UNIVERSIDAD DE BUENOS AIRES
Facultad de Ciencias Exactas y Naturales
Departamento de Matemática

Tesis de Licenciatura

PROGRESIONES ARITMÉTICAS EN SUBCONJUNTOS DE LOS PRIMOS

Miguel N. Walsh

Director: Pablo De Nápoli

Fecha de Presentación: Agosto 2010

Índice general

Introducción **4**
0.1. Objetivo del trabajo 4
0.2. Notación 5
0.3. Estructura de la tesis 7
0.4. Agradecimientos 8

1. Problemas sobre la distribución de los números primos **9**
1.1. La conjetura de Hardy-Littlewood 9
1.2. El estudio de aproximantes 12
1.3. Progresiones aritméticas en subconjuntos de los primos 15
1.4. Primos en intervalos pequeños 17

2. Estimaciones de cribas **20**
2.1. Cribas sobre intervalos 20
2.2. Comportamiento sobre formas lineales 26
2.3. Estimaciones de correlación 31

3. Integrales de contorno **36**
3.1. El caso unidimensional 36
3.2. El caso general 42

4. Distancias acotadas entre primos **44**
4.1. El nivel de distribución de los números primos 44
4.2. Densidad en tuplas 46
4.3. Demostración del Teorema 1.3 50
4.4. Demostración del Teorema 1.2 52

5. Medidas pseudoaleatorias **55**
5.1. El teorema de Szemerédi 55
5.2. Pseudoaleatoriedad 57
5.3. Construcción de medidas pseudoaleatorias 59
5.4. Deducción condicional del Teorema 1.1 60
5.5. Demostración de la condición de formas lineales 63
5.6. Demostración de la condición de correlación 65

6. La teoría de la norma de Gowers **70**
6.1. Las normas de Gowers 70
6.2. Teoría inversa de la norma de Gowers 74
6.3. Normas PCA 77

7. El principio de transferencia y el teorema generalizado de Von Neumann **82**
7.1. Descomposición y la norma CBA 82
7.2. Aproximaciones polinómicas 84
7.3. El teorema de transferencia 85
7.4. El teorema generalizado de Von Neumann 87
7.5. Demostración de la Proposición 5.3 91

8. Deducción de los corolarios **93**

Referencias **97**

Índice **101**

Introducción

0.1. Objetivo del trabajo

En el 2004, en uno de los trabajos más importantes de los últimos tiempos en teoría de números, Ben Green y Terence Tao demostraron que los números primos poseen progresiones aritméticas arbitrariamente largas, resolviendo así una de las conjeturas más antiguas sobre la distribución de los primos.

Lo notable del trabajo de Green y Tao radica no sólo en la importancia del resultado en sí, sino en los métodos que emplean para obtenerlo, introduciendo en su trabajo herramientas de áreas como la teoría ergódica y la combinatoria aditiva que abren nuevas puertas en la teoría analítica de números.

El objetivo de esta tesis es demostrar que los métodos de Green y Tao pueden adaptarse para encontrar progresiones aritméticas arbitrariamente largas en diversos subconjuntos de los primos. Un ejemplo de un conjunto al cual nuestro resultado se aplica son los primos de Chen (primos p para los cuales $p+2$ es el producto de a lo sumo dos primos distintos). En su trabajo, Green y Tao conjeturaron que sus métodos debían poder aplicarse para demostrar que los primos de Chen poseen las progresiones en cuestión y lo demostraron en particular para el caso de progresiones aritméticas de longitud tres. Más en general, dado cualquier conjunto de k enteros $h_1, \ldots, h_k$, demostraremos que existen progresiones aritméticas arbitrariamente largas tales que para cada miembro n de la progresión, el conjunto $n+h_1, \ldots, n+h_k$ acumula una cantidad pequeña de factores primos distintos.

Para muchos de los subconjuntos más interesantes de los primos, su propia infinitud es aún conjetural, por lo que no podemos aspirar a encontrar en ellos progresiones aritméticas de tamaño arbitrariamente grande. Un ejemplo de un tal conjunto son los primos gemelos. Veremos sin embargo que como consecuencia de nuestro resultado es posible obtener algún tipo de control sobre estos conjuntos en lo que a progresiones se refiere. Por ejemplo, veremos que si la densidad de los primos gemelos es al menos una fracción de los valores conjeturados, entonces necesariamente poseerán progresiones aritméticas arbitrariamente largas.

Además, aprovecharemos durante el trabajo para usar ideas recientes de Timothy Gowers que permiten dar demostraciones alternativas de varias proposiciones de Green y Tao, mediante la introducción de un teorema de transferencia abstracto y argumentos del análisis funcional.

Finalmente, las herramientas desarrolladas en esta tesis nos permitirán exponer los revolucionarios resultados de Dan Goldston, János Pintz y Cem Yildirim sobre espaciados pequeños entre los primos. En su trabajo de 2005, estos tres autores demostraron que existen primos consecutivos cuya distancia es arbitrariamente más pequeña que la distancia promedio, resolviendo de esta forma una vieja conjetura. Más aún, mostraron que asumiendo ciertas hipótesis generales es posible demostrar la existencia de una constante absoluta C tal que para infinitos primos p, $p + C$ también es primo. Una demostración de estos resultados será dada en la tesis. Además, veremos que aplicando el teorema de extensión a estos resultados, podremos garantizar la existencia de una constante c tal que para todo entero $r \geq 3$ habrá r primos $p_1, \ldots, p_r$ en progresion aritmética tales que $p_1 + c, \ldots, p_r + c$ será también una progresión aritmética de números primos. Más aún, asumiendo la conjetura de Elliot-Halberstam, podremos tomar $c \leq 20$.

0.2. Notación

Fijaremos ahora un poco de notación. En repetidas ocasiones en éste trabajo nos encontraremos con un parámetro entero N. En éste contexto escribiremos $O(1)$ para referirnos a una cantidad uniformemente acotada a lo largo de todos los valores de N y denotaremos por $o(1)$ a una cantidad que tiende a 0 cuando $N \to \infty$. Las constantes implícitas en estas notaciones podrán depender a veces de otros parámetros que se agregarán como subíndices cuando haya riesgo de confusión. Así, por ejemplo, $O_{c_1,c_2}(1)$ hará referencia a una magnitud acotada uniformemente por una constante $C(c_1, c_2) > 0$ que depende únicamente de c_1 y c_2. De todas formas la mayoría de las veces los parámetros en cuestión estarán fijos, en cuyo caso se tenderá a obviar los correspondientes subíndices para aliviar la notación. Asimismo, dada una magnitud X no negativa, abreviaremos $O(1)X$ por $O(X)$ y $o(1)X$ por $o(X)$.

Los símbolos k, k_i, l, l_i, r se referirán siempre a parámetros fijos, de modo que escribiremos $O_{k,k_i,l,l_i,r}(1) = O(1)$ y $o_{k,k_i,l,l_i,r}(1) = o(1)$. Además, en repetidas ocasiones usaremos el símbolo c para referirnos a una constante $c = c(k, k_i, l, l_i, r)$ cuyo valor puede cambiar en cada ocurrencia.

Denotaremos por $\mathbb{P}$ al conjunto de los números primos. Cuando consideremos $\mathbb{Z}_N$ se asumirá siempre que es $N \in \mathbb{P}$. La demostración de que tal elección es posible será llevada a cabo en los casos que se presten a confusión, aunque en general esto será trivial (ya sea por el postulado de Bertrand o simplemente por la infinitud de los números primos) y en consecuencia se omitirá aclaración. Ésto nos permitirá en particular invertir a todos los elementos de $\mathbb{Z}_N$.

Dado un conjunto no vacío A nos referiremos por $|A|$ a la cardinalidad de A. Dada una función $f : A \to \mathbb{R}$, escribiremos

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

y llamaremos al valor $\mathbb{E}(f)$ la esperanza de f. Escribiremos $\mathbb{E}_x(f)$ en lugar de $\mathbb{E}(f)$ para resaltar la variable de sumación en caso de que esto pueda prestarse a confusión.

Más en general, si $P(x)$ es cualquier proposición referida a un elemento de A que es verdadera para al menos un tal elemento, definimos la esperanza condicional

$$\mathbb{E}(f(x)|P(x)) := \frac{\sum_{\{x \in A : P(x) \text{ es cierta}\}} f(x)}{|\{x \in A : P(x)\}|}$$

y extendemos esta notación para varias variables en la forma obvia. Alternativamente, utilizaremos de ser conveniente la notación $\mathbb{E}_{P(x)} f(x)$ en lugar de $\mathbb{E}(f(x)|P(x))$.

Si B es un subconjunto de A escribiremos $\mathbf{1}_B : A \to \mathbb{R}$ para referirnos a la función que vale 1 si $x \in B$ y 0 en otro caso. Asimismo, si $P(x)$ es como antes, abreviaremos por $\mathbf{1}_{P(x)}$ a $\mathbf{1}_{\{x \in A : P(x)\}}$.

Para todo $1 \leq q < \infty$ y $f : \mathbb{Z}_N \to \mathbb{R}$ definimos las normas L^q como

$$\|f\|_{L^q} := \mathbb{E}\left(|f|^q\right)^{1/q}.$$

Definimos además la norma L^∞ de la forma clásica como

$$\|f\|_{L^\infty} := \sup_{x \in \mathbb{Z}_N} |f(x)|.$$

De esta forma, para todo $1 \leq q \leq \infty$ consideraremos el espacio de Banach $L^q\left(\mathbb{Z}^N\right)$ formado por todas las funciones de $\mathbb{Z}_N$ a $\mathbb{R}$ con la norma L^q. En particular, $L^2\left(\mathbb{Z}^N\right)$ será un espacio de Hilbert con el producto interno

$$\langle f, g \rangle := \mathbb{E}(fg).$$

En general, escribiremos

$$\langle f, g \rangle_{P(X)} := \frac{\sum_{x \in A : P(x)} f(x) g(x)}{|\{x \in A : P(x)\}|}.$$

Sea G un grupo abeliano finito. Dada una función $f : G \to \mathbb{C}$ definimos sobre el grupo dual $\hat{G}$ de caracteres sobre G la transformada de Fourier $\hat{f}$ de f mediante la fórmula

$$\hat{f}(\chi) = \mathbb{E}_x f(x) \overline{\chi(x)} = \mathbb{E}_x f(x) \chi(-x),$$

siendo entonces facilmente verificable la fórmula de inversión de Fourier

$$f(x) = \sum_{\chi \in \hat{G}} \hat{f}(\chi)\chi(x).$$

Dada una función $f : \mathbb{Z} \to \mathbb{R}$ y un parámetro entero N, podemos considerar la restricción $f|_{[1,N]} : [1, N] \cap \mathbb{Z} \to \mathbb{R}$ y de aquí podemos definir de la manera obvia una función $f|_{\mathbb{Z}_N} : \mathbb{Z}_N \to \mathbb{R}$. A lo largo de éste trabajo abusaremos de notación y llamaremos f a estas tres funciones. Asimismo, $f \in \mathbb{R}^N$ podrá referirse tanto a una función de $[1, N] \cap \mathbb{Z}$ en $\mathbb{R}$ como a una función de $\mathbb{Z}_N$ en $\mathbb{R}$, siendo el caso claro del contexto.

Decimos que $A \subseteq \mathbb{Z}$ tiene densidad (superior) positiva si

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

Más en general, decimos que $f : \mathbb{Z} \to \mathbb{R}$ tiene densidad (superior) positiva si

$$\limsup_{N\to\infty} \mathbb{E}_{n\leq N} f(n) > 0.$$

A la hora de evaluar una integral de contorno, utilizaremos el subíndice (a) para referirnos al contorno dado por los valores $a + iT$ con $T \in \mathbb{R}$. De esta forma, $\int_{(1)}$ hará referencia por ejemplo, a la integral de contorno a lo largo de la recta vertical de parte real igual a 1.

Por último, introducimos la siguiente definición que nos será de ayuda. Dados dos conjuntos A y B no vacíos, decimos que $\Phi : A \to B$ es un cubrimiento uniforme de B por A si Φ es sobreyectivo y para todo $b \in B$ es $\left|\left\{\Phi^{-1}(b) : b \in B\right\}\right| = |A|/|B|$. La utilidad de los cubrimientos uniformes así definidos radica en que si Φ es un cubrimiento uniforme de B por A entonces para toda función $f : B \to \mathbb{R}$ se tiene la identidad

$$\mathbb{E}_{a\in A} f(\Phi(a)) = \mathbb{E}_{b\in B} f(b).$$

0.3. Estructura de la tesis

El trabajo está organizado de la siguiente forma.

El Capítulo 1 provee una introducción más extensa a los problemas tratados. En él se repasa el contexto histórico y se introducen algunas definiciones. A su vez, se enuncian con precisión los resultados principales del trabajo y algunos de sus corolarios.

En el Capítulo 2 se demuestran tres importantes proposiciones que proveen información sobre la distribución de ciertas funciones que modelan el comportamiento de los números primos. Estas demostraciones serán condicionales a la evaluación de una familia de integrales de contorno que involucran a la

función zeta de Riemann. Estas evaluaciones serán el objetivo del Capítulo 3. Los métodos aquí empleados se basan en los argumentos utilizados en [17], [18] y [24].

En el Capítulo 4 aprovechamos las herramientas desarrolladas en los capítulos anteriores para demostrar los teoremas de Goldston, Pintz y Yildirim sobre espaciados pequeños entre los primos (ver [17], [18]). Éste capítulo es de exposición y no contiene resultados originales.

En el Capítulo 5 se introduce el concepto de medida pseudoaleatoria y se construye una familia de tales medidas que utilizaremos en el resto del trabajo. Para esto nos basamos en los métodos empleados en [24]. Además, se da una demostración del resultado principal de la tesis condicional a una importante proposición de Green y Tao.

El objetivo de los Capítulos 6 y 7 es dar una demostración de la mencionada proposición de Green y Tao. Para lograr tal objetivo se emprende el estudio de una cierta familia de normas. Además, se demuestra un teorema de transferencia abstracto utilizando argumentos recientes de Gowers (ver [21]) . Esto permite dar una demostración de la proposición en cuestión que presenta algunas ventajas respecto a la demostración original de Green y Tao. Estos capítulos no contienen resultados originales, aunque la presentación es novedosa.

Finalmente, en el Capítulo 8 se deducen varios corolarios del resultado principal.

0.4. Agradecimientos

A Pablo De Nápoli, director de esta tesis, por su tiempo, sus consejos y su excelente disposición.

A los jurados Ariel Pacetti y Román Sasyk, tanto por su tiempo, como por sus comentarios y consejos respectivamente.

A mis compañeros, Ernesto, Javi, Javier, Mauro, Pablo, Sergio y Tom.

A Gustavo Krimker, quien leyó mi primer trabajo cuando era un estudiante de escuela secundaria.

Especialmente a Natalio H. Guersenzvaig, quien con su conocimiento, experiencia y confianza en mí, contribuyó a desarrollar mi pasión por la matemática.

A mis padres Luisa y Eduardo y a mi tía Laura.

Capítulo 1

Problemas sobre la distribución de los números primos

1.1. La conjetura de Hardy-Littlewood

Indiscutiblemente, uno de los problemas abiertos más importantes de la Teoría de Números es el de establecer la conjetura de Hardy-Littlewood concerniente a la frecuencia de ciertos patrones en los números primos. Entre las innumerables consecuencias de tal resultado, se incluye por ejemplo la existencia de infinitos pares de primos gemelos, además de una estimación asintótica de su densidad. Más aún, una leve generalización de esta conjetura permite incluir entre sus implicaciones una fórmula para la cantidad de representaciones de un número par como la suma de dos números primos (conjetura de Goldbach) y para el total de números primos p debajo de una magnitud dada tales que $2p + 1$ es también primo (estos son los llamados primos de Sophie Germain) .

Para comprender la conjetura de Hardy-Littlewood es conveniente tener presente el modelo para la distribución de los primos propuesto por Harald Cramér [8]. El celebrado teorema del número primo nos dice que si $\pi(x)$ denota la cantidad de números primos debajo de x, entonces $\pi(x)$ es aproximadamente $\frac{x}{\log x}$. A su vez, la intuición sugiere que la primalidad de un número no debería afectar en general la probabilidad de otro número de ser también primo. Uniendo estas dos observaciones, Cramér propone que la función indicatriz de los números primos ha de comportarse como un conjunto de variables de Bernoulli $X(n)$ independientes con parámetros $\log^{-1} n$.

La vida, sin embargo, resulta no ser tan sencilla. Es evidente, por ejemplo, que la dupla n, $n+1$ consiste en un par de números primos únicamente cuando es $n = 2$ y sin embargo, el modelo de Cramér nos dice que habrán

de existir aproximadamente $\frac{x}{\log^2 x}$ valores $n \leq x$ para los cuales n y $n+1$ serán primos simultáneamente. Naturalmente, la dificultad reside en que tal dupla abarca todas las clases residuales módulo 2 y en consecuencia, para todo n, alguno de los miembros será divisible por 2. Por supuesto, lo mismo puede pasar con cualquier otro módulo, de modo que por ejemplo, la terna $n, n+2, n+4$ presentará un problema semejante puesto que por análogas razones poseerá siempre un miembro divisible por 3.

En 1932, Godfrey H. Hardy y John E. Littlewood conjeturaron que las únicas fluctuaciones al modelo de Cramér provienen de la cantidad de clases residuales ocupadas. Más concretamente, fijada una tupla $\mathcal{H} = \{h_1, \ldots, h_k\}$ de números enteros (asumiremos siempre $h_i \neq h_j$ si $i \neq j$) y un primo p consideramos el conjunto $\Omega(p)$ de las distintas clases residuales entre los $-h_i (\text{mod } p)$ y abusamos un poco de notación escribiendo $n \in \Omega(p)$ en lugar de $n (\text{mod } p) \in \Omega(p)$. Es evidente entonces que el conjunto $n+h_1, \ldots, n+h_k$ contendrá un múltiplo de p si y sólo si $n \in \Omega(p)$. Luego, la probabilidad de que ningún elemento de tal conjunto sea divisible por p es $\left(1 - \frac{|\Omega(p)|}{p}\right)$, mientras que la probabilidad de que k elementos independientes sean coprimos con p es $\left(1 - \frac{1}{p}\right)^k$.

Dado que el modelo de Cramér considera a los elementos de $n+h_1, \ldots, n+h_k$ como independientes, el razonamiento anterior nos sugiere que esto nos conducirá a una subestimación de la probabilidad de que éste conjunto no contenga múltiplos de p por un factor de $\left(1 - \frac{|\Omega(p)|}{p}\right)\left(1 - \frac{1}{p}\right)^{-k}$. Repitiendo éste argumento para todo primo p, concluimos la necesidad de introducir el factor de corrección

$$\mathfrak{G}(\mathcal{H}) = \prod_p \left(1 - \frac{|\Omega(p)|}{p}\right)\left(1 - \frac{1}{p}\right)^{-k}$$

a la estimación propuesta por Cramér. El valor $\mathfrak{G}(\mathcal{H})$ es conocido como la serie singular de la tupla $\mathcal{H}$. Como mencionamos, Hardy y Littlewood consideraron que esta ha de ser la única discrepancia entre el modelo de Cramér y la realidad, formulando en consecuencia la siguiente

Conjetura 1.1 (Conjetura de Hardy-Littlewood [30]). *Sea $\mathcal{H} = \{h_1, \ldots, h_k\}$, con $h_i \in \mathbb{Z}$. Entonces*

$$|\{n \leq x : n + h_1, \ldots, n + h_k \in \mathbb{P}\}| \sim \mathfrak{G}(\mathcal{H}) \frac{x}{\log^k x}.$$

Notar que cuando p es mayor que todos los elementos de $\mathcal{H}$, se tiene $\Omega(p) = k$, lo que causa que el producto que define a $\mathfrak{G}(\mathcal{H})$ sea siempre convergente. La pregunta relevante es entonces determinar cuándo es $\mathfrak{G}(\mathcal{H}) \neq 0$. Por un lado es claro que si es $\Omega(p) = p$ para algún primo la serie singular se anulará, siendo esto lo que sucedía en los ejemplos $\mathcal{H} = \{0, 1\}$ y $\mathcal{H} = \{0, 2, 4\}$

dados al comienzo. Por otro lado, si es $\Omega(p) \neq p$ para todo primo p, ningún factor del producto será nulo y para todo primo p suficientemente grande tendremos

$$\log\left(1-\frac{|\Omega(p)|}{p}\right)\left(1-\frac{1}{p}\right)^{-k} = \log\left(1-\frac{k}{p}\right)\left(1-\frac{1}{p}\right)^{-k} \sim -\frac{k(k-1)}{2p^2}$$

por lo que en tal caso la suma de los logaritmos converge y en consecuencia la serie singular es no nula. Concluimos entonces bajo la veracidad de la conjetura de Hardy-Littlewood que el conjunto $\{n+h_1,\ldots,n+h_k\}$ consistirá exclusivamente de primos para infinitos valores de n si y sólo si para la tupla $\mathcal{H}$ asociada es $|\Omega(p)| < p$ para todo primo p. Tales tuplas se llaman admisibles .

Para estudiar estos problemas, resulta natural la introducción de la función $\vartheta(n)$ que vale $\log n$ si n es primo y 0 en caso contrario. En realidad, puesto que la densidad de las potencias de primos es insignificante y por razones prácticas que serán mencionadas en breve, resulta más útil trabajar con la función de von Mangoldt $\Lambda(n)$ que vale $\log p$ si n es una potencia de un primo p y 0 en otro caso. En éste nuevo escenario, el teorema del número primo resulta equivalente a la estimación

$$\mathbb{E}(\Lambda(n)|n \leq x) = 1 + o_x(1).$$

Del mismo modo, obtenemos la siguiente estimación equivalente a la conjetura de Hardy-Littlewood:

Conjetura 1.2 (Conjetura de Hardy-Littlewood, de nuevo)**.** *Consideremos $\mathcal{H} = \{h_1,\ldots,h_k\}$, con $h_i \in \mathbb{Z}$. Entonces*

$$\mathbb{E}(\Lambda(n+h_1)\ldots\Lambda(n+h_k)|n \leq x) = (1+o_x(1))\mathfrak{G}(\mathcal{H}).$$

De éste modo, la función $\Lambda(n)$ nos permite ponderar a los primos (y conjeturalmente a las tuplas de primos) de modo tal que su densidad sea 1. De todas formas, la utilidad principal de la función de von Mangoldt proviene de satisfacer la convolución aritmética:

$$\Lambda(n) = \sum_{d|n} \mu(d)\log\frac{n}{d},$$

donde $\mu(n)$ es la función de Möbius cuyo soporte consiste en los números libres de cuadrados, valiendo -1 si n tiene una cantidad impar de factores primos y 1 en otro caso (en particular, $\mu(1) = 1$).

Esta identidad provee muchas ventajas, principalmente por su similitud a lo que es conocido como situación de criba. Para ver esto, considerar la pequeña modificación dada por

$$\Lambda_{(R)}(n) = \sum_{d|n} \mu(d) \log \frac{R}{d}.$$

Esta función tiene muchas similitudes con Λ en su comportamiento, la más relevante siendo que también posee como soporte a las potencias de los números primos. Debido a esta propiedad, el entendimiento del comportamiento de $\Lambda_{(R)}(n)$ resulta extremadamente útil para estudiar la distribución de los números primos.

1.2. El estudio de aproximantes

La principal ventaja de $\Lambda_{(R)}(n)$ sobre $\Lambda(n)$ radica obviamente en que el valor de la variable n desaparece de la suma interna, de modo que toda la información utilizada sobre n pasa a referirse estrictamente a la constitución del conjunto de sus divisores. De esta forma, cuando deseémos estudiar la suma de $\Lambda_{(R)}(n)$ sobre un intervalo I, muchos problemas técnicos que surgen al estudiar $\Lambda(n)$ se reducen a preguntas mucho más simples, como lo es por ejemplo la de estimar la cantidad de múltiplos en I de un divisor d dado.

A pesar de que la modificación mencionada de la función de von Mangoldt claramente nos permite posicionarnos mejor respecto a nuestras intenciones, aún surgen problemas. En particular, la cantidad de divisores d presentes en los intervalos de interés suele ser demasiado grande, volviendo inmanejables los errores propios de cualquier tipo de estimación utilizada. Éste obstáculo resulta ser crucial y la realidad es que no se conocen herramientas que permitan solucionar tal problema.

No todo está perdido sin embargo, aunque para poder avanzar será necesario sacrificar nuestro objetivo inicial de estudiar directamente el comportamiento de $\Lambda(n)$. En su lugar, lo estudiaremos indirectamente mediante la introducción de la truncación de la función de von Mangoldt dada por

$$\Lambda_R(n) = \sum_{\substack{d|n \\ d \leq R}} \mu(d) \log \frac{R}{d}.$$

De esta manera, nos concentraremos únicamente en el estudio de los d menores que R. Ajustando R de manera adecuada, esto nos permitirá obtener errores aceptables. El principio detrás de esta elección es que si R puede tomarse lo suficientemente grande, manteniendo el error manejable, $\Lambda_R(n)$ debería ser una aproximación razonable para $\Lambda(n)$.

Antes de especificar de qué manera las estimaciones de $\Lambda_R(n)$ resultan ser extremadamente útiles para el estudio de los números primos, debemos

recordar que nuestra intención original era el estudio de tuplas de números en forma simultánea. Obviamente, lo más natural sería considerar para una tupla $\mathcal{H} = \{h_1, \ldots, h_k\}$ el producto $\Lambda(n+h_1)\ldots\Lambda(n+h_k)$. Sin embargo, el trabajar con k valores simultáneos de Λ potencia demasiado los errores, que ya notamos significativos en el caso $k = 1$.

Para solucionar esto, introducimos una nueva modificación de la función de von Mangoldt dada por

$$\Lambda(n; \{0\}, r) = \frac{1}{r!} \sum_{d|n} \mu(d) \left(\log \frac{n}{d}\right)^r.$$

Aquí $1/r!$ es un factor de normalización que puede ser ignorado. Así como $\Lambda(n)$ se caracterizaba por tener como soporte a las potencias de primos, $\Lambda(n; \{0\}, r)$ es una función notable por tener como soporte a los números con a lo sumo r factores primos distintos. Esto sugiere inmediatamente el camino a seguir. Dada una tupla $\mathcal{H} = \{h_1, \ldots, h_k\}$ consideramos el polinomio $P(n; \mathcal{H}) = (n+h_1)\ldots(n+h_k)$ y escribimos

$$\Lambda(n; \mathcal{H}, k+l) = \frac{1}{(k+l)!} \sum_{d|P(n;\mathcal{H})} \mu(d) \left(\log \frac{n}{d}\right)^{k+l}.$$

Por los comentarios anteriores, vemos que el soporte de esta función consistirá de los n tales que el conjunto $\{n+h_1, \ldots, n+h_k\}$ acumule a lo sumo $k+l$ factores primos distintos. En particular, tomando $l = 0$ el soporte de n consistirá en los n para los cuales tal conjunto consiste enteramente de potencias de primos, cuya densidad notamos anteriormente difería insignificantemente de la de aquellos n para los cuales el conjunto está formado estrictamente por números primos. La razón por la cual se incluye el parámetro adicional l es debido a que en muchos casos, es más fácil trabajar con los casi-primos (números con una cantidad pequeña de factores primos) en lugar de trabajar con los primos directamente.

Para poder estudiar esto consideramos como antes la alteración

$$\Lambda_{(R)}(n; \mathcal{H}, k+l) = \frac{1}{(k+l)!} \sum_{d|P(n;\mathcal{H})} \mu(d) \left(\log \frac{R}{d}\right)^{k+l}.$$

Dada esta modificación, no es difícil verificar las desigualdades

$$\Lambda_{(R)}(n; \mathcal{H}, k+l) \left(\frac{\log P(n; \mathcal{H})}{\log R}\right)^{k+l} \leq \Lambda(n; \mathcal{H}, k+l) \leq \Lambda_{(R)}(n; \mathcal{H}, k+l),$$

de donde concluimos que el soporte de esta función consiste también en los números con a lo sumo $k+l$ factores primos distintos. Podemos entonces considerar, como lo hicimos para la función de von Mangoldt, la truncación

dada por

$$\Lambda_R(n;\mathcal{H},k+l)=\frac{1}{(k+l)!}\sum_{\substack{d|P(n;\mathcal{H})\\ d\leq R}}\mu(d)\left(\log\frac{R}{d}\right)^{k+l}.$$

Nuevamente, uno esperaría que para una elección adecuada de R esta convolución imite en muchos aspectos el comportamiento de $\Lambda(n;\mathcal{H},k+l)$ mientras el error se mantiene controlado.

Las funciones del tipo mencionado arriba fueron estudiadas intensamente en la última década. En un revolucionario trabajo, Dan Goldston, Janos Pintz y Cem Yildirim [18] establecieron estimaciones asintóticas para la función $\Lambda_R(n;\mathcal{H},k+l)$ como parte de su programa para encontrar espaciados arbitrariamente pequeños entre los primos. Para lograr tal propósito, transformaron tal problema en la evaluación de una integral de contorno que involucra la función zeta de Riemann. Gracias a esto, conocemos ahora en gran medida el comportamiento de una función que suponemos similar a $\Lambda(n;\mathcal{H},k+l)$. Pero dado que al fin y al cabo no podemos probar efectivamente tal similitud, ¿es posible utilizar esta información para obtener resultados concretos sobre la distribución de los números primos? Por supuesto, no haríamos esta pregunta si la respuesta no fuese afirmativa.

En el 2004, en lo que es sin duda alguna uno de los trabajos más importantes de las últimas decadas en la teoría de números, Ben Green y Terence Tao [24] establecieron la existencia de progresiones aritméticas arbitrariamente largas en los primos. Es decir, que para todo $r\in\mathbb{N}$ existirán $x,h\in\mathbb{N}$ tales que $x+jh$ es un número primo para todo $0\leq j\leq r-1$. Tan impresionante como el resultado en sí, es la cantidad de oxígeno que introducen a la teoría de números bajo la forma de herramientas provenientes de diversas áreas de la matemática, particularmente de la teoría ergódica. Gracias a estas notables ideas, consiguen así sacar provecho de la distribución de $\Lambda_R(n)$ para probar su impactante resultado sobre la distribución de los números primos.

Para comprender mejor tal trabajo, es necesario trasladarse al área de la matemática que hoy en día se conoce como combinatoria aditiva (aunque considerando su expansión a preguntas que no son estrictamente aditivas, se ha sugerido que el nombre combinatoria aritmética resultaría más apropiado). En 1973, Endre Szemerédi resolvió una famosa conjetura de Paul Erdös y Paul Turán, estableciendo así un resultado fundacional para la combinatoria aditiva [42]. El teorema de Szemerédi nos dice que todo subconjunto de los enteros de densidad positiva posee progresiones aritméticas arbitrariamente largas. Mientras que la demostración de Szemerédi es esencialmente combinatoria, con el tiempo éste histórico resultado sería demostrado nuevamente en diversos contextos. En 1977, Hillel Furstenberg obtuvo el resultado mediante métodos de teoría ergódica [15], dando lugar así a una relación entre preguntas de la combinatoria aditiva y la teoría ergódica que se volvería

extremadamente productiva con los años. Asimismo, en 1998, Timothy Gowers logró dar una nueva prueba del teorema de Szemerédi, esta vez utilizando análisis armónico [19]. Es en ese trabajo en el que Gowers introdujo las normas U^k de las cuales hoy se sabe que un conocimiento adecuado permitiría dar un paso enorme hacia la conjetura de Hardy-Littlewood.

Incorporando las ideas principales de estos tres trabajos, Green y Tao observaron que el resultado podía extenderse más allá de los conjuntos con densidad positiva. Para esto introducen una noción adecuada de pseudoaleatoriedad. Mientras que es bastante sencillo demostrar que un conjunto pseudoaleatorio posee progresiones aritméticas arbitrariamente largas, Green y Tao van mucho más lejos y establecen que bastará con que el conjunto en cuestión esté dominado por otro pseudoaleatorio y de densidad semejante, para garantizar que posea las deseadas progresiones aritméticas. Es aquí dónde las estimaciones para $\Lambda_R(n)$ entran en juego.

Como vimos anteriormente, en el estudio de los primos la función más útil para caracterizarlos suele ser la función de von Mangoldt $\Lambda(n)$. El objetivo sería entonces poder encontrar una función pseudoaleatoria que mayorice a $\Lambda(n)$ y posea una densidad similar. Es fácil ver que una leve modificación de $\Lambda_R(n)$ funciona de mayorante para $\Lambda(n)$ al mismo tiempo que retiene una frecuencia similar. ¿Podremos demostrar entonces que $\Lambda_R(n)$ es una función pseudoraleatoria, implicando así la existencia de progresiones aritméticas arbitrariamente largas en los primos?

En un principio, la respuesta es no. La función de von Mangoldt, al igual que los números primos, posee la peculiaridad de encontrarse irregularmente distribuida a lo largo de las diversas clases residuales. Esto se debe por supuesto a que los números primos sólo pueden aparecer infinitas veces en las clases residuales que son coprimas con el módulo. Es fácil de probar que una función con una tal irregularidad no puede ser mayorada por una función pseudoaleatoria, afectando esto nuestras esperanzas de obtener una tal mayorización para $\Lambda(n)$. Sin embargo, éste obstáculo resulta no ser esencial. Mediante la introducción de lo que ellos llamaron el truco-W, Green y Tao consiguen construir una variación de $\Lambda(n)$ que sí está distribuida en forma suficientemente regular. Aplicando entonces el mismo truco al mayorante $\Lambda_R(n)$ e implementando las herramientas desarrolladas por Goldston y Yildirim, Green y Tao demuestran que la modificación de $\Lambda_R(n)$ en cuestión define efectivamente una función pseudoaleatoria obteniendo así su resultado sobre progresiones aritméticas en los primos.

1.3. Progresiones aritméticas en subconjuntos de los primos

El resultado principal de esta tesis es una extensión del teorema de Green-Tao. Dado $\delta > 0$ decimos que $f : \mathbb{Z}_N \to \mathbb{R}_{\geq 0}$ es δ-denso si satis-

face $\mathbb{E}(f(n)|\mathbb{Z}_N) \geq \delta$. Asimismo, decimos que una tal f tiene soporte en $\mathcal{A} \subseteq \mathbb{Z}$ si su soporte está contenido en la proyección canónica de $\mathcal{A} \cap [1, N]$ en $\mathbb{Z}_N$. Por último, definimos la normalización

$$\hat{\Lambda}_R(n; \mathcal{H}, k+l) = \frac{\Lambda_R^2(n; \mathcal{H}, k+l)}{\log^{k+2l} R}.$$

Dadas estas definiciones, el enunciado concreto de nuestro resultado es el siguiente:

Teorema 1.1. *Sea $\mathcal{A} \subseteq \mathbb{N}$. Sea $\delta > 0$ arbitrario y $r \geq 3$ un entero también arbitrario. Si para todo N suficientemente grande existe $f_N : \mathbb{Z}_N \to \mathbb{R}_{\geq 0}$ δ-denso y de soporte en $\mathcal{A}$ con*

$$f_N(n) \leq \hat{\Lambda}_R(n; \mathcal{H}, k+l)$$

en $\mathbb{Z}_N$, para $R = N^{r^{-1}2^{-r-4}}$ y cierta elección fija de $\mathcal{H} = \{h_1, \ldots, h_k\}$ y $l > 0$, entonces $\mathcal{A}$ posee progresiones aritméticas de longitud r.

El principal aporte de éste resultado es la demostación de que para todo $\mathcal{H} = \{h_1, \ldots, h_k\}$ y $l > 0$, existe una adaptación apropiada de $\Lambda_R(n; \mathcal{H}, k+l)$ que define una función pseudoaleatoria. Utilizaremos además una observación reciente de Gowers [21], que nos permitirá obtener demostraciones más directas de algunos de los resultados de Green y Tao mediante la implementación de herramientas del análisis funcional, particularmente el teorema de Hahn-Banach.

