Esto es, $f_A = \chi_A - \delta \cdot I$, donde $\chi_A : \mathbb{Z}_N \to \mathbb{R}$ denota a la función característica de $A$, e $I : \mathbb{Z}_N \to \mathbb{R}$ denota a la función unidad.*

Notemos que

$$\widehat{f_A}(0) = \sum_{s \in \mathbb{Z}_N} f_A(s) = 0. \tag{3.1}$$

En efecto,

$$\begin{aligned} \widehat{f_A}(0) &= \sum_{s \in \mathbb{Z}_N} f_A(s)\omega^{-s\cdot 0} \\ &= \sum_{s \in \mathbb{Z}_N} f_A(s) = \sum_{s \in \mathbb{Z}_N} \chi_A(s) - \delta \sum_{s \in \mathbb{Z}_N} I(s) \\ &= |A| - \delta N = 0. \end{aligned}$$

También tenemos

$$\widehat{f_A}(r) = \widehat{\chi_A}(r) \text{ para cada } r \neq 0. \tag{3.2}$$

Veamos, dado $r \neq 0$ tenemos

$$\begin{aligned} \widehat{f_A}(r) &= \sum_{s \in \mathbb{Z}_N} f_A(s)\omega^{-sr} = \sum_s \big(\chi_A(s) - \delta\big)\omega^{-sr} \\ &= \sum_s \chi_A(s)\omega^{-sr} - \delta \sum_s \omega^{-sr} = \widehat{\chi_A}(r). \end{aligned}$$

En lo que sigue $D$ denotará al disco unitario cerrado en $\mathbb{C}$, esto es, $D = \{z \in \mathbb{C};\ |z| \leq 1\}$.

**Lema 3.1.** *Sea $f : \mathbb{Z}_N \to D$. Las siguientes afirmaciones son equivalentes.*

*(i)* $\displaystyle\sum_k \left| \sum_s f(s)\overline{f(s-k)} \right|^2 \leq c_1 N^3.$

*(ii)* $\displaystyle\sum_{a-b=c-d} f(a)\overline{f(b)f(c)}f(d) \leq c_1 N^3.$

*(iii)* $\sum_r \left|\widehat{f}(r)\right|^4 \leq c_1 N^4$.

*(iv)* $\max_r \left|\widehat{f}(r)\right| \leq c_2 N$.

*(v)* $\sum_k \left|\sum_s f(s)\overline{g(s-k)}\right|^2 \leq c_3 N^2 ||g||_2^2$ *para cada función* $g : \mathbb{Z}_N \to \mathbb{C}$.

*Demostración.* Mostremos la equivalencia entre (i) y (ii). De (2.8)

$$\begin{aligned}
\sum_{t\in\mathbb{Z}_N}\left|\sum_{m\in\mathbb{Z}_N} f(m)\overline{f(m-t)}\right|^2 &= \sum_{m,n,t\in\mathbb{Z}_N} f(m)\overline{f(m+t)f(m+n)}f(m+n+t)\\
&= \sum_{m,n} f(m)\overline{f(m+n)}\sum_t \overline{f(m+t)}f(m+n+t)\\
&= \sum_{m,n} f(m)\overline{f(m+n)}\sum_b \overline{f(b)}f(b+n)\\
&= \sum_{a,b} f(a)\overline{f(b)}\sum_n \overline{f(a+n)}f(b+n)\\
&= \sum_{a,b} f(a)\overline{f(b)}\sum_c \overline{f(c)}f(c+b-a)\\
&= \sum_{a,b,c} f(a)\overline{f(b)f(c)}f(c+b-a)\\
&= \sum_{\substack{a,b,c,d\\ d-c=b-a}} f(a)\overline{f(b)f(c)}f(d).
\end{aligned}$$

La equivalencia entre (i) y (iii) se sigue de la identidad (2.8).

Mostremos que (iii) implica (iv). Supongamos que se cumple (iii), esto es $\sum_r \left|\widehat{f}(r)\right|^4 \leq c_1 N^4$. Por lo tanto $\max_r \left|\widehat{f}(r)\right| \leq c_1^{1/4} N$. Así, eligiendo $c_2 = c_1^{1/4}$ obtenemos (iv). En seguida mostremos que (iv) implica (iii). Supongamos que $\max_r \left|\widehat{f}(r)\right| \leq c_2 N$. Puesto que

$$\sum_r |\widehat{f}(r)|^4 \leq \max_t |\widehat{f}(t)|^2 \sum_r |\widehat{f}(r)|^2 \leq N^2 \max_t |\widehat{f}(t)|^2,$$

donde la última desigualdad se da porque $\sum_r |\widehat{g}(r)|^2 = N\sum_r |g(r)|^2 \leq N^2$ para toda función $g : \mathbb{Z}_N \to D$, tenemos que $\sum_r |\widehat{f}(r)|^4 \leq c_2^2 N^4$, y por lo tanto (iv) implica (iii) si elegimos $c_1 = c_2^2$. Es obvio que (v) implica (i), basta elegir $g = f$ y $c_1 = c_3$ en (v), puesto que

$$||f||_2^2 = \sum_r |f(r)|^2 \leq N.$$

Finalmente veamos que (iii) implica (v). Por el lema 2.1 tenemos que para cualquier función $g : \mathbb{Z}_N \to \mathbb{C}$,

$$\sum_k \left|\sum_s f(s)\overline{g(s-k)}\right|^2 = N^{-1}\sum_r |\widehat{f}(r)|^2 |\widehat{g}(r)|^2,$$

y entonces

$$N^{-1}\sum_r |\widehat{f}(r)|^2|\widehat{g}(r)|^2 \leq N^{-1}\left(\sum_r |\widehat{f}(r)|^4\right)^{1/2}\left(\sum_r |\widehat{g}(r)|^4\right)^{1/2}$$

por la desigualdad de Cauchy-Schwarz. Usando además que

$$\left(\sum_r |\widehat{g}(r)|^4\right)^{1/2} \leq \sum_r |\widehat{g}(r)|^2 = N\sum_r |g(r)|^2 = N||g||_2^2,$$

vemos que

$$\sum_k \left|\sum_s f(s)\overline{f(s-k)}\right|^2 \leq \left(\sum_r |\widehat{f}(r)|^4\right)^{1/2}||g||_2^2.$$

Así, si se cumple (iii), esto es, $\sum_r |\widehat{f}(r)|^4 \leq c_1 N^4$, entonces

$$\sum_k \left|\sum_s f(s)\overline{f(s-k)}\right|^2 \leq c_1^{1/2}N^2||g||_2^2.$$

Por ende (iii) implica (v) si elegimos $c_3 = c_1^{1/2}$.

Observamos que cuando decimos que una propiedad involucrando $c_i$ implica otra involucrando $c_j$, significa que si la primera vale, entonces la segunda también vale para una constante $c_j$, que depende de $c_i$, y que tiende a cero cuando $c_i$ tiende a cero; así, si uno va de una afirmación a otra y luego regresamos, no necesariamente recobramos la constante original.

□

**Definición 3.2.** *Sea $f : \mathbb{Z}_N \to D$ una función que satisface la condición (i) del lema anterior con $c_1 = \alpha$. Entonces diremos que $f$ es $\alpha$-***uniforme***. Además, si $f$ es la función balanceada $f_A$ de un conjunto $A \subseteq \mathbb{Z}_N$, entonces diremos que $A$ es $\alpha$-***uniforme***.*

Sea $A \subseteq \mathbb{Z}_N$ un conjunto $\alpha$-uniforme de cardinalidad $\delta N$, y sea $f = f_A$ su función balanceada, entonces

$$\sum_r |\widehat{\chi_A}(r)|^4 = |A|^4 + \sum_r |\widehat{f}(r)|^4 \leq |A|^4 + \alpha N^4.$$

En efecto,

$$\begin{aligned}\sum_r |\widehat{\chi_A}(r)|^4 &= |\widehat{\chi_A}(0)|^4 + \sum_{r\neq 0} |\widehat{\chi_A}(r)|^4 = |A|^4 + \sum_{r\neq 0} |\widehat{f_A}(r)|^4 \\ &= |A|^4 + \sum_{r\in\mathbb{Z}_N} |\widehat{f_A}(r)|^4 \leq |A|^4 + \alpha N^4.\end{aligned}$$

Para mostrar el teorema de Roth, dado $A \subseteq \{0, 1, 2, \ldots, N-1\}$, tal que $|A| = \delta N$, intentamos mostrar que el número de soluciones del siguiente problema,

$$\begin{cases} x + z = 2y \\ x, y, z \in A, \end{cases} \tag{3.3}$$

es positivo para $N$ suficientemente grande. Sea $\eta$ al número de soluciones del problema anterior. Intentemos acotar inferiormente $\eta$. Además llamemos por $\mathbb{Z}$-progresión aritmética a las ternas $(x, y, z)$ que sean solución. Con el fin de aprovechar la estructura de grupo de $\mathbb{Z}_N$, por ejemplo dada una función $f : \mathbb{Z}_{\mathrm{N}} \to \mathbb{R}$ tenemos que $\sum_{\xi \in \mathbb{Z}_{\mathrm{N}}} f(\xi) = \sum_{\xi \in \mathbb{Z}_{\mathrm{N}}} f(\xi + r)$ para cualquier $r \in \mathbb{Z}_{\mathrm{N}}$, así como también $\sum_{\xi \in \mathbb{Z}_{\mathrm{N}}} f(3\xi) = \sum_{\xi \in \mathbb{Z}_{\mathrm{N}}} f(\xi)$ cuando $N$ no es múltiplo de tres; intentamos primero acotar inferiormente el número de soluciones del problema:

$$\begin{cases} x + z \equiv 2y \mod N \\ x, y, z \in A \end{cases} \tag{3.4}$$

Sea $\eta_0$ el número de soluciones del problema anterior, y llamemos por $\mathbb{Z}_N$-progresión aritmética a las ternas $(x, y, z)$ que sean soluciones de tal problema. Entonces $\eta_0 \geq \eta$, puesto que toda solución de (3.3) es también solución de (3.4). Tenemos que

$$\begin{aligned} \eta_0 &= \sum_{\substack{x,y,z \in A \\ x+z \equiv 2y \bmod N}} 1 = \sum_{x \in A} \sum_{y \in A} \sum_{z \in A} \frac{1}{N} \sum_{k=0}^{N-1} \omega^{-(x+z-2y)k} \\ &= \frac{1}{N} \sum_{k=0}^{N-1} \left( \sum_{x} A(x) \omega^{-xk} \right) \left( \sum_{y} A(y) \omega^{-y(-2k)} \right) \left( \sum_{z} A(z) \omega^{-zk} \right) \\ &= \frac{1}{N} \sum_{k=0}^{N-1} \widehat{A}(k)^2 \widehat{A}(-2k). \end{aligned}$$

Así,

$$\eta_0 = \delta^3 N^2 + N^{-1} \sum_{k=1}^{N-1} \widehat{A}(k)^2 \widehat{A}(-2k).$$

Observemos que si $|\widehat{A}(k)| \leq \epsilon N$ para todo $k \neq 0$, entonces

$$\begin{aligned} \left| \sum_{k=1}^{N-1} \widehat{A}(k)^2 \widehat{A}(-2k) \right| &\leq \max_{k \neq 0} |\widehat{A}(-2k)| \sum_{k=1}^{N-1} |\widehat{A}(k)|^2 \\ &\leq \epsilon N \cdot N \sum_{k=0}^{N-1} |A(k)|^2 \\ &= \epsilon \delta N^3. \end{aligned}$$

Así, $\eta_0 \geq \delta^3 N^2 - \epsilon\delta N^2$. Si $\epsilon < \delta^2/2$, entonces $\eta_0 > \delta^3 N^2/2$, esto es, $A$ contiene al menos $\delta^3 N^2/2$ $\mathbb{Z}_N$-progresiones aritméticas de longitud tres, incluyendo las progresiones aritméticas triviales, aquellas con diferencia común cero. Sin embargo, no queremos solamente $\mathbb{Z}_N$-progresiones aritméticas, sino mas bien $\mathbb{Z}$-progresiones aritméticas.

**Observación 3.1.** *Si $(x, y, z) \in \{0, 1, \ldots, N-1\}^3$ es una $\mathbb{Z}_N$-progresión aritmética tal que $x, y \in [N/3, 2N/3)$, entonces $(x, y, z)$ es una $\mathbb{Z}$-progresión aritmética. En efecto,*

$$N/3 \leq y < 2N/3 \quad y \quad |z - y| = |y - x| < N/3,$$

*puesto que $-\dfrac{2N}{3} < z - y < \dfrac{2N}{3}$, $-\dfrac{N}{3} < y - x < \dfrac{N}{3}$ y $y - x \equiv z - y \mod N$. Por lo tanto*

$$N/3 \leq y < 2N/3 \quad y \quad -N/3 < z - y < N/3,$$

*y así, $0 < z < N$.*

**Observación 3.2.** *Si $x, y \in A \cap [N/3, 2N/3)$, entonces*

$$x + z \equiv 2y \mod N \text{ si, y solo si, } x + z = 2y.$$

*En efecto. Consideremos $x, y \in A \cap [N/3, 2N/3)$, y supongamos que $x + z \equiv 2y \mod N$. Entonces $|x - y| < \dfrac{N}{3}$, $|z - y| < \dfrac{2N}{3}$ y $N|x + z - 2y$. Como*

$$|x + z - 2y| \leq |x - y| + |z - y| < \frac{N}{3} + \frac{2N}{3} = N,$$

*necesariamente se debe tener que $x + z = 2y$. La implicancia recíproca es trivial.*

Modifiquemos un poco los argumentos anteriormente usados para acotar inferiormente $\eta_0$, a fin de acotar $\eta$. Sean $A$, $\delta$ y $\eta$ como en el enunciado del problema (3.3). Consideremos el conjunto $B = A \cap [N/3, 2N/3)$.

**Lema 3.2.** *Sean $N \in \mathbb{N}$ un número impar y $A \subseteq \mathbb{Z}_N$. Si $A$ es tal que $|\widehat{A}(k)| \leq \epsilon N$ para todo $k \in \mathbb{Z}_N - \{0\}$. Y además si $\epsilon < \dfrac{\delta^2}{8}$ y $|B| \geq \dfrac{\delta}{4}N$, entonces $\eta \geq \dfrac{\delta^3 N^2}{32}$.*

Obsrvemos que la condición requerida en las hipótesis del lema, $|\widehat{A}(k)| \leq \epsilon N$ para todo $k \in \mathbb{Z}_N - \{0\}$, se satisfará si $A \subseteq \mathbb{Z}_N$ es $\epsilon^4$-uniforme.

*Demostración.* Acotemos inferiormente $\eta$,

$$\eta = \sum_{\substack{x,y,z \in A \\ x+z=2y}} 1 \geq \sum_{x \in B} \sum_{y \in B} \sum_{\substack{z \in A \\ x+z=2y}} 1 = \sum_{x \in B} \sum_{y \in B} \sum_{\substack{z \in A \\ x+z \equiv 2y \text{ mód } N}} 1$$

$$
\begin{aligned}
&= \sum_{x\in B}\sum_{y\in B}\sum_{z\in A}\frac{1}{N}\sum_{k=0}^{N-1}\omega^{-(x+z-2y)k}\\
&= \frac{1}{N}\sum_{k=0}^{N-1}\left(\sum_{x=0}^{N-1}B(x)\omega^{-xk}\right)\left(\sum_{y=0}^{N-1}B(y)\omega^{-y(-2k)}\right)\left(\sum_{z=0}^{N-1}A(z)\omega^{-zk}\right)\\
&= \frac{1}{N}\sum_{k=0}^{N-1}\widehat{B}(k)\widehat{B}(-2k)\widehat{A}(k)\\
&= \frac{1}{N}|B|^2|A| + \frac{1}{N}\sum_{k=1}^{N-1}\widehat{B}(k)\widehat{B}(-2k)\widehat{A}(k),
\end{aligned}
$$

donde la segunda igualdad se da por la observación 3.2. Así,

$$\eta \geq \delta|B|^2 + \frac{1}{N}\sum_{k=1}^{N-1}\widehat{B}(k)\widehat{B}(-2k)\widehat{A}(k).$$

Por la desigualdad de Cauchy-Schwarz y la identidad de Plancherel

$$
\begin{aligned}
\left|\sum_{k=1}^{N-1}\widehat{B}(k)\widehat{B}(-2k)\widehat{A}(k)\right| &\leq \max_{k\neq 0}|\widehat{A}(k)|.\left|\sum_{k=0}^{N-1}\widehat{B}(k)\widehat{B}(-2k)\right|\\
&\leq \epsilon N\cdot\left(\sum_{k=0}^{N-1}|\widehat{B}(k)|^2\right)^{1/2}\left(\sum_{k=0}^{N-1}|\widehat{B}(-2k)|^2\right)^{1/2}\\
&\leq \epsilon N\sum_{k=0}^{N-1}|\widehat{B}(k)|^2\\
&= \epsilon N^2\sum_{k=0}^{N-1}|B(k)|^2\\
&= \epsilon N^2|B|,
\end{aligned}
$$

donde para que se dé la tercera desigualdad usamos el hecho de que $N$ es impar. Si $\epsilon < \dfrac{\delta^2}{8}$ y $|B| \geq \dfrac{\delta}{4}N$,

$$
\begin{aligned}
\eta &\geq \delta|B|^2 - \epsilon N|B| \geq \left(\delta|B| - \epsilon N\right)|B|\\
&\geq \left(\frac{\delta^2}{4}N - \frac{\delta^2}{8}N\right)|B| > \left(\frac{\delta^2 N}{8}\right)\frac{\delta N}{4} = \frac{\delta^3 N^2}{32}.
\end{aligned}
$$

□

El número de progresiones aritméticas triviales $(x, y, z)$, esto es, con $x = y = z$, es $|A| = \delta N$. Por tanto, el número de progresiones aritméticas no triviales, que denotaremos $\eta_1$, es tal que

$$\eta_1 > \frac{\delta^3 N^2}{32} - \delta N.$$

Si $N \geq (8/\delta)^2$, entonces $\eta_1 > 2\delta N - \delta N = \delta N > 0$. Esto es, si los coeficientes de Fourier $\widehat{A}(k)$ son suficientemente pequeños en norma para todo $k \neq 0$, y si $N$ es suficientemente grande tenemos que existe al menos una progresión aritmética no trivial de longitud tres constituida solo por elementos de $A$.

**Proposición 3.1.** *Sea $A \subseteq \{0, 1, \ldots, N-1\}$, tal que $|A| = \delta N$, donde $N$ es impar. Si $A$ no contiene ninguna $\mathbb{Z}$-progresión aritmética de longitud tres no trivial, entonces una de las siguientes afirmaciones es válida:*

*(i)* $N \leq (8/\delta)^2$.

*(ii) Existe una $\mathbb{Z}$-progresión aritmética $P$ de longitud $|P| \geq N/3$ tal que*

$$|A \cap P| \geq \left(\delta + \frac{\delta}{8}\right)|P|.$$

*(iii) Para cada $\epsilon \leq \delta^2/8$ se tiene que $\max\limits_{k \neq 0} |\widehat{A}(k)| > \epsilon N$.*

*Demostración.* Si (iii) no se cumple, entonces existe $\epsilon \leq \delta^2/8$ tal que $\max\limits_{k \neq 0} |\widehat{A}(k)| \leq \epsilon N$. Tenemos dos posibilidades, $|B| \geq \dfrac{\delta N}{4}$ o $|B| < \dfrac{\delta N}{4}$.

- Si $|B| \geq \dfrac{\delta N}{4}$, entonces del lema 3.2 y del hecho que $\epsilon \leq \dfrac{\delta^2}{8}$, tenemos que $N < (8/\delta)^2$, esto es, se cumple (i).
- Si $|B| < \dfrac{\delta N}{4}$, entonces $\max\big\{|A \cap [0, N/3)|, |A \cap [2N/3, N]|\big\} \geq \dfrac{3\delta N}{8} = \big(9\delta/8\big)\big(N/3\big)$, lo que concluye (ii) considerando $P = A \cap [0, N/3)$ o $P = A \cap [2N/3, N]$.

Esto termina la prueba. □

Ahora, dado $A \subseteq \{0, 1, \ldots, N-1\}$ con $|A| = \delta N$, nos planteamos la interrogante: ¿qué ocurre si $A \subseteq \{0, 1, \ldots, N-1\}$ no cumple que $\max\limits_{k \neq 0} |\widehat{A}(k)| \leq \epsilon N$ para $\epsilon$ suficientemente pequeño? Mostraremos que en este caso $A$ goza de la propiedad del *"incremento de densidad"*, esto es, veremos que existe alguna progresión aritmética $P$, incluida en $\{0, 1, \ldots, N-1\}$, tal que la densidad de $A \cap P$, respecto a $P$, aumenta; es decir, $|A \cap P| \geq (\delta + \epsilon')|P|$, donde $\epsilon' > 0$ depende de $\epsilon$.

La siguiente definición reúne algunas condiciones suficientes de modo que podamos hallar $\mathbb{Z}$-progresiones aritméticas apartir de $\mathbb{Z}_N$-progresiones aritméticas.

**Definición 3.3.** *Una $\mathbb{Z}_N$-progresión aritmética $P \subseteq \{0, 1, \ldots, N-1\}$* **no traslapa**, *si su cardinalidad $L$ y su diferencia común $d$ satisfacen $dL < N$.*

Si $P \subseteq \{0, 1, \ldots, N-1\}$ es una $\mathbb{Z}_N$-progresión aritmética que no traslapa, entonces $P$ se descompone en dos $\mathbb{Z}$-progresiones aritméticas disjuntas $P_1$ y $P_2$.

**Lema 3.3.** *Sean $\delta > 0$ y $\epsilon' > 0$. Sea $B'$ una $\mathbb{Z}_N$-progresión aritmética que no traslapa, tal que $\dfrac{|B' \cap A|}{|B'|} \geq \delta + \epsilon'$, entonces existe alguna $\mathbb{Z}$-progresión aritmética $B$ con $|B| \geq \dfrac{1}{2}\epsilon'|B'|$, tal que $\dfrac{|B \cap A|}{|B|} \geq \delta + \dfrac{\epsilon'}{2}$.*

*Demostración.* De las hipótesis dadas, tenemos que $\epsilon', \delta \in ]0, 1[$. Sea $B' = P_1 \cup P_2$, donde $P_1$ y $P_2$ son $\mathbb{Z}$-progresiones aritméticas disjuntas. Supongamos que $|P_1| \leq |P_2|$.

- Si $|P_1| \leq \dfrac{1}{2}\epsilon'|B'|$, entonces

$$\begin{aligned} |A \cap P_2| &= |A \cap B'| - |A \cap P_1| \\ &\geq |A \cap B'| - |P_1| \\ &\geq (\delta + \epsilon')|B'| - \frac{1}{2}\epsilon'|B'| \\ &= \left(\delta + \frac{1}{2}\epsilon'\right)|B'|. \end{aligned}$$

  En este caso, como $|P_1| + |P_2| = |B'|$ y $|P_1| \leq \dfrac{1}{2}\epsilon'|B'|$ y $0 < \epsilon' < 1$, tenemos que

$$|P_2| = |B'| - |P_1| \geq \epsilon'|B'| - \frac{1}{2}\epsilon'|B'| = \frac{1}{2}\epsilon'|B'|.$$

- Si $|P_1| > \dfrac{1}{2}\epsilon'|B'|$, entonces $|P_i| > \dfrac{1}{2}\epsilon'|B'|$ para $i = 1, 2$. Entonces

$$\begin{aligned} |A \cap P_1| + |A \cap P_2| = |A \cap B'| &\geq (\delta + \epsilon')|B'| \\ \Big[|A \cap P_1| - \delta|P_1|\Big] + \Big[|A \cap P_2| - \delta|P_2|\Big] &\geq \epsilon'|B'|. \end{aligned}$$

  Luego

$$|A \cap P_i| - \delta|P_i| \geq (\epsilon'/2)|B'|, \text{ para algún } i \in \{1, 2\}.$$

□

**Lema 3.4.** *Consideremos $\epsilon > 0$ y $A \subseteq \mathbb{Z}_N$. Si $|\widehat{A}(r)| \geq \epsilon N$ para algún $r \neq 0$, entonces existe alguna $\mathbb{Z}_N$-progresión aritmética $B'$ que no traslapa, de cardinalidad al menos $\sqrt{N}/4$ tal que*

$$|A \cap B'| \geq \left(\delta + \frac{\epsilon}{4}\right)|B'|.$$

Buscamos un incremento de densidad relativa, esto es, dado $A \subseteq \{0, 1, \ldots, N-1\}$ con $|A| = \delta N$, requerimos hallar un conjunto $B$, subprogresión aritmética de $\{0, 1, \ldots, N-1\}$, tal que

$$|A \cap (B + x)| \geq \left(\delta + \frac{\epsilon}{4}\right)|B|, \tag{3.5}$$

para algún $x \in \{0, 1, \ldots, N-1\}$, y algún $\epsilon > 0$. Pero la expresión anterior se puede desarrollar como

$$\begin{aligned}
\sum_y A(y)B(y-x) &\geq \delta|B| + (\epsilon/4)|B| \\
\sum_y A(y)B(y-x) &\geq \delta \sum_y B(y) + (\epsilon/4)|B| \\
\sum_y A(y)B(y-x) &\geq \delta \sum_y B(y-x) + (\epsilon/4)|B| \\
\sum_y \Big(A(y) - \delta\Big)B(y-x) &\geq (\epsilon/4)|B|. \qquad (3.6)
\end{aligned}$$

Esto, de algún modo, motiva a definir la función balanceada $f_A(z) = A(z) - \delta$, definición 3.1, con lo que (3.5) y (3.6) son equivalentes a que

$$\sum_y f_A(y)B(y-x) \geq (\epsilon/4)|B| \text{ para algún } x \in \{0, 1, \ldots, N\}. \tag{3.7}$$

De un modo más general, observemos que si se dan dos funciones $f, g : \mathbb{Z}_N \to \mathbb{C}$, entonces para que $\Big|\sum_y f(y)\overline{g(y-x)}\Big| \geq \beta$ para algún $x$, es suficiente que $\sum_x \Big|\sum_y f(y)\overline{g(y-x)}\Big| \geq \beta N$. Además, tenemos que

$$\begin{aligned}
\sum_x \Big|\sum_y f(y)\overline{g(y-x)}\Big| &= \sum_x \Big|\sum_y f(y)\overline{g(y-x)}\omega^{-kx}\Big| \\
&\geq \Big|\sum_x \sum_y f(y)\overline{g(y-x)}\omega^{-kx}\Big| \\
&= \Big|\sum_x \Big(\sum_y f(y)\overline{g(y-x)}\Big)\omega^{-kx}\Big| \\
&= \Big|\sum_x (f * g)(x)\omega^{-kx}\Big| \\
&= |(f*g)\widehat{\ }(k)| = |\widehat{f}(k)|\,|\widehat{g}(k)|, \qquad (3.8)
\end{aligned}$$

para todo $k \in \mathbb{Z}_N$. Así, para mostrar (3.7) es suficiente mostrar que $|\widehat{f_A}(k)|\,|\widehat{B}(k)| \geq (\epsilon/4)|B|N$ para algún $k \in \mathbb{Z}_{\mathrm{N}}$, pues en este caso tendremos que

$$\sum_x \Big|\sum_y f_A(y)B(y-x)\Big| \geq (\epsilon/4)|B|N, \tag{3.9}$$

y por lo tanto existirá algún $x \in \mathbb{Z}_N$ tal que

$$\left|\sum_y f_A(y)B(y-x)\right| \geq (\epsilon/4)|B|.$$

De lo contrario, para cada $x \in \mathbb{Z}_N$

$$\left|\sum_y f_A(y)B(y-x)\right| < (\epsilon/4)|B|,$$

y por lo tanto $\displaystyle\sum_x \left|\sum_y f_A(y)B(y-x)\right| < (\epsilon/4)|B|N$, lo que contradiría (3.9).

*Prueba del Lema 3.4.* Mostremos primero que para cualquier $1 \leq r \leq N-1$ existe alguna $\mathbb{Z}_N$-progresión aritmética $B$, que no traslapa y con cardinalidad de al menos $\sqrt{N}/4$, tal que $|\widehat{B}(r)| \geq \dfrac{1}{2}|B|$. Para ello, fijemos $r \neq 0$. Particionamos $[0, N-1]^2$ en una cantidad menor que $N$ de cuadrados iguales (por ejemplo $\left\lceil \sqrt{N}-1 \right\rceil^2$ cuadrados iguales). Se sigue, del principio de las casillas, considerando la colección de pares

$$\Big\{(0,0),(1,r),(2,2r),\ldots,(N-1,(N-1)r)\Big\},$$

o

$$\Big\{(N-1,0),(N-2,r),(N-3,2r),\ldots,(0,(N-1)r)\Big\},$$

que existen enteros $l$ y $k$, con $0 < l < k < N-1$, tales que

$$k-l \leq \sqrt{N} \quad \text{y} \quad r(k-l) \leq \sqrt{N} \bmod N.$$

Sea $d = k-l$. Definimos $B$ como la progresión siguiente de longitud $L = \left\lfloor \dfrac{\sqrt{N}}{\pi} \right\rfloor$ con diferencia común $d$,

$$B = \left\{-\frac{L-1}{2}d,\ldots,-2d,-d,0,d,2d,\ldots,\frac{L-1}{2}d\right\},$$

en caso que $L$ sea impar; o

$$B = \left\{-\frac{L-2}{2}d,\ldots,-2d,-d,0,d,2d,\ldots,\frac{L}{2}d\right\},$$

en caso que $L$ sea par. Tenemos que

$$\begin{aligned}
\left|\widehat{B}(r) - |B|\right| &\leq \left|\sum_x B(x)\left[\omega^{-xr}-1\right]\right| \\
&\leq \sum_{|l|\leq \frac{1}{2}|B|} \left|\omega^{-ldr}-1\right| \\
&< \sum_{|l|\leq \frac{1}{2}|B|} \left|-\frac{2\pi i}{N}ldr\right| \\
&\leq \frac{1}{2}|B|\left|\frac{2\pi}{N}\cdot\frac{|B|}{2}\cdot\sqrt{N}\right| \leq \frac{1}{2}|B|, \qquad (3.10)
\end{aligned}$$

donde la penúltima desigualdad se da por la desigualdad del valor medio, y la última se da porque $dr < \sqrt{N} \bmod N$ y $|B| \leq \dfrac{\sqrt{N}}{\pi}$. De (3.10) se tiene que

$$\left|\widehat{B}(r)\right| \geq \frac{1}{2}|B|. \tag{3.11}$$

En particular, para el $r \neq 0$ del enunciado del lema, existe una $\mathbb{Z}_N$-progresión aritmética $B$, con $|B| \geq \sqrt{N}/4$ (puesto que $|B| = \lfloor \sqrt{N}/\pi \rfloor \geq \lfloor \sqrt{N}/4 \rfloor$), y tal que $\left|\widehat{B}(r)\right| \geq \dfrac{1}{2}|B|$. La progresión $B$ no traslapa puesto que $d \leq \sqrt{N}$ y $|B| \leq \sqrt{N}/\pi$ y por tanto $d|B| \leq \sqrt{N} \cdot \sqrt{N}/\pi < N$. Como observamos antes, de (3.7),

$$\left|A \cap (B + x)\right| \geq \left(\delta + \frac{1}{4}\epsilon\right)|B| \text{ si, y solo si, } \sum_y f_A(y)B(y - x) \geq \frac{1}{4}\epsilon|B|. \tag{3.12}$$

Consideremos $G(x) = \sum_y f_A(y)B(y - x)$, esto es $G = f_A * B$, de (3.2), (3.8), (3.11) y de la hipótesis $|\widehat{A}(r)| \geq \epsilon N$, tenemos que

$$\sum_x |G(x)| \geq \left|\widehat{G}(r)\right| = \left|\widehat{f_A}(r)\right| \left|\widehat{B}(r)\right| \geq \epsilon N \cdot \frac{1}{2}|B| = \frac{1}{2}\epsilon N|B|.$$

Además, de (3.1) tenemos que

$$\sum_x G(x) = \sum_x \sum_y f_A(y)B(y - x) = \sum_y f_A(y) \sum_x B(y - x) = |B| \sum_y f_A(y) = 0,$$

y por lo tanto

$$\sum_x \Big(|G(x)| + G(x)\Big) \geq \frac{1}{2}\epsilon N|B|,$$

lo cual implica que para algún $x \in \{0, 1, \ldots, N - 1\}$

$$|G(x)| + G(x) \geq \frac{1}{2}\epsilon|B|,$$

y por lo tanto $G(x) \geq 0$ y $G(x) \geq \dfrac{1}{4}\epsilon|B|$. De (3.12) tenemos que

$$|A \cap (B + x)| \geq \left(\delta + \frac{1}{4}\epsilon\right)|B| = \left(\delta + \frac{1}{4}\epsilon\right)|B + x|,$$

considerando $B' = B + x$ tenemos probado el lema. □

**Proposición 3.2.** *Sean $\epsilon > 0$ y $A \subseteq \{1, 2, \ldots, N\}$, y supongamos que $\left|\widehat{A}(r)\right| \geq \epsilon N$ para algún $r \neq 0$, entonces existe una $\mathbb{Z}$-progresión aritmética $P \subseteq \{0, 1, \ldots, N - 1\}$, de cardinalidad al menos $\dfrac{1}{32}\epsilon\sqrt{N}$, tal que*

$$|A \cap P| \geq \left(\delta + \frac{1}{8}\epsilon\right)|P|.$$

*Demostración.* Del lema 3.4, existe una $\mathbb{Z}_N$-progresión aritmética $B$ que no traslapa, de cardinalidad al menos $\sqrt{N}/4$ tal que

$$|A \cap B| \geq \Big(\delta + \frac{1}{4}\epsilon\Big)|B|.$$

Del lema 3.3, existe una subprogresión aritmética de $B$, que denominamos $P$, tal que

$$|A \cap P| \geq \Big(\delta + \frac{1}{8}\epsilon\Big)|P|,$$

donde $|P| \geq \frac{1}{2} \cdot \frac{1}{4}\epsilon \cdot |B| \geq \frac{1}{32}\epsilon\sqrt{N}$. □

**Corolario 3.1.** *Sean $\delta > 0$ y $A$ un subconjunto de $\{1, 2, \ldots, N\}$, con $|A| = \delta N$, y $N > 64\delta^{-2}$ un número impar. Entonces $A$ contiene una $\mathbb{Z}$-progresión aritmética (no trivial) de longitud tres, o existe una $\mathbb{Z}$-progresión aritmética $P$, de longitud $|P| \geq \frac{1}{256}\delta^2\sqrt{N}$, tal que*

$$|A \cap P| \geq \Big(\delta + \frac{1}{64}\delta^2\Big)|P|.$$

*Demostración.* Supongamos que $A$ no contiene ninguna $\mathbb{Z}$-progresión aritmética de longitud tres. Entonces de la proposición 3.1 se cumple alguna de las siguientes posibilidades:

(i) $N \leq (8/\delta)^2$.

(ii) Existe una $\mathbb{Z}$-progresión aritmética $P$ de longitud $|P| \geq N/3$ tal que

$$|A \cap P| \geq \left(\delta + \frac{\delta}{8}\right)|P|.$$

(iii) Para cada $\epsilon \leq \delta^2/8$ se tiene que $\max\limits_{k \neq 0} |\widehat{A}(k)| > \epsilon N$.

La afirmación (i) no se cumple dadas las hipótesis de la proposición. Si se cumple (ii),

$$|A \cap P| \geq \left(\delta + \frac{\delta}{8}\right)|P| \geq \left(\delta + \frac{\delta^2}{64}\right)|P|,$$

con $|P| \geq N/3 \geq \frac{1}{256}\delta^2\sqrt{N}$. Si se cumple (iii), entonces $\max\limits_{k \neq 0} |\widehat{A}(k)| > \epsilon N$ particularmente para $\epsilon = \delta^2/8$; por lo tanto existe $r \neq 0$ tal que $|\widehat{A}(r)| \geq \epsilon N$. Por la proposición 3.2, existe una $\mathbb{Z}$-progresión aritmética $P$, de longitud al menos $\frac{1}{32}\epsilon\sqrt{N} = \frac{1}{256}\delta^2\sqrt{N}$, tal que

$$|A \cap P| \geq \left(\delta + \frac{\epsilon}{8}\right)|P| = \left(\delta + \frac{\delta^2}{64}\right)|P|.$$

□

Esta última proposición es la pieza fundamental que permite hacer un argumento iterativo para mostrar el teorema de Roth. Esencialmente la prueba que daremos es como sigue: inicialmente suponemos por contradicción que dado el conjunto $A$, $A \subseteq \{1, 2, \ldots, N\}$ con $|A| = \delta N$ ($\delta > 0$) y $N$ suficientemente grande, $A$ no contiene ninguna progresión aritmética no trivial de longitud tres. Entonces del corolario 3.1, tenemos que necesariamente existe alguna subprogresión aritmética $A_1 \subseteq \{1, 2, \ldots, N_1\}$ con densidad $\delta_1$, mayor que $\delta$, que tampoco contiene progresiones aritméticas no triviales de longitud tres. Así, nuevamente por el corolario 3.1, existe alguna subprogresión $A_2 \subseteq \{1, 2, \ldots, N_2\}$ con densidad $\delta_2$, con $\delta_2 > \delta_1 > \delta$. Continuando con este proceso iterativo, en algún momento encontraremos una subprogresión aritmética $P$ en $\{1, 2, \ldots, R\}$ con densidad mayor que 1, lo cual será la contradicción requerida.

**Teorema 3.2** (de Roth)**.** *Existe un número real positivo $C$, tal que para todo $\delta_0 \in ]0, 1[$, si $N \geq \exp\exp(C\delta_0^{-1})$ y si $A \subseteq \{1, 2, \ldots, N\}$ tal que $|A| \geq \delta_0 N$, entonces $A$ contiene al menos una progresión aritmética no trivial, cuya diferencia común es diferente de cero, de longitud tres.*

*Demostración.* Puesto que realizaremos un proceso iterativo y pasaremos de una progresión aritmética a una subprogresión más pequeña en cada iteración, no podemos simplemenmte suponer que $N$ es un primo suficientemente grande. Así que inicialmente abordaremos esta dificultad técnica. Sea $N_0 \in \mathbb{Z}^+$ y sea $A_0 \subseteq \{1, 2, \ldots, N_0\}$ tal que $|A_0| \geq \delta_0 N_0$. Por el postulado de Bertrand, existe $p$ primo entre $N_0/3$ y $2N_0/3$; hacemos $q = N_0 - p$. Si

$$\big|A_0 \cap \{1, 2, \ldots, p\}\big| \leq \delta_0(1 - \delta_0/160)p, \tag{3.13}$$

entonces

$$\big|A_0 \cap \{p+1, \ldots, N_0\}\big| \geq \delta_0(1 + \delta_0/320)\big|\{p+1, \ldots, N_0\}\big|. \tag{3.14}$$

En efecto

$$\begin{aligned}
\big|A_0 \cap \{p+1, \ldots, N_0\}\big| &= \big|A_0 \cap \{1, 2, \ldots, N_0\}\big| - \big|A_0 \cap \{1, 2, \ldots, p\}\big| \\
&\geq \delta_0 N_0 - \delta_0(1 - \delta_0/160)p \\
&= \delta_0\big(N_0 - (1 - \delta_0/160)p\big) = \delta_0(q + \delta_0 p/160) \\
&= \delta_0\left(1 + \frac{\delta_0 p}{160q}\right)q \\
&\geq \delta_0\left(1 + \frac{\delta_0}{320}\right)\big|\{p+1, \ldots, N_0\}\big|,
\end{aligned}$$

donde la última desigualdad se da porque $2N_0/3 \geq p \geq N_0/3$ y $q = N_0 - p \leq 2N_0/3$. Este caso será el **caso 0**.

Si no se cumple (3.13), tenemos que

$$\big|A_0 \cap \{1, 2, \ldots, p\}\big| > \delta_0(1 - \delta_0/160)p.$$

Hacemos $N = p$ (así, $N$ es primo), $A = A_0 \cap \{1, 2, \ldots, p\}$ y $\delta = \delta_0(1 - \delta_0/160)$. Con esta nueva notación tenemos que $|A| > \delta N$. Consideremos $B = A \cap [N/3, 2N/3)$. Si se cumple que

$$|B| \leq \delta N/5, \tag{3.15}$$

entonces

$$\big|A \cap [0, N/3)\big| \geq 2\delta N/5 = (6\delta/5)(N/3), \tag{3.16}$$

o

$$\big|A \cap [2N/3, N]\big| \geq 2\delta N/5 = (6\delta/5)(N/3). \tag{3.17}$$

En efecto,

$$\begin{aligned}\big|A \cap [0, N/3)\big| + \big|A \cap [2N/3, N]\big| &= |A| - \big|A \cap [N/3, 2N/3)\big| \\ &> \delta N - \delta N/5 = 4\delta N/5.\end{aligned}$$

Si $\big|A \cap [0, N/3)\big| < 2\delta N/5$ y $\big|A \cap [2N/3, N]\big| < 2\delta N/5$, entonces $4\delta N/5 > 4\delta N/5$, lo cual es una contradición. Este caso se denominará el **caso 1**.

Sea $\alpha = \delta^2/10$ y supongamos que

$$|\widehat{A}(r)| > \alpha N \text{ para algún } r \neq 0. \tag{3.18}$$

En este caso, por la proposición 3.2 existe una progresión aritmética $P \subseteq \{1, 2, \ldots, N\}$ de cardinalidad al menos $\dfrac{\alpha}{32}\sqrt{N}$ tal que

$$|A \cap P| \geq \left(\delta + \frac{\alpha}{8}\right)|P| = \left(\delta + \frac{\delta^2}{80}\right)|P|. \tag{3.19}$$

Esta situación se llamará el **caso 2**.

Si la desigualdad (3.18) no se verifica, entonces

$$\big|\widehat{A}(r)\big| \leq \alpha N \text{ para cada } r \neq 0,$$

lo cual dice que $A$ satisface la condición (iv) del lema 3.1, por tanto, $A$ es $\alpha^4$-uniforme. Así, el número de ternas $(x, y, z) \in A \times B^2$ tal que $x + z = 2y$ es mayor que $\delta|B|^2 - \alpha|B|N$. En efecto,

$$\begin{aligned}
\sum_{\substack{x\in A, y\in B, z\in B,\\ x+z=2y}} 1 &= N^{-1} \sum_{x\in A}\sum_{y\in B}\sum_{z\in B}\sum_{r} \omega^{r(2y-x-z)} \\
&= N^{-1}\sum_{r} \widehat{A}(r)\widehat{B}(-2r)\widehat{B}(r) = |N^{-1}\sum_{r} \widehat{A}(r)\widehat{B}(-2r)\widehat{B}(r)| \\
&\geq |N^{-1}\widehat{A}(0)\widehat{B}(0)\widehat{B}(0)| - |N^{-1}\sum_{r\neq 0} \widehat{A}(r)\widehat{B}(-2r)\widehat{B}(r)| \\
&\geq N^{-1}|A||B|^2 - N^{-1}\sum_{r\neq 0} |\widehat{A}(r)|\,|\widehat{B}(-2r)|\,|\widehat{B}(r)| \\
&\geq N^{-1}|A||B|^2 - N^{-1}\max_{r\neq 0} |\widehat{A}(r)| \sum_{r\neq 0} |\widehat{B}(-2r)|\,|\widehat{B}(r)| \\
&\geq N^{-1}\cdot \delta N\cdot |B|^2 - N^{-1}\max_{r\neq 0} |\widehat{A}(r)| \left(\sum_{r\neq 0} |\widehat{B}(-2r)|^2\right)^{1/2} \left(\sum_{r\neq 0} |\widehat{B}(r)|^2\right)^{1/2} \\
&\geq N^{-1}\cdot \delta N\cdot |B|^2 - N^{-1}\cdot \alpha N \left(\sum_{r\in \mathbb{Z}_{\mathrm{N}}} |\widehat{B}(-2r)|^2\right)^{1/2} \left(\sum_{r\in \mathbb{Z}_{\mathrm{N}}} |\widehat{B}(r)|^2\right)^{1/2} \qquad (3.20) \\
&\geq \delta|B|^2 - N^{-1}\cdot \alpha N\cdot \left(\sum_{r} |\widehat{B}(r)|^2\right) \\
&\geq \delta|B|^2 - \alpha|B|N,
\end{aligned}$$

donde la última desigualdad se da por la identidad de Plancherel,

$$\sum_{r\neq 0} |\widehat{B}(r)|^2 \leq \sum_{r} |\widehat{B}(r)|^2 = N\sum_{r} |B(r)|^2 \leq N|B|.$$

Cabe resaltar que en este punto es importante el hecho de que $N$ sea primo (impar) como se observa en (3.20) puesto que cuando $k$ varía sobre todo $\mathbb{Z}_{\mathrm{N}}$ entonces $2k$ también abarca todo $\mathbb{Z}_{\mathrm{N}}$.

Si la desigualdad (3.15) no se cumple, entonces $|B| > \delta N/5$, y por lo tanto tenemos que

$$\sum_{\substack{x\in A, y\in B, z\in B,\\ x+z=2y}} 1 \geq \delta|B|^2 - \alpha|B|N > \delta^3 N^2/50.$$

En efecto, tenemos que

$$\delta|B|^2 - \alpha|B|N = \delta|B|\Big(|B| - \frac{\delta}{10}N\Big),$$

y puesto que $|B| > \frac{\delta N}{5}$ tenemos que $|B| - \frac{\delta N}{10} > \frac{\delta N}{10}$. Por tanto $\delta|B|\Big(|B| - \frac{\delta N}{10}\Big) > \frac{\delta^3 N^2}{50}$. Así, como $|B| > \delta N/5$, tenemos que el número de ternas $(x, y, z) \in A \times B^2$ en progresión aritmética es mayor que $\delta^3 N^2/50$. Puesto que $B$ está en el tercio medio de $A$, $B = A \cap [N/3, 2N/3)$, estas ternas están en progresión aritmética en $\{1, 2, \ldots, N\}$ (y no sólo en $\mathbb{Z}_N$). Y puesto que existen $N$ progresiones aritméticas degeneradas (progresiones con razón cero), podemos concluir que $A$ contiene al menos una progresión aritmética (no trivial) de longitud tres si $N \geq 50\delta^{-3}$.

En efecto, si $N \geq 50\delta^{-3}$ entonces $\delta^3 N^2/50 - N \geq 0$. Así

$$\text{número de 3-progresiones no degeneradas} > \delta^3 N^2/50 - N \geq 0.$$

Esta situación se denominará el **caso 3**.

Resumiendo, si el caso 3 vale y $N \geq 50\delta^{-3}$, entonces $A$ contiene al menos una progresión aritmética de longitud tres. En el caso 2 podemos hallar, de (3.19), una subprogresión $P$ de $\{1, 2, \ldots, N\}$ de cardinalidad al menos $\frac{\alpha}{32}\sqrt{N}$, tal que $|A \cap P| \geq \delta(1 + \delta/80)|P|$. Puesto que $\{1, 2, \ldots, N\} \subseteq \{1, 2, \ldots, N_0\}$, $A = A_0 \cap \{1, 2, \ldots, N\}$ y $\delta\Big(1 + \frac{\delta}{80}\Big) \geq \delta_0\Big(1 + \frac{\delta_0}{320}\Big)$ (pues $\delta = \delta_0(1 - \delta_0/160)$), podemos concluir que en el caso 2 existe una subprogresión $P$ de $\{1, 2, \ldots, N_0\}$ con cardinalidad de al menos $\frac{\alpha}{32}\sqrt{\frac{N_0}{3}} \geq \frac{\alpha}{64}\sqrt{N_0} = \frac{\delta^2}{640}\sqrt{N_0} > \frac{\delta_0^2}{649}\sqrt{N_0}$, puesto que $\delta > \frac{159\delta_0}{160}$, tal que

$$|A_0 \cap P| \geq \delta_0\Big(1 + \frac{\delta_0}{320}\Big)|P|.$$

En el caso 1 consideramos $P = [0, N/3)$ o $P = [2N/3, N]$ de modo que por (3.16) o (3.17) tenemos

$$|A_0 \cap P| = |A \cap P| \geq \Big(\delta + \frac{\delta}{5}\Big)|P| \geq \Big(\delta + \frac{\delta^2}{80}\Big)|P| \geq \Big(\delta_0 + \frac{\delta_0^2}{320}\Big)|P|,$$

donde la primera igualdad se da puesto que $A \cap P = \Big(A_0 \cap \{1, 2, \ldots, N\}\Big) \cap P = A_0 \cap P$.

En el caso 0 consideramos $P = \{p + 1, \ldots, N_0\}$ y tenemos de (3.14) que

$$|A_0 \cap P| \geq \delta_0\Big(1 + \frac{\delta_0}{320}\Big)|P|.$$

Dado que $\delta > 159\delta_0/160$ y $N \geq N_0/3$ tenemos que $N_0 \geq 150(159\delta_0/160)^{-3}$ implica $N \geq 50\delta^{-3}$. Así, dados $\delta_0 > 0$ y $A \subseteq \{1, 2, \ldots, N_0\}$ tales que $|A| \geq \delta_0 N_0$ y $N_0 \geq 153\delta_0^{-3}$, tenemos, en cualquiera de los casos, dos posibilidades:

(a) Existe una progresión aritmética no trivial de longitud tres.

(b) Existe una subprogresión $P$ de $\{1, 2, \ldots, N_0\}$ con cardinalidad de al menos $\frac{\delta_0^2}{649}\sqrt{N_0}$ tal que

$$|A_0 \cap P| \geq \delta_0\left(1 + \frac{\delta_0}{320}\right)|P|.$$

Supongamos, por contradicción que $A_0$ no contiene ninguna progresión aritmética no trivial de longitud tres y que $N_0$ es suficientemente grande, $N_0 \geq 153\delta_0^{-3}$. Entonces tenemos que, por (b), existe una subprogresión aritmética

$$P_1 = \{a_1 + d_1, a_1 + 2d_1, a_1 + 3d_1, \ldots, a_1 + N_1 d_1\} \subseteq \{1, 2, \ldots, N_0\},$$

con $N_1 = |P_1| \geq \frac{\alpha}{64}\sqrt{N_0} \geq \frac{\delta_0^2}{649}\sqrt{N_0}$, y tal que

$$|A_0 \cap P_1| \geq \delta_0\left(1 + \frac{\delta_0}{320}\right)|P_1|.$$

Identifiquemos $P_1$ con $\{1, 2, \ldots, N_1\}$, bajo la correspondencia biunívoca

$$\psi_1 : P_1 \to \{1, 2, \ldots, N_1\} \text{ definida por } a_1 + id_1 \mapsto i,$$

además identifiquemos $A_0 \cap P_1$ con $A_1$ bajo la misma correspondencia, $\psi_1(A_0 \cap P_1) = A_1$. Como tal correspondencia es biyectiva, tenemos que

$$A_1 \subseteq \{1, 2, \ldots, N_1\} \text{ es tal que } |A_1| = \delta_1 N_1,$$

donde $\delta_1 \geq \delta_0 + \frac{\delta_0^2}{320}$ y $N_1 \geq \frac{\delta_0^2}{649}N_0^{1/2}$.
Observemos que si $A_1$ contiene una progresión aritmética de longitud tres $(x, x + d, x + 2d)$ no trivial, entonces $A_0 \cap P_1$, y por tanto $A_0$, contendrá también una progresión aritmética $\big(a_1 + xd_1, a_1 + (x + d)d_1, a_1 + (x + 2d)d_1\big)$ no trivial, lo cual no puede ocurrir por hipótesis. Nuevamente, como $N_0$ será elegido suficientemente grande tal que $N_1 \geq 153\delta_0^{-3}$, y por lo tanto $N_1 \geq 153\delta_1^{-3}$, y puesto que $A_1$ no contiene ninguna progresión aritmética no trivial, de (b) tenemos que existe una subprogresión aritmética

$$P_2 = \{a_2 + d_2, a_2 + 2d_2, a_2 + 3d_2, \ldots, a_2 + N_2 d_2\} \subseteq \{1, 2, \ldots, N_1\},$$

con $N_2 = |P_2| \geq \frac{\delta_1^2}{649}N_1^{1/2}$, y tal que

$$|A_1 \cap P_2| \geq \delta_1\left(1 + \frac{\delta_1}{320}\right)|P_2|.$$

Como antes, identificamos $P_2$ con $\{1, 2, \ldots, N_2\}$ bajo la correspondencia biyectiva

$$\psi_2 : P_2 \to \{1, 2, \ldots, N_2\} \text{ definida por } a_2 + id_2 \mapsto i,$$

y además identificando $A_1 \cap P_2$ con $A_2$ bajo la misma correspondencia, $\psi_2(A_1 \cap P_2) = A_2$, obtenemos que

$$A_2 \subseteq \{1, 2, \ldots, N_2\} \text{ es tal que } |A_2| = \delta_2 N_2,$$

donde

$$\delta_2 \geq \delta_1 + \frac{\delta_1^2}{320} \geq \left(\delta_0 + \frac{\delta_0^2}{320}\right) + \frac{\delta_0^2}{320} = \delta_0 + 2\left(\frac{\delta_0^2}{320}\right),$$

y además $N_2 \geq \dfrac{\delta_1^2}{649} N_1^{1/2}$.

Si iteramos este argumento $k$ veces, con $k \geq 1$, obtenemos algún conjunto $A_k$ incluido en $\{1, 2, \ldots, N_k\}$ con densidad $\delta_k$, esto es $|A_k| = \delta_k N_k$, tal que

$$\delta_k \geq \delta_0 + k\left(\frac{\delta_0^2}{320}\right).$$

En efecto, si tenemos $A_r \subseteq \{1, 2, \ldots, N_r\}$, $(r \in \mathbb{N})$ con densidad $\delta_r$, esto es $|A_r| = \delta_r N_r$, tal que $N_r \geq 153\delta_0^{-3}$ y $\delta_r \geq \delta_0 + r\left(\dfrac{\delta_0^2}{320}\right)$, entonces debido a (b) y puesto que $A_0$, y por lo tanto $A_r$, no contiene progresiones aritméticas no triviales de longitud tres, existe una subprogresión aritmética

$$P_{r+1} = \{a_{r+1} + d_{r+1}, a_{r+1} + 2d_{r+1}, a_{r+1} + 3d_{r+1}, \ldots, a_{r+1} + N_{r+1}d_{r+1}\} \subseteq \{1, 2, \ldots, N_r\},$$

tal que

$$N_{r+1} = |P_{r+1}| \geq \frac{\delta_r^2}{649} N_r^{1/2} \text{ y } \quad |A_r \cap P_{r+1}| \geq \left(\delta_r + \frac{\delta_r^2}{320}\right)|P_{r+1}|.$$

Así, considerando la biyección

$$\psi_{r+1} : P_{r+1} \to \{1, 2, \ldots, N_{r+1}\} \text{ definida por } a_{r+1} + id_{r+1} \mapsto i,$$

y siendo

$$A_{r+1} = \psi_{r+1}(A_r \cap P_{r+1}) = \Big\{i \in \{1, 2, \ldots, N_r\} : \ a_{r+1} + id_{r+1} \in A_r\Big\},$$

obtenemos $A_{r+1} \subseteq \{1, 2, \ldots, N_{r+1}\}$ tal que

$$N_{r+1} \geq \frac{\delta_r^2}{649} N_r^{1/2} \text{ y } \quad |A_{r+1}| = \delta_{r+1} N_r,$$

donde

$$\delta_{r+1} \geq \delta_r + \frac{\delta_r^2}{320} \geq \left(\delta_0 + r\frac{\delta_0^2}{320}\right) + \frac{\delta_0^2}{320} = \delta_0 + (r+1)\frac{\delta_0^2}{320}.$$

Ahora, si iteramos el argumento anterior al menos $k$ veces, con $k \geq \dfrac{320}{\delta_0}$, obtenemos algún conjunto $A_k \subseteq \{1, 2, \ldots, N_k\}$, con densidad $\delta_k$, esto es $|A_k| = \delta_k N_k$, tal que $\delta_k \geq \delta_0 + k\dfrac{\delta_0^2}{320}$, y así $\delta_k \geq 2\delta_0$. Análogamente, después de al menos $\dfrac{320}{2\delta_0}$ iteraciones adicionales la densidad aumenta de $2\delta_0$ a $4\delta_0$; después de al menos $\dfrac{320}{4\delta_0}$ iteraciones adicionales la densidad se incrementa de $4\delta_0$ a $8\delta_0$. Prosiguiendo de esta manera, una densidad de al menos $2^l\delta_0$ es alcanzada luego de por lo menos

$$\frac{320}{\delta_0} + \frac{320}{2\delta_0} + \cdots + \frac{320}{2^{l-1}\delta_0} = \frac{320}{\delta_0}\left(1 + \frac{1}{2} + \cdots + \frac{1}{2^{l-1}}\right)$$

iteraciones. Puesto que la función $l \mapsto 2^l\delta_0$ es tal que $\lim\limits_{l\to+\infty} 2^l\delta_0 = +\infty$, la densidad excederá 1, para $l$ suficientemente grande, específicamente si $l = \left\lceil \dfrac{\log(1/\delta_0)}{\log 2} \right\rceil + 1$ entonces $2^{l-1}\delta_0 \geq 1$, esto es $2^l\delta_0 > 1$. De este modo, en al menos $\dfrac{320}{\delta_0}\left(1 + \dfrac{1}{2} + \dfrac{1}{4} + \dfrac{1}{8} + \cdots + \dfrac{1}{2^{\lceil \frac{\log(1/\delta_0)}{\log 2} \rceil}}\right)$ y en no más que $\dfrac{640}{\delta_0}\left(1 + \dfrac{1}{2} + \dfrac{1}{4} + \dfrac{1}{8} + \cdots\right) = \dfrac{1280}{\delta_0}$ iteraciones la densidad será mayor que 1.

Es de hecho imposible que cualquier subconjunto $A_r$ de $\{1, 2, \ldots, N_r\}$ tenga densidad mayor que 1, y por lo tanto habremos llegado a una contradicción en $k_1 = \left\lfloor \dfrac{1280}{\delta_0} \right\rfloor$ iteraciones si la inclusión $A_{k_1} \subseteq \{1, 2, \ldots, N_{k_1}\}$ tiene sentido, es decir, si $N_{k_1} \geq 1$, y si también es cierta la condición

$$N_k \geq 153\delta_k^{-3} \text{ para cada } k \in \{1, 2, \ldots, k_1 - 1\},$$

de modo que podamos pasar de una iteración a la siguiente. Dado que $N_1 \geq N_2 \geq \cdots \geq N_{k_1}$, y puesto que $153\delta_0^{-3} \geq 153\delta_k^{-3}$ para todo $k \in \{1, 2, \ldots, k_1 - 1\}$, es suficiente que $N_{k_1-1} \geq 153\delta_0^{-3}$. Después de la primera iteración el tamaño de la subprogresión $P_1$, $N_1$, es al menos

$$\frac{\alpha}{64}N_0^{1/2} = \frac{\delta^2}{640}N_0^{1/2} > \frac{0,9875\,\delta_0^{1/2}}{640}N_0^{1/2} > \frac{\delta_0^2}{649}N_0^{1/2},$$

puesto que $\delta = \delta_0\left(1 - \dfrac{\delta_0}{160}\right) > \dfrac{159}{160}\delta_0$ y así $\delta^2 > 0,9875\,\delta_0^2$. Luego de la segunda iteración el tamaño de la subprogresión $P_2$, $N_2$, es al menos

$$\frac{\alpha_1}{64}N_1^{1/2} > \frac{\delta_1^2}{649}N_1^{1/2} \geq \frac{\delta_0^2}{649}\left(\frac{\delta_0^2}{649}N_0^{1/2}\right)^{1/2} = \frac{\delta_0^{2+1}N_0^{1/4}}{649^{1+\frac{1}{2}}},$$

puesto que $\delta_1 \geq \delta_0$ y donde $\alpha_1 = \dfrac{1}{10}\left[\delta_1\left(1 - \dfrac{\delta_1}{160}\right)\right]^2$. Después de la tercera iteración el tamaño de la subprogresión $P_3$, $N_3$, es al menos

$$\frac{\alpha_2}{64}N_2^{1/2} > \frac{\delta_2^2}{649}N_2^{1/2} \geq \frac{\delta_0^2}{649}\left(\frac{\delta_0^{2+1}}{649^{1+\frac{1}{2}}}N_0^{1/4}\right)^{1/2} = \frac{\delta_0^{2+1+\frac{1}{2}}N_0^{1/8}}{649^{1+\frac{1}{2}+\frac{1}{4}}},$$

puesto que $\delta_2 \geq \delta_0$ y donde $\alpha_2 = \dfrac{1}{10}\left[\delta_2\left(1-\dfrac{\delta_2}{160}\right)\right]^2$. En general, luego de la $k$-ésima iteración la longitud de la subprogresión $P_k$, $N_k$, es al menos

$$\frac{\delta_0^{2+1+\frac{1}{2}+\frac{1}{4}+\cdots+\frac{1}{2^{k-2}}}N_0^{1/2^k}}{649^{1+\frac{1}{2}+\frac{1}{4}+\cdots+\frac{1}{2^{k-1}}}} \geq \frac{\delta_0^4 N_0^{1/2^k}}{649^2},$$

donde la última desigualdad se da porque $\delta_0 \in ]0,1[$. Así, es suficiente mostrar que si $k_1 = \left\lfloor \dfrac{1280}{\delta_0} \right\rfloor$ entonces

$$\frac{\delta_0^4 N_0^{1/2^{k_1}}}{649^2} \geq 1 \quad \text{y} \quad \frac{\delta_0^4 N_0^{1/2^{k_1-1}}}{649^2} \geq 153\delta_0^{-3}. \tag{3.21}$$

La primera condición es para que $N_{k_1} \geq 1$, y por ende para que $N_k \geq 1$ para todo $1 \leq k \leq k_1$, mientras que la segunda condición es para que $N_{k_1-1} \geq 153\delta_0^{-3}$, y por consiguiente para que $N_k \geq 153\delta_0^{-3} \geq 153\delta_k^{-3}$ para cada $1 \leq k \leq k_1 - 1$.

