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

Tesis de Licenciatura de la Universidad de Buenos Aires en el área Ciencias Matemáticas

Fenómeno de Bohr y sumabilidad para funciones en el cubo Booleano

Sequeira, Juan Ignacio

Director: Dr. Daniel E. Galicer

Fecha de defensa: 06/10/2020

Agradecimientos

Agradezco a los Dres. Martín Mazzitelli y Santiago Muro por aceptar ser los jurados de la tesis, leerla con detenimiento y enviarme las correcciones pertinentes.

Agradezco al Dr. Daniel Galicer por cumplir de manera excepcional el rol de director. Siempre atento a todos los detalles, desde la formalidad académica hasta el excelente trato en cada momento hacia mi persona. Valoro mucho esto último.

Agradezco y dedico este trabajo a mi madre, padre, hermano, amigos y familiares por ayudarme y motivarme día a día.

Fenómeno de Bohr y sumabilidad para funciones en el cubo Booleano

Resumen

Esta tesis tiene como objetivo estudiar el comportamiento de los coeficientes de Fourier para funciones a valores reales con dominio en los vértices de un cubo n-dimensional en términos de la norma uniforme. Nuestra motivación proviene de la teoría hecha en el análisis complejo, particularmente para polinomios complejos en varias variables. Realizamos constantemente este paralelismo a lo largo del trabajo y observamos una notable similitud con los resultados, pero también ciertas diferencias que traen como consecuencia nuevas técnicas y preguntas.

Probaremos que la constante de Bohnenblust-Hille para funciones Booleanas de grado d está acotada superiormente por una constante absoluta con exponente $\sqrt{d \log d}$.

Definimos el concepto de radio de Bohr en este contexto y estimamos el comportamiento asintótico del mismo para varias subclases de funciones sobre cubos Booleanos finitos, como por ejemplo: la clase de todas las funciones reales, la subclase de todas las funciones homogéneas y por último la clase de aquellas funciones en que la esperanza se mantiene muy por debajo de la norma uniforme.

Por último, estudiaremos las desigualdades hipercontractivas, las mismas cumplen un rol fundamental en los conceptos mencionados.

Introducción

Las funciones Booleanas son quizás el objeto de estudio más básico en informática teórica, y el análisis de Fourier se ha convertido en una herramienta indispensable en este campo. El tema también ha desempeñado un papel clave en otras áreas de la matemática, desde la combinatoria, la teoría de grafos aleatorios y la física estadística, hasta la geometría Gaussiana, los espacios métricos/de Banach y la teoría de la elección social.

Generalmente (dependiendo el contexto) este tipo de funciones tienen el siguiente aspecto $f : \{0,1\}^n \to \{0,1\}$ que asigna cada vector binario de longitud n, o cadena, en un solo valor binario, o bit. Por lo general hay varias maneras de representar los bits. A veces se escribe verdadero y falso; a veces -1 y 1, considerados como números reales. Otras veces 0 y 1, también como elementos del cuerpo $\mathbb{F}_2$ de tamaño 2, o simplemente como símbolos. Nosotros usaremos en la totalidad del trabajo el -1 y 1, por tanto una función Booleana se verá como $f : \{\pm 1\}^n \to \{\pm 1\}$.

Aquí hay algunos ejemplos de aplicaciones:

- En el diseño de circuitos, una función Booleana puede representar el comportamiento deseado de un circuito con n entradas y una salida.
- En combinatoria extrema, una función Booleana f puede identificarse con un "sistema de conjuntos" $\mathscr{F}$ en $[n] = \{1, 2, ..., n\}$, donde los conjuntos $X \subset [n]$ se identifican con sus indicadores $0-1$ y $X \in \mathscr{F}$ si y solo si $f(X) = 1$.
- En la teoría de codificación, una función Booleana podría ser la función indicadora para el conjunto de mensajes en un código binario de corrección de errores de longitud n.
- En la teoría del aprendizaje, una función Booleana puede representar un "concepto" con n atributos binarios.
- En la teoría de elección social, una función Booleana se puede identificar con una "regla de votación" para una elección con dos candidatos llamados 0 y 1.

El dominio de una función Booleana $\{\pm 1\}^n$ se lo suele llamar cubo de Hamming o hipercubo, cubo Booleano o cubo discreto. El nombre "cubo de Hamming" enfatiza que a menudo estamos interesados en la distancia de Hamming entre cadenas $x, y \in \{\pm 1\}^n$, definido por

$$\Delta(x,y) = \#\{i : x_i \neq y_i\}.$$

Si suponemos que se tiene un problema relacionado con las funciones Booleanas con alguna de las siguientes dos características:

- la distancia de Hamming es relevante;
- se está contando cadenas, o en donde está involucrada la distribución de probabilidad uniforme en $\{\pm 1\}^n$;

entonces puede ayudar el análisis de las funciones Booleanas. En términos generales, esto significa obtener información sobre las funciones mediante el análisis de su expansión de Fourier.

La expansión de Fourier de una función Booleana $f : \{\pm 1\}^n \to \{\pm 1\}$ es simplemente su representación como un polinomio real, tetraedral (tetraedral significa que ninguna variable x_i aparece al cuadrado, en cubos, etc.) por ejemplo, suponga que $n = 2$ y $f = \max_2$, la función "máximo" en 2 bits:

$$\begin{aligned} \max_2(+1, +1) &= +1, \\ \max_2(+1, -1) &= +1, \\ \max_2(-1, +1) &= +1, \\ \max_2(-1, -1) &= -1. \end{aligned}$$

Entonces $\max_2$ se puede representar como un polinomio tetraedral,

$$\max_2(x_1, x_2) = \frac{1}{2} + \frac{1}{2}x_1 + \frac{1}{2}x_2 - \frac{1}{2}x_1x_2.$$

Podemos obtener esta representación polinomial tetraedral para cualquier función Booleana arbitraria y podemos generalizar el concepto al caso más general que son las funciones Booleanas a valores reales $f : \{\pm 1\}^n \to \mathbb{R}$. De hecho, lo que termina ocurriendo es que podemos representar estas últimas funciones como una suma de polinomios tetraedrales. El polinomio tetraedral para f puede tener hasta 2^n términos, correspondiente a los subconjuntos $S \subset [n]$. Escribimos los monomios para cada S como:

$$x^S = \prod_{i \in S} x_i \qquad \text{(con } x^{\emptyset} = 1 \text{ por convención)},$$

que suelen llamarse funciones de Walsh y encontrarse con otra notación χ_S, para todo $S \subset [n]$. En cuanto a los coeficientes de f, tenemos:

$$\hat{f}(S) = \text{coeficiente del monomio } x^S \text{en la representación polinomial de } f.$$

Obteniendo así para toda $f : \{\pm 1\}^n \to \mathbb{R}$ una única expresión como un polinomio tetraedral

$$f(x) = \sum_{S \subset [n]} \hat{f}(S) x^S. \tag{1}$$

Esta expresión se llama expansión de Fourier de f (o también Fourier-Walsh), y el número real $\hat{f}(S)$ se llama coeficiente de Fourier de f en S. Colectivamente, los coeficientes se conocen como el espectro de Fourier de f.

Con todo este desarrollo es bastante natural definir la noción de grado para una función $f : \{\pm 1\}^n \to \mathbb{R}$. Decimos que f tienen *grado d* si $\hat{f}(S) = 0$ para todo $S \subset [n]$ tal que $|S| > d$, i.e., el espectro de Fourier se concentra en niveles menores o iguales a d. Decimos que f es *d-homogénea* si cumple que $\hat{f}(S) = 0$ con $|S| \neq d$, i.e., el espectro de Fourier se concentra en el nivel d.

La magnífica expresión polinomial de las funciones Booleanas a valores reales motivó que se pueda estudiar la sumabilidad de los coeficientes de Fourier en comparación con la norma uniforme de la función. Estas ideas provienen del análisis complejo, más precisamente lo que se conoce como el estudio del radio de Bohr.

La historia del radio de Bohr comienza a principios del siglo XX junto al desarrollo de la teoría de holomorfía en infinitas variables. Bajo esta atmósfera Harald Bohr encuentra un profundo vínculo entre las series de potencias en infinitas variables con las series de Dirichlet. Las regiones de convergencia, convergencia absoluta y uniforme de las series de Dirichlet definen semiplanos en el plano complejo. Bohr estaba principalmente interesado en controlar la región de convergencia absoluta de cada serie. Para esto se concentró en particular en hallar el ancho de la banda más grande en que una serie de Dirichlet puede converger uniformemente pero no absolutamente. A esta cuestión se la conoce como *problema de convergencia absoluta de Bohr*. Este ciclo de ideas llevó a Bohr a preguntarse si era posible comparar el valor absoluto de una serie de potencias en una variable con la suma de los valores absolutos de sus coeficientes. Como respuesta, logró demostrar:

Teorema 0.0.1. *(Bohr) El radio $r = 1/3$ es el valor óptimo para el cual se verifica la desigualdad:*

$$\sum_{k \geqslant 0} |c_r| r^k \leqslant \sup_{z \in \mathbb{D}} \left| \sum_{k \geqslant 0} c_k z^k \right|,$$

para toda función holomorfa $f(z) = \sum_{k \geqslant 0} c_k z^k$ en el disco con $\sup_{z \in \mathbb{D}} |f(z)| < \infty$.

El problema de convergencia absoluta de Bohr fue resuelto 20 años más tarde por Bohnenblust y Hille, ellos mostraron en [BH31] que el ancho máximo de dicha banda es $\frac{1}{2}$. Usando la descomposición de Taylor en polinomios homogéneos para funciones holomorfas e inspirados en la desigualdad 4/3 de Littlelwood logran traducir el problema planteado por Bohr en el estudio de ciertas desigualdades que relacionan la sumabilidad de los coeficientes de polinomios homogéneos con su norma uniforme en la bola ℓ_∞. Así prueban que para cualquier polinomio m-homogéneo en n variables complejas $P = \sum_{\alpha_1 + \ldots + \alpha_n = m} a_\alpha(P) z^\alpha$, vale que la $\frac{m+1}{2m}$-norma de los coeficientes de P se acota superiormente con el producto de $C_m > 0$ y la norma uniforme de P en la bola ℓ_∞, donde la constante $C_m > 0$ es una constante independiente de la cantidad de variables n. Este tipo de constantes adquirió el nombre de constantes de Bohenblust-Hille (al igual que la desigualdad) y se estudió el crecimiento asintótico en función de m, el grado de homogeneidad de los polinomios.

En 1989, bastante tiempo después, se retoma el estudio del Teorema de Bohr 0.0.1 en [DT89], generalizándolo y conectándolo con nociones de la teoría local de espacio de Banach. Comienza un estudio sistemático de lo que se conoce como radio de Bohr n-dimensional. Dada alguna norma sobre $\mathbb{C}^n$, es decir $X = (\mathbb{C}^n, \|\cdot\|_X)$, se estudia la variante del problema de Bohr que proviene de reemplazar al disco unidad por la bola de X. Se busca así el mayor radio $r > 0$ tal que para toda función holomorfa en la bola de X de la forma $f(z) = \sum_{\alpha \in \mathbb{N}_0^n} a_\alpha(f) z^\alpha$ valga

$$\sup_{\|z\|_X < r} \sum_{\alpha \in \mathbb{N}_0^n} |a_\alpha z^\alpha| \leqslant \sup_{\|z\| \leqslant 1} \left| \sum_{\alpha \in \mathbb{N}_0^n} a_\alpha(f) z^\alpha \right|.$$

A este radio maximal se lo llama radio de Bohr de la bola de X y se lo nota $K(B_X)$. De este modo podemos decir que el resultado de Bohr se traduce como $K(\mathbb{D}) = 1/3$. Vale comentar que para la bola de ningún espacio con dimensión más grande que uno, se sabe el valor preciso del radio de Bohr. Los avances en este campo están ligados a entender el comportamiento asintótico de este radio para las bolas de ciertos espacios cuando su dimensión tiende a infinito.

Como ya mencionamos las constantes de Bohnenblust-Hille fueron de total importancia en los avances de teoría de series de Dirichlet. Ron Blei probó en [Ble01] que para todo $d \leqslant n$ existe una constante $C(d) > 0$ tal que para cada función Booleana $f : \{\pm 1\}^n \to \mathbb{R}$ de grado d se tiene

$$\left(\sum_{S \subset [n], |S| \leqslant d} |\widehat{f}(S)|^{\frac{2d}{d+1}} \right)^{\frac{d+1}{2d}} \leqslant C(d) \|f\|_{\{\pm 1\}^n}, \tag{2}$$

y este exponente $\frac{2d}{d+1}$ es óptimo, en el sentido que no se puede reemplazar el factor $C(d)$ por una constante independiente de n para exponentes menores. La mejor constante posible $C(d)$ es denotada por

$$BH_{\{\pm 1\}}^{\leqslant d}.$$

La desigualdad (2) es la versión Booleana análoga a la desigualdad para polinomios complejos ya comentada. La misma afirma (con esta notación) que para todo d existe una constante (la mejor) $BH_{\mathbb{T}}^{\leqslant d}$ tal que para todo polinomio complejo de grado d, $P(z) = \sum_{\alpha \in \mathbb{N}_0^n, |\alpha| \leqslant d} a_\alpha z^\alpha$ con n variables complejas $z_1, ..., z_n$ cumple

$$\left(\sum_{\alpha \in \mathbb{N}_0^n, |\alpha| \leqslant d} |\ a_\alpha|^{\frac{2d}{d+1}} \right)^{\frac{d+1}{2d}} \leqslant BH_{\mathbb{T}}^{\leqslant d} \|P\|_{\mathbb{T}^n}; \tag{3}$$

donde el exponente $\frac{2d}{d+1}$ no se puede mejorar (en el mismo sentido que antes). Originalmente, este resultado fue probado para los polinomios homogéneos de grado d. Podemos entonces denotar la constante en este caso como

$$BH_{\mathbb{T}}^{=d},$$

junto a un truco básico de homogeneización se puede demostrar que

$$BH_{\mathbb{T}}^{=d} = BH_{\mathbb{T}}^{\leqslant d}. \tag{4}$$

En años recientes la constante de Bohnenblust-Hille $BH_{\mathbb{T}}^{\leqslant d}$ tuvo un estudio intensivo. La demostración original de (3) y sus mejoras dan estimaciones superiores que son esencialmente de orden $\sqrt{d}^d$, pero en [DFOC+11] se probó que $BH_{\mathbb{T}}^{\leqslant d}$ crece como mucho $\sqrt{2}^d$, luego esto se mejoró en [PSS14] donde los autores muestran que existe una constante absoluta $C > 0$ tal que

$$BH_{\mathbb{T}}^{\leqslant d} \leqslant C^{\sqrt{d \log d}}. \tag{5}$$

Por lo tanto, la constante $BH_{\mathbb{T}}^{\leqslant d}$ crece subexponencialmente con respecto al grado d. En particular esto implica que

$$\limsup_{d \to \infty} \sqrt[d]{BH_{\mathbb{T}}^{\leqslant d}} = 1. \tag{6}$$

Ahora nos preguntamos cómo es el crecimiento asintótico de $BH_{\{\pm 1\}}^{\leqslant d}$ como función de d, tal que para cada $n \in \mathbb{N}$ el grupo $\mathbb{T}^n$ es reemplazado por el grupo $\{\pm 1\}^n$. Tenemos el siguiente resultado análogo:

Teorema 0.0.2. *Existe una constante absoluta $C > 0$ tal que para cada $d \in \mathbb{N}$*

$$BH_{\{\pm 1\}}^{\leqslant d} \leqslant C^{\sqrt{d \log d}},$$

en particular, $\limsup_{d \to \infty} \sqrt[d]{BH_{\{\pm 1\}}^{\leqslant d}} = 1.$

Uno de los objetivos de esta tesis es recrear la demostración de este resultado, además de comentar algunas aplicaciones y preguntas abiertas que se relacionan con estas ideas. El otro objetivo es hacer un estudio símilar al radio de Bohr pero en el contexto Booleano, es decir estudiar el radio Booleano para distintas familia de funciones $f : \{\pm 1\}^n \to \mathbb{R}$.

Vale destacar que estos trabajos fueron realizados recientemente por Andreas Defant, Mieczysław Mastyło y Antonio Pérez, podemos encontrarlos en [DMP19] y [DMP18].

Con respecto al estudio del radio Booleano comenzaremos con la familia más grande posible, calcularemos exactamente el radio Booleano de todas las funciones en el cubo n-dimensional $\{\pm 1\}^n$ a valores reales, obteniendo un $2^{\frac{1}{n}} - 1$ como resultado. Si observamos el caso análogo, es decir, el radio de Bohr de todos los polinomios complejos en el toro n-dimensional sabemos que sólo se conoce el valor exacto para $n = 1$. Por lo tanto, la comparación que podemos realizar entre estos es una comparación asintótica .

También estudiaremos los radios Booleanos para las familias de funciones m-homogéneas $\mathcal{B}_{=m}^n$ y homogéneas $\mathcal{B}_{\mathrm{hom}}^n$ en $\{\pm 1\}^n$. Demostraremos que para cada $1 \leqslant m \leqslant n$ existe C_m (independiente de n) tal que

$$C_m^{-1} n^{\frac{1}{2m}} \binom{n}{m}^{-\frac{1}{2m}} \leqslant \rho(\mathcal{B}_{=m}^n) \leqslant C_m n^{\frac{1}{2m}} \binom{n}{m}^{-\frac{1}{2m}},$$

donde además $\lim_m C_m = 1$. La demostración de la acotación inferior se basa en el uso de la desigualdad de Bohnenblust-Hille mientras que la estimación superior requiere de un argumento probabilístico del tipo Kahane-Salem-Zygmund. El hecho que C_m converge a uno es la llave para probar

$$\rho(\mathcal{B}^n_{\mathrm{hom}}) = \sqrt{\frac{\log n}{n}}(1 + O(1)).$$

También haremos la comparación con el radio de Bohr de todos los polinomios complejos m-homogéneos en varias variables $K_n^{=m}$ y el radio de Bohr de todos los polinomios homogéneos complejos K_n^{hom}.

La siguiente familia para estimar el radio Booleano será de todas las funciones de grado d en $\{\pm 1\}^n$, el cuál denotaremos $\rho(\mathcal{B}^n_{\leqslant d})$. El resultado (menos preciso que el caso homogéneo) afirma que para cada $1 \leqslant d \leqslant n$, existen constantes c_d y C_d (independientes de n) tal que

$$c_d \frac{1}{n^{\frac{1}{2}}} \leqslant \rho(\mathcal{B}^n_{\leqslant d}) \leqslant C_d \frac{1}{n^{\frac{d-1}{2d}}}.$$

La demostración de la acotación inferior requiere una desigualdad para funciones en $\{\pm 1\}^n$ que podría ser de interés independiente, y afirma lo siguiente: si $f : \{\pm 1\}^n \to \mathbb{R}$ es una función de grado d, entonces

$$\left(\sum_{S \neq \varnothing} |\hat{f}(S)|^2 \right)^{\frac{1}{2}} \leqslant 2e^d (1 - |\hat{f}(\varnothing)|).$$

Por último estimaremos el radio Booleano de la familia de funciones $f : \{\pm 1\}^n \to \mathbb{R}$ tal que $|\mathbb{E}[f]| \leqslant (1 - \delta)\|f\|_\infty$, la misma denotaremos $\mathcal{B}^n_\delta$. El resultado principal muestra la existencia de una constante absoluta $C > 0$ tal que para todo $n \in \mathbb{N}$ y $\frac{1}{2^n} \leqslant \delta \leqslant 1$

$$C^{-1} \frac{1}{\sqrt{n}\sqrt{\log(\frac{2}{\delta})}} \leqslant \rho(\mathcal{B}^n_\delta) \leqslant C \frac{1}{\sqrt{n}\sqrt{\log(\frac{2}{\delta})}}.$$

La demostración de la estimación inferior depende de la desigualdad hipercontractiva, mientras que para la estimación superior realizaremos el estudio del radio Booleano de la siguiente familia de funciones (del tipo threshold)

$$\psi_{\alpha,n} : \{\pm 1\}^n \to \mathbb{R}, \quad \psi_{\alpha,n}(x) = sign(x_1 + \ldots + x_n - \alpha), \qquad 0 \leqslant \alpha < n.$$

Probaremos que existe una constante $C > 0$ tal que para todo $0 \leqslant \alpha \leqslant n$

$$C^{-1} \frac{1}{\alpha + \sqrt{n}} \leqslant \rho(\psi_{\alpha,n}) \leqslant C \frac{1}{\alpha + \sqrt{n}}.$$

La estrategia de la prueba permite dar la precisión asintótica del radio Booleano de la función mayoría $Maj_n = \psi_{0,n}$ para n impar, la misma es un clásico ejemplo en teoría de elección social.

Por último, un factor común que podemos encontrar en el estudio de estas ideas es la comparación de las p-normas para funciones Booleanas a valores reales para distintos valores de p, con ésto se logran muchos de los pasos relevantes para la distintas demostraciones del trabajo. La norma p-ésima en este contexto es

$$\|f\|_p = \mathbb{E}[|f(x)|^p]^{\frac{1}{p}}.$$

Las desigualdades que realizan esta tarea son conocidas comúnmente como *desigualdades hipercontractivas*. Se puede encontrar un largo camino en la historia de las mismas. Sus raíces más tempranas están en el trabajo de Paley [Pal32] de 1932, que usando un teorema de interpolación probó que se puede comparar la norma 2 con la norma $p \in \mathbb{R}_+$ para funciones Booleanas homogéneas de grados 2 a valores reales. En 1968 Bonami en [Bon68] prueba que para funciones de grado k, se tiene $\|f\|_q \leqslant c_k\sqrt{q}\|f\|_2$ para todo $q \geqslant 2$. Independientemente en 1969, Kiener publicó en su tesis doctoral [Kie69], la extensión del resultado de hipercontractividad de Paley que afirma que si $f : \{\pm 1\}^n \to \mathbb{R}$ es k-homogénea entonces $c_{p,k}\|f\|_2 \leqslant \|f\|_p \leqslant C_{p,k}\|f\|_2$ para todo $p \in \mathbb{R}_+$.También de forma independiente en 1969, Schreiber consideró en [Sch69] polinomios tetraedrales sobre una sucesión ortonormal $\mathbf{x}_1$,..., $\mathbf{x}_n$ de variables aleatorias reales (o complejas) centradas. Mostró que si f tiene un grado como mucho k, entonces, para cualquier número entero par $q \geqslant 4$, vale que $\|f\|_q \leqslant C\|f\|_2$, donde C depende solo de k, q y las normas q de los $\mathbf{x}_i$'s. En 1970, Bonami publicó en su tesis doctoral [Bon70] el teorema de hipercontractividad completo:

Teorema 0.0.3. *(Bonami) Sea $f : \{\pm 1\}^n \to \mathbb{R}$ y sean $1 \leqslant p \leqslant q \leqslant \infty$. Entonces*

$$\|T_\rho f\|_q \leqslant \|f\|_q, \qquad \textit{para } 0 \leqslant \rho \leqslant \sqrt{\frac{p-1}{q-1}}. \tag{7}$$

Considerando la expansión de Fourier de f en (1), tenemos la expresión del operador de ruido T_ρ dada por

$$T_\rho f(x) := \sum_{S \subset [n]} \hat{f}(S)\rho^{|S|}x^S,$$

con $x \in \{\pm 1\}^n$. Si bien no veremos la demostración general, probaremos algunos casos particulares con el fin de motivar/comentar las ideas que subyacen a estas desigualdades. Tomamos como principal referencia el acabado libro de Ryan O'Donnell [O'D14].

Esta tesis está organizada de la siguiente manera.

En el Capítulo 1 comenzaremos definiendo conceptos generales que serán útiles en el desarrollo de la monografía. También estimaremos el comportamiento asintótico de la constante de Bohnenblust-Hille para polinomios reales en el cubo real $[-1,1]^n$ y además comentaremos algunas propiedades de comparación de normas de polinomios sobre el toro n-dimensional complejo $\mathbb{T}^n$ y el cubo real $[-1,1]^n$, que sirven como puente entre los resultados del caso complejo y el caso Booleano.

En el Capítulo 2 tenemos como objetivo principal demostrar que la constante de Bohnenblust-Hille $BH^{\leqslant d}_{\{\pm 1\}}$ para funciones Booleanas de grado d (como indica el supraíndice) se acota superiormente por una constante absoluta con potencia $\sqrt{d \log d}$. Realizaremos primero la demostración para el caso homogéneo, es decir para funciones d-homogéneas, con la idea de mostrar con mayor claridad los pasos de la prueba. En ambos casos seguiremos la estrategia conceptual de la famosa desigualdad 4/3 de Littlewood, refinando los argumentos. Finalizando el capítulo comentaremos informalmente algunas aplicaciones y preguntas abiertas. Como por ejemplo la famosa conjetura de Aaronson-Ambainis que de ser cierta implicaría preguntas abiertas como, por ejemplo, una famosa conjetura en teoría de información cuántica.

En el Capítulo 3 estudiaremos el radio Booleano para distintas familias de funciones $f : \{\pm 1\}^n \to \mathbb{R}$. En todo momento haremos un paralelismo comparativo enunciando los resultados análogos para el clásico caso complejo, es decir para polinomios complejos en varias variables.

Para finalizar el trabajo escribimos el Capítulo 4 con el objetivo de demostrar ciertas desigualdades que usamos en los capítulos precedentes. Además de que cada uno de los resultados tienen un interés propio en si mismo, todos son fundamentales para el desarrollo de estas ideas. En la primera sección simplemente se demostrará la famosa desigualdad de Blei y en la última sección estudiaremos las desigualdades hipercontractivas. Si bien los enunciados de estas últimas se pueden genealizar a otros espacios, nuestro estudio será especificamente en el cubo Booleano $\{\pm 1\}^n$. Tendremos como fin mostrar las estrategias e interpretaciones generales que se suelen hacer en este contexto.

Índice general

Capítulo 1

Preliminares

Comenzaremos definiendo las nociones más básicas para el estudio de funciones en el cubo Booleano $\{\pm 1\}^n$. Para una mayor información referimos a las fuentes [O'D08] y [DW08], o el libro más extenso [O'D14].

1.0.1. Funciones Booleanas

En la introducción mostramos el desarrollo polinomial de la función máx_2

$$\text{máx}_2(x_1, x_2) = \frac{1}{2} + \frac{1}{2}x_1 + \frac{1}{2}x_2 - \frac{1}{2}x_1 x_2.$$

Para obtener tales representaciones polinomiales en general podemos proceder de la siguiente manera. Dada una función Booleana arbitraria $f : \{\pm 1\}^n \to \{\pm 1\}$ existe un método para encontrar un polinomio que interpola los 2^n valores que f asigna a los puntos $\{\pm 1\}^n$. Para cada $a = (a_1, ..., a_n) \in \{\pm 1\}^n$ *el polinomio indicador*

$$1_{\{a\}}(x) = \left(\frac{1 + a_1 x_1}{2}\right)\left(\frac{1 + a_2 x_2}{2}\right) ... \left(\frac{1 + a_n x_n}{2}\right),$$

toma 1 cuando $x = a$ y 0 cuando $x \in \{\pm 1\}^n \backslash \{a\}$. Entonces f tiene la siguiente representación polinomial

$$f(x) = \sum_{a \in \{\pm 1\}^n} f(a) 1_{\{a\}}(x).$$

Ilustrando con el ejemplo $f = \max_2$ nuevamente, tenemos

$$\begin{aligned}\text{máx}_2(x_1, x_2) = (+1)\left(\frac{1+x_1}{2}\right)\left(\frac{1+x_1}{2}\right)\\ + (+1)\left(\frac{1-x_1}{2}\right)\left(\frac{1+x_1}{2}\right)\\ + (+1)\left(\frac{1+x_1}{2}\right)\left(\frac{1-x_1}{2}\right)\\ + (-1)\left(\frac{1-x_1}{2}\right)\left(\frac{1-x_1}{2}\right) = \frac{1}{2} + \frac{1}{2}x_1 + \frac{1}{2}x_2 - \frac{1}{2}x_1x_2.\end{aligned}$$

Hacemos dos comentarios sobre este procedimiento de interpolación. Primero, funciona igualmente bien en el caso más general de las funciones Booleanas a valores reales

$$f : \{\pm 1\}^n \to \mathbb{R}.$$

Segundo, dado que los polinomios indicadores son tetraedrales cuando se expanden, la interpolación siempre produce un polinomio tetraedral. De hecho, tiene sentido que podamos representar funciones $f : \{\pm 1\}^n \to \mathbb{R}$ con polinomios multilineales: dado que sólo nos interesan las entradas x donde $x_i = \pm 1$, y cualquier factor de x_i^2 puede ser reemplazado por 1.

Tomemos una función $f : \{\pm 1\}^n \to \mathbb{R}$. Una representación canónica de f se hace en términos de las funciones de Walsh, que están definidas para $S \subset [n] = \{1, ..., n\}$ como:

$$\chi_S : \{\pm 1\}^n \to \{\pm 1\},$$

$$\chi_S(x) := \prod_{k \in S} x_k, \qquad x \in \{\pm 1\}^n,$$

donde $\chi_\varnothing(x) := 1$ para todo $x \in \{\pm 1\}^n$. Vimos que para cada $f : \{\pm 1\}^n \to \mathbb{R}$ la expansión de Fourier-Walsh es

$$f(x) = \sum_{S \subset [n]} \hat{f}(S) \prod_{k \in S} x_k = \sum_{S \subset [n]} \hat{f}(S)\chi_S(x), \qquad x \in \{\pm 1\}^n. \tag{1.1}$$

Es útil explorar esta idea desde la perspectiva del álgebra lineal. El conjunto de todas las funciones $f : \{\pm 1\}^n \to \mathbb{R}$ forma un espacio vectorial V, ya que podemos sumar dos funciones (punto a punto) y podemos multiplicar una función por un escalar real. El espacio vectorial V es 2^n-dimensional: podemos pensar a las funciones en este espacio vectorial como vectores en $\mathbb{R}^{2^n}$, donde apilamos los 2^n valores de $f(x)$ en un vector columna (en un orden fijo). A continuación ilustramos la expansión de Fourier de la función máx_2 con esta perspectiva

$$\text{máx}_2 = \begin{pmatrix} +1 \\ +1 \\ +1 \\ -1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} +1 \\ +1 \\ +1 \\ +1 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} +1 \\ -1 \\ +1 \\ -1 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} +1 \\ +1 \\ -1 \\ -1 \end{pmatrix} + -\frac{1}{2}\begin{pmatrix} +1 \\ -1 \\ -1 \\ +1 \end{pmatrix}. \tag{1.2}$$

En términos más generales, la expansión de Fourier muestra que cada función $f : \{\pm 1\}^n \to \mathbb{R}$ en V es una combinación lineal de las funciones de Walsh; es decir, son un conjunto generador para V. Dado que el número de funciones de Walsh es $2^n = \dim V$, podemos deducir que de hecho son una base para V. En particular, esto justifica la unicidad de la expansión de Fourier.