Obtendremos varias consecuencias del Teorema 1.1, algunas de las cuales enunciamos a continuación. En primer lugar, tomando $\mathcal{H} = \{0\}$, $l = 1$ y $f_N = \mathbf{1}_{\varepsilon N \leq n \leq 2\varepsilon N} c\Lambda(n)$ para ciertas constantes $c, \varepsilon > 0$ suficientemente pequeñas es posible recuperar el

Corolario 1.1 (Teorema de Green-Tao)**.** *Los primos contienen progresiones aritméticas arbitrariamente largas.*

En su monumental trabajo de 1973, Chen Jingrun demostró que para todo $h \in \mathbb{Z}$ existen infinitos primos p tales que $p + 2h$ es el producto de a lo sumo dos factores primos distintos [7]. Llamaremos h-primos de Chen a los que satisfagan la mencionada propiedad. Veremos que del teorema de Chen podemos deducir el siguiente corolario:

Corolario 1.2. *Existen progresiones aritméticas arbitrariamente largas de h-primos de Chen para todo entero h.*

En particular, cuando es $h = 1$ se los suele llamar simplemente primos de Chen. Green y Tao habían predecido en su trabajo que sus métodos debían poder adaptarse para demostrar que los primos de Chen contienen progresiones aritméticas arbitrariamente largas y lo demostraron en particular para el caso de progresiones aritméticas de longitud 3 (ver [23]). El

resultado general fue demostrado en 2009, en forma independiente a éste trabajo, por Binbin Zhou [48].

Asumiendo la conjetura de Hardy-Littlewood es posible deducir también del Teorema 1.1 que para toda tupla $\mathcal{H} = \{h_1, \dots, h_k\}$ con $\mathfrak{G}(\mathcal{H}) \neq 0$ y para todo entero r, existirán r números en progresión aritmética $n_1, \dots, n_r$, tales que las tuplas $\{n_i + h_1, \dots, n_i + h_k\}$ consistirán enteramente de primos $\forall$ $1 \leq i \leq r$. En realidad, no es necesario obtener la estimación asintótica precisa dada por la conjetura de Hardy-Littlewood sino que bastará únicamente con que la densidad sea del orden correcto. El enunciado preciso es el siguiente:

Corolario 1.3. *Sea $\mathcal{H} = \{h_1, \dots, h_k\}$, con $h_i \in \mathbb{Z}$ para todo $1 \leq i \leq k$. Si es*

$$\mathbb{E}(\Lambda(n + h_1) \dots \Lambda(n + h_k) | N/2 \leq n < N) \geq c + o(1),$$

para cierta constante $c = c(\mathcal{H}) > 0$, entonces el conjunto de tuplas primas $\mathbb{P}_{\mathcal{H}} = \{n \in \mathbb{N} | n + h_1, \dots, n + h_k \in \mathbb{P}\}$ contiene progresiones aritméticas arbitrariamente largas.

En particular, si la cantidad de primos gemelos debajo de x es $\gg \frac{x}{\log^2 x}$ entonces habrá progresiones aritméticas arbitrariamente largas de primos gemelos. Es interesante notar que lo que se está afirmando en el Corolario 1.3 es que podemos obtener las progresiones deseadas asumiendo únicamente una hipótesis de densidad.

Incondicionalmente, tenemos el siguiente resultado.

Corolario 1.4. *Sea $\mathcal{H} = \{h_1, \dots, h_k\} \subseteq \mathbb{Z}$ una tupla admisible. Existen entonces progresiones aritméticas arbitrariamente largas tales que, para cada miembro n de la progresión, la tupla $\{n + h_1, \dots, n + h_k\}$ acumula a lo sumo $3k \log k$ factores primos distintos.*

1.4. Primos en intervalos pequeños

Las aplicaciones de conocer la distribución de $\Lambda_R(n; \mathcal{H}, k+l)$ no terminan aquí. Una pregunta fundamental sobre la distribución de los números primos es su densidad en intervalos. En particular, la estimación de las posibles diferencias entre primos consecutivos representa un problema de notable importancia.

En una dirección, el teorema del número primo nos da la desigualdad

$$\limsup_{n \to \infty} \frac{p_{n+1} - p_n}{\log p_n} \geq 1$$

mientras que del modelo de Cramér (como también de la conjetura de Hardy-Littlewood, gracias a una serie de brillantes deducciones llevadas a cabo por Gallagher [16]), uno espera

$$\limsup_{n \to \infty} \frac{p_{n+1} - p_n}{\log p_n} = \infty.$$

Esta suposición es de hecho correcta y fue demostrada por Erik Westzynthius [45] en 1931. Subsecuentemente el valor del límite superior fue estimado de modo más preciso por Erdös [10] y por Robert Rankin [37], quien demostró la existencia de una constante c tal que para infinitos primos p_n se tiene

$$p_{n+1} - p_n > c \log p_n \frac{(\log \log p_n) \log \log \log \log p_n}{(\log \log \log p_n)^2},$$

dando así la mejor estimación conocida. Quizás una medida de la relevancia de éste problema es el curioso hecho de que Erdös haya ofrecido 10,000 dólares por cualquier mejora en el orden de esta cota, siendo este premio el más alto jamás ofrecido por Erdös y 10 veces superior al mayor premio que llegó a entregar (fue justamente Szemerédi quien recolectó aquel premio con la demostración de su teorema ya mencionado).

La otra dirección sea quizás incluso más interesante puesto que óptimamente uno espera poder obtener

$$\liminf_{n\to\infty} p_{n+1} - p_n = 2,$$

puesto que esto es obviamente equivalente a la conjetura de los primos gemelos. Sin embargo, tal objetivo aparece en el presente fuera del alcance. Nuevamente, el teorema del número primo nos dice que

$$\liminf_{n\to\infty} \frac{p_{n+1} - p - n}{\log p_n} \leq 1.$$

Éste resultado fue mejorado en diversos trabajos por Erdös [11], Enrico Bombieri y Harold Davenport [3], Martin Huxley [31] y particularmente Helmut Maier [34], cuyo trabajo resulta muy interesante puesto que para lograr su resultado saca ventaja de la curiosa variación existente entre lo que el modelo de Cramér predice y lo que en realidad sucede en los intervalos pequeños.

De todas formas no fue hasta el 2005 que Dan Goldston, János Pintz y Cem Yildirim en un impresionante trabajo [18] obtuvieron el siguiente

Teorema 1.2 (Goldston-Pintz-Yildirim)**.** *Es*

$$\liminf_{n\to\infty} \frac{p_{n+1} - p_n}{\log p_n} = 0.$$

Los métodos utilizados para obtener éste resultado poseen muchos puntos en común con las herramientas empleadas para deducir el Teorema 1.1, por lo que aprovecharemos la ocasión para incluir una demostración de éste teorema. En éste trabajo, los autores sacan un notable provecho de conocer el comportamiento de $\Lambda_R(n; \mathcal{H}, k+l)$. Como hemos visto con anterioridad, uno esperaría que esta función se comporte en forma similar a $\Lambda(n; \mathcal{H}, k+l)$, que tiene como soporte a las tuplas $\{n + h_1, \ldots, n + h_k\}$ que tienen a lo sumo

$k+l$ factores primos distintos. Debido a esto, uno esperaría que si es $h \in \mathcal{H}$ entonces $\Lambda(n+h)$ debería tener una densidad relativa significativa respecto a $\Lambda_R(n; \mathcal{H}, k+l)$. Precisamente, uno esperaría que para valores altos de N

$$\langle \Lambda(n+h), \Lambda_R(n; \mathcal{H}, k+l) \rangle_{N \leq n \leq 2N}$$

sea particularmente grande. Resulta que efectivamente tal cantidad puede ser calculada, en parte gracias a las mencionadas estimaciones para la función $\Lambda_R(n; \mathcal{H}, k+l)$, pero también gracias a una poderosa herramienta de la teoría analítica de números conocida como el teorema de Bombieri-Vinogradov (una especie de hipótesis generalizada de Riemann en promedio, que puede ser obtenida incondicionalmente). Esta capacidad para localizar a los números primos dentro de una tupla, es la herramienta fundamental que permite demostrar el Teorema 1.2.

A pesar de su terrible poder, el teorema de Bombieri-Vinogradov no se supone óptimo. Éste teorema nos dice que el nivel de distribución de los primos en progresiones aritméticas es $\theta = 1/2$, siendo intencionalmente vagos sobre lo que esto significa (ver §4.1). Sin embargo, se cree en general que éste nivel de distribución es en realidad tan cercano a 1 como se desee. Esta afirmación es conocida como la conjetura de Elliot-Halberstam. Asumiendo tal hipótesis, los métodos desarrollados por Goldston, Pintz y Yildirim se extienden para demostrar el siguiente fantástico resultado:

Teorema 1.3 (Goldston-Pintz-Yildirim)**.** *Supongamos que los primos tienen nivel de distribución $\theta > 1/2$. Entonces existe una constante $C = C(\theta)$ tal que toda tupla $\mathcal{H} = \{h_1, \ldots, h_k\}$ con $k \geq C$ y $\mathfrak{G}(\mathcal{H}) \neq 0$ posee infinitos valores de n para los cuales el conjunto $\{n+h_1, \ldots, n+h_k\}$ contiene al menos dos números primos. En particular, asumiendo la conjetura de Elliot-Halberstam podemos tomar $C = 7$, por lo cual, aplicando el resulado a la tupla $\{0, 2, 6, 8, 12, 18, 20\}$ obtenemos:*

$$\liminf_{n \to \infty} p_{n+1} - p_n \leq 20.$$

Una demostración de esta afirmación será incluida en esta tesis. Aplicando el Teorema 1.1 a éste resultado, vamos a obtener el siguiente

Corolario 1.5. *Supongamos que los primos tienen nivel de distribución $\theta > 1/2$. Existe entonces una constante $C = C(\theta)$ tal que para toda tupla $\mathcal{H} = \{h_1, \ldots, h_k\}$ con $k \geq C$ y $\mathfrak{G}(\mathcal{H}) \neq 0$ existe un par i, j con $1 \leq i < j \leq k$ tal que el conjunto $\mathbb{P}_{ij} = \{p \in \mathbb{P} \mid p + (h_j - h_i) \in \mathbb{P}\}$ contiene progresiones aritméticas arbitrariamente largas. En particular, asumiendo la conjetura de Elliot-Halberstam, existe una constante $0 \leq c \leq 20$ tal que para todo entero $r > 0$ existen r primos $p_1, \ldots, p_r$ en progresión aritmética para los cuales la progresión aritmética $p_1 + c, \ldots, p_r + c$ también está formada en su totalidad por números primos.*

Capítulo 2

Estimaciones de cribas

Sea $\mathcal{H} = \{h_1, \dots, h_k\} \subseteq \mathbb{Z}$. A lo largo de éste capítulo asumiremos siempre que para toda tal tupla $\mathcal{H}$ se tiene $h_i \neq h_j$ para todo par $i \neq j$. Más aún, supondremos también en todos los casos que las tuplas son admisibles, en el sentido de que la serie singular $\mathfrak{G}(\mathcal{H})$ es no nula.

En éste capítulo se demostrarán tres importantes proposiciones que caracterizan el comportamiento de $\Lambda_R(n; \mathcal{H}, k+l)$. Las dos primeras proposiciones estudiarán la distribución de esta función a lo largo de intervalos y de formas lineales. La tercera proposicion nos dará una estimación de correlación. Todas las demostraciones serán condicionales a la evaluación de una integral de contorno que será efectuada en el siguiente capítulo. Los métodos aquí empleados se basan en los argumentos utilizados en [17], [18] y [24].

En todos los casos los parámetros k, k_i, l, l_i se suponen fijos y en consecuencia se omitirán los correspondientes subíndices en las notaciones O, o, en concordancia con lo mencionado en §0.2.

2.1. Cribas sobre intervalos

El primer objetivo será obtener una estimación para la esperanza de $\Lambda_R(n; \mathcal{H}, k+l)$ en intervalos de la forma $[N, 2N]$. Esto fue logrado en [17], [18]. La estrategia consiste en reducir el problema en cuestión a la estimación de una integral de contorno que involucra a la función zeta de Riemann. Tal integral será evaluada más adelante. El enunciado es el siguiente:

Proposición 2.1 ([17], Lema 1; [18], Prop. 1)**.** *Sean $k, l > 0$ enteros y $\mathcal{H} = \{h_1, \dots, h_k\} \subseteq [1, H] \cap \mathbb{Z}$. Sean N, R parámetros, con*

$$H \ll \log N \ll \log R. \tag{2.1}$$

Suponiendo además

$$R \leq N^{1/2}/(\log N)^C \tag{2.2}$$

para una cierta constante C suficientemente grande que depende sólo de k y l, tenemos

$$\mathbb{E}_{N<n\leq 2N}\Lambda_R^2(n;\mathcal{H},k+l)=(\mathfrak{G}(\mathcal{H})+o(1))\frac{1}{(k+2l)!}\binom{2l}{l}(\log R)^{k+2l},$$

donde la constante implícita depende únicamente de k, l y las constantes implícitas en (2.1) .

Nota. Si bien las estimaciones del orden de H respecto a N son superfluas a la hora de estudiar una tupla $\mathcal{H}$ fija, esto nos permitirá estudiar más adelante a todas las k-tuplas del intervalo $[1,H]$ en forma simultánea.

Nota. El hecho de evaluar el cuadrado de $\Lambda_R(n;\mathcal{H},k+l)$ se debe a que lo que en realidad nos interesa es la variación en magnitud de esta función y más aún, a la hora de estudiar la distribución de los números primos, se vuelve de fundamental importancia que el peso en cuestión pueda tomarse no negativo. Más en general, el estudio del cuadrado de una función en lugar de la función misma dentro de la teoría de cribas tiene su orígen en los trabajos de Atle Selberg sobre la función zeta de Riemann. Allí Selberg notó que esto permitía simplificar la optimización de parámetros.

Demostración. Usaremos el símbolo c para referirnos a una constante positiva que dependa únicamente de k y l, pero cuyo valor puede variar en cada caso. Para todo primo p denotamos por $\Omega(p)$ al conjunto de diferentes clases residuales entre $-h(\mathrm{mod}\ p)$, $h\in\mathcal{H}$ y abusamos un poco de notación escribiendo $n\in\Omega(p)$ en lugar de $n(\mathrm{mod}\ p)\in\Omega(p)$. Recordemos que estamos asumiendo la admisibilidad de $\mathcal{H}$ y que esto es equivalente a la desigualdad $|\Omega(p)|<p$ para todo p. Extendemos a Ω multiplicativamente sobre los enteros libres de cuadrados[1], de modo que es $n\in\Omega(d)$ si y sólo si es $n\in\Omega(p)$ para todo divisor primo p de d. En efecto, para $(d,d')=1$ obtenemos

$$|\Omega(dd')|=|\Omega(d)||\Omega(d')|$$

por el teorema chino del resto. Escribimos $P(n;\mathcal{H})=(n+h_1)\dots(n+h_k)$ y observamos que $n\in\Omega(d)$ si y sólo si $d|P(n;\mathcal{H})$.

Consideramos la función

$$\lambda_R(d;k+l):=\begin{cases}\frac{1}{(k+l)!}\mu(d)\left(\log\frac{R}{d}\right)^{k+l} & \text{si } d\leq R,\\ 0 & \text{si } d>R,\end{cases}\tag{2.3}$$

donde μ es la función de Möbius. Con esta definición, obtenemos para

[1]Notar que a lo largo de esta sección sólo necesitaremos tratar los casos de enteros d libres de cuadrados debido a la aparición de la función de Möbius.

$\Lambda_R(n;\mathcal{H},k+l)$ la expresión

$$\begin{aligned}\Lambda_R(n;\mathcal{H},k+l) &= \frac{1}{(k+l)!}\sum_{\substack{d|P(n;\mathcal{H})\\ d\leq R}}\mu(d)\left(\log\frac{R}{d}\right)^{k+l}\\ &= \sum_{d:n\in\Omega(d)}\lambda_R(d;k+l).\end{aligned}$$

Expandiendo el cuadrado tenemos entonces que $\mathbb{E}_{N<n\leq 2N}\Lambda_R^2(n;\mathcal{H},k+l)$ es igual a

$$\sum_{d,d'}\lambda_R(d;k+l)\lambda_R(d';k+l)\mathbb{E}_{N<n\leq 2N}\left(\mathbf{1}_{d,d'|P(n;\mathcal{H})}\right). \tag{2.4}$$

Escribiendo $[d,d']$ para el mínimo común múltiplo de d y d', es claro que la esperanza interna es igual a

$$\mathbb{E}_{N<n\leq 2N}\left(\mathbf{1}_{n\in\Omega([d,d'])}\right) = \frac{|\Omega([d,d'])|}{[d,d']} + O\left(\frac{|\Omega([d,d'])|}{N}\right).$$

Teniendo en cuenta esto, vemos que (2.4) es igual a

$$\mathcal{T} + O\left(\frac{1}{N}\sum_{d,d'}|\Omega([d,d'])|\lambda_R(d;k+l)\lambda_R(d';k+l)\right), \tag{2.5}$$

con

$$\mathcal{T} := \sum_{d,d'}\frac{|\Omega([d,d'])|}{[d,d']}\lambda_R(d;k+l)\lambda_R(d';k+l). \tag{2.6}$$

Debido a la multiplicatividad de $\Omega(d)$ tenemos la desigualdad $|\Omega([d,d'])| \leq |\Omega(d)|\,|\Omega(d')|$, con esto, podemos reescribir (2.5) como

$$\mathcal{T} + O\left(\frac{1}{N}\left(\sum_{d}|\Omega(d)|\,|\lambda_R(d;k+l)|\right)^2\right). \tag{2.7}$$

Sea $\tau_r(d)$ la cantidad de representaciones de d como el producto de r números naturales. Para proseguir necesitaremos el siguiente resultado bien conocido que nos será de utilidad en repetidas ocasiones. Para su demostración, ver [33, Capítulo 1].

Lema 2.1. *Sean $r,s>0$ enteros arbitrarios. Entonces*

$$\mathbb{E}_{d\leq x}(\tau_r(d)^s) \ll_{r,s} (\log x)^{r^s-1}.$$

Supongamos ahora que d es un entero positivo libre de cuadrados que divide a $P(n;\mathcal{H}) = (n+h_1)\dots(n+h_k)$. Podemos escribir entonces $d = d_1,\dots,d_k$ con $d_i|(n+h_i)$. Si a su vez $d|P(m;\mathcal{H})$ para cierto m que da a lugar a la misma descomposición $d_1,\dots,d_k$ de d (es decir, tal que $d_i|(m+h_i)$ para

todo i), se sigue del teorema chino del resto que ha de ser $n \equiv m (\mathrm{mod}\ d)$. Concluimos así la desigualdad $|\Omega(d)| \leq \tau_k(d)$. Aplicando entonces el Lema 2.1 a (2.7) y utilizando que es $|\lambda_R(d; k+l)| \ll (\log R)^c$ si es $d \leq R$ y $\lambda_R(d; k+l) = 0$ en otro caso, obtenemos

$$\begin{aligned}\mathbb{E}_{N<n\leq 2N}\Lambda_R^2(n; \mathcal{H}, k+l) &= \mathcal{T} + O\left(\frac{R^2 (\log R)^c}{N}\right) \\ &= \mathcal{T} + o(1)\end{aligned}$$

en donde hemos usado la hipótesis (2.2).

Sea ahora $\log^+ x = \max(\log x, 0)$. Tenemos entonces la bien conocida identidad integral

$$(\log^+ x)^r = \frac{r!}{2\pi i}\int_{(1)} x^s \frac{ds}{s^{r+1}}. \tag{2.8}$$

Observar además que es $\lambda_R(d; k+l) = \frac{1}{(k+l)!}\mu(d)(\log^+ \frac{R}{d})^{k+l}$. La idea será entonces sacar provecho de esta identidad para evaluar $\mathcal{T}$. Notar de (2.6), expandiendo $\lambda_R(d; k+l)$, que todas las funciones de d involucradas son multiplicativas en d, excepto justamente $(\log^+ \frac{R}{d})^{k+l}$. Sin embargo, se da el notable hecho de que (2.8) nos permite expresar a esto como la integral de una función que también es multiplicativa en d. Es esto lo que nos permitirá llevar el problema a un producto infinito sobre los primos y poder aprovechar que sabemos manejar tales productos relativamente bien gracias a la teoría de la función zeta de Riemann (y dado que después de todo, un contexto multiplicativo es mucho más natural para los números primos).

Utilizando entonces (2.8) tenemos

$$\lambda_R(d; k+l) = \frac{\mu(d)}{2\pi i}\int_{(1)} \left(\frac{R}{d}\right)^s \frac{ds}{s^{k+l+1}}.$$

Insertando esto en (2.6) obtenemos

$$\mathcal{T} = \frac{1}{(2\pi i)^2}\int_{(1)}\int_{(1)} F(s, s'; \Omega)\frac{R^{s+s'}}{(ss')^{k+l+1}} ds ds', \tag{2.9}$$

donde es

$$\begin{aligned}F(s, s'; \Omega) &:= \sum_{d,d'} \mu(d)\mu(d')\frac{|\Omega([d, d'])|}{[d, d']\, d^s d'^{s'}} \\ &= \prod_p \left(1 - \frac{|\Omega(p)|}{p}\left(\frac{1}{p^s} + \frac{1}{p^{s'}} - \frac{1}{p^{s+s'}}\right)\right),\end{aligned}$$

y donde el producto se obtiene directamente de la multiplicatividad de los factores involucrados.

Recordando la definición de la función zeta de Riemann

$$\zeta(s) = \sum_n \frac{1}{n^s} = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1},$$

la expresión producto obtenida para $F(s, s'; \Omega)$ sugiere que muchas de sus propiedades serán semejantes a las de $\left(\frac{\zeta(s+s'+1)}{\zeta(s+1)\zeta(s'+1)}\right)^k$, notando en particular que es $|\Omega(p)| = k$ para $p > H$ y en consecuencia $F(s, s'; \Omega)$ poseerá un cero de orden k en $(0, 0)$. Para estudiar la validez de esta expectativa introducimos la función

$$G(s, s'; \Omega) := F(s, s'; \Omega) \left(\frac{\zeta(s+1)\zeta(s'+1)}{\zeta(s+s'+1)}\right)^k.$$

Observar que el valor de esta función en el orígen está dado por

$$G(0, 0; \Omega) = \prod_p \left(1 - \frac{|\Omega(p)|}{p}\right)\left(1 - \frac{1}{p}\right)^{-k} = \mathfrak{G}(\mathcal{H}), \tag{2.10}$$

que es distinto de cero por hipótesis.

Obtendremos pronto importantes propiedades de esta función en una cierta región $\Re(s), \Re(s') > -c$, pero antes que eso introduciremos el siguiente lema fundamental que será demostrado más adelante:

Lema 2.2. *Sea N un parámetro real positivo. Sea $R = N^{c'}$ con $0 < c' < 1$. Sea $G = G(s, s')$ una función de $2m$ variables complejas $s = (s_1, \ldots, s_m)$, $s' = (s'_1, \ldots, s'_m)$, regular y acotada para $\Re(s_1), \ldots, \Re(s_m), \Re(s'_1), \ldots, \Re(s'_m) > -c''$, con $c'' > 0$. Supongamos además que en tal región satisface*

$$|G(s, s')| < c_1 \exp\left(c_2 (\log N)^{-2c_3\sigma} \log\log\log N\right), \tag{2.11}$$

con $\sigma = \min\left(\Re(s_1), \Re(s'_1), \ldots, \Re(s_m), \Re(s'_m), 0\right)$. Entonces, si N (y en consecuencia R) es suficientemente grande en términos de c'' y c_3, tenemos

$$\frac{1}{(2\pi i)^{2m}} \int_{(1)} \cdots \int_{(1)} G(s, s') \prod_{j=1}^m \left(\frac{\zeta(s_j + s'_j + 1)}{\zeta(s_j + 1)\zeta(s'_j + 1)}\right)^{k_j} \frac{R^{s_j + s'_j}}{(s_j s'_j)^{k_j + l_j + 1}} ds_j ds'_j$$
$$= \left(G(0, \ldots, 0) + o_{c', c_1, c_2, m}(1)\right) \prod_{j=1}^m \frac{(\log R)^{k_j + 2l_j}}{(k_j + 2l_j)!} \binom{2l_j}{l_j}.$$

Nota. La forma general en la que hemos enunciado éste lema nos será útil más adelante. Para la presente demostración, sin embargo, necesitaremos únicamente el caso $m = 1$.

Vemos entonces que para probar la Proposición 2.1 bastará demostrar que $G(s, s'; \Omega)$ satisface las hipótesis del Lema 2.2. En efecto, de (2.9) y la defínción de $G(s, s'; \Omega)$ tenemos

$$\mathcal{T} = \frac{1}{(2\pi i)^2} \int_{(1)} \int_{(1)} G(s, s'; \Omega) \left(\frac{\zeta(s + s' + 1)}{\zeta(s+1)\zeta(s'+1)}\right)^k \frac{R^{s+s'}}{(ss')^{k+l+1}} ds ds',$$

que es la situación $m = 1$ del Lema 2.2. El resultado se sigue entonces de (2.10).

Para esto nos restringiremos a la región $\Re(s), \Re(s') > -c$ con c suficientemente pequeño. Consideremos un primo $p > H$ de modo tal que es $|\Omega(p)| = k$ y estudiemos su correspondiente factor en $G(s, s'; \Omega)$. Éste será de la forma

$$\left(1 - \frac{k}{p}\left(\frac{1}{p^s} + \frac{1}{p^{s'}} - \frac{1}{p^{s+s'}}\right)\right) \frac{\left(1 - \frac{1}{p^{s+s'+1}}\right)^k}{\left(1 - \frac{1}{p^{s+1}}\right)^k \left(1 - \frac{1}{p^{s'+1}}\right)^k}. \tag{2.12}$$

Al igual que en el enunciado del lema escribimos $\sigma = \min(\Re(s), \Re(s'), 0)$. En la región en cuestión tenemos las igualdades

$$\begin{aligned}
\left(1 - \frac{1}{p^{s+s'+1}}\right)^k &= \left(1 - \frac{k}{p^{s+s'+1}} + O\left(\frac{1}{p^{2(2\sigma+1)}}\right)\right), \\
\left(1 - \frac{1}{p^{s+1}}\right)^{-k} &= \left(\sum_r \frac{1}{p^{r(s+1)}}\right)^k = \left(1 + \frac{1}{p^{s+1}} + \frac{1}{p^{s+1}(p^{s+1} - 1)}\right)^k \\
&= \left(1 + \frac{k}{p^{s+1}} + O\left(\frac{1}{p^{2\sigma+2}}\right)\right), \\
\left(1 - \frac{1}{p^{s'+1}}\right)^{-k} &= \left(\sum_r \frac{1}{p^{r(s'+1)}}\right)^k = \left(1 + \frac{1}{p^{s'+1}} + \frac{1}{p^{s'+1}(p^{s'+1} - 1)}\right)^k \\
&= \left(1 + \frac{k}{p^{s'+1}} + O\left(\frac{1}{p^{2\sigma+2}}\right)\right).
\end{aligned} \tag{2.13}$$

Insertando estas tres identidades en (2.12) obtenemos que esto es igual a $1 + O\left(\frac{1}{p^{2\sigma+2}}\right)$ de donde concluimos que la parte de $G(s, s'; \Omega)$ correspondiente a los factores $p > H$ está uniformemente acotada en la región estudiada, siempre que sea $c < 1/2$.

Para estudiar la parte correspondiente a los $p \leq H$ notamos que bastará con considerar $(3k)^{\frac{1}{2c+1}} < p \leq H$ puesto que el intervalo restante es finito y en consecuencia define un producto uniformemente acotado. Nuevamente, tenemos

$$\begin{aligned}
\left|\log\left(1 - \frac{|\Omega(p)|}{p^{s+1}} - \frac{|\Omega(p)|}{p^{s'+1}} + \frac{|\Omega(p)|}{p^{s+s'+1}}\right)\right| &\leq \log\left(1 - \frac{3k}{p^{2\sigma+1}}\right) \\
&= \sum_{n=1}^{\infty} \frac{(3k)^n}{np^{n(2\sigma+1)}} \\
&\leq \frac{3k}{p^{2\sigma+1} - 3k}
\end{aligned}$$

y además es

$$\left|\log\left(1-\frac{1}{p^{s+1}}\right)\right| = \left|\sum_{n=1}^{\infty}\frac{1}{np^{n(s+1)}}\right| \leq \frac{1}{p^{\sigma+1}-1},$$
$$\left|\log\left(1-\frac{1}{p^{s'+1}}\right)\right| = \left|\sum_{n=1}^{\infty}\frac{1}{np^{n(s'+1)}}\right| \leq \frac{1}{p^{\sigma+1}-1},$$
$$\left|\log\left(1-\frac{1}{p^{s+s'+1}}\right)\right| = \left|\sum_{n=1}^{\infty}\frac{1}{np^{n(s+s'+1)}}\right| \leq \frac{1}{p^{2\sigma+1}-1}. \tag{2.14}$$

Con estas identidades, podemos concluir que el valor absoluto del logaritmo de cada p-factor de $G(s,s';\Omega)$ en el intervalo $(3k)^{\frac{1}{2c+1}} < p \leq H$ está acotado por $3k((p^{2\sigma+1}-1)^{-1}+(p^{2\sigma+1}-3k)^{-1}) = O(p^{-1-2\sigma})$. Pero es

$$\begin{aligned}\sum_{p\leq H}\frac{1}{p^{1+2\sigma}} &\ll H^{-2\sigma}\sum_{p\leq H}\frac{1}{p} \ll H^{-2\sigma}\log\log H\\ &\ll (\log N)^{-2\sigma}\log\log\log N,\end{aligned}$$

en donde hemos usado (2.1) y la estimación clásica $\sum_{p\leq x}p^{-1} \ll \log\log x$. Exponenciando esto y puesto que vimos que el resto del producto está uniformemente acotado en la región en cuestión, obtenemos que efectivamente $G(s,s';\Omega)$ satisface (3.1). Esto finaliza la demostración. □

2.2. Comportamiento sobre formas lineales

El siguiente objetivo será extender los métodos desarrollados arriba para estudiar el comportamiento de $\Lambda_R(n;\mathcal{H},k+l)$ a lo largo de varias formas lineales simultáneas. Con éste propósto en mente, implementaremos el truco-W de Green-Tao que consiste en evaluar nuestra función en $Wn+a$ en lugar de en n para una elección apropiada de los parámetros W y a. La principal utilidad de esta sencilla técnica reside en evitar las obstrucciones provenientes de los primos pequeños que causan que $\Lambda_R(n;\mathcal{H},k+l)$ esté distribuido irregularmente en las clases residuales pequeñas. Entre las ventajas de esto se encuentra la simplificación del estudio de la serie singular y el hecho de que todas las formas lineales den lugar esencialmente a la misma distribución.

A lo largo de esta sección $w(N)$ denotará una función que tiende al infinito con N en forma suficientemente lenta, de modo que es $1/w(N) = o(1)$ y escribiremos $W := \prod_{p\leq w(N)}p$. Asimismo, R será una potencia pequeña de N y denotaremos por φ_k a la función de Euler generalizada, es decir

$$\varphi_k(n) := n\prod_{p\mid n}\left(1-\frac{1}{p}\right)^k.$$

El resultado que vamos a probar es el siguiente:

Proposición 2.2. *Sean m y t enteros positivos. Para cada $1 \leq i \leq m$, sean $\psi_i : \mathbb{Z}^t \to \mathbb{Z}$ formas lineales afines, $\psi_i(\mathbf{x}) := \sum_{j=1}^{t} L_{ij}x_j + b_i$, donde L_{ij}, b_i son coeficientes enteros con $|L_{ij}| \leq \sqrt{w(N)}/2$ para todo $i = 1, \dots, m$ y $j = 1, \dots, t$. Asumimos que las tuplas $(L_{i1}, \dots, L_{it})$ son no nulas y ninguna es un múltiplo racional de otra. Sean $k_1, \dots, k_m, l_1, \dots, l_m > 0$ enteros arbitrarios y $\mathcal{H}_1, \dots, \mathcal{H}_m \subseteq [1, H] \cap \mathbb{Z}$ tuplas admisibles (no necesariamente distintas) con $|\mathcal{H}_i| = k_i$. Aquí H es un entero arbitrario pero fijo. Asumimos además que N es suficientmente grande, de modo que se satisfaga $w(N) > H$. Escribimos $\theta_i := W\psi_i + a_i$ con los a_i elegidos de forma tal que $a_i + \mathcal{H}_i$ consista de enteros coprimos con W. Supongamos además que $B \subseteq \mathbb{R}^t$ es el producto de t intervalos de longitud al menos R^{10m}. Entonces, si $w(N)$ crece lo suficientemente despacio con N, tenemos*

$$\begin{aligned}&\mathbb{E}_{\mathbf{x}\in B}\left(\Lambda_R^2(\theta_1(\mathbf{x}); \mathcal{H}_1, k_1 + l_1) \dots \Lambda_R^2(\theta_m(\mathbf{x}); \mathcal{H}_m, k_m + l_m)\right)\\ &= (1 + o_{m,t}(1)) \prod_{i=1}^{m} \frac{W}{\varphi_{k_i}(W)} \frac{(\log R)^{k_i+2l_i}}{(k_i + 2l_i)!} \binom{2l_i}{l_i}.\end{aligned}$$

Nota. La demostración de esta proposición en el caso particular $\mathcal{H}_1, \dots, \mathcal{H}_m = \{0\}$, $k_1, \dots, k_m = 1$, $l_1, \dots, l_m = 0$ fue dada en [24] (Prop. 9.5). La generalización aquí presente adapta los métodos allí utilizados, que a su vez están basados en [18].

Demostración. Usaremos c para denotar una constante positiva que depende de $k_1, \dots, k_m, l_1, \dots, l_m$ cuyo valor podrá variar en cada ocurrenica. Comenzamos expandiendo los cuadrados dentro de la esperanza para expresar a esta como

$$\mathbb{E}_{x\in B}\left(\prod_{i=1}^{m} \sum_{\substack{d_i, d_i' \leq R \\ d_i, d_i' | P(\theta_i(\mathbf{x}); \mathcal{H}_i)}} \frac{\mu(d_i)\mu(d_i')}{(k_i + l_i)!^2} \left(\log^+ \frac{R}{d_i}\right)^{k_i+l_i} \left(\log^+ \frac{R}{d_i'}\right)^{k_i+l_i}\right),$$

lo cual podemos reescribir como

$$\begin{aligned}\sum_{d_1,\dots,d_m,d_1',\dots,d_m' \in \mathbb{N}} &\left(\prod_{i=1}^{m} \frac{\mu(d_i)\mu(d_i')}{(k_i + l_i)!^2} \left(\log^+ \frac{R}{d_i}\right)^{k_i+l_i} \left(\log^+ \frac{R}{d_i'}\right)^{k_i+l_i}\right)\\ &\times \mathbb{E}_{\mathbf{x}\in B}\left(\prod_{i=1}^{m} \mathbf{1}_{d_i, d_i' | P(\theta_i(\mathbf{x}); \mathcal{H}_i)}\right).\end{aligned}$$

Notar que la presencia de la función de Möbius nos permite asumir que los d_i, d_i' son todos libres de cuadrados. Sea $D = [d_1, \dots, d_m, d_1', \dots, d_m']$ el mínimo común múltiplo de los divisores en cuestión, de modo que es $D \leq R^{2m}$. Puesto que $d_i, d_i' | \theta_i(\mathbf{x})$ si y sólo si $d_i, d_i' | \theta_i(\mathbf{x} + D\mathbf{y})$ para cualquier vector $\mathbf{y}$ de $\mathbb{R}^t$, se sigue que

$$\mathbb{E}_{\mathbf{x}\in B}\left(\prod_{i=1}^{m} \mathbf{1}_{d_i, d_i' | P(\theta_i(\mathbf{x}); \mathcal{H}_i)}\right) = \mathbb{E}_{\mathbf{x}\in \mathbb{Z}_D^t}\left(\prod_{i=1}^{m} \mathbf{1}_{d_i, d_i' | P(\theta_i(\mathbf{x}); \mathcal{H}_i)}\right) + O_{m,t}(R^{-8m}).$$

En efecto, si escribimos $B = \prod_{j=1}^{t} I_j$ con I_j un intervalo de $\mathbb{R}$ y consideramos para cada j un subintervalo maximal $I'_j \subseteq I_j$ que cubra uniformemente a $\mathbb{Z}_D$ mediante la proyección canónica, entonces $B' = \prod_{j=1}^{t} I'_j$ cubre uniformemente a $\mathbb{Z}_D^t$. Dado que un elemento de B posee su j-ésima coordenada en $I_j - I'_j$ con probabilidad a lo sumo $D/|I_j|$, el término de error dado se sigue de que por hipótesis es $|I_j| \geq R^{10m}$ para todo j.

