PRIMOS, PARIDAD Y ANÁLISIS

HARALD HELFGOTT Y ADRIÁN UBIS

Resumen. Sea $\lambda(n) = 1$ cuando n tiene un número par de divisores primos (contados con multiplicidad) y $\lambda(n) = -1$ de lo contrario. En general, distinguir entre números con un número par o impar de divisores primos es una de las tareas más difíciles en la teoría analítica de números. Un trabajo reciente de Matomäki y Radziwiłł muestra que, en promedio, ambos existen con la misma frecuencia aún en intervalos muy cortos. Este avance ya ha tenido varias aplicaciones importantes en las manos de Matomäki, Radziwiłł, Tao y Teräväinen. Explicaremos en detalle una prueba completa del resultado original de Matomäki y Radziwiłł, así como de varias aplicaciones.

Índice

1. Introducción

1.1. Primalidad y paridad. Los números primos son uno de los objetos de estudio principales de la teoría de números. Los problemas clásicos acerca de los primos son clásicos en parte por su dificultad: nuestras técnicas nos permiten acercarnos a la meta,

Date: Versión final: 6 de junio de 2019.

Estas notas corresponden al curso dictado por los autores en la escuela AGRA III, Aritmética, Grupos y Análisis, del 9 al 20 de Julio de 2018 en Córdoba, Argentina.

pero, una vez que se trata de encontrar todos los casos primos y sólo ellos, nos vemos a menudo frustrados.

Consideremos un ejemplo entre muchos. La *conjetura fuerte de Goldbach* dice que todo número par $n \geq 4$ puede escribirse como la suma de dos primos. Rényi [Rén47] mostró que todo número par suficientemente grande puede escribirse como la suma de un primo y un número que es el producto de a lo más K primos, donde K es una constante. Luego hubo una sucesión de trabajos que mejoraron la constante. Finalmente, en [Che73], Jing-Run Chen logró mostrar que todo par suficientemente grande puede escribirse como la suma de un primo y un número que es ya sea un primo o el producto de dos primos. Empero, la conjetura fuerte de Goldbach sigue sin resolver.

La misma situación ocurre con la *conjetura de los primos gemelos*, la cual plantea que hay un número infinito de primos p tales que $p+2$ también es primo. Se puede dar una cota superior al número de tales p con $p \leq N$ mediante métodos de *criba* (cómo veremos en los ejercicios en la sección 2.2), pero no tenemos una cota inferior. Como en el caso de Goldbach, hubo una sucesión de trabajos llevando a un resultado similar al de Chen; también hay otras aproximaciones a las dos conjeturas (*Goldbach débil*, *distancias acotadas entre primos*). Empero, las conjeturas en sí siguen abiertas.

En general, distinguir entre un primo y el producto de dos primos es muy difícil. Si se usa un procedimiento de *criba* de tipo tradicional, hacer tal distinción es no sólo difícil sino imposible. En general, las cribas, por sí solas, no llegan a distinguir entre números con un número par o impar de factores primos; se trata del *problema de paridad* (§2.2). Claro está, hay otros procedimientos, generalmente analíticos, que nos permiten probar algunos enunciados sobre los números primos.

Un resultado de base de este tipo es el *teorema de los números primos*. Recordamos que nos dice que el número $\pi(x)$ de primos entre 1 y x es aproximadamente $x/\log x$. Para ser más precisos, en la versión de Hadamard y de la Vallée Poussin (1896), nos dice que

$$\pi(x) = Li(x) + O\left(xe^{-c\sqrt{\log x}}\right), \tag{1.1}$$

donde $Li(x) = \int_2^x \frac{dt}{\log t}$, $c > 0$ es una constante y $O(f(x))$ quiere decir un término cuyo valor absoluto está acotado por $f(x)$ por una constante.

La prueba se basa sobre las propiedades de la función zeta de Riemann $\zeta(s)$, y, en particular sobre el hecho que sabemos que no toma el valor 0 cuando s está dentro de una cierta región. Por los mismos métodos, o usando (1.1), podemos mostrar que el número de enteros $n \leq x$ con un número par o impar de factores primos es asintóticamente el mismo:

$$\sum_{n\leq x} \lambda(n) = O\left(e^{-c\sqrt{\log x}}x\right), \tag{1.2}$$

donde $\lambda(n)$ es la *función de Liouville*, la cual es la función completamente multiplicativa tal que $f(p) = -1$ para todo número primo. El enunciado (1.2) también es cierto si $\lambda(n)$ se reemplaza por la más familiar *función de Möbius* $\mu(n)$, definida como $\mu(n) = \lambda(n)$ para n sin divisores cuadrados (aparte de 1), y como $\mu(n) = 0$ si $d^2|n$ para algún $d > 1$.

De hecho lo más importante es que la cota en (1.2) dividida por la cota trivial x va a cero:

$$\sum_{n\leq x} \lambda(n) = o(x),$$

donde $o(f(x))$ denota cualquier función $g(x)$ tal que $\lim_{x\to\infty} g(x)/f(x) = 0$. (Aquí, $f(x) = x$.) En otras palabras, el promedio de λ en el intervalo $1 \leq n \leq x$ es $o(1)$, es decir, tiende a cero.

En general, es muy interesante saber que λ tiene promedio $o(1)$ en un intervalo o conjunto de números, pues esto quiere decir precisamente que sabemos que en ese conjunto el número de enteros con un número par o impar de factores primos (contados con multiplicidad) es asintóticamente el mismo. Sabíamos, por ejemplo, que

$$\frac{1}{H}\sum_{n=N+1}^{N+H}\mu(n)=o(1)$$

para $H\geq N^{\alpha}$ y α dentro de un cierto rango. El primer resultado con $\alpha<1$ fue probado por Hoheisel (1930); el mejor valor de α conocido en nuestros días es $7/12+\epsilon$, $\epsilon>0$ arbitrario ([Mot76] y [Ram76], basados en parte en [Hux72]). También se tenían resultados "en promedio": por ejemplo, se sabía que

$$\int_X^{2X}\left|\sum_{x<n\leq x+h}\lambda(n)\right|^2 dx=o(Xh^2) \tag{1.3}$$

para $h\geq X^{1/6+\epsilon}$, $\epsilon>0$ arbitrario. La desigualdad (1.3) es equivalente al enunciado siguiente: dado $h=h(x)$ tal que $h(x)\geq x^{1/6+\epsilon}$,

$$\sum_{N<n\leq N+h(x)}\lambda(n)=o(h(x))$$

para todo entero $N\in[x,2x]$ fuera de un conjunto de $o(x)$ excepciones. (La equivalencia es inmediata; el lado izquierdo de (1.3) es una varianza.)

La *conjetura de Chowla* nos dice que, para $h_1,h_2,\ldots,h_k$ enteros distintos,

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n\leq N}\lambda(n+h_1)\cdots\lambda(n+h_k)=0. \tag{1.4}$$

Cuando $k>1$, la conjetura está abierta y se considera muy difícil. En general, se cree que, para todo polinomio $P\in\mathbb{Z}[x]$ no constante,

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n\leq N}\mu(P(x))=0.$$

(Las conjeturas de *Hardy-Littlewood* y de *Schinzel* son enunciados relacionados para los números primos.) Nuevamente, no hay caso probado con $\deg P>1$, si bien se tienen resultados para polinomios en dos variables.

Una puerta fue abierta recientemente por Kaisa Matomäki y Maksym Radziwiłł [MR16]. Ha conducido ya a numerosos resultados.

1.2. Resultados de Matomäki, Radziwiłł y Tao.

El siguiente resultado fue un gran avance, puesto que, como hemos visto, la conclusión se sabía sólo bajo la condición $h(x)\geq x^{1/6+\epsilon}$, mucho más fuerte que $h(x)\to\infty$.

Teorema 1.1. *Sea $h=h(x)$ tal que $h(x)\to\infty$ cuando $x\to\infty$. Entonces, cuando $X\to\infty$,*

$$\left|\sum_{N<n\leq N+h(X)}\lambda(n)\right|=o(h(X)) \tag{1.5}$$

para todo entero $N\in[X,2X]$ fuera de un conjunto de $o(X)$ excepciones.

En verdad, con algo de trabajo adicional, Matomäki y Radziwiłł demuestran un análogo del Teorema 1.1 para todo $f : \mathbb{Z} \to \mathbb{R}$ multiplicativo que satisfaga $|f(n)| \leq 1$ para todo n [MR16, Thm. 1]:

$$\left| \frac{1}{h(X)} \sum_{N<n\leq N+h(X)} f(n) - \frac{1}{X} \sum_{X<n\leq 2X} f(n) \right| = o(1)$$

para todo $N \in [X, 2X]$ fuera de $o(X)$ excepciones. (Hay generalizaciones también para f con valores complejos [MRT15, Thm. A.1].)

(Por cierto, es plausible que (1.5) sea cierto para *todo* $N \in [X, 2X]$, con tal que, digamos, $h(x) \geq C(\log x)^2$. Empero, es fácil construir contraejemplos a tal enunciado con f multiplicativa, $f \neq \lambda$.)

Así como el Teorema 1.1 es un resultado sobre la cancelación en intervalos cortos en promedio, se puede probar la conjetura de Chowla en promedio.

Teorema 1.2. ([MRT15, Thm. 1.1]) *Sea $k \geq 1$ arbitrario. Sea $h = h(x)$ tal que $h(x) \to \infty$ cuando $x \to \infty$. Entonces, cuando $N \to \infty$,*

$$\sum_{n\leq N} \lambda(n+h_1)\cdots\lambda(n+h_k) = o(N)$$

para enteros $(h_1, \ldots, h_k)$, $1 \leq h_i \leq h(N)$, cualesquiera, fuera de un conjunto de $o\left(h(N)^k\right)$ excepciones.

La prueba se basa en la del Teorema 1.1, vía el método del círculo.

Los Teoremas 1.1 y 1.2 admiten variantes cuantitativas. Aún para $h(X)$ pequeño, los métodos de Matomäki, Radziwiłł y Tao pueden dar una cota de la forma

$$O((\log\log h(X))/\log h(X)).$$

Los términos de error de esta forma son típicos de pruebas en las cuales el caso de los primos se ve como una excepción. Las pruebas que veremos no son una excepción; de hecho, ponen a los enteros que no tienen una factorización típica en el término de error.

(En verdad, nosotros probaremos cotas del tipo $O(1/(\log h(x))^\alpha)$, $\alpha > 0$. Nuestro énfasis será en dar una exposición clara y concisa, y no en conseguir necesariamente las mejores cotas que los métodos permiten.)

Los métodos y resultados de Matomäki y Radziwiłł han tenido varias otras aplicaciones [MRT16], [Tao16a], [Tao16b], [MRT19], [MRT17]. Discutiremos una de ellas: la prueba de una versión logarítmica de la conjetura de Chowla con $k = 2$.

Teorema 1.3. ([Tao16b, Thm. 1.1]) *Sean a_1, a_2, b_1, b_2 enteros tales que (a) $a_1, a_2 \geq 1$ y (b) (a_1, b_1), (a_2, b_2) no son múltiplos el uno del otro. Sea $\omega = \omega(x)$ tal que $\omega(x) \to \infty$ cuando $x \to \infty$. Entonces, cuando $x \to \infty$,*

$$\sum_{x/\omega(x)<n\leq x} \frac{\lambda(a_1n+b_1)\lambda(a_2n+b_2)}{n} = o(\log\omega(x)).$$

También este teorema se generaliza a una clase de funciones multiplicativas, incluyendo algunas para las cuales la conjetura de Chowla "estándar" (1.4) no sería cierta.

1.3. Notación. Cuando escribimos $f(n) \ll g(n)$, $g(n) \gg f(n)$ o $f(n) = O(g(n))$, queremos decir la misma cosa, esto es, que hay $N > 0$ y $C > 0$ tales que $|f(n)| \leq C \cdot g(n)$ para todo $n \geq N$. Escribimos $\ll_a$, $\gg_a$, O_a si N y C dependen de a (digamos). Como de costumbre, $f(n) = o(g(n))$ significa que $|f(n)|/g(n)$ tiende a 0 cuando $n \to \infty$.

Dado un subconjunto $A \subset X$, $1_A : X \to \mathbb{C}$ es la función característica de A:

$$1_A(x) = \begin{cases} 1 & \text{si } x \in A, \\ 0 & \text{si no.} \end{cases}$$

Denotamos por $|A|$ el número de elementos de un conjunto finito A.

Escribiremos (a, b) por el máximo comun divisor de dos enteros $a, b \neq 0$, siempre y cuando no haya posibilidad de confusión con el par ordenado (a, b).

La distancia $d(\beta_1, \beta_2)$ entre dos elementos $\beta_1, \beta_2 \in \mathbb{R}/\mathbb{Z}$ se define como $\min_{a\in\mathbb{Z}} |a + \beta_1 - \beta_2|$. Por ejemplo, la distancia entre 0,01 y 0,99 es 0,02. Podemos también definir, para $\beta \in \mathbb{R}$, la distancia $d(\beta, \mathbb{Z})$ como la distancia entre β y el entero más cercano. Está claro que $d(\beta, \mathbb{Z})$ depende sólo de $\beta \bmod \mathbb{Z}$, y que $d(\beta_1, \beta_2) = d(\beta, \mathbb{Z})$ para cualquier β tal que $\beta \equiv \beta_1 - \beta_2 \bmod \mathbb{Z}$.

1.4. Agradecimientos. El viaje y estadía de los autores fueron financiados en parte por la fundación Humboldt. Se deben las gracias a Boris Bukh y Nikos Frantzikinakis por sus valiosos comentarios.

2. Breve repaso y resultados preliminares

2.1. Función zeta. Fórmula de Perron. Sea $f : \mathbb{N} \to \mathbb{C}$. Muchas veces estamos interesados en sumas finitas de f:

$$\sum_{n\leq x} f(n).$$

En general, podemos tratar de obtener resultados sobre una función f estudiando una funcion generatriz de f, como, por ejemplo, $\sum_{n=1}^{\infty} f(n)z^n$. Si f está asociada a un problema multiplicativo, tiene sentido estudiar la siguiente función generatriz, llamada *función zeta* de f:

$$Z_f(s) = \sum_{n=1}^{\infty} f(n)n^{-s} \qquad s \in \mathbb{C}.$$

La idea de usar los factores n^{-s} es que son multiplicativos: $(ab)^{-s} = a^{-s}b^{-s}$. La idea es que, si obtenemos suficiente información sobre $Z_f(s)$, podremos deducir información sobre $f(n)$.

Si f es multiplicativa, entonces, como lo notó ya Euler, podemos factorizar $Z_f(s)$ como un producto sobre los primos:

$$Z_f(s) = \prod_p (1 + f(p)p^{-s} + f(p^2)p^{-2s} + \ldots) \qquad \Re s > 1 \tag{2.1}$$

Por ejemplo, en el caso $f = 1$ tenemos que su función zeta es

$$Z_1(s) = \zeta(s) = \sum_{n=1}^{\infty} n^{-s} = \prod_p \frac{1}{1 - 1/p^s}, \tag{2.2}$$

de donde, por ejemplo, obtenemos que hay infinitos primos, puesto que $\lim_{s\to 1^+} \sum_n n^{-s}$ diverge, y $\lim_{s\to 1^+} 1/(1 - 1/p^s)$ puede divergir sólo si es un producto infinito. Escogiendo s de manera más precisa, podemos obtener cotas útiles sobre los primos (ejercicio 4).

Veamos cómo usar la función zeta de manera elemental para acotar la suma de una función aritmética, $\sum_{n<x} \tau(n^2)n^{-1}$, con τ la función divisor. Para $s > 1$,

$$\sum_{n\leq x} \frac{\tau(n^2)}{n} \leq \sum_{n} \frac{\tau(n^2)}{n} \left(\frac{x}{n}\right)^{s-1} = x^{s-1} \sum_{n} \frac{\tau(n^2)}{n^s} = x^{s-1} \prod_{p} \left(1 + \frac{3}{p^s} + O\left(\frac{1}{p^{2s}}\right)\right)$$
$$\ll x^{s-1} \prod_{p} \left(1 + \frac{3}{p^s}\right) \leq x^{s-1} \prod_{p} \left(1 + \frac{1}{p^s}\right)^3 \leq x^{s-1} (\sum_{n} n^{-s})^3 \ll \frac{x^{s-1}}{(s-1)^3},$$

donde hemos usado el hecho que $\sum_n n^{-s} - \int_1^\infty t^{-s} dt \ll 1$. Tomando $s = 1 + \frac{1}{\log x}$, obtenemos

$$\sum_{n\leq x} \frac{\tau(n^2)}{n} \ll (\log x)^3. \tag{2.3}$$

Éste es el orden de magnitud correcto: en verdad, $\sum_{n\leq x} \tau(n^2)n^{-1}$ es asintótica a una constante por $(\log x)^3$.

Si queremos asintóticas para sumas necesitamos algo más preciso para relacionar las sumas con la función zeta. La herramienta que lo permite es la integral de Perron, que es capaz de expresar que un número y sea mayor o menor que 1 en términos analíticos, y en particular en términos de los «armónicos» y^s, s complejo: para $\sigma > 0$,

$$\frac{1}{2\pi i} \int_{\sigma - i\infty}^{\sigma + i\infty} y^{-s} \frac{ds}{s} = \begin{cases} 1 & \text{si } 0 < y < 1, \\ 1/2 & \text{si } y = 1, \\ 0 & \text{si } y > 1. \end{cases} \tag{2.4}$$

Esta fórmula puede probarse usando el teorema de los residuos, o el teorema de inversión de Fourier para la función $1_{(0,\infty)}(x)e^{-\sigma x}$ en $x = \log y$, ya que $y^{-it} = e^{-it \log x}$. La integral debe comprenderse como el límite $\lim_{T\to\infty} \int_{\sigma-iT}^{\sigma+iT} y^{-s} ds/s$.

Así, usando Perron, obtenemos, para $x > 0$,

$$\sum_{n\leq x} f(n) = \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} Z_f(s) x^s \frac{ds}{s} + \begin{cases} \frac{1}{2} f(x) & \text{si } x \in \mathbb{Z}, \\ 0 & \text{si } x \notin \mathbb{Z}. \end{cases} \tag{2.5}$$

Si bien todos los armónicos x^{it} contribuyen a la integral, típicamente los importantes van a ser los que tienen frecuencia t pequeña. Para demostrarlo, la idea es que es mejor trabajar con una función continua, antes que con la función en el lado derecho de (2.4). Podemos, por ejemplo, trabajar con la siguiente función, continua en $(0, \infty)$:

$$\psi_\delta(x) = 1_{(0,1-\delta)} + \frac{1-x}{\delta} 1_{[1-\delta,1]} \qquad 0 < \delta < \frac{1}{2}. \tag{2.6}$$

Por lo mismo que es igual a $1_{(0,1)}(x)$ cuando $0 \leq x < 1 - \delta$ o $x \geq 1$, está claro que, sí $|f(x)| \leq 1$ para todo x, entonces

$$\sum_{n\leq x} f(n) = O(\delta x) + \sum_{n=1}^{\infty} f(n) \psi_\delta \left(\frac{n}{x}\right) \tag{2.7}$$

Usando el teorema de inversión de Fourier (o la integral de Perron), no es difícil demostrar (ejercicio 5) que

$$\psi_\delta(y) = \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} (M\psi_\delta)(s) y^{-s} ds, \tag{2.8}$$

donde

$$(M\psi_\delta)(s) = \frac{1}{s(s+1)} \frac{1 - (1-\delta)^{s+1}}{\delta}. \tag{2.9}$$

Utilizamos la notación $M\psi$ aquí pues se trata de una *transformada de Mellin* (lo cual no es sino una transformada de Fourier con un cambio complejo de variable). En general,

$$M\psi(s) := \int_0^\infty \psi(x) x^{s-1} dx,$$

y, bajo ciertas condiciones de integrabilidad sobre ψ y $M\psi$,

$$\psi(x) = \frac{1}{2\pi i} \int_{\sigma - i\infty}^{\sigma + i\infty} M\psi(s) x^{-s} ds \qquad \text{(teorema de inversión de Mellin)},$$

lo cual se deduce del teorema de inversión de Fourier.

Lema 2.1. *Si $|f(n)| \leq 1$ para todo n, entonces*

$$\sum_{n \leq x} f(n) = O(\delta x \log x) + \frac{1}{2\pi i} \int_{1 + \frac{1}{\log x} - i\delta^{-2}}^{1 + \frac{1}{\log x} + i\delta^{-2}} x^s (M\psi_\delta)(s) Z_f(s) ds.$$

Demostración. Por (2.7) y (2.8), para $\sigma > 0$,

$$\sum_{n \leq x} f(n) = O(\delta x) + \frac{1}{2\pi i} \int_{\sigma - i\infty}^{\sigma + i\infty} x^s (M\psi_\delta)(s) Z_f(s) ds.$$

Vemos que (2.9) implica que $|M(\psi_\delta)(s)| \leq 2/\delta |s(s+1)|$ cuando $\Re s > 0$. Como $|f(n)| \leq 1$ para todo n, $|Z_f(\sigma + it)| \leq \zeta(\sigma) \leq 1/(\sigma - 1) + O(1)$ para $\sigma > 1$. Por lo tanto,

$$\int_{\sigma + iT}^{\sigma + i\infty} x^s (M\psi_\delta)(s) Z_f(s) ds \ll \frac{|x^\sigma|}{\sigma - 1} \int_T^\infty \frac{1}{\delta |t(t+1)|} dt \ll \frac{|x^\sigma|}{\delta(\sigma - 1)T}. \tag{2.10}$$

Tomamos $\sigma = 1 + 1/\log x$ y $T = \delta^{-2}$, y el lado derecho de (2.10) se vuelve $O(\delta x \log x)$. La cota para la integral de $\sigma - i\infty$ a $\sigma - iT$ es evidentemente la misma. □

Luego efectivamente sólo las frecuencias pequeñas van a ser importantes para controlar el promedio. Veamos cómo usar este hecho para ver que hay cancelación en la suma

$$\sum_{n \leq x} \lambda(n)$$

Para ello, miramos a su función zeta correspondiente a $\lambda(n)$:

$$Z_\lambda(s) = \sum_n \lambda(n) n^{-s} = \prod_p (1 - p^{-s} + p^{-2s} - \ldots) = \prod_p \frac{1}{1 + p^{-s}} = \prod_p \frac{1 - p^{-s}}{1 - p^{-2s}},$$

luego

$$Z_\lambda(s) = \frac{\zeta(2s)}{\zeta(s)}. \tag{2.11}$$

Para acotar el promedio de f queremos extender $Z_f(s)$ a s con parte real < 1 con el propósito de usar la integral en (2.5) para la σ más pequeña que podamos, ya que $|x^s| = x^\sigma$.

Para $\Re s > 1$,

$$\zeta(s) = \frac{1}{s-1} + \sum_{n=1}^\infty n^{-s} - \int_1^\infty u^{-s}\, du = \frac{1}{s-1} + \sum_{n=1}^\infty \int_n^{n+1} (n^{-s} - u^{-s})\, du \tag{2.12}$$

y es muy sencillo ver que la última sumatoria converge en $\Re s > 0$. Por lo tanto, podemos hablar de $\zeta(s)$ como $1/(s-1)$ más una función analítica en la región $\Re s > 0$. Así, hemos extendido el numerador y el denominador en (2.11) a $\Re s > 0$; empero, también debemos controlar cuando el denominador se desvanece. Como el denominador es $\zeta(s)$, controlar cuando es cero en toda la región $\Re s > 0$ equivaldría a la hipótesis de Riemann. Vamos a citar el mejor resultado conocido en dicha dirección, debido a Vinogradov y Korobov:

Teorema 2.2. ([IK04, Thm. 8.29]) *Hay una constante $c > 0$ tal que $\zeta(s) \neq 0$ para $s = \sigma + it$ con $\sigma \geq 1 - c(\log t)^{-2/3}(\log\log t)^{-1/3}$, $|t| \geq 3$, y en dicha zona también se cumplen las cotas*

$$\frac{1}{\zeta(s)} \ll (\log t)^{2/3}(\log\log t)^{1/3},$$
$$\frac{\zeta'(s)}{\zeta(s)} \ll (\log t)^{2/3}(\log\log t)^{1/3}.$$

Con dicho resultado podemos acotar el promedio de λ:

Teorema 2.3.
$$\sum_{n\leq x} \lambda(n) \ll x\exp(-(\log x)^{3/5+o(1)}).$$

Demostración. Aplicamos el Lema 2.1 para $f = \lambda$, y definimos $T = \delta^{-2}$. Por el teorema de Cauchy, podemos desplazar la línea de integración a los segmentos rectos L_-, L_1, L_+, donde L_1 va de $\sigma - iT$ a $\sigma + iT$, L_- va de $1 + 1/\log x - iT$ a $\sigma - iT$ y L_+ de $\sigma + iT$ a $1 + 1/\log x + iT$; aquí $\sigma = 1 - c(\log T)^{-2/3}(\log\log T)^{-1/3}$, y utilizamos el Teorema 2.2 para asegurarnos que $\zeta(2s)/\zeta(s)$ es analítica en la región entre la vieja y la nueva línea de integración. Así,

$$\begin{aligned}\sum_{n\leq x} \lambda(n) &= O(\delta x \log x) + \frac{1}{2\pi i}\int_{L_-+L_1+L_+} x^s M\psi_\delta(s)\frac{\zeta(2s)}{\zeta(s)}\, ds \\ &= O(\delta x\log x) + O(x\delta^3) + \frac{1}{2\pi i}\int_{\sigma-iT}^{\sigma+iT} x^s M\psi_\delta(s)\frac{\zeta(2s)}{\zeta(s)}\, ds,\end{aligned}$$

donde usamos la cota del Teorema 2.2, así como la cota $M\psi_\delta(s) = O(1/\delta|s(s+1)|)$, y el hecho que $\zeta(2s)$ está acotada por $\zeta(2\sigma)$; podemos asumir $2\sigma > 3/2$ (digamos), así que $\zeta(2\sigma) = O(1)$. Como $|x^s| = |x^\sigma| = x\exp\left(-c(\log x)(\log T)^{-2/3}(\log\log T)^{-1/3}\right)$, obtenemos, usando las mismas cotas,

$$\begin{aligned}&\sum_{n\leq x}\lambda(n) = O(\delta x\log x) + O\left(\delta^{-1} x e^{-c(\log x)(\log T)^{-2/3}(\log\log T)^{-1/3}}\log x\right)\\ &\ll e^{-C(\log x)^\alpha(\log\log x)^\beta} x\log x \\ &+ e^{C(\log x)^\alpha(\log\log x)^\beta - c(\log x)(2C(\log x)^\alpha(\log\log x)^\beta)^{-2/3}(\log 2C + \alpha\log\log x + \beta\log\log\log x)^{-1/3}} x\log x\end{aligned}$$

para $\delta = e^{-C(\log x)^\alpha(\log\log x)^\beta}$ y $T = \delta^{-2}$. Está claro que lo óptimo es escoger α y β tales que $\alpha = 1 - 2\alpha/3$ y $\beta = -2\beta/3 - 1/3$, i.e., $\alpha = 3/5$ y $\beta = -1/5$. Asimismo, escogemos C suficientemente pequeño para que $C < c(2C)^{-2/3}\cdot(2\alpha)^{-1/3}$, digamos. Así obtenemos

$$\sum_{n\leq x}\lambda(n) \ll e^{-C(\log x)^{3/5}(\log\log x)^{-1/5}} x\log x \ll e^{-(C/2)(\log x)^{3/5}(\log\log x)^{-1/5}} x$$

para C más grande que una constante. □

Corolario 2.4. *Sean $x \geq 1$, $t \leq \exp^{(\log x)^{3/5-\epsilon}}$, $\epsilon > 0$. Entonces*

$$\sum_{x<n\leq 2x} \frac{\lambda(n)}{n^{1+it}} \ll \exp(-(\log x)^{3/5+o_\epsilon(1)}).$$

Demostración. Por el Teorema 2.3 y sumación por partes (ejercicio 2), obtenemos en verdad que

$$\sum_{x<n\leq 2x}\frac{\lambda(n)}{n^{1+it}} \ll (1+|t|)\exp(-(\log x)^{3/5+o(1)})\log x.$$

□

Es de suponer que la gran mayoría de lectores ya han visto por lo menos una prueba del teorema de los números primos basada sobre la integración compleja y las propiedades de $\zeta(s)$. De todas maneras, demos los primeros pasos de una prueba siguiendo las líneas generales que acabamos de sentar.

Por (2.2), la función $Z(s) = \log \frac{1}{\zeta(s)}$ es de la forma $Z_{1_P} + O(1)$ para $\Re s > 1$, donde $Z_{1_P} = \sum_p p^{-s}$ es la serie que corresponde a la función indicatriz 1_P del conjunto de todos los primos. Un problema con usar esta función $Z(s)$ es que $1/\zeta(1) = 0$, lo cual significa que no podemos definir $Z(s)$ como meromorfa alrededor de $s = 1$. Una manera de solucionar el problema sería integrar por partes en la integral que va de $c - i\infty$ a $c + i\infty$; luego aparecería, en vez de $Z(s)$, su derivada

$$Z'(s) = -\frac{\zeta'(s)}{\zeta(s)} = Z_\Lambda, \tag{2.13}$$

donde $\Lambda(n) = \log p$ para $n = p^\alpha$, p algún primo, y $\Lambda(n) = 0$ para otros n. Debido a (2.12), $Z(s)$ es meromorfa en $\Re s > 0$ con un polo en $s = 1$ de resto 1.

En vez de seguir el procedimiento descrito en el párrafo anterior, es más fácil usar directamente Z_Λ para demostrar el teorema de los números primos en la siguiente forma.

Teorema 2.5. *Para $x > 0$,*

$$\sum_{p \leq x} \log p = x + O(x \exp(-(\log x)^{3/5+o(1)}))$$

.

Demostración. Ejercicio 7. □

Corolario 2.6. *Para $x > 0$,*

$$\pi(x) = Li(x) + O(x \exp(-(\log x)^{3/5+o(1)})), \tag{2.14}$$

$$\sum_{p \leq x} \frac{\log p}{p} = \log x + \kappa_1 + O(\exp(-(\log x)^{3/5+o(1)})), \tag{2.15}$$

$$\sum_{p \leq x} 1/p = \log \log x + \kappa_2 + O(\exp(-(\log x)^{3/5+o(1)})), \tag{2.16}$$

donde κ_1 y κ_2 son constantes.

Por cierto, las fórmulas menos precisas $\sum_{p \leq x} (\log p)/p = \log x + O(1)$ y $\sum_{p \leq x} 1/p = \log \log x + \kappa_2 + O(1/\log x)$ ya eran conocidas antes del teorema de los números primos (Chebyshev-Mertens, años 1848–1874; ver [IK04, §2.2]).

Esbozo de prueba del corolario. Por sumación por partes. En el caso de (2.15), para evitar problemas debidos al hecho que $\exp(-(\log x')^{3/5+o(1)})$ puede ser bastante más grande que $\exp(-(\log x)^{3/5+o(1)})$ para x' mucho más pequeño que x, conviene estimar las sumas

$$\sum_{n > x} \left(1_P(n) \cdot \frac{\log n}{n} - \frac{1}{n} \right), \qquad \sum_{n \leq x} \frac{1}{n}, \tag{2.17}$$

la primera de ellas por sumación por partes, y la segunda por la fórmula $\sum_{n \leq x} 1/n = \log x + \gamma + O(1/x)$, donde γ es una constante (*constante de Euler*). Ésta última fórmula se establece fácilmente: las áreas que quedan entre la hipérbola $y = 1/x$ y las líneas horizontales $y = 1/n$ para $n \leq x \leq n + 1$ pueden ponerse en pila, y entonces está claro que su suma para $n \geq 1$ es una constante, y su suma para $n > x$ es $O(1/x)$. Nótese por último que la primera suma en (2.17), tomada sobre todos los $n \geq 1$, converge a una constante, gracias a la estimación de esa misma suma por sumación por partes para x general. Estas observaciones son suficientes para construir una prueba (ejercicio).

Procedemos de la misma manera para establecer 2.16: estimamos las sumas

$$\sum_{n>x}\left(\frac{1_P(n)}{n}-\frac{1}{n\log n}\right),\qquad \sum_{n\leq x}\frac{1}{n\log n},$$

la primera de ellas por sumación por partes, y la segunda por una comparación con $\int_2^x 1/n\log n$. □

De manera análoga, podemos calcular la probabilidad de que un número no tenga factores primos en un rango dado.

Lema 2.7. *Sea* $1\leq Q^\alpha\leq Q\leq x^{\frac{1}{(\log\log x)^3}}$. *Entonces*

$$\sum_{n<x,p|n\Rightarrow p\notin[Q^\alpha,Q]} 1=\alpha x+x\cdot O\left(\exp\left(-\text{mín}\left((\alpha\log Q)^{3/5+o(1)},\frac{\log x}{3\log Q}\right)\right)\right).$$

Para Q cercano a x, el problema comenzaría a cambiar de cariz (*función de Dickman, función de Buchstab*; ver, por ejemplo, [MV07, §7.1–7.2]).

Demostración. Observemos que la suma del enunciado es $\sum_{n<x} g(n)$, donde g es la función totalmente multiplicativa con $Z_g(s)=\prod_{p\notin[Q^\alpha,Q]}(1-p^{-s})^{-1}=\zeta(s)\prod_{Q^\alpha\leq p\leq Q}(1-p^{-s})$. La función $Z_g(s)$ tiene un polo en $s=1$ con residuo $\prod_{Q^\alpha\leq p\leq Q}(1-p^{-1})$. Gracias a (2.16), vemos que

$$\begin{aligned}\prod_{p\leq z}(1-p^{-1})&=\exp\left(-\sum_{p\leq z}p^{-1}+c+O\left(z^{-1}\right)\right)=e^{-\log\log z+c'+O(\exp(-(\log z)^{3/5+o(1)}))}\\&=\frac{C}{\log z}(1+O(\exp(-(\log z)^{3/5+o(1)})),\end{aligned}$$

donde c, c' y C son constantes. En consecuencia,

$$\prod_{Q^\alpha\leq p\leq Q}(1-p^{-1})=\alpha\cdot(1+O(\exp(-(\alpha\log Q)^{3/5+o(1)}))). \tag{2.18}$$

Por otra parte,

$$|Z_g(\sigma+it)|\leq|\zeta(\sigma+it)|\prod_{p\leq Q}(1+p^{-\sigma}).$$

Para $1-1/\log Q\leq\sigma\leq 1$,

$$\begin{aligned}\prod_{p\leq Q}(1+p^{-\sigma})&\leq\exp\left(\sum_{p\leq Q}p^{-\sigma}\right)\leq\exp\left(Q^{1-\sigma}\sum_{p\leq Q}p^{-1}\right)\\&\leq\exp(e(\log\log Q+\kappa+o(1)))\ll(\log Q)^3,\end{aligned}$$

por el corolario 2.6. Usando (2.12), vemos que, para $s=\sigma+it$ con $\sigma\geq 1-1/\log Q$,

$$\begin{aligned}\zeta(s)&=\frac{1}{s-1}+O\left(\int_1^Q\left(\lfloor u^{-s}\rfloor-u^{-s}\right)du+\int_Q^\infty\left(\lfloor u^{-s}\rfloor-u^{-s}\right)du\right)\\&=\frac{1}{s-1}+O\left(\int_1^Q\frac{du}{u}+|s|\int_Q^\infty u^{-\sigma-1}du\right)=\frac{1}{s-1}+O\left(\log Q+|t|/Q^\sigma\right).\end{aligned}$$

Así, para $|t|\leq Q$, concluímos que $|\zeta(s)|=1/(s-1)+O(\log Q)$. En particular, $|Z_g(\sigma+it)|\ll(\log Q)^4$ para $1\leq|t|\leq Q$.

