

Aplicaciones del análisis armónico

(análisis armónico y de todo un poco)

Fernando Chamizo

Curso 2012-2013

Índice general

Glosario	IV
1. Introducción	1
1.1. Precalentamiento	1
1.1.1. La serie de Fourier original	2
1.1.2. Algunos teoremas	5
1.1.3. Algunos ejemplos y gráficos	9
1.2. Aprendiendo a sumar	13
1.2.1. Sumación por partes	13
1.2.2. La fórmula de sumación de Poisson	16
1.2.3. Técnicas de variable compleja	21
1.3. Análisis armónico discreto	27
1.3.1. Caracteres y transformadas	28
1.3.2. Implementación del análisis de Fourier discreto	32
1.3.3. Análisis de Fourier discreto en la vida cotidiana	38
2. Sumas e integrales oscilatorias	49
2.1. El principio de incertidumbre	49
2.1.1. Diferentes formas del principio de incertidumbre	50
2.1.2. Un poco de física cuántica	55
2.2. Integrales oscilatorias	65
2.2.1. El principio de fase estacionaria	66
2.2.2. Las matemáticas de la difracción	72
2.3. Sumas oscilatorias	80
2.3.1. El método de van der Corput	81
2.3.2. La gran criba	87
2.3.3. Algunas aplicaciones aritméticas	91

Glosario

$[x]$	parte entera de x	15
$\{x\}$	$x - [x] - 1/2$	18
$a \equiv b \pmod{m}$	$m \mid a - b$	32
$f \gg g$	$g \ll f$	13
Δ	Laplaciano	55
$f \ll g$	$\limsup f /g < \infty$	13
∇	Gradiente	75
$\partial\Omega$	Borde de la variedad Ω	75
$f * g$	Convolución $\int f(t)g(x-t)dt$	5
$f \sim g$	$\lim f/g = 1$	13
$f \asymp g$	$f \gg g$ y $f \ll g$	13
C	Funciones continuas	2
C_0^n	C^n con soporte compacto	3
C^n	Funciones con n derivadas continuas	2
δ	Delta de Dirac	3
$d(n)$	Número de divisores	97
D_N	Núcleo de Dirichlet $\sum_{n=-N}^N e(nx)$	5
δ_P	Delta de Dirac 1-periódica	3
$e(x)$	Exponencial compleja $e^{2\pi ix}$	2
F_N	Núcleo de Fejér $\sum_{n=-N}^N (1 - n /N)e(nx)$	6
$\ f\ _p$	La norma $(\int f ^p)^{1/p}$	7
\widehat{f}	Transformada de Fourier $\int f(x)e(-\xi x) dx$	4
\mathcal{F}	Transformada inversa $\int f(\xi)e(x\xi) d\xi$	4
$\Gamma(s)$	Función Gamma $\int_0^\infty x^{s-1}e^{-x} dx$	12
γ	Constante de Euler $\Gamma'(1)$	15

\mathbb{H}	$\{x + iy : x \in \mathbb{R}, y > 0\}$	18
$\Im(z)$	Parte imaginaria de z	22
$\Lambda(n)$	$\log p$ si $n = p^k$ con p primo y 0 en el resto	16
L^p	Espacio de funciones con $\int f ^p < \infty$	8
$\mu(n)$	función de Möbius	90
$f = O(g)$	$\limsup f /g < \infty$	13
$f = o(g)$	$\lim f/g = 0$	13
$\pi(x)$	$\#\{p \leq x : p \text{ primo}\}$	16
$\pi_2(N)$	Primos gemelos hasta N	97
$\psi(x)$	$\sum_{p \leq x} \Lambda(n)$	16
$r_2(n)$	$\#\{(a, b) \in \mathbb{Z}^2 : n = a^2 + b^2\}$	93
$\Re(z)$	Parte real de z	22
S^1	El grupo multiplicativo $\{z \in \mathbb{C} : z = 1\}$	29
$\sigma(n)$	$\sum_{d n} d$	18
$S_N f$	Suma parcial de la serie de Fourier de f	5
\mathbb{T}	El toro unidimensional $\mathbb{R}/\{x \mapsto x + 1\}$	2
$\theta(z)$	$\sum_{n=-\infty}^{\infty} e(n^2 z)$	18
\mathbb{Z}_N	Grupo aditivo de los enteros módulo N	27

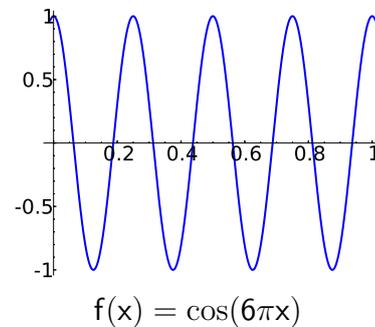
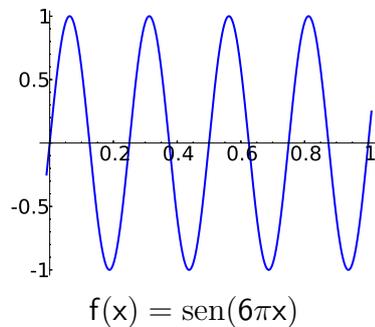
Capítulo 1

Introducción

1.1. Prealentamiento

El término *análisis* deriva de una palabra griega que significa descomposición (y entre otras cosas, ¡incluso muerte!) y es éste sentido etimológico, excluyendo el paréntesis, el que vamos a manejar en el curso, más que el otro significado común en Matemáticas de algo con derivadas e integrales.

El *análisis armónico* se ocupa, a grandes rasgos, de la descomposición de funciones en *tonos puros* que llamaremos *armónicos*. Sin rigor, consideramos tonos puros a ciertos objetos que nos recuerdan a las funciones $\text{sen}(2\pi nx)$ y $\text{cos}(2\pi nx)$ con $n \in \mathbb{Z}$, las cuales aparecen en los desarrollos de Fourier clásicos.



En términos acústicos, en el rango de frecuencias audibles estas funciones suenan como mantener uno de los pitidos de la alarma de un reloj electrónico o una nota de una flauta.

Con esta ambigüedad, el análisis armónico se convierte en un área muy amplia cuyas fronteras son muy subjetivas y están sujetas a qué deseamos denominar armónicos. Incluso con toda esta amplitud no se cubren muchos temas actuales importantes en los que los armónicos aparecen más como motivación que como ingredientes.

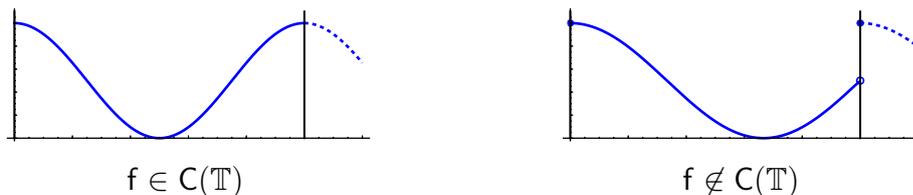
1.1.1. La serie de Fourier original

El punto de partida del análisis armónico fue el *desarrollo en serie de Fourier*. Afirma que cualquier función 1-periódica razonable $f : \mathbb{R} \rightarrow \mathbb{C}$ se puede analizar como

$$(1.1) \quad f(x) = \sum_{n=-\infty}^{\infty} a_n e(nx) \quad \text{con} \quad a_n = \int_0^1 f(x) e(-nx) dx$$

donde se ha empleado la notación sintética $e(x) = e^{2\pi i x}$ para englobar los senos y cosenos que representan los tonos puros. A los a_n se les llama *coeficientes de Fourier*. Las funciones $\{e(nx)\}_{n=-\infty}^{\infty}$ son los armónicos en esta descomposición.

Por la periodicidad, el intervalo de integración $[0, 1]$ se puede sustituir por cualquier otro de longitud 1. A este respecto, en muchos contextos se muestra conveniente introducir el *toro* $\mathbb{T} = \mathbb{R}/\{x \mapsto x + 1\}$ obtenido al enrollar la recta real alrededor de la circunferencia unidad (un toro unidimensional es una circunferencia), o equivalentemente pegar los extremos del intervalo $[0, 1]$. En este sentido, $f : \mathbb{T} \rightarrow \mathbb{C}$ representa una función 1-periódica y a_n viene dado por la integral sobre \mathbb{T} . No hay que dejarse impresionar por toda esta palabrería, simplemente considerarla como una notación conveniente o una taquigrafía. Por ejemplo $C(\mathbb{T})$ y $C^n(\mathbb{T})$ representan las funciones continuas y las funciones de clase C^n que son 1-periódicas.

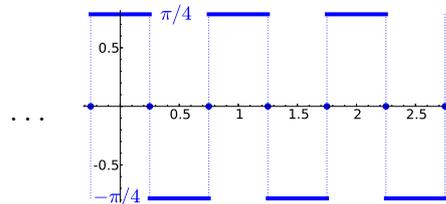


Los matemáticos desafortunadamente demasiado a menudo son muy reacios a hablar de un resultado antes de enunciarlo con rigor y generalidad. En el caso del desarrollo en serie de Fourier, es difícil traducir “cualquier función 1-periódica razonable” en la clase más amplia de funciones para las que se cumple (1.1). El problema es tan complejo y con tantas ramificaciones que en muchas exposiciones desvía la atención del problema realmente natural: incluso suponiendo que partimos de una función 1-periódica muy regular ¿por qué tiene que ser superposición de senos y cosenos? ¿Por qué podemos, teóricamente, imitar el sonido de cualquier instrumento musical tocando muchas flautas al mismo tiempo? Ésta es la pregunta candente, nunca mejor dicho, en la famosa memoria de 1822 de J. Fourier sobre la teoría analítica del calor [Fou88] y que de ningún modo es obvia, de hecho destacados matemáticos de los siglos XVIII y XIX manifestaron opiniones en ambos sentidos.

La primera serie de Fourier que aparece en su famosa memoria es:

$$(1.2) \quad \frac{\pi}{4} = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} \cos((2n+1)x) \quad \text{para} \quad -\frac{\pi}{2} < x < \frac{\pi}{2}.$$

Fourier nota que en $[\pi/2, 3\pi/2]$, el resto del intervalo hasta completar un periodo, el resultado es $-\pi/4$ por las simetrías del coseno, salvo en los extremos donde la función vale cero. Con la notación de (1.1), escribiríamos esto como



$$\dots = \sum_{n=-\infty}^{\infty} \frac{(-1)^n}{4n+2} e((2n+1)x) \quad \text{para todo } x \in \mathbb{R}.$$

Los argumentos de Fourier son muy objetables desde el rigor actual, e incluso con el de su tiempo. Su razonamiento [Fou88, Ch.III, Sec.II.171] para “probar” (1.2) fue suponer $\pi/4 = \sum a_n \cos((2n+1)x)$ e igualar las series de Taylor de ambos miembros en $x=0$. Con ello se llega a un sistema lineal con infinitas incógnitas, los a_n , y coeficientes tendiendo a infinito que Fourier “resuelve” para obtener $a_n = (-1)^n/(2n+1)$, incluso a pesar de que todas las ecuaciones del sistema menos una dan lugar a series divergentes al comprobar las soluciones. La prueba rigurosa de la convergencia de la serie de Fourier bajo condiciones bastante generales tendría que esperar a un famoso trabajo de P.G.L. Dirichlet en 1828.

Después de este desvío histórico, volvamos al problema de por qué nos deberíamos creer que las funciones 1-periódicas razonables coinciden con su serie de Fourier. Comencemos considerando la *delta de Dirac 1-periódica* (también llamada *peine de Dirac*)

$$(1.3) \quad \delta_P(x) = \sum_{n=-\infty}^{\infty} \delta(x-n)$$

donde δ es la *delta de Dirac* habitual. Esta notación, introducida por P.A.M. Dirac en su matematización de la física cuántica, representa algo que los físicos dicen que es una función que satisface $\int_{\mathbb{R}} \delta = 1$ y $\int_{\mathbb{R}} \delta(x)f(x) dx = f(0)$, mientras que los matemáticos prefieren negarle la categoría de función y decir que es una distribución o una sucesión de funciones $\{\phi_n\}_{n=1}^{\infty}$, que llaman *aproximación de la identidad*, tal que $\int_{\mathbb{R}} \phi_n = 1$ y $\lim_{n \rightarrow \infty} \int_{\mathbb{R}} \phi_n(x)f(x) dx = f(0)$. Aquí f es una *función test* típicamente C_0^{∞} . La función δ_P tiene la misma propiedad reemplazando \mathbb{R} por \mathbb{T} . Podemos imaginar estos objetos como la función de densidad correspondiente a una partícula unidad localizada en el origen. La masa total es 1 pero no hay nada fuera del origen, $\delta(x) = 0$ si $x \neq 0$ y $\delta(0) = \infty$, con un infinito tan grande que la integral de δ es 1 en cualquier intervalo que contenga al origen.

Si δ_P coincidiera con su serie de Fourier, es decir, si la igualdad

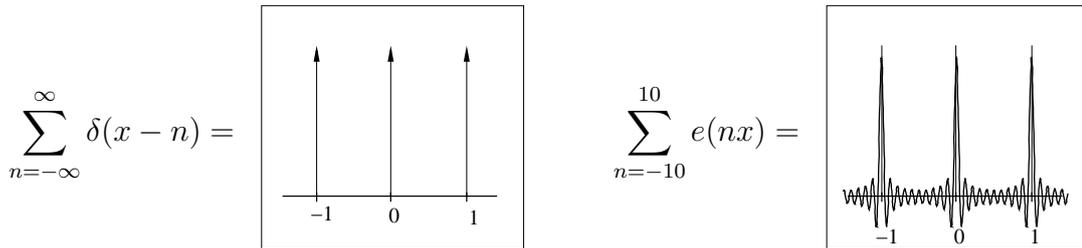
$$(1.4) \quad \delta_P(x) = \sum_{n=-\infty}^{\infty} e(nx)$$

tuviera sentido, entonces formalmente

$$f(x) = \int_{\mathbb{T}} \delta_P(t) f(x-t) dt = \sum_{n=-\infty}^{\infty} \int_{\mathbb{T}} f(x-t) e(nt) dt = \sum_{n=-\infty}^{\infty} e(nx) \int_{\mathbb{T}} f(u) e(-nu) du$$

donde se ha usado el cambio de variable $t = x - u$, y esto es (1.1).

En definitiva, si nos creemos el desarrollo de Fourier de la delta de Dirac periódica, deberíamos creernos en una línea el de las funciones razonables. ¿Por qué albergamos alguna esperanza en hacer riguroso este argumento fantasmagórico? Simplemente porque al truncar el segundo miembro de (1.4) se obtienen cosas que nos recuerdan al primer miembro:



Pasar de funciones 1-periódicas a T -periódicas es tan fácil como hacer el cambio $x \mapsto x/T$, con ello el desarrollo de Fourier de funciones razonables T -periódicas responde a la fórmula

$$(1.5) \quad f(x) = \frac{1}{T} \sum_{n=-\infty}^{\infty} a_n e\left(n \frac{x}{T}\right) \quad \text{con} \quad a_n = \int_{-T/2}^{T/2} f(x) e\left(-n \frac{x}{T}\right) dx$$

donde el $1/T$ proviene del diferencial. Si permitimos $T \rightarrow \infty$ manteniendo $\xi = n/T$ fijo, los coeficientes a_n vendrán dados por la *transformada de Fourier*

$$(1.6) \quad \widehat{f}(\xi) = \int_{-\infty}^{\infty} f(x) e(-\xi x) dx.$$

Por otra parte, imponer que algo tenga periodo infinito es como no imponer nada porque $x + \infty = \infty$, además la suma en (1.5) se puede ver como una suma de Riemann. Con estas ideas, nos arriesgamos a establecer, como hizo Fourier en el capítulo IX de su memoria, una suerte de análogo en \mathbb{R} de las series de Fourier, al que se le llama *fórmula de inversión*:

$$(1.7) \quad (\widehat{f})^{\sim} = f \quad \text{donde} \quad \check{g}(x) = \int_{-\infty}^{\infty} g(\xi) e(x\xi) d\xi.$$

A \check{g} se le llama *transformada inversa (o antitransformada) de Fourier* de g . Nótese en la analogía con (1.1) una vez que se descodifica la notación. La fórmula de inversión es

$f(x) = \int_{\mathbb{R}} \widehat{f}(\xi) e(x\xi) d\xi$ y estamos expresando f como una superposición de armónicos $\{e(\xi x)\}_{\xi \in \mathbb{R}}$ cuya naturaleza no discreta motiva integrar más que sumar al analizar f .

De nuevo surge la pregunta de por qué esto es cierto para “funciones razonables”, ahora mitigada por la deducción anterior. La respuesta análoga a (1.4) es la enigmática fórmula que raramente aparecerá en un libro de análisis matemático:

$$(1.8) \quad \boxed{\delta(x) = \int_{-\infty}^{\infty} e(x\xi) d\xi.}$$

Si aceptamos esta fórmula en la que integramos una función no integrable. (1.7) se sigue de un Fubini sospechoso:

$$\int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f(t) e(-\xi t) dt \right) e(\xi x) d\xi = \int_{-\infty}^{\infty} f(t) \int_{-\infty}^{\infty} e((x-t)\xi) d\xi = \int_{-\infty}^{\infty} f(t) \delta(x-t) dt$$

que coincide con $f(x)$.

1.1.2. Algunos teoremas

Para hacer rigurosa la demostración anterior de (1.1), definimos el sumatorio truncado y la operación integral empleada en la idea intuitiva:

$$D_N(x) = \sum_{n=-N}^N e(nx) \quad y \quad (f * g)(x) = \int_{\mathbb{T}} f(t) g(x-t) dt.$$

A D_N se le llama *núcleo de Dirichlet* y a $f * g$ *convolución de f y g* . Se puede definir análogamente en \mathbb{R} y en ambos casos es conmutativa y asociativa.

Si procedemos como antes pero reemplazando la suma formal $\sum e(nx)$ por $D_N(x)$, obtendremos una fórmula para las sumas parciales de la serie de Fourier, que denotaremos por $S_N f$,

$$S_N f(x) = \sum_{|n| \leq N} a_n e(nx) = (D_N * f)(x).$$

La idea es que δ_P es formalmente la identidad para la convolución, $\delta_P * g = g$, y si en algún sentido $D_N \rightarrow \delta_P$, la fórmula anterior debería dar (1.1). Para buscar sentido a $D_N \rightarrow \delta_P$, usamos el viejo truco de poner y quitar, empleando que $\int_{\mathbb{T}} D_N = 1$,

$$(D_N * f)(x) = \int_{\mathbb{T}} f(x-t) D_N(t) dt = f(x) + \int_{\mathbb{T}} (f(x-t) - f(x)) D_N(t) dt.$$

Si $f \in C(\mathbb{T})$ entonces $f(x-t) \rightarrow f(x)$ uniformemente cuando $t \rightarrow 0$. Si la función D_N estuviera bastante concentrado en el origen y $\lim \int_{\mathbb{T}} |D_N| < \infty$, tendríamos (1.1) para funciones continuas. Sin embargo $\int_{\mathbb{T}} |D_N|$ se comporta como $\log N$ y la última condición

no es cierta. Las cosas son más tangibles de lo que parecen puesto que sumando una progresión geométrica tenemos la expresión explícita

$$(1.9) \quad D_N(x) = \frac{\operatorname{sen}(2\pi(N + 1/2)x)}{\operatorname{sen}(\pi x)} \quad \text{para } x \notin \mathbb{Z} \quad \text{y} \quad D_N(x) = 2N + 1 \quad \text{para } x \in \mathbb{Z}.$$

El *Lema de Riemann-Lebesgue* asegura que para cualquier función integrable g , se cumple $\int_{\mathbb{T}} g(x)e(\lambda x) dx = 0$ si $\lambda \rightarrow \infty$. Por tanto, si para cierto x , $(f(x-t) - f(x))/\operatorname{sen}(\pi t)$ es integrable como función de t , deducimos que $S_N f(x) \rightarrow f(x)$, como deseábamos (habitualmente se llama a este resultado *criterio de Dini*). Si exigimos $f \in C^1(\mathbb{T})$, entonces nuestra hipótesis será satisfecha en todo punto, y hemos probado:

Teorema 1.1.1. *Para cualquier $f \in C^1(\mathbb{T})$ se cumple $\lim_{N \rightarrow \infty} S_N f = f$ uniformemente.*

Empleando que el núcleo de Dirichlet es perfectamente simétrico, especialmente que $\int_0^{1/2} D_N = \int_{-1/2}^0 D_N = 1/2$, se pueden permitir discontinuidades de salto, siempre que haya regularidad a derecha e izquierda, y la serie de Fourier tenderá al punto medio. Esta observación y el hecho de que la *funciones de variación acotada* sean esencialmente las de derivada integrable (permitiendo así usar el criterio de Dini), sugieren que el resultado más general al que podemos llegar con los razonamientos anteriores es [Kat76, II.2]

Teorema 1.1.2. *Si f es de variación acotada en \mathbb{T} , se cumple $\lim_{N \rightarrow \infty} S_N f(x) = \frac{1}{2}(f(x^-) + f(x^+))$.*

Antes hemos dejado caer que el problema de la convergencia de la serie de Fourier sería más simple si la integral $\int_{\mathbb{T}} |D_N|$ estuviera uniformemente acotada en N . Con esta idea, se introduce el *núcleo de Fejér*

$$F_N(x) = \sum_{n=-N}^N \left(1 - \frac{|n|}{N}\right) e(nx) = \frac{\operatorname{sen}^2(\pi N x)}{N \operatorname{sen}^2(\pi x)},$$

que al ser positivo cumple $\int_{\mathbb{T}} |F_N| = \int_{\mathbb{T}} F_N = 1$. Entonces el argumento anterior, cambiando D_N por F_N , nos da que las *sumas de Cesàro* de la serie de Fourier, es decir

$$\tilde{S}_N f = \sum_{|n| < N} a_n \left(1 - \frac{|n|}{N}\right) e(nx),$$

constituyen la forma de sumar correcta al desarrollar por Fourier funciones continuas. Éste es un resultado bien conocido de L. Fejér. Exagerando un poco, el hecho de que queramos seguir sumando de la manera habitual es una manía matemática que no tiene demasiado reflejo en el plano práctico. Ya L. Euler, el gran artista de la sumación, inventó métodos para acelerar series porque tenía que calcular sumas sin ordenador. Considerar

$\tilde{S}_N f$ en vez de $S_N f$ es una forma de aceleración de series en sentido fuerte (véanse las gráficas de la siguiente sección) puesto que incluso hace que algunas series de Fourier divergentes dejen de serlo, permitiendo sustituir “razonable” por “continua” en la frase que introduce (1.1).

Teorema 1.1.3 (de Fejér). *Para cualquier $f \in C(\mathbb{T})$ se cumple $\lim_{N \rightarrow \infty} \tilde{S}_N f = f$ uniformemente.*

De hecho, incluso si f tiene discontinuidades de salto, la convergencia sigue siendo uniforme en cualquier compacto que no las incluya y en las discontinuidades $\tilde{S}_N f$ tiende al punto medio, como antes [Zyg77, III.3].

Supongamos que tenemos una función $f \in C^\infty(\mathbb{T})$, entonces integrando por partes, sus coeficientes de Fourier decaen muy rápido y la convergencia de la serie de Fourier es absoluta y uniforme. Con ello no tendremos ningún reparo en elevar al cuadrado en (1.1) e integrar sobre \mathbb{T} . El resultado es la bella *identidad de Parseval* (también llamada *identidad de Plancherel*)

$$(1.10) \quad \int_{\mathbb{T}} |f|^2 = \sum_{n=-\infty}^{\infty} |a_n|^2.$$

Toda la magia ha venido de que $\int_{\mathbb{T}} e(-mx)e(nx) dx = 0$ si $n \neq m$ y es 1 si $n = m$. Dicho de manera pedante pero útil, $\{e(nx)\}_{n=-\infty}^{\infty}$ es una familia ortonormal con el producto escalar $\langle f, g \rangle = \int_{\mathbb{T}} \bar{f}g$ que es el de $L^2(\mathbb{T})$, las funciones de cuadrado integrable. Ahora bien, C^∞ es denso en L^2 y entonces (1.10) tiene perfecto sentido para $f \in L^2$ independientemente de que su serie de Fourier converja. De (1.10) se sigue $\|f - S_N f\|_2^2 = \sum_{|n| > N} |a_n|^2$ con $\|\cdot\|_2$ la norma en L^2 , esto es, $\|f\|_2 = (\int |f|^2)^{1/2}$. Con todo ello, las series de Fourier convergen siempre, si aceptamos vivir en el mundo L^2 que es completo (podemos tomar límites sin salir de él) y mucho más amplio que las funciones continuas.

Teorema 1.1.4. *Para cualquier $f \in L^2(\mathbb{T})$ se cumple $\lim_{N \rightarrow \infty} \|f - S_N f\|_2 = 0$.*

Por su estructura de espacio de Hilbert (espacio completo con producto escalar), L^2 tiene un regusto de álgebra lineal que nos familiariza con las series de Fourier. Así $S_N f$ es el resultado de proyectar f en el espacio de dimensión finita $\{e(nx)\}_{n=-N}^N$. Desde otro punto de vista, los coeficientes de Fourier vienen del ajuste de mínimos cuadrados al minimizar la varianza. En suma, los polinomios trigonométricos son álgebra lineal pura y en el límite tenemos las series de Fourier.

Por polarización o por un cálculo directo obtenemos una identidad de Parseval (1.10) para el producto escalar que resulta ser dual en cierto modo de otra que da los coeficientes de Fourier de una convolución.

$$(1.11) \quad \sum_{n=-\infty}^{\infty} \bar{a}_n b_n = \int_{\mathbb{T}} \bar{f}g \quad \text{y} \quad c_n = a_n b_n,$$

donde $\{a_n\}$, $\{b_n\}$ y $\{c_n\}$ son los coeficientes de Fourier de f , g y $f * g$.

La teoría L^2 es maravillosa pero no debemos olvidarnos de que con respecto al problema original, lo que hemos hecho es cambiar la definición. Ahora la igualdad de (1.1) hay que entenderla en sentido de que el momento de orden 2 de la diferencia (la varianza) es cero. La pregunta matemática natural es si todavía tenemos convergencia en el sentido habitual. Por supuesto, las funciones de L^2 están definidas salvo conjuntos de medida cero, entonces necesariamente nos tendremos que olvidar de ellos. Una respuesta afirmativa la dio L. Carleson en 1966 [Car66]. Su teorema, muy difícil de probar, es uno de los hitos del análisis armónico:

Teorema 1.1.5 (de Carleson). *Para cualquier $f \in L^2(\mathbb{T})$ se cumple*

$$\lim_{N \rightarrow \infty} S_N f(x) = f(x) \quad \text{para casi todo } x.$$

Este teorema fue extendido por R. Hunt a $f \in L^p(\mathbb{T})$ con $1 < p < \infty$. M. Lacey y C. Thiele dieron en 2000 una prueba mucho más sencilla que la original [LT00].

Respecto a la falta de convergencia, con algunas manipulaciones con series lacunares, se construye [Kat76, II.2] [Kör88, §15] una función continua cuya serie de Fourier no converge en un punto. También es posible conseguir que la convergencia falle en infinitos puntos. De hecho hasta se puede tomar un conjunto denso (este ejemplo es muy antiguo, data de 1873 y se debe a P. du Bois-Reymond), además un resultado de J.-P. Kahane y Y. Katznelson [KK66] asegura que para conjunto de medida cero $E \subset \mathbb{T}$ existe una función continua tal que su serie de Fourier diverge en E . Por otro lado A. Kolmogorov, demostró que existían funciones integrables, esto es, en $L^1(\mathbb{T})$, tales que su serie de Fourier diverge en todo punto. Nótese que ser integrable es lo mínimo que se puede pedir para que tenga sentido calcular los coeficientes de Fourier.

Respecto a la transformada de Fourier, la teoría es más o menos paralela. Una diferencia esencial es que para la propia definición se necesita cierto decaimiento. Esto es notorio cuando intentamos hacer la teoría L^2 . Si bien se tenía $L^2(\mathbb{T}) \subset L^1(\mathbb{T})$, ahora por la no compacidad de \mathbb{R} nos enfrentamos a $L^2(\mathbb{R}) \not\subset L^1(\mathbb{R})$ y no podremos llevar a cabo la integración en (1.6) en general. La solución es usar algún tipo de regularización [DM72, Ch.2]. Una vez abierta la veda de la regularización, es posible dar sentido a (1.7) para cualquier función de L^1 . Dentro de la teoría L^2 se vuelve a tener la identidad de Parseval y la relación con la convolución:

$$(1.12) \quad \int_{-\infty}^{\infty} \bar{f}g = \int_{-\infty}^{\infty} \widehat{\bar{f}}\widehat{g} \quad \text{y} \quad (f * g)^{\wedge} = \widehat{f}\widehat{g}.$$

A efectos más teóricos que prácticos, cabe preguntarse si hay alguna clase sencilla y amplia de funciones en la que podamos tomar transformadas y antitransformadas en

el sentido estricto de la definición, sin regularizaciones. La respuesta son las funciones C^∞ con *decaimiento rápido* o *clase de Schwartz*, en la que todas las derivadas decaen más rápido que el inverso de un polinomio. Esta clase desempeña un papel fundamental en la teoría de distribuciones de L. Schwartz que, entre otras cosas, permite que los matemáticos escriban deltas de Dirac sin sufrir escalofríos.

La invariancia de la clase de Schwartz proviene de la relación entre \widehat{f} y $\widehat{f^{(k)}}$ que se sigue fácilmente integrando por partes y tiene su análogo también en series de Fourier:

$$(1.13) \quad a_n = (2\pi in)^{-k} \int_{\mathbb{T}} f^{(k)}(x) e(-nx) dx \quad \text{y} \quad \widehat{f}(\xi) = (2\pi i\xi)^{-k} \widehat{f^{(k)}}(\xi)$$

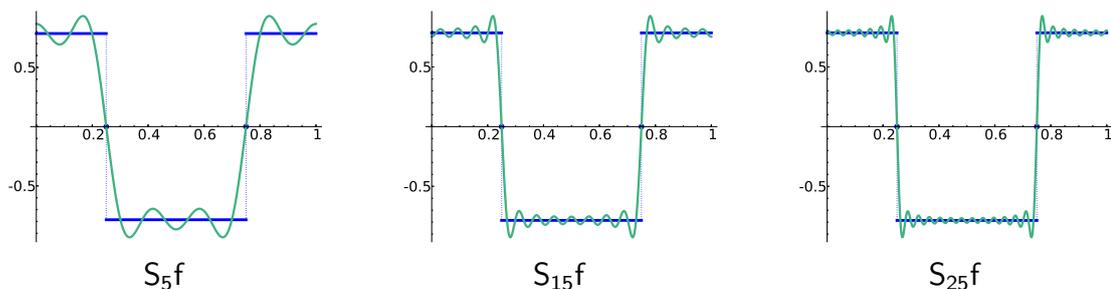
para n y ξ no nulos.

1.1.3. Algunos ejemplos y gráficos

El Teorema 1.1.2 asegura que tenemos la igualdad

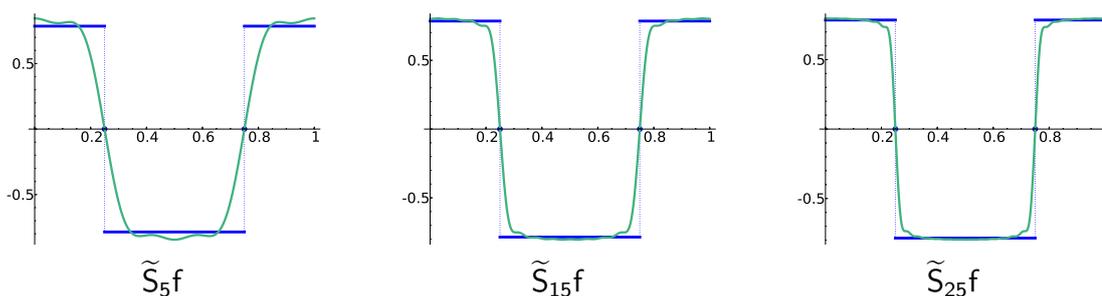
$$f(x) = \sum_{n=-\infty}^{\infty} \frac{(-1)^n}{4n+2} e((2n+1)x) = \sum_{n=-\infty}^{\infty} \frac{(-1)^n}{2n+1} \cos(2\pi(2n+1)x)$$

donde f es la función 1-periódica de onda rectangular considerada por Fourier (escalada), que en $[-1/2, 1/2]$ vale $\pi/4$ si $|x| < 1/4$, $-\pi/4$ si $|x| > 1/4$ y $f(-1/4) = f(1/4) = 0$. La convergencia en la primera serie hay que entenderla, según hemos visto, como límite de $\sum_{n=-N}^N$. Las siguientes gráficas muestran que las cosas funcionan como cabría esperar, tomando valores de N cada vez mayores.



Veremos más adelante que ese primer pico que aparece a ambos lados de las discontinuidades curiosamente no disminuye su tamaño por muy grande que tomemos N .

Las siguientes gráficas permiten comparar lo que se consigue sumando a la manera de Fejér tomando el mismo número de sumandos (realmente uno menos, porque en $\widetilde{S}_N f$ el coeficiente $a_{\pm N}$ está multiplicado por cero). La velocidad de convergencia ha mejorado notablemente aunque, por supuesto, no puede ser uniforme porque la función límite es discontinua.



Ahora vamos a jugar un poco con esta serie de Fourier. Por definición, f^2 es constante $\pi^2/16$ en casi todo punto y por tanto (1.10) es muy sencillo de aplicar y conduce inmediatamente a

$$\frac{\pi^2}{16} = \sum_{n=-\infty}^{\infty} \frac{1}{4(2n+1)^2} \quad \text{que equivale a} \quad \frac{\pi^2}{8} = \frac{1}{1^2} + \frac{1}{3^2} + \frac{1}{5^2} + \frac{1}{7^2} + \dots$$

Si calculamos $f * f$, la convolución de f consigo misma, se obtiene la función 1-periódica diente de sierra que en $[-1/2, 1/2]$ vale $\pi^2(1-4|x|)/16$, la cual es continua y C^1 a trozos. De modo que por la segunda fórmula de (1.11) se tiene la igualdad (con convergencia uniforme)

$$\frac{\pi^2}{16}(1-4|x|) = \sum_{n=-\infty}^{\infty} \frac{e((2n+1)x)}{4(2n+1)^2} \quad \text{para } -\frac{1}{2} \leq x \leq \frac{1}{2}.$$

Lo cual es coherente con el cálculo directo de los coeficientes de Fourier. Aplicando la primera fórmula de (1.11) a esta función y la original, y utilizando (1.10), se obtiene las sumas de otras dos bellas series:

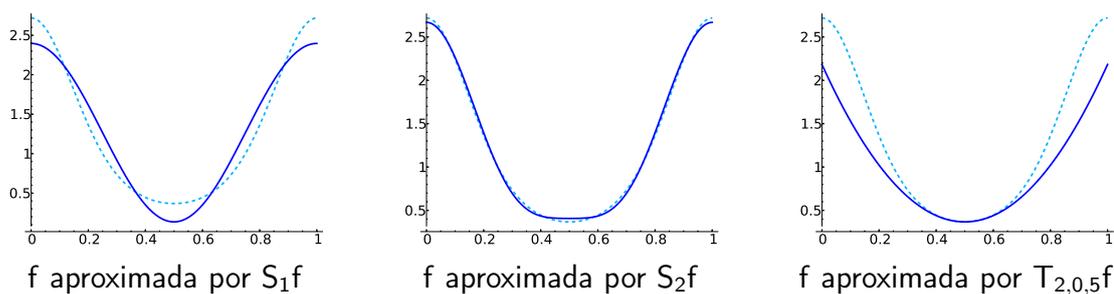
$$\frac{\pi^3}{32} = \frac{1}{1^3} - \frac{1}{3^3} + \frac{1}{5^3} - \frac{1}{7^3} + \dots \quad \text{y} \quad \frac{\pi^4}{96} = \frac{1}{1^4} + \frac{1}{3^4} + \frac{1}{5^4} + \frac{1}{7^4} + \dots$$

En los libros de análisis armónico, a menudo se muestran gráficas como las primeras que hemos visto y se hace poco hincapié en lo espectacularmente bien que aproximan las sumas parciales $S_N f$ cuando la función es suficientemente regular.

Consideremos por ejemplo la función 1-periódica $f(x) = e^{\cos(2\pi x)}$. Con ayuda de un programa o de un paquete matemático, podemos aproximar sus primeros coeficientes de Fourier. Por la simetría, $a_n = a_{-n}$ y no hace falta calcular los de índice negativo.

$$\begin{aligned} a_0 &= 1,266065877 & a_1 &= 0,565159103 & a_2 &= 0,135747669 & a_3 &= 0,022168424 \\ a_4 &= 0,002737120 & a_5 &= 0,000271463 & a_6 &= 2,24886 \cdot 10^{-5} & a_7 &= 1,59921 \cdot 10^{-6} \end{aligned}$$

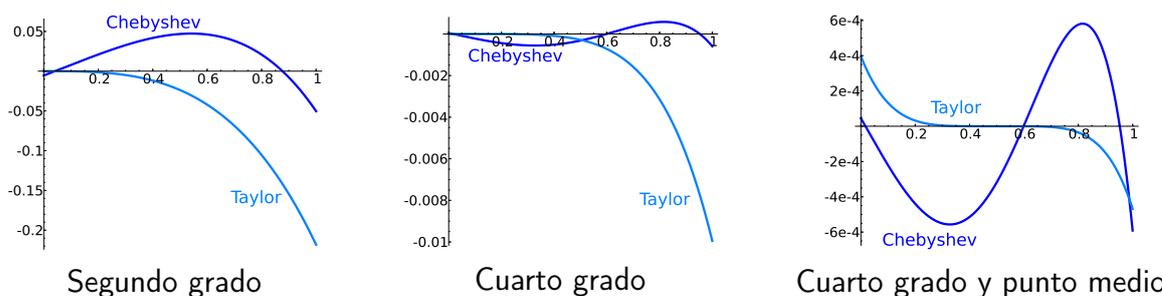
La aproximación es notoria usando muy pocos coeficientes de Fourier. En las figuras siguientes se muestra con línea de puntos la gráfica de f . En la segunda, sólo con a_0 , a_1 y a_2 se consigue una aproximación poco distinguible a simple vista. En la tercera figura se muestra la pobreza en términos globales del polinomio de Taylor de grado 2 en el punto medio.



Para reivindicar el polinomio de Taylor se podría argumentar que al no requerir la periodicidad es más general y sobre todo, desde el punto de vista numérico arcaico, que es un polinomio. Algo que evaluamos a mano sin grandes esfuerzos en contraposición con las funciones trascendentes seno y coseno. Sin embargo con un truco ingenioso se puede forzar a que Fourier siga venciendo a Taylor en muchas situaciones. Digamos por ejemplo que deseamos una aproximación polinómica buena de $g(x) = e^x$ en $[0, 1]$ para hacer cálculos a mano. Si forzamos la periodicidad de g , el resultado ni siquiera estaría en $C(\mathbb{T})$, pero f como antes, sí cumple $f \in C^\infty(\mathbb{T})$ y ambas funciones se relacionan mediante $f(x) = g(\cos(2\pi x))$. Aunque no lo parezca, $\cos(n \arccos x)$ es un polinomio para cualquier $n \in \mathbb{Z}$ [Kör88, §5, §45], son los *polinomios de Chebyshev*, entonces $P_N(x) = a_0 + 2 \sum_{n=1}^N a_n \cos(n \arccos x)$ son aproximaciones polinómicas de e^x , que provienen de un cambio de variable en la serie de Fourier. Por ejemplo,

$$P_4(x) = 0,043793 x^4 + 0,177347 x^3 + 0,499196 x^2 + 0,997307 x + 1,000044$$

Por otro lado, en Cálculo I nos enseñaron $T_N = \sum_{n=0}^N x^n/n!$ como la aproximación de Taylor más destacada. Las dos primeras figuras muestran las gráficas de los errores $P_N(x) - e^x$ y $T_N(x) - e^x$ para $N = 2$ y $N = 4$, respectivamente:



La última figura es un jarro de agua fría. Resulta que Taylor en el punto medio, que tiene una fórmula más fea que T_N , todavía vence a Fourier-Chebyshev. Sin embargo, esta victoria dista mucho de ser total. Pensemos en qué ocurriría si deformamos la gráfica de e^x en $[3/4, 1]$ dejando todavía una función $C^\infty([0, 1])$. Entonces el polinomio de Taylor, que sólo emplea información local, no podrá aproximar más allá de cierta cantidad fija, mientras que Fourier-Chebyshev convergerá a la función por el Teorema 1.1.1.

La transformada de Fourier explícita más famosa es la de las *funciones gaussianas*:

$$(1.14) \quad \boxed{f(x) = e^{-ax^2} \quad \longrightarrow \quad \widehat{f}(\xi) = \sqrt{\frac{\pi}{a}} e^{-\pi^2 \xi^2 / a}} \quad \text{para } a > 0.$$

Partiendo de $\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$ se llega a $\int_{-\infty}^{\infty} e^{-ax^2} e^{-2zx} dx = e^{z^2/a} \sqrt{\pi/a}$ para cualquier $z \in \mathbb{R}$ y la fórmula anterior se deduce justificando que también es posible tomar z complejo, concretamente $z = \pi \xi i$.

Elijiendo $a = \pi$ se tiene una función que es igual que su transformada (y que su antitransformada). En [DM72, §2.5] hay una descripción completa y sencilla de todas las funciones de L^2 con esta propiedad. Éstas son parte de las *funciones de Hermite* que tienen un papel importante en un cálculo básico de física cuántica.