Reemplazando esto en nuestra estimación anterior, obtenemos un error de a lo sumo $O_{m,t}\left(R^{-6m}\log^{mc} R\right) = o_{m,t}(1)$. Usando ahora que los d_i, d'_i se pueden tomar libres de cuadrados, obtenemos del teorema chino del resto la identidad

$$\mathbb{E}_{\mathbf{x}\in\mathbb{Z}_D^t}\left(\prod_{i=1}^{m}\mathbf{1}_{d_i,d'_i|P(\theta_i(\mathbf{x});\mathcal{H}_i)}\right) = \prod_{p|D}\mathbb{E}_{\mathbf{x}\in\mathbb{Z}_p^t}\left(\prod_{i:p|d_id'_i}\mathbf{1}_{p|P(\theta_i(\mathbf{x});\mathcal{H}_i)}\right).$$

Aquí la restricción $p|D$ es superflua puesto que en caso contrario el factor es igual a 1. En particular, escribiendo $X_{d_1,\dots,d_m}(p) := \{1 \leq i \leq m : p|d_i\}$ y

$$\omega_X(p) := \mathbb{E}_{\mathbf{x}\in\mathbb{Z}_p^t}\left(\prod_{i\in X}\mathbf{1}_{p|P(\theta_i(\mathbf{x});\mathcal{H}_i)}\right)$$

para cada $X \subseteq \{1,\dots,m\}$, tenemos

$$\mathbb{E}_{\mathbf{x}\in\mathbb{Z}_D^t}\left(\prod_{i=1}^{m}\mathbf{1}_{d_i,d'_i|P(\theta_i(\mathbf{x});\mathcal{H}_i)}\right) = \prod_{p}\omega_{X_{d_1,\dots,d_m}\cup X_{d'_1,\dots,d'_m}}(p).$$

Nuestro objetivo será entonces probar

$$\begin{aligned}\sum_{d_1,\dots,d_m,d'_1,\dots,d'_m\in\mathbb{N}}&\left(\prod_{i=1}^{m}\frac{\mu(d_i)\mu(d'_i)}{(k_i+l_i)!^2}\left(\log^+\frac{R}{d_i}\right)^{k_i+l_i}\left(\log^+\frac{R}{d'_i}\right)^{k_i+l_i}\right)\\ &\times\prod_{p}\omega_{X_{d_1,\dots,d_m}\cup X_{d'_1,\dots,d'_m}}(p)\\ &= (1+o_{m,t}(1))\prod_{i=1}^{m}\frac{W}{\varphi_{k_i}(W)}\frac{(\log R)^{k_i+2l_i}}{(k_i+2l_i)!}\binom{2l_i}{l_i}.\end{aligned}$$

Procediendo como en la demostración de la Proposición (2.1), utilizamos (2.8) para expresar el lado izquierdo de esta identidad como

$$\frac{1}{(2\pi i)^{2m}}\int_{(1)}\cdots\int_{(1)}F(s,s')\prod_{j=1}^{m}\frac{R^{s_j+s'_j}}{(s_js'_j)^{k_j+l_j+1}}dsds', \tag{2.15}$$

donde hay $2m$ variables complejas $s := (s_1,\dots,s_m)$, $s' := (s_1,\dots,s_m)$ y es

$$F(s,s') := \sum_{d_1,\dots,d_m,d'_1,\dots,d'_m\in\mathbb{N}}\left(\prod_{j=1}^{m}\frac{\mu(d_j)\mu(d'_j)}{d_j^s{d'_j}^{s'}}\right)\prod_{p}\omega_{X_{d_1,\dots,d_m}\cup X_{d'_1,\dots,d'_m}}(p). \tag{2.16}$$

Los índices han sido cambiados de i a j para evitar confusión con la raíz de -1. Notar que cada divisor p de $D = [d_1, \dots, d_m, d'_1, \dots d'_m]$ contribuye un factor de

$$\left(\frac{(-1)^{|X_{d_1,\dots,d_m}|+|X_{d'_1,\dots,d'_m}|}}{\exp\left(\log p \left(\sum_{j\in X_{d_1,\dots,d_m}} s_j + \sum_{j\in X_{d'_1,\dots,d'_m}} s'_j\right)\right)} \right) \omega_{X_{d_1,\dots,d_m}\cup X_{d'_1,\dots,d'_m}}(p)$$

al sumando de (2.16) que es multiplicativo en D. Debido a esto, el teorema fundamental de la aritmética nos da la expresión $F(s,s') = \prod_p E_p(s,s')$ con

$$E_p(s,s') := \sum_{X,X'\subseteq\{1,\dots m\}} \frac{(-1)^{|X|+|X'|}}{p^{\sum_{j\in X} s_j + \sum_{j\in X'} s'_j}} \omega_{X\cup X'}(p) \tag{2.17}$$

allí donde tal producto converge absolutamente.

Para proceder aprovecharemos que ninguna parte lineal de un ψ_i es un múltiplo racional de otra, lo cual nos permitirá deducir cotas en el tamaño de $\omega_X(p)$. A continuación escribiremos $\omega_i(p)$ para referirnos a $\omega_X(p)$ cuando es $X = \{i\}$. El resultado concreto es el siguiente

Lema 2.3. *Si es $p \le w(N)$ entonces tenemos $\omega_X(p) = 0$ para todo X no vacío, de modo que es $E_p = 1$ en este caso. Si en cambio es $p > w(N)$ entonces tenemos $\omega_i(p) = p^{-1}k_i$ cuando es $|X| = 1$ y $\omega_X(p) \le Kp^{-2}$ si es $|X| \ge 2$, con $K = \prod_{1\le i\le m} k_i$.*

Demostración. En primer lugar, es claro que si es $p \le w(N)$ entonces es $\theta_i(\mathbf{x}) \equiv a_i (\mathrm{mod}\ p)$ para todo $x \in \mathbb{Z}^t$, por lo que el primer resultado se sigue. Supongamos entonces $p > w(N)$. Por hipótesis, tenemos entonces $p > H$ y en consecuencia $\theta_i(\mathbf{x}) + \mathcal{H}_i$ abarca exactamente k_i clases residuales módulo p, lo cual nos da el caso $|X| = 1$ del enunciado. Resta ver entonces que sucede cuando es $|X| \ge 2$. Para esto veamos que ninguna de las s formas lineales puras $W(\psi_i - b_i)$ es un múltiplo de ninguna otra módulo p. En efecto, si tal fuese el caso, debería ser $L_{ij}L_{i'j}^{-1} \equiv L_{il}L_{i'l}^{-1} (\mathrm{mod}\ p)$ para todo par $1 \le j \le l \le t$. Pero si dos racionales en forma reducida con numerador y denominador acotados en módulo por $\sqrt{\omega(N)}/2 < \sqrt{p}/2$ son congruentes módulo p, entonces claramente deben ser iguales, de donde se sigue entonces que las formas lineales puras $\psi_i - b_i$, $\psi_{i'} - b_{i'}$ son múltiplos racionales entre sí, contrario a la hipótesis. Ahora bien, si para cada $i \in X$ elegimos un $h_i \in \mathcal{H}_i$ entonces se sigue de lo anterior que las soluciones de $\prod_{i\in X} \mathbf{1}_{p|W(\psi_i(\mathbf{x}))+a_i+h_i} = 1$ en $\mathbb{Z}_p^t$ están contenidas en un subespacio de codimensión al menos 2 (puesto que es $|X| \ge 2$). Luego la esperanza de éste indicador es a lo sumo p^{-2}. Notamos finalmente que si es $\prod_{i\in X} \mathbf{1}_{p|P(\theta_i(\mathbf{x});\mathcal{H}_i)} = 1$ entonces el anterior sistema debe satisfacerse para cierta elección de $h_i \in \mathcal{H}_i$, $i \in X$. Dado que hay a lo sumo K posibles elecciones y puesto que acabamos de ver que la esperanza de cada tal sistema está acotada por p^{-2}, se sigue que es $\omega_X(p) \le Kp^{-2}$, como se quería probar. □

Aplicando esto a (2.17) obtenemos

$$E_p(s,s') = 1 - \mathbf{1}_{p>w(N)} \sum_{j=1}^{m} k_j \left(p^{-1-s_j} + p^{-1-s'_j} - p^{-1-s_j-s'_j}\right) + \mathbf{1}_{p>w(N)} \sum_{\substack{X,X'\subseteq\{1,\dots,m\} \\ |X\cup X'|\geq 2}} \frac{O\left(1/p^2\right)}{p^{\sum_{j\in X} s_j + \sum_{j\in X'} s'_j}}. \tag{2.18}$$

La idea será aprovechar esta expansión para colocarnos en la situación del Lema 2.2. Escribimos entonces $E_p = E_p^{(1)} E_p^{(2)} E_p^{(3)}$ con

$$E_p^{(1)}(s,s') := E_p(s,s') \prod_{j=1}^{m} \frac{\left(1 - \mathbf{1}_{p>w(N)} p^{-1-s_j-s'_j}\right)^{k_j}}{\left(1 - \mathbf{1}_{p>w(N)} p^{-1-s_j}\right)^{k_j} \left(1 - \mathbf{1}_{p>w(N)} p^{-1-s'_j}\right)^{k_j}}$$

$$E_p^{(2)}(s,s') := \prod_{j=1}^{m} \frac{\left(1 - \mathbf{1}_{p\leq w(N)} p^{-1-s_j-s'_j}\right)^{k_j}}{\left(1 - \mathbf{1}_{p\leq w(N)} p^{-1-s_j}\right)^{k_j} \left(1 - \mathbf{1}_{p\leq w(N)} p^{-1-s'_j}\right)^{k_j}}$$

$$E_p^{(3)}(s,s') := \prod_{j=1}^{m} \left(1 - p^{-1-s_j}\right)^{k_j} \left(1 - p^{-1-s'_j}\right)^{k_j} \left(1 - p^{-1-s_j-s'_j}\right)^{-k_j}$$

y definiendo $G_j := \prod_p E_p^{(j)}$ para $j = 1, 2, 3$, obtenemos así la expresión $F = G_1 G_2 G_3$ con

$$G_3(s,s') = \prod_{j=1}^{m} \left(\frac{\zeta(1+s_j+s'_j)}{\zeta(1+s_j)\zeta(1+s'_j)}\right)^{k_j}.$$

Reemplazando esta representación de F en (2.15), vemos que para obtener una estimación del tipo deseado bastará probar que $G_1 G_2$ satisface las condiciones del Lema 2.2. Esto será logrado a través del siguiente resultado:

Lema 2.4. *Los productos $\prod_p E_p^{(j)}$, $j = 1, 2$, convergen absolutamente a G_j en $\Re(s_1), \dots, \Re(s_m), \Re(s'_1), \dots, \Re(s'_m) > -\alpha$ para cierta constante $\alpha > 0$ que depende de m. Además, escribiendo $\sigma = \min\left(\Re(s_1), \Re(s'_1), \dots, \Re(s_m), \Re(s'_m), 0\right)$ y $s = (s_1, \dots, s_m)$, $s' = (s'_1, \dots, s'_m)$, tenemos*

$$\begin{aligned} G_1(s,s') &\ll_m 1 \\ G_1(0,0) &= 1 + o_m(1) \\ G_2(s,s') &\ll \exp(C_m w(N)^{-2\sigma} \log\log w(N)) \\ G_2(0,0) &= \prod_{j=1}^{m} \frac{W}{\varphi_{k_j}(W)} \end{aligned}$$

Demostración. En primer lugar, tenemos que es $E_p^{(1)} = 1$ si $p \leq w(N)$. Asimismo, si es $p > w(N)$ entonces obtenemos de (2.18) y las identidades de (2.13) que es $E_p^{(1)} = 1+O(p^{-2+2m\alpha})$, de donde se sigue la convergencia absoluta del producto y la cota para G_1, mediante una elección suficientemente pequeña de $\alpha > 0$. El valor de $G_1(0,0)$ se sigue de lo anterior, notando que $w(N)$ diverge con N. Para estudiar G_2 utilizamos en cambio las identidades de (2.14) obteniendo la cota

$$\log|G_2| \ll_m w(N)^{-2\sigma} \sum_{p\leq w(N)} \frac{1}{p} \ll_m w(N)^{-2\sigma} \log\log w(N)$$

en la región en cuestión, dónde hemos utilizado como antes la estimación $\sum_{p\leq x} p^{-1} \ll \log\log x$. Finalmente, el valor de $G_2(0,0)$ es inmediato por definición de φ_k y W. □

Puesto que el crecimiento de $w(N)$ se asume lo suficientemente lento respecto a N, podemos suponer en particular $w(N) < \log N$. Con esto, se sigue del Lema 2.4 que $G := G_1G_2$ satisface las condiciones del Lema 2.2. Asimismo, del Lema 2.4 se sigue también que es

$$G(0,0) = (1+o_m(1))\prod_{i=1}^{m} \frac{W}{\varphi_{k_i}(W)}$$

y el resultado queda así demostrado. □

2.3. Estimaciones de correlación

Habiendo estudiado el comportamiento de $\Lambda_R(n;\mathcal{H},k+l)$ a lo largo de formas lineales, concluiremos éste capítulo estudiando ciertas propiedades de correlación de esta función. Esta vez nuestro estudio incluirá formas puras que sí son múltiplos racionales entre sí (de hecho, sérán iguales) de modo que el resultado anterior no será aplicable. De todas formas, gran parte de tal razonamiento podrá adaptarse a éste contexto.

Hasta el final de esta sección $H,k,l > 0$ serán enteros fijos y $\mathcal{H} = \{h_1,\dots,h_k\} \subseteq [1,H]\cap\mathbb{Z}$ denotará una tupla admisible también fija. Como antes, $w(N)$ será una función que tiende al infinito con N suficientemente despacio y escribiremos $W := \prod_{p\leq w(N)} p$. El enunciado es el siguiente:

Proposición 2.3. *Sea $m \geq 1$ un entero y B un intervalo de longitud al menos R^{10m}. Supongamos que $z_1 < \ldots < z_m$ son enteros distintos satisfaciendo $|z_i| \leq N^2$ para todo $1 \leq i \leq m$. Definimos*

$$\Delta := \left| \prod_{1\leq i<j\leq m} \prod_{-H\leq b\leq H} (W(z_j - z_i) + b) \right|.$$

Entonces, si N es suficientemente grande en términos de m y a es un entero tal que $a+\mathcal{H}$ consiste de elementos coprimos con W, tenemos

$$\mathbb{E}_{x\in B}\left(\Lambda_R^2(W(x+z_1)+a;\mathcal{H},k+l)\dots\Lambda_R^2(W(x+z_m)+a;\mathcal{H},k+l)\right)$$
$$\leq (1+o_m(1))\left(\frac{W(\log R)^{k+2l}}{\varphi_k(W)(k+l)!}\binom{2l}{l}\right)^m \prod_{\substack{p|\Delta\\(p,W)=1}}\left(1+O_m(p^{-1/2})\right) \quad (2.19)$$

Nota. Similarmente a la Proposición 2.2, el caso particular $\mathcal{H}=\{0\}$, $H=k=l=0$ fue demostrado en [24]. La siguiente demostración adapta los métodos allí utilizados, que a su vez están basados en [18].

Demostración. Para llevar a cabo esta demostración, comenzamos notando que la principal diferencia entre esta proposición y la Proposición 2.2 radica en que en éste caso las formas lineales puras involucradas son todas iguales a x. Si bien esto nos impide por supuesto sacar provecho de que ninguna forma lineal sea múltiplo racional de otra, tal hipótesis en la demostración de la Proposición 2.2 sólo se comenzó a utilizar a partir del Lema 2.3. Debido a esto, las argumentos anteriores siguen siendo válidos y en particular podemos expresar el lado izquierdo de (2.19) en la forma (2.15) con $F(s,s')$ definido como en (2.16) y con $k_i=k$, $l_i=l$ para todo $1\leq i\leq m$, con la única diferencia de que ahora será

$$\omega_X(p):=\mathbb{E}_{x\in\mathbb{Z}_p}\left(\prod_{i\in X}\mathbf{1}_{p|P(W(x+z_i)+a;\mathcal{H})}\right).$$

Al igual que antes obtenemos la expresión $F(s,s')=\prod_p E_p(s,s')$, con $E_p(s,s')$ definido igual que en (2.17) (tomando en cuenta por supuesto la nueva definición de $\omega_X(p)$). En lugar del Lema 2.3 tenemos ahora el siguiente resultado:

Lema 2.5. *Si es $p\leq w(N)$ entonces es $\omega_X(p)=0$ para todo X no vacío, siendo en tal caso $E_p=1$. Si es $p>w(N)$ entonces tenemos $\omega_X(p)\leq kp^{-1}$ para todo X no vacío, siendo en particular $\omega_X(p)=kp^{-1}$ si es $|X|=1$. Además, si es $|X|\geq 2$, $\omega_X(p)$ será nulo a menos que p divida a Δ.*

Demostración. La primera afirmación es evidente puesto que en tal caso $W(x+z_i)+a+\mathcal{H}$ consistirá de elementos coprimos con p para todo $x\in\mathbb{Z}_p$. Supongamos entonces $p>w(N)$. Como $w(N)$ diverge con N, eligiendo N suficientemente grande podemos asumir $W>H$. Con esto, el caso $|X|=1$ es claro. Consideremos finalmente $|X|\geq 2$. En tal caso, puesto que $P(W(x+z_i)+a;\mathcal{H})$ es divisible por p para exactamente k elementos de $\mathbb{Z}_p$, se sigue que la esperanza del indicador es a lo sumo kp^{-1}. Además, si tal esperanza es no nula, deberá existir un valor de $x\in\mathbb{Z}_p$ para el cual p divida a $W(x+z_i)+a+h_{ij}$ para todo $i\in X$ y cierto $h_{ij}\in\mathcal{H}$. En particular, para todo par $i,i'\in X$, $i<i'$, p dividirá a $(W(x+z_{i'})+a+h_{i'j'})-(W(x+z_i)+a+h_{ij})=$

$W(z_{i'} - z_i) + (h_{i'j'} - h_{ij})$ para cierto par $h_{ij}, h_{i'j'} \in \mathcal{H}$. Pero por definición de $\mathcal{H}$ y H, sabemos que es $|h_{j'} - h_j| \leq H$ para todo par $h_j, h_{j'}$ en $\mathcal{H}$, de donde se sigue que p habrá de dividir a Δ, como se deseaba probar. □

El Lema 2.5 nos permite escribir

$$E_p(s,s') = 1 - \mathbf{1}_{p>w(N)} \sum_{j=1}^{m} k\left(p^{-1-s_j} + p^{-1-s'_j} - p^{-1-s_j-s'_j}\right) + \mathbf{1}_{p>w(N),p|\Delta}\lambda_p(s,s')$$

con

$$\lambda_p(s,s') = \sum_{\substack{X,X'\subseteq\{1,\dots,m\} \\ |X\cup X'|\geq 2}} \frac{O\left(1/p\right)}{p^{\sum_{j\in X} s_j + \sum_{j\in X'} s'_j}}.$$

Consideramos ahora la factorización $E_p = E_p^{(0)} E_p^{(1)} E_p^{(2)} E_p^{(3)}$ con

$$E_p^{(0)}(s,s') := 1 + \mathbf{1}_{p>w(N),p|\Delta}\lambda_p(s,s')$$

$$E_p^{(1)}(s,s') := \frac{E_p(s,s')}{E_p^{(0)}(s,s')} \prod_{j=1}^{m} \frac{\left(1 - \mathbf{1}_{p>w(N)}p^{-1-s_j-s'_j}\right)^k}{\left(1 - \mathbf{1}_{p>w(N)}p^{-1-s_j}\right)^k \left(1 - \mathbf{1}_{p>w(N)}p^{-1-s'_j}\right)^k}$$

$$E_p^{(2)}(s,s') := \prod_{j=1}^{m} \frac{\left(1 - \mathbf{1}_{p\leq w(N)}p^{-1-s_j-s'_j}\right)^k}{\left(1 - \mathbf{1}_{p\leq w(N)}p^{-1-s_j}\right)^k \left(1 - \mathbf{1}_{p\leq w(N)}p^{-1-s'_j}\right)^k}$$

$$E_p^{(3)}(s,s') := \prod_{j=1}^{m} \left(1 - p^{-1-s_j}\right)^k \left(1 - p^{-1-s'_j}\right)^k \left(1 - p^{-1-s_j-s'_j}\right)^{-k}.$$

Escribimos $G_j := \prod_p E_p^{(j)}$ para $j = 1,2,3,4$. Tenemos entonces $F = G_0G_1G_2G_3$ con

$$G_3(s,s') = \prod_{j=1}^{m} \left(\frac{\zeta(1+s_j+s'_j)}{\zeta(1+s_j)\zeta(1+s'_j)}\right)^k.$$

Por supuesto, la idea será utilizar nuevamente el Lema 2.2. Para esto, utilizamos el siguiente resultado análogo al Lema 2.4 que describe el comportamiento de G_0, G_1, G_2.

Lema 2.6. *Los productos $\prod_p E_p^{(j)}$, $j = 0,1,2$, convergen absolutamente a G_j en $\Re(s_1), \Re(s'_1), \dots, \Re(s_m), \Re(s'_m) > -\alpha$ para cierta constante $\alpha > 0$ que depende de m. Además, escribiendo $\sigma = \min\left(\Re(s_1), \Re(s'_1), \dots, \Re(s_m), \Re(s'_m), 0\right)$*

y $s = (s_1, \dots, s_m)$, $s' = (s'_1, \dots, s'_m)$, *tenemos*

$$\begin{aligned}
G_0(s, s') &\ll_m \exp(C_m(\log N)^{-2\sigma m} \log\log\log N) \\
G_1(s, s') &\ll_m 1 \\
G_2(s, s') &\ll \exp(C'_m w(N)^{-2\sigma} \log\log w(N)) \\
G_0(0, 0) &\leq \prod_{\substack{p|\Delta \\ (p,W)=1}} \left(1 + O_m(p^{-1/2})\right) \\
G_1(0, 0) &= 1 + o_m(1) \\
G_2(0, 0) &= (W/\varphi_k(W))^m.
\end{aligned}$$

Demostración. Comenzamos notando que es $E_p^{(1)} = 1$ cuando $p \leq w(N)$. Si en cambio es $p > w(N)$ entonces razonando como en el Lema 2.4 vemos que es

$$E_p^{(1)} = \frac{1 + \lambda_p(s, s') + O_m(p^{2\alpha m - 2})}{1 + \lambda_p(s, s')} = 1 + O_m(p^{2\alpha m - 2})$$

de donde se siguen ambas estimaciones para G_1 de la misma forma que en el Lema 2.4. Asimismo, las estimaciones de G_2 coinciden con las de aquel lema por lo que no hay nada que probar. Nos concentraremos entonces en G_0.

Notando que es $E_p^{(0)} = 1 + O_m(p^{-2\sigma m - 1})$ si $p > w(N)$, $p|\Delta$ y $E_p^{(0)} = 1$ en otro caso, tenemos

$$G_0 = \prod_{\substack{p|\Delta \\ (p,W)=1}} E_p^{(0)} = \prod_{\substack{p|\Delta \\ (p,W)=1}} \left(1 + O_m(p^{-2\sigma m - 1})\right),$$

de donde se sigue la cota para $G_0(0, 0)$. Para acotar esta expresión en la región en cuestión, tomamos logaritmos y reducimos así el problema al de estimar $\sum_{p|\Delta} p^{-2\sigma m - 1}$ (no será necesario considerar la restricción $(p, W) = 1$).

Sea P_x el producto de todos los primos debajo de x. Escribiendo $f(n) = \sum_{p|n} p^{-2\sigma m - 1}$ es claro que es $f(P_x) > f(n)$ para todo entero $n < P_x$, puesto que n poseerá estrictamente menos divisores primos distintos que P_x y si los ordenamos de menor a mayor, la contribución del i-ésimo tal divisor primo de n a la suma no podrá ser mayor que la del i-ésimo primo. Observamos ahora que suponiendo N suficientemente grande tendremos $W > H$ y en consecuencia será

$$\Delta \leq \prod_{1 \leq i < j \leq m} \prod_{-H \leq b \leq H} 2W(z_j - z_i) \leq \prod_{-H \leq b \leq H} (4W)^{\frac{m^2}{2}} N^{m^2} < N^{4Hm^2},$$

en donde hemos usado $|z_i| \leq N^2$ y que $w(N)$ (y en consecuencia W) crece suficientemente despacio con N. Bastará entonces encontrar x con $P_x > N^{4Hm^2}$ y acotar $f(P_x)$.

Del teorema del número primo en la forma $\sum_{p\leq x}\log p \sim x$ se sigue que existe una constante absoluta C con $\prod_{p\leq Cx} p \geq e^x$ para todo x suficientemente grande. Vemos entonces que bastará estudiar $f(P_x)$ con $x = \lfloor 4CHm^2 \log N \rfloor$, pero en tal caso tenemos

$$\sum_{p|P_x} p^{-2\sigma m-1} = \sum_{p\leq 4CHm^2\log N} p^{-2\sigma m-1} \ll_m (\log N)^{-2\sigma m}\log\log\log N.$$

Puesto que por lo anterior es $f(P_x) > f(\Delta)$ obtenemos así la cota deseada. □

Deducimos del Lema 2.6 que $G := G_0G_1G_2$ satisface las condiciones del Lema 2.2. Pero el Lema 2.6 nos dice también que es

$$G(0,0) \leq (1+o_m(1))(W/\varphi_k(W))^m \prod_{\substack{p|\Delta\\(p,W)=1}} \left(1+O_m(p^{-1/2})\right)$$

y en consecuencia aplicando el Lema 2.2 la Proposición 2.3 se sigue. □

Capítulo 3

Integrales de contorno

En éste capítulo llevaremos a cabo la demostración del Lema 2.2. Al igual que antes, mantendremos la convención de denotar por c a una constante positiva que depende a lo sumo de las variables $k, k_1, \dots, k_m, l, l_1, \dots, l_m$ y cuyo valor puede variar en cada ocurrencia. En consecuencia, mantendendremos la convención de suprimir el correspondiente subíndice c de las notaciones O, o.

Los resultados de este capítulo son una simple generalización de los argumentos presentes en [17] y [18].

3.1. El caso unidimensional

El siguiente lema nos dará el caso $m = 1$ del resultado. Una vez demostrado esto, la generalización multidimensional será inmediata.

Lema 3.1. *Sea N un parámetro real positivo. Sea $R = N^{c'}$ con $0 < c' < 1$, y sean $k, l > 0$ parámetros enteros. Sea $G = G(s, s')$ una función de dos variables complejas, regular y acotada en la región $\Re(s), \Re(s') > -c''$, con $c'' > 0$. Supongamos además que en tal región satisface*

$$|G(s, s')| < c_1 \exp\left(c_2 (\log N)^{-2\overline{\sigma} c_3} \log\log\log N\right), \tag{3.1}$$

con $\overline{\sigma} = \min(\Re(s), \Re(s'), 0)$. Si suponemos N (y en consecuencia R) suficientemente grande en términos de c'' y c_3, tenemos entonces

$$\begin{aligned} &\frac{1}{(2\pi i)^2} \int_{(1)} \int_{(1)} G(s, s') \left(\frac{\zeta(s + s' + 1)}{\zeta(s + 1)\zeta(s' + 1)}\right)^k \frac{R^{s+s'}}{(ss')^{k+l+1}} ds ds' \\ &= (G(0, 0) + o_{c', c_1, c_2}(1)) \frac{(\log R)^{k+2l}}{(k + 2l)!} \binom{2l}{l}. \end{aligned} \tag{3.2}$$

Antes de dar la demostración de éste resultado, enunciaremos ciertas propiedades de ζ que utilizaremos contínuamente.

Lema 3.2. *Existe una constante positiva $\bar{c}$ para la cual es $\zeta(\sigma+it)\neq 0$ en la región*

$$\sigma \geq 1 - \frac{4\bar{c}}{\log(|t|+3)}. \tag{3.3}$$

Además, en tal región tenemos las desigualdades

$$\zeta(\sigma+it) - \frac{1}{\sigma-1+it} \ll \log(|t|+3)$$
$$\frac{1}{\zeta(\sigma+it)} \ll \log(|t|+3).$$

Nota. La región dada por (3.3) es conocida como la región libre de ceros clásica de la función zeta de Riemann.

Demostración del Lema 3.2. Ver [44, Capítulo 3]. □

Demostración del Lema 3.1. A lo largo de esta demostración asumiremos siempre que es $G(s,s)\neq 0$ lo cual será superficial (puesto que estimaremos siempre a G mediante (3.1)) excepto en el estudio de ciertos polos. En esos casos será evidente sin embargo que la nulidad de G (y la consecuente reducción en el orden de los polos) afectará los argumentos únicamente en una posible reducción de los términos de error.

Comenzamos definiendo $U=\exp(\sqrt{\log N})$ y moviendo las integrales de contorno de s y s' a las líneas verticales $c_0(\log U)^{-1}+it$ y $c_0(2\log U)^{-1}+it$ respectivamente. Aquí $c_0>0$ es una constante absoluta suficientemente pequeña cuyo valor será especificado más adelante (en particular, nos permitirá trabajar en la región (3.3)). Escribiendo $s=\sigma+it$ y $s'=\sigma'+it'$, tenemos

$$\begin{aligned}
&\frac{1}{(2\pi i)^2}\int_{(\frac{c_0}{\sqrt{\log N}})}\int_{(\frac{c_0}{2\sqrt{\log N}})} G(s,s')\left(\frac{\zeta(s+s'+1)}{\zeta(s+1)\zeta(s'+1)}\right)^k \frac{R^{s+s'}}{(ss')^{k+l+1}}dsds' \\
&\ll_{c_1} \int_{\substack{(\frac{c_0}{\sqrt{\log N}})\\ |t'|\geq U/2}}\int_{(\frac{c_0}{2\sqrt{\log N}})} (\log\log N)^{c_2}\frac{(\log(|t|+3)\log(|t'|+3))^c}{(\log N)^{-k/2}}\frac{R^{\sigma+\sigma'}}{(|ss'|)^{k+l+1}}dsds',
\end{aligned} \tag{3.4}$$

donde hemos aprovechado que es $|s+s'|\geq \frac{3c_0}{2}(\sqrt{\log N})^{-1}$.

Utilizando las desigualades $R\leq N$, $|s'|\geq U/2$ y $\sigma+\sigma'=\frac{3}{2}\frac{c_0}{\sqrt{\log N}}$, vemos que es

$$\begin{aligned}
\frac{R^{\sigma+\sigma'}}{(|ss'|)^{k+l+1}} &\leq \frac{\exp\left(\frac{3\sqrt{\log N}}{2}\right)}{|ss'|^{k+l+1}} \\
&\ll \frac{\exp\left(-0{,}1\sqrt{\log N}\right)}{|s|^{k+l+1}|s'|^{k+l-0{,}6}}.
\end{aligned}$$

Insertando esto en (3.4) obtenemos la cota

$$\begin{aligned}
&\ll_{c_1} \exp\left(-0{,}1\sqrt{\log N}\right)(\log\log N)^{c_2}(\log N)^c \\
&\times \int_{\substack{(\frac{c_0}{\sqrt{\log N}}) \\ |t'|\geq U/2}} \frac{(\log(|t'|+3))^c}{|s'|^{k+l-0{,}6}} \int_{(\frac{c_0}{2\sqrt{\log N}})} \frac{(\log(|t|+3))^c}{|s|^{k+l+1}} dsds' \\
&\ll_{c_1,c_2} \exp\left(-c\sqrt{\log N}\right)
\end{aligned}$$

en donde hemos evaluado la integral interna como $\ll (\log N)^c$, usando las desigualdades $k, l \geq 1$, $|s| \geq \frac{c_0}{\sqrt{\log N}} + |t|$ y $|s'| \geq \frac{c_0}{2\sqrt{\log N}} + |t'|$.

Aplicando exactamente los mismos razonamientos para estimar

$$\frac{1}{(2\pi i)^2} \int_{\substack{(\frac{c_0}{\sqrt{\log N}}) \\ |t'|\leq U/2}} \int_{(\frac{c_0}{2\sqrt{\log N}}) \atop |t|\geq U} G(s,s') \left(\frac{\zeta(s+s'+1)}{\zeta(s+1)\zeta(s'+1)}\right)^k \frac{R^{s+s'}}{(ss')^{k+l+1}} dsds',$$

concluimos que (3.2) es igual a

$$\begin{aligned}
\frac{1}{(2\pi i)^2} \int_{\mathcal{L}_1'} \int_{\mathcal{L}_1} G(s,s') \left(\frac{\zeta(s+s'+1)}{\zeta(s+1)\zeta(s'+1)}\right)^k \frac{R^{s+s'}}{(ss')^{k+l+1}} dsds' \\
+ O_{c_1,c_2}\left(\exp(-c\sqrt{\log N})\right),
\end{aligned}$$

con $\mathcal{L}_1$ y $\mathcal{L}_1'$ los contornos verticales $c_0(\log U)^{-1} + it$ y $c_0(2\log U)^{-1} + it'$ truncados a $|t| \leq U$ y $|t'| \leq U/2$ respectivamente. Lo que hacemos ahora es mover $\mathcal{L}_1$ a la línea vertical $\mathcal{L}_2 : -c_0(\log U)^{-1} + it$, $|t| \leq U$. Al hacer esto, encontramos polos de orden $l+1$ y k en $s = 0$ y $s = -s'$ respectivamente.

Veremos ahora que la elección de c_0 puede efectuarse de modo tal que estos sean los únicos polos que encontremos. En efecto, de aquí en más, las partes reales de nuestras variables satisfacerán siempre $\sigma, \sigma' \geq -\frac{c_0}{\sqrt{\log N}}$ y las partes imaginarias $|t|, |t'| \leq U$. Eligiendo $c_0 \leq \bar{c}$ con $\bar{c}$ como en el Lema 3.2, tenemos

$$\begin{aligned}
-(\sigma+\sigma') &\leq \frac{2c_0}{\sqrt{\log N}} \\
&\leq \frac{4c_0}{\log 3 + \sqrt{\log N}} \\
&\leq \frac{4\bar{c}}{\log(2U+3)} \\
&\leq \frac{4\bar{c}}{\log((|t|+|t'|)+3)},
\end{aligned}$$

de donde se sigue que las variables s, s' y su suma $s+s'$ se encontrarán siempre en la región (3.3) en la cual el Lema 3.2 es aplicable.