También podemos introducir un producto interno para el par de funciones $f, g \in V$. El producto interno habitual en $\mathbb{R}^{2^n}$ correspondería a $\sum_{x\in\{\pm 1\}^n} f(x)g(x)$, pero es más conveniente reescalar esto por un factor de 2^n, convirtiéndolo en un promedio en lugar de una suma. De esta manera, una función Booleana $f : \{\pm 1\}^n \to \{\pm 1\}$ tendrá $\langle f, f\rangle$=1, es decir, será un "vector unitario".

Definición 1.0.1. *Definimos el producto interno $\langle\cdot,\cdot\rangle$ en el par $f, g : \{\pm 1\}^n \to \mathbb{R}$ por*

$$\langle f, g\rangle := 2^{-n} \sum_{x\in\{\pm 1\}^n} f(x)g(x) = \mathbb{E}_{\mathbf{x}\sim\{\pm 1\}}[f(\mathbf{x})g(\mathbf{x})].$$

Usamos la notación $\|f\|_2 = \sqrt{\langle f, f\rangle}$, más en general

$$\|f\|_p = \mathbb{E}[|f(\mathbf{x})|^p]^{\frac{1}{p}}.$$

Notar que hemos introducido la notación probabilística que se utilizará en gran medida a lo largo del trabajo:

Escribimos $\mathbf{x} \sim \{\pm 1\}^n$ para denotar que $\mathbf{x}$ es una cadena aleatoria (o vector aleatorio) elegida uniformemente de $\{\pm 1\}^n$. De manera equivalente, las n coordenadas x_i se eligen independientemente para que sean +1 con probabilidad 1/2 y -1 con probabilidad 1/2. Por lo general escribimos variables aleatorias en negrita y esperanza con $\mathbb{E}$ siempre será con respecto a una variable aleatoria uniforme $\mathbf{x} \sim \{\pm 1\}^n$ a menos que se especifique lo contrario. Por lo tanto, podríamos escribir la esperanza como $\mathbb{E}_x[f(\mathbf{x})g(\mathbf{x})]$ o $\mathbb{E}[f(\mathbf{x})g(\mathbf{x})]$ o incluso $\mathbb{E}[fg]$.

Volviendo a la base de las funciones de Walsh para V, el hecho crucial que subyace en todo análisis de las funciones Booleanas es que esta es una base ortonormal.

Teorema 1.0.2. *Las 2^n funciones de Walsh χ_S forman una base ortonormal para el espacio vectorial V de las funciones $\{\pm 1\}^n \to \mathbb{R}$; es decir,*

$$\langle \chi_S, \chi_T\rangle = \begin{cases} 1 & \text{si } S = T, \\ 0 & \text{si } S \neq T. \end{cases}$$

Recordando la definición $\langle \chi_S, \chi_T\rangle = \mathbb{E}[\chi_S(x)\chi_T(x)]$, el Teorema 1.0.2 se deduce inmediatamente de dos hechos:

Proposición 1.0.3. *Para $x \in \{\pm 1\}^n$ tenemos que $\chi_S(x)\chi_T(x) = \chi_{S\triangle T}(x)$, donde $S \triangle T$ denota la diferencia simétrica de S y T.*

Demostración. Tenemos las siguientes igualdades

$$\chi_S(x)\chi_T(x) = \prod_{i\in S} x_i \prod_{i\in T} x_i = \prod_{i\in S\triangle T} x_i \prod_{i\in S\cap T} x_i^2 = \prod_{i\in S\triangle T} x_i = \chi_{S\triangle T}(x).$$

Luego se concluye lo deseado. □

Proposición 1.0.4. *Sea* $\mathbf{x} \sim \{\pm 1\}^n$, *tenemos*

$$\mathbb{E}[\chi_S(\mathbf{x})] = \mathbb{E}[\prod_{i \in S} x_i] = \begin{cases} 1 & \text{si } S = \emptyset, \\ 0 & \text{si } S \neq \emptyset. \end{cases}$$

Demostración. Si $S = \emptyset$ entonces $\mathbb{E}[\chi_S(\mathbf{x})] = \mathbb{E}[1] = 1$. Caso contrario

$$\mathbb{E}[\prod_{i \in S} \mathbf{x}_i] = \prod_{i \in S} \mathbb{E}[\mathbf{x}_i],$$

porque los bit aleatorios (variables aleatorias uniformes en $\{\pm 1\}$) $\mathbf{x}_1, ..., \mathbf{x}_n$ son independientes. Pero cada uno de los factores $\mathbb{E}[x_i]$ en el producto anterior (no vacío) es (1/2) (+ 1) + (1/2) (-1) = 0. □

Como hemos visto, la expansión de Fourier de $f : \{\pm 1\}^n \to \mathbb{R}$ puede considerarse como la representación de f sobre la base ortonormal de las funciones de Walsh $(\chi_S)_S$, $S \subset [n]$. Sobre esta base, f tiene 2^n "coordenadas", y estos son precisamente los coeficientes de Fourier de f. La "coordenada" de f en la "dirección" χ_S es $\langle f, \chi_S \rangle$; es decir, tenemos la siguiente fórmula para los coeficientes de Fourier:

Proposición 1.0.5. *Sean* $f : \{\pm 1\}^n \to \mathbb{R}$ *y* $S \subset [n]$, *el coeficiente de Fourier de* f *está dado por*

$$\hat{f}(S) = \langle f, \chi_S \rangle = \mathbb{E}[f(\mathbf{x})\chi_S(\mathbf{x}))].$$

Podemos verificar esta fórmula explícitamente:

$$\langle f, \chi_S \rangle = \langle \sum_{T \subset [n]} \hat{f}(T)\chi_T, \chi_S \rangle = \sum_{T \subset [n]} \hat{f}(T)\langle \chi_T, \chi_S \rangle = \hat{f}(S),$$

donde usamos la expansión de Fourier de f, la linealidad de $\langle \cdot, \cdot \rangle$, y finalmente el Teorema 1.0.2. Esta fórmula es la forma más simple de calcular los coeficientes de Fourier de una función dada; también se puede ver como una versión simplificada del método de interpolación ilustrado en la introducción. Alternativamente, esta fórmula puede tomarse como la definición de los coeficientes de Fourier. La base ortonormal de las funciones de Walsh también nos permite medir la "longitud" al cuadrado (norma 2) de $f : \{\pm 1\}^n \to \mathbb{R}$ eficientemente: es solo la suma de los cuadrados de las "coordenadas" de f, es decir, los coeficientes de Fourier. Este hecho simple pero crucial se llama Teorema de Parseval.

Teorema 1.0.6. *(Teorema de Parseval) Para toda* $f : \{\pm 1\}^n \to \mathbb{R}$,

$$\langle f, f \rangle = \mathbb{E}[f(\mathbf{x})^2] = \sum_{S \subset [n]} \hat{f}(S)^2.$$

En particular, si $f : \{\pm 1\}^n \to \{\pm 1\}$ *entonces*

$$\sum_{S \subset [n]} \hat{f}(S)^2 = \mathbb{E}[f(\mathbf{x})^2] = \mathbb{E}[1] = 1.$$

Más generalmente, dadas dos funciones $f, g : \{\pm 1\}^n \to \mathbb{R}$, podemos calcular $\langle f, g\rangle$ tomando el "producto punto" de sus coordenadas en la base ortonormal de Walsh. De esta manera tenemos el Teorema de Plancherel.

Teorema 1.0.7. *(Teorema de Plancherel) Para todas $f, g : \{\pm 1\}^n \to \mathbb{R}$,*

$$\langle f, g\rangle = \mathbb{E}[f(\mathbf{x})g(\mathbf{x})] = \sum_{S\subset[n]} \hat{f}(S)\hat{g}(S).$$

Podemos verificar esta fórmula explícitamente:

$$\langle f, g\rangle = \langle \sum_{S\subset[n]} \hat{f}(S)\chi_S, \sum_{T\subset[n]} \hat{g}(T)\chi_T\rangle = \sum_{S,T\subset[n]} \hat{f}(S)\hat{g}(T)\langle\chi_S, \chi_T\rangle = \sum_{S\subset[n]} \hat{f}(S)\hat{g}(S),$$

y así concluir lo deseado.

1.0.2. Radios de Bohr

Los últimos años mostraron un notable aumento de artículos sobre el fenómeno de Bohr en el análisis complejo. En término del análisis de Fourier, podemos reformular el Teorema de Bohr 0.0.1 con la siguiente notación: el radio $r = 1/3$ es el valor óptimo (más grande) para el cual se verifica

$$\sum_{k=0}^{d} |\hat{f}(k)| r^k \leqslant \|f\|_{\mathbb{T}}, \tag{1.3}$$

donde $\|f\|_{\mathbb{T}}$ es la norma supremo en el toro $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$, $\hat{f}(k)$ como es usual denota el k-ésimo coeficiente de f como función en $\mathbb{T}$. Como ya comentamos en la introducción, este interesante resultado fue mayormente olvidado hasta los trabajos relativamente recientes de Dineen, Dixon, y Boas y Khavinson. Por ejemplo en [BK97] los autores analizaron si se producía un fenómeno similar para series de potencias en varias variables. En lo que sigue escribiremos los radios de Bohr con la notación que usaremos en el desarrollo de la tesis: tomando $N \in \mathbb{N}$, el N-ésimo radio de Bohr K_N es el mejor (i.e. mayor) $0 < r < 1$ tal que para todo polinomio f en N variables complejas se cumple

$$\sum_{\alpha\in\mathbb{N}_0^N} |\hat{f}(\alpha)|\ r^\alpha \leqslant \|f\|_{\mathbb{T}^N}, \tag{1.4}$$

donde ahora el α-ésimo coeficiente de Fourier de f está dado por

$$\hat{f}(\alpha) = \int_{\mathbb{T}^{\mathbb{N}}} f(w)w^{-\alpha}dw,$$

(dw es la medida de Lebesgue normalizada en el toro N-dimensional $\mathbb{T}^{\mathbb{N}}$ y $\|f\|_{\mathbb{T}^N}$ denota la norma supremo de f en $\mathbb{T}^N$). Como ya comentamos sólo se conoce el valor exacto para el caso unidimensional, $K_1 = 1/3$. Se mostró en [BK97] a través de métodos probabilísticos que

$$\limsup_{N\to\infty} \sqrt{\frac{N}{\log N}} K_N \leqslant 1,$$

y en [DFOC$^+$11] basado en la hipercontractividad de la desigualdad de Bohnenblust-Hille tenemos

$$\sqrt{\frac{N}{\log N}} K_N \geqslant \frac{1}{\sqrt{2}} + o(1).$$

La mejora de las técnicas que conducen al resultado anterior podemos encontrarla en [PSS14], donde se muestra

$$\lim_{N\to\infty} \sqrt{\frac{N}{\log N}} K_N = 1. \tag{1.5}$$

Notar que esta igualdad es básicamente un resultado en el decaimiento asintótico del espectro de Fourier $(\hat{f}(\alpha))_{\alpha\in\mathbb{N}_0^N}$ de polinomios f en el grupo abeliano compacto $\mathbb{T}^N$ para dimensiones altas de N.

En la definición de radio Bohr podemos reemplazar la familia de todos los polinomios en N variables por subfamilias de ésta y obtener nuevos radios. Por ejemplo, si nos quedamos con la subfamilia de los polinomios homogéneos, naturalmente definimos el N-radio de Bohr homogéneo

$$K_N^{hom}, \tag{1.6}$$

como la mejor constante r $(0 < r < 1)$ que cumple la desigualdad (1.4) pero ahora para los polinomios homogéneos en N variables.

Si nos quedamos con todos los polinomios en N variables de grado a lo sumo m, definimos

$$K_N^{\leqslant m},$$

es decir el mejor r $(0 < r < 1)$ tal que

$$\sum_{\alpha\in\mathbb{N}_0^N, |\alpha|\leqslant m} |\hat{f}(\alpha)|\ r^\alpha \leqslant \|f\|_{\mathbb{T}}, \tag{1.7}$$

para todo polinomio $f(z) = \sum_{\alpha\in\mathbb{N}_0^N,|\alpha|\leqslant m} \hat{f}(\alpha)z^\alpha$. De igual manera, tomando todos los polinomios m-homogéneos en N variables, obtenemos la definición del N-ésimo radio m-homogéneo de Bohr, al mismo denotaremos como

$$K_N^{=m}. \tag{1.8}$$

1.0.3. Constantes de Bohnenblust-Hille caso real.

Una comparación de la prueba que haremos del Teorema 0.0.2 con (5) muestra que en el caso Booleano nuevas dificultades aparecen. Observando el desarrollo de Fourier-Walsh

(1.1) tenemos que toda función Booleana $f : \{\pm 1\}^n \to \mathbb{R}$ es la restricción de un único polinomio $P_f : \mathbb{R}^n \to \mathbb{R}$ afín en cada variable

$$P_f(x) = \sum_{S \subset [n]} \widehat{f}(S) \prod_{k \in S} x_k, \qquad x \in \mathbb{R}^n.$$

Este tipo especial de polinomio lo llamaremos tetraedral

$$\begin{tikzcd} \{\pm 1\}^n \arrow[rr, hook] \arrow[dr, "f"'] & & \mathbb{R}^n \arrow[dl, "P_f"] \\ & \mathbb{R} & \end{tikzcd}$$

Esta identificación es una isometría para la norma supremo, es decir

$$\|f\|_\infty := \sup_{x \in \{\pm 1\}^n} |f(x)| = \sup_{x \in \{\pm 1\}^n} |P_f(x)| = \sup_{x \in [-1,1]^n} |P_f(x)| =: \|P_f\|_\infty, \tag{1.9}$$

donde la anteúltima igualdad es válida ya que el polinomio es afín en cada variable: pues si suponemos que no, el máximo de f se alcanza en un punto del cubo donde alguna coordenada es menor en módulo a 1, fijando las demás, obtenemos una función lineal no constante en $[-1, 1]$ y el máximo se alcanza en los extremos.

Entonces podemos mirar a este objeto de forma diferente, combinando técnicas del análisis de Fourier en teoría de polinomios y formas multilineales para estudiar el espectro de Fourier de estas funciones. Basado en resultados e ideas que se han desarrollado en los últimos años en el campo de polinomios complejos llevaremos a cabo el trabajo.

¿Qué ocurre con la desigualdad de Bohnenblust-Hille para los polinomios reales en el cubo n-ésimo $[-1, 1]^n$?

Denotamos por

$$BH^{\leqslant d}_{[-1,1]},$$

la mejor constante $C \geqslant 1$ tal que para todo polinomio real de grado d, $P(x) = \sum_{\alpha \in \mathbb{N}_0^n, |\alpha| \leqslant d} a_\alpha x^\alpha$ tenemos

$$\left(\sum_{\alpha \in \mathbb{N}_0^n, |\alpha| \leqslant d} | a_\alpha|^{\frac{2d}{d+1}} \right)^{\frac{d+1}{2d}} \leqslant BH^{\leqslant d}_{[-1,1]} \|f\|_{[-1,1]^n}; \tag{1.10}$$

y análogamente reemplazando los polinomios reales de grado d por los d-homogéneo reales en la definición anterior podemos definir la constante $BH^{=d}_{[-1,1]}$, con esta notación tenemos el siguiente resultado.

Teorema 1.0.8.

$$\operatorname{lím\,sup}_{d \to \infty} \sqrt[d]{BH^{=d}_{[-1,1]}} = 2, \tag{1.11}$$

y

$$\operatorname{lím\,sup}_{d \to \infty} \sqrt[d]{BH^{\leqslant d}_{[-1,1]}} = 1 + \sqrt{2}. \tag{1.12}$$

Vale destacar que la primera igualdad se debe a [CJRMF$^+$15]. Mientras que el teorema anterior muestra que para polinomios reales generales la situación es dramáticamente diferente del caso complejo (6) donde los límites son iguales. El resultado del Teorema 0.0.2 indica que los polinomios tetraedrales parecen estar a medio camino.

Para dar una prueba del Teorema 1.0.8 necesitamos resultados preliminares, recolectamos algunas desigualdades cruciales para polinomios que "regulan el tráfico" entre la norma supremo de un polinomio real en $[-1,1]^n$ con la norma supremo en $\mathbb{T}^n$, como también la misma relación entre los polinomios homogéneos

Lema 1.0.9. *Sea $P(z) = \sum_{|\alpha| \leqslant d} a_\alpha z^\alpha$ un polinomio en $\mathbb{C}^n$ con coeficientes complejos a_α, y con $m \leqslant d$ denotamos por $P_m(z) = \sum_{|\alpha|=m} a_\alpha z^\alpha$ a su parte m-homogénea. Entonces*

1. $\|P_m\|_{\mathbb{T}^n} \leqslant \|P\|_{\mathbb{T}^n}$;
2. $\|P\|_{\mathbb{T}^n} \leqslant (1+\sqrt{2})^d \|P\|_{[-1,1]^n}$;
3. $\|P_d\|_{\mathbb{T}^n} \leqslant 2^{d-1}\|P\|_{[-1,1]^n}$ *para todo P con coeficiente reales. Más aún $C = 2$ es la mejor constante posible que cumple* $\|P_d\|_{\mathbb{T}^n} \leqslant C^{d-1}\|P\|_{[-1,1]^n}$.
4. $\|P_m\|_{[-1,1]^n} \leqslant (1+\sqrt{2})^d \|P\|_{[-1,1]^n}$ *para todo P con coeficientes reales. Más aún $C = 1+\sqrt{2}$ es la mejor constante posible que cumple* $\|P_m\|_{[-1,1]^n} \leqslant C^d \|P\|_{[-1,1]^n}$.

A continuación demostraremos *1*, comentaremos las ideas para probar *4* y en cuanto a la prueba de *2* y *3* podemos encontrarlas en el trabajo de Klimek, sobre métricas en la familia de conjuntos L-regulares compactos polinómicamente convexos en $\mathbb{C}$ [Kli95] y en la generalización por parte de Visser de lo que se conoce como Teorema de Markov [Vis46] respectivamente.

Demostración. Comencemos por el primer ítem. Sabemos por la fórmula de Cauchy que si consideramos para cada $x \in \mathbb{T}^n$ el polinomio $P_x(w) = P(wx_1, ..., wx_n) = \sum_{m=0}^d w^m P_m(x)$, $w \in \mathbb{C}$ tenemos

$$|P_m(x)| = \left| \frac{1}{2\pi} \int_{|w|=1} \frac{P_x(w)}{w^{m+1}} dw \right| \leqslant \sup_{|w|=1} |P_x(w)| \leqslant \sup_{z \in \mathbb{T}^n} |P(z)|,$$

la última desigualdad se debe a que si $|w| = 1$ y $x \in \mathbb{T}^n$ entonces $wx \in \mathbb{T}^n$, luego concluimos la afirmación tomando supremo en todo los $x \in \mathbb{T}^n$.

Por último comentemos las ideas principales para probar el ítem *4*. Recordemos la definición de los polinomios de Chebyshev. Dado $k \in \mathbb{N} \cup \{0\}$ definimos

$$T_k(x) = \cos(k \arccos(x))$$

el k-ésimo polinomio de Chebychev. Mediante propiedades del coseno, se puede obtener una relación de recurrencia entre estos polinomios:

$$\begin{cases} T_0(x) = 1 \\ T_1(x) = x \\ T_{k+1}(x) = 2xT_k(x) - T_{k-1}(x), \end{cases} \tag{1.13}$$

y así una equivalente definición recursiva. Escribimos

$$T_d(x) = \sum_{m=0}^{d} a_m(T_d)x^m, \qquad x \in \mathbb{R},$$

el polinomio de Chebyshev de grado d, del cual se conocen muchas propiedades, por ejemplo, usando la definición no recursiva, se puede ver que $\|T_d\|_\infty = 1$. Podemos considerar también los números de Markov $M_{m,d}$ para $0 \leqslant m \leqslant d$ dados por

$$M_{m,d} = \begin{cases} |a_m(T_d)| & \text{si } m \equiv d(2) \\ |a_m(T_{d-1})| & \text{si } m \equiv d-1(2), \end{cases} \tag{1.14}$$

donde

$$|a_m(T_d)| = 2^{m-1}\frac{\hat{d}}{m!}\frac{(\frac{\hat{d}+m-2}{2})!}{(\frac{\hat{d}-m}{2})!},$$

con $\hat{d} = d$ si $m \equiv d \pmod 2$, y $\hat{d} = d-1$ si $m \equiv d-1 \pmod 2$ (véase [Nat64, p. 56]).

Un punto crucial a lo largo del trabajo es lo que se conoce como el Teorema de Markov (ver [BE95, p. 248]), cuya demostración podemos encontrar en [Nat64, p.56], el cual dice que todo polinomio real $p(t) = \sum_{m=0}^{d} a_m t^m$ de grado d satisface

$$|a_m| \leqslant M_{m,d} \sup_{t\in[-1,1]} |p(t)|, \qquad 0 \leqslant m \leqslant d, \tag{1.15}$$

y tomando T_d se muestra que la estimación es óptima. Sea un polinomio P en $[-1,1]^n$ con coeficientes reales, fijando $x \in \mathbb{R}^n$ tenemos el polinomio

$$p_x(t) = P(tx_1, ..., tx_n) = \sum_{m=0}^{d} t^m P_m(x), \qquad t \in \mathbb{R},$$

entonces usando el Teorema de Markov y tomando supremo en todos los $x \in [-1,1]^n$, obtenemos

$$\|P_m\|_{[-1,1]^n} \leqslant M_{m,d}\|P\|_{[-1,1]^n}.$$

Se puede probar utilizando la fórmula de Stirling y la expresión explícita de $M_{m,d}$ que

$$\limsup_{d\to\infty}(\sup_{m\leqslant d} \sqrt[d]{M_{m,d}}) = 1 + \sqrt{2}. \tag{1.16}$$

Esto prueba la desigualdad del ítem *4* con constante $(1+\sqrt{2})^d$. Finalmente, si asumimos que la desigualdad se logra con alguna constante C^d, tomando el polinomio de Chebyshev T_d conseguimos que $M_{m,d} \leqslant C^d$ para todo $m \leqslant d$, y nuevamente por (1.16) obtenemos $1 + \sqrt{2} \leqslant C$. $\square$

Con estas nuevas herramientas probemos el Teorema 1.0.8. Comencemos con la demostración de la primera parte (1.11), más precisamente la estimación superior, que es una consecuencia inmediata de (5) combinado con el ítem *3* del Lema 1.0.9, pues todo polinomio d-homogéneo con coeficientes reales $P(z) = \sum_{|\alpha|=d} a_\alpha z^\alpha$ cumple la siguiente desigualdad

$$\left(\sum_{\alpha \in \mathbb{N}_0^n, |\alpha|=d} | a_\alpha|^{\frac{2d}{d+1}} \right)^{\frac{d+1}{2d}} \leqslant BH_{\mathbb{T}}^{=d} \|P\|_{\mathbb{T}^n} \leqslant BH_{\mathbb{T}}^{=d} 2^{d-1} \|P\|_{[-1,1]^n}.$$

Simplemente usando la definición de $BH_{[-1,1]}^{=d}$ tenemos que

$$BH_{[-1,1]}^{=d} \leqslant BH_{\mathbb{T}}^{=d} 2^{d-1}.$$

Además como ya comentamos antes (ver (6)),

$$\limsup_{d \to \infty} \sqrt[d]{BH_{\mathbb{T}}^{=d}} = 1,$$

por lo que concluimos la estimación superior

$$\limsup_{d \to \infty} \sqrt[d]{BH_{[-1,1]}^{=d}} \leqslant 2.$$

Para probar la acotación inferior agreguemos un comentario/definición sobre las constantes de Bohnenblust-Hille. Por ejemplo tomemos el caso real. Por definición de $BH_{[-1,1]}^{\leqslant d}$ (1.10) sabemos que es la mejor constante, tal que para todo polinomio real P de grado d $P(x) = \sum_{\alpha \in \mathbb{N}_0^n, |\alpha| \leqslant d} a_\alpha x^\alpha$ se cumple

$$\left(\sum_{\alpha \in \mathbb{N}_0^n, |\alpha| \leqslant d} |a_\alpha|^{\frac{2d}{d+1}} \right)^{\frac{d+1}{2d}} \leqslant BH_{[-1,1]}^{\leqslant d} \|P\|_{[-1,1]^n}.$$

Notar que la constante es independiente de n el número de variables del polinomio. Por lo tanto, podríamos fijar el $n \in \mathbb{N}$ y reducirnos a la subfamilia de polinomios definidos en $[-1, 1]^n$, en este caso la mejor constante $C \geqslant 1$ tal que cumple la desigualdad de Bohnenblust-Hille para todo polinomio de grado d con coeficientes reales en $[-1, 1]^n$, la denotaremos

$$BH_{[-1,1]^n}^{\leqslant d}$$

de esta manera hacemos referencia que la cantidad de variables está previamente fijada. Naturalmente tenemos la siguiente desigualdad

$$BH_{[-1,1]^n}^{\leqslant d} \leqslant BH_{[-1,1]}^{\leqslant d}. \tag{1.17}$$

Ahora estamos en condiciones de probar la estimación inferior de (1.11), especificamente demostremos la siguiente proposición:

Proposición 1.0.10. *Sea $k \in \mathbb{N}$ fijo, entonces*

$$\limsup_{d\to\infty} \left(BH^{=d}_{[-1,1]^{2^k}} \right)^{\frac{1}{d}} \geqslant 2^{1-2^{-k}}. \tag{1.18}$$

Por lo tanto, concluimos

$$\limsup_{d\to\infty} \left(BH^{=d}_{[-1,1]} \right)^{\frac{1}{d}} \geqslant 2. \tag{1.19}$$

En lo que sigue haremos uso de la norma $|\cdot|_\infty$, definida para todo polinomio m-homogéneos en n variables $P = \sum_{\alpha\in\Lambda(m,n)} a_\alpha z^\alpha$ donde $\Lambda(m,n) = \{\alpha \in \mathbb{N}_0^n : |\alpha| := \alpha_1 + ... + \alpha_n = m\}$. Dicha norma se define como

$$|P|_\infty = \max_{\alpha\in\Lambda(m,n)} |a_\alpha|.$$

Además $\|P\|_q = (\sum_{\alpha\in\Lambda} |a_\alpha|^q)^{\frac{1}{q}}$ para $q \in \mathbb{Q}$.

Demostración. Consideramos la sucesión de polinomios (con norma uniforme 1) definida recursivamente por

$$\begin{aligned} Q_2(x_1, x_2) &= x_1^2 - x_2^2, \\ Q_{2^m}(x_1, ..., x_{2^m}) &= Q_{2^{m-1}}(x_1, ..., x_{2^{m-1}})^2 - Q_{2^{m-1}}(x_{2^{m-1}+1}, ..., x_{2^m})^2. \end{aligned}$$

Demostremos (por inducción en m) que

$$|Q_{2^m}^n|_\infty \geqslant \left(\frac{2^n}{n+1}\right)^{2^m-1}, \tag{1.20}$$

para cada número natural n.

Fijemos $n \in \mathbb{N}$ y consideramos $m = 1$, sabemos que

$$2^n = \sum_{k=0}^{n} \binom{n}{k} \leqslant (n+1) \max_{0\leqslant k\leqslant n} \binom{n}{k},$$

el polinomio $2n$-homogéneo Q_2^n admite la siguiente estimación:

$$\|Q_2^n\|_{\frac{4n}{2n+1}} \geqslant |Q_2^n|_\infty = \max_{0\leqslant k\leqslant n} \binom{n}{k} \geqslant \frac{2^n}{n+1}. \tag{1.21}$$

Supongamos ahora que la ecuación (1.20) es cierta para todo $k \in \mathbb{N}$ tal que $k \leqslant m$, probemos la fórmula para $m + 1$. Observemos que

$$Q_{2^{m+1}}^n(x_1, x_2, ..., x_{2^{m+1}}) = \sum_{k=0}^{n} \binom{n}{k} (-1)^{n-k} Q_{2^m}^{2k}(x_1, ..., x_{2^m}) Q_{2^m}^{2(n-k)}(x_{2^m+1}, ..., x_{2^{m+1}}). \tag{1.22}$$

El coeficiente de valor absoluto máximo en un producto de polinomios en conjuntos disjuntos de variables es el producto de los respectivos coeficientes máximos, por lo tanto

$$|Q^n_{2^{m+1}}|_\infty = \max_{0\leqslant k\leqslant n} \binom{n}{k} |Q^{2k}_{2^m}|_\infty |Q^{2(n-k)}_{2^m}|_\infty \geqslant \max_{0\leqslant k\leqslant n} \binom{n}{k} \left(\frac{2^{2n}}{(2k+1)(2n-2k+1)}\right)^{2^m-1},$$

por hipótesis inductiva. Sin embargo, $(2k+1)(2n-2k+1) \leqslant (n+1)^2$ cuando $0 \leqslant k \leqslant n$; así

$$|Q^n_{2^{m+1}}|_\infty \geqslant \left(\frac{2^n}{n+1}\right)^{2^{m+1}-2} \max_{0\leqslant k\leqslant n} \binom{n}{k} \geqslant \left(\frac{2^n}{n+1}\right)^{2^{m+1}-1},$$

por la ecuación (1.21). Por lo tanto, la fórmula dada en (1.20) se cumple para cada entero positivo m. Algo que vale en general es que un polinomio n-homogéneo P cumple

$$\|P\|_{\frac{2n}{n+1}} \geqslant |P|_\infty.$$

Por la ecuación (1.20) tenemos

$$BH^{=n2^m}_{[-1,1]^{2^m}} \geqslant \|Q^n_{2^{m+1}}\|_{\frac{2n}{n+1}} \geqslant |Q^n_{2^{m+1}}|_\infty \geqslant \left(\frac{2^n}{n+1}\right)^{2^m-1}.$$

Considerando $d = n2^k$ con k fijo y usando la estimación anterior obtenemos

$$\left(BH^{=d}_{[-1,1]^{2^k}}\right)^{\frac{1}{d}} \geqslant \left(\frac{2^n}{n+1}\right)^{\frac{2^k-1}{n2^k}} = 2^{1-2^{-k}}(n+1)^{\frac{n(1-2^k)}{2^k}}.$$

y la prueba sigue directamente tomando límite en n. □

Pasemos a probar la igualdad (1.12). Tomemos un polinomio real P de grado d en $[-1,1]^n$, podemos escribirlo como $\sum_{m=0}^d P_m$. Entonces deducimos por la desigualdad de Minkowski y el ítem 1 del Lema 1.0.9 que

$$\begin{aligned}\left(\sum_{\alpha\in\mathbb{N}_0^n, |\alpha|\leqslant d} |a_\alpha|^{\frac{2d}{d+1}}\right)^{\frac{d+1}{2d}} &\leqslant \sum_{m=0}^d \left(\sum_{\alpha\in\mathbb{N}_0^n, |\alpha|=m} |a_\alpha|^{\frac{2m}{m+1}}\right)^{\frac{m+1}{2m}} \\ &\leqslant \sum_{m=0}^d BH^{=m}_{\mathbb{T}} \|P_m\|_{\mathbb{T}^n} \leqslant \sum_{m=0}^d BH^{=m}_{\mathbb{T}} \|P\|_{\mathbb{T}^n}.\end{aligned}$$

Y entonces por el ítem 2 del Lema 1.0.9 y la desigualdad (5) tenemos

$$\begin{aligned}\left(\sum_{\alpha\in\mathbb{N}_0^n, |\alpha|\leqslant d} |a_\alpha|^{\frac{2d}{d+1}}\right)^{\frac{d+1}{2d}} &\leqslant \sum_{m=0}^d BH^{=m}_{\mathbb{T}} (1+\sqrt{2})^d \|P\|_{[-1,1]^n} \\ &\leqslant C^{\sqrt{d\log d}} d(1+\sqrt{2})^d \|P\|_{[-1,1]^n}.\end{aligned}$$

Con esto probamos
$$BH_{[-1,1]}^{\leqslant d} \leqslant C^{\sqrt{d \log d}} d(1+\sqrt{2})^d,$$
y por lo tanto la acotación superior. Por otro lado aplicando esta desigualdad a los polinomios de Chebyshev T_d, por (1.16) observamos
$$1+\sqrt{2} = \limsup_{d\to\infty} \left(\sup_{m \leqslant d} \sqrt[d]{M_{m,d}} \right) \leqslant \limsup_{d\to\infty} \sqrt[d]{BH_{[-1,1]}^{\leqslant d}},$$
y así la conclusión.