La transformada y la transformada inversa de Fourier comparten casi la misma definición. A partir de una se deduce la otra por medio de la relación $\widehat{f}(x) = \check{f}(-x)$. Por ejemplo, un cálculo sencillo prueba que si $f(x) = e^{-|x|}$ entonces $\widehat{f}(\xi) = 2/(1 + 4\pi^2 \xi^2)$, por consiguiente, la transformada de $2/(1 + 4\pi^2 x^2)$ es $e^{-|\xi|}$, lo cual no se obtiene por integración directa. Aplicando (1.12), es posible deducir rápidamente el valor de algunas integrales de aspecto complicado:

$$\int_{-\infty}^{\infty} \frac{dx}{(1+x^2)^2} = \frac{\pi}{2} \int_{-\infty}^{\infty} \frac{4 d\xi}{(1+4\pi^2 \xi^2)^2} = \frac{\pi}{2} \int_{-\infty}^{\infty} e^{-2|x|} dx = \frac{\pi}{2}.$$

Una minitabla de transformadas de Fourier, es la siguiente:

$f(x)$	$\widehat{f}(\xi)$	$f(x)$	$\widehat{f}(\xi)$
$\delta(x)$	1	$e^{-a x }$	$\frac{2a}{a^2 + 4\pi^2 \xi^2}$
x^{-1}	$\pi i \operatorname{sgn}(\xi)$	e^{-ax^2}	$\sqrt{\frac{\pi}{a}} e^{-\pi^2 \xi^2 / a}$
$ x ^{-1/2}$	$ \xi ^{-1/2}$	$(a^2 + x^2)^{-1}$	$\pi a^{-1} e^{-2\pi a \xi }$
$ x ^{-\nu}$	$\frac{\pi 2\pi \xi ^{\nu-1}}{\Gamma(\nu) \cos(\pi \nu / 2)}$	$\operatorname{sech}(ax)$	$\pi a^{-1} \operatorname{sech}(\pi^2 a^{-1} \xi)$

donde $0 < \nu < 1$, $a > 0$ y Γ es la función Gamma que estudiaremos en la próxima sección.

Las transformadas de la izquierda no tienen sentido dentro de la integral de Lebesgue pero aun así son útiles. La más sorprendente es la segunda, que hay que entender como límite de $\int_{\epsilon < |x| < M} x^{-1} e(-\xi x) dx$ cuando $\epsilon \rightarrow 0^+$ y $M \rightarrow \infty$. La utilidad de estas transformadas de funciones malas se hace patente cuando se juntan con otras de funciones buenas. Por ejemplo, a pesar de que $|x|^{-1/2}$ no está siquiera en L^2 , aplicando formalmente (1.12), se infiere $\int_{\mathbb{R}} |x|^{-1/2} g(x) dx = \int_{\mathbb{R}} |\xi|^{-1/2} \widehat{g}(\xi) d\xi$, lo cual es una fórmula correcta si g tiene cierta regularidad y decaimiento.

Lema 1.2.1 (Sumación por partes). *Se cumple la identidad*

$$\sum_{n=1}^N a_n b_n = a_N S_N + \sum_{n=1}^{N-1} (a_n - a_{n+1}) S_n$$

donde $S_n = \sum_{k=1}^n b_k$. Además, si $(a_n)_{n=1}^N$ es real monótona no creciente y positiva, se tiene la acotación

$$\left| \sum_{n=1}^N a_n b_n \right| \leq a_1 \sup_{1 \leq n \leq N} |S_n|.$$

Demostración. Para la identidad, basta escribir $b_n = S_n - S_{n-1}$ y agrupar convenientemente los términos. La acotación se sigue tomando valores absolutos y notando que $|a_N| + \sum_{n=1}^{N-1} |a_n - a_{n+1}| = a_1$ porque la suma es telescópica. \square

A menudo se aplica la acotación anterior después de subdividir en *intervalos diádicos*, es decir, considerando $2^j \leq n < 2^{j+1}$, con ello se busca evitar problemas con el infinito y que otros a_n , no sólo el a_1 , participen en la acotación para hacerla más precisa.

Como primer ejemplo, probaremos la convergencia de la serie $\sum_{n=1}^{\infty} n^{-1/2} \cos(n\sqrt{2})$ que no aparece en los cursos básicos de Cálculo porque no es una serie absolutamente convergente y el *criterio de Leibniz* ($\sum (-1)^n c_n$ converge si c_n decrece a cero) es inaplicable. Tomando $a_n = (n+M)^{-1/2}$ y $b_n = \cos((n+M)\sqrt{2})$ en el Lema 1.2.1, se tiene

$$\sum_{n=M+1}^L n^{-1/2} \cos(n\sqrt{2}) \leq (M+1)^{-1/2} \sup_{M < n \leq L} \frac{1}{2} \left| D_n\left(\frac{\sqrt{2}}{2\pi}\right) - D_M\left(\frac{\sqrt{2}}{2\pi}\right) \right| \ll (M+1)^{-1/2}.$$

Así, cuando $L > M \rightarrow \infty$ se tiene $S_L = S_M + o(1)$ y por tanto la serie converge (las sumas parciales forman una sucesión de Cauchy). El mismo argumento sirve para probar que $\sum n^{-1/2} \cos(2\pi n x)$ converge para todo $0 < x < 1$ (uniformemente sobre compactos) a pesar de que por (1.10) sabemos que ni siquiera es de cuadrado integrable [Zyg77, I,(2.6)].

Parece que las aplicaciones juiciosas de la sumación por partes a series infinitas están limitadas al caso en que $a_n - a_{n+1} \rightarrow 0$, sin embargo algunos métodos de aceleración de series contradicen esta idea. Si c_n decrece a cero, tomando $a_n = (-1)^{n-1}$ y $b_n = c_n - c_{n+1}$ en el Lema 1.2.1, se deduce la igualdad:

$$\sum_{n=1}^{\infty} (-1)^{n-1} c_n = \frac{1}{2} c_1 + \frac{1}{2} \sum_{n=1}^{\infty} (-1)^{n-1} (c_n - c_{n+1}).$$

El método de aceleración de series con *transformadas de Euler* consiste en aplicar indefinidamente esta igualdad. Sin llegar a ese límite, notemos que por ejemplo, para

$c_n = 1/(2n - 1)$ la serie de la derecha da un error 10 veces menor que la segunda al aproximar $\pi/4 = \sum (-1)^{n-1} c_n$ cuando se usan 5 términos, y usando 50 términos el error se hace 100 veces menor. Este tipo de trucos son muy a tener en cuenta para aumentar la rapidez de cálculos masivos con ordenador.

Una versión un poco más compleja de la sumación partes se utiliza primordialmente cuando se quiere ir más allá de una acotación. Lleva asociado el nombre de N.H. Abel, el genial matemático que se escandalizó de que en su tiempo hubiera tan poco hecho sobre convergencia (escribió: “salvo casos muy simples [...] ni una serie infinita ha sido sumada rigurosamente [...] Es cierto que la mayoría son válidas, pero es muy sorprendente”).

Lema 1.2.2 (Lema de Abel). *Sea $(c_n)_{n=1}^{\infty}$ una sucesión arbitraria de números complejos y sea $C(t) = \sum_{n \leq t} c_n$. Dado $x \geq 1$, para cualquier $g : [1, \infty) \rightarrow \mathbb{C}$, $g \in C^1$, se verifica*

$$\sum_{n \leq x} c_n g(n) = C(x)g(x) - \int_1^x C(t)g'(t) dt.$$

Es posible deducir este resultado del Lema 1.2.1 o incluso dar una prueba directa en pocas líneas [Ell75, Th.1.6], sin embargo preferimos aquí una prueba avanzada pero ilustrativa.

Demostración. Por continuidad podemos suponer que x no es entero, entonces el primer miembro es $\int_{1/2}^x h(t)g(t) dt$ con $h(t) = \sum c_n \delta(t - n)$ y δ la delta de Dirac. Como $C'(t) = h(t)$, el resultado se reduce a integrar por partes. El uso de deltas de Dirac se justifica mediante aproximaciones de la identidad. \square

Aplicando el Lema 1.2.2 con $c_n = 1$ y $g(t) = 1/t$ se deduce

$$\sum_{n \leq x} \frac{1}{n} = \frac{[x]}{x} + \int_1^x \frac{[t]}{t^2} dt = \log x + \frac{[x]}{x} + \int_1^x \frac{[t] - t}{t^2} dt.$$

donde $[\cdot]$ indica la parte entera. Escribiendo la integral como $\int_1^x = \int_1^{\infty} - \int_x^{\infty} = \text{cte} + O(x^{-1})$, se concluye el resultado clásico $\sum_{n \leq x} n^{-1} = \log x + \gamma + O(x^{-1})$, donde $\gamma = 0,577\dots$ es una constante llamada *constante de Euler*.

Dos funciones básicas en la teoría de la distribución de primos son $\pi(x)$, el número de primos en $[1, x]$, y $\psi(x)$, el logaritmo del mínimo común múltiplo de los enteros en ese mismo rango. Esta última función fue introducida por P. Chebyshev en sus importantes trabajos sobre los primos que culminaron con su prueba del *postulado de Bertrand* (entre un número mayor que 1 y su doble siempre hay un primo). En principio $\psi(x)$ es muy poco

natural pero de cara a ciertas estimaciones, lo es más que $\pi(x)$. Una fórmula alternativa más manejable para ψ es:

$$\psi(x) = \sum_{n \leq x} \Lambda(n) \quad \text{donde} \quad \Lambda(n) = \begin{cases} \log p & \text{si } n = p^k \text{ con } p \text{ primo} \\ 0 & \text{en otro caso} \end{cases}$$

A la función Λ se le llama *símbolo de von Mangoldt*.

El *teorema de los números primos* asegura $\psi(x) \sim x$ y la sumación por partes traslada este resultado a $\pi(x)$ como $\text{Li}(x) = \int_2^x dt/\log t$, el *logaritmo integral*.

Lema 1.2.3. *Si se cumple $\psi(x) = x + O(E(x))$ para cierta función creciente y positiva E , entonces también se cumple $\pi(x) = \text{Li}(x) + O(x^{1/2} + E(x)/\log x)$. De hecho*

$$\pi(x) = \text{Li}(x) + \frac{\psi(x) - x}{\log x} + \int_2^x \frac{\psi(t) - t}{t(\log t)^2} dt + O(x^{1/2}).$$

Demostración. La primera parte se deduce de la segunda, ya que el integrando está acotado por $-E(x)(1/\log t)'$.

Es fácil ver que $\pi(x) = \sum_{n \leq x} \Lambda(n)/\log n + O(x^{1/2} \log x)$. De hecho con un poco de esfuerzo se reduce el error a $O(x^{1/2})$. El Lema 1.2.2 con $c_n = \Lambda(n+1)$ y $g(x) = 1/\log(x+1)$ prueba $\pi(x) = \psi(x)/\log x - \int_2^x \psi(t)/(t(\log t)^2) dt + O(x^{1/2})$, que es equivalente a la fórmula buscada. \square

1.2.2. La fórmula de sumación de Poisson

Si en (1.4) multiplicamos en ambos miembros por una función e integramos en \mathbb{R} , se obtiene la magnífica fórmula de sumación de Poisson que transforma una suma sobre enteros en otra:

$$\delta_P(x) = \sum e(nx) \Rightarrow \int f(x) \delta_P(x) = \sum \int f(x) e(nx) \Rightarrow \sum f(n) = \sum \hat{f}(n).$$

Lo más difícil a la hora de establecer esta fórmula es dar condiciones que aseguren la convergencia de las series del enunciado y la prueba. Una posibilidad tomada de [DM72] es la siguiente (véase en [Zyg77, I, (13.5)] una versión con hipótesis débiles):

Teorema 1.2.4 (Fórmula de sumación de Poisson). *Para cualquier $f \in C^2(\mathbb{R})$ satisfaciendo $|f(x)| + |f'(x)| + |f''(x)| = O((1 + |x|)^{-2})$, se verifica*

$$\sum_{n=-\infty}^{\infty} f(n) = \sum_{n=-\infty}^{\infty} \hat{f}(n).$$

Demostración. Considérese la función $F(x) = \sum_m f(m+x)$. Sus coeficientes de Fourier son $a_n = \sum \int_m^{m+1} f(x) e(-nx) dx = \hat{f}(n)$ (las hipótesis aseguran convergencia absoluta de la serie de Fourier) y la fórmula se deduce de que $F(0) = \lim S_N f(0)$. \square

Si formalmente elegimos como f la función característica de un conjunto de enteros, este resultado nos daría que su cardinal se expresa en términos de integrales oscilatorias. En muchas ocasiones, estas integrales se aproximan a su vez por funciones oscilatorias (ondas) y parte de la importancia de la fórmula de sumación de Poisson en teoría analítica de números radica en su capacidad para contar enteros a través de interferencias de ondas.

Por supuesto, esta descripción es sólo un boceto porque ya de partida, una función característica no es C^2 y para funciones poco regulares hay problemas de convergencia en el segundo miembro. Por otro lado, la fórmula de sumación de Poisson es involutiva, es decir, de un sólo uso, al aplicarla dos veces recuperamos la suma original. Todo ello no es óbice para que los expertos en teoría analítica de números canten sus virtudes. Así leemos en §5.4 de [Hux96]:

The exponential $e(t)$ is the violin of the mathematical orchestra, and $\rho(t)[= 1/2 - \text{Frac}(t)]$ is the flute, its themes are developed by the violins. Poisson summation is the tuba: very deep, but ridiculous when used too much.

Y en la introducción de [IK04]:

Poisson summation for number theory is what a car is for people in modern communities –it transports things to other places and it takes you back home when applied next time– one cannot live without it.

Veamos un ejemplo concreto del Teorema 1.2.4. En analogía con la historia del pequeño C.F. Gauss y la suma de $1 + 2 + \dots + 100$, imaginemos a un profesor malvado que exige el valor de $S = \sum_{n=-\infty}^{\infty} e^{-n^2/100}$ con unas pocas cifras decimales a una clase de estudiantes armados sólo con calculadoras científicas. Eligiendo en el Teorema 1.2.4 $f(x) = e^{-x^2/100}$ y usando (1.14) se tiene

$$S = \sum_{n=-\infty}^{\infty} \int_{-\infty}^{\infty} f(x)e(nx) dx = 10\sqrt{\pi} \sum_{n=-\infty}^{\infty} e^{-100\pi^2 n^2}$$

y el alumno aventajado escribiría triunfante en su pizarra $0 < S - 10\sqrt{\pi} < 10^{-400}$ donde 10^{-400} es una estimación generosa de la cola de la serie.

Cuando un ejemplo funciona bien es conveniente convertirlo en un resultado y, con suerte, quizá en una teoría. Aquí reformularemos el ejemplo anterior con una notación que nos acerca mínimamente al origen de la teoría de formas modulares. Consideremos la *función θ de Jacobi* (la normalización difiere de la habitual y de la original de C.G.J. Jacobi)

$$(1.16) \quad \theta(z) = \sum_{n=-\infty}^{\infty} e(n^2 z) \quad \text{para } z \in \mathbb{H}$$

donde \mathbb{H} es el semiplano superior $\{x+iy : x \in \mathbb{R}, y > 0\}$, para asegurar la convergencia.

Proposición 1.2.5. *Para todo $z \in \mathbb{H}$, se verifica*

$$\theta(z) = \sqrt{\frac{i}{2z}} \theta\left(\frac{-1}{4z}\right),$$

donde en $\sqrt{}$ se escoge la rama habitual que envía \mathbb{R}^+ en \mathbb{R}^+ .

Demostración. Aplicar el Teorema 1.2.4 a $f(x) = e(zx^2)$. □

Al combinar esta especie de invariancia de θ por $z \mapsto -1/4x$ con la obvia por $z \mapsto z + 1$, se obtienen una serie de simetrías que son relevantes en la teoría de formas modulares. Históricamente, una de sus primeras consecuencias fue que el número de representaciones de un entero n como suma de 4 cuadrados, $r_4(n)$, responde a la fórmula sencilla

$$r_4(n) = 8 \left(\sigma(n) - \sigma\left(\frac{n}{4}\right) \right) \quad \text{donde } \sigma(n) = \sum_{d|n} d$$

y donde se entiende $\sigma(n/4) = 0$ si $4 \nmid n$. El esquema de la prueba clásica es demostrar que las funciones que tienen las mismas simetrías que θ^4 forman un espacio vectorial de dimensión 2 (sobre \mathbb{C}) y demostrar también que

$$f_1(z) = 1 + 24 \sum_{n=1}^{\infty} \left(\sigma(n) - 2\sigma\left(\frac{n}{2}\right) \right) e(nz) \quad \text{y} \quad f_2(z) = 1 + 24 \sum_{n=1}^{\infty} \left(\sigma\left(\frac{n}{2}\right) - 2\sigma\left(\frac{n}{4}\right) \right) e(nz)$$

son dos ejemplos con tales simetrías (con el lenguaje adecuado esto último es más sencillo de lo que parece). Por tanto, existen $\lambda_1, \lambda_2 \in \mathbb{C}$ tales que $\theta^4(z) = \sum_{n=0}^{\infty} r_4(n) e(nz) = \lambda_1 f_1(z) + \lambda_2 f_2(z)$. Ajustando λ_1 y λ_2 para que los dos primeros coeficientes cuadren, se deduce la fórmula para $r_4(n)$.

Otra manera de entender la fórmula de sumación de Poisson es a través del desarrollo de Fourier de la función $\{x\} = x - [x] - 1/2$, que es la parte fraccionaria normalizada para que su promedio sea nulo.

$$(1.17) \quad \{x\} = \frac{i}{2\pi} \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \frac{e(nx)}{n} = -\frac{1}{\pi} \sum_{n=1}^{\infty} \frac{\text{sen}(2\pi nx)}{n},$$

donde la igualdad, como hemos visto, no se da para $x \in \mathbb{Z}$ por la convergencia al punto medio (Teorema 1.1.2).

No es difícil deducir del Lema de Abel (Lema 1.2.2) con $c_n = 1$, la identidad

$$(1.18) \quad \sum_{n=1}^N f(n) = \int_1^N f(x) dx + \frac{1}{2} (f(1) + f(N)) + \int_1^N \{x\} f'(x) dx.$$

Sustituyendo (1.17), la última integral es $2 \sum_{n=1}^{\infty} \int_1^N f(x) \cos(2\pi nx) dx$ después de integrar por partes. Aquí hemos pasado de puntillas sobre el problema de intercambiar suma e integral. Se puede demostrar que si $f \in C^2$, no hay problema. Suponiendo esto, con un poco más de esfuerzo se deduce una fórmula de sumación de Poisson para intervalos finitos.

Proposición 1.2.6. Sean $M, N \in \mathbb{Z}$ con $M > N$ y sea $f \in C^2$. Entonces

$$\sum_{n=N}^M f(n) = \frac{1}{2}(f(N) + f(M)) + \sum_{n=-\infty}^{\infty} \int_N^M f(x) e(nx) dx.$$

Demostración. Según el argumento anterior a partir de (1.18), empleando $e(-nx) + e(nx) = 2 \cos(2\pi nx)$ y $e(0) = 1$, se tiene

$$\sum_{n=1}^M f(n) = \frac{1}{2}(f(1) + f(N)) + \sum_{n=-\infty}^{\infty} \int_1^M f(x) e(nx) dx.$$

Cambiando N por M y restando las identidades, se obtiene el resultado esperado. \square

Una consecuencia indirecta de esto es la evaluación de las *sumas de Gauss*

$$G_N = \sum_{n=1}^N e\left(\frac{n^2}{N}\right) \quad \text{con } N \in \mathbb{Z}^+.$$

A pesar de que esto será una consecuencia más o menos rápida de la Proposición 1.2.6, tiene un contenido aritmético muy profundo y Gauss tuvo que poner mucho esfuerzo hasta conseguir una prueba de la fórmula aparentemente inocente del siguiente resultado.

Proposición 1.2.7. Para cualquier $N \in \mathbb{Z}^+$ se tiene la fórmula

$$G_N = \frac{1 + (-i)^N}{1 - i} \sqrt{N}.$$

Nótese que lo que esto indica es que $N^{-1/2}G_N$ es 4-periódica en \mathbb{Z}^+ tomando los valores 1, 0, i y $1+i$ para los cuatro primeros enteros positivos. Para la prueba seguimos [Dav80]. En [IK04] hay una demostración originalísima utilizando integrales oscilatorias y aritmética, mientras que en [Ros88] se puede ver una prueba clásica sin análisis (en general éstas son más complicadas, véase también [Ter99, §8]).

Demostración. Vamos a aplicar la Proposición 1.2.6 a $f(x) = e(x^2/N)$ en el intervalo $[0, N]$. En principio esto excede en 1 el intervalo de sumación de G_N pero el término extra se cancela con la semisuma de $f(0)$ y $F(N)$ y se obtiene

$$G_N = \sum_{n=-\infty}^{\infty} \int_0^N e\left(\frac{x^2}{N} + nx\right) dx = N \sum_{n=-\infty}^{\infty} e\left(-\frac{1}{4}Nn^2\right) \int_{n/2}^{1+n/2} e(Nx^2) dx$$

donde simplemente se ha hecho el cambio lineal que lleva $[0, N]$ en $[n/2, 1 + n/2]$. Separando en la suma los pares de los impares, se deduce

$$\frac{G_N}{N} = \sum_{n=-\infty}^{\infty} \int_n^{n+1} e(Nx^2) dx + i^{-N} \sum_{n=-\infty}^{\infty} \int_{n-1/2}^{n+1/2} e(Nx^2) dx = (1 + i^{-N}) \int_{-\infty}^{\infty} e(Nx^2) dx$$

donde estamos suponiendo que no hay problema con la convergencia, en particular que la integral existe (en el próximo capítulo estudiaremos esto con detalle). Aplicando (1.14) formalmente o tomando límites cuando $a = \epsilon - 2\pi iN$ y $\epsilon \rightarrow 0^+$, se sigue que el valor de la integral es $\sqrt{2} e(1/8)N^{-1/2} = N^{-1/2}/(1 - i)$. \square

La humilde fórmula (1.18) da más juego de lo que parece, al menos en manos de un genio como Euler. La idea es tan sencilla como pensar en qué obtendríamos al integrar muchas veces por partes. Definimos los *polinomios de Bernoulli* $\{B_m\}_{m=0}^{\infty}$ mediante la recurrencia

$$B'_m(x) = mB_{m-1}(x) \quad \text{con} \quad \int_0^1 B_m = 0 \quad \text{y} \quad B_0 = 1 \quad \text{para } m \in \mathbb{Z}^+.$$

Es decir, esencialmente integrar 1 repetidas veces haciendo mónicos los polinomios resultantes y ajustando la constante de integración para que sus promedios sean cero en $[0, 1]$. También se definen los *números de Bernoulli* como $b_n = B_n(0)$. Algunos ejemplos son:

$$B_0 = 1, \quad B_1 = x - \frac{1}{2}, \quad B_2 = x^2 - x + \frac{1}{6}, \quad B_3 = x^3 - \frac{3}{2}x^2 + \frac{1}{2}x, \quad B_4 = x^4 - 2x^3 + x^2 - \frac{1}{30}.$$

Integrando sucesivas veces (1.17), como $B_1(x) = x$, para $0 \leq x < 1$, se deduce que los polinomios de Bernoulli tienen una serie de Fourier bonita:

$$B_m(x) = -\frac{m!}{(2\pi i)^m} \sum_{n \neq 0} \frac{e(nx)}{n^m} \quad \text{para } m > 1.$$

En particular

(1.19)

$$b_{2m} = 2(-1)^{m+1} \frac{(2m)!}{(2\pi)^{2m}} \left(\frac{1}{1^{2m}} + \frac{1}{2^{2m}} + \frac{1}{3^{2m}} + \frac{1}{4^{2m}} + \dots \right) \quad \text{y} \quad b_{2m+1} = 0 \quad \text{para } m \in \mathbb{Z}^+.$$

Por tanto, en contra de lo que sugieren los primeros ejemplos, $\{|b_{2m}|\}$ es una sucesión que crece extremadamente deprisa.

Proposición 1.2.8 (Fórmula de sumación de Euler–Maclaurin). Sean $N, M, K \in \mathbb{Z}$ tales que $N < M$ y $K > 0$. Para cualquier $f \in C^K([N, M])$, la expresión

$$\sum_{n=N}^M f(n) - \int_N^M f - \frac{1}{2}(f(N) + f(M))$$

es igual a

$$\sum_{n=2}^K \frac{b_n}{n!} (f^{(n-1)}(M) - f^{(n-1)}(N)) + \frac{(-1)^{K+1}}{K!} \int_N^M f^{(K)}(x) B_K(x - [x]) dx.$$

Para $K = 1$ el último sumatorio es vacío y se recupera (1.18). Más interesantes son los casos con $K > 1$. Por ejemplo, si $f(x) = x^3$, $K = 4$ y $N = 1$, recuperamos la bella fórmula para la suma de los cubos:

$$\sum_{n=1}^M n^3 = \frac{1}{4}(M^4 - 1) + \frac{1}{2}(M^3 + 1) + \frac{1/6}{2!}(3M^2 - 3) + 0 = \frac{M^2(M+1)^2}{4}.$$

Véase en [Spi84, 707–708] una explicación “operacional” de la Proposición 1.2.8 para polinomios.

Demostración. La recurrencia implica $B_{K+1}(x - [x]) - b_{K+1} = (K+1) \int_N^x B_K(t - [t]) dt$ donde tomar $t - [t]$ no da ningún problema por la propiedad de promedio nulo. Integrando por partes, la última integral en el enunciado de la proposición es:

$$\frac{b_{K+1}}{K+1} (f^{(K)}(M) - f^{(K)}(N)) - \frac{1}{K+1} \int_N^M f^{(K+1)}(x) B_{K+1}(x - [x]) dx$$

y la prueba se sigue por inducción a partir del caso $K = 1$ que es (1.18). Nótese que $(-1)^{K+1} b_{K+1} = b_{K+1}$ porque para K par, $b_{K+1} = 0$. \square

1.2.3. Técnicas de variable compleja

Nuestros primeros paseos por la variable real nos desvelaron un mundo plagado de funciones monstruosas: continuas pero no diferenciables en ningún punto, C^∞ pero no *analíticas* (no desarrollables como su serie de Taylor), satisfaciendo el teorema de los valores intermedios pero no continuas.

En cambio, la variable compleja desde el principio volvía a poner las cosas en su sitio: las funciones holomorfas (derivables en un entorno de un punto) eran automáticamente C^∞ e incluso analíticas. Además había una serie de hechos milagrosos, basados de una forma u otra en que al integrar una función holomorfa a lo largo de una curva el resultado no depende de pequeñas variaciones con los extremos fijos.

Esto tiene una explicación razonable. Cuando consideramos una función holomorfa, $f'(z) = \lim_{h \rightarrow 0} (f(z+h) - f(z))/h$ establece unas relaciones entre las derivadas parciales de $u = \Re(f)$ y $v = \Im(f)$ como funciones de $x = \Re(z)$ e $y = \Im(z)$, llamadas *ecuaciones de Cauchy-Riemann*. Estas ecuaciones son equivalentes a decir que la forma diferencial $(u + vi) dx + (-v + ui) dy = f(z) dz$ es cerrada, esto es, que el “rotacional” que aparece en el teorema de Green se anula y entonces la integral de línea es siempre cero, el campo

es conservativo. Si una función f es holomorfa en un abierto salvo por una singularidad en cierto $z = s$, puede que ya no integre cero, pero todo vuelve a funcionar cambiando $f(z)$ por $f(z) - r/(z - s)$ donde $r = \text{Res}(f, s)$ es lo que se llama el *residuo* de f en s . Se prueba que esta función tiene una primitiva (a pesar de la posible singularidad el campo vuelve a ser conservativo).

Hay muchos libros en los que aprender los rudimentos de la variable compleja. Un clásico, escrito por el primer o segundo medallista Fields, es [Ahl78]. Más clásica aún es una famosa carta de Gauss de 1811 al astrónomo F. Bessel (reproducida en [Koc91, §8]).

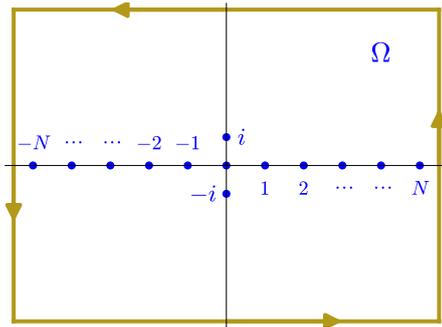
El teorema de los residuos es un resultado maravilloso que permite expresar una integral como una suma, lo cual puede ser útil tanto para hallar integrales como para calcular sumas.

Teorema 1.2.9 (Teorema de los residuos). *Sea $\Omega \subset \mathbb{C}$ una región compacta y simplemente conexa cuya frontera viene dada por una curva rectificable γ orientada positivamente (en contra de las agujas del reloj). Sea f una función holomorfa en Ω salvo por un número finito de singularidades en $z = s_1, s_2, \dots, s_N$ en el interior de Ω , entonces*

$$\int_{\gamma} f(z) dz = 2\pi i \sum_{n=1}^N \text{Res}(f, s_n)$$

donde $\text{Res}(f, s_n)$ es el residuo de f en $z = s_n$.

La función $g(z) = \pi/\text{sen}(\pi z)$ tiene polos en $n \in \mathbb{Z}$ y $\text{Res}(g, n) = (-1)^n$, por ello es una buena máquina sumadora de series alternantes. Tomemos por ejemplo $\Omega = \{z : |\Re(z)| < N + \frac{1}{2}, |\Im(z)| < N + \frac{1}{2}\}$ con $N \in \mathbb{Z}^+$ y $f(z) = (z^2 + 1)^{-1}g(z)$. Las singularidades son polos simples en $z = \pm i$ y en todos los enteros de $[-N, N]$. Los residuos son:



$$\begin{aligned} \text{Res}(f, i) &= \frac{\pi}{\text{sen}(\pi i)} \lim_{z \rightarrow i} \frac{z - i}{z^2 + 1} = \frac{\pi}{e^{-\pi} - e^{\pi}} \\ \text{Res}(f, -i) &= \frac{-\pi}{\text{sen}(\pi i)} \lim_{z \rightarrow -i} \frac{z + i}{z^2 + 1} = \frac{\pi}{e^{-\pi} - e^{\pi}} \\ \text{Res}(f, n) &= \frac{\text{Res}(\pi/\text{sen}(\pi z), n)}{n^2 + 1} = \frac{(-1)^n}{n^2 + 1} \\ \text{entonces } \frac{1}{2\pi i} \int_{\gamma} f &= \frac{2\pi}{e^{-\pi} - e^{\pi}} + \sum_{n=-N}^N \frac{(-1)^n}{n^2 + 1} \end{aligned}$$

Si $N \rightarrow \infty$, f en la frontera de Ω es $O(N^{-2})$ y por tanto $\int_{\gamma} f(z) dz = O(N^{-1})$. en el límite se tiene la evaluación limpia

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n^2 + 1} = \frac{\pi}{e^{\pi} - e^{-\pi}} - \frac{1}{2}.$$

Otra joya, menos básica, de la variable compleja es la teoría de *funciones de orden finito* desarrollada por J. Hadamard (y empleada en las primeras demostraciones del teorema de los números primos) que tiene sus orígenes en un argumento dudoso de Euler.

Lo que hizo Euler esencialmente es notar que para cualquier polinomio $a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$, $a_0 \neq 0$, la suma de los cuadrados de los inversos de sus ceros es $(a_1^2 - 2a_0 a_2)/a_0^2$ (esto es un ejercicio asequible), entonces pensando que $x^{-1} \operatorname{sen} x = 1 - x^2/3! + x^4/5! - \dots$ es un “polinomio infinito” con ceros $x = n\pi$, $n \in \mathbb{Z} - \{0\}$, se tendría

$$\frac{(-2) \cdot 1 \cdot (-1/3!)}{1^2} = \sum_{n \neq 0} \frac{1}{(n\pi)^2} \quad \text{o equivalentemente} \quad \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

La fórmula final es cierta pero el razonamiento es muy dudoso. Euler llegó a probar que $x^{-1} \operatorname{sen} x = \prod_{n=1}^{\infty} (1 - x^2/(n\pi)^2)$, es decir que factoriza como un polinomio, y esta factorización es suficiente para dar un argumento válido. Hadamard demostró que todas las *funciones enteras* (holomorfas en todo \mathbb{C}) de orden $\lambda < \infty$, es decir, satisfaciendo $f(z) = O(e^{C|z|^\lambda})$ para cierta constante C , se factorizan en términos de sus ceros como un polinomio salvo un factor exponencial y salvo correcciones de la convergencia. Concretamente, si denotamos los ceros mediante z_n y llamamos k_0 al orden del posible cero en $z = 0$, se tiene

$$(1.20) \quad f(z) = z^{k_0} e^{g(z)} \prod_n \left(1 - \frac{z}{z_n}\right) e^{z/z_n + \frac{1}{2}(z/z_n)^2 + \dots + \frac{1}{[\lambda]}(z/z_n)^\lambda}$$

donde $g(z)$ es un polinomio de grado a lo más $[\lambda]$. El único propósito de las exponenciales en el producto es asegurar la convergencia. Si no se ha elegido λ óptimo o si hay simetrías especiales, se pueden agrupar en parte con $g(z)$ o simplificar. Por ejemplo, para $f(z) = \operatorname{sen} z$, $\lambda = 1$ y se deduce

$$\operatorname{sen} z = z e^{a+bz} \prod_{n \neq 0} \left(1 - \frac{z}{\pi n}\right) e^{z/(\pi n)} \quad \text{que implica} \quad \operatorname{sen} z = z e^{a+bz} \prod_{n=1}^{\infty} \left(1 - \frac{z^2}{(\pi n)^2}\right).$$

Usando que $z^{-1} \operatorname{sen} z$ es par y que tiene límite 1 cuando $z \rightarrow 0$ se deduce $a = b = 0$, es decir, el producto hallado por Euler. Eligiendo $z = \pi/2$ se tiene la curiosa *fórmula de Wallis*

$$(1.21) \quad \frac{\pi}{2} = \prod_{n=1}^{\infty} \frac{(2n)^2}{(2n-1)(2n+1)} = \frac{2 \cdot 2 \cdot 4 \cdot 4 \cdot 6 \cdot 6 \dots}{1 \cdot 3 \cdot 3 \cdot 5 \cdot 5 \cdot 7 \dots}.$$

La *extensión analítica* es una estrategia que permite sumar cosas que no se pueden sumar. ¿Y para qué deseamos hacer tal cosa? Por ejemplo para aplicar el teorema de

los residuos en regiones más amplias. La clave es el teorema de unicidad, que asegura que si dos funciones holomorfas en una región conexa coinciden en un conjunto con un punto de acumulación, entonces coinciden en todos los puntos. Si hay una extensión holomorfa tiene que ser única, lo cual da una manera unívoca de interpretar algunas series divergentes.

Una de las extensiones analíticas más conocidas y fructíferas es la de una función fundamental para estudiar la distribución de los primos, la *función ζ de Riemann*:

$$(1.22) \quad \zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} \quad \text{para } \Re(s) > 1.$$

Series como $\sum n^{-1/2}$ o $\sum n^{-3/4+i}$ no convergen (para la segunda se puede usar que n^i apenas oscila en intervalos diádicos) lo que sugiere que es imposible dar sentido a $\zeta(1/2)$ o a $\zeta(3/4 - i)$. Sin embargo, por (1.18) (Proposición 1.2.8 para $K = 1$) con $f(x) = x^{-s}$,

$$(1.23) \quad \zeta(s) = \frac{s}{s-1} + \frac{1}{2} - s \int_1^{\infty} \{x\} x^{-s-1} dx$$

y como $\{x\}$ está acotado, la integral converge si $\Re(s) > 0$. Se deduce entonces que $\zeta(s) - 1/(s-1)$ se extiende a una función holomorfa en todo el semiplano derecho. Aplicando la Proposición 1.2.8 en general (integrando (1.23) por partes con cierto cuidado), x^{-s-1} pasa a ser x^{-s-K} y como K es arbitrario, la conclusión es que $\zeta(s)$ tiene una extensión analítica (holomorfa) a todo \mathbb{C} salvo por un polo simple de residuo 1 en $s = 1$. De alguna forma estamos dando sentido a $\sum n^{-s}$ siempre que $s \neq 1$. Pensemos ahora en la fórmula elemental

$$(1 - 2^{1-s})\zeta(s) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^s}$$

que da sentido a $\zeta(s)$ en $\Re(s) > 0$, siempre que $2^{1-s} \neq 1$. Por la unicidad de la extensión analítica, esta fórmula debe definir por ejemplo $\zeta(1/2)$, formalmente $\sum n^{-1/2}$, de manera compatible con (1.23). A partir de ahora llamaremos función ζ a la única extensión analítica de (1.22).

Otro ejemplo destacable de extensión analítica es la de la *función Γ* dada por

$$(1.24) \quad \Gamma(s) = \int_0^{\infty} x^{s-1} e^{-x} dx \quad \text{para } \Re(s) > 0.$$

En cierto modo, Γ generaliza los factoriales para valores complejos porque es fácil probar la ecuación funcional $\Gamma(s+1) = s\Gamma(s)$ y $\Gamma(1) = 1$, de donde $\Gamma(n) = (n-1)!$ para $n \in \mathbb{Z}^+$. Esta ecuación funcional, en la forma $\Gamma(s-1) = \Gamma(s)/(s-1)$ muestra que $\Gamma(s)$ tiene una extensión analítica a $\mathbb{C} - (\mathbb{Z}^- \cup \{0\})$. Además en cada entero no positivo hay un polo

simple. Para un estudio más profundo de la función Γ se necesita una ecuación funcional menos obvia dada por

$$(1.25) \quad \Gamma(s)\Gamma(1-s) = \frac{\pi}{\operatorname{sen}(\pi s)} \quad \text{para } s \in \mathbb{C} - \mathbb{Z}.$$

Inmediatamente, esta relación da una extensión analítica de $\Gamma(s)$ que debe ser coherente con la anterior. Se deduce también que no tiene ceros y por consiguiente $F(s) = 1/\Gamma(s)$ se extiende a una función entera. A partir de (1.24) se tiene que $\Gamma(s) = O(e^{|s|^\lambda})$ en $\Re(s) \geq 1/2$ para todo $\lambda > 1$. De (1.25) se deduce que esta misma acotación es válida para $F(s)$ en $\Re(s) \leq 1/2$ y la primera forma de la ecuación funcional extiende la cota a todo $s \in \mathbb{C}$. Los ceros de F están en los enteros negativos y según (1.20), tenemos (entendiendo $1/\infty = 0$)

$$(1.26) \quad \frac{1}{\Gamma(s)} = se^{\gamma s} \prod_{n=1}^{\infty} \left(1 + \frac{s}{n}\right) e^{-s/n} \quad \text{para } s \in \mathbb{C},$$

donde γ es una constante, la constante de Euler. Aunque su definición habitual es el límite de $\sum_{n=1}^N n^{-1} - \log N$, se puede redefinir como $\Gamma'(1)$. Para probar (1.25), expresamos el primer miembro como una sola integral para $0 < \Re(s) < 1$.

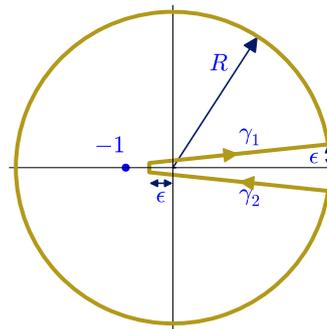
$$\Gamma(s)\Gamma(1-s) = \int_0^\infty \int_0^\infty (x/y)^s x^{-1} e^{-x-y} dy dx \stackrel{y=xt}{=} \int_0^\infty \frac{t^{-s}}{t+1} dt.$$

Ahora vamos a calcular este integral mediante el teorema de los residuos. Tomamos $f(z) = z^{-s}/(z+1)$ donde z^{-s} se define como $e^{-s \log z}$ pero esto da un problema porque $\log z$ es multivaluada, si nos aferramos a $\log z = \log |z| + i \arg(z)$ entonces $\log(a+ei) - \log(a-ei) = 2\pi i$ si $\epsilon \rightarrow 0^+$ y $a \in \mathbb{R}^+$. Por ello escogemos para llegar a cabo la integración, un dominio de tipo *pacman* que evita el eje real positivo:

$$\operatorname{Res}(f, -1) = e^{\pi i s}$$

$$\lim_{\epsilon \rightarrow 0^+} \int_{\gamma_1} f = \int_0^\infty \frac{t^{-s}}{t+1} dt$$

$$\lim_{\epsilon \rightarrow 0^+} \int_{\gamma_2} f = -e^{-2\pi i s} \int_0^\infty \frac{t^{-s}}{t+1} dt$$



Cuando $R \rightarrow \infty$ la contribución de la parte circular de la frontera tiende a cero y el teorema de los residuos nos da para $\epsilon \rightarrow 0^+$ que la integral que queremos evaluar es $2\pi i e^{-\pi i s} / (1 - e^{-2\pi i s})$, que coincide con el segundo miembro de (1.25).

Tomando logaritmos en (1.26) y empleando la Proposición 1.2.8 con $K = 2$ para estimar la parte significativa de la suma infinita, se deduce con algún esfuerzo la *fórmula de Stirling*

$$(1.27) \quad \log \Gamma(s) = \left(s - \frac{1}{2}\right) \log s - s + \frac{1}{2} \log(2\pi) + O(|s|^{-1}) \quad \text{para } |\arg(s)| < \pi - \epsilon,$$

En realidad este esquema no da directamente el valor explícito de la constante $\frac{1}{2} \log(2\pi)$, que se puede evaluar empleando que (1.21) proporciona la asintótica de $\Gamma(2n)/\Gamma^2(n)$ (cf. [Spi84, 711–713]). Hay una prueba completa en [Ahl78, §5.2.5] y otra elemental para $x \in \mathbb{R}^+$ en [Mic08].

Para terminar veamos que $\zeta(s)$ también satisface una ecuación funcional.