Si escribimos $T_1 : \sigma + iU$, $-c_0(\log U)^{-1} \leq \sigma \leq c_0(\log U)^{-1}$ y $T_2 : \sigma - iU$, $-c_0(\log U)^{-1} \leq \sigma \leq c_0(\log U)^{-1}$, vemos que para $s \in T_1 \cup T_2$ y $s' \in \mathcal{L}_1'$

tenemos las desigualdades $|s| \geq U$, $|s+s'| \geq U/2$, $\log|t|, \log|t'| \leq \log U$ y $R^{\sigma+\sigma'} \leq \exp(3\sqrt{\log N}/2)$. Usando entonces $|T_1|, |T_2| = 2c_0(\log U)^{-1}$, obtenemos como antes las desigualdades

$$\begin{aligned}&\frac{1}{(2\pi i)^2}\int_{\mathcal{L}_1'}\int_{T_j} G(s,s')\left(\frac{\zeta(s+s'+1)}{\zeta(s+1)\zeta(s'+1)}\right)^k \frac{R^{s+s'}}{(ss')^{k+l+1}}dsds' \\ &\ll_{c_1,c_2} \left(\exp(-c\sqrt{\log N})\right)\end{aligned}$$

para $j = 1, 2$. Aquí hemos usado también que en la región en cuestión es

$$\begin{aligned}G(s,s') &\ll_{c_1} \exp\left(c_2(\log N)^{\frac{2c_0c_3}{\sqrt{\log N}}}\log\log\log N\right) \\ &\ll_{c_1} (\log\log N)^{c_2} \qquad (3.5)\end{aligned}$$

si N es suficientemente grande en términos de c_3.

Asimismo, si es $s \in \mathcal{L}_2$ y $s' \in \mathcal{L}_1'$, tenemos la desigualdad $|s+s'| \geq \frac{c_0}{2\sqrt{\log N}}$. Además, (3.5) seguirá valiendo y será

$$R^{\sigma+\sigma'} \leq R^{-\frac{c_0}{2\sqrt{\log N}}} \ll \exp(-c\sqrt{\log N}),$$

por lo que no necesitaremos cotas inferiores ni en $|s|$ ni en $|s'|$ para que el término de error sea del orden deseado. Obtenemos así

$$\begin{aligned}&\frac{1}{(2\pi i)^2}\int_{\mathcal{L}_1'}\int_{\mathcal{L}_2} G(s,s')\left(\frac{\zeta(s+s'+1)}{\zeta(s+1)\zeta(s'+1)}\right)^k \frac{R^{s+s'}}{(ss')^{k+l+1}}dsds' \\ &\ll_{c_1,c_2} \exp\left(-c\sqrt{\log N}\right)\int_{\mathcal{L}_1'}\frac{(\log(|t'|+3))^c}{|s'|^{k+l+1}}\int_{\mathcal{L}_2}\frac{(\log(|t|+3))^c}{|s|^{k+l+1}}dsds' \\ &\ll_{c_1,c_2} \exp\left(-c\sqrt{\log N}\right)\end{aligned}$$

en donde hemos evaluado la integral interna de la misma forma que antes.

Uniendo las anteriores desigualdades, nuestra tarea se reduce entonces a estimar

$$\begin{aligned}&\frac{1}{(2\pi i)^2}\int_{\mathcal{L}_1'}\left\{\mathrm{Res}_{s=0}+\mathrm{Res}_{s=-s'}\right\}ds' \\ &\qquad + O_{c_1,c_2}\left(\exp\left(-c\sqrt{\log N}\right)\right) \qquad (3.6)\end{aligned}$$

donde por supuesto, $\mathrm{Res}_{s=0}$ hace referencia al residuo de

$$G(s,s')\left(\frac{\zeta(s+s'+1)}{\zeta(s+1)\zeta(s'+1)}\right)^k \frac{R^{s+s'}}{(ss')^{k+l+1}}$$

en $s = 0$ y lo análogo para $\mathrm{Res}_{s=-s'}$.

Procedemos a estimar

$$\int_{\mathcal{L}_1'} \{\mathrm{Res}_{s=-s'}\}\, ds'. \tag{3.7}$$

Para esto, reescribimos el residuo como

$$\frac{1}{2\pi i}\int_{C(s')} G(s,s')\left(\frac{\zeta(s+s'+1)}{\zeta(s+1)\zeta(s'+1)}\right)^k \frac{R^{s+s'}}{(ss')^{k+l+1}} ds ds'$$

con $C(s'): |s+s'| = (\log N)^{-1}$. Es claro que en $C(s')$ sigue valiendo (3.5), como también que por el Lema 3.2 es $\zeta(s+s'+1) \ll \log N$. Puesto que $\zeta(1+s)$ tiene un polo simple en $s=0$, es $(s\zeta(s+1))^{-1} \ll 1$ para valores suficientemente chicos de $|s|$. Notando además que es $|s'| \ll |s| \ll |s'|$ en $C(s')$, del Lema 3.2 tenemos entonces $(s\zeta(s+1))^{-1} \ll (|s'|+1)^{-1}\log(|s'|+3)$. Por último, observamos que es $R^{s+s'} \ll 1$ y $|C(s')| \ll (\log N)^{-1}$. Deducimos así la cota

$$\begin{aligned}\mathrm{Res}_{s=-s'} \ll_{c_1} (\log N)^{-1}(\log\log N)^{c_2}\left(\log N + \log((\log N)^{-1}+3)\right)^k \\ \times \left(\frac{\log(|s'|+2)}{|s'|+1}\right)^{2k} |s'|^{-2l-2}.\end{aligned}$$

Integrando esto en $\mathcal{L}_1'$ vemos que (3.7) es

$$\ll_{c_1} (\log N)^{k+l}(\log\log N)^{c_2}.$$

Esto nos permite expresar a (3.6) como

$$\frac{1}{2\pi i}\int_{\mathcal{L}_1'} \{\mathrm{Res}_{s=0}\}\, ds' + O_{c_1,c_2}((\log N)^{k+l}(\log\log N)^{c_2}). \tag{3.8}$$

Para evaluar esta última integral definimos

$$Z(s,s') = G(s,s')\left(\frac{(s+s')\zeta(s+s'+1)}{s\zeta(s+1)s'\zeta(s'+1)}\right)^k, \tag{3.9}$$

que es regular en un entorno del $(0,0)$. Notar además que en la región (3.3) obtenemos fácilmente la cota

$$Z(s,s') \ll_{c_1} \exp\left(c_2(\log N)^{-2\overline{\sigma}c_3}\log\log\log N\right)\log^{3k}(|t|+|t'|+3), \tag{3.10}$$

con $\overline{\sigma} = \min(\sigma,\sigma',0)$.

Recordando que teníamos un polo de orden $l+1$ en $s=0$, podemos escribir

$$\mathrm{Res}_{s=0} = \frac{R^{s'}}{s'^{l+1}}\frac{1}{l!}\left(\frac{\partial}{\partial s}\right)^l_{s=0}\left\{\frac{Z(s,s')}{(s+s')^k}R^s\right\}.$$

Insertamos ahora esto en (3.8) y movemos $\mathcal{L}_1'$ a $\mathcal{L}_2'$: $-c_0(2\log U)^{-1} + it'$, $|t'| \le U/2$, encontrando un único polo en el orígen. Si consideramos T_1' : $\sigma' +$

$iU/2, -c_0(2\log U)^{-1} \leq \sigma' \leq c_0(2\log U)^{-1}$ y $T_2' : \sigma' - iU/2, -c_0(2\log U)^{-1} \leq \sigma' \leq c_0(2\log U)^{-1}$, entonces para $s' \in T_1' \cup T_2'$ y $C_0 : |s| = (\log N)^{-1}$ tenemos

$$\left(\frac{\partial}{\partial s}\right)^l_{s=0} \left\{\frac{Z(s,s')}{(s+s')^k} R^s\right\} = \frac{1}{2\pi i}\int_{C_0} \frac{Z(s,s')}{(s+s')^k}\frac{R^s}{s^{l+1}} ds \ll_{c_1,c_2} \exp\left(-c\sqrt{\log N}\right),$$

en donde hemos usado $|s+s'| \gg U$, $R^s \ll 1$ y (3.10). Obtenemos

$$\frac{1}{2\pi i}\int_{T_j'} \mathrm{Res}_{s=0} ds' \ll_{c_1,c_2} \exp\left(-c\sqrt{\log N}\right),$$

donde hemos usado ahora $R^{s'} s'^{-l-1} \ll \exp\left(-c\sqrt{\log N}\right)$. De la misma forma, al ser $R^{s'} \leq \exp\left(-c_0\sqrt{\log N}/2\right)$ para $s' \in \mathcal{L}_2'$, obtenemos también

$$\frac{1}{2\pi i}\int_{\mathcal{L}_2'} \mathrm{Res}_{s=0} ds' \ll_{c_1,c_2} \exp\left(-c\sqrt{\log N}\right).$$

Juntando ahora estas estimaciones podemos reescribir a (3.8) como

$$\begin{aligned} &\mathrm{Res}_{s'=0}\mathrm{Res}_{s=0} + O_{c_1,c_2}((\log N)^{k+l}(\log\log N)^{c_2}) \\ &= \frac{1}{(2\pi i)^2}\int_{C_2}\int_{C_1} \frac{Z(s,s')R^{s+s'}}{(s+s')^k (ss')^{l+1}} ds ds' \\ &+ O_{c_1,c_2}((\log N)^{k+l}(\log\log N)^{c_2}), \end{aligned} \tag{3.11}$$

con $C_1 : |s| = \rho$, $C_2 : |s'| = 2\rho$, para cierta constante pequeña $\rho > 0$. Escribiendo $s' = s\xi$ vemos que esta integral doble es igual a

$$\frac{1}{(2\pi i)^2}\int_{C_3}\int_{C_1} \frac{Z(s,s\xi)R^{s(\xi+1)}}{(\xi+1)^k \xi^{l+1} s^{k+2l+1}} ds d\xi,$$

donde C_3 es el círculo $|\xi| = 2$.

Ahora bien, es

$$\begin{aligned} &\frac{1}{2\pi i}\int_{C_1} \frac{Z(s,s\xi)R^{s(\xi+1)}}{s^{k+2l+1}} ds = \frac{1}{(k+2l)!}\left(\frac{\partial}{\partial s}\right)^{k+2l}_{s=0} \left\{Z(s,s\xi)R^{s(\xi+1)}\right\} \\ &= \frac{1}{(k+2l)!}\sum_{m+n=k+2l}\binom{k+2l}{m}\left(\frac{\partial}{\partial s}\right)^m_{s=0}\{Z(s,s\xi)\}\left(\frac{\partial}{\partial s}\right)^n_{s=0}\left\{R^{s(\xi+1)}\right\}. \end{aligned} \tag{3.12}$$

Para evaluar esto consideramos $C_4 : |s| = (\sqrt{\log N})^{-1}$ en donde tenemos de (3.10) la cota $Z(s,s\xi) \ll_{c_1} (\log\log N)^{c_2}$. Luego, para $m > 0$ es

$$\left(\frac{\partial}{\partial s}\right)^m_{s=0}\{Z(s,s\xi)\} \ll \int_{C_4}\frac{Z(s,s\xi)}{s^{m+1}} ds \ll_{c_1} (\log N)^{\frac{m}{2}}(\log\log N)^{c_2}.$$

Puesto que claramente tenemos también

$$\left(\frac{\partial}{\partial s}\right)^n_{s=0}\left\{R^{s(\xi+1)}\right\} = (\xi+1)^n(\log R)^n,$$

concluimos que (3.12) es igual a

$$\frac{Z(0,0)}{(k+2l)!}(\xi+1)^{k+2l}(\log R)^{k+2l} + o_{c_1,c_2}\left((\log N)^{k+2l}\right).$$

Gracias a esto, podemos reescribir (3.11) como

$$\frac{Z(0,0)}{2\pi i(k+2l)!}(\log R)^{k+2l}\int_{C_3}\frac{(\xi+1)^{2l}}{\xi^{l+1}}d\xi + o_{c_1,c_2}\left((\log N)^{k+2l}\right). \tag{3.13}$$

Ahora bien, como es

$$\frac{1}{l!}\left(\frac{\partial}{\partial\xi}\right)^l_{\xi=0}\left\{(\xi+1)^{2l}\right\} = \binom{2l}{l},$$

obtenemos que (3.13) es igual a

$$Z(0,0)\frac{(\log R)^{k+2l}}{(k+2l)!}\binom{2l}{l} + o_{c_1,c_2}\left((\log N)^{k+2l}\right). \tag{3.14}$$

Luego, para finalizar la demostración, basta con notar que de (3.9) y el hecho de que $\zeta(s)$ tiene en $s=1$ un polo simple de residuo igual a 1, se sigue que es $Z(0,0) = G(0,0)$. Juntando lo anterior obtenemos entonces que (3.14) (y en consecuencia (3.2)) es igual a

$$(1+o_{c',c_1,c_2}(1))G(0,0)\frac{(\log R)^{k+2l}}{(k+2l)!}\binom{2l}{l},$$

en donde hemos usado que es $R = N^{c'}$. El Lema 3.1 queda así demostrado. □

3.2. El caso general

Pasamos ahora al caso general.

Demostración del Lema 2.2. Habiendo demostrado el caso $m = 1$ en el Lema 3.1, aplicaremos inducción. Consideramos entonces $m+1$ y el resultado demostrado para $m \geq 1$. Deseamos evaluar

$$\frac{1}{(2\pi i)^{2(m+1)}}\int_{(1)}\cdots\int_{(1)} G(s,s')\prod_{j=1}^{m+1}\left(\frac{\zeta(s_j+s'_j+1)}{\zeta(s_j+1)\zeta(s'_j+1)}\right)^{k_j}\frac{R^{s_j+s'_j}}{(s_js'_j)^{k_j+l_j+1}}dsds' \tag{3.15}$$

con $k_1, \dots, k_{m+1}, l_1, \dots, l_{m+1} > 0$ enteros positivos y las $2(m+1)$ variables complejas $s = (s_1, \dots, s_{m+1})$ y $s' = (s'_1, \dots, s'_{m+1})$, donde además es

$$G(s_1, \dots, s_{m+1}, s'_1, \dots, s'_{m+1}) < c_1 \exp\left(c_2 (\log N)^{-2\overline{\sigma}c_3} \log\log\log N\right),$$

con $\overline{\sigma} = \min(\Re(s_1), \dots, \Re(s_{m+1}), \Re(s'_1), \dots, \Re(s'_{m+1}), 0)$ y para ciertas constantes c_1, c_2, c_3. Fijando entonces cualquier elección de $s_1, \dots, s_m, s'_1, \dots, s'_m$ en los correspondientes contornos de integración (en particular, con parte real igual a 1), vemos que considerando a $G(s_1, \dots, s_m, s_{m+1}, s'_1, \dots, s'_m, s'_{m+1})$ como función de s_{m+1}, s'_{m+1} se tiene

$$G(s_1, \dots, s_{m+1}, s'_1, \dots, s'_{m+1}) < c_1 \exp\left(c_2 (\log N)^{-2\overline{\sigma}_{m+1}c_3} \log\log\log N\right),$$

con $\overline{\sigma}_{m+1} = \min(\Re(s_{m+1}), \Re(s'_{m+1}), 0)$.

Aplicando entonces el Lema 3.1 a la integral doble de más adentro (la que depende de las variables s_{m+1}, s'_{m+1}), vemos que (3.15) es igual a

$$\begin{aligned}
&\frac{(\log R)^{k_{m+1}+2l_{m+1}}}{(k_{m+1}+2l_{m+1})!}\binom{2l_{m+1}}{l_{m+1}} \\
&\times \frac{1}{(2\pi i)^{2m}} \int_{(1)} \dots \int_{(1)} (G(s_1, \dots, s_m, 0, s'_1, \dots, s'_m, 0) + o_{c',c_1,c_2}(1)) \\
&\times \prod_{j=1}^{m} \left(\frac{\zeta(s_j + s'_j + 1)}{\zeta(s_j+1)\zeta(s'_j+1)}\right)^{k_j} \frac{R^{s_j+s'_j}}{(s_j s'_j)^{k_j+l_j+1}} ds_1 ds'_1 \dots ds_m ds'_m.
\end{aligned}$$

Aplicando a esto la hipótesis inductiva, obtenemos

$$\begin{aligned}
&\frac{1}{(2\pi i)^{2(m+1)}} \int_{(1)} \dots \int_{(1)} G(s, s') \prod_{j=1}^{m+1} \left(\frac{\zeta(s_j + s'_j + 1)}{\zeta(s_j+1)\zeta(s'_j+1)}\right)^{k_j} \frac{R^{s_j+s'_j}}{(s_j s'_j)^{k_j+l_j+1}} ds ds' \\
&= (G(0, \dots, 0) + o_{c',c_1,c_2,m+1}(1)) \prod_{j=1}^{m+1} \frac{(\log R)^{k_j+2l_j}}{(k_j+2l_j)!}\binom{2l_j}{l_j}
\end{aligned}$$

y el Lema 2.2 queda así demostrado. □

Capítulo 4

Distancias acotadas entre primos

Nos embarcaremos ahora en la tarea de demostrar los Teoremas 1.2 y 1.3. Como fue mencionado en §1, nuestro plan para lograr tal propósito será probar que los primos son densos en las tuplas de casi primos (tuplas con una cantidad pequeña de factores primos acumulados); de hecho, asumiendo la conjetura de Elliot-Halberstam, demostraremos que incluso son lo suficientemente densos como para garantizar que muchas tuplas de casi primos contengan al menos dos números primos.

Una de las herramientas fundamentales que utilizaremos ya fue desarrollada y es la Proposición 2.1. Esto nos da una estimación de $\Lambda_R(n; \mathcal{H}, k+l)$ en intervalos. Como hemos visto, uno espera que la distribución de esta función coincida con la distribución de las tuplas $n + \mathcal{H}$ con a lo sumo $k + l$ divisores primos distintos. Teniendo presente esto, es natural conjeturar que si es $h \in \mathcal{H}$, entonces $n + h$ tendrá una probabilidad mucho más alta de ser primo cuando $\Lambda_R(n; \mathcal{H}, k+l)$ sea no nulo.

Resulta ser que esta conjetura puede ser efectivamente probada. La clave para poder lograr esto reside en el Teorema de Bombieri-Vinogradov, que nos permitirá manejar los términos de error que surgen al emplear los métodos de criba y que ahora tienen la dificultad extra de tener que lidiar con la primalidad de uno de los miembros de la tupla.

Éste capítulo es una exposición de los resultados e ideas presentes en [17] y [18].

4.1. El nivel de distribución de los números primos

A lo largo de éste capítulo denotaremos por $\vartheta(n)$ a la función que vale $\log n$ si n es primo y 0 en otro caso. La siguiente definición será de fundamental importancia.

Definición 4.1 (Nivel de distribución)**.** *Sea $0 < \theta \leq 1$. Decimos que los primos tienen nivel de distribución θ si para todo $A > 0$ y todo $0 < \theta' < \theta$ se satisface*

$$\sum_{q \leq x^{\theta'}} \max_{y \leq x} \max_{\substack{a \\ (a,q)=1}} \left| \vartheta(y; a, q) - \frac{y}{\varphi(y)} \right| \ll_A \frac{x}{(\log x)^A}, \tag{4.1}$$

donde es

$$\vartheta(y; a, q) = \sum_{\substack{y < n \leq 2y \\ n \equiv a (mod\ q)}} \vartheta(n).$$

Nota. A veces suele decirse también que los primos tienen nivel de distribución θ si (4.1) se satisface con $\theta' = \theta$ (i.e. para todo $\theta' \leq \theta$). Sin embargo, en éste trabajo nos restringiremos a la Definición 4.1.

Enunciamos ahora el mencionado Teorema de Bombieri-Vinogradov [4, Teorema 17].

Proposición 4.1 (Teorema de Bombieri-Vinogradov)**.** *Para todo $A > 0$ tenemos*

$$\sum_{q \leq \sqrt{x}(\log x)^{-B}} \max_{y \leq x} \max_{\substack{a \\ (a,q)=1}} \left| \vartheta(y; a, q) - \frac{y}{\varphi(y)} \right| \ll_A \frac{x}{(\log x)^A}, \tag{4.2}$$

para una cierta constante $B = B(A)$. En particular, los primos tienen nivel de distribución $1/2$.

Para tener una idea de la importancia de éste resultado, notar que la Hipótesis Generalizada de Riemann implica la estimación

$$\vartheta(x; a, q) = \frac{x}{\varphi(x)} + O(x^{1/2}(\log x)^2), \tag{4.3}$$

de modo que el Teorema de Bombieri-Vinogradov nos provee incondicionalmente una especie de Hipótesis Generalizada de Riemann en promedio (en efecto, insertando (4.3) en el lado izquierdo de (4.2) obtenemos un error del mismo orden).

La característica sobresaliente de poseer información sobre el nivel de distribución de los primos tiene que ver con poder controlar los errores en las aproximaciones *uniformemente* a lo largo de diversos módulos q. La pregunta natural a formular es en consecuencia hasta que magnitud de módulo es posible extender las estimaciones de éste tipo. Originalmente se consideró la posibilidad de que la suma de (4.2) pudiese extenderse hasta todo $q \leq x(\log x)^{-B}$ manteniendo un error del mismo orden, sin embargo fue demostrado por John Friedlander y Andrew Granville en 1989 (elaborando sobre ideas previas de Maier) que tal afirmación es falsa [14]. La más conocida conjetura a éste respecto es la siguiente.

Conjetura 4.1 (Conjetura de Elliot-Halberstam [9])**.** *Los primos tienen nivel de distribución* 1.

Nota. La conjetura original de Peter Elliot y Heini Halberstam (que era además consecuencia de una conjetura de Hugh Montgomery) era en realidad más fuerte que esta afirmación, resultando sin embargo falsa como consecuencia del trabajo de Friedlander y Granville citado arriba.

Es necesario mencionar que tal conjetura va mucho más allá de lo previsto por la Hipótesis Generalziada de Riemann y en consecuencia (a diferencia de lo que sucede con gran parte de las conjeturas sobre la distribución de los números primos) se aparece razonable al menos cuestionar la validez de tal afirmación. Por otra parte, la deducción de la Proposición 4.1 partiendo de la Hipótesis Generalizada de Riemann es totalmente inmediata (en efecto, tal deducción proviene de aplicar burdamente la desigualdad triangular a una suma sobre los ceros de L-funciones, suma en la cual es coherente suponer que exista en realidad un considerable grado de cancelación) por lo cual es muy válido esperar que el nivel de distribución de los primos sea mayor que $1/2$ aún cuando tengamos alguna duda sobre la fuerza total de la Conjetura 4.1. Vale mencionar como último dato a éste respecto que Montgomery [35] conjeturó que (4.3) puede mejorarse hasta

$$\vartheta(x;a,q) = \frac{x}{\varphi(x)} + O_\varepsilon(q^{-\frac{1}{2}}x^{\frac{1}{2}+\varepsilon}), \tag{4.4}$$

para cualquier $\varepsilon > 0$. De ser cierta esta conjetura, insertando (4.4) en (4.1) recuperaríamos la Conjetura 4.1.

4.2. Densidad en tuplas

Veremos más adelante que el Teorema de Bombieri-Vinogradov nos permitirá demostrar que para todo $\varepsilon > 0$ existen infinitos primos p para los cuales el siguiente primo se encuentra a una distancia menor que $\varepsilon \log p$. Veremos también que si los primos tienen nivel de distribución $\theta > 1/2$ entonces existirán infinitos números primos p para los cuales $p + C$ también será primo, con $C = C(\theta)$ una constante absoluta.

Para lograr tal objetivo necesitaremos la siguiente proposición, que para toda tupla $\mathcal{H} = \{h_1, \ldots, h_k\} \subseteq [1, H] \cap \mathbb{Z}$ con

$$H \ll \log N \tag{4.5}$$

nos permite calcular explícitamente la densidad en $[N, 2N]$ de $\vartheta(n + h)$, $h \in \mathcal{H}$, respecto a $\Lambda_R(n; \mathcal{H}, k + l)$.

Proposición 4.2. *Sean $k, l \geq 1$ enteros arbitrarios y N un parámetro entero. Sea*

$$\mathcal{H} = \{h_1, \ldots, h_k\} \subseteq [1, H] \cap \mathbb{Z}$$

y supongamos que vale (4.5). Supongamos además que los primos tienen nivel de distribución θ. Entonces, si es

$$R = N^C \tag{4.6}$$

con $0 < C < \theta/2$, tenemos

$$\sum_{N<n\leq 2N} \vartheta(n+h)\Lambda_R^2(n;\mathcal{H},k+l) \tag{4.7}$$

$$= \begin{cases} (\mathfrak{G}(\mathcal{H}\cup\{h\})+o(1))\,\frac{N}{(k+2l)!}\binom{2l}{l}(\log R)^{k+2l} & \text{si } h\notin\mathcal{H}, \\ (\mathfrak{G}(\mathcal{H})+o(1))\,\frac{N}{(k+2l+1)!}\binom{2(l+1)}{l+1}(\log R)^{k+2l+1} & \text{si } h\in\mathcal{H}. \end{cases} \tag{4.8}$$

Demostración. Al igual que antes, c denotará una constante positiva que dependerá a lo sumo de k y l y cuyo valor podrá variar en cada ocurrencia. La primera observación a realizar es notar que si h pertenece a $\mathcal{H}$ entonces

$$\sum_{N<n\leq 2N} \vartheta(n+h)\Lambda_R^2(n;\mathcal{H},k+l) \tag{4.9}$$

es igual a

$$\sum_{N<n\leq 2N} \vartheta(n+h)\Lambda_R^2(n;\mathcal{H}-\{h\},k+l). \tag{4.10}$$

En efecto, puesto que $\vartheta(n+h)$ es no nulo únicamente si $n+h$ es primo y dado que en el intervalo en cuestión es $n+h > R$ se sigue en tal caso que los divisores de $P(n,\mathcal{H})$ que incluyen al primo $n+h$ no contribuyen a la suma que define a $\Lambda_R^2(n;\mathcal{H},k+l)$.

La igualdad entre (4.9) y (4.10) será de fundamental importancia, puesto que es lo que nos permitirá comparar el crecimiento de la esperanza de $\vartheta(n+h)$ cuando h pertenece a una tupla de casi primos. De todas formas, es válido preguntarse cómo tal identidad puede poseer semejante importancia cuando no es satisfecha por la función $\Lambda(n;\mathcal{H},k+l)$ a la que en teoría $\Lambda_R(n;\mathcal{H},k+l)$ aproxima. La razón para esto es simplemente que al estar truncada la suma a los divisores menores que R, lo que en realidad se espera que $\Lambda(n;\mathcal{H},k+l)$ modele es el comportamiento de las tuplas $n+\mathcal{H}$ con a lo sumo $k+l$ factores primos distintos *pequeños*. Es en tales circunstancias que la anterior identidad es válida y también nuestro plan original, puesto que sigue siendo coherente esperar que la probabilidad de $n+h$ de ser primo incremente si $n+\mathcal{H}$ es una tupla con pocos factores primos pequeños.

Gracias a las observaciones anteriores podemos asumir de ahora en más $h\neq\mathcal{H}$. Expandiendo el cuadrado en el lado izquierdo de (4.7) vemos que esto es igual a

$$\sum_{d,d'} \lambda_R(d;k+l)\lambda_R(d';k+l) \sum_{\substack{[d,d']|P(n;\mathcal{H}) \\ N<n\leq 2N}} \vartheta(n+h), \tag{4.11}$$

donde $\lambda_R(d;k+l)$ fue definido en (2.3). Notar además que es

$$\sum_{\substack{[d,d']|P(n;\mathcal{H})\\ N<n\leq 2N}} \vartheta(n+h) = \sum_{b\in\Omega([d,d'])} \delta((b+h,[d,d']))\vartheta\left(N;b+h,\left[d,d'\right]\right)+O(\log N),$$

donde $\delta(x)$ es igual a 1 si es $x=0$ y se anula en otro caso, puesto que si $[d,d']$ no es coprimo con $b+h$ es $\vartheta\left(N;b+h,[d,d']\right)=0$. Usando $\log N \ll \log R$ concluimos que (4.11) es igual a

$$\sum_{d,d'} \lambda_R(d;k+l)\lambda_R(d';k+l) \tag{4.12}$$
$$\times \sum_{b\in\Omega([d,d'])} \delta((b+h,[d,d']))\vartheta\left(N;b+h,\left[d,d'\right]\right) + O(R^2(\log R)^c),$$

y el error se puede estimar como $o(N)$ usando (4.6).

Para evaluar (4.12) consideramos primero los pares d,d' con $\tau_3([d,d'])<(\log N)^{A/B}$ (aquí τ_3 está definido como en el Lema 2.1) para ciertas constantes $A=A(k,l)$ y $B=B(k,l)$ que serán especificadas más adelante y notamos que es $|\{d,d' : [d,d']=d\}|=\tau_3(d)$. Puesto que es $\tau_1(d)\leq\tau_2(d)\leq\tau_3(d)$ y puesto que es también $\tau_l(m)\leq\tau_l(n)$ si $m|n$ y para todo $l>0$, obtenemos entonces $\tau_k(d)\leq(\tau_3(d))^c$ (subdividiendo sucesivamente los k factores en subconjuntos de 1, 2 o 3 factores).

Con las observaciones de arriba obtenemos la cota $|\Omega([d,d'])|\leq\tau_k([d,d'])\leq(\log N)^{cA/B}$. Debido a esto, si para los pares d,d' en cuestión reemplazamos $\vartheta\left(N;b+h,[d,d']\right)$ por $\frac{N}{\varphi([d,d'])}$ incurrimos en un error de a lo sumo

$$\begin{aligned}
&\sum_{\substack{d,d'\leq R\\ \tau_3([d,d'])<(\log N)^{A/B}}} \lambda_R(d;k+l)\lambda_R(d';k+l)|\Omega([d,d'])| \\
&\times \max_{\substack{a\\ (a,[d,d'])=1}} \left|\vartheta\left(N;a,\left[d,d'\right]\right)-\frac{N}{\varphi([d,d'])}\right| \\
&\ll (\log R)^{2(k+l)}(\log N)^{cA/B} \\
&\times \sum_{\substack{d\leq R^2\\ \tau_3(d)<(\log N)^{A/B}}} \tau_3(d) \max_{\substack{a\\ (a,d)=1}} \left|\vartheta\left(N;a,d\right)-\frac{N}{\varphi(d)}\right| \\
&\ll (\log R)^{2(k+l)}(\log N)^{cA/B}\frac{N}{(\log N)^A} \\
&\ll \frac{N}{(\log N)^{A/2B}}
\end{aligned}$$

si A y B se eligen adecuadamente y en donde hemos usado la Proposición 4.1.

Análogamente, efectuando el mismo reemplazo para los pares d, d' con $\tau_3([d,d']) \geq (\log N)^{A/B}$ y usando la estimación cruda

$$\max_{\substack{a \\ (a,[d,d'])=1}} \left| \vartheta\left(N; a, \left[d, d'\right]\right) - \frac{N}{\varphi([d,d'])} \right| \ll \frac{N}{[d,d']} \log N$$

vemos que el error estará acotado por

$$(\log R)^{2(k+l)} N \log N \sum_{d \leq R^2} \tau_3(d) \frac{|\Omega(d)|}{d} \left(\frac{\tau_3(d)}{(\log N)^{A/B}} \right),$$

en donde hemos usado que el factor entre paréntesis es mayor o igual a 1 en los términos relevantes de la suma. Esto a su vez, va a ser

$$\ll (\log R)^{2(k+l)} \frac{N \log N}{(\log N)^{A/B}} \sum_{d \leq R^2} \frac{(\tau_3(d))^c}{d}. \tag{4.13}$$

Finalmente, del Lema 2.1 se sigue inmediatamente usando sumación parcial que (4.13) es a lo sumo

$$\begin{aligned} &\ll (\log R)^{2(k+l)} \frac{N \log N}{(\log N)^{A/B}} (\log N)^c \\ &\ll \frac{N}{(\log N)^{A/2B}} \end{aligned}$$

si A y B se eligen adecuadamente.

Concluimos entonces de los argumentos anteriores que (4.12) es igual a

$$N\mathcal{T} + o(N)$$

con

$$\mathcal{T} = \sum_{d,d'} \frac{\lambda_R(d, k+l)\lambda_R(d', k+l)}{\varphi([d,d'])} \sum_{b \in \Omega([d,d'])} \delta((b+h, [d,d'])).$$

Notar que por el Teorema chino del resto la suma interna de $\mathcal{T}$ es igual a

$$\prod_{p|[d,d']} \left(\sum_{b \in \Omega(p)} \delta((b+h, p)) \right) = \prod_{p|[d,d']} (|\Omega^+(p)| - 1),$$

donde Ω^+ corresponde al conjunto $\mathcal{H}^+ = \mathcal{H} \cup \{h\}$. En efecto, como $b \in \Omega(p)$, $\delta((b+h,p))$ es nulo si y sólo si $-h \in \Omega(p)$; luego, si esto no sucede la suma es igual a $|\Omega(p)| = |\Omega^+(p)| - 1$, mientras que si sucede es en cambio $|\Omega(p)| = |\Omega^+(p)|$ y la suma devuelve también $|\Omega^+(p)| - 1$, puesto que un término será nulo.

Al igual que en la Proposición 2.1, usando esta vez $\varphi(p) = p - 1$ y la correspondiente multiplicatividad de la función de Euler, obtenemos la expresión

$$\mathcal{T} = \frac{1}{(2\pi i)^2} \int_{(1)} \int_{(1)} \prod_p \left(1 - \frac{|\Omega^+(p)| - 1}{p-1} \left(\frac{1}{p^s} + \frac{1}{p^{s'}} - \frac{1}{p^{s+s'}}\right)\right) \times \frac{R^{s+s'}}{(ss')^{k+l+1}} ds ds'.$$

Al igual que en la Proposición 2.1, el siguiente paso será considerar

$$G(s,s') = \prod_p \left(1 - \frac{|\Omega^+(p)| - 1}{p-1} \left(\frac{1}{p^s} + \frac{1}{p^{s'}} - \frac{1}{p^{s+s'}}\right)\right) \left(\frac{\zeta(s+1)\zeta(s'+1)}{\zeta(s+s'+1)}\right)^k$$

y aplicar el Lema 2.2. La verificación de que $G(s,s')$ así definido satisface las correspondientes hipótesis es exactamente igual a lo hecho en la Proposición 2.1 en el caso en el que la tupla $\mathcal{H}^+$ es admisible, notando que es $|\Omega^+(p)| = k+1$ para $p > H$. Observamos sin embargo que no es necesariamente cierto que $\mathcal{H}^+$ es admisible aún cuando estamos asumiendo que la tupla original $\mathcal{H}$ lo es. De todas formas, notar que todo primo p con $|\Omega^+(p)| = p$ ha de satisfacer $p \leq k+1$. Pero entonces, puesto que la cantidad de tales primos está acotada su tratamiento es trivial y en consecuencia las estimaciones se pueden llevar a cabo nuevamente en la forma dada en la Proposición 2.1. Concluimos entonces que el Lema 2.2 es aplicable a nuestra situación y en consecuencia el resultado se sigue en el caso en que es $h \notin \mathcal{H}$.

Finalmente, si es $h \in \mathcal{H}$, usando la observación al comienzo de la demostración vemos que considerando $\mathcal{H} - \{h\}$ y aplicando la traslación $k \mapsto k-1$, $l \mapsto l+1$, el razonamiento anterior vuelve a ser válido. Esto concluye la demostración. □

4.3. Demostración del Teorema 1.3

Veremos ahora que combinando la Proposición 2.1 y la Proposición 4.2 podemos deducir el Teorema 1.3.