1.0.4. Constantes de Sidon

Podemos pensar las constantes de Bohnenblust-Hille del cubo Booleano (2) y del politoro (3) en términos de ciertos conjuntos llamados conjuntos de Sidon. Más formalmente, tomemos $1 \leqslant p \leqslant \infty$ y un grupo compacto abeliano G. Un subconjunto finito Λ del grupo dual $\hat{G}$ se dice que es un conjunto p-Sidon si existe una constante $C > 0$ (que depende de p y Λ) tal que para todo polinomio trigonométrico $f = \sum_{\gamma\in\Lambda} \hat{f}(\gamma)\gamma$ satisface
$$\left(\sum_{\gamma\in\Lambda} |\hat{f}(\gamma)|^p \right)^{\frac{1}{p}} \leqslant C\|f\|_G, \tag{1.23}$$
donde $\|f\| = \sup_{g\in G} |f(g)|$, además llamamos constante de Sidon $S_p(\Lambda)$ a la mejor C que cumple (1.23). Nos concentraremos en los siguientes grupos: $\mathbb{T}^n$ el politoro n-dimensional, y $\{\pm 1\}^n$ el n-ésimo cubo Booleano. Recordemos que el grupo dual de $\mathbb{T}^n$ consiste en todos los monomios $z^\alpha, \alpha \in \mathbb{Z}^n$, mientras que el grupo dual de $\{\pm 1\}^n$ está formado por todos los monomios x^S, $S \subset [n]$. Ponemos el foco en los siguientes cuatro conjuntos de caracteres en el grupo dual de $\mathbb{T}^n$ y $\{\pm 1\}^n$, respectivamente:
$$\Omega_n^{=d} = \{z^\alpha : \alpha \in \mathbb{N}_0^n, |\alpha| = d\}, \qquad \Omega_n^{\leqslant d} = \{z^\alpha : \alpha \in \mathbb{N}_0^n, |\alpha| \leqslant d\},$$
$$\Lambda_n^{=d} = \{x^S : S \subset [n], |S| = d\}, \qquad \Lambda_n^{\leqslant d} = \{x^S : S \subset [n], |S| \leqslant d\}.$$
Con esta terminología tenemos
$$BH_{\{\pm 1\}}^{\leqslant d} = \sup_n S_{\frac{2d}{d+1}}(\Lambda_n^{\leqslant d}) \leqslant \infty,$$
y
$$BH_{\mathbb{T}}^{\leqslant d} = \sup_n S_{\frac{2d}{d+1}}(\Omega_n^{\leqslant d}) \leqslant \infty,$$
naturalmente tenemos que la reformulación de (2) y (3). Más aún, el Teorema 0.0.2 afirma que existe $C \geqslant 1$ tal que para cada n, d
$$S_{\frac{2d}{d+1}}(\Lambda_n^{\leqslant d}) \leqslant C^{\sqrt{d. \log d}};$$
notemos también que por (5) el análogo para $S_{\frac{2d}{d+1}}(\Omega_n^{\leqslant d})$ también es cierto.

Capítulo 2

Espectro de Fourier en el cubo Booleano

El objetivo de este capítulo será demostrar que la constante de Bohnenblust-Hille para el caso Booleano crece subexponencialmente como comentamos en la introducción en el Teorema 0.0.2. Primero probaremos el caso d-homogéneo para ilustrar la estrategia de la demostración que se generaliza al caso de grado d.

Teorema 2.0.1. *(Caso homogéneo.) Existe una costante absoluta $C > 0$ tal que para todo $d \in \mathbb{N}$*

$$BH_{\{\pm 1\}}^{=d} \leqslant C^{\sqrt{d \log d}}. \tag{2.1}$$

Aunque el Teorema 0.0.2, por supuesto, cubre este caso, preferimos dar una prueba independiente ya que es más simple e ilustra la estrategia general. De todos modos, todas las pruebas de las desigualdades de tipo Bohnenblust-Hille, antiguas o muy recientes, están muy inspiradas por la prueba de la estimación original, que es en sí misma un refinamiento de la prueba de la famosa desigualdad $\frac{4}{3}$ de Littlewood [Lit30]. Recordemos sus cuatro pasos cruciales:

1. Dado un polinomio d-homogéneo $f(z) = \sum_{|\alpha|=d} a_\alpha z^\alpha$ en n variables. Sabemos que este es la reestricción de una d-forma lineal simétrica L_f en $\mathbb{C}^n$ sobre la diagonal $\Delta = \{(z, ..., z) : z \in \mathbb{C}^n\}$.

2. Interpretamos L_f como una "matriz" simétrica $\mathrm{A} = (c_{i_1,\dots,i_m})_{i_1,\dots,i_d=1}^n$, relacionando la $\ell_{\frac{2d}{d+1}}$-norma de los coeficientes de f con la $\ell_{\frac{2d}{d+1}}$-norma en las entradas de A, y dividimos esta norma en los términos mixtos de la forma

$$\sum_{i_k} \left(\sum_{i_1,\dots,i_{k-1},\hat{i}_k,i_{k+1},\dots,i_d} |c_{i_1,\dots,i_d}|^2 \right)^{\frac{1}{2}}.$$

3. Usando la desigualdad de Khinchine podemos estimar los términos mixtos por la norma supremo de L_f en $(\mathbb{T}^n)^d$.

4. Finalmente usamos polarización para estimar $\|L_f\|_\infty$ por $\|f\|_\infty$.

Para la mejora de las constantes, se necesitaron paso a paso refinamientos no triviales de los argumentos anteriores. Modificando varios argumentos de [DFOC$^+$11] y [PSS14], escribimos la prueba del Teorema 2.0.1 y comenzamos a recopilar resultados relevantes que también serán necesarios para la prueba más complicada del Teorema 0.0.2, es decir el caso de grado d.

2.1. El caso d-homogéneo

Antes de comenzar con la demostración tenemos, como mencionamos antes, que definir algunos conceptos para poder enunciar cierto tipo de resultados que serán la clave de la prueba. Estos resultados son conocidos como la desigualdad de Blei, la hipercontractividad de Bonami y la fórmula de polarización. Demostraremos al paso la fórmula de polarización y las dos primeras desigualdades tendrán un apartado especial en otro capítulo del trabajo.

2.1.1. Formas multilineales

Para cada $n \in \mathbb{N}$ y $A \subset \mathbb{N}$ conjunto finito, definimos los conjuntos de funciones

$$\mathcal{I}(A,n) := \{\mathbf{i} : A \to [n]\},$$
$$\mathcal{J}(A,n) := \{\mathbf{j} : A \to [n] \textit{ no decreciente}\}.$$

En el caso $A = [d]$ para $d \in \mathbb{N}$, denotamos a los conjuntos previos por $\mathcal{I}(d,n)$ y $\mathcal{J}(d,n)$ respectivamente. Consideramos la relación de equivalencia en $\mathcal{I}(A,n)$ definida para $\mathbf{i}_1, \mathbf{i}_2 \in \mathcal{I}(A,n)$ como $\mathbf{i}_1 \sim \mathbf{i}_2$ si y sólo si existe una biyección $\sigma : A \to A$ tal que $\mathbf{i}_1 = \mathbf{i}_2 \circ \sigma$.

Notar que para cada $\mathbf{i} \in \mathcal{I}(A,n)$ tenemos un único $\mathbf{j} \in \mathcal{J}(A,n)$ tal que $\mathbf{j} \in [\mathbf{i}]$, donde $[\mathbf{i}]$ es la clase de equivalencia de $\mathbf{i}$. Más aún, el número de elementos de la clase de equivalencia de $\mathbf{i} \in \mathcal{I}(A,n)$ es

$$|[\mathbf{i}]| = \frac{|A|!}{|\{k : \mathbf{i}(k) = 1\}|! \ldots |\{k : \mathbf{i}(k) = n\}|!}. \tag{2.2}$$

Si A_1 y A_2 son dos subconjuntos disjuntos de $\mathbb{N}$, entonces la suma directa de $\mathbf{i}_1 \in \mathcal{I}(A_1,n)$ y $\mathbf{i}_2 \in \mathcal{I}(A_2,n)$ está definida como la siguiente función

$$\mathcal{I}(A_1 \cup A_2, n) \ni \mathbf{i}_1 \oplus \mathbf{i}_2 : A_1 \cup A_2 \to [n],$$

que coincide con $\mathbf{i}_k$ en A_k para cada $k = 1, 2$. Notar que todo elemento de $\mathcal{I}(A_1 \cup A_2, n)$ podemos representarlo de esta manera, esto dice en particular, que si $S \subset [d]$ y denotamos $\hat{S} := [d] \backslash S$, entonces $\mathcal{I}(d,n) = \mathcal{I}(S,n) \oplus \mathcal{I}(\hat{S},n)$. Sin embargo, sólo podemos asumir $\mathcal{J}(d,n) \subset \mathcal{J}(S,n) \oplus \mathcal{J}(\hat{S},n)$. El próximo resultado, normalmente conocido como la desigualdad de Blei, afirma lo siguiente:

Proposición 2.1.1. *Sean $n \in \mathbb{N}$ y $0 \leqslant k \leqslant d$ enteros. Para cada colección de escalares $(a_{\mathbf{i}})_{\mathbf{i}\in\mathcal{I}(d,n)}$ tenemos*

$$\left(\sum_{\mathbf{i}\in\mathcal{I}(d,n)} |\, a_{\mathbf{i}}|^{\frac{2d}{d+1}}\right)^{\frac{d+1}{2d}} \leqslant \left[\prod_{S\subset[d],|S|=k} \left(\sum_{\mathbf{i}_1\in\mathcal{I}(S,n)} \left(\sum_{\mathbf{i}_2\in\mathcal{I}(\hat{S},n)} |\, a_{\mathbf{i}_1\oplus\mathbf{i}_2}|^2\right)^{\frac{1}{2}\frac{2k}{k+1}}\right)^{\frac{k+1}{2k}}\right]^{\frac{1}{\binom{d}{k}}}. \quad (2.3)$$

La idea de la demostración será desarrollada en el capítulo 4 con mayor detalle, en esta sección cumple un rol fundamental en la prueba del resultado principal, como ya veremos.

Toda función d-homogénea $f : \{\pm 1\}^n \to \mathbb{R}$ tiene una única representación

$$f(x) = \sum_{\mathbf{j}\in\mathcal{J}(d,n)} a_{\mathbf{j}} x_{\mathbf{j}} \qquad x_{\mathbf{j}} := x_{j_1} \cdot \ldots \cdot x_{j_d}, \quad (2.4)$$

donde $a_{\mathbf{j}} = 0$ si $\mathbf{j}$ no es inyectiva (i.e no es estrictamente creciente). Podemos observar que f es la restricción a $\{\pm 1\}^n$ de un polinomio d-homogéneo $P_f : \mathbb{R}^n \to \mathbb{R}$ dado por

$$P_f(x) = \sum_{\mathbf{j}\in\mathcal{J}(d,n)} a_{\mathbf{j}} x_{\mathbf{j}}, \quad (2.5)$$

y tenemos además, como ya comentamos en (1.9)

$$\|f\|_\infty := \sup_{x\in\{\pm 1\}^n} |P_f(x)| = \sup_{x\in[-1,1]^n} |P_f(x)| =: \|P_f\|_\infty.$$

El polinomio real P_f tiene asociado una única forma d-lineal simétrica $L_f : (\mathbb{R}^n)^d \to \mathbb{R}$ satisfaciendo $P_f(x) = L_f(x, ..., x)$, además para todo $y^{(1)}, ..., y^{(d)} \in \mathbb{R}^n$ podemos escribir explícitamente que

$$L_f(y^1, ..., y^d) = \sum_{\mathbf{i}\in\mathcal{I}(d,n)} c_{\mathbf{i}} y^1_{i_1} ... y^d_{i_d}, \qquad c_{\mathbf{i}} = \frac{a_{\mathbf{j}}}{d!}, \qquad \mathbf{i} \in [\mathbf{j}]. \quad (2.6)$$

2.1.2. Hipercontractividad

Otro resultado relevante para la demostración es aquel que permite comparar distintas normas. Para cada $1 \leqslant p < \infty$ la L_p-norma de la función $f : \{\pm 1\}^n \to \mathbb{R}$ está dada por

$$\|f\|_p = \mathbb{E}[|f|^p]^{\frac{1}{p}} = \left(\sum_{x\in\{\pm 1\}^n} \frac{1}{2^n} |\, f(x)|^p\right)^{\frac{1}{p}}.$$

Consideramos el siguiente operador conocido como el operador de ruido, definido para cada $-1 < \rho < 1$ como la aplicación T_ρ tal que

$$T_\rho f(x) := \sum_{S\subset[n]} \hat{f}(S)\rho^{|S|} x^S, \qquad x \in \{\pm 1\}^n.$$

Una característica importante de este operador es presentada por el famoso resultado de Bonami y Gross, para el cual demostraremos algunos casos particulares en el Capítulo 4.

Teorema 2.1.2. *Sean* $1 < p \leqslant q \leqslant \infty$ *y* $\rho \leqslant \sqrt{\frac{p-1}{q-1}}$. *Se tiene:*

$$\|T_\rho f\|_q \leqslant \|f\|_p.$$

También tenemos ciertas consecuencias que serán útiles en lo que sigue, como por ejemplo:

Corolario 2.1.3. *Para toda función* $f : \{\pm 1\}^n \to \mathbb{R}$ *de grado* d

$$\|f\|_2 \leqslant \left(\frac{1}{\sqrt{p-1}}\right)^d \|f\|_p, \tag{2.7}$$

y además

$$\|f\|_2 \leqslant e^d \|f\|_1. \tag{2.8}$$

2.1.3. Polarización 1

Usaremos la siguiente fórmula de polarización, inspirada en el resultado de Harris que es el análogo para el caso complejo [Har72, Teorema 1].

Proposición 2.1.4. *Sean* $P : \mathbb{R}^n \to \mathbb{R}$ *un polinomio* d*-homogéneo, y* $L : (\mathbb{R}^n)^d \to \mathbb{R}$ *la única forma* d*-lineal simétrica asociada. Entonces para todo* $1 \leqslant k \leqslant d/2$ *y* $x, y \in [-1,1]^n$

$$|L(\overbrace{x, ..., x}^{k}, \overbrace{y, ..., y}^{d-k})| \leqslant 2d^k \|P\|_{[-1,1]^n}.$$

En la siguiente sección probaremos un resultado de polarización más general y este se deducirá del mismo.

2.1.4. Demostración del caso homogéneo

Estamos en condiciones de dar la demostración del Teorema 2.0.1.

Demostración. Sea $f : \{\pm 1\}^n \to \mathbb{R}$ una función d-homogénea con $d \geqslant 2$, ya que el caso $d = 1$ es claro. Reescribimos a f de la siguiente forma

$$f(x) = \sum_{\mathbf{j} \in \mathcal{J}(d,n)} a_{\mathbf{j}} x_{\mathbf{j}},$$

donde $a_{\mathbf{j}} = 0$ cuando $\mathbf{j}$ no crece estrictamente, y extendemos la definición de los $a_{\mathbf{j}}$ a todo $\mathcal{I}(d,n)$ por $a_{\mathbf{i}} = a_{\mathbf{j}}$ si $\mathbf{i} \in [\mathbf{j}]$. Más aún, para todos los coeficientes $c_{\mathbf{i}}$ de la única forma simétrica L_f, tenemos $c_{\mathbf{i}} = \frac{a_{\mathbf{i}}}{d!}$. Entonces por la desigualdad de Blei (2.3) ($a_{\mathbf{i} \oplus \mathbf{j}} = 0$ siempre que $\mathbf{i}$ o $\mathbf{j}$ no crezcan estrictamente) tenemos para todo $k \leqslant d$

$$\left(\sum_{\mathbf{i}\in\mathcal{J}(d,n)} |\,a_{\mathbf{i}}|^{\frac{2d}{d+1}}\right)^{\frac{d+1}{2d}} \leqslant \left[\prod_{S\subset[d],|S|=k}\left(\sum_{\mathbf{i}\in\mathcal{J}(S,n)}\left(\sum_{\mathbf{j}\in\mathcal{J}(\hat{S},n)} |\,a_{\mathbf{i}\oplus\mathbf{j}}|^2\right)^{\frac{1}{2}\frac{2k}{k+1}}\right)^{\frac{k+1}{2k}}\right]^{\frac{1}{\binom{d}{k}}}.$$

Fijamos $S\subset[d]$ con $|S|=k$ para estimar cada uno de estos factores. Si denotamos $\rho_r:=(r-1)^{-\frac{1}{2}}$, usando (2.2) y (2.7), tenemos

$$\begin{aligned}
\left(\sum_{\mathbf{i}\in\mathcal{J}(S,n)}\left(\sum_{\mathbf{j}\in\mathcal{J}(\hat{S},n)} |\,a_{\mathbf{i}\oplus\mathbf{j}}|^2\right)^{\frac{1}{2}\frac{2k}{k+1}}\right)^{\frac{k+1}{2k}} &\leqslant \rho_{\frac{2k}{k+1}}^{d-k}\left(\sum_{\mathbf{i}\in\mathcal{J}(S,n)}\mathbb{E}_y\left[|\sum_{\mathbf{j}\in\mathcal{J}(\hat{S},n)} a_{\mathbf{i}\oplus\mathbf{j}}y_{\mathbf{j}}|^{\frac{2k}{k+1}}\right]\right)^{\frac{k+1}{2k}}\\
&= \rho_{\frac{2k}{k+1}}^{d-k}\left(\mathbb{E}_y\left[\sum_{\mathbf{i}\in\mathcal{J}(S,n)}|\sum_{\mathbf{j}\in\mathcal{J}(\hat{S},n)} a_{\mathbf{i}\oplus\mathbf{j}}y_{\mathbf{j}}|^{\frac{2k}{k+1}}\right]\right)^{\frac{k+1}{2k}}\\
&= \rho_{\frac{2k}{k+1}}^{d-k}\frac{d!}{(d-k)!}\left(\mathbb{E}_y\left[\sum_{\mathbf{i}\in\mathcal{J}(S,n)}|\sum_{\mathbf{j}\in\mathcal{I}(\hat{S},n)} c_{\mathbf{i}\oplus\mathbf{j}}y_{\mathbf{j}}|^{\frac{2k}{k+1}}\right]\right)^{\frac{k+1}{2k}}\\
&\leqslant \rho_{\frac{2k}{k+1}}^{d-k}\frac{d!}{(d-k)!}\sup_{y\in\{\pm1\}^n}\left(\sum_{\mathbf{i}\in\mathcal{J}(S,n)}|\sum_{\mathbf{j}\in\mathcal{I}(\hat{S},n)} c_{\mathbf{i}\oplus\mathbf{j}}y_{\mathbf{j}}|^{\frac{2k}{k+1}}\right)^{\frac{k+1}{2k}}.
\end{aligned}$$

Por la definición de $BH_{\{\pm1\}}^{=k}$ y la Proposición 2.1.4, tenemos que para todo $y\in\ \{\pm1\}^n$

$$\begin{aligned}
\left(\sum_{\mathbf{i}\in\mathcal{J}(S,n)}|\sum_{\mathbf{j}\in\mathcal{I}(\hat{S},n)} c_{\mathbf{i}\oplus\mathbf{j}}y_{\mathbf{j}}|^{\frac{2k}{k+1}}\right)^{\frac{k+1}{2k}} &\leqslant BH_{\{\pm1\}}^{=k}\sup_{x\in\{\pm1\}^n}|\sum_{\mathbf{i}\in\mathcal{J}(S,n)}\sum_{\mathbf{j}\in\mathcal{I}(\hat{S},n)} c_{\mathbf{i}\oplus\mathbf{j}}x_{\mathbf{i}}y_{\mathbf{j}}|\\
&= BH_{\{\pm1\}}^{=k}\frac{1}{k!}\sup_{x\in\{\pm1\}^n}|\sum_{\mathbf{i}\in\mathcal{I}(S,n)}\sum_{\mathbf{j}\in\mathcal{I}(\hat{S},n)} c_{\mathbf{i}\oplus\mathbf{j}}x_{\mathbf{i}}y_{\mathbf{j}}|\\
&\leqslant BH_{\{\pm1\}}^{=k}\frac{1}{k!}\sup_{x,y\in\{\pm1\}^n}|L_f(\overbrace{x,...,x}^{k},\overbrace{y,...,y}^{d-k})|\\
&\leqslant BH_{\{\pm1\}}^{=k}\frac{1}{k!}2d^k\|f\|_\infty,
\end{aligned}$$

por lo tanto

$$\left(\sum_{\mathbf{i}\in\mathcal{J}(S,n)}\left(\sum_{\mathbf{j}\in\mathcal{J}(\hat{S},n)} |\,a_{\mathbf{i}\oplus\mathbf{j}}|^2\right)^{\frac{1}{2}\frac{2k}{k+1}}\right)^{\frac{k+1}{2k}} \leqslant BH_{\{\pm1\}}^{=k}\rho_{\frac{2k}{k+1}}^{d-k}2d^k\binom{d}{k}\|f\|_\infty,$$

lo que lleva a

$$BH^{=d}_{\{\pm1\}} \leqslant BH^{=k}_{\{\pm1\}}\rho^{d-k}_{\frac{2k}{k+1}}2d^k\binom{d}{k}.$$

Reemplazando el valor de $\rho_{\frac{2k}{k+1}}$ tenemos

$$\begin{aligned}
BH^{=d}_{\{\pm1\}} &\leqslant BH^{=k}_{\{\pm1\}}\left(\frac{k+1}{k-1}\right)^{\frac{d-k}{2}}\binom{d}{k}2d^k \\
&= BH^{=k}_{\{\pm1\}}\left(1+\frac{2}{k-1}\right)^{\frac{d-k}{2}}d^{2k} \\
&\leqslant BH^{=k}_{\{\pm1\}}\exp\left(\frac{d-k}{k-1}+2k\log d\right) \\
&\leqslant BH^{=k}_{\{\pm1\}}\exp\left(2\left(\frac{d}{k+1}+k\log d\right)\right),
\end{aligned}$$

donde usamos que $(1+\frac{1}{x})^x$ es creciente y acotado por e para todo x positivo. Tomando $k \approx \sqrt{\frac{d}{logd}}$ y usando que $BH^{=k}_{\{\pm1\}}$ es no decreciente en k, obtenemos para algún $\alpha > 0$

$$BH^{=d}_{\{\pm1\}} \leqslant BH^{=[\sqrt{d}]}_{\{\pm1\}}\exp\left(\alpha\sqrt{d\log d}\right).$$

Iterando el proceso tenemos

$$\begin{aligned}
BH^{=d}_{\{\pm1\}} &\leqslant e^{\left(\alpha\sqrt{d\log d}\right)}e^{\left(\alpha[\sqrt[4]{d}]\frac{1}{\sqrt{2}}\sqrt{\log d}\right)}BH^{=[\sqrt[4]{d}]}_{\{\pm1\}} \\
&\leqslant e^{\left(\alpha\sqrt{d}\sqrt{\log d}\left(1+\frac{1}{\sqrt{2}}+\frac{1}{(\sqrt{2})^2}+\dots\right)\right)}BH^{=2}_{\{\pm1\}}.
\end{aligned}$$

Finalmente $BH^{=2}_{\{\pm1\}} < \infty$, esto es una consecuencia de la propiedad de isometría (1.9). Para probarlo notemos que $BH^{=2}_{\{\pm1\}} \leqslant BH^{=2}_{[-1,1]}$, ya que esta última constante por definición cumple la desigualdad de Bohnenblust-Hille para polinomios que son extensiones al cubo $[-1,1]^n$ de funciones Booleanas (los que con anterioridad llamamos P_f)

$$\left(\sum_{\alpha\in\mathbb{N}_0^n,|\alpha|=2} |\ a_\alpha|^{\frac{4}{3}}\right)^{\frac{3}{4}} \leqslant BH^{=2}_{[-1,1]}\|P_f\|_{[-1,1]^n},$$

intercambiando la norma de P_f con la norma f (por isometría) tenemos la afirmación. Por último, $BH^{=2}_{[-1,1]}$ es finito por (1.11) y así $BH^{=2}_{\{\pm1\}}$ lo es . □

2.2. El caso de grado *d* (no homogéneo)

Previamente probamos la desigualdad para el caso d-homogéneo, queremos pasar al caso de funciones de grado d, más exactamente.

Teorema 2.2.1. *Existe una constante absoluta $C > 0$ tal que para todo $d \in \mathbb{N}$*

$$BH_{\{\pm 1\}}^{\leqslant d} \leqslant C^{\sqrt{d \log d}}. \tag{2.9}$$

En el caso complejo, sabemos que $BH_{\mathbb{T}}^{=d} = BH_{\mathbb{T}}^{\leqslant d}$. En efecto, para un polinomio de grado d, $P(z) = \sum_{|\alpha| \leqslant d} a_\alpha z^\alpha$ en $\mathbb{T}^n$, consideramos el polinomio $Q(z,w) = \sum_{|\alpha| \leqslant d} a_\alpha z^\alpha w^{d-|\alpha|}$ que es homogéneo en $\mathbb{T}^{n+1}$ con los mismos coeficientes y la misma norma supremo. Veamos en detalle esto último

$$\|Q\|_\infty = \sup_{w \in \mathbb{T}} \sup_{z \in \mathbb{T}^n} \left| \sum_{|\alpha| \leqslant d} a_\alpha z^\alpha w^{d-|\alpha|} \right| = \sup_{w \in \mathbb{T}} \sup_{z \in \mathbb{T}^n} \left| \sum_{|\alpha| \leqslant d} a_\alpha \left(\frac{z}{w}\right)^\alpha w^d \right| \tag{2.10}$$

$$= \sup_{w \in \mathbb{T}} \sup_{z \in \mathbb{T}^n} \left| \left(\sum_{|\alpha| \leqslant d} a_\alpha z^\alpha \right) w^d \right| = \sup_{z \in \mathbb{T}^n} |P(z)| = \|P\|_\infty. \tag{2.11}$$

Donde usamos que si $w \in \mathbb{T}$, $\frac{z}{w} \in \mathbb{T}^n$ para todo $z \in \mathbb{T}^n$. Pero esto no ocurre en el caso real. Otro camino (que también funciona en el caso complejo) es dividir la expansión Fourier de la función de grado d en las partes m-homogéneas f_m

$$\left(\sum_{|S| \leqslant d} |\hat{f}(S)|^{\frac{2d}{d+1}} \right)^{\frac{d+1}{2d}} \leqslant C^{\sqrt{d \log d}} \sum_{m=0}^{d} \|f_m\|_\infty,$$

y controlar la norma de f_m por la norma de la función. Podemos dar esta estimación por el Lema 1.0.9 (4) válida para todo polinomio real, pero esta constante es $(1+\sqrt{2})^n$ que tiene un crecimiento exponencial en el grado (en contraste con el caso complejo que se obtiene constante 1 ver Lema 1.0.9 (1)). Podemos preguntarnos si esta constante se puede mejorar en el caso tetraedral.

Considerando la función mayoría $Maj_d : \{\pm 1\}^d \to \{\pm 1\}$, $Maj_d(x) = sign(x_1 + ... + x_d)$ se puede chequear que la norma de la parte m-homogénea (m impar) satisface

$$\|(Maj_d)_m\|_{\{\pm 1\}^d} = \binom{\frac{d-1}{2}}{\frac{m-1}{2}} \frac{d}{m} \frac{1}{2^{d-1}} \binom{d-1}{\frac{d-1}{2}} \geqslant \binom{\frac{d-1}{2}}{\frac{m-1}{2}} \frac{\sqrt{d}}{m},$$

tomando $m = \frac{d-1}{2}$ tenemos $\|(Maj_d)_m\|_{\{\pm 1\}} = \Omega(\frac{2^{\frac{d}{2}}}{\sqrt{d}})$ (veáse [O'D14, p.121]), es decir la estimación $(1+\sqrt{2})^d$ no es óptima para el caso tetraedral ya que la constante asintóticamente es menor que $\sqrt{2}$. Todos estos problemas obligan a agregar a la estrategia clásica descripta en la sección anterior, algunas técnicas aparentementes nuevas para demostrar que en caso tetraedral podemos todavía obtener constantes $BH_{\{\pm 1\}}^{d}$ que sean subexponenciales como en el caso complejo. A diferencia del Teorema 1.0.8 la clave es que ahora el análisis del cubo Booleano $\{\pm 1\}^n$ está a nuestra disposición, como en el caso complejo tenemos el politoro $\mathbb{T}^n$.