Procedamos como en la demostración del Teorema 2.3, sólo que con

$$\sigma=1-\text{mín}\left(c(\log T)^{-2/3}(\log\log T)^{-1/3},1/\log Q\right).$$

Obtenemos que

$$\sum_{n\le x} g(n) = \alpha x \cdot (1 + O(\exp(-(\alpha \log Q)^{3/5+o(1)}))) + S,$$

donde

$$\begin{aligned} S &= O(\delta x \log x) + O(x\delta^3(\log Q)^4) \\ &\quad + O\left(\delta^{-1} x e^{-(\log x)\min(c(\log T)^{-2/3}(\log\log T)^{-1/3}, 1/\log Q)} \log x\right). \end{aligned}$$

Si $\log Q \le c(\log T)^{2/3}(\log\log T)^{1/3}$, procedemos exactamente como lo hicimos anteriormente, y obtenemos $S = xO(\exp(-(\log x)^{3/5+o(1)}))$. Si $\log Q > c(\log T)^{2/3}(\log\log T)^{1/3}$, tomamos $\delta = x^{-1/2\log Q}$, $T = \delta^{-2} = x^{1/\log Q}$, y así

$$S = O(\delta x \log x) + O(\delta^3 x (\log x)^4) \ll \frac{x \log x}{x^{\frac{1}{2\log Q}}} \ll x^{1-\frac{1}{3\log Q}}$$

para $Q \le x^{1/(\log\log x)^3}$. □

2.1.1. Ejercicios.

1. Una herramienta útil para nosotros va a ser sustituir sumas por integrales. Vamos a demostrar la *Regla del rectángulo*.
   a) Demuestre que $f(n) = \int_n^{n+1} f(n)\,dt = \int_n^{n+1} f(t)\,dt + O(\int_n^{n+1} |f'(t)|\,dt)$.
   b) Usando el apartado anterior, pruebe que, para enteros $a < b$ cualesquiera,

$$\sum_{a<n\le b} f(n) = \int_a^b f(t)\,dt + O\left(\int_a^b |f'(t)|\,dt\right). \qquad \text{(regla del rectángulo)} \tag{2.19}$$

   Por cierto, existen también expresiones similares con términos de error que involucran f'', f''', etc. (*fórmula de Euler-Maclaurin*).
   c) Observe que si f es monótona, entonces la fórmula del apartado anterior da algo similar al *criterio integral* para series:

$$\sum_{a<n\le b} f(n) = \int_a^b f(t)\,dt + O(|f(b) - f(a)|)$$

   d) Use las fórmulas anteriores para demostrar las siguientes aproximaciones:

$$\sum_{n\le x} n^3 = \frac{x^4}{4} + O(x^3), \qquad \sum_{n\le x} \operatorname{sen}\left(\frac{n^2}{x^{3/2}}\right) = cx^{3/4} + O(\sqrt{x}),$$

   con $c > 0$ la constante $c = \int_0^\infty \frac{\operatorname{sen} u}{2\sqrt{u}}\,du$.
2. *a*) Muestre que, para cualquier entero N y $a_1, \dots, a_N \in \mathbb{C}$, $b_1, \dots, b_N \in \mathbb{C}$ arbitrarios,

$$\begin{aligned} \sum_{1\le n\le N} a_n b_n &= \sum_{1\le n\le N} (A(n) - A(n-1)) \cdot b_n \\ &= A(N)\cdot b_N - \sum_{1\le n\le N-1} A(n)\cdot(b_{n+1} - b_n), \end{aligned} \tag{2.20}$$

   donde $A(x) = \sum_{1\le n\le x} a_n$. (Se trata simplemente de recomponer los términos de una suma.) A esta técnica se la denomina *sumación por partes*. Es utilizada cuando sabemos como estimar las sumas de tipo $A(x)$, y queremos estimar una suma $\sum_{n\le N} a_n b_n$.

b) Poner la igualdad de la forma siguiente hace más claro el paralelismo con la integración por partes (la técnica básica aprendida en un primer curso de cálculo integral). Reescriba (2.20) como sigue: para N, $A_0, A_1, \dots, A_N \in \mathbb{C}$ y $b_1, \dots, b_N \in \mathbb{C}$ arbitrarios,

$$\sum_{1\leq n\leq N} (A_n - A_{n-1})b_n = A_N b_N - A_0 b_1 - \sum_{1\leq n\leq N-1} A_n \cdot (b_{n+1} - b_n). \tag{2.21}$$

Alternativamente, muestre que (2.21) es un caso especial de la integración por partes, formulada para integrales de Lebesgue.

3. *a*) Sea $\tau(n)$ la función que cuenta el número de divisores de n, es decir $\tau(n) = \sum_{d|n} 1$. Demuestre, cambiando el orden de sumación, que

$$\sum_{n\leq x} \tau(n) = \sum_{d\leq x} \left\lfloor \frac{x}{d} \right\rfloor = x \sum_{d\leq x} \frac{1}{d} + O(x) = x\log x + O(x),$$

donde $\lfloor x \rfloor$ es la parte entera de x. (Evidentemente, $\lfloor x \rfloor = x + O(1)$.)

b) *(Comentario)* En otras palabras, la esperanza del número de divisores de un entero aleatorio $n \leq x$ es $\log x + O(1)$. Empero, la mayor parte de enteros tienen un número de divisores bastante menor. Como veremos después (ejercicio 4), la esperanza del número de divisores primos de un $n \leq x$ aleatorio es $(1 + o(1))\log\log x$; Verifique que un número con $(1 + o(1))\log\log x$ divisores primos y sin divisores cuadrados aparte de 1 tiene $(\log x)^{(1+o(1))\log 2}$ divisores.
Lo que sucede es que, si bien los números con $\geq C\log\log x$ divisores primos ($C > 1$) resultan estar en la minoría, tienen un número tan grande de divisores (¿cuántos?) que hacen que el promedio (o esperanza) sea bastante superior a la mediana (el valor tal que mitad de los casos son superiores y mitad son inferiores a él).

c) Use sumación por partes e integración por partes para demostrar las siguientes aproximaciones:

$$\sum_{n\leq x} n\tau(n) = \frac{x^2}{2}\log x + O(x^2), \qquad \sum_{n\leq x} \frac{\tau(n)}{n} = \frac{(\log x)^2}{2} + O(\log x).$$

4. En este problema vamos a demostrar la asintótica $\sum_{p\leq x} \frac{1}{p} = (1 + o(1))\log\log x$ a partir del producto de Euler (2.2). Antes de ello, observa que de dicha asintótica es posible deducir que $\sum_{n\leq x} w(n) = x(1 + o(1))\log\log x$, con $w(n) = \sum_{p|n} 1$.

a) Use el criterio integral para series para demostrar que $\zeta(s) = \sum_{n=1}^{\infty} n^{-s} = \frac{1}{s-1} + O(1)$ para $s > 1$.

b) Tomando logaritmos en la identidad de Euler (2.2), usando el apartado anterior y la aproximación de Taylor $\log \frac{1}{1-x} = x + O(x^2)$ para $|x| < 1/2$, demuestre que

$$\sum_{p} p^{-s} = (1 + o(1))\log \frac{1}{s-1} \qquad \text{para } s > 1.$$

En este contexto particular, $o(1)$ quiere decir "una cantidad una cantidad que tiende a 0 cuando $s \mapsto 1$".

c) Tome $s = 1 + \frac{\log\log x}{\log x}$. Usando $\sum_{p>x} p^{-s} \leq \sum_{n>x} n^{-s}$ y el criterio integral, demuestre que $\sum_{p>x} p^{-s} \ll 1$.

d) A partir de los dos apartados anteriores, demuestre que

$$\sum_{p\leq x} p^{-1} \geq \sum_{p\leq x} p^{-1-\frac{\log\log x}{\log x}} = \sum_{p} p^{-1-\frac{\log\log x}{\log x}} + O(1) = (1 + o(1))\log\log x.$$

e) Para $p \leq x$, demuestre que $p^{-1-\frac{1}{\log\log x \log x}} = p^{-1}(1+o(1))$. A partir de ahí, pruebe las desigualdades

$$\sum_{p \leq x} p^{-1} \leq (1+o(1)) \sum_{p} p^{-1-\frac{1}{\log\log x \log x}} = (1+o(1)) \log\log x.$$

5. *a*) Para $y > 0$, la ecuación de Perron (2.4) nos dice que $1_{(0,1)}(y) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} y^{-s}\frac{ds}{s}$, asumiendo que definimos $1_{(0,1)}(1) = 1/2$. Usando dicha fórmula, y teniendo en cuenta que $1_{(0,b)}(y) = 1_{(0,1)}(y/b)$, demuestre que $1_{(0,b)}(y) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} b^s y^{-s}\frac{ds}{s}$.
   b) Usando la fórmula del apartado anterior, demuestre la expresión (2.8).
   c) Demuestre que si ψ_δ es la función definida en (2.6) entonces se cumple que $\int_0^\infty \psi_\delta(u)u^{s-1}\,du$ es igual a la función $M\psi_\delta(s)$ definida en (2.9). Así, el Teorema de inversión de Mellin también demuestra la fórmula (2.8)).
   d) Recuerde que $|M\psi_\delta(s)| \leq 2/\delta|s(s+1)|$. Pruebe que $|M\psi_\delta(s)| \leq 2/s$ para $\delta \leq 1/2(s+1)$. Concluya que $|M\psi_\delta(s)| \leq 4/s$ para todo $\delta > 0$.
6. Decimos que un entero n es *libre de factores cuadrados* si no es divisible por ningun cuadrado d^2 con d un entero mayor que 1. Vamos a demostrar que la proporción de números sin divisores cuadrados es $1/\zeta(2) = 0{,}6079\ldots$ Para ser precisos: sea $Q(x)$ el número de enteros $n \leq x$ libres de factores cuadrados; mostraremos que $\lim_{x\to\infty} Q(x)/x = \frac{1}{\zeta(2)}$.

   Como veremos en el ejercicio 2 de la sección siguiente, es posible dar una prueba elemental de este mismo enunciado (con un mejor término de error). Ahora mostraremos como probarlo usando la función $\zeta(s)$, simplemente para practicar las técnicas que hemos expuesto en esta sección.
   a) Demuestre que $Z_{\mu^2}(s) = \frac{\zeta(s)}{\zeta(2s)}$. Aplique el Lema 2.1 para expresar $Q(x)$ en términos de una integral sobre el segmento V_0 descrito por $s = 1 + \frac{1}{\log x} + it$, con $-\delta^{-c} \leq t \leq \delta^{-c}$, donde $\delta \in (0,1)$ y $c > 0$ son parámetros que luego elegiremos.
   b) Usando la fórmula (2.12), demuestre que

$$\zeta(s) - \frac{1}{s-1} \ll \sum_{n=1}^{\infty} n^{-\sigma} \min\left(1, \frac{1+|t|}{n}\right) \ll (1+|t|)^{1-\sigma}$$

   para $s = \sigma + it$, $\sigma \geq 1/2$. (Es posible probar cotas más fuertes, comenzando por la *cota de convexidad* $\zeta(s) \ll_{\sigma,\epsilon} (1+|t|)^{(1-\sigma)/2+\epsilon}$; no las necesitaremos.)
   c) Deduzca del apartado anterior y del producto de Euler (2.2) para $\zeta(2s)$ que si estamos en la zona $\sigma_0 \leq \sigma < 3$ para algún $\sigma_0 > 1/2$ y a distancia $\gg 1$ de $s = 1$, entonces $Z_{\mu^2}(s) \ll_{\sigma_0} (1+|t|)^{1-\sigma}$.
   d) Usando el Teorema de los residuos y el apartado *a*), demuestre que

$$Q(x) = \frac{x}{\zeta(2)} + O(\delta x \log x) + \frac{1}{2\pi i}\int_{H_+ \cup V \cup H_-} x^s M\psi_\delta(s) Z_{\mu^2}(s) ds.$$

   con V el segmento $\sigma = \sigma_0$ ($\sigma_0 > 1/2$), $-\delta^{-c} \leq t \leq \delta^{-c}$ y H_+, H_- los segmentos horizontales que unen los extremos de V con los de V_0.
   e) Use las cotas que tenemos para Z_{μ^2} y $M\psi_\delta$ para demostrar que

$$Q(x) = \frac{x}{\zeta(2)} + O(\delta x \log x) + O\left(x\delta^{(1+\sigma)c-1}\right) + O\left(x^\sigma \delta^{\sigma-1}\right).$$

   f) Eliga $\delta > 0$, $\sigma > 1/2$ y c adecuadamente para demostrar que $Q(x) = \frac{x}{\zeta(2)} + O_\epsilon(x^{2/3+\epsilon})$ para $\epsilon > 0$ arbitrario.

   El método elemental que veremos en la siguiente seccion da un mejor exponente: 1/2, en vez de 2/3. Es posible obtener algunas mejoras ulteriores combinando el método elemental y el método analítico.

7. Siga los pasos de la prueba del Teorema 2.3, usando la fórmula (2.13) y el Teorema de Vinogradov-Korobov (Teo. 2.2), para demostrar el teorema de los números primos en la siguiente forma:

$$\sum_{n\leq x} \Lambda(n) = x + O(x\exp(-(\log x)^{3/5+o(1)})).$$

Deduzca el Teorema 2.5.

Definición 2.8. Mencionamos la transformada de Fourier, así que debemos precisar las constantes en nuestra definición. La *transformada de Fourier* $\widehat{f} : \mathbb{R} \to \mathbb{C}$ de una función $f : \mathbb{R} \to \mathbb{C}$ se define por

$$\widehat{f}(t) = \int_{-\infty}^{\infty} f(x)e(-xt)dx$$

donde $e(\alpha) = e^{2\pi i\alpha}$. Requerimos siempre que f esté en L^1, es decir, $\int |f(x)|dx < \infty$ ("f es integrable").

2.2. Las cribas y sus limitaciones: el problema de paridad.

Hagamos una breve pausa, no para ver más resultados que necesitamos, sino para comprender mejor por qué los problemas con los que lidiaremos son difíciles, y no pueden ser resueltos sólo mediante cribas.

Digamos que tenemos un conjunto finito A y un conjunto finito de propiedades P. Para cada subconjunto de P, se nos da una fórmula aproximada – con algún término de error – para el número de elementos de A que satisfacen todas las propiedades en el subconjunto. Se nos pide dar una cota, superior o inferior, para el número de elementos de A que no satisfacen ninguna de las propiedades.

La estrategia evidente sería aplicar el principio de inclusión-exclusión (ejercicio 1). Empero, el número de términos en una inclusión-exclusión completa es exponencial en el número de propiedades, por lo cual el error total podría ser enorme. Nos podemos preguntar si existe alguna manera de usar un número mucho más pequeño de términos para obtener cotas superiores o inferiores, o aún expresiones asintóticas. Una *criba* da una respuesta afirmativa a esta pregunta en contextos comunes en la teoría de números. (En particular, A será un conjunto de enteros, y las propiedades a considerar estarán dadas con una biyección natural con los primos en un conjunto finito.)

Una amplia clase de cribas (*cribas pequeñas*) se pueden poner en el formalismo siguiente. Para permitir multiplicidades, en vez de trabajar con un conjunto A, trabajaremos con reales no negativos a_n para $1 \leq n \leq x$. Las propiedades que consideraremos serán la divisibilidad por un conjunto $\mathscr{P}$ de primos $p \leq z$, donde z es un parámetro. Para cada $d \leq D$ (donde $D \geq z$) libre de factores cuadrados y con factores primos sólo en $\mathscr{P}$, se nos da que

$$\sum_{\substack{1\leq n\leq x\\ d|n}} a_n = g(d)X + r_d, \tag{2.22}$$

donde $g : \mathbb{Z}^+ \to \mathbb{R}$ es una función multiplicativa, r_d es el término de error y X depende sólo de x. Aquí $0 \leq g(p) < 1$ para $p \in \mathscr{P}$. Los términos de error r_d son tal que $\sum_{d\leq D} |r_d|$ sea pequeño comparado con X (digamos, $O(X/(\log X)^A)$).

Hay diversas cribas para diversos problemas. Una criba muy simple nos permite calcular $\sum_{n\leq x:\forall d>1, d^2\nmid n} a_n$. Asimismo, es posible usar una criba para dar cotas superiores e inferiores para $\sum_{n\leq x: n \text{ tiene } \leq 3 \text{ factores primos}} a_n$, digamos, o para dar una cota superior para $\sum_{p\leq x} a_p$. ¿Podemos, empero, dar una cota inferior no trivial para $\sum_{p\leq x} a_p$, o una estimación no trivial de $\sum_{n\leq x} \lambda(n)a_n$, puramente a través de una criba?

La respuesta es no, como lo muestra el siguiente argumento de Selberg. Por una parte, la secuencia $a_n = 1/2$ satisface (2.22) con $g(d) = 1/d$, $|r_d| \leq 1$. Por otra parte, consideremos la secuencia $a'_n = (\lambda(n) + 1)/2$. Vemos que, para $d \leq D = x^{1-\epsilon}$,

$$\sum_{\substack{n \leq x \\ d|n}} a'_n = \frac{1}{2} \sum_{n \leq x/d} 1 + \frac{\lambda(d)}{2} \sum_{n \leq x/d} \lambda(n) = \frac{x}{2d} + O\left(\frac{x/d}{(\log x)^{A+1}}\right)$$

gracias a (1.2). Por lo tanto, (2.22) se cumple para $d \leq D$, nuevamente con $g(d) = 1/d$ y $\sum_{d \leq D} |r_d| = O(x/(\log x)^A)$. En otras palabras, una criba con el dato (2.22) para $d \leq D = x^{1-\epsilon}$ no puede distinguir entre a_n y a'_n. Ahora bien, $\sum_{p \leq x} a'_p = 0$ y $\sum_{n \leq x} \lambda(n) a'_n = \sum a'_n \sim x/2$, sumas que difieren drásticamente de $\sum_p a_p = \sum_{p \leq x} 1/2 \sim x/2 \log x$ y $\sum_{n \leq x} \lambda(n) a_n = \sum_{n \leq x} \lambda(n) = o(x)$.

Por lo tanto, una criba, aún con datos válidos para todo $d \leq x^{1-\epsilon}$, no puede, *por si sóla*, darnos una cota inferior para $\sum_p a_p$, o una estimación no trivial para $\sum_n \lambda(n) a_n$. A este hecho se le llama "problema de la paridad"; fue formulado en la forma que acabamos de ver por Selberg.

La razón del nombre se vuelve más clara si consideramos tambén la secuencia $a''_n = (1 - \lambda(n))/2$. Exactamente por el mismo argumento que acabamos de ver, esta secuencia satisface (2.22) con los mismos valores de $g(d)$ para $d \leq x^{1-\epsilon}$ que las secuencias $\{a_n\}$ y $\{a'_n\}$. Por lo tanto, una criba, por si sola, no puede distinguir entre a'_n y a''_n. Ahora bien, $a'_n = 1$ cuando n tiene un número par de factores primos, y $a'_n = 0$ de lo contrario, mientras que $a''_n = 1$ cuando n tiene un número impar de factores primos, y $a''_n = 0$ de lo contrario.

No utilizaremos métodos de cribas en el trabajo presente. Empero, se trata de herramientas útiles y completamente usuales en la teoría de números, por lo cual un cierto conocimiento de ellas es de suma importancia.

2.2.1. Ejercicios.

1. *a*) Sea A un conjunto finito. Entonces el número de elementos de A que no son ni rojos ni grandes es igual a

$$|A| - |\{a \in A : a \text{ es rojo}\}| - |\{a \in A : a \text{ es grande}\}| + |\{a \in A : a \text{ es rojo y grande}\}|.$$

Éste es un caso especial del *principio de inclusión-exclusión.*

b) He aquí un enunciado general del principio de inclusión-exclusión. Sea A un conjunto finito y $\mathscr{P}$ un conjunto finito de propiedades que los elementos de A pueden o no tener. Para $P \subset \mathscr{P}$, denotemos por A_P el conjunto de elementos de A que satisfacen todas las propiedades en P, y por $A_{\backslash \mathscr{P}}$ el conjunto de elementos de A que no satisfacen ninguna de las propiedades en $\mathscr{P}$. Muestre que

$$\left|A_{\backslash \mathscr{P}}\right| = \sum_{i=0}^{|\mathscr{P}|} (-1)^i \sum_{P \subset \mathscr{P} : |P| = i} |A_P| . \tag{2.23}$$

2. Veamos como aún el principio de inclusión-exclusión, sin más elaboración, puede dar una solución a un problema de criba bastante particular. Recordemos que $Q(x)$ denota el número de enteros $n \leq x$ libres de factores cuadrados.

a) Usando (2.23), muestre que, para $\mathbf{P}$ el conjunto de primos $p \leq \sqrt{x}$,

$$\begin{aligned} Q(x) &= \sum_{i=0}^{|\mathbf{P}|} (-1)^i \sum_{P \subset \mathbf{P} : |P| = i} \left|\{n \leq x : p^2 | n \ \ \forall p \in P\}\right| \\ &= \sum_{d | \prod_{p \in \mathbf{P}} p} \mu(d) |\{n \leq x : d^2 | n\}| = \sum_{d \leq \sqrt{x}} \mu(d) \left|\{n \leq x : d^2 | n\}\right| . \end{aligned}$$

Sugerencia: utilice dos veces el hecho que $d^2|n$ implica $d \leq \sqrt{x}$.

b) Tomando en cuenta que, para todo $\ell \in \mathbb{Z}^+$,

$$|\{n \leq x : \ell|n\}| = \frac{x}{\ell} + O(1),$$

muestre que

$$Q(x) = x \cdot \sum_{d \leq \sqrt{x}} \frac{\mu(d)}{d^2} + O(\sqrt{x}).$$

Concluya que

$$Q(x) = x \cdot \sum_{d=1}^{\infty} \frac{\mu(d)}{d^2} + O(\sqrt{x}) = \frac{x}{\zeta(2)} + O(\sqrt{x}).$$

3. Demuestre las dos siguientes desigualdades, las cuales podríamos llamar versiones “incompletas” de (2.23): para $j \leq |\mathscr{P}|$ par,

$$\left|A_{\backslash \mathscr{P}}\right| \leq \sum_{i=0}^{j} (-1)^i \sum_{P \subset \mathscr{P} : |P| = i} |A_P|, \tag{2.24}$$

mientras que, para $j \leq |\mathscr{P}|$ impar, la desigualdad va en la otra dirección:

$$\left|A_{\backslash \mathscr{P}}\right| \geq \sum_{i=0}^{j} (-1)^i \sum_{P \subset \mathscr{P} : |P| = i} |A_P|. \tag{2.25}$$

El primer resultado no trivial de cribas (*criba pura de Brun*) se basó en estas desigualdades. En la teoría de probabilidades, se llaman *desigualdades de Bonferroni*, aunque el trabajo de Bonferroni [Bon36] es posterior a aquel de Brun [Bru15].

Para ver porque estas desigualdades pueden ser útiles para las cribas, acote el número de términos del las sumas en el lado derecho de (2.24) o (2.25), y compárelo, para j pequeño, con el número $2^{|\mathscr{P}|}$ de términos de la suma para $j = |\mathscr{P}|$. Típicamente, cada cantidad $|A_P|$ se estima de manera aproximada, con un error, y por lo tanto el término de error total crece con el número de términos. Así, las desigualdades (2.24) y (2.25), tomadas juntas, pueden terminar dando un resultado más preciso que la igualdad (2.23).

4. Discutamos algunos resultados generales, concretos y no triviales de cribas, utilizando la notación que acabamos de explicar.

Suponemos siempre que hay una constante $\kappa > 0$ (llamada la *dimensión* de la criba) tal que

$$\frac{V(w)}{V(z)} \leq K \left(\frac{\log z}{\log w}\right)^{\kappa} \tag{2.26}$$

para todo $1 < w \leq z$ y alguna constante K, donde $V(y) = \prod_{p \in \mathscr{P} : p \leq y} (1 - g(p))$, g cumpliendo (2.22). Entonces tenemos el siguiente resultado.

Lema 2.9 (Lema fundamental de las cribas)**.** *Sean dados $\{a_n\}_{n \leq x}$, $\mathscr{P}$, z, D, g, X, r_d, κ con las propiedades mencionadas. Entonces, para $s = (\log D)/\log z$,*

$$\sum_{\substack{n \leq x \\ p|n \Rightarrow p \notin \mathscr{P}}} a_n = (1 + O_\kappa(K^{O(1)} e^{-s})) X V(z) + R, \tag{2.27}$$

donde

$$|R| \leq R(D) = \sum_{\substack{d \leq D \\ p|d \Rightarrow p \in \mathscr{P}}} |r_d|.$$

Demostración. Por la criba (no tan pura) de Brun (o varias otras). Ver, por ejemplo, [FI10, Cor. 6.10] o [HR74, §2.8]. □

Para el ejercicio siguiente, es suficiente saber que

$$\sum_{\substack{n\leq x\\ p|n\Rightarrow p\notin\mathscr{P}}} a_n \ll_\kappa K^{O(1)}XV(z)+R(D), \tag{2.28}$$

lo cual se deduce inmediatamente del Lema 2.9. A tal resultado se le puede llamar *criba de cota superior*. Se puede obtener (con definiciones del término de resto $R(D)$ ligeramente distintas) de muchos procedimientos de criba distintos, e.g., la criba de Selberg ([IK04, §6.4] o cualquier otra fuente), la cual da (2.28) sin la dependencia en K, bajo condiciones muy generales.

5. *a*) Sea $k\geq 2$. Veamos como utilizar una criba de cota superior para acotar el número de primos p tal que $p+k$ también es primo.

   b) Como práctica, utilizando (2.28) con $a_n=1$ para $n\leq x$, muestre que
$$\pi(x)\ll\frac{x}{\log x}.$$
   Use la *fórmula de Mertens*
$$\prod_{p\leq y}\left(1-\frac{1}{p}\right)=\frac{c+o(1)}{\log y},$$
   la cual se deduce del Corolario 2.6, pero es históricamente anterior al teorema de los números primos. (Sólo requiere el método de Chebyshev en su prueba; ver [IK04, §2.2]. Por cierto, la constante c es igual a $e^{-\gamma}$, donde γ es la constante de Euler.)

   c) Sean $k,X\in\mathbb{Z}^+$, $x=X(X+k)$. Para $n\leq x$, sea a_n la sucesión indicatriz de los números $n=m(m+k)$ para algún $1\leq m\leq X$. Muestre que la condición (2.22) se cumple para todo $d\geq 1$ libre de factores cuadrados con
$$g(d)=\prod_{\substack{p|d\\ p|k}}\frac{1}{p}\cdot\prod_{\substack{p|d\\ p\nmid k}}\frac{2}{p}$$
   y $|r_d|\leq\tau(d)$. Verifique también que (2.26) se cumple con $\kappa=2$, $K=K(k)=\prod_{p|k}(1-1/p)/(1-2/p)$ y $\mathscr{P}=\{p\leq z: p \text{ primo}\}$, para z arbitrario. Aquí
$$V(y)=\prod_{p\leq y}(1-g(p)).$$

   d) Usando el Lema 2.9 (o (2.28)), deduzca que, para $Y>X>0$ y $k\in\mathbb{Z}^+$ arbitrarios, el número de enteros $Y<m\leq Y+X$ tales que $m(m+k)$ no tiene factores primos de tamaño $\leq\sqrt{X}$ es
$$\sum_{\substack{n\leq x\\ p|n\Rightarrow p>\sqrt{X}}} a_n\ll_\kappa K(k)^{O(1)}XV(z)+\sqrt{X}\log(X)\ll K(k)^{O(1)}\frac{X}{(\log X)^2}.$$
   (Claro está, $\sum_{d\leq D}\tau(d)=\sum_{d\leq D}\sum_{l|d}1=\sum_{l\leq D}\lfloor D/l\rfloor\ll D\log D$.) Concluya que el número de primos $Y<p\leq Y+X$ tal que $p+k$ también es primo es
$$\ll K(k)^C\frac{X}{(\log X)^2} \tag{2.29}$$
   para alguna constante C. Nótese que $K(k)\ll\prod_{p|k}(1+1/p)$.

Por cierto, se cree que la asíntotica correcta es $K(k)X/(\log X)^2$. (Se trata de un caso especial de las *conjeturas de Hardy y Littlewood.*) No sabemos, empero, siquiera si el número de primos p con $p+2$ primo es infinito o no (*problema de los primos gemelos*). Lo mismo vale para p y $p+k$, $k>2$.

2.3. Estimaciones de valor medio.

Lema 2.10. ([Mon71, Thm. 6.1]); *see also* [IK04, Thm. 9.1]) *Sean* $a_1, a_2, \ldots, a_N \in \mathbb{C}$. *Entonces, para todo* $T \geq 0$,

$$\int_0^T \left|\sum_{n\leq N} a_n n^{it}\right|^2 dt = (T + O(N)) \sum_{n\leq N} |a_n|^2. \tag{2.30}$$

Esta es una cota que no deja de recordar a la *gran criba* (de la cual no precisaremos). Como en ciertas pruebas de la gran criba, es más sencillo probar una cota con un factor de log de más. Expliquemos la prueba así, comenzando con una demostración de esa cota más débil.

Demostración. Hagamos un intento directo e ingenuo. Expandiendo el cuadrado, vemos que

$$\int_0^T \left|\sum_{n\leq N} a_n n^{it}\right|^2 dt = \sum_{n_1,n_2\leq N} a_{n_1}\overline{a_{n_2}} \int_0^T (n_1/n_2)^{it} dt. \tag{2.31}$$

Ahora bien, para todo $r>0$,

$$\int_0^T r^{it} dt = \frac{r^{iT}-1}{i\log r} = O\left(\frac{1}{\log r}\right). \tag{2.32}$$

Puesto que $\log r \geq (r-1)/(r+1)$ para $r\geq 1$ (como podemos verificar comparando derivadas), $\log(n_1/n_2) \geq (n_1-n_2)/(n_1+n_2) \geq (n_1-n_2)/2N$ para $1\leq n_2\leq n_1\leq N$, y así $\log(n_1/n_2) \geq |n_1-n_2|/2N$ para $1\leq n_1,n_2\leq N$. Por lo tanto, la expresión en el lado derecho de (2.31) es

$$T \sum_{\substack{n_1,n_2\leq N\\ n_1=n_2}} a_{n_1}\overline{a_{n_2}} + O\left(\sum_{\substack{n_1,n_2\leq N\\ n_1\neq n_2}} |a_{n_1}\overline{a_{n_2}}| \frac{N}{|n_1-n_2|}\right)$$

$$= T\sum_{n\leq N} |a_n|^2 + O\left(N \sum_{\substack{n_1,n_2\leq N\\ n_1\neq n_2}} \frac{|a_{n_1}|^2+|a_{n_2}|^2}{2} \frac{1}{|n_1-n_2|}\right) = (T + O(N\log N)) \sum_{n\leq N} |a_n|^2.$$

Hemos obtenido casi lo que queríamos. Veamos ahora porqué hay un factor de $\log N$ espurio. La integral en (2.32) no es sino $\widehat{1_{[0,T]}}(-(\log r)/2\pi)$, donde 1_I es la función característica del intervalo I. Estamos en verdad en una situación muy similar a la de la sección §2.1, cuando examinamos la fórmula de Perron (2.4).

Como (2.32) nos muestra, la transformada de Fourier de una función $1_I(t)$ decae como $1/|t|$ cuando $t\to\pm\infty$, mientras que, como es bien conocido y fácil de probar, la transformada de Fourier de una función continua f decae por lo menos tan rápidamente como $1/t^2$ (esto es, es $O(1/t^2)$ cuando $t\to\pm\infty$). El factor de $\log N$ viene de una suma $\sum_{n\leq N} 1/n$; nos lo ahorraremos si tenemos una suma de tipo $\sum_{n\leq N} 1/n^2$, que converge.

Por lo tanto, lo que hacemos es elegir una función f continua tal que $f(t)\geq 1_{[0,T]}(t)$ para todo t, y, al mismo tiempo, $\int f(t)dt$ no es mucho mayor que $\int 1_{[0,T]}(t)dt = T$, y la derivada f', aparte de estar definida en casi todas partes, no tiene grandes fluctuaciones (pues esto afectaría la constante en la cota $\widehat{f}(t) = O(1/t^2)$). Algo estándar es la función "trapecio" $f(t)$ igual a 0 cuando $t \leq -N$ o $t \geq T+N$, igual a 1 para $0 \leq t \leq T$, y lineal afín en $[-N, 0]$ y $[T, T+N]$ (igual a $1+t/N$ en el primer intervalo, y a $1-(t-T)/N$ en el segundo). (Es en verdad el mismo tipo de elección que en (2.6). Otras funciones continuas también servirían; ésta es simplemente conveniente.)

Como $f(t) \geq 1_{[0,T]}(t)$,

$$\int_0^T \left|\sum_{n\leq N} a_n n^{it}\right|^2 dt \leq \int_0^T f(t)\left|\sum_{n\leq N} a_n n^{it}\right|^2 dt \leq \sum_{n_1,n_2\leq N} a_{n_1}\overline{a_{n_2}}\widehat{f}\left(-\frac{\log n_1/n_2}{2\pi}\right). \tag{2.33}$$

Ahora bien, para f la función "trapecio", $\widehat{f}(0) = \int f(t)dt = T+2N$, y (ejercicio 2)

$$\widehat{f}(x) = O\left(\frac{1}{Nx^2}\right).$$

Por lo tanto, el valor absoluto del lado derecho de (2.33) es

$$\sum_{n\leq N} |a_n|^2 (T+2N) + O\left(\sum_{\substack{n_1,n_2\leq N\\ n_1\neq n_2}} \frac{|a_{n_1}|^2+|a_{n_2}|^2}{2}\frac{N}{|n_1-n_2|^2}\right) = (T+O(N))\sum_{n\leq N}|a_n|^2.$$

□

Si en la parte izquierda de la igualdad (2.30) integrásemos en un subconjunto $\mathscr{T}$ de $[0,T]$ en vez de en el total $\mathscr{T} = [0,T]$, entonces parece razonable que se siguiera cumpliendo la desigualdad ($\ll$) sustituyendo en la parte derecha T por algo menor. Esto es lo que demostraremos en la Proposición 2.12. El problema es que para conseguirlo no sirve la técnica de la prueba del Lema 2.10, ya que ésta dependía de la cancelación que se obtiene al integrar $(n_1/n_2)^{it}$ en $[0,T]$, y no está claro que se pueda obtener al integrar sobre subconjuntos $\mathscr{T}$ generales de $[0,T]$. Sin embargo, veremos que es posible sustituir dicha cancelación por la de las sumas $\sum_n n^{it}$ para t fijo, que es lo que establecemos a continuación.

Lema 2.11. *Sea $f : \mathbb{R}^+ \to \mathbb{R}$ definido por $f(x) = 1$ para $0 < x \leq 1$, $f(x) = 2-x$ para $1 < x \leq 2$, y $f(x) = 0$ de lo contrario. Entonces, para $x \geq 1$ y $t \in \mathbb{R}$ arbitrario,*

$$\sum_n f(n/x) n^{it} \ll \frac{x}{(1+|t|)^2} + \sqrt{|t|}\log(1+|t|). \tag{2.34}$$

Como de costumbre, la elección de f no tiene nada de particular; cualquier f doblemente diferenciable con f'' en L^1 nos daría una cota de la misma forma que (2.34).

Demonstración, o más bien dicho un comentario; la demonstración está en los ejercicios. En los ejercicios 3–7, para ser precisos. Se trata de una aplicación de la fórmula de sumación de Poisson, seguidas de cotas generales para integrales del tipo $\int_a^b \eta(x)e(\theta(x))dx$, η contínua y θ en C^1. (Se trata de cotas superiores, no de asintóticas, para los dos casos principales – fase no estacionaria y fase estacionaria con $\theta'' \neq 0$.) Algunos notarán que las mismas cotas aparecen en el método de van der Corput. De hecho, los ejercicios están cercanos al tratamiento clásico de tal método ([Mon94, §3.3] o [IK04, §8.3]), con las mejoras que son de esperarse dado que usamos un peso liso f. Si no se usa tal f, se obtiene un resultado ligeramente peor:

$$\sum_{n\leq x} n^{it} \ll \frac{x}{1+|t|} + \sqrt{|t|}\log(1+|t|). \tag{2.35}$$

□

Esbozo de otra demonstración. Por el teorema de inversión de Mellin,

$$\sum_n f(n/x)n^{it} = \frac{1}{2\pi i}\int_{\sigma-i\infty}^{\sigma+i\infty} x^s Mf(s)\zeta(s-it)ds$$

para $\sigma > 1$. Desplazamos la integral hacia la izquierda, pasando por polos en $s = 1$ y en $s = 0$. Las contribuciones de los polos dan los términos del lado derecho de (2.34).

(Aquí estamos utilizando el hecho que la función zeta tiene continuación analítica a todo el plano complejo, así como la cota estándar $|\zeta(it)| \ll \sqrt{t}\log t$, y cotas similares para $|\zeta(\sigma+it)|$, $\Re\sigma < 0$. Para obtener tales cotas, es necesario utilizar la ecuación funcional de $\zeta(s)$ (la cual también da la continuación analítica). Como la prueba usual de la ecuación funcional se basa sobre la formulación de Poisson, esta demostración y la anterior no son tan distintas como parecen a primera vista.) □

Proposición 2.12 (Halász–Montgomery)**.** *Sean $T \geq 2$ y $\mathscr{T} \subset [-T,T]$ medible. Sean $a_1, a_2, \dots, a_N \in \mathbb{C}$. Entonces*

$$\int_{\mathscr{T}} \left|\sum_{n\leq N} a_n n^{it}\right|^2 dt \ll (N + |\mathscr{T}|\sqrt{T}\log T)\sum_{n\leq N}|a_n|^2. \tag{2.36}$$

El lector avisado notará que la prueba tiene gran similitud con una de las pruebas de la *gran criba* (distinta de las *pequeñas cribas* que consideramos en §2.2). La diferencia consistirá en que aplicamos el Lema 2.11 en vez de cotas sobre sumas exponenciales. Parte del procedimiento (utilización de un f continuo) es como en la prueba del Lema 2.10 que acabamos de ver.