Proposición 1.2.10. *La función ζ verifica la ecuación funcional*

$$\pi^{-s/2} \Gamma(s/2) \zeta(s) = \pi^{-(1-s)/2} \Gamma((1-s)/2) \zeta(1-s).$$

Además si multiplicamos ambos miembros por $s(s-1)$ el resultado es una función entera.

Demostración. Seguimos aquí la prueba de la famosa memoria de B. Riemann, que con un sencillo cambio de variable en (1.24) parte de

$$\pi^{-s/2} \Gamma(s/2) \sum_{n=1}^{\infty} n^{-s} = \sum_{n=1}^{\infty} \int_0^{\infty} t^{s/2-1} e^{-\pi n^2 t} dt \quad \text{para } \Re s > 1,$$

o equivalentemente, con la notación de (1.16)

$$\pi^{-s/2} \Gamma(s/2) \sum_{n=1}^{\infty} n^{-s} = \frac{1}{2} \int_0^{\infty} t^{s/2-1} (\theta(it/2) - 1) dt.$$

Separando el rango de integración en $[0, 1] \cup [1, \infty)$ y utilizando la Proposición 1.2.5, el segundo miembro es

$$\frac{1}{2} \int_0^1 t^{s/2-1} (t^{-1/2} \theta(i/2t) - 1) dt + \frac{1}{2} \int_1^{\infty} t^{s/2-1} (\theta(it/2) - 1) dt.$$

Con el cambio $t \mapsto 1/t$ en la primera integral se llega a que para $\Re s > 1$

$$\pi^{-s/2} \Gamma(s/2) \zeta(s) = \frac{1}{s(s-1)} + \frac{1}{2} \int_1^{\infty} (t^{s/2-1} + t^{-1/2-s/2}) (\theta(it/2) - 1) dt.$$

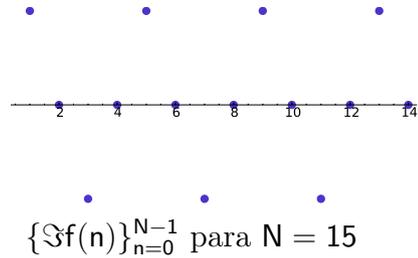
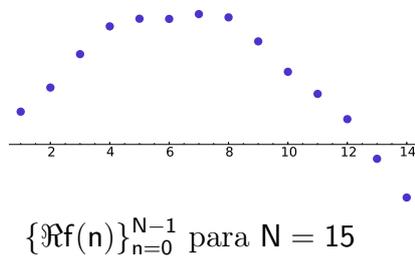
Al multiplicar por $s(s-1)$ el segundo miembro es una función entera e invariante por $s \mapsto 1-s$. Por la unicidad de la extensión analítica, el primer miembro debe tener también tal invariancia. \square

La importancia de esta prueba radica en que se puede “copiar” para demostrar que las funciones ζ que provienen de formas modulares satisfacen una ecuación funcional.

1.3. Análisis armónico discreto

En esta sección “discreto” se refiere más bien a finito, el tema es el análisis armónico en grupos finitos. Aquí nos ocuparemos sólo del caso abeliano, dando un papel protagonista a \mathbb{Z}_N , el grupo aditivo de los enteros módulo N . No obstante, el caso no abeliano tiene una rica teoría y aplicaciones insospechadas [Ter99], [DSV03].

Si consideramos \mathbb{Z}_N como conjunto, son simplemente N elementos y $L^2(\mathbb{Z}_N)$ no es otra cosa que los vectores de \mathbb{C}^N con la norma habitual que viene inducida por el producto escalar $\langle \vec{x}, \vec{y} \rangle = \sum \bar{x}_n y_n$. Intuitivamente es conveniente pensar en la parte real y la parte imaginaria de $f \in L^2(\mathbb{Z}_N)$ como *señales* que en cada tiempo discreto tienen un valor. Por ejemplo, si $f(n) = \text{sen}(n/4) + i^n/20$



Esta idea intuitiva no es gratuita. El análisis armónico discreto tiene sus aplicaciones en el tratamiento de señales digitales. La digitalización en la práctica se produce en “tiempo” y se representa en que la variable ahora pertenece al conjunto finito \mathbb{Z}_N , pero también se produce en la propia señal, en la imagen de la función, en el sentido de que el resultado vendrá dado por un número determinado de bytes. Para el modelo matemático es conveniente despreocuparse en primera instancia de esta segunda digitalización y pensar que las imágenes están en \mathbb{R} o en \mathbb{C} .

Un ejemplo más visual son las imágenes digitales, las cuales están compuestas de *pixels*. Cada uno de ellos tiene habitualmente tres valores (canales) correspondientes a los niveles que contiene de rojo, verde y azul (en ciertas situaciones hay un cuarto correspondiente a la transparencia). Éstos son los tres colores básicos en el sentido de que la fisiología de nuestra visión hace que percibamos el resto de estos colores como combinación de ellos. Entonces cada imagen digital de $N \times M$ pixels se puede entender como tres funciones f_R, f_G y f_B de $\mathbb{Z}_N \times \mathbb{Z}_M$ en \mathbb{R} que dan los niveles de estos colores.



$f_R : \mathbb{Z}_{200} \times \mathbb{Z}_{157} \rightarrow \mathbb{R}$



$f_G : \mathbb{Z}_{200} \times \mathbb{Z}_{157} \rightarrow \mathbb{R}$



$f_B : \mathbb{Z}_{200} \times \mathbb{Z}_{157} \rightarrow \mathbb{R}$

1.3.1. Caracteres y transformadas

Una pregunta natural tras la introducción anterior es dónde participa la estructura de grupo. La respuesta es que en la propia construcción de los armónicos. Cuando estudiamos el análisis de Fourier clásico en \mathbb{T} y en \mathbb{R} , en algún momento fue importante que las exponenciales cumplieran $e(x+y) = e(x)e(y)$, es decir, que aplicaran la operación de grupo de \mathbb{T} o de \mathbb{R} , la suma, en la de $\mathbb{C} - \{0\}$, que es la multiplicación. En \mathbb{Z}_N se muestra natural considerar los armónicos $e_j(n) = e(jn/N)$ que tienen una propiedad similar: $e_j(n+m) = e_j(n)e_j(m)$.

Recordemos que podíamos basar el análisis de Fourier básico en las relaciones (1.4) y (1.8). El problema es que caían bajo sospecha por la falta de convergencia. En el mundo finito estos problemas se desvanecen y la relación análoga es sencillamente:

$$(1.28) \quad \delta_D(n) = \sum_{j=0}^{N-1} e(jn/N) \quad \text{donde} \quad \delta_D(n) = \begin{cases} N & \text{si } N|n \\ 0 & \text{en otro caso.} \end{cases}$$

La prueba más pedestre se reduce a sumar la serie geométrica. El segundo miembro es la “delta de Dirac” en \mathbb{Z}_N porque al integrarla con respecto a la medida de contar normalizada (sumar y dividir por N) el resultado es uno o cero dependiendo de si el conjunto de integración contiene a $\bar{0}$ o no.

La *transformada de Fourier discreta* en \mathbb{Z}_N es el análogo de los coeficientes de Fourier y se define como

$$(1.29) \quad \hat{x}_n = \sum_{m=0}^{N-1} x_m e_n(-m) \quad \text{con} \quad e_n(m) = e(nm/N)$$

para cualquier vector $\vec{x} \in \mathbb{C}^N$ que, como hemos visto, es lo mismo que una función $f: \mathbb{Z}_N \rightarrow \mathbb{C}$. Aunque aquí lo evitaremos, es muy habitual designar a la transformada de Fourier discreta mediante sus siglas en inglés DFT.

La relación (1.28) da inmediatamente un desarrollo de Fourier en armónicos (o si lo preferimos, una fórmula de inversión)

$$(1.30) \quad x_n = \frac{1}{N} \sum_{m=0}^{N-1} \hat{x}_m e_m(n).$$

Hay también una identidad de Parseval y buenas propiedades de la convolución que introduciremos en un contexto más general en el que se muestren menos casuales.

Un *carácter* de un grupo abeliano G (no necesariamente finito) es un *homomorfismo* (función que conserva la ley de grupo) $\chi: G \rightarrow S^1$ donde S^1 denota el grupo multiplicativo de números complejos de módulo 1.

Los caracteres forman a su vez un grupo con la multiplicación $(\chi_1\chi_2)(g) = \chi_1(g)\chi_2(g)$ que suele denotarse mediante \widehat{G} . No es difícil comprobar que $\widehat{\mathbb{T}} = \{e(nx)\}_{n=-\infty}^{\infty}$, $\widehat{\mathbb{R}} = \{e(\xi x)\}_{\xi \in \mathbb{R}}$ y $\widehat{\mathbb{Z}_N} = \{e(jn/N)\}_{j=0}^{N-1}$. Por tanto los caracteres generalizan todos los armónicos que hemos visto hasta ahora. Es posible hacer una teoría general en grupos abelianos localmente compactos [Dei05], [Kat76] en la que las transformadas de Fourier continuas y discretas y los coeficientes de Fourier son diferentes aspectos de un mismo fenómeno: las funciones de $L^2(G)$ se expresan como integrales sobre \widehat{G} con respecto a cierta medida. Incluso si nos limitamos al ámbito de la aritmética, esto no es una generalización innecesaria. El análisis armónico en ciertos grupos (también con estructura de anillo) llamados *adeles* [Rob74] asociados a cuerpos de números, permite deducir análogos muy profundos de la Proposición 1.2.10 a partir de una fórmula de sumación de Poisson que incorpora propiedades acerca de la factorización en ideales primos. Genéricamente se llama a este círculo de ideas *tesis de Tate* [RV99] pues fueron introducidas por J.T. Tate en su tesis doctoral.

Por otro lado, en el caso de grupos no abelianos, la definición anterior de caracteres es demasiado restrictiva y hay que considerar en su lugar *representaciones*, una especie de caracteres matriciales, mientras que se llama caracteres a sus trazas [Ter99], [Art00], [Kna96a], [Kna96b].

Volviendo a nuestro caso, dado G abeliano y finito, por el teorema de clasificación debe existir un isomorfismo $\phi : G \rightarrow \mathbb{Z}_{N_1} \oplus \mathbb{Z}_{N_2} \oplus \cdots \oplus \mathbb{Z}_{N_r}$ y es fácil ver que el grupo de caracteres es

$$(1.31) \quad \widehat{G} = \left\{ \chi : \chi(n) = \prod_{k=1}^r e(j_k \phi(n)/N_k) \quad \text{con } 0 \leq j_k < N_k \right\}.$$

En otras palabras, salvo componer con el isomorfismo, son productos de caracteres de los \mathbb{Z}_{N_k} . Los grupos G y \widehat{G} son duales en el sentido de que los caracteres de \widehat{G} son $\{\chi \mapsto \chi(g)\}_{g \in G}$. Este hecho se extiende al caso infinito y en esta generalidad se llama *teorema de dualidad de Pontryagin*.

El siguiente lema es heredero de (1.28) y admite una demostración directa similar. Vamos, sin embargo, a deducirlo básicamente de las propiedades de grupo.

Lema 1.3.1 (relaciones de ortogonalidad). *Sea G un grupo abeliano finito y $\chi_1, \chi_2 \in \widehat{G}$, entonces*

$$\sum_{g \in G} \overline{\chi_1(g)} \chi_2(g) = \begin{cases} |G| & \text{si } \chi_1 = \chi_2 \\ 0 & \text{si } \chi_1 \neq \chi_2 \end{cases} \quad y \quad \sum_{\chi \in \widehat{G}} \chi(g) = \begin{cases} |G| & \text{si } g = e \\ 0 & \text{si } g \neq e \end{cases}$$

donde $e \in G$ es elemento neutro.

Demostración. Se tiene $\overline{\chi_1(g)}\chi_2(g) = \chi(g)$ con $\chi = \chi_1^{-1}\chi_2$ porque $\chi_1(g) \in S^1$. Si $\chi_1 = \chi_2$ el resultado es trivial y si $\chi_1 \neq \chi_2$ existe $h \in G$ con $\chi(h) \neq 1$, en este caso

$$\sum_{g \in G} \overline{\chi_1(g)}\chi_2(g) = \sum_{g \in G} \chi(g) = \sum_{g \in G} \chi(gh) = \chi(h) \sum_{g \in G} \overline{\chi_1(g)}\chi_2(g)$$

y entonces la suma tiene que ser cero. La relación recién probada cambiando G por \widehat{G} y tomando como χ_1 el elemento neutro, da la segunda fórmula del enunciado. \square

En este contexto se define la *transformada de Fourier discreta* para cada $\chi \in \widehat{G}$ como

$$(1.32) \quad \widehat{f}(\chi) = \sum_{g \in G} f(g)\overline{\chi(g)}.$$

Nótese que para $G = \mathbb{Z}_N$ esto es lo mismo que la definición (1.29) salvo el abuso de notación de representar el carácter por su índice en el argumento de la transformada.

El producto escalar en $L^2(G)$ y la *convolución* vienen definidos respectivamente por

$$\langle f_1, f_2 \rangle = \sum_{g \in G} \overline{f_1(g)}f_2(g) \quad \text{y} \quad (f_1 * f_2)(g) = \sum_{h \in G} f_1(h)f_2(gh^{-1}).$$

Con estas definiciones y el Lema 1.3.1 se prueban en un instante las propiedades cuyo análogo conocíamos en \mathbb{T} y \mathbb{R} . También hay una fórmula de sumación de Poisson [Ter99, §12], aunque no la veremos aquí.

Teorema 1.3.2. *Para G abeliano y finito y $f, f_1, f_2 : G \rightarrow \mathbb{C}$, se verifican la fórmula de inversión*

$$f(g) = \frac{1}{|G|} \sum_{\chi \in \widehat{G}} \widehat{f}(\chi)\chi(g),$$

la identidad de Parseval

$$\langle f, f \rangle = \frac{1}{|G|} \langle \widehat{f}, \widehat{f} \rangle, \quad \text{en general} \quad \langle f_1, f_2 \rangle = \frac{1}{|G|} \langle \widehat{f}_1, \widehat{f}_2 \rangle,$$

y además la propiedad de la convolución

$$\widehat{f_1 * f_2}(\chi) = \widehat{f}_1(\chi)\widehat{f}_2(\chi).$$

Por ejemplo, los elementos invertibles (con la multiplicación) módulo 10, $\mathcal{U}(\mathbb{Z}_{10}) = \{\overline{1}, \overline{3}, \overline{7}, \overline{9}\}$, forman un grupo isomorfo a \mathbb{Z}_4 con el isomorfismo $\phi : \mathcal{U}(\mathbb{Z}_{10}) \rightarrow \mathbb{Z}_4$, dado por $\phi(\overline{1}) = \overline{0}$, $\phi(\overline{3}) = \overline{1}$, $\phi(\overline{7}) = \overline{3}$, $\phi(\overline{9}) = \overline{2}$, que se dice que es un *logaritmo discreto* porque pasa multiplicaciones en sumas. Según el Teorema 1.3.2, los armónicos vienen dados por (1.31) y son

$$\{\chi_0, \chi_1, \chi_2, \chi_3\} \quad \text{con} \quad \chi_j(n) = e(j\phi(n)/4).$$

Todo esto parece demasiado fácil como para ser interesante, pero veremos a continuación que sirve de manera fundamental para resolver un problema difícil en teoría de números.

Empecemos por otro más sencillo. Todos sabemos que hay infinitos primos. La prueba de Euler parte de la importantísima *fórmula producto de Euler*

$$(1.33) \quad \boxed{\zeta(s) = \prod_{p \text{ primo}} (1 - p^{-s})^{-1}} \quad \text{para } \Re(s) > 1,$$

que se sigue de $(1 - p^{-s})^{-1} = 1 + p^{-s} + (p^2)^{-s} + (p^3)^{-s} + \dots$ aplicando el teorema fundamental de la aritmética. Si $s \rightarrow 1^+$, $\zeta(s) \rightarrow +\infty$ y por tanto el producto no puede ser finito.

El problema difícil es probar que hay infinitos primos que acaban por ejemplo en 1. Uno estaría tentado de escribir un análogo de (1.33) con $\sum (10n + 1)^{-s}$ en lugar de $\zeta(s)$ pero tal cosa no funciona porque aunque un número entero acabe en 1, sus factores primos pueden no hacerlo, por ejemplo $221 = 13 \cdot 17$. Extendamos nuestros armónicos a \mathbb{Z} definiendo $\chi_j(n)$ como $\chi_j(\bar{n})$ si $\bar{n} \in \mathcal{U}(\mathbb{Z}_{10})$, esto es si n y 10 son coprimos, y $\chi_j(n) = 0$ en otro caso. En general, a este tipo de extensiones a \mathbb{Z} de los armónicos de $\mathcal{U}(\mathbb{Z}_N)$, se les llama *caracteres de Dirichlet*. Con ello los armónicos se vuelven multiplicativos en \mathbb{Z} . Dada una factorización en primos, $n = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_k^{\alpha_k}$ se tiene $\chi_j(n) = \chi_j^{\alpha_1}(p_1) \chi_j^{\alpha_2}(p_2) \dots \chi_j^{\alpha_k}(p_k)$ y la prueba de (1.33) se repite exactamente para obtener

$$L(s, \chi_j) = \prod_{p \text{ primo}} (1 - \chi_j(p) p^{-s})^{-1} \quad \text{para } \Re(s) > 1, \quad \text{donde } L(s, \chi_j) = \sum_{n=1}^{\infty} \frac{\chi_j(n)}{n^s}.$$

Se dice que $L(s, \chi_j)$ es una *función L de Dirichlet*. Utilizando que los armónicos representan cualquier función, seleccionaremos los primos que acaban en 1 y todo el resultado volverá a reducirse a qué ocurre con $L(s, \chi_j)$ cuando $s \rightarrow 1^+$, lo cual en este caso es algo puramente computacional. Los detalles están en la demostración siguiente.

Proposición 1.3.3. *Hay infinitos primos que acaban en 1. En general hay infinitos primos que acaban en cualquier cifra impar distinta de 5.*

Demostración. Tomando logaritmos y usando el desarrollo de Taylor de $-\log(1 - x)$, se tiene

$$\log L(s, \chi_j) = \sum_{p \text{ primo}} \sum_{k=1}^{\infty} \frac{\chi_j^k(p)}{k p^s} = \sum_{p \text{ primo}} \frac{\chi_j(p)}{p^s} + O(1) \quad \text{para } s > 1,$$

donde simplemente se ha empleado que $\sum p^{-k} < \sum_{n=2}^{\infty} n^{-k} \ll 2^{-k}$. Ahora analizamos la “delta de Dirac” en $\bar{1}$ con los armónicos χ_j (Teorema 1.3.2 o Lema 1.3.1)

$$\chi_0(n) + \chi_1(n) + \chi_2(n) + \chi_3(n) = \begin{cases} 4 & \text{si } n \equiv 1 \pmod{10} \\ 0 & \text{en otro caso.} \end{cases}$$

Así pues

$$(1.34) \quad \log(L(s, \chi_0)L(s, \chi_1)L(s, \chi_2)L(s, \chi_3)) = 4 \sum_{p \equiv 1 \pmod{10}} \frac{1}{p^s} + O(1) \quad \text{para } s > 1.$$

Como $\chi_0(n) = 1$ si n y 10 son coprimos y 0 en el resto, $L(s, \chi_0) > \sum (10n + 1)^{-s} \rightarrow \infty$ cuando $s \rightarrow 1^+$. Por otro lado, tomando 10 términos en las sumas que definen $L(1, \chi_2)$ y $L(1, \chi_1) = \overline{L(1, \chi_3)}$, se tiene

$$L(1, \chi_2) = 0,63 + \epsilon_1 \quad \text{y} \quad \Re(L(1, \chi_3)) = 0,88 + \epsilon_2$$

con $|\epsilon_1|, |\epsilon_2| < 1/10$. Esta cota proviene de que las series son alternadas y el error es menor que el primer término despreciado [Spi84, p.600]. Por tanto $L(1, \chi_1)L(1, \chi_2)L(1, \chi_3)$ es un número real positivo y de (1.34) con $s \rightarrow 1^+$ se deduce que hay infinitos primos que acaban en 1. La prueba es similar para otra cifra admisible k , analizando la “delta de Dirac” en \bar{k} . \square

La generalización natural de este resultado es conocida:

Teorema 1.3.4 (Teorema de Dirichlet). *Si $q \in \mathbb{Z}^+$ y $a \in \mathbb{Z}$ son coprimos, entonces la progresión aritmética $\{qn + a\}_{n=1}^{\infty}$ contiene infinitos primos.*

La demostración anterior permite reducir el resultado a un trabajo computacional para cada q concreto, pero el enunciado general exige que ninguna de las funciones L de Dirichlet tenga un cero en $s = 1$. La prueba de ello es artificiosa (aunque no larga) y puede consultarse por ejemplo en [SS03] o [Dav80].

1.3.2. Implementación del análisis de Fourier discreto

En el mundo real hay transformadas de Fourier por todas partes, sin embargo el mundo virtual es digital y debe reemplazarlas por transformadas de Fourier discretas. Un primer problema es cómo pasar del caso continuo al discreto, en cierto modo (aunque no exactamente), de analógico a digital. Otro problema de índole más técnica es cómo llevar a cabo los cálculos con transformadas discretas de manera eficiente.

En la práctica, nuestros sentidos o nuestros instrumentos de medida sólo operan en un rango de frecuencias. Sería por ejemplo absurdo comprar un altavoz para nuestra cadena de música que reprodujera frecuencias por encima de los 20000 hercios (20000 oscilaciones por segundo), ya que tales sonidos son inaudibles. Por otro lado, es mejor que el mando a distancia de nuestra televisión tenga una frecuencia de funcionamiento por debajo de la de la luz visible para que no se vea afectado cuando encendamos una lámpara.

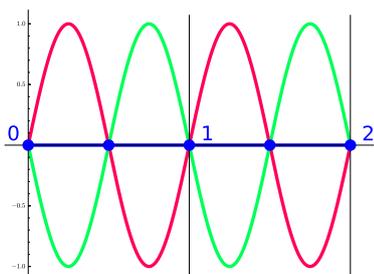
De esta forma, en la práctica es natural considerar que las funciones que representan señales tienen un contenido en frecuencias limitado. Para el caso de una onda periódica,

esto es lo mismo que decir que la serie de Fourier es una suma finita, un polinomio trigonométrico, mientras que para una señal genérica, la hipótesis natural es que sea una *función de banda limitada*, lo que significa que satisface para cierto B

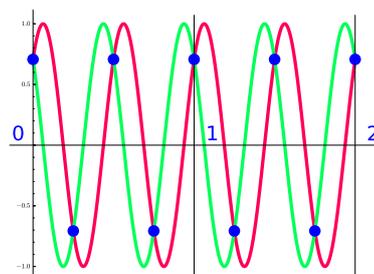
$$(1.35) \quad \widehat{f}(\xi) = 0 \quad \text{cuando } |\xi| \geq B.$$

Recordemos que, a través de (1.7), el valor de $\widehat{f}(\xi)$ indica de alguna forma el “coeficiente” con el que aparece el armónico $e(\xi x)$ de frecuencia ξ al analizar f . Si a uno no le dan mucho miedo las deltas de Dirac, de hecho esta definición recoge también el caso periódico. Dicho sea de paso, el *teorema de Paley-Wiener* describe en cierto modo las funciones de banda limitada como funciones de variable compleja.

Para transformar una señal en un conjunto discreto de datos, digamos que tomamos una muestra (un valor de la función) cada T segundos. En esta situación se dice que la *frecuencia de muestreo* es $\nu_s = 1/T$. Matemáticamente lo que se tiene es una función f y una sucesión $\{f(nT)\}_{n=-\infty}^{\infty}$. Está claro que este conjunto discreto no permite caracterizar la función en general:



$T = 1/2, \nu_s = 2$ no distingue
 $0, \text{sen}(2\pi x)$ y $-\text{sen}(2\pi x)$



$T = 1/4, \nu_s = 4$ no distingue
 $\text{sen}(4\pi x + \frac{\pi}{4})$ y $\cos(4\pi x + \frac{\pi}{4})$

Pensemos en una onda 1-periódica f con frecuencias limitadas al rango $(-B, B)$ y supongamos para simplificar que $B \in \mathbb{Z}^+$. Esta f debe ser un polinomio trigonométrico con a lo más $2B - 1$ coeficientes no nulos. Si $\nu_s \geq 2B$, entonces habrá $2B$ muestras en cada periodo y eso es suficiente para recuperar todos los coeficientes resolviendo un sistema lineal. Lo que dice el *teorema de muestreo de Shannon* es que tal cosa también se cumple para señales no periódicas, de hecho hay una forma fácil de recuperar la señal a partir de la muestra. Por simplicidad, omitiremos las consideraciones con respecto a la regularidad de f en el siguiente resultado, aunque está claro que con $f \in L^1$ continua y $\widehat{f} \in C^2$, se asegura la convergencia absoluta de las series que aparecen en la prueba.

Teorema 1.3.5 (Teorema de muestreo de Shannon). *Sea f satisfaciendo (1.35) con $2B \leq \nu_s$. Entonces*

$$f(x) = \sum_{n=-\infty}^{\infty} f(n/\nu_s) \text{sinc}(\nu_s x - n)$$

donde $\text{sinc}(x) = (\pi x)^{-1} \text{sen}(\pi x)$ si $x \neq 0$ y $\text{sinc}(0) = 1$.

Es decir, dada una frecuencia de muestreo ν_s , la máxima frecuencia que debe contener una función de banda limitada para que la podamos reconstruir es $\nu_s/2$, esto es lo que se llama *frecuencia de Nyquist*. Véase en [Wal96, §5.11] una versión aproximada bajo hipótesis más débiles que (1.35).

Demostración. Por la hipótesis $2B \leq \nu_s$, es posible definir una función g como la extensión ν_s -periódica de \widehat{f} de manera que $g(\xi) = \widehat{f}(\xi)$ para $|\xi| < B$. Su desarrollo de Fourier será, según (1.5)

$$g(\xi) = \nu_s^{-1} \sum_{n=-\infty}^{\infty} e(n\xi/\nu_s) \int_{-\nu_s/2}^{\nu_s/2} \widehat{f}(x) e(-nx/\nu_s) dx.$$

La integral se extiende a todo \mathbb{R} por la condición (1.35), y por la fórmula de inversión y cambiando n por $-n$ se sigue

$$g(\xi) = \nu_s^{-1} \sum_{n=-\infty}^{\infty} f(n/\nu_s) e(-n\xi/\nu_s).$$

Ahora bien, $f(x) = \int \widehat{f}(\xi) e(x\xi) d\xi = \int_{-\nu_s/2}^{\nu_s/2} g(\xi) e(x\xi) d\xi$, también se sigue de la fórmula de inversión, y sustituyendo g en esta última integral se concluye la prueba con un sencillo cálculo. \square

Evidentemente en la práctica la muestra no es una sucesión infinita sino que confiamos en que la señal tenga un principio y un fin (localización en tiempo) o que decaiga hasta límites despreciables. Una manifestación del *principio de incertidumbre* hace que lo primero sea teóricamente imposible para funciones de banda limitada [DM72, §2.9]. Por otro lado, hay un antiguo trabajo poco conocido de A.E. Ingham en el que se dan condiciones necesarias y suficientes para la existencia de una función de banda limitada con decaimiento prefijado [Ing].

Está claro que desde el punto de vista práctico, no podemos guardar todo el desarrollo de Fourier de una función 1-periódica arbitraria en una máquina. En principio sí podríamos hacerlo con los desarrollos de Fourier en \mathbb{Z}_N porque son finitos pero en ambos casos, con vistas a la compresión, sólo deseáramos guardar la información significativa. La mejor situación en \mathbb{T} es cuando tenemos coeficientes de Fourier que decaen rápidamente y basta con almacenar los primeros de ellos. Sin embargo, si tomamos un fragmento de una señal, típicamente cuando la hacemos periódica el resultado es discontinuo en \mathbb{T} y la baja regularidad impide la convergencia rápida de la serie de Fourier. En el caso discreto ocurre algo similar. Sumando por partes (Lema 1.2.1) en (1.29), se tiene para $n \neq 0$

$$\widehat{x}_n = \sum_{m=0}^{N-2} (x_m - x_{m+1}) s_m \quad \text{con} \quad s_m = \sum_{k=0}^m e_n(-k).$$

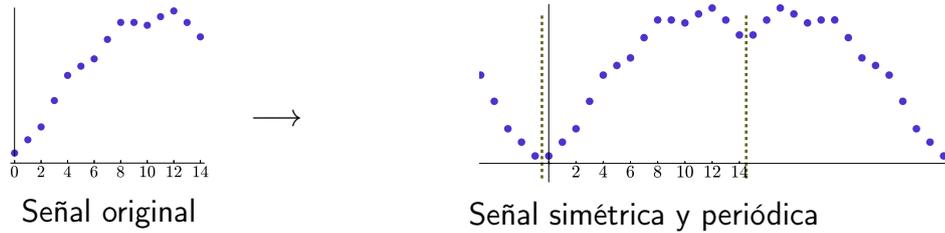
Los s_m oscilan pero se puede comprobar que su promedio no es cero, de hecho se cumple $N^{-1} \sum_{m=0}^{N-1} s_m = 1/2$, por consiguiente esperamos una contribución típica de $\frac{1}{2} \sum_{m=0}^{N-2} (x_m - x_{m+1}) = (x_0 - x_{N-1})/2$ en \hat{x}_n .

Por ejemplo, para la señal $x_n = n - 4,5$ en \mathbb{Z}_{10} , se tiene $\hat{x}_0 = 0$,

$$\hat{x}_1 = -5 + 15,3884i, \quad \hat{x}_2 = -5 + 6,8819i, \quad \hat{x}_3 = -5 + 3,6327i, \quad \hat{x}_4 = -5 + 1,6245i, \quad \hat{x}_5 = -5$$

y $\hat{x}_n = \bar{\hat{x}}_{10-n}$ para $5 < n < 10$.

El problema se soluciona en la práctica sustituyendo la señal por otra de periodo doble obtenida por simetría a través de $-1/2$ (o equivalentemente de $N - 1/2$).



La simetría por $n = -1/2$ sugiere que estamos trasladando los armónicos, y como duplicamos el periodo, deberían pasar a ser $\tilde{\epsilon}_j((n+1/2)/2N)$. Por otra parte, la simetría, hace que las partes imaginarias se cancelen, entonces desde el punto de vista teórico todo funciona como si simplemente hubiéramos cambiado los armónicos exponenciales de (1.29) por unos nuevos armónicos $\cos(\pi(m+1/2)n/N)$. Por las razones antes indicadas, la nueva transformada sustituye en muchas aplicaciones a la transformada de Fourier discreta. El análogo de la definición (1.29) y de (1.30) está incluido en el siguiente resultado:

Proposición 1.3.6. Dado $\vec{x} = (x_n)_{n=0}^{N-1} \in \mathbb{C}^N$ se define su transformada coseno discreta como

$$\hat{x}_n^c = \sum_{m=0}^{N-1} x_m \cos\left(\frac{\pi n}{N}\left(m + \frac{1}{2}\right)\right).$$

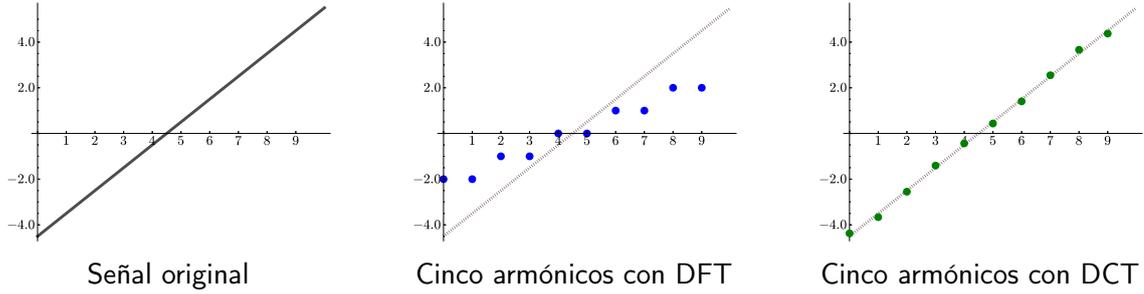
Entonces se tiene la fórmula de inversión

$$x_m = \frac{\hat{x}_0^c}{N} + \frac{2}{N} \sum_{n=1}^{N-1} \hat{x}_n^c \cos\left(\frac{\pi n}{N}\left(m + \frac{1}{2}\right)\right).$$

Para ver que la idea funciona, tomemos de nuevo la señal $x_n = n - 4,5$ en \mathbb{Z}_{10} . Los valores de la transformada coseno discreta son $\hat{x}_n = 0$ si n es par y

$$\hat{x}_1^c = -20,1801, \quad \hat{x}_3^c = -2,1615, \quad \hat{x}_5^c = -0,7071, \quad \hat{x}_7^c = -0,2859, \quad \hat{x}_9^c = -0,0801.$$

Por si esto no fuera suficientemente ilustrativo, veamos lo que ocurre al reconstruir la señal original empleando sólo la primera mitad de los armónicos:



No es ya sólo que la transformada coseno discreta (DCT) da una aproximación casi perfecta con la mitad de la información, es que la transformada de Fourier discreta apenas aproxima en la misma situación.

Demostración de la Proposición 1.3.6. Como las fórmulas son lineales, basta probar la fórmula de inversión en el caso en que $x_{m_0} = 1$ y $x_m = 0$ para $m \neq m_0$. Por sustitución directa y empleando $2 \cos \alpha \cos \beta = \cos(\alpha + \beta) + \cos(\alpha - \beta)$, la fórmula se reduce en este caso, con la notación de (1.28), a

$$\delta_D(m - m_0) = 1 + \sum_{n=1}^{N-1} \left(\cos\left(\frac{\pi n}{N}(m_0 + m + 1)\right) + \cos\left(\frac{\pi n}{N}(m - m_0)\right) \right) \quad \text{para } 0 \leq m < N.$$

La suma de cosenos es par en n y se anula para $n = N$, por tanto se reescribe como

$$\delta_D(m - m_0) = \frac{1}{2} \sum_{n=-N}^{N-1} \left(\cos\left(\frac{\pi n}{N}(m_0 + m + 1)\right) + \cos\left(\frac{\pi n}{N}(m - m_0)\right) \right) \quad \text{para } 0 \leq m < N,$$

que equivale a

$$\delta_D(m - m_0) = \frac{1}{2} \Re \sum_{n=0}^{2N-1} \left(e\left(\frac{n(m_0 + m + 1)}{2N}\right) + e\left(\frac{n(m - m_0)}{2N}\right) \right) \quad \text{para } 0 \leq m < N,$$

y esto se sigue de (1.28) con $2N$ en lugar de N , notando que $0 < m_0 + m + 1 < 2N$. \square

Uno de los algoritmos que sustentan la importancia del análisis de Fourier abstracto en la práctica, y uno de los más importantes en todo el análisis numérico, es la *transformada rápida de Fourier*, más conocida por sus siglas en inglés FFT. Es un algoritmo de tipo *divide y vencerás* (lo que en Matemáticas se llamaría inducción, se reduce recursivamente un problema a otros con parámetros menores) que sirve para calcular transformadas de Fourier discretas en \mathbb{Z}_N mucho más rápido de lo que cabría esperar. Por supuesto, también sirve para calcular transformadas coseno discretas pues éstas se expresan fácilmente en términos de aquellas.

Nuestro objetivo es calcular las sumas

$$y_n = \sum_{m=0}^{N-1} x_m e\left(\frac{mn}{N}\right) \quad \text{para } 0 \leq n < N$$

suponiendo que, como casi siempre en la práctica, N es una potencia de dos.

Si están almacenadas en memoria las raíces N -ésimas de la unidad y no cuesta trabajo calcular o recuperar de la memoria $e(mn/N)$, cada y_n requiere N multiplicaciones, y si queremos hallar todos los y_n , se requieren N^2 . Si uno tiene que hacer estas operaciones “a mano” no es difícil percatarse de que se reducen los cálculos usando la periodicidad de las funciones trigonométricas. No es por ello de extrañar que aunque se suele dar a [CT65] como referencia para la creación del método, algunos autores han señalado que se remonta incluso a Gauss.

La idea clave de la FFT se sintetiza en un línea:

$$(1.36) \quad y_n = \sum_{m=0}^{N'-1} x_{2m} e\left(\frac{mn}{N'}\right) + e\left(\frac{n}{N}\right) \sum_{m=0}^{N'-1} x_{2m+1} e\left(\frac{mn}{N'}\right) \quad \text{con } N' = \frac{N}{2}.$$

Esta igualdad es una obviedad. Lo que hay que comprobar es que con ella es posible reducir el número de operaciones. Llamemos y_n^{par} al primer sumatorio y y_n^{impar} al segundo.

Evidentemente, como sólo involucra la mitad de los términos, cada y_n^{par} requiere N' multiplicaciones, pero además sólo hace falta calcular y_n^{par} para $n = 0, 1, 2, \dots, N' - 1$ porque $y_n^{\text{par}} = y_{n+N'}^{\text{par}}$. En definitiva, tomando la mitad de los términos el número de multiplicaciones por fuerza bruta pasa de N^2 a $(N')^2 = N^2/4$. ¡En vez de dividirse por dos se divide por cuatro! De la misma manera, los y_n^{impar} requieren en total $N^2/4$ multiplicaciones mientras que multiplicar por $e(n/N)$ requiere otras N operaciones. (De hecho con $N/2$ es suficiente por la antisimetría).

En definitiva, al escribir (1.36) hemos transformado las N^2 operaciones que aparecen en la definición de los y_n en $N + 2(N')^2 = N + N^2/2$. Un factor dos no parece una gran cosa pero la fuerza del método está en iterar acumulando muchos de estos doses. Aplicando (1.36) a y_n^{par} e y_n^{impar} , que tienen longitud N' , para descomponerlas en sumas de longitud $N'' = N'/2$, la evolución del número de operaciones tras los dos pasos sería

$$N^2 \xrightarrow{\text{paso 1}} N + 2(N')^2 \xrightarrow{\text{paso 2}} N + 2(N' + 2(N'/2)^2) = 2N + 4(N'')^2.$$

Repitiendo el proceso $k = \log_2 N$ veces se llegará a $kN + 2^k$, con lo que el número total de multiplicaciones es comparable a kN , y esto es muchísimo menor que N^2 cuando N es grande. Por ejemplo $N = 2^{25}$ implica $N^2 = 1,12 \cdot 10^{15}$ y $kN = 8,38 \cdot 10^8$, un valor más de un millón de veces menor. En ordenador actuales (rápidos) el primer número de operaciones podría llevar dos semanas y la FFT reduciría el tiempo a un segundo.

Es fácil entender que tal diferencia es crucial en ingeniería y que un bien de tan primera necesidad y a la vez tan sencillo, no debe pertenecer a nadie. En [Ger99] hay

un extraño juicio acerca de ello: *When Cooley (then at IBM) first presented the FFT, IBM concluded that it was so significant it should be put in the public domain to prevent anyone from trying to patent it, and so it was published openly. Ironically, its very success has made this kind of behavior much less common now.*

En esta exposición de la FFT hemos supuesto que $e(mn/N)$ están dados. Un análisis similar prueba que disponiendo sólo de $e(1/N)$ el número de operaciones sigue siendo comparable al anterior [SS03, §7.13], [Str86, §5.5].

La transformada rápida de Fourier es tan rápida que a veces es conveniente calcular la convolución $a * b$ con la igualdad $a * b = (\widetilde{ab})$ en vez de directamente. Esto tiene una sorprendente consecuencia a la hora de multiplicar números muy grandes. Dados dos enteros N y M escritos en base B , al multiplicarlos aparece una convolución:

$$N = \sum_{r=0}^R a_r B^r, \quad M = \sum_{r=0}^R b_r B^r \quad \text{implica} \quad NM = \sum_{r=0}^{2R} \sum_{s=0}^r a_s b_{r-s} B^r = \sum_{r=0}^{2R} (a * b)_r B^r.$$

Con el método descrito, $a * b$ se calcula bastante rápido. En la práctica B está formado por muchos *bytes*, es como sino nosotros multiplicásemos en base 10 a mano tomando por ejemplo grupos de tres cifras y teniendo cuidado con las llevadas. Elaborar esta idea e incorporar algunas sutilezas da lugar al *algoritmo de Schönhage-Strassen* [SS71] que permite multiplicar números de n cifras en $O(n \log n \log \log n)$ pasos. La constante O es suficientemente grande como para que esto empiece a ser eficiente con respecto a otros algoritmos conocidos sólo para números con al menos decenas de miles de cifras decimales.

1.3.3. Análisis de Fourier discreto en la vida cotidiana

Hoy en día es difícil exagerar la importancia práctica del análisis de Fourier discreto y esto no es una novedad. En [Str86], escrito hace más de 25 años, leemos:

We come now to reality. The truth is that the digital computer has totally defeated the analog computer. The input is a sequence of numbers and not a continuous function. The output is another sequence of numbers, whether it comes from a digital filter or a finite element stress analysis or an image processor. The question is whether the special ideas of Fourier analysis still have a part to play, and the answer is absolutely *yes*. The discrete Fourier transform takes n numbers f_0, \dots, f_{n-1} into n coefficients c_0, \dots, c_{n-1} , and the patterns that are hidden in f stand out more distinctly in c .

Aquí vamos a ver algunas de las aplicaciones sin más propósito que ilustrar que el análisis de Fourier discreto tiene una fuerte motivación práctica.

Comencemos con unas consideraciones sencillas acerca de *filtros lineales*. Matemáticamente no son más que convoluciones y su importancia radica en la propiedad básica que se cumple tanto en \mathbb{R} como en grupos abelianos finitos (véase (1.12) y el Teorema 1.3.2)

$$F = f * g \quad \Rightarrow \quad \widehat{F} = \widehat{f}\widehat{g}.$$

La utilidad práctica es que en teoría de la señal uno quiere habitualmente actuar sobre las frecuencias (por ejemplo para eliminar ruido o crear algunos efectos o reducir el ancho de banda) y la relación anterior dice que actuar sobre las frecuencias es lo mismo que convolver la señal. Muchos matemáticos teóricos posiblemente se sorprenderían de la cantidad de libros de texto para ingenieros basados en esta fórmula que estudian mil y una formas de diseñar e implementar convoluciones útiles.