Demostración del Teorema 1.3. Sea $\mathcal{H} = \{h_1, \ldots, h_k\}$ una tupla admisible. Consideraremos la evaluación de

$$\mathbb{E}_{N<n\leq 2N} \left(\Lambda_R^2(n; \mathcal{H}, k+l) \left(\sum_{j=1}^k \vartheta(n+h_j) - \log 3N\right)\right). \tag{4.14}$$

Lo importante a notar aquí es que el término en paréntesis es positivo si y sólo si la tupla $n + \mathcal{H}$ contiene al menos dos números primos (puesto que es $n + h_j < 3N$ en el intervalo en cuestión). Debido a esto, si podemos mostrar

que (4.14) es positivo, esto implicará la existencia de algún $N < n \leq 2N$ para el cual $n + \mathcal{H}$ contendrá al menos dos números primos.

Para evaluar esta expresión consideramos la Proposición 2.1 y la Proposición 4.2. De ahora en más asumimos que los primos tienen nivel de distribución θ y definimos $R = N^{\frac{\theta'}{2}}$ para cierto $0 < \theta' < \theta$. Insertando las mencionadas estimaciones obtenemos que (4.14) es igual a

$$\begin{aligned}&(\mathfrak{G}(\mathcal{H}) + o(1))\frac{k}{(k+2l+1)!}\binom{2(l+1)}{l+1}(\log R)^{k+2l+1}\\ &-(\mathfrak{G}(\mathcal{H}) + o(1))\frac{\log 3N}{(k+2l)!}\binom{2l}{l}(\log R)^{k+2l}\end{aligned}$$

y esto es igual a

$$\begin{aligned}&\left(\frac{k}{k+2l+1}\frac{2(2l+1)}{l+1}\log R - \log 3N + o(\log N)\right)\\ &\times \mathfrak{G}(\mathcal{H})\frac{(\log R)^{k+2l}}{(k+2l)!}\binom{2l}{l}\\ &= \left(\frac{k}{k+2l+1}\frac{2(2l+1)}{l+1}\frac{\theta'}{2} - 1 + o(1)\right)\\ &\times \log N\mathfrak{G}(\mathcal{H})\frac{(\log R)^{k+2l}}{(k+2l)!}\binom{2l}{l}.\end{aligned}$$

Tomando $l = \lfloor\sqrt{k}\rfloor$, tenemos

$$\lim_{k\to\infty}\frac{k}{k+2l+1}\frac{2(2l+1)}{l+1} = 4. \tag{4.15}$$

Debido a esto, si suponemos $\theta > 1/2$, existirá una constante $C = C(\theta)$ tal que tomando $k > C$, $l = \lfloor\sqrt{k}\rfloor$, θ' suficientemente cercano a θ y N suficientemente grande en términos de k, será

$$\left(\frac{k}{k+2l+1}\frac{2(2l+1)}{l+1}\frac{\theta'}{2} - 1 + o(1)\right) > 0.$$

En particular, tendremos en tal caso que (4.14) será positivo. Concluimos entonces que para toda tupla $\mathcal{H} = \{h_1, \ldots, h_k\}$ con $k > C$ existirá un valor de n en $[N, 2N]$, con N suficientemente grande, para el cual $n + \mathcal{H}$ contendrá al menos dos números primos. Obtenemos así la primer afirmación del teorema.

Para ver la afirmación restante consideramos la tupla de 7 elementos $\mathcal{H} = \{0, 2, 6, 8, 12, 18, 20\}$. Es sencillo ver que tal tupla es admisible. Suponemos además la veracidad de la Conjetura 4.1, de modo que podemos reemplazar a θ' en los razonamientos de arriba por cualquier constante menor que 1. En

particular, consideraremos $\theta' > 20/21$. Entonces, tomando $l = 1$ obtenemos por los argumentos anteriores que (4.14) es igual a

$$\left(\frac{21}{20}\theta' - 1 + o(1)\right) \log N \mathfrak{G}(\mathcal{H}) \frac{(\log R)^{k+2l}}{(k+2l)!} \binom{2l}{l} > 0,$$

por la elección de θ' y asumiendo N suficientemente grande. Concluimos entonces que para tal N existirá $N < n \leq 2N$ para el cual $n+\mathcal{H}$ contendrá al menos dos números primos p y q. En particular, será $|q-p| \leq 20$. El Teorema 1.3 queda así demostrado. □

4.4. Demostración del Teorema 1.2

El objetivo restante será demostrar el Teorema 1.2. En esta ocasión el método empleado arriba no nos será de utilidad así enunciado, puesto que incondicionalmente sólo podemos asumir que los primos tienen nivel de distribución 1/2. El método a emplear para superar tal obstáculo será considerar todas las k-tuplas $\{h_1, \ldots, h_k\} \subseteq [1, H] \cap \mathbb{Z}$ en forma simultánea. Tal idea fue originalmente propuesta por Granville y Kannan Soundararajan. Para aplicarla necesitaremos el siguiente resultado demostrado por Patrick Gallagher [16].

Lema 4.1 (Lema de Gallagher)**.** *Tenemos*

$$\sum_{\substack{\mathcal{H} \subseteq [1,H] \\ |\mathcal{H}|=k}} \mathfrak{G}(\mathcal{H}) = (1 + o(1))H^k,$$

a medida que H tiende al infinito.

Notar que las permutaciones de $\{h_1, \ldots, h_k\}$ son consideradas en el lado izquierdo del enunciado.

Demostración del Teorema 1.2. Consideraremos la evaluación de

$$\sum_{\substack{\mathcal{H} \subseteq [1,H] \\ |\mathcal{H}|=k}} \mathbb{E}_{N<n\leq 2N} \left(\left(\sum_{h \leq H} \vartheta(n+h) - \log 3N \right) \Lambda_R^2(n; \mathcal{H}, k+l) \right). \quad (4.16)$$

Notar como antes que la expresión

$$\left(\sum_{h \leq H} \vartheta(n+h) - \log 3N \right) \quad (4.17)$$

será positiva si y sólo si el intervalo $[1, H]$ contiene al menos dos números primos. En consecuencia, para probar el Teorema 1.3 bastará demostrar

que para todo $\varepsilon > 0$ existirán infinitos valores de N para los cuales (4.16) será positivo, con $H = \varepsilon \log N$.

Fijemos entonces un valor de $\varepsilon > 0$ y consideremos $H = \varepsilon \log N$. Para estimar (4.16) partiremos la suma de (4.17) en dos, considerando por separado los $h \in \mathcal{H}$ y los $h \notin \mathcal{H}$. Separamos también el término $\log 3N$. Por la Proposición 2.1 y el Lema 4.1 tenemos

$$\begin{aligned}&\sum_{\substack{\mathcal{H}\subseteq[1,H]\\|\mathcal{H}|=k}} \mathbb{E}_{N<n\leq 2N}\left((\log 3N)\Lambda_R^2(n;\mathcal{H},k+l)\right)\\ &= (1+o(1))\frac{H^k}{(k+2l)!}\binom{2l}{l}(\log N)(\log R)^{k+2l}. \end{aligned} \tag{4.18}$$

Consideramos ahora

$$\sum_{\substack{\mathcal{H}\subseteq[1,H]\\|\mathcal{H}|=k}} \mathbb{E}_{N<n\leq 2N}\left(\sum_{\substack{h\leq H\\h\notin\mathcal{H}}} \vartheta(n+h)\Lambda_R^2(n;\mathcal{H},k+l)\right).$$

Por la Proposición 4.2 esto es igual a

$$\begin{aligned}&\sum_{\substack{\mathcal{H}\subseteq[1,H]\\|\mathcal{H}|=k+1}} \frac{(\mathfrak{G}(\mathcal{H})+o(1))}{(k+2l)!}\binom{2l}{l}(\log R)^{k+2l}\\ &= (1+o(1))\frac{H^{k+1}}{(k+2l)!}\binom{2l}{l}(\log R)^{k+2l}, \end{aligned} \tag{4.19}$$

en donde hemos usado el Lema 4.1.

Finalmente, consideramos

$$\sum_{\substack{\mathcal{H}\subseteq[1,H]\\|\mathcal{H}|=k}} \mathbb{E}_{N<n\leq 2N}\left(\sum_{\substack{h\leq H\\h\in\mathcal{H}}} \vartheta(n+h)\Lambda_R^2(n;\mathcal{H},k+l)\right).$$

Puesto que para cada $\mathcal{H}$ existen k elecciones posibles de h, usando la Proposición 4.2 y el Lema 4.1, podemos estimar esto como

$$(1+o(1))\frac{kH^k}{(k+2l+1)!}\binom{2(l+1)}{l+1}(\log R)^{k+l+1}. \tag{4.20}$$

Insertando entonces (4.18), (4.19) y (4.20) vemos que (4.16) es igual a

$$\left\{H + \frac{k}{k+2l+1}\frac{2(2l+1)}{l+1}\log R - \log N + o(\log N)\right\}\frac{H^k}{(k+2l)!}\binom{2l}{l}(\log R)^{k+2l}.$$

Bastará entonces ver que la expresión entre llaves es positiva si N es suficientemente grande. Para esto notamos que podemos reescribir a tal factor como

$$\left\{\varepsilon + \frac{k}{k+2l+1}\frac{2(2l+1)}{l+1}\frac{\theta'}{2} - 1 + o(1)\right\}\log N, \tag{4.21}$$

para cualquer $\theta' < 1/2$, en donde hemos usado la definición de H y la Proposición 4.1. Pero entonces, utilizando (4.15) vemos que tomando θ' suficientemente cerca de $1/2$ y k suficientemente grande en términos de ε, la expresión (4.21) será postiiva para N suficientemente grande. En particular, habrá dos números primos distintos en el intervalo $[1, H]$.

Obtenemos entonces

$$\liminf_{n\to\infty} p_{n+1} - p_n \leq H = \varepsilon \log N.$$

Luego, puesto que $\varepsilon > 0$ es arbitrario, concluimos que es

$$\liminf_{n\to\infty} \frac{p_{n+1} - p_n}{\log N} = 0.$$

El Teorema 1.2 queda así demostrado. □

Capítulo 5

Medidas pseudoaleatorias

En éste capítulo definiremos lo que entendemos por medidas pseudoaleatorias y veremos que una variación adecuada de $\Lambda_R(n; \mathcal{H}, k+l)$ satisface tal definición. Para esto adaptaremos los métodos utilizados en [24]. Asimismo, enunciaremos una proposición que nos garantiza progresiones aritméticas en aquellos conjuntos que pueden ser modelados adecuadamente por tales medidas pseudoaleatorias. La demostración de esta proposición abarcará los próximos dos capítulos. Además, condicional a esta afirmación, deduciremos en éste capítulo el Teorema 1.1.

5.1. El teorema de Szemerédi

Como fue mencionado con anterioridad, uno de los instrumentos principales para deducir el Teorema 1.1 es el siguiente resultado fundacional de la combinatoria aditiva demostrado originalmente por Endré Szemerédi [42].

Proposición 5.1 (Teorema de Szemerédi)**.** *Para todo entero $r \geq 1$ y todo número real $0 < \delta \leq 1$, existe un entero $N_0(\delta, r)$ tal que si es $N > N_0(\delta, r)$ entonces todo conjunto $A \subseteq \{1, \ldots, N\}$ con $|A| \geq \delta N$ contiene al menos una progresión aritmética de longitud r.*

Al día de hoy existen diversas demostraciones de esta afirmación en la literatura. Además de la demostración combinatoria dada por Szemerédi, Hillel Furstenberg dio una demostración utilizando teoría ergódica [15] y Timothy Gowers otra empleando análisis de Fourier [19]. Más aún, recientemente Gowers [20] y Vojtech Rödl y Jozef Skokan [39] dieron una demostración de una generalización multidimensional utilizando un lema de regularidad para hipergrafos.

Vale la pena mencionar aquí la siguiente conjetura de Erdös [12] sobre cuánto más que el teorema de Szemerédi puede esperarse.

Conjetura 5.1. *Sea $\mathcal{A} \subseteq \mathbb{N}$ tal que la suma de sus recíprocos diverge. Entonces $\mathcal{A}$ posee progresiones aritméticas arbitrariamente largas.*

Observar que la veracidad de esta conjetura implica la existencia de progresiones aritméticas arbitrariamente largas en los primos (en realidad, esta fue la motivación de Erdös para formular esta conjetura), aunque no así el Teorema 1.1, puesto que por ejemplo, Brun demostró en su famoso trabajo [6] de 1919 que la serie de recíprocos de los primos gemelos converge (en general, sucede lo mismo con cualquier tupla de primos con más de un elemento). La divergencia de la serie de recíprocos en el enunciado de la conjetura de Erdös es en cierta forma decorativa de todos modos puesto que lo que realmente se supone relevante es la densidad del conjunto en cuestión. Al presente no ha habido avances en éste problema que generalicen por completo el teorema de Szemerédi ni se ha logrado demostrar que un conjunto con las propiedades enunciadas en la conjetura contiene infinitas progresiones aritméticas de una longitud fija (no trivial), a pesar de que en éste último caso Jean Bourgain sí logró mejorar el teorema de Szemerédi para progresiones aritméticas de longitud tres (éste caso particular es conocido como el Teorema de Roth) utilizando una ingeniosa variación del método del círculo [5]. En la dirección contraria Félix Behrend [2] construyó en 1946 el ejemplo de mayor densidad conocido hasta la fecha de un conjunto $A \subseteq \{1, \dots, N\}$ sin progresiones aritméticas de longitud tres, siendo $|A| = N \exp\left(-c\sqrt{\log N}\right)$.

En el 2004, Terence Tao [43] dio una prueba ergódico cuantitativa del teorema de Szemerédi. En tal trabajo Tao probó una versión equivalente de éste resultado cuyo enunciado provee un marco ergódico que será de utilidad para las subsecuentes aplicaciones al estudio de los números primos (éste tipo de formulación tiene sus orígenes en el trabajo de Furstenberg de 1977).

Proposición 5.2 (Teorema de Szemerédi, versión ergódica)**.** *Fijemos un número real $0 < \delta \leq 1$ y un número entero $r \geq 1$. Sea N un parámetro entero suficientemente grande y supongamos que $f : \mathbb{Z}_N \to [0, 1]$ es δ-denso. Tenemos entonces*

$$\mathbb{E}_{x,h \in \mathbb{Z}_N} \left(f(x)f(x+h)\dots f(x+(r-1)h)\right) \geq c(r,\delta) - o_{r,\delta}(1) \tag{5.1}$$

para cierta constante $c(r, \delta) > 0$ que no depende ni de f ni de N.

Notar que la Proposición 5.2 aparenta ser más general que la Proposición 5.1 en dos sentidos. En primer lugar, estamos trabajando con una función $f : \mathbb{Z}_N \to [0, 1]$ arbitraria, en lugar de considerar $f : \mathbb{Z}_N \to \{0, 1\}$ como se hace implícitamente en el enunciado original. En segundo lugar, la estimación (5.1) nos está garantizando la existencia de $\gg N^2$ progresiones aritméticas. De todas formas, una deducción de la Proposición 5.2 partiendo de la Proposición 5.1 fue dada por Varnavides [46].

5.2. Pseudoaleatoriedad

Decimos que $\nu : \mathbb{Z}_N \to \mathbb{R}_{\geq 0}$ es una medida si satisface

$$\mathbb{E}_{x \in \mathbb{Z}_N} \nu(x) = 1 + o(1). \tag{5.2}$$

En éste capítulo definiremos lo que entendemos por medida pseudoaleatoria y demostraremos que para toda elección de una tupla admisible $\mathcal{H} = \{h_1, \ldots, h_k\}$ y $l > 0$ existe una variación adecuada de $\Lambda_R(n; \mathcal{H}, k + l)$ que satisface tales condiciones. Esto se logrará mediante la aplicación de los resultados obtenidos en la Proposición 2.2 y la Proposición 2.3.

En éste capítulo enunciaremos también una generalización del teorema de Szemerédi que permite estimar la expresión (5.1) cuando f está acotada por una medida pseudoaleatoria. La demostración de tal resultado será llevada a cabo en los siguientes capítulos. Combinando esto con los resultados de esta sección se deducirá el Teorema 1.1.

Introducimos ahora dos nociones de aleatoriedad originalmente formuladas por Green y Tao. La primera de ellas se conoce como la condición de formas lineales y pide que la función en cuestión se comporte en forma independiente y con idéntica distribución a lo largo de ciertas familias de formas lineales independientes entre sí. La demostración de que una variación apropiada de $\Lambda_R(n; \mathcal{H}, k + l)$ satisface esta condición se obtendrá implementando los resultados de la Proposición 2.2. El enunciado preciso es el siguiente.

Definición 5.1 (Condición de formas lineales; [24], Def. 3.1)**.** *Sea $\nu : \mathbb{Z}_N \to \mathbb{R}_{\geq 0}$ una medida. Sean m_0, t_0, L_0 parámetros. Decimos que ν satisface la condición de formas lineales en (m_0, t_0, L_0) si se cumple lo siguiente. Sean $m \leq m_0$ y $t \leq t_0$ enteros positivos arbitrarios y supongamos que $(L_{ij})_{1 \leq i \leq m, 1 \leq j \leq t}$ son números racionales arbitrarios con numerador y denominador acotados en valor absoluto por L_0, y que b_i, $1 \leq i \leq m$, son elementos arbitrarios de $\mathbb{Z}_N$. Sean $\psi_i : \mathbb{Z}_N^t \to \mathbb{Z}_N$, $1 \leq i \leq m$, formas lineales con $\psi_i(\mathbf{x}) = \sum_{1 \leq j \leq t} L_{ij} x_j + b_i$ donde para identificar a los números racionales L_{ij} en $\mathbb{Z}_N$ en la forma natural asumimos que N es un primo mayor que L_0. Suponemos además que las tuplas $(L_{ij})_{1 \leq j \leq t}$ con $1 \leq i \leq m$ son no nulas y ninguna es un múltiplo racional de la otra. Tenemos entonces*

$$\mathbb{E}_{\mathbf{x} \in \mathbb{Z}_N} \left(\nu(\psi_1(\mathbf{x})) \ldots \nu(\psi_m(\mathbf{x}))\right) = 1 + o_{L_0, m_0, t_0}(1).$$

La segunda noción de aleatoriedad se refiere a la relación de los valores de ν a distancias fijas y es quizás menos natural que la condición de formas lineales. La razón de esto se debe a que en el estudio de los primos, y particularmente de funciones como $\Lambda_R(n; \mathcal{H}, k + l)$, si bien es posible encontrar modificaciones apropiadas de estas que normalizan notablemente su distribución altamente irregular a lo largo de las clases residuales, es imposible eliminar tal irregularidad por completo. En el caso de la condición de

formas lineales, el hecho de que las formas lineales puras involucradas sean independientes permite en cierta forma promediar las variaciones de modo que las irregularidades restantes se cancelen. Esto no sucede sin embargo al estudiar el comportamiento a distancias fijas, donde las formas puras son de hecho iguales, puesto que al no estar promediando sucede que tal distancia puede seleccionarse a modo tal de explotar las relaciones residuales problemáticas que aún existen.

Definición 5.2 (Condición de correlación; [24], Def. 3.2)**.** *Sea $\nu : \mathbb{Z}_N \to \mathbb{R}_{\geq 0}$ una medida. Sea m_0 un parámetro entero positivo. Decimos que ν satisface la condición de m_0-correlación si para todo $1 \leq m \leq m_0$ existe una función de peso $\gamma = \gamma_m : \mathbb{Z}_N \to \mathbb{R}_{\geq 0}$ que satisface las condiciones de momento*

$$\mathbb{E}_{x \in \mathbb{Z}_N} \gamma^q(x) = O_{m,q}(1)$$

para todo $1 \leq q < \infty$ y tal que

$$\mathbb{E}_{x \in \mathbb{Z}_N} \left(\nu(x + z_1) \dots \nu(x + z_m)\right) \leq \sum_{1 \leq i < j \leq m} \gamma(z_j - z_i)$$

para toda elección de $z_1, \dots, z_m \in \mathbb{Z}_N$ (no necesariamente distintos).

Siguiendo la línea de las observaciones hechas en el párrafo anterior, notar que nuevamente la información efectiva que obtenemos viene de promediar, esta vez a través de la estimación de las normas L^q del peso γ utilizado. Para demostrar que esta condición es satisfecha por una adaptación de $\Lambda_R(n; \mathcal{H}, k + l)$ se utilizará la Proposición 2.3.

Ahora sí, definimos lo que entendemos por una medida pseudoaleatoria.

Definición 5.3 (Medidas pseudoaleatorias; [24], Def. 3.3)**.** *Sea $\nu : \mathbb{Z}_N \to \mathbb{R}_{\geq 0}$ una medida. Decimos que ν es r-pseudoaleatoria si satisface la condición de formas lineales en $(r2^{r-1}, 3r-4, r)$ y la condición de 2^{r-1}-correlación.*

La elección de los parámetros en esta definición es en realidad bastante arbitraria. Lo realmente relevante es que tal elección dependa únicamente de r.

Enunciaremos ahora el siguiente resultado vital de Green y Tao cuya demostración nos ocupará los próximos capítulos.

Proposición 5.3 ([24], Teo. 3.5)**.** *Fijemos un número entero $r \geq 3$ y un número real $0 < \delta \leq 1$. Sea $\nu : \mathbb{Z}_N \to \mathbb{R}_{\geq 0}$ una medida r-pseudoaleatoria. Supongamos que $f : \mathbb{Z}_N \to \mathbb{R}_{\geq 0}$ es δ-denso y satisface*

$$0 \leq f(x) \leq \nu(x) \ \forall x \in \mathbb{Z}_N.$$

Entonces

$$\mathbb{E}_{x,h \in \mathbb{Z}_N} \left(f(x) f(x + h) \dots f(x + (r-1)h)\right) \geq c(r, \delta) - o_{r,\delta}(1),$$

donde $c(r, \delta) > 0$ es la misma constante que figura en la Proposición 5.2.

5.3. Construcción de medidas pseudoaleatorias

Asumiendo la Proposición 5.3, el paso restante para deducir el Teorema 1.1 será encontrar para cada $r > 0$ una variación adecuada de $\Lambda_R(n; \mathcal{H}, k+l)$ que defina una medida r-pseudoaleatoria.

Recordemos que la razón por la cual $\Lambda_R(n; \mathcal{H}, k+l)$ así definido no puede definir una medida r-pseudoaleatoria para valores altos de r tiene que ver con que esta función imita en muchos aspectos el comportamiento de $\Lambda(n; \mathcal{H}, k+l)$ y esta función está muy lejos de poseer tal pseudoaleatoriedad debido a su distribución irregular en clases residuales de módulo pequeño. La idea será entonces encontrar una modificación apropiada que evite estas obstrucciones. Para lograr éste objetivo implementaremos nuevamente el truco-W.

De aquí en adelante fijemos una tupla admisible $\mathcal{H} = \{h_1, \dots, h_k\}$, el parámetro $l > 0$ y la longitud $r \geq 3$ de las progresiones aritméticas que deseamos encontrar. Sea $w(N)$ una función que crece suficientemente despacio con N y definamos $W := \prod_{p \leq w(N)} p$. Consideramos la función $\Lambda_R(Wn + a; \mathcal{H}, k+l)$, donde el entero a es elegido de forma tal que $a + \mathcal{H}$ consista enteramente de elementos coprimos con W. De esta manera tan simple, conseguimos una modificación que estará distribuída uniformemente a lo largo de las clases residuales pequeñas. En efecto, para un módulo q pequeño, q será el producto de primos menores que $w(N)$ y en consecuencia los elementos de $Wn + a + \mathcal{H}$ pertenecerán a las mismas clases residuales módulo q para todo valor de $n (\text{mod } q)$, lo cual claramente nos da la uniformidad deseada.

Una vez realizada tal modificación los problemas restantes son menores. El primero es no poder garantizar la positividad de $\Lambda_R(Wn + a; \mathcal{H}, k+l)$ pero esto se soluciona en forma sencilla y estándar tomando el cuadrado de tal función. La segunda objeción tiene que ver con que una medida pseudoaleatoria debe ser en particular una medida y en consecuencia satisfacer (5.2). Para corregir esto, bastará utilizar las estimaciones de la Proposición 2.2 en el caso $m = 1$. Existe un último problema técnico que tiene que ver con que estaremos trabajando en $\mathbb{Z}_N$ en lugar de en $\mathbb{Z}$, por lo cual deberemos restringir el estudio de nuestra medida a un intervalo lo suficientemente pequeño de $\mathbb{Z}_N$ para garantizar que los elementos que estemos estudiando no den vueltas alrededor de $\mathbb{Z}_N$ en lugar de comportarse propiamente como elementos de $\mathbb{Z}$. Teniendo en cuenta estas observaciones, introducimos la siguiente definición.

Definición 5.4. *Sean* $R := N^{r^{-1}2^{-r-4}}$ *y* $\epsilon_r := 1/2^r(r+4)!$. *Sea además* $0 \leq c < \epsilon_r^{-1}$ *un entero. Definimos*

$$\nu_c(n) = \begin{cases} \frac{\varphi_k(W)}{W} \frac{\Lambda_R^2(Wn+a;\mathcal{H},k+l)}{(\log R)^{k+2l}} (k+2l)! \binom{2l}{l}^{-1} & \text{si } c\epsilon_r N \leq n \leq (c+1)\epsilon_r N, \\ 1 & \text{en otro caso,} \end{cases}$$

para todo $0 \leq n < N$. *Queda así definido en forma natural* $\nu_c : \mathbb{Z}_N \to \mathbb{R}_{\geq 0}$.

Veamos que se trata efectivamente de medidas.

Lema 5.1. *Las funciones $\nu_c : \mathbb{Z}_N \to \mathbb{R}_{\geq 0}$ que aparecen en la Definición 5.4 son medidas.*

Demostración. Utilizamos la Proposición 2.2 con $m := 1$, $t := 1$, $\psi(x) := x$ y $B := [c\epsilon_r N, (c+1)\epsilon_r N]$, tomando N suficientemente grande en términos de r de modo que B satisfaga la hipótesis de tal proposición. Tenemos entonces

$$\mathbb{E}_{x \in B} \nu_c(x) = 1 + o(1).$$

Asimismo, por definición es $\nu_c(x) = 1$ si $x \in \mathbb{Z}_N - B$. Juntando ambas estimaciones vemos que ν_c satisface (5.2). □

El objetivo será entonces demostrar la siguiente proposición.

Proposición 5.4. *Las medidas ν_c que aparecen en la Definición 5.4 son r-pseudoaleatorias.*

5.4. Deducción condicional del Teorema 1.1

Pasaremos ahora a dar una demostración del Teorema 1.1 asumiendo la Proposición 5.3 y la Proposición 5.4.

Demostración condicional del Teorema 1.1. Recordemos lo que queremos demostrar. Tenemos un conjunto $\mathcal{A} \subseteq \mathbb{N}$ y sabemos que para todo N suficientemente grande existe $f : \mathbb{Z}_N \to \mathbb{R}_{\geq 0}$ δ-denso, de soporte en $\mathcal{A}$ y acotado puntualmente por $\Lambda_R^2(n; \mathcal{H}, k+l) \log^{-k-2l} R$, con $R = N^{r^{-1}2^{-r-4}}$. Deseamos ver que $\mathcal{A}$ contiene progresiones aritméticas de longitud r.

Es sabido que la sucesión p_n de los números primos satisface la desigualdad $p_{n+1} - p_n < p_n^\theta$ para cierto $\theta < 1$ (en efecto, Revenor Baker, Glyn Harman y János Pintz [1] demostraron que puede tomarse $\theta = 0{,}525$). Con esto, podemos encontrar un menor primo $N' \in [W(N+1), W(N+1) + (W(N+1))^\theta]$, donde N se elige grande de modo que existe $f : \mathbb{Z}_{N'} \to \mathbb{R}_{\geq 0}$ satisfaciendo las hipótesis del párrafo anterior.

Observamos ahora que del Lema 2.1 se concluye fácilmente que es $\tau_2(n) \ll_\varepsilon n^\varepsilon$ para cualquier elección de $\varepsilon > 0$. Usando esto tenemos trivialmente

$$\Lambda_R^2(n; \mathcal{H}, k+l) \ll_\varepsilon n^\varepsilon. \tag{5.3}$$

Fijemos ahora $\varepsilon > 0$ suficientemente pequeño. Usando la cota puntual que poseemos para f, deducimos que es

$$\mathbb{E}_{n \in \mathbb{Z}_{N'}} \left((\mathbf{1}_{[0,W]} + \mathbf{1}_{[W(N+1), W(N+1)+(W(N+1))^\theta]}) f(n) \right) \ll \frac{N^{2\varepsilon}(W + (W(N+1))^\theta)}{W(N+1)} = o(1).$$

Esto nos permite concluir la existencia de una constante $\delta' > 0$ (arbitrariamente cercana a δ) tal que si N es suficientemente grande y $f : \mathbb{Z}_{N'} \to \mathbb{R}_{\geq 0}$ está definido como arriba, entonces $\mathbf{1}_{[W,W(N+1)]}f$ será δ'-denso, tendrá soporte en $\mathcal{A}$ y satisfacerá la misma cota puntual que f.

Denotemos por $\Omega(p)$ al conjunto de los $-h(\text{mod p})$ con $h \in \mathcal{H}$. Vemos entonces que la cantidad de enteros $0 \leq a < W$ para los cuales $a+\mathcal{H}$ consiste enteramente de elementos coprimos con W es $W\prod_{p\leq w(N)}\left(1-\frac{|\Omega(p)|}{p}\right)$. Para cada tal valor, sabemos de la Proposición 2.2 que es

$$\begin{aligned}\mathbb{E}_{n\in[W,W(N+1)]}&\left(\mathbf{1}_{n\equiv a(\text{mod } W)}\Lambda_R^2(n;\mathcal{H},k+l)\right)\\ &= \frac{1}{W}\mathbb{E}_{n\leq N}\left(\Lambda_R^2(Wn+a;\mathcal{H},k+l)\right)\\ &= \frac{(1+o(1))}{\varphi_k(W)}\binom{2l}{l}\frac{(\log R)^{k+2l}}{(k+2l)!}.\end{aligned}$$

Asimismo, sabemos de la Proposición 2.1 que la esperanza de $\Lambda_R^2(n;\mathcal{H},k+l)$ en $[W,W(N+1)]$ es $(1+o(1))\mathfrak{G}(\mathcal{H})\binom{2l}{l}\frac{(\log R)^{k+2l}}{(k+2l)!}$ (en realidad, tal proposición está enunciada para el intervalo $[N,2N]$ pero es evidente que exactamente el mismo argumento se aplica en nuestro caso; en efecto, nuestro intervalo es lo suficientemente grande como para absorber el error de aproximar por $\mathcal{T}$ y tal valor no depende del intervalo sino únicamente de que H se mantenga controlado, lo cual en nuestro caso es trivial puesto que es $H = O(1)$ al estar la tupla fija). Escribimos $n \in \Omega_W$ si y sólo si $n+\mathcal{H}$ posee un elemento que no es coprimo con W. Juntando las dos estimaciones anteriores, vemos que la esperanza de $\mathbf{1}_{n\in\Omega_W}\Lambda_R^2(n;\mathcal{H},k+l)$ en $[W,W(N+1)]$ es igual a

$$\left(\mathfrak{G}(\mathcal{H}) - \frac{W\prod_{p\leq w(N)}\left(1-\frac{|\Omega(p)|}{p}\right)}{\varphi_k(W)}\right)\binom{2l}{l}\frac{(\log R)^{k+2l}}{(k+2l)!} = o((\log R)^{k+2l}),$$

puesto que por definición de $\mathfrak{G}(\mathcal{H})$ es

$$\mathfrak{G}(\mathcal{H}) = \lim_{N\to\infty}\frac{W}{\varphi_k(W)}\prod_{p\leq w(N)}\left(1-\frac{|\Omega(p)|}{p}\right). \tag{5.4}$$

Debido a esto, considerando $f : \mathbb{Z}_{N'} \to \mathbb{R}_{\geq 0}$ como antes y usando la cota puntual, obtenemos

$$\mathbb{E}_{n\in[W,W(N+1)]}\left(\mathbf{1}_{n\in\Omega_W}f(n)\right) = o(1).$$

Luego, recordando que $\mathbf{1}_{[W,W(N+1)]}f : \mathbb{Z}_{N'} \to \mathbb{R}_{\geq 0}$ es δ'-denso y usando el principio de los casilleros, concluimos que existirá $a \notin \Omega_W$, $0 \leq a < W$, con

$$\begin{aligned}\mathbb{E}_{n\in\mathbb{Z}_{N'}}\left(\mathbf{1}_{[W,W(N+1)]}\mathbf{1}_{n\equiv a(\mathrm{mod}\ W)}f(n)\right)&\geq\frac{\delta'+o(1)}{W\prod_{p\leq w(N)}\left(1-\frac{|\Omega(p)|}{p}\right)}\\&\geq\frac{\delta''}{W\prod_{p\leq w(N)}\left(1-\frac{|\Omega(p)|}{p}\right)},\end{aligned}$$

para cierta constante $\delta''>0$ y todo N suficientemente grande.

Fijando tal elección de a e identificando $\mathbb{Z}_N$ en $\mathbb{Z}_{N'}$ en la forma natural, tenemos entonces

$$\mathbb{E}_{n\in\mathbb{Z}_N}f(Wn+a)\geq\frac{\delta''}{\prod_{p\leq w(N)}\left(1-\frac{|\Omega(p)|}{p}\right)}.$$

En particular, utilizando nuevamente el principio de los casilleros, vemos que existe un entero $0\leq c<\epsilon_r^{-1}$ tal que es

$$\mathbb{E}_{n\in\mathbb{Z}_N}\left(\mathbf{1}_{[c\epsilon_rN,(c+1)\epsilon_rN]}f(Wn+a)\right)\geq\frac{\epsilon_r\delta''}{\prod_{p\leq w(N)}\left(1-\frac{|\Omega(p)|}{p}\right)}.$$

Finalmente, fijando un tal valor de c, definimos $F:\mathbb{Z}_N\to\mathbb{R}_{\geq 0}$ como

$$F(n):=\frac{\varphi_k(W)}{W}(k+2l)!\binom{2l}{l}^{-1}\mathbf{1}_{[c\epsilon_rN,(c+1)\epsilon_rN]}f(Wn+a).$$

Es claro entonces que es

$$F(n)\leq\nu_c(n)\tag{5.5}$$

para todo $n\in\mathbb{Z}_N$. Más aún, es

$$\begin{aligned}\mathbb{E}_{n\in\mathbb{Z}_N}F(n)&\geq\frac{\varphi_k(W)}{W}(k+2l)!\binom{2l}{l}^{-1}\frac{\epsilon_r\delta''}{\prod_{p\leq w(N)}\left(1-\frac{|\Omega(p)|}{p}\right)}\\&=(\mathfrak{G}(\mathcal{H})+o(1))(k+2l)!\binom{2l}{l}^{-1}\epsilon_r\delta''\geq\delta''',\end{aligned}$$

para cierto $\delta'''>0$ y todo N suficientemente grande, en donde hemos usado (5.4).

Nota. Hay un pequeño abuso aquí puesto que estamos considerando ν_c con un parámetro $R=N'^{r^{-1}2^{-r-4}}=N^{(1+o(1))r^{-1}2^{-r-4}}$ en lugar de $R=N^{r^{-1}2^{-r-4}}$. Es trivial de todas formas verificar que esto no modifica en nada los argumentos que prueban que se trata de una medida pseudoaleatoria, puesto que de por sí la elección del exponente es bastante arbitraria.