Demostración. Por el principio de dualidad (ejercicio 8d), basta mostrar que, para $a \in L^2(\mathscr{T})$,

$$\sum_{n\leq N}\left|\int_{\mathscr{T}} a(t)n^{it}\right|^2 \ll (N + |\mathscr{T}|\sqrt{T}\log T)\int_{\mathscr{T}} |a(t)|^2 dt.$$

Claro está, para cualquier $f : [0,\infty) \to [0,\infty)$ con $f(x) \geq 1$ para $0 \leq x \leq 1$,

$$\sum_{n\leq N}\left|\int_{\mathscr{T}} a(t)n^{it}\right|^2 \leq \sum_n f\left(\frac{n}{N}\right)\left|\int_{\mathscr{T}} a(t)n^{it}\right|^2.$$

Expandiendo, vemos que

$$\sum_n f\left(\frac{n}{N}\right)\left|\int_{\mathscr{T}} a(t)n^{it}\right|^2 = \int_{\mathscr{T}}\int_{\mathscr{T}} \overline{a(t_1)}a(t_2)\sum_n f\left(\frac{n}{N}\right) n^{i(t_2-t_1)}dt_1dt_2. \tag{2.37}$$

Usando la desigualdad $|a_1a_2| \leq (|a_1|^2 + |a_2|^2)/2$ y el Lema 2.11, concluímos que el lado derecho de (2.37) es a lo más

$$\begin{aligned}
&\int_{\mathscr{T}} |a(t_1)|^2 \int_{\mathscr{T}} \sum_n f\left(\frac{n}{N}\right) n^{i(t_2-t_1)}dt_1dt_2 \\
&\qquad \ll \int_{\mathscr{T}} |a(t_2)|^2 \int_{\mathscr{T}} \left(\frac{N}{(1+|t_1-t_2|)^2} + \sqrt{T}\log T\right) dt_1dt_2 \\
&\qquad \ll \left(N + |\mathscr{T}|\sqrt{T}\log T\right)\int_{\mathscr{T}} |\overline{a(t_2)}|^2 dt_2,
\end{aligned}$$

que era lo que queríamos demostrar. □

2.3.1. Ejercicios.

1. En este problema vamos a ver que es sencillo demostrar el Lema 2.11 en cierto rango para t. Demostrarlo para todo t será más complicado, como veremos en los siguientes ejercicios.
   a) Demuestre que podemos extender la regla del rectángulo (2.19) a funciones complejas $f : \mathbb{R} \to \mathbb{C}$ usando la desigualdad $|a| + |b| \leq \sqrt{2}\sqrt{|a|^2 + |b|^2}$.
   b) Aplicando el apartado anterior, pruebe la desigualdad (2.35) en el rango $|t| \leq \sqrt{x/\log x}$.
   c) Demuestre el Lema 2.11 en el rango $|t| \leq (x/\log x)^{1/3}$.
2. Sean $T, D > 0$. Sea

$$f(t) = \begin{cases} 0 & \text{si } t \leq -D \text{ o } t \geq T + D, \\ 1 + t/D & \text{si } -D \leq t \leq 0, \\ 1 & \text{si } 0 \leq t \leq T, \\ 1 - \frac{t-T}{D} & \text{si } T \leq t \leq T + D. \end{cases} \tag{2.38}$$

   a) Muestre, por integración por partes, que, para $t \neq 0$,

$$\begin{aligned} \widehat{f}(t) &= \frac{1}{2\pi i t}\left(\int_{-D}^{0} + \int_{0}^{T} + \int_{T}^{T+D} f'(x)e(-xt)dx\right) \\ &= \frac{1}{2\pi i D t}\left(\int_{-D}^{0} e(-xt) - \int_{T}^{T+D} e(-xt)dx\right). \end{aligned}$$

   b) Concluya que, para $t \neq 0$,

$$|\widehat{f}(t)| \leq \frac{1/D}{(\pi t)^2}.$$

   c) En general, sea $f : \mathbb{R} \to \mathbb{C}$ tal que (a) $f(x), f'(x) \to 0$ cuando $x \to \pm\infty$, (b) f'' es continua e integrable (i.e., $|f''|_1 = \int_{-\infty}^{\infty} |f''(x)|dx$ es finita). Usando integración por partes dos veces, muestre que, para $t \neq 0$,

$$|\widehat{f}(t)| \leq \frac{|f''|_1}{(2\pi t)^2}.$$

   Nota. La condición que f'' sea continua puede ser reemplazada por condiciones bastante más débiles; basta con usar versiones más generales de la integración por partes (mediante integrales de Riemann-Stieltjes, o distribuciones, medidas, etc.). Basta con que f sea continua, f' sea definida fuera de un conjunto finito de puntos, f' sea de *variación total* finita y $f(x), f'(x) \to 0$ cuando $x \to \pm\infty$. Un enunciado de esta generalidad cubre la función f en (2.38).
3. Sea $f : \mathbb{R} \to \mathbb{C}$ continua e integrable, así como diferenciable fuera de un conjunto finito de puntos. Asumamos también que f' es integrable (en todo intervalo donde está definida) y que $\sum_{n\in\mathbb{Z}} |\widehat{f}(n)| < \infty$. Entonces

$$\sum_{n\in\mathbb{Z}} f(n) = \sum_{n\in\mathbb{Z}} \widehat{f}(n). \qquad \text{(fórmula de sumación de Poisson)} \tag{2.39}$$

   Veamos como probar esta fórmula.
   a) Muestre que la función $F : \mathbb{R} \to \mathbb{C}$ dada por $F(t) = \sum_{n\in\mathbb{Z}} f(t+n)$ está bien definida (es decir, la suma converge). Verifique también que F es de período 1 y que $\int_0^1 |F(t)|dt < \infty$.
   b) El teorema de inversión de Fourier (para funciones de período 1) nos dice que

$$F(t) = \sum_{n\in\mathbb{Z}} a_n e(nt),$$

donde $a_n = \int_0^1 F(t)e(-nt)dt$, bajo la condición que $\sum_n |a_n| < \infty$. Muestre que $a_n = \widehat{f}(n)$.

c) Concluya que

$$\sum_{n\in\mathbb{Z}} f(n) = F(0) = \sum_{n\in\mathbb{Z}} \widehat{f}(n).$$

4. *Fase no estacionaria.* Sean $\theta, \eta : [a,b] \to \mathbb{R}$ tales que η y θ' son continuas. Asumamos que $\theta'(x) \neq 0$ para todo $x \in [a,b]$. (Ésta es la suposición crucial, llamada "fase no estacionaria", pues $\theta(x)$ es la "fase".) Asumamos también que $g(x) = \eta(x)/\theta'(x)$ es monótona en $[a,b]$.

   Muestre, por intregración por partes, que

$$\int_a^b \eta(x)e(\theta(x))dx = \frac{1}{2\pi i}\left(g(x)e(\theta(x))|_a^b + \int_a^b e(\theta(x))dg(x)\right).$$

   Concluya que

$$\left|\int_a^b \eta(x)e(\theta(x))dx\right| \leq \frac{\max(|g(a)|,|g(b)|)}{\pi}. \tag{2.40}$$

   En particular, para θ' monótona con $\theta'(x) \neq 0$ para todo $x \in [a,b]$,

$$\left|\int_a^b e(\theta(x))dx\right| \leq \frac{1/\pi}{\min(|\theta'(a)|,|\theta'(b)|)}. \tag{2.41}$$

   Deduzca de (2.40) la siguiente variante: sea $\theta : [a,b] \to \mathbb{R}$ diferenciable, con θ' continua y $\theta'(x) \neq 0$ para todo $x \in [a,b]$, y $\eta_1, \eta_2 : [a,b] \to \mathbb{R}$ tales que $\eta_1(x)$ y $\eta_2(x)/\theta'(x)$ son monótonas. Entonces, para $\eta = \eta_1 \cdot \eta_2$,

$$\left|\int_a^b \eta(x)e(\theta(x))dx\right| \leq \frac{c}{\pi} \cdot \max\left(\frac{|\eta_2(a)|}{|\theta'(a)|}, \frac{|\eta_2(b)|}{|\theta'(b)|}\right) \tag{2.42}$$

   con $c = \max(|\eta_1(a)|, |\eta_1(b)|, |\eta_1(a) - \eta_1(b)|)$. Es fácil obtener otras variantes: por ejemplo, si, en vez de η_1 monótona en $[a,b]$, tenemos η_1 monótona en $[a,x_0]$ y en $[x_0,b]$ para algún $x_0 \in [a,b]$, y $\eta_1(a) = \eta_1(b) = 0$, entonces (2.42) vale con $c = |\eta_1(x_0)|$.

5. Sea $\theta : [a,b] \to \mathbb{R}$ doblemente diferenciable. Asumamos que $\theta''(x) \geq \rho > 0$ para todo $x \in [a,b]$.

   Para $\delta > 0$ arbitrario, podemos tener $|\theta'(x)| < \delta$ sólo dentro de un intervalo I de longitud $\leq 2\delta/\rho$. (¿Por qué?) Por lo tanto,

$$\int_I e(\theta(x))dx \leq 2\delta/\rho.$$

   (Ésta es la contribución de la región de *fase estacionaria*, es decir, de una vecindad del punto en el cual $\theta'(x) = 0$, si tal punto existe.) Aplique (2.41) para mostrar que

$$\int_{I\setminus[a,b]} e(\theta(x))dx \leq 2/\pi\delta.$$

   Escoja δ de manera óptima para así concluir que

$$\int_a^b e(\theta(x))dx \leq \frac{2}{\sqrt{\pi\rho}}. \tag{2.43}$$

   Aquí también podemos introducir un peso η. Por ejemplo, sea $\eta : [a,b] \to [0,\infty)$ continua tal que η es monótona en $[a,x_0]$ y en $[x_0,b]$ para algún $x_0 \in [a,b]$, y además

$\eta(a) = \eta(b) = 0$. Deduzca de (2.43) que

$$\int_a^b \eta(x)e(\theta(x))dx \leq \frac{2\,\text{máx}_{x\in[a,b]}|\eta(x)|}{\sqrt{\pi\rho}}. \tag{2.44}$$

6. Sean $\theta : [a,b] \to \mathbb{R}$ diferenciable tal que θ' es monótona y $\theta'(x) \neq 0$ para todo $x \in [a,b]$. Sea $\eta : [a,b] \to [0,\infty)$ continua, así como diferenciable fuera de a lo más un conjunto finito de puntos; asumamos también que η' es decreciente y acotada, y que $\eta(a) = \eta(b) = 0$.

   Como el signo de $\theta'(x)$ es constante, y como podemos hacer un cambio de variables $x \mapsto b + a - x$ sin cambiar lo que suponemos sobre η, podemos asumir sin pérdida de generalidad que θ' es creciente y positiva en $[a,b]$. Sea $\alpha = \theta'(a)$.

   Muestre que, por integración por partes,

$$\int_a^b \eta(x)e(\theta(x))dx = -\frac{1}{\alpha}\int_a^b \left(\frac{\eta'(x)}{2\pi i} + \eta(x)(\theta'(x)-\alpha)\right)e(\theta(x))dx. \tag{2.45}$$

   Aplique (2.42) y obtenga que

$$\left|\int_a^b \left(\frac{\eta'(x)}{2\pi i} + \eta(x)(\theta'(x)-\alpha)\right)e(\theta(x))dx\right| \leq \frac{\eta'(a)-\eta'(b)}{2\pi^2\alpha} + \frac{\text{máx}_{x\in[a,b]}|\eta(x)|}{\pi\alpha}.$$

   Por lo tanto

$$\left|\int_a^b \eta(x)e(\theta(x))dx\right| \leq \frac{c}{\text{mín}(|\theta'(a)|,|\theta'(b)|)^2} \tag{2.46}$$

   donde $c = (\eta'(a)-\eta'(b))/2\pi^2 + (\text{máx}_{x\in[a,b]}|\eta(x)|)/\pi$.

7. Sea $\eta : \mathbb{R} \to \mathbb{R}$ dada por $\eta(x) = 2x - 1$ para $1/2 \leq x \leq 1$, $\eta(x) = 2 - x$ para $1 \leq x \leq 2$ y $\eta(x) = 0$ cuando $x < 1/2$ o $x > 2$. Quisiéramos estimar

$$\sum_n \eta(n/X)x^{it}$$

   para $t \neq 0$ arbitrario y $X > 1/2$. (Para $0 < X \leq 1/2$, la suma es trivialmente igual a 0.)

   Por la formula de sumación de Poisson (2.39),

$$\sum_n \eta(n/X)x^{it} = \sum_{n\in\mathbb{Z}}\int_0^\infty \eta(x/X)x^{it}e(-nx)dx,$$

   asumiendo que la suma en el lado derecho converge absolutamente.

   a) Estimemos primero la contribución de $n = 0$. Por integración por partes,

$$\int_0^\infty \eta(x/X)x^{it}dx = -\int_0^\infty \frac{\eta'(x/X)}{X}\frac{x^{it+1}}{it+1}dx.$$

   Utilice integración por partes una vez más para concluir que

$$\int_0^\infty \eta(x/X)x^{it}dx \ll \frac{X}{(|t|+1)^2}.$$

   b) Sea $n \in \mathbb{Z}$ tal que ya sea $|n| \geq 2t/\pi X$ o n es de signo contrario a t. Use (2.46) para mostrar que

$$\int_0^\infty \eta(x/X)x^{it}e(-nx)dx \ll \frac{1}{n^2}. \tag{2.47}$$

   c) Concluya que, para $X \geq 2|t|/\pi$,

$$\sum_n \eta(n/X)n^{it} \ll \frac{X}{(|t|+1)^2} + 1. \tag{2.48}$$

d) Sea ahora $X < 2|t|/\pi$. Muestre que (2.44) nos da que

$$\sum_n \eta(n/X)x^{it}e(-nx) \ll \frac{X}{\sqrt{|t|}}$$

para n arbitrario, y en particular para $|n| \geq 2t/\pi X$. Concluya, usando también (2.47), que

$$\sum_n \eta(n/X)n^{it} \ll \sqrt{|t|}. \tag{2.49}$$

e) Sea $x \geq 1$. Sea f_x la función $t \mapsto f(t/x)$, donde f es como en el enunciado del Lema 2.11. Exprese f_x como una suma de funciones del tipo $\eta(n/X)$ para diversos valores de X, y deduzca de (2.48) y (2.49) que

$$\sum_n f(n/x)n^{it} \ll \frac{x}{(|t|+1)^2} + \sqrt{|t|}(\log|t|+1) + \log x$$

para $t \geq 1$. En otras palabras, el Lema 2.11 es cierto.

8. *Dualidad.*

a) Sean V, W espacios de Hilbert.[1] Definimos la norma $|A|$ de un operador (es decir, una función lineal) $A : V \to W$ por

$$|A| = \sup_{\substack{v\in V\\ v\neq 0}} \frac{|Av|_2}{|v|_2}.$$

Se dice que A es acotado si su norma es finita. Todo A acotado tiene un (único) *operador dual* $A^* : W \to V$, definido como la función lineal tal que

$$\langle A^*v, w\rangle = \langle v, Aw\rangle.$$

b) Sean X, Y son dos espacios de Lebesgue de medida finita. Sea $V = L^2(X)$ y $W = L^2(Y)$. Definamos $A : V \to W$ por

$$(Av)(y) = \int_X K(x,y)v(x)dx,$$

donde $K : X \times Y \to \mathbb{C}$ es una función de imagen acotada. (Se comprende que $v \in V$ es una función $v : X \to \mathbb{C}$, y $Av \in W$ es una función $w : Y \to \mathbb{C}$.) Muestre que la función $A^* : W \to V$ definida por

$$(A^*w)(x) = \int_Y \overline{K(x,y)}w(y)dy$$

es el operador dual de A.

c) Sean V, W espacios de Hilbert. Sea $A : V \to W$ un operador acotado. La desigualdad de Cauchy-Schwarz[2] nos dice que $\langle w_1, w_2\rangle \leq |w_1|_2|w_2|_2$ para $w_1, w_2 \in W$, con igualdad si $w_1 = \lambda w_2$ para algún $\lambda \in \mathbb{C}$. Deduzca que, para $v \in V$ arbitrario, $|Av|_2 = \sup_{w\in W: w\neq 0}\langle w, Av\rangle/|w|_2$. Usando este hecho, muestre que

$$\sup_{\substack{v\in V\\ v\neq 0}} \frac{|Av|_2}{|v|_2} = \sup_{\substack{v\in V\\ v\neq 0}} \sup_{\substack{w\in W\\ w\neq 0}} \frac{\langle w, Av\rangle}{|v|_2|w|_2} = \sup_{\substack{w\in W\\ w\neq 0}} \sup_{\substack{v\in V\\ v\neq 0}} \frac{\langle A^*w, v\rangle}{|v|_2|w|_2} = \sup_{\substack{w\in W\\ w\neq 0}} \frac{|A^*w|_2}{|w|_2}.$$

[1]Recordamos que un *espacio de Hilbert* es un espacio lineal V sobre $\mathbb{R}$ o $\mathbb{C}$ con un producto escalar $\langle\cdot,\cdot\rangle$ tal que V es completo con respecto a la distancia $d(x,y) := |x-y|_2 := \sqrt{\langle x-y, x-y\rangle}$. El único ejemplo que necesitaremos es el siguiente: sea X un espacio de Lebesgue de medida finita; definamos el producto escalar $\langle v_1, v_2\rangle = \int_X \overline{v_1(x)}v_2(x)dx$ para $v_1, v_2 : X \to \mathbb{C}$ medibles; entonces el espacio lineal $L^2(X)$ de funciones $v : X \to \mathbb{C}$ con $|v|_2 < \infty$ es un espacio de Hilbert.

[2]Ya sabemos que el nombre históricamente correcto es *Cauchy* o *Cauchy-Bunyakovsky-Schwarz*, pero todo el mundo sabe lo que *Cauchy-Schwarz* quiere decir.

Concluya que

$$|A| = |A^*| \qquad \text{(principio de dualidad).} \tag{2.50}$$

d) Sean X e Y dos espacios de Lebesgue de medida finita. Sea $K : X \times Y \to \mathbb{C}$ una función acotada y $C \geq 0$ una constante tal que, para todo $v \in L^2(X)$,

$$\int_Y \left| \int_X K(x,y)v(x)dx \right|^2 dy \leq C \int_X |v(x)|^2 dx.$$

Concluya que, para todo $w \in L^2(Y)$,

$$\int_X \left| \int_Y K(x,y)w(y)dy \right|^2 dx \leq C \int_Y |w(y)|^2 dy.$$

3. Cancelación de λ en intervalos cortos, en promedio

3.1. Un primer tratamiento. Queremos demostrar el Teorema 1.1. La demostración pasará por la función $Z_\lambda(s) = \zeta(2s)/\zeta(s)$, asociada con λ. En este contexto, es más fácil usar sumas donde los intervalos estén expresados en términos de n/x.

Teorema 3.1. *Sea $h = h(X)$ tal que $h(X) \to \infty$ cuando $X \to \infty$. Entonces, cuando $X \to \infty$,*

$$\mathbb{E}_{(1-\frac{h}{X})N < n \leq N} \lambda(n) = o_h(1)$$

para todo entero $N \in [X, 2X]$ fuera de un conjunto de $o(X)$ excepciones.

Aquí hemos usado la notación $\mathbb{E}_{n\in A} f(n) = \frac{1}{|A|}\sum_{n\in A} f(n)$. En verdad podemos leer $\mathbb{E}_{X<n\leq 2X}$, $\mathbb{E}_{N<n\leq(1+\frac{h}{X})N}$, etc., como $(1/X)\sum_{X<n\leq 2X}$, $(1/(X/h))\sum_{N<n\leq(1+\frac{h}{X})N}$, etc., ya que $|\{X < n \leq 2X\}| = X + O(1)$, y estamos trabajando asintóticamente.

Es sencillo ver que el Teorema 1.1 es una consecuencia del Teorema 3.1 (ejercicio 1), así que vamos a concentrarnos en demostrar este último. Es una consecuencia directa de la siguiente cota de varianza.

Teorema 3.2. *Sea $h = h(X)$ tal que $h(X) \to \infty$ cuando $X \to \infty$. Entonces, cuando $X \to \infty$,*

$$\mathbb{E}_{X<x\leq 2X} |\mathbb{E}_{(1-\frac{h}{X})x<n\leq x} \lambda(n)|^2 = o_h(1).$$

Nuestra tarea en esta sección será probar este teorema, con, por cierto, una cota precisa en vez de $o_h(1)$ (Teorema 3.17). Para empezar, veamos la relación entre las sumas $\mathbb{E}_{(1-\frac{h}{X})x<n\leq x} f(n)$ y $Z_f(s)$.

Proposición 3.3. *Sea $f : \mathbb{N} \to \mathbb{C}$ una función con soporte en $(X, 2X]$ tal que $|f(x)| \leq 1$ para todo x. Entonces si $x \in (X, 2X]$*

$$\mathbb{E}_{(1-\frac{h}{X})x<n\leq x} f(n) = \frac{1}{2\pi i}\int_{1-i\frac{X}{h\delta^2}}^{1+i\frac{X}{h\delta^2}} x^{s-1} M\psi_{\delta,h}(s) Z_f(s)\, ds + O(\delta)$$

para cierta función $\psi_{\delta,h} : (0, +\infty) \to \mathbb{R}$ con $M\psi_{\delta,h}(s) \ll \min(1, \frac{X/h}{|s|}, \frac{X/h}{\delta|s|^2})$ para $\Re s > 0$.

Demostración. La suma inicial puede escribirse como $\frac{1}{x}\sum_n f(n)\frac{1}{h/X}1_{(1-\frac{h}{X},1]}(\frac{n}{x})$. Lo que vamos a hacer es, como en el Lema 2.1, aproximar la función $\frac{1}{h/X}1_{(1-\frac{h}{X},1)}$ por una función $\psi_{\delta,h}$ que (a) tenga soporte en $[1-\frac{h}{X}, 1]$, (b) valga $\frac{1}{h/X}$ en $[1-(1-\delta)\frac{h}{X}, 1-\delta\frac{h}{X}]$, y entre 0 y $\frac{1}{h/X}$ en el resto del soporte. Podemos tomar una función "trapecio" similar a (2.6). Entonces

$$\psi_{\delta,h}(y) = \frac{1}{2\pi i}\int_{1-i\infty}^{1+i\infty} M\psi_{\delta,h}(s) y^{-s} ds$$

con $M\psi_{\delta,h}$ cumpliendo las condiciones del enunciado. Por lo tanto,

$$\mathbb{E}_{(1-\frac{h}{X})x<n\leq x} f(n) = \frac{1}{2\pi i}\int_{1-i\infty}^{1+i\infty} x^{s-1} M\psi_{\delta,h}(s) Z_f(s)\, ds + O(\delta).$$

Como f es acotada y tiene soporte en $[X,2X]$, tenemos que $|Z_f(1+it)| \ll 1$. Por el decaimiento de $M\psi_{h,\delta}$ podemos cortar la integral a altura $\delta^{-2}X/h$ perdiendo sólo $O(\delta)$. $\square$

Ahora vamos a aplicar esta expresión para evaluar el promedio de sumas cortas de f.

Proposición 3.4. *Sea $f:\mathbb{N}\to\mathbb{C}$ una función con soporte en $(X,2X]$ tal que $|f(x)|\leq 1$ para todo x. Entonces*

$$\mathbb{E}_{X<x\leq 2X} |\mathbb{E}_{(1-\frac{h}{X})x<n\leq x} f(n)|^2 \ll \int_{1-i\frac{X}{h\delta^2}}^{1+i\frac{X}{h\delta^2}} |Z_f(s)|^2 ds + O(\delta)$$

Demostración. Queremos usar la Proposición 3.3, expandir el cuadrado e intentar aprovechar la sumación en x. Para que esto funcione mejor, de nuevo es conveniente introducir una función suave que reemplaze a $1_{[X,2X]}$. En este caso

$$\mathbb{E}_{X<x\leq 2X} |\mathbb{E}_{(1-\frac{h}{X})x<n<x} f(n)|^2 \ll \int \eta\left(\frac{x}{X}\right) |\mathbb{E}_{(1-\frac{h}{X})x<n<x} f(n)|^2 \frac{dx}{x} \tag{3.1}$$

para cierta función η con $0\leq \eta(t)\leq 1$, soporte en $[1/2,3]$, C^2 y tal que η'' es integrable. Ahora, usando la Proposición 3.3 y expandiendo el cuadrado tenemos que la parte derecha de (3.1) es

$$O(\delta) + \frac{1}{(2\pi)^2}\int_{1-i\frac{X}{h\delta^2}}^{1+i\frac{X}{h\delta^2}}\int_{1-i\frac{X}{h\delta^2}}^{1+i\frac{X}{h\delta^2}} (M\psi_{\delta,h})(s_1)\overline{M\psi_{\delta,h}(s_2)} Z_f(s_1)\overline{Z_f(s_2)} I(t_1,t_2) ds_1 ds_2,$$

donde $t_i = \Re s_i$, y con

$$I(t_1,t_2) = \int \eta\left(\frac{x}{X}\right) x^{i(t_1-t_2)}\frac{dx}{x} = X^{i(t_1-t_2)}\int \eta(e^y) e^{i(t_1-t_2)y}\, dy.$$

Por integración por partes, $I(t_1,t_2) \ll (1+|t_1-t_2|)^{-2}$. (En general, para f doblemente diferenciable tal que f y f'' son integrables, tenemos $\widehat{f}(t) \ll (1+|t|)^{-2}$, por integración por partes.) Así, usando la cota $M\psi_{\delta,h}\ll 1$ y $|Z_f(s_1)Z_f(s_2)| \leq |Z_f(s_1)|^2 + |Z_f(s_2)|^2$ obtenemos el resultado. $\square$

Una estimación de valor medio (Lema 2.10) permite mostrar que tenemos una desigualdad en la otra dirección. Esta desigualdad nos dice que no hemos perdido nada con respecto a la cota trivial.

Proposición 3.5. *Si f tiene soporte en $(Y,2Y]$, entonces*

$$\int_{1-iY}^{1+iY} |Z_f(s)|^2\, |ds| \ll \mathbb{E}_{Y<n\leq 2Y} |f(n)|^2.$$

Demostración. Por el Lema 2.10. $\square$

La ganancia sobre esta cota trivial va a venir de factorizar $Z_f(s) = Z_{f_1}(s)Z_{f_2}(s)$.

Proposición 3.6. *Si f_1 es una función con soporte en $(Q,2Q]$ y f_2 es una función con soporte en $\left(\frac{X}{Q},\frac{2X}{Q}\right]$, entonces, para todo intervalo $\mathcal{T}\subset[1-i\frac{X}{Q}, 1+i\frac{X}{Q}]$,*

$$\int_{\mathcal{T}} |Z_{f_1}(s)Z_{f_2}(s)|^2 |ds| \ll M_1^2\, \mathbb{E}_{\frac{X}{Q}<n\leq 2\frac{X}{Q}} |f_2(n)|^2$$

con $M_1 = \max_{s\in\mathcal{T}} |Z_{f_1}(s)| \leq \max|f_1|$.

Demostración. Sacamos el máximo fuera de la integral y aplicamos la proposición 3.5. $\square$

Para usar el resultado anterior, queremos factorizar $Z_f(s)$. La idea es usar la factorización de un número n en $n = pm$ con p su primo más pequeño en cierto intervalo $P_0 < p < Q_0$. Para poder hacer esto primero necesitamos ver que la mayoría de números n tienen algún primo en dicho intervalo, lo cual está asegurado por el Lema 2.7 si tomamos $P_0 = Q_0^{\alpha}$ grande con α pequeño. Ahora, si consideramos los números con primos en un intervalo $(P_0, Q_0]$, $P_0 = Q_0^{\alpha}$, y si sacamos el primo más pequeño en dicho intervalo tenemos la factorización

$$\sum_{\substack{X<n\leq 2X\\ \exists p\in(P_0,Q_0],p|n}} f(n)n^{-s} = \sum_{P_0<p\leq Q_0} f(p)p^{-s} \sum_{\substack{m\\ X<mp\leq 2X\\ p'|m\Rightarrow p'\notin[P_0,p)}} f(m)m^{-s} \tag{3.2}$$

para f totalmente multiplicativa (con p' primo). El problema es que esta no es una factorización en dos funciones zeta, ya que en el sumatorio de dentro aparece la variable p en dos condiciones.

Eliminar la dependencia en p de la condición $X/p < m \leq 2X/p$ es algo completamente de rutina. Eliminar la dependencia en p de la condición $p' \notin (P_0, p)$ también es factible. Lo haremos de manera ligeramente distinta a [MR16] (o al artículo expositorio [Sou17], el cual también da su propia versión).

Primero, un poco de notación. Para $\eta_1, \eta_2 : \mathbb{R}^+ \to \mathbb{C}$, denotaremos por $\eta_1 *_M \eta_2$ la convolución multiplicativa en $\mathbb{R}^+$ ("convolución de Mellin"):

$$(\eta_1 *_M \eta_2)(x) = \int_0^{\infty} \eta_1(x/t)\eta_2(t)\frac{dt}{t}.$$

Para $\upsilon_1, \upsilon_2 : \mathbb{Z}^+ \to \mathbb{C}$, denotaremos simplemente por $\upsilon_1 * \upsilon_2$ la convolución multiplicativa en $\mathbb{Z}^+$ ("convolución de Dirichlet"):

$$(\upsilon_1 * \upsilon_2)(n) = \sum_{d|n} \upsilon_1(n/d)\upsilon_2(d) = \sum_{d|n} \upsilon_1(d)\upsilon_2(n/d).$$

Lema 3.7. *Sean $X \geq 1$, $0 < \alpha, \beta, \delta < 1$, $Q_0 \leq \exp((\log X)^{1-\beta})$ y $P_0 = Q_0^{\alpha}$. Definamos $u_X : \mathbb{Z}^+ \to \mathbb{R}$ por*

$$u_X(n) = \frac{1}{\log(1+\delta)} \int_{P_0}^{Q_0} \left(\upsilon_{1,Q,\delta} * \upsilon_{2,Q(1+\delta),X/Q}\right)(n)\, \frac{dQ}{Q}, \tag{3.3}$$

donde

$$\upsilon_{1,Q,\delta}(n) = \begin{cases} 1 & \textit{si } n \textit{ es primo y } Q < n \leq (1+\delta)Q, \\ 0 & \textit{de otra manera,} \end{cases} \tag{3.4}$$

$$\upsilon_{2,r,Y}(m) = \begin{cases} 1 & \textit{si } m \in (Y, 2Y] \textit{ y } p'|m \Rightarrow p' \notin [P_0, r), \\ 0 & \textit{de otra manera.} \end{cases} \tag{3.5}$$

Entonces existe un conjunto $Err \subset (X, 2(1+\delta)X]$ con

$$\frac{|Err|}{X} \ll \alpha + \delta + \exp(-(\log P_0)^{\min(3/5,\beta)+o(1)})$$

tal que $u_X(n) = 1_{(X,2X]}(n)$ para $n \notin Err$. Más aún, $0 \leq u_X(n) \leq 1$ para todo $n \in \mathbb{Z}^+$.

Demostración. Por el Lema 2.7, existen a lo más

$$\alpha X + X \cdot O\left(\max\left(\exp(-(\log P_0)^{3/5+o(1)}), \exp\left(-\frac{\log X}{3\log Q_0}\right)\right)\right) + \frac{X}{P_0}$$
$$\ll (\alpha + O(\exp(-(\log P_0)^{\min(3/5,\beta)+o(1)})))X$$

elementos de $(X, 2X]$ tales que no hay ningún primo $p \in (P_0, Q_0]$ tal que $p|n$. Incluímos todos estos elementos en el conjunto Err.

Es fácil ver que la diferencia

$$\eta_\Delta(x) = \left(1_{(X,2X]} - \frac{1_{(1,1+\delta]} *_M 1_{(X,2X]}}{\log(1+\delta)}\right)(x)$$

se desvanece cuando $x \leq X$, $(1+\delta)X < x \leq 2X$ o $x > 2(1+\delta)X$. Incluímos en Err, entonces, todos los elementos de $(X, (1+\delta)X]$ y $(2X, 2(1+\delta)X]$.

Por definición y un cambio de variables $t = mQ$, tenemos

$$\begin{aligned}\left(1_{(1,1+\delta]} *_M 1_{(X,2X]}\right)(pm) &= \int_0^\infty 1_{(1,1+\delta]}\left(\frac{p}{Q}\right) 1_{(X,2X]}(mQ)\frac{dQ}{Q} \\ &= \int_{(1-\delta)P_0}^{Q_0} 1_{(Q,(1+\delta)Q]}(p) 1_{\left(\frac{X}{Q},\frac{2X}{Q}\right]}(m)\frac{dQ}{Q}\end{aligned}$$

para $P_0 < p \leq Q_0$ y m arbitrario. Así, para todo n que no hayamos ya incluído en Err,

$$1_{(X,2X]}(n) = \frac{1}{\log(1+\delta)} \int_{(1-\delta)P_0}^{Q_0} \sum_{p|n} 1_{(Q,(1+\delta)Q]}(p) \upsilon_{2,p,X/Q}(n/p)\frac{dQ}{Q}, \tag{3.6}$$

donde $\upsilon_{2,r,Y}$ es como en (3.5). (Está claro que la suma $\sum_{p|n}$ puede tener a lo más un término no nulo, ya que $\upsilon_{2,p,X/Q}(n/p)$ se anulará a menos que p sea el factor primo $\geq P_0$ más pequeño de n.)

Como el conjunto S_2 de enteros $X < n \leq 2(1+\delta)X$ con por lo menos un factor primo en el rango $((1-\delta)P_0, (1+\delta)P_0]$ tiene a lo más

$$\begin{aligned}\sum_{p \in ((1-\delta)P_0,(1+\delta)P_0]} \left(\frac{(1+2\delta)X}{p} + O(1)\right) &= (1+2\delta)X \log\frac{\log(1+\delta)P_0}{\log(1-\delta)P_0} + O\left((1+\delta)P_0\right) \\ &\ll \frac{\delta X}{\log P_0} + P_0\end{aligned}$$

elementos, podemos permitirnos cambiar el rango de la integral en (3.6) de $[(1-\delta)P_0, Q_0]$ a $[P_0, Q_0]$, añadiendo $O(\delta X/\log P_0 + P_0)$ elementos a Err.

Finalmente, debemos examinar lo que sucede cuando cambiamos $\upsilon_{2,p,X/Q}$ en (3.6) (con la integral de P_0 a Q_0) por $\upsilon_{2,Q(1+\delta),X/Q}$. Los únicos enteros afectados están en el conjunto S_3 de enteros $X < n \leq 2(1+\delta)X$ divisibles por algún producto pp' de dos primos p, p' con $P_0 \leq p \leq Q_0$ y $p \leq p' < (1+\delta)p$, y tales que ningún primo $P_0 \leq p'' \leq p$ divide n. Ahora bien,

$$\begin{aligned}|S_3| &\leq \sum_{P_0 \leq p \leq Q_0} \sum_{p_1 < p_1' < (1+\delta)p_1} |\{X < n \leq 2(1+\delta)X : p_1, p_1'|n\}| \\ &\ll \sum_{P_0 \leq p \leq Q_0} \sum_{p < p' < (1+\delta)p} \left(\frac{X}{pp'} + O(1)\right) \\ &\ll X \sum_{P_0 \leq p \leq Q_0} \frac{1}{p}\left(\log\frac{\log(1+\delta)p}{\log p} + O(\exp(-(\log p)^{3/5+o(1)}))\right) + O(Q_0^2) \\ &\ll X \sum_{P_0 \leq p \leq Q_0} \frac{\delta}{p\log p} + X \sum_{P_0 \leq p \leq Q_0} \frac{1}{pe^{(\log p)^{3/5+o(1)}}} + Q_0^2 \\ &\ll \frac{\delta X}{\log P_0} + X\exp(-(\log P_0)^{3/5+o(1)}),\end{aligned}$$

donde acotamos las dos sumas en la penúltima línea por el teorema de los números primos y sumación por partes. Incluímos S_3 en Err. Tenemos ahora

$$1_{(X,2X]}(n) = u_X(n)$$

para todo $n \notin Err$, donde

$$u_X(n) = \frac{1}{\log(1+\delta)} \int_{(1-\delta)P_0}^{Q_0} \sum_{p|n} 1_{(Q,(1+\delta)Q]}(p)\upsilon_{2,Q(1+\delta),X/Q}(n/p)\frac{dQ}{Q}.$$

Por último, como $\upsilon_{2,Q(1+\delta),X/Q}(n/p)$ puede ser no nulo sólo cuando p es el divisor primo $\geq P_0$ más pequeño de n, y $1_{(Q,(1+\delta)Q]}(p) \neq 0$ sólo cuando Q está entre $p/(1+\delta)$ y p, vemos que, para cualquier n, $0 \leq u_X(n) \leq 1$. □

Corolario 3.8. *Sea $f : \mathbb{N} \to \mathbb{C}$ una función completamente multiplicativa tal que $|f(n)| \leq 1$ para todo $n \in \mathbb{N}$. Entonces, para $0 < \alpha, \beta, \delta < 1$, $Q_0 \leq \exp((\log X)^{1-\beta})$ y $P_0 = Q_0^{\alpha}$,*

$$Z_{f\cdot 1_{(X,2X]}} = Z_{err} + \frac{1}{\log(1+\delta)} \int_{P_0}^{Q_0} Z_{f_{1,Q,\delta}} Z_{f_{2,Q(1+\delta),X/Q}} \frac{dQ}{Q}$$

donde

$$f_{1,Q,\delta}(n) = \begin{cases} f(n) & \text{si } n \text{ es primo y } Q < n \leq (1+\delta)Q, \\ 0 & \text{de otra manera,} \end{cases}$$

$$f_{2,Q,Y}(m) = \begin{cases} f(m) & \text{si } m \in (Y, 2Y] \text{ y } p'|m \Rightarrow p' \notin [P_0, Q), \\ 0 & \text{de otra manera,} \end{cases}$$

y $err : \mathbb{N} \to \mathbb{C}$ es una función con soporte en $(X, 2(1+\delta)X]$ tal que

$$\frac{1}{X} \sum_{X<n\leq 2(1+\delta)X} |err(n)|^2 \ll \alpha + \delta + \exp(-(\log P_0)^{\min(3/5,\beta)+o(1)}). \tag{3.7}$$

Demostración. Se sigue inmediatamente del Lema 3.7. □

Ahora que tenemos esencialmente una factorización de $Z_{\lambda 1_{[X,2X]}}$, la idea es mostrar que el máximo del factor

$$Z_{\lambda_{1,Q,\delta}}(1+it) = \sum_{Q<p\leq Q+\delta Q} \lambda(p)p^{-1-it} = -\sum_{Q<p<Q+\delta Q} p^{-1-it}.$$

es pequeño. En realidad esto no es siempre cierto, ya que para t pequeña el signo de $\Re p^{-it}$ va a ser positivo, por lo cual no va a haber cancelación. Afortunadamente, para t pequeña, ya sabemos que $Z_{\lambda 1_{[X,2X]}}(1+it)$ es pequeño por el Corolario 2.4. Cuando t no es pequeña, la cancelación en la suma $\sum_{Q<p<Q+\delta Q} p^{-1-it}$ nos permite acotarla de la manera siguiente.