En el caso de \mathbb{Z}_N una señal es lo mismo que un vector $\vec{x} \in \mathbb{C}^N$. Si lo convolvemos con otro vector \vec{y} que sólo tiene las K primeras coordenadas no nulas, las coordenadas de la convolución serán:

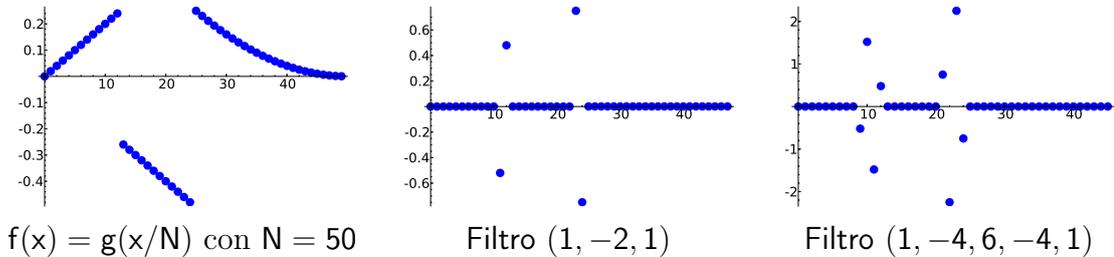
$$(1.37) \quad (\vec{x} * \vec{y})_n = \sum_{k=0}^{K-1} y_k x_{n-k}$$

donde hay que entender $n-k$ módulo N , aunque en algunas aplicaciones conviene definir $x_{n-k} = 0$ si $n-k \notin [0, N)$. Diremos que hemos aplicado el filtro $(y_0, y_1, \dots, y_{K-1})$. En teoría de la señal esto es lo que se llama un FIR (del inglés *Finite Impulse Response*). Un filtro muy sencillo para detectar la “regularidad” de una señal es $(1, -2, 1)$. La idea es que una señal suave satisface $f(n) \approx (f(n-1) + f(n+1))/2$ y por tanto el resultado de aplicar este filtro debe ser cero excepto en los puntos donde hay un problema. Matemáticamente, se tiene

$$\lim_{h \rightarrow 0} \frac{f(x-h) - 2f(x) + f(x+h)}{h^2} = f''(x) \quad \text{y} \quad \lim_{h \rightarrow 0} \frac{f(x-2h) - 4f(x-h) + 6f(x) - 4f(x+h) + f(x+2h)}{h^4} = f^{(4)}(x),$$

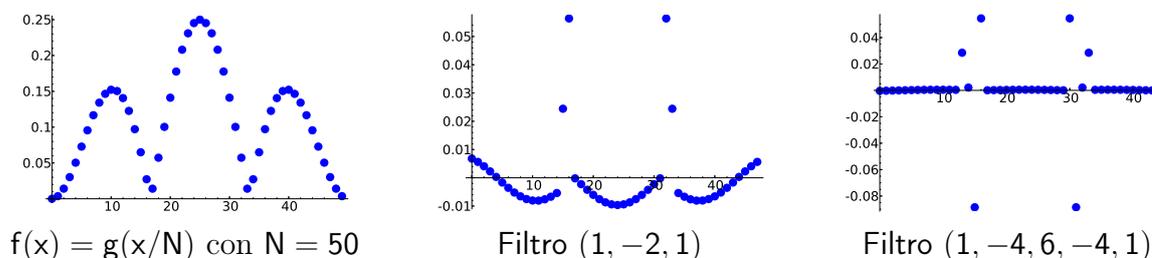
por tanto seguramente $(1, -4, 6, -4, 1)$ es un filtro más sensible para detectar faltas de regularidad.

Con la función $g(x) = x$ en $[0, 1/4)$, $g(x) = -x$ en $x \in [1/4, 1/2)$ y $g(x) = (1-x)^2$ en $[1/2, 1]$, vemos que realmente estos filtros cumplen su cometido.



Por cierto, la multiplicidad de los puntos “detectores” está relacionada con la longitud del filtro y desde el punto de vista práctico no es un inconveniente porque hace más notoria la falta de suavidad.

El segundo filtro no parece dar más información pero eso es debido a que el ejemplo es demasiado simple. Tomando $g(x) = x(1-x)|\sin(3\pi x)|$, ya notamos la diferencia:



La falta de regularidad es un enemigo fundamental para el éxito del análisis de Fourier, como vimos al introducir la transformada coseno discreta, por tanto detectar los puntos malos es interesante de por sí. Sin embargo vamos a aplicar esta idea en unos ejemplos más llamativos que posiblemente nos hayamos encontrado alguna vez sin saber la teoría que escondían.

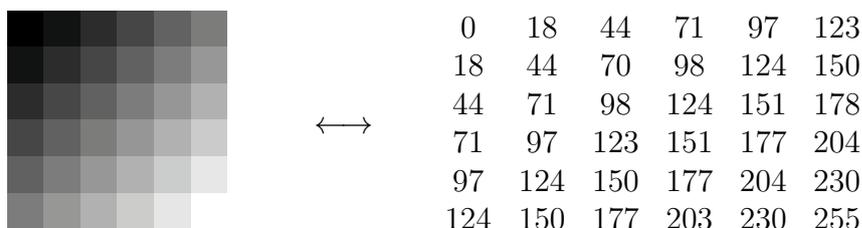
Lo que hemos hecho en \mathbb{Z}_N tiene su traducción inmediata a $\mathbb{Z}_N \times \mathbb{Z}_M$. Los filtros y las señales (las funciones) en $\mathbb{Z}_N \times \mathbb{Z}_M$ los podemos ver como matrices. Un filtro (y_{jk}) transforma la matriz (x_{nm}) haciendo que su elemento nm sea $\sum y_{jk}x_{n-j, m-k}$, es decir utilizando (1.37) en cada uno de los índices.

Consideremos los filtros dados por las matrices

$$H = \begin{pmatrix} 0 & 0 & 0 \\ 1 & -2 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad HV = \begin{pmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad Ob = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

Para detectar faltas de regularidad en las horizontales podemos usar H , y H^t para las verticales, mientras que $HV = H + H^t$ tiene en cuenta ambas direcciones. Si además introducimos las direcciones oblicuas noreste y sudeste, el filtro natural es Ob .

Una imagen en blanco y negro (se sobreentiende que en tonos de gris) de tamaño $N \times M$ viene representada por una función $f : \mathbb{Z}_N \times \mathbb{Z}_M \rightarrow \mathbb{R}$ (porque $f_R = f_G = f_B$ con la notación de la introducción). De hecho su imagen se representa con un solo byte, es decir, $\text{Im } f = \{0, 1, \dots, 255\}$ con 0 correspondiendo al negro y 255 al blanco.



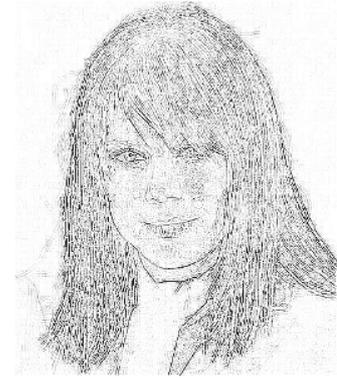
Los filtros HV y Ob aplicados a una imagen f deberían detectar bordes, y lo confirman las siguientes figuras, donde se ha representado el resultado en negativo para que sea más fácil visualizarlo (0 a blanco y 255 a negro). Para los valores de la convolución fuera del rango $[0, 255]$ hay que tomar una decisión, aquí se ha seguido la habitual de asignar 0 a los negativos y 255 a los que son mayores que éste.



Imagen original



Filtro HV (negativo)



Filtro Ob (negativo)

El resultado parece poco denso y un poco aleatorio porque estos filtros detectan variaciones en un rango de 3 pixels, que es una unidad pequeña (en un monitor típicamente $0.75mm$).

Esto sugiere trabajar en bloques de tamaño doble. Es decir, consideramos HV' y Ob' como HV y Ob pero sustituyendo cada elemento por una celda 2×2 con todos sus valores iguales. De esta forma, HV' y Ob' son matrices 6×6 . Con ello se obtiene:

Filtro HV' (negativo)Filtro Ob' (negativo) $\frac{1}{8}Ob'$ (negativo) umbral 240

La mayor norma de Ob' hace que demasiados valores sobrepasen el umbral 255 (negro en el negativo). Para evitarlo, se ha dividido en la última figura el filtro por 8 y se ha bajado el umbral a 240. Con ello hemos conseguido imitar el efecto de transformar una foto en un dibujo a lápiz (véase una variante en [BB09, §4.4]).

La mayor parte de los filtros de los programas de retoque fotográfico están basados en convoluciones.

En el conocido programa libre GIMP, es posible elegir una matriz de convolución para crear un filtro propio (en el menú `Filters>Generic>Convolution Matrix`). En la documentación correspondiente se lee “*Here is a mathematician’s domain. [...] What is a convolution matrix? It’s possible to get a rough idea of it without using mathematical tools that only a few ones know.*” Lo curioso es que a los matemáticos en formación que conocen sobradamente las herramientas, rara vez se les explica su utilidad.

El formato más difundido en la fotografía digital amateur es el formato JPEG. Estas siglas se refieren al comité que introdujo sus especificaciones (*Joint Photographic Experts Group*) pero ha quedado como nombre del propio formato. De nuevo consideraremos una imagen en blanco y negro (tonos de gris), representada por una función $f : \mathbb{Z}_N \times \mathbb{Z}_M \rightarrow \mathbb{R}$ cuya imagen está en realidad en $\{0, 1, \dots, 255\}$. Por cierto, en las fotografías en color, el formato JPEG no trata cada uno de los canales f_R , f_G y f_B por separado sino que hace un cambio simple de coordenadas [Sal02, p.145] porque los experimentos prueban que el ojo humano presenta distinta sensibilidad a los colores dependiendo de su brillo. Después de este cambio de coordenadas todo es similar al caso de blanco y negro.

El objetivo del formato es comprimir la foto perdiendo totalmente cierto número de coeficientes de Fourier y parte de la precisión en el resto. Es por ello un formato de compresión con pérdidas, eso explica por qué las carpetas repletas de ficheros `.jpg` apenas sufren cambios en su tamaño cuando se les aplican programas de compresión y también por qué editar muchas veces una foto (abrir un JPEG y grabarlo como JPEG) reduce su calidad.

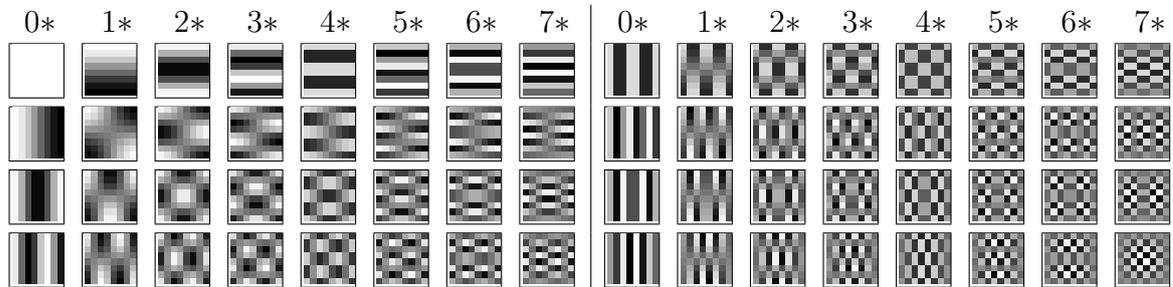
Una foto típica corresponde a una función $f : \mathbb{Z}_N \times \mathbb{Z}_M \rightarrow \mathbb{R}$ nada suave (hay cambios abruptos de color al pasar de un objeto a otro), por ello el análisis de Fourier discreto se muestra poco útil de cara a una aplicación directa. Lo que se hace es descomponer la imagen $N \times M$ en porciones de 8×8 pixels (si $8 \nmid N$ o si $8 \nmid M$, se agranda artificialmente la imagen). Cada porción 8×8 tendrá asignada su función $F : \mathbb{Z}_8 \times \mathbb{Z}_8 \rightarrow \mathbb{R}$ y para una foto típica la inmensa mayoría de estas funciones F será bastante suave, una gradación de color, porque los cambios abruptos de color ocurren en una minoría de los MN pixels de la imagen original (los contornos de los objetos son un conjunto unidimensional). Utilizando la transformada coseno discreta, ya hemos visto que se eliminan los problemas derivados de que después del 7 viene el 0 en \mathbb{Z}_8 . Por ello los armónicos naturales para analizar F son

$$\phi_{nm}(k, l) = \cos\left(\frac{\pi n}{16}(2k + 1)\right) \cos\left(\frac{\pi m}{16}(2l + 1)\right).$$

Por la Proposición 1.3.6 (la fórmula de inversión) aplicada en cada uno de los índices, se deduce que cualquier $F : \mathbb{Z}_8 \times \mathbb{Z}_8 \rightarrow \mathbb{R}$ admite un desarrollo de Fourier discreto

$$(1.38) \quad F = \sum_{n=0}^7 \sum_{m=0}^7 a_{nm} \phi_{nm}.$$

De alguna forma tenemos representada cualquier imagen 8×8 como superposición de imágenes básicas 8×8 . Asignando al negro el -1 y al blanco el 1 , el aspecto de estas imágenes básicas es:



En la mitad izquierda están los armónicos ϕ_{nm} con $0 \leq m < 4$ y en la derecha, los armónicos ϕ_{nm} con $4 \leq m < 8$.

Los coeficientes de Fourier se calculan con cierta precisión (al guardar un fichero JPEG con GIMP, en `Advanced Options > DCT method` hay cierto control sobre ello). Parece claro, y las experiencias lo avalan, que el ojo es poco sensible a cambios rápidos de brillo en áreas pequeñas, es decir, que ve peor los armónicos con frecuencias mayores. La *matriz de cuantificación* es una matriz entera $C = (c_{nm})_{n,m=0}^7$ de naturaleza experimental tal que c_{nm} está en relación inversa a la “visibilidad” del armónico ϕ_{nm} . La matriz más empleada es

$$C = \begin{pmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{pmatrix}$$

En vez de los coeficientes de Fourier a_{nm} , lo que se almacena son los enteros λ_{nm} más cercanos a a_{nm}/c_{nm} . Si F viene dada por 64 valores enteros y ahora almacenamos los λ_{nm} que son otros 64 enteros, cabe preguntarse dónde está la compresión. La respuesta es que típicamente los λ_{nm} son muchas veces cero o valores pequeños, y en estas condiciones, la lista de todos los λ_{nm} para todas las porciones 8×8 de la foto original es fácil de comprimir (sin pérdidas) con métodos habituales (que no dependen del análisis de Fourier y no veremos aquí [Sal02, §3]). Cada vez que F sea suave (una gradación de color) bastarán los primeros coeficientes de Fourier para reconstruir muy bien F con (1.38) y el resto serán prácticamente nulos. Además, la matriz de cuantificación penaliza las frecuencias grandes y hace que los λ_{nm} estén sesgados hacia valores pequeños.

Cuando se abre una imagen JPEG, internamente en el ordenador se descomprime el fichero y se obtienen los λ_{nm} , con ellos se aproximan los a_{nm} por medio de $a_{nm} \approx c_{nm}\lambda_{nm}$. Finalmente se utiliza (1.38) para reconstruir cada porción 8×8 de la imagen.

A pesar de que el formato JPEG logra habitualmente una compresión notoria sin pérdida de calidad aparente, está claro que tiene ciertas deficiencias. La más clara es que la división en porciones 8×8 es arbitraria. Una foto puede tener zonas mucho mayores con gradaciones de color en las que el análisis de Fourier discreto funciona bien.

Lo ideal sería que hubiera un proceso adaptativo en el que el tamaño de los armónicos se ajustase a las diferentes zonas de la imagen. Esta idea se manifiesta en gran medida en el nuevo formato JPEG 2000 que emplea unos armónicos bien diferentes de los que hemos visto hasta ahora y que estudiaremos a continuación.

Las *wavelets* (a veces *ondículas* en español) son unos armónicos que han causado una revolución. Sobre todo en sus comienzos, se resaltó su importancia en algunas cuestiones teóricas pero claramente su mayor influencia ha sido en el campo de las aplicaciones a la hora de tratar y comprimir diversos tipos de datos. Las ventajas de las wavelets en estas aplicaciones son por un lado que la implementación es bastante sencilla porque se basa en un algoritmo iterativo simple y, en contraposición con el análisis de Fourier, que los armónicos están localizados. Pensemos por ejemplo que la lentitud de la convergencia de la primera serie considerada por Fourier (1.2) se debe a que hay una discontinuidad. Todos los armónicos sienten esta discontinuidad que afecta globalmente a la convergencia. Las wavelets permiten un análisis más local que el análisis de Fourier. Aunque el ámbito discreto es el que más aparece en las aplicaciones, las wavelets no están limitadas a él, de hecho nacieron en \mathbb{R} y su propio nombre deriva de que en su forma continua son “ondas breves”. Hay muchas monografías matemáticas acerca de los aspectos teóricos de las wavelets en su forma continua, por ejemplo [HW96].

En esta descripción breve (seguimos en parte [Ger99]), sólo describiremos un miembro sencillo pero destacado de la tribu de las wavelets. Lo que se llama a veces D4 (o Daub4 [Wal08]), donde la nomenclatura hace referencia a I. Daubechies, la campeona de las wavelets.

Vamos a diseñar un filtro (f_0, f_1, f_2, f_3) para detectar la falta de regularidad que normalizaremos de manera que $f_0^2 + f_1^2 + f_2^2 + f_3^2 = 1$ (más adelante se verá la utilidad de ello). Si queremos que anule todas las rectas $f(n) = \alpha n + \beta$, basta con imponer $f_0 + f_1 + f_2 + f_3 = 0$ y $0f_0 + 1f_1 + 2f_2 + 3f_3 = 0$. Además añadimos la condición extraña de que sea ortogonal consigo mismo cuando se traslada dos unidades, esto es, $f_2f_0 + f_3f_1 = 0$. Este sistema no lineal de 4 ecuaciones con 4 incógnitas tiene la solución:

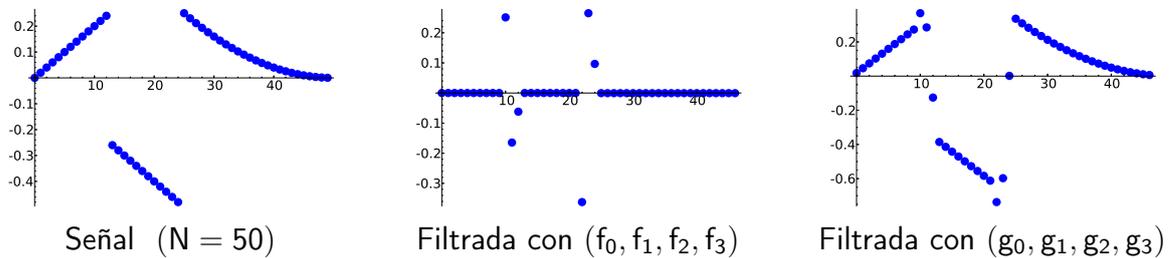
$$f_0 = \frac{1 - \sqrt{3}}{4\sqrt{2}}, \quad f_1 = \frac{\sqrt{3} - 3}{4\sqrt{2}}, \quad f_2 = \frac{3 + \sqrt{3}}{4\sqrt{2}}, \quad f_3 = -\frac{1 + \sqrt{3}}{4\sqrt{2}}.$$

Consideremos ahora el filtro (g_0, g_1, g_2, g_3) más simple que a uno se le ocurriría que sea ortogonal a (f_0, f_1, f_2, f_3) y a su trasladado dos veces:

$$g_0 = \frac{1 + \sqrt{3}}{4\sqrt{2}}, \quad g_1 = \frac{3 + \sqrt{3}}{4\sqrt{2}}, \quad g_2 = \frac{3 - \sqrt{3}}{4\sqrt{2}}, \quad g_3 = \frac{1 - \sqrt{3}}{4\sqrt{2}}.$$

Simplemente se han intercambiado f_0 con f_2 y f_1 con f_3 , modificando los signos para forzar la ortogonalidad. Este nuevo filtro no detecta bordes sino que hace un promedio escalado por $\sqrt{2}$ porque $g_0 + g_1 + g_2 + g_3 = \sqrt{2}$.

Sobre la señal que ya habíamos filtrado antes, vemos que estas ideas se ajustan a la realidad.



Por construcción, la matriz circulante

$$T = \begin{pmatrix} g_0 & g_1 & g_2 & g_3 & 0 & 0 & \dots & 0 & 0 \\ f_0 & f_1 & f_2 & f_3 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & g_0 & g_1 & g_2 & g_3 & & & \\ 0 & 0 & f_0 & f_1 & f_2 & f_3 & & & \\ \vdots & \vdots & & & & & \ddots & & \\ g_2 & g_3 & 0 & 0 & 0 & 0 & \dots & g_0 & g_1 \\ f_2 & f_3 & 0 & 0 & 0 & 0 & \dots & f_0 & f_1 \end{pmatrix} \quad \text{cumple} \quad T^t T = I,$$

es decir, es ortogonal, para ello impusimos algunas de las condiciones sobre los filtros. Cuando se la aplicamos a una señal $\vec{x} \in \mathbb{R}^N$, las coordenadas en lugar impar dan promedios y las de lugar par, faltas en la suavidad. Tras aplicar T , la ortogonalidad implica que la “energía” de la señal, su norma la cuadrado $\|\vec{x}\|^2$, se conserva. Repitiendo el proceso con el vector de promedios en vez de con toda la señal, conseguimos detalles más finos sobre la falta de suavidad (porque los promedios involucran más puntos). Supongamos $N = 2^k$ y definamos la recurrencia partiendo de $\vec{a}_0 = \vec{x}$,

$$(1.39) \quad (\vec{a}_{n+1}, \vec{d}_{n+1}) = \text{dentr}(T\vec{a}_n), \quad n = 0, 1, \dots, k - 2$$

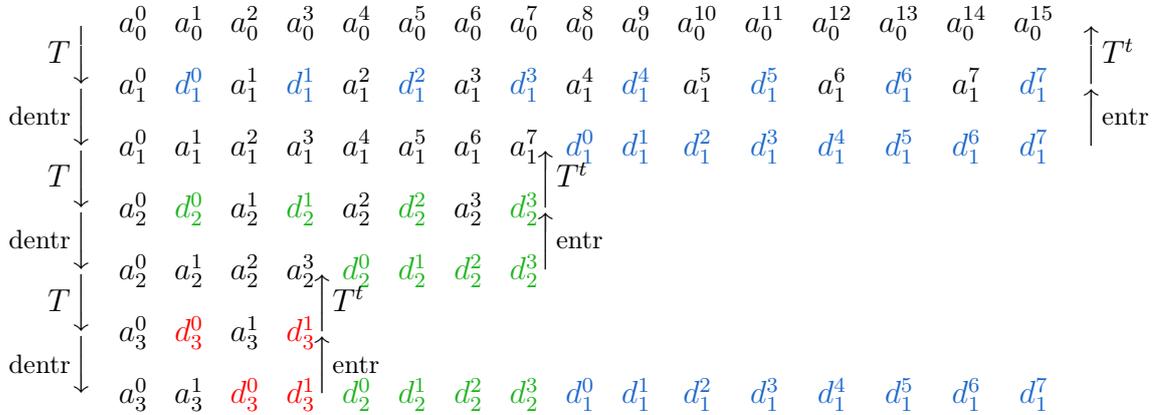
donde dentr indica el operador (lineal) que “desentrelaza” un vector poniendo primero las coordenadas que estaban en lugar impar y después las que estaban en lugar par. Se dice que los \vec{a}_j son los *vectores de aproximación* y que los \vec{d}_j son los *vectores de detalle*. Como $T^{-1} = T^t$, el proceso se puede invertir mediante

$$(1.40) \quad \vec{a}_n = T^t \text{entr}(\vec{a}_{n+1}, \vec{d}_{n+1}), \quad n = 0, 1, \dots, k - 2$$

donde $\text{entr}(\vec{a}_{n+1}, \vec{d}_{n+1})$ es el resultado de entrelazar \vec{a}_{n+1} y \vec{d}_{n+1} , poniendo el primer vector en los lugares impares y el segundo en los pares.

Por iteraciones sucesivas, (1.40) permite recuperar la señal \vec{x} a partir del último vector de aproximación $\vec{a}_{k-1} \in \mathbb{R}^2$, que da una especie de promedio general, y de todos los vectores de detalle $\vec{d}_1, \vec{d}_2, \dots, \vec{d}_{k-1}$ que tienen respectivamente $N/2, N/4, \dots, 2$ coordenadas. En total hay $N = N/2 + N/4 + \dots + 2 + 2$ grados de libertad, como es natural porque $\vec{x} \in \mathbb{R}^N$.

Por ejemplo, para una señal $\vec{x} = \vec{a}_0 \in \mathbb{R}^{16}$, las iteraciones de (1.39) y de (1.40) corresponden, respectivamente, a seguir este esquema en sentido descendente y ascendente:

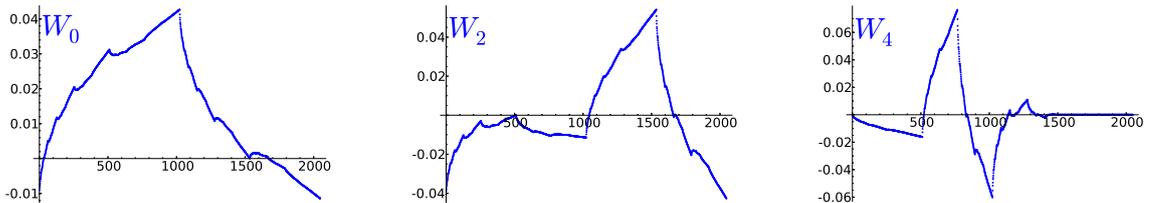


La aplicación W que dada una señal \vec{x} , le asigna $(\vec{a}_{k-2}, \vec{d}_{k-1}, \vec{d}_{k-2}, \dots, \vec{d}_1)$ por medio de (1.39), $N = 2^k$, se llama *transformada wavelet discreta* y a W^{-1} se le llama *transformada wavelet discreta inversa*. Si $\{\vec{e}_0, \vec{e}_1, \dots, \vec{e}_{N-1}\}$ es la base usual de \mathbb{R}^N , entonces por la linealidad

$$(1.41) \quad W(\vec{x}) = \sum_{j=0}^{N-1} \lambda_j \vec{e}_j \quad \implies \quad \vec{x} = \sum_{j=0}^{N-1} \lambda_j W_j \quad \text{con} \quad W_j = W^{-1}(\vec{e}_j).$$

Hemos descompuesto \vec{x} en términos de unos armónicos W_j que llamamos *wavelets*.

Es instructivo, ver el aspecto de algunas de las wavelets cuando N es muy grande. La sorpresa es que en cierto modo sólo hay dos, W_0 y W_2 (a veces llamadas padre y madre de las wavelets) porque W_1 es un trasladado de W_0 y a partir de W_2 el resto son escalados y trasladados; cuanto mayor es el índice mayor es la localización. Por ejemplo, para $N = 2048$,



Otra sorpresa es que estos armónicos no son muy suaves en el límite $N \rightarrow \infty$, más bien tienen un aspecto fractal. Hay otras variantes más complicadas que tienen mayor regularidad [Wal08].

Vamos a ilustrar la utilidad de las wavelets para reducir el ruido de una señal. Digamos que a nuestra señal le sumamos una muestra de la misma longitud correspondiente a una distribución normal $N(0, \sigma)$. Esto es lo que se llama *ruido blanco gaussiano*. En el

lado de la transformada de Fourier, es decir, en cuanto a las frecuencias se refiere, este tipo de ruido se manifiesta como una perturbación aleatoria de todos los coeficientes de Fourier. Bajo la hipótesis de que el ruido no es muy acusado (σ pequeño), una manera de reducirlo es eliminar los coeficientes de Fourier con tamaño menor que cierta cantidad t .

Con la Proposición 1.3.6, este proceso de reducción de ruido sería

$$(1.42) \quad \vec{x} \mapsto \left(\frac{c_0}{N} + \frac{2}{N} \sum_{n=1}^{N-1} c_n \cos \left(\frac{\pi n}{N} \left(m + \frac{1}{2} \right) \right) \right)_{n=0}^{N-1} \quad \text{donde } c_n = \begin{cases} 0 & \text{si } |\hat{x}_n^c| < t \\ \hat{x}_n^c & \text{si } |\hat{x}_n^c| \geq t \end{cases}$$

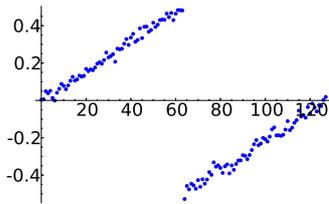
El problema es que si la señal de partida tiene muchos coeficientes de Fourier pequeños que contribuyen efectivamente a ella, estamos perdiendo también la señal. Esto ocurrirá necesariamente cuando la suavidad de la señal es poca (porque los coeficientes de Fourier decaen lentamente).

El análogo con wavelets está basado en (1.41) y es

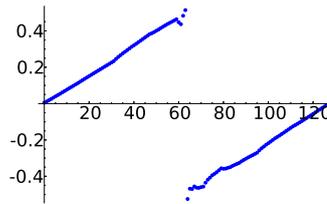
$$(1.43) \quad \vec{x} \mapsto \sum_{j=0}^{N-1} c_j W_j \quad \text{donde } c_j = \begin{cases} 0 & \text{si } |\lambda_j| < t \\ \lambda_j & \text{si } |\lambda_j| \geq t \end{cases}$$

La ventaja de las wavelets es que como sus armónicos están localizados, son menos sensibles a singularidades locales. Desde un punto de vista más teórico, el ruido blanco gaussiano sigue siendo ruido blanco gaussiano del mismo tipo al aplicar W (porque T es ortogonal).

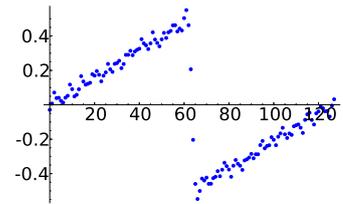
Veamos un ejemplo. Contaminamos la señal $f(n) = \{n/N\}$, $N = 128$ con ruido gaussiano $N(0, 1/50)$, y ponemos como umbral $t = 4,5\sigma = 0,09$ en (1.43) [Wal08], la reducción de ruido es bastante buena. Sin embargo con (1.42) el resultado no es satisfactorio para $t = 0,9$



Señal con ruido ($N = 128$)



(1.43), $t = 0,09$



(1.42), $t = 0,9$

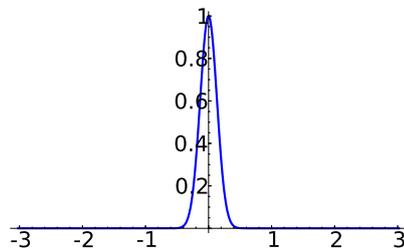
Con otros valores de t , la reducción de ruido con (1.42) no mejora. Esencialmente si t es pequeño se obtiene un resultado como el indicado y si no lo es, se ven unas ondas regulares que oscilan alrededor de la señal.

Capítulo 2

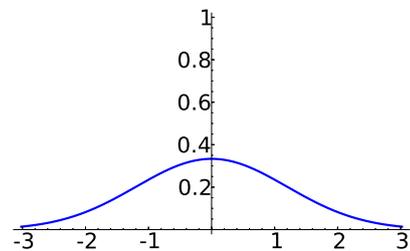
Sumas e integrales oscilatorias

2.1. El principio de incertidumbre

Si una regla tiene divisiones de un milímetro no podemos medir con mucha más precisión. De la misma forma, si tenemos armónicos que dan saltos de longitud λ parece claro que no tendremos capacidad para estudiar con precisión lo que ocurre a escalas menores. Sin embargo hay que poner ciertas acotaciones a esta afirmación porque con dos reglas con divisiones de $6mm$ y $10mm$ podemos medir $2mm$ (porque $2 = 2 \cdot 6 - 10$) y en la explotación de este hecho se basaba en parte el funcionamiento de las reglas de cálculo [BS08, §11] que durante más de cien años fueron un instrumento fundamental para cálculos científicos. Quizá, entonces, con unos pocos armónicos de baja frecuencia podamos representar aspectos locales precisos de las funciones.



Función $f(x) = e^{-9\pi x^2}$



Su transformada de Fourier

Diferentes formas del *principio de incertidumbre* afirman que tal esperanza es vana. En el análisis de Fourier clásico, hay una oposición entre la localización de una función y la de su transformada de Fourier. Es decir, como parece lógico, no podemos ver detalles pequeños sólo con frecuencias bajas. Este hecho aparece en muchos contextos de la Física y las Matemáticas pero, debido a la divulgación y a su importancia, para el gran público el nombre está ligado a la mecánica cuántica y alcanza tintes filosóficos. Cuando oímos que es imposible medir la posición y el momento de una partícula, quizá deberíamos tener presente que en física cuántica no hay partículas como las clásicas (bolitas muy pequeñas) sino que llevan asociadas una especie de ondas.

2.1.1. Diferentes formas del principio de incertidumbre

Una forma muy débil del principio de incertidumbre es que una función y su transformada de Fourier se escalan a través de la siguiente relación cuya prueba es un ejercicio sencillo:

$$(2.1) \quad g(x) = \lambda f(\lambda x) \quad \implies \quad \widehat{g}(\xi) = \widehat{f}(\xi/\lambda) \quad \text{para } \lambda > 0.$$

Digamos para fijar ideas que f y \widehat{f} son funciones no negativas dadas que tienen el 90 % de sus masas (de su norma L^1) en el intervalo $[-2013, 2013]$, entonces g y \widehat{g} tienen el 90 % de sus masas, respectivamente, en los intervalos $[-2013/\lambda, 2013/\lambda]$ y $[-2013\lambda, 2013\lambda]$. Si al variar λ tratamos de localizar g encogiendo el primer intervalo, el segundo se alarga y viceversa.

Otra versión muy débil, esta vez sin escalamientos, es que f y \widehat{f} no pueden tener ambas soporte compacto. Una variante en \mathbb{T} es que una función no nula 1-periódica con sólo un número finito de coeficientes de Fourier no nulos no se puede anular en un intervalo. Esto último se sigue, tras el cambio $e(x) = z$ (los ingenieros hablan de la *transformada* z), de que una función racional no se anula en un intervalo. Lo primero es en cierta manera similar: si f tiene soporte compacto, \widehat{f} define una función entera que no puede anularse en un intervalo (ni en ningún conjunto con un punto de acumulación).

Sabemos que para $f(x) = e^{-\pi x^2}$ se tiene $f = \widehat{f}$. Por otro lado, sin mirar la fórmula explícita, (2.1) asegura que las gaussianas que decaen como $e^{-\alpha\pi x^2}$ tienen transformadas que decaen como $e^{-\pi x^2/\alpha}$. Un bello teorema debido a G.H. Hardy [Har] afirma que esto que ocurre para las gaussianas, establece un límite para lo que sucede con cualquier función.

Teorema 2.1.1 (Hardy). *Si $f(x) = O(e^{-\alpha\pi x^2})$ y $\widehat{f}(x) = O(e^{-\beta\pi x^2})$ con $\alpha, \beta > 0$ y f no es idénticamente nula, entonces $\alpha\beta \leq 1$. Además la igualdad se da únicamente cuando f es un múltiplo de $e^{-\alpha\pi x^2}$.*

Demostración (con una laguna). Por (2.1) podemos suponer $\alpha = 1$. En estas condiciones, $\beta > 1$ implica $\widehat{f}(x) = o(e^{-\pi x^2})$ y la última parte del enunciado probaría que no hay ninguna función con estas características. Es decir, basta probar que si $f(x)$ y $\widehat{f}(x)$ son $O(e^{-\pi x^2})$, f es un múltiplo de $e^{-\pi x^2}$.

Bajo la hipótesis $f(x) = O(e^{-\pi x^2})$, la función

$$F(z) = e^{\pi z^2} \int_{-\infty}^{\infty} f(t)e(-tz) dt \quad \text{con } z = x + iy \in \mathbb{C},$$

está bien definida, porque el integrando es $O(e^{-\pi t^2 + 2\pi ty})$. Además es entera.

Si $y = 0$, usando $\widehat{f}(x) = O(e^{-\pi x^2})$ se tiene $F(z) = e^{\pi z^2} O(e^{-\pi z^2}) = O(1)$. Por otra parte, si $x = 0$, $F(z) = e^{-\pi y^2} O\left(\int_{-\infty}^{\infty} e^{-\pi t^2 + 2\pi ty} dt\right) = O(1)$, gracias al cambio $t \mapsto t + y$.

Si una función entera está acotada en los ejes real e imaginario entonces por el principio del máximo aplicado a cada uno de los cuadrantes (aquí está la laguna), F está acotada en cada uno de ellos, y una función entera acotada es constante. Por consiguiente, para $z = x$ real, $\text{cte} = e^{\pi x^2} \widehat{f}(x)$ y despejando y tomando antitransformadas, se obtiene $f(x) = \text{cte} e^{-\pi x^2}$. \square

La laguna de esta demostración, tomada del blog de T. Tao, está en que se ha aplicado el principio del máximo en una región no acotada. La solución de este problema, que puede verse en dicho blog o en [DM72, §3.2], pasa por utilizar el *principio de Pragmén-Lindelöf*, el cual esencialmente dice que el principio del máximo es cierto en regiones no acotadas siempre que la función tenga un crecimiento controlado. Lo que afea la solución es que la hipótesis sobre el crecimiento no se cumple por un poco, y hay que encoger ligeramente los sectores y multiplicar por un factor $e^{i\epsilon z^2}$ para aplicar el principio de Pragmén-Lindelöf sin hacer trampas.

La versión matemática más conocida del principio de incertidumbre es la *desigualdad de Heisenberg*, que a pesar de la denominación, no se encuentra en el famoso trabajo de W. Heisenberg en mecánica cuántica, sino que es algo posterior.

Teorema 2.1.2 (Desigualdad de Heisenberg). *Para $a, b \in \mathbb{R}$ y $f \in L^2(\mathbb{R})$*

$$16\pi^2 \int_{-\infty}^{\infty} (x - a)^2 |f(x)|^2 dx \cdot \int_{-\infty}^{\infty} (\xi - b)^2 |\widehat{f}(\xi)|^2 d\xi \geq \|f\|_2^4$$

siempre que las integrales existan. Además la igualdad se da sólo si f es un múltiplo de $e(bx)e^{-c(x-a)^2}$ para cierto $c > 0$.

Si consideramos el caso $\|f\|_2 = 1$, que es el que aparece en mecánica cuántica, entonces las medidas $|f(x)|^2 dx$ y $|\widehat{f}(\xi)|^2 d\xi$ son medidas de probabilidad y la desigualdad de Heisenberg afirma que el producto de las varianzas de las medidas correspondientes a f y \widehat{f} es siempre mayor que $1/16\pi^2$.

Demostración. Definiendo $f(x) = g(x - a)e(bx)$, un sencillo cálculo prueba que $\widehat{f}(\xi)$ es $\widehat{g}(\xi - b)e(a(b - \xi))$ y entonces la desigualdad de Heisenberg equivale a

$$16\pi^2 \int_{-\infty}^{\infty} x^2 |g(x)|^2 dx \cdot \int_{-\infty}^{\infty} \xi^2 |\widehat{g}(\xi)|^2 d\xi \geq \|g\|_2^4,$$

es decir, al caso $a = b = 0$.

Supongamos primero que g es de la clase de Schwartz para evitar consideraciones de convergencia. Integrando por partes

$$\int_{-\infty}^{\infty} |g(x)|^2 dx = - \int_{-\infty}^{\infty} x (|g(x)|^2)' dx = -2\Re \int_{-\infty}^{\infty} x g(x) \overline{g'(x)} dx.$$

La desigualdad de Cauchy-Schwarz primero, y después la identidad de Parseval, prueban

$$\|g\|_2^4 \leq 4 \int_{-\infty}^{\infty} x^2 |g(x)|^2 dx \cdot \int_{-\infty}^{\infty} |g'(x)|^2 dx = 4 \int_{-\infty}^{\infty} x^2 |g(x)|^2 dx \cdot \int_{-\infty}^{\infty} |\widehat{g}'(\xi)|^2 d\xi$$

y basta sustituir la segunda fórmula de (1.13). La desigualdad es una igualdad si y sólo si xg y g' son proporcionales y resolviendo una simple ecuación diferencial, esto equivale a que g sea una gaussiana.

Para extender la prueba a una función g que no sea tan regular, se aproxima por funciones g_n de la clase de Schwartz que cumplan $\lim \int_{-\infty}^{\infty} (1 + \xi^2) |\widehat{g}(\xi) - \widehat{g}_n(\xi)|^2 d\xi = 0$ y se hace la prueba para cada una de ellas. Véanse los detalles en [DM72, 116]. \square

El hecho de que las gaussianas sean el caso límite en el principio de incertidumbre, sugiere que dan los mejores filtros en cuanto a localización. Seguramente por ello no debiera sorprendernos que en programas de retoque fotográfico exista el “desenfocaje gaussiano”.

Hay una variante del principio de incertidumbre que es una especie de desigualdad de Heisenberg pero en intervalos finitos. Su aspecto es

$$\arccos \frac{\alpha}{\|f\|_2} + \arccos \frac{\beta}{\|f\|_2} \geq F(ab) \quad \text{con} \quad \alpha = \left(\int_{-a}^a |f|^2 \right)^{1/2} \text{ y } \beta = \left(\int_{-b}^b |f|^2 \right)^{1/2}$$

donde la función F es difícil de describir porque viene dada por el primer autovalor de un operador integral [DM72, §2.9].

