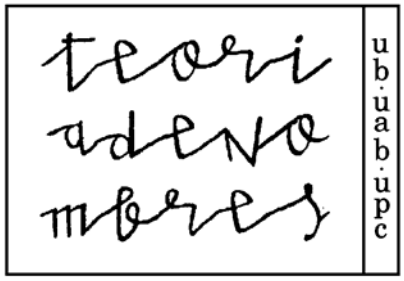
NOTES DEL SEMINARI DE



19è ANY

MÈTODES DE GARBELL

Barcelona, 2005

13

Notes del Seminari de Teoria de Nombres (UB-UAB-UPC)

MÈTODES DE GARBELL

Edició a cura de

J. Guàrdia J. Jiménez

Amb contribucions de

J. González L. Dieulefait J. Quer F. Chamizo

T. Crespo X. Xarles J. Jiménez A. Ubis

J. Guàrdia
Escola Politècnica Superior d'Enginyeria de
Vilanova i la Geltrú
Universitat de Politécnica de Catalunya
Av. Víctor Balaguer s/n
08800 Barcelona

J. Jiménez
ETSETB,
Universitat Politècnica de Catalunya,
Jordi Girona 1-3,
08034 Barcelona

Classificació AMS
Primària: 11A05, 11A41, 11N32,11N35,11N36
Secundària: 11L

Barcelona, 2005

Amb suport parcial de BFM-2003-06768-C02-01, BFM-2003-06768-C02-02,
2002SGR 00148.

ISBN: 84-934244-2-0

Índex

1 Formulación del problema y Criba de Eratóstenes-Legendre.
Josép Gonzalez **1**

1.1 La función $\pi(x)$. 1

1.2 El problema general de criba y el Teorema de Eratóstenes-Legendre 5

1.3 Dimensión de criba 10

1.4 Composición de cribas 12

1.5 Transparencias. 14

2 Criba de Selberg; Una cota superior
Luis Dieulefait. **29**

2.1 La función $\lambda^+(d)$. 29

2.2 El término de error. 32

2.3 Aplicaciones. 35

2.4 Transparencias. 38

3 Garbell de Brun i aplicacions.
Jordi Quer **49**

3.1 Garbells: definicions i notació 49

3.2 Garbells combinatoris 52

3.2.1 Garbell pur de Brun 53
3.2.2 Garbell de Brun-Hooley 55
3.2.3 Garbell de Rosser-Iwaniec 57
3.3 Fórmules de Buchstab 58
3.4 Estimacions amb garbells combinatoris i aplicacions . 61

4 El garbell combinatori.
Fernando Chamizo. **67**

4.1 Cribas combinatorias 67
4.2 La criba de Brun 68
4.3 Las iteraciones de Buchstab 69
4.4 La criba de Rosser 70
4.5 El teorema de criba 73
4.6 El caso lineal. Ejemplos 75
4.7 Transparencias. 78

5 Gran garbell.
Teresa Crespo. **89**

5.1 La desigualtat de grans garbells 89
5.2 La desigualtat de grans garbells per a caràcters additius 90
5.3 Equidistribució entre classes residuals 91
5.4 Generalitzacions 92
5.4.1 Varietats afins 93
5.4.2 Varietats projectives 94
5.5 La desigualtat de grans garbells per a caràcters multiplicatius 94

6 El teorema de Chen
Xavier Xarles **97**

6.1 La conjectura de Goldbach (1742) 97
6.2 Enunciat del teorema 98
6.3 Inici del garbell 100
6.4 La serie singular 102
6.5 El teorema de Bombieri-Vinogradov i el terme d'error 103
6.6 Densitat del garbell 105
6.7 Fitació inferior de $S(A, P(z))$ 105
6.8 L'estratègia de Chen 107
6.9 Fitació superior de $\sum_{z\leq q<y\,,q\in P_N} S(A_q, P(z))$ 110
6.10 La mida de B 114
6.11 Fitació superior de $S(B, P(y))$ 116
6.12 El pas final 117

7 Término de error en la criba lineal.
Jorge Jiménez. **119**

7.1 Preliminares. 119
7.2 Métodos analíticos. 121
7.3 Forma bilineal del término de error. 122

8 Casiprimos representados por polinomios cuadráticos.
Adrián Ubis **129**

8.1 Introducción 129
8.2 Forma bilineal para el término de error 130
8.3 Método de Dispersión 132
8.4 Distribución uniforme 135
8.5 Pesos de Richert mejorados 138

8.6 Notas 142

Introducción.

Estas notas contienen las conferencias sobre métodos de Criba presentadas en la decimonovena edición del Seminari de Teoria de Nombres (UB-UAB-UPC), celebrado del 24 al 27 de Enero de 2005 en Barcelona, en la Facultat de Náutica de la Universitat Politècnica de Catalunya.

El programa general fue elaborado por Jordi Guàrdia i Jorge Jiménez y las sesiones se llevaron a cabo por diferentes personas del seminario, gracias a las cuales tenemos las notas que presentamos aquí.

Resumen de Contenidos.

Un problema de criba consiste en localizar números primos dentro de un conjunto de enteros, finito y fijo, con ciertas propiedades aritméticas muy simples. Por ejemplo, si nuestro conjunto A consiste en los enteros positivos hasta un determinado X, usaremos los métodos de criba para establecer una fórmula, lo más precisa posible, para la cantidad de números primos hasta X. Tal fórmula debe cumplir dos objetivos, por un lado demostrar que existen números primos dentro del conjunto, y por otro determinar cuántos, y entender cómo varía esta cantidad a medida que hacemos X tender a infinito. Escogiendo conjuntos A más generales se pueden tratar problemas como la conjetura de Goldbach, la existencia de infinitos primos gemelos, la cantidad de primos de la forma n^2+1 y en general la representación de primos mediante polinomios, la existencia de primos en intervalos de pequeña longitud o en progresiones aritméticas, etc. En resumen, los métodos de criba permiten analizar cualquier conjunto de naturaleza aritmética relacionado con el estudio de los números primos.

La idea es sencilla: utilizar las propiedades aritméticas básicas del conjunto para cribar de éste aquellos elementos que son múltiplos de primos pequeños, digamos menores que z. Los elementos del conjunto que superan tal criba no pueden tener muchos factores primos, (idealmente uno), por motivos obvios de tamaño, en nuestro ejemplo relacionado con las cantidades z y X. Dependiendo de la aritmética del conjunto A se ha de escoger la estructura del conjunto criba P, formado por primos pequeños, (y determinado por la cantidad z). Dado el conjunto P, el proceso de criba es inicialmente trivial, ya que se sustenta en el principio de inclusión exclusión, de elementos en A con factores en P. Este principio se recoge perfectamente a través de la definición de la función aritmetica $\mu(n)$ de Möbius. Los métodos de criba consisten en aproximar en media $\mu(n)$ mediante funciones continuas. Dependiendo del proceso que relaciona una con las otras, aritmética y análisis, se consiguen diferentes métodos de criba.

Este tipo de herramientas, a pesar de remontarse al método de factorización de Eratóstenes, comienza a tener interés después de los resultados de V. Brun alrededor de 1920, ([B] entre otros), o su aplicación posterior en [Sc] en 1933. Sin embargo la estructura del razonamiento era difícil de entender y de utilizar en problemas prácticos. Más tarde Selberg obtiene en 1947 una relación entre aritmética y análisis mucho mas directa y simple,lo que permitió su posterior uso en conjuntos muy generales de forma casi continuada. Pero no es hasta 1970 aproximadamente en que una serie de artículos de H. Iwaniec revolucionaría los métodos de criba encaminándolos hacia su potencial actual. Concretamente, antes de la aparición de estos resultados el número de factores primos de los elementos que superaban la criba era indeseablemente alto. Iwaniec, con sus mejoras tanto en el propio método, como en su posterior aplicación, consigue demostrar, para conjuntos A de gran generalidad, la existencia de "muchos"elementos casi primos, es decir, con como mucho dos factores primos. Es muy importante señalar que, según un análisis anterior de Selberg, este tipo de resultados eran la cota óptima que se podía esperar utilizando métodos de criba.

El seminario consiste en desarrollar, comenzando desde cero, los métodos de criba necesarios para demostrar uno de los principales re-

sultados de H. Iwaniec sobre casiprimos representados por polinomios cuadráticos, en concreto en la dirección de la famosa conjetura que asegura la existencia de infinitos primos de la forma $n^2 + 1$. En la primera charla se establecerán todos los parámetros necesarios para entender el problema de criba de forma general. En los casos particulares más importantes veremos el valor concreto de tales parámetros. La segunda charla se centra en el método de criba de Selberg mucho más directo y sencillo de establecer. En sus versiones más simples ya es suficientemente poderoso para obtener algunas aplicaciones interesantes que también se mostrarán en la charla. En la tercera charla se desarrolla el método original de Brun que dio paso a lo que hoy se conoce como criba combinatoria. A pesar de su naturaleza enrevesada, este método se pudo generalizar con las ideas de Rosser e Iwaniec, independientemente, que tambíen se verán en esta charla. Se concluye con algunas de sus aplicaciones para ver su funcionamiento. En la cuarta charla se hará una descripción explícita de las funciones continuas que "aproximan"el principio de inclusión y exclusión, junto con una fórmula para el error que se comete. Es importante observar que el análisis de estas funciones debe servir para determinar si efectivamente existen primos en el conjunto, (la función será positiva), y cómo crecen cuando el conjunto se hace infinito. En la quinta charla introducimos una importantísima mejora iniciada por Kuhn, [K], y que ahora se engloba en las que se conocen como cribas pesadas. Este método permitió demostrar a Chen [Ch] un forma débil de la conjetura de Goldbach usando casiprimos en vez de restringirse al conjunto de los números primos. Este será el contenido de la sexta charla. En la séptima charla, de contenido puramente técnico, se introducen las mejoras que el mismo Iwaniec obtuvo en relación con el término de error que se comente al aproximar por funciones continuas mediante el método de criba de Rosser-Iwaniec. La octava y última charla se dedica a presentar el artículo original de Iwaniec, [I5]. Para ello, después de las mejoras en el término de error el artículo utiliza la mejor criba pesada conocida hasta el momento, así como un análisis concreto del conjunto $n^2 + 1$ correspondiente al problema.

Se hace necesaria una explicación al hecho de que todos los problemas mencionados hablan de demostrar la existencia de primos en determinados conjuntos y, sin embargo, hemos comentado que el límite de los métodos de criba esta en localizar elementos casiprimos. De

forma relativamente reciente el mismo Iwaniec, en colaboración con J. Friedlander en [F-I1], han desarrollado un poderoso método de criba que supera las dificultades señaladas por Selberg y consiguen aplicarlo de manera exitosa a uno de los problemas más difíciles dentro de la teoría de números primos como es la existencia de infinitos primos en conjuntos de pequeña densidad. En el caso de [F-I2] tratan los primos representados mediante $x^2 + y^4$. En nuestra opinión, las ideas necesarias para presentar estos dos artículos, comenzando desde cero, superan el tiempo asignado al seminario. Sin embargo, el seminario nos puede dejar a cada uno de los participantes en una posición bastante ventajosa para un posterior análisis de estos métodos punteros.

Jorge Jiménez

Barcelona, 13 de Mayo de 2005.

Capítol 1

Formulación del problema y Criba de Eratóstenes-Legendre.

Josép Gonzalez

Tanto la charla en el seminario, como las transparencias recogidas al final del capítulo fueron realizadas por J. Gonzalez. El texto sólo intenta reconstruir dicha charla. Cualquier error o ambiguedad es responsabilidad única del transcriptor, J. Jiménez.

1.1 La función $\pi(x)$.

Cualquier proceso de criba, en el contexto de la Aritmética, está diseñado para atacar el siguiente problema.

1.1.1 Problema. *Dado un conjunto finito A de enteros positivos, estimar la cantidad de números primos que contiene.*

Existen dos preguntas por excelencia que han llevado a los métodos de criba a desarrollarse hasta los niveles que se encuentran en la actualidad. Estos son

Conjetura de Goldbach, (G). Todo entero positivo par se puede escribir como suma de dos primos.

Primos Gemelos, (PG). Existen infinitos p primos tal que $p+2$ es también un número primo.

La relación de estos problemas con el Problema 1.1.1 estará clara más adelante. A ellos se dedican los primeros artículos de Vigo Brun, máximo percusor del desarrollo de los métodos de criba en la epoca actual. A pesar de ello no es en estos trabajos donde aparece por primera vez algún proceso reconocible como método de criba, sino que es mucho tiempo atras. Típicamente el primer método de criba se atribuye a Eratóstenes de Cirene y ha llegado a nuestros días gracias a la "Introducción a la Aritmética"de Nicomedes escrito alrededor del siglo III a.c. En este caso el conjunto a tratar es el de todos los enteros positivos hasta uno dado N. Curiosamente no está claro que dicho método en su origen pretendiese tanto cuantificar los números primos hasta N, como dar una tabla de los factores primos por los que eran divisibles cada positivo menor que N. En este sentido Eratóstenes presentó una tabla de números como la siguiente

1	~~2~~	~~3~~	~~4~~	~~5~~	~~6~~
7	~~8~~	~~9~~	~~10~~	11	~~12~~
13	~~14~~	~~15~~	~~16~~	17	~~18~~
19	~~20~~	~~21~~	~~22~~	23	~~24~~
~~25~~	~~26~~	~~27~~	~~28~~	29	~~30~~
31	~~32~~	~~33~~	~~34~~	~~35~~	~~36~~

tachando cada número tantas veces como divisores primos tenía. Para transformar la tabla anterior en un método para contar primos, es necesario todavía observar que cualquier positivo menor que N sin divisores primos menores que $\sqrt{N}$ debe ser primo. Dicho de otra forma si $P(\sqrt{N}) = \prod_{p\leq\sqrt{N}} p$, entonces $n \in [\sqrt{N}, N]$ es primo si y sólo si $(n, P(\sqrt{N})) = 1$. Una vez tenemos este ingrediente adicional podemos contar los primos hasta un X dado de la siguiente forma. Sea $\pi(X)$ el número de primos que hay en el intervalo $A = [1, X]$. Entonces, por la observación anterior, $\pi(X) = \pi(\sqrt{X}) + S(A, P(\sqrt{X})$ donde $S(A, P(\sqrt{X})$ es el número de enteros que permanecen sin

tachar en el método de Eratóstenes. Es fácil calcular $S(A, P(\sqrt{X})$ gracias al principio de inclusión exclusión. Concretamente se puede ver que

$$S(A, P(\sqrt{X}) = [X] - \sum_{p \leq \sqrt{X}} \left[\frac{X}{p}\right] + \sum_{p,q \leq \sqrt{X}} \left[\frac{X}{pq}\right] - \dots,$$

es decir, al número total de enteros debemos quitarles los múltiplos de los primos hasta $\sqrt{X}$ y, teniendo en cuenta que los que son múltiplos de dos primos en esas condiciones se restan dos veces, deberán ser sumados otra y asi sucesivamente. En términos de la función de Möbius, $\mu(n)$, función multiplicativa que vale -1 en los números primos y cero en los que no son libres de cuadrados, podemos escribir el principio de inclusión-excluisión como

$$\sum_{d|n} \mu(d) = \begin{cases} 1 & \text{si } n = 1 \\ 0 & \text{si } n \neq 1, \end{cases}$$

con lo que[1]

$$S(A, P) = \sum_{a \in A} \sum_{d|(a,P)} \mu(d). \tag{1.1}$$

Normalmente a $S(A, P)$ se le llama la función de criba asociada a la pareja A, P. Cambiando el orden de sumación obtenemos

$$\sum_{d|P} \mu(d) \sum_{a=dk, a \in A} 1 = \sum_{d|P} \mu(d)|A_d|,$$

donde $A_d = \{a \in A : a \equiv 0 \,(\text{mod } d)\}$. Ahora bien $|A_d| = \left[\frac{X}{d}\right] = \frac{X}{d} - \left\{\frac{X}{d}\right\}$, y por tanto

$$S(A, P) = X \sum_{d|P} \frac{\mu(d)}{d} - \sum_{d|P} \mu(d) \left\{\frac{X}{d}\right\} = X \prod_{p|P} \left(1 - \frac{1}{p}\right) + R(A, P).$$

con un término de error $R(A, P) = O\left(2^{\nu(P)}\right)$ donde $\nu(d)$ cuenta el número de factores primos distintos de d.

[1]Por simplicidad, a partir de ahora, y siempre que no haya confusión, denotaremos $P = P(z)$. En este caso $z = \sqrt{X}$

Desafortunadamente la fórmula anterior no es de ninguna utilidad a la hora de acotar la función $\pi(X)$ ya que el término de error del orden de $2^{\pi(\sqrt{X})}$ es mucho mas grande que el término principal. El problema aparece, en primer lugar, del hecho de que se consideran excesivos términos en $R(A,P)$. El siguiente argumento permite disminuir el número de sumandos de la siguiente forma. Sea $P = P(z)$ para algun $0 < z < x^{1/2}$. En este caso $\pi(X) \leq \pi(z) + S(A,P)$, ya que $S(A,P)$ no solo cuenta números primos, sino que también aquellos con "pocos"factores primos (pues todos sus factores primos son mayores que z). Haciendo el mismo razonamiento anterior obtenemos

$$S(A,P) \leq X \prod_{p<z} \left(1 - \frac{1}{p}\right) + O\left(2^z\right). \tag{1.2}$$

Teniendo en cuenta la fórmula de Mertens

$$\prod_{p<z} \left(1 - \frac{1}{p}\right)^{-1} \sim e^{\gamma} \log z + O(1),$$

con γ la constante de Euler, y escogiendo $z = \log X$, obtenemos

$$\pi(X) \leq \pi(z) + S(A,P) \leq C \frac{X}{\log\log X},$$

para alguna constante C. Es interesante observar que, en la ecuación 1.2, $\prod_{p<z}\left(1 - \frac{1}{p}\right)$ es exactamente la probabilidad de que un entero sea primo con P.

A priori el trato que hemos dado al término de error, $R(A,P)$, estimado a fuerza bruta, pudiera ser que indujese cierta pérdida en la cota final. De hecho, en principio podría ocurrir que la cancelación que existe en la suma hiciese este término despreciable, permitiendo asi una elección de z mucho mayor, con lo que obtendríamos cotas mucho mejores para la función $\pi(X)$. Este no es el caso. Efectivamente supongamos que $R(A,P)$ es despreciable, y escojamos $z = X^c$ para cualquier $1/2 \leq c < 1$. Entonces $\pi(z) < z = o(X/\log X)$ y

$$\pi(X) = \pi(z) + S(A,P) \sim \prod_{p<X^c} \left(1 - \frac{1}{p}\right) \sim e^{-\gamma} X/c \log X,$$

lo cual es imposible que sea cierto para diferentes valores de c a la vez.

1.2 El problema general de criba y el Teorema de Eratóstenes-Legendre

El argumento anterior para obtener (1.2) es fácilmente generalizable a cualquier conjunto A tal que

$$|A_d| = X\frac{\rho(d)}{d} + r(A,d) \tag{1.3}$$

para todo d y para algúna función multiplicativa $\rho(\cdot)$, donde $r(A,d)$ se entiende como el error que se comete al aproximar A por un conjunto uniformemente distribuido en progresiones aritméticas. Obsérvese que, en este caso, tomando $d = 1$, se tiene $|A| \sim X$. En este sentido, siguiendo los pasos utilizados para demostrar (1.2) en el caso de la función $\pi(X)$ se puede demostrar el siguiente resultado

1.2.1 Teorema. *(Criba de Eratóstenes-Legendre) Sea $z > 0$ y A como en (1.3). Entonces*

$$S(A,P) = XV(P) + \theta\sum_{d|P}|r(A,d)|,$$

para algún $|\theta| \leq 1$ donde $V(P) = \prod_{p<z}\left(1 - \frac{\rho(p)}{p}\right)$.

1.2.2 Ejemplo. *Los siguientes son ejemplos de conjuntos que aparecen de forma recurrente en problemas relacionados con métodos de criba.*

1. $A = (Y - X, Y] \cap \mathbb{Z}$. En este caso $|A_d| = \frac{X}{d} + r(A,d)$, con un término de error $r(A,d) = \left\{\frac{Y-X}{d}\right\} - \left\{\frac{Y}{d}\right\} = O(1)$.
2. Sea $f(x) \in \mathbb{Z}[x]$ y $A = \{f(n) : Y - X < n \leq Y\}$. Entonces

$$|A_d| = \sum_{\substack{1\leq l\leq d\\ f(l)\equiv 0 (\mathrm{mod}\ d)}}\ \sum_{\substack{Y-X<n\leq Y\\ n\equiv l(\mathrm{mod}\ d)}} 1 = \frac{X\rho(d)}{d} + O(\rho(d)),$$

donde $\rho(p)$ es el número de soluciones de $f(n) \equiv 0 \,(\mathrm{mod}\ p)$. Tomando $f(n) = an + b$ con $(a, b) = 1$, entonces $\rho(p) = 0$ si $p|a$, $\rho(p) = 1 si p \nmid a$ por lo que

$$V(P) = \prod_{p<z} \left(1 - \frac{1}{p}\right) \prod_{p|a} \left(1 - \frac{1}{p}\right)^{-1} = \frac{a}{\varphi(a)} \prod_{p<z} \left(1 - \frac{1}{p}\right).$$

Tomando $f(n) = n(n+2)$ es posible estudiar el problema de los primos gemelos, y con $n(N-n)$ la conjetura de Goldbach. Sin embargo existe otra manera, quizá mas natural de atacar estos dos problemas, considerando polinomios evaluados en números primos.

3. Sea $f(x) \in \mathbb{Z}[x]$ y $A = \{f(p) : 1 \leq p \leq X\}$. entonces

$$|A_d| = \sum_{\substack{1 \leq l \leq d \\ f(l) \equiv 0 (\mathrm{mod}\ d)}} \sum_{\substack{1 < p \leq X \\ p \equiv l (\mathrm{mod}\ d)}} 1.$$

Si $(m, d) > 1$, entonces la última suma es distinta de cero si y solo si $m = q|d$ con q primo, luego

$$|A_d| = \sum_{\substack{1 \leq l \leq d, (m,d)=1 \\ f(l) \equiv 0 (\mathrm{mod}\ d)}} \pi(X; d, l) + \sum_{\substack{q|d \\ f(q) \equiv 0 (\mathrm{mod}\ d)}} 1.$$

Por el Teorema de Bombieri-Vinogradov sabemos que

$$\pi(X; d, l) = \frac{X}{\varphi(d) \log X} + E(X; d, l)$$

con

$$\sum_{d \leq X^{1/2} / \log^{\beta} X} \max_{(l,d)=1} |E(X; d, l)| \leq \frac{CX}{\log^{\alpha} X}$$

para alguna constante C, X suficientemente grande, α cualquier número positivo y β dependiendo de α. Por tanto queda

$$|A_d| = \frac{X}{\log X} \frac{\rho(d)}{d} + O\left(\frac{X}{\log^{\alpha} X}\right),$$

donde $\rho(d) = \frac{d\psi(d)}{\varphi(d)}$ es una función multiplicativa con $\psi(d) = \{1 \leq l \leq d : f(l) \equiv 0 \,(\mathrm{mod}\ d)\,, (l,d) = 1\}$.

Tomando $f(p) = p+2$ se trata (**PG**), mientras que $f(p) = N-p$ sirve para estudiar (**G**). Es interesante observar también que un estudio análogo al ejemplo anterior nos permitiría estudiar el subconjunto de A definido como $A_{l,k} = \{f(p) \in A : p \equiv l \,(\mathrm{mod}\ k)\}$.

Los métodos de criba pretenden dar herramientas que permitan mejorar el Teorema 1.2.1 en un contexto lo mas general posible. En este sentido es importante recordar que el principal obstáculo a la hora de obtener mejores cotas para $S(A,P)$ se debe al enorme número de términos que hay que considerar en $R(A,P)$. La principal observación detras de cualquier método de criba esta en el hecho de que para dar desigualdades como en el Teorema 1.2.1 no necesitamos la igualdad en 1.1 sino que sería suficiente con dar cotas, superiores e inferiores, de la función de criba. Por otra parte, $S(A,P)$ en realidad cuenta los coprimos con P. En general puede ser conveniente no restringirse a $P = P(z)$, sino escoger $P = \prod_{p\in\mathcal{P}} p$ para algún conjunto de primos $\mathcal{P}$. Asi pues, supongamos que existen una pareja de funciones λ^+, λ^- que cumplen, para todo $m|P$

$$\sum_{d|m} \lambda^-(d) \leq \sum_{d|m} \mu(d) \leq \sum_{d|m} \lambda^+(d), \tag{1.4}$$

y tal que $\lambda^\pm(d) = 0$ para todo $d \geq D$. Entonces, tomando $m = (a,P)$ y sumando en $a \in A$, se obtiene

$$\sum_{d|m} \lambda^-(d)|A_d| \leq S(A,P) \leq \sum_{d|m} \lambda^-(d)|A_d|,$$

y utilizando (1.3) es fácil ver que obtenemos el siguiente teorema

1.2.3 Teorema. *Sea $z > 0$ y A como en (1.3). Entonces*

$$XV^-(P) + R^-(P) \leq S(A,P) \leq XV^+(P) + R^+(P),$$

donde $V^\pm(P) = \sum_{d|P} \frac{\lambda^\pm(d)\rho(d)}{d}$, y $R^\pm(P) = \sum_{d|P} \lambda^\pm(d) r(A,d)$.

A las funciones $\lambda^{\pm}$ se les denomina funciones de criba, superior e inferior respectivamente, de nivel D.

El primer ejemplo sencillo en el que se pueden obtener funciones superiores e inferiores de criba válidas es lo que se denomina **Criba Local**. En este caso hacemos la hipótesis adicional $a \leq X$ para todo $a \in A$ lo que nos permite dar un nivel de criba $D = X$ de forma trivial. El siguiente resultado permite ver una aplicación de este caso particular.

1.2.4 Teorema. *Dado un entero cualquiera m sea $\varphi(X, m)$ el número de enteros menores que X y primos con m. Si m tiene como mucho $X^{1/6 \log\log X}$ factores primos distintos, entonces*

$$\varphi(X, m) = X\frac{\varphi(m)}{m}\left(1 + O(1/(\log X)^2)\right).$$

Demostración. Escoger $\mathcal{P} = \{p|m\}$, $A = [1, X) \cap \mathbb{Z}$, con lo que $\varphi(X, m) = S(A, P)$. En este caso estamos en el Ejemplo 1 de 1.2.2 y por tanto $\rho(d) = 1$ para todo $d|P$, y $|r(A, d)| \leq 1$. Es inmediato ver que las funciones $\lambda^{\pm} = \mu(d)$ si $d \leq X$ y 0 en el resto son funciones de criba válidas. Asi pues

$$V^{\pm}(P) = \sum_{d|P, d<X} \frac{\mu(d)}{d} = V(P) - \sum_{d|P, d\geq X} \frac{\mu(d)}{d}, \text{ y } R^{\pm}(P) \leq \sum_{d|P, d<X} 1,$$

y por tanto, por el Teorema 1.2.3

$$\begin{aligned}\left|\varphi(X, m) - X\frac{\varphi(m)}{m}\right| &\leq X\left|\textstyle\sum_{d|P, d\geq X} \frac{\mu(d)}{d}\right| + \textstyle\sum_{d|P, d<X} 1 \leq \\ &\leq X^{1-\varepsilon} \textstyle\sum_{d|P} \frac{1}{d^{1-\varepsilon}} \leq X^{1-\varepsilon} \prod_{p|P}\left(1 + \frac{1}{p^{1-\varepsilon}}\right).\end{aligned}$$

Tomando z tal que $\pi(z) > |\mathcal{P}|$, y $\varepsilon = 1/\log z$ se tiene

$$\prod_{p|P}\left(1 + \frac{1}{p^{1-\varepsilon}}\right) \leq \prod_{p<z}\left(1 + \frac{e}{p}\right) \leq \exp\sum_{p<z}\frac{e}{p} = O\left((\log z)^e\right).$$

Basta ahora tomar $z = X^{1/5 \log\log X}$ para obtener

$$\left|\varphi(X, m) - X\frac{\varphi(m)}{m}\right| = O(\left(\frac{X}{(\log X)^{5-e}}\right).$$

El resultado es ahora inmediato ya que $m/\varphi(m) \ll \log X$ por la fórmula de Mertens.