Es suficiente mostrar que $\dfrac{\delta_0^4 N_0^{1/2^{k_1}}}{649^2} \geq \delta_0^{-1}$, puesto que en este caso se cumple (3.21). En efecto, si $\dfrac{\delta_0^4 N_0^{1/2^{k_1}}}{649^2} \geq \delta_0^{-1}$ entonces $\dfrac{\delta_0^4 N_0^{1/2^{k_1}}}{649^2} \geq 1$, pues $\delta_0^{-1} > 1$, y $\left(\dfrac{\delta_0^4 N_0^{1/2^{k_1}}}{649^2}\right)^2 \geq \delta_0^{-2}$. Por lo tanto

$$\frac{\delta_0^4 N_0^{1/2^{k_1-1}}}{649^2} \geq 649^2\delta_0^{-3}\cdot\delta_0^{-3} \geq 153\delta_0^{-3}.$$

Además, $\dfrac{\delta_0^4 N_0^{1/2^{k_1}}}{649^2} \geq \delta_0^{-1}$ es equivalente a

$$\begin{aligned} N_0^{1/2^{k_1}} &\geq 649^2\delta_0^{-5} \\ \log N_0 &\geq 2^{k_1}\Big[\log(649^2) + 5\log(\delta_0^{-1})\Big], \end{aligned}$$

y puesto que $k_1 \leq \dfrac{1280}{\delta_0}$, es suficiente probar que

$$\log N_0 \geq 2^{1280\,\delta_0^{-1}}\Big[\log(649^2) + 5\log(\delta_0^{-1})\Big]. \tag{3.22}$$

Por otro lado, como $\log(649^2) + 5\log(\delta_0^{-1}) \leq 2^{4\delta_0^{-1}}$, tenemos que para que se verifique (3.22) basta con que se cumpla que $\log N_0 \geq 2^{1284\,\delta_0^{-1}}$, esto es,

$$N_0 \geq \exp\exp\left(1284(\log 2)\delta_0^{-1}\right).$$

Elegimos $C = 1284(\log 2)$. □

## 3.2. Revisión del teorema de Roth, usando la $U^2$-norma de Gowers

Sea $f : \mathbb{Z}_N \to \mathbb{C}$ una función compleja. En este caso, similarmente al caso real, definimos la $U^2$-*norma de Gowers de* $f$ por

$$||f||_{U^2} = \left(\frac{1}{N^3}\sum_{a,b,c} f(a)\overline{f(a+b)f(a+c)}f(a+b+c)\right)^{1/4}.$$

Veamos que la $U^2$-norma de Gowers controla la distribución de funciones sobre progresiones aritméticas de longitud tres.

**Lema 3.5.** *Sean* $f_1, f_2, f_3 : \mathbb{Z}_N \to \mathbb{C}$ *funciones tales que* $||f_i||_\infty \leq 1$, *para cada* $i = 1, 2, 3$. *Entonces*

$$\left|\sum_{a,d} f_1(a)f_2(a+d)f_3(a+2d)\right| \leq N^2 \min_{i=1,2,3} ||f_i||_{U^2}.$$

*Demostración.* Sea $S = \sum_{a,d} f_1(a)f_2(a+d)f_3(a+2d)$. Por la desigualdad de Cauchy-Schwarz

$$\begin{aligned} |S| &\leq \sum_a \left|\sum_d f_1(a)f_2(a+d)f_3(a+2d)\right| \\ &\leq \left(\sum_a 1^2\right)^{1/2}\left(\sum_a \left|\sum_d f_1(a)f_2(a+d)f_3(a+2d)\right|^2\right)^{1/2}. \end{aligned}$$

Así,

$$\begin{aligned} S^2 &\leq N\sum_a \left|\sum_d f_1(a)f_2(a+d)f_3(a+2d)\right|^2 \\ &= \sum_a\sum_{d,e} f_1(a)f_2(a+d)f_3(a+2d)\overline{f_1(a)f_2(a+e)f_3(a+2e)} \\ &= \sum_a |f_1(a)|^2 \sum_{d,e} f_2(a+d)\overline{f_2(a+e)}f_3(a+2d)\overline{f_3(a+2e)} \\ &\leq N\sum_a\sum_{d,e} f_2(a+d)\overline{f_2(a+e)}f_3(a+2d)\overline{f_3(a+2e)}. \end{aligned}$$

Haciendo los cambios de variables $A = a + d$, $A + k = a + e$, obtenemos

$$\begin{aligned} S^2 &\leq N\sum_A\sum_{d,k} f_2(A)\overline{f_2(A+k)}f_3(A+d)\overline{f_3(A+d+2k)} \\ &= N\sum_{A,k} f_2(A)\overline{f_2(A+k)}\sum_d f_3(A+d)\overline{f_3(A+d+2k)} \\ &\leq \sum_{A,k}\left|\sum_d f_3(A+d)\overline{f_3(A+d+2k)}\right|. \end{aligned}$$

Reemplazando $2k$ por $k$, asumiendo que $N$ es impar, obtenemos que

$$S^2 \leq N \sum_{A,k} \left| \sum_d f_3(A+d)\overline{f_3(A+d+k)} \right|.$$

Nuevamente, por la desigualdad de Cauchy-Schwarz,

$$\begin{aligned} S^2 &\leq N \left( \sum_{A,k} 1^2 \right)^{1/2} \left( \sum_{A,k} \left| \sum_d f_3(A+d)\overline{f_3(A+d+k)} \right|^2 \right)^{1/2} \\ S^4 &\leq N^4 \sum_{A,k} \left| \sum_d f_3(A+d)\overline{f_3(A+d+k)} \right|^2 \\ S^4 &\leq N^4 \sum_{A,k} \sum_{d,e} f_3(A+d)\overline{f_3(A+e)f_3(A+d+k)}f_3(A+e+k). \end{aligned}$$

Haciendo los cambios de variables $a = A+d$, $a+b = A+e$ y $c = k$, tenemos que

$$\begin{aligned} S^4 &\leq N^5 \sum_{a,b,c} f_3(a)\overline{f_3(a+b)f_3(a+c)}f_3(a+b+c) \\ &= N^8 ||f_3||_{U^2}^4. \end{aligned}$$

Así, tenemos que $S \leq N^2 ||f_3||_{U^2}$. Por simetría, análogamente procedemos para $f_1$ y $f_2$. □

**Lema 3.6.** *Sea $A \subseteq [1,N]$, con $|A| = \delta N$ (identifiquemos $A$ con un subconjunto de $\mathbb{Z}_N$). Sea $f$ la función balanceada de $A$. Si $||f||_{U^2} \leq \delta^3/32$, entonces se cumple una de las siguientes afirmaciones:*

*(a) Existen al menos $\delta^3 N^2/32$ progresiones aritméticas de tres términos.*

*(b) Existe una subprogresión $P$ de longitud $N/3$ sobre la que $A$ tiene densidad $9\delta/8$ al menos, esto es,*

$$|A \cap P| \geq (9\delta/8)|P|.$$

*Demostración.* Sean $A_1 = A_2 = A \cap [N/3, 2N/3)$. Si este conjunto tiene cardinalidad menor que $\delta N/4$, entonces

$$|A \cap [1, N/3)| \geq 3\delta N/8 \quad \text{o} \quad |A \cap [2N/3, N]| \geq 3\delta N/8.$$

Por lo tanto

$$|A \cap [1, N/3)| \geq (9\delta/8)|[1, N/3)| \quad \text{o} \quad |A \cap [2N/3, N]| \geq (9\delta/8)|[2N/3, N]|.$$

Supongamos que $|A \cap [N/3, 2N/3)| \geq \delta N/4$, entonces vemos que el número de progresiones aritméticas de tres términos en $A$ es al menos

$$\begin{aligned}\sum_{a,d} A_1(a)A_2(a+d)A(a+2d) &= \sum_{a,d} A_1(a)A_2(a+d)\big(f(a+2d)+\delta\big)\\ &= \delta\sum_{a,d} A_1(a)A_2(a+d) + \sum_{a,d} A_1(a)A_2(a+d)f(a+2d)\\ &\geq \delta|A_1|^2 - \Big|\sum_{a,d} A_1(a)A_2(a+d)f(a+2d)\Big|\\ &\geq \delta^3 N^2/16 - N^2||f||_{U^2},\end{aligned}$$

donde la última desigualdad se da por el lema 3.5. Por tanto, el lema estará probado una vez que $||f||_{U^2} \leq \delta^3/32$. □

Ahora, veamos qué ocurre si la $U^2$-norma de Gowers es grande. Supongamos que $||f||_{U^2} \geq \delta^3/32$. De la igualdad,

$$\sum_k |\widehat{f}(k)|^4 = N\sum_{a,b,c} f(a)\overline{f(a+b)}\,\overline{f(a+c)}f(a+b+c),$$

dada en (2.8), tenemos que

$$||f||_{U^2}^4 = \frac{1}{N^4}\sum_k |\widehat{f}(k)|^4. \tag{3.23}$$

Entonces, usando la identidad de Plancherel

$$\begin{aligned}||f||_{U^2}^4 &\leq \frac{1}{N^4}\max_k |\widehat{f}(k)|^2 \sum_k |\widehat{f}(k)|^2\\ &= \frac{1}{N^3}\max_k |\widehat{f}(k)|^2 \sum_k |f(k)|^2\\ &\leq \frac{1}{N^2}\max_k |\widehat{f}(k)|^2.\end{aligned}$$

Si $||f||_{U^2} \geq \delta^3/32$, entonces

$$(\delta^3/32)^4 \leq ||f||_{U^2}^4 \leq \Big(\frac{1}{N}\max_k |\widehat{f}(k)|\Big)^2,$$

y por ende $\delta^6 N/2^{10} \leq \max_k |\widehat{f}(k)|$. Así, existe $k \neq 0$ (pues $\widehat{f}(0) = 0$) tal que $\delta^6 N/2^{10} \leq |\widehat{f}(k)|$. Por consiguiente, por la proposición 3.2, existe una subprogresión aritmética $P \subseteq [1, N]$, de longitud al menos $\dfrac{\delta^6}{32 \cdot 2^{10}} N^{1/2}$ tal que

$$|A \cap P| \geq \Big(\delta + \frac{\delta^6}{8 \cdot 2^{10}}\Big)|P|.$$

Una observación importante es la siguiente,

**Observación 3.3.** *El incremento de densidad dado por el uso de la $U^2$-norma de Gowers es de $\frac{\delta^6}{8 \cdot 2^{10}}$, que es un incremento mucho menor que el incremento de densidad obtenido en la sección anterior, que fue de $\frac{\delta^2}{64}$. A pesar de ello, el empleo de la $U^2$-norma de Gowers tiene la ventaja de poder ser generalizado para mostrar el teorema de Szemerédi para progresiones aritméticas de longitud $k > 3$, siguiendo aún el argumento del incremento de densidad.*

# Capítulo 4

# Teorema de Szemerédi de longitud cuatro. Perspectiva de Gowers

Como mencionamos antes, el teorema de Szemerédi afirma que cualquier subconjunto con densidad superior positiva de los números naturales contiene alguna progresión aritmética de longitud $k$. El teorema implica el teorema de van der Waerden que afirma que si particionamos el conjunto de los números naturales en una cantidad finita de subconjuntos, entonces alguno de ellos contiene progresiones aritméticas de longitud arbitraria, y fue probado por Szemerédi. El primer progreso dado hacia la prueba del teorema fue debido a Roth [3], quien probó el resultado en el caso especial cuando $k = 3$, usando sumas exponenciales. Posteriormente, Szemerédi halló una prueba diferente, más combinatoria, que pudo extender para probar por primera vez el caso cuando $k = 4$ [1], y posteriormente dando una prueba del caso general [2], para cada $k$. Hubo entonces una prueba adicional debida a Furstemberg [16], quien mostró que técnicas de teoría ergódica podían usarse para probar muchos resultados de teoría de Ramsey que, en general, afirma que en sistemas suficientemente grandes siempre existen subsistemas con estructura, con orden, entre estos el teorema de Szemerédi y ciertas extensiones de este teorema, antes desconocidas.

Una motivación de Gowers para generalizar el argumento analítico de suma de exponenciales de la prueba dada por Roth del teorema de Szemerédi para $k = 3$, al caso general $k > 3$, fue que las cotas dadas por las pruebas conocidas del teorema de Szeméredi eran bastante débiles, y en general para este tipo de problema las mejores cotas vienen del uso

de suma de exponenciales. Por ejemplo, Roth mostró que cuando $k = 3$ podemos tomar $N \geq \exp\exp\big(C/\delta\big)$, donde $C$ es una constante positiva, de modo que cualquier subconjunto $A$ de $\{1, 2, \ldots, N\}$ con cardinalidad mayor que $\delta N$, contiene alguna progresión aritmética no trivial. La cota $\exp\exp\big(C/\delta\big)$ es mejor que las cotas dadas por cualquier argumento combinatorio conocido.

En 1998, Gowers mostró el teorema de Szemerédi para $k = 4$ [1] generalizando el método de Roth. Dado $k = 4$, existe una constante absoluta $C > 0$ tal que si $N \geq \exp\exp\exp\big((1/\delta)^C\big)$ entonces cualquier subconjunto de $\{1, 2, \ldots, N\}$ de tamaño al menos $\delta N$ contiene alguna progresión aritmética (no trivial) de longitud cuatro. Y en 2001, dio una prueba del caso general del teorema de Szemerédi [2], dando una cota del tipo $\exp\exp\big((1/\delta)^C\big)$. Aunque una cota de este tipo puede parecer débil, constituye una mejora respecto a las cotas conocidas anteriormente. Para tener una idea de la magnitud de estas cotas definimos inductivamente la función $T$ por $T(1) = 2$ y $T(n+1) = 2^{T(n)}$, así como también definimos recursivamente la función $W$ por $W(1) = 2$ y $W(n+1) = T\big(W(n)\big)$. Las cotas previas conocidas para el teorema de Szemerédi fueron tan grandes como $W(1/\delta)$.

## 4.1. Uniformidad cuadrática y la $U^3$-norma de Gowers pequeña

Dada una función $f : \mathbb{Z}_N \to \mathbb{C}$ y $k \in \mathbb{Z}_N$, consideremos como antes $\Delta(f;k)(s) = f(s)\overline{f(s-k)}$. Notemos que si $f(s) = \omega^{\phi(s)}$ para alguna función $\phi : \mathbb{Z}_N \to \mathbb{Z}_N$, entonces $\Delta(f;k)(s) = \omega^{\phi(s)-\phi(s-k)}$. Que un conjunto $A \subseteq \{1, 2, \ldots, N\}$ sea *uniforme* nos permite concluir que $A$ contiene progresiones aritméticas de longitud tres. Sin embargo, para determinar progresiones aritméticas de longitud cuatro en $A$ este concepto ya no es suficiente. Requerimos de conjuntos que sean *cuadráticamente uniformes.*

**Lema 4.1.** *Sea la función $f : \mathbb{Z}_N \to D$, donde $D$ es el disco cerrado unitario en $\mathbb{C}$. Las siguientes afirmaciones son equivalentes:*

*(i)* $\displaystyle\sum_u \sum_v \Big|\sum_s f(s)\overline{f(s-u)f(s-v)}f(s-u-v)\Big|^2 \leq c_1 N^4.$

*(ii)* $\displaystyle\sum_k \sum_r \big|\widehat{\Delta(f;k)}(r)\big|^4 \leq c_2 N^5.$

*(iii)* $\left|\widehat{\Delta(f;k)}(r)\right| \geq c_3 N$ *para a lo más* $c_3^2 N$ *pares* $(k,r)$.

*(iv) Para todos los valores de* $k$*, excepto quizás para* $c_4 N$ *valores a lo más, la función* $\Delta(f;k)$ *es* $c_4$*-uniforme.*

**Observación 4.1.** *La afirmación (iii) del lema anterior puede reescribirse como:*

$$T = \Big\{(k,r) \in \mathbb{Z}_N^2;\ |\Delta(f;k)\widehat{\ }(r)| \geq c_3 N\Big\} \text{ es tal que } |T| \leq c_3^2 N.$$

*Demostración.* Veamos la equivalencia entre (i) y (ii). Tenemos que

$$\begin{aligned}
&\sum_u \sum_v \Big|\sum_s f(s)\overline{f(s-u)f(s-v)}f(s-u-v)\Big|^2 \\
&= \sum_u \sum_v \Big|\sum_s \Delta(f;u)(s)\overline{\Delta(f;u)(s-v)}\Big|^2 \\
&= \sum_u \sum_v \Big|\big(\Delta(f;u) * \Delta(f;u)\big)(v)\Big|^2 \\
&= \sum_u N^{-1} \sum_v \Big|\big(\Delta(f;u) * \Delta(f;u)\big)\widehat{\ }(v)\Big|^2 \\
&= N^{-1} \sum_u \sum_v \Big|\Delta(f;u)\widehat{\ }(v)\overline{\Delta(f;u)\widehat{\ }(v)}\Big|^2 \\
&= N^{-1} \sum_u \sum_v \Big|\Delta(f;u)\widehat{\ }(v)\Big|^4.
\end{aligned}$$

Por lo tanto (i) y (ii) son equivalentes con $c_1 = c_2$.

Ahora veamos la equivalencia entre (ii) y (iii). Mostremos que (ii) implica (iii). Supongamos que

$$\sum_k \sum_r \Big|\widehat{\Delta(f;k)}(r)\Big|^4 \leq c_2 N^5 \text{ para algún } c_2 > 0,$$

y supongamos que (iii) no se verifica para algún $c_3$ que depende de $c_2$ y que fijaremos al final, esto es, siendo

$$T = \Big\{(k,r) \in \mathbb{Z}_N^2;\ |\Delta(f;k)\widehat{\ }(r)| \geq c_3 N\Big\},$$

suponemos que $|T| > c_3^2 N$. Entonces

$$\sum_k \sum_r |\Delta(f;k)\widehat{\ }(r)|^4 \geq \sum_{(k,r)\in T} |\Delta(f;k)\widehat{\ }(r)|^4 > |T|(c_3 N)^4 = c_3^6 N^5.$$

Así, fijando $c_3 = c_2^{1/6}$ obtenemos una contradicción.

Mostremos que (iii) implica (ii). Dado $T = \Big\{(k,r) \in \mathbb{Z}_N^2;\ |\Delta(f;k)\widehat{\ }(r)| \geq c_3 N\Big\}$, supongamos

que $|T| \leq c_3^2 N$ para algún $c_3 > 0$. Por contradicción, suponamos que (ii) no se cumple para alguna constante $c_2 > 0$ que depende de $c_3$ y que fijaremos al final, esto es

$$\sum_k \sum_r |\Delta(f;k)\widehat{\ }(r)|^4 > c_2 N^5 \text{ para algún } c_2 > 0, \tag{4.1}$$

entonces existen más que $c_2 N/2$ valores de $k$ para los que $\sum_r |\Delta(f;k)\widehat{\ }(r)|^4 > c_2 N^4/2$. En efecto, sea

$$J = \Big\{ k \in \mathbb{Z}_N;\ \sum_r |\Delta(f;k)\widehat{\ }(r)|^4 \leq c_2 N^4/2 \Big\},$$

y veamos que $|J^c| > c_2 N/2$. De (4.1) tenemos que

$$\sum_{k \notin J} \sum_r |\Delta(f;k)\widehat{\ }(r)|^4 + \sum_{k \in J} \sum_r |\Delta(f;k)\widehat{\ }(r)|^4 > c_2 N^5,$$

y

$$\sum_r |\Delta(f;k)\widehat{\ }(r)|^4 = N \sum_p \Big| \sum_s \Delta(f;k)(s) \overline{\Delta(f;k)(s-p)} \Big|^2 \leq N^4.$$

Por lo tanto

$$|J^c| N^4 + |J| c_2 N^4/2 > c_2 N^5.$$

Así $|J^c| + (N - |J^c|)c_2/2 > c_2 N$, y $|J^c| > |J^c|(1 - c_2/2) > c_2 N/2$. Por ende existen más que $c_2 N/2$ valores de $k$ tales que

$$\sum_r |\Delta(f;k)\widehat{\ }(r)|^4 > c_2 N^4/2.$$

Como

$$\begin{aligned} \sum_r |\Delta(f;k)\widehat{\ }(r)|^4 &\leq \max_r |\Delta(f;k)\widehat{\ }(r)|^2 \sum_r |\Delta(f;k)\widehat{\ }(r)|^2 \\ &\leq \max_r |\Delta(f;k)\widehat{\ }(r)|^2 \cdot N \sum_r |\Delta(f;k)(r)|^2 \\ &\leq N^2 \max_r |\Delta(f;k)\widehat{\ }(r)|^2, \end{aligned}$$

tenemos que existen más que $(c_2/2)N$ valores de $k$ para los cuales $\max_r |\Delta(f;k)\widehat{\ }(r)| > (c_2/2)^{1/2} N$. Así, si tomamos $c_2 = 2c_3^2$, obtenemos una contradicción con la hipótesis (iii). Por lo tanto, para $c_2 = 2c_3^2$ se tiene que $\sum_k \sum_r \big|\widehat{\Delta(f;k)}(r)\big|^4 \leq c_2 N^5$.

Probemos que (ii) implica (iv). Supongamos que se cumple $\sum_k \sum_r \big|\widehat{\Delta(f;k)}(r)\big|^4 \leq c_2 N^5$ para algún $c_2 > 0$, y supongamos que (iv) no se cumple para alguna constante $c_4 > 0$ que depende de $c_2$ y que fijaremos al final, entonces definiendo

$$H = \Big\{ k \in \mathbb{Z}_N;\ \sum_r |\Delta(f;k)\widehat{\ }(r)|^4 > c_4 N^4 \Big\},$$

suponemos que $|H| > c_4 N$. Por lo tanto

$$\sum_k \sum_r |\Delta(f;k)\widehat{\ }(r)|^4 \geq \sum_{k\in H} \sum_r |\Delta(f;k)\widehat{\ }(r)|^4 > |H| c_4 N^4 > c_4^2 N^5.$$

Así, habiendo fijado $c_4 = c_2^{1/2}$ llegamos a una contradicción con (ii). Por ende $|H| \leq c_4 N$, con $c_4 = c_2^{1/2}$. Recíprocamente, supongamos que se cumple (iv) para alguna constante $c_4 > 0$, esto es $|H| \leq c_4 N$, donde

$$H = \Big\{ k \in \mathbb{Z}_N;\ \sum_r |\Delta(f;k)\widehat{\ }(r)|^4 > c_4 N^4 \Big\}.$$

Por lo tanto,

$$\sum_r |\Delta(f;k)\widehat{\ }(r)|^4 \leq c_4 N^4 \text{ para cada } k \in H^c,$$

y

$$N^4 \geq \sum_r |\Delta(f;k)\widehat{\ }(r)|^4 > c_4 N^4 \text{ para cada } k \in H.$$

Así

$$\begin{aligned}\sum_k \sum_r |\Delta(f;k)\widehat{\ }(r)|^4 &= \sum_{k\in H^c} \sum_r |\Delta(f;k)\widehat{\ }(r)|^4 + \sum_{k\in H} \sum_r |\Delta(f;k)\widehat{\ }(r)|^4 \\ &\leq |H^c| \cdot c_4 N^4 + |H| N^4 \leq \big(c_4 |H^c| + |H|\big) N^4 \\ &\leq \big(c_4 N + |H|\big) N^4 \leq 2 c_4 N^5.\end{aligned}$$

Por ende, eligiendo $c_2 = 2c_4$ tenemos que (iv) implica (ii). □

**Definición 4.1.** *Diremos que $f : \mathbb{Z}_N \to D$ es* **cuadráticamente $\alpha$-uniforme** *si*

$$\sum_u \sum_v \Big| \sum_s f(s) \overline{f(s-u) f(s-v)} f(s-u-v) \Big|^2 \leq \alpha N^4.$$

*Un conjunto $A \subseteq \mathbb{Z}_N$ será* **cuadráticamente $\alpha$-uniforme** *si su función balanceada respectiva $f_A$ lo es.*

**Lema 4.2.** *Para cada $1 \leq i \leq k$ sea $f_i : \mathbb{Z}_N \to D$ una función $\alpha_i$-uniforme. Entonces $f_1 + \cdots + f_k$ es $(\alpha_1^{1/4} + \cdots + \alpha_k^{1/4})^4$-uniforme.*

*Demostración.* Se sigue de la definición de $\alpha_i$-uniformidad y del hecho que $\left( \sum_r |\widehat{f}(r)|^4 \right)^{1/4}$ es una norma. □

**Lema 4.3.** *Sea $A \subseteq \mathbb{Z}_N$ un conjunto cuadráticamente $\alpha$-uniforme de tamaño $\delta N$. Entonces para todo $k$, excepto para a lo más $\alpha^{1/2}N$ valores de $k$, $A \cap (A+k)$ es $81\alpha^{1/2}$-uniforme. Además, para todo $k$, excepto a lo más $\alpha^{1/4}N$ valores de $k$,*

$$\Big|\big|A \cap (A+k)\big| - \delta^2 N\Big| \leq \alpha^{1/8} N.$$

*Demostración.* Sea $f = f_A$ la función balanceada de $A$, entonces como $1_{A+k}(x) = 1_A(x-k)$ para cada $x$, tenemos que la función balanceada de $A+k$ es

$$\begin{aligned} f_{A+k}(x) &= 1_{A+k}(x) - \delta \cdot 1(x) = 1_{A+k}(x) - \delta \\ &= 1_A(x-k) - \delta = 1_A(x-k) - \delta \cdot 1(x-k) \\ &= f_A(x-k). \end{aligned}$$

Así,

$$\begin{aligned} 1_{A\cap(A+k)}(s) = 1_A(s)1_{A+k}(s) &= (f+\delta\cdot 1)(s)(f_{A+k}+\delta\cdot 1)(s) \\ &= (f(s)+\delta)(f_{A+k}(s)+\delta) \\ &= (f(s)+\delta)(f(s-k)+\delta) \\ &= \delta^2 + \delta f(s) + \delta f(s-k) + f(s)f(s-k). \end{aligned} \tag{4.2}$$

Las implicaciones de (i) (condición de $c_1$-uniformidad cuadrática) a (ii), tomando $c_2 = c_1$, y de (ii) a (iv) con $c_4 = c_2^{1/2}$ (por tanto $c_4 = c_1^{1/2}$) en el lema 4.1, dan lugar, debido a la $\alpha$-uniformidad cuadrática de $f$, a que las funciones $s \mapsto f(s)f(s-k)$ sean $\alpha^{1/2}$-uniformes para todo $k$, excepto quizás a lo más para $\alpha^{1/2}N$ valores de $k$. Además, como $f$ es cuadráticamente $\alpha$-uniforme, $f$ es $\alpha^{1/2}$-uniforme. En efecto,

$$\begin{aligned} \sum_k \Big|\sum_s f(s)f(s-k)\Big|^2 &= \sum_k \sum_s f(s)f(s-k) \sum_{s'} \overline{f(s')f(s'-k)} \\ &= \sum_k \sum_s \sum_{s'} f(s)f(s-k)\overline{f(s')f(s'-k)} \\ &= \sum_k \sum_s \sum_u f(s)f(s-k)f(s-u)f(s-u-k) \\ &= \sum_k \sum_u \Big[\sum_s f(s)f(s-k)f(s-u)f(s-u-k)\Big] \\ &\leq \sum_k \sum_u \Big|\sum_s f(s)f(s-k)f(s-u)f(s-u-k)\Big| \cdot 1 \\ &\leq \Big(\sum_k \sum_u \Big|\sum_s f(s)f(s-k)f(s-u)f(s-u-k)\Big|^2\Big)^{1/2} \Big(\sum_k \sum_u 1^2\Big)^{1/2}. \end{aligned}$$

Por lo tanto,

$$\sum_k |\sum_s f(s)f(s-k)|^2 \leq N\Big(\sum_k \sum_u \Big|\sum_s f(s)f(s-k)f(s-u)f(s-u-k)\Big|^2\Big)^{1/2}. \qquad (4.3)$$

Si $f$ es cuadráticamente $\alpha$-uniforme, tenemos que

$$\sum_k \sum_u \Big|\sum_s f(s)f(s-k)f(s-u)f(s-u-k)\Big|^2 \leq \alpha N^4,$$

luego de (4.3)

$$\sum_k |\sum_s f(s)f(s-k)|^2 \leq N \cdot \alpha^{1/2} N^2 = \alpha^{1/2} N^3,$$

y así $f$ es $\alpha^{1/2}$-uniforme (por definición). También tenemos que

$$\begin{aligned}
\sum_k |\sum_s f(s)f(s-k)|^2 &= \sum_{k'} |\sum_s f(s)f(s-k')|^2 \\
&= \sum_{k'} |\sum_s f(s-k)f(s-k-k')|^2 \\
&= \sum_{k'} |\sum_s f_{A+k}(s) f_{A+k}(s-k')|^2,
\end{aligned}$$

por tanto $f_{A+k}$ es también $\alpha^{1/2}$-uniforme para cada $k$. Entonces por el lema 4.2, $A \cap (A+k)$ es $81\alpha^{1/2}$-uniforme para al menos $(1-\alpha^{1/2})N$ valores de $k$ (ya que $s \mapsto f(s)f(s-k)$ es $\alpha^{1/2}$-uniforme para todo $k$, excepto quizás para a lo más $\alpha^{1/2}N$ valores de $k$). En efecto, la función balanceada asociada a $A \cap (A+k)$ es $f_{A\cap(A+k)} = A \cap (A+k) - 1 \cdot \rho$, donde $\rho$ es tal que $|A \cap (A+k)| = \rho N$. Pero

$$\begin{aligned}
|A \cap (A+k)| &= \sum_s A \cap (A+k)(s) \\
&= \sum_s \big[\delta^2 + \delta f(s) + \delta f(s-k) + f(s)f(s-k)\big] \\
&= \delta^2 N + \delta \sum_s f(s) + \delta \sum_s f(s-k) + \sum_s f(s)f(s-k) \\
&= \delta^2 N + \sum_s f(s)f(s-k).
\end{aligned}$$

Por tanto $\rho = \delta^2 + N^{-1}\sum_s f(s)f(s-k)$ y

$$f_{A\cap(A+k)} = A \cap (A+k) - \delta^2 - N^{-1} \sum_s f(s)f(s-k).$$

De (4.2),

$$\begin{aligned}
f_{A\cap(A+k)}(s) &= \delta f(s) + \delta f(s-k) + f(s)f(s-k) - N^{-1}\sum_t f(t)f(t-k) \\
f_{A\cap(A+k)} &= \delta f_A + \delta f_{A+k} + f_A f_{A+k} - N^{-1}\sum_t f(t)f(t-k)
\end{aligned}$$

Como $f_A$ es $\alpha^{1/2}$-uniforme, $f_{A+k}$ es $\alpha^{1/2}$-uniforme y $f_A f_{A+k}$ es $\alpha^{1/2}$-uniforme para al menos $(1-\alpha^{1/2})N$ valores de $k$, tenemos que

$$\begin{aligned}
\sum_r \big|\widehat{f}_A(r)\big|^4 &\leq \alpha^{1/2}N^4,\\
\sum_r \big|\widehat{f}_{A+k}(r)\big|^4 &\leq \alpha^{1/2}N^4,\\
\sum_r \big|\widehat{f_A f_{A+k}}(r)\big|^4 &\leq \alpha^{1/2}N^4 \text{ para al menos } (1-\alpha^{1/2})N \text{ valores de } k,\\
\sum_r \big|\widehat{g}(r)\big|^4 &= \big|\widehat{g}(0)\big|^4 = \big|\mathbb{E}(f_A f_{A+k})\big|^4 N^4,
\end{aligned}$$

donde $g$ es la función constante igual a $\mathbb{E}(f_A f_{A+k}) = N^{-1}\sum_t f(t)f(t-k)$. Por tanto

$$\begin{aligned}
\left(\sum_r \big|\widehat{f}_{A\cap(A+k)}(r)\big|^4\right)^{1/4} &= \left(\sum_r |\delta\widehat{f}_A(r) + \delta\widehat{f}_{A+k}(r) + \big(f_A f_{A+k}\big)\widehat{\ }(r) - \widehat{g}(r)|^4\right)^{1/4}\\
&= \left(\sum_{r\neq 0} |\delta\widehat{f}_A(r) + \delta\widehat{f}_{A+k}(r) + \big(f_A f_{A+k}\big)\widehat{\ }(r)|^4\right)^{1/4}\\
&= \left(\sum_{r\neq 0} |\delta\widehat{f}_A(r)|^4\right)^{1/4} + \left(\sum_{r\neq 0} |\delta\widehat{f}_{A+k}(r)|^4\right)^{1/4} + \left(\sum_{r\neq 0} |\widehat{f_A f_{A+k}}(r)|^4\right)^{1/4}\\
&\leq \Big(\delta\alpha^{1/8} + \delta\alpha^{1/8} + \alpha^{1/8}\Big)N \leq (3\alpha^{1/8})N,
\end{aligned}$$

y así

$$\sum_r |\widehat{f}_{A\cap(A+k)}(r)|^4 \leq (81\alpha^{1/2})N^4,$$

esto es, $A\cap(A+k)$ es $81\alpha^{1/2}$-uniforme para al menos $(1-\alpha^{1/2})N$ valores de $k$.

Por otro lado, de (4.2),

$$\Big|\big|A\cap(A+k)\big| - \delta^2 N\Big| = \Big|\sum_s f(s)f(s-k)\Big|,$$

y puesto que $f$ es $\alpha^{1/2}$-uniforme tenemos que $\displaystyle\sum_k \Big|\sum_s f(s)f(s-k)\Big|^2 \leq \alpha^{1/2}N^3$. Así

$$\sum_k \Big|\big|A\cap(A+k)\big| - \delta^2 N\Big|^2 \leq \alpha^{1/2}N^3. \tag{4.4}$$

Definamos $J = \Big\{k;\ \big||A\cap(A+k)| - \delta^2 N\big| \leq \alpha^{1/8}N\Big\} = \Big\{k;\ \big||A\cap(A+k)| - \delta^2 N\big|^2 \leq \alpha^{1/4}N^2\Big\}$, y mostremos que $|J^c| \leq \alpha^{1/4}N$. Supongamos por contradicción que $|J^c| > \alpha^{1/4}N$, entonces

tenemos, de (4.4), que

$$\alpha^{1/2}N^3 \geq \sum_{k\in J^c}\Big||A\cap(A+k)|-\delta^2N\Big|^2 > |J^c|\Big(\alpha^{1/4}N^2\Big),$$

por lo tanto $\alpha^{1/4}N > |J^c| > \alpha^{1/4}N$, lo cual es una contradicción. Así, $|J^c| \leq \alpha^{1/4}N$, esto es

$$\Big||A\cap(A+k)|-\delta^2N\Big| \leq \alpha^{1/8}N$$

para todo $k$, excepto a lo más para $\alpha^{1/4}N$ valores de $k$. □

**Observación 4.2.** *Sea $f : \mathbb{Z}_N \to \mathbb{R}$, entonces la desigualdad de Cauchy-Schwartz implica que $||f||_2 \geq N^{-1/2}||f||_1$.*

**Lema 4.4.** *Sea $f : \mathbb{Z}_N \to \mathbb{R}_+$ una función con $||f||_1 = \omega N$ y supongamos que*

$$||f||_2^2 \leq (1+\epsilon)\omega^2N = N^{-1}||f||_1^2(1+\epsilon).$$

*Si $A \subseteq \mathbb{Z}_N$, entonces $\left|\sum_{s\in A} f(s) - \omega|A|\right| \leq \epsilon^{1/2}\omega N^{1/2}|A|^{1/2}$.*

*Demostración.* La media de $f$ es $\omega$ y su varianza es $\epsilon\omega^2$. Entonces

$$\begin{aligned}
\left|\sum_{s\in A} f(s) - \omega|A|\right| &\leq \sum_{s\in A}|f(s)-\omega| \\
&\leq |A|^{1/2}\left(\sum_{s\in A}\big(f(s)-\omega\big)^2\right)^{1/2} \\
&\leq |A|^{1/2}\left(\sum_{s\in A}\big(f(s)^2 - 2\omega f(s)+\omega^2\big)\right)^{1/2} \\
&\leq |A|^{1/2}\left(\sum_{s}\big(f(s)^2 - 2\omega f(s)+\omega^2\big)\right)^{1/2} \\
&\leq |A|^{1/2}\big((1+\epsilon)\omega^2N + \omega^2N - 2\omega^2N\big)^{1/2} \\
&= |A|^{1/2}(\epsilon\omega^2N)^{1/2}.
\end{aligned}$$

□

**Lema 4.5.** *Sea $N \in \mathbb{N}$ impar. Sean $A, B, C \subseteq \mathbb{Z}_N$ tales que $|A| = \alpha N$, $|B| = \beta N$ y $|C| = \gamma N$. Supongamos que $C$ es $\eta$-uniforme. Entonces*

$$\left|\sum_r |A\cap(B+r)\cap(C+2r)| - \alpha\beta\gamma N^2\right| \leq \eta N^2.$$

*Demostración.* Como antes, identificaremos los conjuntos $A$, $B$ y $C$ con sus respectivas funciones características. Entonces

$$\sum_r \big|A \cap (B+r) \cap (C+2r)\big| = \sum_r \sum_s A(s)B(s-r)C(s-2r).$$

De la fórmula de inversión de $A(s)$, $B(s-r)$ y $C(s-2r)$, tenemos que el lado derecho de la anterior igualdad se puede expresar como

$$\begin{aligned}
&\sum_r \sum_s \left(N^{-1}\sum_x \widehat{A}(x)\omega^{-xs}\right)\left(N^{-1}\sum_y \widehat{B}(y)\omega^{-y(s-r)}\right)\left(N^{-1}\sum_z \widehat{C}(z)\omega^{-z(s-2r)}\right)\\
&\quad = N^{-3}\sum_{r,s}\sum_{x,y,z} \widehat{A}(x)\widehat{B}(y)\widehat{C}(z)\omega^{-xs-y(s-r)-z(s-2r)}\\
&\quad = N^{-3}\sum_{x,y,z} \widehat{A}(x)\widehat{B}(y)\widehat{C}(z)\sum_{r,s}\omega^{-xs-y(s-r)-z(s-2r)}\\
&\quad = N^{-3}\sum_{x,y,z} \widehat{A}(x)\widehat{B}(y)\widehat{C}(z)\sum_r \omega^{r(y+2z)}\sum_s \omega^{-s(x+y+z)}\\
&\quad = N^{-1}\sum_{x,y,z,\ x=z,\ y=-2z} \widehat{A}(x)\widehat{B}(y)\widehat{C}(z)\\
&\quad = N^{-1}\sum_p \widehat{A}(p)\widehat{B}(-2p)\widehat{C}(p)\\
&\quad = N^{-1}\sum_{p\neq 0} \widehat{A}(p)\widehat{B}(-2p)\widehat{C}(p) + N^{-1}|A|\,|B|\,|C|.
\end{aligned}$$

Así,

$$\sum_r \big|A \cap (B+r) \cap (C+2r)\big| = N^{-1}\sum_{p\neq 0} \widehat{A}(p)\widehat{B}(-2p)\widehat{C}(p) + N^{-1}|A|\,|B|\,|C|. \tag{4.5}$$

Sin embargo, por la $\eta$-uniformidad de $C$, y por la desigualdad de Cauchy-Schwarz

$$\left|\sum_{p\neq 0} \widehat{A}(p)\widehat{B}(-2p)\widehat{C}(p)\right| \leq \eta N^3. \tag{4.6}$$

En efecto,

$$\begin{aligned}
\left|\sum_{p\neq 0} \widehat{A}(p)\widehat{B}(-2p)\widehat{C}(p)\right| &\leq \max_{p\neq 0}\big|\widehat{C}(p)\big| \sum_{p\neq 0}\big|\widehat{A}(p)\big|\,\big|\widehat{B}(-2p)\big|\\
&\leq \max_{p\neq 0}\big|\widehat{C}(p)\big| \sum_{p}\big|\widehat{A}(p)\big|\,\big|\widehat{B}(-2p)\big|\\
&\leq \max_{p\neq 0}\big|\widehat{C}(p)\big| \left(\sum_p \big|\widehat{A}(p)\big|^2\right)^{1/2}\left(\sum_p \big|\widehat{B}(-2p)\big|^2\right)^{1/2}\\
&\leq \max_{p\neq 0}\big|\widehat{C}(p)\big|\,||\widehat{A}||_2\,||\widehat{B}||_2\\
&\leq \eta N\cdot N\cdot N.
\end{aligned}$$

El lema se sigue de (4.5) y (4.6). □

**Lema 4.6.** *Sean $A, B, C$ y $D$ subconjuntos de $\mathbb{Z}_N$ con cardinalidades $\alpha N$, $\beta N$ $\gamma N$ y $\delta N$ respectivamente. Supongamos que $C$ sea $\eta$-uniforme y, $C$ y $D$ sean cuadráticamente $\eta$-uniformes para algún $\eta \leq 2^{-20}$. Entonces*

$$\left|\sum_r \big|A \cap (B+r) \cap (C+2r) \cap (D+3r)\big| - \alpha\beta\gamma\delta N^2\right| \leq \frac{3\eta^{1/16}N^2}{\beta\gamma\delta}.$$

*Demostración.* Una vez más identificamos conjuntos con sus funciones características correspondientes. Sea $f(s) = \sum_r B(s-r)C(s-2r)D(s-3r)$, esto es

$$f = \sum_r (B+r) \cap (C+2r) \cap (D+3r).$$

Estimemos las normas $||f||_1$ y $||f||_2$. La prueba del lema 4.3 nos dice que $D$ es $\eta^{1/2}$-uniforme, por consiguiente, por el lema 4.5,

$$\begin{aligned} ||f||_1 &= \sum_s \sum_r B(s-r)C(s-2r)D(s-3r) \\ &= \sum_r \sum_s B(s-r)C(s-2r)D(s-3r) \\ &= \sum_r \big|B \cap (C+r) \cap (C+2r)\big| \\ &\geq N^2\big(\beta\gamma\delta - \eta^{1/2}\big). \end{aligned} \tag{4.7}$$

También por el mismo lema,

$$||f||_1 = \sum_r \big|B \cap (C+r) \cap (D+2r)\big| \leq N^2(\beta\gamma\delta + \eta^{1/2}),$$

y por lo tanto

$$\Big|\frac{||f||_1}{N} - \beta\gamma\delta N\Big| \leq \eta^{1/2}N. \tag{4.8}$$

Para $||f||_2$, tenemos que

$$||f||_2^2 = \sum_s \sum_{r,q} B(s-r)B(s-q)C(s-2r)C(s-2q)D(s-3r)D(s-3q).$$

Si sustituimos $p = q - r$, entonces resulta que

$$\begin{aligned} &\sum_s \sum_{r,p} B(s-r)B(s-r-p)C(s-2r)C(s-2r-2p)D(s-3r)D(s-3r-3p) \\ &\qquad = \sum_s \sum_{r,p} B(s)B(s-p)C(s-r)C(s-r-2p)D(s-2r)D(s-2r-3p) \end{aligned}$$

$$= \sum_{r,p} \sum_{s} B(s)(B+p)(s)(C+r)(s)(C+r+2p)(s)(D+2r)(s)(D+2r+3p)(s)$$
$$= \sum_{r,p} \Big| \Big( B \cap (B+p) \Big) \cap \Big( C \cap (C+2p) + r \Big) \cap \Big( D \cap (D+3p) + 2r \Big) \Big|.$$

Por el lema 4.3, $D \cap (D+3p)$ es $81\eta^{1/2}$-uniforme para todo $p$, excepto a lo más para $\eta^{1/2}N$ valores de $p$. Cuando $D \cap (D+3p)$ es $81\eta^{1/2}$-uniforme, el lema 4.5 implica que

$$\sum_{r} \Big| \Big( B \cap (B+p) \Big) \cap \Big( C \cap (C+2p) + r \Big) \cap \Big( D \cap (D+3p) + 2r \Big) \Big| \tag{4.9}$$

es a lo más

$$N^{-1} |B \cap (B+p)| \, |C \cap (C+2p)| \, |D \cap (D+3p)| + 81\eta^{1/2} N^2.$$

Sumando sobre $p$ tenemos que

$$||f||_2^2 \leq N^{-1} \sum_{p} |B \cap (B+p)| \, |C \cap (C+2p)| \, |D \cap (D+3p)| + 82\eta^{1/2} N^3.$$

En efecto, siendo $P = \big\{ p;\ D \cap (D+3p) \text{ es } 81\eta^{1/2}\text{-uniforme} \big\}$, tenemos que $|P^c| \leq \eta^{1/2} N$, y de (4.9),

$$\begin{aligned}
||f||_2^2 &= \sum_{p} \sum_{r} \big| \big( B \cap (B+p) \big) \cap \big( C \cap (C+2p) + r \big) \cap \big( D \cap (D+3p) + 2r \big) \big| \\
&= \sum_{p \in P} \sum_{r} \big| \big( B \cap (B+p) \big) \cap \big( C \cap (C+2p) + r \big) \cap \big( D \cap (D+3p) + 2r \big) \big| + \\
&\quad + \sum_{p \notin P} \sum_{r} \big| \big( B \cap (B+p) \big) \cap \big( C \cap (C+2p) + r \big) \cap \big( D \cap (D+3p) + 2r \big) \big| \\
&\leq \sum_{p \in P} \Big( N^{-1} |B \cap (B+p)| \, |C \cap (C+2p)| \, |D \cap (D+3p)| + 81\eta^{1/2} N^2 \Big) + \\
&\quad + \sum_{p \notin P} \sum_{r} \big| \big( B \cap (B+p) \big) \cap \big( C \cap (C+2p) + r \big) \cap \big( D \cap (D+3p) + 2r \big) \big| \\
&\leq \sum_{p \in P} \Big( N^{-1} |B \cap (B+p)| \, |C \cap (C+2p)| \, |D \cap (D+3p)| + 81\eta^{1/2} N^2 \Big) + |P^c| N^2 \\
&\leq \sum_{p} \Big( N^{-1} |B \cap (B+p)| \, |C \cap (C+2p)| \, |D \cap (D+3p)| + 81\eta^{1/2} N^2 \Big) + \eta^{1/2} N^3 \\
&\leq N^{-1} \sum_{p} |B \cap (B+p)| \, |C \cap (C+2p)| \, |D \cap (D+3p)| + 82\eta^{1/2} N^3.
\end{aligned}$$

Puesto que $C$ y $D$ son cuadráticamente $\eta$-uniformes, el lema 4.3 implica que

$$|C \cap (C+2p)| \leq \gamma^2 N + \eta^{1/8} N$$

y

$$|D \cap (D+3p)| \leq \delta^2 N + \eta^{1/8} N$$

para todo $p$, excepto a lo más $2\eta^{1/4}N$ valores de $p$. Llamemos $H$ al conjunto de valores de $p$ para los cuales no se cumplen las desigualdades anteriores. Entonces $|H| \leq 2\eta^{1/4}N$, y

$$\begin{aligned}
||f||_2^2 &\leq N^{-1} \sum_{p \notin H} |B \cap (B+p)|\, |C \cap (C+2p)|\, |D \cap (D+3p)| + \\
&\quad + N^{-1} \sum_{p \in H} |B \cap (B+p)|\, |C \cap (C+2p)|\, |D \cap (D+3p)| + 82\eta^{1/2}N^3 \\
&\leq N^{-1} \sum_{p \notin H} |B \cap (B+p)|\, |C \cap (C+2p)|\, |D \cap (D+3p)| + \\
&\quad + N^{-1}|H|N^3 + 82\eta^{1/2}N^3 \\
&\leq N^{-1} \sum_{p \notin H} |B \cap (B+p)|(\gamma^2 + \eta^{1/8})(\delta^2 + \eta^{1/8})N^2 + \\
&\quad + 2\eta^{1/4}N^3 + 82\eta^{1/2}N^3.
\end{aligned}$$

Además,

$$\begin{aligned}
\sum_p |B \cap (B+p)| &= \sum_p \sum_s \big(B \cap (B+p)\big)(s) = \sum_p \sum_s 1_B(s) 1_B(s-p) \\
&= \sum_s 1_B(s) \sum_p 1_B(s-p) = \left(\sum_p 1_B(p)\right)^2 = |B|^2 = \beta^2 N^2,
\end{aligned}$$

y por ende

$$\begin{aligned}
||f||_2^2 &\leq N^{-1}\big(\beta^2 N^2\big)\big(\gamma^2 + \eta^{1/8}\big)\big(\delta^2 + \eta^{1/8}\big)N^2 + 2\eta^{1/4}N^3 + 82\eta^{1/2}N^3 \\
&\leq N^3\Big(\beta^2\gamma^2\delta^2 + \beta^2\big(\gamma^2 + \delta^2\big)\eta^{1/8} + \beta^2\eta^{1/4}\Big) + 2\eta^{1/4}N^3 + 82\eta^{1/2}N^3 \\
&\leq N^3\beta^2\gamma^2\delta^2 + N^3\big(2\eta^{1/8} + \eta^{1/4}\big) + 2\eta^{1/4}N^3 + 82\eta^{1/2}N^3 \\
&\leq N^3\big(\beta^2\gamma^2\delta^2 + 3\eta^{1/8}\big), \qquad (4.10)
\end{aligned}$$

gracias a que $\eta \leq 2^{-20}$, de las hipótesis del lema. De (4.7) y (4.10),

$$\frac{||f||_2^2}{||f||_1^2} \leq \frac{N^3(\beta^2\gamma^2\delta^2 + 3\eta^{1/8})}{N^4(\beta\gamma\delta - \eta^{1/2})^2},$$

y

$$\begin{aligned}
||f||_2^2 &\leq N^{-1}||f||_1^2\left(1 + \frac{3\eta^{1/8}}{\beta^2\gamma^2\delta^2}\right)\left(1 - \frac{\eta^{1/2}}{\beta\gamma\delta}\right)^{-2} \\
||f||_2^2 &\leq N^{-1}||f||_1^2\left(1 + 4\frac{\eta^{1/8}}{\beta^2\gamma^2\delta^2}\right).
\end{aligned}$$

Ahora aplicamos el lema 4.4 con $\epsilon = \dfrac{4\eta^{1/8}}{\beta^2\gamma^2\delta^2}$ y $\left|\omega - \beta\gamma\delta N\right| \leq \eta^{1/2}N$, deduciendo que

$$\begin{aligned}\left|\sum_{s\in A} f(s) - \alpha\beta\gamma\delta N^2\right| &\leq \left|\sum_{s\in A} f(s) - \omega|A|\right| + \left|\omega|A| - \alpha\beta\gamma\delta N^2\right| \\ &= \left|\sum_{s\in A} f(s) - \omega|A|\right| + |A|\cdot|\omega - \beta\gamma\delta N| \\ &\leq \left(\frac{2\alpha^{1/2}\eta^{1/16}}{\beta\gamma\delta}\right)\cdot\omega N + \alpha\eta^{1/2}N^2 \\ &\leq \frac{3\eta^{1/16}N^2}{\beta\gamma\delta},\end{aligned}$$

lo cual es equivalente a la afirmación del lema. □

**Corolario 4.1.** *Sea $A_0 \subseteq \mathbb{Z}_N$ tal que $A_0$ es cuadráticamente $\eta$-uniforme de tamaño $\delta N$, donde $\eta \leq 2^{-208}\delta^{112}$ y $N > 200\delta^{-3}$. Entonces $A_0$ contiene al menos una progresión aritmética de longitud cuatro, o existe una subprogresión $P$ de longitud $N/5$ sobre la cual $A$ tiene densidad $9\delta/8$.*

*Demostración.* En el lema 4.6 tomamos $A = B = A_0 \cap [2N/5, 3N/5)$ y $C = D = A_0$. Tenemos entonces que

$$\sum_r |A \cap (B + r) \cap (C + 2r) \cap (D + 3r)| \geq \alpha\beta\gamma\delta N^2 - \frac{3\eta^{1/16}N^2}{\beta\gamma\delta}, \tag{4.11}$$

donde $\alpha = \beta$ y $\gamma = \delta$ son tales que $|A| = \alpha N$ y $|C| = \gamma N$. Además, podemos suponer que $|A| \geq (\delta/10)N$. Puesto que si $|A| < (\delta/10)N$, entonces alguno de los conjuntos

$$A_1 = A_0 \cap [0, N/5), \quad A_2 = A_0 \cap [N/5, 2N/5), \quad A_3 = A_0 \cap [3N/5, 4N/5) \text{ o } A_4 = A_0 \cap [4N/5, N]$$

será tal que $|A_i| \geq (9\delta/40)N$. Así, la densidad de $A$ sobre alguno de los intervalos $[0, N/5)$, $[N/5, 2N/5)$, $[3N/5, 4N/5)$ o $[4N/5, N]$ será $9\delta/8$. Supongamos que $|A| \geq (\delta/10)N$, entonces $\alpha = \beta \geq \delta/10$. Reemplazando en (4.11) las cotas sobre $\alpha, \beta$ y $\eta$ tenemos

$$\begin{aligned}\sum_r |A \cap (B + r) \cap (C + 2r) \cap (D + 3r)| &\geq \frac{\delta^4}{100}N^2 - \frac{3\cdot 2^{-13}\delta^7}{\delta^3/10}N^2 \\ &= \delta^4 N^2\left(\frac{1}{100} - \frac{30}{2^{13}}\right) \geq \frac{\delta^4 N^2}{200}.\end{aligned} \tag{4.12}$$

Por tanto existen al menos $\delta^4 N^2/200$ sucesiones de la forma $(a, a + d, a + 2d, a + 3d)$ en $A \times B \times C \times D$. De estas, a lo más $\delta N$ sucesiones pueden tener $d = 0$. Como además

$N > 200\delta^{-3}$ (esto es $\delta^3 N/200 > 1$), existe al menos una de tales sucesiones con $d \neq 0$. En efecto, de (4.12)

$$\sum_r |A \cap (B+r) \cap (C+2r) \cap (D+3r)| \quad \geq \quad \frac{\delta^4 N^2}{200} > \delta N.$$

Puesto que $a$ y $a+d$ pertenecen a $[2N/5, 3N/5)$, tenemos que $a+2d \in [N/5, 4N/5)$ y $a+3d \in [0, N)$. Así, podemos considerar a $a, a+d, a+2d$ y $a+3d$ como elementos de $\mathbb{Z}$. Esto es, la sucesión $(a, a+d, a+2d, a+3d)$ es una progresión aritmética en $\mathbb{Z}$ (y no solo en $\mathbb{Z}_N$), cuyas entradas están en $A_0$. □

Similarmente a la definición de la $U^2$-norma de Gowers para funciones conplejas, definimos la $U^3$-*norma de Gowers* de una función $f : \mathbb{Z}_N \to \mathbb{C}$, por

$$\begin{aligned} ||f||_{U^3} \quad = \quad & \left( \frac{1}{N^4} \sum_{a,b,c,d} f(a)\overline{f(a+b)}\,\overline{f(a+c)}\,\overline{f(a+d)} f(a+b+c) f(a+b+d) \times \right. \\ & \left. f(a+c+d)\overline{f(a+b+c+d)} \right)^{1/8}. \end{aligned}$$

Así

$$\begin{aligned} ||f||_{U^3}^8 \quad = \quad & \frac{1}{N^4} \sum_{a,b,c,d} f(a)\overline{f(a+b)}\,\overline{f(a+c)}\,\overline{f(a+d)} f(a+b+c) f(a+b+d) \times \\ & f(a+c+d)\overline{f(a+b+c+d)}. \end{aligned}$$

Denotando $\Delta(f,k)(n) = f(n)\overline{f(n-k)}$ y cambiando $b$ por $-b$ en la sumatoria anterior, tenemos que

$$\begin{aligned} ||f||_{U^3}^8 \quad &= \quad \frac{1}{N^4} \sum_{a,b,c,d} \Delta(f,b)(a)\overline{\Delta(f,b)(a+c)}\,\overline{\Delta(f,b)(a+d)}\Delta(f,b)(a+c+d) \\ &= \quad \frac{1}{N} \sum_b \frac{1}{N^3} \sum_{b,c,d} \Delta(f,b)(a)\overline{\Delta(f,b)(a+c)}\,\overline{\Delta(f,b)(a+d)}\Delta(f,b)(a+c+d) \\ &= \quad \frac{1}{N} \sum_b ||\Delta(f,b)||_{U^2}^4. \end{aligned} \tag{4.13}$$

Así como la $U^2$-norma de Gowers controla las progresiones aritméticas de tres términos, la $U^3$-norma de Gowers controla las progresiones aritméticas de cuatro términos. Además, de (4.13) es clara la relación entre la $U^3$-norma de Gowers pequeña y el concepto de uniformidad cuadrática. En efecto, de (4.13) y de (3.23) tenemos que

$$||f||_{U^3}^8 = \frac{1}{N^5} \sum_b \sum_\xi |\Delta(f,b)\widehat{\ }(\xi)|^4,$$

así que, $f$ es cuadráticamente $\alpha$-uniforme, esto es $\sum_b \sum_\xi |\Delta(f,b)\widehat{\ }(\xi)|^4 \leq \alpha N^5$ si, y solo si, $||f||_{U^3}^8 \leq \alpha$.