Después de estos resultados, cabe preguntarse qué se puede decir en el caso finito. Un enunciado bastante sencillo (tomado de [Ter99, §14]), afirma que en el contexto del Teorema 1.3.2, f y \widehat{f} no se pueden anular ambas muchas veces.

Teorema 2.1.3. *Sea $f : G \rightarrow \mathbb{C}$ no idénticamente nula, con G un grupo abeliano finito. Entonces*

$$|\text{supp } f| |\text{supp } \widehat{f}| \geq |G|$$

donde supp indica el conjunto de puntos donde no se anula una función.

Demostración. Claramente

$$\langle f, f \rangle \leq |\text{supp } f| \max_{g \in G} |f(g)|^2$$

Por la fórmula de inversión (véase el Teorema 1.3.2) y la desigualdad de Cauchy-Schwarz, se tiene

$$\max_{g \in G} |f(g)|^2 \leq \left(\frac{1}{|G|} \sum_{\chi \in \widehat{G}} |\widehat{f}(\chi)| \right)^2 \leq \frac{1}{|G|^2} \sum_{\chi \in \widehat{G}} |\widehat{f}(\chi)|^2 |\text{supp } \widehat{f}|.$$

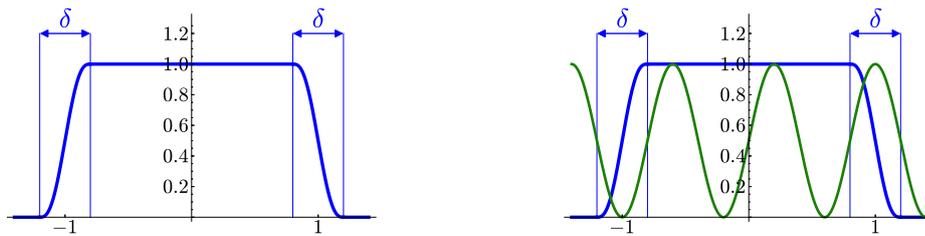
Sustituyendo esta desigualdad en la anterior y aplicando la identidad de Parseval,

$$\langle f, f \rangle \leq |\text{supp } f| |\text{supp } \widehat{f}| \frac{\langle f, f \rangle}{|G|}.$$

Finalmente, los términos $\langle f, f \rangle$ se cancelan porque f no es idénticamente nula. \square

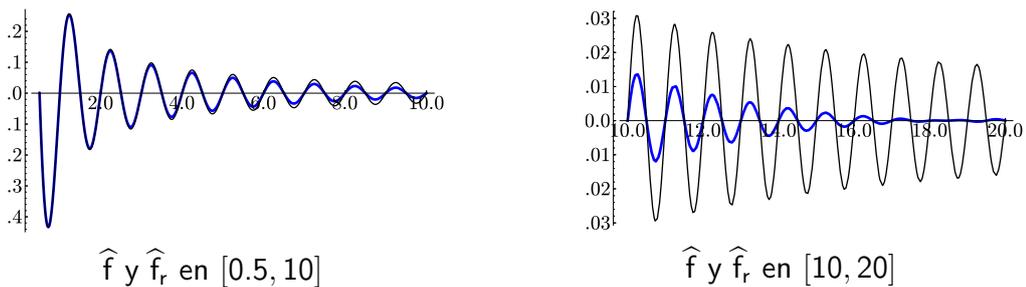
Veamos algunas figuras que muestran la idea del principio de incertidumbre como límite de nuestra precisión.

Supongamos que tenemos la función característica del intervalo $[-1, 1]$, digamos f , y para aplicar en buenas condiciones el análisis de Fourier hacemos una regularización $f_r \in C^\infty$ tal que $f = f_r$ excepto en los intervalos $[-1 - \delta/2, -1 + \delta/2]$ y $[1 - \delta/2, 1 + \delta/2]$. Las ondas de longitud de onda mayor que δ se saltan los intervalos modificados y no perciben el cambio en f . Con una estructura como la del dibujo, $\widehat{f}(\xi)$ y $\widehat{f}_r(\xi)$ serán parecidas para ξ mucho menor que δ^{-1} . Lo mismo ocurrirá con los coeficientes de Fourier en un intervalo que contenga al soporte de f_r . Por otro lado, gracias a (1.13) sabemos que a la larga, \widehat{f}_r decaerá a toda velocidad, más rápido que el inverso de cualquier potencia.



La conclusión es que, usando transformadas de Fourier, no podemos distinguir la función característica de su regularización hasta que no consideramos frecuencias grandes, al menos de tamaño inverso de la precisión δ que buscamos. Ahí está la incertidumbre.

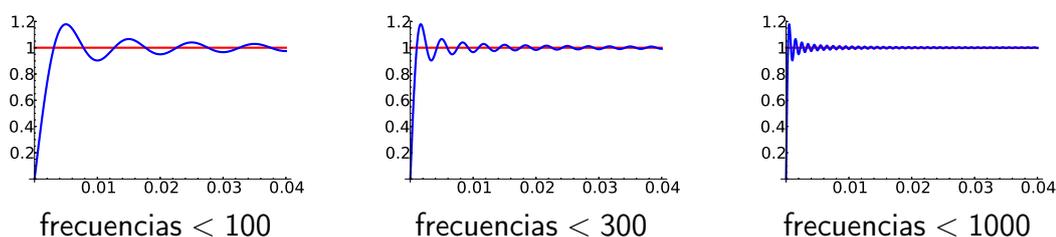
Estos son unos gráficos reales para $\delta = 0,1$ empleando una regularización C^3 .



En el primero apreciamos que hay pocas diferencias entre \widehat{f} y \widehat{f}_r en los primeros valores. Estas diferencias se van acentuando según llegamos a las frecuencias que empiezan a ver a tamaño δ , que son las comparables a $\delta^{-1} = 10$. En el intervalo $[10, 20]$ ya notamos una diferencia clara, con la transformada de Fourier de la regularización mucho menor que la de la función característica.

La localización exacta de una discontinuidad requiere precisión infinita mientras que, según lo anterior, si tomamos frecuencias hasta δ^{-1} nuestra vista está limitada a tamaño δ . En principio a escala menor puede ocurrir cualquier cosa. El *fenómeno de Gibbs* consiste en el curioso hecho de que, por muchos términos que tomemos en la serie de Fourier, en presencia de una discontinuidad aislada de salto, en el intervalo incierto que contiene a la singularidad, siempre hay un “bulto” de altura aproximada del 9% del tamaño del salto.

Por ejemplo, estos son detalles de la serie de Fourier en $[-1/2, 1/2]$ de de la función signo $f(x) = \text{sgn}(x)$.



Cualquier función que sea regular salvo una discontinuidad aislada de salto, se puede escribir como $g(x) + \lambda \text{sgn}(x - \mu)$ con g regular, por tanto basta entender el fenómeno para $f(x) = \text{sgn}(x)$.

Proposición 2.1.4 (Fenómeno de Gibbs). *Para $f(x) = \text{sgn}(x)$, se cumple*

$$\lim_{N \rightarrow \infty} \sup_{|x| < 1/(2N+1)} |f(x) - S_N f(x)| = \int_{-1}^1 \frac{\text{sen}(\pi x)}{\pi x} dx - 1 = 0,17897974 \dots$$

Nótese que, como cabe esperar por lo que hemos visto hasta ahora, el bulto se manifiesta en un intervalo de anchura N^{-1} . Es fácil ver siguiendo la demostración, que el resultado es idéntico cambiando $1/(2N + 1)$ por cualquier cantidad mayor.

Demostración. Por simetría, nos podemos centrar en el caso $x \geq 0$. Sabemos que para cualquier función, $S_N f = D_N * f$. Particularizando en la función signo, se tiene

$$S_N f(x) = - \int_x^{x+1/2} D_N(t) dt + \int_{x-1/2}^x D_N(t) dt.$$

Sustituyendo la fórmula (1.9) para D_N y usando que es par, esto se puede reescribir como

$$S_N f(x) = \int_{-x}^x \frac{\text{sen}((2N+1)\pi t)}{\text{sen}(\pi t)} dt - \int_{-x-1/2}^{x-1/2} \frac{\text{sen}((2N+1)\pi t)}{\text{sen}(\pi t)} dt.$$

Integrando por partes (recuérdese que x es pequeño), la segunda integral es $O(N^{-1})$.

Claramente el supremo de la primera integral se alcanza para $x = 1/(2N + 1)$ porque el integrando es positivo en el rango de x . Finalmente,

$$\int_{-1/(2N+1)}^{1/(2N+1)} \frac{\text{sen}((2N + 1)\pi t)}{\text{sen}(\pi t)} dt = \int_{-1}^1 \frac{\text{sen}(\pi x)}{\pi x} dx$$

se reduce a un simple cambio de variable. \square

Es fácil comprender que el fenómeno de Gibbs es relevante en la práctica pues al reconstruir una señal o una imagen que tenga cambios bruscos, provoca que aparezcan detalles inexistentes, lo que, sobre todo en el tratamiento de imágenes, se llaman *artifacts*.

2.1.2. Un poco de física cuántica

Antes de comenzar, vamos a repasar o introducir alguna de la terminología habitual sobre las ondas en Física [FLS63, §29.3]. Supongamos una onda sinusoidal en una dimensión:

$$u(x, t) = A \text{sen}(2\pi(kx - \nu t)) \quad \text{con } k, \nu > 0.$$

Esta función es solución de la *ecuación de ondas* $u_{tt} = v_p^2 u_{xx}$ con $v_p = \nu/k$. Al argumento de la onda, esto es, a $2\pi(kx - \nu t)$, se le llama *fase* y al coeficiente A , *amplitud*. Además de ello se tienen las siguientes cantidades básicas:

Not.	Significado	Unid.
ν	<i>Frecuencia</i> : el número de oscilaciones por segundo	T^{-1}
T	<i>Periodo</i> : lo que tarda en repetirse un ciclo, coincide con ν^{-1}	T
k	<i>Número de ondas</i> : las oscilaciones que hay para un tiempo fijado por unidad de espacio	L^{-1}
λ	<i>Longitud de onda</i> : separación entre dos oscilaciones para tiempo fijado (periodo espacial), coincide con k^{-1}	L
v_p	<i>Velocidad de fase</i> : lo que avanzan las crestas de la onda por unidad de tiempo, coincide con ν/k	LT^{-1}

Se han indicado las unidades de la manera habitual, L significa longitud y T tiempo. Hay que aclarar que la mayoría de las veces en Física el número de ondas se define como $2\pi k$ y para abreviar del todo la fórmula se define la *frecuencia angular* $\omega = 2\pi\nu$. En tres dimensiones la situación es similar salvo que k y x son vectores porque hay que señalar una dirección, de hecho ya en el caso unidimensional hay dos posibilidades para el signo k , correspondientes a las dos direcciones que hay en \mathbb{R} . La ecuación de ondas en tres dimensiones es $u_{tt} = v_p^2 \Delta u$.

La historia de la física cuántica en sus inicios fue bastante azarosa [SR01]. En un primer periodo, del que deriva su nombre, no era un procedimiento demasiado sistemático, se resolvían problemas clásicamente y se imponía que ciertas cantidades físicas tomaran valores enteros, es decir, que estuvieran cuantizadas (cuantificadas sería quizá más correcto pero apenas se ha usado este término [GP90]).

El gran triunfo de este periodo fue el modelo atómico propuesto por N. Bohr. Imaginemos un átomo de hidrógeno como una bolita pequeña de carga negativa, el electrón, girando alrededor de una grande de carga positiva, el protón. La electrodinámica afirma que una carga en movimiento circular debe radiar energía y esa pérdida de energía se traduce en que el radio de su órbita se reducirá paulatinamente. Haciendo los cálculos, cualquier átomo debería colapsar en un tiempo infinitesimal. Entonces ¿por qué no nos desintegramos? El modelo atómico de Bohr postula que el *momento angular* (masa por velocidad por radio) está cuantizado de forma que al multiplicarlo por $2\pi/h$ da un entero. Aquí y en toda esta sección h es la *constante de Planck*, una cantidad fundamental en física cuántica que en el sistema internacional tiene el valor aproximado:

$$h = 6,62606 \cdot 10^{-34} Js.$$

Sólo cambios en la energía tan grandes como para permitir pasar de un entero a otro provocan cambios en la órbita. Planteado así, parece que podríamos haber hecho innumerables hipótesis alocadas para justificar la estabilidad del átomo. Lo bueno de ésta es que entraba dentro de cierta mínima teoría y que explicaba con notable precisión el *espectro de emisión* de los átomos [Fon05], es decir, las frecuencias emitidas por un átomo cuando se le comunica energía (por cierto, el método usado en astrofísica para saber la composición y velocidad de estrellas y galaxias es analizar estas frecuencias).

Hay varios experimentos que motivaron la física cuántica porque contradecían resultados clásicos. Uno de los más representativos es el efecto fotoeléctrico, descubierto por H. Hertz en el siglo XIX. Consiste en que la luz puede hacer que se emitan electrones (corriente eléctrica) al incidir sobre ciertos metales (por este principio funcionan muchos de nuestros mandos a distancia). Lo curioso es que el efecto no depende tanto de la intensidad de la luz como de su frecuencia (en el espectro visible, de su color). Más concretamente, hasta que no se llega a cierta frecuencia no hay ninguna emisión por grande que sea la intensidad y cuando hay emisión, la energía cinética máxima de cada electrón emitido depende linealmente de la frecuencia pero no depende de la intensidad. La solución propuesta por A. Einstein fue que la luz está compuesta por *cuantos*, partículas sin masa llamadas posteriormente *fotones*, cada uno de los cuales tiene energía

$$(2.2) \quad E = h\nu$$

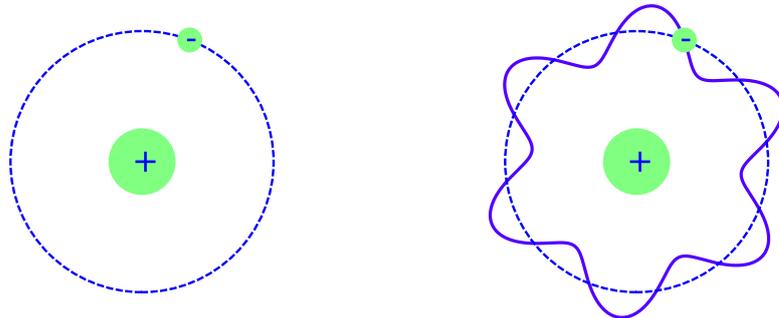
donde, como antes, ν es la frecuencia. Según este modelo, la luz se comporta como un chorro de partículas y cada una de ellas puede comunicar a lo más la energía $h\nu$ al chocar con un electrón. Cuando la intensidad es grande, hay muchos fotones pero si cada uno es poco energético, no arrancarán ningún electrón.

Otro experimento importante, el *efecto Compton* [GP90], mostraba que después de que la luz choca con un electrón libre, cambia su frecuencia de una manera que es compatible con la conservación del momento lineal relativista y (2.2).

Sin embargo, la difracción y la interferencia y, en suma, las *ecuaciones de Maxwell* [Max98], entran en conflicto con que concebamos la luz como un haz de partículas. Aunque suene paradójico, la idea que triunfó es la hipótesis de L. de Broglie de que la materia era de naturaleza ondulatoria. Además esta idea surgía a partir de la física clásica, cuando se comparaban ecuaciones de mecánica y de la *óptica geométrica* (la que considera la luz formada por rayos), algo que tiene sus antecedentes en W.R. Hamilton. De Broglie asociaba a una partícula en movimiento con *momento lineal* p (masa por velocidad) la “onda de materia” con longitud de onda

$$(2.3) \quad \lambda = \frac{h}{p}.$$

Con esto, si pensamos en la onda asociada al electrón al girar alrededor del núcleo describiendo una circunferencia de longitud L , la condición de cuantización de Bohr es equivalente a pedir que su frecuencia sea la de uno de los armónicos (de Fourier) de las funciones L -periódicas. Es decir, que podamos dibujar la onda sobre la trayectoria sin que entre en conflicto con ella misma, que sea estacionaria.



Para radiaciones electromagnéticas (por ejemplo, la luz), la electrodinámica clásica afirma $p = E/c$ donde p es el momento lineal (que ya no se define como masa por velocidad), E la energía y c la velocidad de fase (la velocidad de la luz). Como para cada fotón se cumple (2.2), se tiene $p = h\nu/v_p = hk = h/\lambda$. Entonces (2.3) está en concordancia con la electrodinámica clásica. Otra concordancia notable, es que la mecánica clásica se basa en el *principio de mínima acción*, un principio variacional que afirma que las partículas se mueven de forma que la integral del momento lineal en función del espacio es estacionaria (localmente mínima). Por otro lado, el *principio de Fermat* de la óptica geométrica [FLS63, §26], [BW65, §3] afirma algo similar para λ^{-1} , lo cual sugiere que p y λ^{-1} son proporcionales [GP90].

La fórmula (2.3) abrió un nuevo periodo en la física cuántica y propició la contribución fundamental de E. Schrödinger quien, en estos comienzos de la mecánica cuántica

ondulatoria, introdujo la ecuación básica que rige la evolución de la onda asociada a una partícula.

Esta ecuación no se deriva de principios fundamentales anteriores, es un principio en sí misma. Por ello, todo lo que se puede hacer es mostrar algunos cálculos a modo de motivación (seguimos una versión simplificada de [GP90] y [Fon05]). Es sorprendente, y exasperante para los matemáticos más puros, que muchos textos de Física presenten “deducciones” de ésta y otras ecuaciones que hacen creer al lector que son naturales y las únicas posibles, contradiciendo la propia historia de la Física.

A la onda asociada a una partícula se le llama *función de ondas* y se suele denotar mediante Ψ . En el caso de una dimensión espacial, $\Psi = \Psi(x, t)$ donde x es el espacio y t el tiempo. Supongamos que la onda es monocromática (en frecuencia), es decir, $\Psi(x, t) = \phi(x)e(-\nu t)$ (por razones tipográficas escribimos ϕ en lugar de la ψ habitual), el signo de ν es indiferente y se ha escogido negativo por coherencia con el comienzo de la sección. Esta hipótesis sobre Ψ no es tan arbitraria si pensamos que analizamos por Fourier o que en la ecuación final hacemos separación de variables. Schrödinger supuso que Ψ satisfacía la ecuación de ondas. Recordando $v_p = \nu/k$ y $\lambda = k^{-1}$, se tiene

$$\Psi_{tt} = v_p^2 \Psi_{xx} \quad \implies \quad \phi'' + \left(\frac{2\pi}{\lambda}\right)^2 \phi = 0.$$

En ningún momento se está diciendo que v_p o λ sean constantes, es decir la ecuación de ondas es más complicada que la que se resuelve en los cursos de ecuaciones. La fórmula básica de la mecánica $E = p^2/2m + V(x)$ (la *energía total* es la energía cinética más la *energía potencial*) [FLS63] junto con (2.3), lleva a

$$(2.4) \quad \phi'' + \frac{8\pi^2}{h^2} m(E - V(x))\phi = 0.$$

Fijada la masa y la energía total y dada la energía potencial, podríamos resolver esta ecuación a partir de condiciones iniciales sobre ϕ y ϕ' . Si, en analogía con lo que ocurre con el efecto fotoeléctrico, suponemos que $E = h\nu$, esto determinará Ψ . Por otra parte, de $\Psi_t = -2\pi i E h^{-1} \Psi$ podemos eliminar E de (2.4) y, suponiendo la linealidad, descartar la hipótesis inicial de que la onda sea monocromática. En otras palabras, analizamos Ψ en ondas monocromáticas de diferentes energías y sumamos o integramos los resultados. Lo que se obtiene eliminando la energía es:

$$\frac{ih}{2\pi} \Psi_t = -\frac{h^2}{8\pi^2 m} \Psi_{xx} + V\Psi.$$

El caso con tres coordenadas espaciales es formalmente similar y con esto hemos llegado a la *ecuación de Schrödinger*

$$(2.5) \quad \boxed{\frac{ih}{2\pi} \Psi_t = -\frac{h^2}{8\pi^2 m} \Delta \Psi + V\Psi.}$$

Schrödinger supuso que era válida incluso si la energía potencial V depende de t . Si no es el caso, nótese que por separación de variables, es decir, suponiendo $\Psi(\vec{x}, t) = X(\vec{x})T(t)$, se obtiene el análogo tridimensional de (2.4)

$$(2.6) \quad -\frac{\hbar^2}{8\pi^2m}\Delta\phi + V\phi = E\phi$$

Esto sugiere por qué en algunas situaciones la energía está cuantizada. Los valores que toma E son autovalores de un operador del tipo $-\alpha\Delta + V$ y bajo ciertas condiciones de contorno, formarán un conjunto discreto. Schrödinger, dio con ello una explicación matemática al modelo atómico de Bohr. El hecho de que el momento angular multiplicado por $2\pi/\hbar$ sea entero, se presentaba como consecuencia de que hubiera unos autovalores estrechamente relacionados con los enteros.

Hay varios ejemplos típicos donde se muestra la cuantización de la energía (pozos de potencial [SS90]) pero aquí nos limitaremos a estudiar el caso ridículamente sencillo $V = 0$ en una dimensión, que corresponde a una partícula libre (no afectada por la energía potencial de ningún campo). Es decir, consideramos la ecuación

$$(2.7) \quad i\Psi_t = -\frac{\hbar^2}{4\pi m}\Psi_{xx} \quad \text{con} \quad \Psi(x, 0) = f(x)$$

Proposición 2.1.5. *Si $x \in \mathbb{R}$ y f es de decaimiento rápido, la única solución de (2.7) con decaimiento rápido en x para cada t , es*

$$(2.8) \quad \Psi(x, t) = \int_{-\infty}^{\infty} \widehat{f}(\xi)e\left(-\frac{\hbar\xi^2}{2m}t + x\xi\right) d\xi.$$

Por otro lado, si $x \in \mathbb{T}$ y $f \in C^\infty(\mathbb{T})$, la única solución C^∞ de (2.7) es

$$(2.9) \quad \Psi(x, t) = \sum_{n=-\infty}^{\infty} a_n e\left(-\frac{\hbar n^2}{2m}t + nx\right)$$

donde a_n son los coeficientes de Fourier de f .

En el caso de \mathbb{T} , teniendo en cuenta (2.6), vemos que la energía está cuantizada y toma los valores $\hbar^2 n^2/2m$, lo cual es coherente con la fórmula (2.2). Sin entrar en detalles, uno de los postulados de la llamada *interpretación de Copenhague* de la física cuántica es que las observaciones producen un *colapso de la función de onda* y sólo se pueden detectar “energías propias”, las de una de las ondas elementales (los armónicos) que constituyen la función de ondas.

Aplicando la propiedad de la convolución (1.12) y aplicando “a ciegas” la fórmula (1.14) para a imaginario puro, se obtiene la solución en \mathbb{R} como una integral oscilatoria que degenera cuando $t \rightarrow 0$.

$$(2.10) \quad \Psi(x, t) = (1 - i)\sqrt{\frac{m}{2\hbar t}} \int_{-\infty}^{\infty} f(x - y)e\left(\frac{my^2}{2\hbar t}\right) dy \quad \text{para} \quad t > 0.$$

Seguendo la prueba, está claro que la regularidad requerida se puede relajar bastante. Carleson probó esencialmente que basta con que f tenga un cuarto de derivada en L^2 , concretamente que esté en el espacio de Sobolev $H^{1/4}$ (véase [Tay96] para la definición) y ésta es la mínima regularidad exigible [DK82]. Curiosamente, para \mathbb{T} todavía está abierto el problema de saber la regularidad mínima.

Demostración. Más que comprobar la solución, vamos a deducirla y así la unicidad no requerirá nuevos argumentos. Sea Φ tal que $\widehat{\Phi} = \Psi$ donde se entiende que la transformada es en la primera variable. Por (1.13), la ecuación (2.7) equivale, tomando antitransformadas, a

$$i\Phi_t = -\frac{\pi h}{m}x^2\Phi \quad \text{con} \quad \Phi(x, 0) = \widehat{f}(-x)$$

donde se ha usado que $\check{f}(x) = \widehat{f}(-x)$.

En la variable t esto es una sencilla ecuación diferencial ordinaria, cuya solución es

$$\Phi(x, t) = \widehat{f}(-x)e\left(-\frac{hx^2}{2m}t\right).$$

Tomando transformadas de Fourier y haciendo el cambio $x \mapsto -x$, se tiene el resultado.

El caso de \mathbb{T} es similar, de hecho más sencillo, cambiando transformadas de Fourier por series de Fourier uniformemente convergentes. \square

A pesar de que Schrödinger consiguió matematizar la incipiente física cuántica evitando el misterioso “postulado de números enteros” de Bohr [SR01, p.464], no dio con la interpretación adecuada de la función de ondas. El hecho es que las partículas a nivel atómico se detectaban de una en una y esto chocaba con la interpretación ondulatoria. Basándose en trabajos anteriores, M. Born propuso que Ψ era una *amplitud de probabilidad*, en el sentido de que su *intensidad* $|\Psi|^2$, define una densidad de probabilidad de detectar la partícula que representa [SR01, §12]. Esto significa que si en un instante fijado $|\Psi|^2$ es doble en las cercanías de un punto que en las cercanías de otro, también será doble la probabilidad de detectar allí la partícula. Dicho burdamente, podemos seguir pensando en las partículas como esferitas infinitesimales pero a cambio debemos sacrificar las trayectorias definidas y pensar sólo en términos probabilísticos. Éste es un cambio drástico en la interpretación de los fenómenos físicos que está alejado de nuestra experiencia habitual porque en ella las funciones de ondas decaen más rápido de lo que podamos observar.

Con la interpretación probabilista de Ψ es natural imponer $\Psi(\cdot, t) \in L^2(\mathbb{R})$ (nos limitamos al caso unidimensional aunque el tridimensional es análogo) y que además su norma en este espacio sea 1. De esta forma $|\Psi(x, t)|^2 dx$ define una medida de probabilidad para cada tiempo fijado y la probabilidad de que detectemos una partícula \mathfrak{p} en un conjunto $C \subset \mathbb{R}$ es

$$(2.11) \quad \text{Prob}_t(\mathfrak{p} \in C) = \int_C |\Psi(x, t)|^2 dx.$$

Para ver que todo esto tiene sentido, vamos a comprobar que la normalización es coherente con la solución hallada en el resultado anterior.

Proposición 2.1.6. *Si en las soluciones (2.8) y (2.9) se tiene $\|f\|_2 = 1$, entonces $\|\Psi(\cdot, t)\|_2 = 1$ para todo t .*

Demostración. Según (2.8) y (2.9), Ψ es, respectivamente, la antitransformada de Fourier de $\widehat{f}(x)e(-hx^2t/2m)$ o la serie de Fourier con coeficientes $a_n e(-hn^2t/2m)$. En el primer caso, por la identidad de Parseval $\|\Psi(\cdot, t)\|_2 = \|\widehat{f}\|_2 = \|f\|_2 = 1$, y el segundo es idéntico escribiendo $(\sum |a_n|^2)^{1/2}$ en vez de $\|\widehat{f}\|_2$. \square

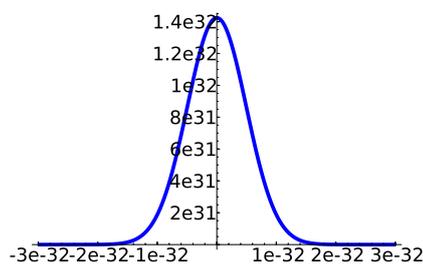
Esta propiedad de “conservación de la energía” en la ecuación de Schrödinger, es general, no se restringe al caso sencillo unidimensional y con $V = 0$ que hemos resuelto, aunque no lo veremos aquí [Fon05, §2.8].

Comprobamos a continuación que para el caso analizado de una partícula libre, lo que nos dice (2.8) es coherente con la mecánica clásica.

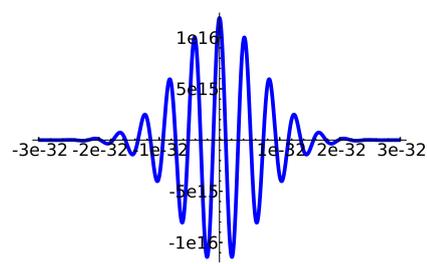
Una partícula en el origen debe tener una función de ondas tal que la medida de probabilidad $|\Psi|^2 dx$ en el instante inicial, esto es, $|f|^2 dx$, tenga casi toda su masa cerca del origen. Para facilitar los cálculos, tomemos $|f(x)| = Ce^{-x^2/a^2}$ con $a > 0$ y C una constante que asegure $\|f\|_2 = 1$, así se tiene decaimiento rápido y “anchura” comparable a a . De acuerdo con (2.3), si la velocidad inicial es v_0 , el número de ondas debería ser mv_0/h , por tanto un modelo natural para una partícula de masa m inicialmente en el origen con velocidad v_0 es

$$(2.12) \quad f(x) = \Psi(x, 0) = \left(\frac{2}{\pi a^2}\right)^{1/4} e^{-x^2/a^2} e\left(\frac{mv_0}{h}x\right).$$

Para a pequeño $|f|^2$ es como una aproximación de la identidad y f tiene una oscilación muy rápida que no afecta a la probabilidad de detectar la partícula en el instante inicial según (2.11).



$|f(x)|^2$ para $a = 15h$



$y = f(x)$ para $a = 15h, m = v = 1$

Ahora tenemos que aplicar la fórmula (2.8) para hallar la función de ondas. Para abreviar escribimos $\alpha = mv_0/h$ y $\beta = ht/2m$. El cambio $\xi \mapsto \xi + \alpha$ con (2.8) para

nuestra f , conduce a

$$\Psi(x, t) = (2\pi a^2)^{1/4} \int_{-\infty}^{\infty} e^{-\pi^2 a^2 \xi^2} e(-\beta(\xi + \alpha)^2 + x(\xi + \alpha)) d\xi,$$

que operando un poco se escribe como

$$\Psi(x, t) = (2\pi a^2)^{1/4} e(\alpha x - \beta \alpha^2) \int_{-\infty}^{\infty} e^{-(\pi^2 a^2 + 2\pi i \beta) \xi^2} e(\xi(x - 2\alpha\beta)) d\xi$$

y con (1.14) llegamos a

$$\Psi(x, t) = (2\pi)^{1/4} e(\alpha x - \beta \alpha^2) \sqrt{\frac{a}{\pi a^2 + 2i\beta}} e^{-\pi(x-2\alpha\beta)^2/(\pi a^2 + 2i\beta)}.$$

Sustituyendo α y β , vemos que la densidad de probabilidad de encontrar la partícula en un punto x en el instante t , viene dada por

(2.13)

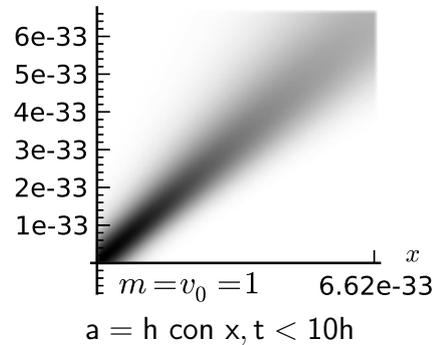
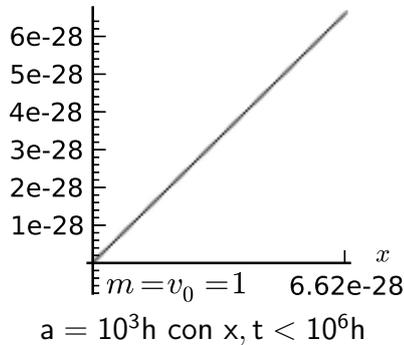
$$|\Psi(x, t)|^2 = \frac{a\sqrt{2\pi}}{\sqrt{\pi^2 a^4 + h^2 t^2 m^{-2}}} e^{-2g(x, t)} \quad \text{con} \quad g(x, t) = \frac{\pi^2 a^2}{\pi^2 a^4 + h^2 t^2 m^{-2}} (x - v_0 t)^2.$$

El valor de h es tan pequeño, que para tiempos razonables y a que no esté en el rango subatómico, se cumple

$$|\Psi(x, t)|^2 \approx \frac{(2/\pi)^{1/2}}{a} e^{-2(x-v_0 t)^2/a^2}.$$

Entonces se concentrará en una bandita de anchura comparable a a alrededor de $x = v_0 t$, con ello recuperamos en el límite la *ley de inercia*, que afirma que para una partícula libre ($V = 0$) las trayectorias son rectilíneas. Quizá parezca un poco paradójico que la velocidad no tenga que ver con la velocidad de fase de Ψ . Cuando una onda no es sinusoidal, puede haber *paquetes de ondas* que se muevan a una *velocidad de grupo* que no esté relacionada con la velocidad interna de las ondas en el interior del paquete [FLS63, §48-4].

Fuera de las escalas clásicas, la función de ondas se dispersará y entonces tendremos una amplia nube de probabilidad donde podríamos detectar la partícula.



Si intentamos que la partícula tienda a ser clásica con una posición en el origen totalmente determinada, deberíamos tomar $a \rightarrow 0^+$ en (2.12) pero entonces (2.13) implica que según avanza el tiempo, las posiciones x con $a^2 h^{-2} t^{-2} m^2 (x - v_0 t)^2 \ll 1$ tienen probabilidad más o menos parecida, esta desigualdad se puede escribir como $|p - p_0| \ll h a^{-1}$ con $p = m x / t$ el momento lineal (clásico) y $p_0 = m v_0$ el momento lineal inicial. Físicamente esto indica que hay una holgura en el momento lineal p de orden $h a^{-1}$. Matemáticamente, ello depende de que no es posible localizar una función y su transformada de Fourier.

Según hemos visto, en la física cuántica no hay posiciones de partículas sino una distribución de probabilidad. Los resultados pueden ser bien diferentes para un mismo experimento. Si hemos de elegir un valor para llamarle espacio o posición, lo más lógico es tomar la esperanza

$$\bar{x} = \int x |\Psi(x, t)|^2 dx = \int \bar{\Psi}(x, t) x \Psi(x, t) dx.$$

El momento tampoco existe como un número. Pensemos por ejemplo en (2.9). Si tomamos como cierto (2.3), esta función de ondas tiene muchos momentos, dados por $p = h k = h n$. La cantidad $|a_n|^2$ debería interpretarse como la probabilidad de que detectemos la onda de momento $h n$, ya que $\int_{\mathbb{T}} |\Psi|^2 dx = \sum |a_n|^2 = 1$ (de nuevo, esto es parte de la interpretación de Copenhague, por la cual sólo se detectan armónicos individuales con probabilidad proporcional al cuadrado del módulo de su amplitud). Entonces la esperanza del momento es

$$\bar{p} = \sum_{n=-\infty}^{\infty} n h |a_n|^2 = \frac{h}{2\pi i} \int \bar{\Psi}(x, t) \Psi_x(x, t) dx.$$

Si posición y momento no son números, ¿cómo deberíamos considerarlos matemáticamente? La última expresión muestra que el momento está relacionado con una operación sobre la función de ondas. J. von Neumann desarrolló la idea de considerar cantidades físicas como operadores y con ello acercó la mecánica cuántica al análisis funcional. De esta forma la dotó de un extenso formalismo matemático que, junto con la contribución de Dirac, ha perdurado hasta nuestros días. En los textos actuales se incluyen muchas veces unos postulados o axiomas de la mecánica cuántica escritos con lenguaje matemático (véase enunciados precisos en [GP90] y algunas ideas básicas en [SS90]). Aquí simplemente señalaremos que el *estado* de un sistema físico viene representado por una función de ondas Ψ , que es un elemento unitario de un espacio de Hilbert complejo separable, mientras que los *observables* son operadores lineales autoadjuntos.

Si definimos los operadores posición \tilde{x} y momento \tilde{p} (siempre en el caso unidimensional) actuando sobre las funciones de ondas de la siguiente manera:

$$(2.14) \quad \tilde{x}(\Psi) = x\Psi \quad \text{y} \quad \tilde{p}(\Psi) = \frac{h}{2\pi i} \frac{\partial \Psi}{\partial x},$$

entonces los valores medios (las esperanzas) de posición y momento deberían definirse como

$$\bar{x} = \langle \Psi, \tilde{x}(\Psi) \rangle = \int x |\Psi(x, t)|^2 dx \quad \text{y} \quad \bar{p} = \langle \Psi, \tilde{p}(\Psi) \rangle = \frac{h}{2\pi i} \int \bar{\Psi}(x, t) \frac{\partial \Psi}{\partial x}(x, t) dx,$$

lo cual es coherente con el ejemplo anterior. Aquí hemos evitado la notación habitual \hat{x} y \hat{p} para evitar confusiones. Es conveniente advertir que en la física cuántica utiliza muchas convenciones notacionales introducidas Dirac (véanse los rudimentos en [SS90]), que difieren de las utilizadas en análisis funcional y en análisis en general, por ejemplo, el desarrollo en serie de Fourier (1.1) se escribiría con la notación cuántica de forma tan pintoresca como $|\rangle = \sum |n\rangle \langle n|$.

Para reconciliar esta extraña visión de cantidades físicas como operadores con la mecánica ondulatoria, podemos pensar que cada función de ondas es superposición de muchas ondas elementales (armónicos) de la forma $e(px/h - Et/h)$, para ser coherentes con (2.2) y (2.3). Para “bajar” el momento de la exponencial y que pase a ser una amplitud, hay que hacer la operación $h(2\pi i)^{-1} \partial / \partial x$, mientras que si queremos hacer lo mismo con la energía, la operación sería $-h(2\pi i)^{-1} \partial / \partial t$, lo cual sugiere que éste es el operador \tilde{E} que le corresponde.

Si añadimos que las funciones actúan por multiplicación, como ya habíamos hecho con x , entonces a la fórmula clásica de conservación de la energía $p^2/2m + V = E$ le corresponde la relación de operadores $\tilde{p}^2/2m + \tilde{V} = \tilde{E}$, que aplicado a una función de ondas Ψ da lugar a

$$\frac{1}{2m} \frac{h}{2\pi i} \frac{\partial}{\partial x} \left(\frac{h}{2\pi i} \frac{\partial \Psi}{\partial x} \right) + V\Psi = -\frac{h}{2\pi i} \frac{\partial \Psi}{\partial t},$$

y operando ¡se obtiene la ecuación de Schrödinger! Es decir, la conservación de la energía clásica permite “deducir” la ecuación de Schrödinger cuando las cantidades físicas, los observables, se sustituyen por sus operadores. Este proceso para obtener fórmulas cuánticas a partir de las clásicas se llama (primera) *cuantización*.

Con todo este lenguaje, el principio de incertidumbre adquiere una formulación matemática bastante general. Dada una función de ondas Ψ , parece natural definir la varianza de un operador autoadjunto A como

$$V(A) = \overline{(A - \bar{A} \text{Id})^2} \quad \text{donde} \quad \bar{A} = \langle \Psi, A\Psi \rangle$$

y el cuadrado se entiende en sentido de operadores, es decir $(A - \bar{A} \text{Id})(A - \bar{A} \text{Id})$.

Proposición 2.1.7. Sean A y B operadores autoadjuntos, entonces

$$V(A)V(B) \geq \frac{1}{4} |\overline{AB - BA}|^2.$$

Demostración. Redefiniendo A como $A - \overline{A} \text{Id}$ y lo mismo con B , nos podemos restringir al caso $\overline{A} = \overline{B} = 0$. En este caso, por la desigualdad de Cauchy-Schwarz,

$$V(A)V(B) = \|A\Psi\|^2\|B\Psi\|^2 \geq |\langle A\Psi, B\Psi \rangle|^2.$$

Por otro lado,

$$\langle A\Psi, B\Psi \rangle = \langle \Psi, AB\Psi \rangle = \frac{1}{2}\langle \Psi, (AB - BA)\Psi \rangle + \frac{1}{2}\langle \Psi, (AB + BA)\Psi \rangle.$$

El operador $AB - BA$ es antiautoadjunto mientras que $AB + BA$ es autoadjunto, por tanto $\overline{AB - BA}$ es imaginario puro y $\overline{AB + BA}$ es real (basta escribir la definición y ver que el conjugado complejo es igual a su negativo o a él mismo). Despreciando la parte real al tomar módulos, se concluye la demostración. \square

Una consecuencia física interesante es que en cuanto dos operadores no conmuten, habrá funciones de onda para las cuales no es posible conseguir varianzas arbitrariamente pequeñas. En el caso de \tilde{x} y \tilde{p} ,

$$\tilde{x}\tilde{p} - \tilde{p}\tilde{x} = x \frac{h}{2\pi i} \frac{\partial \Psi}{\partial x} - \frac{h}{2\pi i} \frac{\partial (x\Psi)}{\partial x} = -\frac{h}{2\pi i} \Psi \quad \implies \quad \overline{\tilde{x}\tilde{p} - \tilde{p}\tilde{x}} = \frac{hi}{2\pi}$$

Por tanto, $h^2/(16\pi^2)$ impone siempre una cota inferior para el producto de las varianzas de \tilde{x} y \tilde{p} . Éste es el *principio de incertidumbre de Heisenberg*, que está más cercano a lo que realmente hizo Heisenberg, aunque él trabajó inicialmente con series de Fourier [SR01].