En la anterior demostración el papel de la variable ε es determinante. Concretamente ésta ha sido elegida de forma que se optimize el compromiso entre el tamaño de los divisores, dado por el nivel D, y el número de divisores a contar en la suma, de forma que $D^{\varepsilon} = C$, constante. Ahora bien, dicho número de divisores se controla a través de z por lo que la elección $\varepsilon = 1/\log z$ es la adecuada. De esta forma se introduce una nueva variable $s = \frac{\log D}{\log z}$ que será de extrema importancia en el desarrollo posterior de la teoría. De forma que si queremos aumentar z, para asi coseguir que $S(A, P)$ cuente sólo primos o casi primos, deberemos aumentar el nivel D para mantener su relación, s, constante. Obsérvese que, de manera trivial, se debe tener $s > 1$.

Dar pues un método de criba consiste en encontrar funciones superiores e inferiores de criba de forma que para conjuntos lo mas generales posibles el Teorema 1.2.3 facilite acotaciones de $S(A, P)$ no triviales, es decir, V^+ sea lo menor posible, V^- lo mayor posible, y podamos estimar $R^{\pm}$ de forma que efectivamente sea un término despreciable para X grandes. Teniendo en cuenta que se espera que $XV(P)$ sea el término principal de la ecuación, (probabilidad por tamaño del conjunto), podemos enunciar el problema anterior de la forma siguiente en términos de la nueva variable s.

1.2.5 Nota. *Un método de criba consiste en encontrar funciones de criba $\lambda^{\pm}$ tal que para conjuntos A lo mas generales posible existan funciones $f(s) = f_{\lambda}(s)$ y $F(s) = F_{\lambda}(s)$, independientes de A, tal que $V^+(P) \leq F(s)V(P)$ y $V^-(P) \geq f(s)V(P)$ produzcan acotaciones no triviales de la función $S(A, P)$ en el sentido anterior.*

Obsérvese que en particular para dar un resultado no trivial se debe conseguir $f(s) > 0$. Por otro lado tanto $V^{\pm}$ como V dependen directamente del conjunto A a través de la función $\rho(d)$. Por tanto, para encontrar un método de criba en el sentido anterior será preciso hacer ciertas restricciones sobre la función multiplicativa que define

el conjunto.

1.3 Dimensión de criba

Es claro que nuestro estudio de la función $\rho(d)$ debe hacerse a través del producto $V(P)$. Ahora bien, supongamos por un momento que $\rho(p) = \kappa$ constante para todo p. Entonces, por la fórmula de Mertens es fácil ver que

$$\prod_{w\le p<z}\left(1-\frac{\rho(p)}{p}\right)^{-1} \sim \left(\frac{\log z}{\log w}\right)^{\kappa}\left(1+\frac{K}{\log w}\right), \tag{1.5}$$

para todo $2 \le w < z$ y para alguna $K > 1$ constante. Asi pues uno podría establecer un método de criba para cualquier conjunto A con $\rho(p) = \kappa$ sin mas que asumir la constante κ en la definición de cada una de las funciones $f(s)$, $F(s)$. Ahora bien, si tenemos estimaciones de criba válidas, con $f(s) > 0$, para un conjunto $\mathcal{P}$, (es decir, que aseguran la existencia de cierta cantidad de casiprimos en el conjunto A, en la línea que exigía el Problema 1.1.1), entonces trivialmente serán ciertas al cribar por un conjunto $\mathcal{P}'$ menor. Por otro lado la función $\rho(d)$ será mas pequeña a medida que decrece el conjunto $\mathcal{P}$. Por tanto, nuestro método de criba sera efectivo cualquiera que sea el conjunto A con $\rho(p) \le \kappa$ para todo p. Estas consideraciones nos llevan a hacer la siguiente definición de dimensión, o densidad, de criba.

1.3.1 Definición. *Se dice que κ es una densidad de criba para cierta función ρ, o que un conjunto A tiene dimensión de criba κ, si existe una constante $K > 1$ tal que*

$$\frac{V(P(w))}{V(P(z))} \le \left(\frac{\log z}{\log w}\right)^{\kappa} K, \quad \textit{si } 2 \le w \le z. \tag{1.6}$$

Se dice que κ es fuerte si la desigualdad es cierta con

$$K = K_w = 1 + \frac{L}{\log w} \tag{1.7}$$

para alguna $L > 1$.

En particular si $\rho(p) \leq \kappa$ para todo p, entonces κ es una densidad furerte para ρ.

Es interesante observar que en el caso en que A tenga dimensión κ entonces tomando $w = p$, $z = p + \varepsilon$ se deduce $\rho(p) \leq (1 - 1/K)p$ $(= \frac{L}{\log w + L} p)$. Asi pues, si K, o L, son muy grandes, no habrá un método de criba eficiente ya que admite conjuntos con casi todos sus elementos múltiplos de p para algún p. Obsérvese que, en el caso de la dimensión fuerte, el tamaño de L puede rebajarse sin más que cribar sólo por primos suficientemente grandes, con lo que w será muy grande, y por tanto $\rho(p) \ll p$. Otra forma posible de rebajar el tamaño de K o L es considerando κ mas grandes. Por otro lado, haciendo $w = 2$ tenemos la estimación directa de $V(P)^{-1} \leq K(2 \log z)^{\kappa}$.

En algunas situaciones es mas útil tener una versión aditiva de la condición de densidad. Concretamente, sea $g(p) = \frac{\rho(p)}{p - \rho(p)} (\sim \frac{|A_p|}{|A| - |A_p|})$, entonces se tiene

1.3.2 Lema. *Sea $g(p)$ tal que*

$$\sum_{w \leq p < z} g(p) \log(p) \leq \kappa \log \frac{z}{w} + C$$

para algún $C > 1$, y para todo $2 \leq w < z$. Entonces κ es una densidad de criba fuerte para $\rho(p)$ con K tal que $\exp(A/\log w) \leq 1 + K/\log w$.

Es fácil ver que si $\rho(p) \leq \kappa$ para todo p entonces $g(p)$ cumple la desigualdad del lema anterior.

En términos directamente de la función $\rho(p)$ se tiene el siguiente resultado.

1.3.3 Lema. *Sea κ una densidad para $\rho(\cdot)$ y $2 \leq w < z$. Entonces*

$$\sum_{w \leq p < z} \frac{\rho(p) \log(p)}{p} < (\kappa + \log K) \log z.$$

Si es una densidad fuerte, entonces

$$\sum_{w\leq p<z}\frac{\rho(p)}{p}\leq \kappa\log\frac{\log z}{\log w}+\frac{K}{\log w}.$$

Ademas $g(p)\leq \frac{K}{\log p}$.

La demostración de ambos resultados se puede encontrar en [G].

Una vez reducidos a una clase de conjuntos de dimensión fija, resolver el Problema 1.1.1 es encontrar un método de criba válido según la Nota 1.2.5, en particular que garantize $f(s)>0$. En este sentido, la definición siguiente es de extrema utilidad.

1.3.4 Definición. *Dado un método de criba para conjuntos de dimensión* κ*, al mayor cero de* $f(s)$*, que se escribe* $\beta(\kappa)$*, se le denomina el límite de criba (del método).*

1.4 Composición de cribas

Ya hemos mencionado a la hora de introducir la dimensión de criba de un conjunto que quizá es necesario considerar sólo primos suficientemente grandes de forma que los elementos del conjunto no fuesen todos múltiplos de uno dado, es decir, permitiendo disminuir L en (1.7). Sin embargo queda el problema de desechar en A aquellos elementos múltiplos de primos pequeños. Este procedimiento es posible gracias a la composición de cribas. En el primer paso, cribamos los primos pequeños, por ejemplo por la criba de Eratostenes-Legendre, y después aplicamos el método de criba mas conveniente al problema en cuestión sólo para cribar primos "grandes". Por último el método de composición de cribas permitirá construir un método de cribas sobre un conjunto $\mathcal{P}$ partiendo de dos definidos sobre sendos subconjuntos disjuntos $\mathcal{P}_1$, $\mathcal{P}_2$ de $\mathcal{P}$ tal que $\mathcal{P}=\mathcal{P}_1\cup\mathcal{P}_2$. En el caso anterior $\mathcal{P}_1$ será el subconjunto de $\mathcal{P}$ formado por los primos pequeños.

Asi pues supongamos que $\lambda_i^{\pm}$ son funciones de criba sobre $\mathcal{P}_i$, con nivel de criba D_i, $i = 1, 2$. Es fácil ver que para cualquier $d|P$ existen únicos $d_1|P_1$, $d_2|P_2$ tal que $d = d_1 d_2$. Asi pues las funciones

$$\lambda^+(d) = \lambda_1^+(d_1)\lambda_2^+(d_2) \qquad \text{y}$$
$$\lambda^-(d) = \lambda_1^+(d_1)\lambda_2^-(d_2) + \lambda_1^-(d_1)\lambda_2^+(d_2) - \lambda_1^+(d_1)\lambda_2^+(d_2),$$

estan bien definidas para cada $d|P$.

1.4.1 Proposición. *Las funciones $\lambda^+(d)$ y $\lambda^-(d)$ son funciones de criba sobre $\mathcal{P}$ de nivel $D = D_1 D_2$. Además se tiene que*

$$V^+(P) = V^+(P_1)V^+(P_2) \qquad y$$
$$\begin{aligned} V^-(P) &= V^-(P_1)V^+(P_2) - V^+(P_1)V^-(P_2) - (V^+(P_1)V^+(P_2)) \\ &= V^-(P_1)V^-(P_2) - (V^+(P_1) - V^-(P_1))(V^+(P_2) - V^-(P_2)). \end{aligned}$$

Demostración. Por (1.4) se tiene[2]

$$(\lambda^+ * 1)(d) = (\lambda_1^+ * 1)(d_1)(\lambda_2^+ * 1)(d_2) \geq (\mu * 1)(d_1)(\mu * 1)(d_2) = (\mu * 1)(d).$$

Por otro lado

$$\begin{aligned} (\lambda^- * 1)(d) &= (\lambda_1^- * 1)(d_1)(\lambda_2^- * 1)(d_2) - \\ &\quad - \big((\lambda_1^+ * 1)(d_1) - (\lambda_1^- * 1)(d_1)\big) \cdot \big((\lambda_2^+ * 1)(d_2) - (\lambda_2^- * 1)(d_2)\big) \\ &\leq (\lambda_1^- * 1)(d_1)(\lambda_2^- * 1)(d_2) \leq (\mu * 1)(d_1)(\mu * 1)(d_2) = (\mu * 1)(d), \end{aligned}$$

ya que $(\lambda_i^+ * 1)(d_i) \geq (\lambda_i^- * 1)(d_i)$ para $i = 1, 2$.

Es interesante observar pues que para dar cotas superiores de $V^+(P)$ necesitamos cotas superiores de $V^+(P_1)$ y $V^+(P_2)$, mientras que para obtener cotas inferiores de $V^-(P)$ hacen falta cotas inferiores de $V^-(P_1), V^-(P_2)$ y cotas superiores de $V^+(P_1)$ y $V^+(P_2)$. Por otro lado observese que en el caso particular en el que $0 \leq \lambda_i^{\pm}(d_i)/\mu(d_i) \leq 1$ entonces $|\lambda^{\pm}(d)| \leq 1$.

[2]En general dadas dos funciones aritméticas f, g, $f * g(n) = \sum_{d|n} f(d)g(n/d)$.

1.5 Transparencias.

Capítol 2

Criba de Selberg; Una cota superior

Luis Dieulefait.

Tanto la charla en el seminario, como las transparencias recogidas al final del capítulo fueron realizadas por L. Dieulefait. El texto sólo intenta reconstruir dicha charla. Cualquier error o ambiguedad es responsabilidad única del transcriptor, J. Jiménez.

2.1 La función $\lambda^+(d)$.

La dificultad intrínseca que aparece en el método de criba desarrollado por Brun en la segunda decena del siglo XX inspiró cierta desconfianza en su utilidad para resolver problemas de caracter aritmético. Sin embargo, a pesar de que varios expertos, (bien es cierto que mas bien pocos), estudiaron y utilizaron dicho método, no fué hasta 1947 que A. Selberg desarrollase un nuevo método basado en la poderosa idea de que $x^2 \geq 0$ para cualquier número real x. Asi de

simple. Concretamente, para cualquier entero m se cumple

$$\sum_{d|m} \mu(d) \leq \left(\sum_{d \leq \sqrt{D}, d|m} \chi(d) \right)^2 = \sum_{\substack{d_1, d_2 \leq \sqrt{D} \\ d_i | m}} \chi(d_1)\chi(d_2),$$

para cualquier función χ tal que $\chi(1) = 1$. Desarrollando el cuadrado y sumando en los elementos de A vemos que, para cualquier $\mathcal{P}$

$$\begin{aligned} S(A,P) &\leq \sum_{a \in A} \sum_{[d_1,d_2]|(a,P)} \chi(d_1)\chi(d_2) = \sum_{[d_1,d_2]|P} \chi(d_1)\chi(d_2) A_{[d_1,d_2]} \\ &= X \sum_{[d_1,d_2]|P} \chi(d_1)\chi(d_2) \tfrac{\rho([d_1,d_2])}{[d_1,d_2]} + \sum_{[d_1,d_2]|P} \chi(d_1)\chi(d_2) r_A([d_1,d_2]) \\ &= XT + E, \end{aligned} \tag{2.1}$$

donde la suma recorre todas las parejas de divisores de P menores que $\sqrt{D}$.

El problema ahora es encontrar la función χ que minimize la forma cuadrática T sujeta a la condición $\chi(1) = 1$. Un método estandar para realizar este cálculo pasa por diagonalizar la forma cuadrática. En este caso se puede hacer fácilmente en términos de la función multiplicativa $g(p) = \rho(p)/(p - \rho(p))$ introducida en el Lema 1.3.2 del capítulo anterior. Concretamente dado $\mathcal{P}$ sea $G(x) = G(P,x) = \sum_{n \leq x, n|P} g(n)$. Entonces

2.1.1 Teorema. *Sea A como en (1.3) del capítulo anterior. Entonces*

$$S(A,P) \leq \frac{X}{G(\sqrt{D})} + E(D,P),$$

donde $E(D,P) = \sum_{[d_1,d_2]|P} \chi(d_1)\chi(d_2) r_A([d_1,d_2])$, con $\chi(d) = 0$ si $d \geq \sqrt{D}$ y

$$\chi(d)\frac{\rho(d)}{d} = \frac{\mu(d)}{G(\sqrt{D})} \sum_{\substack{h \equiv 0 (mod\ d) \\ h|P, h \leq \sqrt{D}}} g(h)$$

en otro caso.

Demostración. Teniendo en cuenta que

$$\frac{d}{g(d)} = \prod_{p|d} \frac{p}{\rho(p)} = \prod_{p|d} \left(1 + \frac{1}{g(p)}\right) = \sum_{\delta|d} \frac{1}{g(\delta)},$$

podemos diagonalizar T en (2.1) como sigue

$$T = \sum_{d_1|P, d_2|P} \chi(d_1)\chi(d_2) \frac{\rho(d_1)\rho(d_2)}{d_1 d_2} \sum_{\delta|(d_1,d_2)} \frac{1}{g(\delta)} =$$
$$= \sum_{\delta|P} \frac{1}{g(\delta)} \left(\sum_{d \equiv 0 (\mathrm{mod}\ \delta)} \chi_d \frac{\rho(d)}{d} \right)^2 = \sum_{\delta|P} \frac{1}{g(\delta)} y^2(\delta).$$

Obsérvese que $y(\delta) = 0$ si $\delta \geq \sqrt{D}$. Podemos despejar el valor de χ de la fórmula para $y(\cdot)$ convolviendo con la función μ. Concretamente

$$\sum_{\delta|P} \mu(\delta) y(\delta d) = \sum_{\delta|P} \mu(\delta) \sum_{c \equiv 0 (\mathrm{mod}\ \delta d)} \chi(c) \frac{\rho(c)}{c} =$$
$$= \sum_{k|P/d} \chi(kd) \frac{\rho(kd)}{kd} \sum_{\delta|k} \mu(\delta) = \chi(d) \frac{\rho(d)}{d},$$

como predecía el teorema. Obsérvese que $(d, \delta) = 1$ en la suma de la izquierda ya que $y(d\delta) = 0$ en otro caso. Tomando $d = 1$ se obtiene

$$\sum_{\delta|P} \mu(\delta) y(\delta) = 1. \tag{2.2}$$

Aplicando la desigualdad de Cauchy-Schwartz a la identidad anterior se obtiene

$$1 \leq \sum_{\delta|P} \frac{\mu(\delta)^2 y(\delta)^2}{g(\delta)} \sum_{\delta|P, \delta \leq \sqrt{D}} g(\delta),$$

y por tanto $T \geq 1/G(\sqrt{D})$, siendo iguales si $g(\delta) = cy(\delta)\mu(\delta)$ para alguna constante c. Basta ahora sumar en δ para obtener $c = G(\sqrt{D})$ por (2.2) como queríamos ver.

2.2 El término de error.

Para que el Teorema 2.1.1 sea útil necesitamos dar una cota del término de error y, por tanto, controlar la función χ.

2.2.1 Lema. *Sea $\chi(d)$ como en el Teorema 2.1.1. Entonces*

$$|\chi(d)| \leq 1.$$

Demostración. Teniendo en cuenta que

$$G(\sqrt{D}) = \sum_{\delta|d} g(\delta) \sum_{\substack{(n,d)=1 \\ n<\sqrt{D}/\delta}} g(n) = \sum_{\delta|d} g(\delta) \sum_{\substack{h\equiv 0 (\mathrm{mod}\ d) \\ n<\sqrt{D}/\delta}} g(n),$$

se deduce

$$\sum_{\delta|d} g(\delta) \sum_{\substack{(n,d)=1 \\ n<\sqrt{D}/d}} g(n) \leq G(\sqrt{D}) \leq \sum_{\delta|d} g(\delta) \sum_{\substack{(n,d)=1 \\ n<\sqrt{D}}} g(n), \qquad (2.3)$$

ya que $g(\cdot)$ es siempre positiva. Por otro lado

$$\sum_{\substack{(n,d)=1 \\ n<\sqrt{D}/d}} g(n) = \sum_{\substack{h\equiv 0 (\mathrm{mod}\ d) \\ h<\sqrt{D}}} g(h),$$

con lo que, por el Teorema 2.1.1,

$$\left|\tfrac{\chi(d)\rho(d)}{d}\mu(d)G(\sqrt{D})\right| = g(d) \sum_{\substack{(n,d)=1 \\ n<\sqrt{D}/d}} g(n) \leq g(d)\tfrac{G(\sqrt{D})}{\sum_{\delta|d} g(\delta)} =$$

$$= \tfrac{G(\sqrt{D})}{\sum_{\delta|d} 1/g(d/\delta)} = \tfrac{G(\sqrt{D})}{\sum_{\delta|d} 1/g(\delta)} = \tfrac{\rho(d)}{d} G(\sqrt{D}),$$

ya que $g(d) = g(d/\delta)g(\delta)$.

2.2.2 Corolario.

$$|E(D,P)| \leq \sum_{[d_1,d_2]|P} |r_A([d_1,d_2])| \leq \sum_{d|P, d<D} 3^{\nu(d)} |r_A(d)|,$$

donde $\nu(d)$ cuenta el número de factores primos distintos de d.

El resultado es inmediato ya que $[d_1, d_2] = p$ sólo es posible para los tres pares $(d_1, d_2) \in \{(1,p),(p,1),(p,p)\}$.

Imponiendo alguna condición adicional sobre el conjunto A se pueden obtener cotas explícitas para el término de error.

2.2.3 Teorema. *Sea (A, P) un problema con dimensión de criba κ y tal que $|r_A(d)| \leq \rho(d)$, para todo $d|P$. Entonces*

1. *Si $\rho(d) \geq 1$ para todo $d|P$ entonces*

$$E(D,P) \ll D \log^{2\kappa} z.$$

2. *Si $\rho(d) = 1$ para todo $d|P$ entonces*

$$E(D,P) \ll \frac{AD}{G^2(\sqrt{D})}$$

para alguna constante A idependiente de κ.

Demostración. Ver transparecias de este capítulo.

2.2.4 Teorema.

$$\pi(y) - \pi(y-x) \leq \frac{2x}{\log x} + O\left(\frac{x}{\log^2 x}\right).$$

Demostración. Basta con dar una cota de $G(\sqrt{D})$ y aplicar los Teoremas 2.1.1 y 2.2.3. Ahora bien

$$G(\sqrt{D}) = \sum_{h<\sqrt{D}} 1/\varphi(h) = \sum_{h<\sqrt{D}} 1/h \prod \left(1 - \frac{1}{p}\right)^{-1} =$$
$$= \sum_{h<\sqrt{D}} 1/\varphi(h) \sum_{p|m \Rightarrow p|h} 1/m \geq \sum_{h<\sqrt{D}} 1/h > \log\sqrt{D}.$$

El teorema se deduce sin mas que escoger $D = X$.

A la vista del teorema anterior uno puede esperar que mejores cotas de $G(\sqrt{D})$ produzcan desigualdades mas finas en las apliciones. En este sentido el siguiente resultado nos permite dar cotas de $G(\sqrt{D})$ en función de la dimensión.

Sea $G_z(\sqrt{D}) = \sum_{d\leq\sqrt{D},d|P(z)} g(d)$, y supongamos que se cumple $B(z) = \frac{1}{\log z}\sum_{p<z}\frac{\rho(p)}{p}\log p \leq B$. Obsérvese que si (A,P) tiene dimensión κ entonces trivialmente $B = \kappa + \log\kappa$ es una constante válida. En estas condiciones podemos dar el siguiente resultado.

2.2.5 Teorema. *Sea $z = D^{1/s}$. Entonces*

$$\frac{1-\exp\left(-\psi_B(s/2)\right)}{V(P(z))} \leq G_z(\sqrt{D}) \leq \frac{1}{V(P(z))},$$

donde $\psi_B(v) = \max\{0, v\log\frac{v}{B} - v + B\}$.

2.2.6 Nota. *Es importante observar que $\psi_B(s/2) \sim \frac{1}{2}s\log s$ cuando $s\to\infty$, y que $\exp(-\psi_B(s/2) \leq e^{7B-s}$ para todo $s > 0$.*

Demostración. Sea $I_z(x) = G_z(\infty) - G_z(x) = \sum_{d\geq x,d|P(z)} g(d)$. El teorema se puede demostrar en términos de esta función como caso particular de las desigualdades $0 \leq V(P(z))I_z(x) \leq \exp(-\psi_B(s/2)$ para $x = \sqrt{D}$. Ahora bien

$$G_z(x) \leq G_z(\infty) = \sum_{d|P(z)} g(d) = \frac{1}{V(P(z))},$$

lo que prueba la desigualdad inferior para la función $I_z(x)$. Para probar la otra desigualdad sea $x = z^v$. Entonces

$$I_z(x) \leq \frac{1}{x^\varepsilon}\sum_{d\geq x,d|P(z)} g(d)d^\varepsilon = \frac{1}{x^\varepsilon}\prod_{p<z}(1+p^\varepsilon g(p)),$$

y por tanto

$$V(P(z))\, I_z(x) \leq \tfrac{1}{x^\varepsilon}\prod_{p<z}\left(1-\tfrac{\rho(p)}{p}+p^\varepsilon\tfrac{\rho(p)}{p}\right) \leq$$
$$\leq \tfrac{1}{x^\varepsilon}\prod_{p<z}\exp\left((p^\varepsilon-1)\tfrac{\rho(p)}{p}\right) = \tfrac{1}{x^\varepsilon}\exp\left(\prod_{p<z}(p^\varepsilon-1)\tfrac{\rho(p)}{p}\right).$$

Este tipo de cotas en términos de la función exponencial hace que la variable $s = \frac{\log D}{\log z}$ sea una forma natural de medir la relación entre D y s. Esta variable, que ya ha aparecido en la definición de límite de criba, será de vital importancia en capítulos posteriores.

Se trata de optimizar el compromiso entre el tamaño de los divisores y su artimética, es decir, el número de factores primos haciendo esta relación constante de alguna manera. En este sentido si tomamos $\varepsilon = c/\log z$ entonces $p^{\varepsilon} - 1 < (e^c - 1)\frac{\log p}{\log z}$, con lo cual

$$I_z(x)V(P(z)) \leq \frac{1}{e^{cv}} \exp\left(\frac{e^c - 1}{\log z} \sum_{p<z} \frac{\rho(p)}{p} \log p\right) \leq \exp\left(-cv + (e^c - 1)B\right).$$

Maximizando la anterior expresión en c se demuestra el teorema. De hecho se puede demostrar la siguiente cota superior para $G_z(z)$.

2.2.7 Teorema. *Supongamos* $\sum_{p<t} g(p) \log p - \kappa \log t < A$ *para cierta constante* $A > 1$ *y para todo* $t < z$. *Entonces*

$$G_z(z) \geq \frac{e^{-\gamma\kappa}}{\Gamma(\kappa + 1)V(P(z))} \left(1 + O(\frac{A}{\log z}\right).$$

Obsérvese que

$$\frac{e^{-\gamma\kappa}}{\Gamma(\kappa + 1)V(P(x))} = \log^{\kappa}(x)\left(1 + O(1/\log x)\right) \prod_{p<x} \left(1 - \frac{1}{p}\right)^{\kappa} \left(1 - \frac{\rho(p)}{p}\right)^{-1}.$$

Para la demostración ver [G,pag 55.]

2.3 Aplicaciones.

Los anteriores teoremas nos permiten dar aplicaciones no triviales tanto para la función $\pi(x)$, como para el problema de los primos gemelos. También permiten dar cotas superiores para la conjetura de Goldbach, (ver transparencias de este capítulo), siendo en este caso de importancia relativa ya que lo que se espera es demostrar la existencia de solución al problema, es decir, una cota inferior y no una cota superior que tiende a infinito.

2.3.1 Teorema. *(Brun-Titchmarsh) Sea $x \leq y$, $k = o(x^{1/2})$. Entonces*

$$\pi(y;l,k) - \pi(y-x;l,k) \leq \frac{2x}{\varphi(k)\log(x/k)} + O\left(x/(k\log^2(x/k))\right),$$

donde la constante involucrada es absoluta.

Demostración. Tomar $\mathcal{P} = \{p < z : p \nmid k\}$, $A = \{kn + l : Y - X \leq n \leq Y\}$ con $X = x/k$, $Y = (y-l)/k$. En este caso se tiene

$$\pi(y;l,k) - \pi(y-x;l,k) \leq S(A,P(z)) + \pi(z;l,k),$$

y por tanto para probar el teorema es suficiente con dar una cota superior para la función $S(A,P)$ con $z = o(x/(k\log^2(x/k))$. Para obtener el teorema debemos dar una cota inferior para $G(\sqrt{D})$ y una superior para el error. Ahora bien por (2.3) se tiene

$$\sum_{\delta|d} \frac{\mu^2(\delta)}{\varphi(\delta)} \sum_{\substack{(n,d)=1 \\ n<\sqrt{D}}} \frac{\mu^2(n)}{\varphi(n)} \geq \sum_{n<\sqrt{D}} \frac{\mu^2(n)}{\varphi(n)}.$$

Teniendo en cuenta que la primera suma es $k/\varphi(k)$ se obtiene

$$G(\sqrt{D}) = \sum_{\substack{(n,d)=1 \\ n<\sqrt{D}}} \frac{\mu^2(n)}{\varphi(n)} \geq \frac{\varphi(k)}{k}\log\sqrt{D}.$$

Basta ahora escoger $D = X$, para obtener el resultado sin mas que aplicar los Teoremas 2.1.1 y 2.2.3.1

2.3.2 Teorema. *Sea $\pi_2(x) = |\{p \leq x : p+2$ es primo.$\}$. Entonces existe una constante c tal que*

$$\pi_2(x) \leq \frac{cx}{\log^2 x}\left(1 + O(\log\log x/\log x)\right).$$

Demostración. Tomar $A = \{n(n+2) : 1 \leq n \leq x\}$. Es fácil ver que $\pi_2(x) \leq S(A,P(z)) + \pi(z)$. Luego de nuevo una cota para $S(A,P)$ es

suficiente para probar el teorema. Ahora bien, por el Teorema 2.2.7 se tiene

$$G(\sqrt{D}) \geq \frac{1}{2} A_D \log^2(\sqrt{D})\,(1 + O(1/\log D))$$

con

$$A_D = \frac{1}{2} \prod_{2<p<\sqrt{D}} \left(1 - \frac{1}{p}\right)^2 \left(1 - \frac{2}{p}\right)^{-1} = \frac{1}{2} \prod_{p>2} \frac{(p-1)^2}{p(p-2)} \left(1 + O\left(1/\sqrt{D}\right)\right),$$

lo que demuestra el teorema de nuevo por los Teoremas 2.1.1 y 2.2.3.2 escogiendo $D = X/\log^7 X$.