**Lema 4.7.** *Sean las funciones $f_1, f_2, f_3, f_4 : \mathbb{Z}_N \to \mathbb{C}$, con $||f_i||_\infty \leq 1$ para $i = 1,2,3,4$. Entonces*

$$\left|\sum_{a,d} f_1(a)f_2(a+d)f_3(a+2d)f_4(a+3d)\right| \leq N^2 \min_{i=1,2,3,4} ||f_i||_{U^3}.$$

*Demostración.* Sea $S = \sum_{a,d} f_1(a)f_2(a+d)f_3(a+2d)f_4(a+3d)$. Por la desigualdad de Cauchy-Schwarz tenemos

$$\begin{aligned}
S^2 &\leq N\sum_a \Big|\sum_d f_1(a)f_2(a+d)f_3(a+2d)f_4(a+3d)\Big|^2 \\
&= N\sum_a \sum_{d,e} f_1(a)\overline{f_1(a)}f_2(a+d)\overline{f_2(a+e)}f_3(a+2d)\overline{f_3(a+2e)}f_4(a+3d)\overline{f_4(a+3e)} \\
&= N\sum_a |f_1(a)|^2 \sum_{d,e} f_2(a+d)\overline{f_2(a+e)}f_3(a+2d)\overline{f_3(a+2e)}f_4(a+3d)\overline{f_4(a+3e)} \\
&\leq N\sum_a \sum_{d,e} f_2(a+d)\overline{f_2(a+e)}f_3(a+2d)\overline{f_3(a+2e)}f_4(a+3d)\overline{f_4(a+3e)} \\
&= N\sum_d \sum_{A,e} f_2(A)\overline{f_2(A+e-d)}f_3(A+d)\overline{f_3(A+2e-d)}f_4(A+2d)\overline{f_4(A+3e-d)} \\
&= N\sum_A \sum_{d,e} f_2(A)\overline{f_2(A+e-d)}f_3(A+d)\overline{f_3(A+2e-d)}f_4(A+2d)\overline{f_4(A+3e-d)} \\
&= N\sum_A \sum_{d,k} f_2(A)\overline{f_2(A+k)}f_3(A+d)\overline{f_3(A+2k+d)}f_4(A+2d)\overline{f_4(A+3k+2d)} \\
&= N\sum_A \sum_{d,k} f_2(A)\overline{f_2(A-k)}f_3(A+d)\overline{f_3(A-2k+d)}f_4(A+2d)\overline{f_4(A-3k+2d)} \\
&= N\sum_k \sum_{A,d} \Delta(f_2,k)(A)\Delta(f_3,2k)(A+d)\Delta(f_4,3k)(A+2d) \\
&= N\sum_k \sum_{a,d} \Delta(f_2,k)(a)\Delta(f_3,2k)(a+d)\Delta(f_4,3k)(a+2d).
\end{aligned}$$

Por otro lado, del lema 3.5, para cada $k$,

$$\left|\sum_{a,d} \Delta(f_2,k)(a)\Delta(f_3,2k)(a+d)\Delta(f_4,3k)(a+2d)\right| \leq N^2||\Delta(f_4,3k)||_{U^2}.$$

Por lo tanto,

$$S^2 \leq N^3 \sum_k ||\Delta(f_4,3k)||_{U^2}.$$

Asumiendo que $N$ no sea divisible ni por 2 ni por 3, tenemos que, reemplazando $3k$ por $k$,

$$S^2 \leq N^3 \sum_k ||\Delta(f_4,k)||_{U^2}.$$

Por la desigualdad de Hölder tenemos que

$$\begin{aligned}\sum_k ||\Delta(f_4,k)||_{U^2} &\leq \Big(\sum_k ||\Delta(f_4,k)||_{U^2}^4\Big)^{1/4}\Big(\sum_k 1^{4/3}\Big)^{3/4}\\ &= N^{3/4}\Big(\sum_k ||\Delta(f_4,k)||_{U^2}^4\Big)^{1/4}.\end{aligned}$$

Además, de (4.13)

$$||f_4||_{U^3}^2 = \frac{1}{N^{1/4}}\Big(\sum_k ||\Delta(f_4,k)||_{U^2}^4\Big)^{1/4},$$

y por lo tanto

$$S^2 \leq N^3 \sum_k ||\Delta(f_4,k)||_{U^2} \leq N^4 ||f_4||_{U^3}^2,$$

y así $|S| \leq N^2||f_4||_{U^3}$. Por simetría, análogamente procedemos con $f_1$, $f_2$ y $f_3$. □

**Lema 4.8.** *Sea $A \subseteq [1,N]$ con $|A| = \delta N$. Identificamos $A$ como un subconjunto de $\mathbb{Z}_N$. Sea $f$ la función balanceada de $A$. Si $||f||_{U^3} \leq \delta^4/144$ entonces se cumple una de las siguientes afirmaciones:*

*(a) Existen $\delta^4 N^2/72$ progresiones aritméticas de cuatro términos en $A$.*

*(b) Existe una subprogresión aritmética de longitud $2N/5$ sobre la cual $A$ tiene densidad mayor que $25\delta/24$.*

*Demostración.* Sean $A_1 = A_2 = A \cap [2N/5, 3N/5)$. Si $|A_1| \leq \delta N/6$, entonces

$$|A \cap [1, 2N/5)| + |A \cap [3N/5, N]| \geq 5\delta N/6,$$

y por lo tanto, $A \cap [1, 2N/5)$ o $A \cap [3N/5, N]$ tendrá cardinalidad de al menos $5\delta N/12$. Así, existirá una subprogresión $P$, a saber $[1, 2N/5)$ o $[3N/5, N]$, de longitud $2N/5$, donde

$$|A \cap P| \geq (25\delta/24)|P|.$$

Supongamos que $|A_1| = |A_2| \geq \delta N/6$. El número de progresiones aritméticas de longitud cuatro en $A$ está acotado inferiormente por

$$\begin{aligned}&\sum_{a,d} A_1(a)A_2(a+d)A(a+2d)A(a+3d)\\ &\quad= \sum_{a,d} A_1(a)A_2(a+d)A(a+2d)\big(f(a+3d)+\delta\big)\\ &\quad= \delta\sum_{a,d} A_1(a)A_2(a+d)A(a+2d) + \sum_{a,d} A_1(a)A_2(a+d)A(a+2d)f(a+3d). \quad (4.14)\end{aligned}$$

Por el lema 4.7

$$\left|\sum_{a,d} A_1(a)A_2(a+d)A(a+2d)f(a+3d)\right| \leq N^2||f||_{U^3}, \tag{4.15}$$

y además tenemos que

$$\begin{aligned}\delta\sum_{a,d} A_1(a)A_2(a+d)A(a+2d) &= \delta\sum_{a,d} A_1(a)A_2(a+d)\big(f(a+2d)+\delta\big)\\ &= \delta^2\sum_{a,d} A_1(a)A_2(a+d) + \delta\sum_{a,d} A_1(a)A_2(a+d)f(a+2d)\\ &\geq \delta^2|A_1|^2 - \delta N^2||f||_{U^2},\end{aligned} \tag{4.16}$$

donde la última desigualdad se da por el lema 3.5. Reemplazando (4.15) y (4.16) en (4.14) obtenemos

$$\sum_{a,d} A_1(a)A_2(a+d)A(a+2d)A(a+3d) \geq \delta^4N^2/36 - \delta N^2||f||_{U^2} - N^2||f||_{U^3}.$$

Como $||f||_{U^2} \leq ||f||_{U^3}$, tenemos que

$$\begin{aligned}\sum_{a,d} A_1(a)A_2(a+d)A(a+2d)A(a+3d) &\geq \delta^4N^2/36 - \delta N^2||f||_{U^3} - N^2||f||_{U^3}\\ &\geq \delta^4N^2/36 - \delta\cdot\delta^4N^2/144 - \delta^4N^2/144\\ &\geq \delta^4N^2/72.\end{aligned}$$

□

## 4.2. $U^3$-norma de Gowers grande

En lo que sigue mostraremos que si $A$ es un subconjunto de $\mathbb{Z}_N$ con cardinalidad $\delta N$, el cual no es cuadráticamente $\alpha$-uniforme, es decir, el cual tiene $U^3$-norma de Gowers grande, entonces existe una progresión aritmética $P$ en $\mathbb{Z}_N$, que más aún es una progresión aritmética considerada como un subconjunto de $\{1, 2, \ldots, N\}$, de tamaño $N^\beta$ sobre la que se da un incremento de densidad, esto es, tal que $|A\cap P| \geq (\delta+\epsilon)|P|$, donde $\beta$ y $\epsilon$ dependen únicamente de $\alpha$ y $\delta$.

### 4.2.1. Comportamiento casi lineal

Dados $A \subseteq \{1, \ldots, N\}$ y $\delta > 0$ tales que $|A| = \delta N$, consideremos la función balanceada de $A$, $f$. En las secciones anteriores vimos que si $||f||_{U^3} \leq \delta^4/144$, entonces se verifica la propiedad del incremento de densidad, o se concluye el teorema de Szemerédi para progresiones de longitud cuatro, esto es, la existencia de progresiones aritméticas de longitud cuatro en $A$. Así, solo resta analizar el caso en el que $||f||_{U^3} > \delta^4/144$.

Supongamos que $f : \mathbb{Z}_N \to [-1, 1]$ sea una función con $||f||_{U^3} \geq \alpha$. Por la definición de la $U^3$-norma de Gowers y (4.13)

$$||f||_{U^3}^8 = \frac{1}{N} \sum_k ||\Delta(f; k)||_{U^2}^4 \geq \alpha^8.$$

Puesto que para cada $k$, $||\Delta(f; k)||_{U^2} \leq 1$, existen al menos $\alpha^8 N/2$ valores de $k$ para los que $||\Delta(f; k)||_{U^2}^4 \geq \alpha^8/2$. En efecto, consideremos el conjunto

$$B = \Big\{k \in \mathbb{Z}_N;\ ||\Delta(f; k)||_{U^2}^4 \geq \alpha^8/2\Big\},$$

y supongamos que $|B| < \alpha^8 N/2$, entonces

$$\begin{aligned}
||f||_{U^3}^8 &= \frac{1}{N} \sum_k ||\Delta(f; k)||_{U^2}^4 \\
&= \frac{1}{N} \sum_{k \in B} ||\Delta(f; k)||_{U^2}^4 + \frac{1}{N} \sum_{k \notin B} ||\Delta(f; k)||_{U^2}^4 \\
&< \frac{1}{N}|B| + \frac{1}{N}|B^c|\big(\alpha^8/2\big) \\
&< \frac{1}{N}\big(\alpha^8 N/2\big) + \frac{1}{N} N\big(\alpha^8/2\big) = \alpha^8,
\end{aligned}$$

lo cual constituye una contradicción. Por tanto $|B| \geq \alpha^8 N/2$. Para cada $k \in B$ tenemos que

$$\begin{aligned}
\frac{\alpha^8}{2} &\leq ||\Delta(f; k)||_{U^2}^4 = \frac{1}{N^4} \sum_l |\Delta(f; k)\widehat{\ }(l)|^4 \\
&\leq \left(\max_l |\Delta(f; k)\widehat{\ }(l)|^2\right) \frac{1}{N^4} \sum_l |\Delta(f; k)\widehat{\ }(l)|^2 \\
&= \left(\max_l |\Delta(f; k)\widehat{\ }(l)|^2\right) \frac{1}{N^3} \sum_l |\Delta(f; k)(l)|^2 \\
&\leq \frac{1}{N^2} \max_l |\Delta(f; k)\widehat{\ }(l)|^2.
\end{aligned}$$

Por lo tanto

$$\alpha^4 N/2 \leq \max_l |\Delta(f; k)\widehat{\ }(l)|.$$

Para cada $k \in B$ elegimos $\phi(k) \in \mathbb{Z}_N$ de entre los $l \in \mathbb{Z}_N$ que cumplen $|\Delta(f;k)\widehat{\ }(l)| \geq \alpha^4 N/2$. Así,

$$\text{existe } \phi : B \to \mathbb{Z}_N, \text{ tal que } \left|\Delta(f;k)\widehat{\ }(\phi(k))\right| \geq \alpha^4 N/2. \tag{4.17}$$

En resumen, hemos mostrado la existencia de un conjunto $B \subseteq \mathbb{Z}_N$, con $|B| \geq \alpha^8 N/2$, tal que para cada $k \in B$, podemos hallar $\phi(k) \in \mathbb{Z}_N$ satisfaciendo (4.17).

Una observación crucial en la prueba de Gowers, del teorema de Szemerédi para progresiones aritméticas de longitud cuatro, es que la función $k \in B \mapsto \phi(k)$ se comporta linealmente para muchos valores de $k$. Para ello veamos la siguiente proposición.

**Proposición 4.1.** *Sean $\alpha > 0$, $B \subseteq \mathbb{Z}_N$ y $D = \{z \in \mathbb{C}; |z| \leq 1\}$. Sean las funciones $f : \mathbb{Z}_N \to D$ y $\phi : B \to \mathbb{Z}_N$ tales que*

$$\sum_{k \in B} |\Delta(f;k)\widehat{\ }(\phi(k))|^2 \geq \alpha N^3. \tag{4.18}$$

*Entonces existen al menos $\alpha^4 N^3$ cuádruplas $(a,b,c,d) \in B^4$ tales que*

$$a + b = c + d \quad y \quad \phi(a) + \phi(b) = \phi(c) + \phi(d).$$

**Lema 4.9.** *Dados $f, g : \mathbb{Z}_N \to \mathbb{C}$, definimos $(f \star g)(s) = \sum_r f(r)g(s-r)$. Tenemos que*

$$(f \star g)\widehat{\ }(s) = \widehat{f}(s)\widehat{g}(s).$$

*Demostración.* En efecto,

$$\begin{aligned}
(f \star g)\widehat{\ }(s) &= \sum_r (f \star g)(r)\omega^{-rs} \\
&= \sum_r \sum_t f(t)g(r-t)\omega^{-rs} \\
&= \sum_r \sum_t f(t)\omega^{-st} g(r-t)\omega^{-s(r-t)} \\
&= \sum_t f(t)\omega^{-st} \sum_r g(r-t)\omega^{-s(r-t)} \\
&= \sum_t f(t)\omega^{-st} \sum_u g(u)\omega^{-su} = \widehat{f}(s)\widehat{g}(s).
\end{aligned}$$

□

*Prueba de la proposición 4.1.* Expandiendo el lado izquierdo de la desigualdad (4.18).

$$\sum_{k \in B} \sum_{s,t} f(s)\overline{f(s-k)f(t)}f(t-k)\omega^{-\phi(k)(s-t)} \geq \alpha N^3.$$

Si introducimos la variable $u = s - t$ podemos reescribir esto como

$$\sum_{k\in B}\sum_{s,u} f(s)\overline{f(s-k)(s-u)}f(s-u-k)\omega^{-\phi(k)u} \geq \alpha N^3.$$

Como $|f(s)| \leq 1$ para cada $s$, se sigue que

$$\sum_{u}\sum_{s}\left|\sum_{k\in B}\overline{f(s-k)}f(s-u-k)\omega^{-\phi(k)u}\right| \geq \alpha N^3,$$

lo cual implica, por la desigualdad de Cauchy-Schwarz

$$\left(\sum_{u}\sum_{s}\left|\sum_{k\in B}\overline{f(s-k)}f(s-u-k)\omega^{-\phi(k)u}\right|^2\right)^{1/2}\left(N^2\right)^{1/2} \geq \alpha N^3,$$

$$\sum_{u}\sum_{s}\left|\sum_{k\in B}\overline{f(s-k)}f(s-u-k)\omega^{-\phi(k)u}\right|^2 \geq \alpha^2 N^4. \tag{4.19}$$

Para $u$ fijo, sea $\gamma(u)$ definido por la ecuación

$$\sum_{s}\left|\sum_{k\in B}\overline{f(s-k)}f(s-u-k)\omega^{-\phi(k)u}\right|^2 = \gamma(u)N^3,$$

esto es,

$$\sum_{s}\left|\sum_{k}\Delta(f;u)(s-k)B(k)\omega^{\phi(k)u}\right|^2 = \gamma(u)N^3,$$

$$\sum_{s}\left|\sum_{k}F(k)G(s-k)\right|^2 = \gamma(u)N^3,$$

donde $F(t) = B(t)\omega^{\phi(t)u}$ y $G(t) = \Delta(f;u)(t)$. Así,

$$\begin{aligned}
\gamma(u)N^3 &= \sum_{s}\left|\sum_{k}F(k)G(s-k)\right|^2 = \sum_{s}\left|(F\star G)(s)\right|^2 \\
&= N^{-1}\sum_{r}\left|(F\star G)\widehat{\ }(r)\right|^2 = N^{-1}\sum_{r}\left|\widehat{F}(r)\widehat{G}(r)\right|^2 \\
&\leq N^{-1}\left(\sum_{r}\left|\widehat{F}(r)\right|^4\right)^{1/2}\left(\sum_{r}\left|\widehat{G}(r)\right|^4\right)^{1/2} \\
&\leq N^{-1}\left(\sum_{r}\left|\widehat{F}(r)\right|^4\right)^{1/2}\left(\sum_{r}\left|\widehat{G}(r)\right|^2\right) \\
&\leq \left(\sum_{r}\left|\widehat{F}(r)\right|^4\right)^{1/2}\left(\sum_{r}\left|G(r)\right|^2\right) \\
&\leq N\left(\sum_{r}\left|\widehat{F}(r)\right|^4\right)^{1/2}.
\end{aligned}$$

Por lo tanto, $\gamma(u)^2N^4 \leq \sum_r \left|\widehat{F}(r)\right|^4$, esto es,

$$\gamma(u)^2N^4 \leq \sum_r \left|\sum_{k\in B} \omega^{\phi(k)u-rk}\right|^4. \tag{4.20}$$

la desigualdad (4.19) implica que $\sum_u \gamma(u) \geq \alpha^2 N$, y por la desigualdad de Cauchy-Schwarz $\sum_u \gamma(u)^2 \geq \alpha^4 N$. Por consiguiente, de la desigualdad (4.20), sumando sobre $u$, tenemos que

$$\sum_u \sum_r \left|\sum_{k\in B} \omega^{\phi(k)u-rk}\right|^4 \geq \alpha^4 N^5.$$

Desarrollando el lado izquierdo tenemos

$$\sum_{u,r} \left|\sum_{a\in B} \omega^{\phi(a)u-ra} \overline{\sum_{c\in B} \omega^{\phi(c)u-rc}}\right|^2$$

$$\sum_{u,r} \left|\sum_{a,c\in B} \omega^{\left(\phi(a)-\phi(c)\right)u-r(a-c)}\right|^2$$

$$\sum_{u,r} \sum_{a,c\in B} \omega^{\left(\phi(a)-\phi(c)\right)u-r(a-c)} \overline{\sum_{d,b\in B} \omega^{\left(\phi(d)-\phi(b)\right)u-r(d-b)}}$$

$$\sum_{u,r} \sum_{a,b,c,d} \omega^{u\left(\phi(a)+\phi(b)-\phi(c)-\phi(d)\right)-r(a+b-c-d)}.$$

por ende,

$$\sum_{u,r} \sum_{a,b,c,d} \omega^{u\left(\phi(a)+\phi(b)-\phi(c)-\phi(d)\right)} \omega^{-r(a+b-c-d)} \geq \alpha^4 N^5.$$

Esto es,

$$\sum_{\substack{a,b,c,d \\ a+b=c+d \\ \phi(a)+\phi(b)=\phi(c)+\phi(d)}} 1 \geq \alpha^4 N^3.$$

□

**Proposición 4.2.** *Sean $f : \mathbb{Z}_N \to [-1,1]$ una función y $\alpha > 0$ tales que $||f||_{U^3} \geq \alpha$, y supongamos que $\phi : B \subseteq \mathbb{Z}_N \mapsto \mathbb{Z}_N$ es definido como en (4.17). Entonces existen al menos $\alpha^{64}2^{-12}N^3$ cuádruplas $(b_1, b_2, b_3, b_4) \in B^4$, tales que*

$$b_1 + b_2 = b_3 + b_4 \quad y \quad \phi(b_1) + \phi(b_2) = \phi(b_3) + \phi(b_4).$$

*Demostración.* De (4.17), de la definición de $B$ y puesto que $|B| \geq \alpha^8 N/2$, tenemos que

$$\begin{aligned} \sum_{k\in B} |\Delta(f;k)\widehat{\ }(\phi(k))|^2 &\geq \sum_{k\in B} (\alpha^4 N/2)^2 = |B|(\alpha^4 N/2)^2 \\ &= \left(\frac{\alpha^{16}}{8}\right) N^3. \end{aligned}$$

De la proposición 4.1, considerando el $\alpha$ que aparece allí como $\alpha^{16}/8$, tenemos que existen al menos $(\alpha^{16}/8)^4 N^3 = \alpha^{64} 2^{-12} N^3$ cuádruplas $(b_1, b_2, b_3, b_4) \in B^4$, tales que $b_1 + b_2 = b_3 + b_4$ y $\phi(b_1) + \phi(b_2) = \phi(b_3) + \phi(b_4)$. □

Llamaremos a las cuádruplas $(b_1, b_2, b_3, b_4) \in B^4$ que cumplan las siguientes condiciones

$$b_1 + b_2 = b_3 + b_4 \quad \text{y} \quad \phi(b_1) + \phi(b_2) = \phi(b_3) + \phi(b_4),$$

como *cuádruplas aditivas*.

### 4.2.2. Cuádruplas aditivas y el teorema de Balog-Szemerédi-Gowers

En el caso de progresiones aritméticas de longitud 4 requerimos de más herramientas. En este parágrafo mostraremos un resultado sobre grafos.

Un grafo $G = G(V, E)$ (no dirigido) consiste de un conjunto finito $V$ de *vértices* y un conjunto finito $E$ de *lados* o *aristas*, donde cada lado es un par $(a, b)$ de vértices.
Si $(a, b) \in E$, decimos que los dos vértices $a$ y $b$ son *adyacentes* o *vecinos*. La colección de todos los vértices adyacentes a $a$ será denotada por $N(a)$. La cardinalidad de $N(a)$ es llamada *el grado del vértice* $a$, y será denotada por $\deg(a)$.
Dado un subconjunto $V'$ de $V$. Consideremos el subgrafo de $G$ generado por $V'$, $G' = G'(V', E')$, donde $E' = \{e \in E;\ e \in V' \times V'\}$.
Un grafo es *bipartito* si se puede particionar su conjunto de vértices en dos conjuntos disjuntos $A$ y $B$ tal que cada lado tiene un vértice en $A$ y el otro en $B$.

Necesitaremos una variante del teorema de Balog-Szemerédi, debida a Gowers. Sean $A$ y $B$ dos conjuntos de $\mathbb{Z}$, con $|A| = |B|$, y $|A + B| \leq C|A|$; denotemos por $r_{A+B}(n)$ al número de maneras de escribir $n$ como $a + b$ con $a \in A$ y $b \in B$, esto es,

$$r_{A+B}(n) = \big|\big\{(a, b) \in A \times B;\ a + b = n\big\}\big|. \tag{4.21}$$

Notemos que

$$\sum_n r_{A+B}(n) = |A||B|, \tag{4.22}$$

y también que $|A + B| = \big|\{n \in \mathbb{Z};\ r_{A+B}(n) \neq 0\}\big|$.

De (4.22), por la desigualdad de Cauchy-Schwarz, tenemos que

$$|A||B| = \sum_{n,\, r_{A+B}(n)\neq 0} r_{A+B}(n) \leq \left( \sum_{n,\, r_{A+B}(n)\neq 0} r_{A+B}(n)^2 \right)^{1/2} \left( \sum_{n,\, r_{A+B}(n)\neq 0} 1^2 \right)^{1/2},$$

por lo tanto

$$\begin{aligned} |A||B| &\leq |A+B|^{1/2} \left( \sum_{n,\, r_{A+B}(n)\neq 0} r_{A+B}(n)^2 \right)^{1/2} \\ |A|^4/|A+B| &\leq \sum_{n,\, r_{A+B}(n)\neq 0} r_{A+B}(n)^2. \end{aligned}$$

Como $|A+B| \leq C|A|$, entonces

$$\begin{aligned} |A|^3/C &\leq \sum_{n,\, r_{A+B}(n)\neq 0} r_{A+B}(n)^2 \\ |A|^3/C &\leq \sum_{n} r_{A+B}(n)^2. \end{aligned}$$

Pero $\sum_{n\in\mathbb{Z}} r_{A+B}(n)^2$ es el número de *cuádruplas aditivas* $(a_1, b_1, a_2, b_2) \in A \times B \times A \times B$ con $a_1 + b_1 = a_2 + b_2$. Así, tenemos que el hecho que $A + B$ sea *"pequeño"* implica la existencia de muchas cuádruplas aditivas. La variante de Gowers es una especie de recíproca de la afirmación anterior. Asume la existencia de muchas cuádruplas aditivas y entonces determina la existencia de subconjuntos $A' \subseteq A$ y $B' \subseteq B$ cuyo conjunto suma $A' + B'$ es pequeño.

Formulamos dos versiones equivalentes del teorema de Balog-Szemerédi-Gowers.

**Teorema 4.1** (Balog-Szemerédi-Gowers, 1$^{\text{era}}$ versión)**.** *Sean $A$ y $B$ dos subconjuntos de un grupo abeliano $(H, +)$, tales que $|A| = |B|$. Supongamos que existen al menos $\alpha|A|^3$ cuádruplas aditivas $(a_1, b_1, a_2, b_2) \in A \times B \times A \times B$ con $a_1 + b_1 = a_2 + b_2$. Entonces existen subconjuntos $A' \subseteq A$ y $B' \subseteq B$ tales que*

$$|A'| \geq \alpha^2|A|/(16\sqrt{2}), \quad |B'| \geq \alpha^2|B|/16 \quad y \quad |A' + B'| \leq 2^{28}\alpha^{-13}|A|.$$

**Teorema 4.2** (Balog-Szemerédi-Gowers, 2$^{\text{da}}$ versión)**.** *Sean $A$ y $B$ dos subconjuntos de un grupo abeliano $(H, +)$, con $|A| = |B|$. Sea $G$ un subgrafo del grafo bipartito completo formado entre $A$ y $B$, tal que $G$ tiene al menos $|A||B|/K$ aristas. Supongamos que $A +_G B = \{a + b;\ (a, b) \in G\}$ es tal que $|A +_G B| \leq K_1|A|$. Entonces existen subconjuntos $A' \subseteq A$ y $B' \subseteq B$ verificando*

$$|A'| \geq |A|/(4\sqrt{2}K), \quad |B'| \geq |B|/(4K) \quad y \quad |A' + B'| \leq 2^{15}K^5K_1^3|A|.$$

Mostremos la equivalencia de las dos versiones del teorema de Balog-Szemerédi-Gowers.

*Teorema 4.2 implica teorema 4.1.* Supongamos que son dados $A$ y $B$ subconjuntos de un grupo abeliano $(H,+)$, con $|A| = |B|$, tales que existen al menos $\alpha|A|^3$ cuádruplas aditivas $(a_1, b_1, a_2, b_2)$ en $A \times B \times A \times B$ con $a_1 + b_1 = a_2 + b_2$. Consideremos para cada $n \in H$, $r_{A+B}(n)$, como en (4.21). Entonces tenemos que $\sum_{n\in H} r_{A+B}(n)^2 \geq \alpha|A|^3$. Por lo tanto existen al menos $\alpha|A|/2$ valores de $n$ en el grupo tales que $r_{A+B}(n) \geq \alpha|A|/2$. En efecto, definamos el conjunto

$$J = \{n \in H;\ r_{A+B}(n) \geq \alpha|A|/2\}.$$

Tenemos que

$$\begin{aligned}\sum_{n\in H} r_{A+B}(n)^2 &= \sum_{n\in J} r_{A+B}(n)^2 + \sum_{n\notin J} r_{A+B}(n)^2 \\ &< \sum_{n\in J} r_{A+B}(n)^2 + (\alpha|A|/2)\sum_{n\notin J} r_{A+B}(n) \\ &\leq \sum_{n\in J} r_{A+B}(n)^2 + (\alpha|A|/2)\cdot|A||B|.\end{aligned}$$

Por tanto,

$$\alpha|A|^3/2 \leq \sum_{n\in J} r_{A+B}(n)^2 \leq |J|\max_n\big(r_{A+B}(n)^2\big).$$

Si $|J| < \alpha|A|/2$, entonces $|A|^2 < \max_n\big(r_{A+B}(n)^2\big)$, esto es, $|A| < \max_n\big(r_{A+B}(n)\big)$, lo cual es una contradicción. Por tanto $|J| \geq \alpha|A|/2$. Definamos el subgrafo $G$ del grafo bipartito completo formado uniendo vértices entre $A$ y $B$,

$$G = \big\{(a,b) \in A \times B;\ \ a+b \in J\big\}.$$

El número de aristas de $G$ es al menos $\alpha^2|A|^2/4$. En efecto,

$$\sum_{n\in J} r_{A+B}(n) \geq \sum_{n\in J} \alpha|A|/2 = |J|(\alpha|A|/2) \geq \alpha^2|A|^2/4.$$

Tenemos también que $|J| \leq 2|A|/\alpha$. En efecto, como

$$|A|^2 = |A||B| = \sum_n r_{A+B}(n) \geq \sum_{n\in J} r_{A+B}(n) \geq |J|(\alpha|A|/2),$$

concluimos que $2|A|/\alpha \geq |J|$. Por ende $|A +_G B| \leq 2|A|/\alpha$. Estando en las hipótesis del teorema 4.2, existen $A' \subseteq A$ y $B' \subseteq B$ tales que

$$|A'| \geq |A|/(4\sqrt{2}K), \quad |B'| \geq |B|/(4K) \quad \text{y} \quad |A'+B'| \leq 2^{15}K^5K_1^3|A|,$$

donde $K = 4/\alpha^2$ y $K_1 = 2/\alpha$. Con esto obtenemos el teorema 4.1. □

*Teorema 4.1 implica teorema 4.2.* Sean $A$ y $B$ subconjuntos de un grupo abeliano $(H,+)$, con $|A| = |B|$. Sea $G$ un subgrafo del grafo bipartito completo formado entre $A$ y $B$, tal que $G$ tenga al menos $|A||B|/K$ aristas y

$$|A +_G B| \leq K_1|A|, \text{ donde } A +_G B = \{a+b \in H;\ (a,b) \in G\},$$

como en las hipótesis del teorema 4.2. Entonces

$$\sum_{n \in H} r_{A+_G B}(n) = |G|, \text{ donde } r_{A+_G B}(n) = \big|\big\{(a,b) \in G;\ a+b=n\big\}\big|.$$

Por la desigualdad de Cauchy-Schwarz,

$$|G| \leq \left( \sum_{n,\, r_{A+_G B}(n) \neq 0} r_{A+_G B}(n)^2 \right)^{1/2} \left( \sum_{n,\, r_{A+_G B}(n) \neq 0} 1^2 \right)^{1/2},$$

así,

$$\sum_{n,\, r_{A+_G B}(n) \neq 0} r_{A+_G B}(n)^2 \geq |G|^2/|A +_G B|.$$

Mas por hipótesis, $|A +_G B| \leq K_1|A|$ y $|G| \geq |A||B|/K = |A|^2/K$. Entonces existen al menos $|G|^2/|A +_G B| \geq \big(1/(K^2K_1)\big)|A|^3$ cuádruplas aditivas. Tomemos $\alpha = 1/(K^2K_1)$. Estando en las hipótesis del teorema 4.1, existen subconjuntos $A' \subseteq A$ y $B' \subseteq B$ tales que

$$|A'| \geq \alpha^2|A|/(16\sqrt{2}), \quad |B'| \geq \alpha^2|B|/16 \quad \text{y} \quad |A'+B'| \leq 2^{28}\alpha^{-13}|A|.$$

Esto es,

$$|A'| \geq |A|/(4\sqrt{2}[2K^2K_1]^2), \quad |B'| \geq |B|/(4[2K^2K_1]^2)$$

$$\text{y} \quad |A'+B'| \leq 2^{15}([2K^2K_1]^2)^5[2K^2K_1]^3|A|.$$

□

A continuación probaremos el teorema 4.2, y por lo tanto el teorema 4.1.

**Lema 4.10.** *Sean $A$ y $B$ subconjuntos de un grupo abeliano. Sea $G$ un grafo bipartito no dirigido teniendo como conjuntos de vértices $A$ y $B$ (esto es, las aristas de $G$ conectan puntos de $A$ con puntos de $B$). Supongamos que el conjunto de aristas de $G$ tenga cardinalidad $|A||B|/K$, para algún $K \geq 1$. Entonces, dado $\epsilon \in (0,1)$, existe un subconjunto $A' \subseteq A$, con $|A'| \geq |A|/(\sqrt{2}K)$, tal que para al menos una fracción $(1-\epsilon)$ de los pares $(a_1,a_2) \in A' \times A'$ tenemos por lo menos $\epsilon|B|/(2K^2)$ caminos de longitud dos en $G$ uniendo $a_1$ y $a_2$.*

*Demostración.* Para cada $a \in A$, definamos

$$B(a) = \{\beta \in B;\ (a, \beta) \in G\}.$$

Análogamente, para cada $b \in B$,

$$A(b) = \{\alpha \in A;\ (\alpha, b) \in G\}.$$

Sean $\epsilon \in ]0, 1[$, y

$$\Omega = \big\{(a_1, a_2) \in A \times A;\ |B(a_1) \cap B(a_2)| \leq \epsilon |B|/(2K^2)\big\}.$$

Tenemos que $\displaystyle\sum_{b\in B} |A(b)| = |A||B|/K$, puesto que el grafo $G$ es bipartito. Por la desigualdad de Cauchy-Schwarz tenemos

$$\sum_{b\in B} |A(b)| \leq \left(\sum_{b\in B} |A(b)|^2\right)^{1/2} \left(\sum_{b\in B} 1^2\right)^{1/2},$$

por tanto

$$\sum_{b\in B} |A(b)|^2 \geq \frac{1}{|B|}\left(\sum_{b\in B} |A(b)|\right)^2 \geq |A|^2|B|/K^2. \tag{4.23}$$

Además, como $x \in B(y)$ si, y sólo si, $y \in A(x)$, tenemos que

$$\begin{aligned}
\sum_{b\in B} \sum_{\substack{a_1,a_2\in A(b)\\(a_1,a_2)\in\Omega}} 1 &= \sum_{b\in B} \sum_{\substack{(a_1,a_2)\in\Omega\\ b\in B(a_1)\cap B(a_2)}} 1 \\
&= \sum_{(a_1,a_2)\in\Omega} \sum_{\substack{b\in B\\ b\in B(a_1)\cap B(a_2)}} 1 \\
&= \sum_{(a_1,a_2)\in\Omega} \sum_{b\in B(a_1)\cap B(a_2)} 1 \\
&= \sum_{(a_1,a_2)\in\Omega} |B(a_1) \cap B(a_2)| \\
&\leq \sum_{(a_1,a_2)\in\Omega} \epsilon|B|/(2K^2) = |\Omega|\epsilon|B|/(2K^2).
\end{aligned}$$

Como $|\Omega| \leq |A|^2$,

$$\sum_{b\in B} \frac{1}{\epsilon}|A(b)^2 \cap \Omega| \leq |A|^2|B|/(2K^2). \tag{4.24}$$

De (4.23) y (4.24), tenemos que

$$\sum_{b\in B} \big(|A(b)|^2 - \frac{1}{\epsilon}|A(b)^2 \cap \Omega|\big) \geq |A|^2|B|/(K^2) - |A|^2|B|/(2K^2) = |A|^2|B|/(2K^2),$$

y

$$\frac{1}{|B|}\sum_{b\in B}\big(|A(b)|^2-\frac{1}{\epsilon}|A(b)^2\cap\Omega|\big)\geq |A|^2/(2K^2).$$

Así, existe algún $b\in B$ (que fijamos) tal que

$$\big(|A(b)|^2-\frac{1}{\epsilon}|A(b)^2\cap\Omega|\big)\geq |A|^2/(2K^2).$$

Tomando $A'=A(b)$ se sigue el lema. En efecto, tenemos que

$$|A'\times A'|-\frac{1}{\epsilon}|(A'\times A')\cap\Omega|\geq |A|^2/(2K^2). \tag{4.25}$$

Por contradicción, supongamos que $|(A'\times A')\cap\Omega^c|<(1-\epsilon)|A'\times A'|$. Como

$$|A'\times A'|=|(A'\times A')\cap\Omega|+|(A'\times A')\cap\Omega^c|,$$

tenemos que

$$\begin{aligned}|A'\times A'| &< |(A'\times A')\cap\Omega|+(1-\epsilon)|A'\times A'|\\ \epsilon|A'\times A'| &< |(A'\times A')\cap\Omega|.\end{aligned}$$

Por lo tanto

$$|A'\times A'|-\frac{1}{\epsilon}|(A'\times A')\cap\Omega|<0,$$

lo cual es una contradicción con (4.25). Así,

$$|(A'\times A')\cap\Omega^c|\geq(1-\epsilon)|A'\times A'|.$$

Esto es, al menos una fracción $(1-\epsilon)$ de pares $(a_1,a_2)\in A'\times A'$ son tales que los caminos de longitud dos en $G$ uniendo $a_1$ con $a_2$ son al menos $\epsilon|B|/(2K^2)$.
Por otro lado, también se tiene de (4.25) que

$$|A'|^2\geq|A'\times A'|-\frac{1}{\epsilon}|(A'\times A')\cap\Omega|\geq|A|^2/(2K^2),$$

y por lo tanto $|A'|\geq|A|/(\sqrt{2}K)$. □

**Lema 4.11.** *Sean $A$ y $B$ subconjuntos de un grupo abeliano. Sea $G$ un grafo bipartito no dirigido teniendo dos conjuntos de vértices, $A$ y $B$. Supongamos que $G$ tenga una cantidad de lados igual a $|A||B|/K$ (esto es, $|G|=|A||B|/K$). Entonces existen $A_1''$ y $A'$ incluidos en $A$ tales que $A_1''\subseteq A'\subseteq A$ y $|A_1''|\geq|A|/(4\sqrt{2}K)$. Además cada vértice en $A'$ tiene al menos grado $|B|/(2K)$, y para cada $a_1\in A_1''$ existen al menos $\big(1-1/(16K)\big)|A'|$ vértices $a_2$ en $A'$*

*tales que $a_1$ y $a_2$ son unidos por al menos $|B|/(256K^3)$ caminos de longitud dos. En términos de conjuntos, para cada $a_1 \in A_1''$*

$$\big|(\{a_1\} \times A') \cap \Omega^c\big| \geq \left(1 - \frac{1}{16K}\right)|A'|,$$

*donde* $\Omega = \left\{(a, a') \in A' \times A';\ |B(a) \cap B(a')| < \frac{|B|}{256K^3}\right\}$.

*Demostración.* Suprimimos de $A$ todos los vértices con grado menor que $|B|/(2K)$. Sea $\tilde{A}$ el conjunto de vértices restantes,

$$\tilde{A} = \{a \in A;\ \text{ el grado de } a \text{ es } > |B|/(2K)\}.$$

Consideremos el subgrafo inducido por $G$ sobre los conjuntos de vértices $\tilde{A}$ y $B$. Puesto que a lo más $|A||B|/(2K)$ lados del grafo $G$ son removidos (puesto que los vértices en $A \setminus \tilde{A}$ tienen grado no mayor que $|B|/(2K)$), el subgrafo inducido tiene al menos tantos lados como

$$|A||B|/K - |A||B|/(2K) = |A||B|/(2K).$$

En particular, $|\tilde{A}||B| \geq |A||B|/(2K)$, y por lo tanto $|\tilde{A}| \geq |A|/(2K)$. El número $K_1 = 2K$ hará el papel de $K$ en el lema anterior.
Del lema 4.10, con $\epsilon = 1/(32K)$, hallamos un subconjunto $A'$ de $\tilde{A}$ con

$$|A'| \geq |\tilde{A}|/(\sqrt{2}K_1) = |A|/(2\sqrt{2}K),$$

tal que al menos una fracción $1 - 1/(32K)$ de los pares $(a_1, a_2) \in A' \times A'$ tienen al menos $\epsilon|B|/(2K_1^2) = |B|/(256K^3)$ caminos de longitud dos conectando $a_1$ y $a_2$. Esto es,

$$|(A' \times A') \cap \Omega^c| \geq \big(1 - 1/(32K)\big)|A' \times A'|,$$

que es equivalente a

$$\frac{1}{32K}|A' \times A'| \geq |(A' \times A') \cap \Omega|, \tag{4.26}$$

donde

$$\Omega = \left\{(a_1, a_2) \in A' \times A';\ |B(a_1) \cap B(a_2)| \leq \frac{|B|}{256K^3}\right\}.$$

Definamos

$$A_1'' = \{a_1 \in A';\ |(\{a_1\} \times A') \cap \Omega| \leq \frac{1}{16K}|A'|\} \quad \text{y} \quad A_2'' = A' \setminus A_1''.$$

De (4.26), tenemos que

$$\begin{aligned}\frac{1}{32K}|A'| &\geq \frac{1}{|A'|}\sum_{a_1\in A'}|(\{a_1\}\times A')\cap\Omega| \\ &= \frac{1}{|A'|}\left[\sum_{a_1\in A_1''}|(\{a_1\}\times A')\cap\Omega| + \sum_{a_1\in A_2''}|(\{a_1\}\times A')\cap\Omega|\right] \\ &> \frac{1}{|A'|}\left[\sum_{a_1\in A_1''}|(\{a_1\}\times A')\cap\Omega| + \sum_{a_1\in A_2''}\frac{1}{16K}\cdot|A'|\right] \\ &= \frac{1}{|A'|}\left[\sum_{a_1\in A_1''}|(\{a_1\}\times A')\cap\Omega| + |A_2''|\cdot\frac{1}{16K}\cdot|A'|\right].\end{aligned}$$

Supongamos que $|A_2''| > |A'|/2$, entonces

$$\frac{1}{32K}|A'| > \frac{1}{|A'|}\sum_{a_1\in A_1''}|(\{a_1\}\times A')\cap\Omega| + \frac{1}{32K}\cdot|A'|,$$

y así, $0 > \dfrac{1}{|A'|}\displaystyle\sum_{a_1\in A_1''}|(\{a_1\}\times A')\cap\Omega|$, lo cual es una contradicción. Por lo tanto $|A_2''| \leq |A'|/2$ y $|A_1''| > |A'|/2$, y así $|A_1''| > |A|/(4\sqrt{2}K)$. □

**Lema 4.12.** *Sean $A$ y $B$ subconjuntos de un grupo abeliano. Sea $G$ un grafo bipartito no dirigido teniendo dos conjuntos de vértices, $A$ y $B$. Supongamos que $|G| = |A||B|/K$. Podemos hallar subconjuntos $A'' \subseteq A$ y $B' \subseteq B$, con $|A''| \geq |A|/(4\sqrt{2}K)$ y $|B'| \geq |B|/(4K)$, tal que para cada $a \in A''$ y cada $b \in B'$ existen al menos $|A||B|/(2^{15}K^5)$ caminos de longitud tres uniendo $a$ y $b$.*

*Demostración.* Como $|G| = |A||B|/K$, del lema 4.11 tenemos que existen conjuntos $A''$ y $A'$, con $A'' \subseteq A' \subseteq A$, tales que $|A''| \geq |A|/(4\sqrt{2}K)$, cada vértice en $A'$ tiene grado al menos $|B|/(2K)$, y para cada $a \in A''$, tenemos que

$$\big|(\{a\}\times A')\cap\Omega^c\big| \geq \left(1-\frac{1}{16K}\right)|A'|,$$

que es equivalente a,

$$\frac{1}{16K}|A'| \geq \big|(\{a\}\times A')\cap\Omega\big|, \tag{4.27}$$

donde $\Omega = \big\{(a_1,a_2)\in A'\times A';\ \big|B(a_1)\cap B(a_2)\big| \leq |B|/(256K^3)\big\}$. Ahora debemos hallar $B' \subseteq B$. Tomaremos $B'$ como el conjunto de vértices en $B$ adyacentes a al menos $|A'|/(8K)$ elementos de $A'$. Esto es,

$$B' = \big\{b\in B;\ |A(b)\cap A'| \geq |A'|/(8K)\big\}.$$

Notemos que el número de lados de $G$ uniendo vértices de $A'$ con vértice en $B$ es al menos $|A'||B|/(2K)$, esto es

$$\left|\left\{(a,b) \in G;\ a \in A',\ b \in B\right\}\right| = \sum_{a \in A'} |B(a)| \geq \frac{|A'||B|}{2K} \tag{4.28}$$

puesto que cada vértice en $A'$ tiene grado al menos $|B|/(2K)$, y el grafo $G$ es bipartito. Entonces, al menos $|B|/(4K)$ vértices de $B$ deben estar conectados a al menos $|A'|/(8K)$ vértices de $A'$, esto es, $|B'| \geq |B|/(4K)$. En efecto, de (4.28),

$$\frac{|A'||B|}{2K} \leq \left|\left\{(a,b) \in G;\ a \in A',\ b \in B\right\}\right| = \sum_{b \in B} |A(b) \cap A'|,$$

y luego $\dfrac{|A'|}{2K} \leq \dfrac{1}{|B|} \displaystyle\sum_{b \in B} |A(b) \cap A'|$. Sea $B'_1 = B \setminus B'$; entonces

$$\begin{aligned}
\frac{|A'|}{2K} &\leq \frac{1}{|B|}\left[\sum_{b \in B'} |A(b) \cap A'| + \sum_{b \in B'_1} |A(b) \cap A'|\right] \\
&< \frac{1}{|B|}\left[\sum_{b \in B'} |A(b) \cap A'| + |B'_1| \cdot \frac{|A'|}{8K}\right] \\
&\leq \frac{1}{|B|}\left[|A'||B'| + |B'_1| \cdot \frac{|A'|}{8K}\right].
\end{aligned}$$

Si $|B'| < \dfrac{|B|}{4K}$, entonces

$$\frac{|A'|}{2K} < \frac{1}{|B|}\left[|A'| \cdot \frac{|B|}{4K} + |B'_1| \cdot \frac{|A'|}{8K}\right] \leq \frac{3|A'|}{8K},$$

lo cual es una contradicción. Por lo tanto $|B'| \geq |B|/(4K)$.

Sean $a \in A''$ y $b \in B'$, entonces, como $b \in B'$,

$$|A(b) \cap A'| \geq |A'|/(8K); \tag{4.29}$$

esto es, tenemos al menos $|A'|/(8K)$ vértices en $A'$ que son adyacentes a $b$. Y como $a \in A''$, de (4.27),

$$\frac{|A'|}{16K} \geq |(\{a\} \times A') \cap \Omega|, \tag{4.30}$$

esto es, a lo más la mitad, $|A'|/(16K)$, de los vértices en $A'$ que son adyacentes a $b$, pueden tener la propiedad de que existan menos que $|B|/(256K^3)$ caminos de longitud dos en $G$ conectando ellos con $a$. Así, existen al menos $|A'|/(16K)$ vértices $a_2$, la otra mitad, que son

adyacentes a $b$, tales que hay al menos $|B|/(256K^3)$ caminos de longitud dos uniendo $a$ y $a_2$. Por lo tanto, existen al menos

$$\frac{|A'|}{16K}\cdot\frac{|B|}{256K^3}\geq\frac{|A''|}{16K}\cdot\frac{|B|}{256K^3}\geq\frac{\big(|A|/(8K)\big)}{16K}\cdot\frac{|B|}{256K^3}=\frac{|A||B|}{2^{15}K^5}$$

caminos de longitud tres uniendo $a$ y $b$.

En términos de desigualdades tenemos de (4.29) que

$$\big|\{a\}\times\big(A(b)\cap A'\big)\big|\geq\frac{|A'|}{8K},$$

y de (4.30),

$$\Big|\big[\{a\}\times\big(A(b)\cap A'\big)\big]\cap\Omega\Big|\leq\frac{|A'|}{16K}.$$

Por lo tanto

$$\Big|\big[\{a\}\times\big(A(b)\cap A'\big)\big]\cap\Omega^c\Big|\geq\frac{|A'|}{8K}-\frac{|A'|}{16K}=\frac{|A'|}{16K}.$$

□

*Demostración del teorema 4.2.* Por el lema 4.12 podemos hallar subconjuntos $A'\subseteq A$ y $B'\subseteq B$, con $|A'|\geq|A|/(4\sqrt{2}K)$ y $|B'|\geq|B|/(4K)$, tales que existen al menos $|A||B|/(2^{15}K^5)$ caminos de longitud tres conectando cualesquiera dos elementos $a\in A'$ y $b\in B'$. Así, dados $a\in A'$ y $b\in B'$, podemos hallar no menos que $|A||B|/(2^{15}K^5)$ pares $(b_1,a_2)\in B\times A$, con $(a,b_1)$, $(b_1,a_2)$ y $(a_2,b)$ siendo lados del grafo $G$. Esto es, $a+b_1=x$, $b_1+a_2=y$ y $a_2+b=z$, son elementos de $A+_G B$. Notemos que

$$a+b=a+b_1-(b_1+a_2)+(a_2+b)=x-y+z. \tag{4.31}$$

Definamos la función

$$\Psi:\ T\to A'+B' \text{ por } \Psi(x,y,z)=x-y+z,$$

donde

$$\begin{aligned}T\ \ =\ \ \Big\{&(x,y,z)\in(A+_G B)^3;\ x=a+b,\ \text{con } (a,b)\in G \text{ y } a\in A';\\ &z=c+d,\ \text{con } (c,d)\in G \text{ y } d\in B'\Big\}.\end{aligned}$$

De (4.31), $\psi$ es sobreyectiva. Por lo tanto $T=\bigcup_{r\in A'+B'}\Psi^{-1}(r)$. Por lo hecho arriba, para cada $r\in A'+B'$, $|\Psi^{-1}(r)|\geq\dfrac{|A||B|}{2^{15}K^5}$. Así

$$|T|=\sum_{r\in A'+B'}|\Psi^{-1}(r)|\geq|A'+B'|\cdot\frac{|A||B|}{2^{15}K^5}.$$

Como $|T| \leq |A +_G B|^3$, tenemos que

$$|A' + B'| \leq 2^{15} K^5 |A +_G B|^3 / (|A||B|).$$

Por hipótesis, $|A +_G B| \leq K_1 |A|$ y $|A| = |B|$, por lo tanto

$$|A' + B'| \leq 2^{15} K^5 K_1^3 |A|.$$

□

### 4.2.3. Linealidad de $\phi$ sobre un subconjunto de $B$

Del inicio de esta sección tenemos que para una función $f : \mathbb{Z}_N \to [-1, 1]$, con $||f||_{U^3} \geq \alpha$, existen un conjunto $B \subseteq \mathbb{Z}_N$ y una función $\phi : B \to \mathbb{Z}_N$, tales que

$$|B| \geq \alpha^8 N/2, \text{ y para cada } k \in B, \, |\Delta(f; k)\widehat{\ }(\phi(k))| \geq \alpha^4 N/2,$$

como mostramos en (4.17). Y además, por la proposición 4.2, existen al menos $\alpha^{64} 2^{-12} N^3$ cuádruplas $(b_1, b_2, b_3, b_4) \in B^4$ tales que

$$b_1 + b_2 = b_3 + b_4 \quad \text{y} \quad \phi(b_1) + \phi(b_2) = \phi(b_3) + \phi(b_4),$$

esto es,

$$\big(b_1, \phi(b_1)\big) + \big(b_2, \phi(b_2)\big) = \big(b_3, \phi(b_3)\big) + \big(b_4, \phi(b_4)\big). \tag{4.32}$$

Sea $\Gamma = \big\{\big(b, \phi(b)\big);\ b \in B\big\}$. $\Gamma$ es un subconjunto del grupo abeliano aditivo $\mathbb{Z}_N^2$ tal que el número de cuádruplas aditivas en $\Gamma$ es al menos $\alpha^{64} 2^{-12} N^3 \geq \alpha^{64} 2^{-12} |\Gamma|^3$. Por el teorema 4.1, considerando $A = B = \Gamma$, tenemos que existen subconjuntos $\Gamma_1 \subseteq \Gamma$ y $\Gamma_2 \subseteq \Gamma$ tales que

$$|\Gamma_1| \geq (\alpha^{64} 2^{-12})^2 |\Gamma| / (16\sqrt{2}), \ \ |\Gamma_2| \geq (\alpha^{64} 2^{-12})^2 |\Gamma| / 16 \quad \text{y} \quad |\Gamma_1 + \Gamma_2| \leq 2^{28} (\alpha^{64} 2^{-12})^{-13} |\Gamma|.$$

Esto es,

$$|\Gamma_1| \geq \alpha^{128} 2^{-29} |\Gamma|, \ \ |\Gamma_2| \geq \alpha^{128} 2^{-28} |\Gamma| \quad \text{y} \quad |\Gamma_1 + \Gamma_2| \leq 2^{184} \alpha^{-832} |\Gamma|.$$

Disminuyendo un poco los conjuntos $\Gamma_1$ y $\Gamma_2$ si es necesario, podemos suponer que

$$|\Gamma_1| = |\Gamma_2| = \alpha^{128} 2^{-30} |\Gamma| \quad \text{y} \quad |\Gamma_1 + \Gamma_2| \leq 2^{200} \alpha^{-840} |\Gamma|.$$

Ahora desearíamos aplicar la siguiente variante del teorema de Freiman, ver [1].

**Teorema 4.3** (Variante de Freiman)**.** *Sea $A \subseteq \mathbb{Z}$ tal que $|A-A| \leq C|A|$ o $|A+A| \leq C|A|$. Entonces existe una progresión aritmética propia $Q_0$ de dimensión $\leq 2^{11}C^{32}$ y de tamaño $\gg \exp(-C^{33})|A|$ tal que*

$$|A \cap Q_0| \gg C^{-16}|Q_0| \gg \exp(-C^{34})|A|.$$

Pero este se aplica a subconjuntos de $\mathbb{Z}$. Identificamos $\mathbb{Z}_N^2$ con $[1,N]^2$, y dividimos este conjunto $[1,N]^2$ en cuatro cuadrados de las mismas dimensiones $Q_1, Q_2, Q_3$ y $Q_4$. Entonces podemos elegir subconjuntos $\Gamma_3 \subseteq \Gamma_1$ y $\Gamma_4 \subseteq \Gamma_2$, con $|\Gamma_3| = |\Gamma_4| \leq \frac{1}{4}|\Gamma_1|$, y cada uno de ellos incluidos en algunos cuadrados $Q_i$ y $Q_j$ $(i,j \in \{1,2,3,4\})$. Los subconjuntos $\Gamma_3$ y $\Gamma_4$ son naturalmente 2-isomorfos a subconjuntos de $\mathbb{Z}$, digamos $\Gamma_5$ y $\Gamma_6$. Aplicando el teorema de Freiman, concluimos que existe una progresión propia $Q$ en $\mathbb{Z}_N^2$, con dimensión $\ll \alpha^{-2^{16}}$ y tamaño $\gg \exp(-\alpha^{2^{-17}})N$, tal que

$$|\Gamma \cap Q| \gg \alpha^{2^{16}}|Q|.$$

Puesto que $Q$ es propia, contiene una progresión aritmética unidimensional de longitud $\gg \exp(-\alpha^{-2^{16}})N^{2^{16}}$. Cubriendo $Q$ por traslaciones de esta progresión, tenemos que existe una progresión aritmética unidimensional $P_0$ (en $\mathbb{Z}_N^2$), con tamaño $\gg \exp(-\alpha^{-2^{16}})N^{\alpha^{2^{16}}}$, tal que

$$|\Gamma \cap P_0| \gg \alpha^{2^{16}}|P_0|.$$

En resumen, manteniendo las notaciones para el conjunto $B$ y la función $\phi : B \to \mathbb{Z}_N$, establecidas al inicio de esta sección, tenemos el siguiente resultado

**Proposición 4.3.** *Existe una progresión aritmética $P \subseteq \mathbb{Z}_N$, con tamaño $\gg \exp(-\alpha^{-2^{16}})N^{\alpha^{2^{16}}}$, y una función $n \mapsto 2\lambda n + \mu$, tal que*

$$|B \cap P| \geq \eta|P|, \quad \text{con } \eta \gg \alpha^{2^{16}},$$

*y*

$$\phi(k) = 2\lambda k + \mu, \quad \text{para cada } k \in B \cap P.$$

### 4.2.4. Extracción de sesgo cuadrático

Con la finalidad de completar la prueba del teorema de Szemerédi para progresiones de longitud cuatro, ahora mostraremos que la función $f$ considerada al inicio de esta sección, sección 2.4., se correlaciona localmente con una función fase cuadrática. Iniciamos con la siguiente proposición.