Obtenemos entonces de la Proposición 5.3 la estimación

$$\mathbb{E}_{x,h\in\mathbb{Z}_N}\left(F(x)F(x+h)\dots F(x+(r-1)h)\right) \geq c(r,\delta''') - o_{r,\delta'''}(1). \quad (5.6)$$

Debido a (5.3) y (5.5) vemos que la contribución del caso degenerado $h=0$ a (5.6) es a lo sumo $o(1)$. Luego, habrá de existir algún par $x,h\in\mathbb{Z}_N$ con $h\neq 0$, para el cual sea

$$F(x)F(x+h)\dots F(x+(r-1)h)\neq 0$$

(en realidad habrá $\gg N^2$ tales pares). Notar que al tener F soporte en $[c\epsilon_r N,(c+1)\epsilon_r N]$ con $\epsilon_r<1/r$, el conjunto $x,x+h,\dots,x+(r-1)h\in\mathbb{Z}_N$ es también una progresión aritmética en $\mathbb{Z}$ (aunque en este caso h puede ser tanto positivo como negativo). Pero entonces, puesto que f tiene soporte en $\mathcal{A}$, se sigue que el conjunto

$$Wx+a, Wx+a+Wh,\dots,Wx+a+(r-1)Wh\in\mathbb{N}$$

pertenece a $\mathcal{A}$. Luego $\mathcal{A}$ posee progresiones aritméticas de longitud r y el Teorema 1.1 queda así demostrado. □

5.5. Demostración de la condición de formas lineales

Para probar la Proposición 5.4 debemos ver que ν_c satisface la condición de formas lineales en $(r2^{r-1},3r-4,r)$ y la condición de 2^{r-1}-correlación. Esto será logrado en las próximas dos proposiciones, mediante una adaptación de los métodos empleados en [24].

Proposición 5.5. *La medida ν_c satisface la condición de formas lineales en $(r2^{r-1},3r-4,r)$.*

Demostración. Para $1\leq i\leq m$ sean $\psi_i(\mathbf{x})=\sum_{j=1}^{t}L_{ij}x_j+b_i$ formas lineales de la forma enunciada en la Defincón 5.1, es decir, con $m\leq r2^{r-1}$, $t\leq 3r-4$ y los L_{ij} números racionales con numerador y denominador de valor absoluto a lo sumo r, tales que las tuplas $(L_{ij})_{1\leq j\leq t}$ son no nulas y ninguna es un múltiplo racional de la otra. Deseamos demostrar

$$\mathbb{E}_{\mathbf{x}\in\mathbb{Z}_N^t}\left(\nu_c(\psi_1(\mathbf{x}))\dots\nu_c(\psi_m(\mathbf{x}))\right)=1+o(1). \quad (5.7)$$

Por supuesto, la idea será aplicar la Proposición 2.2. Para esto, debemos reemplazar los L_{ij} por números enteros, lo cual se consigue tomando común denominador de estos coeficientes, de manera tal que éste denominador se pueda absorber en la variable $\mathbf{x}$ aprovechando que esta pertenece a $\mathbb{Z}_N$ con N primo. Como consecuencia de esto obtenemos coeficientes enteros L_{ij} acotados en valor absoluto por $(r+1)!$. Asimismo, será necesario asumir entonces $(r+1)!<\sqrt{w(N)}/2$, pero es claro que tal desigualdad se sigue de considerar N suficientemente grande.

Sea $Q = Q(N)$ una función que tiende al infinito con N suficientemente despacio. Para todo $(u_1, \dots, u_t) \in \mathbb{Z}_Q^t$ consideramos los intervalos de la forma

$$B_{u_1,\dots,u_t} := \left\{ x \in \mathbb{Z}_N^t : x_j \in [\lfloor u_j N/Q \rfloor, \lfloor (u_j+1)N/Q \rfloor), j = 1, \dots, t \right\}.$$

Con esto, es claro que el lado izquierdo de (5.7) es igual a

$$\mathbb{E}_{(u_1,\dots,u_t)\in\mathbb{Z}_Q^t} \left(\mathbb{E}_{\mathbf{x}\in B_{u_1,\dots,u_t}} \left(\nu_c(\psi_1(\mathbf{x})) \dots \nu_c(\psi_m(\mathbf{x})) \right) \right) + o(1), \tag{5.8}$$

donde el error se debe a que los intervalos en cuestión no tienen un volúmen exactamente igual a N^t/Q^t sino que poseen una variación $O(1)$ respecto a éste valor. Debido a esto, bastará estimar

$$\mathbb{E}_{(u_1,\dots,u_t)\in\mathbb{Z}_Q^t} \left(\nu_c(\psi_1(\mathbf{x})) \dots \nu_c(\psi_m(\mathbf{x})) \right). \tag{5.9}$$

Comenzamos considerando los casos en los que para todo $1 \leq i \leq m$ $\psi_i(B_{u_1,\dots,u_t})$ está completamente contenido en $[c\epsilon_r N, (c+1)\epsilon_r N]$ o posee intersección disjunta con tal intervalo. Decimos en tal caso que la tupla $(u_1, \dots, u_t)$ es uniforme. Dada una tal tupla, podemos efectuar en (5.9) los reemplazos $\nu_c(\psi_i(\mathbf{x})) = \frac{\varphi(W)}{W} \frac{\Lambda_R^2(W\psi_i(\mathbf{x})+a;\mathcal{H},k+l)}{(\log R)^{k+2l}} (k+2l)! \binom{2l}{l}^{-1}$ ó $\nu_c(\psi_i(\mathbf{x})) = 1$ respectivamente.

Puesto que nuestra intención es aplicar la Proposición 2.2 necesitaremos asegurarnos que la discrepancia proveniente de estar trabajando en $\mathbb{Z}_N$ en lugar de $\mathbb{Z}$ no genere ninguna complicación. Claramente esto sólo es relevante para los i con

$$\psi_i(B_{u_1,\dots,u_t}) \subseteq [c\epsilon_r N, (c+1)\epsilon_r N]. \tag{5.10}$$

Notar que aquí se afirma únicamente que la forma lineal va a parar a tal intervalo módulo N. Para solucionar esto, comenzamos notando que la imágen por ψ_i de dos elementos distintos de $B_{u_1,\dots,u_t}$ difieren en a lo sumo

$$t(r+1)!(N/Q + O(1)) < (1-\epsilon_r)N,$$

suponiendo N suficientemente grande. Debido a esto y puesto que estamos asumiendo (5.10) se sigue que existe una constante b_i' con

$$\psi_i(B_{u_1,\dots,u_t}) - b_i' \subseteq [c\epsilon_r N, (c+1)\epsilon_r N] \subseteq \mathbb{Z},$$

es decir, como intervalo de los enteros y no simplemente módulo N. Absorbiendo esta constante b_i' en la constante b_i de la definición de ψ_i, vemos que podemos identificar a $B_{u_1,\dots,u_t}$ en $\mathbb{Z}$ sin obtener ninguna complicación.

Aplicando entonces la Proposición 2.2 con $k_i = k$ y $l_i = l$ para todo i obtenemos entonces que en éste caso (5.9) es igual a 1. Notar que podemos suponer que Q crece lo suficientemente despacio con N como para que sea $N/Q + O(1) \geq N^{\frac{10}{32}}$ para N suficientemente grande, por lo que recordando la definición de R dada en la Definición 5.4 vemos que $B_{u_1,\dots,u_t}$ efectivamente satisface las hipótesis de la Proposición 2.2.

Supongamos ahora que $(u_1, \dots, u_t)$ no es uniforme. En tal caso podemos acotar a ν_c crudamente por $1 + \frac{\varphi(W)}{W} \frac{\Lambda_R^2(W\psi_i(\mathbf{x})+a;\mathcal{H},k+l)}{(\log R)^{k+2l}}(k+2l)!\binom{2l}{l}^{-1}$. Insertando esta estimación en (5.9), expandiendo los productos y aplicando la Proposición 2.2 $2^m - 1$ veces obtenemos la estimación

$$\mathbb{E}_{(u_1,\dots,u_t)\in\mathbb{Z}_Q^t} (\nu_c(\psi_1(\mathbf{x})) \dots \nu_c(\psi_m(\mathbf{x}))) = O_{m,t}(1) + o_{m,t}(1).$$

Bastará entonces probar que la cantidad de tuplas $(u_1, \dots, u_t)$ no uniformes en $\mathbb{Z}_Q^t$ es a lo sumo $O_{m,t}(1/Q) = o_{m,t}(1)$ puesto que en tal caso la contribución de las tuplas no uniformes a (5.8) es insignificante y en consecuencia la presente proposición se sigue de las estimaciones anteriores para tuplas uniformes.

Pasemos entonces a acotar la cantidad de tuplas no uniformes. Fijemos $(u_1, \dots, u_t) \in \mathbb{Z}_Q^t$ y supongamos que existe $1 \leq i \leq m$ con $\psi_i(\mathbf{x}) \in [c\epsilon_r N, (c+1)\epsilon_r N]$ y $\psi_i(\mathbf{x}') \notin [c\epsilon_r N, (c+1)\epsilon_r N]$, para cierto par $\mathbf{x}, \mathbf{x}' \in B_{u_1,\dots,u_t}$. Teniendo en cuenta que $(\lfloor Nu_j/Q \rfloor)_{1\leq j\leq t} \in B_{u_1,\dots,u_t}$ y puesto que hemos visto que las imágenes por ψ_i de dos elementos de tal intervalo difieren en a lo sumo $O_{m,t}(N/Q)$, se sigue que es

$$\psi_i(\mathbf{x}), \psi_i(\mathbf{x}') = \sum_{j=1}^{t} L_{ij}\lfloor Nu_j/Q \rfloor + b_i + O_{m,t}(N/Q).$$

Luego, puesto que $\psi_i(\mathbf{x}')$ está más cerca de uno de los dos bordes de $[c\epsilon_r N, (c+1)\epsilon_r N]$ que de $\psi_i(\mathbf{x})$, debe ser

$$a\epsilon_r N = \sum_{j=1}^{t} L_{ij}\lfloor Nu_j/Q \rfloor + b_i + O_{m,t}(N/Q)$$

con $a = c$ ó $a = c + 1$. Dividiendo por N/Q obtenemos

$$\sum_{j=1}^{t} L_{ij}u_j = a\epsilon_r Q + b_i Q/N + O_{m,t}(1).$$

Recordemos que estamos trabajando en $\mathbb{Z}_Q^t$ y en consecuecia las soluciones del anterior sistema para cada valor fijo del último término estarán contenidas en un subespacio afín de dimensión $t-1$ (recordemos que $(L_{ij})_{1\leq t}$ es no nulo). Luego la cantidad de tuplas $(u_1, \dots, u_t)$ que satisfacen la anterior ecuación es a lo sumo $O_{m,t}(Q^{t-1})$. Finalmente, dejando variar los valores de a e i, se sigue que la proporción de tuplas no uniformes es a lo sumo $O_{m,t}(1/Q)$ como se quería demostrar. Esto concluye la demostración de la Proposición 5.5. □

5.6. Demostración de la condición de correlación

Para concluir la demostración de la Proposición 5.4 bastará probar el siguiente resultado.

Proposición 5.6. *La medida ν_c satisface la condición de 2^{r-1}-correlación.*

Demostración. Análogamente a lo hecho para la Proposición 5.5 la demostración de esta proposición girará entorno a una aplicación de la Proposición 2.3. Antes que esto sin embargo demostraremos un lema que nos ayudará a reemplazar el factor

$$\prod_{\substack{p|\Delta \\ (p,W)=1}} \left(1 + O_m(p^{-1/2})\right)$$

que aparece allí por una cierta suma sobre una función de peso de la forma dada en la Definición 5.2.

Lema 5.2. *Sea $m \geq 1$ un parámetro entero. Existe entonces una función de peso $\gamma = \gamma_m : \mathbb{Z} \to \mathbb{R}_{\geq 0}$ con $\gamma(n) \geq 1$ para todo $n \neq 0$ que satisface*

$$\prod_{\substack{p|\Delta \\ (p,W)=1}} \left(1 + O_m(p^{-1/2})\right) \leq \sum_{1 \leq i < j \leq m} \gamma(z_j - z_i), \tag{5.11}$$

para toda elección de $z_1, \ldots, z_m \in [c\epsilon_r N, (c+1)\epsilon_r N]$ distintos. Aquí Δ está definido como en el enunciado de la Proposición 2.3. Además, γ es tal que satisface la estimación

$$\mathbb{E}_{0<|n|\leq N} \gamma^q(n) = O_{m,q}(1)$$

para todo entero $0 < q < \infty$.

Demostración. Comenzamos notando que es

$$\prod_{\substack{p|\Delta \\ (p,W)=1}} \left(1 + O_m(p^{-1/2})\right) \leq \prod_{1 \leq i < j \leq m} \left(\prod_{-H \leq b \leq H} \prod_{\substack{p|W(z_j - z_i)+b \\ (p,W)=1}} (1 + p^{-1/2}) \right)^{O_m(1)}.$$

Utilizando la desigualdad de las medias aritméticas y geométricas en la forma $x_1 \ldots x_n \leq \frac{x_1^n + \ldots + x_n^n}{n}$ y absorbiendo las diversas constantes en la expresión $O_m(1)$, vemos que eligiendo

$$\gamma_m(n) := O_m(1) \prod_{-H \leq b \leq H} \prod_{\substack{p|Wn+b \\ (p,W)=1}} (1 + p^{-1/2})^{O_m(1)}$$

se satisface (5.11).

En consecuencia, el objetivo restante será obtener la estimación

$$\mathbb{E}_{0<|n|\leq N} \left(\prod_{-H \leq b \leq H} \prod_{\substack{p|Wn+b \\ (p,W)=1}} (1 + p^{-1/2})^{O_m(q)} \right) = O_{m,q}(1).$$

Ahora bien, aplicando la desigualdad de las medias aritméticas y geométricas en la misma forma que antes, vemos entonces que bastará probar

$$\sum_{-H\leq b\leq H} \mathbb{E}_{0<|n|\leq N} \left(\prod_{\substack{p|Wn+b \\ (p,W)=1}} (1+p^{-1/2})^{O_m(q)} \right) = O_{m,q}(1). \tag{5.12}$$

Recordando además que es $H = O(1)$ alcanzará con demostrar que cada sumando es $O_{m,q}(1)$. Notamos ahora que si p es suficientemente grande en términos de $O_{m,q}(1)$ entonces será $(1+p^{-1/2})^{O_m(q)} \leq 1+p^{-1/4}$, por lo cual podemos estimar cada sumando del lado izquierdo de (5.12) como

$$O_{m,q}(1)\mathbb{E}_{0<|n|\leq N} \left(\prod_{\substack{p|Wn+b \\ (p,W)=1}} (1+p^{-1/4}) \right). \tag{5.13}$$

Pero es

$$\prod_{\substack{p|Wn+b \\ (p,W)=1}} (1+p^{-1/4}) \leq \sum_{\substack{d|Wn+b \\ (d,W)=1}} d^{-1/4}$$

y en consecuencia (5.13) estará acotado por

$$\frac{O_{m,q}(1)}{2N} \left(\sum_{n=1}^{N} \sum_{\substack{d|Wn+b \\ (d,W)=1}} d^{-1/4} + \sum_{n=1}^{N} \sum_{\substack{d|Wn-b \\ (d,W)=1}} d^{-1/4} \right) \tag{5.14}$$

en donde hemos usado que W puede tomarse más grande que H asumiendo N suficientemente grande.

Para estimar esto, notamos que todo entero $d < N$ con $(d,W) = 1$ divide $N/d+O(1) = O(N/d)$ valores de $n \equiv b (\text{mod } W)$ y $N/d+O(1) = O(N/d)$ valores de $n \equiv -b (\text{mod } W)$ en cada uno de los intervalos en cuestión. Análogamente, los enteros $d > N$ con $(d,W) = 1$, dividen a lo sumo un elemento de cada forma en tales intervalos. Con esto, (5.14) queda acotado por

$$\begin{aligned} &\frac{O_{m,q}(1)}{2N} \left(\sum_{d=1}^{N} \frac{N}{d} d^{-1/4} + \sum_{d=N+1}^{WN+b} d^{-1/4} \right) \\ &\leq O_{m,q}(1) \left(\zeta(5/4) + \frac{WN}{2N} N^{-1/4} \right) \\ &= O_{m,q}(1) \end{aligned}$$

si $w(N)$ (y en consecuencia W) crece lo suficientemente despacio en respecto a N. El lema queda así demostrado. □

Ahora sí, consideramos $1 \leq m \leq 2^{r-1}$ y $z_1, \ldots, z_m \in \mathbb{Z}_N$. Debemos demostrar

$$\mathbb{E}_{x \in \mathbb{Z}_N} \left(\nu_c(x+z_1) \ldots \nu_c(x+z_m)\right) \leq \sum_{1 \leq i < j \leq m} \gamma(z_j - z_i), \tag{5.15}$$

donde $\gamma = \gamma_m$ es una función acotada en L^q para todo $0 < q < \infty$.

Fijamos ahora $m, z_1, \ldots, z_m$. Tomamos la función γ definida en el Lema 5.2. Recordemos que tal función no está definida en el orígen, por lo cual elegimos

$$\gamma(0) := \exp(Cm \log N / \log\log N) \tag{5.16}$$

para una cierta constante C grande que no depende de m y será especificada luego. Notamos que éste valor de γ contribuye a lo sumo $o_{m,q}(1)$ a la norma L^q, por lo cual se sigue del Lema 5.2 que esta norma será $O_{m,q}(1)$.

Supongamos primero que existen dos valores idénticos de z_i. En tal caso acotamos el lado izquierdo de (5.15) por $||\nu_c||_\infty^m$. Recordemos que $\tau_2(n)$ es la cantidad de divisores de n. Es un resultado conocido de la Teoría de Números la existencia de una constante C' tal que es $\tau_2(n) \leq \exp(C' \log n / \log\log n)$. Precisamente, Severin Wigert [47] demostró en 1906 que

$$\limsup_{n \to \infty} \frac{\log \tau_2(n) \log\log n}{\log n} = \log 2.$$

Insertando esto en la definición de $\Lambda_R(n; \mathcal{H}, k+l)$, obtenemos la cota

$$||\nu_c||_\infty \leq O(1) \exp(C' \log N / \log\log N) \log^{k+l} R \leq \gamma(0),$$

si la constante C en (5.16) se elige lo suficientemente grande en términos de k y l (pero independiente de m). La afirmación se sigue entonces para el caso en el que existen dos índices distintos $1 \leq i < j \leq m$ con $z_i = z_j$.

Supongamos ahora que todos los z_i son distintos. Escribimos

$$g(n) := \mathbf{1}_{[c\epsilon_r N, (c+1)\epsilon_r N]} \nu_c(n),$$

de modo que tenemos la cota

$$\begin{aligned} &\mathbb{E}_{x \in \mathbb{Z}_N} \left(\nu_c(x+z_1) \ldots \nu_c(x+z_m)\right) \\ &\qquad \leq \mathbb{E}_{x \in \mathbb{Z}_N} \left((1+g(x+z_1)) \ldots (1+g(x+z_m))\right). \end{aligned} \tag{5.17}$$

Podemos reescribir el lado derecho de (5.17) como

$$\sum_{A \subseteq \{1,\ldots,m\}} \mathbb{E}_{x \in \mathbb{Z}_N} \left(\prod_{i \in A} g(x+z_i) \right).$$

Notar que para todo $A \subseteq \{1, \ldots, m\}$ podemos asumir $z_i \in [c\epsilon_r N, (c+1)\epsilon_r N]$ (y en consecuencia $g = \nu_c$) para todo $i \in A$, puesto que la esperanza es nula

en otro caso. Aplicando entonces la Proposición 2.3 y el Lema 5.2 obtenemos para todo tal conjunto A la cota

$$\mathbb{E}_{x\in\mathbb{Z}_N}\left(\prod_{i\in A} g(x+z_i)\right) \leq \sum_{1\leq i<j\leq m} \gamma(z_j - z_i) + o_m(1).$$

Sumando sobre todo $A \subseteq \{1,\dots,m\}$ y reajustando γ por una constante que depende sólo de m (recordar que es $\gamma = \gamma_m$) nuestro resultado queda demostrado. □

Demostración de la Proposición 5.4. Esto es inmediato de la Proposición 5.5 y la Proposición 5.6. □

Capítulo 6

La teoría de la norma de Gowers

Emprenderemos ahora la demostración de la Proposición 5.3. Para llevar a cabo tal propósito será necesario estudiar una familia de normas introducidas por Timothy Gowers [19] en 1998 y que hoy en día desempeñan un papel fundamental en la Combinatoria Aditiva.

Los resultados presentados en este capítulo están esencialmente contenidos en [21] y [24].

6.1. Las normas de Gowers

Definición 6.1. *Sea G un grupo abeliano finito y sea $f : G \to \mathbb{C}$. La norma U^2 de Gowers se define como*

$$\|f\|_{U^2}^4 := \mathbb{E}_{x,a,b} f(x)\overline{f(x+a)f(x+b)}f(x+a+b).$$

Pronto daremos la verificación de que esto es en verdad una norma, pero notamos ahora que la propiedad fundamental de esta norma radica en su capacidad de medir la aleatoriedad (o casi-aleatoriedad) de una función. Por ejemplo, se puede demostrar que si g es una función con $\|g\|_\infty \leq 1$ y f posee norma U^2 pequeña, entonces g y $f+g$ tendrán aproximadamente la misma esperanza en progresiones aritméticas de longitud 3. Es decir, será

$$\mathbb{E}_{x,h} g(x)g(x+h)g(x+2h) \sim \mathbb{E}_{x,h}(f+g)(x)(f+g)(x+h)(f+g)(x+2h).$$

En particular, supongamos que tenemos un subconjunto $A \subseteq G$ de densidad δ y sea $A(x)$ la función característica de tal conjunto. Entonces, si escribimos $f(x) = \delta - A(x)$, vemos que si $\|f\|_{U^2}$ es pequeño tendremos

$$\mathbb{E}_{x,h} A(x)A(x+h)A(x+2h) \sim \delta^3,$$

o lo que es lo mismo, la cantidad de progresiones aritméticas de longitud 3 en A es similar a la cantidad esperada de tales progresiones en un conjunto

aleatorio de la misma densidad (es decir, uno en el cual cada elemento de G es elegido independientemente y al azar con probabilidad δ).

Decimos que un conjunto A es *casi-aleatorio* si $\|\delta - A\|_{U^2}$ es pequeño. Siguiendo la línea de las observaciones del párrafo anterior, es posible mostrar que un conjunto casi-aleatorio se comportará en muchos aspectos en forma similar a un conjunto aleatorio, en el sentido de que muchas configuraciones sucederán con frecuencia similar al caso de conjuntos aleatorios con la misma densidad. De todas maneras, existirán ciertos patrones para los cuales esto no sucederá. Por ejemplo, si $A \subseteq \mathbb{Z}_N$ consiste de los elementos cuyos cuadrados pertenecen al intervalo $[-\delta N/2, \delta N/2]$ entonces es posible ver que si N es suficientemente grande la densidad de A es muy cercana a δ, la norma $\|\delta - A\|_{U^2}$ es muy pequeña y sin embargo es

$$\mathbb{E}_{x,h} A(x)A(x+h)A(x+2h)A(x+3h) \geq c\delta^3$$

para cierta constante $c > 0$. Pero puesto que para un conjunto aleatorio de la misma densidad uno espera que esto sea aproximadamente igual a δ^4, se sigue que la norma U^2 es insuficiente para controlar a las progresiones ariméticas de longitud 4. Esto será logrado sin embargo mediante la introducción de las normas de Gowers superiores. Necesitaremos primero la siguiente definición.

Definición 6.2. *Sea $d \geq 1$. Dada una $\{0,1\}^d$-tupla de funciones $(f_\epsilon)_{\epsilon \in \{0,1\}^d}$, $f_\epsilon : G \to \mathbb{C}$, definimos el producto interno*[1] *de Gowers de dimensión d como*

$$\langle (f_\epsilon)_{\epsilon \in \{0,1\}^d} \rangle_{U^d} = \mathbb{E}_{x,h_1,\ldots,h_d} \left(\prod_{\epsilon \in \{0,1\}^d} C^{|\epsilon|} f_\epsilon \left(x + \sum_{1 \leq i \leq d} \epsilon_i h_i \right) \right), \qquad (6.1)$$

donde C denota la operación de conjugación y $|\epsilon|$ la cantidad de coordenadas no nulas de ϵ.

Notar de esta definición que si $(f_\epsilon)_{\epsilon \in \{0,1\}^d}$ no depende de la última coordenada entonces podemos reescribir el lado derecho de (6.1) como

$$\mathbb{E}_{x,h_1,\ldots,h_d} \left(\prod_{\epsilon' \in \{0,1\}^{d-1}} C^{|\epsilon'|} \left(f_{\epsilon'} \left(x + \sum \epsilon'_i h_i \right) \overline{f_{\epsilon'} \left(x + h_d + \sum \epsilon'_i h_i \right)} \right) \right).$$

Esto a su vez puede ser reescrito como

$$\mathbb{E}_{h_1,\ldots,h_{d-1}} \left| \mathbb{E}_y \prod_{\epsilon' \in \{0,1\}^{d-1}} C^{|\epsilon'|} f_{\epsilon'} \left(y + \sum \epsilon'_i h_i \right) \right|^2.$$

[1] Aquí se entiende el producto interno como la generalización natural del producto interno usual al caso de 2^d variables. No necesitaremos esto en el trabajo sin embargo.

Concluimos entonces que en tal caso es $\langle (f_\epsilon)_{\epsilon\in\{0,1\}^d}\rangle_{U^d} \geq 0$. En particular, si es $f_\epsilon = f$ para todo $\epsilon \in \{0,1\}^d$, tenemos

$$\langle (f)_{\epsilon\in\{0,1\}^d}\rangle_{U^d} \geq 0. \tag{6.2}$$

Introducimos entonces la siguiente definición.

Definición 6.3. *Sea $d \geq 2$ y G un grupo abeliano. Para toda $f : G \to \mathbb{C}$ definimos la norma U^d de Gowers como*

$$\|f\|_{U^d}^{2^d} := \langle (f)_{\epsilon\in\{0,1\}^d}\rangle_{U^d} = \mathbb{E}_{x,h_1,\ldots,h_d} \prod_{\epsilon\in\{0,1\}^d} C^{|\epsilon|} f\left(x + \sum h_i\epsilon_i\right).$$

Nota. Consideraremos también la norma U^1 definida análogamente. Sin embargo, en tal caso no se tratará genuinamente de una norma sino de una seminorma, puesto que es

$$\|f\|_{U^1} = |\mathbb{E}_x f(x)|$$

y esto puede ser nulo sin necesidad de que f lo sea.

Procederemos a demostrar que la norma U^d es efectivamente una norma para todo $d \geq 2$. Sabemos de (6.2) que tal cantidad es efectivamente positiva. Análogamente a lo hecho anteriormente, tenemos además

$$\begin{aligned}\langle (f_\epsilon)_{\epsilon\in\{0,1\}^d}\rangle_{U^d} =& \mathbb{E}_{h_1,\ldots,h_{d-1}} \left(\mathbb{E}_y \left(\prod_{\epsilon'\in\{0,1\}^{d-1}} C^{|\epsilon'|} f_{\epsilon',0}\left(y + \sum h_i\epsilon_i\right)\right)\right. \\ &\left. \times \overline{\mathbb{E}_{y'} \left(\prod_{\epsilon'\in\{0,1\}^{d-1}} C^{|\epsilon'|} f_{\epsilon',1}\left(y' + \sum h_i\epsilon_i\right)\right)}\right).\end{aligned}$$

Luego, aplicando la desigualdad de Cauchy-Schwarz respecto a las variables $h_1, \ldots, h_{d-1}$, obtenemos

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

El mismo resultado vale si en lugar de la variable ϵ_d se separa cualquier otra coordenada de ϵ. Aplicando entonces tal desigualdad para cada coordenada obtenemos la siguiente *desigualdad de Gowers-Cauchy-Schwarz*

$$|\langle (f_\epsilon)_{\epsilon\in\{0,1\}^d}\rangle_{U^d}| \leq \prod_{\epsilon\in\{0,1\}^d} \|f_\epsilon\|_{U^d}. \tag{6.3}$$

Usando esto y la multilinealidad del producto interno tenemos

$$\begin{aligned}|\langle (f+g)_{\epsilon\in\{0,1\}^d}\rangle_{U^d}| &\leq \sum_{j\leq 2^d} \binom{2^d}{j} \|f\|_{U^d}^j \|g\|_{U^d}^{2^d-j} \\ &= (\|f\|_{U^d} + \|g\|_{U^d})^{2^d}\end{aligned}$$

y en consecuencia la desigualdad triangular

$$\|f+g\|_{U^d} \leq \|f\|_{U^d} + \|g\|_{U^d}\,. \tag{6.4}$$

Dada f, definiendo ahora $f_\epsilon = 1$ cuando es $\epsilon_d = 1$ y $f_\epsilon = f$ en otro caso, obtenemos también de (6.3) la desigualdad

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

Concluimos entonces que es

$$\|f\|_{U^{d-1}} \leq \|f\|_{U^d} \tag{6.5}$$

para todo $d \geq 2$. En consecuencia, por las observaciones anteriores, para probar que la norma U^d es efectivamente una norma para todo $d \geq 2$ bastará con demostrar que es $\|f\|_{U^2} = 0$ si y sólo si es $f \equiv 0$.

Para probar esto último realizaremos una simple observación que es de gran utilidad a la hora de estudiar la norma U^2.

Lema 6.1. *Sea G un grupo abeliano finito y sea $f : G \to \mathbb{C}$. Entonces $\|f\|_{U^2} = \|\hat{f}\|_4$.*

Demostración. Notamos que es

$$\|f\|_{U^2}^4 = \mathbb{E}_{x+y=z+w} f(x)f(y)\overline{f(z)f(w)}. \tag{6.6}$$

Definiendo la convolución

$$f * g(x) = \mathbb{E}_{y+z=x} f(y)g(z)$$

es fácil verificar que se satisface la identidad

$$\hat{(f * g)} = \hat{f}\hat{g}.$$

Utilizando además la identidad de Parseval

$$\langle f, g\rangle = \langle \hat{f}, \hat{g}\rangle$$

vemos de (6.6) que es

$$\|f\|_{U^2}^4 = \langle f*f, f*f\rangle = \langle \hat{f}^2, \hat{f}^2\rangle = \sum_\psi |\hat{f}(\psi)|^4,$$

donde la suma de la derecha se extiende sobre todos los caracteres $\psi : G \to \mathbb{C}$. Tomando raíces cuartas el resultado se sigue. □

Obtenemos así del Lema 6.1 y las observaciones anteriores, que las normas U^d de Gowers son efectivamente normas para $d \geq 2$.

Como habíamos visto antes, la norma U^2 controlaba varios aspectos de la aleatoriedad de una función. En particular, vimos que si para un conjunto $A \subseteq G$ de densidad δ su función característica satisface que $\|\delta - A\|_{U^2}$ es pequeño, entonces A poseerá aproximadametne la cantidad de progresiones aritméticas de longitud tres que se esperan de un conjunto aleatorio de la misma densidad. Vimos también que esto sin embargo no es cierto para progresiones aritméticas de longitud cuatro. En particular, el ejemplo que dimos posee una propiedad que resulta ser fundamental y es su naturaleza cuadrática.

En su demostración del teorema de Szemerédi, Gowers probó que si A es un subconjunto de $\mathbb{Z}_N$ de densidad δ tal que $\|\delta - A\|_{U^r}$ es suficientemente pequeño, con $r \geq 2$, entonces se satisface

$$\mathbb{E}_{x,h} A(x)A(x+h)\dots A(x+rh) \sim \delta^{r+1}. \tag{6.7}$$

A los conjuntos que satisfacen la mencionada propiedad respecto a la norma U^r se los suele llamar *uniformes de grado* $r-1$. Luego, (6.7) nos dice que un conjunto uniforme de grado $r-1$ posee aproximadamente la cantidad de progresiones aritméticas de longitud $r+1$ que se esperan en un conjunto aleatorio de la misma densidad.

De todas formas, la importancia de las normas de Gowers no radica únicamente en la posibilidad de controlar la cantidad de progresiones aritméticas. El conjunto $\{x, x+h, \dots, x+(r-1)h\}$ puede ser visto como una colección de r formas lineales en las variables x y h. Puede demostrarse que para cualquier colección de formas lineales independientes en cualquier cantidad de variables, existe un r para el cual todo conjunto que sea suficientemente uniforme de grado r poseerá aproximadamente la misma cantidad de las correspondientes configuraciones que un conjunto aleatorio. Tal resultado fue demostrado por Green y Tao [27]. Sin embargo, se desconoce una manera de determinar cuál es el mínimo valor de r que una determinada configuración requiere. A éste respecto, Gowers y Julia Wolf [22] conjeturaron que la respuesta es el menor r para el cual las r-potencias de las formas lineales involucradas son linealmente independientes.

6.2. Teoría inversa de la norma de Gowers

Retornamos ahora a nuestra tarea específica de deducir la Proposición 5.3 de la Proposición 5.2. El plan para llevar a cabo tal tarea será demostrar que si f es una función mayorizada por una medida r-pseudoaleatoria, entonces existirá una función $g : [0,1] \to \mathbb{R}$ cercana a f en la norma U^{r-1} y con la misma densidad que esta (en particular, de densidad positiva). Pero entonces, por observaciones anteriores, la esperanza de f en progresiones

aritméticas será similar a la de g. Pero puesto que para tal g tendremos el teorema de Szemeredi a nuestra disposición, podremos estimar tal cantidad y en consecuencia deducir la información deseada sobre f.

Puesto que lo que deseamos es encontrar una función mayorizada por 1 que aproxime en la norma U^{r-1} a una función mayorizada por una medida r-pseudoaleatoria, será de fundamental importancia que estos dos mayorantes estén efectivamente cerca en la mencionada norma. Esto lo probamos en el siguiente lema.

Lema 6.2. *Supongamos que ν es una medida r-pseudoaleatoria. Entonces, es*

$$\|\nu - 1\|_{U^d} = o(1)$$

para todo $1 \leq d \leq r-1$.

Demostración. Comenzamos notando que por (6.5) bastará con probar el resultado para $d = r-1$. Deseamos mostrar la estimación

$$\mathbb{E}_{x,h_1,\ldots,h_{r-1}} \prod_{\epsilon \in \{0,1\}^{r-1}} \left(\nu\left(x + \sum \epsilon_i h_i\right) - 1\right) = o(1). \tag{6.8}$$

Es claro entonces que bastará probar

$$\mathbb{E}_{x,h_1,\ldots,h_{r-1}} \prod_{\epsilon \in A} \nu\left(x + \sum \epsilon_i h_i\right) = 1 + o(1) \tag{6.9}$$

para todo $A \subseteq \{0,1\}^{r-1}$, puesto que en tal caso (6.8) será igual a

$$\sum_{A \subseteq \{0,1\}^{r-1}} \left((-1)^{|A|} + o(1)\right) = (1-1)^{2^{r-1}} + o(1) = o(1).$$

Ahora bien, puesto que es evidente que ninguna de las formas lineales $x + \sum \epsilon_i h_i$, $\epsilon \in \{0,1\}^{r-1}$, es un múltiplo racional de otra, y dado que ν es una medida r-pseudoaleatoria, podemos aplicar la condición de $(2^{r-1}, r, 1)$-formas lineales para concluir la validez de (6.9). El lema queda así demostrado. □

En la combinatoria aditiva, se llama *teorema directo* a aquel que parte de la descripción de un conjunto para demostrar que éste satisface ciertas propiedades. Un ejemplo sencillo de esto es partir de un conjunto A que consiste de una progresión aritmética y concluir que el conjunto de sumas de A es pequeño en relación a A (precisamente, es $|A + A| = 2|A| - 1$). Por el contrario, se llama *teorema inverso* a aquel que usa como hipótesis que el conjunto en cuestión satisface una cierta propiedad, para concluir de esto una descripción de tal conjunto. En el caso óptimo, tal teorema nos dirá que un conjunto tiene una determinada propiedad si y sólo si tiene cierta forma específica. Un ejemplo (muy importante) de teorema inverso es

el teorema de Freiman (ver [13], [40]). Éste nos dice que para todo C existen constantes k y K tales que si A es un conjunto con $|A+A| \leq C|A|$, entonces A está contenido en una progresión aritmética de dimensión a lo sumo k y cardinalidad a lo sumo $K|A|$.