2.3.3 Corolario. *La serie* $\sum_{p,p+2=q} \frac{1}{p}$ *es convergente.*

Demostración. Veremos una demostración de este hecho en el capítulo siguiente.

2.4 Transparencias.

Capítol 3

Garbell de Brun i aplicacions.

Jordi Quer

Aquest capítol conté una introducció als garbells combinatoris. Es defineixen, se'n donen alguns exemples, i finalment es veu quin tipus de resultats de garbell es poden aconseguir utilitzant el més simple de tots: el garbell pur. En particular es dóna una demostració (elemental) completa i autocontinguda del resultat de Brun, publicat el 1919, que diu que la suma dels inversos dels primers bessons convergeix.

Altres garbells combinatoris més sofisticats, com el de Rosser-Iwaniec, que aquí només es defineix, permeten obtenir resultats molt més potents, el qual serà l'objecte de capítols posteriors.

3.1 Garbells: definicions i notació

Sigui $\mathcal{P}$ un conjunt de nombres primers. Per a cada nombre z es denotarà

$$P = P(z) = \prod_{\substack{p \in \mathcal{P} \\ p < z}} p$$

el producte dels primers de $\mathcal{P}$ menors que z.

Siguin A un conjunt finit de nombres enters. El problema del garbell consisteix a calcular (estimar) el nombre d'elements de A no divisibles per cap primer $p \mid P$

$$S(A,P) = S(A,z) = |\{a \in A : (a,P) = 1\}|.$$

El principi combinatori d'inclusió-exclusió proporciona la fórmula de Legendre per aquest nombre

$$S(A,P) = \sum_{d|P} \mu(d)|A_d|, \qquad A_d = \{a \in A : d \mid a\}. \tag{3.1}$$

Una demostració simple d'aquesta fórmula s'obté observant que la funció

$$\delta(m) = \sum_{d|m} \mu(d) = \begin{cases} 1, & \text{si} \quad m = 1, \\ 0, & \text{si} \quad m > 1. \end{cases}$$

permet definir la funció característica dels elements que es volen comptar com

$$\delta((a,P)) = \sum_{d|(a,P)} \mu(d)$$

de manera que

$$\begin{aligned} S(A,P) &= \sum_{a\in A} \delta((a,P)) = \sum_{a\in A} \sum_{d|(a,P)} \mu(d) = \sum_{d|P} \mu(d) \sum_{\substack{a\in A \\ d|A}} 1 = \\ &= \sum_{d|P} \mu(d)|A_d|. \end{aligned} \tag{3.2}$$

Per intentar estimar el seu valor, el nombre $S(A,P)$ es descompon en un terme principal i un terme d'error de la manera següent. Per a cada $d \mid P(z)$ el cardinal del conjunt A_d s'escriu com

$$|A_d| = X\frac{\rho(d)}{d} + r(A,d), \tag{3.3}$$

on $X = |A|$ és el nombre total d'elements del conjunt A (o una aproximació d'aquest nombre), $\rho(d)$ és una funció multiplicativa, i $r(A,d)$ representa l'error que es comet en aproximar $|A_d|$ per $X\rho(d)/d$. La

idea és que $\rho(d)/d$ representa la densitat dels elements de A que són divisibles per d. La funció $\rho(d)$ es pot construir definint primer $\rho(p)$ per a cada primer $p \mid P$ i després extenent-la per multiplicativitat, esperant que els errors $r(A,d)$ siguin petits no només per als d primers sino per a tots els $d \mid P$, el qual acostuma a passar en tot problema de garbell "raonable."

Substituint aquestes expressions per a $|A_d|$ en la fórmula de Legendre s'obté la descomposició

$$S(A,P) = X\sum_{d|P}\mu(d)\frac{\rho(d)}{d} + \sum_{d|P}\mu(d)r(A,d).$$

Gràcies a que la funció ρ és multiplicativa, el sumatori del primer sumand és igual a

$$\sum_{d|P}\mu(d)\frac{\rho(d)}{d} = \prod_{p|P}\left(1-\frac{\rho(d)}{d}\right),$$

i es denotarà $V(P) = V(z)$. Aquest nombre es pot interpretar com la probabilitat que un nombre enter escollit a l'atzar no sigui divisible per cap dels primers $p \mid P$. Es té la fórmula

$$S(A,P) = XV(P) + R(A,P)$$

on el primer terme representa el terme principal i el segon és el terme d'error $R(A,P) = \sum_{d|P}\mu(d)r(A,d)$.

Tal i com s'ha obtingut aquesta expressió resulta força inútil, fins i tot si tots els errors $r(A,d)$ són petits per a cada d, degut al gran nombre de termes que intervenen al sumatori que dóna aquest error: $2^{|P|}$ termes. Això impedeix que el terme principal $XV(P)$ sigui una bona estimació de $S(A,P)$.

El mètode per intentar millorar-ho consisteix a substituir els coeficients $\mu(d)$ per altres nombres reals $\lambda(d)$, molts dels quals siguin zero, de manera que es puguin obtenir aproximacions de $S(A,P)$ amb expressions semblants a la fórmula de Legendre però que tinguin molts menys termes al sumatori. Concretament, siguin, per a cada $d \mid P$, $\lambda^-(d)$ i $\lambda^+(d)$ nombres reals tals que es tingui la desigualtat

$$\sum_{d|m}\lambda^-(d) \le \sum_{d|m}\mu(d) \le \sum_{d|m}\lambda^+(d) \qquad \text{per a tot} \quad m \mid P. \tag{3.4}$$

Aplicant aquestes desigualtats al tercer terme de (3.2) s'obtenen immediatament les desigualtats

$$\sum_{d|P} \lambda^-(d)|A_d| \leq \sum_{d|P} \mu(d)|A_d| \leq \sum_{d|P} \lambda^+(d)|A_d|. \tag{3.5}$$

És interessant observar que si aquestes desigualtats (3.5) es compleixen per a tot conjunt A aleshores les anteriors (3.4) es dedueixen aplicant aquestes al conjunt $A = \{m\}$, de manera que les funcions de garbell són en cert sentit universals.

Aplicant les aproximacions (3.3) de $|A_d|$ a aquestes fites s'obtenen expressions que són fites superior i inferior del nombre $S(A, D)$

$$\sum_{d|P} \lambda^{\pm}(d)|A_d| = X \sum_{d|P} \lambda^{\pm}(d)\frac{\rho(d)}{d} + \sum_{d|P} \lambda^{\pm}(d) r(A, d).$$

Per poder estimar aquests valors i que el terme d'error no sigui massa gran és convenient que els valors $\lambda^{\pm}(d)$ siguin zero tan com sigui possible. En general, es consideren funcions de garbell tals que $\lambda^{\pm}(d)$ sigui zero per a tot $d > D$, on el nombre D, qua d'anomena *nivell de garbell*, és de la forma $D = z^s$ per un $s \geq 1$. D'aquesta manera, si s és prou petit, la suma que dona el terme d'error en les fites anteriors tindrà pocs termes i potser es podrà fitar convenientment.

3.2 Garbells combinatoris

S'anomenen *garbells combinatoris* els garbells que s'obtenen prenent com a funcions de criba $\lambda^{\pm}(d)$ funcions que siguin truncació de la funció de Möbius $\mu(d)$; és a dir, $\lambda^{\pm}(d)$ és $\mu(d)$ o 0, segons el valor de d. Per tant, un garbell combinatori es pot pensar com una estimació del sumatori

$$S(A, P) = \sum_{d|P} \mu(d)|A_d|$$

obtinguda despreciant uns quants sumands (que correspon a canviar alguns dels $\mu(d)$ per zero).

Formalment es pot escriure de la manera següent: $\mathcal{D}^-$ i $\mathcal{D}^+$ són subconjunts del conjunt de tots els divisors de P, continguts a l'in-

terval $[1, D)$ per un cert nivell de criba D, tals que

$$\sum_{\substack{d|m \\ d\in\mathcal{D}^-}} \mu(d) \leq \sum_{d|m} \mu(d) \leq \sum_{\substack{d|m \\ d\in\mathcal{D}^+}} \mu(d) \tag{3.6}$$

per a tot divisor $m \mid P$. Aleshores s'obtenen fites

$$\sum_{d\in\mathcal{D}^-} \mu(d)|A_d| \leq S(A,P) \leq \sum_{d\in\mathcal{D}^+} \mu(d)|A_d|.$$

Una altra notació que de vegades es fa servir per al mateix és posar

$$\lambda^{\pm}(d) = \mu(d)\chi^{\pm}(d)$$

on $\chi^{\pm}$ són les funcions característiques dels dos subconjunts $\mathcal{D}^{\pm}$.

Per tant, dissenyar garbells combinatoris consisteix a saber trobar conjunts $\mathcal{D}^{\pm}$ que compleixin (3.6). A continuació se'n donaran tres exemples.

3.2.1 Garbell pur de Brun

Aquest garbell va ser introduit per Brun a principis del segle XX. Es tracta de considerar com a conjunts $\mathcal{D}^-$ i $\mathcal{D}^+$ els dels divisors de P que tenen menys d'un nombre fixat $r \geq 1$ de factors primers, parell i senar, respectivament:

$$\begin{aligned} \mathcal{D}^- &= \{d \mid P : \nu(d) < r\}, \qquad r \quad \text{parell} \\ \mathcal{D}^+ &= \{d \mid P : \nu(d) < r\}, \qquad r \quad \text{senar} \end{aligned}$$

on, en endavant, $\nu(d)$ denotarà el nombre de factors primers diferents d'un enter d. Aquest garbell té nivell D per a tot $D \geq z^r$ ja que els divisors $d > D$ han de tenir almenys r factors primers.

Per comprovar que aquests conjunts satisfan la relació (3.6) es te el

3.2.1 Lema. *Si r_0, r_1 són enters positius parell i senar, respectivament, aleshores*

$$\sum_{\substack{d|m \\ \nu(d)<r_0}} \mu(d) \leq \sum_{d|m} \mu(d) \leq \sum_{\substack{d|m \\ \nu(d)<r_1}} \mu(d). \tag{3.7}$$

PROVA: Si $m = 1$ tots tres costats de la desigualtat valen 1. Per $m > 1$ el terme del mig és zero i per tant el que cal veure és que

$$\sum_{\substack{d|m \\ \nu(d)<r_0}} \mu(d) \leq 0 \leq \sum_{\substack{d|m \\ \nu(d)<r_1}} \mu(d),$$

el qual és equivalent a dir que per a tot enter positiu $r > 0$ el nombre

$$\sum_{\substack{d|m \\ \nu(d)<r}} \mu(d)$$

o bé és zero o bé té signe $(-1)^{r-1}$. Per induccio és trivial comprovar que

$$\sum_{\substack{d|m \\ \nu(d)<r}} \mu(d) = (-1)^{r-1}\binom{\nu(m)-1}{r-1},$$

on el coeficient binomial $\binom{\nu(m)-1}{r-1}$ és zero si $r > \nu(m)$. Si $r = 1$ tots dos costats de la identitat valen 1. Suposi's comprovat fins a un valor $r \geq 1$. Aleshores

$$\sum_{\substack{d|m \\ \nu(d)<r+1}} \mu(d) = \sum_{\substack{d|m \\ \nu(d)<r}} \mu(d) + \sum_{\substack{d|m \\ \nu(d)=r}} \mu(d).$$

Per a cada enter $r \geq 0$ el nombre de divisors $d \mid m$ que tenen exactament r factors primers diferents és $\binom{\nu(m)}{r}$ i, per cadascun d'aquests, $\nu(d) = (-1)^r$, per tant

$$\sum_{\substack{d|m \\ \nu(d)=r}} \mu(d) = (-1)^r\binom{\nu(m)}{r},$$

i aplicant hipòtesi d'inducció a l'expressió anterior s'obté

$$\begin{aligned}\sum_{\substack{d|m \\ \nu(d)<r+1}} \mu(d) &= (-1)^{r-1}\binom{\nu(m)-1}{r-1} + (-1)^r\binom{\nu(m)}{r} \\ &= (-1)^r\left(\binom{\nu(m)}{r} - \binom{\nu(m)-1}{r-1}\right) = (-1)^r\binom{\nu(m)-1}{r}.\end{aligned}$$

□

Com a conseqüència s'obté que

$$\sum_{\substack{d|P\\ \nu(d)<r_0}} \mu(d)|A_d| \leq S(A,P) \leq \sum_{\substack{d|P\\ \nu(d)<r_1}} \mu(d)|A_d|,$$

i utilitzant aquesta desigualtat per un parell d'enters consecutius $r, r+1$ es dedueix el

3.2.2 Corol·lari. *Per a tot enter $r > 0$ es té*

$$S(A,P) = \sum_{\substack{d|P\\ \nu(d)<r}} \mu(d)|A_d| + (-1)^r\theta \sum_{\substack{d|P\\ \nu(d)=r}} \mu(d)|A_d|$$

per algun θ amb $0 \leq \theta \leq 1$.

3.2.2 Garbell de Brun-Hooley

És un garbell introduit per Hooley l'any 94 amb el qual s'obtenen immediatament dos resultats que Brun ja havia demostrat de manera molt més complicada: l'existència d'infinits 9-primers bessons i la conjectura de Goldbach per a 9-primers. La idea és descompondre el producte de primers $P = P_1P_2\cdots P_t$ en producte de factors, que equival a donar una partició del conjunt $\{p \in \mathcal{P} : p < z\}$ en t subconjunts disjunts, i aleshores aprofitar les desigualtats (3.7).

Tot factor $d \mid P$ descompon de manera única com $d = d_1d_2\cdots d_t$ amb $d_i \mid P_i$. Siguin $r_1, \ldots, r_t$ enters senars qualssevol. Aleshores aplicant les desigualtats de (3.7), la multiplicativitat de la funció μ, i el fet que el producte de nombres positius és positiu, s'obté immediatament que el conjunt següent

$$D^+ = \{d = d_1d_2\cdots d_t : d_i \mid P_i, \nu(d_i) < r_i\}$$

compleix les condicions (3.6) per poder ser utilitzat com a garbell superior. En efecte, per a tot $m \mid P$, amb $m = m_1m_2\cdots m_t$, es té

$$\sum_{d|m} \mu(d) = \prod_{i=1}^{t} \sum_{d_i|m_i} \mu(d_i) \leq \prod_{i=1}^{t} \sum_{\substack{d_i|m_i\\ \nu(d_i)<r_i}} \mu(d_i) = \sum_{\substack{d|m\\ d\in D^+}} \mu(d).$$

L'intent de fer el mateix per trobar una fita inferir no funciona, ja que el producte de desigualtats no val en aquest cas, en que no es té un control dels signes.

Per obtenir un conjunt adequat per fer un garbell inferior s'utilitza el lema (de demostració trivial per inducció)

3.2.3 Lema. *Siguin* $0 \le x_i \le y_i$ *per* $i = 1, \dots t$. *Aleshores*

$$x_1 x_2 \cdots x_t \ge y_1 y_2 \cdots y_t - \sum_{j=1}^{t} (y_j - x_j) \prod_{i \ne j} y_i$$

Siguin r_i nombres senars, com abans. Per (3.7) es tenen desigualtats

$$\sum_{\substack{d_i \mid p_i \\ \nu(d_i) \le r_i}} \mu(d_i) \le \sum_{d_i \mid p_i} \mu(d_i) \le \sum_{\substack{d_i \mid p_i \\ \nu(d_i) < r_i}} \mu(d_i).$$

Aplicant el lema anterior als nombres

$$x_i = \sum_{d_i \mid p_i} \mu(d_i), \qquad y_i = \sum_{\substack{d_i \mid p_i \\ \nu(d_i) < r_i}} \mu(d_i),$$

i tenint en compte que a partir de les desigualtats anteriors es té

$$0 \le y_i - x_i \le - \sum_{\substack{d_i \mid p_i \\ \nu(d_i) = r_i}} \mu(d_i)$$

resulta que, per a tot $m \mid P$,

$$\sum_{d \mid m} \mu(d) \ge \sum_{\substack{d \mid m \\ \nu(d_i) < r_i}} \mu(d) + \sum_{j=1}^{t} \sum_{\substack{d \mid m \\ \nu(d_i) < r_i, i \ne j \\ \nu(d_j) = r_j}} \mu(d).$$

Aquesta desigualtat demostra que el conjunt $\mathcal{D}^-$ definit per

$$\begin{aligned} D^- = &\{d = d_1 d_2 \cdots d_t : d_i \mid P_i,\ \nu(d_i) < r_i\} \\ &\cup \bigcup_{j=1}^{t} \{d = d_1 d_2 \cdots d_t : d_i \mid P_i,\ \nu(d_i) < r_i \text{ si } i \ne j,\ \nu(d_j) = r_i\} \end{aligned}$$

satisfà la propietat (3.6).

3.2.3 Garbell de Rosser-Iwaniec

Brun va introduir els conjunts següents, per a ser utilitzats en la criba

$$\begin{aligned}\mathcal{D}^- &= \{d = p_1 \cdots p_\ell : \quad p_1 > \cdots > p_\ell, \quad p_r < y_r \quad \text{per a tot } r \text{ parell}\}\\ \mathcal{D}^+ &= \{d = p_1 \cdots p_\ell : \quad p_1 > \cdots > p_\ell, \quad p_r < y_r \quad \text{per a tot } r \text{ senar}\},\end{aligned}$$

on els *paràmetres de truncació* y_r són fites fixades arbitrariament o, fins i tot, poden ser funcions que depenguin dels primers que han sortit anteriorment $y_r = y_r(p_1, \ldots, p_{r-1})$. Per exemple, Brun va estudiar el cas $y_r = D^{\alpha\beta^r}$ on D és el nivell de garbell que es vol assolir i α, β són constants $0 < \alpha, \beta < 1$ a escollir.

Una altra possibilitat, anomenada criba de Rosser o de Rosser-Iwaniec, consisteix a imposar desigualtats

$$p_1 p_2 \ldots p_{r-1} p_r^{\beta+1} < D,$$

que equivalen a considerar paràmetres de truncació que depenen dels primers anteriors

$$y_r = \left(\frac{D}{p_1 \cdots p_{r-1}}\right)^{1/(\beta+1)}.$$

Aquest garbell combinatori és el que serà utilitzat més endavant a les xerrades del seminari per arribar al resultat d'Iwaniec sobre quasi-primers representats per $x2 + 1$.

La justificació que aquests conjunts satisfan (3.6) s'obté a partir del principi d'inclusió exclusió de la manera següent. Per a tot primer p i enter m lliure de quadrats, sigui $m_p = \{q \mid m : q < p\}$ el producte dels factors primers de m menors que p. Tot divisor $d \mid m$ més gran que 1 s'escriu de manera única com $d = pd_p$, on p ha de ser necessàriament el factor primer mes gran dels que divideixen d. Per a tot $d \mid P$ es té la fórmula

$$\begin{aligned}\delta(m) &= \sum_{d|m} \mu(d) = 1 + \sum_{\substack{d|m\\ d>1}} \mu(pd_p) = 1 + \sum_{p|m} \mu(p) \sum_{d_p|m_p} \mu(d_p) = \\ &= 1 - \sum_{p|m} \delta(m_p). \qquad (3.8)\end{aligned}$$

Com que els $\delta(m_p)$ només poden ser zero o u, si es treuen termes del sumatori de la dreta l'únic que pot passar és que el valor augmenti, i es té una desigualtat

$$\delta(m) \leq 1 - \sum_{\substack{p_1 | m \\ p_1 < y_1}} \delta(m_{p_1}).$$

Aplicant (3.8) a cada terme del sumatori de la dreta s'obté

$$\delta(m) \leq 1 - \sum_{\substack{p_1 | m \\ p_1 < y_1}} \delta(m_{p_1}) = 1 - \sum_{\substack{p_1 | m \\ p_1 < y_1}} 1 + \sum_{\substack{p_2 < p_1 | m \\ p_1 < y_1}} \delta(m_{p_2})$$

i ara, com que el sumatori és de termes ≥ 0 no es pot treure res si es vol mantenir la desigualtat. Fent un pas més, es té

$$\delta(m) \leq 1 - \sum_{\substack{p_1 | m \\ p_1 < y_1}} 1 + \sum_{\substack{p_2 < p_1 | m \\ p_1 < y_1}} 1 - \sum_{\substack{p_3 < p_2 < p_1 | m \\ p_1 < y_1}} \delta(m_{p_3})$$

i ara es poden eliminar del sumatori de la dreta tots els sumands que es vulgui, per exemple els que tenen $p_3 < y_3$, mantenint la desigualtat. Seguint aquest procés es veu que $\mathcal{D}^+$ és un garbell superior i amb el procés anàleg per primers en posició parell es veu que $\mathcal{D}^-$ dona lloc a un garbell inferior.

3.3 Fórmules de Buchstab

Sigui z la fita dels nombres primers que es garbellen, de manera que

$$P = P(z) = \{p \in \mathcal{P} : p < z\}.$$

Per a cada divisor $d \mid p$ que sigui $d \neq 1$ es denotarà $p(d)$ el més petit factor primer de d. Aleshores

3.3.1 Lema. (Fórmula de Buchstab) *Per a tot enter $r \geq 1$ es té la identitat*

$$S(A, z) = \sum_{\substack{d | P \\ \nu(d) < r}} \mu(d) |A_d| + (-1)^r \sum_{\substack{d | P \\ \nu(d) = r}} S(A_d, p(d)).$$

PROVA: Per inducció. Quan $r = 1$, tenint en compte que $A_1 = A$, la identitat és

$$S(A, z) = |A| - \sum_{p|P} S(A_p, p).$$

Cada $S(A_p, p)$ compta el nombre d'elements $a \in A$ que són divisibles per p i no són divisibles per cap primer $q < p$; és a dir, els elements de A tals que el seu factor primer més petit és p. Cada element $a \in A$ amb $(a, P) > 1$ intervé a un i només un dels $S(A_p, p)$, aquell amb $p = p(d)$. Això demostra la igualtat.

Suposi's comprovada per un cert nombre r. Aleshores aplicant el cas $r = 1$ a cadascun dels $S(A_d, p(d))$ s'obté

$$S(A_d, p(d)) = |A_d| - \sum_{q<p} S(A_{dq}, q).$$

Utilitzant que $p(dq) = q$, la fórmula per al cas r queda

$$S(A, z) = \sum_{\substack{d|P \\ \nu(d)<r}} \mu(d)|A_d| + (-1)^r \sum_{\substack{d|P \\ \nu(d)=r}} |A_d| + (-1)^{r+1} \sum_{\substack{d|P \\ \nu(d)=r}} S(A_{dq}, p(dq)).$$

Els termes del sumatori del mig són $(-1)^r \mu(d)|A_d|$, de manera que els dos sumatoris inicials donen

$$\sum_{\substack{d|P \\ \nu(d)<r+1}} \mu(d)|A_d|.$$

Tenint en compte que quan d recorre els divisors amb $\nu(d) = r$ factors primers, dq amb $q < p(d)$ recorre els divisors amb $r+1$ factors primers, el darrer sumand és

$$(-1)^{r+1} \sum_{\substack{d|P \\ \nu(d)=r+1}} S(A_d, p(d)).$$

□

El resultat següent és l'anàleg per als termes principals.

3.3.2 Lema. *Per a tot enter $r \geq 1$ es té la identitat*

$$V(z) = \sum_{\substack{d|P \\ \nu(d)<r}} \mu(d)\frac{\rho(d)}{d} + (-1)^r \sum_{\substack{d|P \\ \nu(d)=r}} \frac{\rho(d)}{d} V(p(d)).$$

PROVA: Tenint en compte que quan d recorre els divisors de P amb r factors primers i δ recorre els divisors de $P(p(d))$ els productes $d\delta$ recorren tots els factors de m amb $\geq r$ primers, i aplicant la multiplicativitat de μ i de ρ, es té

$$\begin{aligned}(-1)^r \sum_{\substack{d|P\\ \nu(d)=r}} \frac{\rho(d)}{d} V(p(d)) &= \sum_{\substack{d|P\\ \nu(d)=r}} \mu(d)\frac{\rho(d)}{d} \sum_{\delta|P(p(d))} \mu(\delta)\frac{\rho(\delta)}{\delta} = \\ &= \sum_{\substack{d|P\\ \nu(d)\geq r}} \mu(d)\frac{\rho(d)}{d}.\end{aligned}$$

I la fórmula de l'enunciat s'obté separant el sumatori

$$V(z) = \sum_{d|P} \mu(d)\frac{\rho(d)}{d}$$

en dos troços que continguin els divisors amb $< r$ factors primers i els que tenen $\geq r$ factors. □

Aplicant aquestes fórmules de Buchstab s'obté una estimació per a $S(A,P)$ una mica més precisa que si s'aplica directament el garbell pur:

3.3.3 Corol·lari. *Per a tot $r \geq 1$ es té*

$$S(A,z) = XV(z) + \theta X \left(\sum_{\substack{d|P\\ \nu(d)=r}} \frac{\rho(d)}{d} \right) + \theta R(A,D)$$

per algun nombre θ de valor absolut $|\theta| \leq 1$, on $R(A,D)$ denota l'afitació per al reste corresponent a nivell $D = z^r$

$$R(A,D) = \sum_{\substack{d|P\\ d<D}} |r(A,d)|.$$

PROVA: Aplicant la fórmula de Buchstab i el lema anterior s'obté

$$\begin{aligned}S(A,z) &= XV(z) + \sum_{\substack{d|P\\ \nu(d)<r}} \mu(d) r(A,d) + \\ &+ (-1)^r \sum_{\substack{d|P\\ \nu(d)=r}} \left(S(A_d, p(d)) - \frac{\rho(d)}{d} XV(p(d)) \right).\end{aligned}$$

Tenint en compte que

$$0 \leq S(A_d, p(d)) \leq |A_d| \leq X\frac{\rho(d)}{d} + |r(A,d)|$$

$$0 \leq \frac{\rho(d)}{d}XV(p(d)) \leq \frac{\rho(d)}{d}X \leq X\frac{\rho(d)}{d} + |r(A,d)|$$

resulta que cadascun dels sumands del darrer sumatori està fitat per

$$\left|S(A_d, p(d)) - \frac{\rho(d)}{d}XV(p(d))\right| \leq X\frac{\rho(d)}{d} + |r(A,d)|$$

i l'estimació de l'enunciat s'obté immediatament a partir d'aquesta desigualtat tenint en compte que sempre que $\nu(d) \leq r$ aleshores $d < D$. □

3.4 Estimacions amb garbells combinatoris i aplicacions

A partir del corol·lari anterior es dedueix fàcilment el lema fonamental següent, que val per a qualsevol garbell

3.4.1 Teorema. *Sigui $c = 3.591\ldots$ la solució de $(c/e)^c$. Aleshores per a tot $s \geq 1$ i $D \geq z^{s+c|\log V(z)|}$ es té*

$$S(A,z) = XV(z)\left\{1 + \theta e^{-s}\right\} + \theta R(A,D).$$

Prova: Es considera l'estimació del corol·lari 3.3.3. Aleshores per a tot $r \geq 1$ sigui

$$G_r = \sum_{\substack{d|P \\ \nu(d)=r}} \frac{\rho(d)}{d}$$

i sigui $G = G_1$. Com que G^r conté $r!$ vegades dada sumand $\rho(d)/d$ més altres coses positives es dedueix que

$$G_r \leq \frac{1}{r!}G^r.$$

Aplicant la desigualtat $r! \geq e(r/e)^r$, amb demostració trivial per inducció, i tenint en compte que $G \leq |\log V|$,

$$G_r \leq \frac{1}{e}\left(\frac{eG}{r}\right)^r \leq \frac{1}{e}\left(\frac{e}{r}|\log V|\right)^r = e^{r-1}\left(\frac{|\log V|}{r}\right)^r.$$

Sigui $c = 3.591...$ la solució de l'equació $(c/e)^c = e$. Per a tot $b \geq c$ es té $b^b \geq e^{b+1} \geq e^{2b-c+1}$. Suposi's que $b = r/|\log V| \geq c$. Aleshores es té

$$\left(\frac{r}{|\log V|}\right)^r \geq e^{(2r/|\log V|-c+1)|\log V|} = e^{2r-c|\log V|}V^{-1}$$

i d'aquí es dedueix

$$G_r \leq e^{-r-1+c|\log V|}V.$$

Donat un nombre $s \geq 1$ sigui $r = \lfloor s + c|\log V|\rfloor \geq c|\log V|$. Aquesta elecció de r compleix la condició $r/|\log V| \geq c$ que es necessitava per aplicar les desigualtats anteriors. Aleshores $r \geq s + c|\log V| - 1$ i per tant $-r - 1 + c|\log V| \leq -s$ de manera que es té la desigualtat

$$G_r \leq e^{-s}V$$

tal i com es volia. Pel que fa al reste, l'afitació és la del corol·lari anterior, tenint en compte el valor de r que s'ha considerat. □

Amb aquest resultat és clar que $XV(z)$ aproxima tan bé com es vulgui $S(A, z)$ prenent s prou gran sempre que això no ens impedeixi fitar l'error. Buscant la millor elecció de s en el cas de garbells que tenen densitat feble κ

3.4.2 Corol·lari. *Suposi's que el garbell compleix $|r(A, d)| \leq \rho(d)$ i que té densitat de garbell feble κ. Aleshores*

$$S(A, z) = XV(z)\left\{1 + 4K5\theta(2\log z)^{5\kappa}e^{-\log X/\log ez}\right\}$$

per algun θ de valor absolut $|\theta| \leq 1$.