**Proposición 4.4.** *Sea $f:\ \mathbb{Z}_N \to [-1,1]$ una función, y supongamos que para algún $\lambda \in \mathbb{Z}_N$,*

$$\sum_{k\in\mathbb{Z}_N} |\Delta(f;k)\widehat{\ }(2\lambda k)|^2 \geq \zeta N^3.$$

*Entonces para algún $r \in \mathbb{Z}_N$, $\left|\sum_n f(n)\omega^{\lambda n^2+rn}\right| \geq \sqrt{\zeta}N$.*

*Demostración.* Expandiendo la sumatoria de nuestra hipótesis tenemos que

$$\begin{aligned}\zeta N^3 &\leq \sum_k \sum_{x,y} f(x)f(x-k)f(y)f(y-k)\omega^{2\lambda k(y-x)} \\ &= \sum_k \sum_{x,u} f(x)f(x-k)f(x-u)f(x-u-k)\omega^{2\lambda ku}.\end{aligned}$$

Observemos que $x^2-(x-k)^2-(x-u)^2+(x-k-u)^2 = 2ku$, y reemplazando esto, la última expresión resulta igual a

$$\sum_{x,k,u} f(x)\omega^{\lambda x^2} f(x-k)\omega^{-\lambda(x-k)^2} f(x-u)\omega^{-\lambda(x-u)^2} f(x-u-k)\omega^{\lambda(x-u-k)^2}.$$

Consideremos la función $g(x) = f(x)\omega^{\lambda x^2}$, la expresión anterior resulta ser

$$\sum_{x,k,u} g(x)\overline{g(x-k)}\,\overline{g(x-u)}g(x-u-k) = N^3||g||_{U^2}^4.$$

Por lo tanto

$$\begin{aligned}\zeta N^3 &\leq N^3||g||_{U^2}^4 = \frac{1}{N}\sum_r |\widehat{g}(r)|^4 \\ &\leq \frac{1}{N}\max_r |\widehat{g}(r)|^2 \cdot \sum_r |\widehat{g}(r)|^2 \\ &= \max_r |\widehat{g}(r)|^2 \cdot \sum_r |g(r)|^2 \\ &\leq N \max_r |\widehat{g}(r)|^2.\end{aligned}$$

Así, $\sqrt{\zeta}N \leq \max_r |\widehat{g}(r)|$, y por tanto existe $r \in \mathbb{Z}_N$ tal que

$$\sqrt{\zeta}N \leq |\widehat{g}(r)| = \left|\sum_n g(n)\omega^{-rn}\right| = \left|\sum_n f(n)\omega^{\lambda n^2-rn}\right|.$$

Cambiando $r$ por $-r$ obtenemos el resultado. □

Mantengamos las notaciones de la subsección anterior, subsección 4.2.3. Allí se tenía una función $f:\ \mathbb{Z}_N \to [-1,1]$, (a saber la función balanceada asociada a $A \subseteq [1,N]$, con $|A| = \delta N$), con $||f||_{U^3} \geq \alpha$, un conjunto $B \subseteq \mathbb{Z}_N$ y una función $\phi : B \to \mathbb{Z}_N$, tales que

$$|B| \geq \alpha^8 N/2, \text{ y para cada } k \in B,\ |\Delta(f;k)\widehat{\ }(\phi(k))| \geq \alpha^4 N/2.$$

Sea $P$ la progresión aritmética (en $\mathbb{Z}_N$) dada por la Proposición 4.3. Así, $P$ tiene tamaño $\gg \exp(-\alpha^{-2^{16}})N^{\alpha^{2^{16}}}$, y también se da una función $n \mapsto 2\lambda n + \mu$, tal que

$$|B \cap P| \geq \eta|P|, \quad \text{con } \eta \gg \alpha^{2^{16}}.$$

**Proposición 4.5.** *Manteniendo las notaciones arriba mencionadas. Para cada $x \in \mathbb{Z}_N$, existe $r_x \in \mathbb{Z}_N$ tal que*

$$\sum_x \left| \sum_{k \in P+x} f(k)\omega^{-\lambda k^2 + r_x k} \right| \geq \frac{\eta\alpha^8}{4\sqrt{2}} N|P|.$$

*Demostración.* De la proposición 4.2, y según las notaciones, tenemos que

$$\begin{aligned} \sum_{k\in P} |\Delta(f;k)\widehat{\ }(2\lambda k+\mu)|^2 &\geq \sum_{k\in P\cap B} |\Delta(f;k)\widehat{\ }(2\lambda k+\mu)|^2 \\ &= \sum_{k\in P\cap B} |\Delta(f;k)\widehat{\ }(\phi(k))|^2, \end{aligned}$$

puesto que para cada $k \in P \cap B$, $\phi(k) = 2\lambda k + \mu$. Además, como $|P \cap B| \geq \eta|P|$ y

$$|\Delta(f;k)\widehat{\ }(\phi(k))| \geq \frac{\alpha^4}{2}N, \quad \text{para cada } k \in B,$$

tenemos que

$$\sum_{k\in P} |\Delta(f;k)\widehat{\ }(2\lambda k+\mu)|^2 \geq \frac{\eta\alpha^8}{4}|P|N^2.$$

Expandiendo el lado izquierdo tenemos que

$$\begin{aligned} &\sum_{k\in P}\sum_{x,y} \Delta(f;k)(x)\Delta(f;k)(y)\omega^{(2\lambda k+\mu)(y-x)} \\ &\qquad = \sum_{k\in P}\sum_{x,y} f(x)f(x+k)f(y)f(y+k)\omega^{(2\lambda k+\mu)(y-x)} \\ &\qquad = \sum_{k\in P}\sum_{x,u} f(x)f(x+k)f(x+u)f(x+u+k)\omega^{(2\lambda k+\mu)u}. \end{aligned}$$

Ahora, cada $u \in \mathbb{Z}_N$ puede ser escrito como $l + y$, con $l \in P$ y $y \in \mathbb{Z}_N$, en exactamente $|P|$ modos. entonces la última expresión anterior es igual a

$$\begin{aligned} &\sum_{k\in P}\sum_{x} f(x)f(x+k)\sum_{u} f(x+u)f(x+u+k)\omega^{(2\lambda k+\mu)u} \\ &\qquad = \sum_{k\in P}\sum_{x} f(x)f(x+k)\cdot\frac{1}{|P|}\sum_{y}\sum_{l\in P} f(x+y+l)f(x+y+l+k)\omega^{(2\lambda k+\mu)(l+y)} \\ &\qquad = \frac{1}{|P|}\sum_{k\in P}\sum_{x,y}\sum_{l\in P} f(x)f(x+k)f(x+y+l)f(x+y+l+k)\omega^{(2\lambda k+\mu)(l+y)} \\ &\qquad = \frac{1}{|P|}\sum_{x,y}\sum_{k,l\in P} f(x)f(x+k)f(x+y+l)f(x+y+l+k)\omega^{(2\lambda k+\mu)(l+y)}. \end{aligned}$$

Así,

$$\frac{\eta\alpha^8}{4}|P|^2N \leq \frac{1}{N}\sum_y\sum_x\left|\sum_{k,l\in P} f(x)f(x+k)f(x+y+l)f(x+y+l+k)\omega^{(2\lambda k+\mu)(l+y)}\right|.$$

Por lo tanto, existe $y\in\mathbb{Z}_N$ tal que

$$\begin{aligned}\frac{\eta\alpha^8}{4}|P|^2N &\leq \sum_x\left|\sum_{k,l\in P} f(x)f(x+k)f(x+y+l)f(x+y+l+k)\omega^{(2\lambda k+\mu)(l+y)}\right| \\ &\leq \sum_x\left|\sum_{k,l\in P} f(x+k)f(x+y+l)f(x+y+l+k)\omega^{2\lambda kl+2\lambda ky+\mu l}\omega^{\mu y}\right| \\ &= \sum_x\left|\sum_{k,l\in P} f(x+k)f(x+y+l)f(x+y+l+k)\omega^{2\lambda kl+2\lambda ky+\mu l}\right|. \qquad (4.33)\end{aligned}$$

Como $2\lambda kl = \lambda\big((k+l)^2-k^2-l^2\big)$, tenemos que

$$\frac{\eta\alpha^8}{4}|P|^2N \leq \sum_x\left|\sum_{k,l\in P} f(x+k)\omega^{-\lambda k^2+2\lambda ky}f(x+y+l)\omega^{-\lambda l^2+\mu l}f(x+y+k+l)\omega^{\lambda(k+l)^2}\right|.$$

Definamos para cada $x\in\mathbb{Z}_N$,

$$\begin{aligned} g_{1,x}(k) &= \begin{cases} f(x+k)\omega^{-\lambda k^2+2\lambda ky} & \text{, si} \quad k\in P \\ 0 & \text{, si} \quad k\notin P. \end{cases} \\ g_{2,x}(l) &= \begin{cases} f(x+y+l)\omega^{-\lambda l^2+\mu l} & \text{, si} \quad l\in P \\ 0 & \text{, si} \quad l\notin P. \end{cases} \\ g_{3,x}(k+l) &= \begin{cases} f(x+y+k+l)\omega^{\lambda(k+l)^2} & \text{, si} \quad k+l\in P+P \\ 0 & \text{, si} \quad k+l\notin P+P. \end{cases} \end{aligned}$$

Entonces tendremos que

$$\begin{aligned}\frac{\eta\alpha^8}{4}|P|^2N &\leq \sum_x\left|\sum_{k,l} g_{1,x}(k)g_{2,x}(l)g_{3,x}(k+l)\right| \\ &= \sum_x\left|\frac{1}{N^3}\sum_{k,l}\sum_{r_1}\widehat{g}_{1,x}(r_1)\omega^{-r_1k}\sum_{r_2}\widehat{g}_{2,x}(r_2)\omega^{-r_2l}\sum_{r_3}\widehat{g}_{3,x}(r_3)\omega^{-r_3(k+l)}\right| \\ &= \sum_x\left|\frac{1}{N^3}\sum_{r_1,r_2,r_3}\widehat{g}_{1,x}(r_1)\widehat{g}_{2,x}(r_2)\widehat{g}_{3,x}(r_3)\sum_{k,l}\omega^{-k(r_1+r_3)}\omega^{-l(r_2+r_3)}\right| \\ &= \frac{1}{N}\sum_x\left|\sum_r\widehat{g}_{1,x}(r)\widehat{g}_{2,x}(r)\widehat{g}_{3,x}(-r)\right| \\ &\leq \frac{1}{N}\sum_x\left[\max_r|\widehat{g}_{1,x}(r)|\cdot\sum_r\left|\widehat{g}_{2,x}(r)\widehat{g}_{3,x}(-r)\right|\right].\end{aligned}$$

Por la desigualdad de Cauchy-Schwarz, tenemos

$$\begin{aligned}\frac{\eta\alpha^8}{4}|P|^2N &\leq \sum_x\left[\max_r|\widehat{g}_{1,x}(r)|\cdot\left(\frac{1}{N}\sum_r|\widehat{g}_{2,x}(r)|^2\right)^{1/2}\left(\frac{1}{N}\sum_r|\widehat{g}_{3,x}(-r)|^2\right)^{1/2}\right]\\ &= \sum_x\left[\max_r|\widehat{g}_{1,x}(r)|\cdot\left(\sum_r|g_{2,x}(r)|^2\right)^{1/2}\left(\sum_r|g_{3,x}(-r)|^2\right)^{1/2}\right]\\ &\leq \sum_x\left[\max_r|\widehat{g}_{1,x}(r)|\cdot|P|^{1/2}(2|P|)^{1/2}\right].\end{aligned}$$

Así,

$$\begin{aligned}\frac{\eta\alpha^8}{4\sqrt{2}}|P|N &\leq \sum_x\max_r\left|\widehat{g}_{1,x}(r)\right|\\ &= \sum_x\max_r\left|\sum_{k\in P}g_{1,x}(k)\omega^{-rk}\right|\\ &= \sum_x\max_r\left|\sum_{k\in P}f(x+k)\omega^{-\lambda k^2+2\lambda ky-rk}\right|.\end{aligned}$$

Así, para cada $x$ existe $\overline{r}_x$ (a saber, $\overline{r}_x=-2\lambda y+r$, donde $r$ es un valor fijado en el cual se da el máximo), tal que

$$\sum_x\left|\sum_{k\in P}f(x+k)\omega^{-(\lambda k^2+\overline{r}_xk)}\right|\geq\frac{\eta\alpha^8}{4\sqrt{2}}|P|N.$$

Por lo tanto

$$\begin{aligned}\sum_x\left|\sum_{k\in P+x}f(k)\omega^{-\lambda(k-x)^2-\overline{r}_x(k-x)}\right| &\geq \frac{\eta\alpha^8}{4\sqrt{2}}|P|N\\ \sum_x\left|\sum_{k\in P+x}f(k)\omega^{-\lambda k^2+2\lambda kx-\overline{r}_xk}\omega^{\lambda x^2+\overline{r}_xx}\right| &\geq \frac{\eta\alpha^8}{4\sqrt{2}}|P|N.\end{aligned}$$

Tomando para cada $x$, $r_x=-2\lambda x+\overline{r}_x$, obtenemos el resultado. □

### 4.2.5. Aplicación de la Desigualdad de Weyl, e incremento de densidad

Recordamos de la prueba del teorema de Roth que en el caso de que exista un coeficiente de Fourier grande de $f$, digamos $|\widehat{f}(r)|$ grande para algún $r\neq 0$, procedimos a particionar $[1,N]$ en subprogresiones aritméticas, sobre cada una de las cuales la función $n\mapsto\omega^{rn}$ tuviese una variación controlada. En el caso actual de progresiones aritméticas de longitud 4 aplicaremos la misma técnica.

**Lema 4.13.** *Sea $P$ una progresión aritmética en $\mathbb{Z}_N$ de longitud $R$. Y sea $\psi_1:\mathbb{Z}_N\to\mathbb{Z}_N$, $\psi_1(x)=\alpha x$, una función lineal. Entonces existe una partición de $P$ en progresiones*

*aritméticas (en $\mathbb{Z}_N$), $P_1, \ldots, P_M$, cada una de longitud menor a $R^{1/4}$ (así, $M$ está próximo a $R^{3/4}$), tal que*

$$\left|\omega^{\psi_1(x)} - \omega^{\psi_1(y)}\right| \ll R^{-1/4}, \text{ para cada } x, y \in P_j.$$

*Demostración.* Sin pérdida de generalidad supongamos que $P = [1, R]$. Por el teorema de Dirichlet, existe $q \leq \sqrt{R}$ tal que $||q\alpha/N|| \leq 1/\sqrt{R}$. Particionamos $[1, R]$ en progresiones aritméticas (en $\mathbb{Z}_N$), con diferencia común $q$, y denotemos a estas progresiones por $Q_1, \ldots, Q_t$. Para ello, basta con tomar en $[1, R]$ la relación de equivalencia $\equiv$ (mód $q$). Las clases de equivalencia constituyen una partición de $[1, R]$. Si una de estas progresiones es $a + jq$, con $1 \leq j \leq R/q$, entonces dividimos $[1, R/q]$ en subintervalos de longitud (a lo más) $R^{1/4}$. Así, obtenemos $R^{3/4}$ subprogresiones tales que si $x, y$ están en alguna de estas, digamos $x = a + j_1 q$, $y = a + j_2 q$, con $|j_1 - j_2| \leq R^{1/4}$, entonces, de la desigualdad del valor medio, tenemos que

$$\left|\omega^{\psi_1(x)} - \omega^{\psi_1(y)}\right| \leq \left|\left|\frac{2\pi}{N}\alpha(x-y)\right|\right| = \left|\left|\frac{2\pi}{N}\alpha(j_1 - j_2)q\right|\right| \leq 2\pi R^{1/4}||\alpha q/N|| \leq 2\pi R^{-1/4}.$$

□

**Lema 4.14.** *Sea $P$ una progresión aritmética en $\mathbb{Z}_N$ de longitud $R$. Y sea $\psi_2 : \mathbb{Z}_N \to \mathbb{Z}_N$, $\psi_2(x) = \alpha x^2 + \beta x$, una función cuadrática. Entonces podemos particionar $P$ en progresiones aritméticas (en $\mathbb{Z}_N$), $P_1, \ldots, P_M$, cada una de longitud próxima y menor a $R^{1/128}$ (así, $M$ está próximo a $R^{127/128}$), tal que para cada $x, y \in P$,*

$$\left|\omega^{\psi_2(x)} - \omega^{\psi_2(y)}\right| \ll R^{-\frac{1}{128}}.$$

*Demostración.* Sin pérdida de generalidad supongamos que $P = [1, R]$. Por el teorema de Weyl existe $q \leq R^{1/2}$, tal que $||q^2\alpha/N|| \leq R^{-1/8}$. Particionemos $P$ en subprogresiones $Q_1, \ldots, Q_k$, todas ellas con diferencia común $q$ (basta tomar en $P$ la relación de equivalencia $\equiv$ (mód $q$), como en el lema anterior). Ahora, dividamos cada una de estas progresiones en intervalos de longitud $R^{1/32}$ (como lo hicimos en el lema 4.13). De este modo obtenemos alrededor de $R^{\frac{31}{32}}$ subprogresiones $P_1, \ldots, P_M$, cada una de longitud alrededor de $R^{1/32}$. Consideremos una de estas subprogresiones $P_i$, digamos de la forma $a + jq$, para algún $a$ y $1 \leq j \leq R^{1/32}$. Observemos que

$$\frac{\psi_2(a+jq)}{N} = \frac{\alpha}{N}(a+jq)^2 + \frac{\beta}{N}(a+jq) = \frac{\alpha a^2 + \beta a}{N} + j^2\frac{q^2\alpha}{N} + j\frac{2\alpha aq + \beta q}{N},$$

El primer término $\dfrac{\alpha a^2 + \beta a}{N}$ es constante (no depende de $j$) y no varía sobre esta subprogresión $P_i$. El segundo término, $j^2\dfrac{q^2\alpha}{N}$, visto como un resto (mód 1), varía a lo más $R^{-\frac{1}{16}}$. En

efecto, si $j_1, j_2 \in [1, R^{\frac{1}{32}}]$ entonces $j_1^2, j_2^2 \in [1, R^{1/16}]$, y por lo tanto

$$\left|\left|(j_1^2 - j_2^2)\frac{q^2\alpha}{N}\right|\right| \leq R^{1/16}\left|\left|\frac{q^2\alpha}{N}\right|\right| \leq R^{1/16} \cdot \frac{1}{R^{1/8}} = R^{-1/16}.$$

Finalmente, el último término $\psi_3^i(j) = j\frac{2\alpha aq+\beta q}{N}$, da lugar a un polinomio lineal en $j$. Podemos entonces aplicar el lema 4.13 con la finalidad de hacer a $\psi_3^i$ localmente casi constante. Esto es, por el lema 4.13, podemos descomponer $P_i$ (de longitud $R^{\frac{1}{32}}$) en subprogresiones $P_{i,1}^*, \ldots, P_{i,H}^*$, cada una de longitud $(R^{\frac{1}{32}})^{\frac{1}{4}} = R^{\frac{1}{128}}$, sobre las cuales $\psi_3^i$ varía ( mód 1) a lo más $(R^{\frac{1}{32}})^{-\frac{1}{4}} = R^{-\frac{1}{128}}$. Así, para cada $j_1, j_2 \in P_{i,r}^*$,

$$\left|\left|\frac{\psi_2(a+j_1q)}{N} - \frac{\psi_2(a+j_2q)}{N}\right|\right| \leq R^{-\frac{1}{16}} + R^{-\frac{1}{128}} \leq 2R^{-\frac{1}{128}}.$$

Y por lo tanto,

$$|\omega^{\psi_2(a+j_1q)} - \omega^{\psi_2(a+j_2q)}| \leq \left|\left|\frac{2\pi}{N}\big(\psi_2(a+j_1q) - \psi_2(a+j_2q)\big)\right|\right| \leq 4\pi R^{-\frac{1}{128}}.$$

□

**Lema 4.15.** *Sea $P$ una progresión aritmética en $\mathbb{Z}_N$ de longitud $R$. Entonces podemos particionar $P$ en $4\sqrt{R}$ progresiones aritméticas en $\mathbb{Z}$.*

*Demostración.* Supongamos que la progresión $P$ es de la forma $a + jq$, con $1 \leq j \leq R$. Por el teorema de Dirichlet existe $1 \leq l \leq \sqrt{R}$, tal que $||lq/N|| \leq R^{-1/2}$. Como en el lema 4.13, haciendo uso de la congruencia módulo $l$, descomponemos $[1, R]$ en subprogresiones, cada una de longitud $R^{1/2}/2$ (por tanto la cantidad de subprogresiones es alrededor de $2\sqrt{R}$). Como $l \leq \sqrt{R}$, el diámetro de cada subprogresión es a lo más $R/2$. Dividiendo a su vez, cada una de estas subprogresiones en dos subprogresiones, dividimos $P$ en $4R^{1/2}$ subprogresiones. □

Mantengamos las notaciones de la sección *Linealidad de $\phi$ sobre un subconjunto de $B$*. Allí se tenía una función $f : \mathbb{Z}_N \to [-1, 1]$, (a saber la función balanceada asociada a $A \subseteq [1, N]$, con $|A| = \delta N$), con $||f||_{U^3} \geq \alpha$, un conjunto $B \subseteq \mathbb{Z}_N$ y una función $\phi : B \to \mathbb{Z}_N$, tales que

$$|B| \geq \alpha^8 N/2, \text{ y para cada } k \in B, |\Delta(f; k)\widehat{\ }(\phi(k))| \geq \alpha^4 N/2.$$

Sea $P$ la progresión aritmética (en $\mathbb{Z}_N$) dada por la Proposición 4.3. Así, $P$ tiene tamaño $\gg \exp(-\alpha^{-2^{16}})N^{\alpha^{2^{16}}}$, y también se da una función $n \mapsto 2\lambda n + \mu$, tal que

$$|B \cap P| \geq \eta|P|, \quad \text{con } \eta \gg \alpha^{2^{16}},$$

y

$$\phi(k) = 2\lambda k + \mu, \quad \text{para cada } k \in B \cap P.$$

**Proposición 4.6.** *Sean $P$ y $n \mapsto 2\lambda n + \mu$ como antes, y sea $R$ el tamaño de $P$. Para cada $x \in \mathbb{Z}_N$, podemos particionar $P + x$ en alrededor de $4R^{\frac{255}{256}}$ progresiones aritméticas en $\mathbb{Z}$, $P_{x,1}, \dots, P_{x,M}$, tal que*

$$\sum_{x \in \mathbb{Z}_N} \sum_{j=1}^{M} \left| \sum_{k \in P_{x,j}} f(k) \right| \geq \frac{\eta \alpha^8}{8} NR.$$

*Demostración.* De la proposición 4.5, para cada $x \in \mathbb{Z}_N$ existe $r_x \in \mathbb{Z}_N$ tal que

$$\sum_{x} \left| \sum_{k \in P+x} f(k) \omega^{-(\lambda k^2 + r_x k)} \right| \quad \geq \quad \frac{\eta \alpha^8}{4\sqrt{2}} N|P|.$$

Por el lema 4.14, podemos descomponer $P + x$ en alrededor de $R^{\frac{127}{128}}$ progresiones (en $\mathbb{Z}_N$), $Q_{x,1}, \dots, Q_{x,J}$, sobre cada una de las cuales

$$\left| \omega^{-(\lambda k_1^2 + r_x k_1)} - \omega^{-(\lambda k_2^2 + r_x k_2)} \right| \ll R^{-\frac{1}{128}}, \quad \text{para cada } k_1, k_2 \in Q_{x,i}. \tag{4.34}$$

Entonces, usando el lema 4.15, subdividimos cada una de estas subprogresiones $Q_{x,i}$, de tamaño $R^{\frac{1}{128}}$, en progresiones aritméticas en $\mathbb{Z}$, cada una de longitud $4(R^{\frac{1}{128}})^{1/2} = 4R^{\frac{1}{256}}$. Así, obtenemos una descomposición de $P + x$ en subprogresiones $P_{x,1}, \dots, P_{x,M}$, tales que

$$\left| \omega^{-(\lambda k_1^2 + r_x k_1)} - \omega^{-(\lambda k_2^2 + r_x k_2)} \right| \ll R^{-\frac{1}{128}}, \quad \text{para cada } k_1, k_2 \in P_{x,i}.$$

Para cada $i \in \{1, \dots, M\}$, fijemos $k_{x,i} \in P_{x,i}$. Entonces

$$\begin{aligned}
\sum_{x \in \mathbb{Z}_N} \sum_{j=1}^{M} \left| \sum_{k \in P_{x,j}} f(k) \right| &= \sum_{x \in \mathbb{Z}_N} \sum_{j=1}^{M} \left| \sum_{k \in P_{x,j}} f(k) \omega^{-(\lambda k_{x,j}^2 + r_x k_{x,j})} \right| \\
&= \sum_{x \in \mathbb{Z}_N} \sum_{j=1}^{M} \left| \sum_{k \in P_{x,j}} f(k) \left( \omega^{-(\lambda k_{x,j}^2 + r_x k_{x,j})} - \omega^{-(\lambda k^2 + r_x k)} \right) + \sum_{k \in P_{x,j}} f(k) \omega^{-(\lambda k^2 + r_x k)} \right| \\
&\geq \sum_{x \in \mathbb{Z}_N} \sum_{j=1}^{M} \left| \sum_{k \in P_{x,j}} f(k) \omega^{-(\lambda k^2 + r_x k)} \right| - \sum_{x \in \mathbb{Z}_N} \sum_{j=1}^{M} \left| \sum_{k \in P_{x,j}} f(k) \left( \omega^{-(\lambda k_{x,j}^2 + r_x k_{x,j})} - \omega^{-(\lambda k^2 + r_x k)} \right) \right| \\
&\geq \frac{\eta \alpha^8}{4\sqrt{2}} N|P| - \sum_{x \in \mathbb{Z}_N} \sum_{j=1}^{M} C R^{-\frac{1}{128}} = \frac{\eta \alpha^8}{4\sqrt{2}} N|P| - N|P| C R^{-\frac{1}{128}} \\
&\geq \frac{\eta \alpha^8}{8} NR.
\end{aligned}$$

□

**Corolario 4.2** (Incremento de densidad)**.** *Manteniendo las notaciones de la proposición anterior (proposición 4.6), existe una progresión aritmética (en $\mathbb{Z}$) $Q$, de tamaño al menos $(\eta\alpha^8/128)R^{\frac{1}{256}}$ tal que*

$$|A \cap Q| \geq \left(\delta + \frac{\eta\alpha^8}{128}\right)|Q|.$$

*Demostración.* Notemos que

$$\begin{aligned}\sum_{x\in\mathbb{Z}_N}\sum_{j=1}^{M}\sum_{k\in P_{x,j}} f(k) &= \sum_{x\in\mathbb{Z}_N}\sum_{k\in P+x} f(k)\\ &= |P|\sum_{k\in\mathbb{Z}_N} f(k) = \widehat{f}(0) = 0, \end{aligned} \tag{4.35}$$

puesto que $f$ es la función balanceada asociada a el conjunto $A \subseteq [1, N]$. De la proposición 4.6 tenemos que

$$\sum_{x\in\mathbb{Z}_N}\sum_{j=1}^{M}\left|\sum_{k\in P_{x,j}} f(k)\right| \geq \frac{\eta\alpha^8}{8}NR. \tag{4.36}$$

Sumando (4.35) y (4.36), obtenemos

$$\sum_{x\in\mathbb{Z}_N}\sum_{j=1}^{M}\left[\sum_{k\in P_{x,j}} f(k) + \left|\sum_{k\in P_{x,j}} f(k)\right|\right] \geq \frac{\eta\alpha^8}{8}NR.$$

Así,

$$\sum_{x\in\mathbb{Z}_N}\sum_{j=1}^{M}\text{máx}\left\{0, \sum_{k\in P_{x,j}} f(k)\right\} \geq \frac{\eta\alpha^8}{16}NR. \tag{4.37}$$

Sea $\left\{(x, j) \in \mathbb{Z}_N \times \{1, \ldots, M\};\ |P_{x,j}| \leq (\eta\alpha^8/128)R^{\frac{1}{256}}\right\}$. Como

$$\begin{aligned}\sum_{(x,j)\in J}\text{máx}\left\{0, \sum_{k\in P_{x,j}} f(k)\right\} &\leq |J|\cdot|P_{x,j}| = |J|\cdot(\eta\alpha^8/128)R^{\frac{1}{256}}\\ &\leq NM\cdot(\eta\alpha^8/128)R^{\frac{1}{256}} = (\eta\alpha^8/32)NR,\end{aligned}$$

puesto que $M = 4R^{\frac{255}{256}}$. Por lo tanto, de (4.37),

$$\frac{\eta\alpha^8}{32}NR \leq \sum_{(x,j)\in J^c}\text{máx}\left\{0, \sum_{k\in P_{x,j}} f(k)\right\}. \tag{4.38}$$

Si para cada $(x, j) \in J^c$,

$$\sum_{k\in P_{x,j}} f(k) < \frac{\eta\alpha^8}{128}R^{\frac{1}{256}},$$

entonces

$$\sum_{(x,j)\in J^c}\text{máx}\left\{0, \sum_{k\in P_{x,j}} f(k)\right\} < |J^c|\cdot\frac{\eta\alpha^8}{128}R^{\frac{1}{256}} \leq NM\cdot\frac{\eta\alpha^8}{128}R^{\frac{1}{256}} = \frac{\eta\alpha^8}{32}NR,$$

lo que contradiría (4.38). Por lo tanto, existen $x \in \mathbb{Z}_N$, y $j \in \{1, 2, \ldots, M\}$, tales que

$$\sum_{k \in P_{x,j}} f(k) \geq \frac{\eta \alpha^8}{128} R^{\frac{1}{256}}.$$

Como $f(k) = A(k) - \delta$ para cada $k \in \mathbb{Z}_N$, y $|P_{x,j}| = R^{\frac{1}{256}}$, se tiene que

$$\sum_{k \in P_{x,j}} A(k) - \delta |P_{x,j}| \quad \geq \quad \frac{\eta \alpha^8}{128} |P_{x,j}|,$$

y por lo tanto

$$|A \cap P_{x,j}| \quad \geq \quad \left( \delta + \frac{\eta \alpha^8}{128} \right) |P_{x,j}|.$$

□

# Capítulo 5

# Existencia de progresiones aritméticas de longitud $k \geq 3$. Método ergódico

En 1927 van der Waerden publicó un conocido teorema, el cual afirma que si los enteros positivos son particionados en un número finito de subconjuntos, entonces al menos uno de estos subconjuntos contiene progresiones arbitrariamente largas. En 1936 Erdös y Turán hicieron una conjetura que implica fácilmente el teorema de van der Waerden, afirmando que es posible hallar progresiones aritméticas de longitud $k$ en cualquier subconjunto de los números enteros con densidad positiva, y de este modo la partición hecha de los enteros positivos en el teorema de van der Waerden podía ser, en algún sentido, una distracción. La conjetura fue probada por Szemerédi en 1975, resultado hoy conocido como el *teorema de Szemerédi*, uno de los fundamentales resultados en teoría de números combinatoria.

## 5.1. Aspectos ergódicos

A continuación, antes de probar el teorema de Szemerédi, veremos cuán potentes son los métodos de teoría ergódica en la deducción del teorema de Van der Waerden, (cf. [16, pp. 163–185]).

**Teorema 5.1** (van der Waerden)**.** *Sean* $A = \{a_1, a_2, a_3, \ldots, a_k\}$ *un conjunto finito. Entonces,*

*para cada función sobreyectiva* $\Phi : \mathbb{N} \to A$ *existe algún* $a \in A$ *tal que* $\Phi^{-1}(a)$ *contiene infinitas progresiones aritméticas arbitrariamente largas, esto es, para cada* $r \in \mathbb{N}$*, existen infinitos* $m, n \in \mathbb{N}$ *tales que* $m + in \in \Phi^{-1}(a)$ *para cada* $i \in \{0, 1, 2, \ldots, r-1\}$.

El teorema de Szemerédi involucra conjuntos con densidad superior positiva, concepto que definimos a continuación.

**Definición 5.1.** *Sea* $A \subseteq \mathbb{N}$*, la* densidad superior *de* $A$ *en* $\mathbb{N}$*, denotada por* $\delta(A)$*, se define como*

$$\delta(A) = \limsup_{N\to\infty} \frac{|A \cap [1, N]|}{N},$$

*donde* $[1, N] = \{1, 2, \ldots, N\}$*, y* $|B|$ *denota la cardinalidad del conjunto* $B$.

La siguiente versión del teorema de Szemerédi será llamada versión clásica.

**Teorema 5.2** (Szemerédi, versión clásica)**.** *Sea* $A \subseteq \mathbb{N}$ *tal que* $\delta(A) > 0$*, entonces* $A$ *contiene infinitas progresiones aritméticas de longitud* $k$ *para todo entero* $k \geq 3$.

**Lema 5.1.** *Dado un conjunto finito* $A$*, el producto cartesiano de* $A$ *consigo mismo una cantidad numerable de veces,* $\Omega = A^{\mathbb{N}}$*, es metrizable.*

*Demostración.* Consideremos el conjunto finito $A = \{a_1, a_2, \ldots, a_k\}$, y $\Omega = A^{\mathbb{N}}$. Definamos la métrica $d : \Omega \times \Omega \to \mathbb{R}$ por

$$d(x, y) = \begin{cases} 1/l, & \text{si} \quad x \neq y, \text{ donde } l = \text{mín}\{i \in \mathbb{N};\ x_i \neq y_i\} \\ 0, & \text{si} \quad x = y, \end{cases}$$

donde $x = (x_i)_{i\in\mathbb{N}}$ y $y = (y_i)_{i\in\mathbb{N}}$ pertenecen a $\Omega$.

En efecto, $d$ es una métrica, $d$ es no negativa por definición. Y si $d(x, y) = 0$ entonces $x = y$, en virtud de la definición. El hecho que $d$ sea simétrica, esto es, $d(x, y) = d(y, x)$ para todo $x, y \in \Omega$, también se sigue de la definición. Ahora procedamos a mostrar la desigualdad triangular

$$d(x, z) \leq d(x, y) + d(y, z), \text{ para todo } x, y, z \in \Omega.$$

Consideremos $x, y, z \in \Omega$. Si $x = y$, $y = z$ o $x = z$ la desigualdad anterior es cierta. Supongamos entonces que $x$, $y$ y $z$ son distintos dos a dos. Sean

$$l_1 = \text{mín}\{i \in \mathbb{N};\ x_i \neq y_i\}, \text{ y}$$

$$l_2 = \text{mín}\{i \in \mathbb{N};\, y_i \neq z_i\},$$

para $j < \text{mín}\{l_1, l_2\}$ se tiene que $x_j = y_j = z_j$, luego $\text{mín}\{i \in \mathbb{N};\, x_i \neq z_i\} \geq \text{mín}\{l_1, l_2\}$, por lo tanto

$$\text{máx}\{1/l_1, 1/l_2\} = 1/\,\text{mín}\{l_1, l_2\} \geq d(x, z).$$

Como $1/l_1 + 1/l_2 \geq \text{máx}\{1/l_1, 1/l_2\}$, tenemos que

$$d(x, y) + d(y, z) \geq d(x, z).$$

□

**Lema 5.2.** *Sea $A$ un conjunto finito, y $\Omega = A^{\mathbb{N}}$ dotado con la métrica $d$ definida en el lema 5.1. Entonces $\big(\Omega, d\big)$ es compacto.*

*Demostración.* Ahora probemos que el espacio métrico $(\Omega, d)$ es compacto. Considerando el conjunto finito $A$ dotado de la topología discreta (aquella en la que cada elemento de $A$ constituye un abierto), tenemos que $A$ es compacto (por ser finito). Entonces $\Omega = A^{\mathbb{N}}$ es compacto en la topología producto debido al teorema de Tychonov. Recordemos que la topología producto es la menor topología en la que cada una de las proyecciones $p_n :\, A^{\mathbb{N}} \to A$, $p_n((x_1, x_2, \ldots)) = x_n$ es continua, esto es, $\Omega$ tiene como una subbase topológica a la colección

$$\{p_n^{-1}(a);\ n \in \mathbb{N},\, a \in A\}.$$

Veamos que la topología inducida en $\Omega$ por la métrica $d$ es la misma que la topología producto. Sea $B_R(x) = \{y \in \Omega;\, d(y, x) < R\}$ la bola centrada en $x$ y de radio $R$ según la métrica $d$.

- Si $R > 1$, $B_R(x) = \Omega = p_1^{-1}(A)$.
- Si $R = 1$, $B_R(x) = \{y \in \Omega;\, d(y, x) < 1\} = \{y \in \Omega;\, y_1 = x_1\} = p_1^{-1}(x_1)$.
- Si $0 < R < 1$ consideremos $l = \lfloor 1/R \rfloor \geq 1$, $l \leq 1/R < l + 1$, así $1/(l + 1) < R \leq 1/l$. Entonces

$$\begin{aligned} B_R(x) &= \big\{y \in \Omega;\, d(y, x) < R\big\} \\ &= \big\{y \in \Omega;\, d(y, x) \leq 1/(l + 1)\big\} \\ &= \big\{y \in \Omega;\, y_1 = x_1, y_2 = x_2, \ldots, y_l = x_l\big\} \\ &= p_1^{-1}(x_1) \cap p_2^{-1}(x_2) \cap \cdots \cap p_l^{-1}(x_l). \end{aligned}$$

Así, todo abierto en la topología métrica es abierto en la topología producto.

Recíprocamente, sea $p_j^{-1}(a)$ un elemento típico de la subbase de la topología producto antes mencionada.
Si $x \in p_j^{-1}(a)$ entonces $x_j = a$. Luego

$$B_{1/(j+1)}(x) \subset p_j^{-1}(a),$$

y así, $p_j^{-1}(a)$ es abierto en la topología métrica $(\Omega, d)$. □

Ahora enunciamos el teorema clave en la prueba del teorema de van der Waerden.

**Teorema 5.3** (recurrencia múltiple topológica de Furstenberg-Weiss)**.** *Sean $(X, d)$ un espacio métrico compacto y $T : X \to X$ una aplicación continua. Entonces para cada $k \in \mathbb{N}$ y cada $\epsilon > 0$ existen $x \in X$ y $n \in \mathbb{N}$ tales que*

$$d(T^{in}(x), x) < \epsilon \text{ para cada } i \in \{1, 2, \ldots, k\}.$$

*Más aún, dado $Z \subset X$ denso en $X$, podemos escoger $x$ en $Z$.*

**Lema 5.3.** *Sea $\Omega = A^{\mathbb{N}}$, donde $A$ es un conjunto finito, $\Omega$ dotado con la métrica $d$ del lema anterior. Definamos la función desplazamiento por*

$$\begin{array}{rcl} T : \Omega & \longrightarrow & \Omega \\ x = (x_1, x_2, x_3, \ldots) & \longmapsto & T(x) = (x_2, x_3, x_4, \ldots) \end{array}$$

*Entonces $T$ es una aplicación continua.*

*Demostración.* Veamos que $T$ es una aplicación continua en el espacio métrico $(\Omega, d)$. dados $x = (x_i)$ y $y = (y_i)$ en $\Omega$,

- si $x = y$, entonces de hecho $d\big(T(x), T(y)\big) \leq 2d(x, y)$. Puesto que $d\big(T(x), T(y)\big) = d(x, y) = 0$.
- Supongamos que $x \neq y$; entonces $d(x, y) = 1/l$, donde $l = \text{mín}\{i \in \mathbb{N};\, x_i \neq y_i\}$.
  - si $l = 1$, $x_1 \neq y_1$ y por definición $d(Tx, Ty) \leq 1$, luego $d(Tx, Ty) \leq 2d(x, y)$;
  - si $l = 2$, $x_1 = y_1$ y $x_2 \neq y_2$. así, $d(Tx, Ty) = 1$ y $d(x, y) = 1/2$. Por lo tanto

$$d(Tx, Ty) \leq 2d(x, y);$$

- si $l > 2$, $x_l \neq y_l$ y $x_i = y_i$ para $i \in \{1, 2, \ldots, l-1\}$. Así tenemos

$$(Tx)_{l-1} \neq (Ty)_{l-1} \text{ y } (Tx)_i = (Ty)_i \text{ para } i \in \{1, 2, \ldots, l-2\},$$

  por lo tanto $d(Tx, Ty) = 1/(l-1)$ y $d(x, y) = 1/l$.

En cualquier caso $d(Tx, Ty) \leq 2d(x, y)$, y por lo tanto $T$ es continua. □

*Prueba del teorema de van der Waerden.* Fijemos una función $\Phi : \mathbb{N} \to A = \{a_1, \ldots, a_k\}$, esta implica una manera de particionar los naturales, $\mathbb{N} = \Phi^{-1}(a_1) \cup \cdots \cup \Phi^{-1}(a_k)$. Denotaremos $\Phi = (c_i)_{i \in \mathbb{N}} = \big(c_1, c_2, c_3, \ldots\big)$, donde $c_i = \Phi(i)$. Consideremos la función desplazamiento

$$\begin{array}{rcl} T : \Omega & \longrightarrow & \Omega \\ x = (x_1, x_2, x_3, \ldots) & \longmapsto & T(x) = (x_2, x_3, x_4, \ldots), \end{array}$$

que es continua según el lema 5.3.
Dados $x = (x_i)$, $y = (y_i) \in \Omega = A^{\mathbb{N}}$ y $m, n \geq 0$ conviene resaltar que

$$d\big(T^m(x), T^n(y)\big) < 1 \quad \text{si, y solamente si,} \quad x_{m+1} = y_{n+1}, \tag{5.1}$$

puesto que, $d\big((x_{m+1}, x_{m+2}, \ldots), (y_{n+1}, y_{n+2}, \ldots)\big) = 1$ si, y solo si, $x_{m+1} \neq y_{n+1}$.

Consideremos el subespacio de $\Omega$, $X = \overline{\{T^m(c)\}_{m=1}^{\infty}}$. Este espacio es compacto puesto que es un subconjunto cerrado del espacio compacto $\Omega$. Como $T\big(\{T^m(c)\}_{m=1}^{\infty}\big) \subseteq \{T^m(c)\}_{m=1}^{\infty}$, por la continuidad de $T$ tenemos que $T\left(\overline{\{T^m(c)\}_{m=1}^{+\infty}}\right) \subseteq \overline{T\big(\{T^m(c)\}_{m=1}^{+\infty}\big)}$, y por lo tanto $T(X) \subseteq X$. Podemos considerar ahora la restricción de $T$ a $X$, que también denotaremos por $T : X \to X$. Por el teorema de recurrencia múltiple (teorema 5.3), dados $0 < \epsilon < 1$ y $k \geq 1$ un entero arbitrario, existen $p \in \big\{T^m(c)\big\}_{m=1}^{\infty}$ y $n \in \mathbb{N}$ tales que

$$d\big(T^{in}(p), p\big) < \epsilon \text{ para todo } i = 1, 2, \ldots, k,$$

esto es, existe $m_0 \in \mathbb{N}$ tal que $p = T^{m_0}(c)$, y

$$d\big(T^{in+m_0}(c), T^{m_0}(c)\big) < \epsilon, \text{ para todo } i = 1, 2, \ldots, k.$$

De la equivalencia dada en (5.1) tenemos que

$$c_{in+m_0+1} = c_{m_0+1}, \text{ para todo } i = 1, \ldots, k,$$

es decir

$$in + m_1 \in \Phi^{-1}(c_{m_1}), \text{ para todo } i = 1, \ldots, k,$$

donde $m_1 = m_0 + 1$. Esto demuestra el teorema. □

Para mostrar el teorema de Szemerédi basandonos en argumentos de teoría ergódica, necesitamos estudiar las iteraciones de una función continua $T : X \to X$ sobre un espacio topológico $X$ provisto de una medida boreliana $\mu$.

**Definición 5.2.** *Sea $\Omega$ un espacio topológico y $\mu$ una medida boreliana sobre $\Omega$. Una función continua $T : \Omega \to \Omega$ es $\mu$-invariante si*

$$\int_\Omega \psi d\mu = \int_\Omega \psi \circ T d\mu \text{ para cada } \psi \in C^0(\Omega). \tag{5.2}$$

Dada una aplicación continua $T : \Omega \to \Omega$, la presencia de una medida invariante $\mu$ respecto a $T$ da bastante información estadística sobre la estructura de las órbitas de la aplicación $T$, esto es, de los conjuntos $\{T^n(x)\}_{n=0}^\infty$ para casi todo $x \in X$ (respecto de la medida $\mu$). Por ejemplo tenemos el teorema de Poincaré.

**Teorema 5.4** (Poincaré)**.** *Sea $X$ un espacio topológico provisto de una medida boreliana $\mu$. Si $T : X \to X$ es una aplicación continua $\mu$-invariante y $A \subset X$ tal que $\mu(A) > 0$, entonces $\mu$-casi todo punto $x \in A$, existe $n(x) \in \mathbb{N}$ tal que $T^{n(x)}(x) \in A$*

Con las hipótesis y notaciones del teorema de Poincaré, existe $E \subseteq A$ con $\mu(E) = 0$, tal que para todo $x \in A \setminus E$ existe $n(x) \in \mathbb{N}$ tal que $x \in T^{-n(x)}(A)$. Así

$$A \setminus E \subset \bigcup_{N \in \mathbb{N}} \big[T^{-N}(A) \cap A\big].$$

Como $\mu(A \setminus E) = \mu(A) > 0$, tenemos que existe algún $N \in \mathbb{N}$ tal que

$$\mu(T^{-N}(A) \cap A) > 0.$$

Procediendo de modo análogo con $T^{-N}(A) \cap A$ en vez de $A$, obtenemos que existe $N_1 \in \mathbb{N}$ tal que

$$\mu(T^{-N_1}(A_1) \cap A_1) > 0, \text{ donde } A_1 = T^{-N}(A) \cap A.$$

En particular tenemos que

$$\mu\big(T^{-N_1-N}(A) \cap T^{-N}(A) \cap A\big) > 0.$$

Continuando de este modo obtenemos una sucesión creciente (estrictamente) $(n_k)_{k\in\mathbb{N}}$ de naturales tales que para cada $k \in \mathbb{N}$

$$\mu\big(T^{-n_k}(A) \cap T^{-n_{k-1}}(A) \cap \cdots \cap T^{-n_1}(A) \cap A\big) > 0. \tag{5.3}$$

Cada punto $x \in T^{-n_k}(A) \cap T^{-n_{k-1}}(A) \cap \cdots \cap T^{-n_1}(A) \cap A$ es *recurrente* en el sentido que $T^{n_i}(x) \in A$ para cada $i = 1, \ldots, k$, con $n_1 < n_2 < \cdots < n_k$. Más aún tenemos el siguiente teorema que nos dice que la secuencia $n_1 < n_2 < \cdots < n_k$, puede formar una progresión aritmética.

**Teorema 5.5** (recurrencia múltiple de Furstenberg)**.** *Sean $(X, \tau)$ un espacio topológico y $\mu$ una medida boreliana en $(X, \tau)$. Sea $T : X \to X$ una aplicación continua $\mu$-invariante, y consideremos $k \geq 3$ y $A \subseteq X$ con $\mu(A) > 0$, entonces existe $N \in \mathbb{N}$ tal que*

$$\mu\Big(A \cap T^{-N}(A) \cap T^{-2N}(A) \cap \cdots \cap T^{-(k-1)N}(A)\Big) > 0.$$

Una pregunta natural es que si dado un espacio topológico $(X, \tau)$, provisto de una medida boreliana $\nu$, y dada un aplicación continua $T : X \to X$, ¿existe alguna medida boreliana $\mu$ tal que la aplicación $T$ sea $\mu$-invariante? Cuando el espacio $X$ es compacto metrizable y $T$ es continua la respuesta es afirmativa. La idea de la prueba es simple: tomamos una medida boreliana cualquiera $\nu$ y veamos cómo esa medida cambia por la acción de $T$. Definamos $\nu_k = (T^k)^*(\nu)$, para cada $k \geq 0$ entero, como

$$\nu_k(A) = \nu\big(T^{-k}(A)\big)$$

Tomemos las medidas promedio $\mu_k = \dfrac{1}{k}\displaystyle\sum_{i=0}^{k-1} \nu_i$. Veremos que a medida que $k$ aumenta esta medida promedio $\mu_k$ tiende a quedar menos sensible a la acción de $T$. Es así que requerimos de alguna noción de convergencia de medidas.

**Teorema 5.6** (Teorema de Riesz-Markov)**.** *Sean $X$ un espacio topológico compacto de Hausdorff y $M(X)$ el conjunto de todas las medidas complejas borelianas (finitas) sobre $X$, con norma $||\mu|| = |\mu|(X)$, $\mu \in M(X)$, (aquí, $|\mu|(X)$ es la variación total de $\mu$). Entonces $C(X)^* = M(X)$, donde $C(X) = \{f : X \to \mathbb{R};\ f$ es continua $\}$. Especificamente, la aplicación*

$$\begin{aligned} G : M(X) &\to C(X)^* \\ \mu &\mapsto G_\mu \end{aligned}$$

*donde $G_\mu(\psi) = \displaystyle\int_X \psi d\mu$ para cada $\psi \in C(X)$, es una isometría lineal sobreyectiva, cumpliendo además que cualquier elemento positivo $f \in C(X)^*$ (esto es, $\phi \geq 0$ implica $f(\phi) \geq 0$) está asociado a una única medida positiva boreliana finita $\mu$ sobre $X$.*

**Observación 5.1.** *Sean $X$ un espacio de medida y $\mu$ una medida compleja sobre $X$. La variación total de $\mu$, que denotaremos por $|\mu|(X)$, se definido por*

$$|\mu|(X) = \sup\left\{\sum_{n=1}^{+\infty} |\mu(E_n)|;\ \{E_n\}_{n=1}^{+\infty} \text{ es una partición de conjuntos } \mu\text{-medibles de } X\right\}.$$

Como el espacio vectorial normado $M(X)$ es identificado isométricamente con el dual de $C(X)$, es natural usar la topología débil estrella (sobre $C(X)^*$) una vez que por el análisis funcional tenemos desde ya resultados de compacidad en cualquier bola cerrada en $C(X)^*$ que ayudará a garantizar la existencia de un límite de una sucesión de medidas en $M(X)$.

## 5.2. Prueba del teorema de Szemerédi

Ahora probemos el teorema de Szemerédi usando el teorema de recurrencia múltiple de Furstenberg.

*Prueba del teorema de Szemerédi 5.2.* Sea $A \subset \mathbb{N}$ un conjunto con densidad positiva, es decir,

$$\limsup_{N\to+\infty} \frac{|A \cap [1, N]|}{N} = \delta(A) > 0.$$

Así existe una sucesión estrictamente creciente $(k_n)_{n\in\mathbb{N}}$, tal que

$$\lim_{n\to+\infty} \frac{|A \cap [1, k_n]|}{k_n} = \delta(A) > 0. \tag{5.4}$$

Consideramos el conjunto finito $\{0, 1\}$ dotado con la topología discreta. Entonces $X = \{0, 1\}^{\mathbb{N}}$ con la topología producto $\tau$ es compacto debido al teorema de Tychonov, y $(X, \tau)$ es metrizable con la métrica $d : X \times X \to \mathbb{R}$ definida por

$$d(x, y) = \begin{cases} 1/\operatorname{mín}\{i \in \mathbb{N};\ x_i \neq y_i\} & , \text{si } x \neq y \\ 0 & , \text{si } x = y, \end{cases}$$

donde $x = (x_i), y = (y_i) \in X$. Así, $X$ es métrico compacto.
Sea $T : X \to X$ la aplicación desplazamiento definida por

$$T\big((z_1, z_2, z_3, \ldots)\big) = (z_2, z_3, z_4, \ldots), \text{ donde } (z_n)_{n\in\mathbb{N}} \in X.$$

Como mostramos antes, $T$ es continua. Sea $x = (x_n)_{n\in\mathbb{N}}$ definido por

$$x_n = 1_A(n), \text{ para cada } n \in \mathbb{N}.$$

Dado $\xi \in X$ definamos $\tilde{\delta}_\xi \in C(X)^*$, por

$$\tilde{\delta}_\xi(\psi) = \psi(\xi), \text{ para todo } \psi \in C(X).$$

Como $|\tilde{\delta}_\xi(\psi)| = |\psi(\xi)| \leq \sup_{\alpha \in X} |\psi(\alpha)| = ||\psi||$, en efecto tenemos que $\tilde{\delta}_\xi \in C(X)^*$. Más aún

$$||\tilde{\delta}_\xi|| \leq 1, \text{ para cada } \xi \in X.$$

Gracias a la isometría $G : M(X) \to C(X)^*$ (dada por el teorema de Riesz-Markov), definida por $G(\mu) = G_\mu$ para cada $\mu \in M(X)$, y

$$G_\mu(\psi) = \int_X \psi d\mu, \text{ para cada } \psi \in C(X),$$

tenemos que para cada $\xi \in X$, existe una medida $\delta_\xi \in M(X)$ tal que $G(\delta_\xi) = \tilde{\delta}_\xi$, esto es,

$$\int_X \psi d\delta_\xi = G_{\delta_\xi}(\psi) = \tilde{\delta}_\xi(\psi) = \psi(\xi), \text{ para todo } \psi \in C(X).$$

Dado $B \in \tau$, tenemos que si $1_B$, la función característica de $B$, pertenece a $C(X)$, entonces

$$\delta_\xi(B) = \int_X 1_B d\delta_\xi = G_{\delta_\xi}(1_B) = \tilde{\delta}_\xi(1_B) = 1_B(\xi).$$

Así,

$$\delta_\xi(B) = \begin{cases} 1 & \text{, si} \quad \xi \in B \\ 0 & \text{, si} \quad \xi \notin B. \end{cases}$$

Para cada $n \in \mathbb{N}$ consideremos las medidas complejas borelianas (finitas)

$$\mu_n = \frac{1}{n} \sum_{j=0}^{n-1} \delta_{T^j(x)}.$$

Como $G$ es una isometría y $||\tilde{\delta}_\xi|| \leq 1$ para cada $\xi \in X$, tenemos que $||\delta_{T^j(x)}|| \leq 1$, para cada $j \in \mathbb{N}$. En consecuencia,

$$||\mu_n|| \leq \frac{1}{n} \sum_{j=0}^{n-1} ||\delta_{T^j(x)}|| \leq 1, \text{ para cada } n \in \mathbb{N}.$$

Además, por la linealidad de $G$ tenemos que $G(\mu_n) = \frac{1}{n} \sum_{j=0}^{n-1} \tilde{\delta}_{T^j(x)}$. Sea $\tilde{\mu}_n = G(\mu_n)$. Por el corolario 2.1, $\overline{B}_{C(X)^*}(0;1)$ es secuencialmente compacto en la topología débil estrella de $C(X)^*$. Así, existe una subsucesión $(\tilde{\mu}_{i_n})_{n \in \mathbb{N}}$ de $(\tilde{\mu}_{k_n})_{n \in \mathbb{N}}$, donde $(k_n)_n$ es la sucesión creciente (estrictamente) que aparece en (5.4), y existe $\tilde{\nu} \in C(X)^*$ (y por tanto una medida boreliana

$\nu \in M(X)$ tal que $G(\nu) = \tilde{\nu}$), tales que $(\tilde{\mu}_{i_n})_{n\in\mathbb{N}}$ converge a $\tilde{\nu}$ en la topología débil estrella de $C(X)^*$, esto es,

$$\begin{aligned}\int_X \psi d\nu &= [G(\nu)](\psi) = \tilde{\nu}(\psi) = \widehat{\psi}(\tilde{\nu}) = \lim_{n\to+\infty} \widehat{\psi}(\tilde{\mu}_{i_n}) = \lim_{n\to+\infty} \tilde{\mu}_{i_n}(\psi) \\ &= \lim_{n\to+\infty} [G(\mu_{i_n})](\psi) = \lim_{n\to+\infty} \int_X \psi d\mu_{i_n}.\end{aligned}$$

Por lo tanto

$$\lim_{n\to+\infty} \int_X \psi d\mu_{i_n} = \int_X \psi d\nu \text{ para todo } \psi \in C(X). \tag{5.5}$$

Además, por el teorema de Riesz-Markov $T$ es $\nu$-invariante. Sea

$$Y = \big\{y = (y_n)_{n\in\mathbb{N}} \in X;\, y_1 = 1\big\} = p_1^{-1}\big(\{1\}\big).$$

$Y$ es un conjunto abierto de $(X, \tau)$, también $Y^c = p_1^{-1}\big(\{0\}\big)$ pertenecen a $\tau$. Por lo tanto $1_Y \in C(X)$, y de (5.5),

$$\lim_{n\to+\infty} \mu_{i_n}(Y) = \nu(Y), \tag{5.6}$$

esto es,

$$\nu(Y) = \lim_{n\to+\infty} \frac{1}{i_n} \sum_{j=0}^{i_n-1} \delta_{T^j(x)}(Y) = \lim_{n\to+\infty} \frac{1}{i_n} \sum_{j=1}^{i_n} \delta_{T^{j-1}(x)}(Y).$$

Como $T^{j-1}(x) = (x_j, x_{j+1}, x_{j+2}, \ldots)$ y $Y = p_1^{-1}(\{1\})$, tenemos que

$$\delta_{T^{j-1}(x)}(Y) = 1 \text{ si, y solo si } x_j = 1,$$

esto es,

$$\delta_{T^{j-1}(x)}(Y) = 1 \text{ si, y solo si } j \in A,$$

puesto que $x_j = 1_A(j)$. Por lo tanto

$$\mu_{i_n}(Y) = \frac{1}{i_n} \sum_{j=0}^{i_n-1} \delta_{T^j(x)}(Y) = \frac{1}{i_n} \sum_{j=1}^{i_n} \delta_{T^{j-1}(x)}(Y) = \frac{|A \cap [1, i_n]|}{i_n},$$

y así, de (5.4) y (5.6)

$$0 < \delta(A) = \nu(Y). \tag{5.7}$$

Entonces, por el teorema de recurrencia múltiple de Furstenberg (teorema 5.5), y puesto que $T$ es $\nu$-invariante, existe $N \in \mathbb{N}$ tal que

$$\nu\Big(Y \cap T^{-N}(Y) \cap T^{-2N}(Y) \cap \cdots \cap T^{-(k-1)N}(Y)\Big) = \delta' > 0.$$

Por ende

$$\lim_{n\to+\infty} \mu_{i_n}\Big(Y \cap T^{-N}(Y) \cap T^{-2N}(Y) \cap \cdots \cap T^{-(k-1)N}(Y)\Big) = \delta' > 0.$$

Entonces existe $n_0 \in \mathbb{N}$ tal que para $n \geq n_0$ tenemos

$$\mu_{i_n}\Big(Y \cap T^{-N}(Y) \cap T^{-2N}(Y) \cap \ldots \cap T^{-(k-1)N}(Y)\Big) > \delta'/2 > 0,$$

esto es,

$$\frac{1}{i_n}\sum_{j=0}^{i_n-1} \delta_{T^j(x)}\Big(Y \cap T^{-N}(Y) \cap T^{-2N}(Y) \cap \ldots \cap T^{-(k-1)N}(Y)\Big) > \delta'/2 > 0.$$

Por tanto, existe $j \in \mathbb{N}$ (más aún, existen tantos $j \in \{0, 1, \ldots, i_n - 1\}$ como $i_n \cdot \delta'/2$) tal que

$$T^j(x) \in Y \cap T^{-N}(Y) \cap T^{-2N}(Y) \cap \cdots \cap T^{-(k-1)N}(Y),$$

esto es,

$$x_{j+1} = 1,\, x_{j+N+1} = 1,\, x_{j+2N+1} = 1,\, \cdots,\, x_{j+(k-1)N+1} = 1,$$

es decir,

$$j_0, j_0 + N, j_0 + 2N, \cdots, j_0 + (k-1)N \in A,$$

donde $j_0 = j + 1$. □

# Capítulo 6

# El teorema de Szemerédi relativo a una medida pseudoaleatoria

Probaremos que existen progresiones aritméticas propias, esto es con diferencia común diferente de cero, y de longitud arbitraria, formadas solamente por números primos. Para ello abordaremos el camino seguido por Green y Tao, quienes formularon, en [7], una variante del teorema de Szemerédi en un ámbito más general, el *teorema de Szemerédi relativo a medidas pseudoaleatorias*, el cual probaremos al final de esta sección. Esta versión del teorema de Szemerédi permite trabajar con subconjuntos de los números naturales que, a pesar de que tengan densidad superior nula respecto a $\mathbb{N}$, posean densidad relativa positiva respecto a algún subconjunto de $\mathbb{N}$. Este es precisamente el caso de los números primos $\mathcal{P}$, que por el teorema del número primo, tiene densidad superior nula respecto a $\mathbb{N}$. Otro ingrediente importante para probar el teorema de Green-Tao es mostrar la existencia de una medida pseudoaleatoria que mayore a cierto subconjunto de los números primos, hecho que mostraremos en la siguiente sección, y que fue motivado principalmente por los trabajos de Goldston y Yıldırım, (cf. [8]).

## 6.1. Introducción

En lo que resta de la presente sección el teorema de Szemerédi relativo a una medida pseudoaleatoria.

**Teorema 6.1** (Green-Tao)**.** *Los números primos contienen infinitas progresiones aritméticas*

*de longitud $k$ para todo $k \in \mathbb{N}$, esto es, para todo $k \in \mathbb{N}$, el conjunto*

$$\left\{ \{a + i \cdot r\}_{i=0}^{k-1} \subset \mathcal{P} : \ a, r \in \mathbb{N} \right\}$$

*es infinito, donde $\mathcal{P} \subset \mathbb{N}$ es el conjunto de números primos.*

De hecho podemos mostrar un resultado un poco más fuerte que involucra a subconjuntos de los números primos con densidad relativa positiva.

**Teorema 6.2** (Szemerédi en los primos)**.** *Sea $A$ un subconjunto cualquiera de los números primos, $\mathcal{P}$, con densidad superior relativa positiva, esto es,*

$$\limsup_{N \to \infty} \frac{\left| A \cap [1, N] \right|}{\left| \mathcal{P} \cap [1, N] \right|} > 0.$$

*Entonces $A$ contiene infinitas progresiones aritméticas de longitud $k$ para todo $k \geq 3$ entero.*

Observemos que el teorema 6.2 no se sigue directamente del teorema de Szemerédi clásico. En efecto, el conjunto de los números primos $\mathcal{P}$ está demasiado esparcido como para deducir la existencia de progresiones aritméticas en $\mathcal{P}$. Del teorema del número primo, tenemos que

$$\lim_{n \to +\infty} \frac{|\mathcal{P} \cap [1, N]| \ln(N)}{N} = 1,$$

y por lo tanto $\lim\limits_{n \to +\infty} \dfrac{|\mathcal{P} \cap [1, N]|}{N} = 0$, esto es, $\mathcal{P}$ tiene densidad superior nula respecto a $\mathbb{N}$.
Si reemplazamos el conjunto de los números primos $\mathcal{P}$ en la afirmación del teorema 6.2 por el conjunto de enteros positivos $\mathbb{Z}^+$, entonces resultaría el famoso teorema de Szemerédi (versión clásica).
Las siguientes versiones del teorema de Szemerédi las llamaremos versión finitaria, y funcional, respectivamente.

**Proposición 6.1** (Szemerédi, versión finitaria)**.** *Sean $N$ un entero positivo y $\mathbb{Z}_N = \mathbb{Z}/N\mathbb{Z}$. Sea $1 \geq \delta > 0$ un número real fijado, y sea $k \geq 3$ un entero. Entonces existe un menor natural $N_0(\delta, k) \in \mathbb{Z}^+$ con la propiedad de que, si $N \geq N_0(\delta, k)$ y $A \subseteq \mathbb{Z}_N$ con $|A| \geq \delta N$, entonces tenemos que $A$ contiene alguna progresión aritmética de longitud $k$.*

**Proposición 6.2** (Szemerédi, versión funcional)**.** *Sea $\nu_{\text{const}} : \mathbb{Z}_N \to \mathbb{R}^+$ la función constante unidad, $\nu_{\text{const}} \equiv 1$. Sean $0 < \delta \leq 1$ y $k \geq 1$ un entero fijado. Consideremos también $N \in \mathbb{Z}^+$ un parámetro entero grande y una función $f : \mathbb{Z}_N \to \mathbb{R}^+$ tal que*

$$0 \leq f(x) \leq \nu_{\text{const}}(x), \text{ para todo } x \in \mathbb{Z}_N, \tag{6.1}$$

*y*

$$\mathbb{E}(f(x)|x \in \mathbb{Z}_N) \geq \delta. \tag{6.2}$$

*Entonces tenemos que*

$$\mathbb{E}\Big(f(x)f(x+r)\cdots f\big(x+(k-1)r\big)\Big|x, r \in \mathbb{Z}_N\Big) \geq c(k,\delta) - o_{k,\delta}(1)$$

*para alguna constante $c(k,\delta) > 0$ que no depende ni de $f$ ni de $N$.*

**Observación 6.1.** *Desconociendo por un momento la curiosa notación para la función constante $\nu_{\text{const}}$, existen dos principales diferencias entre esta última proposición y la proposición 5.2. Una es el hecho que estamos tratando con funciones en vez que con conjuntos. Y por otro lado, si desentrañamos el significado de la notación $\mathbb{E}$ vemos que ahora estamos afirmando la existencia de $\gg N^2$ progresiones aritméticas, y no solamente una.*

Esta proposición es equivalente al teorema de Szemerédi en su formulación finitaria. En efecto, consideremos un entero $k \geq 3$, y $\delta \in ]0,1]$. Dado $A \subseteq \mathbb{Z}_N$ tal que $|A|/N \geq \delta$, podemos considerar la *función característica* de $A$, $\chi_A : \mathbb{Z}_N \to \mathbb{R}$, definida por

$$\chi_A(a) = \begin{cases} 1 & \text{, si } a \in A, \\ 0 & \text{, si } a \notin A. \end{cases}$$

Así, puesto que se tiene que $0 \leq \chi_A(a) \leq 1$, para todo $a \in \mathbb{Z}_N$, y $\mathbb{E}\big(\chi_A(x)\big|x \in \mathbb{Z}_N\big) = \dfrac{|A|}{N} \geq \delta$, tenemos que existe $c(k,\delta) > 0$ tal que

$$\mathbb{E}\Big(\chi_A(x)\chi_A(x+r)\cdots\chi_A\big(x+(k-1)r\big)\Big|x, r \in \mathbb{Z}_N\Big) \geq c(k,\delta) - o_{k,\delta}(1).$$

Por lo tanto, existe $N_0 \in \mathbb{N}$, tal que si $N \geq N_0$ entonces

$$\mathbb{E}\Big(\chi_A(x)\chi_A(x+r)\cdots\chi_A\big(x+(k-1)r\big)\Big|x, r \in \mathbb{Z}_N\Big) \geq c(k,\delta)/2 > 0.$$

Por tanto, para cada $N \geq N_0$, el número de pares $(x,r)$ en $\mathbb{Z}_N \times \mathbb{Z}_N$ tales que

$$x, x+r, x+2r, \ldots, x+(k-1)r \in A,$$

es mayor que $N^2\big(c(k,\delta)/2\big)$.

**Definición 6.1.** *La función $\nu : \mathbb{Z}_N \to \mathbb{R}^+$ es* **una medida** *si*

$$\mathbb{E}(\nu) = 1 + o(1), \tag{6.3}$$

*donde $\mathbb{R}^+$ denota al conjunto de los números reales no negativos.*

## 6.2. Medidas pseudoaleatorias y notaciones

En esta sección introduciremos la noción de medida pseudoaleatoria sobre $\mathbb{Z}_N$.