Si escribimos explícitamente $V(\tilde{x})$ y $V(\tilde{p})$, limitándonos al caso $\tilde{x} = \tilde{p} = 0$, para simplificar, se tiene

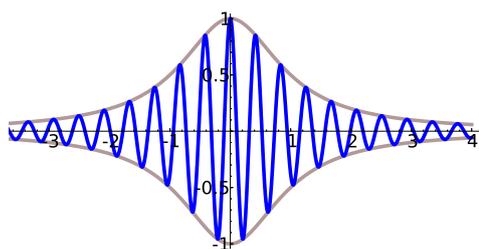
$$V(\tilde{x}) = \int x^2 |\Psi|^2 \quad \text{y} \quad V(\tilde{p}) = -\frac{h^2}{4\pi^2} \int \overline{\Psi} \frac{\partial^2 \Psi}{\partial x^2} = -\frac{h^2}{4\pi^2} \int \overline{\widehat{\Psi}} (2\pi i \xi)^2 \widehat{\Psi} = h^2 \int \xi^2 |\widehat{\Psi}|^2.$$

Entonces la cota inferior $h^2/(16\pi^2)$ para $V(\tilde{x})V(\tilde{p})$ se deduce del Teorema 2.1.2, esto es, el principio de incertidumbre de Heisenberg es consecuencia de la desigualdad de Heisenberg. Más allá de la importante interpretación física, matemáticamente es bello que la falta de localización de una función y su transformada se relacione con la falta de conmutatividad de dos operadores.

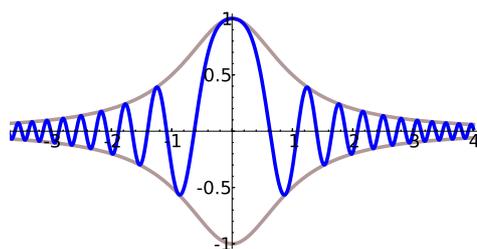
2.2. Integrales oscilatorias

Las integrales oscilatorias aparecen de manera natural en el análisis armónico abstracto pero también en numerosas aplicaciones. Muchas funciones especiales que provienen de problemas físicos concretos se definen como integrales oscilatorias.

La transformada de Fourier es uno de los ejemplos más básicos de integral oscilatoria. Si una función se anula en el infinito y sus derivadas decaen lo suficiente, entonces su transformada de Fourier tiende rápidamente a cero. A pesar de que esto es muy sencillo de probar integrando por partes, parece milagroso geoméricamente (mirando a la gráfica) que las partes positivas y negativas de senos y cosenos se cancelen con tanta precisión. ¿Qué ocurre con las integrales oscilatorias que no son transformadas de Fourier? La idea es que siempre que la fase se aproxime por una función lineal no constante, todo debería funcionar como con la transformada de Fourier.



$$f(x) = (x^2 + 1)^{-1} \cos(15x)$$



$$f(x) = (x^2 + 1)^{-1} \cos(4x^2)$$

Por otro lado, si en algún punto la fase es estacionaria (tiene derivada cero) la falta de oscilación a ritmo constante restringe la cancelación y permite extraer un término principal.

2.2.1. El principio de fase estacionaria

Vamos a considerar integrales oscilatorias de amplitud A y fase λF con $\lambda \in \mathbb{R}$, es decir, del tipo

$$(2.15) \quad I(\lambda) = \int_{-\infty}^{\infty} A(x)e(\lambda F(x)) dx.$$

El propósito del parámetro λ es graduar la oscilación y ver la influencia en el resultado. En términos generales, una integral oscilatoria es menor cuanto mayor es la oscilación.

Consideremos primero el caso $F(x) = -x$ y $A(x) = f(x)$ que es la transformada de Fourier \widehat{f} . Si $f \in C^n$ y $f^{(k)} \in L^1$ para $0 \leq k \leq n$ (la derivada 0 es la propia función), entonces integrando por partes se tiene:

$$(2.16) \quad |\widehat{f}(\xi)| \leq \frac{1}{(2\pi|\xi|)^n} \|f^{(n)}\|_1 \quad \text{para } |\xi| \neq 0.$$

Es decir, con regularidad suficiente, se cumple $\widehat{f}(\xi) = O(|\xi|^{-n})$.

Esta sencilla idea es menos específica de lo que pudiera parecer y da lugar a un resultado general siempre que la fase no sea estacionaria y las integrales tengan sentido.

Proposición 2.2.1. Consideremos (2.15) con $A \in C^n$ y $A^{(k)} \in L^1$ para $0 \leq k \leq n$. Si $G = 1/F' \in C^n$ y $G^{(k)}$ está acotada en el soporte de A para $0 \leq k \leq n$, entonces

$$I(\lambda) = O(|\lambda|^{-n}) \quad \text{cuando} \quad \lambda \rightarrow \infty.$$

Demostración. Integrando por partes

$$I(\lambda) = \int_{-\infty}^{\infty} (A(x)G(x))F'(x)e(\lambda F(x)) dx \ll \lambda^{-1} \int_{-\infty}^{\infty} (A(x)G(x))'e(\lambda F(x)) dx.$$

Si $n = 1$ se obtiene el resultado porque $A, A' \in L^1$ y G y G' están acotadas. En cualquier caso, bajo nuestras hipótesis, $(AG)' \in C^n$ y $(AG)^{(k)} \in L^1$ para $0 \leq k \leq n - 1$ y el resultado se sigue por inducción. \square

Nada similar ocurre si F' se anula en un punto del soporte de la amplitud. Por ejemplo, por extensión analítica, (1.14) debe ser también válida para $\Re(a) > 0$ donde, como en la Proposición 1.2.5, la rama de la raíz es la que aplica \mathbb{R}^+ en \mathbb{R}^+ . Para $a = 1 + 2\pi i\lambda$ se deduce de la evaluación de $\hat{f}(0)$

$$\left| \int_{-\infty}^{\infty} e^{-x^2} e(\lambda x^2) dx \right| = \sqrt{\pi} (1 + 4\pi^2 \lambda^2)^{-1/4} \sim (2|\lambda|)^{-1/2}.$$

Para estudiar los casos en que la fase llega a ser estacionaria, vamos a hacer una reducción previa. Si existe un número finito de puntos x_1, x_2, \dots, x_s en el soporte de A tales que $F'(x_j) = 0$, siempre podemos escoger funciones $\varphi_1, \varphi_2, \dots, \varphi_s \in C_0^\infty$ con soportes en intervalos disjuntos tales que φ_j coincide con A en un entorno de x_j . Entonces

$$(2.17) \quad I(\lambda) = \sum_{j=1}^s \int_{-\infty}^{\infty} \varphi_j(x) e(\lambda F(x)) dx + \int_{-\infty}^{\infty} B(x) e(\lambda F(x)) dx$$

donde $B = A - \sum \varphi_j$. Como F no tiene puntos críticos en el soporte de B , es posible aplicar la Proposición 2.2.1 bajo las condiciones especificadas de regularidad. Por otro lado, en cada una de las integrales hay exactamente un punto crítico de F dentro del soporte de φ_j . Con ello y una traslación vemos que no hay pérdida de generalidad si nos restringimos al caso en que la amplitud cumple $A \in C_0^\infty$ y $x = 0$ es el único punto del soporte de A en el que F' se anula.

Teorema 2.2.2 (Principio de fase estacionaria). Sea $A \in C_0^\infty$ y $F \in C^\infty$ con un solo punto crítico en $x = 0$ dentro del soporte de A y $F''(0) > 0$. Entonces para cada entero positivo n se cumple

$$I(\lambda) = \frac{e(\lambda F(0) + 1/8)}{\sqrt{\lambda F''(0)}} \left(A(0) + \frac{a_1}{\lambda} + \frac{a_2}{\lambda^2} + \dots + O(\lambda^{-n}) \right) \quad \text{cuando} \quad \lambda \rightarrow +\infty$$

donde los a_i dependen de los coeficientes de Taylor de A y F en $x = 0$.

La elección de los signos de λ y $F''(0)$ no impone ninguna restricción porque conjugando todos los casos, bajo $F''(0) \neq 0$, se reducen al indicado.

Los coeficientes a_j no tienen una fórmula general sencilla, aunque siguiendo la prueba se pueden determinar individualmente. Un caso bastante explícito del que deduciremos el resultado general, es el caso cuadrático.

Teorema 2.2.3. *Si $A \in C_0^\infty$ y $F(x) = x^2$, para cada entero positivo n se cumple*

$$I(\lambda) = \frac{1+i}{2\sqrt{\lambda}} \left(\sum_{j=0}^{n-1} \frac{A^{(2j)}(0)}{j!(-8\pi i\lambda)^j} + O(\lambda^{-n}) \right) \quad \text{cuando } \lambda \rightarrow +\infty.$$

Demostración. Por convergencia uniforme, se tiene

$$I(\lambda) = \lim_{\epsilon \rightarrow 0^+} I_\epsilon(\lambda) \quad \text{con} \quad I_\epsilon(\lambda) = \int_{-\infty}^{\infty} A(x) e^{-(\epsilon-2i\lambda)\pi x^2} dx.$$

Usando la identidad de Parseval (por extensión analítica de (1.14) como antes)

$$I_\epsilon(\lambda) = \frac{1}{\sqrt{\epsilon-2i\lambda}} \int_{-\infty}^{\infty} \widehat{A}(\xi) e^{-\pi\xi^2/(\epsilon-2i\lambda)} d\xi.$$

Gracias al decaimiento rápido de \widehat{A} , no hay problema al calcular el límite, y desarrollando por Taylor hasta grado $2(n-1)$ el segundo factor del integrando,

$$I(\lambda) = \frac{1+i}{2\sqrt{\lambda}} \int_{-\infty}^{\infty} \widehat{A}(\xi) e^{-\pi i \xi^2/2\lambda} d\xi = \frac{1+i}{2\sqrt{\lambda}} \left(\sum_{j=0}^{n-1} \frac{1}{j!} \left(\frac{-\pi i}{2\lambda} \right)^j \int_{-\infty}^{\infty} \widehat{A}(\xi) \xi^{2j} d\xi + O(\lambda^{-n}) \right).$$

La prueba se concluye apelando a (1.13) y a la fórmula de inversión. \square

Demostración del Teorema 2.2.2. Cambiando $F(x)$ por $F(x) - F(0)$, nos podemos restringir al caso $F(0) = 0$. Considerando el desarrollo de Taylor de $F(x)/x^2$, es fácil ver que $G(x) = \sqrt{2F(x)/x^2 F''(0)}$ define una función C^∞ en un entorno del origen con $G(0) = 1$. Entonces, por el teorema de la función inversa, $y(x) = xG(x)$ tiene una inversa $x = x(y)$ definida como una función C^∞ en un entorno del origen con $x(0) = 0$ y $x'(0) = 1$. Procediendo como en (2.17) no se restringe generalidad suponiendo que el soporte de A está dentro de dicho entorno del origen. El cambio $x \mapsto x(y)$ lleva a

$$I(\lambda) = \int_{-\infty}^{\infty} A(x(y)) x'(y) e^{(\lambda F''(0)y^2/2)} dy$$

porque $F(x) = y^2(x)F''(0)/2$. Aplicando el Teorema 2.2.3 con $\lambda F''(0)/2$ en lugar de λ , se obtiene el resultado. Nótese que por la regla de derivación de la inversa, las derivadas sucesivas de $A(x(y))$ en $y = 0$ son funciones racionales de las derivadas de A y F en cero. \square

El Teorema 2.2.2 da un desarrollo asintótico de (2.15) como un polinomio en $\lambda^{-1/2}$ si $F'(0) = 0$ y $F''(0) \neq 0$. Cabe preguntarse qué ocurriría si también la derivada segunda se anulase. Una prueba en las mismas líneas demuestra que, en general, si $F^{(k)}(0) = 0$ para $0 < k \leq K$ entonces el resultado será un polinomio en $\lambda^{-1/K}$.

Aunque la condición de que en (2.15) la amplitud sea regular y de L^1 parezca natural, en muchas ocasiones no es así. Por ejemplo, en la evaluación de las sumas de Gauss usando sumación de Poisson o en la ecuación de Schrödinger aparecen transformadas de Fourier de “gaussianas” con exponente imaginario puro. Cabe preguntarse si tiene significado como integral impropia de Riemann, es decir, si existe

$$\lim_{N, M \rightarrow \infty} \int_{-M}^N e(\lambda x^2 - \xi x) dx \quad \text{para cualquier } \lambda, \xi \in \mathbb{R} \text{ con } \lambda \neq 0,$$

y en caso afirmativo si es lícito aplicar la fórmula de inversión (1.7).

Otra cuestión relacionada es qué ocurre cuando trabajamos en un intervalo en lugar de en todo \mathbb{R} . Esto equivale a multiplicar en (2.15) la amplitud por una función característica y la falta de regularidad impide utilizar los resultados anteriores.

Una acotación sencilla bien conocida que es útil tanto en análisis armónico como en teoría analítica de números se recoge en la siguiente proposición. Nótese la asimetría del caso $k = 1$. Las amplitudes se han eliminado porque típicamente en un intervalo se pueden extraer por el segundo teorema del valor medio para integrales (en [Spi84] es el tercero) o integrando por partes. Véase [Mon94] para una versión del caso $k = 1$ con amplitudes.

Proposición 2.2.4 (Lemas de van der Corput). *Dado un entero $k \geq 2$ existe una constante c_k tal que si $F \in C^k([a, b])$ y $F^{(k)}(x) \geq \lambda_k > 0$ para todo $x \in (a, b)$, entonces*

$$\left| \int_a^b e(F(x)) dx \right| \leq c_k \lambda_k^{-1/k}.$$

El resultado también se aplica al caso $k = 1$ bajo la hipótesis adicional de que F' sea monótona.

Demostración. En el caso $k = 1$ se procede simplemente escribiendo

$$\int_a^b e(F(x)) dx = \frac{1}{2\pi i} \int_a^b \frac{1}{F'(x)} 2\pi i F'(x) e(F(x)) dx,$$

e integrando por partes para obtener

$$\left| \int_a^b e(F(x)) dx \right| \leq \frac{1}{2\pi F'(b)} + \frac{1}{2\pi F'(a)} + \frac{1}{2\pi} \left| \int_a^b \left(\frac{1}{F'(x)} \right)' dx \right| \leq \frac{1}{\pi \lambda_1}.$$

El resto de los casos se prueban por inducción. Para $k \geq 2$, la condición $F^{(k)}(x) > \lambda_k > 0$ asegura que $F^{(k-1)}$ es monótona creciente. Por el teorema del valor medio, dado $\delta > 0$, la desigualdad $|F^{(k-1)}(x)| \leq \delta$ se verifica para x incluido en un intervalo I de longitud a lo más $2\delta\lambda_k^{-1}$. En los a lo más dos intervalos que componen $[a, b] - I$ se tiene $F^{(k-1)}(x) > \delta$ o $-F^{(k-1)}(x) > \delta$. Por consiguiente la hipótesis de inducción, quizá conjugando, asegura

$$\left| \int_a^b e(F(x)) dx \right| \leq 2\delta\lambda_k^{-1} + 2c_{k-1}\delta^{-1/(k-1)}.$$

Eligiendo $\delta = \lambda_k^{(k-1)/k}$ se obtiene el resultado esperado. \square

Por ejemplo, de la Proposición 2.2.4 se tiene que siempre que $N_1 > \xi$,

$$\left| \int_{N_1}^{N_2} e(x^2 - \xi x) dx \right| \leq c_1 N_1^{-1}$$

para $0 < N_1 < N_2$ y de aquí no es difícil deducir que la transformada de Fourier de las gaussianas con exponente imaginario puro tiene sentido como integral impropia de Riemann.

Damos a continuación una forma ligeramente más explícita de la misma idea en un intervalo fijado.

Proposición 2.2.5. *Sea $\lambda > 0$ y $a < 0 < b$ entonces*

$$\int_a^b e(\lambda x^2) dx - \frac{1+i}{2\sqrt{\lambda}} \ll \min((|a|\lambda)^{-1}, \lambda^{-1/2}) + \min((b\lambda)^{-1}, \lambda^{-1/2})$$

donde la constante \ll es absoluta, es decir, no depende de ninguno de los parámetros.

Demostración. Descomponemos la integral en tres partes:

$$\int_a^b e(\lambda x^2) dx = \int_{-\infty}^{\infty} e(\lambda x^2) dx - \int_b^{\infty} e(\lambda x^2) dx - \int_{-\infty}^a e(\lambda x^2) dx.$$

La primera, que acabamos de ver que tiene sentido, se puede evaluar confiando en (1.14) o más claramente usando el teorema de los residuos para mover la línea de integración y transformar el integrando en $e^{-2\pi\lambda x^2}$. En cualquier caso, el resultado es $(1+i)/2\sqrt{\lambda}$.

Cada una de las dos últimas integrales son $O(\lambda^{-1/2})$ por la Proposición 2.2.4 con $k = 2$. Por otro lado, este mismo resultado con $k = 1$ asegura que la primera es $O((|a|\lambda)^{-1})$ y la segunda $O((b\lambda)^{-1})$. \square

Completando cuadrados, no es difícil deducir con este resultado que la fórmula de inversión $f(x) = \int \widehat{f}(\xi)e(x\xi) d\xi$ es válida, en el sentido de integrales impropias de Riemann, para $f(x) = e(\lambda x^2)$ a pesar de que esta función no está en L^1 .

Para ilustrar las ideas anteriores vamos a estudiar la asintótica de una de las *funciones de Bessel*, las cuales son transformadas de Fourier de funciones singulares. Originariamente fueron introducidas en relación con el movimiento de los planetas pero aparecen como armónicos en muchos problemas con simetría radial. Por ejemplo, así como al estudiar una cuerda vibrante los armónicos naturales son los de Fourier, para una membrana circular, como la de un tambor, son funciones de Bessel.

Aquí nos ocuparemos de J_1 , definida como

$$(2.18) \quad J_1(\lambda) = \frac{1}{\pi} \int_{-1}^1 \frac{x}{\sqrt{1-x^2}} \operatorname{sen}(\lambda x) dx.$$

Nuestro objetivo es hallar la asintótica de $J_1(2\pi\lambda)$ cuando $\lambda \rightarrow +\infty$. Tras el cambio de variable $x \mapsto \cos(\pi x)$ se tiene

$$(2.19) \quad J_1(2\pi\lambda) = \int_0^1 \cos(\pi x) \operatorname{sen}(2\pi\lambda \cos(\pi x)) dx.$$

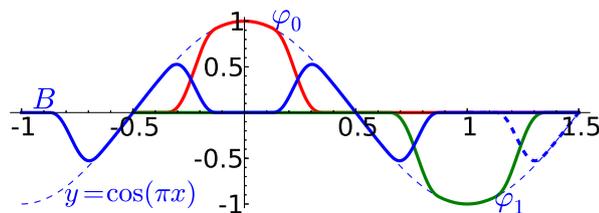
Esta expresión resulta más manejable que (2.18) y la tomaremos como punto de partida. Por la periodicidad,

$$J_1(2\pi\lambda) = \frac{1}{2} \int_0^2 \cos(\pi x) \operatorname{sen}(2\pi\lambda \cos(\pi x)) dx = \frac{1}{2} \Im \int_{-1/2}^{3/2} \cos(\pi x) e(i\lambda \cos(\pi x)) dx.$$

Con la notación de (2.15), $F(x) = \cos(\pi x)$ y la fase tiene puntos estacionarios en $x = 0$ y en $x = 1$. Primero los aislamos como en (2.17). Con esta fin introducimos $\varphi_0 \in C_0^\infty$ tal que $\varphi_0|_{[-0.1, 0.1]} = \cos(\pi x)$, con soporte incluido en $[-0.5, 0.5]$ y $\varphi_1(x) = -\varphi_0(1-x)$ (que coincide con $\cos(\pi x)$ cerca de $x = 1$). La diferencia $b(x) = \cos(\pi x) - \varphi_0(x) - \varphi_1(x)$ se transforma en una función C_0^∞ “moviendo” la parte con $x \in [1, 1.5]$ a $[-1, -0.5]$. Es decir, considerando

$$B(x) = \begin{cases} b(x) & \text{si } x \in [-1/2, 1] \\ b(x+2) & \text{si } x \in [-1, -1/2] \\ 0 & \text{en el resto} \end{cases}$$

En un dibujo,



De esta forma, por la Proposición 2.2.1, para cualquier $N > 0$

$$J_1(2\pi\lambda) = \frac{1}{2}\Im \int_{-\infty}^{\infty} (\varphi_0(x) + \varphi_1(x))e(\lambda F(x))dx + O(\lambda^{-N}).$$

Por otro lado el Teorema 2.2.2 prueba (conjugando para ajustar el signo de F'')

$$\int \varphi_0(x)e(\lambda F(x))dx = \frac{e(\lambda - 1/8)}{\pi\sqrt{\lambda}} + O(\lambda^{-3/2}).$$

Por la simetría o repitiendo el cálculo, $\int \varphi_1(x)e(\lambda F(x))dx$ da el mismo resultado y se concluye

$$(2.20) \quad J_1(2\pi\lambda) = \frac{1}{\pi\sqrt{\lambda}} \operatorname{sen} \left(2\pi\lambda - \frac{\pi}{4} \right) + O(\lambda^{-3/2}).$$

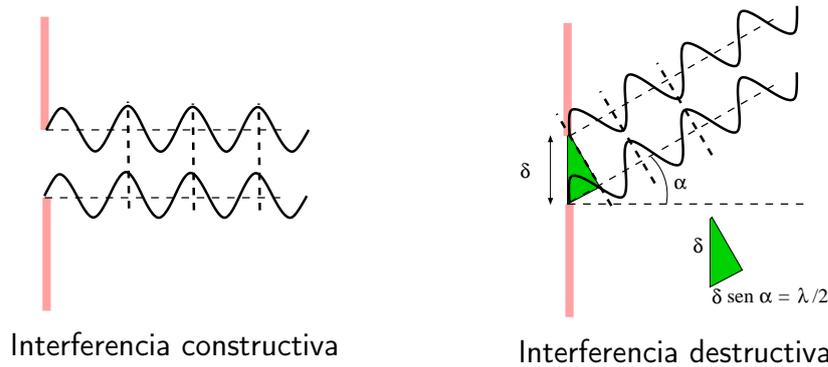
Ya para $\lambda = 1$ el error relativo al aproximar por el término principal es menor que el 6 %, mientras que para 2013 es menor que $4 \cdot 10^{-5}$.

2.2.2. Las matemáticas de la difracción

La *difracción* es un fenómeno puramente ondulatorio por el cual las ondas tienden a bordear los obstáculos. Esto es bastante creíble en el agua y mucho menos en la luz. En nuestra experiencia habitual, la luz viaja en línea recta en forma de rayos sin embargo experimentos precisos (en el límite de lo que podríamos hacer en casa, a pesar de que hay referencias bastante antiguas a la difracción [BW65, §8.1]) muestran que cuando un rayo de luz pasa por un orificio muy pequeño, al situar una pantalla frente a él a cierta distancia, se pueden ver unos curiosos anillos de luz y oscuridad. Por culpa de este fenómeno, es imposible conseguir sombras perfectamente delineadas y hay un límite en ciertos instrumentos ópticos. Por ejemplo, una *cámara oscura* (y por extensión una cámara fotográfica) es en principio mejor (más nítida) cuanto menor sea su apertura pero a partir de cierto límite produce imágenes borrosas.

Aunque matemáticamente el modelo sea complejo, no es difícil hacerse una idea intuitiva. Según el *principio de Huygens*, propuesto por el físico C. Huygens en el siglo XVII, una onda se comporta como si en cada punto generase nuevas ondas esféricas, llamadas *ondas secundarias*, que se mueven “hacia adelante” y al combinarse dan lugar a la nueva forma de la onda en el instante siguiente. Supongamos que la luz, o cualquier otra onda, pasa por una rendija de tamaño comparable a su longitud de onda. Para ciertos ángulos con la normal a la rendija, la superposición de las ondas secundarias que parten de cada punto de la rendija es destructiva y para otros no. Esto se materializa en que al pasar la luz por una rendija de este tipo se obtienen los anillos de luz y oscuridad antes mencionados. Si la rendija es muy grande en comparación con la longitud de onda (que es lo que ocurre en nuestra experiencia cotidiana), las ondas secundarias tienen todos los posibles desfases y es de prever que la superposición de todas ellas tenga gran cancelación

excepto cuando el ángulo es muy pequeño. En otras palabras, una integral como (2.15) debería ser siempre pequeña una vez que evitemos el caso en que λ es casi cero. Por eso nos parece que la luz viaja en línea recta. En otro contexto (el de la electrodinámica cuántica), esta idea está magníficamente explicada en [Fey85]. Una manera teórica de entender por qué la óptica geométrica (la luz considerada como rayos en lugar de ondas) es correcta en primera aproximación, que no exploraremos aquí, es que para longitudes de onda muy pequeñas, la ecuación de ondas, se aproxima por otra ecuación en derivadas parciales, la *ecuación eikonal* que admite soluciones que sí determinan rayos [Tay96, §6.6] [BW65, §3].



Esta explicación, aunque cualitativamente satisfactoria, es deficiente porque el principio de Huygens no es cierto sin modificaciones [BW65, §8.2], de hecho sólo da una primera aproximación correcta para ángulos pequeños. Además lo que buscamos es una justificación de la difracción matemática y físicamente más básica, que no apele a principios injustificados de este tipo.

A pesar de que la difracción ocurre con todo tipo de ondas (y partículas, según la interpretación mecánico cuántica), aquí tendremos en mente el caso de la luz visible para hacer aproximaciones. Por ello, es conveniente conocer algunos datos. En primer lugar, la velocidad (de fase) de la luz y de toda radiación electromagnética es

$$c = 299\,792\,458 \text{ ms}^{-1}$$

según se deduce de las ecuaciones de Maxwell (en realidad en los años 80 del siglo XX se redefinió el metro para que este valor fuera exacto e incluso antes, las constantes en las ecuaciones de Maxwell eran menos conocidas que el valor de c). La longitud de onda de la luz visible varía aproximadamente entre los límites (en metros)

$$3,8 \cdot 10^{-7} < \lambda < 7,4 \cdot 10^{-7} \quad \text{y por tanto} \quad 1,4 \cdot 10^6 < k < 2,6 \cdot 10^6,$$

que según la fórmula $v_p = \nu/k$, corresponde al rango de frecuencias

$$4,0 \cdot 10^{14} < \nu < 7,9 \cdot 10^{14}.$$

Las frecuencias menores son los infrarrojos y las mayores los ultravioletas. Para hacerse una idea, las frecuencias en un receptor de radio (FM) varían entre $8,75 \cdot 10^7$ y $1,08 \cdot 10^8$ y en un horno microondas son del orden de $2,5 \cdot 10^9$ (por cierto, con una sencilla cuenta se deduce que las microondas no son tan “micro”, pues tienen longitudes de unos cuantos centímetros).

La difracción es más evidente para tamaños comparables a la longitud de onda, por ello no es fácil observarla con luz visible. La dependencia de la difracción con la frecuencia, sugiere que para simplificar consideremos ondas monocromáticas, es decir, de la forma

$$(2.21) \quad u(x, y, z, t) = U(x, y, z)e(-\nu t).$$

Por representar una onda electromagnética (la luz), debe satisfacer la ecuación de ondas

$$(2.22) \quad \frac{\partial^2 u}{\partial t^2} = c^2 \Delta u.$$

Para nuestro análisis serán importantes las ondas con simetría esférica. De algún modo esto es una herencia del principio de Huygens.

Proposición 2.2.6. *Supongamos que $U = U(x, y, z)$ tiene simetría esférica, es decir, $U(x, y, z) = f(r)$ con $r = \sqrt{x^2 + y^2 + z^2}$. Entonces (2.21) satisface la ecuación de ondas (2.22) en $\mathbb{R}^3 - \{\vec{0}\}$ si y sólo si*

$$f(r) = A_1 r^{-1} e(kr) + A_2 r^{-1} e(-kr)$$

donde A_1 y A_2 son constantes y $k = \nu/c$.

Demostración. Un cálculo con la regla de la cadena prueba

$$\frac{\partial^2 f}{\partial x^2} = r^{-2} x^2 f(r) + r^{-3} (y^2 + z^2) f'(r).$$

Permutando las variables y sumando las tres parciales segundas, se obtiene

$$\Delta f = f'' + 2r^{-1} f'.$$

Por hipótesis u satisface la ecuación de ondas (2.22) con $c = \nu/k$, lo cual se puede escribir como

$$r f'' + 2f' = -4\pi^2 k^2 r f.$$

El primer miembro es $(rf)''$ y basta emplear que la solución general de $y'' + 4\pi^2 k^2 y = 0$ es $y(r) = A_1 e(kr) + A_2 e(-kr)$. \square

En varias fórmulas aparecerá la variación espacial de una onda esférica, sintetizada como el sencillo cálculo

$$(2.23) \quad \nabla(\|\mathbf{x}\|^{-1}e(k\|\mathbf{x}\|)) = \|\mathbf{x}\|^{-2}e(k\|\mathbf{x}\|)(2\pi ik - \|\mathbf{x}\|^{-1})\mathbf{x}.$$

Nuestra intuición y cursos anteriores de análisis, sugieren que una solución de la ecuación de ondas viene determinada por unos valores iniciales y velocidades iniciales. En el caso de (2.21) las derivadas con respecto al tiempo en $t = 0$ están de nuevo relacionadas con U , por tanto cabe esperar una situación como la del *problema de Dirichlet* $\Delta u = 0$, en el que la solución en un dominio con frontera compacta viene dada por los valores en dicha frontera.

El resultado relevante en este contexto se debe a G. Kirchhoff.

Teorema 2.2.7. *Sea $\Omega \subset \mathbb{R}^3$ un dominio acotado cuyo borde $\partial\Omega$ es una superficie regular S y sea \mathbf{q} un punto interior de Ω . Si (2.21) es solución de la ecuación de ondas (2.22) en un dominio que contiene a Ω ; entonces, con la orientación positiva de S ,*

$$U(\mathbf{q}) = \frac{1}{4\pi} \int_S \left(\frac{e(k\|\mathbf{x} - \mathbf{q}\|)}{\|\mathbf{x} - \mathbf{q}\|} \nabla U(\mathbf{x}) - U(\mathbf{x}) \nabla \left(\frac{e(k\|\mathbf{x} - \mathbf{q}\|)}{\|\mathbf{x} - \mathbf{q}\|} \right) \right) \cdot d\mathbf{S}(\mathbf{x}).$$

Demostración. La ecuación de ondas (2.22) es invariante por traslaciones, por consiguiente se puede suponer $\mathbf{q} = \vec{0}$. Sea $\Omega' = \Omega - B_\epsilon$ donde B_ϵ es la bola centrada de radio ϵ (pequeño). Por la *identidad de Green* (integración por partes en \mathbb{R}^3)

$$\int_{\Omega'} \left(U \Delta \left(\frac{e(k\|\mathbf{x}\|)}{\|\mathbf{x}\|} \right) - \frac{e(k\|\mathbf{x}\|)}{\|\mathbf{x}\|} \Delta U \right) dV = \int_{\partial\Omega'} \left(U \nabla \left(\frac{e(k\|\mathbf{x}\|)}{\|\mathbf{x}\|} \right) - \frac{e(k\|\mathbf{x}\|)}{\|\mathbf{x}\|} \nabla U \right) \cdot d\mathbf{S}.$$

Tanto $U(\mathbf{x})e(-\nu t)$ como $\|\mathbf{x}\|^{-1}e(k\|\mathbf{x}\| - \nu t)$ satisfacen la ecuación de ondas (por la Proposición 2.2.6), por tanto la primera integral es nula. La frontera $\partial\Omega'$ está compuesta por $\partial\Omega = S$ y por ∂B_ϵ , esta última negativamente orientada. Así pues con orientaciones positivas y tomando límites

$$\lim_{\epsilon \rightarrow 0^+} \int_{\partial B_\epsilon} \left(U \nabla \left(\frac{e(k\|\mathbf{x}\|)}{\|\mathbf{x}\|} \right) - \frac{e(k\epsilon)}{\epsilon} \nabla U \right) \cdot d\mathbf{S} = \int_S \left(U \nabla \left(\frac{e(k\|\mathbf{x}\|)}{\|\mathbf{x}\|} \right) - \frac{e(k\|\mathbf{x}\|)}{\|\mathbf{x}\|} \nabla U \right) \cdot d\mathbf{S}.$$

El segundo término del primer integrando no contribuye, porque $|\partial B_\epsilon| = 4\pi\epsilon^2$ (el área de la esfera) y $|\partial B_\epsilon|\epsilon^{-1} \rightarrow 0$. Entonces sólo resta probar

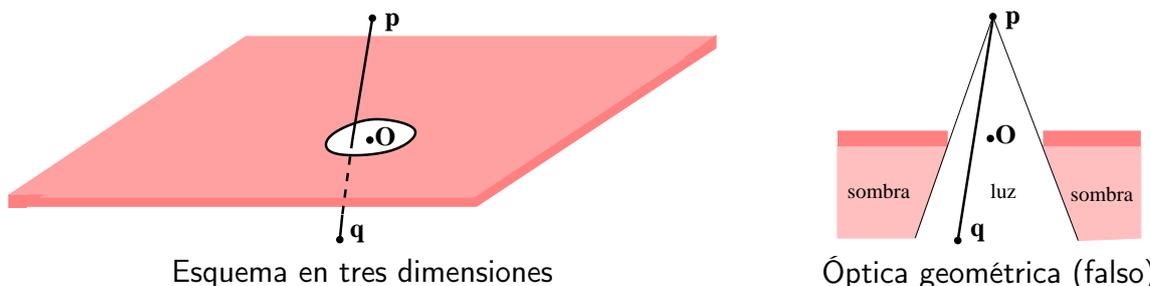
$$4\pi U(\vec{0}) = - \lim_{\epsilon \rightarrow 0^+} \int_{B_\epsilon} U(\mathbf{x}) \nabla \left(\frac{e(k\|\mathbf{x}\|)}{\|\mathbf{x}\|} \right) \cdot d\mathbf{S}.$$

Empleando (2.23), para los $\mathbf{x} \in \partial B_\epsilon$ se tiene

$$-\nabla(\|\mathbf{x}\|^{-1}e(k\|\mathbf{x}\|)) = e(k\epsilon)(\epsilon^{-2} - 2\pi ik\epsilon^{-1})\epsilon^{-1}\mathbf{x}.$$

Se cumple $\int_{\partial B_\epsilon} \epsilon^{-1}\mathbf{x} \cdot d\mathbf{S} = |\partial B_\epsilon| = 4\pi\epsilon^2$ porque $\epsilon^{-1}\mathbf{x}$ es la normal unitaria exterior de ∂B_ϵ , y para finalizar basta apelar al teorema del valor medio para integrales. \square

La situación física en la que analizaremos la difracción es la siguiente: Consideramos el plano $z = 0$ como una pantalla opaca en la que hay una abertura \mathcal{A} en un entorno del origen. En $z > 0$ tenemos una fuente de luz puntual en \mathbf{p} y queremos saber cuál es la intensidad de la luz en un punto arbitrario \mathbf{q} de $z < 0$. Basándonos en nuestra experiencia habitual, la óptica geométrica, el punto \mathbf{q} estará iluminado si el segmento \mathbf{pq} atraviesa \mathcal{A} , mientras que el resto de los puntos de $z < 0$ permanecerá en sombras.



El modelo de difracción que veremos, muestra que realmente la previsión de la óptica geométrica no es correcta por la naturaleza ondulatoria de la luz y también explica por qué en primera aproximación nos parece cierta. Dicho modelo, no es del todo exacto [BW65, XI] pues conlleva algunas aproximaciones pero son naturales desde el punto de vista físico y los resultados coinciden con lo observado en los experimentos.

En primer lugar, aplicamos el Teorema 2.2.7 a un dominio acotado Ω muy grande incluido en $z \leq 0$ tal que \mathcal{A} esté en su borde $\partial\Omega = S$. Este dominio intenta aproximar $z \leq 0$ e idealmente es todo el semiespacio. En los puntos de S que no están en \mathcal{A} , es natural suponer $U = 0$, o bien porque están demasiado lejos de la fuente de luz, o bien porque están en el borde inferior de $z = 0$ pero no en \mathcal{A} y nuestra experiencia sugiere allí una sombra total (pensamos que la óptica geométrica requiere correcciones pero no que sea completamente falsa haciendo increíble que los rayos se curven 90 grados). Con esta aproximación,

$$U(\mathbf{q}) = \frac{1}{4\pi} \int_{\mathcal{A}} \left(\frac{e(k\|\mathbf{x} - \mathbf{q}\|)}{\|\mathbf{x} - \mathbf{q}\|} \nabla U(\mathbf{x}) - U(\mathbf{x}) \nabla \left(\frac{e(k\|\mathbf{x} - \mathbf{q}\|)}{\|\mathbf{x} - \mathbf{q}\|} \right) \right) \cdot d\mathbf{S}(\mathbf{x}).$$

donde \mathcal{A} está orientada con la normal hacia arriba $\vec{n} = (0, 0, 1)$.

Una onda esférica que parte de la fuente puntual \mathbf{p} corresponde a la fórmula indicada en la Proposición 2.2.6 salvo una traslación del origen a \mathbf{p} . Para simplificar, sólo consideraremos la parte con k positivo (ondas que avanzan desde la fuente). Entonces, si no estuviera la pantalla que contiene a la abertura \mathcal{A} , la solución sería

$$(2.24) \quad U(\mathbf{x}) = K \|\mathbf{x} - \mathbf{p}\|^{-1} e(k\|\mathbf{x} - \mathbf{p}\|) \quad \text{con } K \text{ una constante.}$$

Supongamos que esta fórmula sigue siendo correcta en $z > 0$ cuando ponemos la pantalla. Matemáticamente esto parece un poco dudoso porque “pegar” dos soluciones de la ecuación de ondas, una en cada semiespacio, no está claro que pueda hacerse de forma

regular. Sin embargo, físicamente sería muy extraño pensar que interponer una pantalla opaca completamente negra (para evitar reflexiones), tenga alguna influencia en las ondas que parten de \mathbf{p} . Entonces la fórmula anterior debería ser válida en la apertura \mathcal{A} . Así pues,

$$U(\mathbf{q}) = K \int_{\mathcal{A}} \left(\frac{e(k\|\mathbf{x} - \mathbf{q}\|)}{\|\mathbf{x} - \mathbf{q}\|} \nabla \left(\frac{e(k\|\mathbf{x} - \mathbf{p}\|)}{\|\mathbf{x} - \mathbf{p}\|} \right) - \frac{e(k\|\mathbf{x} - \mathbf{p}\|)}{\|\mathbf{x} - \mathbf{p}\|} \nabla \left(\frac{e(k\|\mathbf{x} - \mathbf{q}\|)}{\|\mathbf{x} - \mathbf{q}\|} \right) \right) \cdot d\mathbf{S}(\mathbf{x}),$$

donde se ha redefinido el valor de la constante K .

Recordemos que para la luz visible el número de ondas k es del orden de millones, por tanto en (2.23) podemos despreciar el término $\|\mathbf{x}\|^{-1}$ dentro del paréntesis para tamaños macroscópicos de $\|\mathbf{x}\|$. Suponemos que \mathbf{p} y \mathbf{q} están separados de la pantalla y que por tanto $\|\mathbf{x} - \mathbf{p}\|$ y $\|\mathbf{x} - \mathbf{q}\|$ tienen tamaños macroscópicos cuando $\mathbf{x} \in \mathcal{A}$, con lo cual nuestra aproximación pasa a ser

$$(2.25) \quad U(\mathbf{q}) = \frac{K}{\lambda} \int_{\mathcal{A}} \frac{e(k\|\mathbf{x} - \mathbf{p}\| + k\|\mathbf{x} - \mathbf{q}\|)}{\|\mathbf{x} - \mathbf{p}\|\|\mathbf{x} - \mathbf{q}\|} \left(\frac{\mathbf{x} - \mathbf{p}}{\|\mathbf{x} - \mathbf{p}\|} - \frac{\mathbf{x} - \mathbf{q}}{\|\mathbf{x} - \mathbf{q}\|} \right) \cdot d\mathbf{S}(\mathbf{x}),$$

donde se ha redefinido de nuevo el valor de la constante K , que ahora es una constante compleja, y se ha usado que $k = \lambda^{-1}$.

Finalmente, supondremos que \mathcal{A} tiene un tamaño muy reducido en comparación con su distancia a \mathbf{p} y a \mathbf{q} y que los puntos \mathbf{p} y \mathbf{q} están cerca de estar alineados con ella. Todo esto está de nuevo motivado porque tenemos en mente una corrección pequeña sobre la idea de que los rayos viajan en línea recta.

Bajo estas hipótesis, el ángulo que forma $\vec{n} = (0, 0, 1)$ con $\mathbf{x} - \mathbf{q}$ para $\mathbf{x} \in \mathcal{A}$ es muy similar al que forma con $\mathbf{p} - \mathbf{x}$ o con $\mathbf{p} - \mathbf{q}$, digamos $\beta = \beta(\mathbf{p}, \mathbf{q})$. Entonces

$$\left(\frac{\mathbf{x} - \mathbf{p}}{\|\mathbf{x} - \mathbf{p}\|} - \frac{\mathbf{x} - \mathbf{q}}{\|\mathbf{x} - \mathbf{q}\|} \right) \cdot \vec{n} \approx \cos(\pi - \beta) - \cos \beta = -2 \cos \beta.$$

Además $\|\mathbf{x} - \mathbf{p}\| \approx \|\mathbf{p}\|$ y $\|\mathbf{x} - \mathbf{q}\| \approx \|\mathbf{q}\|$ sin presentar apenas variación cuando se integra sobre \mathcal{A} . Introduciendo estas aproximaciones en (2.25) y redefiniendo una vez más la constante K , se llega a

$$(2.26) \quad U(\mathbf{q}) = \frac{K \cos \beta}{\lambda \|\mathbf{p}\| \|\mathbf{q}\|} \iint_{\mathcal{A}} e(k\|\mathbf{x} - \mathbf{p}\| + k\|\mathbf{x} - \mathbf{q}\|) dx dy$$

donde $\mathbf{x} = (x, y, 0)$ y A es \mathcal{A} considerada en el plano XY en lugar de en el espacio.