Proposición 3.9. *Para $\exp((\log x)^a) \leq t \leq \exp((\log x)^{(3/2)(1-a)})$, $a > 0$, $0 < \delta \leq 1$,*

$$\sum_{x<p\leq(1+\delta)x} p^{-1-it} \ll \exp(-(\log x)^{a+o(1)}).$$

Es, por cierto, necesario poner no sólo una cota inferior sobre t como condición, sino también alguna cota superior, tal y como lo hemos hecho aquí. [3]

[3] Para ver (de manera informal) la necesidad de una cota superior para obtener un resultado no trivial, notamos que tendremos cancelación en la suma $\sum_{x<p\leq(1+\delta)x} p^{-1-it}$ sólo si el número complejo

$$p^{-it} = \exp(it\log p)$$

cambia de argumento lo suficiente. No es plausible que esto pase para todo t grande, por el siguiente argumento heurístico. Deberíamos tener que $\{t\log p/2\pi\} < 1/8$ con probabilidad $1/8$, para t tomado al azar en un intervalo grande. Si tales eventos probabilísticos (uno por primo) son aproximadamente independientes – lo cual es generalmente visto como plausible – entonces, con probabilidad $\gtrsim (1/8)^{\delta Q/\log Q}$, *Prueba de la Proposición 3.9.* Por sumación por partes es suficiente demostrar la desigualdad

$$\sum_{n\leq x} \Lambda(n) n^{-it} \ll x \exp(-(1/2+o(1))(\log x)^a).$$

Como la función zeta correspondiente es $-\zeta'(s)/\zeta(s)$, esta desigualdad puede demostrarse como en la prueba del Teorema 2.5, pero evitando el polo mediante la elección $T = |t|/2$ (digamos), $\delta = 1/\sqrt{T}$. Usando las cotas en el Teorema 2.2, así como la cota $|M\psi_\delta(s)| \ll 1/s$, obtenemos

$$\begin{aligned}\sum_{n\leq x} \Lambda(n) n^{-it} &\ll \frac{x(\log T)^{2/3+o(1)}}{\sqrt{T}} + x\exp\left(-(\log x)(\log T)^{-2/3+o(1)}\right)(\log T)^{2/3+o(1)} \\ &\ll \frac{x}{e^{(1+o(1))(\log x)^a/2}} + \frac{x}{e^{(\log x)^{a+o(1)}}} = \frac{x}{e^{(\log x)^{a+o(1)}}}.\end{aligned}$$

□

Proposición 3.10. *Sean $Q_0 \leq \exp((\log X)^{1-\beta})$, $P_0 = Q_0^\alpha$, $0 < \alpha, \beta < 1$. Sean $h \geq \delta^{-2} Q_0$ y $0 < \delta < 1$. Entonces*

$$\mathbb{E}_{X<x\leq 2X} |\mathbb{E}_{(1-\frac{h}{X})x<n\leq x} \lambda(n)|^2 \ll \alpha + \frac{\eta^2}{\delta^2\alpha^2} + \delta + \exp(-(\log P_0)^{\text{mín}(3/5,\beta)+o(1)}),$$

donde

$$\eta = \max_{\substack{t\in[\exp(\sqrt{\log X}), X/h\delta^2] \\ P_0\leq Q\leq Q_0}} |Z_{\lambda_{1,Q}}(1+it)|. \tag{3.8}$$

Demostración. Aplicamos la Proposición 3.4 con $f = \lambda 1_{[X,2X]}$. Luego aplicamos el Corolario 2.4 para concluir que $Z_{\lambda 1_{[X,2X]}}(1+it) \ll \exp(-(\log X)^{3/5})$ para $t \leq \exp(\sqrt{\log X})$ (digamos). Así, podemos quitar esa parte de la integral, y en el resto usamos el Corolario 3.8 para factorizar $Z_{\lambda 1_{[X,2X]}}$. Aplicando la Proposición 3.5 para controlar Z_{err}, llegamos a

$$\begin{aligned}\mathbb{E}_{X<x\leq 2X} |\mathbb{E}_{(1-\frac{h}{X})x<n<x} \lambda(n)|^2 \ll &\int_{1+i\exp(\sqrt{\log X})}^{1+i\frac{X}{h\delta^2}} \left| \frac{1}{\delta} \int_{P_0}^{Q_0} Z_{\lambda_{1,Q}} Z_{\lambda_{2,Q,X/Q}} \frac{dQ}{Q} \right|^2 ds \\ &+ \alpha + \delta + \exp(-(\log P_0)^{\text{mín}(3/5,\beta)+o(1)}).\end{aligned}$$

Por Cauchy-Schwarz y la Proposición 3.6,

$$\begin{aligned}&\int_{1+i\exp(\sqrt{\log X})}^{1+i\frac{X}{h\delta^2}} \left| \frac{1}{\delta} \int_{P_0}^{Q_0} Z_{\lambda_{1,Q}} Z_{\lambda_{2,Q,X/Q}} \frac{dQ}{Q} \right|^2 ds \\ &\leq \frac{1}{\delta^2} \left(\log \frac{Q_0}{P_0}\right) \int_{P_0}^{Q_0} \int_{1+i\exp(\sqrt{\log X})}^{1+i\frac{X}{h\delta^2}} \left| Z_{\lambda_{1,Q}} Z_{\lambda_{2,Q,X/Q}} \right|^2 ds \frac{dQ}{Q} \\ &\leq \frac{\eta^2}{\delta^2} \left(\log \frac{Q_0}{P_0}\right) \int_{P_0}^{Q_0} \int_{1}^{1+i\frac{X}{h\delta^2}} \left| Z_{\lambda_{2,Q,X/Q}} \right|^2 ds \frac{dQ}{Q} \\ &\leq \frac{\eta^2}{\delta^2} \left(\log \frac{Q_0}{P_0}\right) \int_{P_0}^{Q_0} \mathbb{E}_{\frac{X}{Q}<n\leq 2\frac{X}{Q}} |\lambda_{2,Q,X/Q}|^2 \frac{dQ}{Q},\end{aligned} \tag{3.9}$$

donde η es como en (3.8). (En la última línea, estamos usando la suposición $h\delta^2 \geq Q$.) Usando la cota trivial $\mathbb{E}_{\frac{X}{Q}<n\leq 2\frac{X}{Q}} |\lambda_{2,Q,X/Q}|^2 \leq 1$, obtenemos que la expresión en la última

---

tendremos que $\{t \log p/2\pi\} < 1/8$ para todo $Q < p \leq (1+\delta)Q$. Por tanto, debería haber algún $t \leq 8^Q$ (digamos) para el cual este es el caso, lo cual implicaría que $\Re p^{-it} \geq 1/\sqrt{2}$ para todo $x < p \leq (1+\delta)x$, y por lo tanto no habría suficiente cancelación: $\Re \sum_{x<p\leq(1+\delta)x} p^{-1-it}$ sería $\geq (1/\sqrt{2})|\{x < p \leq (1+\delta)x\}|$.

línea de (3.9) es

$$\ll \frac{\eta^2}{\delta^2}\left(\log\frac{Q_0}{P_0}\right)^2 = \frac{\eta^2}{\delta^2\alpha^2}.$$

□

Usando este resultado y la Proposición 3.10 con $Q_0 = \delta^3 h$ y $\delta = (\log h)^{-1/2}$, podemos completar nuestro objetivo en cierto rango.

Teorema 3.11. *Para* $\exp((\log X)^{2/3+\epsilon}) \leq h \leq \exp((\log X)^{1-\epsilon})$, $0 < \epsilon \leq 1/6$, *tenemos*

$$\mathbb{E}_{X<x\leq 2X}|\mathbb{E}_{(1-\frac{h}{X})x<n<x}\lambda(n)|^2 \ll_\epsilon \frac{(\log X)^{\frac{2}{3}+\frac{\epsilon}{2}}}{\log h} \leq \frac{1}{(\log X)^{\epsilon/2}}. \tag{3.10}$$

Demostración. Aplicaremos la Proposición 3.9 para acotar la cantidad η definida en (3.8). Escogemos $a = 1 - 1/(1+3\epsilon/4)$, $\delta \geq h^{-1/3}$, $Q_0 = \delta^2 h$, $P_0 = \exp((\log X)^{2/3+\epsilon/2})$ y $\alpha = (\log P_0)/\log Q_0 = (\log X)^{2/3+\epsilon/2}/\log Q_0$. De esta manera,

$$(\log P_0)^{\frac{3}{2}(1-a)} = (\log X)^{\left(\frac{2}{3}+\frac{\epsilon}{2}\right)\cdot\frac{3}{2}(1-a)} = \log X.$$

Así, $t \leq X$ implica $t \leq \exp((\log P_0)^{(3/2)(1-a)})$. Como $\epsilon \leq 1/6$, vemos que $a \leq 1/9 < 1/2$, y por lo tanto $t \geq \exp(\sqrt{\log X})$ implica $t \geq \exp((\log Q_0)^a)$. Concluímos que sí podemos aplicar la Proposición 3.9, la cual nos da que

$$\eta \ll \exp(-(\log P_0)^{a+o(1)}) = \exp(-(\log X)^{(2/3+\epsilon/2)(a+o(1))}).$$

Resulta sensato escoger $\delta = \exp(-(\log X)^{2a/3})$. (La condición $\delta \geq h^{-1/3}$ se cumple para X más grande que una constante, ya que $h \geq \exp((\log X)^{2/3+\epsilon})$ y $a \leq 1/9$.)

Ahora aplicamos la Proposición 3.10 con $\beta = \epsilon$, y obtenemos que

$$\begin{aligned}\mathbb{E}_{X<x\leq 2X}|\mathbb{E}_{(1-\frac{h}{X})x<n<x}\lambda(n)|^2 \ll\ & \frac{(\log X)^{\frac{2}{3}+\frac{\epsilon}{2}}}{\log h} + \frac{\exp(-(\log X)^{(2/3+\epsilon/2)(a+o(1))})}{\exp(-(\log X)^{2a/3})(\log X)^{-2/3}} \\ & + \exp(-(\log X)^{2a/3}) + \exp\left(-(\log X)^{2\epsilon/3+o(1)}\right) \ll_\epsilon \frac{(\log X)^{\frac{2}{3}+\frac{\epsilon}{2}}}{\log h}.\end{aligned}$$

□

Es fácil deducir el resultado para $h > \exp((\log X)^{1-\epsilon})$ del Teorema 3.11 (ejercicio 3).

3.1.1. Ejercicios.

1. Queremos demostrar que el Teorema 3.1 implica el Teorema 1.1.
   a) Sea $|f| \leq 1$. Demuestre que para $h, Y > 1$, $0 < \delta < 1$ cualesquiera se cumple

   $$\int_Y^{(1+\delta)Y}\left|\sum_{x<n\leq x+h} f(n)\right| dx = O(\delta)\delta Y h + \int_{Y+h}^{Y+h+\delta Y}\left|\sum_{(1-\frac{h}{Y+h})x<n\leq x} f(n)\right| dx.$$

   b) Aplique el apartado anterior para demostrar que

   $$\int_X^{2X}\left|\sum_{x<n\leq x+h} f(n)\right| dx \ll \delta X h + \frac{1}{\delta}\max_{X\leq X'\leq 3X}\int_{X'}^{2X'}\left|\sum_{(1-\frac{h}{X'})x<n\leq x} f(n)\right| dx.$$

   c) Tomando el δ adecuado, concluya que el Teorema 3.1 implica el Teorema 1.1.

2. Para cierto $B > 1$ y todo $h \geq X^{1/6}(\log X)^B$, vamos a ver una demostración del resultado
$$\mathbb{E}_{X<x\leq 2X}|\mathbb{E}_{(1-\frac{h}{X})x<n\leq x}\Lambda(n) - 1|^2 = o(1),$$
usando diferentes resultados conocidos sobre la función $\zeta(s)$. Se trata de un resultado clásico, conocido ya mucho antes de Matomäki y Radziwiłł.

   Nótese que este resultado implica que hay primos en casi todos los intervalos de longitud h. Usando técnicas similares (pero un poco más complicadas) es posible demostrar lo mismo para $\lambda(n)$ en vez de $\Lambda(n)-1$. Por otra parte: ¿por qué es que la demostración que hemos hecho en esta sección para $\lambda(n)$ no vale para $\Lambda(n)-1$? (Hay, por otra parte, resultados sobre Λ que pueden ser probados de manera indirecta usando los trabajos de Matomäki y Radziwiłł sobre $\lambda(n)$: ver [MRT19, Thm. 1.3].)

   a) Como en la prueba de la Proposición 3.3, demuestre que, si $x \in (X, 2X]$,
$$\mathbb{E}_{(1-\frac{h}{X})x<n\leq x}\Lambda(n) = \frac{1}{2\pi i}\int_{1+\frac{1}{\log X}-i\frac{X}{h\delta^2}}^{1+\frac{1}{\log X}+i\frac{X}{h\delta^2}} x^{s-1}M\psi_{\delta,h}(s)\frac{-\zeta'(s)}{\zeta(s)}\,ds + O(\delta\log X)$$
   para cierta función analítica $M\psi_{\delta,h}(s) \ll \text{mín}(1, \frac{X/h}{|s|}, \frac{X/h}{\delta|s|^2})$ para $\Re s > -1$, con $M\psi_{\delta,h}(1) = 1 + O(\delta)$.

   b) En la zona $-1 < \Re s < 2$, los únicos ceros de la función $\zeta(s)$ están en $0 \leq \Re s \leq 1$, y si $N(T)$ es el número de ceros con $|\Im s| \leq T$, entonces $N(T) \sim \frac{T}{\pi}\log T$ y $N(T+1) - N(T) \ll \log T$. Además, si $|s-1| \gg 1$ y s está a distancia ϵ del cero de $\zeta(s)$ más cercano, entonces $\frac{\zeta'(s)}{\zeta(s)} \ll \frac{\log(2+|s|)}{\text{mín}(1,\epsilon)}$. Usando esa información, el apartado anterior y el teorema de los residuos demuestre que, para algún $|w| \leq 1$,
$$\mathbb{E}_{(1-\frac{h}{X})x<n\leq x}\Lambda(n) = 1 - \sum_\rho x^{\rho-1}M\psi_{\delta,h}(\rho) + O(\delta(\log X)^2),$$
   donde ρ son los ceros de $\zeta(s)$ que satisfacen $|\Im\rho| < \frac{X}{h\delta^2} + w$.

   c) Usando la fórmula del apartado anterior y procediendo como en la demostración de la Proposición 3.4, demuestre que
$$\mathbb{E}_{X<x\leq 2X}|\mathbb{E}_{(1-\frac{h}{X})x<n\leq x}\Lambda(n) - 1|^2 \ll \log X\sum_\rho X^{-2(1-\Re\rho)} + O(\delta(\log X)^2).$$

   d) Se sabe que el número $N(\sigma, T)$ de ceros de $\zeta(s)$ con $\sigma \leq \Re\rho \leq 1$ y $|\Im\rho| \leq T$ satisface $N(\sigma,T) \ll T^{\frac{12}{5}(1-\sigma)}(\log T)^A$, para cierto $A > 1$ (*teorema de densidad*). Use ese hecho para acotar $\sum_{\sigma\leq\Re\rho\leq\sigma+\frac{1}{\log X}} X^{-2(1-\Re\rho)}$ por cierta función $S(\sigma)$ creciente si $h\delta^2 > X^{1/6}$.

   e) Usando los dos apartados anteriores y la región libre de ceros del Teorema 2.2, demuestre el resultado que se menciona al principio del problema. Observe también que si se cumpliera la hipótesis de Riemann, podríamos demostrarlo para $h > (\log X)^6$.

3. Usando el Teorema 3.11, vamos a mostrar que, para $\exp((\log X)^{1-\epsilon}) \leq h \leq X/2$, $0 < \epsilon \leq 1/6$,
$$\mathbb{E}_{X<x\leq 2X}|\mathbb{E}_{(1-\frac{h}{X})x<n\leq x}\lambda(n)|^2 \ll_\epsilon \frac{1}{(\log X)^{\frac{1}{3}-\epsilon}}.$$

   a) Sea $j \in \mathbb{N}$. Si definimos h_j por $1 - \frac{h_j}{X} = (1 - \frac{h}{X})^{1/j}$, dividiendo la suma interior en j trozos, demuestre que para $|f| \leq 1$ se cumple
$$\mathbb{E}_{X<x\leq 2X}|\mathbb{E}_{(1-\frac{h}{X})x<n\leq x}f(n)|^2 \leq j^2 \max_{X'\in[X-h,X]}\mathbb{E}_{X'<x\leq 2X'}|\mathbb{E}_{(1-\frac{h_j}{X})x<n\leq x}f(n)|^2.$$

b) Usando el apartado anterior y el Teorema 3.11, demuestre la desigualdad del enunciado del problema.

3.2. El caso general: valores excepcionales. Cuando h es más pequeño que un cierto nivel – digamos, $\exp((\log X)^{2/3})$ – nos encontramos con un obstáculo. No podemos hacer que Q sea más pequeño que $\exp((\log x)^{2/3+\epsilon})$, pues estaríamos fuera del rango en el cual podemos aplicar la Proposición 3.9. Por otra parte, tampoco parecería que podemos tomar h más pequeño que Q: el rango de integración en

$$\int_0^{X/h} |Z_{\lambda_{2,Q,X/Q}}(1+it)|^2 dt$$

es demasiado grande como para que podamos aplicar la Proposición 3.6 (es decir, la estimación de valor medio que viene del Lema 2.10).

Existe la alternativa siguiente. Nuestra ganancia viene del factor $|Z_{\lambda_{1,Q}}|^2 \leq M_{1,Q}^2$ en (3.9). Podemos estudiar por separado los valores de t para los cuales $|Z_{\lambda_{1,Q}}|$ es excepcionalmente grande – digamos $|Z_{\lambda_{1,Q}}| > Q^{-1/9}$. Si demostramos que tal cosa sucede sólo para muy pocos valores de $s = 1 + it$, podremos usar el teorema del valor medio de Halász-Montgomery (Proposición 2.12) en vez del Lema 2.10 para acotar

$$M_{1,Q}^2 \int_{\substack{0<t<X/h \\ |Z_{\lambda_{1,Q}}|>Q^{-1/9}}} |Z_{\lambda_{2,Q,X/Q}}(1+it)|^2 dt.$$

Mostremos, entonces, que el conjunto de valores de t para los cuales $|Z_{\lambda_{1,Q}}(1+it)| > Q^{-1/9}$ es suficientemente pequeño.

Lema 3.12. *Sea* $|f| \leq 1$ *con soporte en los primos de* $(Q, 2Q]$ *y* $(\log T)^9 < Q < T^{1/3}$*. En ese caso, el conjunto de* t *en* $[0, T]$ *tales que* $|Z_f(1+it)| > Q^{-1/9}$ *tiene medida* $O(T^{4/9})$*.*

El exponente 4/9 no es óptimo, pero nos será suficiente.

Demostración. El teorema del valor medio (Lema 2.10) aplicado a $|Z_f(1+it)|^2$ no controla bien los valores grandes de $|Z_f(1+it)|$. Para conseguir tal control, es mejor usarlo sobre potencias superiores de Z_f. Lo aplicaremos a $Z_f(1+it)^\ell$ con $\ell \in \mathbb{Z}^+$ tal que $Q^{\ell-1} < T \leq Q^\ell$. Así, como

$$Z_f(1+it)^\ell = \sum_n a_n n^{-1-it}$$

donde a_n tiene soporte en $(Q^\ell, (2Q)^\ell]$ y $|a_n| \leq \ell!$, por el Lema 2.10 tenemos

$$\int_0^T |Z_f(1+it)^\ell|^2 = \int_0^T |\sum_{Q^\ell<n\leq(2Q)^\ell} a_n n^{-1-it}|^2 dt \ll (2Q)^\ell \sum_n \frac{|a_n|^2}{n^2}$$
$$\ll \ell! 2^\ell Q^{-\ell} (\sum_{Q<p<2Q} 1)^\ell \ll \ell! 2^\ell \ll \ell^\ell.$$

Así, si A es el conjunto del enunciado, tenemos que $(Q^{-1/9})^{2\ell}|A| \ll \ell^\ell$. Como $Q > (\log T)^9$, sabemos que $\ell < Q^{1/9}$, y por lo tanto

$$|A| \ll (Q^\ell)^{3/9} \ll Q^{3/9} T^{3/9} \ll T^{4/9}.$$

□

En el siguiente resultado mostramos que podemos controlar la integral sobre los valores excepcionales de $Z_{\lambda 1_{(X,2X]}}$.

Teorema 3.13. *Sea $h \geq 1$ y $Q_0 = \exp((\log X)^\beta)$, donde $2/3 < \beta < 1$. Entonces, para $0 < \rho < 1 - \frac{2}{3\beta}$, $\alpha = 1/(\log Q_0)^\rho$ y $1/(\log X) \leq \delta \leq 1/(\log Q_0)^\rho$,*

$$\int_{[1,1+i\frac{X}{h}]\setminus \mathscr{T}_{\lambda,Q_0,\frac{1}{9},\alpha,\delta}} |Z_{\lambda\cdot 1_{(X,2X]}}(s)|^2 ds \ll_{\beta,\rho} \frac{1}{(\log Q_0)^\rho}$$

con $\mathscr{T}_{f,A,\gamma,\alpha,\delta}$ el conjunto de todos los valores de $s = 1 + it$, $0 \leq t \leq X/h$, tales que $|Z_{f_{1,Q,\delta}}(s)| \leq Q^{-\gamma}$ para todo $Q \in [A^\alpha, A] \cap \mathbb{Z}$, donde $f_{1,Q,\delta}(n) = f(n) \cdot 1_{(Q,(1+\delta)Q]}(n)$ si n es primo y 0 en otro caso.

Demostración. Comenzamos como en la demostración de la Proposición 3.10; es decir, aplicamos el Corolario 2.4 para concluir que $Z_{\lambda 1_{(X,2X]}}(1 + it) \ll \exp(-(\log X)^{3/5+o(1)})$ para $t \leq \exp(\sqrt{\log X})$ (digamos). Así, podemos quitar esa parte de la integral, y en el resto usamos el Corolario 3.8 para factorizar $Z_{\lambda 1_{(X,2X]}}$. Aplicando (3.7) y la Proposición 3.5 para controlar Z_{err}, obtenemos

$$\begin{aligned}&\int_{[1,1+iX/h]\setminus \mathscr{T}_{\lambda,Q_0,1/9,(\log Q_0)^{-\rho},(\log Q_0)^{-\rho'}}} |Z_{\lambda 1_{(X,2X]}}(s)|^2 ds \\ &\ll O_\rho\left(\frac{1}{(\log Q_0)^\rho}\right) + \frac{1}{\delta^2}\int_B \left|\int_{Q_0^\alpha}^{Q_0} Z_{f_{1,Q,\delta}}(s) Z_{f_{2,Q(1+\delta),X/Q}}(s)\frac{dQ}{Q}\right|^2 ds\end{aligned}$$

con $B = [1 + i\exp(\sqrt{\log X}), 1 + iX/h] \setminus \mathscr{T}_{\lambda,Q_0,1/9,\alpha,\delta}$. Como $\rho < 1$, el término de error $\exp\left(-(\log P_0)^{3/5+o(1)}\right)$ en (3.7) es absorbido por $O_\rho\left(1/(\log Q_0)^\rho\right)$. Por Cauchy-Schwarz, de manera similar a (3.9),

(3.11)

$$\int_B \left|\int_{Q_0^\alpha}^{Q_0} Z_{f_{1,Q,\delta}}(s) Z_{f_{2,Q(1+\delta),X/Q}}(s)\frac{dQ}{Q}\right|^2 ds \leq (\log Q_0^{1-\alpha})^2 \max_Q \int_B \left|Z_{f_{1,Q,\delta}}(s) Z_{f_{2,Q(1+\delta),X/Q}}(s)\right|^2 ds,$$

donde el máximo de Q se toma en el intervalo $[Q_0^\alpha, Q_0]$. (Después de aplicar Cauchy-Schwarz, invertimos el orden de las integrales, y luego reemplazamos la integral ahora exterior por un máximo.)

Ahora sacamos el máximo de $Z_{\lambda_{1,Q}}$ en B, aplicando la Proposición 3.9 con $a = 1 - 1/((3/2)\beta(1-\rho))$ (Nótese que $\beta(1-\rho)\cdot(3/2)\cdot(1-a) = 1$, por lo cual $\exp((\log Q)^{(3/2)(1-a)}) \geq X$). Obtenemos que

$$\int_B |Z_{\lambda_{1,Q,\delta}} Z_{\lambda_{2,(1+\delta)Q,X/Q}}|^2 ds \ll \exp(-2(\log X)^a) \int_B |Z_{\lambda_{2,Q(1+\delta),X/Q}}|^2 ds.$$

Finalmente, tenemos que $B = \cup_{Q'} B_{Q'}$ con $Q' \in [Q_0^\alpha, Q_0] \cap \mathbb{Z}$ y $B_{Q'}$ contenido en el conjunto de $t \in [0, X]$ tales que $|Z_{\lambda_{1,Q',\delta}}(1 + it)| > Q'^{-1/9}$. Por el Lema 3.12 tenemos que $|B_{Q'}| \ll X^{4/9}$, lo cual implica que $|B| \ll Q_0 X^{4/9} = X^{4/9+o(1)}$. Luego, por Halász-Montgomery (Proposición 2.12), concluímos que

$$\int_B |Z_{\lambda_{2,Q(1+\delta),X/Q}}|^2 ds \ll \left(\frac{X}{Q} + X^{4/9+o(1)}\sqrt{X}\log X\right) \cdot \sum_{n \leq X/Q} (Q/X)^2 \ll 1,$$

lo que demuestra el teorema.

□

3.2.1. Ejercicios.

1. Sea $|f| \leq 1$, $X \geq 1$ y $\epsilon > 0$.
   a) Use el Teorema del valor medio para demostrar que la medida del conjunto de $t \in [0, X]$ para los que $|\sum_{X \leq p \leq 2X} f(p)p^{-it}| > X^{1-\epsilon}$ es $O(X^{2\epsilon})$.

b) Use el Teorema del valor medio, junto con una elevación a una potencia (como en la prueba del Lema 3.12) para demostrar que, para todo $k \geq 1$, la medida del conjunto de $t \in [0, X^k]$ para los que $|\sum_{X \leq p \leq 2X} f(p)p^{-it}| > X^{1-\epsilon}$ es $O_k(X^{2k\epsilon})$.

3.3. El caso general: valores típicos. Es imposible controlar una integral

$$\int_{\mathscr{T}_{\lambda,Q_0,\gamma,\alpha,\delta}} |Z_{\lambda 1_{[X,2X]}}(s)|^2 ds$$

factorizándola con un factor del tamaño Q_0, ya que Q_0 es mucho mayor que h. (Recuérdese que $\mathscr{T}_{\lambda,Q_0,\gamma,\alpha,\delta}$ es el conjunto de los valores "típicos" dentro del intervalo $[0, X/h]$.) La idea crucial ahora es usar que es posible sustituir dicha integral por otra del tipo

$$\int_{\mathscr{T}_{\lambda,q_0,\gamma-\epsilon,\alpha,\delta}} |Z_{\lambda 1_{[X,2X]}}(s)|^2 ds,$$

con q_0 sustancialmente más pequeño que Q_0, si pagamos aumentado el tamaño que permitimos para el factor zeta correspondiente. Dicho resultado está basado en el siguiente lema, que de nuevo usa el teorema del valor medio.

Lema 3.14. *Sea* $0 < \varepsilon < \gamma < \frac{1}{2}$, $(\log Q)^{4/\varepsilon} < q < Q^{\varepsilon/4} < X^{\frac{1}{(\log\log X)^3}}$. *Sean* $f_q, f_Q, g_{X/Q} : \mathbb{Z} \to \mathbb{C}$ *funciones con* $|f_q(n)|, |f_Q(n)|, |g_{X/Q}(n)| \leq 1$ *para todo* n. *Asumamos que* $g_{X/Q}$, f_Q *y* f_q *tienen soporte en* $[X/Q, 2X/Q]$, *en* $[Q, 2Q]$ *y en los primos de* $[q, 2q]$, *respectivamente. Entonces*

$$\int\limits_{\substack{s\in[1,1+iX] \\ |Z_{f_Q}(s)| \ll Q^{-\gamma} \\ |Z_{f_q}(s)| \geq q^{-\gamma+\varepsilon}}} |Z_{f_Q}(s) Z_{g_{X/Q}}(s)|^2\, ds \ll Q^{-\varepsilon}.$$

Demostración. Podemos acotar la integral por

$$\ll Q^{-2\gamma} \int\limits_{\substack{s\in[1,1+iX] \\ |Z_{f_q}(s)| \geq q^{-\gamma+\varepsilon}}} |Z_{g_{X/Q}}(s)|^2\, ds \ll Q^{-2\gamma} \int_1^{1+iX} \left| \left(\frac{Z_{f_q}(s)}{q^{-\gamma+\varepsilon}} \right)^{\ell} \right|^2 |Z_{g_{X/Q}}(s)|^2\, ds.$$

Tomando ℓ natural tal que $q^\ell < Q \leq q^{\ell+1}$, como $(q^{\gamma-\varepsilon})^{2\ell} Q^{-2\gamma} \leq Q^{2\gamma-2\varepsilon} Q^{-2\gamma} = Q^{-2\varepsilon}$, vemos que la integral del enunciado es

$$\ll Q^{-2\varepsilon} \int_1^{1+iX} |Z_{f_q}^{\ell}(s) Z_{g_{X/Q}}(s)|^2\, ds.$$

Ahora se trata de acotar la integral por el teorema del valor medio (Lema 2.10). Tenemos

$$\int_1^{1+iX} |Z_{f_q}^{\ell}(s) Z_{g_{X/Q}}(s)|^2\, ds \ll 2^\ell X \sum_{X/q \leq n \leq 2^{\ell+1} X} \frac{a_n^2}{n^2}$$

con

$$a_n = \sum_{\substack{p_1 p_2 \ldots p_\ell m = n \\ q \leq p_j \leq 2q \\ X/Q \leq m \leq 2X/Q}} 1 \leq \ell! \sum_{\substack{rm=n \\ p|r \Rightarrow q \leq p \leq 2q}} 1 = \ell!\, w(n)$$

con $w(n)$ la función totalmente multiplicativa con $w(p^k) = 1$ si $p \notin [q, 2q]$ y $w(p^k) = k+1$ si $p \in [q, 2q]$. Como $Z_{w^2}(s) = \zeta(s) \prod_{q \leq p \leq 2q} \frac{1+4p^{-s}+9p^{-2s}+\ldots}{1+p^{-s}+p^{-2s}+\ldots}$, luego con residuo en $s = 1$ de la forma $\prod_{q \leq p \leq 2q}(1 + O(1/p)) \ll 1$, podemos proceder como en la demostración del Lema 2.7 para demostrar que $\sum_{n<x} w^2(n) \ll x$ para $x \geq X/q$. Por lo tanto

$$\int_1^{1+iX} |Z_{f_q}^{\ell}(s) Z_{g_{X/Q}}(s)|^2\, ds \ll q^2 2^{2\ell} (\ell!)^2 \ll q^2 \ell^{2\ell}.$$

Como $q < Q^{\varepsilon/4}$ y $\ell^{2\ell} \ll (\log Q)^{2\ell} \ll (q^{\varepsilon/4})^{2\ell} \ll Q^{\varepsilon/2}$, hemos acabado. □

Ahora veamos que efectivamente podemos sustituir la integral de Z_f sobre $\mathscr{T}_{f,Q_0,c,\alpha,\delta}$ ("valores típicos con respecto a Q_0") por la integral de Z_f sobre $\mathscr{T}_{f,q_0,c,\alpha,\delta}$ ("valores típicos con respecto a q_0"), donde q_0 es mucho más pequeño que Q_0.

Proposición 3.15. *Sea f totalmente multiplicativa con $|f(n)| \leq 1$ para todo n. Sea $\mathscr{T}_{f,A,\gamma,\alpha,\delta}$ definido como en el Teorema 3.13. Sea $Q_0 \leq \exp((\log X)^{1-\epsilon})$, $\epsilon \in (0,1)$. Entonces, para $\gamma \in (1/\log\log Q_0, 1/2)$, $0 < \epsilon < 1$, $0 < \rho < 1$, $\rho < \kappa < 1 - \rho$, así como $\alpha \leq 1/(\log Q_0)^\rho$, $\alpha' = 1/(\log q_0)^{1-\epsilon}$, $1/(\log Q_0) \leq \delta \leq 1/(\log Q_0)^\rho$, $0 < \delta' \leq 1$ y $\log q_0 = (\log Q_0)^\kappa$, tenemos que*

$$\int_{\mathscr{T}_{f,Q_0,\gamma,\alpha,\delta}} |Z_{f_X}(s)|^2 ds \leq \int_{\mathscr{T}_{f,q_0,\gamma',\alpha',\delta'}} |Z_{f_X}(s)|^2 ds + O_{\rho,\kappa,\epsilon}\left(\frac{1}{(\log Q_0)^\rho}\right).$$

con $f_X = f \cdot 1_{(X,2X]}$ y $\gamma' = \gamma - 1/\log\log Q_0$.

Cualquier función del tipo $(\log Q_0)^{o(1)}$ podría usarse en vez de $\log\log Q_0$.

Demostración. Podemos suponer que $\alpha = 1/(\log Q_0)^\rho$, ya que $\alpha \leq \alpha_0 = 1/(\log Q_0)^\rho$ implica $\mathscr{T}_{f,Q_0,\gamma,\alpha,\delta} \subset \mathscr{T}_{f,Q_0,\gamma,\alpha_0,\delta}$.