**Definición 6.2.** *Sea $A$ un conjunto finito no vacío (que usualmente consideraremos como $\mathbb{Z}_N$) y sea una función $f : A \to \mathbb{R}$. El* **valor promedio** *de $f$, el cual denotaremos con $\mathbb{E}(f)$, es definido por*

$$\mathbb{E}(f) = \frac{1}{|A|}\sum_{x\in A} f(x),$$

*donde $|A|$ denota la cardinalidad de $A$. De manera más general, si $P(x)$ es una afirmación que concierne a elementos de $A$ y que es cierta para al menos un elemento de $A$, entonces definimos el* **valor promedio de $f$ sujeto a $P$** *por*

$$\mathbb{E}\big(f(x)\big|P(x)\big) = \frac{\sum_{x\in A, P(x)} f(x)}{\big|\{x \in A : P(x)\}\big|}.$$

**Definición 6.3** (Condición de formas lineales)**.** *Sea $\nu : \mathbb{Z}_N \to \mathbb{R}^+$ una medida. Sean $m_0$, $t_0$ y $L_0$ enteros positivos.*

*Diremos que $\nu$ satisface la $(m_0, t_0, L_0)$-***condición de formas lineales** *si se cumple lo siguiente: sean $m \le m_0$ y $t \le t_0$ cualesquiera enteros positivos y supongamos que $(L_{i,j})_{1\le i\le m, 1\le j\le t}$ sea una colección de números racionales arbitrarios con numerador y denominador menores que $L_0$ en valor absoluto. Sean también $b_i$, $1 \le i \le m$, elementos arbitrarios de $\mathbb{Z}_N$. Para cada $1 \le i \le m$ sean las formas lineales $\psi_i : \mathbb{Z}_N^t \to \mathbb{Z}_N$ definidas por $\psi_i(x) = \sum_{j=1}^t L_{i,j}x_j + b_i$, donde $x = (x_1, x_2, \ldots, x_t)$ y los números racionales $L_{i,j}$ son interpretados como elementos de $\mathbb{Z}_N$ en la manera usual (asumiendo que $N$ es un número primo suficientemente grande, $N > L_0$). Supongamos que para $i \in \{1, \ldots, m\}$ las $t$-uplas $(L_{i,j})_{1\le j\le t} \in \mathbb{Q}^t$, $1 \le i \le m$, son no nulas y ninguna $t$-upla es múltiplo racional de ninguna otra. Entonces tenemos que*

$$\mathbb{E}\Big(\nu\big(\psi_1(x)\big)\cdots\nu\big(\psi_m(x)\big)\Big|x \in \mathbb{Z}_N^t\Big) = 1 + o_{L_0,m_0,t_0}(1). \tag{6.4}$$

*Donde el término $o(1)$ no depende de la elección de $b_1, \ldots, b_m$.*

**Observación 6.2.** *El parámetro $m_0$ que controla el número de formas lineales es el más importante y será mantenido relativamente pequeño. El caso $m = 1$ de la condición de formas lineales contiene la condición de que $\nu$ en la ecuación (6.3) sea una medida. Otros ejemplos de la condición de formas lineales que encomtraremos después son*

$$\mathbb{E}\big(\nu(x)\nu(x+h_1)\nu(x+h_2)\nu(x+h_1+h_2)\big|x, h_1, h_2 \in \mathbb{Z}_N\big) = 1 + o(1), \tag{6.5}$$

*aquí* $(m_0, t_0, L_0) = (4, 3, 1)$.

$$\mathbb{E}\big(\nu(x+h_1)\nu(x+h_2)\nu(x+h_1+h_2)\big|h_1,h_2\in\mathbb{Z}_N\big)=1+o(1), \tag{6.6}$$

*para todo* $x \in \mathbb{Z}_N$, *aquí* $(m_0, t_0, L_0) = (3, 2, 1)$. *Y*

$$\begin{aligned}\mathbb{E}\Big(\nu\big((x-y)/2\big)\nu\big((x-y+h_2)/2\big)\nu(-y)\nu(-y+h_1)\times\\ \times\nu\big((x-y')/2\big)\nu\big((x-y'+h_2)/2\big)\nu(-y')\nu(-y'+h_1)\times\\ \times\,\nu(x)\nu(x+h_1)\nu(x+h_2)\nu(x+h_1+h_2)\Big|x,h_1,h_2,y,y'\in\mathbb{Z}_N\Big)=1+o(1),\end{aligned} \tag{6.7}$$

*aquí* $(m_0, t_0, L_0) = (12, 5, 2)$.

**Definición 6.4** (Condición de correlación)**.** *Sea* $\nu : \mathbb{Z}_N \to \mathbb{R}^+$ *una medida y sea* $m_0 \in \mathbb{Z}^+$. *Diremos que* $\nu$ *satisface la* $m_0$**-condición de correlación** *si para cada* $1 < m \leq m_0$ *existe una función peso* $\tau = \tau_m : \mathbb{Z}_N \to \mathbb{R}^+$ *tal que se satisface la condición de momentos*

$$\mathbb{E}(\tau^q) = O_{m,q}(1) \tag{6.8}$$

*para todo* $q \geq 1$, *y tal que*

$$\mathbb{E}\Big(\nu\big(x+h_1\big)\nu\big(x+h_2\big)\cdots\nu\big(x+h_m\big)\Big|x\in\mathbb{Z}_N\Big)\leq\sum_{1\leq i<j\leq m}\tau(h_i-h_j) \tag{6.9}$$

*para todo* $h_1, h_2, \ldots, h_m \in \mathbb{Z}_N$ *(no necesariamente distintos).*

**Definición 6.5** (Medida pseudoaleatoria)**.** *Sea* $\nu : \mathbb{Z}_N \to \mathbb{R}^+$ *una medida. Decimos que* $\nu$ *es una* **medida** $k$**-pseudoaleatoria** *si satisface la* $(k2^{k-1}, 3k-4, k)$*-condición de formas lineales y la* $2^{k-1}$*-condición de correlación.*

La función $\nu_{\text{const}} \equiv 1$ es de hecho una medida $k$-pseudoaleatoria para cualquier $k \in \mathbb{N}$.

**Lema 6.1.** *Sea* $\nu$ *una medida* $k$*-pseudoaleatoria. Entonces* $\nu_{1/2} = (\nu + \nu_{\text{const}})/2 = (\nu + 1)/2$ *es también una medida* $k$*-pseudoaleatoria (aunque posiblemente con límites* $O(\cdot)$ *y* $o(\cdot)$ *un poco diferentes).*

*Demostración.* Es claro que $\nu_{1/2}$ es no negativa, y como $\mathbb{E}(\nu) = 1 + o(1)$ tenemos que $\nu_{1/2}$ es una medida, esto es, $\mathbb{E}(\nu_{1/2}) = 1 + o(1)$. Ahora verifiquemos la condición de formas lineales.

$$\mathbb{E}\left(\prod_{i=1}^{m}\big((\nu+1)/2\big)\big(\phi_i(x)\big)\Big|x\in\mathbb{Z}_N^t\right)=\mathbb{E}\left(\frac{1}{2^m}\sum_{A\subseteq\{1,\ldots,m\}}\prod_{i\in A}\nu\big(\phi_i(x)\big)\Big|x\in\mathbb{Z}_N^t\right)$$

$$= \frac{1}{2^m} \sum_{A\subseteq\{1,\dots,m\}} \mathbb{E}\Big( \prod_{i\in A} \nu\big(\phi_i(x)\big) \Big| x \in \mathbb{Z}_N^t \Big)$$
$$= \frac{1}{2^m} \sum_{A\subseteq\{1,\dots,m\}} \big(1 + o(1)\big)$$
$$= 1 + o(1),$$

puesto que cada término $\mathbb{E}\Big( \prod_{i\in A} \nu\big(\phi_i(x)\big) \Big| x \in \mathbb{Z}_N^t \Big)$ es $1 + o(1)$ para cada $A \subseteq \{1, \dots, m\}$, debido a la $(k2^{k-1}, 3k-4, k)$-condición de formas lineales, ecuación (6.4). Similarmente verificamos la condición de correlación para $\nu_{1/2}$. Sean $1 \leq q < \infty$ y $h_1, \dots, h_m \in \mathbb{Z}_N$

$$\begin{aligned}
\mathbb{E}\Big(\nu_{1/2}\big(x+h_1\big)\cdots\nu_{1/2}\big(x+h_m\big)\Big|x\in\mathbb{Z}_N\Big) &= \frac{1}{2^m}\sum_{A\subseteq\{1,\dots,m\}} \mathbb{E}\Big(\prod_{i\in A}\nu(x+h_i)\Big|x\in\mathbb{Z}_N\Big)\\
&\leq \frac{1}{2^m}\sum_{A\subseteq\{1,\dots,m\}}\ \sum_{i,j\in A, i<j} \tau(h_i-h_j)\\
&\leq \frac{1}{2^m}\sum_{A\subseteq\{1,\dots,m\}}\ \sum_{1\leq i<j\leq m} \tau(h_i-h_j)\\
&\leq \sum_{1\leq i<j\leq m} \tau(h_i-h_j).
\end{aligned}$$

□

**Teorema 6.3** (Szemerédi relativo a medidas pseudoaleatorias)**.** *Sean $k \geq 3$ un entero y $0 < \delta \leq 1$ un parámetro fijado. Supongamos que $\nu : \mathbb{Z}_N \to \mathbb{R}^+$ es una medida $k$-pseudoaleatoria. Sean $N \geq 1$ un parámetro entero grande y $f : \mathbb{Z}_N \to \mathbb{R}^+$, una función cualquiera no negativa tal que*

$$0 \leq f(x) \leq \nu(x) \text{ para todo } x \in \mathbb{Z}_N \tag{6.10}$$

*y*

$$\mathbb{E}(f) \geq \delta. \tag{6.11}$$

*Entonces tenemos que*

$$\mathbb{E}\Big(f(x)f(x+r)\cdots f\big(x+(k-1)r\big)\Big|x,r\in\mathbb{Z}_N\Big) \geq c(k,\delta) - o_{k,\delta}(1), \tag{6.12}$$

*donde $c(k,\delta) > 0$ es la misma constante que aparece en la proposición 6.2.*

A continuación estableceremos algunas notaciones que usaremos.

**Definición 6.6.** *Para cada* $1 \leq q \leq \infty$ *y* $f : \mathbb{Z}_N \to \mathbb{R}$ *definimos la* **norma** $L^q$ **sobre** $\mathbb{Z}_N$ *como*

$$||f||_{L^q} = \mathbb{E}\big(|f|^q\big)^{1/q},$$

*con la convención* $||f||_{L^\infty} = \sup_{x \in \mathbb{Z}_N} |f(x)|$.

Denotaremos con $L^q(\mathbb{Z}_N)$ al espacio de Banach de todas las funciones de $\mathbb{Z}_N$ en $\mathbb{R}$, equipado con la norma $L^q$; de hecho como $\mathbb{Z}_N$ es finito todas las normas $L^q$, $q \geq 1$ son equivalentes. Observamos también que $L^2(\mathbb{Z}_N)$ es un espacio de Hilbert real con el producto interno usual

$$\langle f, g \rangle = \mathbb{E}(fg).$$

Si $\Omega$ es un subconjunto de $\mathbb{Z}_N$, denotaremos con $1_\Omega : \mathbb{Z}_N \to \mathbb{R}$ a la función característica de $\Omega$, esto es $1_\Omega(x) = 1$ si $x \in \Omega$, y $1_\Omega(x) = 0$ en otro caso. Similarmente si $P(x)$ es una proposición lógica concerniente a elementos $x \in \mathbb{Z}_N$, escribiremos $1_{P(x)}$ en vez de $1_{\{x \in \mathbb{Z}_N : P(x)\}}$.
Emplearemos frecuentemente cambio de variables lineales. Para facilitar esto, consideramos la siguiente definición.

**Definición 6.7.** *Sean* $A$ *y* $B$ *conjuntos finitos no vacíos y* $\Phi : A \to B$ *una función. Decimos que* $\Phi$ *es un* **cubrimiento uniforme** *de* $B$ *por* $A$ *si* $\Phi$ *es sobreyectiva y todas las fibras* $\Phi^{-1}(b)$*, con* $b \in B$*, tienen la misma cardinalidad (a saber,* $|A|/|B|$*).*

Observe que si $\Phi : A \to B$ es un cubrimiento uniforme de $B$ por $A$ entonces para cualquier función $f : B \to \mathbb{R}$ tenemos

$$\mathbb{E}\Big(f\big(\Phi(a)\big)\Big|a \in A\Big) = \mathbb{E}\big(f(b)\big|b \in B\big). \tag{6.13}$$

En efecto, como $\Phi : A \to B$ es un cubrimiento. Entonces $A = \bigcup_{b \in B} \Phi^{-1}(b)$, y por lo tanto $|A| = |B|r$, donde $r = \big|\Phi^{-1}(b)\big|$ para cada $b \in B$. Así,

$$\begin{aligned}\mathbb{E}\left(f\big(\Phi(a)\big)\Big|a \in A\right) &= \frac{1}{|A|}\sum_{a \in A} f\big(\Phi(a)\big) = \frac{1}{|A|}\sum_{b \in B}\sum_{a \in \Phi^{-1}(b)} f\big(\Phi(a)\big) \\ &= \frac{1}{|A|}\sum_{b \in B}\big|\Phi^{-1}(b)\big|f\big(b\big) = \frac{r}{|A|}\sum_{b \in B} f(b) = \mathbb{E}\left(f(b)\big|b \in B\right).\end{aligned}$$

## 6.3. Normas de Gowers uniformes y el teorema generalizado de von Neumann

La prueba del teorema 6.3 depende de la descomposición de una función dada $f$ en una componente uniforme según Gowers y otra componente anti-uniforme según Gowers. Lle-

garemos a esta descomposición en las siguientes secciones. En esta sección definiremos la noción de uniformidad según Gowers. El principal resultado de esta sección será el teorema generalizado de von Neumann (proposición 6.3), el cual afirma que las funciones uniformes según Gowers son despreciables respecto al cálculo de sumas tales como las que aparecen en la ecuación (6.12).

Sea $\nu$ una medida $k$-pseudoaleatoria y sea $1 \leq d \leq k-1$. Se tiene que

$$||\nu||_{U^d}^{2^d} = \mathbb{E}\left( \prod_{\omega \in \{0,1\}^d} \nu(x + h \cdot \omega) \middle| x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^d \right) = 1 + o(1),$$

esto se sigue de la $(k2^{k-1}, 3k-4, k)$-condición de formas lineales. Más aún, se tiene el siguiente resultado.

**Lema 6.2.** *Supongamos que $\nu$ sea una medida $k$-pseudoaleatoria. Entonces*

$$||\nu - \nu_{\text{const}}||_{U^d} = ||\nu - 1||_{U^d} = o(1) \tag{6.14}$$

*para todo $1 \leq d \leq k-1$.*

*Demostración.* De (2.17) es suficiente probar la afirmación para $d = k-1$. De la definición de norma de Gowers (2.12), elevando a la potencia $2^{k-1}$, es suficiente mostrar que

$$\mathbb{E}\left( \prod_{\omega \in \{0,1\}^{k-1}} \big(\nu(x + \omega \cdot h) - 1\big) \middle| x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \right) = o(1).$$

Desarrollando el producto $\prod_{\omega \in \{0,1\}^{k-1}} \Big(\nu(x + \omega \cdot h) - 1\Big)$ que aparece en el lado izquierdo de la igualdad anterior, obtemos

$$\begin{aligned} \sum_{A \subseteq \{0,1\}^{k-1}} \prod_{\omega \in A} \nu(x + \omega \cdot h) \prod_{\omega \in \{0,1\}^{k-1} \setminus A} (-1) \\ &= \sum_{A \subseteq \{0,1\}^{k-1}} (-1)^{2^{k-1} - |A|} \prod_{\omega \in A} \nu(x + \omega \cdot h) \\ &= \sum_{A \subseteq \{0,1\}^{k-1}} (-1)^{|A|} \prod_{\omega \in A} \nu(x + \omega \cdot h), \end{aligned}$$

por lo tanto $\mathbb{E}\left( \prod_{\omega \in \{0,1\}^{k-1}} \big(\nu(x + \omega \cdot h) - 1\big) \middle| x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \right)$ es igual a

$$\sum_{A \subseteq \{0,1\}^{k-1}} (-1)^{|A|} \mathbb{E}\left( \prod_{\omega \in A} \nu(x + \omega \cdot h) \middle| x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \right). \tag{6.15}$$

Observemos que la expresión

$$\mathbb{E}\Big( \prod_{\omega\in A} \nu(x+\omega\cdot h)\Big| x\in\mathbb{Z}_N, h\in\mathbb{Z}_N^{k-1}\Big) \tag{6.16}$$

para $A\subseteq\{0,1\}^{k-1}$ fijado ($A$ tiene a lo más $2^{k-1}$ elementos), es de la forma

$$\mathbb{E}\Big(\nu\big(\psi_1(\xi)\big)\cdots\nu\big(\psi_{|A|}(\xi)\big)\Big|\xi\in\mathbb{Z}_N^k\Big),$$

donde $\xi=(x,h_1,\ldots,h_{k-1})$ y las $\psi_i$ son las $|A|$ formas lineales $\xi\mapsto x+\omega_1h_1+\cdots+\omega_{k-1}h_{k-1}$ dadas en algún orden. Cada $\omega=(\omega_1,\ldots,\omega_{k-1})\in A$ define solo una de estas formas lineales. Es claro que ninguna de estas formas lineales es un múltiplo racional de alguna otra porque para dos elementos distintos de $A$, $\omega=(\omega_1,\ldots,\omega_{k-1})$ y $\omega'=(\omega'_1,\ldots,\omega'_{k-1})$, existe algún $i$, $1\leq i\leq k-1$, tal que o $\omega_i=0$ y $\omega'_i=1$, o $\omega_i=1$ y $\omega'_i=0$. Por lo tanto, invocando la $(2^{k-1},k,1)$-condición de formas lineales, que es consecuencia del hecho que $\nu$ es una medida $k$-pseudoaleatoria, concluimos que

$$\mathbb{E}\Big( \prod_{\omega\in A} \nu(x+\omega\cdot h)\Big| x\in\mathbb{Z}_N, h\in\mathbb{Z}_N^{k-1}\Big) = 1+o(1).$$

Por lo tanto, de (6.15) y del hecho que

$$\sum_{A\subseteq\{0,1\}^{k-1}}(-1)^{|A|}=\sum_{n=0}^{2^{k-1}}\binom{2^{k-1}}{n}(-1)^n=(1-1)^{2^{k-1}}=0,$$

tenemos que

$$\mathbb{E}\Big( \prod_{\omega\in\{0,1\}^{k-1}} \nu(x+\omega\cdot h)\Big| x\in\mathbb{Z}_N, h\in\mathbb{Z}_N^{k-1}\Big) = o(1).$$

Observemos que $o(1)=o_k(1)$. □

Es el momento de establecer el teorema generalizado de von Neumann, el cual explica como la expresión $\mathbb{E}\Big(f(x)f(x+r)\cdots f\big(x+(k-1)r\big)\Big|x,r\in\mathbb{Z}_N\Big)$, la cual realiza el conteo de $k$-progresiones aritméticas, es dominada por la norma de Gowers uniforme de $f$.

**Proposición 6.3** (von Neumann, generalizado)**.** *Sea $\nu$ una medida $k$-pseudoaleatoria. Sean $f_1,\ldots,f_{k-1}\in L^1(\mathbb{Z}_N)$, funciones que son puntualmente limitadas por $\nu+\nu_{\text{const}}$, esto es,*

$$|f_j(x)|\leq\nu(x)+1 \text{ para todo } x\in\mathbb{Z}_N,\ 1\leq j\leq k-1. \tag{6.17}$$

*Sea $c_0,\ldots,c_{k-1}$ una permutación de $k$ elementos consecutivos de*

$$\{-k+1,\ldots,-1,0,1,\ldots,k-1\}$$

*(en la práctica tomaremos $c_j = j$). Entonces*

$$\mathbb{E}\Big(\prod_{j=0}^{k-1} f_j(x + c_j r)\Big| x, r \in \mathbb{Z}_N\Big) = O\Big(\inf_{0\le j\le k-1} ||f_j||_{U^{k-1}}\Big) + o(1).$$

Antes de dar la prueba general, primero analizaremos el caso $k = 3$, con $c_j = j$.

*Demostración de la proposición 6.3, caso particular.* Sean $k = 3$ y $c_j = j$ para $j = 0, \dots, k-1$. Reemplazando $\nu$ por $(\nu+1)/2$ (y dividiendo $f_j$ por 2), usando el lema 6.1, podemos asumir sin pérdida de generalidad que

$$|f_j(x)| \le \nu(x) \text{ para todo } x \in \mathbb{Z}_N, 0 \le j \le k-1. \tag{6.18}$$

Como $\nu$ fue reemplazado por $(\nu + 1)/2$, $\nu$ es estrictamente positiva. Permutanto los $f_j$ y $c_j$, si es necesario, podemos suponer que el ínfimo

$$\inf_{0\le j\le k-1} ||f_j||_{U^{k-1}}$$

es alcanzado cuando $j = 0$. Desplazando $x$ por $c_0 r$, podemos asumir que $c_0 = 0$. Así, nuestra tarea es mostrar que

$$\mathbb{E}\Big(\prod_{j=0}^{k-1} f_j(x + c_j r)\Big| x, r \in \mathbb{Z}_N\Big) = O\big(||f_0||_{U^{k-1}}\big) + o(1). \tag{6.19}$$

Tenemos que mostrar que

$$\mathbb{E}\Big(f_0(x) f_1(x + r) f_2(x + 2r)\Big| x, r \in \mathbb{Z}_N\Big) = O\big(||f_0||_{U^2}\big) + o(1).$$

Sea $J_0 = \mathbb{E}\Big(f_0(x) f_1(x+r) f_2(x+2r)\Big| x, r \in \mathbb{Z}_N\Big)$. Será conveniente reparametrizar la progresión $(x, x + r, x + 2r)$ como $(y_1 + y_2, y_2/2, -y_1)$, a través del cambio de variables

$$\begin{aligned} T : \ \mathbb{Z}_N{}^2 &\longrightarrow \Big\{\bar{x} = (a, b, c) \in \mathbb{Z}_N{}^3;\ b - a = c - b\Big\} \\ (y_1, y_2) &\longmapsto (y_1 + y_2, y_2/2, -y_1). \end{aligned}$$

Así, tenemos que

$$\begin{aligned} J_0 &= \mathbb{E}\Big(f_0(x) f_1(x + r) f_2(x + 2r)\Big| x, r \in \mathbb{Z}_N\Big) \\ &= \mathbb{E}\Big(f_0(x)\mathbb{E}\big(f_1(x + r) f_2(x + 2r)\big| r \in \mathbb{Z}_N\big)\Big| x \in \mathbb{Z}_N\Big) \\ &= \mathbb{E}\Big(f_0(x)\mathbb{E}\Big(f_1\big(x + (r - x)\big) f_2\big(x + 2(r - x)\big)\Big| r \in \mathbb{Z}_N\Big)\Big| x \in \mathbb{Z}_N\Big) \end{aligned} \tag{6.20}$$

$$
\begin{aligned}
&= \mathbb{E}\Big( f_0(x)\mathbb{E}\big(f_1(r)f_2(2r-x)\big|r\in\mathbb{Z}_N\big)\Big|x\in\mathbb{Z}_N\Big)\\
&= \mathbb{E}\Big( f_0(x)\mathbb{E}\Big(f_1(r/2)f_2\big(2(r/2)-x\big)\Big|r\in\mathbb{Z}_N\Big)\Big|x\in\mathbb{Z}_N\Big)\\
&= \mathbb{E}\Big( f_0(x)\mathbb{E}\big(f_1(r/2)f_2(r-x)\big|r\in\mathbb{Z}_N\big)\Big|x\in\mathbb{Z}_N\Big)\\
&= \mathbb{E}\Big( f_0(x)f_1(r/2)f_2(r-x)\Big|x,r\in\mathbb{Z}_N\Big)\\
&= \mathbb{E}\Big( f_1(r/2)\mathbb{E}\big(f_0(x)f_2(r-x)\big|x\in\mathbb{Z}_N\big)\Big|r\in\mathbb{Z}_N\Big)\\
&= \mathbb{E}\Big( f_1(r/2)\mathbb{E}\Big(f_0(x+r)f_2\big(r-(x+r)\big)\Big|x\in\mathbb{Z}_N\Big)\Big|r\in\mathbb{Z}_N\Big)\\
&= \mathbb{E}\Big( f_1(r/2)\mathbb{E}\big(f_0(x+r)f_2(-x)\big|x\in\mathbb{Z}_N\big)\Big|r\in\mathbb{Z}_N\Big)\\
&= \mathbb{E}\Big( f_1(r/2)f_0(x+r)f_2(-x)\Big|x,r\in\mathbb{Z}_N\Big)\\
&= \mathbb{E}\Big( f_0(y_1+y_2)f_1(y_2/2)f_2(-y_1)\Big|y_1,y_2\in\mathbb{Z}_N\Big). \qquad (6.21)
\end{aligned}
$$

El hecho que el primer término, $y_2/2$, no depende de $y_1$, y que el segundo término, $-y_1$, no depende de $y_2$ permitirá usar la desigualdad de Cauchy-Schwarz sin necesidad de algún cambio de variables. Como suponemos que $N \geq 1$ es un primo grande, entonces podemos calcular

$$
\begin{aligned}
J_0 &= \mathbb{E}\Big( f_0(y_1+y_2)f_1(y_2/2)f_2(-y_1)\Big|y_1,y_2\in\mathbb{Z}_N\Big)\\
&= \mathbb{E}\Big( \mathbb{E}\big(f_0(y_1+y_2)f_1(y_2/2)\big|y_2\in\mathbb{Z}_N\big)f_2(-y_1)\Big|y_1\in\mathbb{Z}_N\Big).
\end{aligned}
$$

Como $|f_2| \leq \nu$, tenemos que

$$
\begin{aligned}
|J_0| &\leq \mathbb{E}\Big( \big|\mathbb{E}\big(f_0(y_1+y_2)f_1(y_2/2)\big|y_2\in\mathbb{Z}_N\big)\big|\nu(-y_1)\Big|y_1\in\mathbb{Z}_N\Big)\\
&= \mathbb{E}\Big( \big|\mathbb{E}\big(f_0(y_1+y_2)f_1(y_2/2)\big|y_2\in\mathbb{Z}_N\big)\big|\nu^{1/2}(-y_1)\cdot\nu^{1/2}(-y_1)\Big|y_1\in\mathbb{Z}_N\Big).
\end{aligned}
$$

De la desigualdad de Cauchy-Schwarz,

$$
|J_0| \leq \mathbb{E}\Big( \big|\mathbb{E}\big(f_0(y_1+y_2)f_1(y_2/2)\big|y_2\in\mathbb{Z}_N\big)\big|^2\nu(-y_1)\Big|y_1\in\mathbb{Z}_N\Big)^{1/2}\mathbb{E}\Big(\nu(-y_1)\big|y_1\in\mathbb{Z}_N\Big)^{1/2}.
$$

Puesto que $\mathbb{E}(\nu) = 1 + o(1)$, tenemos que

$$
|J_0| \leq \big(1+o(1)\big)^{1/2}\cdot J_1^{1/2}, \qquad (6.22)
$$

donde

$$\begin{aligned} J_1 &= \mathbb{E}\Big( |\mathbb{E}\big(f_0(y_1+y_2)f_1(y_2/2)\big|y_2\in\mathbb{Z}_N\big)|^2\nu(-y_1)\Big|y_1\in\mathbb{Z}_N\Big) \\ &= \mathbb{E}\Big( \mathbb{E}\big(f_0(y_1+y_2)f_1(y_2/2)\big|y_2\in\mathbb{Z}_N\big)\times \\ &\quad \times\mathbb{E}\big(f_0(y_1+y_2')f_1(y_2'/2)\big|y_2'\in\mathbb{Z}_N\big)\nu(-y_1)\Big|y_1\in\mathbb{Z}_N\Big) \\ &= \mathbb{E}\Big( f_0(y_1+y_2)f_0(y_1+y_2')f_1(y_2/2)f_1(y_2'/2)\nu(-y_1)\Big|y_1,y_2,y_2'\in\mathbb{Z}_N\Big) \end{aligned}$$

Así,

$$J_1 = \mathbb{E}\Big( \mathbb{E}\big(f_0(y_1+y_2)f_0(y_1+y_2')\nu(-y_1)\big|y_1\in\mathbb{Z}_N\big)f_1(y_2/2)f_1(y_2'/2)\Big|y_2,y_2'\in\mathbb{Z}_N\Big),$$

y como $|f_1|\leq\nu$

$$\begin{aligned} |J_1| &\leq \mathbb{E}\Big( \big|\mathbb{E}\big(f_0(y_1+y_2)f_0(y_1+y_2')\nu(-y_1)\big|y_1\in\mathbb{Z}_N\big)\big|\nu(y_2/2)\nu(y_2'/2) \\ &\quad \Big|y_2,y_2'\in\mathbb{Z}_N\Big), \\ &= \mathbb{E}\Big( \big|\mathbb{E}\big(f_0(y_1+y_2)f_0(y_1+y_2')\nu(-y_1)\big|y_1\in\mathbb{Z}_N\big)\big|\nu^{1/2}(y_2/2)\nu^{1/2}(y_2'/2)\times \\ &\quad \times\nu^{1/2}(y_2/2)\nu^{1/2}(y_2'/2)\Big|y_2,y_2'\in\mathbb{Z}_N\Big). \end{aligned}$$

Nuevamente, por la desigualdad de Cauchy-Schwarz

$$\begin{aligned} |J_1| \leq \mathbb{E}\Big( \big|\mathbb{E}\big(f_0(y_1+y_2)f_0(y_1+y_2')\nu(-y_1)\big|y_1\in\mathbb{Z}_N\big)\big|^2\nu(y_2/2)\nu(y_2'/2)\Big|y_2,y_2'\in\mathbb{Z}_N\Big)^{1/2}\times \\ \times\,\mathbb{E}\Big(\nu(y_2/2)\nu(y_2'/2)\Big|y_2,y_2'\in\mathbb{Z}_N\Big)^{1/2}, \end{aligned}$$

esto es,

$$|J_1|\leq J_2^{1/2}\mathbb{E}(\nu),$$

con

$$J_2 = \mathbb{E}\Big( \big|\mathbb{E}\big(f_0(y_1+y_2)f_0(y_1+y_2')\nu(-y_1)\big|y_1\in\mathbb{Z}_N\big)\big|^2\nu(y_2/2)\nu(y_2'/2)\Big|y_2,y_2'\in\mathbb{Z}_N\Big),$$

de donde

$$|J_1|^{1/2} \quad \leq \quad J_2^{1/4}\big(1+o(1)\big)^{1/2}.$$

Luego, de (6.22)

$$|J_0| \leq J_2^{1/4}\big(1+o(1)\big). \tag{6.23}$$

Hagamos una estimación de $J_2$,

$$\begin{aligned} J_2 &= \mathbb{E}\bigg(\mathbb{E}\big(f_0(y_1+y_2)f_0(y_1+y_2')\nu(-y_1)\big|y_1\in\mathbb{Z}_N\big)\mathbb{E}\big(f_0(y_1'+y_2)f_0(y_1'+y_2')\nu(-y_1')\big|y_1'\in\mathbb{Z}_N\big) \\ &\qquad \bigg| y_2, y_2' \in \mathbb{Z}_N\bigg) \\ &= \mathbb{E}\bigg(f_0(y_1+y_2)f_0(y_1+y_2')f_0(y_1'+y_2)f_0(y_1'+y_2')\nu(-y_1)\nu(-y_1')\nu(y_2/2)\nu(y_2'/2) \\ &\qquad \bigg| y_1, y_1', y_2, y_2' \in \mathbb{Z}_N\bigg). \end{aligned}$$

Si no fuese por los términos $\nu$, $J_2$ sería $||f_0||_{U^2}^4$. Si reemplazamos el cubo

$$\Big\{\big(y_1+y_2, y_1'+y_2, y_1+y_2', y_1'+y_2'\big);\ y_1, y_1', y_2, y_2' \in \mathbb{Z}_N\Big\},$$

por $\Big\{(x, x+h_1, x+h_2, x+h_1+h_2);\ x, h_1, h_2 \in \mathbb{Z}_N\Big\}$, (a través de un cambio de variables como el hecho en (6.21)), resulta que

$$J_2 = \mathbb{E}\bigg(f_0(x)f_0(x+h_1)f_0(x+h_2)f_0(x+h_1+h_2)W(x,h_1,h_2)\bigg|x,h_1,h_2\in\mathbb{Z}_N\bigg),$$

donde

$$W(x,h_1,h_2) = \mathbb{E}\bigg(\nu(-y)\nu(-y-h_1)\nu\big((x-y)/2\big)\nu\big((x-y+h_2)/2\big)\bigg|y\in\mathbb{Z}_N\bigg). \tag{6.24}$$

Para comparar $J_2$ con $||f_0||_{U^2}^4$ requerimos comparar $W(x,h_1,h_2)$ con 1.

$$\begin{aligned} J_2 - ||f_0||_{U^2}^4 &= \mathbb{E}\bigg(f_0(x)f_0(x+h_1)f_0(x+h_2)f_0(x+h_1+h_2)W(x,h_1,h_2)\bigg|x,h_1,h_2\in\mathbb{Z}_N\bigg) \\ &\qquad - \mathbb{E}\bigg(f_0(x)f_0(x+h_1)f_0(x+h_2)f_0(x+h_1+h_2)\bigg|x,h_1,h_2\in\mathbb{Z}_N\bigg) \\ &= \mathbb{E}\bigg(f_0(x)f_0(x+h_1)f_0(x+h_2)f_0(x+h_1+h_2)\big(W(x,h_1,h_2)-1\big)\bigg|x,h_1,h_2\in\mathbb{Z}_N\bigg), \end{aligned}$$

así,

$$\begin{aligned} \Big|J_2 - ||f_0||_{U^2}^4\Big| &\leq \mathbb{E}\bigg(\nu(x)\nu(x+h_1)\nu(x+h_2)\nu(x+h_1+h_2)\big|W(x,h_1,h_2)-1\big|\bigg|x,h_1,h_2\in\mathbb{Z}_N\bigg) \\ &= \mathbb{E}\bigg(\nu^{1/2}(x)\nu^{1/2}(x+h_1)\nu^{1/2}(x+h_2)\nu^{1/2}(x+h_1+h_2)\big|W(x,h_1,h_2)-1\big|\times \\ &\qquad \times \nu^{1/2}(x)\nu^{1/2}(x+h_1)\nu^{1/2}(x+h_2)\nu^{1/2}(x+h_1+h_2)\bigg|x,h_1,h_2\in\mathbb{Z}_N\bigg). \end{aligned}$$

Por la desigualdad de Cauchy-Schwarz

$$
\begin{aligned}
&\big|J_2 - ||f_0||_{U^2}^4\big| \\
&\quad \le \mathbb{E}\bigg(\nu(x)\nu(x+h_1)\nu(x+h_2)\nu(x+h_1+h_2)\bigg|x,h_1,h_2\in\mathbb{Z}_N\bigg)^{1/2}\times \\
&\qquad \times \mathbb{E}\bigg(\nu(x)\nu(x+h_1)\nu(x+h_2)\nu(x+h_1+h_2)\big(W(x,h_1,h_2)- \\
&\qquad -1\big)^2\bigg|x,h_1,h_2\in\mathbb{Z}_N\bigg)^{1/2} \\
&\quad \le \mathbb{E}\bigg(\nu(x)\nu(x+h_1)\nu(x+h_2)\nu(x+h_1+h_2)\bigg|x,h_1,h_2\in\mathbb{Z}_N\bigg)^{1/2}\times \\
&\qquad \times \bigg[\mathbb{E}\bigg(\nu(x)\nu(x+h_1)\nu(x+h_2)\nu(x+h_1+h_2)W^2(x,h_1,h_2)\bigg|x,h_1,h_2\in\mathbb{Z}_N\bigg)- \\
&\qquad -2\mathbb{E}\bigg(\nu(x)\nu(x+h_1)\nu(x+h_2)\nu(x+h_1+h_2)W(x,h_1,h_2)\bigg|x,h_1,h_2\in\mathbb{Z}_N\bigg)+ \\
&\qquad +\mathbb{E}\bigg(\nu(x)\nu(x+h_1)\nu(x+h_2)\nu(x+h_1+h_2)\bigg|x,h_1,h_2\in\mathbb{Z}_N\bigg)\bigg]^{1/2}.
\end{aligned}
$$

Así, para mostrar que $\big|J_2 - ||f_0||_{U^2}^4\big|$ es convenientemente pequeño es suficiente mostrar que

$$\mathbb{E}\bigg(\nu(x)\nu(x+h_1)\nu(x+h_2)\nu(x+h_1+h_2)W^q(x,h_1,h_2)\bigg|x,h_1,h_2\bigg) = 1+o(1)$$

para $q=0,1,2$. Pero esto último se sigue de la condición de formas lineales, por ejemplo, el caso $q=2$ es justamente la expresión (6.7). □

Retornemos ahora a la prueba de (6.19) en el caso general. Supongamos que $0\le d\le k-1$, y que tenemos los vectores $y=(y_1,\dots,y_{k-1})\in\mathbb{Z}_N^{k-1}$ y $y'=(y'_{k-d},\dots,y'_{k-1})\in\mathbb{Z}_N^d$. Para cualquier $S\subseteq\{k-d,\dots,k-1\}$ definamos el vector $y^{(S)}=(y_1^{(S)},\dots,y_{k-1}^{(S)})\in\mathbb{Z}_N^{k-1}$ como

$$y_i^{(S)} = \begin{cases} y_i, & \text{si } i\notin S \\ y'_i, & \text{si } i\in S. \end{cases}$$

El conjunto $S$ indica las componentes de $y^{(S)}$ que provienen de $y'$.

**Lema 6.3** (Lema de Cauchy-Schwarz)**.** *Sea $\nu:\mathbb{Z}_N\to\mathbb{R}^+$ una medida. Sean $\phi_0,\phi_1,\dots,\phi_{k-1}:\mathbb{Z}_N^{k-1}\to\mathbb{Z}_N$ funciones de $k-1$ variables $y_1,y_2,\dots,y_{k-1}$ tales que $\phi_i$ no depende de $y_i$ para $1\le i\le k-1$. Supongamos que $f_0,f_1,\dots,f_{k-1}\in L^1(\mathbb{Z}_N)$ sean funciones satisfaciendo $|f_i(x)|\le\nu(x)$ para todo $x\in\mathbb{Z}_N$ y para todo $i$, $0\le i\le k-1$. Para cada $0\le d\le k-1$,*

*definimos las cantidades*

$$J_d = \mathbb{E}\Bigg( \prod_{S\subseteq\{k-d,\dots,k-1\}} \Big( \prod_{i=0}^{k-d-1} f_i(\phi_i(y^{(S)})) \Big) \Big( \prod_{i=k-d}^{k-1} \nu^{1/2}(\phi_i(y^{(S)})) \Big) \Big| y \in \mathbb{Z}_N^{k-1},\, y' \in \mathbb{Z}_N^d \Bigg) \quad (6.25)$$

*y*

$$P_d = \mathbb{E}\Bigg( \prod_{S\subseteq\{k-d,\dots,k-1\}} \nu(\phi_{k-d-1}(y^{(S)})) \Big| y \in \mathbb{Z}_N^{k-1},\, y' \in \mathbb{Z}_N^d \Bigg). \quad (6.26)$$

*Entonces para cada* $0 \le d \le k-2$*, tenemos la desigualdad*

$$|J_d|^2 \le P_d J_{d+1}. \quad (6.27)$$

**Observación 6.3.** *La aparición de* $\nu^{1/2}$ *en la ecuación (6.25) puede parecer extraña. Sin embargo, como* $\phi_i$ *no depende de la* $i$*-ésima variable, cada factor* $\nu^{1/2}$ *en la ecuación (6.25) ocurre dos veces. Si tomamos* $k = 3$ *y*

$$\phi_0(y_1, y_2) = y_1 + y_2, \;\; \phi_1(y_1, y_2) = y_2/2, \;\; \phi_2(y_1, y_2) = -y_1, \quad (6.28)$$

*obtenemos las cantidades* $J_0$*,* $J_1$ *y* $J_2$ *que coinciden con los de la precedente discusión.*

*Demostración.* Sea $0 \le d \le k-2$ (aquí $1 \le k-d-1$). Consideremos

$$\begin{aligned} J_d \;=\; & \mathbb{E}\Bigg( \prod_{S\subseteq\{k-d,\dots,k-1\}} \Bigg[ \prod_{i=0}^{k-d-1} f_i(\phi_i(y^{(S)})) \prod_{i=k-d}^{k-1} \nu^{1/2}(\phi_i(y^{(S)})) \Bigg] \\ & \Big| y \in \mathbb{Z}_N^{k-1}, y' = (y'_{k-d}, \dots, y'_{k-1}) \in \mathbb{Z}_N^d \Bigg). \end{aligned}$$

El producto del cual se toma el promedio es de la forma

$$\begin{aligned} & \prod_{S\subseteq\{k-d,\dots,k-1\}} \Bigg[ \prod_{i=0}^{k-d-1} f_i\big(\phi_i(y^{(S)})\big) \prod_{i=k-d}^{k-1} \nu^{1/2}\big(\phi_i(y^{(S)})\big) \Bigg] \\ & \quad = \prod_{S\subseteq\{k-d,\dots,k-1\}} f_{k-d-1}\big(\phi_{k-d-1}(y^{(S)})\big) \prod_{S\subseteq\{k-d,\dots,k-1\}} \Bigg[ \prod_{i=0}^{k-d-2} f_i\big(\phi_i(y^{(S)})\big) \times \\ & \qquad \times \prod_{i=k-d}^{k-1} \nu^{1/2}\big(\phi_i(y^{(S)})\big) \Bigg] \\ & \quad = \prod_{S\subseteq\{k-d,\dots,k-1\}} \Big[ f_{k-d-1}\big(\phi_{k-d-1}(y^{(S)})\big) \nu^{-1/2}\big(\phi_{k-d-1}(y^{(S)})\big) \Big] \times \\ & \qquad \times \prod_{S\subseteq\{k-d,\dots,k-1\}} \Bigg[ \prod_{i=0}^{k-d-2} f_i\big(\phi_i(y^{(S)})\big) \prod_{i=k-d-1}^{k-1} \nu^{1/2}\big(\phi_i(y^{(S)})\big) \Bigg]. \end{aligned}$$

Aquí hemos dividido y multiplicado por factores de la forma $\nu^{1/2}(\phi_{k-d-1}(y^{(S)}))$, para cada $S \subseteq \{k-d,\dots,k-1\}$. Como $\phi_{k-d-1}$ no depende de $y_{k-d-1}$, tomando promedios sobre las demás variables $y_1,\dots,y_{k-d-2},y_{k-d},\dots,y_{k-1},y'_{k-d},\dots,y'_{k-1}$, tenemos

$$\begin{aligned} J_d &= \mathbb{E}\Bigg( \prod_{S\subseteq\{k-d,\dots,k-1\}} f_{k-d-1}\big(\phi_{k-d-1}(y^{(S)})\big)\nu^{-1/2}\big(\phi_{k-d-1}(y^{(S)})\big) \times \\ &\quad \times \mathbb{E}\Big( \prod_{S\subseteq\{k-d,\dots,k-1\}} \Big[ \prod_{i=0}^{k-d-2} f_i\big(\phi_i(y^{(S)})\big) \prod_{i=k-d-1}^{k-1} \nu^{1/2}\big(\phi_i(y^{(S)})\big) \Big] \Big| y_{k-d-1}\in\mathbb{Z}_N \Big) \\ &\quad \Big| y_1,\dots,y_{k-d-2},y_{k-d},\dots,y_{k-1},y'_{k-d},\dots,y'_{k-1}\in\mathbb{Z}_N \Bigg). \end{aligned}$$

Así

$$J_d = \mathbb{E}\Big( G(y,y')H(y,y') \Big| y_1,\dots,y_{k-d-2},y_{k-d},\dots,y_{k-1},y'_{k-d},\dots,y'_{k-1}\in\mathbb{Z}_N \Big),$$

donde

$$G(y,y') = \prod_{S\subseteq\{k-d,\dots,k-1\}} f_{k-d-1}\big(\phi_{k-d-1}(y^{(S)})\big)\nu^{-1/2}\big(\phi_{k-d-1}(y^{(S)})\big)$$

y

$$H(y,y') = \mathbb{E}\Bigg( \prod_{S\subseteq\{k-d,\dots,k-1\}} \Big[ \prod_{i=0}^{k-d-2} f_i\big(\phi_i(y^{(S)})\big) \prod_{i=k-d-1}^{k-1} \nu^{1/2}\big(\phi_i(y^{(S)})\big) \Big] \Big| y_{k-d-1}\in\mathbb{Z}_N \Bigg).$$

Por la desigualdad de Cauchy-Schwarz

$$\begin{aligned} |J_d| &\le \mathbb{E}\Big( |G(y,y')|^2 \Big| y_1,\dots,y_{k-d-2},y_{k-d},\dots,y_{k-1},y'_{k-d},\dots,y'_{k-1}\in\mathbb{Z}_N \Big)^{1/2} \times \\ &\quad \mathbb{E}\Big( |H(y,y')|^2 \Big| y_1,\dots,y_{k-d-2},y_{k-d},\dots,y_{k-1},y'_{k-d},\dots,y'_{k-1}\in\mathbb{Z}_N \Big)^{1/2}, \end{aligned}$$

y así

$$\begin{aligned} |J_d|^2 \le \mathbb{E}\Big( |G(y,y')|^2 \Big| y_1,\dots,y_{k-d-2},y_{k-d},\dots,y_{k-1},y'_{k-d},\dots,y'_{k-1}\in\mathbb{Z}_N \Big) \times \\ \mathbb{E}\Big( |H(y,y')|^2 \Big| y_1,\dots,y_{k-d-2},y_{k-d},\dots,y_{k-1},y'_{k-d},\dots,y'_{k-1}\in\mathbb{Z}_N \Big). \end{aligned}$$

Como $|f_{k-d-1}(x)| \le \nu(x)$ para todo $x\in\mathbb{Z}_N$, vemos de (6.26) que

$$\mathbb{E}\Big( |G(y,y')|^2 \Big| y_1,\dots,y_{k-d-2},y_{k-d},\dots,y_{k-1},y'_{k-d},\dots,y'_{k-1}\in\mathbb{Z}_N \Big) \le P_d.$$

En efecto,

$$\mathbb{E}\Big(|G(y,y')|^2\Big|y_1,\dots,y_{k-d-2},y_{k-d},\dots,y_{k-1},y'_{k-d},\dots,y'_{k-1}\in\mathbb{Z}_N\Big)$$
$$\leq \mathbb{E}\Bigg(\prod_{S\subseteq\{k-d,\dots,k-1\}}\nu\big(\phi_{k-d-1}(y^{(S)})\big)\Bigg|y_1,\dots,y_{k-d-2},y_{k-d},\dots,y_{k-1},y'_{k-d},\dots,y'_{k-1}\in\mathbb{Z}_N\Bigg)$$
$$= \mathbb{E}\Bigg(\prod_{S\subseteq\{k-d,\dots,k-1\}}\nu\big(\phi_{k-d-1}(y^{(S)})\big)\Bigg|y\in\mathbb{Z}_N^{k-1},y'\in\mathbb{Z}_N^d\Bigg)=P_d$$

La última igualdad se da ya que en el promedio de (6.26) es irrelevante la variable $y_{k-d-1}$, puesto que $\phi_{k-d-1}$ no depende de esta variable. Además,

$$\mathbb{E}\Big(|H(y,y')|^2\Big|y_1,\dots,y_{k-d-2},y_{k-d},\dots,y_{k-1},y'_{k-d},\dots,y'_{k-1}\in\mathbb{Z}_N\Big)$$
$$= \mathbb{E}\Bigg(\mathbb{E}\Big(\prod_{S\subseteq\{k-d,\dots,k-1\}}\Big[\prod_{i=0}^{k-d-2}f_i\big(\phi_i(y^{(S)})\big)\prod_{i=k-d-1}^{k-1}\nu^{1/2}\big(\phi_i(y^{(S)})\big)\Big|y_{k-d-1}\in\mathbb{Z}_N\Big)\Big]\times \tag{6.29}$$

$$\times\,\mathbb{E}\Bigg(\mathbb{E}\Big(\prod_{S\subseteq\{k-d,\dots,k-1\}}\Big[\prod_{i=0}^{k-d-2}f_i\big(\phi_i(y^{(S)})\big)\prod_{i=k-d-1}^{k-1}\nu^{1/2}\big(\phi_i(y^{(S)})\big)\Big|y'_{k-d-1}\in\mathbb{Z}_N\Big)\Big] \tag{6.30}$$
$$\Bigg|y_1,\dots,y_{k-d-2},y_{k-d},\dots,y_{k-1},y'_{k-d},\dots,y'_{k-1}\in\mathbb{Z}_N\Bigg).$$

En (6.29) $y^{(S)}_{k-d-1}=y_{k-d-1}$, mientras que en (6.30), $y^{(S)}_{k-d-1}=y'_{k-d-1}$. Luego

$$\mathbb{E}\Big(|H(y,y')|^2\Big|y_1,\dots,y_{k-d-2},y_{k-d},\dots,y_{k-1},y'_{k-d},\dots,y'_{k-1}\in\mathbb{Z}_N\Big)$$
$$= \mathbb{E}\Bigg(\prod_{S\subseteq\{k-d,\dots,k-1\}}\Big[\prod_{i=0}^{k-d-2}f_i\big(\phi_i(y^{(S)})\big)\prod_{i=k-d-1}^{k-1}\nu^{1/2}\big(\phi_i(y^{(S)})\big)\Big]\times$$
$$\times\prod_{S\subseteq\{k-d-1,k-d,\dots,k-1\};\,k-d-1\in S}\Big[\prod_{i=0}^{k-d-2}f_i\big(\phi_i(y^{(S)})\big)\prod_{i=k-d-1}^{k-1}\nu^{1/2}\big(\phi_i(y^{(S)})\big)\Big]$$
$$\Bigg|y_1,\dots,y_{k-d-2},y_{k-d-1},y_{k-d},\dots,y_{k-1},y'_{k-d-1},y'_{k-d},\dots,y'_{k-1}\in\mathbb{Z}_N\Bigg)$$
$$= \mathbb{E}\Bigg(\prod_{S\subseteq\{k-d-1,\dots,k-1\}}\Big[\prod_{i=0}^{k-d-2}f_i\big(\phi_i(y^{(S)})\big)\prod_{i=k-d-1}^{k-1}\nu^{1/2}\big(\phi_i(y^{(S)})\big)\Big]$$
$$\Bigg|y\in\mathbb{Z}_N^{k-1},y'\in\mathbb{Z}_N^{d+1}\Bigg)=J_{d+1}.$$

Así

$$J_{d+1}=\mathbb{E}\Big(|H(y,y')|^2\Big|y_1,\dots,y_{k-d-2},y_{k-d},\dots,y_{k-1},y'_{k-d},\dots,y'_{k-1}\in\mathbb{Z}_N\Big),$$

y por lo tanto

$$|J_d|^2 \leq P_d J_{d+1}, \tag{6.31}$$

con lo cual el lema queda demostrado. □

Aplicando el lema anterior $k-1$ veces obtenemos

$$|J_0|^{2^{k-1}} \leq J_{k-1} \prod_{d=0}^{k-2} P_d^{2^{k-2-d}}. \tag{6.32}$$

En efecto,

$$\begin{aligned}
|J_0|^2 &\leq P_0 J_1 & |J_0|^4 &\leq P_0^2 J_1^2 \leq P_0^2 P_1 J_2 \\
|J_1|^2 &\leq P_1 J_2 & |J_0|^8 &\leq P_0^{2^2} P_1^2 J_2^2 \leq P_0^{2^2} P_1^2 P_2 J_3 \\
|J_2|^2 &\leq P_2 J_3 & |J_0|^{16} &\leq P_0^{2^3} P_1^{2^2} P_2^2 J_2^2 \leq P_0^{2^3} P_1^{2^2} P_2^2 P_3 J_4 \\
&\vdots & &\vdots \\
|J_{k-2}|^2 &\leq P_{k-2} J_{k-1} & |J_0|^{2^{k-1}} &\leq J_{k-1} \prod_{d=0}^{k-2} P_d^{2^{k-2-d}}.
\end{aligned}$$

De (6.25) tenemos que

$$J_0 = \mathbb{E}\left( \prod_{i=0}^{k-1} f_i\big(\phi_i(y)\big) \middle| y \in \mathbb{Z}_N^{k-1} \right) \tag{6.33}$$

*Demostración de la proposición 6.3.* Para $y = (y_1, \ldots, y_{k-1}) \in \mathbb{Z}_N^{k-1}$, sean

$$\phi_i(y) = \sum_{j=1}^{k-1} \left(1 - \frac{c_i}{c_j}\right) y_j$$

para $i = 0, \ldots, k-1$. Recordemos que $c_j = 0$ sólo cuando $j = 0$. Entonces $\phi_0(y) = y_1 + \cdots + y_{k-1}$ y $\phi_i$ no depende de $y_i$ para $1 \leq i \leq k-1$. Notemos que si $k = 3$ y $c_j = j$ obtenemos los $\phi_i$ como en (6.28).

$$\phi_i(y) = \sum_{j=1}^{k-1} y_j - c_i \left( \sum_{j=1}^{k-1} \frac{y_j}{c_j} \right).$$

Sea $\Phi : \mathbb{Z}_N^{k-1} \to \mathbb{Z}_2^2$, definida por

$$\Phi(y) = \left( y_1 + \cdots + y_{k-1}, \frac{y_1}{c_1} + \cdots + \frac{y_{k-1}}{c_{k-1}} \right).$$

La función $\Phi$ es una aplicación recubridora. En efecto, dados $x, r \in \mathbb{Z}_N$ y $y \in \mathbb{Z}_N^{k-1}$ tenemos que $\Phi(y) = (x, r)$ si, y sólo si

$$y_1 + \cdots + y_{k-1} = x \text{ e } \frac{y_1}{c_1} + \cdots + \frac{y_{k-1}}{c_{k-1}} = r,$$

esto es, si y sólo si

$$y_2 = \left(\frac{1}{c_1} - \frac{1}{c_2}\right)^{-1}\left(\frac{x}{c_1} - r - \sum_{j=3}^{k-1}(\frac{1}{c_1} - \frac{1}{c_j})y_j\right) \text{ e } y_1 = x - y_2 - \cdots - y_{k-1}.$$

Así, para cada $z = (x, r) \in \mathbb{Z}_N^2$, $\Phi^{-1}(z)$ tiene $N^{k-3}$ elementos.

Consideremos las aplicaciones $T_j : \mathbb{Z}_N^2 \to \mathbb{Z}_N$ definida por

$$T_j(a, b) = a - c_j b,$$

para $1 \leq j \leq k-1$. Tenemos que

$$\begin{aligned} T_j \circ \Phi(y) &= T_j\Big(y_1 + \cdots + y_{k-1}, \frac{y_1}{c_1} + \cdots + \frac{y_{k-1}}{c_{k-1}}\Big) \\ &= y_1 + \cdots + y_{k-1} - c_j\Big(\frac{y_1}{c_1} + \cdots + \frac{y_{k-1}}{c_{k-1}}\Big) \\ &= \sum_{r=1}^{k-1}\Big(1 - \frac{c_j}{c_r}\Big)y_r \\ &= \phi_j(y). \end{aligned}$$

Así,

$$\begin{aligned} \mathbb{E}\left(\prod_{j=0}^{k-1} f_j(x + c_j r)\middle| x, r \in \mathbb{Z}_N\right) &= \mathbb{E}\left(\prod_{j=0}^{k-1} f_j(x - c_j r)\middle| x, r \in \mathbb{Z}_N\right) \\ &= \mathbb{E}\left(\prod_{j=0}^{k-1} f_j\big(T_j(x, r)\big)\middle| x, r \in \mathbb{Z}_N\right) \\ &= \mathbb{E}\left(\prod_{j=0}^{k-1} f_j\Big(T_j\big(\Phi(y)\big)\Big)\middle| y \in \mathbb{Z}_N^{k-1}\right) \\ &= \mathbb{E}\left(\prod_{j=0}^{k-1} f_j\big(\phi_j(y)\big)\middle| y \in \mathbb{Z}_N^{k-1}\right) \\ &= J_0, \end{aligned} \tag{6.34}$$

gracias a (6.33). Por otro lado $P_d = 1 + o(1)$ para todo $0 \leq d \leq k-2$ (así, $1 \leq k-d-1$). En efecto,

$$P_d = \mathbb{E}\left(\prod_{S \subseteq \{k-d, \ldots, k-1\}} \nu\big(\phi_{k-d-1}(y^{(S)})\big)\middle| y \in \mathbb{Z}_N^{k-1}, y' \in \mathbb{Z}_N^d\right),$$

$\nu$ satisface la $(k2^{k-1}, 3k-4, k)$-condición de formas lineales, $2^d \leq k2^{k-1}$ y $d + k - 1 \leq 3k - 4$ (pues $d \leq k-2$), y los $c_j$ son tales que $|c_j| \leq k-1$. En realidad, es suficiente que $\nu$ satisfaga

la $(2^d, k-1+d, k)$-condición de formas lineales. Aplicando (6.32) obtenemos

$$J_0^{2^{k-1}} \leq \big(1+o(1)\big) J_{k-1}. \tag{6.35}$$

Estimemos ahora

$$\begin{aligned} J_{k-1} &= \mathbb{E}\bigg( \prod_{S\subseteq\{1,\dots,k-1\}} \Big[ \prod_{i=0}^{0} f_i\big(\phi_i(y^{(S)})\big) \prod_{i=1}^{k-1} \nu^{1/2}\big(\phi_i(y^{(S)})\big) \Big] \bigg| y \in \mathbb{Z}_N^{k-1}, y' \in \mathbb{Z}_N^{k-1} \bigg) \\ &= \mathbb{E}\bigg( \prod_{S\subseteq\{1,\dots,k-1\}} \Big[ f_0\big(\phi_0(y^{(S)})\big) \prod_{i=1}^{k-1} \nu^{1/2}\big(\phi_i(y^{(S)})\big) \Big] \bigg| y \in \mathbb{Z}_N^{k-1}, y' \in \mathbb{Z}_N^{k-1} \bigg). \end{aligned}$$

Ahora consideremos $y \in \mathbb{Z}_N$ fijo. Cuando $S$ varía sobre todos los subconjuntos de $\{1,\dots,k-1\}$, $\phi_0(y^{(S)})$ varía sobre el cubo $(k-1)$-dimensional $\Big\{x+\omega\cdot h : \omega \in \{0,1\}^{k-1}\Big\}$, donde

$$x = y_1 + \cdots + y_{k-1} \ \text{ y } \ h_i = y_i' - y_i,\ i = 1,2,\dots,k-1.$$

Y $y^{(S)}$ varía sobre $\Big\{y+\omega h : \omega \in \{0,1\}^{k-1}\Big\}$, donde $y = (y_1,\dots,y_{k-1})$ y $(\omega h)_j = \omega_j h_j$, $1 \leq j \leq k-1$.

$$\begin{aligned} J_{k-1} &= \mathbb{E}\bigg( \prod_{S\subseteq\{1,\dots,k-1\}} \Big[ f_0\big(\phi_0(y^{(S)})\big) \prod_{i=1}^{k-1} \nu^{1/2}\big(\phi_i(y^{(S)})\big) \Big] \bigg| y \in \mathbb{Z}_N^{k-1}, y' \in \mathbb{Z}_N^{k-1} \bigg) \\ &= \mathbb{E}\bigg( \mathbb{E}\Big( \prod_{S\subseteq\{1,\dots,k-1\}} \Big[ f_0\big(\phi_0(y^{(S)})\big) \prod_{i=1}^{k-1} \nu^{1/2}\big(\phi_i(y^{(S)})\big) \Big] \Big| y' \in \mathbb{Z}_N^{k-1} \Big) \bigg| y \in \mathbb{Z}_N^{k-1} \bigg) \\ &= \mathbb{E}\bigg( \mathbb{E}\Big( \prod_{\omega\in\{0,1\}^{k-1}} \Big[ f_0\big(\phi_0(y)+\omega\cdot h\big) \prod_{i=1}^{k-1} \nu^{1/2}\big(\phi_i(y+\omega h)\big) \Big] \Big| h \in \mathbb{Z}_N^{k-1} \Big) \bigg| y \in \mathbb{Z}_N^{k-1} \bigg) \\ &= \mathbb{E}\bigg( \prod_{\omega\in\{0,1\}^{k-1}} f_0\big(\phi_0(y)+\omega\cdot h\big) \prod_{\omega\in\{0,1\}^{k-1}} \prod_{i=1}^{k-1} \nu^{1/2}\big(\phi_i(y+\omega h)\big) \bigg| y, h \in \mathbb{Z}_N^{k-1} \bigg) \\ &= \mathbb{E}\bigg( \mathbb{E}\Big( \prod_{\omega\in\{0,1\}^{k-1}} f_0\big(\phi_0(y)+\omega\cdot h\big) \prod_{\omega\in\{0,1\}^{k-1}} \prod_{i=1}^{k-1} \nu^{1/2}\big(\phi_i(y+\omega h)\big) \Big| y_{k-1} \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \Big) \\ &\qquad \bigg| y_1,\dots,y_{k-2} \in \mathbb{Z}_N \bigg) \\ &= \mathbb{E}\bigg( \mathbb{E}\Big( \prod_{\omega\in\{0,1\}^{k-1}} f_0(x+\omega\cdot h) \prod_{\omega\in\{0,1\}^{k-1}} \prod_{i=1}^{k-1} \nu^{1/2}\big(\phi_i((y_1,\dots,y_{k-2},x-y_1-\cdots \\ &\qquad -y_{k-2})+\omega h)\big) \Big| x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \Big) \bigg| y_1,\dots,y_{k-2} \in \mathbb{Z}_N \bigg) \\ &= \mathbb{E}\bigg( \prod_{\omega\in\{0,1\}^{k-1}} f_0(x+\omega\cdot h) \prod_{\omega\in\{0,1\}^{k-1}} \prod_{i=1}^{k-1} \nu^{1/2}\big(\phi_i((y_1,\dots,y_{k-2},x-y_1 \end{aligned}$$

$$- y_2 - \cdots - y_{k-2}) + \omega h)\Big)\Big| x, y_1, \ldots, y_{k-2} \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1}\Big)$$

$$= \mathbb{E}\Big( \prod_{\omega \in \{0,1\}^{k-1}} f_0(x + \omega \cdot h) \mathbb{E}\Big( \prod_{\omega \in \{0,1\}^{k-1}} \prod_{i=1}^{k-1} \nu^{1/2}\big(\phi_i((y_1, \ldots, y_{k-2}, x$$

$$- y_1 - y_2 - \cdots - y_{k-2}) + \omega h)\big)\Big| y_1, \ldots, y_{k-2} \in \mathbb{Z}_N\Big)\Big| x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1}\Big)$$

$$J_{k-1} = \mathbb{E}\Big( \prod_{\omega \in \{0,1\}^{k-1}} f_0(x + \omega \cdot h) W(x,h)\Big| x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1}\Big), \tag{6.36}$$

donde

$$\begin{aligned}
&W(x,h)\\
&= \mathbb{E}\Big( \prod_{\omega \in \{0,1\}^{k-1}} \prod_{i=1}^{k-1} \nu^{1/2}\Big(\phi_i\big((y_1, \ldots, y_{k-2}, x - y_1 - \cdots - y_{k-2}) + \omega h\big)\Big)\Big| y_1, \ldots, y_{k-2} \in \mathbb{Z}_N\Big)\\
&= \mathbb{E}\Big( \prod_{i=1}^{k-1} \prod_{\omega \in \{0,1\}^{k-1}} \nu^{1/2}\Big(\phi_i\big((y_1, \ldots, y_{k-2}, x - y_1 - \cdots - y_{k-2}) + \omega h\big)\Big)\Big| y_1, \ldots, y_{k-2} \in \mathbb{Z}_N\Big)\\
&= \mathbb{E}\Big( \prod_{i=1}^{k-1} \Big[ \prod_{\omega \in \{0,1\}^{k-1}, \omega_i = 0} \nu^{1/2}\big(\phi_i(y + \omega h)\big) \prod_{\omega \in \{0,1\}^{k-1}, \omega_i = 1} \nu^{1/2}\big(\phi_i(y + \omega h)\big)\Big]\Big| y_1, \ldots, y_{k-2} \in \mathbb{Z}_N\Big)\\
&= \mathbb{E}\Big( \prod_{i=1}^{k-1} \prod_{\omega \in \{0,1\}^{k-1}, \omega_i = 0} \nu^{1/2}\big(\phi_i(y + \omega h)\big)\Big| y_1, \ldots, y_{k-2} \in \mathbb{Z}_N\Big),
\end{aligned}$$

donde la componente $j$-ésima de $y$ es $y_j$ para $1 \leq j \leq k-2$, e $y_{k-1} = x - y_1 - \cdots - y_{k-2}$.