2.2.1. Formas multiafines.

Sea $Q : \mathbb{R}^n \to \mathbb{R}$ un polinomio de grado d. Sabemos que Q admite una única representación conocida como la expresión monomial. Existe una única familia de escalares $a_{j_1 \dots j_m}$ tal que para cada $x \in \mathbb{R}^n$

$$Q(x) = \sum_{m=0}^{d} \sum_{1 \leqslant j_1 \leqslant \dots \leqslant j_m \leqslant n} a_{j_1 \dots j_m} x_{j_1} x_{j_2} \dots x_{j_m}. \tag{2.12}$$

Para cada $0 \leqslant m \leqslant d$, la parte m-homogénea de Q es

$$Q_m(x) = \sum_{1 \leqslant j_1 \leqslant \dots \leqslant j_m \leqslant n} a_{j_1 \dots j_m} x_{j_1} x_{j_2} \dots x_{j_m}, \qquad x \in \mathbb{R}^n.$$

Decimos que Q es un polinomio tetraedral si tiene la siguiente forma

$$Q(x) = \sum_{m=0}^{d} \sum_{1 \leqslant j_1 < \dots < j_m \leqslant n} a_{j_1 \dots j_m} x_{j_1} x_{j_2} \dots x_{j_m}, \qquad x \in \mathbb{R}^n. \tag{2.13}$$

En otras palabras, los polinomios tetraedrales son caracterizados por ser afines en cada variable $x_j \in \mathbb{R}$. Esto es equivalente a decir que $Q(x)$ preserva combinaciones convexas en cada variable, i.e., si $\mathbf{x}_1, \dots, \mathbf{x}_n$ son variables aleatorias independientes (reales), entonces

$$\mathbb{E}(Q(\mathbf{x}_1, \dots, \mathbf{x}_n)) = Q(\mathbb{E}[\mathbf{x}_1], \dots, \mathbb{E}[\mathbf{x}_n]). \tag{2.14}$$

Como ya mencionamos, lo importante de estos polinomios es que tenemos una correspondencia isométrica natural entre estos y funciones reales en el cubo Booleano. En efecto, si tomamos una función $f : \{\pm 1\}^n \to \mathbb{R}$ de grado d tenemos

$$a_{j_1 \dots j_m} := \hat{f}(S) \text{ donde } S = \{j_1 < \dots < j_m\}, \qquad 0 \leqslant m \leqslant d, \tag{2.15}$$

entonces la expansión Fourier-Walsh de f la podemos reescribir

$$f(x) = \sum_{m=0}^{d} \sum_{1 \leqslant j_1 < \dots < j_m \leqslant n} a_{j_1 \dots j_m} x_{j_1} x_{j_2} \dots x_{j_m}, \qquad x \in \{\pm 1\}^n.$$

Denotamos por Q_f el único polinomio tetraedral que coincide con f en $\{\pm 1\}^n$, con sus respectivos monomios y coeficientes de Fourier relativos.

Siguiendo [Kwa87, p. 1065-66], todo $Q : \mathbb{R}^n \to \mathbb{R}$ de grado d tiene una única d-forma $L_Q : (\mathbb{R}^n)^d \to \mathbb{R}$ que satisface las siguientes propiedades:

I. L_Q es d-afín, i.e. es una función afín en cada variable $x^{(j)} \in \mathbb{R}^n$.

II. L_Q es simétrico, i.e, la evaluación de $L_Q(x^{(1)}, \dots, x^{(d)})$ es invariante por permutaciones en las coordenadas $x^{(1)}, \dots, x^{(d)}$.

III. $L_Q(x, \dots, x) = Q(x)$ para todo $x \in \mathbb{R}^n$.

Para cada $x^{(1)}, ..., x^{(d)} \in \mathbb{R}$ podemos escribir explícitamente vía fórmula de polarización

$$L_Q(x^{(1)}, ..., x^{(d)}) = \mathbb{E}_\xi \left[\sum_{m=0}^{d} \frac{\xi_1 ... \xi_d}{d!} (\xi_1 + ... + \xi_n)^{d-m} Q_m(\xi_1 x^{(1)} + ... + \xi_d x^{(d)}) \right],$$

donde consideramos la esperanza sobre $\xi \in \{\pm 1\}^n$. Además L_Q cumple las siguientes observaciones:

I. Si Q es d-homogéneo, entonces L_Q es la usual d-forma lineal asociada a Q.

II. Para $f : \{\pm 1\}^n \to \mathbb{R}$ denotamos por L_f a la d- forma asociada al polinomio tetraedral Q_f.

2.2.2. Polarización 2

El siguiente resultado será crucial. Básicamente es un análogo de la fórmula de polarización de Harris para polinomios complejos homogéneos y su variante real, la Proposición 2.1.4. Nuestro caso es mucho más limitado, ya que para un polinomio Q de grado d su forma multiafín asociada L_Q en cada variables conserva combinaciones convexas y no lineales.

Proposición 2.2.2. *Sea $Q : \mathbb{R}^n \to \mathbb{R}$ un polinomio de grado d y L_Q su d-forma afín asociada. Entonces, para todo $0 \leqslant m \leqslant \frac{d}{2}$ tenemos que*

$$\sup_{x,y \in [-1,1]^n} |L_Q(\overbrace{x, ..., x}^{m}, \overbrace{y, ..., y}^{d-m})| \leqslant 2d^m \sup_{x \in [-1,1]^n} |Q(x)|. \tag{2.16}$$

Antes de comenzar con la demostración, daremos un resultado auxiliar de interpolación, como hicimos antes con el clásico Teorema de Markov 1.15. Veremos otra propiedad de los coeficientes de los polinomios de Chebyshev. En contraste con el resultado de Markov, consideramos una base diferente en el espacio de polinomios de grado d. Para cada $m \leqslant d$, tenemos

$$\psi_{d,m}(t) = 2^{-d}(1+t)^m(1-t)^{d-m} = \left(\frac{1+t}{2}\right)^m \left(\frac{1-t}{2}\right)^{d-m}.$$

Proposición 2.2.3. *Sea $Q(t)$ polinomio de grado d. Entonces existen únicos escalares $a_n = a_n(Q)$ $(0 \leqslant n \leqslant d)$ tal que*

$$Q(t) = \sum_{n=0}^{d} a_n \psi_{d,n}(t), \qquad t \in \mathbb{R}.$$

Más aún, para todo $0 \leqslant n \leqslant d$:

$$|a_n(Q)| \leqslant |a_n(T_d)| \sup_{t \in [-1,1]^n} |Q(t)|, \tag{2.17}$$

donde T_d es el polinomio de Chebyshev de grado d. En particular, tenemos la estimación óptima

$$|a_n(Q)| \leqslant \sum_{m=0}^{min\{d-n,n\}} 4^m \binom{d}{2m} \binom{d-2m}{n-m} \sup_{t\in[-1,1]^n} |Q(t)|. \tag{2.18}$$

Demostración. Comenzaremos probando que todo polinomio $Q(t)$ de grado d, es una combinación de los elementos $\psi_{d,m}$ $(0 \leqslant m \leqslant d)$ donde la unicidad se desprende del álgebra lineal. En el proceso tendremos una expresión de los coeficientes, con eso probaremos la desigualdad (2.17) y finalmente calcularemos explícitamente los coeficientes $a_n(T_d)$ donde concluiremos que por (2.17) vale (2.18).

Comenzamos tomando $d+1$ puntos arbitrarios y fijos (a pesar que luego los elegiremos cuidadosamente)

$$-1 \leqslant t_0 < t_1 < ... < t_d \leqslant 1.$$

Podemos comprobar fácilmente que tenemos asociada, lo que se conoce como la base polinomial de Lagrange, que viene dada por

$$\Delta_m(t) = \prod_{j=0, j\neq m}^{d} \frac{t-t_j}{t_m - t_j}, \qquad 0 \leqslant m \leqslant d, \tag{2.19}$$

y satisface que

$$Q(t) = \sum_{m=0}^{d} Q(t_m)\Delta_m(t), \qquad t \in [-1,1], \tag{2.20}$$

ya que si restamos los polinomios, conseguimos otro de grado a lo sumo d con $d+1$ raíces distintas (las cuales son t_j con $0 \leqslant j \leqslant d+1$), por lo tanto dicha resta debe ser el polinomio nulo.

Ahora escribimos el polinomio $\Delta_m(t)$ en términos de $\psi_{d,n}(t)$. Notar que

$$\prod_{j=0, j\neq m}^{d} (t - t_j) = \prod_{j=0, j\neq m}^{d} \left(\left(\frac{1+t}{2}\right)(1-t_j) - \left(\frac{1-t}{2}\right)(1+t_j) \right).$$

Expandiendo el producto y agrupando términos, tenemos la siguiente expresión

$$\prod_{j=0, j\neq m}^{d} (t - t_j) = \sum_{n=0}^{d} \alpha_{m,n}(-1)^{d-n} \left(\frac{1+t}{2}\right)^n \left(\frac{1-t}{2}\right)^{d-n}, \tag{2.21}$$

donde los coeficientes $\alpha_{m,n}$ son

$$\alpha_{m,n} = \sum_{S\subset[d]_m, |S|=n} \prod_{j\in S}(1+t_j) \prod_{j\in[d]_m\setminus S} (1-t_j), \tag{2.22}$$

con $[d]_m = \{1, ..., d\}\setminus\{m\}$. Reemplazando (2.21) en (2.19), junto a la expresión (2.20) concluimos que

$$Q(t) = \sum_{n=0}^{d} \left(\sum_{m=0}^{d} Q(t_m) \frac{\alpha_{m,n}(-1)^{d-n}}{\prod_{j\neq m}(t_m - t_j)} \right) \psi_{d,n}(t), \qquad t \in [-1,1]. \tag{2.23}$$

Esto muestra que los coeficientes a_n existen y están dados explícitamente por

$$a_n = a_n(Q) = \sum_{m=0}^{d} Q(t_m) \frac{\alpha_{m,n}(-1)^{d-n}}{\prod_{j\neq m}(t_m - t_j)}, \qquad 0 \leqslant n \leqslant d, \tag{2.24}$$

consiguiendo así la primer parte de la proposición.

Ahora mostremos (2.17). Notar primero que los $t_j's$ se ordenan acorde a la partición tomada antes, donde

$$\prod_{j\neq m}(t_m - t_j) = (-1)^{d-m} \prod_{j\neq m} |t_m - t_j|,$$

reescribimos (2.24) como

$$a_n = a_n(Q) = (-1)^n \sum_{m=0}^{d} Q(t_m) \frac{\alpha_{m,n}(-1)^m}{\prod_{j\neq m}|t_m - t_j|}.$$

Otra observación útil que se desprende de (2.22), es que $\alpha_{m,n} \geqslant 0$ para todo m, n. Sin pérdida de generalidad podemos asumir que $\sup_{t\in[-1,1]} |Q(t)| = 1$ normalizando el polinomio y también podemos elegir los siguientes puntos de interpolación

$$t_m = \cos\left(\frac{m\pi}{d}\right), \qquad m = 0, ..., d,$$

que satisfacen la condición de la partición. Recordando que el polinomio de Chebyshev T_d de grado d, satisface que $T_d(t_m) = (-1)^m$ para todo $0 \leqslant m \leqslant d$, pues cumple la propiedad $T_n(\cos(\theta)) = \cos(n\theta)$. Podemos plantear lo siguiente

$$|a_m(Q)| \leqslant \sum_{m=0}^{d} \frac{\alpha_{m,n}}{\prod_{j\neq m}|t_m - t_j|} = \sum_{m=0}^{d} \frac{T_d(t_m)\alpha_{m,n}(-1)^m}{\prod_{j\neq m}|t_m - t_j|} = (-1)^n a_m(T_d),$$

esto implica (2.17). Finalmente, calculemos los coeficiente $a_n(T_d)$ usando la expresión del polinomio de Chebyshev:

$$\begin{aligned}
T_d(t) &= \sum_{m=0}^{[\frac{d}{2}]} \binom{d}{2m}(t^2-1)^m t^{d-2m}\\
&= \sum_{m=0}^{[\frac{d}{2}]} \binom{d}{2m}(t-1)^m(t+1)^m \left(\frac{t-1}{2}+\frac{t+1}{2}\right)^{d-2m}\\
&= \sum_{m=0}^{[\frac{d}{2}]} \binom{d}{2m}(t-1)^m(t+1)^m \sum_{k=0}^{d-2m}\binom{d-2m}{k}\left(\frac{t-1}{2}\right)^k\left(\frac{t+1}{2}\right)^{d-2m-k}\\
&= \sum_{m=0}^{[\frac{d}{2}]}\sum_{k=0}^{d-2m}\binom{d}{2m}\binom{d-2m}{k}2^{2m}\left(\frac{t-1}{2}\right)^{k+m}\left(\frac{t+1}{2}\right)^{d-m-k}\\
&= \sum_{n=0}^{d}\left(\sum_{m=0}^{min\{d-n,n\}} 4^m\binom{d}{2m}\binom{d-2m}{n-m}\right)\left(\frac{t-1}{2}\right)^n\left(\frac{t+1}{2}\right)^{d-n}.
\end{aligned}$$

Por lo tanto

$$a_n(T_d) = (-1)^n \sum_{m=0}^{min\{d-n,n\}} 4^m\binom{d}{2m}\binom{d-2m}{n-m}, \tag{2.25}$$

y así termina de demostración. □

Lo siguiente será demostrar la Proposición 2.2.2.

Demostración. Fijemos $x, y \in \{\pm 1\}^n$. Notar que como L_Q es afín en cada variable, para todo $t \in [-1,1]$ tenemos que

$$P(t) = Q\left(\frac{1+t}{2}x+\frac{1-t}{2}y\right) = \sum_{m=0}^{d}\binom{d}{m}\left(\frac{1+t}{2}\right)^m\left(\frac{1-t}{2}\right)^{d-m} L_Q(\overbrace{x,...,x}^{m},\overbrace{y,...,y}^{d-m}). \tag{2.26}$$

Más aún, $P(t)$ es la evaluación de Q en una combinación convexa de elementos en $[-1,1]^n$, así

$$\sup_{t\in[-1,1]} |P(t)| \leqslant \|Q\|_\infty.$$

Sin perdidad de generalidad podemos asumir que $\|Q\|_\infty = 1$. Aplicando (2.18) en (2.26) tenemos que, si $0 \leqslant m \leqslant \frac{d}{2}$, entonces conseguimos la acotación

$$|L_Q(\overbrace{x,...,x}^{m},\overbrace{y,...y}^{d-m})| \leqslant \frac{1}{\binom{d}{m}}\sum_{k=0}^{m}4^k\binom{d}{2k}\binom{d-2k}{m-k}.$$

Usando la fórmula de aproximación de Stirling

$$(2k)! = \binom{2k}{k}(k!)^2 \geqslant \frac{4^k}{2\sqrt{k}}(k!)^2, \quad k \geqslant 1.$$

Entonces

$$\begin{aligned}\frac{1}{\binom{d}{m}}\sum_{k=0}^{m}4^k\binom{d}{2k}\binom{d-2k}{m-k} &= \sum_{k=0}^{m}4^k\frac{m!(d-m)!}{(2k)!(m-k)!(d-m-k)!}\\ &\leqslant 1+\sum_{k=1}^{m}2\sqrt{k}\binom{m}{k}\binom{d-m}{k}\\ &\leqslant 2\sum_{k=0}^{m}\binom{m}{k}(d-1)^k = 2d^m,\end{aligned}$$

esto completa la demostración. $\square$

2.2.3. Demostración del caso de grado d

Para la demostración del caso general 0.0.2 seguimos la estrategia como en el caso homogéneo, pero antes debemos introducir algunas nociones más.

Para cada $n \in \mathbb{N}$ y cada subconjunto finito $A \subset \mathbb{N}$ definimos los conjuntos

$$\begin{aligned}\mathcal{I}_0(A,n) &:= \{\mathbf{i} : A \to [n] \cup \{0\}\}\\ \mathcal{J}_0(A,n) &:= \{\mathbf{j} : A \to [n] \cup \{0\} \text{ no decreciente}\}.\end{aligned}$$

En el caso $A = [d]$ para $d \in \mathbb{N}$, denotamos los conjuntos $\mathcal{I}_0(d,n)$ y $\mathcal{J}_0(d,n)$ respectivamente. Tomamos $\mathbf{i} \in \mathcal{J}_0(A,n)$, y escribimos $supp(\mathbf{i}) = \{k : i(k) \neq 0\}$ que llamaremos el soporte de $\mathbf{i}$. Consideramos la relación de equivalencia definida para $\mathbf{i}_1, \mathbf{i}_2 \in \mathcal{I}_0(A,n)$ como $\mathbf{i}_1 \sim \mathbf{i}_2$ si y sólo si existe una biyección $\sigma : A \to A$ tal que $\mathbf{i}_1 = \mathbf{i}_2 \circ \sigma$.

Notar que para cada $\mathbf{i} \in \mathcal{I}_0(A,n)$ tenemos un único $\mathbf{j} \in \mathcal{J}_0(A,n)$ tal que $\mathbf{j} \in [\mathbf{i}]$ donde $[\mathbf{i}]$ es la clase de equivalencia de $\mathbf{i}$. Con esta identificación reescribimos la expansión monomial (2.12) de un polinomio de grado d $Q : \mathbb{R}^n \to \mathbb{R}$ como

$$Q(x) = \sum_{\mathbf{j}\in\mathcal{J}_0(d,n)} a_{\mathbf{j}} \prod_{k\in supp(\mathbf{j})} x_{\mathbf{j}(k)}, \qquad x \in \mathbb{R}^n. \tag{2.27}$$

Recordar que el polinomio tetraedral es afín en cada variable, usando la notación de (2.27) vemos que los coeficientes distintos de cero $a_{\mathbf{j}}$ deben satisfacer que $\mathbf{j}$ es inyectivo en su soporte.

Decimos que un elemento $\mathbf{i} \in \mathcal{I}_0(A,n)$ es afín si $\mathbf{i}(k) \neq \mathbf{i}(\hat{k})$ donde $k, \hat{k} \in supp(\mathbf{i})$ son diferentes; notar que todo elemento de $[\mathbf{i}]$ también es afín. Como consecuencia el polinomio tetraedral tiene una expansión monomial de la forma

$$Q(x) = \sum_{\mathbf{j}\in\mathcal{J}_0(d,n) \text{ afín}} a_{\mathbf{j}} \prod_{k\in supp(\mathbf{j})} x_{\mathbf{j}(k)}, \qquad x \in \mathbb{R}^n. \tag{2.28}$$

De otra manera, la asociación de la forma d-afín de un polinomio Q con la expansión monomial (2.27) está dada por

$$L_Q(x^{(1)}, ..., x^{(d)}) = \sum_{\mathbf{i}\in\mathcal{I}_0(A,n)} c_{\mathbf{i}} \prod_{k\in supp(\mathbf{i})} x^{(k)}_{\mathbf{i}(k)}, \tag{2.29}$$

donde $c_{\mathbf{i}} = \frac{a_{\mathbf{j}}}{|[\mathbf{j}]|}$ si $\mathbf{i} \in [\mathbf{j}]$. De nuevo, usaremos la cardinalidad de $[\mathbf{i}]$, $\mathbf{i} \in \mathcal{I}_0(A, n)$ que está dada por

$$|[\mathbf{i}]| = \frac{|A|!}{|\{k : \mathbf{i}(k) = 0\}|!|\{k : \mathbf{i}(k) = 1\}|!...|\{k : \mathbf{i}(k) = n\}|!}. \tag{2.30}$$

Más aún, si A_1 y A_2 son dos subconjuntos disjuntos de $\mathbb{N}$, entonces la suma directa de $\mathbf{i}_1 \in \mathcal{I}_0(A_1, n)$ y $\mathbf{i}_2 \in \mathcal{I}_0(A_2, n)$ está definida con la siguientes función

$$\mathcal{I}_0(A_1 \cup A_2, n) \ni \mathbf{i}_1 \oplus \mathbf{i}_2 : A_1 \cup A_2 \to [n] \cup \{0\}.$$

En particular, $\mathcal{I}_0(d, n) = \mathcal{I}_0(S, n) \oplus \mathcal{I}_0(\hat{S}, n)$ para $S \subset [d]$ y $\hat{S} := [d] \backslash S$.

En lo que sigue demostraremos el resultado principal del capítulo, es decir el Teorema 0.0.2.

Demostración. Fijo $0 \leqslant m \leqslant d$. Sea $f : \{\pm 1\}^n \to \mathbb{R}$ una función de grado d y Q_f el polinomio teatraedral asociado. Sabemos por (2.28) que tiene una expansión monomial

$$Q_f(x) = \sum_{\mathbf{j} \in \mathcal{J}(d,n)} a_{\mathbf{j}} \prod_{k \in supp(\mathbf{j})} x_{\mathbf{j}}(k),$$

donde $a_{\mathbf{j}} = 0$ cuando $\mathbf{j}$ no es afín, extendemos esta definición poniendo $a_{\mathbf{i}} = a_{\mathbf{j}}$ si $\mathbf{i} \in [\mathbf{j}]$ para $\mathbf{i} \in \mathcal{I}_0(d, n)$, $\mathbf{j} \in \mathcal{J}_0(d, n)$. Sea $L_f : (\mathbb{R}^n)^d \to \mathbb{R}$ es la forma d-afín de Q_f descripta anteriormente. Para todo $S \subset [d]$ y $x^{(1)}, ..., x^{(d)} \in \mathbb{R}^n$

$$L_f(x^{(1)}, ..., x^{(d)}) = \sum_{\mathbf{i}_1 \in \mathcal{I}_0(S,n)} \sum_{\mathbf{i}_2 \in \mathcal{I}_0(\hat{S},n)} c_{\mathbf{i}_1 \oplus \mathbf{i}_2} \prod_{k \in supp(\mathbf{i}_1)} x^{(k)}_{\mathbf{i}_1(k)} \prod_{k \in supp(\mathbf{i}_2)} x^{(k)}_{\mathbf{i}_2(k)},$$

donde

$$c_{\mathbf{i}_1 \oplus \mathbf{i}_2} = \frac{a_{\mathbf{i}_1 \oplus \mathbf{i}_2}}{|[\mathbf{i}_1 \oplus \mathbf{i}_2]|}.$$

Para la igualdad se puede verificar que la parte derecha satisface las tres condiciones I, II, III (2.2.1) y recordar que L_f es la única forma que cumple las condiciones. Para $|S| = m$ y $x, y \in [-1, 1]^n$, evaluando el mapa L_f en los elementos $x^{(k)} = x$ si $k \in S$ y $x^{(k)} = y$ si $k \notin S$ y por la simetría de L_f tenemos que

$$\begin{aligned} L_f(\overbrace{x, ..., x}^{m}, \overbrace{y, ..., y}^{d-m}) &= \sum_{\mathbf{i}_1 \in \mathcal{I}_0(S,n)} \sum_{\mathbf{i}_2 \in \mathcal{I}_0(\hat{S},n)} c_{\mathbf{i}_1 \oplus \mathbf{i}_2} \prod_{k \in supp(\mathbf{i}_1)} x_{\mathbf{i}_1(k)} \prod_{k \in supp(\mathbf{i}_2)} y_{\mathbf{i}_2(k)} \\ &= \sum_{\mathbf{j}_1 \in \mathcal{J}_0(S,n)} \sum_{\mathbf{j}_2 \in \mathcal{J}_0(\hat{S},n)} |[\mathbf{j}_1]| \cdot |[\mathbf{j}_2]| \cdot c_{\mathbf{j}_1 \oplus \mathbf{j}_2} x_{\mathbf{j}_1} y_{\mathbf{j}_2} \\ &= \sum_{\mathbf{j}_1 \in \mathcal{J}_0(S,n)} \sum_{\mathbf{j}_2 \in \mathcal{J}_0(\hat{S},n)} A_{\mathbf{j}_1 \oplus \mathbf{j}_2} x_{\mathbf{j}_1} y_{\mathbf{j}_2}, \end{aligned}$$

donde

$$A_{\mathbf{j}_1 \oplus \mathbf{j}_2} = \frac{|[\mathbf{j}_1]| \cdot |[\mathbf{j}_2]|}{|[\mathbf{j_1} \oplus \mathbf{j_2}]|} a_{\mathbf{j}_1 \oplus \mathbf{j}_2}. \tag{2.31}$$

Notar que el coeficiente $a_{\mathbf{j}_1 \oplus \mathbf{j}_2}$ distinto de cero satisface que $\mathbf{j}_1 \oplus \mathbf{j}_2$ es afín, implica en particular que $\mathbf{j}_1, \mathbf{j}_2$ son afines y por (2.30) en este caso, podemos acotar

$$\frac{|[\mathbf{j}_1 \oplus \mathbf{j}_2]|}{|[\mathbf{j}_1]| \cdot |[\mathbf{j}_2]|} = \frac{d! \cdot |\{k : \mathbf{j}_1(k) = 0\}|! \cdot |\{k : \mathbf{j}_2(k) = 0\}|!}{(|\{k : (\mathbf{j}_1 \oplus \mathbf{j}_2)(k) = 0\}|)! \cdot m! \cdot (d-m)!} \leqslant \binom{d}{m}. \tag{2.32}$$

Usando la desigualdad de Blei (2.3), tenemos que

$$\left(\sum_{\mathbf{j} \in \mathcal{J}_0(d,n)} |\, a_{\mathbf{j}}|^{\frac{2d}{d+1}} \right)^{\frac{d+1}{2d}} \leqslant \left[\prod_{S \subset [d], |S|=k} \left(\sum_{\mathbf{j}_1 \in \mathcal{J}_0(S,n)} \left(\sum_{\mathbf{j}_2 \in \mathcal{J}_0(\hat{S},n)} |\, a_{\mathbf{j}_1 \oplus \mathbf{j}_2}|^2 \right)^{\frac{1}{2}\frac{2k}{k+1}} \right)^{\frac{k+1}{2k}} \right]^{\frac{1}{\binom{d}{k}}}. \tag{2.33}$$

Para $S \subset [d]$ fijo tal que $|S| = m$, estimaremos cada factor del lado derecho de la desigualdad anterior. Si denotamos $\rho_r = (r-1)^{\frac{-1}{2}}$ para $r > 1$, entonces usando (2.7) y (2.32), para cada $\mathbf{j_1} \in \mathcal{J}_0(S,n)$ tenemos

$$\left(\sum_{\mathbf{j}_2} |a_{\mathbf{j}_1 \oplus \mathbf{j}_2}|^2 \right)^{\frac{1}{2}} = \frac{|[\mathbf{j}_1 \oplus \mathbf{j}_2]|}{|[\mathbf{j}_1]| \cdot |[\mathbf{j}_2]|} \left(\sum_{\mathbf{j}_2} |A_{\mathbf{j}_1 \oplus \mathbf{j}_2}|^2 \right)^{\frac{1}{2}} \tag{2.34}$$

$$\leqslant \binom{d}{m} \left(\sum_{\mathbf{j}_2} |A_{\mathbf{j}_1 \oplus \mathbf{j}_2}|^2 \right)^{\frac{1}{2}} = \binom{d}{m} \left(\mathbb{E}_x \left| \sum_{\mathbf{j}_2} A_{\mathbf{j}_1 \oplus \, \mathbf{j}_2} x_{\mathbf{j}_2} \right|^2 \right)^{\frac{1}{2}} \tag{2.35}$$

$$\leqslant \binom{d}{m} \rho_{\frac{2m}{m+1}}^{d-m} \left(\mathbb{E}_x \left| \sum_{\mathbf{j}_2} A_{\mathbf{j}_1 \oplus \, \mathbf{j}_2} x_{\mathbf{j}} \right|^{\frac{2m}{m+1}} \right)^{\frac{m+1}{2m}}. \tag{2.36}$$

Sumando sobre todo $\mathbf{j}_1 \in \mathcal{J}_0(S,n)$ obtenemos

$$\begin{aligned}
\left(\sum_{\mathbf{j}_1} \left(\sum_{\mathbf{j}_2} |a_{\mathbf{j}_1 \oplus \mathbf{j}_2}|^2 \right)^{\frac{1}{2}\frac{2m}{m+1}} \right)^{\frac{m+1}{2m}} &\leqslant \binom{d}{m} \rho_{\frac{2m}{m+1}}^{d-m} \left(\sum_{\mathbf{j}_1} \mathbb{E}_x \left| \sum_{\mathbf{j}_2} A_{\mathbf{j}_1 \oplus \mathbf{j}_2} x_{\mathbf{j}_1} \right|^{\frac{2m}{m+1}} \right)^{\frac{m+1}{2m}} \\
&= \binom{d}{m} \rho_{\frac{2m}{m+1}}^{d-m} \left(\mathbb{E}_x \left[\sum_{\mathbf{j}_1} \left| \sum_{\mathbf{j}_2} A_{\mathbf{j}_1 \oplus \mathbf{j}_2} x_{\mathbf{j}_1} \right|^{\frac{2m}{m+1}} \right] \right)^{\frac{m+1}{2m}} \\
&\leqslant \binom{d}{m} \rho_{\frac{2m}{m+1}}^{d-m} \sup_{x \in \{\pm 1\}^n} \left(\sum_{\mathbf{j}_1} \left| \sum_{\mathbf{j}_2} A_{\mathbf{j}_1 \oplus \, \mathbf{j}_2} x_{\mathbf{j}_1} \right|^{\frac{2m}{m+1}} \right)^{\frac{m+1}{2m}}.
\end{aligned}$$

Y así el último supremo lo acotamos usando el segundo resultado de Polarización (la Proposición 2.2.2)

$$\begin{aligned}
\sup_{x\in\{\pm1\}^n}\left(\sum_{\mathbf{j}_1}\left|\sum_{\mathbf{j}_2} A_{\mathbf{j}_1\oplus\mathbf{j}_2}x_{\mathbf{j}_1}\right|^{\frac{2m}{m+1}}\right)^{\frac{m+1}{2m}} &\leqslant BH^{\leqslant m}_{\{\pm1\}}\sup_{x,y\in\{\pm1\}^n}\left|\sum_{\mathbf{j}_1}\sum_{\mathbf{j}_2}A_{\mathbf{j}_1\oplus\mathbf{j}_2}x_{\mathbf{j}_1}y_{\mathbf{j}_2}\right| \\
&= BH^{\leqslant m}_{\{\pm1\}}\sup_{x,y\in\{\pm1\}^n}|L_f(\overbrace{x,...x}^{m},\overbrace{y,...,y}^{d-m})| \\
&= BH^{\leqslant m}_{\{\pm1\}}2d^m\|f\|_\infty.
\end{aligned}$$

Por lo tanto

$$\left(\sum_{\mathbf{j}_1}\left(\sum_{\mathbf{j}_2}|\ a_{\mathbf{j}_1\oplus\mathbf{j}_2}|^2\right)^{\frac{1}{2}\frac{2m}{m+1}}\right)^{\frac{m+1}{2m}} \leqslant BH^{\leqslant m}_{\{\pm1\}}\binom{d}{m}\rho^{d-m}_{\frac{2m}{m+1}}2d^m\|f\|_\infty.$$

Como S es arbitrario, podemos aplicar esta cota para cada factor del lado derecho de (2.33) obteniendo

$$\left(\sum_{\mathbf{j}\in\mathcal{J}_0(d,n)}|\ a_{\mathbf{j}}|^{\frac{2d}{d+1}}\right)^{\frac{d+1}{2d}} \leqslant BH^{\leqslant m}_{\{\pm1\}}\binom{d}{m}\rho^{d-m}_{\frac{2m}{m+1}}2d^m\|f\|_\infty.$$

Consecuentemente

$$\begin{aligned}
BH^{\leqslant d}_{\{\pm1\}} &\leqslant BH^{\leqslant m}_{\{\pm1\}}\rho^{d-m}_{\frac{2m}{m+1}}\binom{d}{m}2d^m = BH^{\leqslant m}_{\{\pm1\}}\left(\frac{m+1}{m-1}\right)^{\frac{d-m}{2}}\binom{d}{m}2d^m \\
&\leqslant BH^{\leqslant m}_{\{\pm1\}}\left(1+\frac{2}{m-1}\right)^{\frac{d-m}{2}}d^{2m} \leqslant BH^{\leqslant m}_{\{\pm1\}}\exp\left(\frac{d-m}{m-1}+2m\log d\right) \\
&\leqslant BH^{\leqslant m}_{\{\pm1\}}\exp\left(2\left(\frac{d}{m+1}+m\log d\right)\right).
\end{aligned}$$

Tomando $m=[\sqrt{d/\log d}]$ en la desigualdad previa tenemos

$$BH^{\leqslant d}_{\{\pm1\}} \leqslant BH^{\leqslant[\sqrt{d/\log d}]}_{\{\pm1\}}\exp\left(4\sqrt{d\log d}\right),$$

e iterando (como en el caso homogéneo) concluimos la afirmación. □

2.3. Observaciones y Problemas Abiertos

2.3.1. La conjetura de Aaronson-Ambainis

El Teorema 0.0.2 prueba que la constante $BH^{\leqslant d}_{\{\pm1\}}$ crece a lo sumo subexponencial en el grado d, pero podría ser cierto que crece a lo sumo polinomialmente, i.e., $BH^{\leqslant d}_{\{\pm1\}} \leqslant poly(d)$.