En particular, $S(A, z) \sim XV(z)$ quan $X \to \infty$ uniformement per $z \leq X^{1/5\kappa \log\log X}$.

Prova: Sigui K la constant associada al fet que la densitat feble és κ. En particular es compleix la desigualtat

$$V(z)^{-1} \leq K(2\log z)^{\kappa}.$$

Es parteix de l'estimació del teorema anterior, vàlida per a tot $s \geq 1$, d'on es dedueix

$$S(A, z) = X\left\{V(z) + \theta e^{1-s}\right\} + \theta R(A, D)$$

per a tot s. En efecte, si $s \geq 1$ és conseqüència del teorema observant que $V(z)e^{-s} < ee^{-s} = e^{1-s}$ i per valors $s \leq 1$ alesrhores $e^{1-s} \geq 1$ i l'estimació és trivial.

El resultat es demostrarà prenent el valor

$$s = c \log V(z) + \log X / \log ez$$

a aquesta expressió, que correspon a tenir

$$\begin{aligned} D &= z^{\log X/\log ez} = e^{\log z \log X/\log ez} = X^{\log z/\log ez} = X^{1-1/\log ez} = \\ &= Xe^{-\log X/\log ez}. \end{aligned}$$

Si els restes estan fitats de la manera de l'enunciat, es té l'afitació

$$\begin{aligned} R(A,D) &= \sum_{\substack{d|P \\ d<D}} |r(A,d)| \leq D \sum_{\substack{d|P \\ d<D}} \frac{\rho(d)}{d} = D \prod_{p|P} \left(1 + \frac{\rho(p)}{p}\right) \leq \\ &\leq D \prod_{p|P} \left(1 - \frac{\rho(p)}{p}\right)^{-1} = DV(z)^{-1}. \end{aligned}$$

i per al valor de D que es considera, es té

$$R(A,D) \leq Xe^{-\log X/\log ez} V(z)^{-1} \leq Xe^{-\log X/\log ez} V(z)^{-c}.$$

D'altra banda, per aquest valor de s es té

$$e^{1-s} = ee^{-c\log V(z) - \log X/\log ez} = eV(z)^{-c} e^{-\log X/\log ez}.$$

Tenint en compte que $1 + e < 4$ s'obté

$$S(A,P) = XV(z)\left\{1 + 4\theta V(z)^{-c-1} e^{-\log X/\log ez}\right\}$$

i ara aplicant la desigualtat $V(z)^{-1} \leq K(2\log z)^{\kappa}$ s'obté el resultat.

Per veure la observació final, si $z \leq X^{1/5\kappa \log\log X}$ aleshores

$$\log ez \leq 1 + \frac{\log X}{5\kappa \log\log X} < \frac{\log X}{5\kappa \log\log X}$$

i per tant

$$-\frac{\log X}{\log ez} \leq \frac{-\log X}{\frac{\log X}{5\kappa \log\log X}} = -5\kappa \log\log X$$

d'on es dedueix que

$$e^{-\log X/\log ez} \leq e^{-5\kappa \log\log X} = (\log X)^{-5\kappa}.$$

D'altra banda

$$(2\log z)^{5\kappa} \leq \left(\frac{2\log X}{5\kappa \log\log X}\right)^{5\kappa}$$

de manera que el producte és

$$(2\log z)^{5\kappa} e^{-\log X/\log ez} \leq \left(\frac{2}{5\kappa \log\log X}\right)^{5\kappa}$$

que tendeix a zero quan $X \to \infty$. □

3.4.3 Corol·lari. *Aplicant-lo al garbell dels nombres primers s'obté*

$$\pi(x) \ll x\frac{\log\log x}{\log x}.$$

PROVA: Es tracta d'un garbell lineal, amb $\kappa = 1$, i

$$V(z) = \prod_{p<z}\left(1 - \frac{1}{p}\right) \ll \frac{1}{\log z}.$$

Aplicant el corol·lari anterior pel valor $z = X^{1/5\log\log X}$, de manera que

$$\frac{1}{\log z} = \frac{5\log\log X}{\log X}$$

s'obté immediatament el resultat. □

3.4.4 Corol·lari. *Aplicant-lo al garbell dels nombres primers bessons s'obté*

$$\pi_2(x) \ll x\left(\frac{\log\log x}{\log x}\right)2.$$

PROVA: El mateix que en el cas anterior, tenint en compte que el garbell dels primers bessons té densitat 2. □

3.4.5 Corol·lari. (Brun) *La suma dels inversos dels primers bessons és convergent*

$$\sum_{p,p+2 \text{ primers}} \left(\frac{1}{p} + \frac{1}{p+2}\right) = 1.9021602393....$$

El valor del límit s'anomena constant de Brun.

PROVA: Sigui C la constant associada a la desigualtat del corol·lari anterior. Aleshores

$$\begin{aligned}\sum_{p,p+2 \text{ primers}} \left(\frac{1}{p} + \frac{1}{p+2}\right) &< 2 \sum_{p,p+2 \text{ primers}} \frac{1}{p} = 2\sum_{k=0}^{\infty} \left(\sum_{\substack{p,p+2 \text{ primers} \\ e^k \le p < e^{k+1}}} \frac{1}{p} \right) \\ &< 2\sum_{k=0}^{\infty} \frac{1}{e^k} \pi_2(e^{k+1}).\end{aligned}$$

Fent servir la fita per la funció π_2 obtinguda es té

$$\begin{aligned}\sum_{p,p+2 \text{ primers}} \left(\frac{1}{p} + \frac{1}{p+2}\right) &< 2C \sum_{k=0}^{\infty} \frac{e^{k+1}(\log(k+1))2}{e^k(k+1)2} = \\ &= 2eC \sum_{k=1}^{\infty} \frac{(\log k)2}{k2} < \infty,\end{aligned}$$

degut a que $(\log x)2$ creix més poc a poc que qualsevol potència positiva de x. □

Capítol 4

El garbell combinatori.

Fernando Chamizo.

4.1 Cribas combinatorias

La fórmula exacta de la criba de Eratóstenes-Legendre (inclusión-exlusión),

$$S(A,z)=\sum_{d|P(z)}\mu(d)|A_d|, \tag{4.1}$$

tiene una seria deficiencia, y es que el número de sumandos crece exponencialmente con z y por tanto en las aproximaciones $|A_d|=X\rho(d)/d+r(A,d)$ la acumulación de los términos de error $r(A,d)$ arruina el término principal excepto para z excesivamente pequeño.

Supongamos que sólo tenemos un control adecuado de los términos de error cuando $d<D$ con $D=z^s$, para algún $s>1$; típicamente una acotación de la forma

$$\sum_{\substack{d|P(z)\\ d<D}}|r(A,d)|\ll\frac{X}{\log^C X} \tag{4.2}$$

para $C>0$ suficientemente grande, una vez establecida una cota para D en función de X.

La criba combinatoria consiste en despreciar en (4.1) los sumandos correspondientes a los valores de d que estén fuera de cierto sub-

conjunto de $[1, D]$ (a D se le llama nivel de criba). Con ello perderemos la igualdad pero procediendo adecuadamente todavía podremos conseguir desigualdades.

Concretamente, dados dos subconjuntos $\mathcal{D}^+, \mathcal{D}^- \subset [1, D]$ para los que se verifique

$$\sum_{\substack{d|n \\ d\in\mathcal{D}^-}} \mu(d) \leq \sum_{d|n} \mu(d) \leq \sum_{\substack{d|n \\ d\in\mathcal{D}^+}} \mu(d)$$

se obtiene

$$\sum_{\substack{d|P(z) \\ d\in\mathcal{D}^-}} \mu(d)|A_d| \leq S(A,z) \leq \sum_{\substack{d|P(z) \\ d\in\mathcal{D}^+}} \mu(d)|A_d|. \tag{4.3}$$

La aproximación $A_d \approx X\rho(d)/d$ lleva comúnmente a términos principales comparables a $XV(z)$, salvo constantes dependiendo de s, y se siguen desigualdades del tipo

$$f(s)XV(z) + \texttt{error} \leq S(A,z) \leq F(s)XV(z) + \texttt{error}, \tag{4.4}$$

donde, como antes, $s = \log D/\log z$ (esto es, $D = z^s$), $s \geq 1$, y el error viene de (4.2).

4.2 La criba de Brun

La criba de Brun se basa en las desigualdades

$$\sum_{\substack{d|n \\ \omega(d)\leq 2k+1}} \mu(d) \leq \sum_{d|n} \mu(d) \leq \sum_{\substack{d|n \\ \omega(d)\leq 2k}} \mu(d),$$

donde $\omega(d)$ es el número de factores primos de d, de forma que $\mathcal{D}^-$ y $\mathcal{D}^+$ son subconjuntos de enteros con un número impar y par de factores primos, respectivamente.

Brun perfeccionó esta elección y probó algunos resultados notables (el más conocido es que la suma de los inversos de los primos gemelos converge). Como es de esperar, cuando $s \to \infty$, es decir

cuando la restricción dada por el tamaño de D no se hace sentir, los términos principales en (4.4) se acercan al término esperado desde el punto de vista probabilista: $XV(z)$. A la cuantificación de esta propiedad se le suele llamar *lema fundamental*. Explícitamente, si K y κ son como en la definición de dimensión de criba:

$$\prod_{w\le p<z}\big(1-\rho(p)/p\big)^{-1}\le K(\log z/\log w)^{\kappa}, \tag{4.5}$$

una forma del lema fundamental afirma que para s suficientemente grande en comparación con κ, se cumple que $f(s)$ y $F(s)$ son $1+O\big(K^{10}e^{9\kappa-s}\big)$ con una constante O absoluta.

4.3 Las iteraciones de Buchstab

Buchstab introdujo la fórmula

$$S(A,z)=|A|-\sum_{p<z}S(A_p,p), \tag{4.6}$$

cuya demostración se reduce a notar que los elementos cribados están en alguno de los conjuntos contados por los $S(A_p,p)$ y que éstos son disjuntos. Una de las virtudes de esta fórmula es que permite transformar cotas inferiores en cotas superiores y viceversa. Pero sobre todo, permite en algunas ocasiones mejorar resultados de criba.

Procediendo sin rigor, se pueden anticipar los resultados de la criba de Rosser aplicando sucesivamente (4.6), éstas son las iteraciones de Buchstab. Antes de ilustrar este punto, nótese la fórmula

$$V(z)=1-\sum_{p<z}\frac{\rho(p)}{p}V(p) \tag{4.7}$$

que, aunque inmediata, podría considerarse una consecuencia de (4.6).

Supongamos que en (4.4) despreciamos el error (lo que indicamos empleando $\lesssim$ en lugar de $\le$). Por (4.6)

$$S(A,z)\lesssim X-\sum_{p<z}f(s_p)\frac{X}{p}V(p)$$

donde $s_p = \log(D/p)/\log p = \log D/\log p - 1$. Empleando (4.7) esto se puede reescribir como

$$S(A,z) \lesssim XV(z)\left(1+\sum_{p<z}\big(1-f\big(\frac{\log D}{\log p}-1\big)\big)\frac{\rho(p)}{p}\frac{V(p)}{V(z)}\right).$$

Si pudiéramos aplicar la definición de dimensión (4.5) con $K=1$ e igualdad, por (4.7) se tendría que $-\rho(t)V(t)/t$ es el incremento de $\log^{-\kappa} t$ y $V(z)$ es como $\log^{-\kappa} z$ de modo que el paréntesis exterior se debería aproximar por

$$1+\int_0^z \big(1-f\big(\frac{\log D}{\log t}-1\big)\big)\frac{d(\log^{-\kappa} t)}{\log^{-\kappa} z} = 1+\kappa s^{-\kappa}\int_s^\infty (1-f(t-1))t^{\kappa-1}\,dt$$

(esta igualdad se sigue del cambio $\log D/\log t \mapsto t$).

Lo mismo se aplica con la cota inferior. Si este proceso fuera "contractivo" en el límite se obtendría (4.4) con funciones f y F cumpliendo

$$F(s) = 1+\kappa s^{-\kappa}\int_s^\infty (1-f(t-1))t^{\kappa-1}\,dt,$$
$$f(s) = 1+\kappa s^{-\kappa}\int_s^\infty (1-F(t-1))t^{\kappa-1}\,dt,$$

o lo que es lo mismo

$$\begin{cases} (s^\kappa f(s))' = \kappa s^{\kappa-1}F(s-1) \\ (s^\kappa F(s))' = \kappa s^{\kappa-1}f(s-1) \end{cases} \tag{4.8}$$

Si todo este proceso se pudiera justificar, y supiéramos resolver las ecuaciones (4.8) con $f(\infty)=F(\infty)=1$ (el lema fundamental), tendríamos un candidato para desigualdad de criba óptima. Sin embargo, la acumulación de términos de error y las simplificaciones incorrectas, hacen de ello una tarea poco realista. En su lugar, veremos una criba combinatoria que permite alcanzar el límite de las iteraciones de Buchstab.

4.4 La criba de Rosser

Ciertamente no parece útil contemplar en las iteraciones de Buchstab rangos en los que las desigualdades de criba (4.4) sean triviales.

Supongamos que existe un $\beta > 1$ a partir del cual la cota inferior en (4.4) deja de ser trivial, esto es, $f(\beta) = 0$ y $f(s) > 0$ para $s > \beta$. A este valor se le llama límite de criba (*sieving limit*).

El s correspondiente a $S(A_p, p)$ es $s_p = \log(D/p)/\log p$, y lo anterior sugiere desconsiderar en (4.6) los términos con $\beta \geq s_p$, obteniéndose

$$S(A,z) \leq |A| - \sum_{\substack{p<z \\ p^{\beta+1}<D}} S(A_p, p).$$

En la segunda iteración de Buchstab no podemos eliminar términos sin perder la desigualdad, por tanto tenemos simplemente

$$S(A,z) \leq |A| - \sum_{\substack{p_1<z \\ p_1^{\beta+1}<D}} |A_{p_1}| + \sum_{\substack{p_2<p_1<z \\ p_1^{\beta+1}<D}} S(A_{p_1p_2}, p_2).$$

Pero en la tercera iteración podemos eliminar los términos con $\beta \geq s_{p_1p_2p_3}$, donde $s_{p_1p_2p_3} = \log(D/p_1p_2p_3)/\log p_3$, y se sigue

$$\begin{aligned} S(A,z) \;\leq\; & |A| - \sum_{\substack{p_1<z \\ p_1^{\beta+1}<D}} |A_{p_1}| + \sum_{\substack{p_2<p_1<z \\ p_1^{\beta+1}<D}} |A_{p_1p_2}| \\ & - \sum_{\substack{p_3<p_2<p_1<z \\ p_1^{\beta+1}<D,\ p_3^{\beta+1}p_2p_1<D}} S(A_{p_1p_2p_3}, p_3). \end{aligned}$$

Razonamientos análogos dan lugar a cotas inferiores. Con ello hemos creado una criba combinatoria determinada por

$$\mathcal{D}^+ = \{p_1 \cdots p_m \;:\; p_m < \cdots < p_1 \text{ y } p_{2r+1}^{\beta+1} p_{2r} \cdots p_1 < D \text{ si } 2r+1 \leq m\}$$
$$\mathcal{D}^- = \{p_1 \cdots p_m \;:\; p_m < \cdots < p_1 \text{ y } p_{2r}^{\beta+1} p_{2r-1} \cdots p_1 < D \text{ si } 2r \leq m\}.$$

Ésta es la criba de Rosser (también llamada a veces de Rosser-Iwaniec).

Al ser una criba combinatoria, se pueden controlar los términos de error bajo la condición (4.2) y se obtienen cotas del tipo

$$X \sum_{\substack{d|P(z) \\ d\in\mathcal{D}^-}} \mu(d)\frac{\rho(d)}{d} + \texttt{error} \leq S(A,z) \leq X \sum_{\substack{d|P(z) \\ d\in\mathcal{D}^+}} \mu(d)\frac{\rho(d)}{d} + \texttt{error}.$$

Ahora esperamos extraer de estos sumandos un factor $V(z)$ y acumular las cantidades sobrantes en las funciones f y F. Con tal fin escribimos cada sumatorio como $V(z)$ suprimiendo los productos correspondientes a los elementos que no estén en $\mathcal{D}^-$ o en $\mathcal{D}^+$, y estos productos eliminados los clasificamos en diferentes sumatorios V_m, donde m indica que el m-ésimo mayor primo es el primero para el que falla la condición que define $\mathcal{D}^-$ y $\mathcal{D}^+$. Concretamente:

$$\sum_{\substack{d|P(z)\\ d\in\mathcal{D}^-}} \mu(d)\frac{\rho(d)}{d} = V(z) - \sum_{2|m} V_m(z) \quad \text{y} \quad \sum_{\substack{d|P(z)\\ d\in\mathcal{D}^+}} \mu(d)\frac{\rho(d)}{d} = V(z) + \sum_{2\not|m} V_m(z)$$

con

$$V_m(z) = \sum_{(p_1,p_2,\dots,p_m)\in\mathcal{N}_m} \frac{\rho(p_1p_2\cdots p_m)}{p_1p_2\cdots p_m} V(p_m)$$

donde

$$\mathcal{N}_m = \big\{(p_1,p_2,\dots,p_m) \,:\, p_m<\cdots<p_1<z,\\ p_m^{\beta+1}p_{m-1}\cdots p_1 \ge D, p_n^{\beta+1}p_{n-1}\cdots p_1 < D \text{ si } m-n=\text{par}>0\big\}$$

Como habíamos visto, la definición de dimensión de criba (4.5), suponiendo $K\approx 1$ y usando (4.7), sugiere que $-\rho(t)V(t)/t$ se comporta como el incremento de $\log^{-\kappa} t$, $V(z)$ como $\log^{-\kappa} z$ y $\rho(t)/t$ como el incremento de $\kappa\log\log t$. Así que una conjetura plausible es que

$$\frac{V_m(z)}{\log^\kappa z V(z)} \approx \int\limits_{(t_1,\dots,t_m)\in\mathcal{N}_m} \frac{\kappa\, dt_1}{t_1\log t_1}\cdot\frac{\kappa\, dt_2}{t_2\log t_2}\cdots\frac{\kappa\, dt_{m-1}}{t_{m-1}\log t_{m-1}}\, d(-\log^{-\kappa} t_m).$$

Con el cambio $\log t_j/\log D \mapsto u_j$ se sigue $V_m(z)/V(z)\approx f_m(s)$ donde

$$f_m(s) = \kappa^m s^{-\kappa} \int\limits_{\substack{0<t_m<\cdots<t_1<1/s\\ t_1+\cdots+t_{n-1}+(\beta+1)t_n<1 \text{ si } m-n=\text{par}>0\\ t_1+\cdots+t_{m-1}+(\beta+1)t_m\ge 1}} t_1^{-1}t_2^{-1}\cdots t_{m-1}^{-1}t_m^{-\kappa-1}\, dt_1\, dt_2\ldots\, dt_m.$$

En definitiva, cabe esperar que se cumplan las desigualdades de criba (4.4) con

$$f(s) = 1 - \sum_{2|m} f_m(s) \qquad \text{y} \qquad F(s) = 1 + \sum_{2\not|m} f_m(s) \tag{4.9}$$

para $s > \beta$ con β el "último cero" de f.

4.5 El teorema de criba

Transformar las ideas del apartado anterior en un teorema es una tarea ardua.

Un punto básico es cuantificar la precisión de la aproximación de $V_m(z)/V(z)$ por $f_m(s)$ lo que lleva al problema aparentemente irresoluble de que el error acumulado tiende exponencialmente a infinito con el número de términos. En concreto, se puede probar que para $M \in \mathbb{Z}^+$

$$\sum_{\substack{m\le M\\ m\equiv M \ (2)}} V_m(z) < V(z)\Big(\sum_{\substack{m\le M\\ m\equiv M \ (2)}} f_m(s)+(K-1)M^2K^M\big(\frac{\beta+M}{\beta-1}\big)^{\kappa M}\Big) \tag{4.10}$$

donde K es como en (4.5).

La forma de llegar a (4.10) pasa por notar primero que escribiendo $z_m = D^{1/(\beta+m)}$ e $y_m = \min\big(z, D^{1/(\beta+\delta_m)}\big)$ con $\delta_m = (1-(-1)^m)/2$, se tiene

$$V_m(z) = \sum_{z_m\le p<y_m} \frac{\rho(p)}{p} V_{m-1}(D/p,p) \qquad \text{si } \beta-\delta_m \le s < \beta+m,$$

donde $V_{m-1}(D/p,p)$ es $V_{m-1}(p)$ cambiando D por D/p. Análogamente definiendo $s_m = \max(s, \beta+\delta_m)$ se sigue

$$s^\kappa f_m(s) = \kappa \int_{s_m}^{\beta+m} f_{m-1}(t-1)t^{\kappa-1}\,dt \qquad \text{si } \beta-\delta_m \le s < \beta+m. \tag{4.11}$$

Evidentemente $V_m(z) = f_m(s) = 0$ si $s \ge \beta+m$. Para demostrar (4.10) se aplican estas fórmulas de recurrencia y sumación por partes completándose un proceso de inducción.

Incluso olvidando el crecimiento del error en (4.10), hay un problema de naturaleza más técnica pero en absoluto trivial, y es la convergencia de las series (4.9). Una vez supuesta para $s > \beta$, derivando

(4.11) (nótese que $s_m = s$ para $s_m \geq \beta + \delta_m$) se puede deducir que las funciones (4.9) verifican (4.8). De hecho la convergencia se prueba definiendo unas soluciones de ecuaciones de este tipo que actúan como mayorantes. La determinación de β en función de κ, el único parámetro del que depende (4.11) y por tanto f y F, no es fácil. Para $\kappa = 1$ (criba lineal) se tiene $\beta = 2$, mientras que $\beta \approx 1 + 3'591\kappa$ para κ grande. De la segunda ecuación (4.8) se sigue $F(s) = A/s^\kappa$ donde la constante A también está determinada por κ (debido a la homogeneidad de (4.8) y la condición $F(\infty) = f(\infty) = 1$). De nuevo, no hay una fórmula sencilla para A pero se puede probar que para $\kappa = 1$ se tiene $A = 2e^\gamma = 3'562\ldots$ y $A \approx 2'556(\beta-1)^\kappa$ para κ grande. Una vez hallados $\beta = \beta(\kappa)$ y $A = A(\kappa)$, es fácil calcular numéricamente $f(s)$ y $F(s)$ para cualquier valor de s recursivamente por medio de:

$$\begin{cases} (s^\kappa f(s))' = \kappa s^{\kappa-1} F(s-1) & \text{si } s > \beta \\ (s^\kappa F(s))' = \kappa s^{\kappa-1} f(s-1) & \text{si } s > \beta + 1 \\ f(s) = 0 & \text{si } s \leq \beta \\ F(s) = A/s^\kappa & \text{si } s \leq \beta + 1 \end{cases} \tag{4.12}$$

Para compensar el crecimiento exponencial en M del último término de (4.10), debería ser K muy próximo a 1 pero en (4.5) esto no va a ser posible si w es pequeño. Típicamente se tiene $K = 1 + L/\log w$ con L una constante. Repitiendo la construcción de $\mathcal{D}^-$ y $\mathcal{D}^+$ pero ahora añadiendo la condición de que todos los factores sean mayores que w, podríamos escribir en (4.10) $L/\log w$ en lugar de $K - 1$. Por otra parte, $p_j > w$ implica $v_m(z) = 0$ para $m + \beta \geq \log D/\log w$ de modo que podemos fijar en el término de error $M = \log D/\log w - \beta$ mientras que M continúa siendo arbitrario en los sumatorios. Eligiendo $w = D^{-\epsilon/\log\epsilon}$ (w "poco mayor" que D^ϵ), $\log D > L$ y $\log\log D > \epsilon^{-1}\log^2\epsilon$, se consigue que el término de error sea comparable a ϵL con ϵ tan pequeño como se desee. En definitiva

$$\sum_{\substack{d|P(z) \\ d\in\mathcal{D}_*^-}} \mu(d)\frac{\rho(d)}{d} = V(z)\big(f(s) + O(\epsilon L)\big) \quad \text{y}$$

$$\sum_{\substack{d|P(z) \\ d\in\mathcal{D}_*^+}} \mu(d)\frac{\rho(d)}{d} = V(z)\big(F(s) + O(\epsilon L)\big),$$

donde $\mathcal{D}_*^-$ y $\mathcal{D}_*^+$ son como $\mathcal{D}^-$ y $\mathcal{D}^+$ pero con la condición de que los factores primos sean mayores que $w = D^{-\epsilon/\log\epsilon}$.

Evidentemente esto no es directamente aplicable, porque a la hora de cribar hemos descartado simplemente los primos $p_j \le w$. Pero el lema fundamental asegura que podemos construir una criba de nivel $\widetilde{D} = D^\epsilon$ y $\widetilde{z} = D^{-\epsilon/\log\epsilon}$ para cribar los primos $p_j \le \widetilde{z} = w$ en la que se cumpla (4.4) con $f(s)$ y $F(s)$ iguales a $1 + O(\epsilon L^{10})$.

Finalmente, componiendo ambas cribas se tendría una criba de nivel $D^{1+\epsilon}$ y $F(s)$ y $f(s)$ vendrían dadas por las series (4.9), que a su vez son soluciones de (4.12) (una vez determinados β y A), salvo un error $(1 + O(\epsilon L))(1 + O(\epsilon L^{10})) = 1 + O(\epsilon L^{11})$.

Con algunos cambios, por ejemplo renombrar $D^{1+\epsilon}$ como D y elegir un ϵ adecuado, se tiene el teorema de criba:

4.5.1 Teorema. *Si L es una constante tal que* (4.5) *es válido con $K = 1 + L/\log w$. Para $\log D > 2L$ y suponiendo* (4.2)*, se tiene para $s > \beta = \beta(\kappa)$*

$$\big(f(s)+O(\Delta)\big)XV(z)+O(E) \le S(A,z) \le \big(F(s)+O(\Delta)\big)XV(z)+O(E)$$

donde $E = X/\log^C X$, $\Delta = L^{11}(\log\log\log D)^3/\log\log D$ y $f(s)$ y $F(s)$ son las soluciones de (4.12)*.*

4.6 El caso lineal. Ejemplos

Como hemos mencionado antes, para $\kappa = 1$ (criba lineal) se tiene $\beta = 2$ y $A = 2e^\gamma$. Gracias a (4.12), en el intervalo $[2, 3]$

$$f(s) = \frac{2e^\gamma}{s}\log(s-1) \qquad \text{y} \qquad F(s) = \frac{2e^\gamma}{s}.$$

Iterando se puede extender este rango de definición (con fórmulas cada vez más complejas).