Por definición de norma $U^{k-1}$, tenemos que

$$\mathbb{E}\Big( \prod_{\omega \in \{0,1\}^{k-1}} f_0(x + \omega \cdot h)\Big| x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1}\Big) = ||f_0||_{U^{k-1}}^{2^{k-1}}.$$

Para mostrar (6.19) es suficiente, por (6.34), (6.35) y (6.36), probar que

$$\mathbb{E}\Big( \big(W(x,h) - 1\big) \prod_{\omega \in \{0,1\}^{k-1}} f_0(x + \omega \cdot h)\Big| x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1}\Big) = o(1).$$

Como $|f_j| \leq \nu$ es suficiente mostrar que

$$\mathbb{E}\Big( \big|W(x,h) - 1\big| \prod_{\omega \in \{0,1\}^{k-1}} \nu(x + \omega \cdot h)\Big| x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1}\Big) = o(1).$$

Tenemos que el lado izquierdo de la anterior desigualdad se puede expresar como

$$\mathbb{E}\Big( \big|W(x,h) - 1\big| \prod_{\omega \in \{0,1\}^{k-1}} \nu^{1/2}(x + \omega \cdot h) \prod_{\omega \in \{0,1\}^{k-1}} \nu^{1/2}(x + \omega \cdot h)\Big| x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1}\Big).$$

Y por la desigualdad de Cauchy-Schwarz esta expresión es menor o igual que

$$\mathbb{E}\Big(\big|W(x,h)-1\big|^2\prod_{\omega\in\{0,1\}^{k-1}}\nu(x+\omega\cdot h)\Big|x\in\mathbb{Z}_N,h\in\mathbb{Z}_N^{k-1}\Big)^{1/2}\mathbb{E}\Big(\prod_{\omega\in\{0,1\}^{k-1}}\nu(x+\omega\cdot h)\Big|x\in\mathbb{Z}_N,h\in\mathbb{Z}_N^{k-1}\Big)^{1/2}.$$

Así, es suficiente mostrar el siguiente lema.

**Lema 6.4.** *Para $n=0,2$ tenemos*

$$\mathbb{E}\Big(|W(x,h)-1|^n\prod_{\omega\in\{0,1\}^{k-1}}\nu(x+\omega\cdot h)\Big|x\in\mathbb{Z}_N,h\in\mathbb{Z}_N^{k-1}\Big)=0^n+o(1).$$

*Aquí convenimos que $0^0=1$.*

*Demostración.* Desarrollando $|W(x,h)-1|^2=W^2(x,h)-2W(x,h)+1$, vemos que es suficiente mostrar que

$$\mathbb{E}\Big(W(x,h)^q\prod_{\omega\in\{0,1\}^{k-1}}\nu(x+\omega\cdot h)\Big|x\in\mathbb{Z}_n,h\in\mathbb{Z}_N^{k-1}\Big)=1+o(1)$$

para $q=0,1,2$.

- Cuando $q=0$ usamos la $(2^{k-1},k,1)$-condición de formas lineales. En este caso aparecen formas lineales en las variables $x,h_1,\dots,h_{k-1}$, del tipo

  $$(x,h_1,\dots,h_{k-1})\mapsto x+\omega\cdot h,\ \text{donde } \omega\in\{0,1\}^{k-1}.$$

  Existen $2^{k-1}$ formas de este tipo.

- En caso que $q=1$ usamos la $\big(2^{k-2}(k+1),2k-2,k\big)$-condición de formas lineales. En este caso se presentan formas lineales en las variables $x,h_1,\dots,h_{k-1},y_1,\dots,y_{k-2}$, de los siguientes tipos: para cada $1\le i\le k-1$ y cada $\omega\in\{0,1\}^{k-1}$, con $\omega_i=0$,

  $$(y_1,\dots,y_{k-2},x,h_1,\dots,h_{k-1})\mapsto\phi_i(y+\omega h);$$

  y para cada $\omega\in\{0,1\}^{k-1}$,

  $$(y_1,\dots,y_{k-2},x)\mapsto x+\omega\cdot h.$$

Del primer tipo hay $2^{k-2}(k-1)$ formas, y del segundo hay $2^{k-1}$ formas. Luego, en total hay $2^{k-2}(k+1)$ formas. Las formas $\phi_i$ involucran términos fraccionarios del tipo $\frac{1}{c_j}$ donde $c_i \in \{-k+1, \dots, -1, 0, 1, \dots, k-1\}$. Así, $L_0 = k$.

- Finalmente, cuando $q = 2$ usamos la $(k2^{k-1}, 3k-4, k)$-condición de formas lineales. En este caso aparecen formas en las variables $x, h_1, \dots, h_{k-1}, y_1, \dots, y_{k-2}, y'_1, \dots, y'_{k-2}$ (hay $k + 2(k-2) = 3k-4$ variables). Se presentan tres tipos de formas lineales. Las del primer tipo son

$$(y_1, \dots, y_{k-2}, x, h_1, \dots, h_{k-1}) \mapsto \phi_i(y + \omega h),$$

para cada $1 \leq i \leq k-1$ y cada $\omega \in \{0,1\}^{k-1}$, con $\omega_i = 0$, (de este tipo hay $2^{k-2}(k-1)$ formas). Las del segundo tipo son

$$(y'_1, \dots, y'_{k-2}, x, h_1, \dots, h_{k-1}) \mapsto \phi_i(y' + \omega h),$$

para cada $1 \leq i \leq k-1$ y cada $\omega \in \{0,1\}^{k-1}$, con $\omega_i = 1$, (de este tipo hay $2^{k-2}(k-1)$ formas). Y las del tercer tipo son

$$(h_1, \dots, h_{k-1}) \mapsto x + \omega \cdot h,$$

donde $\omega \in \{0,1\}^{k-1}$ (hay $2^{k-1}$ formas de este tipo). En total hay $k2^{k-1}$ formas. Además, las formas $\phi_i$ involucran términos de la forma $\frac{1}{c_j}$, con $|c_j| \leq k$, y por lo tanto $L_0 = k$. Recordemos que $y_{k-1} = x - y_1 - \cdots - y_{k-2}$ y $y'_{k-1} = x - y'_1 - \cdots - y'_{k-2}$.

□

Con lo cual, concluimos la prueba de la proposición 6.3. □

## 6.4. Funciones Gowers anti-uniformes

Habiendo estudiado la $U^{k-1}$-norma, ahora estudiaremos la norma dual $(U^{k-1})^*$.

**Definición 6.8.** *Dada una función $g : \mathbb{Z}_N \to \mathbb{R}$ definimos su* **norma dual de Gowers** *como*

$$||g||_{(U^{k-1})^*} = \sup\Big\{|\langle f, g\rangle| : f \in U^{k-1}(\mathbb{Z}_N),\ ||f||_{U^{k-1}} \leq 1\Big\}. \tag{6.37}$$

*Y decimos que $g$ es* **anti-uniforme según Gowers** *si $||g||_{(U^{k-1})^*} = O(1)$ y $||g||_{L^\infty} = O(1)$.*

Si $g$ es anti-uniforme según Gowers y si $\big|\langle f,g\rangle\big|$ es grande, entonces $f$ no puede ser uniforme según Gowers (esto es, $f$ no puede tener norma de Gowers pequeña) puesto que

$$\left|\left\langle \frac{f}{||f||_{U^{k-1}}}, g\right\rangle\right| \leq ||g||_{(U^{k-1})^*} \text{ para } f \neq 0,$$

y entonces

$$\big|\langle f,g\rangle\big| \leq ||f||_{U^{k-1}}||g||_{(U^{k-1})^*}.$$

La $(U^{k-1})^*$-norma dual es una norma genuina para $k \geq 3$, puesto que la $U^{k-1}$-norma es una auténtica norma para $k \geq 3$. En el caso $k = 3$ tenemos la fórmula explícita,

$$||g||_{(U^2)^*} = \left(\sum_{\xi\in\mathbb{Z}_N} \big|\widehat{g}(\xi)\big|^{4/3}\right)^{3/4} = ||\widehat{g}||_{l^{4/3}(\mathbb{Z}_N)}.$$

Por ejemplo, de la fórmula de inversión y la desigualdad de Hölder,

$$\begin{aligned}
\big|\langle f,g\rangle\big| &= \Big|\mathbb{E}\big(f(x)g(x)\big|x\in\mathbb{Z}_N\big)\Big| = \left|\mathbb{E}\left(\sum_{\xi\in\mathbb{Z}_N}\widehat{f}(\xi)\omega^{x\xi}g(x)\middle|x\in\mathbb{Z}_N\right)\right| \\
&= \left|\sum_{\xi\in\mathbb{Z}_N}\widehat{f}(\xi)\mathbb{E}\big(\omega^{x\xi}g(x)\big|x\in\mathbb{Z}_N\big)\right| \\
&= \left|\sum_{\xi\in\mathbb{Z}_N}\widehat{f}(\xi)\widehat{g}(-\xi)\right| \leq \left(\sum_{\xi\in\mathbb{Z}_N}\big|\widehat{f}(\xi)\big|^4\right)^{1/4}\left(\sum_{\xi\in\mathbb{Z}_N}\big|\widehat{g}(-\xi)\big|^{4/3}\right)^{3/4} \\
&= ||f||_{U^2}\left(\sum_{\xi\in\mathbb{Z}_N}\big|\widehat{g}(\xi)\big|^{4/3}\right)^{3/4}
\end{aligned}$$

Así tenemos que $||g||_{(U^2)^*} \leq \left(\sum_{\xi\in\mathbb{Z}_N}\big|\widehat{g}(\xi)\big|^{4/3}\right)^{3/4} = ||\widehat{g}||_{l^{4/3}(\mathbb{Z}_N)}$.

Además, por la definición de la $(U^{k-1})^*$-norma dual,

$$||g||_{(U^{k-1})^*} = \sup\Big\{\big|\langle f,g\rangle\big| : f \in U^{k-1}(\mathbb{Z}_N),\ ||f||_{U^{k-1}} \leq 1\Big\},$$

Así tenemos que, para $f \neq 0$, $\left|\left\langle \frac{f}{||f||_{U^{k-1}}}, g\right\rangle\right| \leq ||g||_{(U^{k-1})^*}$, y por lo tanto

$$\big|\langle f,g\rangle\big| \leq ||f||_{U^{k-1}}||g||_{(U^{k-1})^*}, \text{ para toda } f : \mathbb{Z}_N \to \mathbb{R}.$$

En esta sección veremos cómo generar una clase de funciones anti-uniformes según Gowers. En la siguiente sección veremos que toda función se puede descomponer en una parte uniforme según Gowers, más otra parte anti-uniforme según Gowers que es acotada.

Un modo sencillo de generar funciones anti-uniformes según Gowers es la siguiente. Para cada función $F \in L^1(\mathbb{Z}_N)$ definimos la ***función dual*** $\mathcal{D}F$ de $F$ como

$$\mathcal{D}F(x) = \mathbb{E}\left( \prod_{\omega \in \{0,1\}^{k-1},\, \omega \neq 0} F(x+\omega\cdot h) \middle| h \in \mathbb{Z}_N^{k-1} \right) \tag{6.38}$$

donde 0 denota al elemento de $\{0,1\}^{k-1}$ consistente de coordenadas nulas.

**Lema 6.5.** *Sea $\nu$ una medida $k$-pseudoaleatoria, y sea $F \in L^1(\mathbb{Z}_N)$. Entonces, tenemos las identidades*

$$\big\langle F, \mathcal{D}F \big\rangle = ||F||_{U^{k-1}}^{2^{k-1}} \tag{6.39}$$

*y*

$$||\mathcal{D}F||_{(U^{k-1})^*} = ||F||_{U^{k-1}}^{2^{k-1}-1}. \tag{6.40}$$

*Si además asumimos que*

$$\big|F(x)\big| \leq \nu(x) + 1 \text{ para todo } x \in \mathbb{Z}_N,$$

*entonces tenemos que*

$$||\mathcal{D}F||_{L^\infty} \leq 2^{2^{k-1}-1} + o(1). \tag{6.41}$$

*Demostración.* De (2.12) y (6.38), tenemos que

$$\begin{aligned}
\big|\langle \mathcal{D}F, F\rangle\big| &= \left| \mathbb{E}\Big( F(x)\mathbb{E}\Big( \prod_{\omega\in\{0,1\}^{k-1},\omega\neq 0} F(x+\omega\cdot h)\Big| h\in\mathbb{Z}_N^{k-1}\Big)\Big| x\in\mathbb{Z}_N\Big)\right| \\
&= \left| \mathbb{E}\Big( F(x) \prod_{\omega\in\{0,1\}^{k-1},\omega\neq 0} F(x+\omega\cdot h)\Big| x\in\mathbb{Z}_N, h\in\mathbb{Z}_N^{k-1}\Big)\right| \\
&= \left| \mathbb{E}\Big( \prod_{\omega\in\{0,1\}^{k-1}} F(x+\omega\cdot h)\Big| x\in\mathbb{Z}_N, h\in\mathbb{Z}_N^{k-1}\Big)\right| \\
&= ||F||_{U^{k-1}}^{2^{k-1}},
\end{aligned}$$

esto es,

$$\big|\langle \mathcal{D}F, F\rangle\big| = ||F||_{U^{k-1}}^{2^{k-1}}.$$

En caso que $F = 0$, de (6.38) tenemos que (6.40) se verifica de hecho. Supongamos que $F \neq 0$. De (6.37) y (6.39) tenemos que

$$\begin{aligned}
||\mathcal{D}F||_{(U^{k-1})^*} &= \sup\Big\{ \big|\langle \mathcal{D}F, f\rangle\big| : f\in U^{k-1}, ||f||_{U^{k-1}} \leq 1\Big\} \\
&\geq \left| \left\langle \mathcal{D}F, \frac{F}{||F||_{U^{k-1}}} \right\rangle \right| = \frac{1}{||F||_{U^{k-1}}} \langle \mathcal{D}F, F\rangle \\
&= ||F||_{U^{k-1}}^{2^{k-1}-1}.
\end{aligned}$$

Por otro lado, dado una función arbitraria $f : \mathbb{Z}_N \to \mathbb{R}$,

$$\begin{aligned} \big|\langle \mathcal{D}F, f\rangle\big| &= \left|\mathbb{E}\Big(f(x)\mathbb{E}\Big(\prod_{\omega\in\{0,1\}^{k-1},\omega\neq 0} F(x+\omega\cdot h)\Big|h\in\mathbb{Z}_N^{k-1}\Big)\Big|x\in\mathbb{Z}_N\Big)\right| \\ &= \left|\mathbb{E}\Big(\prod_{\omega\in\{0,1\}^{k-1}} f_\omega(x+\omega\cdot h)\Big|x\in\mathbb{Z}_N, h\in\mathbb{Z}_N^{k-1}\Big)\right|, \end{aligned}$$

donde

$$f_\omega = \begin{cases} f & \text{si } \omega = 0, \\ F & \text{si } \omega \neq 0. \end{cases}$$

Así, por la desigualdad de Gowers Cauchy-Schwarz,

$$\begin{aligned} \big|\langle \mathcal{D}F, f\rangle\big| &= \Big|\big\langle (f_\omega)_{\omega\in\{0,1\}^{k-1}}\big\rangle_{U^{k-1}}\Big| \\ \big|\langle \mathcal{D}F, f\rangle\big| &\leq \prod_{\omega\in\{0,1\}^{k-1}} ||f_\omega||_{U^{k-1}} = ||f||_{U^{k-1}}||F||_{U^{k-1}}^{2^{k-1}-1}, \end{aligned}$$

esto es,

$$\big|\langle \mathcal{D}F, f\rangle\big| \leq ||f||_{U^{k-1}}||F||_{U^{k-1}}^{2^{k-1}-1} \text{ para cada } f \in U^{k-1},$$

por lo tanto $||\mathcal{D}F||_{(U^{k-1})^*} \leq ||F||_{U^{k-1}}^{2^{k-1}-1}$, y así

$$||\mathcal{D}F||_{(U^{k-1})^*} \leq ||F||_{U^{k-1}}^{2^{k-1}-1}.$$

Finalmente, mostremos (6.41). Por hipótesis $|F| \leq 2\Big(\dfrac{\nu+1}{2}\Big) = 2\nu_{1/2}$. Así, es suficiente mostrar que

$$\mathcal{D}\nu_{1/2}(x) \leq 1 + o(1)$$

uniformemente en la elección de $x \in \mathbb{Z}_N$. El lado izquierdo puede ser expandido, de (6.38), como

$$\mathbb{E}\Big(\prod_{\omega\in\{0,1\}^{k-1},\omega\neq 0} \nu_{1/2}(x+\omega\cdot h)\Big|h\in\mathbb{Z}_N^{k-1}\Big).$$

Usando la condición de formas lineales (6.4) (y el lema 6.1) esta expresión es $1 + o(1)$. (este es el único lugar en donde usamos la condición de formas lineales en el caso no homogéneo, donde alguno de los $b_i$ es no nulo, $b_i \neq 0$; aquí todos los $b_i$ son iguales a $x$). Observe que (6.8) corresponde al caso $k = 3$ de la aplicación de la condición de formas lineales a la situación actual. □

Llamaremos a las funciones de la forma $\mathcal{D}F$, donde $F$ es puntualmente acotada por $\nu + 1$ (esto es, $F(x) \leq \nu(x) + 1$ para todo $x \in \mathbb{Z}_N$) como *funciones anti-uniformes básicas según*

*Gowers*. Observe que de la ecuación (6.41) para $N$ suficientemente grande las funciones de este tipo toman valores sobre el intervalo $I = \big[ -2^{2^{k-1}}, 2^{2^{k-1}} \big]$.

**Proposición 6.4.** *Supongamos que $\nu$ sea una medida $k$-pseudoaleatoria. Sean $K \geq 1$ un entero fijo, $I = \big[ -2^{2^{k-1}}, 2^{2^{k-1}} \big]$ y $\Phi : I^K \to \mathbb{R}$ una función continua y $\mathcal{D}F_1, \ldots, \mathcal{D}F_K$ funciones anti-uniformes básicas según Gowers, y definamos la función $\psi : \mathbb{Z}_N \to \mathbb{R}$ por*

$$\psi(x) = \Phi\big(\mathcal{D}F_1, \ldots, \mathcal{D}F_K\big).$$

*Entonces*

$$\big\langle \nu - 1, \psi \big\rangle = o_{K,\Phi}(1).$$

*Además, si $\Phi$ varía sobre un subconjunto compacto $E \subset C^0(I^K)$ de funciones continuas sobre $I^K$ (en la topología uniforme) entonces la cota es uniforme en $\Phi$ (esto es, podemos reemplazar $o_{K,\Phi}(1)$ por $o_{K,E}(1)$ en este caso).*

Probaremos este resultado en dos pasos. Primero estableceremos el resultado para $\Phi$ polinomial y entonces usaremos un argumento de aproximación de Weierstrass para deducir el caso general.

Sea $K \geq 1$, y sean $F_1, \ldots, F_K \in L^1(\mathbb{Z}_N)$ funciones tales que

$$\big|F_j(x)\big| \leq \nu(x) + 1 \text{ para todo } x \in \mathbb{Z}_N,\ 1 \leq j \leq K.$$

Reemplazando $\nu$ por $(\nu+1)/2$, dividiendo los $F_j$ por 2, y usando el lema 6.1 podemos asumir sin pérdida de generalidad que

$$\big|F_j(x)\big| \leq \nu(x) \text{ para todo } x \in \mathbb{Z}_N,\ 1 \leq j \leq K. \tag{6.42}$$

**Lema 6.6.** *Sea $d \geq 1$ un entero. Para cualquier polinomio $P$ de $K$ variables y grado $d$ con coeficientes reales (independientes de $N$), tenemos que*

$$\big|\big|P(\mathcal{D}F_1, \ldots, \mathcal{D}F_K)\big|\big|_{(U^{k-1})^*} = O_{K,d,P}(1).$$

*Demostración.* Por la linealidad es suficiente probar esto cuando $P$ es un monomio. Redefiniendo el monomio $P$ si es necesario, podemos suponer que $P$ es de la forma

$$P(x_1, x_2, \ldots, x_{K-1}, x_K) = x_1 x_2 \cdots x_{K-1} x_K.$$

En efecto, si

$$P(x_1, \ldots, x_K) = x_1^{e_1} \cdots x_K^{e_K}, \text{ donde } e_1 + \cdots + e_K = d,$$

tenemos que $P(\mathcal{D}F_1,\dots,\mathcal{D}F_K)=\mathcal{D}F_1^{e_1}\cdots\mathcal{D}F_K^{e_K}$. Si consideramos el polinomio $Q$ definido por $Q(y_1,y_2\dots,y_d)=y_1y_2\cdots y_d$, entonces

$$P(\mathcal{D}F_1,\dots,\mathcal{D}F_K)=Q(\mathcal{D}F_1,\dots,\mathcal{D}F_1,\dots,\mathcal{D}F_K,\dots,\mathcal{D}F_K),$$

donde $\mathcal{D}F_i$ en $Q$ se repite $e_i$ veces.

Requerimos probar que $\left\langle f,\prod_{j=1}^{K}\mathcal{D}F_j\right\rangle=O_K(1)$ para toda $f:\mathbb{Z}_N\to\mathbb{R}$ tal que $||f||_{U^{k-1}}\leq 1$.

Por definición $\left\langle f,\prod_{j=1}^{K}\mathcal{D}F_j\right\rangle$ es igual a

$$\mathbb{E}\bigg(f(x)\prod_{j=1}^{K}\mathbb{E}\bigg(\prod_{\omega\in\{0,1\}^{k-1},\omega\neq 0}F_j(x+\omega\cdot h^{(j)})\bigg|h^{(j)}\in\mathbb{Z}_N^{k-1}\bigg)\bigg|x\in\mathbb{Z}_N\bigg).$$

Hacemos el cambio de variables $h^{(j)}=h+H^{(j)}$ para cada $h\in\mathbb{Z}_N^{k-1}$ y reescribimos la expresión anterior como

$$\begin{aligned}&\mathbb{E}\bigg(f(x)\prod_{j=1}^{K}\mathbb{E}\bigg(\prod_{\omega\in\{0,1\}^{k-1},\omega\neq 0}F_j(x+\omega\cdot h^{(j)})\bigg|h^{(j)}\in\mathbb{Z}_N^{k-1}\bigg)\bigg|x\in\mathbb{Z}_N,h\in\mathbb{Z}_N^{k-1}\bigg)\\&=\mathbb{E}\bigg(f(x)\prod_{j=1}^{K}\mathbb{E}\bigg(\prod_{\omega\in\{0,1\}^{k-1},\omega\neq 0}F_j(x+\omega\cdot H^{(j)}+\omega\cdot h)\bigg|H^{(j)}\in\mathbb{Z}_N^{k-1}\bigg)\\&\qquad\qquad\bigg|x\in\mathbb{Z}_N,h\in\mathbb{Z}_N^{k-1}\bigg)\end{aligned}\tag{6.43}$$

Expandiendo la productoria en $j$ e intercambiando esperanzas (promedios $\mathbb{E}$), esto último se reescribirá como

$$\mathbb{E}\bigg(\Big\langle\big(f_{\omega,H}\big)_{\omega\in\{0,1\}^{k-1}}\Big\rangle_{U^{k-1}}\bigg|H\in\big(\mathbb{Z}_N^{k-1}\big)^K\bigg),\tag{6.44}$$

donde $H=(H^{(1)},\dots,H^{(K)})$, $f_{0,H}=f$ y $f_{\omega,H}=g_{\omega\cdot H}$ para $\omega\neq 0$, donde

$$\omega\cdot H=\big(\omega\cdot H^{(1)},\dots,\omega\cdot H^{(K)}\big)$$

y

$$g_{u^{(1)},\dots,u^{(K)}}=\prod_{j=1}^{K}F_j(x+u^{(j)})\text{ para cualesquiera }u^{(1)},\dots,u^{(K)}\in\mathbb{Z}_N.\tag{6.45}$$

Por ejemplo, en el caso en el que $K=2$ se tiene de (6.43),

$$\begin{aligned}&\mathbb{E}\bigg(f(x)\mathbb{E}\bigg(\prod_{\omega\in\{0,1\}^{k-1},\omega\neq 0}F_1(x+\omega\cdot H^{(1)}+\omega\cdot h)\bigg|H^{(1)}\in\mathbb{Z}_N^{k-1}\bigg)\times\\&\quad\times\mathbb{E}\bigg(\prod_{\omega\in\{0,1\}^{k-1},\omega\neq 0}F_2(x+\omega\cdot H^{(2)}+\omega\cdot h)\bigg|H^{(2)}\in\mathbb{Z}_N^{k-1}\bigg)\bigg|x\in\mathbb{Z}_N,h\in\mathbb{Z}_N^{k-1}\bigg)\end{aligned}$$

$$= \mathbb{E}\Bigg( \mathbb{E}\Big( f(x) \prod_{\omega\in\{0,1\}^{k-1},\omega\neq 0} F_1(x+\omega\cdot H^{(1)}+\omega\cdot h)\Big| H^{(1)}\in\mathbb{Z}_N^{k-1}\Big)\times$$

$$\times\, \mathbb{E}\Big( \prod_{\omega\in\{0,1\}^{k-1},\omega\neq 0} F_2(x+\omega\cdot H^{(2)}+\omega\cdot h)\Big| H^{(2)}\in\mathbb{Z}_N^{k-1}\Big)\Bigg| x\in\mathbb{Z}_N, h\in\mathbb{Z}_N^{k-1}\Bigg)$$

$$= \mathbb{E}\Bigg( \mathbb{E}\Big( f(x) \prod_{\omega\in\{0,1\}^{k-1},\omega\neq 0} \Big[F_1(x+\omega\cdot H^{(1)}+\omega\cdot h)F_2(x+\omega\cdot H^{(2)}+\omega\cdot h)\Big]\Big| H^{(1)},$$

$$H^{(2)}\in\mathbb{Z}_N^{k-1}\Big)\Bigg| x\in\mathbb{Z}_N, h\in\mathbb{Z}_N^{k-1}\Bigg)$$

$$= \mathbb{E}\Bigg( f(x) \prod_{\omega\in\{0,1\}^{k-1},\omega\neq 0}\prod_{i=1}^{2} F_i(x+\omega\cdot H^{(i)}+\omega\cdot h)\Bigg| x\in\mathbb{Z}_N, h, H^{(1)}, H^{(2)}\in\mathbb{Z}_N^{k-1}\Bigg)$$

$$= \mathbb{E}\Bigg( \mathbb{E}\Big( f(x) \prod_{\omega\in\{0,1\}^{k-1},\omega\neq 0}\prod_{i=1}^{2} F_i(x+\omega\cdot H^{(i)}+\omega\cdot h)\Big| x\in\mathbb{Z}_N, h\in\mathbb{Z}_N^{k-1}\Big)\Bigg| H^{(1)}, H^{(2)}\in\mathbb{Z}_N^{k-1}\Bigg)$$

$$= \mathbb{E}\Bigg( \mathbb{E}\Big( \prod_{\omega\in\{0,1\}^{k-1}} F_{\omega,H^{(1)},H^{(2)}}(x+\omega\cdot h)\Big| x\in\mathbb{Z}_N, h\in\mathbb{Z}_N^{k-1}\Big)\Bigg| H^{(1)}, H^{(2)}\in\mathbb{Z}_N^{k-1}\Bigg),$$

donde

$$F_{0,H^{(1)},H^{(2)}}(\xi) = f(\xi) \text{ para todo } \xi\in\mathbb{Z}_N,$$

y

$$F_{\omega,H^{(1)},H^{(2)}}(\xi) = F_1(\xi+\omega\cdot H^{(1)})F_2(\xi+\omega\cdot H^{(2)}) \text{ para todo } \omega\neq 0,$$

y todo $\xi\in\mathbb{Z}_N$. Por lo tanto

$$\mathbb{E}\Bigg( \mathbb{E}\Big( \prod_{\omega\in\{0,1\}^{k-1}} F_{\omega,H^{(1)},H^{(2)}}(x+\omega\cdot h)\Big| x\in\mathbb{Z}_N, h\in\mathbb{Z}_N^{k-1}\Big)\Bigg| H^{(1)}, H^{(2)}\in\mathbb{Z}_N^{k-1}\Bigg)$$

$$= \mathbb{E}\Bigg( \Big\langle \big(F_{\omega,H^{(1)},H^{(2)}}\big)_{\omega\in\{0,1\}^{k-1}}\Big\rangle_{U^{k-1}}\Bigg| H^{(1)}, H^{(2)}\in\mathbb{Z}_N^{k-1}\Bigg).$$

Volviendo a (6.44), por la desigualdad de Cauchy-Schwarz tenemos

$$\Big\langle f, \prod_{j=1}^{K}\mathcal{D}F_j\Big\rangle = \mathbb{E}\Bigg( \Big\langle (g_{\omega\cdot H})_{\omega\in\{0,1\}^{k-1}}\Big\rangle_{U^{k-1}}\Bigg| H\in\big(\mathbb{Z}_N^{k-1}\big)^K\Bigg)$$

$$\leq \mathbb{E}\Bigg( \prod_{\omega\in\{0,1\}^{k-1}} ||g_{\omega\cdot H}||_{U^{k-1}}\Bigg| H\in\big(\mathbb{Z}_N^{k-1}\big)^K\Bigg)$$

$$\leq \mathbb{E}\Bigg( ||f||_{U^{k-1}} \prod_{\omega\in\{0,1\}^{k-1},\,\omega\neq 0} ||g_{\omega\cdot H}||_{U^{k-1}}\Bigg| H\in\big(\mathbb{Z}_N^{k-1}\big)^K\Bigg)$$

$$\leq \mathbb{E}\Bigg( \prod_{\omega\in\{0,1\}^{k-1},\,\omega\neq 0} ||g_{\omega\cdot H}||_{U^{k-1}}\Bigg| H\in\big(\mathbb{Z}_N^{k-1}\big)^K\Bigg)$$

$$\leq \mathbb{E}\Bigg( \prod_{\omega\in\{0,1\}^{k-1},\,\omega\neq 0} \Big(||g_{\omega\cdot H}||_{U^{k-1}}^{2^{k-1}-1}\Big)^{1/(2^{k-1}-1)} \Bigg| H\in \big(\mathbb{Z}_N^{k-1}\big)^K\Bigg)$$
$$\leq \mathbb{E}\Bigg( \Bigg[ \prod_{\omega\in\{0,1\}^{k-1},\,\omega\neq 0} ||g_{\omega\cdot H}||_{U^{k-1}}^{2^{k-1}-1}\Bigg]^{1/(2^{k-1}-1)} \Bigg| H\in \big(\mathbb{Z}_N^{k-1}\big)^K\Bigg).$$

Ahora, por la desigualdad de la media geométrica y media aritmética, tenemos

$$\begin{aligned}\Big\langle f, \prod_{j=1}^K \mathcal{D}F_j\Big\rangle &\leq \mathbb{E}\Bigg(\frac{1}{2^{k-1}-1}\sum_{\omega\in\{0,1\}^{k-1},\,\omega\neq 0} ||g_{\omega\cdot H}||_{U^{k-1}}^{2^{k-1}-1}\Bigg| H\in \big(\mathbb{Z}_N^{k-1}\big)^K\Bigg)\\ &= \mathbb{E}\Big(||g_{\omega\cdot H}||_{U^{k-1}}^{2^{k-1}-1}\Big|\omega\in\{0,1\}^{k-1},\omega\neq 0, H\in \big(\mathbb{Z}_N^{k-1}\big)^K\Big)\\ &= \mathbb{E}\Big(\mathbb{E}\Big(||g_{\omega\cdot H}||_{U^{k-1}}^{2^{k-1}-1}\Big| H\in \big(\mathbb{Z}_N^{k-1}\big)^K\Big)\Big|\omega\in\{0,1\}^{k-1},\omega\neq 0\Big).\end{aligned}$$

Luego es suficiente probar que

$$\mathbb{E}\Big(||g_{\omega\cdot H}||_{U^{k-1}}^{2^{k-1}-1}\Big| H\in \big(\mathbb{Z}_N^{k-1}\big)^K\Big) = O_K(1)$$

para cada $\omega\in\{0,1\}^{k-1}$, $\omega\neq 0$.

Fijemos $\omega\in\{0,1\}^{k-1}$, $\omega\neq 0$. Por la desigualdad de Hölder (para $p_1,\dots,p_m\in\mathbb{Z}$ enteros positivos tales que $1/p_1+\cdots+1/p_m=1$, se tiene que $||f_1\cdots f_m||_1\leq ||f_1||_{p_1}\cdots||f_m||_{p_m}$), tenemos que

$$\begin{aligned}\mathbb{E}\Big(||g_{\omega\cdot H}||_{U^{k-1}}^{2^{k-1}-1}\Big| H\in \big(\mathbb{Z}_N^{k-1}\big)^K\Big) &= \mathbb{E}\Big(1\cdot ||g_{\omega\cdot H}||_{U^{k-1}}\cdots||g_{\omega\cdot H}||_{U^{k-1}}\Big| H\in \big(\mathbb{Z}_N^{k-1}\big)^K\Big)\\ &= \big|\big|\phi_0\phi_1\cdots\phi_{2^{k-1}-1}\big|\big|_1,\end{aligned}$$

donde $\phi_0(H)=1$ y $\phi_n(H)=||g_{\omega\cdot H}||_{U^{k-1}}$ para cada $H\in\big(\mathbb{Z}_N^{k-1}\big)^K$, $1\leq n\leq 2^{k-1}-1$. Así

$$\begin{aligned}\big|\big|\phi_0\phi_1\cdots\phi_{2^{k-1}-1}\big|\big|_1 &\leq ||\phi_0||_{2^{k-1}}\prod_{n=1}^{2^{k-1}-1}||\phi_n||_{2^{k-1}}\\ &\leq ||\phi_1||_{2^{k-1}}^{2^{k-1}-1}\\ &= \mathbb{E}\Big(\phi_1(H)^{2^{k-1}}\Big| H\in \big(\mathbb{Z}_N^{k-1}\big)^K\Big)^{(2^{k-1}-1)/2^{k-1}}\\ &\leq \mathbb{E}\Big(\phi_1(H)^{2^{k-1}}\Big| H\in \big(\mathbb{Z}_N^{k-1}\big)^K\Big)\\ &= \mathbb{E}\Big(||g_{\omega\cdot H}||_{U^{k-1}}^{2^{k-1}}\Big| H\in \big(\mathbb{Z}_N^{k-1}\big)^K\Big).\end{aligned}$$

Por lo tanto, es suficiente mostrar que

$$\mathbb{E}\Big(||g_{\omega\cdot H}||_{U^{k-1}}^{2^{k-1}}\Big| H\in \big(\mathbb{Z}_N^{k-1}\big)^K\Big) = O_K(1).$$

Como $\omega \neq 0$ (es fijo), la función

$$\Phi : H = (H^{(1)}, \dots, H^{(K)}) \mapsto \omega \cdot H = (\omega \cdot H^{(1)}, \dots, \omega \cdot H^{(K)})$$

es un cubrimiento uniforme de $\mathbb{Z}_N^K$ por $\left(\mathbb{Z}_N^{k-1}\right)^K$. Así, por (6.13) podemos modificar el lado izquierdo de la igualdad anterior

$$\begin{aligned}
&\mathbb{E}\Big( ||g_{\Phi(H)}||_{U^{k-1}}^{2^{k-1}} \Big| H \in \left(\mathbb{Z}_N^{k-1}\right)^K \Big) \\
&\qquad = \mathbb{E}\Big( ||g_J||_{U^{k-1}}^{2^{k-1}} \Big| J \in \mathbb{Z}_N^K \Big) \\
&\qquad = \mathbb{E}\Big( ||g_{u^{(1)},\dots,u^{(K)}}||_{U^{k-1}}^{2^{k-1}} \Big| u^{(1)}, \dots, u^{(K)} \in \mathbb{Z}_N \Big) \\
&\qquad = \mathbb{E}\Big( \mathbb{E}\Big( \prod_{\tilde{\omega} \in \{0,1\}^{k-1}} g_{u^{(1)},\dots,u^{(K)}}(x + \tilde{\omega} \cdot h) \Big| x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \Big) \Big| u^{(1)}, \dots, u^{(K)} \in \mathbb{Z}_N \Big) \\
&\qquad = \mathbb{E}\Big( \prod_{\tilde{\omega} \in \{0,1\}^{k-1}} \prod_{j=1}^{K} F_j\big(x + \tilde{\omega} \cdot h + u^{(j)}\big) \Big| x, u^{(1)}, \dots, u^{(K)} \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \Big) \\
&\qquad = \mathbb{E}\Big( \prod_{j=1}^{K} \prod_{\tilde{\omega} \in \{0,1\}^{k-1}} F_j\big(x + \tilde{\omega} \cdot h + u^{(j)}\big) \Big| x, u^{(1)}, \dots, u^{(K)} \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \Big) \\
&\qquad = \mathbb{E}\Big( \prod_{j=1}^{K} \mathbb{E}\Big( \prod_{\tilde{\omega} \in \{0,1\}^{k-1}} F_j\big(x + \tilde{\omega} \cdot h + u^{(j)}\big) \Big| u^{(j)} \in \mathbb{Z}_N \Big) \Big| x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \Big) \\
&\qquad = \mathbb{E}\Big( \prod_{j=1}^{K} \mathbb{E}\Big( \prod_{\tilde{\omega} \in \{0,1\}^{k-1}} F_j\big(x + \tilde{\omega} \cdot h + u\big) \Big| u \in \mathbb{Z}_N \Big) \Big| x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \Big).
\end{aligned}$$

Aplicando (6.42), es suficiente mostrar que

$$\mathbb{E}\Big( \mathbb{E}\Big( \prod_{\tilde{\omega} \in \{0,1\}^{k-1}} \nu(x + \tilde{\omega} \cdot h + u) \Big| u \in \mathbb{Z}_N \Big)^K \Big| x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \Big) = O_K(1).$$

Haciendo el cambio de variables $y = u + x$ obtenemos

$$\begin{aligned}
&\mathbb{E}\Big( \mathbb{E}\Big( \prod_{\tilde{\omega} \in \{0,1\}^{k-1}} \nu(x + \tilde{\omega} \cdot h + u) \Big| u \in \mathbb{Z}_N \Big)^K \Big| x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \Big) \\
&\qquad = \mathbb{E}\Big( \mathbb{E}\Big( \prod_{\tilde{\omega} \in \{0,1\}^{k-1}} \nu(y + \tilde{\omega} \cdot h) \Big| y \in \mathbb{Z}_N \Big)^K \Big| x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \Big) \\
&\qquad = \mathbb{E}\Big( \mathbb{E}\Big( \prod_{\tilde{\omega} \in \{0,1\}^{k-1}} \nu(y + \tilde{\omega} \cdot h) \Big| y \in \mathbb{Z}_N \Big)^K \Big| h \in \mathbb{Z}_N^{k-1} \Big).
\end{aligned}$$

Por lo tanto basta mostrar que

$$\mathbb{E}\Big( \mathbb{E}\Big( \prod_{\tilde{\omega} \in \{0,1\}^{k-1}} \nu(y + \tilde{\omega} \cdot h) \Big| y \in \mathbb{Z}_N \Big)^K \Big| h \in \mathbb{Z}_N^{k-1} \Big) = O_K(1).$$

Para ello aplicamos la condición de correlación (definición 6.4). La $2^{k-1}$-condición de correlación nos dice que

$$\mathbb{E}\Bigg(\prod_{\tilde{\omega}\in\{0,1\}^{k-1}}\nu(y+h\cdot\tilde{\omega})\Bigg|y\in\mathbb{Z}_N\Bigg)\leq\sum_{\tilde{\omega},\tilde{\omega}'\in\{0,1\}^{k-1},\tilde{\omega}\neq\tilde{\omega}'}\tau\big(h\cdot(\tilde{\omega}-\tilde{\omega}')\big),$$

donde $\tau$ es la función peso que satisface

$$\mathbb{E}\big(\tau^q(m)\big|m\in\mathbb{Z}_N\big)=O_q(1)\text{ para todo }q\geq 1. \tag{6.46}$$

Aplicando la desigualdad triangular en $L_K\big(\mathbb{Z}_N^{k-1}\big)$, obtenemos

$$\begin{aligned}\mathbb{E}\Bigg(\mathbb{E}\Big(\prod_{\tilde{\omega}\in\{0,1\}^{k-1}}\nu(y+\tilde{\omega}\cdot h)\Big|y\in\mathbb{Z}_N\Big)^K\Bigg|h\in\mathbb{Z}_N^{k-1}\Bigg)^{1/K}&\\ \leq\mathbb{E}\Bigg(\Bigg[\sum_{\tilde{\omega},\tilde{\omega}'\in\{0,1\}^{k-1},\tilde{\omega}\neq\tilde{\omega}'}\tau\big(h\cdot(\tilde{\omega}-\tilde{\omega}')\big)\Bigg]^K\Bigg|h\in\mathbb{Z}_N^{k-1}\Bigg)^{1/K}&\\ \leq\sum_{\tilde{\omega},\tilde{\omega}'\in\{0,1\}^{k-1},\tilde{\omega}\neq\tilde{\omega}'}\Bigg[\mathbb{E}\Big(\tau\big(h\cdot(\tilde{\omega}-\tilde{\omega}')\big)^K\Big|h\in\mathbb{Z}_N^{k-1}\Big)\Bigg]^{1/K}.&\end{aligned}$$

Luego, queda por mostrar que

$$\mathbb{E}\Big(\tau\big(h\cdot(\tilde{\omega}-\tilde{\omega}')\big)^K\Big|h\in\mathbb{Z}_N^{k-1}\Big)=O_K(1)$$

para cualesquiera $\tilde{\omega},\tilde{\omega}'\in\{0,1\}^{k-1}$, tales que $\tilde{\omega}\neq\tilde{\omega}'$. Fijemos $\tilde{\omega},\tilde{\omega}'\in\{0,1\}^{k-1}$, tales que $\tilde{\omega}\neq\tilde{\omega}'$. Como $h\mapsto h\cdot(\tilde{\omega}-\tilde{\omega}')$ es un cubrimiento uniforme de $\mathbb{Z}_N$ por $\mathbb{Z}_N^{k-1}$. Por lo tanto de (6.13)

$$\mathbb{E}\Big(\tau\big(h\cdot(\tilde{\omega}-\tilde{\omega}')\big)^K\Big|h\in\mathbb{Z}_N\Big)=\mathbb{E}\Big(\tau(m)^K\Big|m\in\mathbb{Z}_N\Big).$$

Así, de (6.46) con $q=K$, el lema queda demostrado. □

## 6.5. Conjuntos de Bohr generalizados y $\sigma$-álgebras

Para usar la proposición 6.4 asociaremos un $\sigma$-álgebra a cada función anti-uniforme según Gowers, tal que toda función medible en cada $\sigma$-álgebra pueda ser aproximada por una función del tipo considerado en la proposición 6.4.

**Definición 6.9.** *Un* **$\sigma$-álgebra** $\mathcal{B}$ *en* $\mathbb{Z}_N$ *es cualquier colección de subconjuntos de* $\mathbb{Z}_N$ *que contiene el conjunto vacío* $\emptyset$ *y* $\mathbb{Z}_N$*, y que es cerrado bajo complemento, unión e intersección de conjuntos. Definimos los* **átomos** *de la* $\sigma$*-álgebra* $\mathcal{B}$ *como la familia de los elementos no*

*vacíos minimales de $\mathcal{B}$ (con respecto a la inclusión de conjuntos). Es claro que los átomos de $\mathcal{B}$ forman una partición de $\mathbb{Z}_N$, y $\mathcal{B}$ consiste de uniones arbitrarias de sus átomos (incluyendo la unión vacía, $\emptyset$). Una función $f \in L^q(\mathbb{Z}_N)$ es medible con respecto al $\sigma$-álgebra $\mathcal{B}$ si los conjuntos de nivel $\left\{f^{-1}\big(\{x\}\big) : x \in \mathbb{R}\right\}$ de $f$ pertenecen a $\mathcal{B}$, o equivalentemente si $f$ es constante en los átomos de $\mathcal{B}$.*

Definimos $L^q(\mathcal{B})$, $L^q(\mathcal{B}) \subset L^q(\mathbb{Z}_N)$, como el subespacio de $L^q(\mathbb{Z}_N)$ consistente de las funciones $\mathcal{B}$-medibles, equipadas con la norma $L^q$. Podemos definir el operador esperanza condicional $f \mapsto \mathbb{E}(f|\mathcal{B})$ mediante

$$\mathbb{E}(f|\mathcal{B})(x) = \mathbb{E}\big(f(y)\big|y \in \mathcal{B}(x)\big)$$

para todo $x \in \mathbb{Z}_N$, donde $\mathcal{B}(x)$ es el único átomo en $\mathcal{B}$ que contiene a $x$.
Si $\mathcal{B}_1, \ldots, \mathcal{B}_K$ son $\sigma$-álgebras, entonces denotamos por $\bigvee_{j=1}^K \mathcal{B}_j = \mathcal{B}_1 \vee \cdots \vee \mathcal{B}_K$ la $\sigma$-álgebra generada por estas álgebras, es decir, la $\sigma$-álgebra cuyos átomos son las intersecciones de los átomos en $\mathcal{B}_1, \ldots, \mathcal{B}_K$. Adoptaremos la convención de que cuando $K = 0$, $\bigvee_{j=1}^K \mathcal{B}_j$ representa la $\sigma$-álgebra trivial $\{\emptyset, \mathbb{Z}_N\}$.

**Proposición 6.5** (Cada función genera una $\sigma$-álgebra)**.** *Sea $\nu$ una medida $k$-pseudoaleatoria, sean $0 < \epsilon < 1$, $0 < \eta < 1/2$ y $G \in L^\infty(\mathbb{Z}_N)$ una función tomando valores en el intervalo $I = [-2^{2^{k-1}}, 2^{2^{k-1}}]$. Entonces existe un $\sigma$-álgebra $\mathcal{B}_{\epsilon,\eta}(G)$ con las siguientes propiedades:*

- *($G$ pertenece a este $\sigma$-álgebra) Para cualquier $\sigma$-álgebra $\mathcal{B}$, tenemos*

$$||G - \mathbb{E}(G|\mathcal{B} \cap \mathcal{B}_{\epsilon,\eta}(G))||_{L^\infty(\mathbb{Z}_N)} \leq \epsilon. \tag{6.47}$$

- *$\mathcal{B}_{\epsilon,\eta}(G)$ es generada por una cantidad de átomos del tipo $O(1/\epsilon)$, esto es, existe una constante $C$, que solo depende de $\epsilon$ y $\eta$ , tal que la cantidad de átomos que genera la $\sigma$-álgebra es menor que $C/\epsilon$ para $N$ suficientemente grande.*

- *(Aproximación por funciones continuas) Si $A$ es cualquier átomo en $\mathcal{B}_{\epsilon,\eta}(G)$, entonces existe una función continua $\Psi_A : I \to [0,1]$ tal que*

$$\big|\big|\big(1_A - \Psi_A(G)\big)(\nu + 1)\big|\big|_{L^1(\mathbb{Z}_N)} = O(\eta). \tag{6.48}$$

  *Además $\Psi_A$ pertenece a un conjunto compacto $E = E_{\epsilon,\eta}$ de $C^0(I)$ (que es independiente de $F, \nu, N$ y $A$).*

*Demostración.* Del teorema de Fubini y de (6.3) tenemos que

$$\int_0^1 \sum_{n\in\mathbb{Z}} \mathbb{E}\Big(1_{G(x)\in[\epsilon(n-\eta+\alpha),\epsilon(n+\eta+\alpha)]}\big(\nu(x)+1\big)\Big|x\in\mathbb{Z}_N\Big)d\alpha = 2\eta\mathbb{E}\big(\nu(x)+1\big|x\in\mathbb{Z}_N\big) = O(\eta).$$

En efecto,

$$\int_0^1 \sum_{n\in\mathbb{Z}} \mathbb{E}\Big(1_{G(x)\in[\epsilon(n-\eta+\alpha),\epsilon(n+\eta+\alpha)]}\big(\nu(x)+1\big)\Big|x\in\mathbb{Z}_N\Big)d\alpha$$
$$= \mathbb{E}\left(\sum_{n\in\mathbb{Z}}\int_0^1 1_{G(x)\in[\epsilon(n-\eta+\alpha),\epsilon(n+\eta+\alpha)]}\big(\nu(x)+\alpha\big)d\alpha\Big|x\in\mathbb{Z}_N\right).$$

Como $G$ toma valores en $I = \big[-2^{2^{k-1}}, 2^{2^{k-1}}\big]$, la sumatoria anterior (al parecer infinita) es finita. Además, para cada $x\in\mathbb{Z}_N$, solo interesan los valores $n\in\mathbb{Z}$ tales que

$$G(x)\in\big[\epsilon(n-\eta+\alpha),\epsilon(n+\eta+\alpha)\big] \text{ para algún } \alpha\in[0,1],$$

esto es, interesan los $n\in\mathbb{Z}$ tales que

$$G(x)\in\bigcup_{\alpha\in[0,1]}\big[\epsilon(n-\eta+\alpha),\epsilon(n+\eta+\alpha)\big] = \big[\epsilon(n-\eta),\epsilon(n+1+\eta)\big].$$

Sea $m\in\mathbb{Z}$ tal que $G(x)\in\big[\epsilon(m-\eta),\epsilon(m+1+\eta)\big]$, entonces

(i) $G(x)\in\big[\epsilon(n-\eta),\epsilon(n+1+\eta)\big]$ solo para $n=m$ cuando

$$G(x)\in\big(\epsilon(m+\eta),\epsilon(m+1-\eta)\big).$$

(ii) $G(x)\in\big[\epsilon(n-\eta),\epsilon(n+1+\eta)\big]$ solo para $n=m$ y $n=m-1$ cuando

$$G(x)\in\big[\epsilon(m+1-\eta),\epsilon(m+1+\eta)\big].$$

(iii) $G(x)\in\big[\epsilon(n-\eta),\epsilon(n+1+\eta)\big]$ solo para $n=m$ y $n=m+1$ cuando

$$G(x)\in\big[\epsilon(m-\eta),\epsilon(m+\eta)\big].$$

En cada uno de los tres casos anteriores se tiene que

$$\sum_{n\in\mathbb{Z}}\int_0^1 1_{G(x)\in[\epsilon(n-\eta+\alpha),\epsilon(n+\eta+\alpha)]}\big(\nu(x)+1\big)d\alpha$$
$$= \sum_{n\in\{m-1,m,m+1\}}\int_0^1 1_{G(x)\in[\epsilon(n-\eta+\alpha),\epsilon(n+\eta+\alpha)]}\big(\nu(x)+1\big)d\alpha$$
$$= 2\eta\big(\nu(x)+1\big).$$

Por ejemplo consideremos el caso (iii), en que $t = G(x) \in \big[\epsilon(m-\eta), \epsilon(m+\eta)\big]$, tenemos

$$\sum_{n\in\mathbb{Z}} \int_0^1 1_{G(x)\in[\epsilon(n-\eta+\alpha),\epsilon(n+\eta+\alpha)]}\big(\nu(x)+1\big)d\alpha$$

$$= \int_0^1 1_{G(x)\in[\epsilon(m-1-\eta+\alpha),\epsilon(m-1+\eta+\alpha)]}\big(\nu(x)+1\big)d\alpha +$$

$$+ \int_0^1 1_{G(x)\in[\epsilon(m-\eta+\alpha),\epsilon(m+\eta+\alpha)]}\big(\nu(x)+1\big)d\alpha.$$

Calculemos

$$\int_0^1 1_{G(x)\in[\epsilon(m-1-\eta+\alpha),\epsilon(m-1+\eta+\alpha)]}\big(\nu(x)+1\big)d\alpha.$$

Tememos que $G(x) \in \big[\epsilon(m-1-\eta+\alpha), \epsilon(m-1+\eta+\alpha)\big]$ si, y solo si $\alpha \in [t/\epsilon - m - \eta + 1, 1]$.

Por lo tanto

$$\int_0^1 1_{G(x)\in[\epsilon(m-1-\eta+\alpha),\epsilon(m-1+\eta+\alpha)]}\big(\nu(x)+1\big)d\alpha$$

$$= \int_{t/\epsilon-m-\eta+1}^1 1_{G(x)\in[\epsilon(m-1-\eta+\alpha),\epsilon(m-1+\eta+\alpha)]}\big(\nu(x)+1\big)d\alpha$$

$$= \big(1-(t/\epsilon-m-\eta+1)\big)\big(\nu(x)+1\big) = (m+\eta-t/\epsilon)\big(\nu(x)+1\big).$$

Similarmente, $G(x) \in \big[\epsilon(m-\eta+\alpha), \epsilon(m+\eta+\alpha)\big]$ si, y solo si $\alpha \in [0, t/\epsilon - m + \eta]$. Así

$$\int_0^1 1_{G(x)\in[\epsilon(m-\eta+\alpha),\epsilon(m+\eta+\alpha)]}\big(\nu(x)+1\big)d\alpha$$

$$= \int_0^{t/\epsilon-m+\eta} 1_{G(x)\in[\epsilon(m-\eta+\alpha),\epsilon(m+\eta+\alpha)]}\big(\nu(x)+1\big)d\alpha$$

$$= \big((t/\epsilon-m+\eta)-0\big)\big(\nu(x)+1\big) = (t/\epsilon-m+\eta)\big(\nu(x)+1\big).$$

Por lo tanto

$$\sum_{n\in\mathbb{Z}} \int_0^1 1_{G(x)\in[\epsilon(n-\eta+\alpha),\epsilon(n+\eta+\alpha)]}\big(\nu(x)+1\big)d\alpha = \big((m+\eta-t/\epsilon)+(t/\epsilon-m+\eta)\big)\big(\nu(x)+1\big)$$

$$= 2\eta\big(\nu(x)+1\big).$$

Tomando promedios

$$\mathbb{E}\left(\sum_{n\in\mathbb{Z}} \int_0^1 1_{G(x)\in[\epsilon(n-\eta+\alpha),\epsilon(n+\eta+\alpha)}\big(\nu(x)+1\big)d\alpha \,\middle|\, x\in\mathbb{Z}_N\right)$$

$$= \mathbb{E}\Big(2\eta\big(\nu(x)+1\big)\,\Big|\, x\in\mathbb{Z}_N\Big)$$

$$= 2\eta\big(\mathbb{E}(\nu)+1\big)$$

$$= 2\eta\big(2+o(1)\big)$$

$$= O(\eta),$$

por ende

$$\int_0^1 \sum_{n\in\mathbb{Z}_N} \mathbb{E}\Big(1_{G(x)\in[\epsilon(n-\eta+\alpha),\epsilon(n+\eta+\alpha)]}\big(\nu(x)+1\big)\Big|x\in\mathbb{Z}_N\Big)d\alpha = O(\eta).$$

Entonces, existe $c>0$ tal que

$$\int_0^1 \sum_{n\in\mathbb{Z}_N} \mathbb{E}\Big(1_{G(x)\in[\epsilon(n-\eta+\alpha),\epsilon(n+\eta+\alpha)]}\big(\nu(x)+1\big)\Big|x\in\mathbb{Z}_N\Big)d\alpha \leq c\eta. \tag{6.49}$$

Si para todo $\alpha\in[0,1]$

$$\sum_{n\in\mathbb{Z}} \mathbb{E}\Big(1_{G(x)\in[\epsilon(n-\eta+\alpha),\epsilon(n+\eta+\alpha)]}\big(\nu(x)+1\big)\Big|x\in\mathbb{Z}_N\Big) \geq (c+1)\eta,$$

entonces

$$\int_0^1 \sum_{n\in\mathbb{Z}} \mathbb{E}\Big(1_{G(x)\in[\epsilon(n-\eta+\alpha),\epsilon(n+\eta+\alpha)]}\big(\nu(x)+1\big)\Big|x\in\mathbb{Z}_N\Big)d\alpha \geq (c+1)\eta,$$

lo que contradiría (6.49). Así, existe $\alpha\in[0,1]$ (fijemos tal $\alpha$) tal que

$$\sum_{n\in\mathbb{Z}} \mathbb{E}\Big(1_{G(x)\in[\epsilon(n-\eta+\alpha),\epsilon(n+\eta+\alpha)]}\big(\nu(x)+1\big)\Big|x\in\mathbb{Z}_N\Big) < (c+1)\eta,$$

entonces

$$\sum_{n\in\mathbb{Z}} \mathbb{E}\Big(1_{G(x)\in[\epsilon(n-\eta+\alpha),\epsilon(n+\eta+\alpha)]}\big(\nu(x)+1\big)\Big|x\in\mathbb{Z}_N\Big) = O(\eta). \tag{6.50}$$

Sea $\mathcal{B}_{\epsilon,\eta}(G)$ el $\sigma$-álgebra cuyos átomos son los conjuntos $G^{-1}\Big(\big[\epsilon(n+\alpha),\epsilon(n+1+\alpha)\big)\Big)$, donde $n\in\mathbb{Z}$. Esta $\sigma$-álgebra está bien definida puesto que los intervalos $\big[\epsilon(n+\alpha),\epsilon(n+1+\alpha)\big)$, $n\in\mathbb{Z}$, son disjuntos dos a dos y $\bigcup_{n\in\mathbb{Z}}\big[\epsilon(n+\alpha),\epsilon(n+1+\alpha)\big)=\mathbb{R}$.