En lo siguiente intentaremos explicar cómo el ciclo de ideas de Bohnenblust-Hille y lo anterior está vinculado con la teoría de información cuántica.

Tomando $n, k \in \mathbb{N}$, una función $f : (\{\pm 1\}^n)^k \to \mathbb{R}$ se dice que es una forma multilineal si ésta es lineal en cada entrada, i.e., para $x_1 \in \mathbb{R}^n$ e $y \in \mathbb{R}^n$ (y similarmente para las demás posiciones)

$$f(x_1 + y, x_2, ..., x_k) = f(x_1, x_2, ..., x_k) + f(y, x_2, ..., x_k).$$

Toda función Booleana k-homogénea la podemos escribir como

$$f(x^1, ..., x^k) = \sum_{1 \leqslant i_1, ..., i_k \leqslant n} a_{i_1, ... i_k} x^1_{i_1} ... x^k_{i_k}, \qquad x^1, ..., x^k \in \{\pm 1\}^n.$$

Montanaro probó en [Mon12] que

$$\left(\sum_{1 \leqslant i_1, ..., i_k \leqslant n} |a_{i_1 ... i_k}|^{\frac{2k}{k+1}} \right)^{\frac{k+1}{2k}} \leqslant k^C \|f\|_\infty, \tag{2.37}$$

donde $C > 0$ es una constante absoluta y

$$\|f\|_\infty = \max_{x^1, ..., x^k \in \{\pm 1\}^n} |f(x^1, ..., x^k)|.$$

Además aplicó esta estimación para juegos XOR en la teoría de informatica cuántica, así como para resolver un caso especial de la conjetura de Aaronson-Ambainis en [AA14] que comentaremos a continuación.

Conjetura 2.3.1. (Aaronson-Ambainis) Existen constantes absolutas $c_1, c_2 > 0$ tal que para toda función $f : \{\pm 1\}^n \to [-1, 1]$ de grado d se satisface

$$c_1 (Var(f)/d)^{c_2} \leqslant \max_{1 \leqslant j \leqslant n} Inf_j(f), \tag{2.38}$$

donde $Var(f) := \sum_{S \neq \emptyset} \hat{f}(S)^2$ e $Inf_j(f) := \sum_{j \in S} \hat{f}(S)^2$ para todo $j \in [n]$.

Notar que si la varianza de f es grande en proporción al grado d (por el d dividiendo en la desigualdad) estaríamos diciendo que existe un $j \in [n]$ tal que la j-influencia de f es un valor significativo.

Aaronson y Ambainis en 2008 mostraron que de ser cierto este resultado implicaría una conjetura bien conocida sobre el poder de la computación cuántica, que se remonta a una pregunta de 1999 hecha por Fortnow y Rogers en [FR99]. Informalmente la conjetura afirma que todos los algoritmos de consulta cuántica podrían simularse eficientemente mediante algoritmos de consultas clásicas en la mayoría de las entradas.

El trabajo de Aaronson y Ambainis [AA14] donde se enuncia la conjetura, lleva de título "La necesidad de estructura en aceleraciones cuánticas" dicho título se relaciona con la siguiente situación comentada por los autores: si la conjetura fuera cierta seguiría siendo coherente con la existencia de algoritmos cuánticos como el de Simon y Shor. Estos son algoritmos que se pueden resolver exponencialmente más rápido en una computadora cuántica que en una computadora clásica. Simplemente lo que se demostraría es la importancia de la estructura global para los algoritmos de Simon y Shor, es decir cierta cualidad en estos algoritmos es extremadamente no aleatoria, por ejemplo, codificar una función periódica (en el caso del algoritmo de Shor) o codificar una función que oculta una cadena secreta s (en el caso de Simon). Esto subrayaría que las aceleraciones cuánticas superpolinómicas dependen de la estructura.

No está claro si existe o no un vínculo directo entre el supuesto crecimiento polinomial en d de la constante de Bohnenblust-Hille $BH^{\leqslant d}_{\{\pm 1\}}$ con la Conjetura de Aaronson-Ambainis. Lo relevante de la sección es que podemos destacar dos hechos que conectan ambas preguntas.

Siguiendo el enfoque de Montanaro en [Mon12, Corolario 20] tenemos en primer lugar:

Caso A 2.3.2. Supongamos que $BH^{\leqslant d}_{\{\pm 1\}}$ no crece más que polinomialmente en d. Entonces para toda función $f : \{\pm 1\}^n \to [-1, 1]$ de grado d tal que $|\hat{f}(S)| = \alpha$ para todo $S \subset [n]$. Se satisface que existen $c_1, c_2 > 0$ constantes absolutas tal que

$$c_1 (Var(f)/d)^{c_2} \leqslant \max_{1 \leqslant j \leqslant n} Inf_j(f),$$

para todo $j \in [n]$.

Observemos que esto da un caso particular de la conjetura de Aaronson-Ambainis, si suponemos cierto el crecimiento como mucho polinomial de la constante de Bohnenblust-Hille.

Demostración. Recordemos que $Var(f) = \sum_{S \neq \varnothing} \hat{f}(S)^2$ e $Inf_j(f) = \sum_{j \in S} \hat{f}(S)^2$ para todo $j \in [n]$. Entonces si consideramos $f : \{\pm 1\}^n \to [-1, 1]$ como en el enunciado. Obtenemos

$$Var(f) = \alpha^2 \sum_{m=1}^{d} \binom{n}{m},$$

y para cada $j \in [n]$

$$Inf_j(f) = \alpha^2 \sum_{m=1}^{d} \binom{n-1}{m-1} = \frac{\alpha^2}{n} \sum_{m=1}^{d} m \binom{n}{m}.$$

Más aún, usando que $\binom{n}{d} \geqslant (\frac{n}{d})^d$ tenemos

$$\alpha \left[\frac{n}{d} \sum_{m=1}^{d} \binom{n}{m} \right]^{\frac{1}{2}} \leqslant \alpha \left[\sum_{m=1}^{d} \binom{n}{m} \right]^{\frac{d+1}{2d}} \leqslant \left(\sum_{S \neq \varnothing} |\hat{f}(S)|^{\frac{2d}{d+1}} \right)^{\frac{d+1}{2d}} \leqslant BH^{\leqslant d}_{\{\pm 1\}}.$$

Juntando todo obtenemos

$$\frac{Var(f)^2}{\max_{1 \leqslant j \leqslant n} Inf_j(f)} \leqslant \alpha^2 n \sum_{m=1}^{d} \binom{n}{m} \leqslant d \left(BH^{\leqslant d}_{\{\pm 1\}} \right)^2 \leqslant b\ d^l, \tag{2.39}$$

con $b > 0$ y $l \geqslant 2$ constantes absolutas, la última desigualdad se debe a la suposición que $BH_{\{\pm 1\}}^{\leqslant d}$ no crece más que polinomialmente en d. Por otro lado, como $0 < Var(f) \leqslant 2$ ya que $|f(x)| \leqslant 1$ para todo $x \in \{\pm 1\}^n$, tenemos que

$$\frac{1}{2^{l-2}}\frac{1}{b}\left(\frac{Var(f)}{d}\right)^l = \frac{1}{2^{l-2}}\frac{1}{b}\frac{Var(f)^2 Var(f)^{l-2}}{d^l} \leqslant \frac{1}{b}\frac{Var(f)^2}{d^l}. \tag{2.40}$$

Luego de (2.39) y (2.40) concluimos

$$\frac{1}{2^{l-2}}\frac{1}{b}\left(\frac{Var(f)}{d}\right)^l \leqslant \max_{1\leqslant j\leqslant n} Inf_j(f),$$

lo deseado. □

Por otro lado, por [IDO06] podemos enunciar el siguiente resultado:

Caso B 2.3.3. Existen constantes absolutas $c_1, c_2 > 0$ tal que para toda función $f : \{\pm 1\}^n \to [-1, 1]$ de grado d se satisface

$$c_1(Var(f)/2^d)^{c_2} \leqslant \max_{1\leqslant j\leqslant n} Inf_j(f),$$

para todo $j \in [n]$.

Notemos que esta afirmación se consigue reemplazando en la conclusión de la Conjetura 2.3.1 el d por el 2^d. Se puede encontrar una breve prueba de este hecho basada en la hipercontractividad en [OZ16]. Lo relevante es que existe una estrecha similitud entre la idea principal de esta prueba y la demostración de las desigualdades de tipo Bohnenblust-Hille, como detallaremos a continuación.

Demostración. Tomando $f : \{\pm 1\}^n \to [-1, 1]$ de grado d tenemos que

$$Var(f) \leqslant \sum_{j=1}^{n} Inf_j(f) \leqslant \sqrt{\max_{1\leqslant j\leqslant n} Inf_j(f)} \sum_{j=1}^{n} \sqrt{Inf_j(f)}. \tag{2.41}$$

Siguiendo la notación y los pasos de la demostración del Teorema 0.0.2, tenemos que, si $S = \{d\}$ y $S' = [d-1]$, entonces

$$\sum_{j=1}^{n} \sqrt{Inf_j(f)} \leqslant \sum_{j_1\in\mathcal{J}(S,n)} \left(\sum_{j_2\in\mathcal{J}(S',n)} |a_{j_1\oplus j_2}|^2\right)^{\frac{1}{2}}.$$

Acotando esto último de la misma manera que en (2.34), con la distinción que en este caso usaremos el resultado de hipercontractividad que compara la norma dos con la norma uno, más precisamente (2.8), y así obtenemos

$$\sum_{j_1\in\mathcal{J}(S,n)} \left(\sum_{j_2\in\mathcal{J}(S',n)} |a_{j_1\oplus j_2}|^2\right)^{\frac{1}{2}} \leqslant \binom{d}{1} e^{d-1} 2d \leqslant d^2 e^d.$$

Por lo tanto, aplicando esta estimación para (2.41) tenemos

$$\frac{Var(f)^2}{\max_{1\leqslant j\leqslant n} Inf_j(f)} \leqslant d^4 e^{4d}. \tag{2.42}$$

Por otro lado, existe una constantes absolutas $l > 0$ lo suficientemente grande tal que

$$\frac{d^4 e^{4d}}{2^{dl}} \leqslant 1.$$

Juntando esto con (2.42), se obtiene

$$\frac{1}{2^{l-2}}\left(\frac{Var(f)}{2^d}\right)^l = \frac{Var(f)^2}{d^4e^{4d}}\frac{Var(f)^{l-2}d^4\ e^{4d}}{2^{l-2}2^{dl}} \leqslant \frac{Var(f)^2}{d^4e^{4d}},$$

que es el resultado deseado. □

Para finalizar comentamos que este estilo de problema parece estar bien popularizado en ciertas comunidades, por ejemplo, Y. Filmus, H. Hatami, S. Heilman, E. Mossel, R. O'Donnell, S. Sachdeva, A. Wan y K. Wimmer enumeran preguntas en su colección de problemas abiertos "Análisis real en informática: una colección de problemas abiertos" (ver [FHH$^+$14]), donde se menciona que la conjetura de Aaronson y Ambainis se resuelve para algunas clases especiales de funciones, un caso particular es cuando se toman las funciones $f : \{\pm 1\}^n \to \{\pm 1\}$ es decir, cuando f sólo alcanza dos valores. También la conjetura se puede probar para polinomios simétricos $f : \{\pm 1\}^n \to [-1, 1]$, más precisamente:

Definición 2.3.4. *Sea la función $f : \{\pm 1\}^n \to \mathbb{R}$ se dice simétrica si $f(x_1, .., x_n)$ es independiente por permutaciones en las variables $x_1, ..., x_n$.*

Enunciemos el resultado exacto:

Teorema 2.3.5. *Existe una constante universal $c_1 > 0$ tal que para cualquier función simétrica $f : \{\pm 1\}^n \to [-1, 1]$ de grado d, existe $j \in [n]$ tal que*

$$c_1 \left(Var(f)/d\right)^4 \leqslant Inf_j(f).$$

Podemos encontrar la prueba en [Bac12]. Un comentario adicional es que la prueba en el caso simétrico es factible porque la expresión de la j-influencia de f no depende de j, para aclarar esta afirmación se puede probar (ver [O'D14, pág. 48]) que

$$Inf_j(f) = \mathbb{E}|D_j(f)|^2,$$

donde $D_j(f)(x_1, ..., x_n) := \frac{f(x_1,...,x_j,...,x_n)-f(x_1,...,-x_j,...,x_n)}{2}$, y la esperanza se toma en $\{\pm 1\}^n$. Entonces uno puede reemplazar la j-ésima influencia por $\frac{1}{n}\mathbb{E}\left(\sum_{j=1}^n |D_j(f)|^2\right)$, luego la pregunta se convierte según los expertos en "un problema típico de multiplicadores de Fourier" y conocen técnicas para abordarlo.

2.3.2. Norma de Lorentz

Otro espacio en el que se han estudiado las desigualdades de Bohnenblust-Hille es el espacio de sucesiones de Lorentz, en lo que sigue trataremos de comentar algunos resultados y cómo por este camino podemos generalizar la constante de Bohenblust-Hille para funciones en el cubo Booleano.

Definición 2.3.6. *Dado $K = \mathbb{R}$ o $\mathbb{C}$ consideramos $K^{\mathbb{N}}$, el espacio vectorial de todas las sucesiones $(x_n)_{n\in\mathbb{N}}$ con $x_n \in K$. Un espacio de sucesiones de Banach es un subespacio X de $K^{\mathbb{N}}$ dotado de una norma completa tal que, si $x = (x_n)_n \in K^{\mathbb{N}}$ e $y = (y_n)_n \in X$ son sucesiones que cumplen $|x_n| \leqslant |y_n|$ para todo $n \in \mathbb{N}$, entonces $x \in X$ y $\|x\| \leqslant \|y\|$.*

Observemos que aquí indexamos las sucesiones en $\mathbb{N}$, pero también podriamos indexar las sucesiones en un subconjunto $I \subset \mathbb{N}$, como haremos luego.

Definición 2.3.7. *Un espacio de sucesiones de Banach es simétrico si $x_\sigma = (x_{\sigma(n)})_n \in X$ y $\|x_\sigma\| = \|x\|$ para toda permutación σ de N.*

Consideremos el espacio de sucesiones de Banach $(X(I), \|\cdot\|_X)$ con sucesiones a $\mathbb{C}$-valores $(x_i)_{i\in I}$. En lo que sigue, los espacios de Lorentz jugarán un rol importante.

Definición 2.3.8. *Tomando $1 \leqslant p < \infty$ y $1 \leqslant q \leqslant \infty$, el espacio de Lorentz $\ell_{p,q}(I)$ ($\ell_{p,q}$ para abreviar), consiste de todas las sucesiones $x = (x_i)_{i\in\mathbb{N}}$ para el cual la expresión*

$$\|x\|_{\ell_{p,q}} = \begin{cases} \left(\sum_{k\in J} x_k^{*q} k^{\frac{q}{p}-1}\right)^{\frac{1}{q}} & si \quad q < \infty \\ \\ \sup_{k\in J} k^{\frac{1}{p}} x_k^* & si \quad q = \infty, \end{cases} \tag{2.43}$$

es finita. Aquí, como es usual para $x = (x_i)_{i\in I} \in \ell_\infty(I)$, denotamos por $x^* = (x_j^*)_{j\in J}$ el reordenamiento decreciente de x definido por

$$x_j^* = \text{ínf}\{\lambda > 0, \#\{i \in I, |x_i| > \lambda\} \leqslant j\}, \qquad j \in J,$$

donde $J = \{1, ..., n\}$ cuando card$(I) = n$ y $J = \mathbb{N}$ cuando I es infinito.

La expresión (2.43) es una norma si $q \leqslant p$, y es una cuasi-norma si $q > p$, es decir no cumple la desigualdad triangular pero sí $\|x + y\| \leqslant k(\|x\| + \|y\|)$ para algún $k > 0$. En el segundo caso $\|\cdot\|_{\ell_{p,q}}$ es equivalente a una norma. Por supuesto $\ell_{p,p}$ es el espacio ℓ_p ya que la aplicación $x \to x^*$ es una isometría. Andreas Defant y Mieczysław Masytiło en [DM16] probaron el siguiente resultado:

Teorema 2.3.9. *Para cada entero positivo d el espacio de Lorentz $\ell_{\frac{2d}{d+1},1}$ es el espacio X de sucesiones de Banach simétrico más pequeño para el cual existe una desigualdad de Bohnenblust-Hille (reemplazando en el lado izquierdo de la desigualdad la norma $\ell_{\frac{2d}{d+1}}$ por la norma de X).*

La expresión "más pequeño" hace referencia al suceso más general que encontramos con dellate en [BKS88, Teorema 2.5.2]: Si $1 \leqslant p < \infty$ y X es un espacio de sucesiones de Banach simétrico con norma $\| \cdot \|_X$ entonces tenemos que $\ell_{p,1} \hookrightarrow X$ con

$$\|x\|_X \leqslant \|x\|_{\ell_{p,1}}, \qquad x \in \ell_{p,1}.$$

El siguiente teorema motiva la generalización de la desigualdad de Bohenblust-Hille para funciones $f : \{\pm 1\}^n \to \mathbb{R}$ de grado d.

Teorema 2.3.10. *Existe $L \geqslant 1$ tal que para toda función $f : \{\pm 1\}^n \to \mathbb{R}$ de grado d se satisface*

$$\sum_{k=1}^{D(n,d)} f^*(k) \frac{1}{k^{\frac{d-1}{2d}}} \leqslant L^d \|f\|_\infty, \tag{2.44}$$

donde $D(n,d) = \sum_{m=0}^{d} \binom{n}{m}$ y $(f^*(k))_{k=1}^{D(n,d)}$ es el reordenamiento decreciente de

$$\left(|\hat{f}(S)| \right)_{S \subset [n], |S| \leqslant d}.$$

Notar que la norma en la parte izquierda de la desigualdad (2.44) es la norma del espacio de Lorentz

$$\ell_{\frac{2d}{d+1},1}\left(\{S : S \subset [n], |S| \leqslant d\}\right) = \left(\prod_{S \subset [n], |S| \leqslant d} \mathbb{R}, \qquad \| \cdot \|_{\frac{2d}{d+1},1} \right),$$

y como ya mencionamos, podemos entender el resultado como una generalización de la desigualdad Bohnenblust-Hille para funciones en el cubo Booleano con la restricción de que no sabemos si la constante en (2.44) todavía se comporta subexponencialmente. Si $C(d)$ denota la mejor constante posible en (2.44) (reemplazando L^d), entonces la siguiente idea de la demostración muestra que

$$\limsup_{d \to \infty} C(d)^{\frac{1}{d}} \leqslant \sqrt{2}(1 + \sqrt{2}).$$

Demostración. Sea $f : \{\pm 1\}^n \to \mathbb{R}$ un polinomio de grado d. Si Q_f es el correspondiente polinomio tetraedral, luego considerando un polinomio complejo (de grado d) se deduce de [DM16, Teorema 16] que para cada $\epsilon > 0$ existe una constante $C(\epsilon)$ que depende sólo de ϵ tal que

$$\left\| \left(\hat{f}(S) \right)_{S \subset [n], |S| \leqslant d} \right\|_{\ell_{\frac{2d}{d+1},1}} \leqslant C(\epsilon)(\sqrt{2} + \epsilon)^d \sup_{z \in \mathbb{D}^n} |Q_f(z)|. \tag{2.45}$$

Tener en cuenta que el resultado en [DM16, Teorema 16] se realiza para funciones homogéneas de grado d. Pero el resultado se extiende fácilmente al caso de grado d; de hecho, cada polinomio grado d, Q_f en $[-1, 1]^n$ puede verse como un polinomio d-homogéneo en

$\mathbb{T}^{n+1}$ con la misma norma compleja y los mismos coeficientes. Por otro lado, recordemos el resultado de Klimek del Lema 1.0.9 (2)

$$\sup_{z\in\mathbb{D}^n} |Q_f(z)| \leqslant (1+\sqrt{2})^d \sup_{x\in[-1,1]} |Q_f(x)|.$$

Combinando esto con (2.45) obtenemos (2.44). $\square$

Capítulo 3

Fenómeno de Bohr para funciones en el cubo Booleano

Para comenzar presentaremos la definición principal del capítulo, el radio Booleano para una familia de funciones $\mathcal{F}$ en $\{\pm 1\}^N$, pero primero lo definiremos para una función cualquiera $f : \{\pm 1\}^N \to \mathbb{R}$:

Definición 3.0.1. *El radio Booleano de una función $f : \{\pm 1\}^N \to \mathbb{R}$ es el número positivo $\rho = \rho(f)$ que satisface*

$$\sum_{S \subset [N]} |\hat{f}(S)| \rho^{|S|} = \|f\|_\infty.$$

En general, calcular el radio Booleano exacto de una función dada puede ser complicado, tanto como determinar los coeficientes de Fourier. Sin embargo, es posible dar estimaciones cuando conocemos alguna característica de la función, como el número de variables de las que depende, la distribución del espectro de Fourier (ya sea que se concentre en niveles bajos o en conjuntos con el mismo cardinal) o el tamaño de $\mathbb{E}[f]$ en relación con $\|f\|_\infty$. Para esto consideremos el siguiente concepto:

Definición 3.0.2. *Tomemos una clase $\mathcal{F}$ de funciones en el cubo Booleano $\{\pm 1\}^N$, el radio Booleano de $\mathcal{F}$ es definido como*

$$\rho(\mathcal{F}) = \sup \left\{ \rho \geqslant 0 : \sum_{S \subset [N]} |\hat{f}(S)| \rho^{|S|} \leqslant \|f\|_\infty, \quad \text{para toda } \ f \in \mathcal{F} \right\},$$

o equivalentemente

$$\rho(\mathcal{F}) = \text{ínf}\{\rho(f) : f \in \mathcal{F}\}.$$

Con el fin de estudiar este radio nos concentraremos en las siguientes cinco clases, que naturalmente surgen del desarrollo de la expansión de Fourier-Walsh de una función cualquiera. Es decir, para cada $f : \{\pm 1\}^N \to \mathbb{R}$ tenemos

$$f(x) = \sum_{S \subset [n]} \hat{f}(S)\chi_S, \qquad x \in \{\pm 1\}^N. \tag{3.1}$$

Definimos para cada $N \in \mathbb{N}$, $0 \leqslant d \leqslant N$ y $0 \leqslant \rho \leqslant 1$:

$$\begin{aligned}
\mathcal{B}^N &:= \text{ toda función f} : \{\pm 1\}^N \to \mathbb{R}, \\
\mathcal{B}^N_{=d} &:= \text{ toda función } d\text{-homogénea } f : \{\pm 1\}^N \to \mathbb{R}, \\
\mathcal{B}^N_{\text{hom}} &:= \text{ toda función homogénea } f : \{\pm 1\}^N \to \mathbb{R}, \\
\mathcal{B}^N_{\leqslant d} &:= \text{ toda función } f : \{\pm 1\}^N \to \mathbb{R} \text{ con deg } f \leqslant d, \\
\mathcal{B}^N_{\delta} &:= \text{ toda función } f : \{\pm 1\}^N \to \mathbb{R} \text{ satisfaciendo } |\mathbb{E}[f]| \leqslant (1-\delta)\|f\|_\infty.
\end{aligned}$$

En la primera sección de este capítulo se probará que el radio Booleano de $\mathcal{B}^N$ obtiene la fórmula

$$\rho(\mathcal{B}^N) = 2^{\frac{1}{N}} - 1 = \frac{\log 2}{N}(1 + o(1)).$$

Luego, en la siguiente sección trataremos con las familias de funciones homogéneas $\mathcal{B}^N_{=m}$ y $\mathcal{B}^N_{hom}$. En el Teorema 3.2.1 obtenemos que para cada $1 \leqslant m \leqslant N$ existe C_m (independiente de N) tal que

$$C_m^{-1} N^{\frac{1}{2m}} \binom{N}{m}^{-\frac{1}{2m}} \leqslant \rho(\mathcal{B}^N_{=m}) \leqslant C_m N^{\frac{1}{2m}} \binom{N}{m}^{-\frac{1}{2m}},$$

donde además $\lim_m C_m = 1$. Esto muestra que para m grande, la estimación previa es asintóticamente óptima. La demostración de la acotación inferior se basa en un trabajo reciente de la desigualdad (de tipo Bohnenblust-Hille (3.7)) mientras que la estimación superior requiere de un argumento probabilístico (ver Lema 3.2.3). El hecho que C_m converge a uno es la llave para probar

$$\rho(\mathcal{B}^N_{\text{hom}}) = \sqrt{\frac{\log N}{N}}(1 + O(1)),$$

ver Corolario 3.2.2.

En la siguiente sección, trabajaremos con el caso de funciones de grado d, aquí la estimación para el radio Booleano de $\mathcal{B}^N_{\leqslant d}$ es menos precisa que en el caso homogéneo. Mostraremos en el Teorema 3.3.1 que para cada $1 \leqslant d \leqslant N$ existen constantes c_d y C_d (independientes de N) tal que

$$c_d \frac{1}{N^{\frac{1}{2}}} \leqslant \rho(\mathcal{B}^N_{\leqslant d}) \leqslant C_d \frac{1}{N^{\frac{d-1}{2d}}}.$$

La demostración de la acotación inferior requiere una desigualdad para funciones en $\{\pm 1\}^N$ que podría ser de interés independiente, y afirma lo siguiente: si $f : \{\pm 1\}^n \to \mathbb{R}$ es una función de grado d, entonces

$$\left(\sum_{S\neq\varnothing} |\hat{f}(S)|^2\right)^{\frac{1}{2}} \leqslant 2e^d(1-|\hat{f}(\varnothing)|).$$

El trabajo previo para probarlo recuerda las desigualdades de F. Wiener y Caratheodory para polinomios complejos (véase los Corolarios 3.3.3, 3.3.5 y el Teorema 3.3.2). Finalmente, la última sección está apuntada a la estimación del radio Booleano de $\mathcal{B}_\delta^N$. El resultado principal es el Teorema 3.4.1 el cual muestra que existe una constante absoluta $C > 0$ tal que para todo $N \in \mathbb{N}$ y $\frac{1}{2^N} \leqslant \delta \leqslant 1$

$$C^{-1}\frac{1}{\sqrt{N}\sqrt{\log(\frac{2}{\delta})}} \leqslant \rho(\mathcal{B}_\delta^N) \leqslant C\frac{1}{\sqrt{N}\sqrt{\log(\frac{2}{\delta})}}.$$

Notar que $\delta = \frac{1}{2^N}$ es el valor "sensible" más pequeño, ya que para todo $0 \leqslant \delta < \frac{1}{2^N}$ tenemos consecuentemente que

$$\Omega(\frac{1}{N}) = \rho(\mathcal{B}^N) \leqslant \rho(\mathcal{B}_\delta^N) \leqslant \rho(\mathcal{B}_{\frac{1}{2^N}}^N) = \Theta(\frac{1}{N}),$$

donde $\Omega(f(n)) = g(n)$ significa que $g(n)$ está acotado inferiormente por un múltiplo de $f(n)$ a partir de un cierto $n \in \mathbb{N}$ y $\Theta(f(n)) = g(n)$ es la notación para decir que $g(n)$ está acotada inferior y superiormente por múltiplos de $f(n)$, cuando n es grande.

La demostracón de la estimación inferior del teorema depende de la desigualdad hipercontractiva, mientras que para la estimación superior tenemos el estudio del radio Booleano de la siguiente familia de funciones (del tipo threshold)

$$\psi_{\alpha,N} : \{\pm1\}^N \to \mathbb{R}, \quad \psi_{\alpha,N}(x) = sign(x_1 + ... + x_N - \alpha), \qquad 0 \leqslant \alpha < N.$$

En efecto, probaremos en el Teorema 3.4.4 que existe una constante $C > 0$ tal que para todo $0 \leqslant \alpha \leqslant N$

$$C^{-1}\frac{1}{\alpha+\sqrt{N}} \leqslant \rho(\psi_{\alpha,N}) \leqslant C\frac{1}{\alpha+\sqrt{N}}.$$

La estrategia de la prueba permite dar la precisión asintótica del radio Booleano de la función mayoría $Maj_N = \psi_{0,N}$ para N impar (Corolario 3.4.10).

3.1. Caso 1: Todas las funciones en el cubo Booleano

El siguente teorema dará con precisión el radio Booleano para todas las funciones en el cubo Booleano N-dimensional $\{\pm1\}^N$.