Para comenzar,

$$\int_{\mathscr{T}_{f,Q_0,\gamma,\alpha,\delta}} |Z_{f_X}(s)|^2 ds \leq \int\limits_{\mathscr{T}_{f,q_0,\gamma',\alpha',\delta'}} |Z_{f_X}(s)|^2 ds + \int\limits_{\mathscr{T}_{f,Q_0,\gamma,\alpha,\delta}\setminus\mathscr{T}_{f,q_0,\gamma',\alpha',\delta'}} |Z_{f_X}(s)|^2 ds.$$

Apliquemos el Corolario 3.8 para factorizar Z_{f_X}, usando (3.7) y la Proposición 3.5 para controlar Z_{err}. Obtenemos, procediendo como en (3.9) o (3.11),

$$\int\limits_{\mathscr{T}_{f,Q_0,\gamma,\alpha,\delta}\setminus\mathscr{T}_{f,q_0,\gamma',\alpha',\delta'}} |Z_{f_X}(s)|^2 ds \ll O_\rho\left(\frac{1}{(\log Q_0)^\rho}\right)$$

$$+\frac{\left(\log Q_0^{1-\alpha}\right)^2}{\delta^2} \max_{Q\in[Q_0^\alpha,Q_0]} \int\limits_{\mathscr{T}_{f,Q_0,\gamma,\alpha,\delta}\setminus\mathscr{T}_{f,q_0,\gamma',\alpha',\delta'}} \left|Z_{f_{1,Q,\delta}}(s) Z_{f_{2,Q(1+\delta),X/Q}}(s)\right|^2 ds.$$

Ahora bien, para todo $s \in \mathscr{T}_{f,Q_0,\gamma,\alpha,\delta}$, tenemos, por definición, que $|Z_{f_{1,Q,\delta}}(s)| < Q^{-\gamma}$ para todo $Q \in [Q_0^\alpha, Q_0] \cap \mathbb{Z}$, luego $|Z_{f_{1,Q,\delta}}| \leq \frac{2}{Q} + Q^{-\gamma} \ll Q^{-\gamma}$ para todo $Q \in [Q_0^\alpha, Q_0]$; más aún, si $s \notin \mathscr{T}_{f,q_0,\gamma',\alpha,\delta}$, entonces $|Z_{f_{1,q,\delta}}(s)| \geq q^{-\gamma'}$ para algún $q \in [q_0^\alpha, q_0] \cap \mathbb{Z}$. Así, vemos que

$$\int\limits_{\mathscr{T}_{f,Q_0,\gamma,\alpha,\delta}\setminus\mathscr{T}_{f,q_0,\gamma',\alpha',\delta'}} \left|Z_{f_{1,Q,\delta}}(s) Z_{f_{2,Q(1+\delta),X/Q}}(s)\right|^2 ds$$

$$\leq 2q_0 \max_{q\in[q_0^\alpha,q_0]} \int\limits_{\substack{s\in[1,1+iX] \\ |Z_{f_{1,Q,\delta}}|\ll Q^{-\gamma} \\ |Z_{f_{1,q,\delta'}}|\geq q^{-\gamma'}}} |Z_{f_{1,Q,\delta}}(s) Z_{f_{2,Q(1+\delta),X/Q}}(s)|^2 ds,$$

donde recordamos que $\mathscr{T}_{f,Q_0,\gamma,\alpha,\delta} \subset [1, 1+iX/h] \subset [1, 1+iX]$. Aplicando el Lema 3.14, tenemos la cota

$$q_0 \int\limits_{\substack{s\in[1,1+iX] \\ |Z_{f_{1,Q,\delta}}|\ll Q^{-\gamma} \\ |Z_{f_{1,q,\delta'}}|\geq q^{-\gamma'}}} |Z_{f_{1,Q}}(s) Z_{f_{2,Q,X/Q}}(s)|^2 ds \ll q_0 Q^{-\varepsilon} \leq q_0 (Q_0^\alpha)^{-\varepsilon} < Q_0^{-3\alpha\varepsilon/4}$$

para $\gamma' = \gamma - \varepsilon$ y cualquier $\varepsilon \in (0,\gamma)$ tal que $(\log Q_0)^{4/\varepsilon} < q_0^{\alpha}$ y $q_0 < Q_0^{\varepsilon\alpha/4}$. Para $\varepsilon = 1/\log\log Q_0$, estas desigualdades se cumplen (debido a $\kappa + \rho < 1$), si asumimos que Q_0 es más grande que una constante c que depende sólo de ρ y κ. Bajo las mismas condiciones, $Q_0^{-3\alpha\varepsilon/4} < 1/(\log Q_0)^{\rho}$ (por un amplio margen). □

Teorema 3.16. *Sean* $0 < \epsilon < 1/3$, $Q_0 = \exp((\log X)^{1-\epsilon})$ *y* $1 \le h \le \exp((\log X)^{2/3-\epsilon})$. *Entonces, para* $\alpha = \delta = 1/(\log Q_0)^{\rho}$, $1/6 < \rho < 1/3$,

$$\int_{\mathscr{T}_{\lambda,Q_0,\frac{1}{9},\alpha,\delta}} |Z_{\lambda\cdot 1_{(X,2X]}}(s)|^2 ds \ll_\epsilon \text{máx}\left(\frac{1}{(\log Q_0)^{\rho}}, \frac{1}{(\log h)^{1-\epsilon}}\right).$$

Demostración. Supongamos primero que $h \ge \exp(\sqrt{\log Q_0})$. Aplicamos la Proposición 3.15 una vez con ρ, $\gamma = 1/9$, $\delta' = 1/\log h$, $\alpha' = 1/(\log h)^{1-\epsilon}$ y $\kappa = (\log\log h)/(\log\log Q_0)$, de tal manera que $q_0 = h$. Es fácil ver que $1/2 \le \kappa \le (2/3-\epsilon)/(1-\epsilon) < 2/3$, así que $\rho < \kappa < 1-\rho$. Obtenemos

$$\int_{\mathscr{T}_{\lambda,Q_0,\frac{1}{9},\alpha,\delta}} |Z_{\lambda 1_{[X,2X]}}(s)|^2 ds \le O_\epsilon\left(\frac{1}{(\log Q_0)^{\rho}}\right) + \int_{\mathscr{T}_{\lambda,h,\frac{1}{18},\alpha',\delta'}} |Z_{\lambda 1_{[X,2X]}}(s)|^2 ds,$$

puesto que podemos suponer que $\gamma - 1/\log\log Q_0 \ge 1/18$.

Ahora utilizamos el Corolario 3.8 para factorizar $Z_{\lambda 1_{[X,2X]}}$ con h en vez de Q_0, $\alpha = (\log h)^{-(1-\epsilon)}$. Por Cauchy-Schwarz, la Proposición 3.6 y la definición de $\mathscr{T}_{\lambda,h,\frac{1}{18},\alpha',\delta'}$, concluímos que

$$\int_{\mathscr{T}_{\lambda,h,\frac{1}{18},\alpha',\delta'}} |Z_{\lambda 1_{[X,2X]}}(s)|^2 ds \ll O_\epsilon\left(\frac{1}{(\log h)^{1-\epsilon}}\right) + (\log h)^4 (h^{\alpha})^{-\frac{2}{18}} = O_\epsilon\left(\frac{1}{(\log h)^{1-\epsilon}}\right).$$

Supongamos ahora que $h < \exp(\sqrt{\log Q_0})$. Procederemos exactamente como antes, excepto que iteraremos el uso de la Proposición 3.15. Definimos $Q_{i+1} = \exp(\sqrt{\log Q_i})$ para $i = 0, 1, \ldots$, hasta que llegamos a un i tal que $(\log h)^2 < \log Q_i \le (\log h)^4$. Entonces definimos $m = i+1$, $Q_m = h$.

Comenzamos aplicando la Proposición 3.15 con ρ, $\gamma = 1/9$, $\delta' = \delta_1 = 1/\log Q_1$, $\alpha' = \alpha_1 = 1/(\log Q_1)^{1-\epsilon}$, $\kappa = (\log\log Q_1)/(\log\log Q_0) = 1/2$ y los valores iniciales de α y δ. Obtenemos

$$\int_{\mathscr{T}_{\lambda,Q_0,\frac{1}{9},\alpha,\delta}} |Z_{\lambda 1_{[X,2X]}}(s)|^2 ds \le O_\epsilon\left(\frac{1}{(\log Q_0)^{\rho}}\right) + \int_{\mathscr{T}_{\lambda,Q_1,\gamma_1,\alpha_1,\delta_1}} |Z_{\lambda 1_{[X,2X]}}(s)|^2 ds,$$

donde $\gamma_1 = \gamma - 1/\log\log Q_0$. Luego aplicamos la Proposición 3.15 para $i = 1, 2, \ldots, m-1$, con $\rho = 1/2-\epsilon$, $\gamma = \gamma_i$, $\delta = \delta_i$, $\alpha = \alpha_i$, $\delta' = \delta_{i+1} = 1/\log Q_{i+1}$, $\alpha' = \alpha_{i+1} = 1/(\log Q_{i+1})^{1-\epsilon}$ y $\kappa = (\log\log Q_{i+1})/\log\log Q_i = 1/2$. Así,

$$\int_{\mathscr{T}_{\lambda,Q_i,\gamma_i,\alpha_i,\delta_i}} |Z_{\lambda 1_{[X,2X]}}(s)|^2 ds \le O_\epsilon\left(\frac{1}{(\log Q_i)^{1/2}}\right) + \int_{\mathscr{T}_{\lambda,Q_{i+1},\gamma_{i+1},\alpha_{i+1},\delta_{i+1}}} |Z_{\lambda 1_{[X,2X]}}(s)|^2 ds,$$

donde $\gamma_{i+1} = \gamma_i - 1/\log\log Q_i$. Está claro que $1/(\log Q_i)^{1/2} \le 1/(\log Q_{i+1})$. Verificamos que

$$\sum_{i<m} \frac{1}{\log\log Q_i} \le \frac{1}{\log\log Q_m} \sum_{k\ge 0} \frac{1}{2^k} = \frac{2}{\log\log h} < \frac{1}{18},$$

ya que podemos asumir que h es más grande que una constante. Por lo tanto, $\gamma_m \ge 1/18$.

Terminamos usando el Corolario 3.8 exactamente como antes (h en vez de Q_0, $\alpha = (\log h)^{-(1-\epsilon)}$). Por Cauchy-Schwarz y la Proposición 3.6, concluímos que

$$\begin{aligned}\int_{\mathscr{T}_{\lambda,Q_m,\gamma_m,\alpha_m,\delta_m}} |Z_{\lambda 1_{[X,2X]}}(s)|^2 ds &\ll O_\epsilon\left(\frac{1}{(\log h)^{1-\epsilon}}\right) + (\log h)^4 \left(\exp\left((\log h)^\epsilon\right)\right)^{-\frac{2}{18}} \\ &= O_\epsilon\left(\frac{1}{(\log h)^{1-\epsilon}}\right).\end{aligned}$$

□

Ahora ya podemos concluir nuestro resultado principal.

Teorema 3.17. *Sea $1 \leq h \leq X$. Entonces, para $\epsilon > 0$ arbitrariamente pequeño,*

$$\mathbb{E}_{X<x\leq 2X}|\mathbb{E}_{(1-\frac{h}{X})x<n\leq x}\lambda(n)|^2 \ll_\epsilon \text{máx}\left(\frac{1}{(\log X)^{\frac{1-\epsilon}{3}}}, \frac{1}{(\log h)^{1-\epsilon}}\right). \tag{3.12}$$

Demostración. Podemos asumir que $h \leq \exp((\log X)^{2/3-\epsilon})$ (como en el Ejercicio 3). Aplicando la Proposición 3.4 tenemos la cota

$$\mathbb{E}_{X<x\leq 2X}|\mathbb{E}_{(1-\frac{h}{X})x<n\leq x}\lambda(n)|^2 \ll \int_1^{1+i\frac{X}{h\delta^2}} |Z_{\lambda 1_{[X,2X]}}(s)|^2 ds + O(\delta).$$

Tomemos $\delta = h^{-1/4}$ (digamos). El resultado (con un múltiplo constante de ϵ en vez de ϵ, lo cual es claramente inofensivo) se sigue del Teorema 3.13 y el Teorema 3.16, aplicados con $h\delta^2$ en vez de h, $\alpha = \delta = 1/(\log Q_0)^\rho$, $Q_0 = \exp((\log X)^{1-\epsilon})$ y $\rho = 1 - \frac{2}{3(1-2\epsilon)}$. □

Es posible obtener una cota con $(\log\log h)/\log h$ en vez de $1/(\log h)^{1-\epsilon}$ [MR16]. Hemos optado por nuestra cota por simplicidad.

El Teorema 3.17 implica inmediatamente el Teorema 3.1, por un breve y conocido argumento (Chebyshev / cota de varianza) que ahora mostraremos. Escribamos la cota en (3.12) de la manera $\leq c_\epsilon/\text{máx}(\log h, (\log X)^{1/3})^{1-\epsilon}$. Entonces, para todo C, la proporción de valores de $X < x \leq 2X$ tales que

$$\left|\mathbb{E}_{(1-\frac{h}{X})x<n\leq x}\lambda(n)\right| \geq \frac{C\sqrt{c_\epsilon}}{\text{máx}(\log h, (\log X)^{1/3})^{(1-\epsilon)/2}}$$

es a lo más $1/C^2$.

Si, alternativamente, siguiéramos [MR16], podríamos mostrar que, para h en cierto rango, una cota más fuerte vale: la proporción de x tales que

$$|\mathbb{E}_{(1-h/X)x<n\leq x}\lambda(n)| \geq C/(\log h)^{1-\epsilon} \tag{3.13}$$

es pequeña. La diferencia principal consiste en que [MR16] separa desde un principio el conjunto $\mathcal{S}$ de enteros $X < n \leq 2X$ con factores primos del tamaño deseado, y prueba una cota de varianza para ellos; luego reintroduce los enteros que quedaron fuera de $\mathcal{S}$ al último momento. De esta manera consigue mostrar que pocos x satisfacen (3.13), aún si la forma de la cota de varianza es en esencia la misma.

3.3.1. Ejercicios.

1. Sea $2 \leq h = o(X)$, $h \in \mathbb{N}$, $h \to \infty$.
   a) Sea $f(n) = \cos(\frac{2\pi n}{100h})$.
      1) Demuestre que, si bien oscila en intervalos largos (es decir, $\sum_{X<n\leq 2X} f(n) = o(X)$), no oscila en intervalos cortos:

$$|\sum_{x<n\leq x+2h} f(n)| \gg h$$

para todo $x \in [X, 2X]$. Usando la Proposición 3.4, deduzca que

$$\int_0^{cX/h} \Big|\sum_{X<n\leq 2X} f(n)n^{-1-it}\Big|^2\, dt \gg 1$$

para alguna constante $c > 1$.

2) Muestre que

$$\sum_{X<n\leq u} f(n) \leq 100h$$

para todo u. (Consejo: la función f es periódica, con período 100.) Deduzca, por sumación por partes, que

$$\sum_{X<n\leq 2X} f(n)n^{-1-it} \ll th/X$$

para todo t.

3) Deduzca que, para h tal que $(\log h)^{150} \leq X/h$,

$$\int_{(\log h)^{100}}^{cX/h} \Big|\sum_{X<n\leq 2X} f(n)n^{-1-it}\Big|^2\, dt \gg 1$$

para alguna constante $c > 1$.

$b)$ Sea S el conjunto de números que son producto de dos primos, uno en $[P, 2P]$ y otro en $[X/P, 2X/P]$. Sea $g(n)$ la función multiplicativa tal que $g(p) = \cos(\frac{2\pi p}{100h})$. Asumiendo que $h = X^{1/10}$ y $P = \sqrt{h}$, demuestre que

$$\int_{(\log h)^{100}}^{cX/h} \Big|\sum_{n\in S} g(n)n^{-1-it}\Big|^2\, dt \ll_c \frac{1}{(\log X)^{20}}.$$

para cualquier constante $c > 1$. Sugerencia: use el Teorema del valor medio (Lema 2.10) y la Proposición 3.9 (o más bien su prueba, pues consideramos valores de t por debajo de lo que la Proposición 3.9 cubre).

2. Sea $4 \leq h \leq (\frac{X}{2})^{1/5}$. Sea S el conjunto de números que son producto de tres primos, uno en $[h, 2h]$, otro en $[h^2, 2h^2]$ y otro en $[X/h^3, 2X/h^3]$. Sea $f(n)$ una función multiplicativa tal que $Z(t) = \sum_{q\in[h^2,2h^2]} f(q)q^{-1-it} \ll (h^2)^{-1/9}$ para todo $t \in [(\log h)^{100}, X/h]$.

$a)$ Muestre que

$$\sum_{n\in S} f(n)n^{-1-it} = \sum_{p\in[h,2h]} f(p)p^{-1-it} \sum_{q\in[h^2,2h^2]} f(q)q^{-1-it} \sum_{r\in[X/h^3,2X/h^3]} f(r)r^{-1-it}$$

con p, q, r primos.

$b)$ Use la cota para $Z(t)$ y el Teorema del valor medio (Lema 2.10) para probar que

$$\int_{\substack{(\log h)^{100} \\ |\sum_{p\in[h,2h]} f(p)p^{-1-it}|>h^{-1/18}}}^{X/h} \Big|\sum_{n\in S} f(n)n^{-1-it}\Big|^2\, dt \ll \frac{1}{(h^{2/9})^2}\cdot\frac{1}{(h^{-1/18})^4} = \frac{1}{h^{2/9}}.$$

c) Use el Teorema del valor medio nuevamente para concluír que
$$\int_{(\log h)^{100}}^{X/h} |\sum_{n\in S} f(n)n^{-1-it}|^2\, dt \ll \frac{1}{h^{2/9}}.$$

4. Coeficientes de Fourier de λ en intervalos cortos, en promedio

Nuestra tarea en esta sección es acotar promedios de coeficientes de Fourier de $\lambda\cdot 1_{(x,x+h]}$:
$$\int_0^X \left|\sum_{x<n\leq x+h} \lambda(n)e(\alpha n)\right| dx.$$
Al final de la sección, veremos como tales cotas implica que el promedio de expresiones como $\lambda(n)\lambda(n+c)$ tiende a 0 para una proporción que tiende a 1 de todos los c en un pequeño intervalo ("Chowla en promedio"). Las líneas generales del argumento siguen las que se seguían para el mismo problema cuando c varía en un intervalo grande, así como para problemas análogos. (El artículo [MRT15] menciona las fuentes [DD82], [Kát86], [MV77] y [BSZ13].)

El tratamiento será distinto dependiendo de si $\alpha \in \mathbb{R}$ está en un *arco mayor* o en un *arco menor*. (En verdad la definición depende sólo de $\alpha \bmod \mathbb{Z}$.) Tomemos $Q = \sqrt{h}$. Por el teorema de aproximación de Dirichlet (Ejercicio 1), existen $a, q \in \mathbb{Z}^+$, $1 \leq q \leq Q$, $(a, q) = 1$, tales que $|\alpha - a/q| \leq 1/qQ$. Si tales a, q existen con $q \leq R$ para cierto R pequeño (será una potencia de $\log h$), decimos que α está en un *arco mayor*; de lo contrario, existen tales a, q con $q > R$, y decimos que α está en un *arco menor*.

4.0.1. Ejercicios.

1. (Teorema de aproximación de Dirichlet)
   a) Sean dados $\alpha \in \mathbb{R}$ y $M \in \mathbb{Z}^+$. Muestre que existen $0 \leq m_1 < m_2 \leq M$ tales que αm_1 y αm_2 están a distancia $\leq 1/(M+1)$ en $\mathbb{R}/\mathbb{Z}$ el uno del otro. Deduzca que $|(m_2 - m_1)\alpha - a| \leq 1/(M+1)$ para algún $a \in \mathbb{Z}^+$.
   b) Sean dados $\alpha \in \mathbb{R}$ y un real $Q \geq 1$. Concluya que existen $a, q \in \mathbb{Z}^+$, $1 \leq q \leq Q$, $(a, q) = 1$, tales que
   $$\left|\alpha - \frac{a}{q}\right| \leq \frac{1}{qQ}.$$
   (Aplique la primera parte del ejercicio con $M = \lfloor Q \rfloor$.)
2. En este problema vamos a discutir, a vuelo de pájaro, una variante relativamente sencilla del problema ternario de Goldbach, para ilustrar el uso de sumas exponenciales $\sum_n \lambda(n)e(\alpha n)$ similares a las que estudiaremos en esta sección (si bien en verdad más antiguas).

   Para $N \in \mathbb{Z}^+$, consideraremos la suma
   $$S_N = \sum_{\substack{a,b,c\geq 1\\ a+b+c=N}} \lambda(a)\lambda(b)\lambda(c).$$
   Si cambiáramos λ por la función indicatriz de los primos estaríamos calculando el número de maneras de escribir N como suma de tres primos. Nosotros queremos demostrar que $S_N = o(N^2)$, lo que nos diría que entre las tripletas de números que suman N hay la misma probabilidad de tener un número par que uno impar de primos.
   a) Demuestre la identidad $\int_0^1 e(j\alpha)\, d\alpha = 1_{j=0}$ para $j \in \mathbb{Z}$, con $e(x) = e^{2\pi i x}$. (Es la identidad sobre la cual se basa el análisis de Fourier sobre $\mathbb{R}/\mathbb{Z}$.)

b) Usando el apartado anterior, demuestre que
$$S_N = \int_0^1 (A_\lambda(\alpha))^3 e(-N\alpha)\, d\alpha,$$
con $A_\lambda(\alpha) = \sum_{1\leq n\leq N} \lambda(n)e(n\alpha)$.

c) Es posible demostrar que $\max_{\alpha\in[0,1]} |A_\lambda(\alpha)| = o(N)$. (Ésta es, claro está, la parte difícil, la cual omitimos.) Usando esta cota, demuestre que
$$S_N \leq o(N) \int_0^1 |A_\lambda(\alpha)|^2 d\alpha = o(N)N = o(N^2).$$

d) Sea $T_N = \sum_{a,b\geq 1: a+b=N} \lambda(a)\lambda(b)$ la suma correspondiente a Goldbach con dos primos. Observe que si intentamos demostrar que $T_N = o(N)$ con las técnicas de los apartados anteriores, no funciona. ¿Por qué? Esto es parecido a lo que ocurre en la integral de la Proposición 3.4, y por eso allí fue necesario factorizarla.

4.1. Arcos menores.

Proposición 4.1. *Sean $X \geq 1$, $1 \leq h \leq X$, $1 \leq R \leq \log h$. Sea $\alpha \in \mathbb{R}$ tal que existen $a, q \in \mathbb{Z}^+$, $R \leq q \leq \sqrt{h}$, $(a,q) = 1$, tales que $|\alpha - \frac{a}{q}| \leq \frac{1}{q\sqrt{h}}$. Entonces*
$$\frac{1}{hX} \int_X^{2X} \left| \sum_{x<n\leq x+h} \lambda(n)e(\alpha n) \right| dx \ll \frac{(\log R)^3}{R^{1/4}}. \tag{4.1}$$

Demostración. Probar (4.1) es equivalente a probar que, para todo $\theta : \mathbb{R} \to \mathbb{C}$ medible, con soporte en $[X, 2X]$, y tal que $|\theta(x)| \leq 1$ para todo x,
$$\int_{\mathbb{R}} \theta(x) \sum_{x<n\leq x+h} \lambda(n)e(\alpha n)dx \ll hX\frac{(\log R)^3}{R^{1/4}}. \tag{4.2}$$

Sean $\beta = 1/2$, $\delta = R^{-1/4}$, $Q_0 = \exp(R^{1/4})$ y $P_0 = \exp((\log R)^3)$. Por el Lema 3.7, el lado izquierdo de (4.2) es igual a
$$O\left(\frac{(\log R)^3}{R^{1/4}}\right) \cdot hX + \int_{\mathbb{R}} \theta(x) \sum_{x<n\leq x+h} \lambda(n)u_X(n)e(\alpha n)dx, \tag{4.3}$$
donde u_X es como en (3.3) (el error proviene de que cada valor de n satisface $x < n \leq x+h$ sólo para x dentro de un intervalo de largo $\leq h$). Entonces, por la definición (3.3) de u_X, es suficiente demostrar que
$$\frac{1}{\log(1+\delta)} \int_{P_0}^{Q_0} I(X,Q,\delta)\frac{dQ}{Q} \ll \frac{(\log R)^3}{R^{1/4}} \cdot hX, \tag{4.4}$$
donde
$$\begin{aligned}I(X,Q,\delta) &= \int_{\mathbb{R}} \theta(x) \sum_{x<n\leq x+h} \lambda(n) \left(\upsilon_{1,Q,\delta} * \upsilon_{2,Q(1+\delta),X/Q}\right)(n)e(\alpha n)dx \\ &= -\sum_m \upsilon_{2,Q(1+\delta),X/Q}(m)\lambda(m) \sum_{Q<p\leq(1+\delta)Q} e(\alpha mp) \int_{\mathbb{R}} \theta(x)1_{(x,x+h]}(mp)dx,\end{aligned}$$
y $\upsilon_{1,Q,\delta}$, $\upsilon_{2,r,Y}$ están definidos como en (3.4) y (3.5). Por Cauchy-Schwarz,
$$I(X,Q,\delta)^2 \ll \frac{X}{Q} \sum_{X/Q<m\leq 2X/Q} \left| \sum_{Q<p\leq(1+\delta)Q} e(\alpha mp) \int_{\mathbb{R}} \theta(x)1_{(x,x+h]}(mp)dx \right|^2, \tag{4.5}$$
ya que $\upsilon_{2,Q,(1+\delta),X/Q}$ tiene soporte en $[X/Q, 2X/Q]$ y $|\upsilon_{2,Q,(1+\delta),X/Q}(m)| \leq 1$ para todo m.

Ahora expandimos el cuadrado, y cambiamos el orden de la suma:[4]

$$I(X,Q,\delta)^2 \ll \frac{X}{Q} \sum_{Q<p_1,p_2\leq(1+\delta)Q} \int_{\mathbb{R}} \int_{\mathbb{R}} \theta(x_1)\overline{\theta(x_2)} \sum_{\substack{X/Q<m\leq 2X/Q \\ \frac{x_i}{p_i}<m\leq\frac{x_i+h}{p_i}}} e(\alpha(p_1-p_2)m)dx_1dx_2.$$

La suma sobre m es una suma sobre un intervalo I contenido en $(x_1/p_1, x_1/p_1+h/p_1] \subset (x_1/p_1, x_1/p_1+h/Q)$. Puede ser acotada trivialmente por $\leq |I|+1 \leq h/Q+1$; también vemos que se anula a menos que $(x_1/p_1,(x_1+h)/p_1] \cap (x_2/p_2,(x_2+h)/p_2] \neq \emptyset$, lo cual implica

$$\frac{p_2}{p_1}x_1 - h < x_2 < \frac{p_2}{p_1}x_1 + (1+\delta)h. \tag{4.6}$$

También sabemos que, para cualquier $\beta \in \mathbb{R}$ no entero y cualquier intervalo I escrito en la forma $[m_0,m_1]$, $m_0,m_1 \in \mathbb{Z}$,

$$\sum_{m\in I} e(\beta m) = \frac{e(\beta(m_1+1)) - e(\beta m_0)}{e(\beta)-1} \qquad \text{(suma geométrica)}$$

y, por lo tanto,

$$\left|\sum_{m\in I} e(\beta m)\right| \leq \frac{2}{|e(\beta)-1|} = \frac{1}{\sin \pi\beta} \leq \frac{2/\pi}{d(\beta,\mathbb{Z})}, \tag{4.7}$$

donde $d(\beta,\mathbb{Z})$ es la distancia entre β y el entero más cercano. Por lo tanto,

$$\sum_{\substack{X/Q<m\leq 2X/Q \\ \frac{x_i}{p_i}<m\leq\frac{x_i+h}{p_i}}} e(\alpha(p_1-p_2)m) \ll \text{mín}\left(\frac{h}{Q}, \frac{1}{d(\alpha(p_1-p_2),\mathbb{Z})}\right),$$

y así, por la condición (4.6),

$$I(X,Q,\delta)^2 \ll \frac{X}{Q}Xh \sum_{Q<p_1,p_2\leq(1+\delta)Q} \text{mín}\left(\frac{h}{Q}, \frac{1}{d(\alpha(p_1-p_2),\mathbb{Z})}\right).$$

Ahora bien, para todo entero l, el número de parejas $Q \leq p_1,p_2 \leq Q+\delta Q$ tales que $p_1-p_2 = l$ es $O(\delta Q)$. Vemos entonces que

$$I(Q,X,\delta)^2 \ll \delta X^2 h \sum_{\substack{l\in\mathbb{Z} \\ |l|\leq\delta Q}} \text{mín}\left(\frac{h}{Q}, \frac{1}{d(\alpha l,\mathbb{Z})}\right).$$

Ahora bien, por un lema de Vinogradov (ejercicio 1),

$$I(Q,X,\delta)^2 \ll \delta X^2 h \left(\frac{\delta Q}{q}+1\right)\left(\frac{h}{Q}+q\log q\right).$$

Como $R \leq q \leq \sqrt{h}$, $\delta = R^{-1/4}$ y $\frac{R}{\delta} \ll P_0 \leq Q \leq Q_0 = \exp(R^{1/4}) \leq h^{1/4}$, tenemos que $q\log q \ll \frac{h}{Q}$ y

$$I(Q,X,\delta)^2 \ll \delta X^2 h \left(\frac{\delta h}{q} + \frac{h}{Q}\right) \ll \frac{\delta^2 X^2 h^2}{R}.$$

Por lo tanto

$$\frac{1}{\log(1+\delta)} \int_{P_0}^{Q_0} I(X,Q,\delta)\frac{dQ}{Q} \ll (\log Q_0)\,\frac{Xh}{R^{1/2}} = \frac{Xh}{R^{1/4}}$$

[4]A partir de este momento, algunos lectores reconocerán ciertos paralelos con la versión de Linnik de la prueba del trabajo de Vinogradov sobre los tres primos.

lo que demuestra (4.4). □

4.1.1. Ejercicios.

1. (Lema de Vinogradov)
   a) Para $a, q \in \mathbb{Z}^+$ coprimos y $n = n_0, n_0+1, \dots, n_0+q-1$, n_0 arbitrario, muestre que las fracciones $n/q \bmod \mathbb{Z}$ no son sino
   $$0, 1/q, 2/q, \dots, (q-1)/q$$
   en algún orden.
   b) Sean $\alpha \in \mathbb{R}$, $a, q \in \mathbb{Z}^+$, $1 \leq q \leq Q$, $(a,q) = 1$, tales que $|\alpha - a/q| \leq 1/qQ$. Muestre que, para $n = n_0, n_0+1, \dots, n_0+q-1$, n_0 arbitrario, las distancias $d(n\alpha, \mathbb{Z})$ son a lo más
   $$0, 0, \frac{1/2}{q}, \frac{1}{q}, \frac{3/2}{q}, \dots, \frac{(q-2)/2}{q},$$
   en algún orden.
   c) Sean $\alpha \in \mathbb{R}$, $a, q \in \mathbb{Z}^+$, $1 \leq q \leq Q$, $(a,q) = 1$, tales que $|\alpha - a/q| \leq 1/qQ$. Deduzca que, para n_0 y X arbitrarios,
   $$\sum_{n=n_0}^{n_0+q-1} \min\left(X, \frac{1}{d(n\alpha, \mathbb{Z})}\right) \leq 2X + 2q(\log(q-2)+1) \ll X + q\log q.$$
   Concluya que, para N arbitrario,
   $$\sum_{|n|\leq N} \min\left(X, \frac{1}{d(n\alpha, \mathbb{Z})}\right) \ll \left(\frac{N}{q} + 1\right)(X + q\log q).$$
2. En este problema vamos a reproducir el esquema descrito en el ejercicio 2 de §4.0.1 para estimar la suma
   $$S_N = \sum_{p+q+b=N} \log p \log q$$
   que cuenta el número de maneras (ponderadas) de escribir N como suma de dos primos y un número natural cualquiera. Esta es una versión muy sencilla del problema ternario de Goldbach, y podríamos estimarla sin introducir exponenciales $e(n\alpha)$, pero por una cuestión didáctica veamos que es posible hacerlo con ellas.
   a) Procediendo como en el ejercicio 2 de §4.0.1, muestre que
   $$S_N = \int_{-1/2}^{1/2} F(\alpha)G(\alpha)^2 e(-N\alpha)\, d\alpha$$
   con $F(\alpha) = \sum_{b=1}^{N} e(b\alpha)$ y $G(\alpha) = \sum_{p\leq N} \log p\, e(p\alpha)$.
   b) En este caso los arcos menores van a ser la zona donde F es pequeña. Tomemos como arcos menores $\mathfrak{m} = [-1/2, 1/2] \setminus (-\frac{1}{N\delta}, \frac{1}{N\delta})$, con $0 < \delta < 1/2$. Demuestre usando (4.7) que si $\alpha \in \mathfrak{m}$ entonces
   $$|F(\alpha)| \ll \delta N.$$
   c) Use el apartado anterior, el razonamiento expuesto en el ejercicio 2*c* de §4.0.1 y el TNP en la forma del Teorema 2.5 para demostrar que
   $$\int_{\mathfrak{m}} F(\alpha)G(\alpha)^2 e(-N\alpha)\, d\alpha \ll \delta N^2 \log N.$$
   En los ejercicios de la siguiente sección veremos que la integral sobre los arcos mayores $\mathfrak{M} = (-\frac{1}{N\delta}, \frac{1}{N\delta})$ da la contribución principal para S_N (para cierto δ).

4.2. Arcos mayores y conclusión. Examinemos ahora el caso de $\alpha = a/q$ con q un entero pequeño. La idea es entonces que $e(\alpha n)$ es una función q-periódica. Así, lo que nos gustaría es sustituir $\lambda(n)e(\alpha n)$ por $\lambda(n)f(n)$ con $f(n)$ alguna función multiplicativa, para intentar usar las técnicas de la sección anterior. Esto es posible, usando análisis de Fourier para funciones $F : G \to \mathbb{C}$, con $G = (\mathbb{Z}/q\mathbb{Z})^\times$ el grupo (abeliano) de residuos módulo q coprimos con q, con la operación de multiplicación. Dicho análisis permite expresar F como suma de varios *carácteres módulo q*. Un carácter módulo q no es sino un homomorfismo $\chi_* : G \to \mathbb{C}^\times$. Podemos extender tal homomorfismo a una función multiplicativa y q-periódica $\chi : \mathbb{Z} \to \mathbb{C}$, simplemente poniendo $\chi(n) = 0$ para n no coprimo con q.

Nosotros sólo vamos a usar que $|\chi_*(g)| = 1$ y que los carácteres χ_* módulo q forman una base ortonormal $\widehat{G}$ del espacio de funciones de $G = (\mathbb{Z}/q\mathbb{Z})^\times$ a $\mathbb{C}$. En verdad, $\widehat{G}$ es también es un grupo. Todos estos hechos son ciertos en general para grupos abelianos finitos G. Pueden, por cierto, ser probados fácilmente (ejercicio 1).

Si α está cerca de un racional con denominador pequeño, vamos a ver cómo pasar de sumas cortas de $\lambda(n)e(\alpha n)$ a sumas cortas de $\lambda(n)\chi(n)$ para algún carácter χ de módulo pequeño.

Proposición 4.2. *Sea $1 < R^4 < h \leq X$, $|\alpha - \frac{a}{q}| \leq \frac{1}{q\sqrt{h}}$ con $q \leq R$. Entonces*

$$\frac{1}{hX}\int_X^{2X}\left|\sum_{x<n\leq x+h}\lambda(n)e(\alpha n)\right|dx \ll \frac{1}{R} + R\max_{\chi,h',X'}\frac{1}{h'X'}\int_{X'}^{3X'}\left|\sum_{x<n\leq x+h'}\lambda(n)\chi(n)\right|dx$$

donde el máximo se toma sobre los carácteres χ de módulo menor o igual que R y sobre $X' \in [X/R, X]$, $h' \in [\sqrt{h}/R^2, \sqrt{h}/R]$.