Como ejemplo ilustrativo cribemos en el conjunto

$$A = \{n^2 + 1 \,:\, n \le X\}.$$

De la fórmula $\left(\frac{-1}{p}\right) = (-1)^{(p-1)/2}$, se tiene para $p \neq 2$

$$|A_p| = |\{n \leq X \ : \ p|n^2+1\}| = \begin{cases} 2X/p + O(1) & \text{si } p \equiv 1 \ (4) \\ 0 & \text{si } p \equiv 3 \ (4) \end{cases}$$

Por tanto $\rho(p) = 2$ si $p \equiv 1$ (4), y $\rho(p) = 0$ si $p \equiv 3$ (4). La criba es consecuentemente lineal.

En general se tiene por el teorema chino de resto

$$|A_{p_1 p_2 \cdots p_r}| = \frac{\rho(p_1 p_2 \cdots p_r)X}{p_1 p_2 \cdots p_r} + O(2^r).$$

Así que siempre que $D \leq X/\log^{C+1} X$ está asegurado (4.2).

La condición $s > 2$ que se necesita para conseguir una cota inferior no trivial lleva a que el mayor z aceptable es $X^{1/2-\epsilon}$ con ϵ arbitrariamente pequeño. Con esta elección se tiene

$$S(A,z) \geq C_\epsilon \frac{X}{\log X} \qquad \text{para } X \text{ grande y cierta } C_\epsilon > 0.$$

Los elementos de A están acotados por X^2+1 y como para $\epsilon < 0'1$ se cumple $5(1/2-\epsilon) > 2$, en estas condiciones, los elementos que subsisten tras la criba tienen menos de 5 factores primos. Es decir, hemos probado:

Hay infinitos números de la forma n^2+1 con a lo más cuatro factores primos.

Veamos un ejemplo doble de Selberg que tiene gran interés teórico. Consideremos

$$A^{\text{par}} = \{n \leq 2X \ : \ \lambda(n) = 1\} \quad \text{y} \quad A^{\text{impar}} = \{n \leq 2X \ : \ \lambda(n) = -1\}$$

donde λ es la función de Liouville que vale 1 si el número de factores primos (contando multiplicidades) es par y -1 si es impar.

Se tiene

$$|A_d^{\text{par}}| = |\{n \leq 2X/d \ : \ \lambda(d)\lambda(n) = 1\}| = \frac{1}{2} \sum_{n \leq 2X/d} (1 + \lambda(d)\lambda(n)).$$

Se conoce por métodos de teoría analítica de números que $\sum_{n\leq N}\lambda(n) = O(N/\log^C N)$ para cualquier $C > 0$, por tanto $\rho(d) = 1$ con un error admisible, y la criba es lineal. Lo mismo se aplica a A^{impar}.

Según el teorema, salvo términos de error, $S(A^{\text{par}}, z)$ y $S(A^{\text{impar}}, z)$ están entre $f(s)XV(z)$ y $F(s)XV(z)$. Pero por otra parte, aplicando las iteraciones de Buchstab directamente a $S(A^{\text{par}}, z)$ y $S(A^{\text{impar}}, z)$, desarrollando un proceso iterativo, se puede probar que para $D = X$ y cada $s > 1$

$$S(A^{\text{par}}, z) = f(s)XV(z) + O\big(\frac{X}{\log^2 X}\big)$$

y

$$S(A^{\text{impar}}, z) = F(s)XV(z) + O\big(\frac{X}{\log^2 X}\big).$$

De aquí se pueden deducir dos consecuencias importantes:

- Si $\kappa = 1$, el teorema de criba es óptimo en el sentido de que los términos principales no se pueden mejorar, ya que para A^{par} y A^{impar} se alcanzan.
- Los conjuntos A^{par} y A^{impar} son indistinguibles desde el punto de vista de la criba (ya que $|A_d^{\text{par}}|$ y $|A_d^{\text{impar}}|$ son similares), y sin embargo son bien distintos y $S(A^{\text{par}}, z)$ y $S(A^{\text{impar}}, z)$ tienen diferente asintótica. Hay un límite teórico para separar con métodos de criba números con una cantidad par o impar de factores (fenómeno de paridad).

4.7 Transparencias.

Capítol 5

Gran garbell.

Teresa Crespo.

5.1 La desigualtat de grans garbells

El gran garbell va ser inventat en un article curt per Yu V. Linnik l'any 1941 [L]. Encara que basat en principis molt diferents que els mètodes de garbell convencionals, el mètode de gran garbell s'aplica al problema de garbell usual, és a dir, és una eina per a estudiar conjunts del tipus

$$(\mathcal{M}, \mathcal{P}, \Omega) = \{m \in \mathcal{M} : m(mod\, p) \notin \Omega_p \text{ per a qualsevol } \ p \in \mathcal{P}\}.$$

Els mètodes de gran garbell són més potents que els de garbell combinatori quan el conjunt $\mathcal{M}$ està contingut en un interval curt i el nombre $\omega(p) = |\Omega_p|$ de classes residuals que s'excloen és gran en comparació al mòdul p. El nom de gran garbell fa justament referència al fet que el nombre de classes residuals que s'excloen és gran. La desigualtat de grans garbells també té aplicacions fora de la teoria de garbells.

La desigualtat de grans garbells acota el polinomi trigonomètric

$$S(\alpha) = \sum_n a_n e(\alpha n)$$

on posem $e(x) := e^{2\pi i x}$ i $A = (a_n)$ són nombres complexos en un segment $M < n \leq M + N$. Suposem ara que agafem punts α_r que siguin δ-espaiats mòdul 1, és a dir $\|\alpha_r - \alpha_s\| \geq \delta$ si $r \neq s$, on $\|x\|$ indica la distància de x a l'enter més proper. La desigualtat dels grans garbells afima que

$$\sum_r |S(\alpha_r)|^2 \leq D(\delta, N) \sum_n |a_n|^2$$

on $D(\delta, N)$ depen només de δ i N. Més concretament, s'obté

5.1.1 Teorema. *Per a qualsevol conjunt de punts δ-espaiats $\alpha_r \in \mathbb{R}/\mathbb{Z}$ i a_n nombres complexos qualssevol amb $M < m \leq M + N$, on $0 < \delta \leq \frac{1}{2}$ i N és un enter positiu, tenim*

$$\sum_r |S(\alpha_r)|^2 \leq (\delta^{-1} + N - 1) \sum_n |a_n|^2.$$

La desigualtat del teorema és la millor possible, va ser provada independentment per A. Selberg i H. Montgomery-R.C. Vaughan (1973). Amb valors no tant acurats per a $D(\delta, N)$ es pot provar més facilment la desigualtat (cf. [Gr]) i pot ser suficient per a moltes aplicacions. Per al teorema, Iwaniec [I6] dóna la prova de Montgomery i Vaughan que fa servir propietats de matrius hermítiques i d'operadors en espais de Banach.

5.2 La desigualtat de grans garbells per a caràcters additius

En aplicacions de la desigualtat de grans garbells a la teoria de nombres, s'agafa sovint els punts α_r que siguin racionals a/q amb $1 \leq q \leq Q$ i $(a, q) = 1$. Aquests punts estan espaiats per $\delta = Q^{-2}$, ja que, si $a/q \neq a'/q'$, tenim

$$\|\frac{a}{q}-\frac{a'}{q'}\| = \|\frac{aq'-a'q}{qq'}\| \geq \frac{1}{qq'} \geq \frac{1}{Q^2}.$$

El teorema 5.1.1 dóna

5.2.1 Teorema. *Per a nombres complexos a_n qualssevol amb $M < n \leq M+N$, on N és un enter positiu, tenim*

$$\sum_{q\leq Q} \sum_{a(mod\, q)}^{*} |S(\frac{a}{q})|^2 \leq (Q^2+N-1)\sum_n |a_n|^2.$$

Si $A=(a_n)$ té suport contingut en una progressió aritmètica $n \equiv l\,(mod\, k)$ i $(k,q)=1$, podem fer un canvi de variables i obtenir

5.2.2 Corol·lari. *Per a nombres complexos a_n qualssevol amb $M < n \leq M+N$, tenim*

$$\sum_{\substack{q\leq Q\\(q,k)=1}} \sum_{a(mod\, q)}^{*} |\sum_{n\equiv l(mod\, k)} a_n e(\frac{an}{q})|^2 \leq (Q^2+k^{-1}N)\sum_n |a_n|^2.$$

5.3 Equidistribució entre classes residuals

Com a aplicació del teorema 5.2.1 s'obté que un conjunt general d'enters diferents entre ells $\mathcal{M} \subset (M, M+N]$ representa quasi totes les classes residuals per a quasi tots els mòduls primers $p \leq \sqrt{N}$ sempre que $\mathcal{M}$ sigui gran. Més exactament, agafant per a $A=(a_n)$ la funció característica del conjunt $\mathcal{M}$, s'obté

$$\sum_{p\leq\sqrt{N}} p \sum_{\nu(mod\, p)} |X(p,\nu) - p^{-1}X|^2 \leq 2NX \tag{5.1}$$

on $X=|\mathcal{M}|$ és el nombre d'elements de $\mathcal{M}$, $X(p,\nu)$ és el nombre d'elements m de $\mathcal{M}$ amb $m \equiv \nu\, mod\, p$.

Linnik aplicà aquest resultat per evaluar el menor residu no quadràtic $mod\, p$, és a dir el menor enter positiu $q(p)$ tal que $\left(\frac{q(p)}{p}\right) = -1$. Notem que $q(p)$ és primer. Es conjectura que

$$q(p) <<_\varepsilon p^\varepsilon$$

(és a dir $q(p) \leq K(\varepsilon)p^\varepsilon$, on $K(\varepsilon)$ és un cert factor constant depenent de ε) per a qualsevol $\varepsilon > 0$.

5.3.1 Teorema. (Linnik) *El nombre de primers $p \leq N$ tals que $q(p) > N^\varepsilon$ està acotat per una constant depenent de ε.*

La prova s'obté considerant el problema de garbell $(\mathcal{M}, \mathcal{P}, \Omega)$ amb

$$\begin{aligned} \mathcal{M} &= \{1, 2, \ldots, N\} \\ \mathcal{P} &= \{p \leq \sqrt{N} : \left(\frac{n}{p}\right) = 1 \text{ per a tot } \ n \leq N^\varepsilon\} \\ \Omega_p &= \{\nu mod\, p : \left(\frac{\nu}{p}\right) = -1\} \end{aligned}$$

i aplicant(5.1) amb X el nombre d'elements a

$$(\mathcal{M}, \mathcal{P}, \Omega) = \{1 \leq m \leq N : \left(\frac{m}{p}\right) = 1 \text{ per a qualsevol } \ p \in \mathcal{P}\}.$$

5.4 Generalitzacions

El teorema 5.1.1 es pot generalitzar agafant punts α_r en el tor $T = \mathbb{R}^n/\mathbb{Z}^n$ o més encara considerant un $\mathbb{Z}$-mòdul lliure Λ de rang m, el seu dual $\Lambda' = Hom(\Lambda, \mathbb{Z})$ i el tor $T_\Lambda = \Lambda'_{\mathbb{R}}/\Lambda'$, on $\Lambda'_{\mathbb{R}} = \mathbb{R} \otimes \Lambda'$ (cf. [Se1,2]). Si $\lambda \in \Lambda$, denotem per χ_λ el caràcter corresponent de T_Λ:

$$\chi_\lambda(x) = exp(2\pi i \langle \lambda, x \rangle), x \in T_\Lambda = \Lambda'_{\mathbb{R}}/\Lambda'.$$

5.4.1 Teorema. *Sigui Λ un $\mathbb{Z}$-mòdul lliure de rang m i sigui T_Λ el tor dual. Escollim normes a $\Lambda_\mathbb{R}$ i $\Lambda'_\mathbb{R}$. Existeix una constant c, depenent només de les normes escollides, tal que, si $\delta > 0$ i $x_1, \cdots, x_r \in T_\Lambda$ són δ-espaiats, en el sentit que*

$$\|x_i - x_j\| > \delta, \quad \textit{per a tot } i \neq j,$$

i, si f és una funció complexa sobre T_Λ que és combinació lineal de caràcters χ_λ amb $\lambda \in \Lambda$ en una bola de diàmetre $\leq N$, amb $N \geq 1$, aleshores

$$\sum_{i=1}^{r} |f(x_i)|^2 \leq c \sup(N, \delta^{-1})^m \|f\|_2^2.$$

Observacions. 1) La norma $x \mapsto \|x\|$ sobre T_Λ es defineix per pas al quocient a partir de la norma sobre $\Lambda'_\mathbb{R}$:

$$\|x\| = \inf_{y \mapsto x} \|y\|.$$

2) La L^2-norma $\|f\|_2$ de f relativa a la mesura de Haar dx amb volum 1 és:

$$\|f\|_2^2 = \int_T |f(x)|^2 dx = \sum |a_\lambda|^2,$$

on $f = \sum a_\lambda \chi_\lambda$.

Com a conseqüència del teorema 5.4.1 s'obtenen cotes per al nombre de punts enters de varietats.

5.4.1 Varietats afins

Sigui K un cos de nombres de grau d i $\mathcal{O}_K$ l'anell d'enters de K. Per a $x \in \mathcal{O}_K$, posem

$$\overline{|\, x \,|} = \max_\sigma |\sigma x|$$

on σ recorre les immersions de K en $\mathbb{C}$. Siguin n un enter ≥ 1 i $x = (x^{(1)}, \ldots, x^{(n)}) \in \mathcal{O}_K^n$. Posem

$$\overline{|\ x\ |} = \max_{1 \leq i \leq n} \overline{|\ x^{(i)}\ |}$$

Sigui N un enter ≥ 1 i V una varietat algebraica irreductible de dimensió n a l'espai afí $\mathbf{A}_N$ sobre K. Sigui $V(X)$ el nombre de punts $x \in V$ amb $x^{(i)} \in \mathcal{O}_K$ i $\overline{|\ x\ |} \leq X$. Suposem que el grau de V és ≥ 2.

5.4.2 Teorema. *$V(X) = O(X^{(n-1/2)d}(logX)^{\gamma})$ amb $\gamma < 1$.*

5.4.2 Varietats projectives

Sobre l'espai projectiu $\mathbf{P}_N$ sobre el cos K, sigui H l'altura standard normalitzada. Sigui V una varietat projectiva irreductible de dimensió n a l'espai projectiu $\mathbf{P}_N$. Sigui $V_K(X)$ el nombre de punts $x \in V(K)$ amb $H(x) \leq X$. Suposem que V no és una varietat lineal.

5.4.3 Teorema. *$V_K(X) = O(X^{(n+1/2)d}(logX)^{\gamma})$ amb $\gamma < 1$.*

5.5 La desigualtat de grans garbells per a caràcters multiplicatius

Donat un caràcter multiplicatiu $\chi(mod\, q)$, considerem la suma de Gauss

$$\tau(\chi) = \sum_{a(mod\, q)} \chi(a) e(\frac{a}{q}).$$

Per a un caràcter primitiu $\chi(mod\, q)$, tenim

$$|\ \tau(\chi)\ |^2 = q.$$

Si χ no és primitiu i s és el conductor de χ, posem $q = rs$. Suposem que tenim $(r, s) = 1$. Denotem χ_s el caràcter primitiu induit per χ. Tenim

$$\sum_{a(mod\, q)} \chi(a)e(\frac{an}{q}) = \overline{\chi}(n)\chi_s(r)c_r(n)\tau(\chi_s) \tag{5.2}$$

on $c_r(n) := \sum_{b(mod\, r)} e(\frac{bn}{r})$ és la suma de Ramanujan. En efecte, per la identitat de Bézout, tenim $a = bs + cr$ i

$$\begin{aligned}
\sum_{a(mod\, q)} \chi(a)e(\frac{an}{q}) &= \sum_{b(mod\, r)}{}^{*} \sum_{c(mod\, s)}{}^{*} \chi(bs+cr)e(\frac{bn}{r}+\frac{cn}{s}) \\
&= \sum_{b(mod\, r)}{}^{*} \sum_{c(mod\, s)}{}^{*} \chi(cr)e(\frac{bn}{r}).e(\frac{cn}{s}) \\
&= \chi(r)(\sum_b e(\frac{bn}{r}))(\sum_c \chi(c)e(\frac{cn}{s})) \\
&= \overline{\chi}(n)\chi_s(r)c_r(n)\tau(\chi_s).
\end{aligned}$$

Ara, es compleix $c_r(1) = \mu(r)$, on μ indica la funció μ de Möbius i, per tant, per a $n = 1$, tenim

$$\tau(\chi) = \mu(r)\chi_s(r)\tau(\chi_s)$$

i d'aquí

$$|\, \tau(\chi)\, |^2 = \mu^2(r)s \quad \text{si} \;\; (r,s) = 1.$$

Donada una successió finita $A = (a_n)$ de nombres complexos, denotem, per a qualsevol funció aritmètica $f : \mathbb{N} \to \mathbb{C}$

$$T(f) = \sum_n a_n f(n).$$

En particular, si $f(n) = e(an/q)$, aleshores $T(f) = S(a/q)$. Per (5.2), obtenim

$$T(\overline{\chi}c_r) = \overline{\chi}_s(r)\tau(\chi_s)^{-1} \sum_{a(mod\, q)} \chi(a)S(\frac{a}{q}).$$

D'aquí per la ortogonalitat de caràcters

$$\sum_{\substack{rs\le Q\\(r,s)=1}} \frac{s}{\varphi(rs)} \sum_{\chi(mod\, s)}{}^{*} \mid T(\chi c_r)\mid^2 \le$$

$$\le \sum_{q\le Q}\frac{1}{\varphi(q)} \sum_{\chi(mod\, q)} \mid \sum_{a(mod\, q)} \chi(a)S(\frac{a}{q})\mid^2 = \sum_{q\le Q}\sum_{a(mod\, q)}{}^{*} \mid S(\frac{a}{q})\mid^2 .$$

Aplicant el teorema 5.2.1, obtenim

5.5.1 Teorema. *Donats nombres complexos qualssevol a_n amb $M < n \le M+N$, on N és un enter positiu, tenim*

$$\sum_{\substack{rs\le Q\\(r,s)=1}} \frac{s}{\varphi(rs)} \sum_{\chi(mod\, s)}{}^{*} \mid T(\chi c_r)\mid^2 \le (Q^2+N-1)\sum_n \mid a_n\mid^2 .$$

Aquest resultat és degut a Bombieri i Davenport. Amb $r=1$, obtenim

$$\sum_{q\le Q}\frac{q}{\varphi(q)} \sum_{\chi(mod\, q)}{}^{*} \mid T(\chi)\mid^2 \le (Q^2+N)\sum_n \mid a_n\mid^2 .$$

Si posem $\mathcal{G}(s,\chi)=\sum \dfrac{a_n\chi(n)}{n^s}$, amb $s=\sigma+it$, obtenim

$$\sum_{q\le Q}\frac{q}{\varphi(q)} \sum_{\chi(mod\, q)}{}^{*} \mid \mathcal{G}(s,\chi)\mid^2 \le (Q^2+N)\sum_n \frac{\mid a_n\mid^2}{n^{2\sigma}}.$$

Aquesta última desigualtat es fa servir en la prova del Teorema de Chen.

Teresa Crespo
Departament d'Algebra i Geometria
Universitat de Barcelona
08007 Barcelona
teresa.crespo@ub.edu

Capítol 6

El teorema de Chen

Xavier Xarles

6.1 La conjectura de Goldbach (1742)

En una carta de Goldbach a Euler de l'any 1742, aquest va dir a Euler que creia, després d'unes quantes proves numèriques, que:

Tot nombre senar (> 5) és suma de tres nombres primers.

(el que ara es coneix com a la conjectura de Goldbach Ternària).

Euler li respongué en una altra carta que potser era cert un resultat més fort, i es que:

Tot nombre parell (> 2) és suma de dos nombres primers.

(o conjectura de Golbach Binària).

A més d'aquesta conjectura, també es coneguda una altra conjectura, anomenada dels primers bessons:

Hi ha infinits nombres primers p tals que $p + 2$ és primer.

Fem un repàs ràpid de que es coneix respecte aquestes conjectures.

- Chen (1966): Tot nombre parell prou gran és suma de un primer i un quasiprimer (i.e. de $\mathbb{P}_2$) [Ch].
- Vinogradov (1937): Tot nombre senar prou gran és suma de tres primers. Prou gran: $> 10^{43000}$ (aquest resultat es demostra utilitzant el mètode del cercle de Hardy i Littlewood) [Vi].
- Ramaré (1999): Tot nombre és suma de com molt 7 primers (que es demostra utilitzant la densitat de Schnirelman i afitacions per la densitat de la suma de dos primers) [Ram].
- Deshouillers, Effinger, Te Riele and Zinoviev (1997): L'Hipòtesis de Riemann Generalitzada (GRH) implica que tot nombre senar és suma de tres primers [DERZ].

6.2 Enunciat del teorema

L'objectiu d'aquest capitol és explicar algunes de les idees que hi ha en la demostració del famós teorema de J. Chen, provat l'any 1966. De fet, Chen va anunciar el seu teorema l'any 1966 però no va publicar la demostració fins l'any 1973, aparentment per dificultats degudes a la Revolució Cultural a la Xina. Curiosament, va aparèixer abans la demostració en el llibre de Halberstam i Richert [H-R] que l'article "original"[Ch].

6.2.1 Teorema. *(Chen 1966, 1973) Tot nombre parell suficientment gran és suma d'un nombre primer i un nombre que té com a molt dos factors primers.*

Concretament, si N és un nombre parell i

$$r_2(N) := \sharp\{p \in \mathbb{P} \mid p \leq N \ , \ N - p \in \mathbb{P}_2\}$$

aleshores

$$r_2(N) \geq C \cdot \mathcal{S}(N) \frac{N}{\log^2(N)},$$

on el terme $\mathcal{S}(N)$ és la serie singular del problema:

$$\mathcal{S}(N) := 2 \prod_{p>2} \left(1 - \frac{1}{(p-1)^2}\right) \prod_{2<p|N} \frac{p-1}{p-2}$$

i

$$C = 0.0848$$

(Chen diu $C = 0.33$*).*

Observem que la constant que apareix en l'enunciat:

$$B_2 := 2\prod_{p>2}\left(1 - \frac{1}{(p-1)^2}\right) \sim 1.32045665$$

és l'anomenada constant dels nombres primers bessons.

De fet, en el mateix article Chen va demostrar també el següent resultat sobre la conjectura dels primers bessons.

6.2.2 Teorema. *(Chen) Hi ha infinits primers* p *tals que* $p+2 \in \mathbb{P}_2$*. Concretament, tenim que*

$$|\{p \in \mathbb{P} \mid p \leq N\ ,\ p+2 \in \mathbb{P}_2\}| \geq C \cdot B_2 \frac{N}{\log^2(N)},$$

on $C = 1/31 \sim 0.0323$*.*

Podem comparar aquest resultat amb el que un esperaria obtenir (per exemple, utilitzant el mètode del cercle (*circle method*) de Hardy i Littlewood).

6.2.3 Conjectura *(Goldbach, Hardy-Littlewood) Tot nombre parell suficientment gran (*> 2*) és suma de dos nombres primers.*

Concretament, si N *és un nombre parell i*

$$r(N) := \sharp\{p \in \mathbb{P} \mid p \leq N\ ,\ N - p \in \mathbb{P}\}$$

aleshores

$$r(N) \sim B_2 \prod_{2<p|N} \frac{p-1}{p-2} \frac{N}{\log^2(N)}.$$

Podeu veure que la única diferència, a part de que l'enunciat hauria de ser valid per a sumes de primers i no per a sumes d'un primer i un quasi-primer, és que la constant C hauria de valdre 1 asimptòticament.

Tenim també una fita superior que s'obté fàcilment del garbell de Selberg:

6.2.4 Teorema. *Si N es un nombre parell (prou gran) i*

$$r(N) := \sharp\{p \in \mathbb{P} \mid p \leq N\ ,\ N - p \in \mathbb{P}\}$$

aleshores

$$r(N) \leq 8 \cdot \mathcal{S}(N) \frac{N}{\log^2(N)},$$

on el terme $\mathcal{S}(N)$ és la serie singular del problema:

$$\mathcal{S}(N) := 2 \prod_{p>2} \left(1 - \frac{1}{(p-1)^2}\right) \prod_{2<p|N} \frac{p-1}{p-2}.$$

6.3 Inici del garbell

Sigui N un nombre parell. Considerem el conjunt A del qual anem a intentar garbellar els nombres (quasi)-primers:

$$A := \{N - p \mid p \in P_N, 1 \leq p < N\},$$

on, com es usual, hem denotat

$$P_N := \{p \in \mathbb{P} \mid (p, N) = 1\}.$$

Tenim aleshores que

$$|A| = \pi(N) - \omega(N) = \frac{N}{\log(N)} \left(1 + O(\frac{1}{\log(N)})\right),$$

ja que

$$\omega(N) = O(\log(N))$$

(Exercici: Penseu en els nombres N de la forma $2 \cdot 3 \cdot 5 \cdots p_n$ producte dels n-primers primers).

Fixem un $z < N$ (de fet $z < \sqrt{N}$) i l'objectiu és garbellar de A tots els nombres que són divisibles per primers menors que z. Així considerem

$$P(z) := \prod_{p<z, p\in\mathbb{P}} p$$

i

$$S(A, P(z)) := |\{a \in A \mid (a, P(z)) = 1\}|\,.$$

6.3.1 Observació. Si posem $z = N^{1/k}$ i calculem $S(A, P(z))$, ens queden essencialment els nombres de $\mathbb{P}_{k-1}$. O sigui que són producte de com a molt $k-1$ nombres primers (pseudo-primes d'ordre $k-1$).

El següent conjunt a estudiar és

$$A_d := \{a \in A \mid a \equiv 0 \mod d\} = \{N - p \mid p \in P_N \ , \ p \equiv N \mod d\}.$$

Així el nombre d'elements de A_d és aproximadament igual al nombre de primers menors que N que són congruents amb N mòdul d (aproximadament ja que hem de treure els primers que divideixen a N), i per tant

$$|A_d| = \pi(N; d, N) + O(\omega(N)) = \frac{|A|}{\phi(d)} + r(d)$$

on $\phi(d)$ és la ϕ d'Euler, i $r(d)$ es un cert terme d'error que fitarem més endavant.

Aquí hem utilitzat que, si denotem com és usual per

$$\pi(x; d, a) := \{p \in \mathbb{P} \mid p \equiv a \mod d\}$$

aleshores

$$\pi(x; d, a) = \frac{\pi(x)}{\phi(d)} + \delta(x; a, d)$$

sempre que $(a, d) = 1$, on $\delta(x; a, d)$ és un cert terme d'error que haurem d'afitar (per exemple utilitzant el teorema de Bombieri-Vinogradov).

Utilitzant així l'estratègia dels garbells ja estudiada, tindrem que

$$S(A, P(z)) \simeq X \, V(z) + \text{ Error}$$

on

$$X = \frac{N}{\log(N)} \sim |A|$$

i

$$V(z) = \sum_{d|P(z)} \frac{\mu(d)}{\phi(d)} = \prod_{p|P(z)} \left(1 - \frac{1}{\phi(p)}\right) = \prod_{p|P(z)} \left(1 - \frac{1}{p-1}\right)$$

6.4 La serie singular

Anem primer a veure d'on surt el terme

$$\mathcal{S}(N) := 2\prod_{p>2}\left(1-\frac{1}{(p-1)^2}\right)\prod_{2<p|N}\frac{p-1}{p-2}$$

que apareix a l'enunciat del teorema.

Observem que el terme principal que ens ha sortit abans (i.e. $X\,V(z)$) és de la forma $V(z)N/\log(N)$, així que el que podem esperar és que

$$V(z) \simeq K\mathcal{S}(N)/\log(z)$$

per una certa constant K (que és la "prova" que el garbell "naïf" no ens podrà funcionar de manera òptima). Per a demostrar-ho sols cal utilitzar el lema de Mertens.