Teorema 3.1.1. *Para cada* $N \in \mathbb{N}$

$$\rho(\mathcal{B}_N) = 2^{\frac{1}{N}} - 1. \tag{3.2}$$

En particular,

$$\lim_{N\to\infty} N\rho(\mathcal{B}_N) = \log 2 + o(N).$$

Comparando este resultado con el caso complejo, es decir para todos los polinomios $f(z) = \sum_{\alpha \in \mathbb{N}_0^N} \hat{f}(\alpha) z^\alpha$, en el N-ésimo toro dimensional $\mathbb{T}^N$ definido en (1.4), en el caso Booleano tenemos la estimación precisa de $\rho(\mathcal{B}_N)$ para todo N, y en el caso de complejo sólo se conoce para $N = 1$, como vimos en el Teorema 0.0.1. Un ingrediente relevante de la prueba original de $K_1 = \frac{1}{3}$ (más precisamente la modificación de la demostración de Bohr por M. Riesz, I. Schur, y F. Wiener) es un resultado de F. Wiener el cual establece que para toda función holomorfa f con coeficientes c_n en el disco abierto $\mathbb{D}$ y con $\|f\|_\infty \leqslant 1$ se tiene

$$|c_n| \leqslant 1 - |c_o|^2 \leqslant 2(1 - |c_0|), \qquad n \in \mathbb{N}. \tag{3.3}$$

El punto clave de la demostración del Teorema 3.1.1 es que para el caso Booleano tenemos un resultado más fuerte del tipo F. Wiener, como muestra el siguiente lema:

Lema 3.1.2. *Para toda función $f : \{\pm 1\}^N \to [-1, 1]$ y par de subconjuntos $A, B \subset [N]$ con $A \neq B$ tenemos que*

$$|\hat{f}(A)| + |\hat{f}(B)| \leqslant 1.$$

Demostración. Sin pérdida de generalidad podemos asumir que existe $k \in B \backslash A$. Entonces, para cada $x \in \{\pm 1\}^N$ tenemos que

$$f(x_1, ..., x_k, ..., x_N) = \sum_{S \subset [N] \backslash \{k\}} \hat{f}(S) x^S + \sum_{S \subset [N], k \in S} \hat{f}(S) x^S,$$

y

$$f(x_1, ..., -x_k, ..., x_N) = \sum_{S \subset [N] \backslash \{k\}} \hat{f}(S) x^S - \sum_{S \subset [N], k \in S} \hat{f}(S) x^S.$$

Como $\|f\|_\infty \leqslant 1$, combinando ambas igualdades conseguimos

$$\left| \sum_{S \subset [N] \backslash k} \hat{f}(S) x^S \right| + \left| \sum_{S \subset [N], k \in S} \hat{f}(S) x^S \right| \leqslant 1.$$

Luego la afirmación se cumple, pues

$$\begin{aligned} |\hat{f}(A)| + |\hat{f}(B)| &= |\mathbb{E}[f x^A]| + |\mathbb{E}[f x^B]| \\ &= \left| \mathbb{E}[\sum_{S \subset [N] \backslash k} \hat{f}(S) x^S x^A] \right| + \left| \mathbb{E}[\sum_{S \subset [N], k \in S} \hat{f}(S) x^S x^B] \right| \\ &\leqslant \mathbb{E} \left| \sum_{S \subset [N] \backslash k} \hat{f}(S) x^S \right| + \mathbb{E} \left| \sum_{S \subset [N], k \in S} \hat{f}(S) x^S \right| \leqslant 1, \end{aligned}$$

donde en la penúltima desigualdad introdujimos el módulo en la esperanza y sumamos sobre los conjunos $S \subset [N]$ respetando la inclusión/exclusión de k en S.

$\square$

Estamos listos para probar el Teorema 3.1.1.

Demostración. Sea $f : \{\pm 1\}^N \to \mathbb{R}$ con $\|f\|_\infty \leqslant 1$. Aplicando la estimación del Lema 3.1.2, para todo $\rho > 0$ tenemos

$$\sum_{S\subset[N]} |\hat{f}(S)|\rho^{|S|} = |\hat{f}(\varnothing)| + \sum_{S\neq\varnothing} |\hat{f}(S)|\rho^{|S|}$$

$$\leqslant |\hat{f}(\varnothing)| + (1 - |\hat{f}(\varnothing)|) \sum_{k=1}^{N} \binom{N}{k} \rho^k$$

$$= |\hat{f}(\varnothing)| + (1 - |\hat{f}(\varnothing)|)((1+\rho)^N - 1).$$

Si tomamos $\rho = 2^{\frac{1}{N}} - 1$, entonces el último término es igual a 1 y concluimos que $\rho(\mathcal{B}_N) \geqslant 2^{\frac{1}{N}} - 1$. Estimemos ahora ρ_N por arriba. Denotamos por $\mathbf{1} \in \{\pm 1\}^N$ el elemento cuyas entradas son todas iguales a 1. Observemos que la función $F : \{\pm 1\}^N \to \{\pm 1\}$ dada por

$$F(x) = \begin{cases} -1 & si \quad x = \mathbf{1} \\ 1 & si \quad x \neq \mathbf{1}, \end{cases}$$

la expansión de Fourier de F está dada por

$$F(x) = 1 - \frac{1}{2^{N-1}} + \sum_{S\neq\varnothing} \frac{-1}{2^{N-1}} x^S, \qquad x \in \{\pm 1\}^N.$$

Para ver esto, notar que la función $f : \{\pm 1\}^N \to \mathbb{R}$ definida como $f(x) = 1$ para $x = \mathbf{1}$ y $f(x) = 0$ para $x \neq \mathbf{1}$ tiene la siguiente expansión Fourier

$$f(x) = \prod_{n=1}^{N} \left(\frac{1 + x_n}{2} \right) = \frac{1}{2^N} \sum_{S\subset N} x^S, \qquad x \in \{\pm 1\}^N,$$

como $F(x) = 1 - 2f(x)$ para todo $x \in \{\pm 1\}^N$, la expansión de Fourier deseada para F se sigue haciendo un despeje.

Finalmente, por la definición del radio Booleano $\rho_N = \rho(\mathcal{B}_N)$ obtenemos que

$$1 = \|F\|_\infty \geqslant 1 - \frac{1}{2^{N-1}} + \sum_{S\neq\varnothing} \frac{1}{2^{N-1}} \rho_N^{|S|} = 1 - \frac{1}{2^{N-1}} + \frac{1}{2^{N-1}}((1+\rho_N)^N - 1).$$

Así debe ocurrir $(1+\rho_N)^N - 1 \leqslant 1$, despejando conseguimos la acotación superior de ρ_N y esto prueba que $\rho_N = 2^{\frac{1}{N}} - 1$. $\square$

3.2. Caso 2: Las funciones homogéneas

El problema de estimación del radio Booleano en las clases de funciones homogéneas $\mathcal{B}^N_{=m}$ y $\mathcal{B}^N_{hom}$ en el cubo Booleano $\{\pm 1\}^N$ es más delicado que el caso $\mathcal{B}^N$, como podemos notar en el siguiente resultado.

Teorema 3.2.1. *Existe una constante $C > 1$ tal que para todo $1 \leqslant m \leqslant N$*

$$c_m N^{\frac{1}{2m}} \binom{N}{m}^{\frac{-1}{2m}} \leqslant \rho(\mathcal{B}^N_{=m}) \leqslant C_m N^{\frac{1}{2m}} \binom{N}{m}^{\frac{-1}{2m}}, \tag{3.4}$$

donde

$$c_m = \frac{1}{m^{\frac{1}{2m}} C^{\sqrt{\frac{\log m}{m}}}} \qquad y \qquad C_m = C^{\frac{1}{m}}.$$

Antes de pasar a las consecuencias de este resultado y su demostración, comparemos el resultado análogo que conocemos para el toro N-dimensional $\mathbb{T}^N$ (en lugar del cubo Booleano $\{\pm 1\}^N$). Recordando la definición del radio m-homogéneo de Bohr $K_N^{=m}$ (ver (1.8)) se sabe que cumple la siguiente estimación (podemos ver la prueba implícitamente en [PSS14] y [DFOC+11]): existe una constante absoluta $D > 1$ tal que para todo m, N satisface

$$\frac{1}{D}\left(\frac{m}{N}\right)^{\frac{m-1}{2m}} \leqslant K_N^{=m} \leqslant D\left(\frac{m}{N}\right)^{\frac{m-1}{2m}}, \tag{3.5}$$

por lo tanto para m fijo, el decaimiento asintótico del radio de Bohr $K_N^{=m}$ y el radio Booleano $\rho(\mathcal{B}^N_{=m})$ coinciden.

El control del decaimiento de c_m y C_m cuando $m \to \infty$ dará lo siguiente.

Corolario 3.2.2.

$$\lim_{N\to\infty} \sqrt{\frac{N}{\log N}}\, \rho(\mathcal{B}^N_{hom}) = 1.$$

Basado en la fórmula anterior llegamos a la conclusión esencial de que el estudio del radio Booleano y radio de Bohr muestra una notable diferencia. De hecho si tenemos en cuenta que

$$\lim_{N\to\infty} \sqrt{\frac{N}{\log N}}\, K^N_{hom} = 1, \tag{3.6}$$

donde la definición de K^N_{hom} está hecha en (1.6) (veáse [PSS14] y [DFOC+11]), entonces en el caso complejo, el radio de Bohr de todos los polinomios (1.5) y el radio de Bohr de todos los polinomios homogéneos (3.6) no se distinguen. Pero para las funciones en el cubo Booleano $\{\pm 1\}^N$, por el precedente corolario y el resultado del Teorema 3.1.1 muestra una diferencia dramática.

La razón más profunda de esta diferencia es la distorsión entre una función $f : \{\pm 1\}^N \to \mathbb{R}$ y las partes homogenéas de esta, es decir la siguiente expresión representa la parte m-homogénea de una función $f : \{\pm 1\}^N \to \mathbb{R}$

$$f_m(x) = \sum_{S\subset[N],|S|=m} \hat{f}(S)x^S, \qquad 0 \leqslant m \leqslant N,$$

entonces se vuelve visible si comparamos con la norma supremo, la estimación

$$\|f_m\|_\infty \leqslant C^m \|f\|_\infty,$$

donde la mejor constante C por lo menos es mayor o igual a $\sqrt{2}+1$ por el Lema 1.0.9, mientras que por la desigualdad de Cauchy el resultado análogo para polinomios complejos en $\mathbb{T}^N$, obtenemos una constante igual a 1.

Para la demostración del Teorema 3.2.1 y Corolario 3.2.2 necesitamos resultados preliminares. Por un lado necesitamos el resultado principal del capítulo anterior, el cual dice que existe una constante absoluta $C > 0$ tal que para toda función $f : \{\pm1\}^N \to \mathbb{R}$ de grado d tenemos que

$$\left(\sum_{|S|\leqslant d} |\hat{f}(S)|^{\frac{2d}{d+1}}\right)^{\frac{d+1}{2d}} \leqslant C^{\sqrt{d\log d}}\|f\|_\infty. \tag{3.7}$$

La característica relevante para nosotros es que la constante involucrada tiene crecimiento subexponencial en el sentido que

$$\lim_{d\to\infty} \left(C^{\sqrt{d\log d}}\right)^{\frac{1}{d}} = 1,$$

esto jugará un rol fundamental en la demostración de Teorema 3.2.1. El segundo resultado es de naturaleza probabilística. En particular, es consecuencia de un hecho popularmente conocido como la desigualdad de tipo Kahane-Salem-Zygmund, y será crucial para estimar $\rho(\mathcal{B}_m^N)$ superiormente.

Lema 3.2.3. *Para todo $N \in \mathbb{N}$ y para toda familia de números reales $(c_S)_{S\subset[N]}$, existe una elección de signos $(\xi_S)_{S\subset[N]}$ en $\{\pm1\}$ tal que*

$$\left\|\sum_{S\subset[N]} \xi_S c_S x_S\right\|_\infty \leqslant 6\sqrt{\log 2}\sqrt{N}\left(\sum_{S\subset[N]} |c_S|^2\right)^{\frac{1}{2}}. \tag{3.8}$$

Antes de probar el lema enunciemos el siguiente resultado de tipo Kahane-Salem-Zygmund, podemos encontrar la demostración en [KK93, Teorema 1, p. 68].

Teorema 3.2.4. *Sea E un espacio de medida con medida positiva μ, y $\mu(E) < \infty$. Sea B el espacio lineal de funciones medibles acotadas en E, cerrado bajo conjugaciones complejas, y supongamos que existe $\rho > 0$ con la siguiente propiedad: si $f \in B$ y f es real, existe un conjunto medible $I = I(f) \subset E$ tal que $\mu(I) \geqslant \mu(E)/\rho$ y $|f(t)| \geqslant \frac{1}{2}\|f\|_\infty$ para $t \in I$. Considerando una suma finita de variables aleatorias*

$$P = \sum \xi_n f_n,$$

donde ξ_n es una sucesión de variables aleatorias subnormal y $f_n \in B$. Entonces, para todo $k > 2$,

$$\mathbb{P}\left(\|P\|_\infty \geqslant 3\left(\sum \|f_n\|_\infty^2 \log(2\rho k)\right)^{\frac{1}{2}}\right) \leqslant \frac{2}{k}. \tag{3.9}$$

Demostración. (del Lema 3.2.3) Toda $f : \{\pm 1\}^N \to \mathbb{R}$ satisface que

$$\mathbb{P}\left(\{x \in \{\pm 1\}^N : |f(x)| \geqslant \frac{1}{2}\|f\|_\infty\}\right) \geqslant \frac{1}{2^N}. \tag{3.10}$$

Pues caso contrario, dado que la cantidad de elemento de $\{\pm 1\}^N$ es 2^N, concluimos que $|f(x)| < \frac{1}{2}\|f\|_\infty$ para todo x elemento en el cubo Booleano, especialmente para el que realiza el máximo, y esto es absurdo.

Si $(\xi_S)_{S\subset[N]}$ son variables aleatorias independientes en un espacio de probabilidad $(\Omega, \Sigma, \mathbb{P})$ tomando las evaluaciones de 1 y de -1 con igual probabilidad, luego usando la desigualdad anterior en el Teorema 3.2.4, es decir tomando $P = \sum_{S\subset[N]} \xi^S c_S \chi_S$, $\rho = 2^N$ y $k = 4$ tenemos

$$\mathbb{P}\left\{\left\|\sum_{S\subset[N]} \xi^S c_S \chi_S\right\|_\infty \geqslant 3\left(\log(2^{N+3})\sum_{S\subset[N]} |c_S|^2\right)^{\frac{1}{2}}\right\} \leqslant \frac{1}{2},$$

lo que conlleva a la conclusión esperada. □

Ya familiarizados con los dos conceptos anteriores podemos demostrar el Teorema 3.2.1.

Demostración. Comenzamos dando una acotación superior para $\rho = \rho(\mathcal{B}_m^N)$, para el cual hacemos uso del Lema 3.2.3. Consideramos la familia $(c_S)_S \subset [N]$ donde $c_S = 1$ si $|S| = m$ y $c_S = 0$ de otra manera, entonces existe $(\xi_S)_{S\subset[N]}$ en $\{\pm 1\}$ tal que

$$\rho^m \binom{N}{m} = \sum_{|S|=m} \rho^m \leqslant \left\|\sum_{|S|=m} \xi_S x_S\right\|_\infty \leqslant 6\sqrt{\log 2}\sqrt{N}\binom{N}{m}^{\frac{1}{2}},$$

donde la primera desigualdad corresponde a la definición de ρ. Así la estimación superior se desprende de esto último con una constante absoluta C. Ahora probemos la estimación inferior, combinando la desigualdad (3.7) con la desigualdad de Hölder, obtenemos para cualquier función m-homogénea $f : \{\pm 1\}^N \to \mathbb{R}$

$$\sum_{|S|=m} |\hat{f}(S)| \leqslant \left(\sum_{|S|=m} |\hat{f}(S)|^{\frac{2m}{m+1}}\right)^{\frac{m+1}{2m}} \binom{N}{m}^{\frac{m-1}{2m}} \leqslant C^{\sqrt{m\log m}}\binom{N}{m}^{\frac{m-1}{2m}} \|f\|_\infty.$$

Como consecuencia tenemos

$$\sum_{|S|=m} |\hat{f}(S)|\rho^m \leqslant \|f\|_\infty \quad para \quad \rho^m = C^{-\sqrt{m \log m}} \binom{N}{m}^{\frac{1-m}{2m}}.$$

Sabiendo que $\binom{N}{m} \geqslant \left(\frac{N}{m}\right)^m$ para todo entero $1 \leqslant m \leqslant N$, concluimos

$$\rho(\mathcal{B}^N_{=m}) \geqslant C^{-\sqrt{\frac{\log m}{m}}} \binom{N}{m}^{\frac{1}{2m^2}} \binom{N}{m}^{-\frac{1}{2m}} \geqslant \frac{1}{m^{\frac{1}{2m}} C^{\sqrt{\frac{\log m}{m}}}} N^{\frac{1}{2m}} \binom{N}{m}^{-\frac{1}{2m}},$$

esto completa la demostración. □

Por último daremos la demostración del Corolario 3.2.2

Demostración. Notar que $\mathcal{B}^N_{hom}$ es la unión de todas las familias $\mathcal{B}^N_{=m}$esto inmediatamente lleva a la relación

$$\rho(\mathcal{B}^N_{hom}) = \inf_{0 \leqslant\ m \leqslant N} \rho(\mathcal{B}^N_{=m}).$$

Usando la estimación $\sqrt{2\pi} n^{n+\frac{1}{2}} e^{-n} \leqslant n! \leqslant e n^{n+\frac{1}{2}} e^{-n}$ válida para todo $n \geqslant 1$, podemos acotar para todo $1 \leqslant m \leqslant N$

$$\binom{N}{m} \geqslant \frac{\sqrt{2\pi}}{e^2} \left(\frac{N}{m}\right)^m \left(\frac{N}{N-m}\right)^{N-m} \sqrt{\frac{N}{m(N-m)}}.$$

Aplicando está desigualdad en el Teorema 3.2.1, obtenemos que

$$\rho(\mathcal{B}^N_{=m}) \leqslant C_m N^{\frac{1}{2m}} \left(\frac{e^2}{\sqrt{2\pi}}\right)^{\frac{1}{2m}} \sqrt{\frac{m}{N}} \left(1 - \frac{m}{N}\right)^{\frac{1}{2}(\frac{N}{m}-1)} \left(\frac{m(N-m)}{N}\right)^{\frac{1}{4m}}.$$

Tomando $m = [\log N]$, la previa desigualdad lleva a

$$\rho(\mathcal{B}^N_{hom}) \leqslant \rho(\mathcal{B}^N_{=[\log N]}) \leqslant \sqrt{\frac{\log N}{N}}(1 + o(1)).$$

Para probar la estimación inversa, recordemos ahora la estimación inferior dada en el Teorema 3.2.1, que combinada con la desigualdad $\binom{N}{m} \leqslant (\frac{Ne}{m})^m$ válida para $1 \leqslant m \leqslant N$, produce

$$\frac{\rho(\mathcal{B}^N_{=m})}{\sqrt{\frac{\log N}{N}}} \geqslant \frac{N^{\frac{1}{2m}}\sqrt{m}}{C_m \sqrt{\log N}\sqrt{e}} = \frac{1}{C_m} \exp\left(\frac{\log N}{2m} - \frac{1}{2} + \frac{\log m}{2} - \frac{\log\log N}{2}\right). \qquad (3.11)$$

Sea $\epsilon > 0$, fijo $m_0 \in N$ tal que $C_m^{-1} > 1 - \epsilon$ cuando $m \geqslant m_0$. Usando 3.11 y que la función $N \to \frac{\log N}{2m} - \frac{1}{2} + \frac{\log m}{2} - \frac{\log\log N}{2}$ tiende a infinito cuando N tiende a infinito y m está fijo, deducimos la existencia de algún $N_0 \in \mathbb{N}$ tal que

$$\inf_{0 \leqslant m \leqslant m_0} \rho(B^N_{=m}) \geqslant \sqrt{\frac{\log N}{N}}(1 - \epsilon), \qquad N \geqslant N_0. \qquad (3.12)$$

Por otro lado obtenemos, la función

$$x \in \mathbb{R}_+ \to \frac{\log N}{2x} - \frac{1}{2} + \frac{\log x}{2} - \frac{\log \log N}{2}.$$

Alcanza su mínimo en $x = \log N$ con valor igual a cero. Así, de nuevo por (3.11)

$$\inf_{m_0 \leqslant m \leqslant N} \rho(\mathcal{B}_{=m}^N) \geqslant \sqrt{\frac{\log N}{N}} \inf_{m_0 \leqslant m \leqslant N} \frac{1}{C_m} \geqslant \sqrt{\frac{\log N}{N}}(1-\epsilon), \tag{3.13}$$

para cada $N \in \mathbb{N}$. Combinando (3.12) y (3.13) concluimos que

$$\rho(\mathcal{B}_{hom}^N) \geqslant \inf_{0 \leqslant m \leqslant N} \rho(\mathcal{B}_{=m}^N) \geqslant \sqrt{\frac{\log N}{N}}(1-\epsilon), \qquad N \geqslant N_0,$$

y con esto probamos el corolario. □

3.3. Caso 3: Las funciones de grado d

La desigualdad de Cauchy permite extender el resultado (3.5) hasta el caso de grado d

$$\frac{1}{D}\left(\frac{d}{N}\right)^{\frac{d-1}{2d}} \leqslant K_N^{\leqslant d} \leqslant D\left(\frac{d}{N}\right)^{\frac{d-1}{2d}}, \tag{3.14}$$

donde la definición de $K_N^{\leqslant d}$ está dada en (1.7). Desafortunadamente, la información que tenemos para el cubo Booleano es menos precisa.

Teorema 3.3.1. *Sea $1 \leqslant d \leqslant N$. Entonces, existen constantes absolutas $c_d, C_d > 0$ que satisfacen*

$$c_d \frac{1}{N^{\frac{1}{2}}} \leqslant \rho(\mathcal{B}_{\leqslant d}^N) \leqslant C_d \frac{1}{N^{\frac{d-1}{2d}}}.$$

Señalamos que la acotación superior se conseguirá fácilmente por los resultados de la sección anterior, la estimación inferior requiere de algún trabajo preliminar que tiene también un interés independiente.

Teorema 3.3.2. *Para toda función $f : \{\pm 1\}^N \to \mathbb{R}$ y cada $x \in \{\pm 1\}^N$ tenemos que*

$$\left|\hat{f}(\varnothing) + \frac{1}{2}\sum_{S \neq \varnothing} \hat{f}(S) x^S\right| + \left|\frac{1}{2}\sum_{S \neq \varnothing} \hat{f}(S) x^S\right| \leqslant \|f\|_\infty.$$

Demostración. Fijo $x \in \{\pm 1\}^N$, tomemos $A \subset [N]$ y si denotamos por $\tilde{x}$ el elemento definido como $\tilde{x}_n = x_n$ si $n \notin A$ y $\tilde{x}_n = -x_n$ si $n \in A$; entonces tenemos que

$$|f(x)| = \left|\sum_{|S \cap A| \ par} \hat{f}(S) x^S + \sum_{|S \cap A| \ impar} \hat{f}(S) x^S\right|,$$

$$|f(\tilde{x})| = \left| \sum_{|S\cap A|\ par} \hat{f}(S)x^S - \sum_{|S\cap A|\ impar} \hat{f}(S)x^S \right|.$$

Ambas cosas juntas conducen a

$$\left| \sum_{|S\cap A|\ par} \hat{f}(S)x^S \right| + \left| \sum_{|S\cap A|\ impar} \hat{f}(S)x^S \right| \leqslant \|f\|_\infty.$$

Si sumamos sobre todos los $A \subset [N]$ y dividimos por 2^N, por desigualdad triangular tenemos que

$$\left| \frac{1}{2^N} \sum_{A\subset[N]} \sum_{|S\cap A|\ par} \hat{f}(S)x^S \right| + \left| \frac{1}{2^N} \sum_{A\subset[N]} \sum_{|S\cap A|\ impar} \hat{f}(S)x^S \right| \leqslant \|f\|_\infty. \tag{3.15}$$

Podemos reescribir la doble suma en una sola

$$\sum_{A\subset[N]} \sum_{|S\cap A|\ impar} \hat{f}(S)x^S = \sum_{S\neq\emptyset} \hat{f}(S)x^S |\{A \subset [N] : |A\cap S|\ impar\}|.$$

Para contar el número de subconjuntos $A \subset [N]$ tal que $|A \cap S|$ es impar, notar que tal conjunto es de la forma $A = A_1 \cup A_2$ donde $A_1 \subset [N]\backslash S$ y $A_2 \subset S$ satisface que $|A_2|$ es impar. Como el número de subconjuntos $A_2 \subset S \neq \emptyset$ con un número impar de elementos es precisamente $2^{|S|-1}$, se deduce que

$$|\{A \subset [N] : |A\cap S|\ impar\}| = 2^{N-|S|}2^{|S|-1} = 2^{N-1}.$$

Por lo tanto

$$\frac{1}{2^N} \sum_{A\subset[N]} \sum_{|S\cap\ A|\ impar} \hat{f}(S)x^S = \frac{1}{2} \sum_{S\neq\emptyset} \hat{f}(S)x^S. \tag{3.16}$$

Siguiendo la misma estrategia tenemos

$$\frac{1}{2^N} \sum_{A\subset[N]} \sum_{|S\cap A|\ par} \hat{f}(S)x^S = \hat{f}(\emptyset) + \frac{1}{2} \sum_{S\neq\emptyset} \hat{f}(S)x^S. \tag{3.17}$$

Reemplazando (3.16) y (3.17) en (3.15), obtenemos el resultado. □

Si tomamos esperanza en la desigualdad del Teorema 3.3.2 obtenemos el siguiente resultado.

Corolario 3.3.3. *Para toda función $f : \{\pm1\}^N \to [-1,1]$ tenemos que*

$$\mathbb{E} \left| \sum_{S\neq\emptyset} \hat{f}(S)x^S \right| \leqslant 2(1 - |\hat{f}(\emptyset)|).$$

Observemos que el corolario previo no puede mejorarse, en el sentido que no podemos encontrar $p > 1$ y $\gamma > 0$ tal que

$$\left(\mathbb{E}|\sum_{S\neq\varnothing}\hat{f}(S)x_S|^p\right)^{\frac{1}{p}} \leqslant \gamma(1-|\hat{f}(\varnothing)|), \tag{3.18}$$

para toda función $f : \{\pm 1\}^N \to \mathbb{R}$. En efecto, tomando $f : \{\pm 1\}^N \to \{\pm 1\}$ con $\sigma(f = 1) = \lambda$ y $\sigma(f = -1) = 1 - \lambda$ para algún $0 < \lambda < \frac{1}{2}$, entonces $\hat{f}(\varnothing) = 2\lambda - 1$ y si definimos

$$\theta(x) := f(x) - \hat{f}(\varnothing) = \begin{cases} 2-2\lambda & si \quad f(x) = 1, \\ -2\lambda & si \quad f(x) = -1. \end{cases}$$

Obtenemos

$$\mathbb{E}|\theta(x)|^p = (2-2\lambda)^p\lambda + (2\lambda)^p(1-\lambda).$$

Reemplazando este valor en (3.18) del lado izquierdo y del lado derecho conocemos exactamente $|\hat{f}(\varnothing)|$, conseguimos lo siguiente

$$(2-2\lambda)^p\lambda + (2\lambda)^p(1-\lambda) \leqslant \gamma^p(2\lambda)^p, \tag{3.19}$$

lo que lleva a

$$(1-\lambda)^p \leqslant (1-\lambda)^p + \lambda^{p-1}(1-\lambda) \leqslant \gamma^p(\lambda)^{p-1}.$$

Para N lo suficientemente grande, podemos tomar λ tan chico como se necesite para contradecir la desigualdad previa para un γ dado.

Observación 3.3.4. También se sigue de (3.19) que para el caso $p = 1$ la acotación $2 - 2\lambda \leqslant \gamma$ es válida, esto dice que la constante 2 en (3.3.3) es óptima.

Corolario 3.3.5. *Sea $f : \{\pm 1\}^N \to [-1, 1]$ una función de grado d. Entonces*

$$\left(\sum_{0\leqslant|S|\leqslant d}|\hat{f}(S)|^2\right)^{\frac{1}{2}} \leqslant 2e^d(1-|\hat{f}(\varnothing)|). \tag{3.20}$$

Demostración. Para demostrarlo hacemos uso de un caso particular de la desigualdad hipercontractiva, cuya demostración se realizara en otro capítulo (ver 4.2.29). Dicho resultado afirma que

$$\|f\|_2 \leqslant e^d\|f\|_1,$$

para toda función $f : \{\pm 1\}^N \to \mathbb{R}$ de grado d. Combinando esto con el Corolario 3.3.3 obtenemos la conclusión deseada. □

Estamos en condiciones de probar el Teorema 3.3.1

Demostración. La acotación superior es consecuencia de la desigualdad $\rho(\mathcal{B}^N_{\leqslant d}) \leqslant \rho(\mathcal{B}^N_{=d})$ y el Teorema 3.2.1. Para probar la acotación inferior, fijemos (normalizando) una función $f : \{\pm 1\}^N \to [-1,1]$ de grado d con $\|f\|_\infty = 1$. Tomando $\rho = \frac{C}{\sqrt{N}}$ donde $0 < C < 1$ es una constante, usando la desigualdad de Hölder junto con el Corolario 3.3.5 obtenemos

$$\begin{aligned}
\sum_{S\subset[N]} |\hat{f}(S)|\rho^{|S|} &\leqslant |\hat{f}(\varnothing)| + \left(\sum_{0\leqslant|S|\leqslant d} |\hat{f}(S)|^2\right)^{\frac{1}{2}} \left(\sum_{m=1}^{d} \rho^{2m}\binom{N}{m}\right)^{\frac{1}{2}} \\
&\leqslant |\hat{f}(\varnothing)| + 2e^d(1-|\hat{f}(\varnothing)|)\left(\sum_{m=1}^{d}(N\rho^2)^m\right)^{\frac{1}{2}} \\
&\leqslant |\hat{f}(\varnothing)| + 2e^d(1-|\hat{f}(\varnothing)|)\frac{C}{\sqrt{1-C^2}}.
\end{aligned}$$

Entonces podemos encontrar una constante $C = c_d$ tal que la suma previa sea menor o igual a uno. Esto produce que

$$\rho(\mathcal{B}^N_{\leqslant d}) \geqslant \frac{c_d}{\sqrt{N}}.$$

De esta manera tenemos el radio Booleano para funciones de grado menor o igual a d acotado inferiormente. $\square$

La relación entre la exponencial de N en la estimación superior e inferior de 3.3.1 es notoria. Algunos autores creen firmemente que $\frac{1}{N^{\frac{d-1}{2d}}}$ es el verdadero orden asintótico en vista del resultado para el caso d-homogéneo. En este sentido planteamos la siguiente pregunta.

Pregunta 3.3.6. Dado $d \in \mathbb{N}$, ¿existe alguna constante $C_d > 0$ tal que para toda función $f : \{\pm 1\}^N \to [-1,1]$ de grado d se tenga que

$$\left\| \sum_{S\neq 0} \hat{f}(S)x^S \right\|_\infty \leqslant C_d(1-|\hat{f}(\varnothing)|)? \tag{3.21}$$

Una respuesta positiva para la pregunta previa podría implicar el reemplazo de $N^{\frac{1}{2}}$ por $N^{\frac{d-1}{2d}}$ en la acotación inferior de 3.3.1. En efecto, podríamos refinar la segunda parte de la demostración del teorema usando la desigualdad

$$\sum_{0\leqslant|S|\leqslant d} |\hat{f}(S)|\rho^{|S|} \leqslant \left(\sum_{0<|S|\leqslant d} |\hat{f}(S)|^{\frac{2d}{d+1}}\right)^{\frac{d+1}{2d}} \left(\sum_{m=1}^{d} \rho^{m\frac{2d}{d-1}}\binom{N}{m}\right)^{\frac{d-1}{2d}},$$

entonces aplicando (3.7) para acotar el primer factor junto con (3.21) obtenemos lo querido. Notar que la pregunta 3.3.6 es equivalente a la siguiente:

Pregunta 3.3.7. Dado $d \in \mathbb{N}$, ¿existe alguna constante $C_d > 0$ tal que para toda función $f : \{\pm 1\}^N \to [-1, 1]$ de grado d y cada $1 \leqslant m \leqslant d$ se tenga

$$\left\| \sum_{|S|=m} \hat{f}(S) x^S \right\|_\infty \leqslant C_d(1 - |\hat{f}(\varnothing)|)?$$

La condición previa es muy similar a la afirmación del Teorema de Wiener para polinomios complejos en varias variables $Q(z) = \sum_{m=0}^{d} Q_m(z)$, donde Q_m denota la parte m-homogénea, el cual para cada $0 \leqslant m \leqslant d$

$$\|Q_m\|_\infty \leqslant 1 - |Q_0|^2 \leqslant 2(1 - |Q_0|);$$

véase (3.3) el caso de dimensión 1. Sin embargo, en el caso del cubo Booleano no podemos esperar que la constante C_d en la Pregunta 3.3.7, satisfaga $C_d \leqslant C$ para C constante absoluta. De lo contrario

$$\begin{aligned} \sum_{0<|S|\leqslant N} |\hat{f}(S)|\rho^{|S|} &\leqslant \sum_{m=1}^{N} \left(\sum_{|S|=m} |\hat{f}(S)|^{\frac{2m}{m+1}} \right)^{\frac{m+1}{2m}} \rho^m \binom{N}{m}^{\frac{m-1}{2m}} \\ &\leqslant C(1 - |\hat{f}(\varnothing)|) \sum_{m=1}^{N} \left(\frac{\rho\sqrt{N}\sqrt{e}}{\sqrt{m}} \right)^m, \end{aligned}$$

y entonces $\rho(\mathcal{B}^N) \geqslant \Omega(\frac{1}{\sqrt{N}})$. Pero por el Teorema 3.1.1 tenemos que $\rho(\mathcal{B}^N) \leqslant O(\frac{1}{N})$, una contradicción.