Dicho sea de paso, si sólo usamos que \mathbf{p} está muy lejano, digamos que $\mathbf{p} = (0, 0, p_3)$ con $p_3 \rightarrow +\infty$, de forma que la onda esférica (2.24) degenera en una onda plana paralela a \mathcal{A} , entonces la aproximación sin quitar el denominador $\|\mathbf{x} - \mathbf{q}\|$ sería

$$(2.27) \quad U(\mathbf{q}) = \frac{K e(k\|\mathbf{p}\|) \cos \beta}{\lambda \|\mathbf{p}\|} \iint_{\mathcal{A}} \frac{e(k\|\mathbf{x} - \mathbf{q}\|)}{\|\mathbf{x} - \mathbf{q}\|} (1 + \cos \gamma) dx dy$$

donde γ es el ángulo entre \vec{n} y $\mathbf{q} - \mathbf{x}$, que es muy similar al formado entre \mathbf{q} y $-\vec{n}$. Ésta es una realización del principio de Huygens porque $e(k\|\mathbf{x} - \mathbf{q}\|)/\|\mathbf{x} - \mathbf{q}\|$ son ondas esféricas que parten de \mathcal{A} . El factor $1 + \cos \gamma$ favorece que las ondas viajen “hacia delante” (es cero para $\gamma = \pi$ y máximo para $\gamma = 0$). Esta forma corregida y bastante precisa del principio de Huygens fue introducida por A. Fresnel.

Incluso tras todas estas simplificaciones, la integral oscilatoria en (2.26) es demasiado complicada para llevar a cabo cálculos explícitos. Afortunadamente con aproximaciones lineales de la fase ya se tienen resultados importantes en óptica. Cuando se toman estas aproximaciones lineales, se habla de *difracción de Fraunhofer* (J. Fraunhofer fue el primero en analizar las frecuencias de la luz de los astros) y cuando se toman aproximaciones cuadráticas, se habla de *difracción de Fresnel*. Este último análisis sólo aporta novedades cuando \mathbf{p} o \mathbf{q} están cerca de la pantalla [BW65, §8.3.3].

El desarrollo de Taylor hasta grado uno con resto, asegura

$$\begin{cases} \|\mathbf{x} - \mathbf{p}\| = \|\mathbf{p}\| - \frac{p_1x + p_2y}{\|\mathbf{p}\|} + O\left(\frac{x^2 + y^2}{\|\mathbf{p}\|}\right) & \text{con } \mathbf{p} = (p_1, p_2, p_3) \\ \|\mathbf{x} - \mathbf{q}\| = \|\mathbf{q}\| - \frac{q_1x + q_2y}{\|\mathbf{q}\|} + O\left(\frac{x^2 + y^2}{\|\mathbf{q}\|}\right) & \text{con } \mathbf{q} = (q_1, q_2, q_3) \end{cases}$$

Entonces la difracción de Fraunhofer responde a la fórmula

$$(2.28) \quad U(\mathbf{q}) = e(k(\|\mathbf{p}\| + \|\mathbf{q}\|)) \frac{K \cos \beta}{\lambda \|\mathbf{p}\| \|\mathbf{q}\|} \widehat{\chi}_A\left(\frac{kp_1}{\|\mathbf{p}\|} + \frac{kq_1}{\|\mathbf{q}\|}, \frac{kp_2}{\|\mathbf{p}\|} + \frac{kq_2}{\|\mathbf{q}\|}\right)$$

donde $\widehat{\chi}_A$ es la transformada de Fourier en \mathbb{R}^2 de la función característica de A , esto es

$$\widehat{\chi}_A(\xi_1, \xi_2) = \iint_A e(-\xi_1x - \xi_2y) dx dy.$$

La mayor contribución en (2.28) se producirá cuando la integral no sea oscilatoria, es decir, cuando

$$\frac{kp_1}{\|\mathbf{p}\|} + \frac{kq_1}{\|\mathbf{q}\|} = 0 \quad \text{y} \quad \frac{kp_2}{\|\mathbf{p}\|} + \frac{kq_2}{\|\mathbf{q}\|} = 0.$$

Esto es lo mismo que decir que las dos primeras coordenadas de los vectores unitarios $\mathbf{p}/\|\mathbf{p}\|$ y $\mathbf{q}/\|\mathbf{q}\|$ son opuestas. Como ambos son unitarios y uno está a cada lado del plano XY , entonces las terceras coordenadas también serán opuestas.

En definitiva, fuera del caso en que \mathbf{p} , el origen y \mathbf{q} estén alineados o casi alineados en (2.28) habrá una integral oscilatoria que induce cancelación. Con esto tenemos una explicación matemática de por qué en primera aproximación nos parece que la luz se propaga en línea recta. Se podría argüir que en realidad esto apareció como hipótesis aproximada en la propia deducción pero en realidad sólo se usó para quitar términos no oscilatorios.

Después de este desarrollo teórico, analizaremos dos ejemplos básicos (véase [Wal96] para otros con datos concretos).

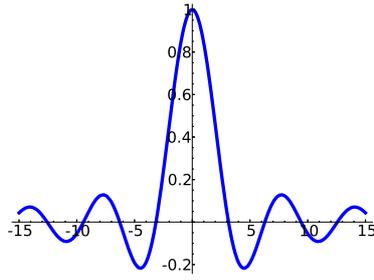
Para una abertura cuadrada de lado δ , $A = [-\delta/2, \delta/2] \times [-\delta/2, \delta/2]$, se tiene

$$\widehat{\chi}_A(\xi_1, \xi_2) = \frac{\text{sen}(\pi\xi_1\delta) \text{sen}(\pi\xi_2\delta)}{\pi^2\xi_1\xi_2}.$$

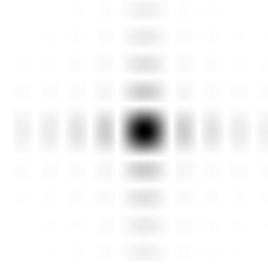
Si la fuente de luz \mathbf{p} está en $(0, 0, p_3)$, se obtiene

$$\left| \widehat{\chi}_A\left(\frac{kq_1}{\|\mathbf{q}\|}, \frac{kq_2}{\|\mathbf{q}\|}\right) \right|^2 = \|\mathbf{q}\|^2 \frac{\text{sen}^2(\pi\delta kq_1/\|\mathbf{q}\|) \text{sen}^2(\pi\delta kq_2/\|\mathbf{q}\|)}{\pi^2 k^2 q_1^2 q_2^2}.$$

Si ponemos una segunda pantalla en $z = q_3$ fijado, entonces en experimentos reales se ve un esquema de luces y sombras correspondientes a los valores de la función anterior (véase [Wal96] para una foto real). La fase de la luz no se ve en una pantalla, lo que se ve es su *intensidad* que físicamente representa la densidad de potencia por unidad de área y en el caso de la luz es proporcional al cuadrado del módulo de la función (2.21) que define la onda.



$f(x) = x^{-1} \text{sen}(x)$



"Intensidad" de $f(x)f(y)$

Un caso más simétrico es cuando A es una abertura circular de radio δ , para el cual

$$\widehat{\chi}_A(\xi_1, \xi_2) = \iint_{x^2+y^2 < \delta^2} e(-\xi_1 x - \xi_2 y) dx dy.$$

La simetría del problema implica que $\widehat{\chi}_A$ es radial (y es un interesante ejercicio probarlo matemáticamente). En particular $\widehat{\chi}_A(\xi_1, \xi_2) = \widehat{\chi}_A(0, R)$ con $R = \sqrt{\xi_1^2 + \xi_2^2}$ y se tiene

$$\widehat{\chi}_A(\xi_1, \xi_2) = \int_{-\delta}^{\delta} \int_{-\sqrt{\delta^2-x^2}}^{\sqrt{\delta^2-x^2}} e(-Ry) dy dx = \frac{1}{\pi R} \int_{-\delta}^{\delta} \text{sen}(2\pi R\sqrt{\delta^2-x^2}) dx.$$

El cambio $x \mapsto \pm\delta\sqrt{1-x^2}$ con \pm el signo de x , prueba

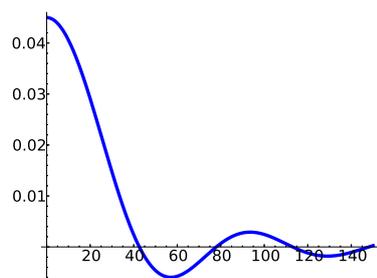
$$\widehat{\chi}_A(\xi_1, \xi_2) = \frac{\delta}{\pi R} \int_{-1}^1 x \frac{\text{sen}(2\pi R\delta x)}{\sqrt{1-x^2}} dx = \frac{\delta}{R} J_1(2\pi R\delta),$$

donde se ha usado la definición (2.18) de J_1 .

Según (2.20),

$$\left| \widehat{\chi}_A\left(\frac{kq_1}{\|\mathbf{q}\|}, \frac{kq_2}{\|\mathbf{q}\|}\right) \right|^2 \sim \frac{\delta}{\pi^2 R^3} \text{sen}^2\left(2\pi R\delta - \frac{\pi}{4}\right) \quad \text{con } R = k \frac{\sqrt{x^2 + y^2}}{\|\mathbf{p}\|}$$

cuando $R\delta$ es grande. Esto revela un esquema de difracción en anillos (véanse fotos reales en [BW65] y [Wal96, §6])



$$f(x) = x^{-1} J_1(0,09x)$$



$$\text{"Intensidad" de } f(\sqrt{x^2 + y^2})$$

Los anillos se separan más cuanto menor es δ , haciendo el efecto bien distinto de nuestra experiencia cotidiana. Cuando $R\delta$ es pequeño, con Taylor sobre (2.19) es fácil deducir

$$\left| \widehat{\chi}_A\left(\frac{kq_1}{\|\mathbf{q}\|}, \frac{kq_2}{\|\mathbf{q}\|}\right) \right|^2 \sim \pi^2 \delta^4$$

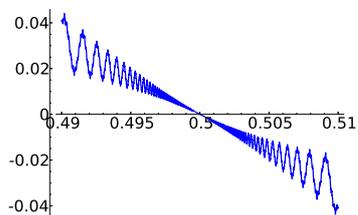
Si queremos que la intensidad de los primeros anillos sea comparable a la intensidad en el centro de la pantalla para verlos bien, entonces δ^4 debería ser comparable a δ/R^3 . Recordando que para la luz visible $\lambda = k^{-1}$ es menor que una *micra* (una milésima de milímetro), se sigue que este tamaño de δ sería conveniente para observar cómodamente la difracción.

2.3. Sumas oscilatorias

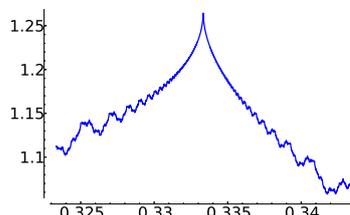
Las sumas oscilatorias que vamos a tratar son sumas trigonométricas, los parientes discretos de las integrales oscilatorias que ya hemos estudiado. Esta naturaleza discreta hace que se manifiesten en ocasiones fenómenos aritméticos que no han aparecido en nuestro análisis de las integrales oscilatorias. Aunque en lo sucesivo nos ocuparemos de sumas finitas, damos aquí una ilustración de este hecho a través de una serie que según K. Weierstrass [Edg04] fue considerada por Riemann como un ejemplo de función continua no derivable en ningún punto. En realidad sí lo es en infinitos puntos, aunque esto no se probó hasta 1970 ¡y lo hizo un estudiante! La serie en cuestión es

$$R(x) = \sum_{n=1}^{\infty} \frac{\text{sen}(2\pi n^2 x)}{n^2}.$$

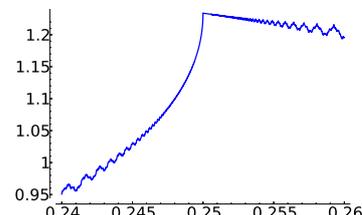
Una mirada cercana alrededor de los puntos $x = 1/2$, $1/3$ y $1/4$ muestra las siguientes curiosas gráficas



$R(x)$ en $[0.490, 0.510]$



$R(x)$ en $[0.323, 0.343]$



$R(x)$ en $[0.240, 0.260]$

Pues bien, para $x = a/b$ fracción irreducible, cualitativamente se obtiene, con un “zoom” suficientemente grande, una de estas tres situaciones. Si $b \equiv 2 \pmod{4}$, al igual que en la primera figura, hay una oscilación que se estrecha tanto que llega a haber tangente en x (por tanto es derivable en ese punto), la pendiente es de hecho siempre $-\pi$. Si $b \equiv 1, 3 \pmod{4}$, entonces siempre hay una tangente vertical (pendiente infinita) como en la segunda figura. En el caso restante, $b \equiv 0 \pmod{4}$, hay siempre una tangente vertical por un lado y no por el otro, como en la tercera figura. Saber por qué lado aparece la tangente vertical, o en el segundo caso, si hay una cuspide o no, requiere considerar propiedades aritméticas más finas de a y b [Dui91] relacionadas con la ley de reciprocidad cuadrática. Finalmente, en los irracionales, no hay patrones claros. Según varía la escala surgen diferentes configuraciones fractales que no se estabilizan y dependen de la aproximación por racionales del punto considerado.

La relación entre sumas oscilatorias y teoría de números actúa en ambos sentidos, a veces se usan métodos aritméticos para estudiar funciones como la anterior y a veces obtenemos resultados aritméticos con técnicas analíticas. Este último esquema es parte de un área muy amplia llamada *teoría analítica de números* de la que veremos algunos ejemplos.

2.3.1. El método de van der Corput

Si una función varía lentamente, todos sabemos por los cursos básicos de cálculo o de manera más precisa por la Proposición 1.2.8 que las sumas se aproximan por integrales. Sin embargo parece increíble que se salve de algún modo esta idea trabajando con sumas oscilatorias en las que cada término presenta una variación apreciable con respecto al anterior. El método introducido por J. van der Corput alrededor de 1920 se basa en llevar a cabo esta aproximación, aparentemente imposible, de una suma oscilatoria por una o varias integrales oscilatorias. Además se combina con un ingenioso argumento elemental debido a H. Weyl que actúa sobre las fases.

La idea de aproximar sumas oscilatorias por integrales oscilatorias es fácil de entender

cuando se consideran sumas regularizadas

$$S = \sum_n \varphi(n)e(f(n)) \quad \text{con} \quad \varphi \in C_0^\infty(\mathbb{R}).$$

La fórmula de sumación de Poisson (Teorema 1.2.4) implica

$$S = \sum_n \int_{-\infty}^{\infty} \varphi(x)e(f(x) - nx) dx$$

y en primera instancia tenemos S escrita como suma de infinitas integrales. En el soporte de φ , la derivada de la fase $F(x) = f(x) - nx$ será grande en valor absoluto cuando $|n|$ es grande y un argumento similar al empleado en la Proposición 2.2.1 permite deducir que la contribución es pequeña. En la cantidad finita de integrales que restan, se pueden usar los lemas de van der Corput (Proposición 2.2.4) o incluso alguna variante del principio de fase estacionaria para obtener una nueva suma trigonométrica.

En muchas aplicaciones en teoría de números, las sumas trigonométricas aparecen sin regularizar, a pesar de que en muchas ocasiones esto es más una costumbre que una necesidad teórica real. Ello causa que el argumento antes indicado requiera modificaciones fundamentales.

Proposición 2.3.1. *Sea $f \in C^2([a, b])$, $a, b \in \mathbb{Z}$ con f' monótona, y $\alpha, \beta \in \mathbb{Z}$ tales que $\alpha < f' < \beta$; entonces*

$$\sum_{\alpha \leq n \leq \beta} e(f(n)) = \sum_{\alpha \leq n \leq \beta} \int_a^b e(f(x) - nx) dx + O(\log(\beta - \alpha + 1)).$$

Demostración. Cambiando $f(n)$ por $f(n) - \alpha n$, lo cual no modifica el valor de la suma, podemos suponer $\alpha = 0$.

Integrando por partes en la Proposición 1.2.6 (cambiando n por $-n$), se tiene

$$(2.29) \quad \sum_{\alpha \leq n \leq \beta} e(f(n)) = \int_a^b e(f(x)) dx + \sum_{n \neq 0} \frac{1}{n} \int_a^b f'(x)e(g(x)) dx + O(1),$$

con $g(x) = f(x) - nx$.

Si $n > \beta$ o $n < \alpha = 0$, se tiene $g'(x) \neq 0$ en $[a, b]$ y se puede integrar por partes para obtener

$$2\pi \left| \int_a^b f'(x)e(f(x) - nx) dx \right| \leq \left| \frac{f'(a)}{g'(a)} \right| + \left| \frac{f'(b)}{g'(b)} \right| + \int_a^b |(f'/g')'| \leq \frac{2\beta}{|\beta - n|},$$

donde para la última desigualdad se ha usado que f'/g' es monótona (ya que es composición de f' y una función creciente) y por tanto se pueden meter la última integral

dentro del valor absoluto, como en la prueba de la Proposición 2.2.4. Nótese que esto asegura la convergencia absoluta de la serie en (2.29). Según lo anterior, la contribución de los términos con $n > \beta$ y $n < 0$ está acotada por

$$\sum_{n=-\infty}^{-1} \frac{\beta}{n(n-\beta)} + \sum_{n=\beta+1}^{\infty} \frac{\beta}{n(n-\beta)} \ll \sum_{|n|>2\beta} \frac{1}{n^2} + \sum_{0 \neq |n| \leq 2\beta} \frac{1}{|n|} \ll \log(\beta+2),$$

donde se han aproximado sumas por integrales. Sustituyendo en (2.29) se tiene

$$\sum_{a \leq n \leq b} e(f(n)) = \int_a^b e(f(x)) dx + \sum_{0 < n \leq \beta} \frac{1}{n} \int_a^b f'(x) e(g(x)) dx + O(\log(\beta+2)).$$

Ahora bien,

$$\int_a^b f'(x) e(g(x)) dx - n \int_a^b e(g(x)) dx = \frac{e(g(b)) - e(g(a))}{2\pi i} = O(1)$$

y despejando y sustituyendo $\int f'(x) e(g(x)) dx$ se obtiene el resultado. \square

La condición $\alpha, \beta \in \mathbb{Z}$ se ha introducido en el enunciado para simplificarlo (véase en [IK04] la forma habitual) y tiene como contrapartida que alguna de las integrales pueden ser incorporadas fácilmente en el término de error. Para incidir sobre este punto separamos un caso particular.

Corolario 2.3.2. *Con la notación de la Proposición 2.3.1, si $|f'| \leq 1/2$ en $[a, b]$, entonces*

$$\sum_{a \leq n \leq b} e(f(n)) = \int_a^b e(f(x)) dx + O(1).$$

Demostración. Basta aplicar la Proposición 2.3.1 con $\alpha = -1$, $\beta = 1$ y estimar las integrales correspondientes a $n = \pm 1$ con el caso $k = 1$ de la Proposición 2.2.4. \square

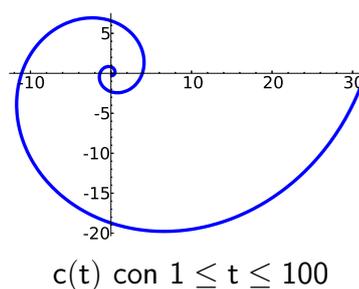
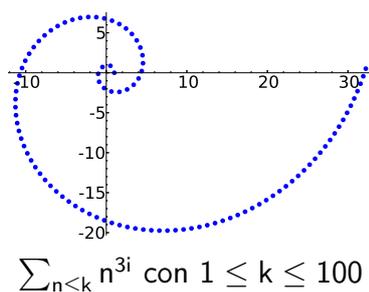
Un ejemplo correspondiente a $f(n) = \frac{3}{2\pi} \log n$ es

$$\sum_{n=1}^N n^{3i} = \frac{N^{1+3i}}{1+3i} + O(1).$$

Entonces la gráfica de las sumas parciales de $\sum n^{3i}$ hasta N debe aproximarse por la curva parametrizada

$$c(t) = \left(\frac{t}{10} \cos(3 \log t) + \frac{3t}{10} \operatorname{sen}(3 \log t), -\frac{3t}{10} \cos(3 \log t) + \frac{t}{10} \operatorname{sen}(3 \log t) \right)$$

que es una *espiral equiangular* (reparametrizada en polares, $r = Ce^\theta$).



Una vez que con la Proposición 2.3.1 sabemos aproximar sumas oscilatorias por integrales oscilatorias, estamos en situación de aplicar técnicas como las de la anterior sección. El empeño en no regularizar sólo deja como opciones directas aplicar los lemas de van der Corput (Proposición 2.2.4). El siguiente resultado es el más emblemático en este sentido. Es posible dar una prueba elemental sin pasar por la Proposición 2.3.1 con un uso ingenioso de la sumación por partes, aunque no seguiremos esa línea aquí (véase [GK91]).

Teorema 2.3.3. *Sea $f \in C^2([a, b])$ con $a, b \in \mathbb{Z}$. Si $0 < \lambda \ll |f''| \ll \lambda$ entonces*

$$\sum_{a \leq n \leq b} e(f(n)) \ll (b-a)\lambda^{1/2} + \lambda^{-1/2}.$$

Demostración. El teorema del valor medio aplicado a f' , asegura que en la Proposición 2.3.1 se pueden elegir α y β con $\beta - \alpha \ll \lambda(b-a) + 1$. Según la Proposición 2.2.4 con $k = 2$, cada una de las integrales es $O(\lambda^{-1/2})$, de forma que

$$\sum_{a \leq n \leq b} e(f(n)) \ll (\lambda(b-a) + 1)\lambda^{-1/2} + \log(\lambda(b-a) + 2).$$

Si $\lambda \leq 1$ el primer sumando domina al logaritmo, mientras que para $\lambda > 1$ el teorema es trivial. \square

Si por ejemplo aplicamos el resultados con $[a, b] = [1, N]$ y $f(n) = n^2/N$, se tiene una acotación para las sumas de Gauss del orden correcto según la Proposición 1.2.7, además se deduce que las sumas parciales de las sumas de Gauss también son $O(\sqrt{N})$, lo cual no es intuitivo en absoluto.

La condición $\lambda \ll |f''| \ll \lambda$, es decir, que f'' presente valores comparables en todo el intervalo es típica y muchas forzada tras una subdivisión en intervalos diádicos. La debilidad del resultado aparece cuando λ es muy pequeño o cuando es poco pequeño, concretamente si $\lambda \ll (b-a)^{-2}$ o $\lambda \gg 1$. el primer caso no es tan importante porque indica que f es prácticamente una función lineal. A pesar de que podemos dar fases como $f(n) = n^2$ que se ajustan al segundo caso para las que la única acotación es la trivial, en principio no hay razón para suponer que si $e(f(n))$ oscila realmente, la

condición $\lambda \gg 1$ impida la cancelación. Así por ejemplo, esperamos que haya cancelación en $\sum e(n^{5/2})$. La idea para tratar estas sumas que oscilan demasiado como para que el Teorema 2.3.3 sea útil proviene de un trabajo anterior de Weyl de principios del siglo XX. Se basa simplemente en subdividir el intervalo de sumación en subintervalos y calcular su módulo al cuadrado. Con ello se obtendrán incrementos de las fases y la oscilación será menor. Graduando la longitud de los intervalos, el parámetro H en el siguiente resultado, tendremos cierto control sobre la magnitud de la oscilación.

Lema 2.3.4. *Dada una función $f : [a, b] \rightarrow \mathbb{R}$ con $a, b \in \mathbb{Z}$, para cualquier entero positivo $H \leq b - a$ se tiene*

$$\left| \sum_{a \leq n \leq b} e(f(n)) \right|^2 \leq \frac{4(b-a)^2}{H} + \frac{4(b-a)}{H} \sum_{1 \leq r < H} \left| \sum_{a \leq n \leq b-r} e(f(n+r) - f(n)) \right|.$$

Demostración. Para cada $n \in \mathbb{Z}$ consideremos el intervalo $I_n = [1, H] \cap [a - n, b - n]$. El paso inicial, a veces llamado *Weyl shift*, es el artificio aparentemente inútil de repetir la suma H veces

$$|S| = \left| \sum_{a \leq n \leq b} e(f(n)) \right| = \left| \frac{1}{H} \sum_n \sum_{m \in I_n} e(f(m+n)) \right| \leq \frac{1}{H} \sum_{a-H \leq n \leq b-1} \left| \sum_{m \in I_n} e(f(m+n)) \right|.$$

La desigualdad de Cauchy-Schwarz produce los incrementos de f

$$|S|^2 \leq \frac{b-a+H}{H^2} \sum_{a-H \leq n \leq b-1} \sum_{m \in I_n} \sum_{l \in I_n} e(f(m+n) - f(l+n)).$$

Separando el término diagonal $l = m$ y notando que intercambiar l y m sólo conjuga los sumandos; al cambiar el orden de sumación se tiene:

$$|S|^2 \leq \frac{b-a+H}{H^2} \left(H(b-a+H) + 2 \sum_{1 \leq l < m \leq H} \left| \sum_{a-l \leq n \leq b-m} e(f(m+n) - f(l+n)) \right| \right).$$

Ahora basta renombrar $l+n \mapsto n$, $m-l \mapsto r$ y emplear $H \leq b-a$. \square

La estrategia es ahora clara, si una suma oscila demasiado, entonces le aplicamos unas cuantas veces el Lema 2.3.4 para terminar usando el Teorema 2.3.3. Con esta idea se obtiene el siguiente resultado que se reduce a Teorema 2.3.3 cuando $k = 2$.

Teorema 2.3.5. *Sea $f \in C^k([a, b])$ para cierto $k \geq 2$ con $a, b \in \mathbb{Z}$. Supongamos que $\lambda \leq |f^{(k)}(x)| \leq h\lambda$ con $h, \lambda \in \mathbb{R}^+$, entonces*

$$\sum_{a \leq n \leq b} e(f(n)) \ll h^{2/K} (b-a) \lambda^{1/(2K-2)} + (b-a)^{1-2/K} \lambda^{-1/(2K-2)} \quad \text{donde } K = 2^{k-1}.$$

Demostración. Suponemos $\lambda > 1$ ya que en otro caso la cota es trivial.

Procedemos por inducción en k . Como ya hemos observado, el caso $k = 2$ es el Teorema 2.3.3. Si suponemos el resultado cierto hasta $k - 1$, se tendrá

$$\sum_{a \leq n \leq b-r} e(f(n+r) - f(n)) \ll h^{2/K'} (b-a)(rh\lambda)^{1/(2K'-2)} + (b-a)^{1-2/K'} (r\lambda)^{-1/(2K-2)}$$

con $K' = K/2$, porque la derivada $k - 1$ -ésima de $f(n+r) - f(n)$ está entre $r\lambda$ y $hr\lambda$, por el teorema del valor medio. Introduciendo esta cota en el Lema 2.3.4 y eligiendo como H la parte entera de $\lambda^{-1/(K-1)}$, después de hacer algunas operaciones se obtiene el resultado. \square

Por ejemplo con $k = 3$ y $k = 4$ se obtiene

$$\sum_{N \leq n \leq 2N} e(n^{5/2}) \ll N^{11/12} \quad \text{y} \quad \sum_{N \leq n \leq 2N} e(n^{5/2}) \ll N^{25/28}$$

mientras que el Teorema 2.3.3 daba la acotación trivial. Valores mayores de k no producen ninguna mejora sobre la última cota.

Es posible rescatar la idea del principio de fase estacionaria para integrales en intervalos acotados. Por dar una idea del tipo de resultados (véase [Ivi03], [IK04] y [Mon94]), mencionaremos por ejemplo que si $f \in C^4([a, b])$ con $0 < \lambda \leq f'' < 1$ tiene un punto estacionario $f'(c) = 0$ con $\min(b-c, c-a) \gg b-a$ y $|f^{(k)}|(b-a)^{k-2} \ll \lambda$ para $k = 3$ y $k = 4$, entonces

$$\int_a^b e(f(x)) dx = \frac{e(f(c) + 1/8)}{\sqrt{f''(c)}} + O(\lambda^{-1}(b-a)^{-1}).$$

Si utilizamos resultados de este tipo en la Proposición 2.3.1 habremos transformado la suma inicial en una nueva suma. Esto es lo que se llama *Proceso B*. Por otra parte, también el Lema 2.3.4 permite pasar de una suma a otra, es lo que se llama *Proceso A*. Combinando ambos procesos para terminar aplicando la estimación trivial, completa el método de van der Corput. De cara a la presentación de esta idea, el método de *pares de exponentes* [GK91] describe de forma sencilla y elegante cómo actúan estos procesos en una suma oscilatoria bajo ciertas condiciones (bastante estrictas). De hecho existe un algoritmo que permite hallar la mejor manera de combinar ambos procesos para una suma dada. Las mejoras son habitualmente cualitativamente pequeñas con respecto de la cota del Teorema 2.3.5. Por ejemplo, la suma $S = \sum_{n < \sqrt{t}} n^{it}$ es relevante en el estudio de la función ζ (véase [Ivi03]). Si aplicamos para acotarla el Teorema 2.3.5 con $k = 3$, tras dividir en intervalos diádicos se tiene

$$S = \sum_{N=2^j \leq \sqrt{t}} \sum_{\substack{N \leq n < 2N \\ n < \sqrt{t}}} e\left(\frac{t}{2\pi} \log n\right) \ll \sum_{N=2^j \leq \sqrt{t}} (N(t/N^3)^{1/6} + N^{1/2}(t/N^3)^{-1/6}) \ll t^{5/12}.$$

Por otro lado, la mejor combinación de los procesos A y B reduce el exponente $5/12$ en poco más de 2 milésimas

2.3.2. La gran criba

Si se consideran sumas oscilatorias con cierta amplitud suave, $\sum A(n)e(f(n))$, podemos sumar por partes y trabajar con sumas sin amplitud. Sin embargo, en diversas situaciones, especialmente en teoría de números, las amplitudes son funciones que varían de forma incontrolable, muchas veces con contenido aritmético. La aspiración sería una teoría de sumas oscilatorias con coeficientes arbitrarios $\sum a_n e(f(n))$ pero tal cosa no puede existir porque si los coeficientes al variar entran en “resonancia” con las fases acabarán con la cancelación, el caso extremo es $a_n = e(-f(n))$. En muchos problemas, aparecen familias de sumas con unos mismos coeficientes. Bajo la denominación de *gran criba* se recogen una serie de técnicas que explotan la idea de que si hay cierta “independencia” entre las fases de las diferentes sumas, entonces los coeficientes no pueden entrar en resonancia en todas ellas y aunque alguna suma pueda ser grande, en promedio serán pequeñas.

El nombre de gran criba requiere cierta explicación. Fue introducido en un trabajo de Yu.V. Linnik y seguramente su difusión debe mucho a que apareció en el título de un libro de E. Bombieri. El hecho es que en algunas de las aplicaciones originales (en especial en la de Linnik) la acotación de sumas oscilatorias se usaba para diseñar un método para estimar el tamaño de un conjunto que evitase muchas clases de congruencia (véase el Teorema 2.3.8), lo que se llama en general un *método de criba*, en recuerdo de la criba de Eratóstenes. El nombre ha permanecido pero es poco indicativo y confuso porque se aplica por igual a las acotaciones de sumas trigonométricas y a los propios métodos de criba.

Aquí trataremos el caso más habitual, el de fases lineales, en el que la independencia se expresa como una separación entre las frecuencias. La terminología en uso es bastante intuitiva: Se dice que x_1, x_2, \dots están δ -espaciados en \mathbb{T} , con $0 < \delta < 1$, si al considerarlos como números reales, cumplen $\delta \leq \|x_\nu - x_\mu\|$ para $\nu \neq \mu$ donde $\|\cdot\|$ indica la distancia al entero más cercano. Esto es como decir que estén a distancia al menos δ unos de otros cuando pensamos \mathbb{T} enrollado como una circunferencia de longitud 1.

Proposición 2.3.6 (Desigualdad de gran criba). *Existe una constante absoluta C tal que si $x_1, x_2, \dots \in \mathbb{T}$ están δ -espaciados se cumple*

$$\sum_{\nu} \left| \sum_{n=1}^N a_n e(nx_\nu) \right|^2 \leq C(N + \delta^{-1}) \sum_{n=1}^N |a_n|^2$$

para cualesquiera $a_1, a_2, \dots, a_n \in \mathbb{C}$. De hecho, se puede tomar $C = 1$.

La complicación artificial del enunciado en lo que respecta a la constante, tiene una explicación. Deducir la desigualdad con constante uno, que es óptima, es bastante más duro y no es necesario para muchas aplicaciones, por ello probaremos en primer lugar la Proposición 2.3.6 para una constante indeterminada y después obtendremos el caso $C = 1$ dando por supuesto cierto resultado.

La prueba con la constante indeterminada la obtendremos con un argumento analítico elemental que se debe a P.X. Gallagher. Siguiendo el razonamiento con un poco de cuidado, se podría obtener $C = \pi$ como constante válida.

Demostración de la Proposición 2.3.6 con C indeterminada. Si $I \subset \mathbb{R}$ es un intervalo, para todo $x, y \in I$ se tiene $|f(x)| \leq |f(y)| + \int_I |f'|$, simplemente por la regla de Barrow. Al tomar y de modo que $|f|$ alcance un mínimo, se sigue

$$|f(x)| \leq \frac{1}{|I|} \int_I |f| + \int_I |f'|$$

que puede considerarse como una desigualdad de Sobolev.

Identificando \mathbb{T} con el intervalo $[0, 1)$ y subdividiéndolo en intervalos semiabiertos I_k de longitud $\delta/2$, en cada uno de ellos a lo más hay un x_ν . Ahora elegimos $f = g^2$ con $g(x) = \sum_{n=1}^N a_n e(nx)$ en la desigualdad anterior para $I = I_k$ y sumamos en k .

$$\sum_{\nu} \left| \sum_{n=1}^N a_n e(nx_{\nu}) \right|^2 \ll \sum_k \left(\delta^{-1} \int_{I_k} |g|^2 + \int_{I_k} |gg'| \right) = \delta^{-1} \int_0^1 |g|^2 + \int_0^1 |gg'|.$$

La desigualdad de Cauchy-Schwarz implica

$$\sum_{\nu} \left| \sum_{n=1}^N a_n e(nx_{\nu}) \right|^2 \ll \delta^{-1} \int_0^1 |g|^2 + \left(\int_0^1 |g'|^2 \right)^{1/2} \left(\int_0^1 |g|^2 \right)^{1/2}$$

y la identidad de Parseval concluye la prueba. \square

Una forma de obtener $C = 1$ es por medio de una variante de la *desigualdad de Hilbert* obtenida por H.L. Montgomery y R.C Vaughan [MV74]. Otra forma, desarrollada por A. Selberg, se basa en cierta manera en encontrar la mejor mayorante de una función característica que sea una función entera con ciertas propiedades (véase [Mon78] y [Mur08]).

En su forma original la desigualdad de Hilbert establece

$$\left| \sum_{n \neq m} \sum \frac{a_n \bar{a}_m}{n - m} \right| \leq \pi \quad \text{cuando} \quad \sum |a_n|^2 = 1.$$

Aunque la prueba es elemental, es bastante ingeniosa. Una aplicación (no directa) debida a M. Riesz es la acotación del operador conjugado en el disco. Dada una función holomorfa en el disco unidad, si el valor en el borde de su parte real está en $L^p(\mathbb{T})$ para cierto

$1 < p < \infty$, entonces el de su parte imaginaria también. La variante de Montgomery y Vaughan es:

Lema 2.3.7. *Si $x_1, x_2, \dots \in \mathbb{T}$ están δ -espaciados, se cumple*

$$\left| \sum_{\nu \neq \mu} \frac{c_\nu \bar{c}_\mu}{\text{sen}(\pi(x_\nu - x_\mu))} \right| \leq \delta^{-1} \quad \text{cuando} \quad \sum |c_\nu|^2 = 1.$$

Para la prueba, véase [MV74] o [IK04].

Demostración de la Proposición 2.3.6 con $C = 1$. Consideremos un vector cuya coordenada ν es $\sum_n \bar{a}_n e(-nx_\nu)$. Sea b_ν el valor de esta coordenada cuando se normaliza el vector, entonces

$$\sum_{\nu} \left| \sum_{n=1}^N a_n e(nx_\nu) \right|^2 = \left(\sum_{\nu} b_\nu \sum_{n=1}^N a_n e(nx_\nu) \right)^2 \leq \sum_{n=1}^N |a_n|^2 \cdot \sum_{n=1}^N \left| \sum_{\nu} b_\nu e(nx_\nu) \right|^2.$$

Desarrollando el cuadrado lo que resta probar es

$$\sum_{\nu \neq \mu} \sum_{n=1}^N b_\nu \bar{b}_\mu e(n(x_\nu - x_\mu)) \leq \delta^{-1}.$$

Para ello utilizamos la identidad

$$\sum_{n=1}^N e(n(x_\nu - x_\mu)) = \frac{e((N+1/2)x_\nu) e(-(N+1/2)x_\mu)}{2i \text{sen}(\pi(x_\nu - x_\mu))} - \frac{e(x_\nu/2) e(-x_\mu/2)}{2i \text{sen}(\pi(x_\nu - x_\mu))}.$$

Por el Lema 2.3.7 con $c_\nu = b_\nu e((N+1/2)x_\nu)$, la primera fracción contribuye a la suma triple con una cantidad acotada en módulo por $(2\delta)^{-1}$ y lo mismo ocurre con la segunda fracción tomando $c_\nu = b_\nu e(x_\nu/2)$. \square

Ahora ya estamos preparados para enunciar la versión combinatoria de la gran criba. Recordamos de la teoría de números elemental que se dice que un entero positivo es *libre de cuadrados* si no es divisible por ningún cuadrado perfecto mayor que 1. La función característica de los números libres de cuadrados es $\mu^2(n)$ donde μ es la *función de Möbius* definida como

$$\mu(n) = \begin{cases} (-1)^k & \text{si } n = p_1 p_2 \cdots p_k \text{ con } p_j \text{ primos distintos} \\ 1 & \text{si } n = 1 \\ 0 & \text{en otro caso} \end{cases}$$

Si no se usa la desigualdad criba con $C = 1$, entonces la conclusión en el siguiente resultado se mantiene reemplazando \leq por \ll .

Teorema 2.3.8 (La gran criba como método de criba). *Sea \mathcal{Z} un conjunto de enteros $\mathcal{Z} \subset [M+1, M+N]$, $N, M \in \mathbb{Z}^+$. Supongamos que para cada primo $p \leq Q$, hay $\omega(p)$ clases módulo p que no aparecen en \mathcal{Z} , entonces*

$$|\mathcal{Z}| \leq \frac{N+Q^2}{\sum_{q \leq Q} h(q)} \quad \text{con} \quad h(q) = \mu^2(q) \prod_{p|q} \frac{\omega(p)}{p - \omega(p)}.$$

Demostración. Lo que vamos a probar es

$$(2.30) \quad h(q)|\mathcal{Z}|^2 \leq \sum_{\substack{a=1 \\ (a,q)=1}}^q \left| \sum_{n \in \mathcal{Z}} e\left(\frac{an}{q}\right) \right|^2 \quad \text{para } q \leq Q \text{ libre de cuadrados.}$$

Veamos que esto es suficiente para conseguir nuestro objetivo. Sea $a_n = 1$ si $n+M \in \mathcal{Z}$ y $a_n = 0$ en otro caso, entonces

$$\sum_{q \leq Q} \sum_{\substack{a=1 \\ (a,q)=1}}^q \left| \sum_{n \in \mathcal{Z}} e\left(\frac{an}{q}\right) \right|^2 = \sum_{\nu} \left| \sum_{n=1}^N a_n e((n+M)x_{\nu}) \right|^2 = \sum_{\nu} \left| \sum_{n=1}^N a_n e(nx_{\nu}) \right|^2$$

donde x_{ν} recorre las fracciones irreducibles $0 < a/q < 1$, y estas fracciones están claramente Q^{-2} -espaciadas. Entonces el resultado se sigue de la Proposición 2.3.6 sumando en (2.30).

Ahora deduciremos (2.30) por inducción en el número de factores primos de q . Si $q = 1$, es decir, si no tiene factores primos, el resultado es trivial (interpretando $h(1) = 1$ por el producto vacío). Supongamos ahora que el resultado es cierto para un libre de cuadrados $q \leq Q$ y vamos a deducirlo para qp con $p \nmid q$ un primo $p \leq Q$.

Según las hipótesis, sólo se obtienen $p - \omega(p)$ clases al reducir módulo p los elementos de \mathcal{Z} , sean $b_1, b_2, \dots, b_{p-\omega(p)}$ representantes de estas clases. Por la desigualdad de Cauchy-Schwarz

$$\left| \sum_{n \in \mathcal{Z}} e\left(\frac{bn}{q}\right) \right|^2 = \left| \sum_{j=1}^{p-\omega(p)} \sum_{\substack{n \in \mathcal{Z} \\ n \equiv b_j \pmod{p}}} e\left(\frac{bn}{q}\right) \right|^2 \leq (p - \omega(p)) \sum_{k=1}^p \left| \sum_{\substack{n \in \mathcal{Z} \\ n \equiv k \pmod{p}}} e\left(\frac{bn}{q}\right) \right|^2.$$

Recordando (1.28), al expandir el cuadrado el último término es

$$(p - \omega(p)) \sum_{n \in \mathcal{Z}} \sum_{m \in \mathcal{Z}} \frac{1}{p} \sum_{a=1}^p e\left(\frac{a(n-m)}{p}\right) e\left(\frac{b(n-m)}{q}\right) = \frac{p - \omega(p)}{p} \sum_{a=1}^p \left| \sum_{n \in \mathcal{Z}} e\left(\frac{an}{p}\right) e\left(\frac{bn}{q}\right) \right|^2.$$

Separando la contribución de $a = p$, despejando y sumando en b , se deduce

$$h(p) \sum_{\substack{b=1 \\ (b,q)=1}}^q \left| \sum_{n \in \mathcal{Z}} e\left(\frac{bn}{q}\right) \right|^2 \leq \sum_{\substack{a=1 \\ (a,p)=1}}^p \sum_{\substack{b=1 \\ (b,q)=1}}^q \left| \sum_{n \in \mathcal{Z}} e\left(\frac{an}{p}\right) e\left(\frac{bn}{q}\right) \right|^2 = \sum_{\substack{a=1 \\ (a,pq)=1}}^{pq} \left| \sum_{n \in \mathcal{Z}} e\left(\frac{an}{pq}\right) \right|^2$$

donde la igualdad se debe al teorema chino del resto. Por la hipótesis de inducción, esto implica

$$h(p)h(q)|\mathcal{Z}|^2 \leq \sum_{\substack{a=1 \\ (a,pq)=1}}^{pq} \left| \sum_{n \in \mathcal{Z}} e\left(\frac{an}{pq}\right) \right|^2$$

que es (2.30) para qp . □

En seguida veremos aplicaciones significativas del resultado anterior. Para apreciar su sentido, consideraremos un ejemplo trivial. Supongamos que queremos contar los enteros positivos en $[1, N]$ que no son ni múltiplos de 2 ni de 3. Un razonamiento elemental lleva a que hay $N/6 + O(1)$ de ellos. Con la notación del Teorema 2.3.8, $\omega(2) = \omega(3) = 1$ que implica $h(2) = 1$ y $h(3) = 1/2$, y $Q = 3$ da que el cardinal buscado es a lo más $2N/5 + O(1)$. Esto es más débil que la asintótica elemental pero pensemos que cuando incrementamos la cantidad de primos el argumento elemental se complica y empieza a fallar si esta cantidad depende de N . La “gran” criba es superior cuando se considera una “gran” cantidad de primos.