Demostración. Comenzamos con la desigualdad

$$\frac{1}{hX}\int_X^{2X}\left|\sum_{x<n\leq x+h}f(n)\right|dx \leq O\left(\frac{h_1}{h}\right) + \frac{1}{h_1X}\int_X^{3X}\left|\sum_{x<n\leq x+h_1}f(n)\right|dx$$

para $1 \leq h_1 \leq h$, $|f| \leq 1$, que simplemente proviene de dividir la suma interior en sumas de longitud h_1. La usamos con $f(n) = \lambda(n)e(\alpha n)$ y $h_1 = \sqrt{h}/R$. Ahora bien,

$$\left|\sum_{x<n\leq x+h_1}\lambda(n)e(\alpha n)\right| = \left|\sum_{x<n\leq x+h_1}\lambda(n)e\left(\frac{an}{q}\right)e\left(\left(\alpha - \frac{a}{q}\right)(n-x)\right)\right|$$

y como $|(\alpha - a/q)(n-x)| \leq \frac{1}{q\sqrt{h}}h_1 \leq R^{-1}$ tenemos que

$$|\sum_{x<n\leq x+h_1}\lambda(n)e(\alpha n)| = \left|\sum_{x<n\leq x+h_1}\lambda(n)e\left(\frac{a}{q}n\right)\right| + O(R^{-1})h_1.$$

Así

$$\frac{1}{hX}\int_X^{2X}\left|\sum_{x<n\leq x+h}\lambda(n)e(\alpha n)\right|dx \ll R^{-1} + \frac{1}{h_1X}\int_X^{3X}\left|\sum_{x<n\leq x+h_1}\lambda(n)e\left(\frac{a}{q}n\right)\right|dx.$$

Como $n \mapsto e(an/q)$ es q-periódica, podemos escribir (ejercicio 2)

$$e\left(\frac{a}{q}n\right) = \sum_{d|q}1_{d|n}\sum_{\chi\in(\mathbb{Z}/((q/d)\mathbb{Z}))^\times}a_\chi\chi(n/d),$$

donde $|a_\chi| \leq 1$. Podríamos tener una mejor cota para a_χ (*sumas de Gauss*), pero no nos importa. Obtenemos que

$$\frac{1}{X}\int_X^{3X}\left|\sum_{x<n\leq x+h_1}\lambda(n)e\left(\frac{a}{q}n\right)\right|dx \leq \frac{1}{X}\sum_{d|q}\sum_{\chi\in(\mathbb{Z}/((q/d)\mathbb{Z}))^\times}\int_X^{3X}\left|\sum_{\frac{x}{d}<m\leq\frac{x}{d}+\frac{h_1}{d}}\lambda(m)\chi(m)\right|dx$$

$$\leq q \max_{\substack{d|q \\ (\mathbb{Z}/((q/d)\mathbb{Z}))^\times}} \frac{1}{X/d}\int_{X/d}^{3X/d}\left|\sum_{x'<m\leq x'+\frac{h_1}{d}}\lambda(m)\chi(m)\right|dx'.$$

□

Con la proposición anterior vemos que sólo nos queda controlar el promedio de sumas cortas de $\lambda(n)\chi(n)$. Para ello, vamos a ver que funcionarían las técnicas de la sección anterior, igual que para $\lambda(n)$. Lo único que necesitaríamos usar en la prueba es cancelación para las sumas

$$\sum_{n<x}\lambda(n)\chi(n) \qquad \sum_{p<Q}\chi(p)p^{it_0},$$

y el control de dicha sumas depende de las funciones $Z_{\lambda\chi}(s) = Z_{\chi^2}(2s)/Z_\chi(s)$ y $Z'_\chi(s)/Z_\chi(s)$ con $Z_\chi(s) = L(s,\chi)$, donde $s \mapsto L(s,\chi)$ son las así llamadas funciones L de Dirichlet:

$$L(s,\chi) = \sum_n \chi(n)n^{-s}.$$

Por lo tanto, necesitaremos control sobre los ceros y el tamaño de dichas funciones zeta. Para el caso de λ usamos [IK04, Thm. 8.29] (que es nuestro Teorema 2.2) y en el lema anterior a dicho teorema puede verse que dicho control proviene de cotas superiores para $|L(s,\chi)|$ en dicha zona. Como

$$L(s,\chi) = q^{-s}\sum_{b(q)}\chi(b)\sum_{m=0}^{\infty}\left(m+\frac{b}{q}\right)^{-s}$$

y $\sum_{m=0}^{\infty}(m+\alpha)^{-s}$ es muy similar a $\zeta(s)$, esencialmente podemos conseguir para $|L(s,\chi)|$ las mismas cotas que para $|\zeta(s)|$, y así tenemos un equivalente al Teorema 2.2 para este caso.

Teorema 4.3. *Hay una constante $c > 0$ tal que $L(s,\chi) \neq 0$ para $s = \sigma + it$ con $\sigma \geq 1 - \frac{c}{\log q + (\log t)^{2/3}(\log\log t)^{1/3}}$, $|t| \geq 3$, y en dicha zona también se cumplen las cotas*

$$\frac{1}{L(s,\chi)} \ll \log q + (\log t)^{2/3}(\log\log t)^{1/3},$$

$$\frac{L'(s,\chi)}{L(s,\chi)} \ll \log q + (\log t)^{2/3}(\log\log t)^{1/3}.$$

Además, en $|t| \leq 3$ se cumple que $L(s,\chi) \neq 0$ en $\sigma > 1 - \frac{c}{\sqrt{q}(\log q)^2}$ y en esa zona $\frac{1}{L(s,\chi)} \ll \sqrt{q}(\log q)^2$, $\frac{L'(s,\chi)}{L(s,\chi)} \ll \sqrt{q}(\log q)^2$.

Las cotas para $|t| \leq 3$ son clásicas y provienen de otras técnicas [MV07, Capítulo 11]. Éstos resultados son todos efectivos; no esconden constantes no especificables – en particular, no lidiamos con los *ceros de Siegel*.

Para demostrar el Teorema 2.3 usamos las cotas del Teorema 2.2 con t escogido de manera tal que $(\log t)^{2/3}(\log\log t)^{1/3}$ sea igual a $(\log x)^{2/5}(\log\log x)^{1/5}$, por lo que, si

$\sqrt{q}(\log q)^2 \ll (\log x)^{2/5}(\log\log x)^{1/5}$, tendremos las mismas cotas en el Teorema 4.3, y luego serán ciertos los resultados análogos al Teorema 2.3 y al Corolario 2.4 para $\lambda(n)\chi(n)$. En particular

Corolario 4.4. *Sean $x \geq 1$, $t \leq e^{(\log x)^{3/5-\epsilon}}$, $\epsilon > 0$, y χ un carácter de módulo $q \leq (\log x)^{4/5-\epsilon}$. Entonces*

$$\sum_{x<n\leq 2x} \frac{\lambda(n)\chi(n)}{n^{1+it}} \ll \exp(-(\log x)^{3/5+o_\epsilon(1)}).$$

Además, en la zona $\log t > (\log Q)^a$ con $a > 0$, si $q \leq (\log Q)^2$ tenemos que las cotas del Teorema 4.3 serían las mismas que sin $\log q$, y luego el resultado análogo a la Proposición 3.9 también será cierto.

Proposición 4.5. *Sea χ un carácter de módulo $q \leq (\log Q)^2$. Para $\exp((\log Q)^a) \leq t_0 \leq \exp((\log Q)^{(3/2)(1-a)})$, $a > 0$, $0 < \delta \leq 1$,*

$$\sum_{Q<p\leq(1+\delta)Q} \chi(p)p^{-1-it_0} \ll \exp(-(\log Q)^{a+o(1)}).$$

Como en la demostración del Teorema 3.17 se usa el equivalente al Corolario 4.4 con $x \approx X$ y el equivalente a la Proposición 4.5 con $\log Q = (\log X)^{1-\epsilon}$, tenemos que

Teorema 4.6. *Sea $1 < h \leq X$, $\epsilon > 0$. Para $q \leq (\log X)^{4/5-\epsilon}$,*

$$\mathbb{E}_{X<x\leq 2X}|\mathbb{E}_{(1-\frac{h}{X})x<n\leq x}\lambda(n)\chi(n)|^2 \ll_\epsilon \frac{1}{(\log h)^{1-\epsilon}} + \frac{1}{(\log X)^{1/3-\epsilon}}.$$

Corolario 4.7. *Sea $1 < h \leq X$, $\epsilon > 0$. Para $q \leq (\log X)^{4/5-\epsilon}$,*

$$\mathbb{E}_{X<x\leq 2X}\,|\mathbb{E}_{x<n\leq x+h}\lambda(n)\chi(n)| \ll_\epsilon \frac{1}{(\log h)^{1/3-\epsilon}} + \frac{1}{(\log X)^{1/9-\epsilon}}.$$

Demostración. Notemos que es suficiente demostrar la desigualdad

$$\mathbb{E}_{X<x\leq 2X}\,|\mathbb{E}_{x-h<n\leq x}\lambda(n)\chi(n)| \ll_\epsilon \frac{1}{(\log h)^{1/3-\epsilon}} + \frac{1}{(\log X)^{1/9-\epsilon}}. \tag{4.8}$$

Ahora, para $0 < \delta \leq 1$ tenemos que

$$\mathbb{E}_{X<x\leq 2X}\,|\mathbb{E}_{x-h<n\leq x}\lambda(n)\chi(n)| \ll \mathbb{E}_{j:1<(1+\delta)^j\leq 2}\,\mathbb{E}_{\substack{Y<x\leq Y+\delta Y\\ Y=(1+\delta)^j X}}\,|\mathbb{E}_{x-h<n\leq x}\lambda(n)\chi(n)|$$

y como para cualquier $x \in (Y, Y+\delta Y]$ se cumple que

$$\mathbb{E}_{x-h<n\leq x}\lambda(n)\chi(n) = O(\delta) + \mathbb{E}_{x-h\frac{x}{Y}<n\leq x}\lambda(n)\chi(n)$$

deducimos que

$$\mathbb{E}_{X<x\leq 2X}\,|\mathbb{E}_{x-h<n\leq x}\lambda(n)\chi(n)| \ll \delta + \max_{Y\in[X,2X]} \mathbb{E}_{Y<x\leq Y+\delta Y}|\mathbb{E}_{x-h\frac{x}{Y}<n\leq x}\lambda(n)\chi(n)|.$$

Ahora aplicamos Cauchy-Schwarz:

$$\mathbb{E}_{Y<x\leq Y+\delta Y}|\mathbb{E}_{x-h\frac{x}{Y}<n\leq x}\lambda(n)\chi(n)| \ll \sqrt{\mathbb{E}_{Y<x\leq Y+\delta Y}|\mathbb{E}_{x-h\frac{x}{Y}<n\leq x}\lambda(n)\chi(n)|^2},$$

y luego, por el Teorema 4.6 (acotando brutalmente la integral de una cantidad no negativa sobre $(Y, Y+\delta Y]$ por la integral de la misma cantidad sobre $(Y, 2Y]$) vemos que

$$\mathbb{E}_{X<x\leq 2X}\,|\mathbb{E}_{x-h<n\leq x}\lambda(n)\chi(n)| \ll \delta + \max_{Y\in[X,2X]} \sqrt{\delta^{-1}\left(\frac{1}{(\log h)^{1-\epsilon}} + \frac{1}{(\log Y)^{1/3-\epsilon}}\right)}.$$

Finalmente, tomando

$$\delta = \text{mín}\left(1, \text{máx}\left(\frac{1}{(\log h)^{1/3-\epsilon}}, \frac{1}{(\log X)^{1/9-\epsilon}}\right)\right)$$

demostramos (4.8).

$\square$

Usando los resultados anteriores, obtenemos lo siguiente.

Proposición 4.8. *Sea* $1 < R^5 \leq h \leq X$, $|\alpha - \frac{a}{q}| \leq \frac{1}{q\sqrt{h}}$ *con* $q \leq R \leq (\log X)^{4/5-\epsilon}$. *Entonces, para* $\epsilon > 0$,

$$\frac{1}{hX}\int_X^{2X}\left|\sum_{x<n\leq x+h}\lambda(n)e(\alpha n)\right|dx \ll_\epsilon \frac{1}{R} + \frac{R}{(\log h)^{\frac{1}{3}-\epsilon}} + \frac{R}{(\log X)^{\frac{1}{9}-\epsilon}}.$$

Demostración. Por la Proposición 4.2 y el Corolario 4.7. $\square$

Finalmente, usando el Teorema de aproximación de Dirichlet y tomando R igual a $\text{mín}((\log h)^{1/3}, (\log X)^{1/9})^{4/5}$ en la Proposiciones 4.1 y 4.8, concluímos que

Teorema 4.9. *Sean* $X > 1$, $1 < h \leq X$. *Entonces, para todo* $\alpha \in \mathbb{R}$ *y* $\epsilon > 0$,

$$\frac{1}{hX}\int_X^{2X}\left|\sum_{x<n\leq x+h}\lambda(n)e(\alpha n)\right|dx \ll_\epsilon \frac{1}{(\log h)^{\frac{1}{15}-\epsilon}} + \frac{1}{(\log X)^{\frac{1}{45}-\epsilon}}.$$

Los exponentes $1/15$, $1/45$ no son de ninguna manera óptimos; el lector interesado puede divertirse mejorándolos.

Vamos a ver como el teorema anterior implica la conjetura de Chowla en promedio (Teorema 1.2). Para ello, veamos la relación entre las sumas cortas de $f(n)$ y sumas largas de $\overline{f(n)}f(n+h)$. Supongamos que $f(n)$ tiene soporte finito. Entonces, expandiendo el cuadrado y cambiando el orden de sumación obtenemos

$$\sum_x |\sum_{x<n\leq x+h} f(n)|^2 = \sum_{|n-m|<h} f(m)\overline{f(n)}(h - |n-m|)$$

Por tanto, escribiendo $m = n + j$, tenemos que

$$\sum_x |\sum_{x<n\leq x+h} f(n)|^2 = \sum_{|j|<h}(h-|j|)\sum_n f(n+j)\overline{f(n)}.$$

Ahí vemos directamente que las sumas cortas de $f(n)$ controlan un promedio de sumas largas de $\overline{f(n)}f(n+j)$. El problema es que podría ser que hubiera cancelación en la suma en j en vez de en la suma en n. Para evitarlo, usamos la misma identidad con $f(n) = f_0(n)e(n\alpha)$, que da

$$\sum_x |\sum_{x<n\leq x+h} f_0(n)e(n\alpha)|^2 = \sum_{|j|<h}\left[(h-|j|)\sum_n f_0(n+j)\overline{f_0(n)}\right]e(j\alpha).$$

Ahora, teniendo en cuenta la identidad de Parseval

$$\int_0^1 |\sum_j a_j e(j\alpha)|^2 d\alpha = \sum_j |a_j|^2, \tag{4.9}$$

la cual simplemente proviene de que $\int_0^1 e(m\alpha)\,d\alpha = 1_{m=0}$, tenemos que

$$\sum_{|j|<h}|(h-|j|)|^2\left|\sum_n f_0(n+j)\overline{f_0(n)}\right|^2 = \int_0^1\left|\sum_x|\sum_{x<n\leq x+h} f_0(n)e(n\alpha)|^2\right|^2 d\alpha.$$

Así, si

$$M = \max_{\alpha} \sum_x |\sum_{x<n\leq x+h} f_0(n)e(n\alpha)|^2$$

entonces sacándolo fuera de la integral tenemos

$$\sum_{|j|<h} |(h-|j|)|^2 \left|\sum_n f_0(n+j)\overline{f_0(n)}\right|^2 \leq M \sum_x \int_0^1 |\sum_{x<n\leq x+h} f_0(n)e(n\alpha)|^2 \, d\alpha,$$

y usando de nuevo (4.9) sobre la parte derecha conclüímos que

$$\sum_{|j|<h} |(h-|j|)|^2 \left|\sum_n f_0(n+j)\overline{f_0(n)}\right|^2 \leq M \sum_x \sum_{x<n\leq x+h} |f_0(n)|^2.$$

Finalmente, aplicando esta desigualdad con $f_0(n) = \lambda(n)1_{(X,2X]}(n)$ y usando el Teorema 4.9, obtenemos

Corolario 4.10. *Para $1 \leq h \leq X$ y para todo $\epsilon > 0$*

$$\frac{1}{hX^2} \sum_{j\leq h/2} |\sum_{X<n,n+j\leq 2X} \lambda(n+j)\lambda(n)|^2 \ll_\epsilon \frac{1}{(\log h)^{\frac{1}{15}-\epsilon}} + \frac{1}{(\log X)^{\frac{1}{45}-\epsilon}}.$$

A partir de ahí es muy sencillo eliminar la condición $X < n+j \leq 2X$, y así obtener el Teorema 1.2 en el caso de dos factores ($k=2$). El caso de más factores se sigue del siguiente lema (ejercicio 3)

Lema 4.11. *Sea $1 \leq H \leq X$. Para funciones $f, g : \mathbb{Z}^+ \to \mathbb{C}$ cualesquiera con $|g(n)| \leq 1$, $|f(n)| \leq 1$ para todo n y soporte en $[1, X]$, se cumple la desigualdad*

$$\frac{1}{HX^2} \sum_{h\leq H} |\sum_n f(n+h)g(n)|^2 \leq \sqrt{\frac{1}{HX^2} \sum_{|h|\leq H} |\sum_n f(n+h)\overline{f(n)}|^2}.$$

Reflexiones finales. Como hemos visto, una vez que se tiene una cierta cota sobre sumas exponenciales en promedio (Teorema 4.9) podemos obtener el resultado de tipo "Chowla en promedio" que deseábamos (Teorema 1.2) con bastante facilidad.

Es bueno reflexionar sobre que tipo de cotas sobre sumas exponenciales necesitamos para obtener otros resultados sobre las autocorrelaciones de λ. ¿Qué pasa si deseamos saber el promedio de $\lambda(n)\lambda(n+h)\lambda(n+2h)$ para la mayor parte de valores de h en un intervalo pequeño, digamos?

Muchas tales preguntas se reducen a acotar *seminormas de Gowers* $|\lambda|_{U^k}$. Ya sería un muy buen comienzo, por ejemplo, poder acotar la *segunda norma de Gowers* $|\lambda|_{U^2}$, definida por

$$|\lambda|_{U^2}^2 = \frac{1}{H^2X} \sum_{1\leq h_1,h_2\leq H} \sum_{1\leq n\leq X} \lambda(n)\lambda(n+h_1)\lambda(n+h_2)\lambda(n+h_1+h_2) \tag{4.10}$$

para todo $H = H(X) \to \infty$. En efecto, si pudiéramos mostrar que $|\lambda|_{U^2} = o(1)$, tendríamos que

$$\frac{1}{HX} \sum_{h\leq H} \sum_{n\leq X} \lambda(n)\lambda(n+h)\lambda(n+2h) = o(1)$$

despues de algunas aplicaciones de Cauchy-Schwarz. Más aún, si pudiéramos mostrar que la tercera norma de Gowers $|\lambda|_{U^3}$ es $o(1)$, podríamos deducir, de manera similar, que

$$\frac{1}{HX^2} \sum_{h\leq H} \left|\sum_{n\leq X} \lambda(n)\lambda(n+h)\lambda(n+2h)\right|^2 = o(1).$$

Acotar (4.10) por $o(1)$ resulta ser equivalente a probar la *conjetura de uniformidad de Fourier*

$$\frac{1}{hX}\int_0^X \sup_{\alpha\in\mathbb{R}/\mathbb{Z}}\left|\sum_{x<n\leq x+h}\lambda(n)e(\alpha n)\right| dx = o(1)$$

para $h = h(N) \to \infty$. Esta conjetura, planteada por primera vez en [MRT15], parece más difícil que el Teorema 4.9, en el cual el orden de la integral y del supremo $\sup_{\alpha\in\mathbb{R}/\mathbb{Z}}$ (implícito en el Teorema 4.9) es el inverso.

4.2.1. Ejercicios.

1. Sea G un grupo abeliano finito.
   a) Pruebe que, para todo carácter χ de G y todo $g \in G$, $|\chi(g)| = 1$. (Sugerencia: muestre que $\chi(g)^{|G|} = 1$.)
   b) Sea χ un carácter no trivial de G, es decir, un carácter tal que existe un $g \in G$ para el cual $\chi(g) \neq 1$. Muestre que
   $$\psi(g)\sum_{h\in G}\psi(h) = \sum_{h\in G}\psi(gh) = \sum_{h\in G}\psi(h).$$
   Concluya que $\sum_{h\in G}\psi(h) = 0$.
   c) Sean χ, χ' dos carácteres distintos de G. Muestre que $\psi = \overline{\chi}\cdot\chi'$ es un carácter no trivial. Entonces, por (1*b*), $\sum_{g\in G}\overline{\chi}(g)\chi'(g) = 0$.
   d) Muestre que hay $|G|$ carácteres distintos $\chi : G \to \mathbb{C}$. Concluya que los carácteres de G forman una base ortonormal del espacio de funciones de G a $\mathbb{C}$ con el producto escalar
   $$\langle f_1, f_2\rangle = \frac{1}{|G|}\sum_{g\in G}\overline{f_1}(g)f_2(g).$$
2. *a*) Muestre que, para toda función q-periódica $f : \mathbb{Z} \to \mathbb{C}$ con soporte en los enteros coprimos con q,
   $$f(n) = \sum_{\chi\in\widehat{G}_q}\widehat{f}(\chi)\chi(n),$$
   donde $G_q = (\mathbb{Z}/q\mathbb{Z})^\times$, $\widehat{G}_q$ es el grupo de carácteres de G (ejercicio 1) y
   $$\widehat{f}(\chi) = \frac{1}{|G|}\sum_{m\in(\mathbb{Z}/q\mathbb{Z})^\times}\overline{\chi(m)}f(m).$$
   Para este propósito, utilice el hecho que $\widehat{G}_q$ es una base ortonormal (ejercicio 1*d*). Está claro que, si $|f(m)| \leq 1$ para todo m, $|\widehat{f}(\chi)| \leq 1$.
   b) Para toda función q-periódica $f : \mathbb{Z} \to \mathbb{C}$,
   $$f(n) = \sum_{d|q} f(n)1_{(n,q)=d} = \sum_{d|q} 1_{d|n}f_d(n/d),$$
   donde $f_d : \mathbb{Z} \to \mathbb{C}$ es la función tal que $f_d(m) = f(md)$ si $(m, q/d) = 1$ y $f_d(m) = 0$ de otra manera. Concluya que para ciertos $|a_\chi| \leq 1$
   $$f(n) = \sum_{d|q}\sum_{\chi\in\widehat{G}_{q/d}} a_\chi 1_{d|n}\chi(n/d).$$
3. Demuestre el Lema 4.11. Para hacerlo: expanda el cuadrado, meta adentro la suma en h, aplique Cauchy-Schwarz, expanda los cuadrados y reagrúpelos de forma que queden cuadrados de sumas en n. Aplique dicho lema para demostrar el Teorema 1.2 a partir del Corolario 4.10.

4. En este problema vamos a concluir la estimación de $S_N = \sum_{p+q+b=N} \log p \log q$ que comenzamos en el ejercicio 2 de §4.1.1 . Allí vimos que para cualquier $0 < \delta < 1/2$

$$S_N = O(\delta N^2 \log N) + \int_{\mathfrak{M}} F(\alpha) G^2(\alpha) e(-N\alpha)\, d\alpha$$

con $\mathfrak{M} = (-\frac{1}{N\delta}, \frac{1}{N\delta})$ los arcos mayores, $F(\alpha) = \sum_{b=1}^N e(b\alpha)$, $G(\alpha) = \sum_{p\leq N} \log p\, e(p\alpha)$.

a) Use la regla del rectángulo (2.19) para obtener la estimación

$$F(\alpha) = \frac{e(N\alpha) - 1}{2\pi i \alpha} + O(\delta^{-1}) \tag{4.11}$$

para $\alpha \in \mathfrak{M}$.

b) Escriba $G(\alpha) = \sum_{n\leq N} (1_{n \text{ primo}} \log n) e(n\alpha)$. Use sumación por partes (ejercicio 2 de §2.1.1), el teorema 2.5 y la estimación (4.11) para obtener que

$$G(\alpha) = \frac{e(N\alpha) - 1}{2\pi i \alpha} + O_A\left(\frac{\delta^{-1} N}{(\log N)^A}\right)$$

para $\alpha \in \mathfrak{M}$ y $A > 0$ arbitrario.

c) Por los apartados anteriores y el cambio de variable $\alpha = t/N$, demuestre que

$$S_N = (C + O(\delta^2))N^2 + O(\delta N^2 \log N) + O_A\left(\frac{N^2}{\delta^2 (\log N)^A}\right)$$

con C definida como la constante $\int_{-\infty}^{\infty} \frac{(e(t)-1)^3 e(-t)}{(2\pi i t)^3}\, dt$. Tomando $\delta = (\log N)^{-2}$, $A = 5$, obtenemos que

$$S_N = (C + o(1))N^2.$$

d) Para demostrar que $C = 1/2$ sin dolor, verifique que el mismo razonamiento muestra que $T_N = \sum_{a+b+c=N} 1$ satisface $T_N = (C+o(1))N^2$, y luego muestre que $T_N = (1/2 + o(1))N^2$ de otra manera. Alternativamente, muestre que $C = 1/2$ como prefiera.

5. La autocorrelación de λ en escala logarítmica

5.1. Inicio y esbozo del argumento. Quisiéramos ahora probar el Teorema 1.3. Para simplificar la notación, nos concentraremos en el caso $a_1 = a_2 = b_2 = 1$, $b_1 = 0$; es decir, probaremos que, para $w = w(x)$ tal que $w \to \infty$ cuando $x \to \infty$,

$$\sum_{\frac{x}{w} < n \leq x} \frac{\lambda(n)\lambda(n+1)}{n} = o(\log w). \tag{5.1}$$

El tratamiento del caso general (a_i, b_i arbitrarios) es prácticamente idéntico.

De hecho, probaremos (5.1) en la siguiente forma cuantitativa.

Teorema 5.1. *Sean $w > e^e$, $x > e^{e^e}$. Entonces*

$$\sum_{x/w < n \leq x} \frac{\lambda(n)\lambda(n+1)}{n} \ll \frac{\log w}{\text{mín}(\log_3 w, \log_4 x)^{1/5}}.$$

Cuando escribimos $\log_k$, queremos decir el logaritmo iterado k veces: $\log_2 x = \log\log x$, $\log_3 x = \log\log\log x$, $\log_4 x = \log\log\log\log x$.

El primer paso hacia el Teorema 5.1 consiste en usar la multiplicatividad de λ para escribir la suma como una suma de sumas con una condición de divisibilidad.

Lema 5.2. *Sea $1 \leq w \leq x$. Sean $1 \leq K_0 \leq K_1 < x/w$. Entonces*

$$\sum_{\frac{x}{w}<n\leq x} \frac{\lambda(n)\lambda(n+1)}{n} = \frac{1}{\ell}\left(\sum_{K_0<p\leq K_1} \sum_{\substack{\frac{x}{w}<n\leq x \\ p|n}} \frac{\lambda(n)\lambda(n+p)}{n} + O(\log K_1)\right), \tag{5.2}$$

donde $\ell = \sum_{K_0<p\leq K_1} p^{-1}$.

La idea principal de la prueba es que el intervalo $x/w < n \leq x$ con el peso $1/n$ es casi invariante bajo desplazamientos multiplicativos $p\cdot$, p pequeño.

Demostración. Está claro que

$$\sum_{K_0<p\leq K_1} \frac{1}{p} \sum_{\frac{x}{w}<n\leq x} \frac{\lambda(n)\lambda(n+1)}{n} = \sum_{K_0<p\leq K_1} \sum_{\frac{x}{w}<n\leq x} \frac{\lambda(pn)\lambda(pn+p)}{pn}$$

Ahora bien, para $p \leq K_1$,

$$\begin{aligned}\sum_{\frac{x}{w}<n\leq x} \frac{\lambda(pn)\lambda(pn+p)}{pn} &= \sum_{\frac{x}{pw}<n\leq\frac{x}{p}} \frac{\lambda(pn)\lambda(pn+p)}{pn} + O\left(\frac{1}{p}\sum_{\frac{x}{pw}<n\leq\frac{x}{w}} \frac{1}{n} + \frac{1}{p}\sum_{\frac{x}{p}<n\leq x} \frac{1}{n}\right) \\ &= \sum_{\substack{\frac{x}{w}<n\leq x \\ p|n}} \frac{\lambda(n)\lambda(n+p)}{n} + \frac{O(\log p)}{p}.\end{aligned}$$

Aplicando el Corolario 2.6 y dividiendo por $\sum_{K_0<p\leq K_1} p^{-1}$, obtenemos el resultado. □

Tras este resultado, para demostrar (5.1), será suficiente mostrar que para algún par (K_0, K_1) con $\log K_1 = o(\log w \sum_{K_0<p\leq K_1} \frac{1}{p})$ se cumple

$$\sum_{K_0<p\leq K_1} \sum_{\substack{\frac{x}{w}<n\leq x \\ p|n}} \frac{\lambda(n)\lambda(n+p)}{n} = o\left((\log w)\cdot \sum_{K_0<p\leq K_1} \frac{1}{p}\right). \tag{5.3}$$

La idea principal es la siguiente. Los métodos de Matomäki y Radziwiłł (en particular, los que acabamos de ver en §4) nos bastarán para probar que

$$\sum_{K_0<p\leq K_1} \frac{1}{p} \sum_{\frac{x}{w}<n\leq x} \frac{\lambda(n)\lambda(n+p)}{n} = o\left((\log w)\cdot \sum_{K_0<p\leq K_1} \frac{1}{p}\right). \tag{5.4}$$

Ahora bien, si vemos al hecho de ser divisible por p como un evento aleatorio de probabilidad $1/p$, tiene sentido que los lados izquierdos de (5.3) y (5.4) sean aproximadamente iguales.

En verdad, que n sea divisible por p *es* un evento aleatorio, si tomamos n al azar entre x/w y x. El problema reside en que se trata de un evento no independiente de $\lambda(n)\lambda(n+p)$.

Ahora bien, resulta ser que – para hablar de manera aproximada – si la dependencia entre los eventos $p|n$ $(K_0 < p \leq K_1)$ y $(\lambda(n), \lambda(n+1), \dots)$ $(x/w < n \leq x)$ es fuerte para muchos valores de (K_0, K_1), entonces existe un valor de (K_0, K_1) en la cual no lo es tanto. Existe una medida de dependencia – la *información mutua*, definida en términos de la *entropía* – la cual, si bien es un tanto burda, goza de una propiedad de aditividad. Esto conlleva que se pueda tratar a la entropía como un recurso agotable; podemos pensar en ella como una sopa con una cantidad finita de lentejas, de tal manera que, si la gente se va sirviendo, eventualmente a alguien le tendrá que tocar pocas lentejas (es decir, poca información mutua).[5] En el momento que nos toca pocas lentejas, la dependencia es débil, y sabremos proceder.

5.1.1. Ejercicios.

1. Queremos ver que el resultado (5.1) no es tan fuerte como la conjetura de Chowla. Sea $(a_n)_{n\in\mathbb{N}}$ una sucesión con $|a_n| \leq 1$.
   a) Demuestre que $\sum_{n\leq x} a_n = o(x)$ implica $\sum_{x/w\leq n\leq x} \frac{a_n}{n} = o(\log w)$ cuando $x \to \infty$, con $x \geq w = w(x) \to \infty$. *Sugerencia: use sumación por partes.*
   b) Observe, usando la regla del rectángulo (2.19), que la sucesión $a_n = n^i$ satisface $\sum_{x/w\leq n\leq x} a_n/n = O(1) = o(\log w)$ cuando $x \to \infty$ para cualquier $x \geq w = w(x) \to \infty$, pero que $\sum_{n\leq x} a_n \neq o(x)$. Muestre que lo mismo ocurre para $a_n = \Re n^i = \cos(\log n)$.
   c) Demuestre que si tuviéramos cancelación para $w = 2$, es decir $\sum_{x/2<n\leq x} \frac{a_n}{n} = o(1)$, entonces sí podríamos deducir que $\sum_{n\leq x} a_n = o(x)$. *Sugerencia: use sumación por partes.*
2. El Teorema 5.1 vale en el rango $2 \leq w \leq x^{1/8}$. Usando ese resultado, deduzca que el teorema también es cierto en el rango $x^{1/8} < w \leq x$.

5.2. Sumas y esperanzas. En esta sección vamos a formalizar la relación entre las sumas en (5.3) y los conceptos de probabilidad e independencia que hemos comentado. Para ello, es conveniente primero partir la suma en n en segmentos cortos (de longitud H). Esto es lo que hacemos en el siguiente resultado.

Lema 5.3. *Sea H un número natural tal que $K_1 < H \leq x/w$. Entonces para cualquier $p \leq K_1$ tenemos*
$$\sum_{\substack{\frac{x}{w}<n\leq x\\ p|n}} \frac{\lambda(n)\lambda(n+p)}{n} = \frac{1}{H}\sum_{\frac{x}{w}<n\leq x}\frac{1}{n}\sum_{j\leq H-p}\lambda(n+j)\lambda(n+j+p)1_{p|n+j} + O\left(\frac{\log w}{H}+\frac{1}{p}\right).$$

La idea es simplemente utilizar el hecho de que el peso $1/n$ y el intervalo $(x/w, x]$ son aproximadamente invariantes bajo pequeños desplazamientos aditivos.

Demostración. Para cualquier $j \leq H$, por cambio de variable
$$\sum_{\substack{\frac{x}{w}<n\leq x\\ p|n}} \frac{\lambda(n)\lambda(n+p)}{n} = \sum_{\substack{\frac{x}{w}-j<n\leq x-j\\ p|n+j}} \frac{\lambda(n+j)\lambda(n+j+p)}{n+j}.$$

Como
$$\sum_{\substack{\frac{x}{w}-j<n\leq \frac{x}{w}\\ p|n+j}} \frac{1}{n+j} + \sum_{\substack{x-j<n\leq x\\ p|n+j}} \frac{1}{n+j} = \sum_{\substack{\frac{x}{w}<n\leq \frac{x}{w}+j\\ p|n}} \frac{1}{n} + \sum_{\substack{x<n\leq x+j\\ p|n}} \frac{1}{n} \ll \frac{1}{p}\left(1+\log\frac{x/w+H}{x/w}\right) \ll \frac{1}{p}$$
y
$$\begin{aligned}\sum_{\substack{\frac{x}{w}<n\leq x\\ p|n+j}} \left(\frac{1}{n}-\frac{1}{n+j}\right) &\leq \sum_{\substack{n>x/w\\ p|n+j}} \frac{j}{n(n+j)} \leq 2H\sum_{\substack{n>x/w\\ p|n+j}} \frac{1}{(n+j)^2}\\ &\leq 2H\sum_{\substack{n>x/w\\ p|n}} \frac{1}{n^2} \ll \frac{H}{p^2}\cdot\frac{1}{x/wp} \leq \frac{1}{p},\end{aligned}$$

[5] En la versión oral de estas charlas, se mencionó a un ollón de locro, pero se hizo aparente que parte de la audiencia no sabía qué era el locro.

vemos que

$$(H-p)\sum_{\substack{\frac{x}{w}<n\leq x\\ p|n}}\frac{\lambda(n)\lambda(n+p)}{n}=\sum_{j\leq H-p}\sum_{\frac{x}{w}<n\leq x}\frac{\lambda(n+j)\lambda(n+j+p)1_{p|n+j}}{n}+O\left(\frac{1}{p}\right)\cdot(H-p).$$

Dividamos todo por H. Para concluír, notemos que

$$\frac{p}{H}\left|\sum_{\substack{\frac{x}{w}<n\leq x\\ p|n}}\frac{\lambda(n)\lambda(n+p)}{n}\right|\leq\frac{p}{H}\sum_{\substack{\frac{x}{w}<n\leq x\\ p|n}}\frac{1}{n}\ll\frac{p}{H}\cdot\frac{\log w+1}{p}\leq\frac{\log w}{H}+\frac{1}{p}.$$

□

Tomemos $K_0=\epsilon H/2$, $K_1=\epsilon H$, con $0<\epsilon<1$ pequeño. Nuestro objetivo – el cual nos permitirá demostrar (5.3), gracias al Lema 5.3 – será acotar de forma no trivial la suma

$$S=\frac{1}{H}\sum_{\frac{x}{w}<n\leq x}\frac{1}{n}\sum_{K_0<p\leq K_1}\sum_{j\leq H-p}\lambda(n+j)\lambda(n+j+p)1_{p|n+j}. \tag{5.5}$$

Para ver la conexión con las probabilidades, observemos que si N es la variable aleatoria que toma valores en el conjunto de enteros en el intervalo $(x/w,x]$ con probabilidad

$$\mathbb{P}(N=n)=\frac{1/n}{L}\quad\text{si}\quad n\in(x/w,x], \tag{5.6}$$

donde $L=\sum_{x/w<n\leq x}\frac{1}{n}$, entonces podemos escribir

$$S=\frac{L}{H}\cdot\mathbb{E}\left(\sum_{K_0<p\leq K_1}\sum_{j\leq H-p}\lambda(N+j)\lambda(N+j+p)1_{N\equiv -j(p)}\right),$$

es decir, S es la esperanza de una variable aleatoria que es una suma doble de variables aleatorias. Para examinar la dependencia entre la condición de divisibilidad y los términos con λ, definimos la variables aleatorias

$$X_H=(\lambda(N+1),\lambda(N+2),\ldots,\lambda(N+H)),\qquad Y_H=(N\bmod p)_{K_0<p\leq K_1}, \tag{5.7}$$

con X_H tomando valores en $\{-1,1\}^H$ e Y_H en $\Omega=\prod_{K_0<p\leq K_1}\frac{\mathbb{Z}}{p\mathbb{Z}}$. Así podemos escribir

$$S=\frac{L}{H}\cdot\mathbb{E}(F(X_H,Y_H)) \tag{5.8}$$

con $F:\{-1,1\}^H\times\Omega\to\mathbb{R}$ la función

$$F(\vec{x},\vec{y})=\sum_{K_0<p\leq K_1}\sum_{j\leq H-p}x_jx_{j+p}1_{y_p\equiv -j(p)}. \tag{5.9}$$

Obtenemos el siguiente resultado. El interés en estimar la expresión en el lado izquierdo de (5.10) viene, claro está, del hecho que aparece en el lado derecho de (5.2).