3.4. Caso 4: La clase $\mathcal{B}_\delta^N$

Para finalizar el capítulo, el resultado principal de esta sección es el análisis del radio Booleano para la familia $\mathcal{B}_\delta^N$, la cual está formada por funciones $f : \{\pm 1\}^N \to \mathbb{R}$ que cumplen la siguiente estimación:

$$|\mathbb{E}[f]| \leqslant (1 - \delta)\|f\|_\infty.$$

Sabemos que la desigualdad $|\mathbb{E}[f]| \leqslant \|f\|_\infty$ es cierta para toda $f : \{\pm 1\}^N \to \mathbb{R}$, entonces la idea de considerar esta familia es quedarnos con funciones que cumplan una estimación más ajustada entre la esperanza y la norma supremo (dependiendo de δ) y medir como cambia el radio Booleano para esta restricción.

Teorema 3.4.1. *Existe una constante absoluta $C > 0$ tal que para todo $N \in \mathbb{N}$ y para todo $\frac{1}{2^N} \leqslant \delta \leqslant 1$, tenemos*

$$\frac{C^{-1}}{\sqrt{N}\sqrt{\log(\frac{2}{\delta})}} \leqslant \rho(\mathcal{B}_\delta^N) \leqslant \frac{C}{\sqrt{N}\sqrt{\log(\frac{2}{\delta})}}. \tag{3.22}$$

La demostración del teorema se separa en dos partes. Primero abordaremos el lado izquierdo de la desigualdad, que se basa en el uso de la desigualdad hipercontractiva de Bonami. Recordar que para $-1 < \rho < 1$ el operador de ruido T_ρ asigna para cada función $f : \{\pm 1\}^N \to \mathbb{R}$ en el cubo Booleano otra función

$$T_\rho(f) : \{\pm 1\}^N \to \mathbb{R}, \qquad\qquad T_\rho f(x) = \sum_{S \subset [N]} \hat{f}(S)\rho^{|S|}x^S.$$

Como ya comentamos en el Capítulo 2, en el Lema de Bonami 4.2.17, este operador cumple una característica importante que es la siguiente:

Teorema 3.4.2. *Para todo $1 \leqslant p \leqslant q \leqslant \infty$*

$$\|T_\rho f\|_q \leqslant \|f\|_p \qquad\qquad \rho \leqslant \sqrt{\frac{p-1}{q-1}}.$$

Para la otra desigualdad del Teorema 3.4.1, daremos una estimación precisa de cierto tipo de funciones pertenecientes a $\mathcal{B}_\delta^N$.

3.4.1. Acotación inferior

Teorema 3.4.3. *Sea $f : \{\pm 1\}^N \to \mathbb{R}$ y $0 \leqslant \delta \leqslant 1$ con $|\mathbb{E}[f]| \leqslant (1-\delta)\|f\|_\infty$. Entonces*

$$\rho(f) \geqslant \frac{1}{5\sqrt{N}\sqrt{\log(\frac{2}{\delta})}}.$$

Demostración. Sin pérdida de generalidad podemos asumir que $f : \{\pm 1\}^N \to \mathbb{R}$ satisface $\|f\|_\infty = 1$ y $\mathbb{E}[f] = 1 - \delta$. Afirmamos que para cada $0 < \varepsilon < 1$ y $0 \leqslant m \leqslant N$ tenemos que

$$\left(\sum_{|S|=m} \left|\hat{f}(S)\right|^2 \right)^{\frac{1}{2}} \leqslant 2\varepsilon^{\frac{-m}{2}} \left(\frac{\delta}{2}\right)^{\frac{1}{1+\varepsilon}}. \tag{3.23}$$

Para ver esto, definimos $g : \{\pm 1\}^N \to [0, 1]$ como $g(x) = \frac{(1-f(x))}{2}$. Notar que

$$\hat{g}(S) = -\frac{1}{2}\hat{f}(S), \qquad\qquad \varnothing \neq S \subset [N]. \tag{3.24}$$

Más aún, para todo $p \geqslant 1$,

$$(\mathbb{E}[|g(x)|^p])^{\frac{1}{p}} \leqslant (\mathbb{E}[g(x)])^{\frac{1}{p}} = \left(\frac{\delta}{2}\right)^{\frac{1}{p}}. \tag{3.25}$$

Combinando la relación entre los coeficientes de Fourier (3.24), el teorema hipercontractivo 2.7 (con $p = 1 + \varepsilon$, $q = 2$, $\rho = \sqrt{\varepsilon}$) y (3.25), obtenemos que

$$\begin{aligned}
\sum_{|S|=m} \left|\hat{f}(S)\right|^2 = 4 \sum_{|S|=m} |\hat{g}(S)|^2 &\leqslant \frac{4}{\varepsilon^m} \sum_{S \subset [N]} |\hat{g}(S)|^2 \varepsilon^{|S|} \\
&\leqslant \frac{4}{\varepsilon^m} (\mathbb{E}|g(x)|^p)^{\frac{2}{p}} \leqslant \frac{4}{\varepsilon^m} \left(\frac{\delta}{2}\right)^{\frac{2}{1+\varepsilon}},
\end{aligned}$$

lo que prueba la afirmación. Aplicando (3.23) para $0 < \rho$ y $0 < \varepsilon < 1$, conseguimos lo siguiente

$$\begin{aligned}
\sum_{S\neq\varnothing} |\hat{f}(S)|\rho^{|S|} &= \sum_{m=1}^{N} \rho^m \sum_{|S|=m} |\hat{f}(S)| \leqslant \sum_{m=1}^{N} \rho^m \sqrt{\binom{N}{m}} (\sum_{|s|=m}^{N} |\hat{f}(S)|^2)^{\frac{1}{2}} \\
&\leqslant \sum_{m=1}^{N} \rho^m \sqrt{\binom{N}{m}} 2\varepsilon^{\frac{-m}{2}} \left(\frac{\delta}{2}\right)^{\frac{1}{1+\varepsilon}} \\
&\leqslant \sum_{m=1}^{N} \rho^m e^{\frac{m}{2}} \frac{N^{\frac{m}{2}}}{m^{\frac{m}{2}}} 2\varepsilon^{\frac{-m}{2}} \left(\frac{\delta}{2}\right)^{\frac{1}{1+\varepsilon}} = 2\left(\frac{\delta}{2}\right)^{\frac{1}{1+\varepsilon}} \sum_{m=1}^{N} \left(\frac{\rho\sqrt{e}\sqrt{N}}{\sqrt{\varepsilon}\sqrt{m}}\right)^m .
\end{aligned}$$

Tomando

$$\varepsilon = \frac{\log 2}{\log(\frac{2}{\delta})} \qquad y \qquad \rho = \frac{1}{5\sqrt{N}\sqrt{\log(\frac{2}{\delta})}},$$

deducimos que

$$\left(\frac{\delta}{2}\right)^{\frac{1}{1+\varepsilon}} = \frac{\delta}{2}\exp\left(\frac{\varepsilon}{1+\varepsilon}\log\frac{2}{\delta}\right) = \frac{\delta}{2}\exp\left(\frac{\log(\frac{2}{\delta})\log 2}{\log(\frac{2}{\delta}) + \log 2}\right) \leqslant \delta.$$

Usando ahora que $\hat{f}(\varnothing) = \mathbb{E}[f] = 1-\delta$, la estimación previa permite concluir

$$\sum_{S\neq\varnothing} |\hat{f}(S)|\rho^{|S|} \leqslant 2\delta \sum_{m=1}^{N} \left(\frac{\sqrt{e}}{5\sqrt{m}\sqrt{\log 2}}\right)^m \leqslant \delta = 1 - |\hat{f}(\varnothing)|.$$

Y así pudimos acotar inferiormente el radio Booleano de cada función perteneciente a $\mathcal{B}_\delta^N$, con δ positivo y menor que uno. □

3.4.2. Funciones de tipo threshold

Mostraremos que la desigualdad en el Teorema 3.4.3, es óptima. Para eso daremos estimaciones precisas del radio Booleano para las siguientes funciones dos-paramétricas:

$$\psi_{N,\alpha} : \{\pm 1\}^N \to \{\pm 1\}, \qquad \psi_{N,\alpha}(x) = sign(x_1 + ... + x_N - \alpha),$$

donde $\alpha \in \mathbb{R}$, $N \in \mathbb{N}$ y como es usual $sign(y) = 1$ para $y \geqslant 0$ y -1 para $y < 0$. Observemos que diferentes valores de α puede generar la misma función. Nos restringimos al caso en que $\alpha \geqslant 0$. Para algunos resultados además asumimos que $N - \alpha$ es un número natural impar. De esta manera, garantizamos que diferentes valores de α conduce a funciones diferentes y también evitamos que aparezca $sign(0)$. Señalamos que las funciones $\psi_{N,\alpha}$ pertenece a la amplia clase de funciones conocidas como funciones de tipo threshold o umbral (donde se detalla mayor información en el capítulo 5 de [O'D14]). El objetivo es mostrar el siguiente resultado.

Teorema 3.4.4. *Para cada $N \in \mathbb{N}$ y $0 \leqslant \alpha < N$ tenemos que*

$$C^{-1}\frac{1}{\alpha+\sqrt{N}} \leqslant \rho(\psi_{N,\alpha}) \leqslant C\frac{1}{\alpha+\sqrt{N}},$$

donde C es una constante absoluta independiente de α y N.

La demostración involucra una serie de resultados preliminares. Comencemos con una estimación técnica para el radio Booleano con algunas suposiciones en α. Para $N \in \mathbb{N}$ con $0 \leqslant \alpha < N$, sea $G : \mathbb{R}_+ \to \mathbb{R}_+$ la función dada por

$$G(r) := \sup_{z\in\mathbb{T}} |1+zr|^{\frac{N+\alpha-1}{2}} |1-zr|^{\frac{N-\alpha-1}{2}}, \qquad r \in \mathbb{R}+, \tag{3.26}$$

y usemos esto para definir la función $I : \mathbb{R}_+ \to \mathbb{R}_+$

$$I(\rho) = \int_0^\rho G(r)dr, \;\; \rho \in \mathbb{R}_+. \tag{3.27}$$

Antes de comenzar con los resultados preliminares requeridos definimos, siguiendo el trabajo de Guadarrama [Gua05], el radio de Bohr de la clase P_N de todos los polinomio complejo de grado a lo sumo N por

$$R_N := \sup\left\{ r > 0 : \sum_{k=0}^N |c_k(p)| r^k \leqslant \|p\|_{\mathbb{D}}, \quad p(z) = \sum_{k=0}^N c_k(p) z^k, \;\; z \in \mathbb{D} \right\}. \tag{3.28}$$

En dicho trabajo se prueba que existen constantes absolutas $c_1, c_2 > 0$ tal que para N suficientemente grande,

$$\frac{1}{3} + \frac{c_1}{3^{\frac{N}{2}}} < R_N < \frac{1}{3} + c_2\frac{\log N}{N}.$$

Dryanov y Fournier [DF02] prueban la siguiente estimación: para todo polinomio $p \in P_N$ $p(z) = \sum_{k=0}^N c_k z^k$ tenemos

$$|c_k| \leqslant 2\cos\left(\frac{\pi}{[\frac{N}{k}]+2}\right)(\|p\|_{\mathbb{D}} - |c_0|), \qquad 1 \leqslant k \leqslant N,$$

donde $[\frac{N}{k}]$ denota la parte entera de $\frac{N}{k}$. Esta estimación implica que la única raíz t_N en (0,1] de la ecuación

$$\sum_{k=1}^N \cos\left(\frac{\pi}{[\frac{N}{k}]+2}\right) t^k = \frac{1}{2},$$

satisface $0 < t_N \leqslant R_N$. Un comentario adicional de la observarción, se puede ver en el artículo de Fournier [Fou08], donde entre otras cosas se muestra que de hecho $t_N < R_N$ para todo N.

Teorema 3.4.5. *Para $N \in \mathbb{N}$ y $0 \leqslant \alpha < N$ tal que $N-\alpha$ es impar, sea $\rho = \rho_{\alpha,N}$ el radio Booleano de $\psi = \psi_{\alpha,N}$. Entonces*

$$I(\rho) \leqslant \frac{1}{N\binom{N-1}{\frac{N-\alpha-1}{2}}} \sum_{n \leqslant \frac{N-\alpha-1}{2}} \binom{N}{n} \leqslant R_N I\left(\frac{\rho}{R_N}\right),$$

donde R_N es el radio de Bohr definido en (3.28).

Demostración. Comenzamos con la siguiente afirmación, para cada número complejo z obtenemos

$$\sum_{n=1}^{N} \binom{N-1}{n-1} \hat{\psi}([n]) z^{n-1} = \binom{N-1}{\frac{N-\alpha-1}{2}} \frac{1}{2^{N-1}} (1+z)^{\frac{N+\alpha-1}{2}} (1-z)^{\frac{N-\alpha-1}{2}}. \tag{3.29}$$

Para probar esto, hacemos uso de la N-ésima derivada discreta de una función en el cubo Booleano, aplicado a ψ da una función $D_N\psi : \{-1,1\}^{N-1} \to \mathbb{R}$ que definimos como

$$D_N\psi(x_1, ..., x_{N-1}) := \frac{\psi(x_1, ..., x_{N-1}, 1) - \psi(x_1, ..., x_{N-1}, -1)}{2} = \begin{cases} 1 & \text{si } x_1 + ... + x_{N-1} = \alpha \\ 0 & \text{si no.} \end{cases}$$

La expansión de Fourier de $D_N\psi$ podemos calcularla explícitamente en término de los coeficientes de ψ (véase [O'D14, p. 30, Prop. 2.9] para detalles de la demostración). Usando más que nunca que ψ es invariante por permutación de sus coordenadas, para cada $\varnothing \neq S \subset [N]$ satisface $\hat{\psi}(S) = \hat{\psi}([n])$ donde $|S| = n$, además obtenemos que

$$D_N\psi(x) = \sum_{N \in S \subset [N]} \hat{\psi}(S) x^{S \setminus \{N\}} = \sum_{n=1}^{N} \hat{\psi}([n]) \sum_{N \in S \subset [N], |S|=n} x^{S \setminus \{N\}}. \tag{3.30}$$

Por lo tanto, aplicando el operador de ruido T_r $(-1 < r < 1)$ a $D_N\psi$ obtenemos que

$$T_r D_N \psi(1, ..., 1) = \sum_{n=1}^{N} \binom{N-1}{n-1} \hat{\psi}([n]) r^{n-1}. \tag{3.31}$$

Por otra parte, $T_r D_N \psi$ tiene la siguiente definición probabilística (ver [O'D14, sección 2.4]), si $\boldsymbol{x}_i$ son variables aleatorias independientes a valores 1 y -1 con probabilidad $\frac{1}{2} + \frac{1}{2}r$ y $\frac{1}{2} - \frac{1}{2}r$ respectivamente, entonces

$$\begin{aligned} T_r D_N \psi(1, ..., 1) &= \mathbb{E}[D_N \psi(\boldsymbol{x}_1, ..., \boldsymbol{x}_{N-1})] = \mathbb{P}(\boldsymbol{x}_1 + ... + \boldsymbol{x}_{N-1} = \alpha) && (3.32) \\ &= \binom{N-1}{\frac{N-\alpha-1}{2}} \frac{1}{2^{N-1}} (1+r)^{\frac{N+\alpha-1}{2}} (1-r)^{\frac{N-\alpha-1}{2}}. && (3.33) \end{aligned}$$

Comparando (3.31) y (3.32) llegamos a la conclusión de que (3.29) es cierta para polinomios complejos con $r \in (-1, 1)$. Se sigue de (3.29) que

$$\sup_{z\in\mathbb{T}}\left|\sum_{n=1}^{N}\binom{N-1}{n-1}\hat{\psi}([n])(rz)^{n-1}\right| = \binom{N-1}{\frac{N-\alpha-1}{2}}\frac{G(r)}{2^{N-1}}, \qquad r\geqslant 0, \tag{3.34}$$

donde G está definida en (3.26). Usando ahora la definición de R_N, deducimos que para todo $r>0$

$$\sum_{n=1}^{N}\binom{N-1}{n-1}|\hat{\psi}([n])|(rR_N)^{n-1} \leqslant \binom{N-1}{\frac{N-\alpha-1}{2}}\frac{G(r)}{2^{N-1}} \leqslant \sum_{n=1}^{N}\binom{N-1}{n-1}|\hat{\psi}([n])|r^{n-1}.$$

Integrando la desigualdad previa sobre el intervalo $[0,R]$ con $R>0$, produce

$$\frac{1}{NR_N}\sum_{n=1}^{N}\binom{N}{n}|\hat{\psi}([n])|(RR_N)^{n} \leqslant \binom{N-1}{\frac{N-\alpha-1}{2}}\frac{I(R)}{2^{N-1}} \leqslant \frac{1}{N}\sum_{m=1}^{N}\binom{N}{n}|\hat{\psi}([n])|R^{n}. \tag{3.35}$$

Reemplazando en la expresión anterior la variable R primero por $\frac{\rho}{R_N}$ y luego por ρ, para conseguir respectivamente, la segunda y luego la primera desigualdad de las siguiente expresión

$$\frac{N}{2^{N-1}}\binom{N-1}{\frac{N-\alpha-1}{2}}I(\rho) \leqslant \sum_{n=1}^{N}\binom{N}{n}|\hat{\psi}([n])|\rho^{n} \leqslant \frac{N}{2^{N-1}}\binom{N-1}{\frac{N-\alpha-1}{2}}R_N I(\frac{\rho}{R_N}). \tag{3.36}$$

Dividiendo por $\frac{1}{2^{N-1}}\binom{N-1}{\frac{N-\alpha-1}{2}}$ tenemos la desigualdad del enunciado casi lista, sólo queda estimar la parte del medio.

Notar que por la buena definición del radio Booleano, el número ρ satisface

$$\sum_{n=1}^{N}\binom{N}{n}|\hat{\psi}([n])|\rho^{n} = 1-|\hat{\psi}(\varnothing)|. \tag{3.37}$$

La esperanza de ψ está dada por $\hat{\psi}(\varnothing)$, podemos fácilmente calcularla como

$$\begin{aligned}\hat{\psi}(\varnothing) &= \sigma(x_1+...+x_N>\alpha)-\sigma(x_1+...+x_N<\alpha)\\ &= 2\sigma(x_1+...+x_N>\alpha)-1,\end{aligned}$$

usando esto último en (3.37) escribimos

$$\sum_{n=1}^{N}\binom{N}{n}|\hat{\psi}([n])|\rho^{n} = 2\sigma(x_1+...+x_N>\alpha). \tag{3.38}$$

Observar que $x_1+...+x_N$ depende de la cantidad de $1's$ y $-1's$: si m variables son iguales a -1 y $N-m$ son iguales a 1 entonces $x_1+...+x_N=N-2m$, luego

$$\sigma(x_1+...+x_N>\alpha) = \frac{1}{2^N}\sum_{N-2m>\alpha}\binom{N}{m},$$

y por lo tanto de (3.38) tenemos que

$$\sum_{n=1}^{N} \binom{N}{n} |\hat{\psi}([n])| \rho^n = \frac{1}{2^{N-1}} \sum_{m \leqslant \frac{N-\alpha-1}{2}} \binom{N}{m}.$$

Aplicando esto último en la desigualdad (3.36), concluimos que la afirmación del Teorema 3.4.5 es válida. □

Como consecuencia inmediata del Teorema 3.4.5 tenemos el siguiente corolario.

Corolario 3.4.6. *Para $N \in \mathbb{N}$ y $0 \leqslant \alpha < N$ tal que $N - \alpha$ es impar, sea $\rho = \rho_{\alpha,N}$ es el radio Booleano de $\psi = \psi_{\alpha,N}$. Entonces*

$$I(\rho) \leqslant \frac{1}{N\binom{N-1}{\frac{N-\alpha-1}{2}}} \sum_{n \leqslant \frac{N-\alpha-1}{2}} \binom{N}{n} \leqslant \frac{1}{3} I(3\rho). \tag{3.39}$$

Demostración. Como $\frac{1}{3} \leqslant R_N$ y G es una función estrictamente creciente en $\mathbb{R}_+$, la función $r \to \frac{I(r)}{r}$ es creciente en $(0, \infty)$. En consecuencia

$$R_N I\left(\frac{\rho}{R_N}\right) \leqslant \frac{1}{3} I(3\rho),$$

y entonces aplicamos el Teorema 3.4.5. □

Comencemos el análisis del Corolario 3.4.6. Para controlar el orden de los números combinatorios en 3.4.6 necesitamos un resultado auxiliar [McK89], y para presentar este resultado introducimos la función $Y : [0, \infty) \to [0, \infty)$ definida por

$$Y(x) = e^{\frac{1}{2}x^2} \int_x^\infty e^{\frac{-1}{2}t^2} dt, \qquad x \geqslant 0. \tag{3.40}$$

Lema 3.4.7. *Para cada $N \in \mathbb{N}$ y para $0 \leqslant \alpha < N$ con $N - \alpha$ siendo un entero impar, existe un número real $0 \leqslant c_{\alpha,N} \leqslant \sqrt{\frac{\pi}{2}}$ tal que*

$$\sum_{n \leqslant \frac{N-\alpha-1}{2}} \binom{N}{n} = \sqrt{N} \binom{N-1}{\frac{N-\alpha-1}{2}} Y\left(\frac{\alpha+1}{\sqrt{N}}\right) \exp\left(\frac{c_{\alpha,N}}{\sqrt{N}}\right).$$

Demostración. El siguiente resultado se encuentra con detalle en [McK89], tomando $N \geqslant 1$ y $\frac{N}{2} \leqslant k \leqslant N$ tenemos

$$\sum_{j=k}^{N} \binom{N}{j} = \sqrt{N} \binom{N-1}{N-k} Y\left(\frac{2k-N}{\sqrt{N}}\right) \exp\left(\frac{E(k,N)}{\sqrt{N}}\right), \tag{3.41}$$

para algún número real

$$0 \leqslant E(k,N) \leqslant \text{mín}\left(\sqrt{\frac{\pi}{2}}, \frac{2\sqrt{N}}{2k-N}\right).$$

Usando simples propiedades de los coeficientes binomiales y un cambio de variables, tenemos

$$\sum_{n \leqslant \frac{N-\alpha-1}{2}} \binom{N}{n} = \sum_{n \leqslant \frac{N-\alpha-1}{2}} \binom{N}{N-n} = \sum_{n = \frac{N+\alpha+1}{2}} \binom{N}{n}.$$

Por lo tanto, reemplazando $k = \frac{N+\alpha+1}{2}$ en (3.41) obtenemos el resultado. □

Finalmente necesitamos otro resultado técnico sobre las funciones G e I que se definieron en (3.26) y (3.27).

Lema 3.4.8. *Para* $0 \leqslant \alpha \leqslant N-1$ *tenemos*

$$G(r) = \begin{cases} (1+r)^{\frac{N+\alpha-1}{2}}(1-r)^{\frac{N-\alpha-1}{2}} & si \quad r \in [0, r_\alpha], \\ (1+r^2)^{\frac{N-1}{2}}(1+\frac{\alpha}{N-1})^{\frac{N+\alpha-1}{4}}(1-\frac{\alpha}{N-1})^{\frac{N-\alpha-1}{4}} & si \quad r \in [r_\alpha, 1], \end{cases}$$

donde

$$r_\alpha := \frac{\alpha}{(N-1)+\sqrt{(N-1)^2-\alpha^2}}.$$

Demostración. Tomando $z = c + it \in \mathbb{T}$ con $c^2 + t^2 = 1$ podemos escribir

$$G(r) = \sup_{t \in [-1,1]} (1+r^2+2rt)^{\frac{N+\alpha-1}{4}}(1+r^2-2rt)^{\frac{N-\alpha-1}{4}}, \qquad r \in [0,1].$$

Para mostrar la fórmula requerida para G, definimos dado un $r \in [0,1]$ la función $g_r : [-1,1] \to \mathbb{R}+$ por

$$g_r(t) = (1+r^2+2rt)^{\frac{N+\alpha-1}{4}}(1+r^2-2rt)^{\frac{N-\alpha-1}{4}}, \qquad t \in [-1,1]. \tag{3.42}$$

El cálculo de la primera derivada muestra que el signo de g_r' depende solamente del signo de $t_{\alpha,r} - t$, pues

$$g_r'(t) = r^2(N-1)(1+r^2+2rt)^{\frac{N+\alpha-1}{4}-1}(1+r^2-2rt)^{\frac{N-\alpha-1}{4}-1}(t_{\alpha,r}-t), \qquad t \in (-1,1), \tag{3.43}$$

donde

$$t_{\alpha,r} := \frac{\alpha(1+r^2)}{2r(N-1)}.$$

Ahora observermos que $t_{\alpha,r} \leqslant 1$ si y sólos si $r \in [r_\alpha, 1]$, y que g_r alcanza el máximo para $t_{\alpha,r}$. Así

$$G(r) = g_r(t_{\alpha,r}) = (1+r^2)^{\frac{N-1}{2}}\left(1+\frac{\alpha}{N-1}\right)^{\frac{N+\alpha-1}{4}}\left(1-\frac{\alpha}{N-1}\right)^{\frac{N-\alpha-1}{4}}.$$

Por otro lado cuando $r \in [0, r_\alpha]$, se tiene que g_r es creciente en [-1,1] y

$$G(r) = g_r(1) = (1+r)^{\frac{N+\alpha-1}{2}}(1-r)^{\frac{N-\alpha-1}{2}},$$

y esto completa la demostración. $\square$

Combinamos los precedentes dos lemas para explotar la estimación explícita de $\rho(\psi_{\alpha,N})$ dada en el Teorema 3.4.5.

Proposición 3.4.9. *Existe una constante absoluta $C > 0$ tal que para cada $\alpha, N \in \mathbb{N}$ con $N - \alpha$ entero impar tenemos*

$$Ce^{\frac{2}{\sqrt{N}}}\frac{1}{\sqrt{N}}Y\left(\frac{\alpha+1}{\sqrt{N}}\right) \leqslant \rho(\psi_{\alpha,N}) \leqslant e^{\frac{2}{\sqrt{N}}}\frac{1}{\sqrt{N}}Y\left(\frac{\alpha+1}{\sqrt{N}}\right).$$

Demostración. Aplicamos el Teorema 3.4.5. Si notamos $\rho = \rho(\psi_{\alpha,N})$ para simplificar, entonces combinando el Corolario 3.4.6 y el Lema 3.4.7 obtenemos

$$e^{\frac{-2}{\sqrt{N}}}I(\rho) \leqslant \frac{1}{\sqrt{N}}Y\left(\frac{\alpha+1}{\sqrt{N}}\right) \leqslant \frac{1}{3}I(3\rho). \tag{3.44}$$

Ahora observemos que por la definición dada en (3.40), Y es una función decreciente en $\mathbb{R}_+$. Como la función G definida en (3.26) satisface $G(r) \geqslant 1$ para todo $r \geqslant 0$, esto se obtiene reemplazando z por i en la expresión del supremo en la definición de $G(r)$, lo cual cede la parte derecha de la estimación requerida:

$$\rho \leqslant \int_0^\rho G(r)dr = I(\rho) \leqslant e^{\frac{2}{\sqrt{N}}}\frac{1}{\sqrt{N}}Y\left(\frac{\alpha+1}{\sqrt{N}}\right). \tag{3.45}$$

Para la acotación inferior de ρ, notemos que, como G es una función creciente en [0,1] (esto puede deducirse por el Teorema de módulo máximo) tenemos que

$$\frac{1}{3}I(3\rho) = \frac{1}{3}\int_0^{3\rho} G(r)dr \leqslant \rho G(3\rho). \tag{3.46}$$

Ahora acotemos $G(3\rho)$ por una constante independiente de N y α. Para poder hace uso del Lema 3.4.8, donde una explícita expresión para $G(r)$ es presentada, antes distinguimos dos casos para $r \in [0,1]$ de acuerdo a si $r = 3\rho \leqslant r_\alpha$ o $r = 3\rho \geqslant r_\alpha$. Asumimos que $3\rho \leqslant r_\alpha$ y aplicamos la caracterización de G, vista en el Lema 3.4.8 para deducir

$$G(3\rho) = (1+3\rho)^\alpha(1-9\rho^2)^{\frac{N-\alpha-1}{2}} \leqslant (1+3\rho)^\alpha \leqslant \exp(3\rho\alpha).$$

La fórmula anterior para Y en (3.45) muestra junto al hecho de que $Y(x) \leqslant \frac{1}{x}$ para todo $x > 0$, la siguiente acotación

$$\rho \leqslant e^2\frac{1}{\sqrt{N}}Y\left(\frac{\alpha+1}{\sqrt{N}}\right) \leqslant \frac{e^2}{\alpha+1}.$$

Usando esto para terminar de acortar $G(3\rho)$, tenemos

$$G(3\rho) \leqslant \exp(3e^2),$$

en el caso que $3\rho \leqslant r_\alpha$.