Debido a las mencionadas propiedades de las normas U^d, un tal teorema inverso para estas normas que nos permita describir las características de una función con norma U^d grande, sería de fundamental importancia. Recordar que nuestra intención es mostrar que podemos, mediante perturbaciones pequeñas en la norma U^{r-1}, transformar una función de cierto tipo en otra más manejable. Si esto fuese imposible, un teorema inverso adecuado nos permitiría deducir la forma específica de tales perturbaciones y en consecuencia (con suerte) derivar un absurdo.

Sin embargo, el problema del teorema inverso para la norma de Gowers no es en lo absolto sencillo. En el caso de la norma U^2 es fácil ver, utilizando el Lema 6.1, que si la norma U^2 de una función f es grande (y la función f está acotada en la norma L^2), entonces f ha de estar muy correlacionado con algún caracter. De todas formas, esto ya no es cierto en el caso de la norma U^3. Como vimos antes, los comportamientos cuadráticos juegan un rol relevante en éste caso. Por ejemplo, si consideramos la función $f(x) = \exp(4\pi i x^2/N)$, entonces de la identidad

$$\begin{aligned} &x^2 - (x+a)^2 - (x+b)^2 - (x+c)^2 + (x+a+b)^2 \\ &+ (x+a+c)^2 + (x+b+c)^2 - (x+a+b+c)^2 = 0 \end{aligned}$$

deducimos fácilmente que es $\|f\|_{U^3} = 1$. Sin embargo, f así definido no se correlaciona en forma significante con ningún caracter. Vemos entonces que será necesario incluir en consideración las funciones de éste tipo, que suele llamárselas funciones de fase cuadrática . A éste respecto, Green y Tao demostraron que una función con norma U^3 grande (y controlada en otros aspectos) ha de correlacionarse con ciertos objetos que son generalizaciones de tales funciones de fase cuadrática [25]. Más aún, conjeturaron precisamente el resultado a esperar para las normas de Gowers superiores [27]. La veracidad de tal conjetura tendría implicaciones importantísimas, entre ellas una estimación precisa para la frecuencia en los primos de una gran cantidad de patrones lineales. Muy recientemente, Green, Tao y Tamar Ziegler [26] lograron demostrar tal conjetura para la norma U^4 con un argumento que parece generalizarse para los casos superiores, por lo que es probable que tal problema quede completamente resuelto en el corto plazo.

No necesitaremos en nuestro trabajo resultados de semejante fuerza. Por el contrario, nos contentaremos con deducir un teorema inverso suave (y algo artificial) que nos permitirá concluir que una función con norma de Gowers grande se correlacionará con algún miembro de una amplia familia de funciones sobre las cuales podemos obtener información no trivial. Con tal propósito en mente introducimos la siguiente definición para un valor de r fijo

Definición 6.4. *Dada $f : G \to \mathbb{C}$, sea $\mathcal{D}f$ la función*

$$\mathcal{D}f(x) = \mathbb{E}_{h_1,\dots,h_{r-1}} \prod_{\substack{\epsilon\in\{0,1\}^{r-1}\\ \epsilon\neq 0}} f\left(x + \sum \epsilon_i h_i\right).$$

Si X es un conjunto de funciones de G en $\mathbb{C}$, entonces una función básica anti-uniforme (respecto a X) es una función de la forma $\mathcal{D}f$ con $f \in X$.

Es trivial entonces que si $\|f\|_{U^{r-1}}$ es grande para cierta función $f \in X$, entonces f se correlacionará significativamente con una función básica anti-unfiorme respecto a X (precisamente, con $\mathcal{D}f$). Lo importante será entonces ver que podemos decir sobre estas últimas funciones. En particular, veremos que podemos controlar los productos de tales funciones, lo cual probará ser de fundamental importancia para la demostración de la Proposición 5.2.

6.3. Normas PCA

Una norma en $\mathbb{C}^N$ se dice *algebraica* si satisface $\|\mathbf{1}\| = 1$ y $\|fg\| \leq \|f\|\|g\|$ para todo par de funciones $f, g \in \mathbb{C}^N$. Una de las propiedades especiales de las normas algebraicas es que son particularmente dúctiles para demostrar teoremas de transferencia en los cuales se desea trasladar funciones de un espacio en otro sin alterar demasiado la norma en cuestión. Más precisamente, tales teoremas resultan más fáciles de conseguir cuando la norma *dual* a la estudiada es algebraica.

Puesto que ya hemos señalado que nuestra intención es obtener un teorema de transferencia, sería muy útil que las normas duales a las de Gowers resulten ser algebraicas. Si bien esto no es cierto, una observación crucial de Green y Tao es que tal norma dual sí posee suficientes propiedades en común con las normas algebraicas como para poder llevar a cabo tal teorema de transferencia. La siguiente definición encapsula las mencionadas propiedades (es fácil deducir las correspondientes analogías para el caso en que la norma dual efectivamente es algebraica). Notar que de ahora en más nos restringiremos a funciones de $\mathbb{Z}_N$ en $\mathbb{R}$ identificándolas con elementos de $\mathbb{R}^N$ en la forma evidente.

Definición 6.5. *Sea $\|.\|$ una norma en $\mathbb{R}^N$ y sea X un subconjunto acotado de $\mathbb{R}^N$ con span$(X) = \mathbb{R}^N$. Decimos que $\|.\|$ es una norma predual cuasi algebraica, o norma PCA, respecto a X si existe un operador (no lineal) $\mathcal{D} : \mathbb{R}^N \to \mathbb{R}^N$, una función estrictamente decreciente $c : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, una función creciente $C : \mathbb{N} \to \mathbb{R}$ y una constante absoluta $\overline{C}$ tales que se satisfacen las siguientes condiciones:*

1. *$\langle f, \mathcal{D}f\rangle \leq 1$ para toda $f \in X$;*
2. *$\langle f, \mathcal{D}f\rangle \geq c(\varepsilon)$ para toda $f \in X$ con $\|f\| \geq \varepsilon$;*

3. $\|\mathcal{D}f\|_\infty \leq \overline{C}$ *para toda* $f \in X$;

4. $\|\mathcal{D}f_1 \dots \mathcal{D}f_K\|^* \leq C(K)$ *para cualquier elección de* $f_1, \dots, f_K \in X$.

El resto de éste capítulo estará dedicado a demostrar que las normas de Gowers son preduales cuasi algebraicas con respecto a un conjunto X apropiado. En el próximo capítulo demostraremos el deseado teorema de transferencia para las normas PCA. Estas dos afirmaciones, combinadas con un teorema generalizado de von Neumann que será probado también en el próximo capítulo, nos permitirán obtener la Proposición 5.3 y finalizar así la demostración del Teorema 1.1.

De aquí hasta el final del capítulo fijaremos un valor $r \geq 3$. Fijaremos también una medida r-pseudoaelatoria $\nu \in \mathbb{R}^N$ y nos concentraremos en el estudio de la norma U^{r-1}. Definimos X como el conjunto de funciones $f \in \mathbb{R}^N$ acotadas puntualmente por $\nu + 1$ (notar que trivialmente es $\text{span}(X) = \mathbb{R}^N$). Como los operadores $\mathcal{D}$ que aparecen en la definición de norma PCA tomaremos por supuesto a las funciones básicas anti-uniformes respecto a X que fueron definidas en la Definición 6.4, pero normalizadas por un factor de 2^{-2^r}. Pasaremos entonces a demostrar el mencionado resultado.

Proposición 6.1. *En* $\mathbb{R}^N$*, la norma* U^{r-1} *es una norma PCA respecto a* X*. Más aún,* $\mathcal{D}, c, C, \overline{C}$ *no dependen de* N*.*

Demostración. Sea $f \in X$. Por definición de X es $f \leq \nu + 1$. En consecuencia, tenemos

$$\begin{aligned}
\langle f, \mathcal{D}f \rangle &= 2^{-2^r} \|f\|_{U^{r-1}}^{2^{r-1}} \\
&\leq 2^{-2^r} \left(\|\nu\|_{U^{r-1}} + \|1\|_{U^{r-1}} \right)^{2^{r-1}} \\
&= 2^{-2^r} \left(2 + o(1) \right)^{2^{r-1}} \\
&\leq 1,
\end{aligned}$$

donde hemos usado la condición de formas lineales para ν. Vemos entonces que se cumple la propiedad (1) de la Definición 6.5.

Análogamente, la condición (2) de tal definición nos pide que sea $\langle f, \mathcal{D}f \rangle \geq c(\varepsilon)$ para todo $f \in X$ con $\|f\|_{U^{r-1}} \geq \varepsilon$. Esto sin embargo es trivial en nuestro caso por definición de $\mathcal{D}$, tomando $c(\varepsilon) = 2^{-2^r} \varepsilon^{2^{r-1}}$.

Para ver la propiedad (3) utilizaremos la condición de formas lineales. Pero antes, necesitaremos el siguiente sencillo lema que nos dice que las medidas pseudoaleatorias forman una estrella alrededor de la función constantemente igual a 1.

Lema 6.3. *Si* ν *es una medida* r*-pseudoaleatoria, entonces también lo es* $(\nu + 1)/2$*.*

Demostración. Es evidente que tal función es no negativa y posee esperanza igual a $1 + o(1)$. Que satisface la condición de formas lienales se sigue inmediatamente de la multilinealidad del producto interno de Gowers. La condición de correlación se verifica de la misma forma. □

Acotando f por $2(\nu + 1)/2 = 2\nu_{1/2}$ vemos que bastará probar

$$\|\mathcal{D}\nu_{1/2}\|_\infty \leq 1 + o(1).$$

Ahora bien, es

$$\mathcal{D}\nu_{1/2} = \mathbb{E}_{h_1,\dots,h_{r-1}} \prod_{\substack{\epsilon\in\{0,1\}^{r-1}\\ \epsilon\neq 0}} \nu_{1/2}(x + \sum \epsilon_i h_i).$$

Aplicando entonces el Lema 6.3 y la condición de formas lineales (con todos los b_i iguales a x), vemos que esto es igual a $1 + o(1)$ y en consecuencia, concluimos que se satisface la condición (3).

La condición (4) es la más difícil de verificar. Deseamos probar la existencia de una función creciente $C : \mathbb{N} \to \mathbb{R}$ que satisfaga

$$\|\mathcal{D}f_1 \dots \mathcal{D}f_K\|^*_{U^{r-1}} \leq C(K)$$

para cualquier elección de las funciones $f_1, \dots, f_K \in X$.

Consideremos entonces $f \in \mathbb{R}^N$ con $\|f\|_{U^{r-1}} \leq 1$. Bastará probar que para toda tal f es

$$\langle f, \prod_{j=1}^{K} \mathcal{D}f_j \rangle = O_K(1).$$

Notar que el lado izquierdo de esto puede ser escrito como

$$\mathbb{E}_x f(x) \prod_{j=1}^{K} \mathbb{E}_{h_1^{(j)},\dots,h_{r-1}^{(j)}} \prod_{\substack{\epsilon\in\{0,1\}^{r-1}\\ \epsilon\neq 0}} f_j(x + \sum \epsilon_i h_i^{(j)}). \tag{6.10}$$

Dados $h_1, \dots, h_{r-1} \in \mathbb{Z}_N$ escribimos $h_i^{(j)} = h_i + H_i^{(j)}$. Promediando sobre todas las tales $k-1$-tuplas obtenemos que (6.10) es igual a

$$\mathbb{E}_x f(x) \mathbb{E}_{h_1,\dots,h_{r-1}} \prod_{j=1}^{K} \mathbb{E}_{H_1^{(j)},\dots,H_{r-1}^{(j)}} \prod_{\substack{\epsilon\in\{0,1\}^{r-1}\\ \epsilon\neq 0}} f_j(x + \sum \epsilon_i H_i^{(j)} + \sum \epsilon_i h_i).$$

Si expandimos el producto interior e intercambiamos las esperanzas, en particular llevando afuera las esperanzas que dependen de los $H_i^{(j)}$, vemos que podemos reescribir esto en términos del producto interno de Gowers como

$$\mathbb{E}_{H\in\left(\mathbb{Z}_N^{r-1}\right)^K} \langle (f_{\epsilon,H})_{\epsilon\in\{0,1\}^{r-1}} \rangle_{U^{r-1}}, \tag{6.11}$$

donde es $H := \left(\left(H_1^{(1)}, \dots, H_{r-1}^{(1)}\right), \dots, \left(H_1^{(K)}, \dots, H_{r-1}^{(K)}\right)\right)$, $f_{0,H} := f$ y $f_{\epsilon,H} := g_{\epsilon.H}$ para $\epsilon \neq 0$, donde es $\epsilon.H = \left(\sum \epsilon_i H_i^{(1)}, \dots, \sum \epsilon_i H_i^{(K)}\right)$ y

$$g_{u^{(1)},\dots,u^{(K)}}(x) := \prod_{j=1}^{K} f_j(x + u^{(j)}). \tag{6.12}$$

Aplicando la desigualdad de Gowers-Cauchy-Schwarz (6.3) podemos acotar (6.11) por

$$\mathbb{E}_{H \in (\mathbb{Z}_N^{r-1})^K} \|f\|_{U^{r-1}} \prod_{\substack{\epsilon \in \{0,1\}^{r-1} \\ \epsilon \neq 0}} \|g_{\epsilon.H}\|_{U^{r-1}}.$$

Puesto que por hipótesis es $\|f\|_{U^{r-1}} \leq 1$, bastará entonces probar la estimación

$$\mathbb{E}_{H \in (\mathbb{Z}_N^{r-1})^K} \prod_{\substack{\epsilon \in \{0,1\}^{r-1} \\ \epsilon \neq 0}} \|g_{\epsilon.H}\|_{U^{r-1}} = O_K(1).$$

Pero a su vez, aplicando a esto la desigualdad de las medias aritméticas y geométricas en la misma forma en que se ha utilizado previamente, vemos que bastará probar

$$\mathbb{E}_{H \in (\mathbb{Z}_N^{r-1})^K} \|g_{\epsilon.H}\|_{U^{r-1}}^{2^{r-1}} = O_K(1),$$

para cada $\epsilon \in \{0,1\}^{r-1}$, $\epsilon \neq 0$. Notar que en tal caso, el mapa $H \mapsto \epsilon.H$ es un cubrimiento uniforme de $\mathbb{Z}_N^K$ por $(\mathbb{Z}_N^{r-1})^K$. Luego, el lado izquierdo de la última expresión es igual a

$$\mathbb{E}_{u^{(1)},\dots,u^{(K)}} \|g_{u^{(1)},\dots,u^{(K)}}\|_{U^{r-1}}^{2^{r-1}}.$$

Expandiendo la norma de Gowers y usando (6.12), vemos que esto es igual a

$$\mathbb{E}_{u^{(1)},\dots,u^{(K)}} \mathbb{E}_{x,h_1,\dots,h_{r-1}} \prod_{\epsilon' \in \{0,1\}^{r-1}} \prod_{j=1}^{K} f_j(x + u^{(j)} + \sum \epsilon_i' h_i). \tag{6.13}$$

Usando que es $f \leq \nu + 1$, el Lema 6.3 y (6.13), vemos que bastará probar

$$\mathbb{E}_{x,h_1,\dots,h_{r-1}} \left(\mathbb{E}_u \prod_{\epsilon \in \{0,1\}^{r-1}} \nu(x + u + \sum \epsilon_i' h_i) \right)^K = O_K(1).$$

Aplicando el cambio de variables $y := x + u$ una esperanza se vuelve redundante y en consecuencia el problema se reduce a mostrar

$$\mathbb{E}_{h_1,\dots,h_{r-1}} \left(\mathbb{E}_y \prod_{\epsilon \in \{0,1\}^{r-1}} \nu(y + \sum \epsilon_i' h_i) \right)^K = O_K(1).$$

Es aquí donde la condición de correlación cumplirá su propósito. En efecto, tal condición implica la existencia de una función de peso $\gamma = \gamma_{2^{r-1}}$ con $\mathbb{E}(\gamma^q) = O_q(1)$ para todo q y que satisface la desigualdad

$$\mathbb{E}_y \prod_{\epsilon \in \{0,1\}^{r-1}} \nu(y + \sum \epsilon_i' h_i) \leq \sum_{\substack{\epsilon', \epsilon'' \in \{0,1\}^{r-1} \\ \epsilon' \neq \epsilon''}} \gamma\left(\sum h_i(\epsilon_i' - \epsilon_i'')\right).$$

Aplicando entonces la desigualdad triangular vemos que bastará probar

$$\mathbb{E}_{h_1,\ldots,h_{r-1}} \left(\gamma\left(\sum h_i(\epsilon_i' - \epsilon_i'')\right)\right)^K = O_K(1),$$

para todo par $\epsilon' \neq \epsilon''$ en $\{0,1\}^{r-1}$. Pero dado que en tal caso el mapa $h_1, \ldots, h_{r-1} \mapsto \sum h_i(\epsilon_i' - \epsilon_i'')$ es un cubrimiento uniforme de $\mathbb{Z}_N$ por $(\mathbb{Z}_N)^{r-1}$ se sigue que el lado izquierdo es simplemente $\mathbb{E}(\gamma^K)$. Dado que sabemos que esto es $O_K(1)$, el resultado queda así demostrado. □

Capítulo 7

El principio de transferencia y el teorema generalizado de Von Neumann

El primer objetivo de éste capítulo será demostrar el mencionado teorema de transferencia que dada una función mayorada por una medida pseudoaleatoria nos permite, mediante una perturbación pequeña en la norma de Gowers, obtener una función acotada con la misma densidad. Esto será logrado a través de un teorema abstracto que nos ofrece tal conclusión para cualquier norma PCA. Tal teorema abstracto fue formulado por Timothy Gowers [21] y a continuación expondremos la demostración dada por este autor. Cabe destacar que un resultado análogo fue dado en forma independiente por Omer Reingold, Luca Trevisan, Madhul Tulsiani y Salil Vadhan en el contexto de la teoría de juegos [38].

Para la demostración de éste resultado emplearemos el teorema de Hahn-Banach en la formulación dada a continuación. Notar que al igual que antes utilizaremos durante éste capítulo la convención de interpretar a los elementos de $\mathbb{R}^N$ como funciones de $\mathbb{Z}_N$ en $\mathbb{R}$ en la forma evidente.

Proposición 7.1 (Teorema de Hahn-Banach)**.** *Sea $K \subseteq \mathbb{R}^N$ un cuerpo convexo y $f \in \mathbb{R}^N$ una función que no está contenida en K. Entonces, existe una constante α y un funcional lineal no nulo ϕ, tal que se satisfacen las desigualdades $\langle f, \phi\rangle \geq \alpha$ y $\langle g, \phi\rangle \leq \alpha$ para todo $g \in K$.*

7.1. Descomposición y la norma CBA

El siguiente corolario del teorema de Hahn-Banach nos será de gran utilidad.

Corolario 7.1. *Sean $K_1, \ldots, K_m$ cuerpos convexos cerrados de $\mathbb{R}^N$ que contienen al 0. Sean $c_1, \ldots, c_m \in \mathbb{R}_{\geq 0}$ y supongamos que $f \in \mathbb{R}^N$ es una función*

que no puede ser escrita en la forma $f = f_1 + \ldots + f_m$ *con* $f_i \in c_i K_i$. *Entonces, existe un funcional lineal* ϕ *satisfaciendo las desigualdades* $\langle f, \phi\rangle > 1$ *y* $\langle g, \phi\rangle \leq c_i^{-1}$ *para todo* $1 \leq i \leq m$ *y todo* $g \in K_i$.

Demostración. Escribimos $K = \sum_{1\leq i\leq m} c_i K_i$. Entonces, K es un cuerpo convexo cerrado que por hipótesis no contiene a f. El hecho de que K sea cerrado nos permite concluir la existencia de un $0 < \delta < 1$ tal que $\delta f \notin K$. Aplicando entonces la Proposición 7.1, vemos que existe un funcional lineal no nulo ϕ con $\delta\langle f, \phi\rangle \geq \alpha$ y $\langle g, \phi\rangle \leq \alpha$ para todo $g \in K$ y para cierta constante α.

Deseamos ver que α puede tomarse distinto de cero. Para esto notamos que por hipótesis $0 \in K$ y que además, puesto que K es cerrado, existirá una bola euclídea B con $\delta f \notin K + B$. Vemos entonces que efectivamente podemos suponer $\alpha \neq 0$. Dividiendo por α concluimos de lo anterior la existencia de un funcional lineal ϕ con $\langle f, \phi\rangle \geq \delta^{-1} > 1$ y $\langle g, \phi\rangle \leq 1$ para todo $g \in c_i K_i \subseteq K$. En particular, será $\langle g, \phi\rangle \leq c_i^{-1}$ para todo $g \in K_i$ y el resultado queda así demostrado. □

El anterior corolario nos permitirá deducir que si una función f mayorada por una medida pseudoaleatoria no puede escribirse de la forma $g + h$ con g acotada y h pequeña en la norma de Gowers, entonces deberá existir un funcional lineal ϕ con propiedades especiales que trataremos de explotar.

Para llevar a cabo éste plan necesitaremos algunos lemas. De ahora en más, $X, \mathcal{D}, c, C$ y $\overline{C}$ estarán dados por la Definición 6.5. Además, siguiendo la convención dada en la segunda parte de la Definición 6.4, llamaremos a una función de la forma $\mathcal{D}f$ con $f \in X$ una función basica anti-uniforme respecto a X. Puesto que la correlación de una función f con tales funciones básicas anti-uniformes jugará un papel preponderante, introducimos ahora una norma que nos da una medida de tal correlación. Definimos entonces la norma $\|.\|_{CBA}$ (correlación básica anti-uniforme) como

$$\|f\|_{CBA} := \text{máx}\left\{|\langle f, \mathcal{D}g\rangle| : g \in X\right\}.$$

Para estudiar el dual de esta norma utilizaremos el siguiente lema. Recordar que el dual de una norma $\|.\|$ está definido por

$$\|g\|^* := \text{máx}\left\{\langle f, g\rangle : \|f\| \leq 1\right\}.$$

Lema 7.1. *Sea* $\Sigma \subseteq \mathbb{R}^N$ *un conjunto que genera a* $\mathbb{R}^N$. *Si definimos en* $\mathbb{R}^N$ *una norma* $\|.\|$ *mediante la fórmula*

$$\|f\| = \text{ínf}\left\{\sum_{i=1}^{k} |\lambda_i| : f = \sum_{i=1}^{k} \lambda_i \sigma_i, \sigma_1, \ldots, \sigma_k \in \Sigma\right\}.$$

entonces esto define efectivamente una norma cuya norma dual $\|.\|^*$ *viene dada por la fórmula*

$$\|f\|^* = \text{máx}\left\{|\langle f, \sigma\rangle| : \sigma \in \Sigma\right\},$$

Demostración. Es fácil verificar que esto define una norma. Supongamos entonces que es $|\langle z, \sigma\rangle| \geq 1$ para algún $\sigma \in \Sigma$. Puesto que claramente es $\|\sigma\| \leq 1$, ha de ser también $\|z\|^* \geq 1$, de donde concluimos que $\|z\|^* \geq \text{máx}\{|\langle f, \sigma\rangle| : \sigma \in \Sigma\}$.

Supongamos ahora que es $\|z\|^* > 1$, por lo que existe un x con $\|x\| \leq 1$ y $|\langle x, z\rangle| \geq 1 + \varepsilon$ para cierto $\varepsilon > 0$. Por definición de $\|.\|$ podemos encontrar $\sigma_1, \ldots, \sigma_k \in \Sigma$ con $x = \sum_{i=1}^k \lambda_i \sigma_i$ y $\sum_{i=1}^k |\lambda_i| < 1 + \varepsilon$. Luego, es

$$\sum_{i=1}^{k} |\lambda_i| \, |\langle \sigma_i, z\rangle| > \sum_{i=1}^{k} |\lambda_i|.$$

Concluimos que ha de ser $|\langle \sigma_i, z\rangle| > 1$ para cierto $1 \leq i \leq k$ y en consecuencia $\text{máx}\{|\langle z, \sigma\rangle| : \sigma \in \Sigma\}$ es al menos 1, por lo que el resultado se sigue. □

Por dualidad, obtenemos entonces que en nuestro caso es

$$\|f\|^*_{CBA} = \text{ínf}\left\{\sum_{i=1}^{k} |\lambda_i| : f = \sum_{i=1}^{k} \lambda_i \mathcal{D} f_i, f_1, \ldots, f_k \in X\right\}. \tag{7.1}$$

Notar que la condición (2) de la Definición 6.5 es entonces equivalente a que para todo $f \in X$ con $\|f\| \geq \varepsilon$, es $\|f\|_{CBA} \geq c(\varepsilon)$.

7.2. Aproximaciones polinómicas

Si $P(x) = a_n x^n + \ldots + a_1 x + a_0$ es un polinomio, definimos R_P como el polinomio $R_P(x) = C(n)\,|a_n|\,x^n + \ldots + C(1)\,|a_1|\,x + |a_0|$ (recordar que C está dado por la Definición 6.5). Por otra parte, si $J : \mathbb{R} \to \mathbb{R}$ es una función contínua y C_1, C_2, δ con constantes positivas, definimos $\rho(C_1, C_2, \delta, J)$ como el doble del ínfimo de $R_P(C_2)$ sobre todos los polinomios P con $|P(x) - J(x)| \leq \delta$ para todo $x \in [-C_1, C_1]$.

Dadas estas definiciones procedemos a probar el siguiente resultado sobre aproximaciones polinómicas que nos será de gran ayuda.

Lema 7.2. *Sean $\|.\|$ una norma PCA, $J : \mathbb{R} \to \mathbb{R}$ una función contínua y C_1, C_2, δ constantes positivas con $C_1 = C_2\overline{C}$. Entonces, existe un polinomio P tal que es $\|P\phi - J\phi\|_\infty \leq \delta$ y $\|P\phi\|^* < \rho(C_1, C_2, \delta, J)$ para todo $\phi \in \mathbb{R}^n$ con $\|\phi\|^*_{CBA} \leq C_2$.*

Nota. Observar que por la expresión (7.1) del dual de la norma CBA, una función ϕ con $\|\phi\|^*_{CBA}$ pequeño se correlacionará con unas pocas funciones básicas anti-uniformes. Puesto que tales funciones poseen un valor acotado en la norma $\|.\|^*$, uno espera que lo mismo suceda con ϕ. La idea de éste lema es aprovechar la similitud de $\|.\|^*$ con una norma algebraica para comprobar tal suposición para todo polinomio que aproxime a ϕ.

Demostración. En primer lugar, es inmediato de la definición de $\rho(C_1, C_2, \delta, J)$ la existencia de un polinomio P satisfaciendo las desigualdades $R_P(C_2) < \rho(C_1, C_2, \delta, J)$ y $|P(x) - J(x)| \leq \delta$ para todo $x \in [-C_1, C_1]$.

Sea ahora $\phi \in \mathbb{R}^N$ una función con $\|\phi\|^*_{CBA} \leq C_2$. De (7.1) sabemos entonces que ϕ es una combinacion lineal acotada de funciones básicas anti-uniformes. Precisamente, dado cualquier $\varepsilon > 0$, existirá un cierto conjunto de funciones $f_1, \ldots, f_k \in X$ con

$$\phi = \sum_{i=1}^{k} \lambda_i \mathcal{D} f_i \tag{7.2}$$

y $\sum_{i=1}^{k} |\lambda_i| < C_2 + \varepsilon$. Puesto que la condición (3) de las normas PCA nos garantiza la desigualdad $\|\mathcal{D}f\|_\infty \leq \overline{C}$ para todo $f \in X$, aplicando la desigualdad triangular obtenemos $\|\phi\|_\infty \leq C_2\overline{C} = C_1$. Vemos entonces que la imágen de ϕ está contenida en el intervalo en el cual sabemos que P aproxima a J. Concluimos así que es $\|P\phi - J\phi\|_\infty \leq \delta$.

Observamos ahora que de (7.2) se sigue que, para todo entero $m \geq 1$ y todo $\varepsilon > 0$, ϕ^m puede escribirse como una combinación lineal de productos de m funciones básicas anti-uniformes, con los valores absolutos de los coeficientes de tal combinación sumando a lo sumo $C_2^m + \varepsilon$. Utilizando la condición (4) de la Definición 6.5 vemos que la norma $\|.\|^*$ de cada uno de estos productos está acotada por $C(m)$, lo cual combinado con la observación anterior nos devuelve la desigualdad $\|\phi^m\|^* \leq C_2^m C(m)$.

Deducimos entonces que si P como arriba se escribe en la forma

$$P(x) = a_n x^n + \ldots + a_1 x + a_0$$

tenemos

$$\begin{aligned} \|P\phi\|^* &= |a_n|\|\phi^n\|^* + \ldots + |a_1|\|\phi\|^* + |a_0| \\ &\leq |a_n|C(n)C_2^n + \ldots + |a_1|C(1)C_2 + |a_0| \\ &= R_P(C_2) \\ &< \rho(C_1, C_2, \delta, J). \end{aligned}$$

Con esto, el lema queda demostrado. □

7.3. El teorema de transferencia

Empleando las observaciones anteriores, pasamos ahora a demostrar el mencionado resultado de transferencia.

Proposición 7.2. *Sean $\mu, \nu \in \mathbb{R}^N$ funciones no negativas con $\mathbb{E}(\mu), \mathbb{E}(\nu) \leq 1$. Sean δ, η constantes positivas. Sea $\|.\|$ una norma PCA en $\mathbb{R}^N$ respecto al conjunto*

$$X = \left\{ f \in \mathbb{R}^N : f(x) \leq \max\left\{\mu(x), \nu(x)\right\}, \forall x \in \mathbb{Z}_N \right\}.$$

Sea $J : \mathbb{R} \to \mathbb{R}$ *la función contínua* $J(x) = (x + |x|)/2$ *y sea*

$$\varepsilon = \delta/2\rho(c(\eta)^{-1}\overline{C}, c(\eta)^{-1}, \delta/4, J).$$

Supongamos además que es $\|\mu - \nu\| \leq \varepsilon$. *Entonces, para toda función* f *con* $0 \leq f \leq \nu$, *existe una función* g *con* $0 \leq g \leq \mu(1-\delta)^{-1}$ *satisfaciendo la desigualdad* $\|f - g\| \leq \eta$.

Demostración. Notar que la afirmación del enunciado es equivalente a la posibilidad de descomponer a toda tal f en la forma $g + h$, con $0 \leq g \leq \mu(1-\delta)^{-1}$ y $\|h\| \leq \eta$. Fijemos entonces $0 \leq f \leq \nu$ y supongamos que tal descomposición no existe. Es claro que por la definición de la norma CBA esto implica también que no existirá una tal descomposición con $0 \leq g \leq \mu(1-\delta)^{-1}$ y $\|h\|_{CBA} \leq c(\eta)$. Aplicando entonces el Corolario 7.1 obtenemos un funcional lineal ϕ satisfaciendo $\langle f, \phi\rangle > 1$, $\langle \phi, g\rangle \leq 1$ para todo g con $0 \leq g \leq \mu(1-\delta)^{-1}$ y $\|\phi\|^*_{CBA} \leq c(\eta)^{-1}$ (en éste último caso el cuerpo convexo consiste por supuesto en los h con $\|h\|_{CBA} \leq 1$).

Escribiendo $\phi_+ = \max\{\phi, 0\}$ notamos que la función g como arriba que maximiza $\langle \phi, g\rangle$ es la que es igual a $\mu(1-\delta)^{-1}$ cuando es $\phi > 0$ e igual a 0 en otro caso. En consecuencia, la primer condición dada sobre ϕ es equivalente a la desigualdad $\langle \mu(1-\delta)^{-1}, \phi_+\rangle \leq 1$. Notando que es $\phi_+ = J\phi$, tenemos entonces $\langle \mu, J\phi\rangle \leq (1-\delta)$.

Aplicamos ahora el Lema 7.2 con $C_2 = c(\eta)^{-1}$ y la constante $\delta/4$. Obtenemos así un polinomio P con $\|J\phi - P\phi\|_\infty \leq \delta/4$ y

$$\|P\phi\|^* < \rho = \rho(c(\eta)^{-1}\overline{C}, c(\eta)^{-1}, \delta/4, J).$$

Luego, es

$$\langle \mu, P\phi\rangle \leq \langle \mu, \phi_+\rangle + \mathbb{E}(\mu)\|J\phi - P\phi\|_\infty \leq 1 - \frac{3\delta}{4}.$$

Dado que μ y ν están cerca en la norma $\|.\|$ y $P\phi$ está acotado en la norma dual, tenemos además

$$\langle \nu, P\phi\rangle \leq \langle \mu, P\phi\rangle + \|P\phi\|^*\|\mu - \nu\| < 1 - \frac{3\delta}{4} + \varepsilon\rho$$

y en consecuencia, es también

$$\langle \nu, \phi_+\rangle \leq \langle \nu, P\phi\rangle + \mathbb{E}(\nu)\|J\phi - P\phi\|_\infty < 1 - \frac{\delta}{2} + \varepsilon\rho.$$

Pero entonces, puesto que por hipótesis es $0 \leq f \leq \nu$, obtenemos

$$1 < \langle f, \phi\rangle \leq \langle f, \phi_+\rangle < 1 - \frac{\delta}{2} + \varepsilon\rho.$$

Pero esto es absurdo, puesto que por la definición de ε dada en el enunciado el lado derecho es igual a 1. Concluimos que f puede descomponerse en la forma deseada y el resultado queda así demostrado. □

Aplicaremos ahora esta proposición al caso que nos concierne. Sea entonces μ la función constantemente igual a 1 y ν una medida r-pseudoaleatoria. La norma PCA en cuestión será por supuesto la norma de Gowers $\|.\|_{U^{r-1}}$. Notar que hemos demostrado que tal norma es PCA respecto al conjunto de la funciones acotadas por $\nu + 1$, por lo cual lo es en particular respecto al conjunto X dado en el enunciado de la Proposición 7.2.

Observar además que si es $f = g + h$ con $\|h\|_{U^{r-1}} \leq \eta$, se sigue de (6.5) y la definición de la norma U^1 que es también

$$\mathbb{E}(g) = \mathbb{E}(f) - \mathbb{E}(h) \geq \mathbb{E}(f) - \eta.$$

Por el Lema 6.2, el $\varepsilon > 0$ del enunciado de la Proposición 7.2 puede tomarse arbitrariamente pequeño (asumiendo N suficientemente grande). Esto nos permite elegir a δ y η también arbitrariamente pequeños (aquí estamos usando que $\mathcal{D}, c, C, \overline{C}$, y en consecuencia ρ, no dependen de N). Finalmente, notamos que reescalando a ν (por un factor que tiende a 1 a medida que N tiende al infinito) podemos suponer que la esperanza de ν es menor o igual a 1 sin alterar con esto la pseudoaleatoriedad de esta función. Uniendo estas observaciones obtenemos de la Proposición 7.2 el siguiente corolario.

Corolario 7.2. *Sea $\delta > 0$. Sea $\nu \in \mathbb{R}^N$ una medida r-pseudoaleatoria. Supongamos que $f \in \mathbb{R}^N$ es δ-denso y satisface $0 \leq f \leq \nu$. Sea $\varepsilon > 0$ arbitrario. Entonces, si N es suficientemente grande, tenemos la descomposición $f = g + h$ con $g \leq 1 + o_\varepsilon(1)$, $\mathbb{E}(g) \geq \delta - o_\varepsilon(1)$ y $\|h\|_{U^{r-1}} \leq \varepsilon$.*

7.4. El teorema generalizado de Von Neumann

El paso final hacia la demostración de la Proposición 5.3 (y en consecuencia del Teorema 1.1) será mostrar que una perturbación pequeña en la norma de Gowers preserva aproximadamente la cantidad de progresiones aritméticas. Esto se logrará a través de la siguiente proposición.