Si $\mathcal{B}$ es una $\sigma$-álgebra cualquiera, entonces para cada uno de los átomos de $\mathcal{B}\vee\mathcal{B}_{\epsilon,\eta}(G)$, $G$ toma valores sobre un intervalo de diámetro $\epsilon$, de lo cual se sigue (6.47). Ahora verifiquemos la propiedad de aproximación por funciones continuas. Sea $A=G^{-1}\Big(\big[\epsilon(n+\alpha),\epsilon(n+1+\alpha)\big)\Big)$ un átomo de $\mathcal{B}_{\epsilon,\eta}(G)$. Como $G$ toma valores en $I$, podemos asumir que $n=O(1/\epsilon)$ ($A$ es vacío para $n\in\mathbb{Z}$ tal que $\epsilon|n|\geq 2^{2^k}$). Sea $\psi_\eta:\mathbb{R}\to[0,1]$ una función continua fijada tal que

$$\psi_\eta(x)=1,\ \text{si } x\in[\eta,1-\eta];\ \text{y } \psi_\eta(x)=0,\ \text{si } x\notin[-\eta,1+\eta],$$

y definamos $\Psi_A(x)=\psi_\eta(x/\epsilon-\eta-\alpha)$. Es claro que $\Psi_A$ varía sobre un subconjunto compacto $F_{\epsilon,\eta}$ de $C^0(I)$ (pues $n$ y $\alpha$ están acotadas). Además de (6.50) se tiene que

$$\mathbb{E}\Big(\big|\big(1_A(x)-\Psi_A(G(x))\big)\big|\big(\nu(x)+1\big)\Big|x\in\mathbb{Z}_N\Big) = \Big|\Big|\big(1_A-\Psi_A(G)\big)(\nu+1)\Big|\Big|_{L^1(\mathbb{Z}_N)} = O(\eta).$$

En efecto, $1_A(x) - \Psi_A\big(G(x)\big)$ toma valores en el intervalo $[0,1]$ y se anula para aquellos $x$ tales que $G(x) \in \big[\epsilon(n+\eta+\alpha), \epsilon(n+1-\eta+\alpha)\big]$, por lo tanto

$$\begin{aligned}\mathbb{E}\bigg(\big|\big(1_A(x) - \Psi_A(G(x))\big)\big|\big(\nu(x)+1\big)\bigg|x \in \mathbb{Z}_N\bigg) \\ &\leq \mathbb{E}\bigg(\sum_{m\in\mathbb{Z}} 1_{G(x)\in[\epsilon(m-\eta+\alpha),\epsilon(m+\eta+\alpha)}\big(\nu(x)+1\big)\bigg|x\in\mathbb{Z}_N\bigg) \\ &= \sum_{m\in\mathbb{Z}} \mathbb{E}\bigg(1_{G(x)\in[\epsilon(m-\eta+\alpha),\epsilon(m+\eta+\alpha)}\big(\nu(x)+1\big)\bigg|x\in\mathbb{Z}_N\bigg),\end{aligned}$$

lo que muestra (6.48). □

**Proposición 6.6.** *Sea $\nu$ una medida $k$-pseudoaleatoria. Consideremos $K \geq 1$ un entero fijado y $\mathcal{D}F_1, \ldots, \mathcal{D}F_K \in L^\infty(\mathbb{Z}_N)$ funciones anti-uniformes básicas según Gowers.*
*Sean $0 < \epsilon < 1$, $0 < \eta < 1/2$ y $\mathcal{B}_{\epsilon,\eta}(\mathcal{D}F_j)$, $j = 1, \ldots, K$ construidos como en la proposición 6.5. Sea $\mathcal{B} = \mathcal{B}_{\epsilon,\eta}(\mathcal{D}F_1)\vee\cdots\vee\mathcal{B}_{\epsilon,\eta}(\mathcal{D}F_K)$. Entonces si $\eta < \eta_0(\epsilon, K)$ es suficientemente pequeño y $N > N_0(\epsilon, K, \eta)$ es suficientemente grande, tenemos que*

$$||\mathcal{D}F_j - \mathbb{E}(\mathcal{D}F_j|\mathcal{B})||_{L^\infty(\mathcal{Z}_N)} \leq \epsilon \text{ para todo } 1 \leq j \leq K \tag{6.51}$$

*Además existe un conjunto $\Omega \in \mathcal{B}$ tal que*

$$\mathbb{E}\big((\nu+1)1_\Omega\big) = O_{K,\epsilon}(\eta^{1/2}) \tag{6.52}$$

*y tal que*

$$||(1-1_\Omega)\mathbb{E}(\nu-1|\mathcal{B})||_{L^\infty(\mathbb{Z}_N)} = O_{K,\epsilon}(\eta^{1/2}). \tag{6.53}$$

*Demostración.* La afirmación (6.51) se sigue de (6.47). Mostremos (6.52) y (6.53). Como cada uno de los $\mathcal{B}(\mathcal{D}F_j)$ es generado por $O(1/\epsilon)$ átomos, vemos que $\mathcal{B}$ es generado por $O_{K,\epsilon}(1)$ átomos. Un átomo $A$ de $\mathcal{B}$ será llamado *pequeño* si $\mathbb{E}\big((\nu+1)1_A\big) \leq \eta^{1/2}$. Sea $\Omega$ la unión de todos los átomos pequeños. Es claro que $\Omega \in \mathcal{B}$ verifica (6.52). Resta mostrar (6.53), para lo cual será suficiente mostrar que

$$\frac{\mathbb{E}\big((\nu-1)1_A\big)}{\mathbb{E}(1_A)} = \mathbb{E}(\nu-1|A) = o_{K,\epsilon,\eta}(1) + O_{K,\epsilon}(\eta^{1/2}) \tag{6.54}$$

para todos los átomos $A \in \mathcal{B}$ no pequeños. Pero por definición de átomo pequeño tenemos para $A$ no pequeño,

$$\mathbb{E}\big((\nu-1)1_A\big) + 2\mathbb{E}(1_A) = \mathbb{E}\big((\nu+1)1_A\big) > \eta^{1/2}.$$

Luego, para probar (6.54) resta mostrar que

$$\mathbb{E}\big((\nu-1)1_A\big) = o_{K,\epsilon,\eta}(1) + O_{K,\epsilon}(\eta). \tag{6.55}$$

En efecto, de (6.55) existe $L_{K,\epsilon} > 0$ tal que

$$\big|\mathbb{E}\big((\nu-1)1_A\big)\big| \leq L_{K,\epsilon}\eta \tag{6.56}$$

para todo átomo $A$ no pequeño. Como $A$ es no pequeño

$$\mathbb{E}\big((\nu-1)1_A\big) + 2\mathbb{E}(1_A) \geq \eta^{1/2}. \tag{6.57}$$

Como $0 < \eta < 1$, $0 < \eta < \eta^{1/2}$; así, para $\eta$ suficientemente pequeño $\eta^{1/2} - L_{K,\epsilon}\eta > \eta^{1/2}/2$. Entonces, de (6.56) y (6.57) tenemos

$$\begin{aligned} 2\mathbb{E}(1_A) &\geq \eta^{1/2} - \mathbb{E}\big((\nu-1)1_A\big) \\ &\geq \eta^{1/2} - L_{K,\epsilon}\eta \\ &\geq \eta^{1/2}/2. \end{aligned}$$

Así, de (6.55)

$$\begin{aligned} \frac{\mathbb{E}\big((\nu-1)1_A\big)}{\mathbb{E}(1_A)} &= \frac{4}{\eta^{1/2}}\Big(o_{K,\epsilon,\eta}(1) + O_{K,\epsilon}(\eta)\Big) \\ &= o_{K,\epsilon,\eta}(1) + O_{K,\epsilon}(\eta^{1/2}). \end{aligned}$$

Por lo tanto (6.54) se sigue de (6.55). Por otro lado, como $A$ es la intersección de $K$ átomos $A_1, \ldots, A_K$ de $\mathcal{B}_{\epsilon,\eta}(\mathcal{D}F_1), \ldots, \mathcal{B}_{\epsilon,\eta}(\mathcal{D}F_K)$, respectivamente, vemos de (6.48) que existe una función continua $\Psi_A : I^K \to [0,1]$ tal que

$$\Big|\Big|(\nu+1)\big(1_A - \Psi_A(\mathcal{D}F_1, \ldots, \mathcal{D}F_K)\big)\Big|\Big|_{L^1(\mathbb{Z}_N)} = O(\eta).$$

En efecto, $A = \bigcap_{i=1}^{K} A_i$, $A_i \in \mathcal{B}_{\epsilon,\eta}(\mathcal{D}F_i)$. De la proposición 6.5 existen funciones continuas $\Psi_{A_i} : I \to [0,1]$ tales que

$$\Big|\Big|\big(1_{A_i} - \Psi_{A_i}(\mathcal{D}F_i)\big)(\nu+1)\Big|\Big|_{L^1(\mathbb{Z}_N)} = O(\eta).$$

Para $i = 1, 2$, tenemos

$$\Big|\Big|\big(1_{A_1}1_{A_2} - 1_{A_2}\Psi_{A_1}(\mathcal{D}F_1)\big)(\nu+1)\Big|\Big|_{L^1(\mathbb{Z}_N)} = O(\eta)$$

y

$$\Big|\Big|\big(1_{A_1}1_{A_2} - 1_{A_1}\Psi_{A_2}(\mathcal{D}F_2)\big)(\nu+1)\Big|\Big|_{L^1(\mathbb{Z}_N)} = O(\eta).$$

Así, por desigualdad triangular,

$$\Big|\Big|\Big(1_{A_1\cap A_2} - \big(1_{A_1}\Psi_{A_2}(\mathcal{D}F_2) + 1_{A_2}\Psi_{A_1}(\mathcal{D}F_1)\big)\Big)(\nu+1)\Big|\Big|_{L^1(\mathbb{Z}_N)} = 2O(\eta),$$

continuando de este modo, obtenemos $\Psi_A$ tal que

$$\Big|\Big|\big(1_A - \Psi_A(\mathcal{D}F_1,\dots,\mathcal{D}F_K)\big)(\nu+1)\Big|\Big|_{L^1(\mathcal{Z}_N)} = O_K(\eta).$$

En particular, como $|\nu - 1| \leq \nu + 1$,

$$\Big|\Big|(\nu-1)\big(1_A - \Psi_A(\mathcal{D}F_1,\dots,\mathcal{D}F_K)\big)\Big|\Big|_{L^1(\mathcal{Z}_N)} = O_K(\eta).$$

De la proposición 6.4, tenemos que

$$\mathbb{E}\big((\nu-1)\Psi_A(\mathcal{D}F_1,\dots,\mathcal{D}F_K)\big) = o_{K,\epsilon,\eta}(1).$$

Entonces, por la desigualdad triangular

$$\begin{aligned}\big|\mathbb{E}\big((\nu-1)1_A\big)\big| &\leq \Big|\mathbb{E}\Big((\nu-1)\big(1_A - \Psi_A(\mathcal{D}F_1,\dots,\mathcal{D}F_K)\big)\Big)\Big| + \Big|\mathbb{E}\big((\nu-1)\Psi_A(\mathcal{D}F_1,\dots,\mathcal{D}F_K)\big)\Big| \\ &\leq O_{K,\epsilon}(\eta) + o_{K,\epsilon,\eta}(1).\end{aligned}$$

□

## 6.6. Prueba del teorema de Szemerédi relativo a medidas pseudoaleatorias

Ahora tenemos toda la maquinaria necesaria para demostrar el teorema de Szemerédi relativo a medidas pseudoaleatorias, teorema 6.3, a partir de la proposición 6.2. La proposición clave es la siguiente descomposición, la cual separa una función arbitraria en una suma de una componente uniforme según Gowers, y otra componente anti-uniforme según Gowers, además de un error despreciable.

**Proposición 6.7** (Teorema de estructura generalizado de Koopman-von Neumann)**.** *Sea $\nu$ una medida $k$-pseudoaleatoria, y sea $f \in L^1(\mathbb{Z}_N)$ una función no negativa satisfaciendo $0 \leq f(x) \leq \nu(x)$ para todo $x \in \mathbb{Z}_N$. Sea $0 < \epsilon \ll 1$ un parámetro pequeño, y sea $N > N_0(\epsilon)$ suficientemente grande. Entonces existen una $\sigma$-álgebra $\mathcal{B}$ y un conjunto excepcional $\Omega \in \mathcal{B}$ tal que*

- *(condición de pequeñez)*

$$\mathbb{E}(\nu 1_\Omega) = o_\epsilon(1) \tag{6.58}$$

- *($\nu$ es uniformemente distribuido fuera de $\Omega$)*

$$\big|\big|(1-1_\Omega)\mathbb{E}(\nu - 1|\mathcal{B})\big|\big|_{L^\infty} = o_\epsilon(1) \tag{6.59}$$

*y*

- *(estimación de uniformidad de Gowers)*

$$\big|\big|(1-1_\Omega)\big(f - \mathbb{E}(f|\mathcal{B})\big)\big|\big|_{U^{k-1}} \leq \epsilon^{1/2^k}. \tag{6.60}$$

*Prueba del teorema 6.3 asumiendo la proposición 6.7.* Sean $f, \delta$ como en el teorema 6.3, y sea $0 < \epsilon \ll \delta$ un parámetro que será elegido convenientemente después. Sea $\mathcal{B}$ como en la anterior descomposición, y definamos $f_U = (1-1_\Omega)\big(f - \mathbb{E}(f|\mathcal{B})\big)$ y $f_{U^\perp} = (1-1_\Omega)\mathbb{E}(f|\mathcal{B})$ (el subíndice $U$ indica la componente uniforme según Gowers, y $U^\perp$ la componente anti-uniforme según Gowers). De las ecuaciones (6.10), (6.11) y (6.58) y del hecho que $\Omega$ es medible tenemos que

$$\begin{aligned}
\mathbb{E}\big(f_{U^\perp}\big) &= \mathbb{E}\big((1-1_\Omega)\mathbb{E}(f|\mathcal{B})\big) = \mathbb{E}\big(\mathbb{E}(f|\mathcal{B})\big) - \mathbb{E}\big(1_\Omega\mathbb{E}(f|\mathcal{B})\big) \\
&= \mathbb{E}\big(\mathbb{E}(f|\mathcal{B})\big) - \mathbb{E}\big(1_\Omega\mathbb{E}(f|\mathcal{B})\big) = \mathbb{E}(f) - \mathbb{E}\big(1_\Omega\mathbb{E}(f|\mathcal{B})\big) \\
&\geq \mathbb{E}(f) - \mathbb{E}(\nu 1_\Omega) \geq \delta - o_\epsilon(1).
\end{aligned} \tag{6.61}$$

También por la ecuación (6.59) vemos que $f_{U^\perp}$ es limitada superiormente por $1 + o_\epsilon(1)$,

$$\begin{aligned}
||f_{U^\perp}||_{L^\infty} &= \big|\big|(1-1_\Omega)\mathbb{E}(f|\mathcal{B})\big|\big|_{L^\infty} \\
&\leq \big|\big|(1-1_\Omega)\big(\mathbb{E}(f-1|\mathcal{B})\big)\big|\big|_{L^\infty} + \big|\big|(1-1_\Omega)\mathbb{E}(1|\mathcal{B})\big|\big|_{L^\infty} \\
&\leq \big|\big|(1-1_\Omega)\big(\mathbb{E}(\nu-1|\mathcal{B})\big)\big|\big|_{L^\infty} + \big|\big|(1-1_\Omega)\mathbb{E}(1|\mathcal{B})\big|\big|_{L^\infty} \\
&\leq o_\epsilon(1) + 1.
\end{aligned} \tag{6.62}$$

Como $f$ es no negativa, así también lo es $f_{U^\perp}$, y por lo tanto

$$0 \leq f_{U^\perp}(x) \leq 1 + o_\epsilon(1) \text{ para todo } x \in \mathbb{Z}_N. \tag{6.63}$$

De la proposición 6.2 y de las ecuaciones (6.61) y (6.63)

$$\mathbb{E}\Big(f_{U^\perp}(x)f_{U^\perp}(x+r)\cdots f_{U^\perp}\big(x+(k-1)r\big)\Big|x, r \in \mathbb{Z}_N\Big) \geq c(k,\delta) - o_\epsilon(1) - o_{k,\delta}(1). \tag{6.64}$$

Por otro lado, de (6.60) tenemos que $||f_U||_{U^{k-1}} \leq \epsilon^{1/2^k}$; como $(1-1_\Omega)f$ está acotada por $\nu$ (puesto que $0 \leq |f| \leq \nu$) y $|f_{U^\perp}|$ está acotada por $1 + o_\epsilon(1)$ (de la ecuación (6.62)), vemos que

$$\begin{aligned} |f_U(x)| &= \big|\big((1-1_\Omega)f - f_{U^\perp}\big)(x)\big| \\ &\leq \big|\big((1-1_\Omega)f\big)(x)\big| + \big|f_{U^\perp}(x)\big| \\ &\leq \nu(x) + 1 + o_\epsilon(1) \end{aligned} \tag{6.65}$$

para todo $x \in \mathbb{Z}_N$. Aplicando el teorema de Von Neumann generalizado (proposición 6.3) obtenemos

$$\mathbb{E}\Big(f_0(x)f_1(x+r)\cdots f_{k-1}\big(x+(k-1)r\big)\Big|x, r \in \mathbb{Z}_N\Big) = O(\epsilon^{1/2^k}) + o_\epsilon(1),$$

donde cada $f_j$ es igual a $f_U$ o $f_{U^\perp}$, con al menos un $f_j$ igual a $f_U$. Así, en particular existe una constante $t$ (independiente de $\epsilon$ y $N$) tal que

$$\mathbb{E}\Big(f_0(x)f_1(x+r)\cdots f_{k-1}\big(x+(k-1)r\big)\Big|x, r \in \mathbb{Z}_N\Big) \geq -t\epsilon^{1/2^k} - o_\epsilon(1). \tag{6.66}$$

Sumando la ecuación (6.64) con $2^k - 1$ ecuaciones del tipo (6.66) obtenemos

$$\begin{aligned} &\mathbb{E}\Big(\widehat{f}(x)\widehat{f}(x+r)\cdots\widehat{f}\big(x+(k-1)r\big)\Big|x, r \in \mathbb{Z}_N\Big) \\ &\qquad \geq c(k,\delta) - o_\epsilon(1) - o_{k,\delta}(1) - (2^k-1)\big(t\epsilon^{1/2^k} + o_\epsilon(1)\big), \end{aligned} \tag{6.67}$$

esto es

$$\mathbb{E}\Big(\widehat{f}(x)\widehat{f}(x+r)\cdots\widehat{f}\big(x+(k-1)r\big)\Big|x, r \in \mathbb{Z}_N\Big) \geq c(k,\delta) - O(\epsilon^{1/2^k}) - o_\epsilon(1) - o_{k,\delta}(1),$$

donde $\widehat{f} = f_U + f_{U^\perp} = (1-1_\Omega)f$. Pero, puesto que $0 \leq (1-1_\Omega)f \leq f$ tenemos que

$$\mathbb{E}\Big(f(x)f(x+r)\cdots f\big(x+(k-1)r\big)\Big|x, r \in \mathbb{Z}_N\Big) \geq c(k,\delta) - O(\epsilon^{1/2^k}) - o_\epsilon(1) - o_{k,\delta}(1).$$

Dado que $\epsilon$ puede ser tomado arbitrariamente pequeño (siempre y cuando $N$ sea tomado suficientemente grande), el término de error del lado derecho de la desigualdad anterior puede ser tomado arbitrariamente pequeño, eligiendo $N$ suficientemente grande, dependiendo de $k$ y $\delta$. □

Para obtener la prueba completa del teorema 6.3, es suficiente probar la proposición 6.7. Para construir la $\sigma$-álgebra $\mathcal{B}$ dada en la proposición, la idea es a grandes rasgos como sigue.

Iniciamos con $\mathcal{B}$ siendo la $\sigma$-álgebra trivial $\mathcal{B} = \{\emptyset, \mathbb{Z}_N\}$. Si la función $f - \mathbb{E}(f|\mathcal{B})$ es uniforme según Gowers (en el sentido de (6.60)), entonces terminamos el algoritmo. En otro caso, usamos la maquinaria de la funciones duales, para localizar una función anti-uniforme según Gowers $\mathcal{D}F_1$ que tenga correlación no trivial con $f$, y añadimos los conjuntos de nivel de $\mathcal{D}F_1$ a la $\sigma$-álgebra $\mathcal{B}$. La correlación no trivial garantizará que la $L^2$ norma de $\mathbb{E}(f|\mathcal{B})$ aumente en una cantidad no trivial durante este procedimiento, mientras que la pseudoaleatoriedad asegurará que $\mathbb{E}(f|\mathcal{B})$ permanezca uniformemente acotada. Repetimos el anterior algoritmo hasta que $f - \mathbb{E}(f|\mathcal{B})$ resulte una función suficientemente uniforme según Gowers, es decir con norma de Gowers suficientemente pequeña.

**Proposición 6.8** (Paso iterativo)**.** *Sean $\nu$ una medida $k$-pseudoaleatoria y $f \in L^1(\mathbb{Z}_N)$ una función no negativa tal que $0 \leq f(x) \leq \nu(x)$ para todo $x \in \mathbb{Z}_N$. Sean $0 < \eta \ll \epsilon \ll 1$ números pequeños, y $K \geq 0$ un entero. Supongamos que $\eta \ll \eta_0(\epsilon, K)$ sea suficientemente pequeño y que $N > N_0(\epsilon, K, \eta)$ sea suficientemente grande. Sea $F_1, \ldots, F_K \in L^1(\mathbb{Z}_N)$ una colección de funciones tales que*

$$\big|F_j(x)\big| \leq \big(1 + O_{K,\epsilon}(\eta^{1/2})\big)\big(\nu(x) + 1\big) \tag{6.68}$$

*para todo $1 \leq j \leq K$ y todo $x \in \mathbb{Z}_N$. Sea $\mathcal{B}_K$ una $\sigma$-álgebra*

$$\mathcal{B}_K = \mathcal{B}_{\epsilon,\eta}(\mathcal{D}F_1) \vee \cdots \vee \mathcal{B}_{\epsilon,\eta}(\mathcal{D}F_1) \tag{6.69}$$

*donde $\mathcal{B}_{\epsilon,\eta}(\mathcal{D}F_j)$ es como en la proposición 6.5, y supongamos que existe un conjunto $\Omega_K$ en $\mathbb{Z}_K$ tal que*

*(i)*

$$\mathbb{E}\big((\nu + 1)1_{\Omega_K}\big) = O_{K,\epsilon}\big(\eta^{1/2}\big), \ \ y \tag{6.70}$$

*(ii)*

$$\big|\big|(1 - 1_{\Omega_K})\mathbb{E}(\nu - 1|\mathcal{B}_K)\big|\big|_{L^\infty(\mathbb{Z}_N)} = O_{K,\epsilon}\big(\eta^{1/2}\big). \tag{6.71}$$

*Sea*

$$F_{k+1} = (1 - 1_{\Omega_K})(f - \mathbb{E}(f|\mathcal{B}_K)) \tag{6.72}$$

*y supongamos que $F_{k+1}$ verifica la estimación de no uniformidad de Gowers*

$$||F_{K+1}||_{U^{k-1}} > \epsilon^{1/2^k}. \tag{6.73}$$

*Entonces tenemos que*

$$\big|\big|(1-1_{\Omega_K})\mathbb{E}(f|\mathcal{B}_K)\big|\big|_{L^\infty(\mathbb{Z}_N)} \leq 1 + O_{K,\epsilon}\big(\eta^{1/2}\big) \tag{6.74}$$

*y*

$$|F_{K+1}(x)| \leq \big(1 + O_{K,\epsilon}(\eta^{1/2})\big)\big(\nu(x)+1\big). \tag{6.75}$$

*Además, si* $\mathcal{B}_{K+1}$ *es la* $\sigma$*-álgebra*

$$\mathcal{B}_{K+1} = \mathcal{B}_K \vee \mathcal{B}_{\epsilon,\eta}(\mathcal{D}F_{K+1}) = \mathcal{B}_{\epsilon,\eta}(\mathcal{D}F_1) \vee \cdots \vee \mathcal{B}_{\epsilon,\eta}(\mathcal{D}F_{K+1}) \tag{6.76}$$

*entonces existe un conjunto* $\Omega_{K+1} \supseteq \Omega_K$ *talque*

*(i)*

$$\mathbb{E}\big((\nu+1)1_{\Omega_{K+1}}\big) = O_{K,\epsilon}\big(\eta^{1/2}\big), \tag{6.77}$$

*(ii)*

$$\big|\big|(1-1_{\Omega_{K+1}})\mathbb{E}(\nu-1|\mathcal{B}_{K+1})\big|\big|^2_{L^\infty(\mathbb{Z}_N)} = O_{K,\epsilon}\big(\eta^{1/2}\big), \tag{6.78}$$

*(iii) (Propiedad del incremento de energía)*

$$\big|\big|(1-1_{\Omega_{K+1}})\mathbb{E}(f|\mathcal{B}_{K+1})\big|\big|^2_{L^2(\mathbb{Z}_N)} \geq \big|\big|(1-1_{\Omega_K})\mathbb{E}(f|\mathcal{B}_K)\big|\big|^2_{L^2(\mathbb{Z}_N)} + 2^{-2^k+1}\epsilon. \tag{6.79}$$

Suponiendo verdadera la proposición 6.8 probemos la proposición 6.7.

*Demostración de la proposición 6.7.* Fijemos $\epsilon > 0$ y sea $K_0$ el mínimo entero mayor que $2^{2^k}/\epsilon + 1$, $K_0$ será una cota superior para el número de iteraciones del algoritmo que describiremos en breve. Requeriremos de un parámetro $\eta$, $0 < \eta \ll \epsilon$, que será escogido posteriormente (tomamos $\eta < \eta_0(\epsilon, K_0)$, y entonces tomaremos $N > N_0(\eta, \epsilon)$ suficientemente grande).

Para construir $\mathcal{B}$ y $\Omega$, para cada $K \in [0, K_0]$, hallaremos una secuencia de funciones anti-uniformes básicas según Gowers, $\mathcal{D}F_1, \ldots, \mathcal{D}F_K$ en $\mathbb{Z}_N$, además de *conjuntos excepcionales*, $\Omega_0 \subseteq \Omega_1 \subseteq \cdots \subseteq \Omega_K \subseteq \mathbb{Z}_N$, a través del siguiente algoritmo.

**Algoritmo 6.1.**

*paso 0 Iniciamos con* $K = 0$ *y* $\Omega_0 = \emptyset$ *(luego incrementaremos el valor de* $K$*).*

*paso 1 Sea* $\mathcal{B}_K$ *y* $F_{K+1}$ *definidos como en (6.69) y (6.72) respectivamente. Para* $K = 0$ *tendremos* $\mathcal{B}_0 = \{\emptyset, \mathbb{Z}_N\}$ *y* $F_1 = f - \mathbb{E}(f)$*. Para* $K = 0$ *tenemos que (6.68), (6.70) y (6.71) se verifican directamente puesto que* $\mathbb{E}(\nu - 1) = o(1)$*. Como mostraremos depués, estas tres estimaciones seguirán siendo válidas, al pasar de* $K$ *a* $K+1$*, a lo largo del algoritmo.*

*paso 2 Si la estimación (6.73) falla, esto es, si*

$$||F_{K+1}||_{U^{k-1}} \leq \epsilon^{1/2^k},$$

*entonces elegimos* $\Omega = \Omega_K$ *y* $\mathcal{B} = \mathcal{B}_K$*, y el algoritmo acaba.*

*paso 3 Si (6.73) es verdadero, definimos* $\mathcal{B}_{K+1}$ *como en (6.76) (aquí requerimos que* $K \leq K_0$*, esto será garantizado en el paso 4). Por la proposición 6.8 existe un conjunto excepcional* $\Omega_{K+1} \supseteq \Omega_K$ *en* $\mathcal{B}_{K+1}$*, tal que se verifican las condiciones (6.77), (6.78), y la propiedad del incremento de energía (6.79).*

*paso 4 Incrementamos* $K$ *a* $K+1$*; observamos de la construcción que las desigualdades (6.68), (6.70) y (6.71) serán aún válidas al hacer este incremento de* $K$ *a* $K+1$*. (6.68) se sigue de (6.75), (6.70) de (6.77), y (6.71) de (6.78). Si* $K > K_0$ *entonces terminamos el algoritmo con un error; en otro caso retornamos al paso 1.*

**Observación 6.4.** *El entero* $K$ *indicará el número de iteraciones del algoritmo. A pesar de que* $O_{K,\epsilon}(\eta^{1/2})$ *depende de* $K$*, como veremos a continuación, este algoritmo finalizará antes que* $K$ *alcance el valor* $K_0$*.*

Asumiendo la proposición 6.8, veremos que este algoritmo finalizará en un número finito de pasos con dos resultados: termina satisfactoriamente en el *paso 2* para algún $K \leq K_0$, o termina con un error en el *paso 4* cuando $K$ excede $K_0$. Supongamos por el momento que el algoritmo culmina satisfactoriamente en la $K$-ésima iteración. Entonces es claro que el algoritmo en este caso producirá una $\sigma$-álgebra $\mathcal{B}$ y un conjunto *excepcional* $\Omega$ verificando las propiedades requeridas por la proposición 6.7 con un término de error $O_{K,\epsilon}(\eta^{1/2})$ en vez de $o_\epsilon(1)$. El estimado (6.60) se sigue del hecho que $F_{K+1} = (1 - 1_{\Omega_K})(f - \mathbb{E}(f|\mathcal{B}_K))$ no verifica (6.73). Los estimados (6.58) y (6.59) se siguen de (6.70) y (6.71) respectivamente. Haciendo $\eta$ suficientemente pequeño, podemos reemplazar $O_{K,\epsilon}(\eta^{1/2})$ por $o_{K,\epsilon}(1)$. La dependencia del error respecto a $K$ no es relevante puesto que $K$ está acotado por $K_0$ y $K_0$ sólo depende de $\epsilon$.

Para concluir la prueba de la proposición resta mostrar que el algoritmo no culmina en error. Supongamos por contradicción que el algoritmo alcanza la $K_0$-ésima iteración antes de finalizar con error en el paso 4. Entonces, si definimos la *energía* para $0 \leq K \leq K_0 + 1$ por

$$E_K = \big|\big|(1 - 1_{\Omega_K})\mathbb{E}(f|\mathcal{B}_K)\big|\big|^2_{L^2(\mathbb{Z}_N)},$$

entonces vemos de (6.79) que

$$E_{K+1} \geq E_K + 2^{-2^k+1}\epsilon \text{ para cada } 0 \leq K \leq K_0. \tag{6.80}$$

Así,

$$E_{K_0} \geq K_0 2^{-2^k+1}\epsilon \geq \big(2^{2^k}/\epsilon + 1\big)2^{-2^k+1}\epsilon \geq 2 + 2^{-2^k+1}\epsilon.$$

También, por (6.74) tenemos que

$$0 \leq E_K \leq 1 + O_{K,\epsilon}\big(\eta^{1/2}\big) \text{ para todo } 0 \leq K \leq K_0;$$

en particular,

$$E_{K_0} \leq 1 + O_{K,\epsilon}\big(\eta^{1/2}\big).$$

Tomando $\eta < \eta_0(K, \epsilon)$ suficientemente pequeño llegamos a una contradicción. Por lo tanto, el algoritmo no alcanza la $K_0$-ésima iteración, y por el contrario termina satisfactoriamente en el *paso 2*. □

Sólo queda demostrar la proposición 6.8.

*Demostración de la proposición 6.8.* Sean $\nu, f, K, \epsilon, \eta, F_1, \ldots, F_K, F_{K+1}, \mathcal{B}_K, \Omega_K, \Omega_{K+1}$ como en el enunciado de la proposición. Comenzamos probando (6.74) y (6.75). De (6.71) tenemos

$$\big|\big|(1 - 1_{\Omega_K})\mathbb{E}(\nu|\mathcal{B}_K)\big|\big|_{L^\infty} \leq 1 + O_{K,\epsilon}(\eta^{1/2}),$$

en efecto

$$\begin{aligned}
\big|\big|(1 - 1_{\Omega_K})\mathbb{E}(\nu|\mathcal{B}_K)\big|\big|_{L^\infty} &\leq \big|\big|(1 - 1_{\Omega_K})\mathbb{E}(\nu - 1|\mathcal{B}_K)\big|\big|_{L^\infty} \\
&\leq \big|\big|(1 - 1_{\Omega_K})\mathbb{E}(1|\mathcal{B}_K)\big|\big|_{L^\infty} \\
&\leq O_{K,\epsilon}\big(\eta^{1/2}\big) + 1.
\end{aligned}$$

Como $f$ es no negativa y $|f| \leq \nu$, obtenemos (6.74). De (6.72) y (6.74) tenemos (6.75),

$$\begin{aligned}
|F_{K+1}(x)| &\leq \big|(1 - 1_{\Omega_K})(x)\big(f - \mathbb{E}(f|\mathcal{B}_K)\big)(x)\big| \\
&\leq \big|(1 - 1_{\Omega_K})(x)f(x)| + |(1 - 1_{\Omega_K})(x)\mathbb{E}(f|\mathcal{B}_K)(x)\big| \\
&\leq \big|(1 - 1_{\Omega_K})(x)\nu(x)\big| + \big(1 + O_{K,\epsilon}(\eta^{1/2})\big) \\
&\leq \nu(x)\big(1 + O_{K,\epsilon}(\eta^{1/2})\big) + \big(1 + O_{K,\epsilon}(\eta^{1/2})\big) \\
&= \big(1 + O_{K,\epsilon}(\eta^{1/2})\big)\big(\nu(x) + 1\big).
\end{aligned}$$

De este modo, en particular mostramos que $\mathcal{D}F_1, \ldots, \mathcal{D}F_{K+1}$ son funciones anti-uniformes básicas según Gowers.

Aplicando el lema 6.5 y usando (6.68) y (6.75) tenemos que

$$||\mathcal{D}F_j||_{L^\infty(\mathbb{Z}_N)} \leq 2^{2^{k-1}-1} + O_{K,\epsilon}(\eta^{1/2}) \text{ para todo } 1 \leq j \leq K+1, \tag{6.81}$$

toda vez que $N$ sea suficientemente grande (dependiendo de $K, \epsilon, \eta$). Aplicando la proposición 6.6 concluimos que existe $\Omega \in \mathcal{B}_{K+1}$ tal que

$$\mathbb{E}\big((\nu+1)1_\Omega\big) = O_{K,\epsilon}(\eta^{1/2}),$$

y

$$\big|\big|(1-1_\Omega)\mathbb{E}(\nu-1|\mathcal{B}_{K+1})\big|\big|_{L^\infty(\mathbb{Z}_N)} = O_{K,\epsilon}(\eta^{1/2}).$$

Sea $\Omega_{K+1} = \Omega_K \cup \Omega$, entonces podemos verificar (6.77) y (6.78) de (6.70) y (6.71), respectivamente, puesto que

$$1_\Omega \leq 1_{\Omega_K} + 1_\Omega \;\text{ y }\; 1 - 1_{\Omega_{K+1}} \leq (1-1_\Omega)/2 + (1-1_{\Omega_K})/2.$$

Resta verificar (6.79), la propiedad del incremento de energía que dice que

$$\big|\big|(1-1_{\Omega_{K+1}})\mathbb{E}(f|\mathcal{B}_{K+1})\big|\big|^2_{L^2(\mathbb{Z}_N)} \geq \big|\big|(1-1_{\Omega_K})\mathbb{E}(f|\mathcal{B}_K)\big|\big|_{L^2(\mathbb{Z}_N)} + 2^{-2^k+1}\epsilon. \tag{6.82}$$

Para este fin explotaremos la hipótesis (6.73). Por el lema 6.5 y la definición (6.72), tenemos que

$$\Big|\Big\langle (1-1_{\Omega_K})\big(f - \mathbb{E}(f|\mathcal{B}_K)\big), \mathcal{D}F_{K+1}\Big\rangle\Big| = \big|\langle F_{K+1}, \mathcal{D}F_{K+1}\rangle\big| = ||F_{K+1}||_{U^{k-1}} \geq \epsilon^{1/2}. \tag{6.83}$$

Por otro lado, de las expresiones (6.68), (6.70) y (6.81) tenemos que

$$\begin{aligned}
\Big|\Big\langle (1_{\Omega_{K+1}} - 1_{\Omega_K})\big(f - \mathbb{E}(f|\mathcal{B}_K)\big), \mathcal{D}F_{K+1}\Big\rangle\Big| &\leq ||\mathcal{D}F_{K+1}||_{L^\infty}\mathbb{E}\Big((1_{\Omega_{K+1}} - 1_{\Omega_K})\big|f - \mathbb{E}(f|\mathcal{B}_K)\big|\Big) \\
&\leq O_{K,\epsilon}(1)\mathbb{E}\big((1_{\Omega_{K+1}} - 1_{\Omega_K})(\nu+1)\big) \\
&\leq O_{K,\epsilon}(1)\mathbb{E}\big(1_{\Omega_{K+1}}(\nu+1)\big) \\
&\leq O_{K,\epsilon}(1)O_{K,\epsilon}\big(\eta^{1/2}\big) = O_{K,\epsilon}\big(\eta^{1/2}\big),
\end{aligned} \tag{6.84}$$

mientras que de (6.47) y (6.74) tenemos

$$\begin{aligned}
&\Big|\Big\langle (1-1_{\Omega_{K+1}})\big(f - \mathbb{E}(f|\mathcal{B}_K)\big), \mathcal{D}F_{K+1} - \mathbb{E}(\mathcal{D}F_{K+1}|\mathcal{B}_{K+1})\Big\rangle\Big| \\
&\qquad \leq \big|\big|\mathcal{D}F_{K+1} - \mathbb{E}(\mathcal{D}F_{K+1}|\mathcal{B}_{K+1})\big|\big|_{L^\infty}\mathbb{E}\Big((1-1_{\Omega_{K+1}})\big|f - \mathbb{E}(f|\mathcal{B}_K)\big|\Big) \\
&\qquad \leq O(\epsilon)\mathbb{E}\big((1-1_{\Omega_{K+1}})(\nu+1)\big) = O(\epsilon).
\end{aligned} \tag{6.85}$$

Por la desigualdad triangular, de (6.83), (6.84) y (6.85) tenemos

$$\begin{aligned}\Big|\Big\langle(1-1_{\Omega_{K+1}})\big(f-\mathbb{E}(f|\mathcal{B}_K)\big),\mathbb{E}(\mathcal{D}F_{K+1}|\mathcal{B}_{K+1})\Big\rangle\Big| \\ &\geq \Big|\Big\langle(1-1_{\Omega_{K+1}})\big(f-\mathbb{E}(f|\mathcal{B}_K)\big),\mathcal{D}F_{K+1}\Big\rangle\Big|- \\ &\quad-\Big|\Big\langle(1-1_{\Omega_{K+1}})\big(f-\mathbb{E}(f|\mathcal{B}_K)\big),\mathcal{D}F_{K+1}-\mathbb{E}(\mathcal{D}F_{K+1}|\mathcal{B}_{K+1})\Big\rangle\Big| \\ &\geq \Big|\Big\langle(1-1_{\Omega_K})\big(f-\mathbb{E}(f|\mathcal{D}F_K)\big),\mathcal{D}F_{K+1}\Big\rangle\Big|- \\ &\quad-\Big|\Big\langle(1_{\Omega_{K+1}}-1_{\Omega_K})\big(f-\mathbb{E}(f|\mathcal{D}F_K)\big),\mathcal{D}F_{K+1}\Big\rangle\Big|- \\ &\quad-\Big|\Big\langle(1-1_{\Omega_{K+1}})\big(f-\mathbb{E}(f|\mathcal{D}F_K)\big),\mathcal{D}F_{K+1}-\mathbb{E}(\mathcal{D}F_{K+1}|\mathcal{B}_{K+1})\Big\rangle\Big| \\ &\geq \epsilon^{1/2}-O_{K,\epsilon}\big(\eta^{1/2}\big)-O(\epsilon).\end{aligned}$$

Pero como $1-1_{\Omega_{K+1}}$, $\mathbb{E}(\mathcal{D}F_{K+1}|\mathcal{B}_{K+1})$ y $\mathbb{E}(f|\mathcal{B}_K)$ son medibles en $\mathcal{B}_{K+1}$, podemos reemplazar $f$ por $\mathbb{E}(f|\mathcal{B}_{K+1})$, y así

$$\Big|\Big\langle(1-1_{\Omega_{K+1}})\big(\mathbb{E}(f|\mathcal{B}_{K+1})-\mathbb{E}(f|\mathcal{B}_K)\big),\mathbb{E}(\mathcal{D}F_{K+1}|\mathcal{B}_{K+1})\Big\rangle\Big|\geq\epsilon^{1/2}-O_{K,\epsilon}\big(\eta^{1/2}\big)-O(\epsilon).$$

Así, por la desigualdad de Cauchy-Schwarz

$$\begin{aligned}\big|\big|(1-1_{\Omega_{K+1}})\big(\mathbb{E}(f|\mathcal{B}_{K+1})-\mathbb{E}(f|\mathcal{B}_K)\big)\big|\big|_{L^2(\mathbb{Z}_N)}\big|\big|\mathbb{E}(\mathcal{D}F_{K+1}|\mathcal{B}_{K+1})\big|\big|_{L^\infty(\mathbb{Z}_N)} \\ \geq\epsilon^{1/2}-O_{K,\epsilon}\big(\eta^{1/2}\big)-O(\epsilon),\end{aligned}$$

y por lo tanto, de (6.81) tenemos que

$$\big|\big|(1-1_{\Omega_{K+1}})\big(\mathbb{E}(f|\mathcal{B}_{K+1})-\mathbb{E}(f|\mathcal{B}_K)\big)\big|\big|_{L^2(\mathbb{Z}_N)}\geq 2^{-2^{k-1}+1}\epsilon^{1/2}-O_{K,\epsilon}\big(\eta^{1/2}\big)-O(\epsilon). \tag{6.86}$$

Recordando de (6.74) que $\mathbb{E}(f|\mathcal{B}_K)\leq 1+O_{K,\epsilon}\big(\eta^{1/2}\big)$ fuera de $\Omega_K$, observamos que si $\eta<\eta_0(\epsilon,K)$ es suficientemente pequeño, entonces

$$\big|\big|(1_{\Omega_{K+1}}-1_{\Omega_K})\mathbb{E}(f|\mathcal{B}_K)\big|\big|_{L^2}^2\leq 2||1_{\Omega_{K+1}}-1_{\Omega_K}||_{L^2}^2\leq 2||1_{\Omega_{K+1}}-1_{\Omega_K}||_{L^1}\leq 2\mathbb{E}(1_{\Omega_{K+1}}), \tag{6.87}$$

donde, por (6.70), $\mathbb{E}(1_{\Omega_{K+1}})=O_{K,\epsilon}\big(\eta^{1/2}\big)$. Por la desigualdad triangular y (6.74) vemos que para probar (6.79) es suficiente probar que

$$\begin{aligned}\big|\big|(1-1_{\Omega_{K+1}})\mathbb{E}(f|\mathcal{B}_{K+1})\big|\big|_{L^2}^2\geq\big|\big|(1-1_{\Omega_{K+1}})\mathbb{E}(f|\mathcal{B}_K)\big|\big|_{L^2}^2+ \\ 2^{-2^k+2}\epsilon-O_{K,\epsilon}\big(\eta^{1/2}\big)-O(\epsilon^{3/2}),\end{aligned} \tag{6.88}$$

puesto que podemos absorber el término de error $-O_{K,\epsilon}(\eta^{1/2})-O(\epsilon^{3/2})$ en el término $2^{-2^k+2}\epsilon$ mediante la elección de $\epsilon$ suficientemente pequeño, dependiendo de $k$, y $\eta$ suficientemente pequeño dependiendo de $K$ y $\epsilon$. En efecto, desarrollando el primer término del lado derecho de la desigualdad anterior tenemos

$$\begin{aligned}
&\big|\big|(1-1_{\Omega_{K+1}})\mathbb{E}(f|\mathcal{B}_K)\big|\big|^2_{L^2(\mathbb{Z}_{\mathrm{N}})}\\
&\quad=\big|\big|(1-1_{\Omega_K})\mathbb{E}(f|\mathcal{B}_K)-(1_{\Omega_{K+1}}-1_{\Omega_K})\mathbb{E}(f|\mathcal{B}_K)\big|\big|^2_{L^2(\mathbb{Z}_{\mathrm{N}})}\\
&\quad=\big|\big|(1-1_{\Omega_K})\mathbb{E}(f|\mathcal{B}_K)\big|\big|^2_{L^2(\mathbb{Z}_{\mathrm{N}})}+\big|\big|(1_{\Omega_{K+1}}-1_{\Omega_K})\mathbb{E}(f|\mathcal{B}_K)\big|\big|^2_{L^2(\mathbb{Z}_{\mathrm{N}})}\\
&\qquad-2\Big\langle(1-1_{\Omega_K})\mathbb{E}(f|\mathcal{B}_K),(1_{\Omega_{K+1}}-1_{\Omega_K})\mathbb{E}(f|\mathcal{B}_K)\Big\rangle\\
&\quad=\big|\big|(1-1_{\Omega_K})\mathbb{E}(f|\mathcal{B}_K)\big|\big|^2_{L^2(\mathbb{Z}_{\mathrm{N}})}+\big|\big|(1_{\Omega_{K+1}}-1_{\Omega_K})\mathbb{E}(f|\mathcal{B}_K)\big|\big|^2_{L^2(\mathbb{Z}_{\mathrm{N}})}\\
&\qquad-2\Big\langle(1-1_{\Omega_{K+1}})\mathbb{E}(f|\mathcal{B}_K),(1_{\Omega_{K+1}}-1_{\Omega_K})\mathbb{E}(f|\mathcal{B}_K)\Big\rangle\\
&\qquad-2\Big\langle(1_{\Omega_{K+1}}-1_{\Omega_K})\mathbb{E}(f|\mathcal{B}_K),(1_{\Omega_{K+1}}-1_{\Omega_K})\mathbb{E}(f|\mathcal{B}_K)\Big\rangle\\
&\quad=\big|\big|(1-1_{\Omega_K})\mathbb{E}(f|\mathcal{B}_K)\big|\big|^2_{L^2(\mathbb{Z}_{\mathrm{N}})}-\big|\big|(1_{\Omega_{K+1}}-1_{\Omega_K})\mathbb{E}(f|\mathcal{B}_K)\big|\big|^2_{L^2(\mathbb{Z}_{\mathrm{N}})},
\end{aligned}\tag{6.89}$$

donde la última igualdad se da ya que $\Big\langle(1-1_{\Omega_{K+1}})\mathbb{E}(f|\mathcal{B}_K),(1_{\Omega_{K+1}}-1_{\Omega_K})\mathbb{E}(f|\mathcal{B}_K)\Big\rangle=0$. Así, de (6.87) y reemplazando (6.89) en (6.88) tenemos que

$$\begin{aligned}
\big|\big|(1-1_{\Omega_{K+1}})\mathbb{E}(f|\mathcal{B}_{K+1})\big|\big|^2_{L^2}\geq\big|\big|(1-1_{\Omega_K})\mathbb{E}(f|\mathcal{B}_{K+1})\big|\big|^2_{L^2}-O_{K,\epsilon}(\eta^{1/2})\\
+2^{-2^k+2}\epsilon-O_{K,\epsilon}\big(\eta^{1/2}\big)-O(\epsilon^{3/2}).
\end{aligned}$$

Por ende, es suficiente probar (6.88).

Podemos reescribir el término del lado izquierdo de la desigualdad (6.88) como

$$\begin{aligned}
&\Big|\Big|(1-1_{\Omega_{K+1}})\mathbb{E}(f|\mathcal{B}_K)+(1-1_{\Omega_{K+1}})\big(\mathbb{E}(f|\mathcal{B}_{K+1})-\mathbb{E}(f|\mathcal{B}_K)\big)\Big|\Big|^2_{L^2(\mathbb{Z}_N)}\\
&\qquad=\Big|\Big|(1-1_{\Omega_{K+1}})\mathbb{E}(f|\mathcal{B}_K)\Big|\Big|^2_{L^2(\mathbb{Z}_N)}+\Big|\Big|(1-1_{\Omega_{K+1}})\big(\mathbb{E}(f|\mathcal{B}_{K+1})-\mathbb{E}(f|\mathcal{B}_K)\big)\Big|\Big|^2_{L^2(\mathbb{Z}_N)}\\
&\qquad\quad+2\Big\langle(1-1_{\Omega_{K+1}})\mathbb{E}(f|\mathcal{B}_K),(1-1_{\Omega_{K+1}})\big(\mathbb{E}(f|\mathcal{B}_{K+1})-\mathbb{E}(f|\mathcal{B}_K)\big)\Big\rangle.
\end{aligned}$$

Entonces por (6.86) será suficiente mostrar que

$$\Big\langle(1-1_{\Omega_{K+1}})\mathbb{E}(f|\mathcal{B}_K),(1-1_{\Omega_{K+1}})\mathbb{E}\big(f|\mathcal{B}_{K+1})-\mathbb{E}(f|\mathcal{B}_K)\big)\Big\rangle=O_{K,\epsilon}\big(\eta^{1/2}\big).$$

Como $\big(1-1_{\Omega_{K+1}}\big)^2=\big(1-1_{\Omega_{K+1}}\big)$,

$$\begin{aligned}
&\Big\langle(1-1_{\Omega_{K+1}})\mathbb{E}(f|\mathcal{B}_K),(1-1_{\Omega_{K+1}})\mathbb{E}\big(f|\mathcal{B}_{K+1})-\mathbb{E}(f|\mathcal{B}_K)\big)\Big\rangle\\
&\qquad\qquad=\Big\langle\big(1-1_{\Omega_{K+1}}\big)\mathbb{E}(f|\mathcal{B}_K),\mathbb{E}(f|\mathcal{B}_{K+1})-\mathbb{E}(f|\mathcal{B}_K)\Big\rangle.
\end{aligned}\tag{6.90}$$

Ahora notemos que $(1-1_{\Omega_K})\mathbb{E}(f|\mathcal{B}_K)$ es medible con respecto a $\mathcal{B}_K$, y por consiguiente ortogonal a $\mathbb{E}(f|\mathcal{B}_{K+1})-\mathbb{E}(f|\mathcal{B}_K)$, esto es,

$$\Big\langle \big(1-1_{\Omega_K}\big)\mathbb{E}(f|\mathcal{B}_K), \mathbb{E}(f|\mathcal{B}_{K+1})-\mathbb{E}(f|\mathcal{B}_K)\Big\rangle = 0 \tag{6.91}$$

puesto que $\mathcal{B}_K$ es un sub-$\sigma$-álgebra de $\mathcal{B}_{K+1}$. Así, de (6.90) y (6.91) resulta

$$\begin{aligned}\Big\langle (1-1_{\Omega_{K+1}})\mathbb{E}(f|\mathcal{B}_K),&(1-1_{\Omega_{K+1}})\mathbb{E}\big(f|\mathcal{B}_{K+1})-\mathbb{E}(f|\mathcal{B}_K)\big)\Big\rangle \\ &= -\Big\langle \big(1_{\Omega_{K+1}}-1_{\Omega_K}\big)\mathbb{E}(f|\mathcal{B}_K), \mathbb{E}(f|\mathcal{B}_{K+1})-\mathbb{E}(f|\mathcal{B}_K)\Big\rangle. \end{aligned}\tag{6.92}$$

Como antes, del hecho que $\big(1_{\Omega_{K+1}}-1_{\Omega_K}\big)\mathbb{E}(f|\mathcal{B}_K)$ es medible con respecto a $\mathcal{B}_{K+1}$, tenemos que

$$\Big\langle \big(1_{\Omega_{K+1}}-1_{\Omega_K}\big)\mathbb{E}(f|\mathcal{B}_K), f-\mathbb{E}(f|\mathcal{B}_{K+1})\Big\rangle = 0,$$

y por lo tanto, de (6.92),

$$\begin{aligned}\Big\langle (1-1_{\Omega_{K+1}})\mathbb{E}(f|\mathcal{B}_K),&(1-1_{\Omega_{K+1}})\mathbb{E}\big(f|\mathcal{B}_{K+1})-\mathbb{E}(f|\mathcal{B}_K)\big)\Big\rangle \\ &= -\Big\langle \big(1_{\Omega_{K+1}}-1_{\Omega_K}\big)\mathbb{E}(f|\mathcal{B}_K), f-\mathbb{E}(f|\mathcal{B}_K)\Big\rangle. \end{aligned}\tag{6.93}$$

Como $\mathbb{E}(f|\mathcal{B}_K)(x)\leq 2$ si $\eta<\eta_0(\epsilon,K)$ es suficientemente pequeño y $x\notin\Omega_K$ (puesto que se cumple (6.74)), podemos mayorar (6.93) por

$$2\mathbb{E}\big((1_{\Omega_{K+1}}-1_{\Omega_K})|f-\mathbb{E}(f|\mathcal{B}_K)|\big).$$

Dado que estamos suponiendo que $0\leq f(x)\leq\nu(x)$, tenemos que

$$\begin{aligned}2\mathbb{E}\big((1_{\Omega_{K+1}}-1_{\Omega_K})|f-\mathbb{E}(f|\mathcal{B}_K)|\big) &\leq 2\mathbb{E}\big((1_{\Omega_{K+1}}-1_{\Omega_K})|\nu+\mathbb{E}(\nu|\mathcal{B}_K)|\big). \\ &\leq 2\mathbb{E}\big((1_{\Omega_{K+1}}-1_{\Omega_K})|2\nu+\mathbb{E}(\nu|\mathcal{B}_K)|\big) \\ &\leq 4\mathbb{E}\big((1_{\Omega_{K+1}}-1_{\Omega_K})(\nu+1)\big),\end{aligned}$$

donde la última desigualdad se da porque $\mathbb{E}(\nu|\mathcal{B}_K)(x)\leq 2$ para $x\notin\Omega_K$ (de (6.71)).
Puesto que $\mathbb{E}\big((1_{\Omega_{K+1}}-1_{\Omega_K})(\nu+1)\big)=O_{K,\epsilon}(\eta^{1/2})$ como requeríamos, por (6.70). Esto concluye la demostración de la proposición 6.8, y por consiguiente del teorema 6.3. □

# Capítulo 7

# Una medida pseudoaleatoria que acota a los números primos

Una vez probado el teorema 6.3, para mostrar que los números primos contienen infinitas progresiones aritméticas de longitud arbitraria solo nos queda probar que existe una medida pseudoaleatoria $\nu : \mathbb{Z}_N \to \mathbb{R}^+$ que acota superiormente a una función adecuada $f : \mathbb{Z}_N \to \mathbb{R}^+$ soportada en el conjunto de los números primos, es decir, una función que es nula fuera del conjunto de los números primos y no nula en gran parte de los primos, donde identificamos $\{1, 2, \ldots, N\}$ con $\mathbb{Z}_N$.
Así, el resultado principal por demostrar es el siguiente teorema.

**Proposición 7.1.** *Los números primos contienen infinitas progresiones aritméticas de longitud $k$ para todo $k \geq 3$ entero.*

Como en casi cualquier problema aditivo involucrando números primos, iniciamos considerando la función de von Mangoldt.

**Definición 7.1.** *Definimos la* **función de von Mangoldt** $\Lambda : \mathbb{Z}^+ \to \mathbb{R}$ *por*

$$\Lambda(n) = \begin{cases} \log p & , \text{ si } n = p^k, \text{ para algún } k \in \mathbb{N}, \\ 0 & , \text{ en otro caso.} \end{cases}$$

Del teorema del número primo, tenemos que

$$\frac{1}{N} \sum_{n \leq N} \Lambda(n) = 1 + o(1). \tag{7.1}$$

Así, para probar el teorema 7.1 es suficiente exhibir una medida $k$-pseudoaleatoria $\nu : \mathbb{Z}_N \to \mathbb{R}^+$, tal que $\nu(n) \geq c(k)\Lambda(n)$, para alguna constante $c(k) > 0$ dependiendo solamente de $k$, ya que

$$\mathbb{E}\big(c(k)\Lambda(n)\big|n \in \mathbb{Z}_N\big) = c(k) > 0,$$

y $\Lambda$ está soportada en los números primos. Mas tal medida $k$-pseudoaleatoria no puede existir porque los primos (y la función de von Mangoldt) están concentrados sobre cierta clase de residuos. Para cada entero $q > 1$, $\Lambda$ es solamente no nula sobre aquellas $\phi(q)$ clases de residuos $a \pmod q$ para los que $(a, q) = 1$; mientras que una medida pseudoaleatoria está uniformemente distribuida sobre todas las $q$ clases de residuos.

Para salvar esta dificultad modificaremos la función de von Mangoldt $\Lambda$.

**Definición 7.2.** *Definimos la* **función de von Mangoldt modificada** $\tilde{\Lambda} : \mathbb{Z}_N \to \mathbb{R}$ *(dependiendo de $N$) como*

$$\tilde{\Lambda}(n) = \begin{cases} \frac{\Phi(W)}{W}\log(Wn+1) & \text{, si } Wn+1 \text{ es primo,} \\ 0 & \text{, en otro caso,} \end{cases}$$

*donde* $W = W(N) = \prod_{\substack{p \leq \omega(N) \\ p \text{ primo}}} p$, *y* $\omega : \mathbb{Z}^+ \to \mathbb{R}$ *es una función que tiende lentamente a* $+\infty$ *cuando $N$ tiende a* $+\infty$ *(por ejemplo* $\omega(N) = \log\log N$*). Obsérvese que estamos identificando* $\mathbb{Z}_N$ *con* $\{1, 2, \ldots, N\}$.

Del teorema del número primo también tenemos que $\dfrac{1}{N}\sum_{n \leq N}\tilde{\Lambda}(n) = 1 + o(1)$.

**Proposición 7.2.** *Sean* $\epsilon_k = \dfrac{1}{2^k(k+4)!}$ *y $N$ un primo suficientemente grande. Entonces existe una medida $k$-pseudoaleatoria* $\nu : \mathbb{Z}_N \to \mathbb{R}^+$ *tal que*

$$\nu(n) \geq k^{-1}2^{-k-5}\tilde{\Lambda}(n), \text{ para todo } \epsilon_k N \leq n \leq 2\epsilon_k N.$$

*Prueba del teorema 7.1.* Asumiendo la proposición 7.2 consideremos $N$, un número primo grande. Definamos la función $f \in L^1(\mathbb{Z}_N)$ como

$$f(n) = \begin{cases} k^{-1}2^{-k-5}\tilde{\Lambda}(n) & \text{, si } \epsilon_k N \leq n \leq 2\epsilon_k N, \\ 0 & \text{, en otro caso.} \end{cases}$$

Del teorema de Dirichlet obtenemos que

$$\begin{aligned}\mathbb{E}(f) &= \frac{k^{-1}2^{-k-5}}{N}\sum_{\epsilon_k N\le n\le 2\epsilon_k N}\tilde{\Lambda}(n)\\ &= k^{-1}2^{-k-5}\epsilon_k\frac{1}{\epsilon_k N}\sum_{\epsilon_k N\le n\le 2\epsilon_k N}\tilde{\Lambda}(n)\\ &= k^{-1}2^{-k-5}\epsilon_k\big(1+o(1)\big)=\delta\big(1+o(1)\big),\end{aligned}$$

donde $\delta = k^{-1}2^{-k-5}\epsilon_k$. Ahora, aplicando la proposición 7.2 y el teorema 6.3 concluimos que

$$\mathbb{E}\bigg(f(x)f(x+r)\cdots f\big(x+(k-1)r\big)\bigg|x,r\in\mathbb{Z}_N\bigg)\ge c(k,\delta)-o(1).$$

Observe que el caso degenerado $r=0$ puede contribuir a lo más con $\frac{1}{N}\log^k(WN+1)$, y gracias a que $\omega(N)$ crece lentamente y para $N$ suficientemente grande, este caso puede contribuir con a lo más $O\bigg(\frac{1}{N}\log^k\big(N^2\big)\bigg)$. Pero este término es despreciable, siendo $O\bigg(\frac{1}{N}\log^k N\bigg)=o(1)$, y puede ser absorbido por el término $-o(1)$ del lado derecho de la desigualdad. Además, cada progresión contada por la expresión del lado izquierdo no es sólo una progresión en $\mathbb{Z}_N$, sino es una genuina progresión en $\mathbb{Z}$ puesto que $\epsilon_k < 1/k$. Observemos que $f(x)\neq 0$ únicamente cuando $x = Wn+1$ para algún $n$, y $Wn+1$ es primo. Como el lado derecho de la desigualdad es positivo para $N$ suficientemente grande, el teorema se sigue de la definición de $f$ y de $\tilde{\Lambda}$. □

Así, sólo resta probar la proposición 7.2. Para obtener una cota para $\tilde{\Lambda}(n)$, consideramos la fórmula $\Lambda(n)=\sum_{d|n}\mu(d)\log(n/d)=\sum_{d|n}\mu(d)\log(n/d)_+$ para la función de von Mangoldt, donde $\mu$ es la función de Möbius y $\log(x)_+ = \max\big\{\log(x),\,0\big\}$. Motivado por esto, definimos lo siguiente.