6.4.1 Lema. *Per a tot $z < N$ tenim que*

$$V(z) = \prod_{p<z,(p,N)=1}\left(1-\frac{1}{p-1}\right) = \mathcal{S}(N)\frac{e^{-\gamma}}{\log z}\left(1+O\left(\frac{1}{\log N}\right)\right).$$

Demostració (Idea): Sigui

$$W(z) := \prod_{2<p<z}\left(1-\frac{1}{p-1}\right).$$

Veurem que

$$\frac{V(z)}{W(z)} = \prod_{2<p,p|N}\frac{p-1}{p-2}\left(1+O(\frac{1}{\log N})\right)$$

$$W(z) = 2\prod_{p>2}\left(1-\frac{1}{(p-1)^2}\right)\frac{e^{-\gamma}}{\log z}\left(1+O(\frac{1}{\log z})\right).$$

Tenim que

$$\begin{aligned}\frac{V(z)}{W(z)} &= \prod_{2<p<z,p|N}\left(1-\frac{1}{p-1}\right)^{-1} = \\ &= \prod_{2<p,p|N}\left(1-\frac{1}{p-1}\right)^{-1}\prod_{p\geq z,p|N}\left(1-\frac{1}{p-1}\right).\end{aligned}$$

Ara be,

$$\prod_{2<p,p|N}\left(1-\frac{1}{p-1}\right)^{-1}=\prod_{2<p,p|N}\frac{p-1}{p-2}$$

i

$$\prod_{p\geq z,p|N}\left(1-\frac{1}{p-1}\right)=1+O(\frac{1}{\log N})\left(>1-\frac{8\log(N)}{N^{1/8}}\right).$$

D'altre banda

$$W(z)=2\prod_{p>2}\left(1-\frac{1}{(p-1)^2}\right)\prod_{p\geq z}\left(1-\frac{1}{p(p-2)}\right)\prod_{p<z}\left(1-\frac{1}{p}\right)$$

i tenim que (Mertens)

$$\prod_{p<z}\left(1-\frac{1}{p}\right)=\frac{e^{-\gamma}}{\log z}\left(1+O(\frac{1}{\log z})\right)$$

i que

$$\prod_{p\geq z}\left(1-\frac{1}{p(p-2)}\right)=1+O(\frac{1}{z})\left(<1+\frac{2}{z}\right).$$

El resultat s'obté ajuntant totes les fitacions anteriors. □

6.5 El teorema de Bombieri-Vinogradov i el terme d'error

Recordem que volem fitar el terme d'error que farem al aplicar la fórmula

$$S(A,P(z))\simeq X\,V(z)+\text{ Error}$$

o una fórmula similar, on

$$X=\frac{N}{\log(N)}$$

i

$$V(z)=\prod_{p|P(z)}\left(1-\frac{1}{\phi(p)}\right).$$

Per a fer això el que ens interessa és fitar el terme d'error

$$\delta(x;a,d) = \frac{\pi(x)}{\phi(d)} - \pi(x;d,a)$$

en valor absolut.

Ara bé, en general l'error que fem és massa gran i/o massa incontrolable, de manera que al sumar respecte d obtenim un error massa gran. El que podem fer és fitar directament l'error al sumar tots els $\delta(x;a,d)$ per a un d prou petit respecte x. Això és justament el que ens dóna el teorema següent.

6.5.1 Teorema. *(Bombieri-Vinogradov) Per a tot $A > 0$ existeix una constant $B = B(A)$ tal que, si denotem per*

$$D(A) := \frac{x^{\frac{1}{2}}}{(\log x)^{B(A)}}$$

aleshores

$$\sum_{d<D(A)} \max_{(d,a)=1} |\delta(x;d,a)| \ll \frac{x}{(\log x)^A}.$$

De fet, tot i que no ho necessitarem per a res, hi ha fitacions explícites d'aquest $B(A)$ en funció de A. Bombieri demostra que $B(A) = 3A + 22$ va bé, però Vaughau ha demostrat que podem prendre $B(A) = A + 5/2$. També se sap que si és certa l'Hipòtesis de Riemann Generalitzada, aleshores podem prendre $B = A + 1$.

De fet, el resultat ideal seria poder tenir el teorema però amb

$$D(A) := \frac{x^{\delta}}{(\log x)^{B}}$$

per algun $\delta > 1/2$. El fet que només ho sabem per a $\delta = 1/2$ fa que el limit del garbell que obtindrem no es suficient per a demostrar el teorema directament (vegeu més endavant).

Fixem-nos que per a poder aplicar el teorema necessitarem sumar només pels d's prou petits. Per altre banda, donat que el terme principal que tindrem ha de ser de la forma $N/(\log N)^2$, voldrem que el terme d'error sigui més petit que això, per exemple de la forma $N/(\log N)^3$, i.e. $A = 3$ en el teorema, i per tant $B = 6$ ens anirà bé.

6.6 Densitat del garbell

Anem a veure que efectivament estem en un cas de garbell lineal. Cal veure així el següent lema.

6.6.1 Lema. *Sigui N un nombre parell fixat. Aleshores, per a tot $w \geq 2$ i per a tot $z > w$ tenim que*

$$\prod_{w\leq p<z, p\in P_N} \left(1-\frac{1}{p-1}\right)^{-1} < \frac{\log(z)}{\log(w)}\left(1+O(\frac{1}{\log w})\right)$$

Demostració: Observem primer que

$$\prod_{w\leq p<z} \left(1-\frac{1}{p-1}\right)^{-1} = \prod_{w\leq p<z} \frac{(p-1)^2}{p(p-2)} \prod_{w\leq p<z} \left(1-\frac{1}{p}\right)^{-1}$$

L' avantatge d'escriure-ho així és que el primer producte està fitat (de manera absoluta) i tendeix a 1, mentre que el segon producte el podem fitar superiorment fent servir el lema de Mertens.

En efecte, tenim per Mertens que, si $2 \leq w < z$

$$\prod_{w\leq p<z} \left(1-\frac{1}{p}\right)^{-1} < \frac{\log(z)}{\log(w)}\left(1+O(\frac{1}{\log w})\right)$$

D'altre banda, el producte

$$\prod_{w\leq p} \frac{(p-1)^2}{p(p-2)} = \prod_{w\leq p}\left(1+\frac{1}{p(p-2)}\right) < \prod_{w\leq n}\left(1+\frac{1}{n(n-2)}\right) = \frac{w-1}{w-2} < 2$$

esta fitat per una constant que tendeix a 1 quan w augmenta. El resultat és immediat a partir d'aquí. □

6.7 Fitació inferior de $S(A, P(z))$

Primer comencem posant una aplicació directe del teorema de Jurkart-Richert.

6.7.1 Proposició. *Sigui* $D \geq z^2$ *i*

$$s = \frac{\log(D)}{\log(z)}.$$

Aleshores

$$(F(s)+\epsilon)V(z)X + R > S(A, P(z)) > (f(s)+\epsilon)V(z)X - R$$

on

$$R = \sum_{d<D, d|P(z)} |r(d)|,$$

i $\epsilon \longrightarrow 0$ *quan* $D \longrightarrow \infty$.

Ara, donat k i $z = N^{1/k}$, el que volem es poder trobar s com més gran millor ($f(s)$ és creixent) de manera que el terme d'error estigui controlat per Bombieri-Vinogradov i sigui fitat per $N/(\log N)^3$.

Recordem que

$$r(d) = \pi(N; d, N) - \frac{\pi(N)}{\phi(d)} + O(\log N) = \delta(N; d, N) + O(\log N)$$

(l'últim terme prové del teorema dels nombres primers). Així el que volem és que

$$R = \sum_{d<D, d|P(z)} |r(d)| \sim \sum_{d<D} |\delta(N; d, N)| \ll \frac{N}{\log(N)^3}$$

i per tant el que hem de prendre és

$$D = D(3) := \frac{N^{1/2}}{(\log N)^{B(3)}} \ll \frac{N}{\log(N)^3}$$

per una certa constant $B(3)$ (e.g. $B(3) = 6$).

Així, si posem $z = N^{1/k}$, i $D = D(3)$ tenim que

$$s := \frac{\log D}{\log z} = \frac{k}{2} - kB(3)\frac{\log\log N}{\log N},$$

i per tant augmentant N suficientment podem acostar-nos a $k/2$ tant com volem. Observem que només obtindrem alguna fita inferior positiva si $s > 2$ (ja que $f(2) = 0$), i per tant si $k > 4$. Així podem deduir, posant $k = 5$ i utilitzant que si $2 \leq s \leq 4$ aleshores

$$f(s) = \frac{2e^{\gamma}\log(s-1)}{s}$$

el següent primer resultat.

6.7.2 Corol·lari. *Si N és un nombre parell prou gran, aleshores N es pot expressar com la suma d'un nombre primer més un nombre que és com a molt producte de 4 primers (i.e. de $\mathbb{P}_4$). Concretament tenim que*

$$|\{p \in \mathbb{P} \mid p \leq N \ , \ N - p \in \mathbb{P}_4\}| > S(A, P(N^{1/5})) >$$

$$\left(5e^{\gamma} f(\frac{5}{2})\right) \mathcal{S}(N) \frac{N}{\log(N)^2} > .5778 \mathcal{S}(N) \frac{N}{\log(N)^2}$$

Més endavant utilitzarem el cas $k = 8$ i per tant, utilitzant que

$$f(4) = \frac{e^{\gamma} \log(3)}{2}.$$

6.7.3 Corol·lari. *Si $z = N^{1/8}$, aleshores per a N prou gran tenim que*

$$S(A, P(z)) > \left(\frac{e^{\gamma} \log(3)}{2} + O(\epsilon)\right) \frac{N}{\log(N)} V(z)$$

$$> (e^{\gamma} 4 \log(3)) \, \mathcal{S}(N) \frac{N}{\log(N)^2}.$$

6.8 L'estratègia de Chen

Com hem vist, el problema d'aplicar el garbell de la manera com ho hem fet és que, si ho apliquem a $N^{1/k}$ amb k massa petit (voldríem $k = 3$ per obtenir el teorema de Chen, o bé $k = 2$ per a obtenir la conjectura de Goldbach), la fita inferior que obtenim és negativa, i per tant no ens serveix de res.

El problema és que, si bé és cert que tot nombre no divisible per a cap primer menor que la seva arrel k-èssima és automàticament producte de com a molt $(k-1)$-primers, no tots els nombres de $\mathbb{P}_{k-1}$ són d'aquesta forma.

Dit d'una altre manera, no estem considerant nombres que de fet ens podrien anar bé per al nostre teorema.

La idea és considerar els nombres que no són divisibles per cap nombre menor que $N^{1/k}$ ($k = 8$ ens anirà bé, tot i que el Chen ho fa amb $k = 10$), i després anar traient d'aquí d'una altra manera els nombres que no són producte de dos nombres primers.

El que farem és considerar el següent: comptarem els nombres primers $p \leq N$, amb $(p, N) = 1$, tals que $N - p$

1. no és divisible per a cap nombre menor que $N^{1/3}$ (i per tant és com a molt producte de dos nombres primers), o
2. és divisible per un únic nombre primer p_1, amb $N^{1/8} \leq p_1 < N^{1/3}$, i per tant $N - p = p_1 m$, amb m no divisible per a cap nombre primer menor que $N^{1/3}$, i a més m és primer.

Us preguntareu, com podem comptar això realment? Aquí és on es veu (en el meu entendre) el geni de Chen; el que fem és agafar la fita inferior que hem calculat abans en el cas $z = N^{1/8}$, i ara li restem cotes superiors de certs conjunts que podem fitar utilitzant els mètodes anterior de garbells.

Anem a passos: Considerem primer la següent suma: si $z = N^{1/8}$,

$$S(A, P(z)) - \frac{1}{2} \sum_{N^{1/8} \leq p_1 < N^{1/3}, p_1 \in P_N} S(A_{p_1}, P(z)))$$

Aquí el que estem contant són els nombres primers p tals que $N - p$ no és divisible per cap nombre primer menor que z però amb un cert pes (que pot ser i de fet serà en molts casos negatiu). Concretament, estem sumant

1. 1 si $N - p$ no és divisible per cap nombre primer més petit que $N^{1/3}$.
2. $1/2$ si $N - p$ és divisible per exactament un nombre primer més petit que $N^{1/3}$ i més gran que $N^{1/8}$.
3. 0 si $N - p$ és divisible per exactament dos nombres primers més petits que $N^{1/3}$ i més grans que $N^{1/8}$.
4. $-1/2$ si $N - p$ és divisible per exactament tres nombres primers més petits que $N^{1/3}$ i més grans que $N^{1/8}$.

5. i en general $(2-r)/2$ si $N-p$ és divisible per exactament r nombres primers més petits que $N^{1/3}$ i més grans que $N^{1/8}$ (com a molt r valdrà 7).

Si aquesta suma és >0 aleshores :

- $N = p + \mathbb{P}_2$ amb pes 1, o bé
- $N = p + p_1 m$ amb pes $1/2$, on
  - $N^{1/8} \leq p_1 < N^{1/3}$
  - m no divisible per a cap primer $< N^{1/3}$ (per tant $m \in \mathbb{P}_2$).

Ara només ens faltarà treure (amb pes $1/2$) els casos en que el m anterior no és primer.

Observem que si $m = p_2 p_3$, aleshores

$$p_2^2 < p_2 p_3 = \frac{N-p}{p_1} < \frac{N}{p_1}$$

i per tant que

$$N^{\frac{1}{3}} < p_2 < \left(\frac{N}{p_1}\right)^{\frac{1}{2}}.$$

El que fem es considerar el conjunt

$$B := \left\{ N - p_1 p_2 p_3 \mid \begin{array}{c} N^{1/8} \leq p_1 < N^{1/3} \leq p_2 \leq p_3 \ , \\ \\ p_1 p_2 p_3 < N \ , (p_1 p_2 p_3, N) = 1 \end{array} \right\}$$

i intentem fitar el nombre de primers que hi ha a B; per exemple prenen la suma $S(B, P(y))$, on $y = N^{1/3}$:

$$Primers(B) < S(B, P(y)) + y = S(B, P(y)) + N^{1/3}.$$

(aquest últim terme l'ignorarem doncs quedarà incorporat a l'error fet, ja que serà més petit que l'error).

Observem que el que hem fet ha estat intercanviar els papers de p i de $p_1 p_2 p_3$. Enlloc de calcular els nombres de la forma $N-p$ d'una certa forma, el que fem és calcular els nombres de la forma

$N - p_1p_2p_3$ no divisibles per cap primer menor que y. Aquest "intercanvi"(*switch*) és el nucli de la demostració, i la seva gràcia és que B té el mateix ordre que A però és molt més petit. De fet

$$B \sim Const.|A|$$

on

$$Const. = 0.3631$$

Ara tenim que la següent suma ens afita inferiorment el que volem comptar en el teorema: Si denotem per $z = N^{1/8}$ i per $y = N^{1/3}$,

$$S(A, P(z)) - \frac{1}{2} \sum_{z \le p_1 < y\,, p_1 \in P_N} S(A_{p_1}, z)) - \frac{1}{2} S(B, P(y)) - y$$

La part més difícil de la demostració serà la fitació superior de $S(B, P(y))$.

6.9 Fitació superior de $\sum_{z \le q < y\,, q \in P_N} S(A_q, P(z))$

El nostre objectiu ara es trobar una fitació superior del terme

$$\sum_{N^{1/8} \le p_1 < N^{1/3}\,, p_1 \in P_N} S(A_{p_1}, P(z))$$

de manera que no sigui massa gran (comparada amb la fitació inferior que em donat abans del terme $S(A, P(z))$).

6.9.1 Teorema. *Si $z = N^{1/8}$ i $y = N^{1/3}$, aleshores*

$$\sum_{z \le q < y\,, q \in P_N} S(A_q, P(z)) < \left(\frac{e^\gamma \log 6}{2} + \epsilon \right) \frac{N}{\log N} V(z).$$

El que hem de fer es aplicar el teorema de Jurkat-Richert però ara buscant una fita superior. Com abans prenem

$$S(A_q, z) < Constant|A_q|V(z) + Error$$

on la constant i el terme d'error els hem de buscar de manera òptima.

El terme d'error en qüestió vindrà donat per una suma respecte d's petits (coprimers amb tots els primers menors que z, z petit) de

$$r_q(d) = |(A_q)_d| - \frac{|A_q|}{\phi(d)} = |A_{qd}| - \frac{|A_q|}{\phi(d)} =$$

(on hem utilitzat que $(q,d)=1$ ja que $z\leq q$)

$$= \left(|A_{qd}| - \frac{|A|}{\phi(qd)}\right) + \left(\frac{|A|}{\phi(qd)} - \frac{|A_q|}{\phi(d)}\right) =$$

i per definició de $r(d)$ d'abans

$$= r(qd) - \frac{r(q)}{\phi(d)}.$$

Així, podrem aplicar el teorema de Bombieri-Vinogradov prenen

$$D_q := \frac{D(4)}{q} := \frac{N^{1/2}}{(\log N)^{B(4)}}$$

(per tal de tenir error de la forma $N/\log(N)^4$), on si voleu podeu prendre $B(4)=4+3=7$.

El teorema de Jurkat-Richert ens diu aleshores que

$$S(A_q,z) < (F(s_q)+O(\epsilon))|A_q|V(z) + R_q$$

on

$$s_q := \frac{\log D_q}{\log z}$$

i

$$R_q = \sum_{d<D_q, d|P(z)} |r_q(d)| \leq \sum_{d<D_q,d|P(z)} |r(qd)| + |r(q)| \sum_{d<D_q,d|P(z)} \frac{1}{\phi(d)}.$$

Abans de començar a estimar el error, observem que nosaltres volem calcular una fita superior de una suma de $S(A_q,z)$, i no d'un de sol. Així els posem tots junts i veiem que es el que volem calcular:

$$\sum_{z\leq q<y\,,q\in P_N} S(A_q,P(z)) < \sum_{z\leq q<y\,,q\in P_N} (F(s_q)+O(\epsilon))|A_q|V(z) + \sum_{z\leq q<y\,,q\in P_N} R_q$$

Primer el terme d'error:

$$\sum_{z\le q<y\,,q\in P_N} R_q \le$$

$$\le \sum_{z\le q<y\,,q\in P_N}\ \sum_{d<D/q,d|P(z)} r(qd) + \sum_{z\le q<y\,,q\in P_N} |r(q)| \sum_{d<D/q,d|P(z)} \frac{1}{\phi(d)} \le$$

$$\sum_{d'<D,d|P(z)} r(d') + \sum_{z\le q<y\,,q\in P_N} |r(q)| \sum_{d<N^{1/2}} \frac{1}{\phi(d)} \ll$$

$$\ll \frac{N}{(\log N)^4} + \frac{N}{(\log N)^4}\log N \ll \frac{N}{(\log N)^3}$$

on hem utilitzat Bombieri-Vinogradov a dues de les sumes i que $y < N^{1/3} < D$ en la segona si N és prou gran, ja que

$$D = \frac{N^{1/2}}{(\log N)^{B(4)}};$$

i hem utilitzat que (exercici!)

$$\sum_{d<N} \frac{1}{\phi(d)} \ll \log N.$$

Ara estimem el terme principal:

$$s_q := \frac{log(D/q)}{\log z} = 8\frac{N^{1/2}/q}{\log N} - 8(B)\frac{\log\log N}{\log N}$$

Com que $N^{1/8} = z \le q < y = N^{1/3}$, tenim que

$$\frac{4}{3} < 8\frac{N^{1/2}/q}{\log N} \le 3$$

i així $1 \le s_q \le 3$. Però justament per aquests valors de s sabem que el valor de $F(s)$ és

$$F(s_q) = \frac{2e^\gamma}{s_q} = \frac{e^\gamma \log(N)}{4\log(N^{1/2}/q)} + O(\frac{\log\log N}{\log N}).$$

D'altre banda

$$|A_q| = \pi(N;q,N)+O(\log N) = \frac{N}{\phi(q)\log N}\left(1+O(\frac{1}{\log N})\right)+\delta(N;q,N).$$

Així tenim que

$$\sum_{z\leq q<y\,,q\in P_N}(F(s_q)+O(\epsilon))|A_q| = \frac{e^\gamma N}{4}\sum_{z\leq q<y\,,q\in P_N}\frac{1}{\phi(q)\log(N^{1/2}/q)} + O(\frac{N}{\log N})$$

gràcies a aplicar el teorema de Bombieri-Vinagrodov de nou pels termes

$$\sum_{z\leq q<y\,,q\in P_N}\delta(N;q,N) = O(\frac{N}{(\log N)^3})$$

(i a altres termes que es pot veure queden dins el terme d'error).

Ara ens cal calcular

$$\sum_{z\leq q<y\,,q\in P_N}\frac{1}{\phi(q)\log(N^{1/2}/q)} < \sum_{z\leq q<y}\frac{1}{q\log(N^{1/2}/q)} + O(\frac{1}{z\log N})$$

que és un càlcul bastant estàndard, utilitzant que

$$\sum_{p<t}\frac{1}{p} = \log\log t + B + O(\frac{1}{t}).$$

per una certa constant B. En efecte tenim que (l'error es pot ignorar)

$$\sum_{z\leq q<y}\frac{1}{q\log(N^{1/2}/q)} \sim \int_z^y \frac{d\log\log t}{log(N^{1/2}/t)} = \int_z^y \frac{dt}{t\log(t)\log(N^{1/2}/t)} =$$

i fent el canvi $t = N^\alpha$ és igual a

$$= \frac{1}{\log N}\int_{1/8}^{1/3}\frac{d\alpha}{\alpha(\frac{1}{2}-\alpha)} = \frac{2\log 6}{\log N}.$$

En resum:

$$\sum_{z\leq q<y\,,q\in P_N}(F(s_q)+O(\epsilon))|A_q|V(z) + O(\frac{N}{(\log N)^3}) < \frac{e^\gamma\log 6}{2}\frac{NV(z)}{\log N}$$

Com a conseqüència tenim així que

6.9.2 Corol·lari. *Si N és un nombre parell prou gran, aleshores N es pot expressar com la suma d'un nombre primer més un nombre que és com a molt producte de 3 primers (i.e. de $\mathbb{P}_3$). Concretament tenim que*

$$|\{p \in \mathbb{P} \mid p \leq N\ ,\ N-p \in \mathbb{P}_3\}| >$$

$$> S(A, P(N^{1/8})) - \frac{1}{2} \sum_{N^{1/8} \leq q < N^{1/3}\ , q \in P_N} S(A_q, P(N^{1/8})) >$$

$$\left(\frac{e^\gamma \log 3}{2} - \frac{1}{2}\frac{e^\gamma \log 6}{2} + \epsilon \right) \frac{NV(N^{1/8})}{\log(N)} >$$

$$> 8 \left(\frac{\log 3}{2} - \frac{\log 6}{4} + \epsilon \right) \mathcal{S}(N) \frac{N}{\log(N)^2} >$$

$$> 0.811 \mathcal{S}(N) \frac{N}{\log(N)^2}$$

6.10 La mida de B

6.10.1 Lema.

$$|B| = b\frac{N}{\log(N)} + O\left(\frac{N}{(\log(N))^2}\right)$$

on

$$\begin{aligned} b &:= \iint\limits_{\substack{\frac{1}{8} < \alpha < \frac{1}{3} < \beta \\ \alpha + 2\beta < 1}} (\alpha\beta)^{-1}(1-\alpha-\beta)^{-1} d\alpha d\beta = \int_{\frac{1}{8}}^{\frac{1}{3}} \frac{\log(2-3\alpha)}{\alpha(1-\alpha)} d\alpha \\ &= 0.3631 \end{aligned}$$

Demostració: Recordem que

$$B := \left\{ N - p_1p_2p_3 \mid \begin{array}{c} N^{1/8} \leq p_1 < N^{1/3} \leq p_2 \leq p_3\ , \\ \\ p_1p_2p_3 < N\ , (p_1p_2p_3, N) = 1 \end{array} \right\}$$

Denotem $z = N^{1/8}$ i $y = N^{1/3}$. Així tenim que

$$|B| \leq \sum_{\substack{z \leq p_1 < y \leq p_2 \leq p_3 \\ p_1 p_2 p_3 < N}} 1 \leq \sum_{\substack{z \leq p_1 < y \leq p_2 \\ p_1 p_2^2 < N}} \pi\left(\frac{N}{p_1 p_2}\right)$$

Aplicant el teorema dels nombres primers tenim que aquesta última suma esta fitada per (amb un cert error, que ignoraré)

$$< \sum_{\substack{z \leq p_1 < y \leq p_2 \\ p_1 p_2^2 < N}} \frac{N}{p_1 p_2} \log\left(\frac{N}{p_1 p_2}\right)^{-1}$$

$$= N \sum_{z \leq p_1 < y} \frac{1}{p_1} \sum_{y \leq p_2 < (N/p_1)^{1/2}} \frac{1}{p_2 \log(\frac{N}{p_1 p_2})}.$$

Ara el que cal es estimar aquestes sumes utilitzant que

$$\sum_{p<t} \frac{1}{p} = \log\log t + B + O(\frac{1}{t}).$$

per una certa constant B. Per exemple, per a la suma interior tenim que

$$\sum_{y \leq p_2 < (N/p_1)^{1/2}} \frac{1}{p_2 \log(\frac{N}{p_1 p_2})} \sim \int_y^{(N/p_1)^{1/2}} \frac{1}{\log(\frac{N}{p_1 t})} d\log\log t$$

I substituint això en la suma global tenim que és aproximadament igual a

$$\sim N \int_{N^{1/8}}^{N^{1/3}} \int_y^{(N/u)^{1/2}} \frac{1}{\log(\frac{N}{ut})} d\log\log t d\log\log u$$

Finalment el canvi $t = N^\alpha$ i $u = N^\beta$ ens dóna que aquesta última integral és igual a

$$= \frac{N}{\log N} \int_{\frac{1}{8}}^{\frac{1}{3}} \int_{\frac{1}{3}}^{\frac{1-\beta}{2}} \frac{d\alpha d\beta}{\alpha\beta(1-\alpha-\beta)}.$$

□

6.11 Fitació superior de $S(B, P(y))$

Bé, per acabar de demostrar el teorema el que volem es trobar una fitació per sobre del terme $S(B, P(y))$ que no sigui gaire gran, de manera que juntament amb els altres encara ens surti un nombre positiu. El que veurem és el següent resultat:

6.11.1 Teorema.

$$S(B, P(y)) < \left(\frac{be^\gamma}{2} + \epsilon\right) \frac{NV(z)}{\log N} + O(\frac{N}{(\log N)^3})$$

on b és la constant del lema anterior, $b = 0.3631$.

De fet, per a demostrar això només ens cal veure que:

$$S(B, P(y)) < \left(\frac{e^\gamma}{2} + \epsilon\right) |B|V(z) + O(\frac{N}{(\log N)^3})$$

ja que abans hem calculat el cardinal de B.

El terme principal és fàcil: Prenem per un cert A

$$D = \frac{N^{1/2}}{(\log N)^A}$$

i tenim

$$S(B, P(y)) < (F(s) + \epsilon)|B|V(y) + Error$$

on

$$s = \frac{\log D}{\log y} = \frac{3}{2} + O(\frac{\log\log N}{\log N}) \in [1, 3]$$

ja que $y = N^{1/3}$. Així

$$F(s) = \frac{4e^\gamma}{3} + O(\frac{\log\log N}{\log N})$$

i per tant

$$S(B, P(y)) < \left(\frac{4e^\gamma}{3} + \epsilon\right) |B|V(y) + Error = \left(\frac{e^\gamma}{2} + \epsilon\right) |B|V(z) + Error$$

utilitzant que

$$V(y) = V(z) \left(\frac{\log z}{\log y} + O(\frac{1}{\log N})\right) = V(z) \left(\frac{3}{8} + O(\frac{1}{\log N})\right)$$

(i.e.

$$\frac{4e^\gamma}{3}\frac{3}{8} = \frac{e^\gamma}{2}).$$

El problema és la fitació del error; i és que no podem aplicar el teorema de Bombieri-Vinogradov. Veure que el terme d'error és el que toca ens portaria massa lluny i ho deixarem per una altra ocasió. Només us diré que s'ha d'aplicar en un moment donat la teoria del gran Garbell, que s'ha de fer un estudi delicat utilitzant certs caràcters de Dirichlet, etc. Si ho voleu veure podeu mirar o bé el llibre [H-R] o be el llibre [Nat].