Proposición 7.3 (Teorema generalizado de von Neumann; [24], Prop. 5.3)**.** *Sea ν una medida r-pseudoaleatoria y sean $f_0, \ldots, f_{r-1}$ funciones en $\mathbb{R}^n$ satisfaciendo $|f_j| \leq \nu + 1$ para todo $0 \leq j \leq r-1$. Entonces*

$$\mathbb{E}_{x,h \in \mathbb{Z}_N} \prod_{j=1}^{r-1} f_j(x + jr) = O\left(\inf_{0 \leq j \leq r-1} \|f\|_{U^{r-1}} \right) + o(1). \tag{7.3}$$

Nota. La cota de $\nu + 1$ se debe en éste caso a que el resultado será aplicado al estudio de funciones h como en el enunciado del Corolario 7.2, para las cuales la única cota que dispondremos será de la forma $|h| \leq |f| + |g| \leq \nu + 1$ con un error $o(1)$ que puede ser absorbido por ν.

Demostración. Comenzamos notando que adhiriendo una magnitud $o(1)$ podemos suponer que ν es estrictamente positivo. Notar también que utilizando el Lema 6.3 (dividiendo y multiplicando a cada f_j por 2) podemos

asumir $|f_j| \leq \nu$ a cambio de incrementar la constante implícita en el lado derecho de (7.3) por un factor de 2^r.

Lo primero que haremos es reescribir la progresión aritmética en cuestión en términos de un mayor número de variables, de modo tal que a cada coordenada de la progresión aritmética le corresponda una variable de la cual es independiente. Esto nos otorgará un cierto espacio para trabajar con cada f_j en forma independiente.

Reescribiendo el lado izquierdo de (7.3) como

$$\mathbb{E}_{x,h\in\mathbb{Z}_N} \prod_{j=1}^{r-1} f_j(x + c_j r) \tag{7.4}$$

para ciertos elementos $c_1, \ldots, c_j \in \{-r-1, \ldots, 0, \ldots, r-1\}$ podemos suponer sin pérdida de generalidad que el ínfimo en el lado izquierdo de (7.3) se alcanza en $j = 0$. Asumiremos esto en el resto de la demostración.

Consideramos entonces las variables $y_1, \ldots, y_{r-1} \in \mathbb{Z}_N$ y escribimos

$$h := \sum_{i=1}^{r-1} \frac{y_i}{c_i}. \tag{7.5}$$

Entonces, si consideramos las funciones

$$\phi_i(y_1, \ldots, y_{r-1}) := \sum_{j=1}^{r-1} \left(1 - \frac{c_i}{c_j}\right) y_j$$

para $0 \leq i \leq r-1$ y escribimos

$$x = y_1 + \ldots + y_r, \tag{7.6}$$

obtenemos las igualdades $\phi_i(y_1, \ldots, y_{r-1}) = x + c_i h$ para todo i, donde además la función ϕ_i no depende de la variable y_i. Notar que (7.5) y (7.6) definen un mapa $\Phi : \mathbb{Z}_N^{r-1} \to \mathbb{Z}_N^2$ dado por $\Phi(y_1, \ldots, y_{r-1}) = (x, h)$. Pero puesto que éste mapa es un cubrimiento uniforme, obtenemos que (7.4) es igual a

$$\mathbb{E}_{y\in\mathbb{Z}_N^{r-1}} \prod_{i=1}^{r-1} f_i(\phi_i(y)), \tag{7.7}$$

con $y = (y_1, \ldots, y_{r-1})$. Aprovechando que ϕ_{r-1} no depende de y_{r-1} podemos reescribir esto (multiplicando y dividiendo por $\nu^{1/2}(\phi_{r-1}(y_1, \ldots, y_{r-1}))$) como

$$\mathbb{E}_{y_1,\ldots,y_{r-2}} G_0(y) H_0(y)$$

con

$$G_0(y) := f_{r-1}(\phi_{r-1}(y))\nu^{-1/2}(\phi_{r-1}(y))$$

y

$$H_0(y) := \mathbb{E}_{y_{r-1}} \nu^{1/2}(\phi_{r-1}(y)) \left(\prod_{i=0}^{r-2} f_i(\phi_i(y)) \right).$$

Con esto, usando la desigualdad de Cauchy-Schwarz podemos acotar el cuadrado de (7.7) por

$$\left(\mathbb{E}_{y_1,\dots,y_{r-2}} |G_0(y)|^2 \right) \left(\mathbb{E}_{y_1,\dots,y_{r-2}} |H_0(y)|^2 \right).$$

Ahora bien, el primero de estos factores va a estar acotado por

$$\mathbb{E}_{y_1,\dots,y_{r-2}} \nu(\phi_{r-1}(y)) = 1 + o(1),$$

donde hemos usado como siempre la condición de formas lineales. Concluimos entonces que el cuadrado de (7,7) estará acotado por

$$(1+o(1))\mathbb{E}_{y_1,\dots,y_{r-2},y_{r-1},h_{r-1}} \prod_{\epsilon\in\{0,1\}} \nu^{1/2}(\phi_{r-1}(y_1,\dots,y_{r-1}+\epsilon h_{r-1})) \times \left(\prod_{i=0}^{r-2} f_i(\phi_i(y_1,\dots,y_{r-1}+\epsilon h_{r-1})) \right).$$

Hemos logrado así eliminar el papel jugado por f_{r-1} en una manera que pronto veremos probará satisfactoria a nuestras intenciones. Esto sugiere llevar a cabo un razonamiento análogo para las demás funciones f_i mediante una generalización del argumento anterior. Esta vez, necesitaremos estimar expresiones de la forma

$$\mathbb{E}_{y,h\in\mathbb{Z}_N^{r-1}} \prod_{\substack{\epsilon\in\{0,1\}^{r-1} \\ \epsilon_1=\dots=\epsilon_{r-d-1}=0}} \left(\prod_{i=r-d}^{r-1} \nu^{1/2}(\phi_i(y+\epsilon h)) \right) \left(\prod_{i=0}^{r-d-1} f_i(\phi_i(y+\epsilon h)) \right), \tag{7.8}$$

con $y = (y_1,\dots,y_{r-1})$, $h = (h_1,\dots,h_{r-1})$ y $\epsilon h = (\epsilon_1 h_1,\dots,\epsilon_{r-1}h_{r-1})$ (notar que el rango de ϵ hace redundante el promediar sobre las $r-d-1$ primeras variabes de h, lo cual sin embargo alivia la notación).

Para estimar esto procedemos como antes, utilizando que ϕ_{r-d-1} no depende de la variable y_{r-d-1} para reescribir (7.8) como

$$\mathbb{E}\left(G(y,h)H(y,h)\right) \tag{7.9}$$

con la esperanza yendo sobre las variables $y_1,\dots,y_{r-d-2},y_{r-d},\dots,y_{r-1}$, $h_1,\dots,h_{r-d-2},h_{r-d},\dots,h_{r-1} \in \mathbb{Z}_N$ y las funciones G y H dadas por

$$G(y,h) := \prod_{\substack{\epsilon\in\{0,1\}^{r-1} \\ \epsilon_1=\dots=\epsilon_{r-d-1}=0}} f_{r-d-1}(\phi_{r-d-1}(y+\epsilon h))\nu^{-1/2}(\phi_{r-d-1}(y+\epsilon h))$$

y

$$H(y,h) := \mathbb{E}_{y_{r-d-1},h_{r-d-1}} \prod_{\substack{\epsilon\in\{0,1\}^{r-1}\\ \epsilon_1=\ldots=\epsilon_{r-d-1}=0}} \left(\prod_{i=r-d-1}^{r-1} \nu^{-1/2}(\phi_i(y+\epsilon h))\right) \times \left(\prod_{i=0}^{r-d-2} f_i(\phi_i(y+\epsilon h))\right).$$

Como antes, acotamos el cuadrado de (7.9) utilizando la desigualdad de Cauchy-Schwarz. Ahora bien, usando $f_{r-d-1} \leq \nu$, vemos que es

$$\mathbb{E}\left(|G(y,h)|^2\right) \leq \prod_{\substack{\epsilon\in\{0,1\}^{r-1}\\ \epsilon_1=\ldots=\epsilon_{r-d-1}=0}} \nu(\phi_{r-d-1}(y+\epsilon h)) = 1+o(1),$$

por la condición de formas lineales. Concluimos entonces que (7.8) está acotado por

$$(1+o(1))\mathbb{E}_{y,h\in\mathbb{Z}_N^{r-1}} \prod_{\substack{\epsilon\in\{0,1\}^{r-1}\\ \epsilon_1=\ldots=\epsilon_{r-d-2}=0}} \left(\prod_{i=r-d-1}^{r-1} \nu^{1/2}(\phi_i(y+\epsilon h))\right) \times \left(\prod_{i=0}^{r-d-2} f_i(\phi_i(y+\epsilon h))\right).$$

Notar que aquí el promediar sobre la variable h_{r-d-1} deja de ser redundante. Partiendo entonces de (7.7) e iterando el anterior procedimiento $r-1$ veces (es decir, hasta que la única función que no haya sido reemplazada por ν sea f_0), concluimos de la identidad

$$\phi_0(y+\epsilon h) = x + \sum \epsilon_i h_i$$

que el lado izquierdo de (7.3) está acotado por

$$(1+o(1))\left(\mathbb{E}_{x\in\mathbb{Z}_N}\mathbb{E}_{h\in\mathbb{Z}_N^{r-1}} W(x,h) \prod_{\epsilon\in\{0,1\}^{r-1}} f_0(x+\sum \epsilon_i h_i)\right)^{\frac{1}{2^{r-1}}}$$

donde $W(x,h)$ está dado por

$$\begin{aligned} W(x,h) &= \mathbb{E}_{y_1,\ldots,y_{r-1}} \prod_{\epsilon\in\{0,1\}^{r-1}} \prod_{i=1}^{r-1} \nu^{1/2}(\phi_i(y+\epsilon h)) \\ &= \mathbb{E}_{y_1,\ldots,y_{r-1}} \prod_{i=1}^{r-1} \prod_{\substack{\epsilon\in\{0,1\}^{r-1}\\ \epsilon_i=0}} \nu^{1/2}(\phi_i(y+\epsilon h)) \end{aligned}$$

cpn $y = (y_1, \dots, y_{r-2}, x - y_1 - \dots y_{r-2})$.

Puesto que por la definición de la norma de Gowers es

$$\left(\mathbb{E}_{x \in \mathbb{Z}_N} \mathbb{E}_{h \in \mathbb{Z}_N^{r-1}} \prod_{\epsilon \in \{0,1\}^{r-1}} f_0(x + \sum \epsilon_i h_i) \right)^{\frac{1}{2^{r-1}}} = \|f_0\|_{U^{r-1}},$$

vemos que para probar la Proposición 7.3 bastará con mostrar la estimación

$$\mathbb{E}_{x \in \mathbb{Z}_N} \mathbb{E}_{h \in \mathbb{Z}_N^{r-1}} |W(x,h) - 1| \prod_{\epsilon \in \{0,1\}^{r-1}} f_0(x + \sum \epsilon_i h_i) = o(1).$$

Acotando f_0 por $\nu = \nu^{1/2}\nu^{1/2}$ y aplicando la desigualdad de Cauchy-Schwarz reducimos la tarea a demostrar

$$\mathbb{E}_{x \in \mathbb{Z}_N} \mathbb{E}_{h \in \mathbb{Z}_N^{r-1}} |W(x,h) - 1|^2 \prod_{\epsilon \in \{0,1\}^{r-1}} \nu(x + \sum \epsilon_i h_i) = o(1), \tag{7.10}$$

puesto que el otro factor que obtenemos es igual a $\|\nu\|_{U^{r-1}} = 1 + o(1)$, lo último por la condición de formas lineales.

Expandiendo el cuadrado en (7.10) vemos que bastará probar que es

$$\mathbb{E}_{x \in \mathbb{Z}_N} \mathbb{E}_{h \in \mathbb{Z}_N^{r-1}} W(x,h)^q \prod_{\epsilon \in \{0,1\}^{r-1}} \nu(x + \sum \epsilon_i h_i) = o(1)$$

para $q = 0, 1, 2$. Esto se logra, por supuesto, aplicando la condición de formas lineales. En el caso $q = 0$ usamos la condición de $(2^{r-1}, r, 1)$-formas lineales en las variables $x, h_1, \dots, h_{r-1}$ y las formas $x + \sum \epsilon_i h_i$, $\epsilon \in \{0,1\}^{r-1}$. En el caso $q = 1$ en cambio usamos la condición de $(2^{r-2}(r+1), 2r-2, r)$-formas lineales con las variables $x, h_1, \dots, h_{r-1}, y_1, \dots, y_{r-1}$ y las formas lineales usadas en el caso $q = 0$ más las dadas por $\phi_i(y + \epsilon h)$, $1 \leq i \leq r-1$, con $\epsilon \in \{0,1\}^{r-1}$ y $\epsilon_i = 0$. Finalmente, para $q = 2$, aplicamos la condición de $(r2^{r-1}, 3r-4, r)$-formas lineales en las variables

$$x, h_1, \dots, h_{r-1}, y_1, \dots, y_{r-1}, y'_1, \dots, y'_{r-1}$$

(aquí las variables y_i, y'_i provienen de la expansión de $|W(x,h)|^2$), con las mismas formas lineales del caso $q = 1$ más las dadas por $\phi_i(y' + \epsilon h)$, $1 \leq i \leq r-1$, con $\epsilon \in \{0,1\}^{r-1}$ y $\epsilon_i = 0$ (aquí como antes hemos escrito $y = (y_1, \dots, y_{r-2}, x - y_1 - \dots - y_{r-2})$ y lo análogo para y'). Con esto, la Proposición 7.3 queda demostrada. □

7.5. Demostración de la Proposición 5.3

Pasaremos ahora a aplicar las herramientas desarrolladas para dar demostrar la Proposición 5.3, concluyendo así la demostración del Teorema 1.1.

Demostración de la Proposición 5.3. Sean f, δ como en el enunciado y sea $\varepsilon > 0$ arbitrario. Aplicando el Corolario 7.2 y asumiendo N suficientemente grande, obtenemos una descomposición de la forma $f = g + h$ con $g \leq 1 + o_\varepsilon(1)$, $\mathbb{E}(g) \geq \delta - o_\varepsilon(1)$ y $\|h\|_{U^{r-1}} \leq \varepsilon$. Luego, para estimar

$$\mathbb{E}_{x,h} f(x) f(x+r) \dots f(x+(r-1)h)$$

bastará con estimar las 2^r expresiones de la forma

$$\mathbb{E}_{x,h} f_0(x) f_1(x+r) \dots f_{r-1}(x+(r-1)h),$$

donde para cada $0 \leq i \leq r-1$ puede ser $f_i = g$ o $f_i = h$.

Notamos primero que usando la Proposición 5.2 y la estimación que poseemos sobre la densidad de g, tenemos

$$\mathbb{E}_{x,h} g(x) g(x+r) \dots g(x+(r-1)h) \geq c(r,\delta) - o_\varepsilon(1) - o_{r,\delta}(1),$$

lo cual cubre el caso en que es $f_i = g$ para todo i. Pero en los casos restantes será $f_i = h$ para algún i y en consecuencia, aplicando la Proposición 7.3, obtenemos

$$\begin{aligned} \mathbb{E}_{x,h} f_0(x) f_1(x+r) \dots f_{r-1}(x+(r-1)h) &= O_\varepsilon\left(\|h\|_{U^{r-1}}\right) + o_\varepsilon(1) \\ &= O(\varepsilon) + o_\epsilon(1). \end{aligned}$$

donde hemos usado que es $|h| \leq |f| + |g| \leq \nu + 1$ (absorbiendo el error $o_\varepsilon(1)$ de la cota de g en ν). Concluimos entonces que es

$$\mathbb{E}_{x,h} f(x) f(x+r) \dots f(x+(r-1)h) \geq c(r,\delta) - O(\varepsilon) - o_\varepsilon(1) - o_{r,\delta}(1). \quad (7.11)$$

Pero dado que $\varepsilon > 0$ es arbitrario, se sigue que el lado izquierdo de (7.11) es al menos $c(r,\delta) - o_{r,\delta}(1)$ y el resultado queda así demostrado. □

Capítulo 8

Deducción de los corolarios

Habiendo concluido la demostración del Teorema 1.1 deduciremos en éste capítulo varios corolarios de éste resultado. De aquí en adelante siempre que estemos trabajando en $\mathbb{Z}_N$ con un valor fijo de r, tomaremos $R = N^{r^{-1}2^{-r-4}}$. Comenzamos viendo que del Teorema 1.1 se recupera fácilmente el Teorema de Green-Tao.

Demostración del Corolario 1.1. Fijemos la longitud r de la progresión aritmética que deseamos encontrar. Sea N suficientemente grande y $\varepsilon < 1/2$ una constante positiva arbitraria (en particular, independiente de N). Definimos $f_N : \mathbb{Z}_N \to \mathbb{R}_{\geq 0}$ como

$$f_N(n) = \mathbf{1}_{[\varepsilon N, 2\varepsilon N]} \frac{\vartheta(n)}{4} r^{-1} 2^{-r-4},$$

donde como de costumbre $\vartheta(n)$ denota la función que vale $\log n$ si n es primo y 0 en otro caso. Del teorema del número primo obtenemos la estimación

$$\mathbb{E}(f_N) \geq (1 + o(1)) \frac{\varepsilon}{4} r^{-1} 2^{-r-4}.$$

Tomemos $k = l = 1$ y $\mathcal{H} = \{0\}$. Para N suficientemente grande obtenemos entonces

$$f_N(n) \leq \hat{\Lambda}_R (n; \mathcal{H}, k + l) . \tag{8.1}$$

En efecto, la desigualdad es trivial si n no es primo o no pertenece al intervalo $[\varepsilon N, 2\varepsilon N]$. Si n es primo y está contenido en el mencionado intervalo el único divisor de n menor o igual a R será 1 (puesto que si N es suficientemente grande será $\varepsilon N > R$) y en consecuencia el lado derecho de (8.1) será igual a

$$\left(\frac{1}{2} \log^2 R \right)^2 \frac{1}{\log^3 R} = \frac{1}{4} \log R$$

y la afirmación se sigue de la definición de R. Puesto que el soporte de f_N claramente está contenido en $\mathbb{P}$, fijando δ como cualquier constante menor a

$\varepsilon r^{-1}2^{-r-4}/4$ se sigue que para todo N suficientemente grande f_N satisface las hipótesis del Teorema 1.1. Aplicando entonces tal resultado, el corolario se obtiene inmediatamente. □

Procederemos ahora a demostrar que hay progresiones aritméticas arbitrariamente largas de h-primos de Chen para todo entero h. Necesitaremos el siguiente famoso resultado (ver [7]).

Proposición 8.1 (Teorema de Chen)**.** *Para todo N suficientemente grande la cantidad de h-primos de Chen en el intervalo $[N/2, N)$ es $\geq c_h N/\log^2 N$ para cierta constante absoluta $c_h > 0$. Más aún, podemos obtener la misma cota asumiendo adicionalmente que para cada tal h-primo de Chen p todos los divisores primos de $p+h$ exceden $N^{1/10}$.*

Nota. La cota de $N^{1/10}$ es la demostrada originalmente por Chen y ha sido mejorada. Para una demostración del teorema de Chen con los factores primos excediendo $N^{3/11}$ ver [32].

Demostración del Corolario 1.2. Fijemos el valor de h para el cual deseamos hallar progresiones de h-primos de Chen. Como antes, fijemos también la longitud r de la progresión aritmética a encontrar.

Definimos $f_N : \mathbb{Z}_N \to \mathbb{R}_{\geq 0}$ como

$$f_N(n) = \mathbf{1}_{[N/2,N)} \frac{\vartheta_h(n)}{36} r^{-2} 2^{-2r-8}.$$

donde $\vartheta_h(n)$ es igual a $\log^2 n$ si n es un h-primo de Chen con los divisores primos de $p+h$ excediendo $N^{1/10}$ y es igual a 0 en otro caso. Aplicando el teorema de Chen obtenemos

$$\mathbb{E}(f_N) \geq (1 + o(1))c'_h,$$

para cierta constante c'_h que depende únicamente de c_h y r.

Tomamos ahora $k = 2$, $l = 1$ y $\mathcal{H} = \{0, h\}$. Obtenemos para N suficientemente grande

$$f_N(n) \leq \hat{\Lambda}_R\left(n; \mathcal{H}, k + l\right). \tag{8.2}$$

En efecto, la desigualdad es trivial si es $n < N/2$ o $\vartheta_h(n) = 0$. En caso contrario, todos los divisores primos de $P(n; \mathcal{H})$ excederán a R (puesto que para N suficientmente grande será mín $\left\{N/2, N^{1/10}\right\} \geq R$) y en consecuencia el lado derecho de (8.2) será igual a

$$\left(\frac{1}{6}\log^3 R\right)^2 \frac{1}{\log^4 R} = \frac{1}{36}\log^2 R$$

de donde la afirmación se sigue por definición de R. Usando entonces (8.2), que f_N tiene soporte en los h-primos de Chen, tomando cualquier constante $0 < \delta < c'_h$ y asumiendo N suficientemente grande, el resultado se obtiene del Teorema 1.1. □

Demostración del Corolario 1.3. A esta altura el razonamiento a seguir es claro. Fijados el valor de r y la tupla $\mathcal{H} = \{h_1, \dots, h_k\}$ en cuestión definimos

$$f_N(n) = \mathbf{1}_{[N/2,N)} \frac{\left(r^{-1}2^{-r-4}\right)^k}{(k+l)!^2} \vartheta(n+h_1)\dots\vartheta(n+h_k).$$

Por hipótesis, sabemos que existe una constante $\delta > 0$ con

$$\mathbb{E}(f_N) \geq \delta$$

para todo N. Al igual que en las otras demostraciones, la desigualdad

$$f_N(n) \leq \hat{\Lambda}_R\left(n; \mathcal{H}, k+1\right),$$

se obtiene notando que si el lado izquierdo no se anula para cierto $n \in [N/2, N)$, entonces todos los factores primos de $P(n; \mathcal{H})$ excederán R. Vemos así que f_N satisface las hipótesis del enunciado del Teorema 1.1 y el corolario queda demostrado. □

Para probar el Corolario 1.4 necesitaremos el siguiente resultado

Proposición 8.2. *Sea $\mathcal{H} = \{h_1, \dots, h_k\} \subseteq \mathbb{Z}$ una tupla admisible. Existe entonces una constante absoluta c tal que, si $\pi_{\mathcal{H},c}^{3k\log k}(N)$ denota la cantidad de valores $n \in [N/2, N)$ para los cuales $\{n+h_1, \dots, n+h_k\}$ acumula a lo sumo $3k\log k$ factores primos distintos, todos los cuales exceden N^c, se tiene la estimación*

$$\pi_{\mathcal{H},c}^{3k\log k}(N) \gg \frac{N}{\log^k N}.$$

Demostración. La demostración de esta proposición con una cota de $(k+1)\log v_k + k$ en la cantidad de factores primos distintos, para un cierto parámetro v_k, está dada en [29, Capítulo 10]. La cota de $3k\log k$ se sigue de la estimación de v_k dada en [28]. □

Demostración del Corolario 1.4. El resultado se sigue utilizando los mismos razonamientos que en la demostración del Corolario 1.2, considerando en éste caso $\hat{\Lambda}_R\left(n; \mathcal{H}, k+3k\log k\right)$ y aplicando la Proposición 8.2. □

En forma simultánea e independiente al presente estudio y basándose en el trabajo de Binbin Zhou [48], János Pintz [36] dio una demostración alternativa de los Corolarios 1.3 y 1.4. Además, adaptando los argumentos de su colaboración con Dan Goldston y Cem Yildirim expuesta en el Capítulo 4, obtiene la siguiente:

Proposición 8.3. *Supongamos que los primos tienen nivel de distribución $\vartheta > 1/2$. Existe entonces una constante $C = C(\vartheta)$ tal que para toda tupla admisible $\mathcal{H}$ con $|\mathcal{H}| = k \geq C$ existen al menos*

$$c_1(\mathcal{H})\frac{N}{\log^k N}$$

enteros $n \in (N/2, N]$ *para los cuales la* k*-tupla* $n + \mathcal{H}$ *contiene al menos dos números primos y no posee factores primos menores que* $N^{c_2(k)}$*. En particular, asumiendo la conjetura de Elliot-Halberstam, lo anterior vale con* $C = 7$*.*

Utilizando éste resultado, Pintz deduce el Corolario 1.5. Una deducción alternativa de tal corolario puede obtenerse inmediatamente aplicando el Teorema 1.1.

Demostración del Corolario 1.5. Supongamos que los primos tienen nivel de distribución ϑ y fijemos una k-tupla $\mathcal{H}$, con $k \geq C$ y la constante C como en la Proposición 8.3. Aplicando el principio de los casilleros a esta proposición concluimos la existencia de enteros distintos $h_i, h_j \in \mathcal{H}$ tales que para al menos

$$\frac{c_1(\mathcal{H})}{\binom{k}{2}} \frac{N}{\log^k N}$$

enteros $n \in (N/2, N]$, $n + h_i$ y $n + h_j$ serán ambos primos y la tupla $n + \mathcal{H}$ no poseerá factores primos menores que $N^{c_2(k)}$.

Es evidente entonces que mediante los razonamientos empleados en la deducción de los anteriores corolarios, la primera afirmación del Corolario 1.5 se deduce del Teorema 1.1 y las observaciones del párrafo anterior.

Asumiendo la conjetura de Elliot-Halberstam, obtenemos el resultado anterior para toda tupla admisible de al menos 7 elementos. Considerando entonces la tupla admisible

$$\mathcal{H} = \{0, 2, 6, 8, 12, 18, 20\}$$

el resultado se sigue, aplicando nuevamente el principio de los casilleros. □

Bibliografía

[1] R. C. Baker, G. Harman y J. Pintz, *The difference between consecutive primes, II*, Proc. London Math. Soc. **83** (2001), 532-562.

[2] F. A. Behrend, *On the sets of integers which contain no three in arithmetic progression*, Proc. Nat. Acad. Sci. U.S.A. **32** (1946), 331-332.

[3] E. Bombieri y H. Davenport, *Small differences between prime numbers*, Proc. Roy. Soc. Ser. A **293** (1966), 1-18.

[4] E. Bombieri, *On the large sieve*, Mathematika **12** (1965), 201-225.

[5] J. Bourgain, *On triples in arithmetic progression*, GAFA **9** (1999), 968-984.

[6] V. Brun, *La série* $1/5+1/7+1/11+1/13+1/17+1/19+1/29+1/31+1/41+1/43+1/59+1/61\ldots$, *les dénominateurs sont nombres premiers jumeaux est convergente où finie*, Bull. Sci. Math. **43** (1919), 124-128.

[7] J.-R. Chen, *On the representation of a large even integer as the sum of a prime and a product of at most two primes*, Sci. Sinica **16** (1973), 157-176.

[8] H. Cramér, *On the order of magnitude of the difference between consecutive prime numbers*, Acta Arith. **2** (1936), 23-46.

[9] P. D. T. A. Elliot y H. Halberstam, *A conjecture in prime number theory*, Symp. Math. **4** (1968-1969), 59-72.

[10] P. Erdös, *On the difference of consecutive primes*, Quart. J. Math. Oxford **6** (1935), 124-128.

[11] P. Erdös, *The difference between consecutive primes*, Duke Math J. **6** (1940), 438-441.

[12] P. Erdös, *On the combinatorial problems which I would most like to see solved*, Combinatorica **1** (1981), 28.

[13] G. R. Freiman, *Foundations of a Structural Theory of Set Addition*, Kazan Gos. Ped. Inst., Kazan, 1966.

[14] J. Friedlander y A. Granville, *Limitations to the equi-distribution of primes I*, Ann. Math. **129** (1989), 363-382.

[15] H. Furstenberg, *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. Analyse Math. **31** (1977), 204-256.

[16] P. X. Gallagher, *On the distribution of primes in short intervals*, Mathematika **23** (1976), 4-9.

[17] D. A. Goldston, Y. Motohashi, J. Pintz y C. Y. Yildirim, *Small gaps between primes exist*, Proc. Japan Acad. Ser. A Math. Sci. **82** (2006), no. 4, 61-65.

[18] D. A. Goldston, J. Pintz y C. Y. Yildirim, *Primes in tuples I*, Ann. Math. **170** (2009), no. 2, 819-962.

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

[20] W. T. Gowers, *Hypergraph regularity and the multidimensional Szemerédi theorem*, Ann. Math. **166** (2007), no. 3, 897-946.

[21] W. T. Gowers, *Decompositions, approximate structure, transference, and the Hahn-Banach theorem*, Bull. London Math. Soc., por aparecer.

[22] W. T. Gowers y J. Wolf, *The true complexity of a system of linear equations*, Proc. London Math. Soc. **100** (2010), 155-176.

[23] B. J. Green y T. Tao, *Restriction theory of the Selberg sieve, with applications*, J. Th. Nombres Bordeaux **18** (2006), 147-182.

[24] B. J. Green y T. Tao, *The primes contain arbitrarily long arithmetic progressions*, Ann. Math. **167** (2008), no. 2, 481-547.

[25] B. J. Green y T. Tao, *An inverse theorem for the Gowers $U^3(G)$ norm*, Proc. Edinburgh Math. Soc. **51** (2008), no. 1, 73-153.

[26] B. J. Green, T. Tao y T. Ziegler, *An inverse theorem for the Gowers U^4 norm*, preprint.

[27] B. J. Green y T. Tao, *Linear equations in primes*, Ann. Math., por aparecer.

[28] H. W. Hagedorn, *Sieve methods and polynomial sequences*, Acta Arith. **28** (1975), 245-252.

[29] H. H. Halberstam y H.-E. Richert, *Sieve methods*, New York, London: Academic Press, 1974.

[30] G. H. Hardy y J. E. Littlewood, *Some problems of "Partitio Numerorum"; III: On the expression of a number as a sum of primes*, Acta Math. **44** (1923), 1-70.

[31] M. Huxley, *Small differences between consecutive primes. II.*, Mathematika **24** (1977), 142-152.

[32] H. Iwaniec, *Sieve methods*, graduate course, Rutgers, 1996.

[33] H. Iwaniec y E. Kowalski, *Analytic Number Theory*, Providence, RI: American Mathematical Society, 2004.

[34] H. Maier, *Small differences between prime numbers*, Michigan Math. J. **35** (1988), 323-344.

[35] H. L. Montgomery, *Problems concerning prime numbers*, Proc. Symp. Pure Math. **28** (1976), 307-310.

[36] J. Pintz, *Are there arbitrarily long arithmetic progressions in the sequence of twin primes?*, preprint.

[37] R. Rankin, *The difference between consecutive primes*, J. London Math. Soc. **13** (1938), 242-244.

[38] O. Reingold, L. Trevisan, M. Tulsiani y S. Vadhan, *Dense subsets of pseudorandom sets*, FOCS (2008), 76-85.

[39] V. Rödl y J. Skokan, *Regularity lemma for k-uniform hypergraphs*, Random Structures and Algorithms **28** (2006), no. 2, 180-194.

[40] I. Z. Ruzsa, *Generalized arithmetic progressions and sumsets*, Acta Math. Hungar. **65** (1994), 379-388.

[41] K. Soundararajan, *Small gaps between prime numbers: The work of Goldston-Pintz-Yildirim*, Bull. Amer. Math. Soc. **44** (2007), 1-18.

[42] E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. **27** (1975), 299-345.

[43] T. Tao, *A quantitative ergodic theory proof of Szemerédi's theorem*, Electron. J. Combin. **13** (2006), no. 1, Research Paper 99, 49 pp. (electronic).

[44] E. C. Titchmarsh, *The theory of the Riemann zeta function*, Oxford University Press, 1ra ed, 1951.

[45] E. Westzynthius, *Über die Verteilung der Zahlen, die zu der n ersten Primzhalen teilerfremd sind*, Comm. Phys. Math. Helsingfors **25** (1931), 1-37.

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

[47] S. Wigert, *Sur l'ordre de grandeur du nombre diviseurs d'un entier*, Ark. Math. **3** (1906-1907), no. 18, 1-9.

[48] B. Zhou, *The Chen primes contain arbitrarily long arithmetic progressions*, Acta Arith. **138** (2009), 301-315.

Índice alfabético

L^q, 6
$P(n;\mathcal{H})$, 13
R_P, 84
U^2, 70
U^d, 72
W, 26
Δ, 31
Λ, 11
$\Lambda(n;\mathcal{H},k+l)$, 13
Λ_R, 12
$\Lambda_R(n;\mathcal{H},k+l)$, 14
Ω, 10
δ-denso, 15
λ_R, 21
λ_p, 33
$\log^+$, 23
$\mathcal{H}$, 13
$\mathfrak{G}(\mathcal{H})$, 10
μ, 11
ν_c, 59
ω_X, 28
π, 9
$\rho(C_1,C_2,\delta,J)$, 84
τ_r, 22
φ_k, 26
ϑ, 11
$\vartheta(y;a,q)$, 45
$\zeta(s)$, 24
$w(N)$, 26

casi-primos, 13
condición de correlación, 58
condición de formas lineales, 57
conjetura de Elliot-Halberstam, 46
conjetura de Erdös, 55
conjetura de Goldbach, 9
conjetura de Hardy-Littlewood, 10, 11
conjetura de Montgomery, 46
conjunto aleatorio, 71
conjunto casi-aleatorio, 71
conjuntos uniformes de grado $r-1$, 74
correlación cuadrática, 76
cubrimiento uniforme, 7

densidad positiva, 7
desigualdad de Gowers-Cauchy-Schwarz, 72

ejemplo de Behrend, 56
esperanza, 6
esperanza condicional, 6

fórmula de inversión de Fourier, 7
función básica anti-uniforme, 77, 83
función de Euler generalizada, 26
función de Möbius, 11
función de von Mangoldt, 11
función indicatriz, 6
función zeta de Riemann, 24

grupo dual, 6

h-primo de Chen, 16
hipótesis generalizada de Riemann, 45

lema de Gallagher, 52

medida, 57
medidas pseudoaleatorias, 58
modelo de Cramér, 9

nivel de distribución, 44

norma U^2 de Gowers, 70
norma U^d de Gowers, 72
norma algebraica, 77
norma de correlación básica anti-uniforme (CBA), 83
norma predual cuasi algebraica (PCA), 77

postulado de Bertrand, 5
primos de Chen, 16
primos de Sophie Germain, 9
primos gemelos, 17
problema inverso de la norma de Gowers, 76
producto interno de Gowers, 71

región libre de ceros clásica, 37

serie singular, 10
soporte, 16

teorema de Bombieri-Vinogradov, 45
teorema de Brun, 56
teorema de Chen, 94
teorema de Freiman, 76
teorema de Goldston-Pintz-Yildirim, 18
teorema de Green-Tao, 16
teorema de Hahn-Banach, 82
teorema de Roth, 56
teorema de Szemerédi, 55
teorema de Szemerédi, versión ergódica, 56
teorema de transferencia, 85
teorema del número primo, 9, 11, 35
teorema directo, 75
teorema generalizado de von Neumann, 87
teorema inverso, 75
transformada de Fourier, 6
truco-W, 26, 59
tupla admisible, 11
tupla uniforme, 64
tuplas de casi primos, 44

variables de Bernoulli, 9