**Definición 7.3.** *Sea $R$ un parámetro fijado. Definimos la función $\Lambda_R:\mathbb{Z}^+\to\mathbb{R}^+$, donde $\mathbb{R}^+=[0;+\infty)$, como*

$$\Lambda_R(n)=\sum_{\substack{d|n\\ d\le R}}\mu(d)\log(R/d)=\sum_{d|n}\mu(d)\log(R/d)_+.$$

**Definición 7.4.** *Sean $R=N^{k^{-1}2^{-k-4}}$ y $\epsilon_k=\dfrac{1}{2^k(k+4)!}$. Definimos la función $\nu:\mathbb{Z}_N\to\mathbb{R}^+$ por*

$$\nu(n)=\begin{cases}\frac{\phi(W)}{W}\frac{\Lambda_R(Wn+1)^2}{\log R} & ,\ \textit{cuando } \epsilon_k N\le n\le 2\epsilon_k N;\\ 1 & ,\ \textit{en otro caso,}\end{cases}$$

*para todo $0<n\le N$, donde identificamos $\{1,2,\ldots,N\}$ con $\mathbb{Z}_N$ en la manera usual.*

La función $\nu$ será nuestro mayorante requerido en la proposición 7.2, como lo muestra el siguiente lema.

**Lema 7.1.** *Se cumple que $\nu(n) \geq 0$ para todo $n \in \mathbb{Z}_N$, además $\nu(n) \geq k^{-1}2^{-k-5}\tilde{\Lambda}(n)$ para todo $\epsilon_k N \leq n \leq 2\epsilon_k N$ (si $N$ es suficientemente grande dependiendo de $k$).*

*Demostración.* La primera afirmación es trivial pues por definición $\nu$ es no negativo. La segunda afirmación (la desigualdad) es trivial para aquellos $n$ tales que $Wn+1$ no es primo. Por otro lado, si $Wn+1$ es primo, los divisores de este número son únicamente 1 y el propio $Wn+1$; como $W$ es grande, $Wn+1 > W\epsilon_k N > R$ , siempre que $N$ sea suficientemente grande. Luego tenemos que $\Lambda_R(Wn+1) = \sum_{d|Wn+1} \mu(d)\log(R/d)_+ = \mu(1)\log R = \log R$. Por lo tanto

$$\begin{aligned}\nu(n) &= \frac{\phi(W)}{W}\log R = \frac{\phi(W)}{W}k^{-1}2^{-k-4}\log N \\ &= k^{-1}2^{-k-5}\frac{\phi(W)}{W}\log N^2 \geq k^{-1}2^{-k-5}\frac{\phi(W)}{W}\log(Wn+1),\end{aligned}$$

donde la última desigualdad se da asumiendo que $\omega(N)$ tiene un crecimiento suficientemente lento de modo que $\log N^2 \geq \log(WN+1)$, para lo que basta que $W \leq N-1$. □

Requeriremos de los dos resultados siguientes, que son esencialmente debidos a Goldston y Yıldırım, necesarios para mostrar que $\nu$ es una medida pseudoaleatoria.

**Proposición 7.3** (Golston-Yıldırım)**.** *Sean $m, t \in \mathbb{Z}^+$. Para cada $1 \leq i \leq m$, sean $\psi_i(x) = \sum_{j=1}^t L_{ij}x_j + b_i$, formas lineales con coeficientes enteros $L_{ij}$ tales que $|L_{ij}| \leq \sqrt{\omega(N)}/2$ para cada $i = 1, 2, \ldots, m$ y $j = 1, 2, \ldots, t$. Asumiendo que las $t$-uplas $(L_{ij})_{j=1}^t$ no son idénticamente nulas, y que cualquier $t$-upla no es múltiplo racional de cualquier otra. Escribimos $\theta_i = W\psi_i + 1$. Supongamos que $B = \prod_{i=1}^t I_i \subseteq \mathbb{R}^t$, donde los $I_i$ son intervalos con $|I_i| \geq R^{10m}$. Entonces (asumiendo que la función $\omega(N)$ tiene crecimiento suficientemente lento en $N$)*

$$\mathbb{E}\Big(\Lambda_R\big(\theta_1(x)\big)^2 \cdots \Lambda_R\big(\theta_m(x)\big)^2 \Big| x \in B\Big) = \big(1 + O_{m,t}(1)\big)\left(\frac{W\log R}{\phi(W)}\right)^m.$$

La siguiente proposición ayudará a mostrar que $\nu$ satisface la condición de correlación.

**Proposición 7.4** (Goldston-Yıldırım)**.** *Sea $m \geq 1$ entero, y sea $B$ un intervalo de al menos $R^{10m}$. Supongamos que $h_1, \ldots, h_m$ son enteros distintos entre sí satisfaciendo $|h_i| \leq N^2$ para todo $1 \leq i \leq m$, y sea $\Delta = \prod_{1 \leq i < j \leq m} |h_i - h_j|$. Entonces, para $N$ suficientemente grande,*

*dependiendo de $m$, y admitiendo que la función $\omega(N)$ tiene un crecimiento suficientemente lento en $N$*

$$\mathbb{E}\Big(\Lambda_R\big(W(x+h_1)+1\big)^2\cdots\Lambda_R\big(W(x+h_m)+1\big)^2\Big|x\in B\Big) \\ \leq \big(1+o_m(1)\big)\Big(\frac{W\log R}{\phi(W)}\Big)^m\prod_{p|\Delta}\big(1+O_m(p^{-1/2})\big), \qquad (7.2)$$

*donde $p$ recorre el conjunto de los números primos.*

Asumiendo ambas proposiciones 7.3 y 7.4 podemos concluir la prueba de la proposición 7.2. Iniciamos mostrando que $\nu$ es una medida.

**Lema 7.2.** *La función $\nu:\mathbb{Z}_{\mathrm{N}}\to\mathbb{R}^+$ construida en la definición 7.4 es una medida, esto es, verifica $\mathbb{E}(\nu)=1+o(1)$.*

*Demostración.* Aplicando la proposición 7.2 con $m=t=1$, $\psi_1(x_1)=x_1$ y $B=[\epsilon_k N, 2\epsilon_k N]$. Tomando $N$ suficientemente grande (dependiendo de $k$) tenemos que $|B|=\epsilon_k N\geq R^{10}$. Por lo tanto

$$\mathbb{E}\Big(\Lambda_R(Wx+1)^2\Big|x\in B\Big) \;=\; \big(1+o(1)\big)\Big(\frac{W\log R}{\phi(W)}\Big),$$

y así

$$\begin{aligned}\mathbb{E}\Big(\frac{\phi(W)}{W}\frac{\Lambda_R(Wx+1)^2}{\log R}\Big|x\in B\Big) &= 1+o(1),\\ \mathbb{E}\big(\nu(x)|x\in B\big) &= 1+o(1).\end{aligned}$$

Pero de la misma definición, $\nu(x)=1$ si $x\in\mathbb{Z}_N\setminus B$, y por lo tanto

$$\mathbb{E}\big(\nu(x)|x\in\mathbb{Z}_N\setminus B\big)=1.$$

Así,

$$\begin{aligned}N\mathbb{E}\big(\nu(x)|x\in\mathbb{Z}_N\big) &= \sum_{x\in B}\nu(x)+\sum_{x\in\mathbb{Z}_N\setminus B}\nu(x)\\ &= \epsilon_k N\mathbb{E}\big(\nu(x)|x\in B\big)+(N-\epsilon_k N)\mathbb{E}\big(\nu(x)|x\in\mathbb{Z}_N\setminus B\big),\end{aligned}$$

y

$$\mathbb{E}\big(\nu(x)|x\in\mathbb{Z}_N\big) \;=\; \epsilon_k\big(1+o(1)\big)+(1-\epsilon_k)\big(1+o(1)\big)=1+o(1),$$

como queríamos. □

Ahora verificaremos que $\nu$ satisface la condición de formas lineales.

**Proposición 7.5.** *La medida $\nu$ satisface la $(k \cdot 2^{k-1}, 3k-4, k)$-condición de formas lineales.*

*Demostración.* Sean $\psi_i(x) = \sum_{j=1}^{t} L_{ij}x_j + b_i$ formas lineales como las que aparecen en la definición 6.3, esto es, cumpliendo que $m \leq k \cdot 2^{k-1}$, $t \leq 3k-4$ y donde los $L_{ij}$ son números racionales con numerador y denominador a lo más $k$ en valor absoluto, y ninguna de las $t$-uplas $(L_{ij})_{j=1}^{t}$ es cero o es múltiplo racional de alguna otra. Requerimos mostrar que

$$\mathbb{E}\big(\nu(\psi_1(x)) \cdots \nu(\psi_m(x))\big| x \in \mathbb{Z}_N^t\big) = 1 + o(1). \tag{7.3}$$

Podemos eliminar los denominadores de los $L_{ij}$ y asumir que todos los $L_{ij}$ son enteros, a expensas de incrementar la cota superior de los $L_{ij}$ a $|L_{ij}| \leq (k+1)!$. Puesto que $\omega(N)$ tiende a $+\infty$ cuando $N$ tiende a $+\infty$, podemos asumir que $(k+1)! < \sqrt{\omega(N)}/2$ siempre que consideremos $N$ suficientemente grande. De este modo tenemos que $|L_{ij}| \leq \sqrt{\omega(N)}/2$. Debido a que $\nu$ está definida a través de dos reglas de correspondencia, no podemos aplicar directamente la proposición 7.3. Dividimos el rango de sumación $\mathbb{Z}_N^t$ en $Q^t$ bloques, todos de casi igual tamaño, donde $Q = Q(N)$ es una función que crece lentamente a $+\infty$ con $N$, que será fijada después. Sean los bloques

$$B_{u_1u_2\cdots u_t} = \Big\{x \in \mathbb{Z}_N^t : x_j \in \big[\lfloor u_j N/Q \rfloor, \lfloor (u_j+1)N/Q \rfloor\big[\,,\ j = 1, 2, \ldots, t\Big\},$$

donde $u_1, u_2, \ldots, u_t \in Z_Q$, e identificamos $Z_Q$ con $\{0, 1, 2, \ldots, Q-1\}$. Obsservemos que salvo un error multiplicativo de $1 + o(1)$, podemos escribir el lado izquierdo de (7.3) como

$$\mathbb{E}\bigg(\mathbb{E}\big(\nu(\psi_1(x)) \cdots \nu(\psi_m(x))\big| x \in B_{u_1u_2\cdots u_t}\big)\bigg| u_1, u_2, \ldots, u_t \in Z_Q\bigg).$$

En efecto,

$$\begin{aligned}
&\mathbb{E}\Big(\nu\big(\psi_1(x)\big) \cdots \nu\big(\psi_m(x)\big)\Big| x \in \mathbb{Z}_N^t\Big) \\
&\qquad = \frac{1}{N^t} \sum_{x. \in \mathbb{Z}_N^t} \nu\big(\psi_1(x)\big) \cdots \nu\big(\psi_m(x)\big) \\
&\qquad = \frac{1}{N^t} \sum_{u_1,\ldots,u_t \in Z_Q} \sum_{x \in B_{u_1u_2\cdots u_t}} \nu\big(\psi_1(x)\big) \cdots \nu\big(\psi_m(x)\big) \\
&\qquad = \sum_{u_1,\ldots,u_t \in Z_Q} \frac{|B_{u_1u_2\cdots u_t}|}{N^t} \mathbb{E}\Big(\nu\big(\psi_1(x)\big) \cdots \nu\big(\psi_m(x)\big)\Big| x \in B_{u_1u_2\cdots u_t}\Big),
\end{aligned}$$

donde $|B_{u_1u_2\cdots u_t}|$ es aproximadamente $N^t/Q^t$. Llamamos a una $t$-upla $(u_1,\dots,u_t) \in Z_Q^t$ *buena*, si para cada $1 \le i \le m$ los conjuntos $\psi_i(B_{u_1u_2\cdots u_t})$ están completamente contenidos en $[\epsilon_k N, 2\epsilon_k N]$, o son disjuntos de este intervalo. De la proposición 7.3 y de la definición 7.4 tenemos que

$$\mathbb{E}\big(\nu(\psi_1(x))\cdots\nu(\psi_m(x))\big|x \in B_{u_1u_2\cdots u_t}\big) = 1 + o_{m,t}(1),$$

siempre que la $t$-upla $(u_1,u_2,\dots,u_t)$ fuese buena, puesto que podemos sustituir $\nu(\psi_i(x))$ por $\dfrac{\phi(W)}{W\log R}\Lambda_R^2(\theta_i(x))$, o por 1, si el crecimiento de $Q$ es suficientemente lento de modo que $N/Q > R^{10m}$ por la definición de $R$. Cuando $(u_1,\dots,u_t)$ es una $t$-upla *no buena* podemos acotar superiormente $\nu$ por $1 + \dfrac{\phi(W)}{W\log R}\Lambda_R^2(\theta_i(x))$. Multiplicamos, expandimos y aplicamos la proposición 7.3 a cada término para obtener una cota superior del tipo

$$\mathbb{E}\big(\nu(\psi_1(x))\cdots\nu(\psi_m(x))\big|x \in B_{u_1u_2\cdots u_t}\big) = O_{m,t}(1) + o_{m,t}(1).$$

Veremos a continuación que la proporción de $t$-uplas no buenas es a lo más $O_{m,t}(1/Q)$. Por lo tanto el lado derecho de (7.3) es $1 + o_{m,t}(1) + O_{m,t}(1/Q) = 1 + o_{m,t}(1)$, lo cual concluye la demostración una vez elegido $Q$ con crecimiento suficientemente lento en $N$.

En lo que resta verificaremos la afirmación acerca de la proporción de $t$-uplas no buenas.

Supongamos que $(u_1,u_2,\dots,u_t)$ es una $t$-upla no buena, entonces existen $1 \le i \le m$ y $x, x' \in B_{u_1u_2\cdots u_t}$ tal que $\psi_i(x) \in [\epsilon_k N, 2\epsilon_k N]$, pero $\psi_i(x') \notin [\epsilon_k N, 2\epsilon_k N]$. Pero de la definición de $B_{u_1u_2\cdots u_t}$ (y por la acotación de los $L_{ij}$) tenemos que

$$\psi_i(x),\, \psi_i(x') = \sum_{j=1}^{t} L_{ij}\lfloor Nu_j/Q\rfloor + b_i + O_{m,t}(N/Q).$$

Por lo tanto, para algún $a \in \{1,2\}$

$$a\epsilon_k N = \sum_{j=1}^{t} L_{ij}\lfloor Nu_j/Q\rfloor + b_i + O_{m,t}(N/Q).$$

Dividiendo por $N/Q$ obtenemos

$$\sum_{j=1}^{t} L_{ij}u_j = a\epsilon_k Q + b_i Q/N + O_{m,t}(1) \bmod Q.$$

Como $(L_{ij})_{j=1}^t$ es no nulo, el número de $t$-uplas $(u_1,u_2,\dots,u_t)$ que satisfacen esta ecuación es a lo más $O_{m,t}(Q^{t-1})$. Haciendo que $a$ e $i$ varíen concluimos que la proporción de $t$-uplas no buenas es a lo más $O_{m,t}(1/Q)$. □

A continuación usaremos la proposición 7.4 para mostrar que $\nu$ satisface la condición de correlación; para ello estimaremos primero el factor $\prod_{p|\Delta,\ p \text{ primo}} \big(1 + O_m(p^{-1/2})\big)$ que aparece en tal proposición.

**Lema 7.3.** *Sea $m \geq 1$. Existe una función peso $\tau = \tau_m : \mathbb{Z} \to \mathbb{R}^+$ tal que $\tau(n) \geq 1$ para todo $n \neq 0$, y para cada $h_1, h_2, \ldots, h_m \in \big[\epsilon_k N, 2\epsilon_k N\big]$ distintos, tenemos que*

$$\prod_{p|\Delta,\ p\ primo} \big(1 + O_m(p^{-1/2})\big) \leq \sum_{1\leq i<j\leq m} \tau(h_i - h_j),$$

*donde $\Delta = \prod_{1\leq i<j\leq m} |h_i - h_j|$, de modo que $\mathbb{E}\big(\tau^q(n)\big|0 < |n| \leq N\big) = O_{m,q}(1)$ para todo $0 < q < +\infty$.*

*Demostración.* Notemos que

$$\prod_{p|\Delta,\ p \text{ primo}} \big(1 + O_m(p^{-1/2})\big) \leq \prod_{1\leq i<j\leq m} \left( \prod_{p|h_i-h_j,\ p \text{ primo}} (1 + p^{-1/2}) \right)^{O_m(1)}.$$

Podemos entonces, usando la desigualdad entre la media aritmética y la media geométrica (y absorbiendo las constantes en el factor y en el exponente $O_m(1)$) tomar

$$\tau_m(n) = O_m(1) \prod_{p|n,\ p \text{ primo}} \big(1 + O_m(p^{-1/2})\big)^{O_m(1)}$$

para cada $n \neq 0$ (el valor de $\tau$ en 0 es irrelevante para el lema, puesto que estamos considerando los $h_i$ distintos). Para concluir la prueba del lema basta mostrar que

$$\mathbb{E}\left( \prod_{p|n,\ p \text{ primo}} (1 + p^{-1/2})^{O_m(1)} \middle| 0 < |n| \leq N \right) = O_{m,q}(1)$$

para $0 < q < +\infty$. Como $(1+p^{-1/2})^{O_m(q)} \leq 1 + p^{-1/4}$ para todos los primos $p$, con excepción de a lo más $O_{m,q}(1)$ primos, tenemos que

$$\mathbb{E}\left( \prod_{p|n,\ p \text{ primo}} (1+p^{-1/2})^{O_m(q)} \middle| 0 < |n| \leq N \right) \leq O_{m,q}(1) \cdot \mathbb{E}\left( \prod_{p|n,\ p \text{ primo}} (1+p^{-1/4}) \middle| 0 < |n| \leq N \right).$$

Por otro lado, $\prod_{p|n,\ p \text{ primo}} (1 + p^{-1/4}) \leq \sum_{d|n} d^{-1/4}$, y por lo tanto

$$\begin{aligned} \mathbb{E}\left( \prod_{p|n,\ p \text{ primo}} (1 + p^{-1/2})^{O_m(q)} \middle| 0 < |n| \leq N \right) &\leq O_{m,q}(1) \cdot \frac{1}{N} \sum_{1\leq n\leq N} \sum_{d|n} d^{-1/4} \\ &\leq O_{m,q}(1) \cdot \frac{1}{N} \sum_{d=1}^{N} \frac{N}{d} d^{-1/4} = O_{m,q}(1), \end{aligned}$$

lo último puesto que $\sum_{d=1}^{+\infty} d^{-5/4} < +\infty$. □

Probemos ahora que $\nu$ satisface la condición de correlación.

**Proposición 7.6.** *La medida $\nu$ satisface la $2^{k-1}$-condición de correlación.*

*Demostración.* Deseamos mostrar que para $1 \leq m \leq 2^{k-1}$ y $h_1, h_2, \ldots, h_m \in \mathbb{Z}_N$ tenemos que

$$\mathbb{E}\big(\nu(x+h_1)\nu(x+h_2)\cdots\nu(x+h_m)\big) \leq \sum_{1\leq i<j\leq m} \tau(h_i - h_j),$$

donde la función peso $\tau = \tau_m$ es acotada en $L^q$ para todo $q$.

Fijemos $m$, $h_1, \ldots, h_m$. Vamos a considerar la función peso construida en el lema 7.3 (identificando $\mathbb{Z}_N$ con los enteros entre $-N/2$ y $N/2$) multiplicado por un factor constante $O_m(1)$ conveniente, además definimos $\tau(0) = \exp(Cm \log N/\log\log N)$ para alguna constante absoluta y grande $C$. Por el lema anterior concluimos que $\mathbb{E}(\tau^q) = O_{m,q}(1)$ para todo $q$, pues $\tau(0)$ solo contribuye con $O_{m,q}(1)$ al valor de $\mathbb{E}(\tau^q)$.

Trataremos inicialmente el caso en que dos de los $h_i$ son iguales. En este caso podemos usar la estimación

$$||\nu||_{L^\infty} \leq \exp\left(\frac{2\log N}{\log\log N}\right),$$

que se sigue de la definición de $\nu$. En efecto, obtenemos de la definición que $|\nu(n)| \leq \log N \cdot d(n)$, donde $d(n)$ es el número de divisores de $n$. Tenemos por otro lado que

$$d(n) = O\left(\exp\left(\frac{3\log N}{\log\log N}\right)\right),$$

de donde se sigue que

$$\mathbb{E}\big(\nu(x+h_1)\cdots\nu(x+h_m)\big|x\in\mathbb{Z}_N\big) \leq ||\nu||_{L^\infty} \leq \tau(0) \leq \sum_{1\leq i<j\leq m} \tau(h_i - h_j),$$

ya que dos de los $h_i$ son iguales.

Supongamos ahora que los $h_i$ son todos distintos. Sea

$$g(n) = \frac{\phi(W)}{W}\frac{\Lambda_R^2(Wn+1)}{\log R}\cdot 1_{[\epsilon_k N, 2\epsilon_k N]}.$$

Por la construcción de $\nu$, tenemos que

$$\mathbb{E}\big(\nu(x+h_1)\cdots\nu(x+h_m)\big|x\in\mathbb{Z}_N\big) \leq \mathbb{E}\Big(\big(1+g(x+h_1)\big)\cdots\big(1+g(x+h_m)\big)\Big|x\in\mathbb{Z}_N\Big).$$

El lado derecho de la última desigualdad puede ser expresado como

$$\sum_{A\subseteq\{1,\dots,m\}} \mathbb{E}\left(\prod_{i\in A} g(x+h_i)\middle| x\in\mathbb{Z}_N\right).$$

Notemos que para $i,j\in A$ podemos suponer que $|h_i-h_j|\leq \epsilon_k N$, pues en caso contrario la esperanza correspondiente se anula. Por la proposición 7.4 y por el lema 7.3 obtenemos que

$$\mathbb{E}\left(\prod_{i\in A} g(x+h_i)\middle| x\in\mathbb{Z}_N\right) \leq \big(1+O_m(1)\big)\sum_{1\leq i<j\leq m}\tau_m(h_i-h_j),$$

y sumando sobre todos los subconjuntos $A$ de $\{1,2,\dots,m\}$, obtenemos el resultado deseado luego de multiplicar $\tau_m$ por una constante adecuada. □

## 7.1. Estimaciones de correlación para $\Lambda_R$

Para concluir la prueba del teorema 7.1 nos es suficiente verificar la proposición 7.3 y la proposición 7.4. Esto será alcanzado en esta sección asumiendo la estimación (lema 7.6) para una cierta clase de integral de contorno involucrando la función $\zeta$. Las técnicas en esta sección son semejantes a aquellas en [8].

### 7.1.1. La condición de formas lineales para $\Lambda_R$

Probaremos la proposición 7.3. Consideremos para cada $1\leq i\leq m$ las formas lineales $\psi_i(x)=\sum_{j=1}^t L_{ij}x_j+b_i$ en $t$ variables $x_1,\dots,x_t$, tales que los coeficientes $L_{ij}$ satisfacen $|L_{ij}|\leq\sqrt{\omega(N)}/2$, donde $\omega(N)$ es una función que tiende a $+\infty$ cuando $N$ así lo hace. Suponemos además que ninguna de las $t$-uplas $(L_{ij})_{j=1}^t$ es nula y que ninguna es múltiplo racional de alguna otra. Definamos $\theta_i=W\psi_i+1$.

Sea $B=\prod_{j=1}^t I_j$ el producto de los intervalos $I_j$, cada uno de los cuales con $|I_j|\geq R^{10m}$. Queremos mostar que

$$\mathbb{E}\Big(\Lambda_R\big(\theta_1(x)\big)^2\cdots\Lambda_R\big(\theta_m(x)\big)^2\Big|x\in B\Big)=\big(1+O_{m,t}(1)\big)\left(\frac{W\log R}{\phi(W)}\right)^m.$$

El primer paso es eliminar el papel de $B$ en el lado izquierdo de la igualdad deseada precedente. De la definición, $\Lambda_R(n)=\sum_{\substack{d|n\\ d\leq R}}\mu(d)\log(R/d)$, tenemos que

$$\mathbb{E}\Big(\Lambda_R\big(\theta_1(x)\big)^2\cdots\Lambda_R\big(\theta_m(x)\big)^2\Big|x\in B\Big)=\mathbb{E}\left(\prod_{i=1}^m\sum_{d_i,d_i'\leq R}\mu(d_i)\mu(d_i')\log\frac{R}{d_i}\log\frac{R}{d_i'}\middle| x\in B\right),$$

y el lado derecho puede ser reescrito como

$$\sum_{d_1,\dots,d_m,d_1',\dots,d_m'} \left( \prod_{i=1}^m \mu(d_i)\mu(d_i') \log\frac{R}{d_i} \log\frac{R}{d_i'} \right) \mathbb{E}\left( \prod_{i=1}^m 1_{d_i,d_i'|\theta_i(x)} \middle| x \in B \right). \tag{7.4}$$

De la presencia de la función de Möbius podemos asumir que todos los $d_i$ y $d_i'$ son libres de cuadrados. Sea $D = [d_1, \dots, d_m, d_1', \dots, d_m']$, el mínimo común múltiplo de los $d_i$ y $d_i'$, así $D \leq R^{2m}$. Observamos que la expresión $\prod_{i=1}^m 1_{d_i,d_i'|\theta_i(x)}$ es periódica con periodo $D$ en cada una de las componentes de $x$, y puede entonces ser definida sobre $\mathbb{Z}_D^t$. Puesto que $B$ es un producto de intervalos de longitud al menos $R^{10m}$, tenemos que

$$\mathbb{E}\left( \prod_{i=1}^m 1_{d_i,d_i'|\theta_i(x)} \middle| x \in B \right) = \mathbb{E}\left( \prod_{i=1}^m 1_{d_i,d_i'|\theta_i(x)} \middle| x \in \mathbb{Z}_D^t \right) + O_{m,t}(R^{-8m}).$$

La contribución del término de error $O_m(R^{-8m})$ a (7.4) puede ser crudamente estimada por $O_{m,t}(R^{-6m}\log^{2m} R)$. Así que nuestra tarea es ahora mostrar que

$$\begin{aligned} \sum_{d_1,\dots,d_m,d_1',\dots,d_m' \leq R} &\left( \prod_{i=1}^m \mu(d_i)\mu(d_i') \log\frac{R}{d_i} \log\frac{R}{d_i'} \right) \mathbb{E}\left( \prod_{i=1}^m 1_{d_i,d_i'|\theta_i(x)} \middle| x \in \mathbb{Z}_D^t \right) \\ &= \big(1 + o_{m,t}(1)\big) \left( \frac{W \log R}{\phi(W)} \right)^m. \end{aligned} \tag{7.5}$$

Para probar la igualdad anterior procederemos de manera estándar (como en [8]) para reescribir el lado izquierdo de la igualdad (7.5) como una integral de contorno de un producto de Euler, el cual puede ser puesto en términos de la función $\zeta$ Riemann y algunos otros factores simples. Iniciamos usando el teorema chino del resto (y los $d_i$, $d_i'$ libres de cuadrados) para obtener

$$\mathbb{E}\left( \prod_{i=1}^m 1_{d_i,d_i'|\theta_i(x)} \middle| x \in \mathbb{Z}_D^t \right) = \prod_{p|D} \mathbb{E}\left( \prod_{i:\, p|d_i d_i'} 1_{\theta_i(x)\equiv 0 (\mathrm{mod}\, p)} \middle| x \in \mathbb{Z}_p^t \right).$$

Notemos que la condición $p|D$ puede ser eliminada puesto que el multiplicando es 1 en otro caso. En particular, si escribimos $X_{d_1,\dots,d_m}(p) = \{1 \leq i \leq m : p|d_i\}$ y

$$\omega_X(p) = \mathbb{E}\left( \prod_{i\in X} 1_{\theta_i(x)\equiv 0 (\mathrm{mod}\, p)} \middle| x \in \mathbb{Z}_p^t \right), \tag{7.6}$$

para cada subconjunto $X \subseteq \{1, \dots, m\}$, entonces tenemos que

$$\mathbb{E}\left( \prod_{i=1}^m 1_{d_i,d_i'|\theta_i(x)} \middle| x \in \mathbb{Z}_D^t \right) = \prod_p \omega_{X_{d_1,\dots,d_m}(p) \cup X_{d_1',\dots,d_m'}(p)}(p).$$

Así, podemos escribir el lado izquierdo de (7.5) como

$$\sum_{d_1,\dots,d_m,d_1',\dots,d_m'\in\mathbb{Z}^+}\left(\prod_{i=1}^m\mu(d_i)\mu(d_i')\left(\log\frac{R}{d_i}\right)_+\left(\log\frac{R}{d_i'}\right)_+\right)\prod_p\omega_{X_{d_1,\dots,d_m}(p)\cup X_{d_1',\dots,d_m'}(p)}(p).$$

Ahora necesitamos expresar el logaritmo en términos de alguna función multiplicativa de los $d_i$ y $d_i'$. Con esta finalidad introducimos la recta vertical $\Gamma_1$

$$\Gamma_1(t)=\frac{1}{\log R}+it,\ t\in\mathbb{R}, \tag{7.7}$$

y observamos la igualdad

$$\frac{1}{2\pi i}\int_{\Gamma_1}\frac{x^z}{z^2}dz=(\log x)_+,$$

válida para cualquier $x>0$. Usando esta identidad podemos reescribir el lado izquierdo de (7.5) como

$$(2\pi i)^{-2m}\int_{\Gamma_1}\cdots\int_{\Gamma_1}F(z,z')\prod_{j=1}^m\frac{R^{z_j+z_j'}}{z_j^2{z_j'}^2}dz_jdz_j', \tag{7.8}$$

donde $2m$ integrales de contorno en las variables $z_1,\dots,z_m,z_1',\dots,z_m'$ sobre $\Gamma_1$, con $z=(z_1,\dots,z_m)$ y $z'=(z_1',\dots,z_m')$, y

$$F(z,z')=\sum_{d_1,\dots,d_m,d_1',\dots,d_m'\in\mathbb{Z}^+}\left(\prod_{j=1}^m\frac{\mu(d_j)\mu(d_j')}{d_j^{z_j}{d_j'}^{z_j'}}\right)\prod_p\omega_{X_{d_1,\dots,d_m}(p)\cup X_{d_1',\dots,d_m'}(p)}(p). \tag{7.9}$$

Observe que el sumando en (7.9) es una función multiplicativa de $D=[d_1,\dots,d_m,d_1',\dots,d_m']$, y así tenemos (al menos formalmente) la representación del producto de Euler $F(z,z')=\prod_p E_p(z,z')$, donde

$$E_p(z,z')=\sum_{X,X'\subseteq\{1,\dots,m\}}\frac{(-1)^{|X|+|X'|}\omega_{X\cup X'}(p)}{p^{\sum_{j\in X}z_j+\sum_{j\in X'}z_j'}}. \tag{7.10}$$

De (7.6) tenemos que $\omega_\emptyset(p)=1$ y $\omega_X(p)\le 1$, y así $E_p(z,z')=1+O_\sigma(1/p^\sigma)$ cuando $\mathcal{R}(z_j),\mathcal{R}(z_j')>\sigma$ (obtenemos estimaciones más precisas más adelante). Así, este producto de Euler es absolutamente convergente a $F(z,z')$ en el dominio $\{\mathcal{R}(z_j),\mathcal{R}(z_j')>1\}$ al menos.

Ahora explotaremos las hipótesis sobre la parte lineal de $\psi_1,\dots,\psi_m$, que son no nulas, y ninguna múltiplo racional de alguna otra. Esto será hecho vía las siguientes estimaciones elementales sobre $\omega_X(p)$.

**Lema 7.4.** *Si $p\le\omega(N)$, entonces $\omega_X(p)=0$ para $X\neq\emptyset$; en particular, $F_p=1$ cuando $p\le\omega(N)$. Por otro lado, si $p>\omega(N)$, entonces $\omega_X(p)=p^{-1}$ cuando $|X|=1$, y $\omega_X(p)\le p^{-2}$ cuando $|X|\ge 2$.*

*Demostración.* La primera afirmación es clara, puesto que los mapas $\theta_j : \mathbb{Z}_p^t \to \mathbb{Z}_p$ son idénticamente 1 cuando $p \leq \omega(N)$. La segunda afirmación (cuando $p > \omega(N)$ y $|X| = 1$) es similar puesto que en este caso $\theta_j$ cubre uniformemente $\mathbb{Z}_p$. Ahora supongamos que $p > \omega(N)$ y $|X| = 2$. Afirmamos que ninguna de las $s$ formas lineales puras $W(\psi_i - b_i)$ es múltiplo racional de alguna otra $(\mathrm{mod}\, p)$, en verdad, si esto fuese así entonces tendríamos $L_{ij}L_{i'j}^{-1} \equiv \lambda \,(\mathrm{mod}\, p)$ para alguna $\lambda$, y para todo $j = 1, \ldots, t$. Pero si $a/q$ y $a'/q'$ son dos números racionales irreducibles, con $|a|, |a'|, q, q' < \sqrt{\omega(N)}/2$, entonces claramente $a/q \not\equiv a'/q' \,(\mathrm{mod}\, p)$, a menos que $a = a'$ y $q = q'$. De esto se sigue que dos formas lineales puras $\psi_i - b_i$ y $\psi_{i'} - b_{i'}$ son tales que una sea múltiplo racional de la otra, lo que contradiría la hipótesis. Así, el conjunto de $x \in (\mathbb{Z}/p\,\mathbb{Z})^t$ para los cuales $\theta_i(x) \equiv 0 \,(\mathrm{mod}\, p)$ para todo $i \in X$, está contenido en la intersección de dos subespacios afines de $(\mathbb{Z}/p\,\mathbb{Z})^t$, y como tal tiene cardinalidad a lo más $p^{t-2}$. □

Este lema implica, comparando con (7.10), que

$$\begin{aligned} E_p(z, z') =& 1 - 1_{p>\omega(N)} \sum_{j=1}^{m} \left(p^{-1-z_j} + p^{-1-z'_j} - p^{-1-z_j-z'_j}\right) \\ &+ 1_{p>\omega(N)} \sum_{\substack{X, X' \subseteq \{1,\ldots,m\} \\ |X \cup X'| \geq 2}} \frac{O(1/p^2)}{p^{\sum_{j\in X} z_j} + \sum_{j \in X'} z'_j}, \end{aligned} \tag{7.11}$$

donde los términos $O(1/p^2)$ no dependen de $z, z'$. Para tomar ventaja de esta expansión, factorizamos $E_p = E_p^{(1)} E_p^{(2)} E_p^{(3)}$, donde

$$\begin{aligned} E_p^{(1)}(z, z') &= \frac{E_p(z, z')}{\prod_{j=1}^m (1 - 1_{p>\omega(N)} p^{-1-z_j})(1 - 1_{p>\omega(N)} p^{-1-z'_j})(1 - 1_{p>\omega(N)} p^{-1-z_j-z'_j})^{-1}} \\ E_p^{(2)}(z, z') &= \prod_{j=1}^{m} (1 - 1_{p\leq\omega(N)} p^{-1-z_j})^{-1}(1 - 1_{p\leq\omega(N)} p^{-1-z'_j})^{-1}(1 - 1_{p\leq\omega(N)} p^{-1-z_j-z'_j}) \\ E_p^{(3)}(z, z') &= \prod_{j=1}^{m} (1 - p^{-1-z_j})(1 - p^{-1-z'_j})(1 - p^{-1-z_j-z'_j})^{-1}. \end{aligned}$$

Escribiendo $G_j = \prod_p E_p^{(j)}$ para $j = 1, 2, 3$, tenemos que $F = G_1 G_2 G_3$ (al menos para $\mathcal{R}(z_j), \mathcal{R}(z'_j)$ suficientemente grande). Si introducimos la función zeta de Riemann,

$$\zeta(s) = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1},$$

tenemos que

$$G_3(z, z') = \prod_{j=1}^{m} \frac{\zeta(1 + z_j + z'_j)}{\zeta(1 + z_j)\zeta(1 + z'_j)}. \tag{7.12}$$

En particular, $G_3$ puede ser continuado meromórficamente a todo $\mathbb{C}^{2m}$.

**Definición 7.5.** *Para cada $\sigma > 0$ sea $\mathcal{D}_\sigma^m \subseteq \mathbb{C}^{2m}$, el dominio*

$$\mathcal{D}_\sigma^m = \Big\{ z_j, z_j' : -\sigma < \mathcal{R}(z_j), \mathcal{R}(z_j') < 100,\ j = 1, \dots, m \Big\}.$$

*Si $G = G(z, z')$ es una función analítica de $2m$ variables complejas sobrre $\mathcal{D}_\sigma^m$, consideramos la $C^k(\mathcal{D}_\sigma^m)$ norma de $G$ para cualquier entero $k \geq 0$ como*

$$||G||_{C^k(\mathcal{D}_\sigma^m)} = \sup_{a_1,\dots,a_m,a_1',\dots,a_m'} \left|\left| \left(\frac{\partial}{\partial z_1}\right)^{a_1} \cdots \left(\frac{\partial}{\partial z_m}\right)^{a_m} \left(\frac{\partial}{\partial z_1'}\right)^{a_1'} \cdots \left(\frac{\partial}{\partial z_m'}\right)^{a_m'} G \right|\right|_{L^\infty(\mathcal{D}_\sigma^m)},$$

*donde $a_1, \dots, a_m, a_1', \dots, a_m'$ son tales que $a_j \geq 0$, $a_j' \geq 0$ para cada $j = 1, \dots, m$, y $a_1 + \cdots + a_m + a_1' + \cdots + a_m' \leq k$.*

**Lema 7.5.** *Los productos de Euler $\prod_p E_p^{(j)}$ para $j = 1, 2$ son absolutamente convergentes en el dominio $\mathcal{D}_{1/6m}^m$. En particular, $G_1$ y $G_2$ pueden ser continuados analíticamente a este dominio. Además*

$$\begin{aligned} ||G_1||_{C^m(\mathcal{D}_{1/6m}^m)} &\leq O_m(1) \\ ||G_2||_{C^m(\mathcal{D}_\sigma^m)} &\leq O_{m,\omega(N)}(1) \\ G_1(0,0) &= 1 + o_m(1) \\ G_2(0,0) &= \big(W/\phi(W)\big)^m. \end{aligned}$$

*Demostración.* Primero consideremos $j = 1$. De (7.11) y de la expansión de Taylor tenemos la cota $E_p^{(1)}(z, z') = 1 + O_m(p^{-2+4/6m})$ en $\mathcal{D}_{1/6m}^m$, la cual da la convergencia deseada y también la $C^m(\mathcal{D}_{1/6m}^m)$ cota sobre $G_1$; la estimación para $G(0,0)$ también se sigue puesto que los factores de Euler $E_p^{(1)}(z, z')$ son idénticamente 1 cuando $p \leq \omega(N)$. La cota sobre $G_2$ es fácil, puesto que esta es sólo un producto finito de Euler involucrando a lo más $\omega(N)$ términos; la fórmula para $G_2(0,0)$ sigue de un cálculo directo puesto que $\dfrac{\phi(W)}{W} = \prod_{p<\omega(N)} \left(1 - \dfrac{1}{p}\right)$. □

Para hacer una estimación de las cotas de (7.8), invocamos el siguiente lema de integración de contorno.

**Lema 7.6.** *Sea $R$ un número real positivo. Sea $G = G(z, z')$ una función analítica de $2m$ variables complejas sobre el dominio $\mathcal{D}_\sigma^m$ para algún $\sigma > 0$, y supongamos que*

$$||G||_{C^m(\mathcal{D}_\sigma^m)} = \exp\big(O_{m,\sigma}(\log^{1/3} R)\big). \tag{7.13}$$

*Entonces*

$$\frac{1}{(2\pi i)^{2m}}\int_{\Gamma_1}\cdots\int_{\Gamma_1} G(z,z')\prod_{j=1}^{m}\frac{\zeta(1+z_j+z'_j)}{\zeta(1+z_j)\zeta(1+z'_j)}\frac{R^{z_j+z'_j}}{z_j^2 {z'_j}^2}dz_j dz'_j$$
$$= G(0,\dots,0)\log^m R + \sum_{j=1}^{m} O_{m,\sigma}\big(||G||_{C^j(\mathcal{D}_\sigma^m)}\log^{m-j} R\big) + O_{m,\sigma}\big(\exp(-\delta\sqrt{\log R})\big),$$

*para algún* $\delta = \delta(m) > 0$.

*Prueba de la proposición 7.3.* Aplicamos el lema 7.6 con $G = G_1G_2$ y $\sigma = 1/6m$. Del lema 7.5 y la regla de Leibnitz tenemos las cotas

$$||G||_{C^j(\mathcal{D}^m_{1/6m})} \le O_{j,m,\omega(N)}(1) \text{ para cada } 1 \le j \le m,$$

y en particular obtenemos (7.13) al elegir $\omega(N)$ creciendo suficientemente lento con $N$. También tenemos $G(0,0) = \big(1+o_m(1)\big)\left(\dfrac{W}{\phi(W)}\right)^m$ de este lema. Concluimos (de nuevo considerando $\omega(N)$ con crecimiento suficientemente lento en $N$) que la cantidad en (7.8) es $\big(1+o_m(1)\big)\left(\dfrac{W\log R}{\phi(W)}\right)^m$, como requeríamos. De este modo concluimos la prueba de la proposición 7.3. □

### 7.1.2. Correlación de orden superior para $\Lambda_R$

Ahora probaremos la proposición 7.4 usando argumentos similares a los usados para mostrar la proposición 7.3. La principal diferencia está en que el número de variables $t$ es en este caso solo 1, además todas las formas lineales son iguales entre sí, $\psi_i(x_1) = x_1$. En particular, cada forma lineal es ahora múltiplo racional de las otras, y por ende el lema 7.4 no es aplicable. Sin embargo, los argumentos anteriores a este lema, lema 7.4, son aún válidos; así podemos escribir el lado izquierdo de (7.2) como una expresión de la forma (7.8) más un error pequeño, donde $F$ está nuevamente definido por (7.9) y $E_p$ por (7.10); la diferencia es ahora que $\omega_X(p)$ está dado por

$$\omega_X(p) = \mathbb{E}\left(\prod_{i\in X} 1_{W(x+h_i)+1\equiv 0 \,(\mathrm{mod}\, p)}\,\middle|\, x\in\mathbb{Z}_p\right).$$

Tenemos también que $\omega_\emptyset(p) = 1$ para todo $p$. El análogo del lema 7.4 es el siguiente.

**Lema 7.7.** *Si* $p \le \omega(N)$, *entonces* $\omega_X(p) = 0$ *para todo conjunto* $X \ne \emptyset$; *en particular* $E_p = 1$ *cuando* $p \le \omega(N)$. *Si* $p > \omega(N)$, *entonces* $\omega_X(p) = p^{-1}$ *cuando* $|X| = 1$, *y* $\omega_X(p) \le p^{-1}$ *cuando* $|X| \ge 2$. *Además, si* $|X| \ge 2$ *entonces* $\omega_X(p) = 0$, *a menos que* $p$ *divida a* $\Delta = \prod_{1\le i<j\le s}|h_i - h_j|$.

*Demostración.* Cuando $p \leq \omega(N)$ tenemos que $W(x+h_i)+1 \equiv 1 \pmod p$ para cada $x \in \mathbb{Z}_p$, y cada $i \in X$, puesto que $W \equiv 0 \pmod p$ (cuando $X \neq \emptyset$), y por lo tanto $\omega_X(p) = 0$. Cuando $p > \omega(N)$ y $|X| \geq 1$, $\omega_X(p)$ es igual a $1/p$ cuando las clases de residuos $\{h_i \pmod p;\ i \in X\}$ son todas iguales entre sí, esto sucede en particular cuando $|X| = 1$. En efecto, sea $X \subseteq \{1,\dots,m\}$; puesto que $p > \omega(N)$ tenemos que $W \neq 0$ en $\mathbb{Z}_p$, y como $\{h_i \pmod p;\ i \in X\} = \{h_0\}$ para algún $h_0 \in \mathbb{Z}_p$,

$$\omega_X(p) = \mathbb{E}\Big(1_{W(x+h_0)+1\equiv 0 \,(\mathrm{mod}\, p)} \Big| x \in \mathbb{Z}_p\Big) = \frac{1}{p},$$

donde la última igualdad se da ya que $\psi : \mathbb{Z}_p \to \mathbb{Z}_p$ definido por $x \mapsto W(x+h_0)+1$ es biyectiva, pues $W \neq 0$. Por otro lado, si $p > \omega(N)$ y $|X| \geq 1$, $\omega_X(p) = 0$ cuando las clases de residuos $\{h_i \pmod p;\ i \in X\}$ no son todas iguales entre sí. En efecto, en este caso existen $h_0$ y $h_0'$ distintos en $\mathbb{Z}_p$ tales que $\prod_{i\in X} 1_{W(x+h_i)+1\equiv 0 \,(\mathrm{mod}\, p)}$ contiene al factor $1_{W(x+h_0)+1\equiv 0 \,(\mathrm{mod}\, p)} 1_{W(x+h_0')+1\equiv 0 \,(\mathrm{mod}\, p)}$, el cual es siempre nulo puesto que si $1_{W(x+h_0)+1\equiv 0 \,(\mathrm{mod}\, p)}(y) = 1$ para algún $y \in \mathbb{Z}_p$, esto es si $W(y+h_0)+1 \equiv 0 \pmod p$, entonces $W(y+h_0')+1 \not\equiv 0 \pmod p$ ya que $h_0$ y $h_0'$ son elementos distintos en $\mathbb{Z}_p$ y $W \neq 0$, y así $1_{W(x+h_0')+1\equiv 0 \,(\mathrm{mod}\, p)}(y) = 0$. Análogamente, si $1_{W(x+h_0')+1\equiv 0 \,(\mathrm{mod}\, p)}(y) = 1$ entonces si $1_{W(x+h_0)+1\equiv 0 \,(\mathrm{mod}\, p)}(y) = 0$. En cualquier caso

$$1_{W(x+h_0)+1\equiv 0 \,(\mathrm{mod}\, p)}(y) 1_{W(x+h_0')+1\equiv 0 \,(\mathrm{mod}\, p)}(y) = 0 \text{ para cada } y \in \mathbb{Z}_p,$$

por lo tanto $\prod_{i\in X} 1_{W(x+h_i)+1\equiv 0 \,(\mathrm{mod}\, p)}(y) = 0$ para cada $y \in \mathbb{Z}_p$, y así $\omega_X(p) = 0$. □

A la luz del lema anterior, el análogo de (7.11) es ahora

$$E_p(z,z') = 1 - 1_{p>\omega(N)} \sum_{j=1}^{m} \big(p^{-1-z_j} + p^{-1-z_j'} - p^{-1-z_j-z_j'}\big) + 1_{p>\omega(N),\, p|\Delta} \lambda_p(z,z'), \qquad (7.14)$$

donde $\lambda_p(z,z') = \sum_{\substack{X,X'\subseteq\{1,\dots,m\}\\ |X\cup X'|\geq 2}} \frac{O(1/p)}{p^{\sum_{j\in X} z_j + \sum_{j\in X'} z_j'}}$, y las cantidades $O(1/p)$ no dependen de $z, z'$. Podemos así factorizar

$$E_p = E_p^{(0)} E_p^{(1)} E_p^{(2)} E_p^{(3)},$$

donde

$$E_p^{(0)} = 1 + 1_{p>\omega(N),\, p|\Delta} \lambda_p(z,z')$$

$$E_p^{(1)} = \frac{E_p}{E_p^{(0)} \prod_{j=1}^{m} (1 - 1_{p>\omega(N)} p^{-1-z_j})(1 - 1_{p>\omega(N)} p^{-1-z_j'})(1 - 1_{p>\omega(N)} p^{-1-z_j-z_j'})^{-1}}$$

$$E_p^{(2)} = \prod_{j=1}^{m} (1 - 1_{p\le\omega(N)} p^{-1-z_j})^{-1}(1 - 1_{p\le\omega(N)} p^{-1-z'_j})^{-1}(1 - 1_{p\le\omega(N)} p^{-1-z_j-z'_j})$$
$$E_p^{(3)} = \prod_{j=1}^{m} (1 - p^{-1-z_j})(1 - p^{-1-z'_j})(1 - p^{-1-z_j-z'_j})^{-1}.$$

Sean $G_j = \prod_p E_p^{(j)}$. Entonces, consideremos como antes $F = G_0G_1G_2G_3$ y $G_3$ dado por (7.12). Para $G_0$, $G_1$ y $G_2$ tenemos el siguiente lema, análogo al lema 7.5.

**Lema 7.8.** *Sea $0 < \sigma < 1/6m$. Entonces los productos de Euler $\prod_p E_p^{(l)}$, para $l = 0, 1, 2$, son absolutamente convergentes en el dominio $\mathcal{D}_\sigma^m$. En particular, $G_0$, $G_1$ y $G_2$ pueden ser continuados analíticamente a este dominio. Además tenemos*

$$||G_0||_{C^r(\mathcal{D}_\sigma^m)} \le O_m\Big(\frac{\log R}{\log\log R}\Big)^r \prod_{p|\Delta}\big(1 + O_m(p^{2m\sigma-1})\big) \quad \textit{para } 0 \le r \le m, \tag{7.15}$$

$$||G_0||_{C^m(\mathcal{D}_{1/6m}^m)} \le \exp\big(O_m(\log^{1/3} R)\big), \tag{7.16}$$

$$||G_1||_{C^m(\mathcal{D}_{1/6m}^m)} \le O_m(1),$$

$$||G_2||_{C^m(\mathcal{D}_{1/6m}^m)} \le O_{m,\omega(N)}(1),$$

$$G_0(0,0) = \prod_{p|\Delta}\big(1 + O_m(p^{-1/2})\big), \tag{7.17}$$

$$G_1(0,0) = 1 + o_m(1),$$

$$G_2(0,0) = \big(W/\phi(W)\big)^m.$$

*Demostración.* Las estimaciones sobre $G_1$ y $G_2$ son análogas a las de $G_1$ y $G_2$ en el lema 7.5 (los factores adicionales $\lambda_p(z, z')$ que aparecen en el numerador y denominador de $E_p^{(1)}$ se cancelan para el primer orden y no representan mayor dificultad). Ahora haremos las estimaciones para $G_0$.

Probaremos (7.15). Fijemos $l$. Observemos que $G_0 = \prod_{p|\Delta} E_p^{(0)}$. Los números primos que dividen a $\Delta$ son a lo más $O(\log\Delta/\log\log\Delta)$. Usando la cota

$$\Delta = \prod_{1\le i<j\le m} |h_i - h_j| \le N^{m^2} \le R^{O_m(1)}, \tag{7.18}$$

vemos que el número de factores en el producto de Euler para $G_0$ es $O_m\Big(\frac{\log R}{\log\log R}\Big)$. Derivando $r$ veces para cada $1 \le r \le m$ y usando la regla de Leibnitz, obtenemos una suma de $O_m\big((\log R/\log\log R)^r\big)$ términos, cada uno de los cuales consiste de $O_m(\log R/\log\log R)$

factores, cada uno a su vez igual a alguna derivada de $1+\lambda_p(z,z')$ de orden entre 0 y $r$. Sobre $\mathcal{D}_\sigma^m$, cada factor es acotado por $1+O_m(p^{2m\sigma-1})$: los términos conteniendo un número no nulo de derivadas serán muy pequeños puesto que el término constante 1 es eliminado. De ello resulta (7.5).

Ahora mostremos (7.6). De (7.15), es sufiicciente mostrar que

$$\prod_{p|\Delta}\big(1+O_m(p^{2m\sigma-1})\big)\le \exp\big(O_m(\log^{1/3} R)\big).$$

Tomando logaritmos y usando la hipótesis $\sigma<1/6m$ y (7.18), basta mostrar que

$$\sum_{p|\Delta} p^{-2/3}\le O(\log^{1/3} R).$$

Pero existen a lo más $O(\log\Delta/\log\log\Delta)$ primos dividiendo $\Delta$, por lo quela suma al lado izquierdo puede ser acotada por

$$\sum_{1\le n\le O\left(\frac{\log\Delta}{\log\log\Delta}\right)} n^{-2/3}=O(\log^{1/3}\Delta),$$

cono requeríamos.

La cota en (7.17) ahora se sigue de la estimación $E_p^{(0)}=1+O_m(p^{-1/2})$. □

*Prueba de la proposición 7.4.* Aplicaremos el lema 7.6 con $\sigma=1/6m$ y $G=G_0G_1G_2$. Nuevamente por la regla de Leibnitz tenemos la cota (7.14), y además

$$||G||_{C^r(\mathcal{D}_\sigma^m)}\le O_m(1)O_{m,\omega(N)}(1)\left(\frac{\log R}{\log\log R}\right)^r\prod_{p|\Delta}\Big(1+O_m\big(p^{-1/2}\big)\Big),$$

para todo $0\le r\le m$. Del lema 7.8 y del lema 7.6 podemos estimar (7.8) siendo

$$\begin{aligned}(2\pi i)^{-2m}\int_{\Gamma_1}\cdots\int_{\Gamma_1}F(z,z')\prod_{j=1}^m\frac{R^{z_j+z_j'}}{z_j^2z_j'^2}dz_jdz_j' &\le \big(1+o_m(1)\big)\left(\frac{W}{\phi(W)}\right)^m\log^m R\prod_{p|\Delta}\big(1+O_m(p^{-1/2})\big)\\ &+O_{m,\omega(N)}\left(\frac{\log^m R}{\log\log R}\right)\prod_{p|\Delta}\big(1+O_m(p^{-1/2})\big)+O_m\Big(\exp\big(-\delta\sqrt{\log R}\big)\Big).\end{aligned}$$

Entonces la afirmación (7.2) se sigue eligiendo $\omega(N)$ (y por consiguiente $W$) con crecimiento lo suficientemente lento en $N$ (y por consiguiente en $R$). Así deducimos la proposición 7.4. □

# Capítulo 8

# Conclusiones

Consideramos algunas extensiones y refinamientos de la prueba dada sobre la existencia de progresiones aritméticas arbitrariamente largas formadas solamente por números primos. Notemos que la prueba dada muestra la existencia de una constante $\gamma(k)$ tal que el número de progresiones aritméticas de longitud $k$ constituidas por primos menores que $N$ es al menos $\Big(\gamma(k)+o(1)\Big)\dfrac{N^2}{\log^k N}$. Esto es porque el término de error en (6.12) no necesita ser del tipo $o(1)$ para garantizar que existen progresiones aritméticas, sino basta que sea menor que $\dfrac{1}{2}c(k,\delta)+o(1)$ por ejemplo. Esto permite que la expresión $\omega(N)$ considerada antes con crecimiento suficientemente lento puede ser considerada como una constante suficientemente grande solamente dependiendo de $k$ (e independiente de $N$).

El método empleado para probar el teorema 1.4 se puede extender para probar el teorema 1.5, teorema de Szemerédi en los primos, esto es, que cualquier subconjunto $A$ de los números primos con densidad (superior) relativa positiva contiene alguna progresión aritmética de longitud $k$. El único cambio significativo es que debemos usar el principio de las casillas para reemplazar la clase de residuos $n \equiv 1(\mathrm{mod}\, W)$, usada en el presente trabajo, por $n \equiv b(\mathrm{mod}\, W)$ para algún $b$ coprimo con $W$. Todo esto puesto que el conjunto $A$ no necesariamente obedece un teorema del tipo del teorema de Dirichlet, a diferencia del conjunto de los números primos que verifica el teorema del Dirichlet sobre primos en progresiones aritméticas. Además, puesto que suponemos que $A$ tiene densidad superior positiva $\delta(A)$, tenemos que existe una sucesión de naturales $N_1 < N_2 < \cdots$, no necesariamente primos, tales que $\lim\limits_{n\to+\infty}\dfrac{|A\cap[1,N_n]|}{|\mathcal{P}\cap[1,N_n]|}=\delta(A)$. Sin embargo, usando el postulado de Bertrand podemos

considerar los $N_n$ primos salvo un factor $O(1)$ a lo más.

Recientemente se probó más aún que dadas funciones polinomiales $F_1, F_2, \ldots, F_r$ de $\mathbb{N}$ en $\mathbb{N}$, tales que $F_i(0) = 0$, entonces existen infinitas $k$-uplas $\Big(a + F_1(d), \ldots, a + F_k(d)\Big)$ en las que cada coordenada es un número primo, (cf. [21]). Particularmente, si $F_i(x) = ix$ para cada $i \in \{0, 1, \ldots, k-1\}$, concluimos el teorema de Szemerédi en los primos. De este modo, este reciente resultado constituye una generalización del teorema de Green-Tao. También cabe mencionar que la validez de la conjetura de Erdös sobre progresiones aritméticas, conjetura 1.1, concluiría inmediatamente el teorema de Szemerédi en los primos puesto que la suma de las inversas de los primos es infinito.

# Bibliografía


[1] Gowers W.T., A new proof of Szemerédi's theorem for arithmetic progressions of length four, GAFA 8 (1998), 529-551.

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

[3] Roth K.F., On certain sets of integers, J. London Math. Soc. 28 (1953), 245-252.

[4] Granville, Andrew; An introduction to aditive combinatorics
`http://www.dms.umontreal.ca/~andrew/PDF/ProcAddPap.pdf`

[5] Gowers, Timothy; Additive and combinatorial number theory (2007)
`http://www.dpmms.cam.ac.uk/~wtg10/addnoth.notes.ps`

[6] Arbieto A., Matheus C. y Moreira C. G., Aspectos Ergódicos da Teoria dos Números, Publicações Matemáticas, IMPA, Brasil (2010).

[7] Green B. and Tao T., The primes contain arbitrarily long arithmetic progressions, Annals of Mathematics 167 (2008), 481-547.

[8] Goldston D. and Yıldırım C.Y., Small gaps between primes I, preprint available at
`http://www.arxiv.org/abs/math/0504336`

[9] Soundararajan K., Small gaps between prime numbers: the work of Goldston-Pintz-Yıldırım, Bulletin (New Series) of the American Mathematical Society, Volumen 44, Number 1, January 2007, Pages 1-18.

[10] Goldston D.A., Pintz J. and Yıldırım C. Y., The path to recent progress on small gaps between primes, Clay Mathematics Proceedings Volume 7, 2007

[11] Rudín W., Real and Complex Analysis, McGraw-Hill, New York, 1966.

[12] Titchmarsh E.C., The theory of the Riemann zeta function, Oxford University Press, 2nd ed, 1986.

[13] Apostol T.M., Introduction to Analytic Number Theory, Springer-Verlag, NewYork-Heidelberg-Berlin, 1976.

[14] Apostol T.M., Introducción a la Teoría analítica de números, Editorial Reverté S.A., 1984.

[15] Tao T. and Vu V., Additive combinatorics, Editorial Board, Cambridge studies in advanced mathematics 105.

[16] Petersen K., Ergodic theory, Cambridge University Press, Cambridge studies in advanced mathematics 2.

[17] Furstenberg H., Katznelson Y. and Ornstein D., The ergodic theoretical proof of Szemerédi's theorem, Bull. Amer. Soc. 7 (1982), 527-552.

[18] Gowers T., Hypergraph regularity and the multidimensional Szemerédi's theorem, preprint.

[19] Brezis H., Functional Analysis, Sobolev Spaces and Partial Differential Equations, Editorial Board, 2010.

[20] De Oliveira C.R., Introdução à análise funcional, Projeto Euclides, IMPA, 2010.

[21] Terence Tao and Tamar Ziegler, *The primes contain arbitrarily long polynomial progressions*, preprint available at
`http://arxiv.org/pdf/math/0610050v2.pdf`