2.3.3. Algunas aplicaciones aritméticas

Comencemos con un problema con sabor geométrico consistente en contar los puntos de coordenadas enteras bajo la gráfica de una función no negativa. Concretamente, queremos aproximar

$$\mathcal{N}_f(A, B) = \{(n, m) \in \mathbb{Z}^2 : A \leq n \leq B, 0 \leq m \leq f(n)\}.$$

Este problema tiene raíces muy antiguas que llegan hasta Gauss y Dirichlet pero desde el punto de vista moderno, están ligados a I.M. Vinogradov que es probablemente el matemático que más ha contribuido a la estimación de sumas trigonométricas.

No se pierde generalidad suponiendo $A, B \in \mathbb{Z}$. Es fácil intuir que la integral es una buena aproximación, con alguna modificación para tener en cuenta los bordes rectos. Por supuesto, se necesita alguna condición sobre f porque en otro caso los resultados pueden ser muy variados. Por ejemplo, si f es constante $\mathcal{N}_f(A, B)$ da saltos de $B - A + 1$ según esa constante es algo menor o algo mayor que un entero positivo. Algo crucial para evitar esta situación es que la gráfica de f tenga cierta curvatura. Nosotros aquí tendremos en mente el caso de una familia de gráficas que se va expandiendo por una homotecia de razón λ , es decir, $f(x) = \lambda g(x/\lambda)$. Si g tiene una curvatura acotada, esto implica que f'' es comparable a λ^{-1} . Como los extremos del intervalo, A y B , se dilatan de la misma manera la condición natural con este ejemplo en mente es que f'' sea comparable a $(B - A)^{-1}$.

Teorema 2.3.9. Sea $f \in C^2([A, B])$ no negativa con $A < B$ enteros. Supongamos que $1 \ll (B - A)f'' \ll 1$, entonces

$$\mathcal{N}_f(A, B) = \int_A^B f(x) dx + \frac{1}{2}(B - A + f(A) + f(B)) + O((B - A)^{2/3}).$$

Demostración. Partimos de la identidad

$$\mathcal{N}_f(A, B) = \frac{1}{2}(B - A + 1) + \sum_{n=A}^B f(n) - \sum_{n=A}^B \{f(n)\} \quad \text{donde } \{x\} = x - [x] - \frac{1}{2}.$$

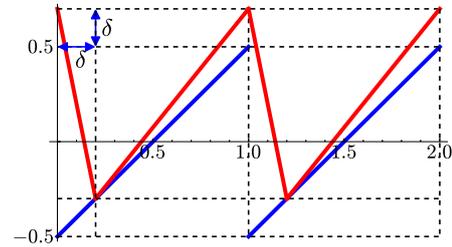
Por la fórmula de sumación de Euler-Maclaurin, Proposición 1.2.8, con $K = 2$, aplicada a la primera suma, se deduce

$$\mathcal{N}_f(A, B) = \int_A^B f(x) dx + \frac{1}{2}(B - A + f(A) + f(B)) - \sum_{n=A}^B \{f(n)\} + O(1).$$

Entonces todo lo que hay que ver es que la suma restante es $O((B - A)^{2/3})$. Sustituyendo $\{x\}$ por su serie de Fourier, se obtiene una suma trigonométrica pero la convergencia es un problema y es conveniente regularizar antes de aplicar ningún método.

Sea ϕ la función 1-periódica que conecta con segmentos $(0, 1/2 + \delta)$, $(\delta, -1/2 + \delta)$ y $(1, 1/2 + \delta)$ para cierto $0 < \delta \leq 1/2$.

$$\phi(x) = \begin{cases} -\frac{x}{\delta} + \frac{1}{2} + \delta & \text{si } 0 \leq x \leq \delta \\ \frac{x-1}{1-\delta} + \frac{1}{2} + \delta & \text{si } \delta \leq x \leq 1 \end{cases}$$



La fórmula para ϕ en realidad no es importante, lo único que emplearemos es que $\int_0^1 \phi = \delta$ y que ϕ' es $-\delta^{-1}$ para $0 < x < \delta$ y $(1 - \delta)^{-1}$ para $\delta < x < 1 - \delta$, lo cual se deduce del dibujo. Con ello, integrando por partes, se tiene que los coeficientes de Fourier son

$$a_0 = \delta \quad \text{y} \quad a_n = \frac{1}{2\pi i n} \int_0^1 \phi'(x) e(-nx) dx = \frac{1 - e(-n\delta)}{4\pi^2 n^2 \delta (1 - \delta)} \quad \text{si } n \neq 0.$$

Está claro que $-\phi(-x) \leq \{x\} \leq \phi(x)$, por tanto para cierta elección del signo \pm independiente de las variables de sumación,

$$\sum_{n=A}^B \{f(n)\} \ll (B - A)\delta + \left| \sum_{m \neq 0} \frac{1 - e(-m\delta)}{m^2 \delta} \sum_{n=A}^B e(\pm m f(n)) \right|.$$

Por el Teorema 2.3.3 con $\lambda = m(B - A)^{-1}$, se tiene

$$\sum_{n=A}^B \{f(n)\} \ll (B - A)\delta + (B - A)^{1/2}\delta^{-1} \sum_{m=1}^{\infty} \frac{|1 - e(-m\delta)|}{m^{3/2}}.$$

Si empleamos $|1 - e(-m\delta)| \ll m\delta$ para $\delta|m| < 1$ (Taylor) y $|1 - e(-m\delta)| \leq 2$ en el resto, deducimos que la suma es $O(\delta^{1/2})$. Escogiendo $\delta = \frac{1}{2}(B - A)^{-1/3}$ se termina la prueba. \square

En teoría de números es muy común estimar sumas o promedios de funciones aritméticas. Si nos fijamos en $r_2(n)$, el número de representaciones como suma de dos cuadrados, entonces la suma se relaciona con el número de puntos de coordenadas enteras en un círculo.

$$\sum_{n \leq R^2} r_2(n) = \mathcal{N}(R) \quad \text{con} \quad \mathcal{N}(R) = \#\{(n, m) \in \mathbb{Z}^2 : n^2 + m^2 \leq R^2\}.$$

Razonando geoméricamente, es fácil ver que $\mathcal{N}(R)$ es como el área del círculo salvo un error $O(R)$. Con el Teorema 2.3.9, considerando la parte del círculo en $[0, R/\sqrt{2}] \times \mathbb{R}^+$ y usando las simetrías, se refina este resultado hasta

$$(2.31) \quad \mathcal{N}(R) = \pi R^2 + O(R^{2/3}).$$

Un antiguo problema llamado el *problema del círculo de Gauss* consiste en determinar el exponente óptimo en el término de error. La conjetura (*conjetura de Hardy*) es que $2/3$ se puede sustituir por cualquier $\sigma > 1/2$ y se conoce que el exponente $1/2$ no es válido.

Utilizando métodos más precisos que los empleados en el Teorema 2.3.9 vamos a mejorar (2.31).

Teorema 2.3.10. *Sea $\mathcal{N}(R)$ el número de puntos de \mathbb{Z}^2 en el interior de un círculo de radio R , entonces*

$$\mathcal{N}(R) = \pi R^2 + O(R^{27/41}).$$

Nótese que pasar de $2/3$ a $27/41$, cuantitativamente es una mejora de menos del 2%. El menor exponente conocido hasta la fecha supone una mejora de menos del 6% sobre $2/3$, lo cual es descorazonador, habida cuenta que entre ambos resultados han pasado prácticamente 100 años de estudio de sumas trigonométricas.

Los cálculos hechos para la difracción de Fraunhofer de una abertura circular prueban que si χ_R es la función característica de un círculo de radio R se tiene

$$(2.32) \quad \widehat{\chi}_R(n, m) = R \frac{J_1(2\pi R\sqrt{n^2 + m^2})}{\sqrt{n^2 + m^2}} \quad \text{para} \quad n^2 + m^2 \neq 0.$$

Si aplicamos “a ciegas” la versión en dos variables de la fórmula de sumación de Poisson (Teorema 1.2.4) entonces obtenemos la *serie de Hardy*

$$(2.33) \quad \mathcal{N}(R) = \pi R^2 + R \sum_{n^2+m^2 \neq 0} \frac{J_1(2\pi R\sqrt{n^2+m^2})}{\sqrt{n^2+m^2}},$$

la cual fue establecida por Hardy para $R > 0$ no entero, entendiendo la suma como límite sobre $n^2 + m^2 \leq M$.

La idea para demostrar el teorema anterior es sustituir J_1 por su fórmula asintótica (2.20) y utilizar el Teorema 2.3.5 con $k = 5$ para estimar la suma en n o en m . El problema para seguir esta idea es que (2.33) no converge absolutamente, de hecho la prueba de esta igualdad es bastante complicada y la lenta convergencia hace que sea poco útil. El problema proviene de que $\widehat{\chi}_R$ decae demasiado despacio y la solución pasa por regularizarla, así el esquema funciona como se ha explicado si se cambia (2.33) por la fórmula del Lema 2.3.13, que es más fea pero converge absolutamente. Antes de enunciar este resultado, separamos la estimación de la suma trigonométrica fundamental y un cálculo explícito de una transformada de Fourier que nos servirá para regularizar.

Lema 2.3.11. *Supongamos $N \leq a < b \leq 2N$ y $R^{-1/2}N^{33/16} \leq m \leq a$, entonces*

$$\sum_{a \leq n \leq b} e(R\sqrt{n^2+m^2}) \ll m^{1/15} R^{1/30} N^{4/5}.$$

Demostración. Un cálculo prueba que la derivada quinta de $f(n) = R\sqrt{n^2+m^2}$ satisface $\lambda \ll f^{(5)} \ll \lambda$ donde $\lambda = Rm^2N^{-6}$. El Teorema 2.3.5 con $k = 5$ implica en esta situación

$$\sum_{a \leq n \leq b} e(R\sqrt{n^2+m^2}) \ll m^{1/15} R^{1/30} N^{4/5} + m^{-1/15} R^{-1/30} N^{43/40}$$

y la condición $m \geq R^{-1/2}N^{33/16}$ asegura que el segundo sumando es superfluo. \square

Lema 2.3.12. *Para $\delta > 0$ sea la función $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por $\frac{3}{2\pi}\delta^{-3}(\delta^2 - x^2 - y^2)^{1/2}$ si $x^2 + y^2 \leq \delta^2$ y 0 en el resto, y sea $F(t) = 3t^{-3}(\sin t - t \cos t)$ para $t > 0$. Entonces*

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e(-x\xi - y\eta) dx dy = F(2\pi\delta\sqrt{\xi^2 + \eta^2}).$$

En particular, la integral de f es $\lim_{t \rightarrow 0^+} F(t) = 1$.

Demostración. Como ya se vio en el caso de J_1 , geoméricamente parece claro que la integral debe ser una función radial de (ξ, η) . Se deja al lector curioso el interesante (y nada inmediato) ejercicio de probarlo analíticamente. Teniendo esto en cuenta, hay que demostrar

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e(-x\delta^{-1}R) dx dy = F(2\pi R) \quad \text{para } R > 0.$$

Con el cambio $(x, y) \mapsto (\delta x, \delta y)$, la integral es

$$\frac{3}{2\pi} \int_{-1}^1 e(-Rx) \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \sqrt{1-x^2-y^2} dy dx = \frac{3}{2\pi} \int_{-1}^1 e(-Rx) \frac{\pi}{2} (1-x^2) dx,$$

donde se ha empleado la integral que da el área del semicírculo. La igualdad se sigue ahora integrando por partes dos veces de la manera habitual. \square

Ahora vamos por fin con la versión regularizada de (2.33). Bajo la filosofía del principio de incertidumbre, consideramos una serie que imita a (2.33) para frecuencias menores que δ^{-1} pagando el precio de ver borroso en una corona de anchura δ del círculo de radio R . Su área es la que crea el término de error $O(R\delta)$. Cuando menor sea δ , menor será este error pero la serie tendrá más términos significativos.

Lema 2.3.13. *Dado $R > 2$, para cualquier $R^{-1/2} < \delta < 1$ existe $r \in [R - 2\delta, R + 2\delta]$ tal que*

$$\mathcal{N}(R) = \pi R^2 + \frac{r^{1/2}}{\pi} \sum_{n^2+m^2 \neq 0} \frac{\text{sen}(2\pi r \sqrt{n^2+m^2} - \pi/4)}{(n^2+m^2)^{3/4}} F(2\pi\delta\sqrt{n^2+m^2}) + O(R\delta).$$

Demostración. Como f en el Lema 2.3.12 tiene soporte en el círculo centrado de diámetro 2δ y su integral es 1, se cumple $\chi_{R-2\delta} * f \leq \chi_R \leq \chi_{R+2\delta} * f$, donde se ha usado la notación de (2.32) para la función característica del círculo y el asterisco significa la convolución en \mathbb{R}^2 (como en \mathbb{R} pero en ambas variables). Por la continuidad de $\chi_{R \pm 2\delta} * f$, el teorema de los valores intermedios asegura

$$\mathcal{N}(R) = \sum_{n,m \in \mathbb{Z}} \chi_R(n, m) = \sum_{n,m \in \mathbb{Z}} (\chi_r * f)(n, m) \quad \text{para algún } r \in [R - 2\delta, R + 2\delta].$$

Usando la fórmula de sumación de Poisson (Teorema 1.2.4) y la propiedad de la convolución (1.12) en sus versiones bidimensionales,

$$\mathcal{N}(R) = \sum_{n,m \in \mathbb{Z}} \widehat{\chi}_r(n, m) \widehat{f}(n, m) = \pi R^2 + O(R\delta) + \sum_{n^2+m^2 \neq 0} \widehat{\chi}_r(n, m) F(2\pi\delta\sqrt{n^2+m^2}).$$

Ahora basta recordar (2.32) y (2.20). El término de error es absorbido por $O(R\delta)$. \square

Demostración del Teorema 2.3.10. Vamos a elegir $\delta = R^{-14/41}$ en el Lema 2.3.13, con ello lo que buscamos es

$$\sum_{n^2+m^2 \neq 0} \frac{e(2\pi r \sqrt{n^2+m^2})}{(n^2+m^2)^{3/4}} F(2\pi\delta\sqrt{n^2+m^2}) \ll R^{13/82}.$$

Por la simetría, podemos restringir la suma a $0 < n < m$, ya que $m = 0$ o $n = m$ claramente sólo contribuyen $O(1)$. Sea $\mathfrak{N} = \{2^j m : j \geq 0\}$, entonces lo que hay que probar es

$$(2.34) \quad \sum_{m=1}^{\infty} \sum_{N \in \mathfrak{N}} S_N \ll R^{13/82} \quad \text{con} \quad S_N = \sum_{N < n \leq 2N} \frac{e(2\pi r \sqrt{n^2 + m^2})}{(n^2 + m^2)^{3/4}} F(2\pi \delta \sqrt{n^2 + m^2}).$$

Se deja como ejercicio comprobar que podemos restringir la suma a los N que satisfacen $m < R^{-1/2} N^{33/16}$, ya que para el resto la estimación trivial es suficiente.

La función $F(t)/t^{3/2}$ es decreciente para $0 < t < 4$, de donde para $N \leq \delta^{-1}/6$ se puede sumar por partes con el Lema 1.2.1 y después aplicar el Lema 2.3.11 obteniendo, para algún $N' \leq 2N$,

$$S_N \ll N^{-3/2} \sum_{N < n \leq N'} e(2\pi r \sqrt{n^2 + m^2}) \ll m^{1/15} R^{1/30} N^{-7/10}.$$

Si $N > \delta^{-1}/6$, escribiendo $\sin t = (e^{it} - e^{-it})/2i$ y $\cos t = (e^{it} + e^{-it})/2$ en $F(t)$, podemos agrupar estas funciones oscilatorias con la exponencial compleja en S_N reemplazando r por $r' = r \pm \delta$. Entonces, notando $F(t) \ll t^{-2}$, la sumación por partes da en este caso

$$S_N \ll N^{-3/2} (\delta N)^{-2} \sum_{N < n \leq N'} e(2\pi r' \sqrt{n^2 + m^2}) \ll m^{1/15} R^{1/30} N^{-27/10} \delta^{-2}.$$

Al sumar sobre la progresión geométrica $N \in \mathfrak{N}$, el resultado será comparable al primer término en cada caso, por tanto la contribución de toda la suma en (2.34) es

$$\sum_{m \ll \delta^{-1}} m^{1/15} R^{1/30} m^{-7/10} + \sum_{m \gg \delta^{-1}} m^{1/15} R^{1/30} m^{-27/10} \delta^{-2} \ll R^{1/30} \delta^{-11/30},$$

lo cual prueba (2.34) para el valor escogido de δ . □

Pasemos ahora a ver algunas aplicaciones de la gran criba. Comenzamos estimando la cantidad de primos en un intervalo.

Teorema 2.3.14. *Sea $\pi(N)$ el cardinal de los primos menores o iguales que N . Para todo $N, M \in \mathbb{Z}^+$ se cumple*

$$\pi(M + N) - \pi(M) \leq (2 + o(1)) \frac{N}{\log N}$$

donde $o(1)$ tiende a cero cuando $N \rightarrow \infty$.

El teorema de los números primos afirma que en el caso $M = O(1)$ se puede cambiar la igualdad con la constante 1 en lugar de 2. M.N. Huxley probó que esta igualdad también es correcta si $M = O(N^\sigma)$ para algún $\sigma < 12/7$.

Corolario 2.3.15. *Se tiene $\pi(N) = O(N/\log N)$.*

Demostración. Consideremos

$$\mathcal{Z} = \{M \leq n \leq M + N : n \not\equiv 0 \pmod{p} \quad \forall p \leq Q\}.$$

Es obvio que \mathcal{Z} contiene todos los primos que son mayores que Q , mientras que hay menos de Q que son menores que Q , por tanto

$$\pi(M + N) - \pi(M) \leq Z + Q.$$

Como se excluye una clase por cada primo, $h(p) = 1/(p-1) = p^{-1} + p^{-2} + p^{-3} + \dots$, con la notación del Teorema 2.3.8. Por tanto, para q libre de cuadrados, $h(q)$ es la suma de los inversos de todos los números con los mismos factores primos que q . Así pues

$$\sum_{q \leq Q} h(q) \geq \sum_{q \leq Q} \frac{1}{q} = \log Q + O(1).$$

Eligiendo por ejemplo $Q = \sqrt{N}/\log N$ en el Teorema 2.3.8 se obtiene el resultado. \square

Siguiendo esta idea de primos en intervalos, una conjetura muy antigua es si hay infinitos *primos gemelos*, esto es, primos cuya diferencia es 2. La gran criba permite dar una cota que es del orden de magnitud esperado y obtener lo que se llama el *teorema de Brun* que afirma que la serie de los inversos de los primos gemelos converge. V. Brun fue el pionero de los métodos de criba y su prueba es muy anterior a la introducción de la gran criba.

Teorema 2.3.16. *Sea \mathcal{G} el conjunto de los $n \in \mathbb{Z}^+$ tales que n y $n + 2$ son primos (gemelos) y sea $\pi_2(N)$ el cardinal de $\mathcal{G} \cap [1, N]$. Existen dos constantes positivas C_1 y C_2 tales que*

$$\pi_2(N) \leq C_1 \frac{N}{(\log N)^2} \quad y \quad \sum_{p \in \mathcal{G}} \frac{1}{p} = C_2.$$

En la demostración apelaremos al siguiente lema. Con técnicas bien conocidas de teoría analítica de números [IK04, §1] se deduce una fórmula asintótica.

Lema 2.3.17. *Sea $d(n)$ la función que cuenta el número de divisores (positivos) de n , entonces*

$$\sum_{n \leq N} \mu^2(n) d(n) \gg N \log N.$$

Demostración. Vamos a emplear que

$$(2.35) \quad \sum_{n \leq N} \mu^2(n) d(n) \geq \sum_{n \leq N} d(n) - \sum_{p \leq \sqrt{N}} \sum_{n \leq N/p^2} d(p^2 n).$$

Se tiene

$$\sum_{n \leq N} d(n) = \sum_{kl \leq N} 1 = \sum_{k \leq N} \left(\frac{N}{k} + O(1) \right) = N \log N + O(N).$$

Un razonamiento similar sirve para tratar $\sum_{n \leq N/p^2} d(p^2 n)$ imponiendo que k o l sean divisibles por p^2 o cada una de ellas por p . Módulo p^2 esto es $0 \cdot l$, $k \cdot 0$ y $p \cdot p$, lo que hacen $2p^2$ posibilidades entre las p^4 totales. En los casos $p = 2, 3$ se obtiene

$$\sum_{n \leq N/4} d(n) = \frac{1}{2} N \log N + O(N). \quad \text{y} \quad \sum_{n \leq N/9} d(n) = \frac{2}{9} N \log N + O(N).$$

Para el resto de los primos usaremos simplemente que $d(p^2 n) \leq 3d(n)$, porque cada divisor de $p^2 n$ dividido por 1, p o p^2 da un divisor de n . Entonces, deducimos a partir de (2.35)

$$\sum_{n \leq N} \mu^2(n) d(n) \geq \left(1 - \frac{1}{2} - \frac{2}{9} - 3 \sum_{p \geq 5} p^{-2} \right) N \log N + O(N).$$

Todo lo que resta comprobar es que $\sum_{p \geq 5} p^{-2} \leq 5/54 = 0.09259 \dots$. Para ello usamos que $p^{-2} \leq \log(1 + p^{-2}) + \frac{1}{2} p^{-4}$ y por tanto

$$\sum_{p \geq 5} p^{-2} \leq \log \prod_{p \geq 5} \frac{1 - p^{-4}}{1 - p^{-2}} + \frac{1}{2} \sum_{n \geq 5} n^{-4} = \log \frac{\zeta(2)}{\zeta(4)} - \log \frac{5}{4} - \log \frac{10}{9} + \frac{1}{2} \left(\zeta(4) - \sum_{n=1}^4 n^{-4} \right),$$

donde se ha empleado la fórmula producto de Euler (1.33). Con (1.19) podemos evaluar $\zeta(2)$ y $\zeta(4)$ y comprobar que la última expresión es algo menor que 0.092. \square

Demostración del Teorema 2.3.16. Sea

$$\mathcal{Z} = \{n \in [1, N] : n \not\equiv 0, -2 \pmod{p} \quad \forall p \leq \sqrt{N}\}.$$

Entonces, con la notación del Teorema 2.3.8, $h(2) = 1$ y para cada $2 < p \leq \sqrt{N}$ se tiene $h(p) = 2/(p-2)$. Entonces para $q \leq \sqrt{N}$

$$h(q) \geq \mu^2(q) \prod_{p|q} \frac{2}{p} = \frac{\mu^2(q) d(q)}{q}.$$

Sumando por partes en el Lema 2.3.17, se tiene

$$\sum_{q \leq \sqrt{N}} h(q) \gg (\log N)^2.$$

Evidentemente $\pi_2(N) \leq Z + \sqrt{N}$ y el Teorema 2.3.8 con $Q = [\sqrt{N}]$ prueba la primera parte del enunciado.

La segunda parte se deduce de la primera mediante sumación por partes. Concretamente, escogiendo en el Lema 1.2.2 $g(n) = 1/n$ y $c_n = 1$ si $n \in \mathcal{G}$ y $c_n = 0$ en otro caso. De esta forma $C(N) = \pi_2(N)$. \square

Bibliografía

- [Ahl78] L. V. Ahlfors. *Complex analysis*. McGraw-Hill Book Co., New York, third edition, 1978. An introduction to the theory of analytic functions of one complex variable, International Series in Pure and Applied Mathematics.
- [Art00] J. Arthur. Harmonic analysis and group representations. *Notices Amer. Math. Soc.*, 47(1):26–34, 2000.
- [BB09] S. A. Broughton and K. Bryan. *Discrete Fourier analysis and wavelets*. John Wiley & Sons Inc., Hoboken, NJ, 2009. Applications to signal and image processing.
- [BS08] J. Bryant and C. Sangwin. *How round is your circle?* Princeton University Press, Princeton, NJ, 2008. Where engineering and mathematics meet.
- [BW65] M. Born and E. Wolf. *Principles of optics: Electromagnetic theory of propagation, interference and diffraction of light*. With contributions by A. B. Bhatia, P. C. Clemmow, D. Gabor, A. R. Stokes, A. M. Taylor, P. A. Wayman and W. L. Wilcock. Third revised edition. Pergamon Press, Oxford, 1965.
- [Car66] L. Carleson. On convergence and growth of partial sums of Fourier series. *Acta Math.*, 116:135–157, 1966.
- [CT65] J. W. Cooley and J. W. Tukey. An algorithm for the machine calculation of complex Fourier series. *Math. Comp.*, 19:297–301, 1965.
- [Dav80] H. Davenport. *Multiplicative number theory*, volume 74 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1980. Revised by Hugh L. Montgomery.
- [Dei05] A. Deitmar. *A first course in harmonic analysis*. Universitext. Springer-Verlag, New York, second edition, 2005.
- [DK82] B. E. J. Dahlberg and C. E. Kenig. A note on the almost everywhere behavior of solutions to the Schrödinger equation. In *Harmonic analysis (Minneapolis,*

- Minn., 1981*), volume 908 of *Lecture Notes in Math.*, pages 205–209. Springer, Berlin, 1982.
- [DM72] H. Dym and H. P. McKean. *Fourier series and integrals*. Academic Press, New York, 1972. Probability and Mathematical Statistics, No. 14.
- [DSV03] G. Davidoff, P. Sarnak, and A. Valette. *Elementary number theory, group theory, and Ramanujan graphs*, volume 55 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 2003.
- [Dui91] J. J. Duistermaat. Self-similarity of “Riemann’s nondifferentiable function”. *Nieuw Arch. Wisk. (4)*, 9(3):303–337, 1991.
- [Edg04] G. A. Edgar, editor. *Classics on fractals*. Studies in Nonlinearity. Westview Press. Advanced Book Program, Boulder, CO, 2004.
- [Ell75] W. J. Ellison. *Les nombres premiers*. Hermann, Paris, 1975. En collaboration avec Michel Mendès France, Publications de l’Institut de Mathématique de l’Université de Nancago, No. IX, Actualités Scientifiques et Industrielles, No. 1366.
- [Fey85] R. P. Feynman. *QED: The Strange Theory of Light and Matter*. Princeton University Press, 1985.
- [FLS63] R. P. Feynman, R. B. Leighton, and M. Sands. *The Feynman lectures on physics. Vol. 1: Mainly mechanics, radiation, and heat*. Addison-Wesley Publishing Co., Inc., Reading, Mass.-London, 1963.
- [Fon05] P. Fong. *Elementary quantum mechanics*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, expanded edition, 2005.
- [Fou88] J. Fourier. *Théorie analytique de la chaleur*. Éditions Jacques Gabay, Paris, 1988. Reprint of the 1822 original.
- [Ger99] N. Gershenfeld. *The nature of mathematical modeling*. Cambridge University Press, Cambridge, 1999.
- [GK91] S. W. Graham and G. Kolesnik. *van der Corput’s method of exponential sums*, volume 126 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1991.
- [GP90] A. Galindo and P. Pascual. *Quantum mechanics. I*. Texts and Monographs in Physics. Springer-Verlag, Berlin, 1990. Translated from the Spanish by J. D. García and L. Alvarez-Gaumé.

- [Har] G. H. Hardy. A Theorem Concerning Fourier Transforms. *J. London Math. Soc.*, S1-8(3):227.
- [Hux96] M. N. Huxley. *Area, lattice points, and exponential sums*, volume 13 of *London Mathematical Society Monographs. New Series*. The Clarendon Press Oxford University Press, New York, 1996. Oxford Science Publications.
- [HW96] E. Hernández and G. Weiss. *A first course on wavelets*. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1996. With a foreword by Yves Meyer.
- [IK04] H. Iwaniec and E. Kowalski. *Analytic number theory*, volume 53 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2004.
- [Ing] A. E. Ingham. A Note on Fourier Transforms. *J. London Math. Soc.*, S1-9(1):29.
- [Ivi03] A. Ivić. *The Riemann zeta-function*. Dover Publications Inc., Mineola, NY, 2003. Theory and applications, Reprint of the 1985 original [Wiley, New York; MR0792089 (87d:11062)].
- [Kat76] Y. Katznelson. *An introduction to harmonic analysis*. Dover Publications Inc., New York, corrected edition, 1976.
- [KK66] J.-P. Kahane and Y. Katznelson. Sur les ensembles de divergence des séries trigonométriques. *Studia Math.*, 26:305–306, 1966.
- [Kna96a] A. W. Knaapp. Group representations and harmonic analysis from Euler to Langlands. I. *Notices Amer. Math. Soc.*, 43(4):410–415, 1996.
- [Kna96b] A. W. Knaapp. Group representations and harmonic analysis. II. *Notices Amer. Math. Soc.*, 43(5):537–549, 1996.
- [Koc91] H. Koch. *Introduction to classical mathematics. I*, volume 70 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1991. From the quadratic reciprocity law to the uniformization theorem, Translated and revised from the 1986 German original, Translated by John Stillwell.
- [Kör88] T. W. Körner. *Fourier analysis*. Cambridge University Press, Cambridge, 1988.
- [LT00] M. Lacey and C. Thiele. A proof of boundedness of the Carleson operator. *Math. Res. Lett.*, 7(4):361–370, 2000.

- [Max98] J. C. Maxwell. *A treatise on electricity and magnetism. Vol. 1, 2.* Oxford Classic Texts in the Physical Sciences. The Clarendon Press Oxford University Press, New York, 1998. With prefaces by W. D. Niven and J. J. Thomson, Reprint of the third (1891) edition.
- [Mic08] R. Michel. The $(n + 1)$ th proof of Stirling's formula. *Amer. Math. Monthly*, 115(9):844–845, 2008.
- [Mon78] H. L. Montgomery. The analytic principle of the large sieve. *Bull. Amer. Math. Soc.*, 84(4):547–567, 1978.
- [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, 1994.
- [Mur08] M. R. Murty. *Problems in analytic number theory*, volume 206 of *Graduate Texts in Mathematics*. Springer, New York, second edition, 2008. Readings in Mathematics.
- [MV74] H. L. Montgomery and R. C. Vaughan. Hilbert's inequality. *J. London Math. Soc. (2)*, 8:73–82, 1974.
- [PWZ96] M. Petkovšek, H. S. Wilf, and D. Zeilberger. *A = B*. A K Peters Ltd., Wellesley, MA, 1996. With a foreword by Donald E. Knuth. computer disk. See also <http://www.math.upenn.edu/~wilf/AeqB.html>.
- [Rob74] A. Robert. Des adèles: pourquoi? *Enseignement Math. (2)*, 20:133–145, 1974.
- [Ros88] H. E. Rose. *A course in number theory*. Oxford Science Publications. The Clarendon Press Oxford University Press, New York, 1988.
- [RV99] D. Ramakrishnan and R. J. Valenza. *Fourier analysis on number fields*, volume 186 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1999.
- [Sal02] D. Salomon. *A guide to data compression methods*. Springer-Verlag, 2002.
- [Spi84] M. Spivak. *Calculus Vol. I, II*. Editorial Reverté, Barcelona, 1984.
- [SR01] J. M. Sánchez-Ron. *Historia de la física cuántica: el periodo fundacional (1860–1926)*. Crítica, 2001.
- [SS71] A. Schönhage and V. Strassen. Schnelle Multiplikation grosser Zahlen. *Computing (Arch. Elektron. Rechnen)*, 7:281–292, 1971.

- [SS90] D. Stauffer and H. E. Stanley. *From Newton to Mandelbrot: A Primer in Theoretical Physics*. Springer-Verlag, Berlin, 1990.
- [SS03] E. M. Stein and R. Shakarchi. *Fourier analysis*, volume 1 of *Princeton Lectures in Analysis*. Princeton University Press, Princeton, NJ, 2003. An introduction.
- [Str86] G. Strang. *Introduction to applied mathematics*. Wellesley-Cambridge Press, Wellesley, MA, 1986.
- [Tay96] M. E. Taylor. *Partial differential equations*, volume 23 of *Texts in Applied Mathematics*. Springer-Verlag, New York, 1996. Basic theory.
- [Ter99] A. Terras. *Fourier analysis on finite groups and applications*, volume 43 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 1999.
- [Wal96] J. S. Walker. *Fast Fourier transforms*. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, second edition, 1996. With 1 IBM-PC floppy disk (3.5 inch; HD).
- [Wal08] J. S. Walker. *A primer on wavelets and their scientific applications*. Studies in Advanced Mathematics. Chapman & Hall/CRC, Boca Raton, FL, second edition, 2008.
- [Zyg77] A. Zygmund. *Trigonometric series. Vol. I, II*. Cambridge University Press, Cambridge, 1977. Reprinting of the 1968 version of the second edition with Volumes I and II bound together.

Índice alfabético

- Abel, N.H. (1802–1829), 15
aceleración de series, 6, 14
adeles, 29
amplitud, 55
amplitud de probabilidad, 60
análisis, 1
análisis armónico, 1
aproximación de la identidad, 3
armónicos, 1
artifact, 55
- banda limitada, función de, 33
Bernoulli, números de, 20
Bernoulli, polinomios de, 20
Bertrand, postulado de, 15
Bessel, F. (1784–1846), 22
Bessel, funciones de, 71
Bohr, átomo de, 56
Bohr, N. (1885–1962), 56
Bombieri, E. (1940–), 87
Born, M. (1882–1972), 60
Brun, teorema de, 97
Brun, V. (1885–1978), 97
- cámara oscura, 72
carácter, 28
Carleson, L. (1928–), 8
Cauchy–Riemann, ecuaciones de, 21
Chebyshev, P. (1821–1894), 15
Chebyshev, polinomios de, 11
colapso de la función de onda, 59
Compton, efecto, 57
convolución, 5, 30
coseno, transformada discreta, 35
- criba, 87
cuantificación, matriz de, 43
cuantización, 64
cuanto, 56
- Daubechies, I. (1954–), 44
DCT, 36
de Broglie, L. (1892–1987), 57
delta de Dirac, 3
delta de Dirac 1-periódica, 3
DFT, 28
diente de sierra, 10
difracción, 72
Dini, criterio de, 6
Dirac, P.A.M. (1902–1984), 3
Dirichlet, carácter de, 31
Dirichlet, función L de, 31
Dirichlet, P.G.L. (1805–1859), 3
Dirichlet, problema de, 75
distribuciones, teoría de, 9
divide y vencerás, algoritmo, 36
du Bois-Reymond, P. (1831–1889), 8
- ecuación de ondas, 55
eikonal, ecuación, 73
Einstein, A. (1879–1955), 56
electrón, 56
energía cinética, 58
energía potencial, 58
energía total, 58
espacio completo, 7
espacio de Hilbert, 7
espectro de emisión, 56
espiral equiangular, 83

- estado, 63
 Euler, constante de, 15, 25
 Euler, L. (1707–1783), 6
 Euler, producto de, 31
 Euler, transformada de, 14
 Euler–Maclaurin, fórmula de sumación de, 20
 extensión analítica, 23
 fase, 55
 fase estacionaria, principio de, 67
 Fejér, L. (1880–1959), 6
 Fermat, principio de, 57
 FFT, 36
 filtro lineal, 39
 FIR, 39
 fórmula de inversión, 4, 28, 30, 35
 fotón, 56
 Fourier, antitransformada de, 4
 Fourier, coeficientes de, 2
 Fourier, J. (1768–1830), 2
 Fourier, serie de, 2
 Fourier, transformada de, 4
 Fourier, transformada discreta de, 28, 30
 Fourier, transformada inversa de, 4
 Fourier, transformada rápida de, 36
 Fraunhofer, difracción de, 78
 Fraunhofer, J. (1787–1826), 78
 frecuencia, 55
 frecuencia angular, 55
 frecuencia de muestreo, 33
 Fresnel, A. (1788–1827), 78
 Fresnel, difracción de, 78
 función analítica, 21
 función de decaimiento rápido, 9
 función de ondas, 58
 función de orden finito, 23
 función entera, 23
 función multivaluada, 25
 función test, 3
 Gamma, función, 12, 24
 Gauss, C. F. (1777–1855), 17
 Gauss, problema del círculo de, 93
 Gauss, sumas de, 19
 gaussiana, función, 12
 Gibbs, fenómeno de, 54
 gran criba, 87
 gran criba, desigualdad de, 87
 Green, identidad de, 75
 Green, teorema de, 21
 Hadamard, J. (1865–1963), 23
 Hamilton, W.R. (1805–1865), 57
 Hardy, conjetura de, 93
 Hardy, G.H. (1877–1947), 50
 Hardy, serie de, 94
 Heisenberg, principio de incertidumbre de, 65
 Heisenberg, W. (1901–1976), 51
 hercio, 32
 Hermite, funciones de, 12
 Hertz, H. (1857–1894), 56
 Hilbert, desigualdad de, 88
 homomorfismo, 28
 Huxley, M.N. (1944–), 96
 Huygens, C. (1629–1695), 72
 Huygens, principio de, 72, 78
 identidad de Parseval, 8
 incertidumbre, principio de, 34, 49
 inercia, ley de, 62
 Ingham, A.E. (1900–1967), 34
 intensidad, 79
 interpretación de Copenhague, 59
 intervalos diádicos, 14
 Jacobi, C.G.J. (1804–1851), 17
 Jacobi, función θ de, 17
 JPEG, 42
 JPEG 2000, 44
 Kirchhoff, G. (1824–1887), 75
 Kolmogorov, A. (1903–1987), 8

- Landau, notación de, 13
 Leibniz, criterio de, 14
 Leibniz, G.W. (1646–1718), 13
 libre de cuadrados, 89
 Linnik, Yu.V. (1915–1972), 87
 logaritmo discreto, 30
 logaritmo integral, 16
 longitud de onda, 55
- Maxwell, ecuaciones de, 57
 micra, 80
 mínima acción, principio de, 57
 Möbius, función de, 89
 momento angular, 56
 momento lineal, 57
 Montgomery, H.L. (1944–), 88
- núcleo de Dirichlet, 5
 núcleo de Fejér, 6
 número de ondas, 55
 números primos, teorema de los, 16, 96
 Nyquist, frecuencia de, 34
- observable, 63
 onda rectangular, 9
 onda secundaria, 72
 ondas, ecuación de, 55, 74
 ondícula, 44
 óptica geométrica, 57
- Paley-Wiener, teorema de, 33
 paquete de ondas, 62
 pares de exponentes, 86
 Parseval, identidad de, 7, 30
 peine de Dirac, 3
 periodo, 55
 pixel, 27
 Plancherel, identidad de, 7
 Planck, constante de, 56
 Poisson fórmula de sumación de, 16
 Pontryagin, teorema de dualidad de, 29
 Prágmén-Lindelöf, principio de, 51
- primos gemelos, 97
 Proceso A, 86
 Proceso B, 86
 protón, 56
- regla de cálculo, 49
 representación, 29
 residuo, 22
 residuos, teorema de los, 22
 Riemann, B. (1826–1866), 26
 Riemann, función ζ de, 24
 Riemann-Lebesgue, Lema de, 6
 Riesz, M. (1886–1969), 88
 ruido blanco gaussiano, 46
- Schönhage-Strassen, algoritmo de, 38
 Schrödinger, ecuación de, 58
 Schrödinger, E. (1887–1961), 57
 Schwartz, clase de, 9
 Schwartz, L. (1915–2002), 9
 Selberg, A. (1917–2007), 88
 señal, 27
 Shannon, teorema de muestreo de, 33
 Stirling, fórmula de, 26
 sumación por partes, 13
 sumas de Cesàro, 6
- Tao, T. (1975–), 51
 Tate, J.T. (1925–), 29
 Tate, tesis de, 29
 tonos puros, 1
 toro, 2
 transformada z , 50
 truco telescópico, 13
- van der Corput, J. (1890–1975), 81
 van der Corput, lemas de, 69
 variación acotada, 6
 Vaughan, R.C. (1945–), 88
 vector de aproximación, 45
 vectores de detalle, 45
 velocidad de fase, 55

velocidad de grupo, 62
Vinogradov, I.M. (1891–1983), 91
Vinogradov, notación de, 13
von Mangoldt, símbolo de, 16
von Neumann, J. (1903–1957), 63

Wallis fórmula de, 23
wavelet, 44, 46
wavelet, transformada discreta, 46
wavelet, transformada discreta inversa, 46
Weierstrass, K. (1815–1897), 80
Weyl shift, 85
Weyl, H. (1885–1955), 81