6.12 El pas final

Posem tot el que tenim junt:

$$|\{p \in \mathbb{P} \mid p \leq N\ ,\ N-p \in \mathbb{P}_2\}| >$$

$$S(A,P(z)) - \frac{1}{2}\sum_{z\leq p_1<y\,,p_1\in P_N} S(A_{p_1},z)) - \frac{1}{2}S(B,P(y)) >$$

$$\left(\frac{e^\gamma \log(3)}{2} - \frac{1}{2}\frac{e^\gamma \log 6}{2} - \frac{1}{2}\frac{be^\gamma}{2} + \epsilon\right)\frac{N}{\log(N)}V(z) =$$

$$= 8\left(\frac{\log(3)}{2} - \frac{\log 6}{4} - \frac{b}{4} + \epsilon\right)\mathcal{S}(N)\frac{N}{\log(N)^2}$$

$$> 0.0848\mathcal{S}(N)\frac{N}{\log(N)^2}$$

si N és gran. I per tant hem demostrat el teorema.

X. Xarles
Departament de Matemàtiques
Edifici C,
Universitat Autònoma de Barcelona
08193 Bellaterra, Barcelona,
xarles@mat.uab.es

Capítol 7

Término de error en la criba lineal.

Jorge Jiménez.

7.1 Preliminares.

Por el Teorema de Jurkat-Richert, Teorema 4.5.1 del Capítulo 4, sabemos que

$$XV(z)(f(s)+O(\Delta))+R^- \leq S(A,z) \leq XV(z)(F(s)+O(\Delta))+R^+,$$

donde $R^\pm = R(A,P,\Lambda^\pm)$ es el término de error, para cierto Δ explícito en términos de las hipótesis de dimensión de criba y el nivel. De hecho sabemos que las funciones $f(s)$ y $F(s)$ son óptimas debido a los ejemplos extremales de Selberg. Sin embargo, todavía se puede mejorar el resultado agrandando D, lo que nos permitirá hacer mayor z a la vez que mantenemos $s > 2$. Recuérdese que a mayor z, menor es el número de factores primos en los elementos contados en $S(A,z)$. Este esquema se puede llevar a cabo siempre y cuando sepamos controlar el error, $R(A,P,\Lambda^\pm)$ para D grandes.

En general para estimar el error se utilizan cotas del estilo

$$R(A,P,\Lambda) \leq \sum_{d\leq D} \gamma^{\nu(d)}|r(A,d)|,$$

para alguna constante γ. Es claro que este tipo de fórmulas no nos permite tomar D mayor que X, ya que es imposible controlar individualmente el término de error $r(A,d)$ para enteros mas grandes que el número de elementos del conjunto. Obsérvese que, en ese caso, en general no habrá múltiplos de d en el conjunto, o como mucho habrá uno lo cual haría $r(A,d) = 0$ ó 1 tan grande como el término principal. Por tanto necesitamos dar cotas directamente de $R(A,P,\Lambda) = \sum_{d\leq D}\lambda(d)r(A,d)$, midiendo la cancelación que se produce al sumar términos de signo, en principio, aleatorio. En este sentido es interesante observar como los ejemplos de Selberg producen sucesiones sesgadas, en cuyos errores individuales predomina un signo determinado. Concretamente, si A es la sucesión de enteros con un número par de factores primos, se tiene, para ciertos rangos de M y d en función de X,

$$2r(A,d) = \lambda(d)\sum_{m\leq M}\lambda(m) - \{X/d\}\,.$$

Eligiendo M tal que $\sum_{m\leq M}\lambda(m) \neq 0$, es fácil ver que el producto $\lambda(d)r(A,d)$ tiene signo constante para cualquier d que cumpla la desigualdad $\lambda(d)\sum_{m\leq M}\lambda(m) < 0$. En conjuntos generales este tipo de desviación en el signo no se produce, lo cual nos permitirá acotar mejor el error y así ganar en el tamaño de los niveles de criba aceptables. El estudio del signo será posible gracias al caracter multiplicativo del término de error. Concretamente probaremos que éste, para las funciones de criba lineal, se puede escribir, para cada pareja $MN = D$, como

$$R(A,P,\Lambda) \sim \sum_{\substack{m<M\\ n<N}} a_m b_n r(A,mn), \tag{7.1}$$

para algunos $|a_m| < 1, |b_n| < 1$. El significado explícito del simbolo $\sim$ aparecerá en el enunciado del Teorema 7.3.7 en la última página del capítulo.

Es interesante observar que, en el caso de la criba de Selberg, el término de error viene dado por el Teorema 2.1.1 que es del tipo (7.1) de manera inmediata. Sin embargo hay que señalar que, en ese caso, el método no cuenta con la flexibilidad que se obtiene de (7.1) en los parámetros M y N, ya que $M = N = \sqrt{D}$.

7.2 Métodos analíticos.

Una vez obtenida la fórmula (7.1), la estimación del error se obtiene gracias a métodos en Teoría Analítica de Números. En el caso del Teorema de Chen, (Ver Capítulo 6 y [H-R]), aparecen los polinomios de Dirichlet y las desigualdades de gran criba, vistas en el Capítulo 5. En el siguiente ejemplo se puede observar como entra en juego la transformada de Fourier y las estimaciones de sumas exponenciales. Sea $A = [Y - X, Y) \cap \mathbb{Z}$. Denotando por

$$\psi(t) = \begin{cases} 1/2 - \{t\} & \text{Si } t \text{ no es entero} \\ 0 & \text{resto,} \end{cases}$$

podemos escribir $|A_d|$ como

$$|A_d| = \frac{X}{d} + \psi(Y/d) - \psi((Y - X)/d).$$

Teniendo en cuenta la transformada de Fourier de la función ψ,

$$\psi(t) = \frac{1}{2\pi i} \sum_{h \neq 0} \frac{e(ht)}{h},$$

y la fórmula (7.1), el término de error quedará como suma de dos series del estilo

$$\sum_{\substack{m<M \\ n<N}} a_m b_n \sum_{h \neq 0} c_h e(hf(mn)).$$

Aproximando por una función suficientemente suave podemos cambiar el orden de sumación para obtener

$$\sum_{h \neq 0} |c_h| | \sum_{\substack{m<M \\ n<N}} a_m b_n e(hf(mn))|.$$

Aplicando a la última suma, $S(h,M,N) = \sum_{\substack{m<M \\ n<N}} a_m b_n e(hf(mn))$, la desigualdad de Cauchy-Schwartz se obtiene

$$|S(h,M,N)|^2 \leq \sum_{m<M} |a_m|^2 \sum_{m<M} |\sum_{n<N} b_n e(hf(mn))|^2.$$

La aportación principal vendrá del término diagonal $n_1 = n_2$ que se acota trivialmente por M^2N. Por tanto $|S(h,M,N)| \leq M\sqrt{N} < D$ con lo que se puede considerar un nivel mayor en función de X. Para acotar los términos no diagonales se utilizan lemas tipo Van der Corput o estimaciones de Weyl de sumas de Kloosterman.

Existen casos en los que ninguno de los dos métodos anteriores se puede aplicar. En este caso un tercer método, el método de dispersión de Linnik, puede ser que nos de los resultados deseados. Este método es el que utiliza Iwaniec en el artículo [I5] en el que estudia, entre otros, el conjunto $n^2 + 1$ y que será objeto del siguiente capítulo.

7.3 Forma bilineal del término de error.

En esta sección vamos a obtener la fórmula (7.1) para el término de error cometido con la función superior de criba de Rosser-Iwaniec en el caso lineal. La fórmula para la función inferior se obtiene de manera análoga.

Es necesario recordar que, en este caso, la función superior de criba viene dada como composición de dos cribas, una λ_1 de nivel D^ε para cribar los primos pequeños, concretamente sobre $P(w)$ tal que $\frac{\log D}{\log w} = \varepsilon^{-1} \log \varepsilon^{-1}$, y la otra λ^+, de nivel D, en $P(z,w) = \prod_{w \leq p < z} p$,

definida mediante la fórmula

$$\lambda^+(d) = \mu(d) \sum_{n\geq 1} \sum_{p_1\cdots p_n = d} \psi^+(p_1, \ldots, p_n), \tag{7.2}$$

donde ψ^+ vale 1 si

$$p_1 > \cdots > p_n \quad \text{y} \quad p_1 \cdots p_m p_m^2 < D \text{ para todo } m \text{ impar}, \tag{7.3}$$

y cero en otro caso. Es importante observar que la suma en (7.2) es en realidad una suma finita hasta $\nu = \varepsilon^{-1} \log \varepsilon^{-1}$ ya que $w^n < p_1 \cdots p_n < D$ con lo que $n \leq \frac{\log D}{\log w} = \nu$.

La forma bilineal aparece relacionada con la siguiente definición.

7.3.1 Definición. *Una función aritmética $f(d)$ se dice que es bien factorizable de nivel D si para cada pareja $M > 1$, $N > 1$ tal que $MN = D$, existen funciones $|g(m)| \leq 1$, $|h(n)| \leq 1$ con soporte en $[1, M]$ y $[1, N]$ respectivamente, y tal que $f(d) = \sum_{mn=d} g(m)h(n) = g * h$.*

En este sentido, nuestro objetivo es demostrar que $\lambda = \lambda_1 \lambda^+$ es una función bien factorizable. El siguiente lema es inmediato.

7.3.2 Lema. *Sea $f = g * h$ con h bien factorizable de nivel D y $|g(c)| \leq 1$, $g(c) = 0$ para todo $c > C$ y para algún $C \leq D$. Entonces f es bien factorizable de nivel CD.*

Asi pues, para obtener el resultado, es suficiente demostrar que la función λ^+ es bien factorizable. Que una función aritmética sea bien factorizable quiere decir que, en algún sentido, la función tiene un caracter multiplicativo, es decir, se comporta de manera independiente en cada uno de los factores primos. Asi pues, el primer paso que daremos para obtener una representación en el sentido de la definición será separar cada uno de los factores primos que aparecen en (7.2) en función de su tamaño. Para ello sea $\mathfrak{b} = (b_1, \ldots, b_n)$ tal que $b_i = 1, 2, 4, \ldots$ recorre, de manera independiente, las potencias de 2.

Entonces

$$\lambda^+(d) = \mu(d) \sum_{n \le \nu} \sum_{\mathfrak{b}} \sum_{\substack{p_1 \cdots p_n = d \\ b_i < p_i \le 2b_i}} \psi^+(p_1, \ldots, p_n).$$

Es fácil ver que (7.3) implica

$$b_1 \ge \cdots \ge b_n, \tag{7.4}$$

y

$$b_1 \cdots b_m b_m < D, \tag{7.5}$$

para todo m. Efectivamente si $b_2 > b_1$, entonces por ser potencias de 2 se tiene $p_2 > b_2 \ge 2b_1 \ge p_1$ lo cual es imposible. Por otro lado (7.4) es consecuencia directa de (7.3). Por tanto, cualquier vector de potencias de 2 que aparece en la suma pertenece al conjunto $\mathcal{B}^n = \{\mathfrak{b} : \text{ que cumplen (7.4) y (7.5)}\}$. Es fácil dar la cota $|\mathcal{B}^n| \le (\log D)^n$ para el conjunto teniendo en cuenta que $2^k = b_i < p_i < D$.

Que la función λ^+ sea bien factorizable es, de alguna manera, consecuencia del caracter multiplicativo del conjunto $\mathcal{B}^n$.

7.3.3 Lema. *Sea $D = D_1 D_2$ con $D_1 > 1$, $D_2 > 1$. Para todo $\mathfrak{b} \in \mathcal{B}^n$ existe una partición $\mathfrak{b} = (\mathfrak{b}_1, \mathfrak{b}_2)$ tal que $||\mathfrak{b}_i|| < D_i$, para $i = 1, 2$ donde $||\mathfrak{b}|| = b_1 \cdots b_n$.*

La demostración, por inducción, es consecuencia inmediata de (7.5), (ver [I6]).

Una vez separados cada factor, dividimos la condición conjunta (7.3), en condiciones independientes para cada primo gracias al siguiente método de separación de variables, que recogemos en forma de dos lemas.

7.3.4 Lema. *Sea $x \ge 1$. Existe una función $g(t)$ tal que*

$$\int_{-\infty}^{\infty} |g(t)| dt < \log 6x,$$

y tal que se cumple para todo entero positivo l

$$\int_{-\infty}^{\infty} g(t)l^{it} = \begin{cases} 1 & \text{si } l \leq x \\ 0 & \text{resto.} \end{cases}$$

Demostración. Sea $f(u) = \min\{u, 1, [x] + 1 - u\}$ si $u \in [0, [x] + 1]$ y $f(u) = 0$ en el resto, y $G(s)$ su transformada de Mellin. Entonces, integrando por partes

$$\begin{aligned} G(s) = & \int_0^\infty f(u)u^{s-1}du = \tfrac{1}{s}\int_0^1 u^s du - \tfrac{1}{s}\int_{[x]}^{[x+1]} u^s du \\ = & \quad \tfrac{1}{s(s+1)}\left(1 + [x]^{s+1} - [x+1]^{s+1}\right), \end{aligned}$$

de donde se obtiene

$$|G(s)| \leq \min\left\{1 + \log x, \frac{2}{|s|}, \frac{2(x+1)}{|s(s+1)|}\right\}.$$

El resultado se deduce integrando $|g(t)| = |G(it)|$ en cada uno de los intervalos $[0, 1), [1, x), [x, \infty)$ por separado.

7.3.5 Lema. *Sea* $y \geq 1$. *Existe una función* $h(t)$ *tal que*

$$\int_{-\infty}^{\infty} |h(t)|dt < \log 6y^2,$$

y tal que se cumple para todo par de enteros positivos $m, n \leq y$

$$\int_{-\infty}^{\infty} h(t)(m/n)^{it} = \begin{cases} 1 & \text{si } m \leq n \\ 0 & \text{resto.} \end{cases}$$

La demostración es análoga a la del Lema 7.3.4 teniendo en cuenta que cualquier pareja de racionales de denominador menor que y estan separados, como poco, por $1/y^2$, (ver [I6]).

Escogiendo $x = D$ en el Lema 7.3.4 e $y = \sqrt{D}$ en el Lema 7.3.5, podemos escribir

$$\psi^+(p_1, \ldots, p_n) = \prod_{\substack{1<m\leq n \\ m \text{ impar}}} \int_{-\infty}^{\infty} g(t)(p_1 \cdots p_m p_m^2)^{it}dt \prod_{1\leq l<n} \int_{-\infty}^{\infty} h(t)\left(\frac{p_{l+1}}{p_l}\right)^{it}.$$

Obsérvese que d es libre de cuadrados, con lo que se da la desigualdad estricta en el Lema 7.3.5. Basta ahora hacer un cambio de variable lineal para obtener

$$\psi^+ = \int_{\mathbb{R}^n} \varphi^+(t_1, \ldots, t_n) p_1^{it_1} \cdots p_n^{it_n} dt_1 \cdots dt_n,$$

para alguna φ^+ tal que

$$\int_{\mathbb{R}R^n} |\varphi^+| d\mathfrak{t} < (\log D)^{2n},$$

teniendo en cuenta los Lemas 7.3.4 y 7.3.5, y el hecho de que aparecen menos de $\frac{3}{2}n$ integrales en ψ^+. De esta forma se tiene

$$\lambda^+(d) = \sum_{n \leq \nu} \sum_{\mathcal{B}^n} \int_{\mathbb{R}} \varphi^+(\mathfrak{t}) \lambda_d(\mathfrak{b}, \mathfrak{t}) d\mathfrak{t},$$

con

$$\lambda_d(\mathfrak{b}, \mathfrak{t}) = \mu(d) \sum_{\substack{p_1 \cdots p_n = d \\ \mathfrak{b} < (p_1, \ldots, p_n) \leq 2\mathfrak{b}}} p_1^{it_1} \cdots p_n^{it_n}.$$

Es fácil ver que $\lambda_d(\mathfrak{b}, \mathfrak{t}) \leq n!$ y $\lambda_d(\mathfrak{b}, \mathfrak{t}) = 0$ si $d > 2^n ||\mathfrak{b}||$. Ambas propiedades nos permiten demostrar el siguiente resultado.

7.3.6 Lema. *Para todo $\mathfrak{b} \in \mathcal{B}^n$, $\mathfrak{t} \in \mathbb{R}^n$, $\frac{1}{n!}\lambda_d(\mathfrak{b}, \mathfrak{t})$ es una función bien factorizable de nivel $4^n D$.*

Demostración. Supongamos $4^n D = MN$ con $M \geq N > 1$ y supongamos, en primer lugar, $N > 2^n$. Sean $D_1 = 2^{-n}M$, $D_2 = 2^{-n}N$. Por el Lema 7.3.3 sabemos que $\mathfrak{b} = (\mathfrak{b}_1, \mathfrak{b}_2)$ tal que $||\mathfrak{b}_i|| < D_i$. Para este par de vectores es fácil ver que

$$\frac{1}{n!}\lambda_d(\mathfrak{b}, \mathfrak{t}) = \sum_{d_1 d_2 = d} \frac{1}{l!}\lambda_{d_1}(\mathfrak{b}_1, \mathfrak{t}_1) \frac{l!}{n!}\lambda_{d_2}(\mathfrak{b}_2, \mathfrak{t}_2),$$

ya que, a cada orden de $p_1 \cdots p_n = d$ contado en $\lambda_d(\mathfrak{b}, \mathfrak{t})$, le corresponden dos divisores d_1, d_2 de d, y viceversa. Por otro lado, si $\mathfrak{b}_1$ tiene l componentes, entonces $\frac{1}{l!}\lambda_{d_1}(\mathfrak{b}_1, \mathfrak{t}_1)| \leq 1$ por definición, mientras que $|\frac{l!}{n!}\lambda_{d_2}(\mathfrak{b}_2, \mathfrak{t}_2)| = \left|\frac{\lambda_{d_2}(\mathfrak{b}_2, \mathfrak{t}_2)}{\binom{n}{l}(n-l)!}\right| \leq 1/\binom{n}{l} \leq 1.$

Por último para todo $d_1 > M = 2^n D_1 > 2^n ||b_1||\ \lambda_{d_1} = 0$ por definición, siendo análogo el caso $d_2 > N$, lo cual termina la demostración en el caso $N > 2^n$. Si $N \leq 2^n$, entonces $M \geq 2^n D$ y la convolución trivial

$$\frac{1}{n!}\lambda_d(\mathfrak{b},\mathfrak{t}) = \sum_{d_1 d_2 = d} \frac{1}{n!}\lambda_d(\mathfrak{b},\mathfrak{t}) \sum_{k|d_2} \mu(k),$$

demuestra el resultado.

Asi pues, los Lemas 7.3.2 y 7.3.6 nos permiten escribir el término de error como

$$R(A,P,\Lambda^+) = \sum_{c|P(w)} \lambda_1(c) \sum_{d|P(z,w)} \lambda(d)^+ r(A,cd) = \sum_{n\leq\nu} \sum_{\mathcal{B}^n} \int_{\mathbb{R}^n} \varphi^+(\mathfrak{t}) E(\mathfrak{b},\mathfrak{t}) d\mathfrak{t},$$

donde

$$E(\mathfrak{b},\mathfrak{t}) = \nu! \sum_{d|P(z)} \rho(d) r(A,d),$$

para alguna ρ bien factorizable de nivel $4^\nu D^{1+\varepsilon}$. Por tanto, teniendo en cuenta que $|\lambda_d(\mathfrak{b},\mathfrak{t})| \leq |\lambda_d(\mathfrak{b},0)|$, podemos escoger la peor de las ρ involucradas en la fórmula para obtener

7.3.7 Teorema. *Sea $\varepsilon \leq e^{-10}$, $\nu = \varepsilon^{-1}\log\varepsilon^{-1}$. Para cada pareja $MN = D$ se tiene*

$$|R(A,P,\Lambda)| \leq (\nu \log D)^{3\nu} \Big| \sum_{\substack{m<M, n<N \\ mn|P}} a_m b_n r(A,mn)\Big|,$$

para algunos $|a_m| \leq 1$, $|b_n| \leq 1$.

J. Jiménez Urroz
Departamento de Matemática Aplicada IV
Edifici C-3, Campus Nord
Universitat Politècnica de Catalunya
08034, Barcelona,
jjimenez@ma4.upc.edu

Capítol 8

Casiprimos representados por polinomios cuadráticos.

Adrián Ubis

8.1 Introducción

El teorema de Dirichlet asegura que cualquier polinomio de grado uno con coeficientes en $\mathbb{Z}$ coprimos representa infinitos primos. Nada similar se conoce para un polinomio de grado superior. Si G es un polinomio con coeficientes en $\mathbb{Z}$ irreducible sobre $\mathbb{Q}$ y sobre $\mathbb{F}_p$ para todo primo p, argumentando de manera probabilística podríamos conjeturar que, (ver [Sc-Si] y [Ba-Ho]),

$$|\{n \leq x : G(n) \text{ primo }\}| \sim \Gamma_G \frac{x}{\log x}$$

donde $\Gamma_G = (\deg G)^{-1} \prod_p (1 - \rho(p)/p)(1 - 1/p)^{-1}$, $\rho(p)$ el número de soluciones de la ecuación $G(n) = 0$ en $\mathbb{F}_p$.

Una posible aproximación al problema es ver que $G(n)$ toma infinitos valores en P_r (conjunto de números con a lo más r factores primos). En esta dirección vamos a ver el siguiente resultado [I5]

8.1.1 Teorema. *Sea $G(n) = an^2 + bn + c$ un polinomio irreducible sobre $\mathbb{Q}$ y sobre $\mathbb{F}_p$ para todo p primo, con $a > 0$. Entonces, para x suficientemente grande se cumple que*

$$|\{n \leq x : G(n) \in P_2\}| > \frac{1}{77}\Gamma_G \frac{x}{\log x}.$$

Por simplicidad vamos a realizar la exposición para el caso $G(n) = n^2 + 1$. Estamos en el contexto de la criba lineal, de hecho $\rho(q) = 1 + \chi_4(q)$ (donde χ_4 es el carácter no trivial módulo 4) para q primo y en general $\rho(n) = \prod_{p|n}(1 + \chi_4(p)) = \sum_{d|n} \chi_4(d)\mu^2(d)$ $(4 \nmid n)$, $\rho(4n) = 0$. La cota trivial $\rho(n) \leq r(n)$ (donde $r(n)$ es el número de puntos de coordenadas enteras en la circunferencia centrada en el cero de radio $n^{1/2}$) muestra que un nivel de criba posible es $D = o(x/\log x)$. De hecho podemos ver que manteniendo el error habitual $R(A; D)$, el máximo nivel posible es $D = x$, porque si $D > x$

$$\sum_{d \asymp D} |r(A, d)| \geq x \sum_{d \asymp D} \rho(d) d^{-1} \gg x.$$

Esto nos permite probar el Teorema 8.1.1 para P_4 directamente y para P_3 usando pesos (ver sección 4), pero el caso P_2 es inalcanzable.

8.2 Forma bilineal para el término de error

La forma usual de tratar el término de error

$$\sum_{d<D, d|P(z)} \varrho_d r(A, d) \tag{8.1}$$

en los métodos de criba era simplemente poniendo valores absolutos, es decir acotarla por

$$\sum_{d<D, d|P(z)} |\varrho_d||r(A, d)|.$$

Esto era debido a que, aun teniendo cierto control sobre $r(A, d)$, los coeficientes ϱ_d van a depender de la función μ de Möbius la cual no sabemos controlar. El primer intento de superar estas dificultades al

menos en promedio se debe a C. Hooley [Ho1], [Ho2]. La idea es que si estamos cribando sobre varios conjuntos $A_1, A_2, \dots A_l$ a la vez, el término de error conjunto será

$$\sum_d \varrho_d \sum_k r(A_k, d).$$

De esta forma, si tenemos cierta información sobre los A_k podemos obtener cancelación en la suma interior sin necesitar conocimiento de ϱ_d. Así, considerando como A_k la progresión aritmética de módulo k y resto a fijo, Hooley mejoró la constante en el teorema de Brun-Tichsmarsh para casi todo $x^{1/2} \leq k \leq x^{1-\varepsilon}$ (usando la criba de Selberg).

Pero el primer matemático en aprovechar la estructura específica de los coeficientes ρ_d fue Y. Motohashi [Mo], mejorando la constante del teorema de Brun-Tichmarsh esta vez para todo módulo k en el rango $1 \leq q \leq x^{3/7}$. De forma explícita, lo que hizo fue expresar el término de error en la criba de Selberg mediante funciones generatrices, consiguiendo cancelación a través de desigualdades de tipo gran criba para sumas con caracteres.

Lo que Iwaniec supo ver es que la información usada por Motohashi sobre $\varrho(d)$ era sólo el hecho de que (8.1) en la criba de Selberg se escribe de forma natural como

$$\sum_{\substack{d_1 \leq D^{1/2} \\ d_1 | P(z)}} \sum_{\substack{d_1 \leq D^{1/2} \\ d_1 | P(z)}} \lambda_{d_1} \lambda_{d_2} r(A, [d_1, d_2]),$$

es decir que tenemos una separación multiplicativa de d en variables d_1 y d_2. Motohashi no usaba ninguna información sobre λ_d (aparte de la cota $\lambda_d \ll 1$).

Iwaniec intentó ver si esta separación de variables era posible en la criba de Rosser. En el caso lineal, probó (ver Capítol 7) que no sólo era posible, sino que se podía hacer de una forma más general. En otras palabras, demostró el siguiente lema fundamental de criba:

8.2.1 Lema. *Para todo $\epsilon > 0$,$M, N \geq 2$ y $MN = D$ suficientemente grande con respecto a ϵ se cumplen las desigualdades*

$$XV(z)(f(s) - \epsilon) - E \leq S(A, z) \leq XV(z)(F(s) + \epsilon) + E$$

donde $E \ll R(A; M, N)$ la constante implícita dependiente sólo de s y de las constantes que definen la densidad de criba, y

$$R(A; M, N) = \sum_{m<M,\, n<N,\, mn|P(z)} a_m b_n r(A, mn) \tag{8.2}$$

a_n, b_n acotados por 1 en valor absoluto y dependientes sólo de M, N, z y ϵ (no dependen de A).

El error en la criba de Selberg correspondería a $M = N = D^{1/2}$. Que en el resultado anterior podamos escoger M, N con mayor flexibilidad es un factor a favor de la criba de Rosser. Esto explica los resultados positivos obtenidos por Iwaniec.

La forma del error (8.2) nos permite deshacernos de la dependencia de una de las variables y estimar entonces

$$\sum_{m<M} B(A; m, N), \tag{8.3}$$

donde $B(A; m, N) = |\sum_{n<N} b_n r(A, mn)|$. Por tanto, si hay cierta independencia en los $M - 1$ vectores $(r(A, mn))_{n<N}$ vamos a poder conseguir cancelación sea cual sea el comportamiento de b_n, porque $B(A; m, N)$ no puede ser grande para muchos valores de m (idea de gran criba).

8.3 Método de Dispersión

Esta nueva forma para el término de error se puede aprovechar en la práctica de diversas maneras, dependiendo del problema a tratar. En nuestro caso para M pequeña usaremos la estimación trivial $|r(A, mn)| \leq \rho(mn) \leq r(mn)$ y por tanto

$$\sum_{m<M} B(A; m, N) \ll MN, \tag{8.4}$$

(En este paso no es necesaria la separación de variables). Cuando M es grande aplicamos Cauchy-Schwarz (después de dividir en intervalos diádicos) obteniendo la cota $M^{1/2}\mathfrak{D}^{1/2}$, con

$$\mathfrak{D} = \sum_{M<m<2M} B(A; m, N)^2$$

midiendo la dispersión (varianza) del error.