Lema 5.4. *Sea* $1\leq w<x$. *Sean* $K_0=\epsilon H/2$ *y* $K_1=\epsilon H$, *con* $H\leq x/w$ *y* $\text{máx}\left(\frac{1}{\log w},\frac{1}{\sqrt{H}}\right)\leq\epsilon<1$. *Entonces*

$$\sum_{K_0<p\leq K_1}\sum_{\substack{\frac{x}{w}<n\leq x\\ p|n}}\frac{\lambda(n)\lambda(n+p)}{n}=\frac{L}{H}\mathbb{E}(F(X_H,Y_H))+O\left(\epsilon\frac{\log w}{\log H}\right), \tag{5.10}$$

donde $L=\sum_{x/w<n\leq x}n^{-1}$ *y* F *es como en* (5.9).

Demostración. Por el Lema 5.3 y las estimaciones del Corolario 2.6,

$$\begin{aligned}
\sum_{K_0<p\le K_1}\sum_{\substack{\frac{x}{w}<n\le x\\ p|n}}\frac{\lambda(n)\lambda(n+p)}{n} &= S+\sum_{K_0<p\le K_1} O\left(\frac{1}{p}+\frac{\log w}{H}\right)\\
&= S+\left(\log\log K_1-\log\log\frac{K_1}{2}+O\left(\frac{1}{\log K_1}\right)\right)+O\left(\frac{\log w}{H}\cdot\frac{K_1}{\log K_1}\right)\\
&= S+O\left(\frac{\epsilon\log w+1}{\log K_1}\right),
\end{aligned}\tag{5.11}$$

donde S es como en (5.5). Gracias a $\epsilon\ge \text{máx}(1/\log w, 1/\sqrt{H})$, vemos que $\epsilon\log w+1\le 2\epsilon\log w$ y $\log K_1\le(\log H)/2$. Usamos la expresión (5.8) para S. □

El plan es mostrar que, para algún H, las variables X_H e Y_H son más o menos independientes, y usar este hecho para obtener $\mathbb{E}(F(X_H,Y_H)) = o(H/\log H)$, lo cual es exactamente lo necesario para poder concluír, por el Lema 5.2, que (5.1) se cumple. Para ello aprovecharemos que la función F puede escribirse como una suma de variables aleatorias:

$$F(\vec{x},\vec{y})=\sum_{\frac{\epsilon H}{2}<p\le\epsilon H}F_p(\vec{x},y_p)\tag{5.12}$$

con $\vec{y}=(y_p)_p$ y $F_p(\vec{x},\cdot):\mathbb{Z}/p\mathbb{Z}\to\mathbb{R}$ definida por $F_p(\vec{x},t)=\sum_{j\le H-p,\, j\equiv -t(p)}x_jx_{j+p}$. Como

$$\|F_p\|_\infty\le\frac{H}{p}\le\frac{2}{\epsilon},\tag{5.13}$$

por el teorema de los números primos vemos que

$$\|F\|_\infty\ll\frac{H}{\log H}\tag{5.14}$$

por lo que sólo necesitamos mejorar un poco esa cota para obtener nuestro objetivo para $\mathbb{E}(F(X_H,Y_H))$.

Ahora bien, tenemos el siguiente resultado.

Lema 5.5. *Sean $C,n\ge 1$ y $\Omega=\Omega_1\times\ldots\times\Omega_n$ con Ω_m conjuntos finitos, y G_m una función real sobre Ω_m para $1\le m\le n$, con $\|G_m\|_\infty\le C$. Definamos $G:\Omega\to\mathbb{R}$ por*

$$G(\vec{t})=G(t_1,t_2,\ldots t_n)=G_1(t_1)+G_2(t_2)+\ldots+G_n(t_n),$$

y sea $\overline{G}$ su promedio sobre Ω. Entonces, para cualquier $0<\mu<1$, se cumple que

$$|G(\vec{t})-\overline{G}|\le\mu\cdot Cn$$

para todo $\vec{t}\in\Omega$ excepto para un conjunto de tamaño a lo más $2|\Omega|^{1-\frac{\mu^2/2}{\log R}}$, con $R=|\Omega|^{1/n}$.

Demostración. Si consideramos una variable aleatoria $T=(T_1,\ldots,T_n)$, T_m con distribución uniforme en Ω_m e independientes, el enunciado del lema equivale a decir que

$$\mathbb{P}(|G(T)-\mathbb{E}(G(T))|\ge\mu\cdot Cn)\le 2e^{-\frac{\mu^2 n}{2}}.$$

Pero esto se deduce directamente de la desigualdad de Hoeffding, que nos dice que si $X_1,X_2,\ldots,X_n$ son variables aleatorias independientes tomando valores en el intervalo $[-C,C]$ entonces para $S=X_1+\ldots+X_n$ se cumple que

$$\mathbb{P}(|S-\mathbb{E}(S)|\ge s)\le 2e^{-\frac{s^2}{2C^2n}}.\tag{5.15}$$

Un breve comentario: si bien (5.15) es de forma claramente similar al teorema central del límite, se trata de un resultado de *grandes desviaciones*, pues es válido para s arbitrario, mientras que el teorema central del límite se ocupa del caso en el cual s está acotado por un múltiplo constante de la desviación estándar, es decir, el caso $s \ll C\sqrt{n}$. □

Así, teniendo en cuenta (5.12) y (5.13), podemos aplicar el Lema 5.5 con $F(\vec{x}, \cdot)$ para obtener (usando el teorema de los números primos) que, para H más grande que una constante y $\epsilon \geq 2/\sqrt{H}$,

$$|F(\vec{x}, \vec{y}) - \overline{F(\vec{x}, \cdot)}| \leq \mu \frac{4H}{\log H} \tag{5.16}$$

para todo $\vec{y} \in \Omega$ excepto en un conjunto $E_{\vec{x}} \subset \Omega$ que satisface

$$|E_{\vec{x}}| \leq 2|\Omega|^{1-\frac{\mu^2}{2\log H}}. \tag{5.17}$$

Esta cota, junto con (5.13), nos ayudará a estimar $\mathbb{E}(F(X_H, Y_H))$, ya que sólo tendremos que preocuparnos en mostrar que, la mayor parte del tiempo, Y_H tiende a evitar un conjunto relativamente pequeño de valores E_{X_H}, dado por X_H. Claro está, como Y_H está casi equidistribuída, tal aseveración se deduciría de inmediato si X_H y Y_H fueran variables independientes. Como veremos, bastará probar una forma muy débil de independencia, para algún H.

5.2.1. Ejercicios. Los siguientes ejercicios dan un ejemplo muy básico de como deducir un enunciado sobre los enteros de un enunciado probabilístico general sobre sumas de variables aleatorias.

1. Pruebe la *desigualdad de Chebyshev* para una variable aleatoria X:
$$\mathbb{P}(|X - \mu| \geq \lambda\sigma^2) \leq \frac{1}{\lambda^2}$$
para todo $\lambda \in \mathbb{R}$, con $\mu = \mathbb{E}[X]$ la esperanza de X y $\sigma^2 = \mathbb{E}[(X-\mu)^2]$ su varianza.
2. En este ejercicio queremos demostrar que en $[1, x]$ casi todo entero (es decir, todos, salvo $o(x)$ de ellos) tiene $(1 + o(1)) \log\log x$ divisores primos distintos.
   Podemos asumir que x es entero. Sea N una variable aleatoria en el espacio $\Omega = \{n \leq x\}$ tal que $\mathbb{P}(N = n) = 1/x$ para todo $n \in \Omega$. Para todo primo $p \leq D = x^{1/4}$, considere la variable aleatoria $X_p = 1_{N\equiv 0(p)}$.
   a) Demuestre que X_p es una variable de Bernoulli con media $\mu_p = \frac{1}{p} + O(x^{-1})$ y por lo tanto con varianza $\sigma_p^2 = \mu_p(1-\mu_p) = \frac{1}{p}(1-\frac{1}{p}) + O(x^{-1})$.
   b) Pruebe que para $p \neq q$ las variables X_p y X_q tienen covarianza casi nula: para $Y_p = X_p - \mu_p$ tenemos $\mathbb{E}[Y_p Y_q] \ll x^{-1}$.
   c) Muestre que la variable $w = \sum_{p\leq D} X_p$ tiene media $\mu = \sum_{p\leq D} \mu_p$ y que su varianza $\sigma^2 = \mathbb{E}[(\sum_{p<D} Y_p)^2]$ satisface $\sigma^2 = \sum_{p\leq D} \sigma_p^2 + O(x^{-1/2})$. (En otras palabras, las variables X_p se comportan como si fueran independientes.)
   d) Usando el teorema de los números primos (o, para ser precisos, (2.16)), concluya que $\sigma^2 = \mu + O(1) = \log\log x + O(1)$.
   e) Use la desigualdad de Chebyshev para demostrar el enunciado del ejercicio, observando que un entero en $[1, x]$ tiene a lo sumo 3 divisores primos mayores que D.

5.3. Entropía e información mutua. Ahora comenzamos la labor de mostrar que, para algún H, las variables aleatorias X_H e Y_H no son muy dependientes. Esto lo mediremos mediante el concepto de «información mutua», que pasamos a definir.

5.3.1. Definiciones. Sea X una variable aleatoria con un número finito de valores posibles x. La *entropía* $\mathbb{H}(X)$ de X es

$$\mathbb{H}(X) = -\sum_x p_x \log p_x,$$

donde p_x es la probabilidad $\mathbb{P}(X = x)$ de que X tome el valor x. La *entropía condicional* de X con respecto a una variable aleatoria Y (que suponemos tener también un número finito de valores, o por lo menos ser discreta) es

$$\mathbb{H}(X|Y) = \sum_y \mathbb{H}(X|Y = y)\mathbb{P}(Y = y), \tag{5.18}$$

donde, para E un evento probabilístico (como $Y = y$), $\mathbb{H}(X|E)$ se define por $\mathbb{H}(X|E) = -\sum_x p_{x,E} \log p_{x,E}$, donde $p_{x,E} = \mathbb{P}(X = x|E)$.

Las siguientes propiedades básicas son fáciles de probar (ver los ejercicios). Escribimos $\mathbb{H}(X, Y)$ para denotar la entropía $\mathbb{H}((X, Y))$ de la variable aleatoria (X, Y), donde X e Y son variables aleatorias.

1. La entropía $\mathbb{H}(X)$ y la entropía condicional $\mathbb{H}(X|Y)$ son no negativas.
2. $\mathbb{H}(X, Y) = \mathbb{H}(X|Y) + \mathbb{H}(Y) = \mathbb{H}(Y|X) + \mathbb{H}(X)$,
3. $\mathbb{H}(X|Y) \leq \mathbb{H}(X)$,
4. $\mathbb{H}(X, Y) \leq \mathbb{H}(X) + \mathbb{H}(Y)$ (subaditividad de la entropía).
5. Si X toma $\leq N$ valores distintos, $\mathbb{H}(X) \leq \log N$.

También podemos acotar con facilidad la diferencia entre las entropías de dos variables aleatorias X, Y que toman los mismos valores con probabilidades distintas (ejercicio 2).

Definimos la *información mutua* $\mathbb{I}(X, Y)$:

$$\mathbb{I}(X, Y) = \mathbb{H}(X) + \mathbb{H}(Y) - \mathbb{H}(X, Y). \tag{5.19}$$

Por la subaditividad de la entropía, $\mathbb{I}(X, Y) \geq 0$. Está claro por 2. que

$$\mathbb{H}(X|Y) = \mathbb{H}(X) - \mathbb{I}(X, Y), \qquad \mathbb{H}(Y|X) = \mathbb{H}(Y) - \mathbb{I}(X, Y). \tag{5.20}$$

5.3.2. El argumento por agotamiento de información mutua. Consideramos ahora las variables aleatorias X_H, Y_H definidas en (5.7) en términos de la variable aleatoria N cuya distribución fue dada en (5.6). La meta es mostrar que existe un $H \in [H_-, H_+]$ (donde H_-, H_+ serán especificados más tarde) tal que $\mathbb{I}(X_H, Y_H)$ es pequeña en comparación con H.

Podemos definir para H_1, H_2 arbitrarios,

$$X_{H_1,H_1+H_2} = (\lambda(N + j))_{H_1<j\leq H_1+H_2}. \tag{5.21}$$

Podemos asumir sin pérdida de generalidad que $w \leq x^{1/8}$. También asumiremos que $H_1, H_2 \leq x^{1/8}$. Entonces, por el ejercicio $3c$,

$$\mathbb{H}(X_{H_1,H_1+H_2}) = \mathbb{H}(X_{H_2}) + O\left(1/x^{5/8}\right).$$

Por la subaditividad de la entropía, deducimos que

$$\mathbb{H}(X_{H_1+H_2}) \leq \mathbb{H}(X_{H_1}) + \mathbb{H}(X_{H_1,H_1+H_2}) \leq \mathbb{H}(X_{H_1}) + \mathbb{H}(X_{H_2}) + O(1/x^{5/8}). \tag{5.22}$$

Vemos por el ejercicio $3d$ (suponiendo que $\delta \asymp (\prod_{K_0<p\leq K_1} p)^{-1}$ satisface $\delta^{-1} \leq x^{1/8}$) que el mismo razonamiento vale para la entropía condicional:

$$\mathbb{H}(X_{H_1,H_1+H_2}|(N + H_1 \bmod p)_{K_0<p\leq K_1}) = \mathbb{H}(X_{H_2}|(N \bmod p)_{K_0<p\leq K_1}) + O(1/\sqrt{x}).$$

Por el teorema de los números primos (en la forma [RS62, Thm. 9]),

$$\prod_{K_0<p\leq K_1} p \leq e^{\sum_{p\leq K_1} \log p} \leq e^{1{,}02 K_1}. \tag{5.23}$$

Supondremos entonces que $K_1 \leq (\log x)/9$.

Ahora bien, $N + H_1 \bmod p$ codifica la misma información que $N \bmod p$. Por lo tanto,

$$\mathbb{H}(X_{H_1,H_1+H_2}|(N \bmod p)_{K_0<p\leq K_1}) = \mathbb{H}(X_{H_1,H_1+H_2}|(N + H_1 \bmod p)_{K_0<p\leq K_1})$$

Recordamos la definición (5.7) de Y_H. De nuevo por subaditividad, como en (5.22), concluímos que

$$\mathbb{H}(X_{H_1+H_2}|Y_H) \leq \mathbb{H}(X_{H_1}|Y_H) + \mathbb{H}(X_{H_2}|Y_H) + O(1/\sqrt{x}).$$

Iterando con $H_1 = H, 2H, 3H, \ldots$ y $H_2 = H$, vemos que

$$\mathbb{H}(X_{kH}|Y_H) \leq k\mathbb{H}(X_H|Y_H) + O(k/\sqrt{x})$$

para $kH \leq x^{1/8}$, y en consecuencia, otra vez por subaditividad,

$$\mathbb{H}(X_{kH}) \leq \mathbb{H}(X_{kH}|Y_H) + \mathbb{H}(Y_H) \leq k\mathbb{H}(X_H|Y_H) + \mathbb{H}(Y_H) + O(k/\sqrt{x}),$$

así que, por (5.20),

$$\frac{\mathbb{H}(X_{kH})}{kH} \leq \frac{\mathbb{H}(X_H)}{H} - \frac{\mathbb{I}(X_H, Y_H)}{H} + \frac{\mathbb{H}(Y_H)}{kH} + O\left(\frac{1}{\sqrt{x}}\right).$$

Por la propiedad 5 de la entropía y (5.23),

$$\mathbb{H}(Y_H) \leq \log\left(\prod_{K_0<p\leq K_1} p\right) \leq 1{,}02\epsilon H.$$

Concluímos que,

$$\frac{\mathbb{H}(X_{kH})}{kH} \leq \frac{\mathbb{H}(X_H)}{H} - \frac{\mathbb{I}(X_H, Y_H)}{H} + 1{,}02\frac{\epsilon}{k} + O\left(\frac{1}{\sqrt{x}}\right). \tag{5.24}$$

He aquí el sujeto de nuestra metáfora: la tasa de entropía $\mathbb{H}(X_{H'})/H'$ (con $H' = H, kH, \ldots$) es "las lentejas", y la tasa $\mathbb{I}(X_H, Y_H)/H$ es la cucharada que nos servimos. Para cuantificar que tan poca lenteja nos terminaremos sirviendo en algún momento, basta con un sencillo lema.

Lema 5.6. *Sea $h_1 \geq 15$ arbitrario y, para $j \geq 1$, $h_{j+1} = \lfloor 4\log h_j \log_3 h_j \rfloor \cdot h_j$. Entonces*

$$\sum_{j=1}^{J} \frac{1}{\log h_j \log_3 h_j} \geq 100$$

para algún J cumpliendo $\log J \ll (\log_2 h_1)^2$.

Demostración. Ejercicio 4. □

Corolario 5.7. *Sean X_H, Y_H y N como en (5.6) y (5.7), con $K_0 = \epsilon H/2$, $K_1 = \epsilon H$, $0 < \epsilon \leq 1$ y $w \leq x^{1/8}$. Sea $H_- \geq 3$. Entonces hay un $H_+ > H_-$, dependiendo sólo de H_- y cumpliendo $\log_3 H_+ \leq 2\log_3 H_- + O(1)$, tal que*

$$\mathbb{I}(X_H, Y_H) \leq \frac{H}{\log H \log_3 H} \tag{5.25}$$

para algún entero $H \in [H_-, H_+]$, con tal que $x \geq \exp(H_+^9)$.

Es fácil ver que la condición $x \geq \exp(H_+^9)$ se deduce de $\log_3 H_+ \leq 2\log_3 H_- + O(1)$ para x más grande que una constante y $H_- \leq \exp(\exp(\sqrt{\log_3 x}/C))$, donde $C > 0$ es otra constante.

Demostración. Podemos asumir sin pérdida de generalidad que H_- es un entero más grande que una constante apropiada. Sea $h_1 = H_-$; definamos $h_2, h_3, \ldots$ como en el Lema 5.6, y sea $k_j = \lfloor 4 \log h_j \log_3 h_j \rfloor$. Utilizando la suposición que $h_j \geq H_-$ es más grande que una constante, así como la desigualdad $k_j \leq k_j h_j \leq x^{1/8}$, la cual debemos asumir de todas maneras, simplificamos (5.24), obteniendo

$$\frac{\mathbb{H}(X_{h_{j+1}})}{h_{j+1}} \leq \frac{\mathbb{H}(X_{h_j})}{h_j} - \frac{\mathbb{I}(X_{h_j}, Y_{h_j})}{h_j} + \frac{1}{2 \log h_j \log_3 h_j}. \tag{5.26}$$

Deducimos entonces de las propiedades (1) y (5) de la entropía que

$$\sum_{j \leq J} \left(\frac{\mathbb{I}(X_{h_j}, Y_{h_j})}{h_j} - \frac{1}{2 \log h_j \log_3 h_j} \right) \leq \frac{\mathbb{H}(X_{H_-})}{H_-} \leq \log 2.$$

Por el Lema 5.6, vemos que hay algún J dependiendo sólo de $H_- = h_1$, tal que

$$\frac{\mathbb{I}(X_{h_j}, Y_{h_j})}{h_j} - \frac{1}{2 \log h_j \log_3 h_j} \leq \frac{1}{2 \log h_j \log_3 h_j}$$

para algún $1 \leq j \leq J$. (¿Por qué?) Concluímos que

$$\frac{\mathbb{I}(X_{h_j}, Y_{h_j})}{h_j} \leq \frac{1}{\log h_j \log_3 h_j}.$$

Definimos $H_+ = h_j$ y obtenemos el resultado. La condición $H_+ \leq (\log x)/9$ implica $H_+ \leq x^{1/8}$, y así también la suposición $k_j h_j = h_{j+1} \leq x^{1/8}$ para todo $j \leq J - 1$. □

5.3.3. Ejercicios.

1. Sean X e Y variables aleatorias que toman un número finito de valores. Pruebe las propiedades (1)–(5) de la entropía. Sugerencias para cada propiedad:
   (1) Use simplemente las definiciones de entropía y entropía condicional.
   (2) De nuevo por las definiciones.
   (3) Por la concavidad de $x \mapsto -x \log x$ (primera desigualdad) y por el hecho que $\log x$ es creciente (segunda desigualdad; también se deduce inmediatamente de las propiedades (1) y (2)).
   (4) Use las propiedades (2) y (3).
   (5) Por la concavidad de $x \mapsto \log x$.
2. Sean X e Y variables aleatorias con valores en un conjunto S de N elementos. Denotemos por ν_X, ν_Y sus funciones de distribución. La *distancia de variación total* $|\nu_X - \nu_Y|_{TV}$ se define como $\max_{S' \subset S} |\nu_X(S') - \nu_Y(S')|$. (Es fácil ver que es igual a $\frac{1}{2}|\nu_X - \nu_Y|_1$.)
   a) Muestre que $p = \sum_{x \in S} \min(\nu_X(x), \nu_Y(x))$ es igual a $1 - |\nu_X - \nu_Y|_{TV}$. Si $p = 1$, entonces X e Y tienen la misma distribución, y estamos en el caso trivial. Si $p = 0$, X e Y tienen soportes disjuntos, y la cota que probaremos al final es muy sencilla (muéstrelo llegado el momento). Asumamos de ahora en adelante que $0 < p < 1$.
   b) Sea Z una variable aleatoria con función de distribución
   $$\nu_Z(x) = \frac{1}{p} \min(\nu_X(x), \nu_Y(x)).$$
   Sean X' e Y' variables aleatorias independientes con distribuciones
   $$\nu_{X'}(x) = \begin{cases} \frac{\nu_X(x) - \nu_Y(x)}{1-p} & \text{si } \nu_X(x) > \nu_Y(x), \\ 0 & \text{de otra manera,} \end{cases} \qquad \nu_{Y'}(x) = \begin{cases} \frac{\nu_Y(x) - \nu_X(x)}{1-p} & \text{si } \nu_Y(x) > \nu_X(x), \\ 0 & \text{de otra manera.} \end{cases}$$
   Sea B una variable que toma el valor 0 con probabilidad p y el valor 1 con probabilidad $1 - p$. Construyamos la variable aleatoria C de la siguiente forma:

si $B = 0$, C toma el valor (Z, Z); si $B = 1$, C toma el valor (X', Y'). Muestre que la primera coordenada $C_1 = \pi_1(C)$ de C es una variable con distribución ν_X, mientras que la segunda coordenada $C_2 = \pi_2(C)$ tiene distribución ν_Y. Muestre también que $\mathbb{P}(C_1 \neq C_2) = |\nu_X - \nu_Y|_{TV}$. Se dice que la variable C es un *acoplamiento óptimo* de X e Y.

c) Por las propiedades (3) de la entropía,

$$\mathbb{H}(C_1|B) \leq \mathbb{H}(C_1) \leq \mathbb{H}(C_1|B) + \mathbb{H}(B).$$

Sabemos que

$$\mathbb{H}(C_1|B) = p\mathbb{H}(C_1|B = 0) + (1-p)\mathbb{H}(C_1|B = 1) = p\mathbb{H}(Z) + (1-p)\mathbb{H}(X').$$

De la misma manera,

$$\mathbb{H}(C_2|B) = p\mathbb{H}(Z) + (1-p)\mathbb{H}(Y').$$

Por lo tanto,

$$\begin{aligned}|\mathbb{H}(X) - \mathbb{H}(Y)| &= |\mathbb{H}(C_1) - \mathbb{H}(C_2)| \\ &\leq \mathbb{H}(B) + |\mathbb{H}(C_1|B) - \mathbb{H}(C_2|B)| \\ &\leq \mathbb{H}(B) + (1-p)\,|\mathbb{H}(X') - \mathbb{H}(Y')| \\ &\leq \mathbb{H}(B) + (1-p)\,\mathrm{mín}(\mathbb{H}(X'), \mathbb{H}(Y')).\end{aligned}$$

d) Concluya, por la propiedad (5) de la entropía, que

$$|\mathbb{H}(X) - \mathbb{H}(Y)| \leq |\nu_X - \nu_Y|_{TV} \cdot \log N + \mathbb{H}(B).$$

En verdad, podemos dar una cota ligeramente menor: la variable X' tiene soporte en un subconjunto estricto de S (¿por qué?) y lo mismo es cierto de Y'; concluya que

$$|\mathbb{H}(X) - \mathbb{H}(Y)| \leq |\nu_X - \nu_Y|_{TV} \cdot \log(N-1) + \mathbb{H}(B). \tag{5.27}$$

Aquí, por supuesto,

$$\mathbb{H}(B) = -|\nu_X - \nu_Y|_{TV} \log |\nu_X - \nu_Y|_{TV} - (1 - |\nu_X - \nu_Y|_{TV}) \log(1 - |\nu_X - \nu_Y|_{TV}).$$

3. Sea $I = (x_0, x_1]$ un intervalo en $\mathbb{R}^+$. Sea N una variable aleatoria que toma el valor entero $n \in I$ con probabilidad $(1/n)/L$, donde $L = \sum_{m\in I} 1/m$. Para $h \in \mathbb{Z}^+$, sea $N + h$ la variable que toma el valor $N + h$ con probabilidad $(1/n)/L$. Denote ν_X la distribución de probabilidad de una variable X.

a) Muestre que $|\nu_{N+h} - \nu_N|_{TV} \leq h/x_0L$.

b) Sea f una función sobre $\mathbb{Z}^+$. Deduzca que

$$\left|\nu_{f(N+h)} - \nu_{f(N)}\right|_{TV} \leq h/x_0L. \tag{5.28}$$

c) Sean X_H y X_{H_1,H_1+H_2} como en (5.7) y (5.21). Concluya, usando (5.27) y (5.28), que, para $L = \sum_{x/w<n\leq x} 1/n \geq 1$ y $1 \leq H_1 \leq x_0/2$, $x_0 = x/w$,

$$\begin{aligned}|\mathbb{H}(X_{H_1,H_1+H_2}) - \mathbb{H}(X_{H_2})| &\leq \frac{H_1}{x_0L} \log\left(2^{2H_2} - 1\right) + h\left(\frac{H_1}{x_0L}\right) \\ &\leq \frac{H_1}{x_0L} H_2 \log 4 + h\left(\frac{H_1}{x_0}\right) \leq \frac{H_1}{x_0}\left(H_2 \log 4 + 2\log x_0\right),\end{aligned}$$

donde $h(\epsilon) = -\epsilon \log \epsilon - (1-\epsilon)\log(1-\epsilon) \leq -2\epsilon \log \epsilon$ para $0 < \epsilon \leq 1/2$. En particular, para $1 \leq H_1, H_2 \leq x_0^\alpha$, $\alpha > 0$,

$$|\mathbb{H}(X_{H_1,H_1+H_2}) - \mathbb{H}(X_{H_2})| \ll_\alpha \frac{1}{x_0^{1-2\alpha}}.$$

d) Sea $S \subset (x_0, x_1]$ y $\delta = \mathbb{P}(N \in S) = (\sum_{n\in S} 1/n)/L$. Muestre que la distancia de variación total entre las distribuciones de probabilidad condicional de $N + h$ y N con la condición $N \in S$ es $\leq h/\delta x_0 L$. Concluya que, para $1 \leq H_1 \leq \delta x_0/2$,

$$|\mathbb{H}(X_{H_1,H_1+H_2}|N + H_1 \in S) - \mathbb{H}(X_{H_2}|N \in S)| \leq \frac{H_1}{\delta x_0} \left(H_2 \log 4 + 2 \log x_0\right),$$

con $x_0 = x/w$ y $x_1 = x$. En particular, si $1 \leq H_1, H_2, \delta^{-1} \leq x_0^\alpha$, $0 < \alpha < 1$,

$$|\mathbb{H}(X_{H_1,H_1+H_2}|N + H_1 \in S) - \mathbb{H}(X_{H_2}|N \in S)| \ll_\alpha \frac{1}{x_0^{1-3\alpha}}.$$

4. Sea h_j como en el Lema 5.6.
   a) Muestre que $\log h_j \leq 2(j + \log h_1)\log(j + \log h_1)$ para todo $j \geq 1$.
   b) Pruebe que, para $j \geq \log h_1$,

$$\frac{1}{\log h_j \log_3 h_j} \geq \frac{1/8}{j \log j \log_3 j}.$$

   c) Demuestre que

$$\sum_{\log h_1 < j < (\log h_1)^{C \log_2 h_1}} \frac{1}{j \log j \log_3 j} > 800$$

   para C suficientemente grande, independiente de h_1. (Utilice, por ejemplo, un test de integrales.) Concluya que el Lema 5.6 es cierto.

5.4. Información mutua y dependencia. En la sección anterior hemos visto que la información mutua entre X_H e Y_H es pequeña para algún H. En esta sección vamos a ver cómo usar ese hecho para deducir que X_H e Y_H son más o menos independientes. Todo va a sustentarse sobre nuestra cota para la información mutua.

El esquema sería el siguiente. La variable Y_H para H pequeño es esencialmente uniforme. Por lo tanto, su entropía está muy cerca del máximo posible, y, como la información mutua con X_H es pequeña, deducimos que la entropía $\mathbb{H}(Y_H|X_H)$ también va a estar muy cerca del máximo posible. Esto va a implicar que, con probabilidad casi 1, X_H toma un valor $\vec{x}$ para el cual la entropía de la variable Y_H condicionada a $X_H = \vec{x}$ está cerca del máximo posible. A su vez, eso debería implicar que la distribución de la variable Y_H condicionada a $X_H = \vec{x}$ está cerca de ser uniforme, por lo que habríamos demostrado esencialmente la independencia de X_H e Y_H.

En realidad, este esquema sólo demostrará la independencia “a efectos de esperanza”, es decir,

$$\mathbb{E}(F(X_H, Y_H)) \sim \mathbb{E}(F(X_H, Y_H^*))$$

donde Y_H^* es una variable con la misma distribución que Y_H pero independiente de X_H. Empero, esta igualdad aproximada entre esperanzas es justo lo que necesitamos para acotar nuestras sumas.

Como ya hemos indicado (final de §5.2), la equidistribución de Y_H condicionada a $X_H = \vec{x}$ que requerimos es bastante débil: sólo necesitamos mostrar que Y_H tiende a evitar un pequeño conjunto, de tamaño acotado por la desigualdad (5.17). Debido a la forma de (5.17), el hecho que nuestra cota para la información (5.25) es mejor que la cota trivial por un factor de algo más que $\log H$ será crucial.

El primer paso es muy simple: ver que Y_H es esencialmente uniforme. Sabemos por (5.7) que Y_H toma valores en el conjunto

$$\Omega = \prod_{\epsilon H/2 < p \leq \epsilon H} \mathbb{Z}/p\mathbb{Z}. \tag{5.29}$$

Por el teorema chino de los restos, Ω es isomorfo a $\mathbb{Z}/M\mathbb{Z}$, donde $M = |\Omega| = \prod_{\epsilon H/2<p\leq\epsilon H} p$. Por lo tanto, si $|\Omega| < x/w$ entonces

$$\mathbb{P}(Y_H = \vec{y}) = \mathbb{P}\left(N \equiv y_p \bmod p \ \ \forall p \in \left(\frac{\epsilon H}{2}, \epsilon H\right]\right) = \frac{1}{L} \sum_{\substack{\frac{x}{w}<n\leq x \\ n\equiv y_p(p)\forall p\in(\frac{\epsilon H}{2},\epsilon H]}} \frac{1}{n}$$

$$= \frac{1}{L} \sum_{\substack{\frac{x}{w}<n\leq x \\ n\equiv y_*(|\Omega|)}} \frac{1}{n} = \frac{1}{|\Omega|} \frac{\log w + O(\frac{1}{x/w|\Omega|})}{\log w + O(\frac{1}{x/w})} = \frac{1}{|\Omega|} + O\left(\frac{1}{x/w}\right),$$

donde $L = \sum_{x/w<n\leq n} 1/n$ e y_* es cualquier entero congruente a y_p para cada $p \in (\epsilon H/2, \epsilon H]$. Por otra parte, por el teorema de los números primos,

$$|\Omega| = e^{\frac{\epsilon H}{2}} \left(1 + O\left((\log \epsilon H)^{-50}\right)\right) \ll e^{\frac{\epsilon H}{2}}. \tag{5.30}$$

Luego, tomando $\epsilon H < \log(x/w)$ vemos que

$$\mathbb{P}(Y_H = \vec{y}) = \frac{1}{|\Omega|}\left(1 + O\left(\frac{1}{|\Omega|}\right)\right), \tag{5.31}$$

es decir, la distribución de Y_H es casi uniforme.

Deducimos mediante la cota general (5.27) que

$$\mathbb{H}(Y_H) = \left(1 + O\left(\frac{1}{|\Omega|}\right)\right) \log |\Omega|, \tag{5.32}$$

(El máximo posible sería $\log |\Omega|$.)

Apliquemos ahora el Corolario 5.7, con H_- de forma que $H_+ \leq \frac{1}{9}\log x$. Obtenemos que la información mutua con X_H es pequeña para cierto $H \in [H_-, H_+]$, y, por (5.20), concluímos que

$$\mathbb{H}(Y_H|X_H) \geq \mathbb{H}(Y_H) - \frac{H}{\log H \log_3 H}$$

para dicho H. Así, usando (5.32) y la cota (5.30), vemos que

$$\mathbb{H}(Y_H|X_H) \geq \left(1 - \frac{4}{\epsilon \log H \log_3 H}\right) \log |\Omega| \tag{5.33}$$

para H_- más grande que una constante que depende sólo de ϵ. En otras palabras, para algún $H \in [H_-, H_+]$, la entropía condicional $\mathbb{H}(Y_H|X_H)$ está cerca del máximo posible.

Digamos que un elemento $\vec{x} \in \{-1, 1\}^H$ es *bueno* si

$$\mathbb{H}(Y_H|X_H = \vec{x}) \geq \left(1 - \frac{4}{(\epsilon \log_3 H)^{3/4} \log H}\right) \log |\Omega|, \tag{5.34}$$

y en otro caso decimos que $\vec{x}$ es *malo*. Por la definición (5.18) de $\mathbb{H}(Y_H|X_H)$ en términos de $\mathbb{H}(Y_H|X_H = \vec{x})$, y por la desigualdad (5.33), vemos que

$$\left(1 - \frac{4}{\epsilon \log_3 H \log H}\right) \log |\Omega| \leq \sum_{\vec{x} \text{ malo}} \left(1 - \frac{4}{(\epsilon \log_3 H)^{3/4} \log H}\right) \log |\Omega| \cdot \mathbb{P}(X = \vec{x})$$
$$+ \sum_{\vec{x} \text{ bueno}} \log |\Omega| \cdot \mathbb{P}(X = \vec{x})$$

ya que $\mathbb{H}(Y_H|X_H = \vec{x})$ es como máximo $\log |\Omega|$. De esta desigualdad deducimos la cota

$$\mathbb{P}(X_H \text{ malo}) \leq \frac{1}{(\epsilon \log_3 H)^{1/4}}. \tag{5.35}$$

Luego, es muy probable que X_H sea igual a un valor bueno de $\vec{x}$, i.e., un valor de $\vec{x}$ tal que la entropía de la variable Y_H condicionada a $X_H = \vec{x}$ estará cerca del máximo posible.

Para continuar, querríamos ver que la única distribución que está cerca de alcanzar el máximo de entropía es la uniforme. Esto sería cierto si su entropía está extremadamente cerca de dicho máximo, pero cuando hay un poco más de diferencia dicha unicidad no es cierta (ejercicio 1). Aun así, se puede decir que la variable no concentra mucho su masa, en el siguiente sentido.

Lema 5.8. *Sea Y una variable aleatoria tomando valores es un conjunto Ω finito tal que $\mathbb{H}(Y) \geq (1-\delta)\log|\Omega|$, con $\frac{1}{\log|\Omega|} \leq \delta < 1$. Si un subconjunto $E \subset \Omega$ tiene tamaño*

$$|E| \leq |\Omega|^{1-M\delta} \tag{5.36}$$

para algún $M > 0$, entonces

$$\mathbb{P}(Y \in E) \leq \frac{2}{M}.$$

Demostración. Por la concavidad de la función $x \to -x\log x$, el ejercicio 2 nos da

$$\mathbb{H}(Y) \leq \mathbb{P}(E)\log\frac{|E|}{\mathbb{P}(E)} + \mathbb{P}(E^c)\log\frac{|E^c|}{\mathbb{P}(E^c)} \tag{5.37}$$

donde $\mathbb{P}(A) = \mathbb{P}(Y \in A)$. Escribiendo $p = \mathbb{P}(E)$ y usando $\mathbb{H}(Y) \geq (1-\delta)\log|\Omega|$, vemos que

$$(1-\delta)\log|\Omega| \leq p\log|E| + (1-p)\log|\Omega| - p\log p - (1-p)\log(1-p).$$

Ahora, usando de nuevo la concavidad con los dos últimos sumandos, así como la cota (5.36) sobre el tamaño de E, llegamos a

$$(1-\delta)\log|\Omega| \leq p(1-M\delta)\log|\Omega| + (1-p)\log|\Omega| + \log 2,$$

lo cual da

$$\delta(pM-1)\log|\Omega| \leq \log 2.$$

Luego, si se cumpliera $p > 2/M$, tendríamos $\delta \leq \log 2/\log|\Omega|$, en contradicción a una de las hipótesis del lema. □

El lema anterior indica que, para un $\vec{x}$ bueno, la variable Y_H condicionada a $X_H = \vec{x}$ tiene una distribución que no concentra mucho su masa. Ésta es una forma muy débil de cuasiuniformidad, pero va a ser suficiente para estimar la esperanza de $F(X_H, Y_H)$, debido a (5.16) y (5.17).

Teorema 5.9. *Existe un $c > 0$ tal que lo siguiente se cumple. Sean $x > e^e$, $1 \leq w \leq x^{1/8}$ y $0 < \epsilon \leq 1$. Sea F como en (5.9). Para todo $3 \leq H_- \leq \exp(\exp(c\sqrt{\log_3 x}))$, hay un $H > H_-$, dependiendo sólo de H_- y cumpliendo $\log_3 H \leq 2\log_3 H_- + O(1)$, tal que, para X_H, Y_H, N y Ω definidos como en (5.6), (5.7) y (5.29) con $K_0 = \epsilon H/2$ y $K_1 = \epsilon H$,*

$$\mathbb{E}(F(X_H, Y_H)) = \mathbb{E}(F(X_H, Y_H^*)) + O\left(\frac{H/\log H}{(\epsilon\log_3 H)^{1/4}}\right)$$

para Y_H^* una variable aleatoria con distribución uniforme en Ω e independiente de X_H.

Demostración. Podemos asumir que x y $\epsilon\log_3 H_-$ son más grandes que una constante, ya que si no el resultado es trivial. Aplicamos el Corolario 5.7 para obtener un H_+ con $\log_3 H_+ \leq 2\log_3 H_- + O(1)$ y cierto $H \in [H_-, H_+]$ tal que (5.25) se cumple, i.e., tal que la información mutua entre X_H e Y_H es pequeña. Por la definición de esperanza,

$$\mathbb{E}(F(X_H, Y_H)) = \sum_{\substack{\vec{x}\in\{-1,1\}^H \\ \vec{y}\in\Omega}} F(\vec{x}, \vec{y})\,\mathbb{P}(X_H = \vec{x}, Y_H = \vec{y})$$

$$= \sum_{\vec{x}\in\{-1,1\}^H} \mathbb{P}(X_H = \vec{x}) \sum_{\vec{y}\in\Omega} F(\vec{x},\vec{y})\, \mathbb{P}(Y_H = \vec{y}|X_H = \vec{x}).$$

Ahora dividimos el rango de $\vec{x}$ entre buenos y malos, acorde a la definición previa a (5.34). Acotamos la suma sobre los $\vec{x}$ malos fácilmente:

$$\sum_{\vec{x}\text{ malo}} \mathbb{P}(X_H = \vec{x}) \sum_{\vec{y}\in\Omega} F(\vec{x},\vec{y})\, \mathbb{P}(Y_H = \vec{y}|X_H = \vec{x}) \ll \|F\|_\infty \mathbb{P}(X_H \text{ malo}) \ll \frac{H/\log H}{(\epsilon \log_3 H)^{1/4}}$$

por (5.14) y (5.35). Ahora, para cada $\vec{x}$ bueno, dividimos el rango de $\vec{y}$ en $E_{\vec{x}}$ y $E_{\vec{x}}^c$, con $E_{\vec{x}}$ el conjunto de excepciones en (5.17) con $\mu^2 = (\epsilon \log_3 H)^{-1/2}$. Teniendo en cuenta (5.34), aplicamos el Lema 5.8 para obtener

$$\mathbb{P}(Y \in E_{\vec{x}}|X_H = \vec{x}) \ll \frac{1}{(\epsilon \log_3 H)^{1/4}},$$

de donde

$$\sum_{\vec{x}\text{ bueno}} \mathbb{P}(X_H = \vec{x}) \sum_{\vec{y}\in E_{\vec{x}}} F(\vec{x},\vec{y})\, \mathbb{P}(Y_H = \vec{y}|X_H = \vec{x}) \ll \frac{\|F\|_\infty}{(\epsilon \log_3 H)^{1/4}} \ll \frac{H/\log H}{(\epsilon \log_3 H)^{1/4}}.$$

Así, sólo nos quedan los $\vec{y}$ no excepcionales y los $\vec{x}$ buenos. En este caso usamos la estimación (5.16), y teniendo en cuenta las cotas ya obtenidas vemos que

$$\mathbb{E}(F(X_H, Y_H)) = O\left(\frac{H/\log H}{(\epsilon \log_3 H)^{1/4}}\right) + \sum_{\vec{x}\text{ bueno}} \mathbb{P}(X_H = \vec{x}) \sum_{\vec{y}\notin E_{\vec{x}}} \overline{F(\vec{x},\cdot)}\, \mathbb{P}(Y_H = \vec{y}|X_H = \vec{x}).$$

Ahora usamos de nuevo las cotas para $\mathbb{P}(Y \in E_{\vec{x}}|X_H = \vec{x})$, $\mathbb{P}(X_H \text{ malo})$ y $\|F\|_\infty$ para suplir los $(\vec{x},\vec{y})$ que faltan y así obtener

$$\mathbb{E}(F(X_H, Y_H)) = O\left(\frac{H/\log H}{(\epsilon \log_3 H)^{1/4}}\right) + \sum_{\vec{x}\in\{-1,1\}^H} \mathbb{P}(X_H = \vec{x}) \sum_{\vec{y}\in\Omega} \overline{F(\vec{x},\cdot)}\, \mathbb{P}(Y_H = \vec{y}|X_H = \vec{x}).$$

Pero

$$\sum_{\vec{x}\in\{-1,1\}^H} \mathbb{P}(X_H = \vec{x}) \sum_{\vec{y}\in\Omega} \overline{F(\vec{x},\cdot)}\, \mathbb{P}(Y_H = \vec{y}|X_H = \vec{x}) = \sum_{\vec{x}\in\{-1,1\}^H} \mathbb{P}(X_H = \vec{x})\overline{F(\vec{x},\cdot)}$$

$$= \sum_{\vec{x}\in\{-1,1\}^H} \mathbb{P}(X_H = \vec{x}) \sum_{\vec{y}\in\Omega} F(\vec{x},\vec{y}) \frac{1}{|\Omega|} = \mathbb{E}(F(X_H, Y_H^*)).$$

□

Ahora, por (5.12), tenemos que la esperanza del Teorema 5.9 se puede escribir como

$$\mathbb{E}(F(X_H, Y_H^*)) = \sum_{\epsilon H/2 < p \le \epsilon H} \mathbb{E}(F_p(X_H, (Y_H^*)_p))$$

donde $Y_H^* = ((Y_H^*)_p)_{\epsilon H/2<p\le\epsilon H}$. Como $(Y_H^*)_p$ es independiente de X_H y uniforme en $\mathbb{Z}/p\mathbb{Z}$ tenemos

$$\mathbb{E}(F_p(X_H, (Y_H^*)_p)) = \sum_{t\in\mathbb{Z}/p\mathbb{Z}} \mathbb{E}(F_p(X_H, t))\mathbb{P}((Y_H^*)_p = t) = \frac{1}{p} \sum_{t\in\mathbb{Z}/p\mathbb{Z}} \mathbb{E}(F_p(X_H, t))$$

$$= \frac{1}{p} \sum_{t\in\mathbb{Z}/p\mathbb{Z}} \sum_{\vec{x}\in\{-1,1\}^H} F_p(\vec{x}, t)\mathbb{P}(X_H = \vec{x}) = \frac{1}{p} \sum_{\vec{x}\in\{-1,1\}^H} \mathbb{P}(X_H = \vec{x}) \sum_{j\le H-p} x_j x_{j+p} \sum_{t\in\mathbb{Z}/p\mathbb{Z}} 1_{j\equiv -t(p)}.$$

Luego

$$\mathbb{E}(F(X_H, Y_H^*)) = \mathbb{E}(G(X_H))$$

con

$$G(\vec{x}) = \sum_{\epsilon H/2 < p \leq \epsilon H} \frac{1}{p} \sum_{j \leq H-p} x_j x_{j+p}$$

y por la definición de X_H

$$\mathbb{E}(G(X_H)) = \sum_{\frac{x}{w} < n \leq x} \frac{1}{nL} \sum_{\epsilon H/2 < p \leq \epsilon H} \frac{1}{p} \sum_{j \leq H-p} \lambda(n+j)\lambda(n+j+p).$$

Ahora, procediendo como en la prueba del lema 5.3, podemos reescribir esta suma como

$$\frac{1}{L} \sum_{\epsilon H/2 < p \leq \epsilon H} \frac{1}{p} \left(H \sum_{x/w < n \leq x} \frac{\lambda(n)\lambda(n+p)}{n} + O\left(H + p \log w\right) \right)$$

por lo que finalmente, por el teorema 5.9 (con $H_- = \exp(\exp(\text{mín}(\log_3 x, \log_2 w)^{1/3}))$, digamos) y el lema 5.4, obtenemos el resultado siguiente, el cual es lo que queríamos.

Corolario 5.10. *Sea $w \leq x^{1/8}$, w más grande que una constante. Sea $(\log_3 w)^{-1} \leq \epsilon \leq 1$. Entonces existe H con $\log_3 H \gg \text{mín}(\log_4 x, \log_3 w)$ y $\log_3 H \leq (3/4)\log_3 w$ tal que*

$$\begin{aligned} \sum_{\epsilon H/2 < p \leq \epsilon H} \sum_{\substack{x/w < n \leq x \\ p \mid n}} \frac{\lambda(n)\lambda(n+p)}{n} - \sum_{\epsilon H/2 < p \leq \epsilon H} \frac{1}{p} \sum_{x/w < n \leq x} \frac{\lambda(n)\lambda(n+p)}{n} \\ = O\left(\epsilon + \frac{1}{(\epsilon \log_3 H)^{1/4}}\right) \cdot \frac{\log w}{\log H}. \end{aligned} \tag{5.38}$$

La cota superior sobre $\log_3 H$ (la cual, por cierto, es de lejos más fuerte de lo que necesitamos) nos será útil cuando tengamos que aplicar el Lema 5.2 y el Corolario 5.15.

5.4.1. Ejercicios.

1. Sea Ω un conjunto finito y $E \subset \Omega$ con $|E| = |\Omega|^{1-\delta} \leq |\Omega|/2$. Sea Y la variable aleatoria en Ω que satisface $\mathbb{P}(Y \in E) = \mathbb{P}(Y \in E^c) = 1/2$ y tal que Y es uniforme en E y también en E^c. Observe que Y está muy lejos de ser uniforme en Ω pero que sin embargo su entropía es grande:
$$\mathbb{H}(Y) = (1-\delta)\log|\Omega| + O(1).$$
2. Una función cóncava f cumple $f((1-t)a + tb) \geq (1-t)f(a) + tf(b)$.
   a) Demuestre la desigualdad
   $$\frac{f(a_1) + f(a_2) + \ldots + f(a_n)}{n} \leq f\left(\frac{a_1 + a_2 + \ldots + a_n}{n}\right) \tag{5.39}$$
   para cualquier función cóncava y $n \in \mathbb{N}$.
   b) Sea $f : I \to \mathbb{R}$ una función doblemente diferenciable en un intervalo $I \subset \mathbb{R}$. Asuma que $f''(t) \leq 0$ para todo $t \in I$. Muestre que f es cóncava.
   c) Muestre que la función $x \mapsto x\log(1/x)$ es cóncava en $(0, \infty)$, y por lo tanto en $[0, \infty)$ (definiendo que el valor de $x\log(1/x)$ para $x = 0$ es 0). Usando la desigualdad (5.39), pruebe que
   $$\sum_{y \in A} \mathbb{P}(Y = y) \log \frac{1}{\mathbb{P}(Y = y)} \leq \mathbb{P}(Y \in A) \log \frac{|A|}{\mathbb{P}(Y \in A)}$$
   para cualquier variable aleatoria Y sobre Ω y subconjunto $A \subset \Omega$.
   d) Demuestre la desigualdad (5.37).
3. Sea Y una variable aleatoria en Ω tal que $\mathbb{H}(Y) = \log|\Omega| - o(1/|\Omega|)$.

a) Sea $y \in Y$. Muestre usando (5.37) con $E = \{y\}$ que

$$\log|\Omega| - o\left(\frac{1}{|\Omega|}\right) \leq -p\log p + (1-p)\log\frac{|\Omega|-1}{1-p}$$

con $p = \mathbb{P}(Y = y)$, y que, por lo tanto,

$$p\log|\Omega| + \frac{1-p}{|\Omega|} - o\left(\frac{1}{|\Omega|}\right) \leq -p\log p - (1-p)\log(1-p). \tag{5.40}$$

b) Utilizando que la parte derecha de (5.40) está acotada por $\log 2$, pruebe que $p = o(1)$, y además, usando también que $p = o(1)$, pruebe que

$$t\log t + 1 - o(1) \leq t(1 + o(1))$$

con $t = p|\Omega|$.

c) Demuestre que para $t \neq 1$ tenemos $t\log t + 1 - t > 0$. Concluya que

$$p = \frac{1}{|\Omega|}(1 + o(1)),$$

es decir, Y se comporta asintóticamente como una variable uniforme.

4. Sea $w(n)$ el número de divisores primos distintos de n. Demuestre que

$$S = \frac{1}{x}\sum_{n\leq x} w(n)\lambda(n+1) = o(\log\log x) = o\left(\frac{1}{x}\sum_{n\leq x} w(n)\right)$$

usando el Teorema 2.3 y el ejercicio 2 de 5.2.1. Para ponerlo de otra manera,

$$S = \mathbb{E}(w\lambda_+) + O(1)$$

con w la variable $w = \sum_{p\leq x^{1/4}} 1_{N\equiv 0(p)}$ y $\lambda_+ = \lambda(N+1)$, donde $\mathbb{P}(N = n) = 1/x$ para todo $n \in \Omega = [1, x]$.

5. Sean w, λ_+ las variables aleatorias del problema anterior.

a) λ_+ toma valores en el conjunto $\{-1, 1\}$. Use el Teorema 2.3 para demostrar que su entropía está cerca de la máxima posible, en el sentido que

$$\mathbb{H}(\lambda_+) = (1 + o(1))\log 2.$$

b) Demuestre usando el teorema de los números primos que w toma valores enteros en un intervalo $\Omega = [0, z)$, con $z \sim \frac{\log x}{\log\log x}$. Observe que la máxima entropía que w podría tener es $\log|\Omega| \sim \log\log x$.

c) Use (5.37) con $E = \{w < 2\log\log x\}$ y el ejercicio 2 de 5.2.1 para demostrar

$$\mathbb{H}(w) = o(\log|\Omega|) = o(\log\log x)$$

y que por lo tanto su entropía está lejos de la máxima posible.

5.5. Sumas de $\lambda(n)\lambda(n+p)$, en promedio sobre p. Conclusión. Por lo visto en la sección anterior, para obtener (5.1) sólo nos queda controlar sumas del tipo

$$\sum_{p\leq h}\sum_{X<n\leq 2X} \lambda(n)\lambda(n+p).$$

Estas son más complicadas que las del Corolario 4.10

$$\sum_{j\leq h}|\sum_{X<n\leq 2X} \lambda(n)\lambda(n+j)|^2$$

en el sentido de ahora sólo sumamos sobre primos, pero más sencillas ya que no tenemos el cuadrado. Vamos a ver que de nuevo es posible comprenderlas en términos de los coeficiente de Fourier de $\lambda(n)$ en intervalos cortos. La manera de hacerlo va a ser usar el método del círculo, el cual, en breve, consiste en usar la identidad

$$1_{k=0} = \int_0^1 e(k\alpha)\, d\alpha \qquad k \in \mathbb{Z} \tag{5.41}$$

para reescribir una ecuación aditiva (como $m = n + p$) de forma analítica, usando los armónicos $e(k\alpha)$.

Lema 5.11. *Si* $|w_j| \leq 1$ *tenemos que*

$$\sum_{p\leq h}\sum_{j\leq h} w_{j+p}\overline{w_j} \ll \epsilon\frac{h^2}{\log h} + \frac{h^2}{\log h}\int_{\mathfrak{M}_\epsilon}\left|\sum_{b\leq h} w_b e(b\alpha)\right| d\alpha$$

para cualquier $\epsilon \in (0,1)$, *con* $\mathfrak{M}_\epsilon = \{\alpha \in [0,1] : |\sum_{p\leq h} e(p\alpha)| > \epsilon\frac{h}{\log h}\}$.

Demostración. Usando la identidad (5.41) para $k = m - j - p$ tenemos

$$\sum_{p\leq h}\sum_{j\leq h}\overline{w_j}w_{j+p} = \sum_{m\leq 2h}\sum_{p\leq h}\sum_{j\leq h} w_m\overline{w_j}\int_0^1 e((m-j-p)\alpha)\, d\alpha,$$

y sacando fuera la integral y factorizando obtenemos

$$\sum_{p\leq h}\sum_{j\leq h}\overline{w_j}w_{j+p} = \int_0^1 W_{2h}(\alpha)\overline{W_h(\alpha)P_h(\alpha)}\, d\alpha,$$

con $W_d(\alpha) = \sum_{b\leq d} w_b e(b\alpha)$ y $P_d(\alpha) = \sum_{p\leq d} e(p\alpha)$. Ahora dividimos la integral entre los «arcos mayores» $\mathfrak{M}_\epsilon$ y los «arcos menores» $\mathfrak{m}_\epsilon = [0,1) \setminus \mathfrak{M}_\epsilon$. Para la parte de los arcos mayores tenemos

$$\int_{\mathfrak{M}_\epsilon} W_{2h}(\alpha)\overline{W_h(\alpha)}P_h(-\alpha)\, d\alpha \ll \frac{h^2}{\log h}\int_{\mathfrak{M}_\epsilon}|W_h(\alpha)|\, d\alpha = \frac{h^2}{\log h}\int_{\mathfrak{M}_\epsilon}|\sum_{b\leq h} w_b e(b\alpha)|\, d\alpha.$$

Para la parte de los arcos menores tenemos que

$$\left|\int_{\mathfrak{m}_\epsilon} W_{2h}(\alpha)\overline{W_h(\alpha)}P_h(-\alpha)\, d\alpha\right| \leq \frac{\epsilon h}{\log h}\int_0^1 |W_{2h}(\alpha)||W_h(\alpha)|\, d\alpha.$$

Usando la desigualdad $|ab| \leq |a|^2 + |b|^2$ vemos que

$$\int_0^1 |W_{2h}(\alpha)||W_h(\alpha)|\, d\alpha \leq \int_0^1 |W_{2h}(\alpha)|^2\, d\alpha + \int_0^1 |W_h(\alpha)|^2\, d\alpha$$

y por Parseval (ecuación (4.9))

$$\int_0^1 |W_h(\alpha)|^2\, d\alpha = \sum_{j\leq h}|w_j|^2 \leq h$$

y lo mismo para W_{2h}, luego obtenemos la cota $O(\epsilon h^2/\log h)$ para la parte de los arcos menores.

□

Ahora vamos a ver que la parte que queda (arcos menores) es muy pequeña, por lo que la integral sobre esa parte también va a ser pequeña.

Lema 5.12. *Sea* $0 < \epsilon < 1$, $h > 1$. *Con las definiciones del lema anterior tenemos que*

$$|\mathfrak{M}_\epsilon| \ll \frac{1}{\epsilon^4}\frac{1}{h}.$$

Demostración. La idea, igual que en la demostración del Lema 3.12, es controlar algún promedio de $P_h(\alpha) = \sum_{p\le h} e(p\alpha)$ para concluir que $\mathfrak{M}_\epsilon$ es un conjunto pequeño. La media cuadrática no será suficiente (ver problema 1), pero sí la potencia cuarta. Tenemos que

$$\int_0^1 |P_h(\alpha)|^4\,d\alpha = \int_0^1 |P_h^2(\alpha)|^2\,d\alpha = \int_0^1 |\sum_{|j|\le h}(\sum_{\substack{q-p=j\\ p,q\le h}} 1)e(j\alpha)|^2\,d\alpha,$$

donde p, q se mueven sobre los primos (hemos agrupado las frecuencias $e(q\alpha)\overline{e(p\alpha)} = e((q-p)\alpha)$). Ahora aplicamos Parseval (ecuación 4.9) para obtener

$$\int_0^1 |P_h(\alpha)|^4\,d\alpha = \sum_{|j|\le h} |\sum_{\substack{q-p=j\\ p,q\le h}} 1|^2.$$

Para $j = 0$ tenemos que la suma interior es $O(h/\log h)$ por el teorema de los números primos. Para el resto de js podemos usar la cota de criba (2.29) y así obtener

$$\sum_{|j|\le h} |\sum_{\substack{q-p=j\\ p,q\le h}} 1|^2 \ll \frac{h^2}{(\log h)^2} + \frac{h^2}{(\log h)^4}\sum_{j=1}^{h}\prod_{p|j}\left(1+\frac{1}{p}\right)^C$$

para alguna constante $C > 1$ fija. Como $\sum_{j=1}^h \prod(1+1/p)^C \ll h$ (ejercicio 2), obtenemos

$$\int_0^1 |P_h(\alpha)|^4\,d\alpha \ll \frac{h^3}{(\log h)^4},$$

de donde se deduce la cota para $|\mathfrak{M}_\epsilon|$.

□

Juntando los dos últimos lemas podemos concluir que los promedios de $\lambda(n+p)\lambda(n)$ están controlados por los coeficientes de Fourier de $\lambda(n)$ en intervalos cortos.

Proposición 5.13. *Sea* $0 < \epsilon < 1$ *y* $1 \le h \le \epsilon X$. *Entonces*

$$\sum_{p\le h}\sum_{X<n\le 2X} \lambda(n+p)\lambda(n) \ll \frac{\epsilon h X}{\log h} + \frac{1}{\epsilon^4 \log h}\max_{0\le\alpha\le 1}\int_X^{2X}\left|\sum_{x<m\le x+h}\lambda(m)e(m\alpha)\right|dx.$$

Demostración. Comenzamos por la observación de que si desplazamos el intervalo de sumación en n un poco la suma total casi no varía:

$$\begin{aligned}\sum_{X<n\le 2X}\lambda(n+p)\lambda(n) &= O(j) + \sum_{X+j<n\le 2X+j}\lambda(n+p)\lambda(n)\\ &= O(j) + \sum_{X<n\le 2X}\lambda(n+j+p)\lambda(n+j).\end{aligned}$$

Ahora usamos esa ecuación para todo $j \le h$, obteniendo

$$\sum_{X<n\le 2X}\lambda(n+p)\lambda(n) = O(h) + \frac{1}{h}\sum_{j\le h}\sum_{X<n\le 2X}\lambda(n+j+p)\lambda(n+j).$$

Así, sumando en p e intercambiando el orden de sumación, por TNP tenemos

$$\sum_{p\le h}\sum_{X<n\le 2X}\lambda(n+p)\lambda(n) = O\left(\frac{h^2}{\log h}\right) + \frac{1}{h}\sum_{X<n\le 2X}\sum_{p\le h}\sum_{j\le h}\lambda(n+j+p)\lambda(n+j). \tag{5.42}$$

Ahora usamos el Lema 5.11 con $w_j = \lambda(n+j)$ y llegamos a

$$\sum_{p\le h}\sum_{X<n\le 2X}\lambda(n+p)\lambda(n) \ll \frac{\epsilon h X}{\log h} + \frac{h}{\log h}\int_{\mathfrak{M}_\epsilon}\sum_{X<n\le 2X}|\sum_{b\le h}\lambda(n+b)e(b\alpha)|\,d\alpha. \tag{5.43}$$

Teniendo en cuenta que

$$\left|\sum_{b\leq h}\lambda(n+b)e(b\alpha)\right| = \left|\sum_{n<m\leq n+h}\lambda(m)e(m\alpha)\right|$$

y el Lema 5.12 obtenemos el resultado buscado. □

Usando las cotas para los coeficientes de Fourier de $\lambda(n)$ en intervalos cortos de la sección 4, vemos ahora que hay cancelación en las sumas de $\lambda(n+p)\lambda(n)$.

Corolario 5.14. *Sea* $\log h \leq (\log X)^{1/3}$. *Entonces*

$$\sum_{p\leq h}\sum_{X<n\leq 2X}\lambda(n)\lambda(n+p) \ll \frac{1}{(\log h)^{\frac{1}{75}-o(1)}}\cdot\frac{hX}{\log h}.$$

Demostración. Por la Proposición 5.13 y el Teorema 4.9 tenemos que

$$\sum_{p\leq h}\sum_{X<n<2X}\lambda(n)\lambda(n+p) \ll \frac{\epsilon hX}{\log h} + \frac{hX}{\epsilon^4\log h}\cdot\frac{1}{(\log h)^{\frac{1}{15}-o(1)}}$$

para cualquier $0<\epsilon<1$ y cualquier $1\leq h\leq \epsilon X$ con $\log h \leq (\log X)^{1/3}$. Tomando $\epsilon = (\log h)^{-\frac{1}{75}}$ obtenemos el resultado. □

El resultado que necesitamos se deduce fácilmente del Corolario 5.14.

Corolario 5.15. *Sean* $1\leq w\leq\sqrt{x}$ *y* $1<h\leq w$ *tal que* $\log h\leq(\log\frac{x}{w})^{1/3}$. *Entonces*

$$\sum_{h/2<p\leq h}\frac{1}{p}\sum_{x/w<n\leq x}\frac{\lambda(n)\lambda(n+p)}{n} \ll \frac{1}{(\log h)^{\frac{1}{75}-o(1)}}\cdot\frac{\log w}{\log h}.$$

Demostración. Ejercicio 4. □

Llegamos a la prueba del resultado principal.

Prueba del teorema 5.1. Podemos suponer que $w\leq x^{1/8}$ (pues podemos reducir a este caso subdividiendo el rango $x/w<n\leq x$) y también que w es más grande que una constante (pues, de lo contrario, la cota que debemos probar es trivial). Apliquemos el Corolario 5.10 con $\epsilon=(\log_3 H)^{-1/5}$ y el Corolario 5.15 con $h=\epsilon H$ para obtener que

$$\sum_{\epsilon H/2<p\leq\epsilon H}\sum_{\substack{x/w<n\leq x\\ p|n}}\frac{\lambda(n)\lambda(n+p)}{n} \ll \frac{1}{(\log_3 H)^{1/5}}\cdot\frac{\log w}{\log H}.$$

Luego aplicamos el Lema 5.2 con $K_0=\epsilon H/2$, $K_1=\epsilon H$ para concluír que

$$\sum_{\frac{x}{w}<n\leq x}\frac{\lambda(n)\lambda(n+1)}{n} \ll \log\epsilon H\cdot\frac{1}{(\log_3 H)^{1/5}}\cdot\frac{\log w}{\log H}+\log\epsilon H \ll \frac{\log w}{(\log_3 H)^{1/5}}.$$

□

5.5.1. Ejercicios.

1. Demuestre que

$$\int_0^1 |\sum_{p<h}e(p\alpha)|^2\,d\alpha = \frac{h}{\log h}(1+o(1)).$$

   Deduzca que $|\mathfrak{M}_\epsilon| \ll \frac{\log h}{\epsilon^2 h}$, con $\mathfrak{M}_\epsilon$ definido como en el Lema 5.11. Observe que esta cota es peor que la obtenida en el Lema 5.12 para $1/\sqrt{\log h}<\epsilon<1$.

2. Vamos a demostrar que
$$\sum_{j\leq h}\prod_{p|j}\left(1+\frac{1}{p}\right)^C \ll_C h.$$
Para ello, pruebe las siguientes desigualdades e identidades, con $w(d)=\sum_{p:p|d}1$:
$$\begin{aligned}\sum_{j\leq h}\prod_{p|j}\left(1+\frac{1}{p}\right)^C &\ll_C \sum_{j\leq h}\prod_{p|j}\left(1+\frac{2C}{p}\right)\\ &=\sum_{d\leq h}\frac{(2C)^{w(d)}}{d}\sum_{\substack{j\leq h\\ d|j}}1\leq h\sum_{d=1}^{\infty}\frac{(2C)^{w(d)}}{d^2}\ll_C h.\end{aligned}$$

3. También es factible demostrar el Lema 5.12 estudiando la suma $P(\alpha)=\sum_{p\leq h}e(p\alpha)$ para todo α y viendo cuando es grande. Es posible ver que el conjunto $\mathfrak{M}_\epsilon$ de los valores para los cuales $P(\alpha)$ es grande consiste en los α cercanos a racionales con denominador pequeño. En una dirección (mostrar que los α que no estan cerca a tales racionales no están en $\mathfrak{M}_\epsilon$), esto no es nada fácil; se trata de la parte principal de la estrategia de Vinogradov para el problema ternario de Goldbach. Veamos como demostrar la otra dirección, por lo menos para algunos racionales de denominador pequeño.
   a) Demuestre que si $\alpha=\delta/h$, $|\delta|<1$, entonces $|P(\alpha)|=\frac{h}{\log h}(1+o_h(1)+O(\delta))$. Observe que esto demuestra que $|\mathfrak{M}_\epsilon|\gg h$, lo cual coincide con la cota superior que se obtiene en el Lema 5.12 para $\epsilon\gg 1$.
   b) Demuestre que, de todos los carácteres χ módulo 5, el único cuya función $L(s,\chi)$ tiene un polo en $s=1$ es el carácter trivial χ_0 que vale 1 para todo número no divisible por 5. (Sugerencia: para los otros carácteres $\chi \bmod 5$, utilice sumación por partes para estimar $\sum_n \chi(n)n^{-\sigma}$, $\sigma=1+\epsilon$, recordando que $\sum_{n\leq u}\chi(n)<5$ para todo u (¿por qué?).)
   c) Usando el apartado anterior y el Teorema 4.3, muestre que
$$\sum_{p\leq h, p\equiv b(5)}1=\frac{1}{4}\frac{h}{\log h}(1+o_h(1))=\frac{1}{4}(1+o_h(1))\sum_{p\leq h}1$$
   para $b=1,2,3,4$.
   d) Pruebe usando los apartados anteriores que si $\alpha=\frac{1}{5}+\frac{\delta}{h}$ con $|\delta|<1$, entonces
$$P(\alpha)=-\frac{1}{4}\frac{h}{\log h}(1+O(\delta)+o_h(1)).$$

4. Deseamos mostrar que el Corolario 5.15 se deduce del Corolario 5.14. El procedimiento es sencillo y muy general.
   a) Sea $f:\mathbb{Z}^+\to\mathbb{C}$ arbitrario. Sea $F(t)=\sum_{p\leq t}f(p)$. Muestre que
$$\sum_{\frac{h}{2}<p\leq h}\frac{f(p)}{p}=\int_{h/2}^{h}\frac{F(t)}{t^2}dt+\frac{1}{h}F(h)-\frac{2}{h}F(h/2).$$
   b) Sea $g:\mathbb{Z}^+\to\mathbb{C}$ tal que $|g(n)|\leq 1$ para todo n. Sea $G(t)=\sum_{t<n\leq 2t}g(n)$. Muestre que, para $1\leq x_0\leq x_1$,
$$\sum_{x_0<n\leq x_1}\frac{g(n)}{n}=\int_{x_0}^{x_1}\frac{G(t)}{t^2}dt+O(1).$$

c) Usando (4*b*) y el Corolario 5.14, pruebe que

$$\sum_{p\leq h}\sum_{x/w<n\leq x}\frac{\lambda(n)\lambda(n+p)}{n} \ll \frac{1}{(\log h)^{\frac{1}{75}-o(1)}}\frac{h\log w}{\log h}+\frac{h}{\log h} \tag{5.44}$$

para $1\leq w\leq x$ y $\log h \geq (\log \frac{x}{w})^{1/3}$. Está claro que el término $h/\log h$ puede omitirse para $h\leq w$.

d) Use (5.44) y el apartado (4*a*) para deducir el Corolario 5.15. (Podemos, por cierto, suponer que h es más grande que una constante, pues de lo contrario lo que queremos probar es trivial.)

Referencias

[Bon36] C. E. Bonferroni. Teoria statistica delle classi e calcolo delle probabilita. (Pubbl. d. R. Ist. Super. di Sci. Econom. e Commerciali di Firenze. 8) Firenze: Libr. Internaz. Seeber. 62 S. (1936)., 1936.

[Bru15] V. Brun. Über das *Goldbach*sche Gesetz und die Anzahl der Primzahlpaare. *Arch. Math. Naturvid.*, 34(8):3–19, 1915.

[BSZ13] J. Bourgain, P. Sarnak, and T. Ziegler. Disjointness of Moebius from horocycle flows. In *From Fourier analysis and number theory to Radon transforms and geometry*, volume 28 of *Dev. Math.*, pages 67–83. Springer, New York, 2013.

[Che73] J. R. Chen. On the representation of a larger even integer as the sum of a prime and the product of at most two primes. *Sci. Sinica*, 16:157–176, 1973.

[DD82] H. Daboussi and H. Delange. On multiplicative arithmetical functions whose modulus does not exceed one. *J. London Math. Soc. (2)*, 26(2):245–264, 1982.

[FI10] J. Friedlander and H. Iwaniec. *Opera de cribro*, volume 57 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2010.

[HR74] H. Halberstam and H.-E. Richert. *Sieve methods*. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], London-New York, 1974. London Mathematical Society Monographs, No. 4.

[Hux72] M. N. Huxley. On the difference between consecutive primes. *Invent. Math.*, 15:164–170, 1972.

[IK04] H. Iwaniec and E. Kowalski. *Analytic number theory*, volume 53 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2004.

[Kát86] I. Kátai. A remark on a theorem of H. Daboussi. *Acta Math. Hungar.*, 47(1-2):223–225, 1986.

[Mon71] H. L. Montgomery. *Topics in multiplicative number theory.*, volume 227. Springer, 1971.

[Mon94] H. L. Montgomery. *Ten lectures on the interface between analytic number theory and harmonic analysis*, volume 84 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 1994.

[Mot76] Y. Motohashi. On the sum of the Möbius function in a short segment. *Proc. Japan Acad.*, 52(9):477–479, 1976.

[MR16] K. Matomäki and M. Radziwiłł. Multiplicative functions in short intervals. *Ann. of Math. (2)*, 183(3):1015–1056, 2016.

[MRT15] K. Matomäki, M. Radziwiłł, and T. Tao. An averaged form of Chowla's conjecture. *Algebra Number Theory*, 9(9):2167–2196, 2015.

[MRT16] K. Matomäki, M. Radziwiłł, and T. Tao. Sign patterns of the Liouville and Möbius functions. *Forum Math. Sigma*, 4:e14, 44, 2016.

[MRT17] K. Matomäki, M. Radziwiłł, and T. Tao. Correlations of the von Mangoldt and higher divisor functions II. Divisor correlations in short ranges. *Mathematische Annalen*, pages 1–48, 2017.

[MRT19] K. Matomäki, M. Radziwiłł, and T. Tao. Correlations of the von Mangoldt and higher divisor functions I. Long shift ranges. *Proc. Lond. Math. Soc. (3)*, 118(2):284–350, 2019.

[MV77] H. L. Montgomery and R. C. Vaughan. Exponential sums with multiplicative coefficients. *Invent. Math.*, 43(1):69–82, 1977.

[MV07] H. L. Montgomery and R. C. Vaughan. *Multiplicative number theory. I. Classical theory*, volume 97 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2007.

[Ram76] K. Ramachandra. Some problems of analytic number theory. *Acta Arith.*, 31(4):313–324, 1976.

[Rén47] A. A. Rényi. On the representation of an even number as the sum of a single prime and a single almost-prime number. *Doklady Akad. Nauk SSSR (N.S.)*, 56:455–458, 1947.

[RS62] J. B. Rosser and L. Schoenfeld. Approximate formulas for some functions of prime numbers. *Illinois J. Math.*, 6:64–94, 1962.

[Sou17] K. Soundararajan. The Liouville function in short intervals. *Astérisque*, (390):Exp. No. 1119, 453–479, 2017. Séminaire Bourbaki. Vol. 2015/2016. Exposés 1104–1119.

[Tao16a] T. Tao. The Erdős discrepancy problem. *Discrete Anal.*, pages Paper No. 1, 29, 2016.

[Tao16b] T. Tao. The logarithmically averaged Chowla and Elliott conjectures for two-point correlations. *Forum Math. Pi*, 4:e8, 36, 2016.

Georg-August Universität Göttingen, Mathematisches Institut, Bunsenstrasse 3-5, D-37073 Göttingen, Alemania

Université Paris Diderot, UFR de Mathématiques, case 7012, 75205 Paris CEDEX 13, Francia

Email address: harald.helfgott@gmail.com

Departamento de Matemáticas, Universidad Autónoma de Madrid, Madrid 28049 SPAIN

Email address: adrian.ubis@uam.es