La idea para tratar $\mathfrak{D}$ es escribir

$$r(A, mn) = \sum_{\substack{k<x \\ k^2+1\equiv 0 \pmod{mn}}} 1 - x\frac{\rho(m)\rho(n)}{mn} \tag{8.5}$$

para después expandir el cuadrado y estimar cada término promediando sólo en la variable m, usando información sobre la distribución de A. Nosotros vamos a simplificar ese promedio con un paso intermedio: Podemos reescribir (8.5) como

$$r(A, mn) = \sum_{\substack{0\le v<m \\ v^2+1\equiv 0 \pmod m}} \Big(\sum_{\substack{k<x, k\equiv v \pmod m \\ k^2+1\equiv 0 \pmod n}} 1 - \frac{x}{m}\frac{\rho(n)}{n}\Big)$$

y así $B(A; m, N) = \sum_{v^2+1\equiv 0 \pmod m} S_v$, donde

$$S_v = \sum_{n<N} b_n \sum_{\substack{k<x, k\equiv v \pmod m \\ k^2+1\equiv 0 \pmod n}} 1 - \frac{x}{m}\sum_{n<N} b_n \frac{\rho(n)}{n}.$$

Por Cauchy-Schwarz,

$$B(A; m, N)^2 \le \rho(m) \sum_{v^2+1\equiv 0 \pmod m} S_v^2$$

y como $\rho(m) \ll m^{\varepsilon}$, vemos que

$$\mathfrak{D} M^{-\epsilon} \ll \sum_{M<m<2M} \sum_{v^2+1\equiv 0 \pmod m} S_v^2. \tag{8.6}$$

Ahora sí, expandiendo el cuadrado y reordenando podemos escribir la expresión de la derecha en (8.6) como

$$\sum_{n_1, n_2<N} b_{n_1} b_{n_2} (W - 2xV + x^2 U), \tag{8.7}$$

donde W, U, V dependen de n_1, n_2, A y M, pero no de los coeficientes b_n. Vamos a tratar por separado las expresiones W, U, V, consiguiendo para cada una de ellas que

$$W \sim xV \sim x^2 U \sim \frac{3}{2\pi}\frac{x^2}{M}\frac{\rho(n_1 n_2)}{n_1 n_2} a(n_1 n_2), \tag{8.8}$$

donde $a(n) = (2,n)\prod_{p|n}\frac{p-1}{p+1}$. Así los términos principales se cancelan en (8.7) y si el error en las estimaciones (8.8) es E_{n_1,n_2} la cota para (8.7) vendrá dada por

$$\sum_{n_1,n_2\leq N} b_{n_1}b_{n_2}E_{n_1,n_2} \ll \sum_{n_1,n_2} E_{n_1,n_2}. \tag{8.9}$$

En este proceso no hemos usado conocimiento alguno de los coeficientes b_n. La clave ha sido que la separación de variables nos ha permitido eliminar los coeficientes a_m y así promediar en la variable m, consiguiendo un resultado positivo si tenemos cierto conocimiento aritmético del conjunto A. Por esa misma razón, para que esto sea útil deberemos tomar M grande, lo que da sentido al hecho de que Iwaniec usara la criba de Rosser en vez de la de Selberg (al final cogeremos $M = D^{15/16}$). Como E_{n_1,n_2} va a ser

$$\ll (n_1n_2)^{\varepsilon}x + (n_1n_2M)^{3/4+\varepsilon}(xM^{-1})^2$$

entonces por (8.9) habremos conseguido deducir que

$$\mathfrak{D} \ll (1 + N^{7/2}M^{-5/4}x)x^{1+\varepsilon}. \tag{8.10}$$

Tomando $N = x^{1/15-\epsilon}$, aplicando (8.4) cuando $M \leq x^{14/15-\epsilon}$ y (8.10) cuando $x^{14/15-\epsilon} < M < x^{1-4\epsilon}$ deducimos que:

8.3.1 Lema.

$$\sum_{m<x^{1-4\varepsilon}} B(A;m,x^{1/15-\varepsilon}) \ll x^{1-\varepsilon}.$$

Notar que este resultado nos permite de forma directa elevar el nivel de criba hasta $D = x^{16/15-5\epsilon}$. Así, por el Lema 8.2.1 tendríamos que

$$S(A,x^{8/15-5\epsilon}) \gg xV(x^{8/15-5\epsilon})f(2+\epsilon) + O(x^{1-\epsilon}) \gg x(\log x)^{-1}$$

y como $4 \times \frac{8}{15} > 2$ probamos que hay infinitos números n^2+1 con a lo más 3 factores primos. Para llegar al caso P_2 deberemos usar pesos (ver sección 8.5).

8.4 Distribución uniforme

El tratamiento de los términos U, V y W aumenta de dificultad de U a W. Esto es porque en U sólo tenemos que controlar la función ρ en media, en V un factor proviene de ρ y otro del número de soluciones de ecuaciones en congruencias, y en W los dos factores son de este último tipo. Pero aun en el peor caso, tras unos cambios de variable y estimaciones del tipo

$$[x] = x + O(1)$$

podemos ver que todo se reduce de forma natural al siguiente resultado sobre la distribución uniforme de las soluciones de la ecuación $\Omega^2 + 1 \equiv 0 \pmod{mq}$:

8.4.1 Lema. *Sea $0 < \alpha < \beta < 1$, $M < M_1 < 2M$, q libre de cuadrados. Entonces para todo $\varepsilon > 0$ se cumple que*

$$\sum_{\substack{M<m<M_1\\(m,q)=1}} \rho_{\alpha,\beta}(mq) = \frac{3}{2\pi}(M_1-M)(\beta-\alpha)\rho(q)(2,q)\prod_{p|q}\frac{p-1}{p+1}+O((qM)^{\frac{3}{4}+\varepsilon}),$$

donde $\rho_{\alpha,\beta}(n)$ es el número de soluciones de la ecuación $\Omega^2 + 1 \equiv 0 \pmod{n}$ con $\alpha n < \Omega < \beta n$.

Nota: En realidad en el caso de W hay que imponer la condición adicional $\Omega \equiv w(\mod d)$ para un divisor d de q. Pero la prueba y el resultado son similares.

En [Ho3] Hooley ya necesitó este tipo de resultado. Usando Análisis de Fourier todo el problema se concentra en estudiar sumas trigonométricas de la forma

$$\sum_{\Omega^2+1\equiv 0 \pmod{D}} e(\Omega/D), \tag{8.11}$$

Ω recorriendo una progresión aritmética cualquiera. Para tratarla podemos transformar la suma mediante el siguiente resultado de Lagrange:

8.4.2 Lema. *Sea $D \in \mathbb{N}$. La aplicación*

$$\{(r,s) : r^2+s^2 = D, (r,s) = 1, |r| < s\} \longrightarrow \{0 \le \Omega < D : \Omega^2+1 \equiv 0 \pmod{D}\}$$

definida por

$$\Omega = \frac{\overline{r}}{s}(r^2 + s^2) - \frac{r}{s}$$

es una biyección ($\overline{r}$ es el inverso de r módulo s).

Este lema nos permite escribir

$$\frac{\Omega}{D} = \frac{\overline{r}}{s} + \epsilon$$

donde $\epsilon = \epsilon(r, s)$ es una cantidad muy pequeña, y por tanto vamos a poder estimar (8.11) en función de sumas de Kloosterman. Para estas sumas Hooley obtuvo el siguiente lema a partir de la conjetura de Riemann para curvas algebraicas sobre cuerpos finitos:

8.4.3 Lema. *Si h y s son enteros y $0 < r_2 - r_1 < 2s$ entonces*

$$\sum_{\substack{r_1<r<r_2,(r,s)=1\\ r\equiv\lambda \pmod{\Lambda}}} e(h\frac{\overline{r}}{s}) \ll s^{\frac{1}{2}+\epsilon}(h, s)^{1/2}.$$

Con estas mismas herramientas vamos a probar el resultado de distribución que necesitamos

Demostración del Lema 8.4.1: Si $\alpha = 0, \beta = 1$ entonces $\rho_{\alpha,\beta} = \rho$ y por la fórmula $\rho(n) = \sum_{ab=m, b\not\equiv 0 \pmod 4} \chi_4(a)\mu^2(a)$ deduciríamos por el truco de la hipérbola que

$$\begin{aligned}
\sum_{m<x,(m,q)=1} \rho(m) &= \sum_{\substack{a\le x^{1/2}\\ (a,q)=1}} \chi_4(a)\mu^2(a) \sum_{\substack{b\le \frac{x}{a},(b,q)=1\\ b\not\equiv 0 \pmod 4}} 1 \\
&+ \sum_{\substack{b\le x^{1/2},(b,q)=1\\ b\not\equiv 0 \pmod 4}} \sum_{\substack{x^{1/2}<a\le \frac{x}{b}\\ (a,q)=1}} \chi_4(a)\mu^2(a) \\
&= \frac{1}{2}(1 + \frac{(2,q)}{2})\frac{\varphi(q)}{q}L(1)x + O(x^{3/4+\varepsilon})
\end{aligned}$$

con

$$L(s) = \sum_{(a,q)=1} \chi_4(a)\mu^2(a)a^{-s}$$

Usando el producto de Euler vemos que

$$L(s) = L(s, \chi_4)\zeta(2s)^{-1}(1 - 2^{-2s})^{-1} \prod_{p|q} (1 + \chi_4(p)p^{-s})^{-1}$$

Como $\zeta(2) = \pi^2/6$ y $L(1, \chi_4) = \pi/4$ obtenemos

$$\sum_{m<x,(m,q)=1} \rho(mq) = \frac{3}{2\pi} x\rho(q)(2, q) \prod_{p|q} \frac{p-1}{p+1} + O(q^{\varepsilon} x^{3/4+\varepsilon})$$

En el caso $0 < \alpha < \beta < 1$ el método es más elaborado. La forma de proceder es aproximar $\psi(t)$ (la función indicatriz del intervalo (α, β)) por funciones suaves, de tal forma que sepamos tratar mejor el error. Explícitamente dado $\Delta < \alpha < \beta < 1 - \Delta$ existen funciones $A(t), B(t)$ tales que

$$|\psi(t) - A(t)| = B(t), \tag{8.12}$$

y cuyas series de Fourier son de la forma $A(t) = \beta - \alpha + \sum_{h\neq 0} A_h e(ht)$, $B(t) = \Delta + \sum_{h\neq 0} B_h e(ht)$ con $|A_h|, |B_h| \leq \min(|h|^{-1}, \Delta^{-2}|h|^{-3}) = C_h$. (Para construirlas tomamos $\eta \in C^{\infty}(\mathbb{R})$, función par y positiva con $\eta(0) = 1/2$ y soporte de η igual a $[-\Delta, \Delta]$. Definimos $B(t) = \eta(t - \alpha) + \eta(t - \beta)$ y $A(t)$ la única función en $C^{\infty}(\mathbb{R})$ cumpliendo (8.12)).

Como $\rho_{\alpha,\beta}(mq) = \sum_{\Omega^2+1\equiv 0 \pmod{mq}} \psi(\Omega/mq)$, por (8.12)

$$\sum_{M<m<M_1,(m,q)=1} \rho_{\alpha,\beta}(mq) = (\beta - \alpha)\rho(q) \sum_{M<m<M_1,(m,q)=1} \rho(m) +$$

$$O(\rho(q)\Delta M + \sum_{h\neq 0} C_h | \sum_{\substack{M<m<M_1,(m,q)=1 \\ \Omega^2+1\equiv 0 \pmod{mq}}} e(h\Omega/mq)|).$$

Sólo nos queda ocuparnos del error. Primero eliminamos la condición $(m, q) = 1$

$$\sum_{\substack{m,\Omega \\ (m,q)=1}} e(h\Omega/mq) = \sum_{d|q} \mu(d) \sum_{\substack{qM<D<qM_1, D\equiv 0 \pmod{qd} \\ 0\leq\Omega<D, \Omega^2+1\equiv 0 \pmod{D}}} e(h\Omega/D).$$

Una vez aquí usamos los lemas 8.4.2 y 8.4.3 en la suma interior:

$$\begin{aligned}
\sum_{D,\Omega} e(h\Omega/D) &\leq \sum_{\substack{qM<r^2+s^2<qM_1,|r|<s \\ (r,s)=1, r^2\equiv -s^2 \pmod{dq}}} e(\frac{h\overline{r}}{s} - \frac{hr}{s(r^2+s^2)}) \\
&\leq \rho(dq) \sum_{(\frac{1}{2}qM)^{1/2}<s<(2qM)^{1/2}} \max_{\substack{r_1,r_2 \\ \lambda,\Lambda}} \Big| \sum_{\substack{r_1<r<r_2,(r,s)=1 \\ r\equiv\lambda \pmod{\Lambda}}} e(\frac{h\overline{r}}{s} - \frac{hr}{s(r^2+s^2)})\Big| \\
&\ll \sum_{(\frac{1}{2}qM)^{1/2}<s<(2qM)^{1/2}} s^{1/2+\varepsilon}(h,s)^{1/2}(1+\frac{h}{s^2}) \\
&\ll (qM)^{3/4+\varepsilon}(1+\frac{h}{qM})
\end{aligned}$$

Usando las cotas sobre C_h vemos que el error total es

$$\ll (qM)^{3/4+\varepsilon}(1+\frac{1}{\Delta qM})\log(1/\Delta)$$

y tomando $\Delta = (qM)^{-1}$ hemos terminado. □

8.5 Pesos de Richert mejorados

Ya hemos visto que ni usando la cota mejorada (Lema 8.3.1) para el término de error conseguimos probar el Teorema 8.1.1. ¿A qué se debe?: Lo que usaríamos para ver que hay infinitos casiprimos en A es que

$$S(A, x^{2/3}) = \sum_{a\in A,(a,P(x^{2/3}))=1} 1 > 0, \tag{8.13}$$

pero no sabemos probarlo. Esto ocurre porque la condici$\frac{3}{4}$n de que esa suma sea mayor que cero es mucho más restrictiva que la de que existan infinitos casiprimos en A. Por ejemplo, en esta suma hemos descartado cualquier elemento $a = pq$ con p, q primos y $p < x^{2/3}$, y estos son muchos elementos. Lo que podemos hacer para flexibilizar esta situación es suavizar $S(A, x^{2/3})$, considerando sumas

$$\sum_{a\in A,(a,P(z_*))=1} w(a) \tag{8.14}$$

donde $w(a)$ va a ser una cierta función (peso) y z_* mucho más pequeño que $x^{2/3}$, de forma que de la positividad de la suma podamos deducir que hay infinitos casiprimos en A. Una opción para el peso, que nos va a permitir tratar esta nueva suma a través de varias sumas $S(A_q, z)$, es considerar

$$w(a) = 1 - \sum_{p|a} w_p(a),$$

para ciertas funciones w_p. Esta elección es natural: Nuestro problema en (8.13) era que quitábamos cualquier múltiplo de p, con $p \leq x^{2/3}$. Con los pesos $w(a)$ ya no va a ocurrir esto, sino que el valor de $w(a)$ va a ser mayor o menor dependiendo del número de primos y de qué primos dividan a. Parece adecuado hacer que $w_p(a)$ dependa del tamaño de p. Esto es lo que ocurre en los pesos de Richert, donde

$$w_p(a) = \begin{cases} 1 - \log p / \log x & \text{si } p \leq x \\ 0 & \text{si } p > x. \end{cases} \tag{8.15}$$

Si $p = x^\alpha$, el valor de esta función es $1 - \alpha$ con $0 < \alpha < 1$, y cero para $\alpha > 1$. Es una función que decrece de forma lineal con el exponente de cero a uno. Por tanto lo que estamos haciendo es dar menos peso a los primos pequeños y más a los grandes, que es lo mismo que hacíamos con la criba pero de una manera más regular. Con esta elección, vemos de forma directa que

$$a \text{ libre de cuadrados y } w(a) > 0 \Rightarrow a \text{ es casiprimo.} \tag{8.16}$$

Así, si demostrásemos que la suma (8.14) es positiva para algún $z_* = x^\gamma$, podríamos deducir que en A hay infinitos casiprimos. Vamos a poder tratar esta suma de igual forma que $S(A, z)$, pero el método en este caso va a seguir sin funcionar, es decir no sabremos demostrar que es mayor que cero. Iwaniec consiguió el resultado con una pequeña rebaja de estos pesos (sugerida por Richert): Si p_a es el primo más pequeño que divide a a definimos

$$\widetilde{w}_p(a) = \begin{cases} \log p_a / \log x & \text{si } p_a < p < x^{1/2} \\ w_p(a) & \text{en otro caso.} \end{cases}$$

Está claro que $\widetilde{w}_p(a) \leq w_p(a)$, y por tanto $\widetilde{w}(a) = 1 - \sum_{p|a} \widetilde{w}_p(a) \geq w(a)$. Luego hay posibilidades de que esta vez consigamos que la suma

asociada a $\widetilde{w}$ sea positiva. Pero, ¿se sigue cumpliendo la propiedad fundamental (8.16)? Sí, porque si $\widetilde{w}(a) > 0$ entonces $\widetilde{w}_p(a) = w_p(a)$ para todo $p \mid a$, y por tanto $w(a) = \widetilde{w}(a) > 0$. Esto es debido a que no existen divisores primos de a en el intervalo $(p_a, x^{1/2})$, porque en ese caso

$$\widetilde{w}(a) \leq 1 - (1 - \log p_a/\log x) - \log p_a/\log x \leq 0.$$

Por tanto

$$|\{a \in A : a \in P_2\}| \geq \sum_{\substack{a\in A,(a,P(z_*))=1\\ \mu(a)\neq 0}} \widetilde{w}(a) = W(A, z_*) + O(xz_*^{-1/2}),$$

donde

$$W(A, z_*) = \sum_{a\in A,(a,P(z_*))=1} \widetilde{w}(a).$$

Luego tomando $z_* = x^\gamma$ con γ una constante adecuada (pequeña), para demostrar el Teorema 8.1.1 será suficiente ver que

$$W(A, z_*) > \frac{1}{77}\Gamma_G \frac{x}{\log x}.$$

Pero podemos reescribir $W(A, z_*)$ como

$$\begin{aligned} W(A, z_*) &= S(A, z_*) - \sum_{a\in A,(a,P(z_*))=1} \sum_{p\mid a} w_p(a) \\ = \; & S(A, z_*) - \sum_{z_*\leq p<x^{1/2}} \left(1 - \frac{\log p}{\log x}\right) S(A_p, p) \\ & - \sum_{z_*\leq p_1<p<x^{1/2}} \frac{\log p_1}{\log x} S(A_{pp_1}, p_1) - \sum_{x^{1/2}\leq p<x} \left(1 - \frac{\log p}{\log x}\right) S(A_p, z_*) \end{aligned} \tag{8.17}$$

La aportación realizada por la suma

$$\sum_{x^{1-\varepsilon}\leq p<x} \left(1 - \frac{\log p}{\log x}\right) S(A_p, z_*)$$

es $O(\varepsilon x/\log x)$ debido a la acotación

$$S(A_p, z_*) \leq S(A_p, \frac{x}{p}) \ll \frac{x}{p}\frac{1}{\log(x/p)},$$

la cual proviene del Teorema 4.5.1. Para el resto de términos en (8.17), estimando $S(A_q, z_q)$ mediante las funciones f y F (lema 8.2.1) con nivel D/q ($D = x^\alpha$) y aproximando las sumas así obtenidas por integrales tenemos que

$$\begin{aligned} W(A, z_*) > xV(z_*)\Big\{ & f\left(\frac{\alpha}{\gamma}\right) - \int_\gamma^{\frac{1}{2}} (1-u)F\left(\frac{\alpha-u}{u}\right)\frac{du}{u} \\ - & \int_\gamma^{\frac{1}{2}} \int_\gamma^u u\frac{\gamma}{t} f\left(\frac{\alpha-u-t}{t}\right)\frac{du}{u}\frac{dt}{t} - \int_{\frac{1}{2}}^1 (1-u)F\left(\frac{\alpha-u}{\gamma}\right)\frac{du}{u}\Big\} \\ - & E + O(\frac{\varepsilon x}{\log x}) \end{aligned} \tag{8.18}$$

donde $E = E(\alpha, \gamma)$ es la acumulación de los términos de error de cada estimación de $S(A_q, z_q)$. Por acotaciones elementales de las funciones f y F o por métodos de integración (ordenador) podemos ver que para $\alpha = 16/15 - 5\varepsilon$ y $\gamma = 1/5$ se obtiene que

$$W(A, x^\gamma) > \frac{x}{\gamma \log x}\left\{\frac{2e^C\gamma}{154} + O(\varepsilon)\right\} - E,$$

lo que nos da el resultado (pues $2e^C/154 > 1/77$) si demostramos que

$$E \ll \frac{\varepsilon\, x}{\log x}. \tag{8.19}$$

En realidad, como los coeficientes a_n, b_m del Lema 8.2.1 dependen de z_q (pero no de A_q), no podemos controlar E de forma adecuada. Lo que hacemos para remediarlo es lo siguiente: Para cualesquiera q, z_q con

$$Q \le q < 2Q, Z < z_q < 2Z, \tag{8.20}$$

usamos el Lema 8.2.1 (en vez de con $S(A_q, z_q)$) con $S(A_q, 2Z)$ o $S(A_q, Z)$ dependiendo de si $S(A_q, z_q)$ está sumando o restando en (8.17). De esta forma, como f y F son diferenciables las estimaciones de $S(A_q, Z)$ y $S(A_q, 2Z)$ son similares a la de $S(A_q, z_q)$, luego (8.18) sigue siendo válido pero ahora el error lo podemos controlar

por (con $MN = x^{16/15-5\varepsilon}$)

$$\sum_{\substack{Q<q<2Q\\(q,P(Z))=1}} \sum_{\substack{m<M/Q\\m|P(Z)}} |\sum_{\substack{n<N,(m,n)=1\\n|P(Z)}} b_n r(A, qmn)| \leq$$

$$\leq \max_{k\leq M} d(k) \sum_{m<M} |\sum_{\substack{n<N,(m,n)=1\\n|P(Z)}} b_n r(A, mn)|,$$

debido a que $r(A_q, mn) = r(A, qmn)$ y a que b_n no depende de q. Teniendo en cuenta que hay $O((\log x)^2)$ intervalos de la forma (8.20), tomando $N = x^{1/15-\varepsilon}$ y $M = x^{1-4\varepsilon}$ la acotación (8.19) se sigue de aplicar el lema 8.3.1.

8.6 Notas

El método de dispersión fue elaborado por Linnik para tratar problemas de teoría aditiva. Hooley [Ho2] usó éste método en el contexto de la criba en promedio. Iwaniec adaptó este método para acotar el término de error en general, tras haber hecho la separación de variables.

Con el término de error usual también es posible usar las acotaciones obtenidas por Iwaniec al para $S(A_q, z_q)$ cuando $q > x^{14/15}$. Esto es debido a que en (8.17) promediamos en q los errores:

$$\sum_q \sum_d \varrho_d r(A, dq).$$

También podríamos haber expandido los $r(A, mn)$ en serie de Fourier, pero los términos diagonales nos darían ya $MN^{1/2}$. Teniendo en cuenta que para los errores de $S(A_q, z_q)$ debemos tomar $M_q = M/q \geq 2$, entonces M va a ser como x para q cercanos a x. Esto hace que $MN^{1/2} \gg x$.

En vez de haber usado las funciones suaves A, B, si hubiésemos usado sólo A para estimar ψ, obtendríamos un error $O((qM)^{3/4}(1 + q^{-3/8}M^{1/8}))$ lo que nos llevaría al nivel 32/31 que sería insuficiente para demostrar el resultado.

Bibliografia

[Ba-Ho] *P. T. Bateman and R. A. Horn*, A heuristic asymptotic formula concerning the distribution of prime numbers, Math. Comput. 16, (1962), 363–367.

[Bo] *E. Bombieri*, Le grand crible dans la théorie analytique des nombres, Asterisque 18 Societé matematique de Paris, 1987.

[B] *V. Brun*, Le crible d'Erathostène et le théorème de Goldbach, C. R. Acad. Sci. Paris, 168 (1919), 544–546.

[Ch] *J. Chen*, On the representation of a large even integer as the sum of a prime and the product of at most two primes, Sci. Sinica, 16 (1973), 157–176.

[D] *J. M. Deshouillers*, Progrès récents des petits cribles arithmétiques, Séminaire Bourbaki in Lecture notes in Mathematics, 710 (1979), 248–262.

[DERZ] *J.-M. Deshouillers, G. Effinger, H. te Riele and D. Zinoviev* A complete Vinogradov 3-primes theorem under the Riemann hypothesis. Electron. Res. Announc. Amer. Math. Soc. 3 (1997), 99–104.

[F-H] *K. Ford and H. Halberstam*, The Brun-Hooley sieve, Jour. Num. Theory, 81 (2000), 335–350.

[F-I1] *J. Friedlander and H. Iwaniec, Asymptotic sieve for primes*, Annals of Math. ,148,3 (1998), 1041–1065.

[F-I2] *J. Friedlander and H. Iwaniec*, The polynomial $x^2 + y^4$ captures its primes, Annals of Math., 148,3 (1998), 945–1040.

[G] *G. Greaves*, Sieve in nuber theory, in "A series of modern surveys in Mathematics", Springer-Berlin, 2001.

[Gr] *B. Green*, Notas de exposición en su página web, http://www.dpmms.cam.ac.uk/˜bjg23/expos.html.

[H] *H. Halberstam*, Review on Sieves in Number theory by G. Greaves, Bull. Amer. Math. Soc, 40.1 (2001).

[H-R] *H. Halberstam and H. -E. Richert*, Sieve methods, Academic Press, London, 1974.

[H-Ro] *H. Halberstam and K. F. Roth*, Sequences, Oxford, 1966.

[HB] *D. R. Heath-Brown*, Lectures on sieves. Proceedings of the Session in Analytic Number Theory and Diophantine Equations, Bonner Math. Schriften, 360, Univ. Bonn, Bonn, 2003.

[Ho1] *C. Hooley* On the Brun-Titchmarsh theorem. *J. Reine Angew. Math.* **255**, (1972), 60–79 .

[Ho2] *C. Hooley*, On the Brun-Titchmarsh theorem. II. *Proc. London Math. Soc.* **3** (30),(1975), 114–128 .

[Ho3] *C. Hooley*, On the greatest prime factor of a quadratic polynomial, Acta Math., 117, (1967), 281–299.

[I1] *H. Iwaniec*, Rosser´s sieve, Acta Arith. 36 (1980), 171–202

[I2] *H. Iwaniec*, Rosser´s sieve-Bilinear forms of the Remainder terms-Some applications, Academic Press in Recent progress in Number Theory 1, London, 1981.

[I3] *H. Iwaniec*, On the error term in the linear sieve, Acta Arith., 19 (1971), 1–30.

[I4] *H. Iwaniec*, A new form of the error term in the linear sieve, Acta Arith., 37 (1980), 307–320.

[I5] *H. Iwaniec*, Almost primes represented by quadratic polynomials, Inv. Math., 47 (1978), 171–188.

[I6] *H. Iwaniec*, Notes, Rutgers, 1996.

[J-R] *W. B. Jurkat and H. -E. Richert*, An improvement of Selberg's sieve method I, Acta Arith., 11, (1965), 217–240.

[K] *P. Kuhn*, Zur Viggo Brunschen Siebmethode, I, Norske Vid. Selsk. Forh. Trondheim, 14 (1941), 145–148.

[L] *Y. V. Linnik*, The large sieve, Dokl. Akad. Nauk SSSR 30 (1941), 292–294.

[M] *H. Montgomery*, Topics in Multiplicative Number Theory in Lecture notes in Mathematics, 227, Springer- Berlin, 1986.

[Mo] *Y. Motohashi*, Sieve methods and prime number theory, Springer, Tata institute of fundamental research, 1983.

[Nat] *M. B. Nathanson*, Additive number theory. The classical bases, Graduate Texts in Mathematics, 164. Springer-Verlag, New York, 1996.

[Ram] *O. Ramaré*, On Šnirelman's constant. Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 22 , no. 4,(1995), 645–706.

[Sc-Si] *A. Schinzel and W. Sierpinski*, Sur certaines hypothèses concernant les nombres premieres, Acta Arith. (4) (1958), 185–208.

[Sc] *L. Schnirelmann*, Űber additive eigenschaften von zahlen, Math. Ann., 107 (1933), 649–690.

[S] *A. Selberg*, Lectures on Sieves, in Collected papers, vol. 2, Springer-Berlin, 1991.

[Se1] *J-P. Serre*, Topics in Galois Theory, Notes written by Henri Darmon, Jones and Bartlett Publishers, 1992.

[Se2] *J-P. Serre*, Lectures on the Mordell-Weil Theorem, from notes by Michel Waldschmidt, Vieweg & Sohn, 1990.

[Vi] *I. Vinogradov*, Some Theorems Concerning the Theory of Primes. Recueil Math. 2, (1937) 179–195.