Consideramos ahora el caso $3\rho \geqslant r_\alpha$. Como Y es una función decreciente en $[0,\infty)$, concluimos

$$\rho \leqslant \frac{e^2}{\sqrt{N}} Y\left(\frac{\alpha+1}{\sqrt{N}}\right) \leqslant \frac{e^2}{\sqrt{N}} Y(0) = \frac{C_0}{\sqrt{N}},$$

donde $C_0 = e^2\sqrt{\frac{\pi}{2}}$. Combinando la fórmula para $r_\alpha := \frac{\alpha}{(N-1)+\sqrt{(N-1)^2-\alpha^2}}$ junto con el Lema 3.4.8 obtenemos

$$\alpha \leqslant 2(N-1)r_\alpha \leqslant 6\rho(N-1) \leqslant 6(N-1)\frac{C_0}{\sqrt{N}} \leqslant 6C_0\sqrt{N-1}.$$

Aplicando esta última estimación y el Lema 3.4.8 podemos afirmar

$$\begin{aligned} G(3\rho) &\leqslant \left(1+\frac{\alpha}{N-1}\right)^{\frac{\alpha}{2}} (1+9\rho^2)^{\frac{N-1}{2}} \leqslant \exp\left(\frac{\alpha^2}{2(N-1)} + 9\rho^2\frac{N-1}{2}\right) \\ &\leqslant \exp\left(\frac{36C_0^2}{2} + \frac{9C_0^2}{2}\right) \leqslant \exp(23C_0^2). \end{aligned}$$

Combinando el Corolario 3.4.6, (3.44) y (3.46) se produce la estimación deseada. □

Centremos el foco ahora en la demostración en el Teorema 3.4.4. Usando de nuevo el hecho que Y es una función decreciente e $Y(x) \leqslant \frac{1}{x}$ en $\mathbb{R}_+$, puede probarse además que (véase [McK89])

$$\frac{x}{1+x^2} \leqslant Y(x) \leqslant \frac{1}{x}, \qquad x > 0. \tag{3.47}$$

Demostración. (del Teorema 3.4.4). Sea $N \in \mathbb{N}$ y $0 \leqslant \alpha < N$. Comenzamos observando de la monotonía de Y, para todo $\alpha \leqslant \sqrt{N}$ la acotación

$$\frac{Y(2)}{\sqrt{N}} \leqslant \frac{1}{\sqrt{N}} Y\left(\frac{\alpha+1}{\sqrt{N}}\right) \leqslant \frac{Y(0)}{\sqrt{N}};$$

mientras para $\alpha \geqslant \sqrt{N}$ usando la estimación (3.47) deducimos

$$\frac{1}{\sqrt{N}+\alpha+1} \leqslant \frac{\alpha+1}{N+(\alpha+1)^2} \leqslant \frac{1}{\sqrt{N}} Y\left(\frac{\alpha+1}{\sqrt{N}}\right) \leqslant \frac{1}{\alpha+1} \leqslant \frac{2}{\alpha+\sqrt{N}}.$$

Podemos resumir las desigualdades anteriores diciendo que existe una constante absoluta $C_0 > 0$ (independiente de α y N) que satisface

$$\frac{C_0^{-1}}{\alpha+\sqrt{N}} \leqslant \frac{1}{\sqrt{N}} Y\left(\frac{\alpha+1}{\sqrt{N}}\right) \leqslant \frac{C_0}{\alpha+\sqrt{N}}. \tag{3.48}$$

Dividimos ahora la demostración del teorema en dos casos.

Caso 1: Asumimos que $N-\alpha$ es un número natural impar. Entonces por (3.47) y la Proposición (3.4.9), existe una constante absoluta $C_0>0$ tal que

$$C_0^{-1}\frac{1}{\alpha+\sqrt{N}} \leqslant \rho(\psi_{\alpha,N}) \leqslant C_0\frac{1}{\alpha+\sqrt{N}}. \tag{3.49}$$

Caso 2: Sin ninguna reestricción en $0\leqslant\alpha<N$, sea $0\leqslant n<\frac{N}{2}$ un número entero tal que $\alpha\in[N-2n-2,N-2n)$. Notar que para cada $x\in\{\pm1\}^N$ la suma $x_1+...+x_N$ es igual a $N-2k$ para algún entero k. Por lo tanto, si $\alpha'\in[N-2n-2,N-2n)$ entonces $\psi_{N,\alpha}=\psi_{N,\alpha'}$. En particular, esto es cierto para $\alpha'=N-2n-1$, luego $N-\alpha'$ es un número natural impar. Podemos aplicar entonces (3.49) para obtener que

$$C_0^{-1}\frac{1}{\alpha'+\sqrt{N}} \leqslant \rho(\psi_{\alpha,N})=\rho(\psi_{\alpha',N}) \leqslant C_0\frac{1}{\alpha'+\sqrt{N}}.$$

Finalmente usando que $|\alpha-\alpha'|\leqslant 1$ concluimos el resultado. □

El argumento en la Proposición (3.4.9) puede realizarse para las función Mayoritaria Maj_N (N impar) y así dar la estimación precisa del radio Booleano cuando N tiende a infinito.

Corolario 3.4.10. *El radio Booleano de la función Mayoritaria satisface*

$$\rho(Maj_N)=\frac{\gamma}{\sqrt{N}}(1+o(1)),$$

donde $\gamma>0$ es el único número real satisfaciendo $\int_0^\gamma e^{\frac{u^2}{2}}du=\sqrt{\frac{\pi}{2}}$.

Demostración. En el caso de la función Mayoritaria $Maj_N=\psi_{N,0}$ donde N es impar, reemplazando z por ri en (3.29) y comparando el coeficiente de r^m ($m\in\mathbb{N}$) en ambos lados tenemos

$$\sum_{n=1}^{N}\binom{N-1}{n-1}|\hat{\psi}_N([n])|r^{n-1}=\binom{N-1}{\frac{N-1}{2}}\frac{1}{2^{N-1}}(1+r^2)^{\frac{N-1}{2}}.$$

Notamos por ρ_N el radio Booleano de ψ_N. Entonces, integrando entre 0 y ρ_N en la expresión previa y usando que $\hat{\psi}_N(\varnothing)=0$, deducimos que

$$\frac{1}{N}=\frac{1}{N}\sum_{m=1}^{N}\binom{N}{n}|\hat{\psi}([n])|\rho_N^n=\binom{N-1}{\frac{N-1}{2}}\frac{1}{2^{N-1}}\int_0^{\rho_N}(1+r^2)^{\frac{N-1}{2}}dr,$$

donde la primera igualdad es justamente la definición del radio Booleano. Usando la fórmula de aproximación de Stirling y haciendo el cambio de variable $r=\frac{u}{\sqrt{N}}$, obtenemos que si $C_N:=\frac{\rho_N\sqrt{N}}{\gamma}$ (donde γ es como antes) entonces

$$\int_0^{C_N\gamma}\left(1+\frac{u^2}{N}\right)^{\frac{N-1}{2}}du=\frac{1}{\sqrt{N}}\frac{2^{N-1}}{\binom{N-1}{\frac{N-1}{2}}}=\sqrt{\frac{\pi}{2}}(1+o(1)).$$

Por lo tanto C_N está acotado. Usando la desigualdad $(1+\frac{1}{x})^x \leqslant e \leqslant (1+\frac{1}{x})^{x+1}$ válida para $x>0$ obtenemos que

$$\int_0^{C_N\gamma} e^{\frac{u^2}{2}}du = \sqrt{\frac{\pi}{2}}(1+o(1)).$$

Por la definición de γ se sigue que

$$\int_\gamma^{C_N\gamma} e^{\frac{u^2}{2}}du = o(1).$$

Por último aplicando el Teorema de Valor Medio, conseguimos $C_N = 1 + o(1)$ como queríamos probar. $\square$

3.4.3. Acotación superior

Finalmente, volvemos a la demostración del Teorema 3.4.1, donde necesitamos un resultado auxiliar

Lema 3.4.11. *Para todo $0 \leqslant \alpha < N$ tenemos que*

$$\sigma(x_1 + ... + x_N > \alpha) \geqslant \exp\left[-6(1+\frac{\alpha}{\sqrt{N}})^2\right]. \tag{3.50}$$

Demostración. Hagamos la distinción de dos casos. Si $\alpha \geqslant \frac{N}{2}$, entonces la acotación es trivial

$$\sigma(x_1 + ... + x_N > \alpha) \geqslant 2^{-N} \geqslant \exp(-N) \geqslant \exp\left[-6(1+\frac{\alpha}{\sqrt{N}})^2\right].$$

En cambio si asumimos que $\alpha < \frac{N}{2}$. Por el Lema 3.4.7 tenemos que

$$\sigma(x_1 + ... + x_N > \alpha) = \frac{1}{2^N}\sum_{n \geqslant \frac{N-\alpha-1}{2}} \binom{N}{n} \geqslant \frac{\sqrt{N}}{2^N}\binom{N-1}{\frac{N-\alpha-1}{2}} Y\left(\frac{\alpha+1}{\sqrt{N}}\right).$$

Para acotar el número combinatorio en la expresión anterior recordemos algunas acotaciones establecidas para coeficientes binomiales. Por [LPV03, pág 58] podemos conseguir para cada $0 \leqslant m \leqslant n$

$$\binom{2n}{n-m} \geqslant \binom{2n}{n}\exp\left(\frac{-m^2}{n-m+1}\right) \geqslant \frac{2^{2n-1}}{\sqrt{n}}\exp\left(\frac{-m^2}{n-m+1}\right),$$

donde la desigualdad anterior es consecuencia de la estimación de Stirling. Esto dice que si N es un número impar y α es como antes, entonces

$$\binom{N-1}{\frac{N-\alpha-1}{2}} \geqslant \frac{\sqrt{2}}{4}\frac{2^N}{\sqrt{N}}\exp\left(\frac{-\alpha^2}{2(N-\alpha+1)}\right) \geqslant \frac{\sqrt{2}}{4}\frac{2^N}{\sqrt{N}}\exp\left(-\frac{\alpha^2}{N}\right),$$

donde la desigualdad anterior sigue por la hipótesis $\alpha \leqslant \frac{N}{2}$. Si N es par, usando la relación estándar entre los coeficientes binomiales y la previa acotación tenemos

$$\binom{N-1}{\frac{N-\alpha-1}{2}} \geqslant \binom{N-2}{\frac{N-\alpha-1}{2}} \geqslant \frac{\sqrt{2}}{4}\frac{2^{N-1}}{\sqrt{N-1}}\exp\left(\frac{-(\alpha-1)^2}{N-1}\right).$$

Para evitar esta distinción entre N par o impar podemos escribir la común acotación inferior

$$\binom{N-1}{\frac{N-\alpha-1}{2}} \geqslant \frac{\sqrt{2}}{8}\frac{2^{N}}{\sqrt{N}}\exp\left(\frac{-2\alpha^2}{N}\right).$$

Por lo tanto, tenemos

$$\sigma(x_1+...+x_n>\alpha) \geqslant \frac{\sqrt{2}}{8}\exp\left(\frac{-2\alpha^2}{N}\right)Y\left(\frac{\alpha+1}{\sqrt{N}}\right). \tag{3.51}$$

Para acotar el factor anterior simplemente se recurre a la definición (3.40) de la función Y para obtener que

$$\begin{aligned} Y\left(\frac{\alpha+1}{\sqrt{N}}\right) &= \exp\left(\frac{(\alpha+1)^2}{2N}\right)\int_{\frac{\alpha+1}{\sqrt{N}}}^{\infty} e^{\frac{-t^2}{2}}dt \\ &\geqslant \exp\left(\frac{\alpha^2}{2N}\right)\int_{\frac{\alpha+1}{\sqrt{N}}}^{\frac{\alpha+1}{\sqrt{N}}+1} e^{-\frac{t^2}{2}}dt \\ &\geqslant \exp\left(\frac{\alpha^2}{2N}\right)\exp\left[-\frac{1}{2}\left(1+\frac{\alpha+1}{\sqrt{N}}\right)^2\right] \\ &\geqslant \exp\left(\frac{\alpha^2}{2N}\right)\exp\left[-2\left(1+\frac{\alpha}{\sqrt{N}}\right)^2\right]. \end{aligned}$$

Aplicando la estimación anterior a (3.51) concluimos que

$$\begin{aligned} \sigma(x_1+...+x_N>\alpha) &\geqslant \frac{\sqrt{2}}{8}\exp\left(-\frac{3\alpha^2}{2N}\right)\exp\left[-2\left(1+\frac{\alpha}{\sqrt{N}}\right)^2\right] \\ &\geqslant \exp\left[-6\left(1+\frac{\alpha}{\sqrt{N}}\right)^2\right], \end{aligned}$$

esto completa la demostración. □

Estamos en condiciones de terminar la sección demostrando el Teorma 3.4.1.

Demostración. Sea $N \in \mathbb{N}$ y $\frac{1}{2^N} \leqslant \delta \leqslant 1$. Lo siguiente fue probado en el Teorema 3.4.3

$$\rho(\mathcal{B}_\delta^N) \geqslant \frac{1}{5\sqrt{N}\sqrt{log(\frac{2}{\delta})}}.$$

Probemos la acotación superior, haremos uso de las funciones $\psi_{\alpha,N}$ $(0 \leqslant \alpha < N)$ que fueron estudiadas en la sección previa. Recordar que

$$|\mathbb{E}[\psi_{\alpha,N}]| = 1 - 2\sigma(x_1 + ... + x_N > \alpha) = [1 - 2\sigma(x_1 + ... + x_N > \alpha)]\|\psi_{\alpha,N}\|_\infty.$$

Entonces, por el Lema 3.4.11 deducimos que $\psi_{\alpha,N} \in \mathcal{B}_\delta^N$ cuando

$$\left(1 + \frac{\alpha}{\sqrt{N}}\right)^2 = \frac{\log(\frac{2}{\delta})}{6}. \tag{3.52}$$

Si $\delta < \frac{2}{e^6}$, podemos encontrar $0 \leqslant \alpha < N$ que satisface la condición anterior, luego

$$\rho(\mathcal{B}_\delta^N) \leqslant \rho(\psi_{N,\alpha}) \leqslant \frac{C}{\alpha + \sqrt{N}} = \frac{C}{\sqrt{N}\left(1 + \frac{\alpha}{\sqrt{N}}\right)} = \frac{C\sqrt{6}}{\sqrt{N}\sqrt{\log(\frac{2}{\delta})}}.$$

Si $\delta \geqslant \frac{2}{e^6}$, como $\psi_{N,0} \in \mathcal{B}_\delta^N$ trivialmente, conseguimos que

$$\rho(\mathcal{B}_\delta^N) \leqslant \rho(\psi_{N,0}) \leqslant O\left(\frac{1}{\sqrt{N}}\right),$$

y con esto concluimos la afirmación. $\square$

Capítulo 4

Desigualdades

Este capítulo tiene como objetivo desarrollar con cierto detalle la desigualdad de Blei y las desigualdades hipercontractivas. La relevancia de la primera radica en que la mayoría de las demostraciones modernas de la desigualdad de Bohnenblust-Hille dependen en algún grado de la desigualdad Blei. Por otro lado, las desigualdades hipercontractivas fueron fundamentales para acotar las constantes de Bohnenblust-Hille en el Capítulo 2, también fueron claves en el Capítulo 3, por ejemplo para acotar inferiormente el radio Booleano para la clase $\mathcal{B}_\delta^N$. De todas maneras las desigualdes hipercontractivas tiene un interés propio en el campo del análisis Booleano, ya que es una herramienta poderosa para comparar normas a través de ciertos operadores y goza de múltiples aplicaciones, como se explica por ejemplo en el trabajo de Montanaro [Mon12].

4.1. Desigualdad de Blei

Una de las desigualdades que usamos para acotar el crecimiento de la constante de Bohnenblust-Hille para el caso del cubo Booleano es la desigualdad de Blei (2.3), en lo que sigue daremos la demotración. Comenzamos con un caso particular para motivar los pasos de la demostración.

Proposición 4.1.1. *Para cada $m \geqslant 2$, $n \in \mathbb{N}$, y para cualquier colección $(a_{i_1,\dots,i_m})_{1 \leqslant i_1,\dots,i_m \leqslant n}$ de números complejos tenemos*

$$\left(\sum_{i_1,\dots,i_m=1} |a_{i_1,\dots,i_m}|^{\frac{2m}{m+1}} \right)^{\frac{m+1}{2m}} \leqslant \left[\prod_{k=1}^{m} \sum_{i_k=1}^{n} \left(\sum_{i_1,\dots,i_{k-1},i_{k+1},\dots,i_m}^{n} |a_{i_1,\dots,i_m}|^2 \right)^{\frac{1}{2}} \right]^{\frac{1}{m}}. \tag{4.1}$$

Podemos reescribir la afirmación, usando la notación que desarrollamos en el Capítulo 2 (veáse la Sección 2.1.1), de la siguiente manera:

$$\sum_{\mathbf{i} \in \mathcal{I}(m,n)} |a_{\mathbf{i}}|^{\frac{2m}{m+1}} \leqslant \prod_{k=1}^{m} \left(\sum_{j=1}^{n} \left(\sum_{\mathbf{j} \in \mathcal{I}(m-1,n)} |a_{\mathbf{j} \oplus_k j}|^2 \right)^{\frac{1}{2}} \right)^{\frac{2}{m+1}}. \tag{4.2}$$

Demostración. Procedemos las demostración por inducción en m. Comencemos con el caso $m = 2$:

$$\sum_{i,j} |a_{i,j}|^{\frac{4}{3}} = \sum_i \left(\sum_j |a_{i,j}|^{\frac{2}{3}} |a_{i,j}|^{\frac{2}{3}} \right) \leqslant \sum_i \left(\sum_j |a_{i,j}|^2 \right)^{\frac{1}{3}} \left(\sum_j |a_{i,j}| \right)^{\frac{2}{3}},$$

en el último paso aplicamos la desigualdad de Hölder con $p = 3$ y $p' = \frac{3}{2}$, nuevamente para el próximo, pero ahora con $p = \frac{3}{2}$ y $p' = 3$

$$\leqslant \left[\sum_i \left(\sum_j |a_{i,j}|^2 \right)^{\frac{1}{2}} \right]^{\frac{2}{3}} \left[\left(\sum_i \left(\sum_j |a_{i,j}| \right)^2 \right)^{\frac{1}{2}} \right]^{\frac{2}{3}}.$$

Por la desigualdad de Minkowski, podemos acotar el primer factor

$$\left(\sum_i \left(\sum_j |a_{i,j}| \right)^2 \right)^{\frac{1}{2}} \leqslant \sum_j \left(\sum_i |a_{i,j}|^2 \right)^{\frac{1}{2}},$$

y de todo esto obtenemos que:

$$\left(\sum_{i,j} |a_{i,j}|^{\frac{4}{3}} \right)^{\frac{3}{4}} \leqslant \left[\sum_i \left(\sum_j |a_{i,j}|^2 \right)^{\frac{1}{2}} \right]^{\frac{1}{2}} \left[\sum_j \left(\sum_i |a_{i,j}|^2 \right)^{\frac{1}{2}} \right]^{\frac{1}{2}}.$$

Para el caso general seguimos esencialmente la misma estructura que para $m = 2$. Empezamos aplicando la desigualdad de Hölder dos veces, primero con $p = \frac{m+1}{2}$ y $p' = \frac{m+1}{m-1}$, luego con $p = m + 1$ y $p' = \frac{m+1}{m}$:

$$\begin{aligned}
\sum_{\mathbf{i} \in \mathcal{I}(m,n)} |a_{\mathbf{i}}|^{\frac{2m}{m+1}} &= \sum_{\mathbf{j} \in \mathcal{I}(m-1.n)} \sum_{j=1}^{n} |a_{\mathbf{j} \oplus_m j}|^{\frac{2}{m+1}} |a_{\mathbf{j} \oplus_m j}|^{\frac{2(m-1)}{m+1}} \\
&\leqslant \sum_{\mathbf{j} \in \mathcal{I}(m-1.n)} \left(\sum_j^n |a_{\mathbf{j} \oplus_m j}| \right)^{\frac{2}{m+1}} \left(\sum_j^n |a_{\mathbf{j} \oplus_m j}|^2 \right)^{\frac{m-1}{m+1}} \\
&\leqslant \left(\sum_{\mathbf{j} \in \mathcal{I}(m-1.n)} \left(\sum_j^n |a_{\mathbf{j} \oplus_m j}| \right)^2 \right)^{\frac{1}{m+1}} \times \left(\sum_{\mathbf{j} \in \mathcal{I}(m-1.n)} \left[\left(\sum_j^n |a_{\mathbf{i} \oplus_m j}|^2 \right)^{\frac{1}{2}} \right]^{\frac{2(m-1)}{m}} \right)^{\frac{m}{m+1}}.
\end{aligned}$$

Ahora acotemos el primer factor con la desigualdad de Minkowski

$$\left(\sum_{\mathbf{j} \in \mathcal{I}(m-1.n)} \left(\sum_j^n |a_{\mathbf{j} \oplus_m j}| \right)^2 \right)^{\frac{1}{2}} \leqslant \sum_{j=1}^{n} \left(\sum_{\mathbf{j} \in \mathcal{I}(m-1.n)} |a_{\mathbf{j} \oplus_m j}|^2 \right)^{\frac{1}{2}}.$$

Para el segundo factor, usaremos la hipótesis inductiva (para $m-1$). Con el fin de mantener la notación simple, escribimos $\alpha_{\mathbf{j}} = \left(\sum_j^n |a_{\mathbf{j}\oplus_m j}|^2\right)^{\frac{1}{2}}$, luego

$$\begin{aligned}
\left(\sum_{\mathbf{j}\in\mathcal{I}(m-1.n)} \alpha_{\mathbf{j}}^{\frac{2(m-1)}{m}}\right)^{\frac{m}{m+1}} &\leqslant \left[\prod_{k=1}^{m-1}\sum_{l=1}^{n}\left(\sum_{\mathbf{k}\in\mathcal{I}(m-2,n)} \alpha_{\mathbf{k}\oplus_k l}^2\right)^{\frac{1}{2}}\right]^{\frac{2}{m}\frac{m}{m+1}}\\
&= \left[\prod_{k=1}^{m-1}\sum_{l=1}^{n}\left(\sum_{\mathbf{k}\in\mathcal{I}(m-2,n)} \left(\sum_{j=1}^{m} |a_{(\mathbf{k}\oplus_k l)\oplus_m j}|^2\right)^{\frac{1}{2}2}\right)^{\frac{1}{2}}\right]^{\frac{2}{m+1}}\\
&= \left[\prod_{k=1}^{m-1}\sum_{l=1}^{n}\left(\sum_{\mathbf{k}\in\mathcal{I}(m-2,n)} \sum_{j=1}^{m} |a_{(\mathbf{k}\oplus_k l)\oplus_m j}|^2\right)^{\frac{1}{2}}\right]^{\frac{2}{m+1}}\\
&= \left[\prod_{k=1}^{m-1}\sum_{l=1}^{n}\left(\sum_{\mathbf{j}\in\mathcal{I}(m-1,n)} |a_{\mathbf{j}\oplus_k l}|^2\right)^{\frac{1}{2}}\right]^{\frac{2}{m+1}}.
\end{aligned}$$

Juntando todo obtenemos

$$\begin{aligned}
\sum_{\mathbf{i}\in\mathcal{I}(m,n)} |a_{\mathbf{i}}|^{\frac{2m}{m+1}} &\leqslant \left[\sum_{j=1}^{m}\left(\sum_{\mathbf{j}\in\mathcal{I}(m-1.n)} |a_{\mathbf{j}\oplus_m j}|^2\right)^{\frac{1}{2}}\right]^{\frac{2}{m+1}} \left[\prod_{k=1}^{m-1}\sum_{l=1}^{n}\left(\sum_{\mathbf{j}\in\mathcal{I}(m-1,n)} |a_{\mathbf{j}\oplus_k l}|^2\right)^{\frac{1}{2}}\right]^{\frac{2}{m+1}}.\\
&= \left[\prod_{k=1}^{m}\sum_{l=1}^{n}\left(\sum_{\mathbf{j}\in\mathcal{I}(m-1,n)} |a_{\mathbf{j}\oplus_k l}|^2\right)^{\frac{1}{2}}\right]^{\frac{2}{m+1}},
\end{aligned}$$

que es lo que se quiere concluir. □

Con el caso particular probado, podemos generalizar la afirmación para obtener la desigualdad de Blei.

Teorema 4.1.2. *(Desigualdad de Blei) Sea una colección de números complejos* $a = (a_{\mathbf{i}})_{\mathbf{i}\in\mathcal{I}(m,n)}$*, tenemos que para cada* $1 \leqslant k \leqslant m$*,*

$$\left(\sum_{\mathbf{i}\in\mathcal{I}(m,n)} |a_{\mathbf{i}}|^{\frac{m+1}{2m}}\right)^{\frac{m+1}{2m}} \leqslant \left[\prod_{S\in[m]_k}\left(\sum_{\mathbf{i}\in\mathcal{I}(S,n)}\left(\sum_{\mathbf{j}\in\mathcal{I}(\hat{S},n)} |a_{\mathbf{i}\oplus\mathbf{j}}|^2\right)^{\frac{1}{2}\frac{2k}{k+1}}\right)^{\frac{k+1}{2k}}\right]^{\frac{1}{\binom{m}{k}}}. \tag{4.3}$$

Notar que para el caso que $k = 1$, obtenemos $S = \{i_k\}$, $\frac{2k}{k+1} = 1$ y $\binom{m}{k} = m$, por lo tantoo estamos en la desigualdad anterior. Recordemos que el ingrediente principal fue la desigualdad de Hölder, que nos permitió separar la suma en un producto de dos sumas. Para esta generalización mixta de la desigualdad, necesitaremos la versión general de la desigualdad de Hölder para normas mixtas de Minkowski. Establezcamos algunas notaciones básicas.

Para toda colección $(a_{\mathbf{i}})_{\mathbf{i}\in\mathcal{I}(m,n)}$ y para todo vector $r = (r(k))_{k=1}^{m}$ con $1 \leqslant r(k) < \infty$, la norma mixta de Minkowski $\|a\|_r$ esta dada por la siguiente expresión:

$$\|a\|_r = \left(\sum_{i_1=1}^{n}\left(\sum_{i_2=1}^{n}\left(\dots\left(\sum_{i_{m-1}=1}^{n}\left(\sum_{i_m=1}^{n}|a_{\mathbf{i}}|^{r(m)}\right)^{\frac{r(m-1)}{r(m)}}\right)^{\frac{r(m-2)}{r(m-1)}}\dots\right)^{\frac{r(2)}{r(3)}}\right)^{\frac{r(1)}{r(2)}}\right)^{\frac{1}{r(1)}}. \tag{4.4}$$

Lema 4.1.3. *Sean* $\mathbf{p}_1, ..., \mathbf{p}_N, \mathbf{r}$ *vectores con* $1 \leqslant \mathbf{p}_j(k), \mathbf{r}(k) < \infty$ *para todo j= 1,...,N y* $k = 1, ..., m$ *tal que*

$$\frac{1}{\mathbf{r}(k)} = \sum_{j=1}^{N}\frac{1}{\mathbf{p}_j(k)}.$$

Entonces, para cada elección de colecciones $a_1, ..., a_N$ *indexada en* $\mathcal{I}(m,n)$ *(con* $a_j = (a_j(\mathbf{i}))_{\mathbf{i}\in\mathcal{I}(m,n)}$ *para* $j = 1, ..., N$*) tenemos*

$$\|\prod_{j=1}^{N} a_j\|_{\mathbf{r}} \leqslant \prod_{j=1}^{N}\|a_j\|_{\mathbf{p}_j},$$

donde el producto de la izquierda se toma de forma coordinada.

Podemos observar que el caso $m = 1$ es la clásica desigualdad de Hölder.

Demostración. (del Lema 4.1.3) Hagamos la demostración para $m = N = 2$ y n arbitrario. Fijo $1 \leqslant i \leqslant n$, por la desigualdad de Hölder

$$\left(\sum_{j=1}^{n}|a_1(i,j)a_2(i,j)|^{r(2)}\right)^{\frac{1}{r(2)}} \leqslant \left(\sum_{j=1}^{n}|a_1(i,j)|^{\mathbf{p_1(2)}}\right)^{\frac{1}{\mathbf{p}_1(2)}}\left(\sum_{j=1}^{n}|a_2(i,j)|^{\mathbf{p_2(2)}}\right)^{\frac{1}{\mathbf{p}_2(2)}}$$

Nuevamente usando la desigualdad de Hölder, obtenemos

$$\begin{aligned}
\|a_1a_2\|_r &= \left(\sum_{i=1}^n \left(\sum_{j=1}^n |a_1(i,j)a_2(i,j)|^{r(2)}\right)^{\frac{r(1)}{r(2)}}\right)^{\frac{1}{r(1)}} \\
&\leqslant \left(\sum_{i=1}^n \left(\sum_{j=1}^n | \ a_1(i,j)|^{\mathbf{p}_1(2)}\right)^{\frac{r(1)}{\mathbf{p}_1(2)}} \left(\sum_{j=1}^n |a_2(i,j)|^{\mathbf{p}_2}\right)^{\frac{r(1)}{\mathbf{p}_2(2)}}\right)^{\frac{1}{r(1)}} \\
&\leqslant \left(\sum_{i=1}^n \left(\sum_{j=1}^n |a_1(i,j)|^{\mathbf{p}_1(2)}\right)^{\frac{\mathbf{p}_1(1)}{\mathbf{p}_1(2)}}\right)^{\frac{1}{\mathbf{p}_1(1)}} \left(\sum_{i=1}^n \left(\sum_{j=1}^n |a_2(i,j)|^{\mathbf{p}_2(2)}\right)^{\frac{\mathbf{p}_2(1)}{\mathbf{p}_2(2)}}\right)^{\frac{1}{\mathbf{p}_2(1)}} \\
&= \|a_1\|_{\mathbf{p}_1}\|a_2\|_{\mathbf{p}_2}.
\end{aligned}$$

El caso general es simplemente la iteración de este argumento. □

Y ahora la estructura de la demostración del Teorema de Blei 4.1.2, es más o menos la misma que el caso particular: primero dividimos la suma en un producto usando Hölder y acotamos cada factor usando la desigualdad de Minkowski.

Demostración. Para cada $S \in [m]_k$ definimos, para $1 \leqslant i \leqslant m$,

$$\mathbf{q}_S(i) = \begin{cases} \frac{2k}{k+1} & \text{si } \ i \in S, \\ 2 & \text{si } \ i \notin S. \end{cases}$$

Considerando la m-tupla

$$\mathbf{q} = (\frac{2m}{m+1}, ..., \frac{2m}{m+1})$$

y

$$\theta = \frac{1}{\binom{m}{k}} = \frac{k!(m-k)!}{m!}.$$

Para cada i, tenemos

$$\begin{aligned}
\sum_{S\in[m]_k} \frac{\theta}{\mathbf{q}_S(i)} &= \theta\left(\sum_{S:i\in S} \frac{1}{\mathbf{q}_S(i)} + \sum_{S:i\notin S} \frac{1}{\mathbf{q}_S(i)}\right) \\
&= \binom{m}{k}^{-1} \left(\binom{m-1}{k-1}\frac{k+1}{2k} + \binom{m-1}{k}\frac{1}{2}\right) \\
&= \frac{k!(m-k)!}{m!}\left(\frac{(m-1)!}{(k-1)!(m-k)!}\frac{k+1}{2k} + \frac{(m-1)!}{k!(m-k-1)!}\frac{1}{2}\right) \\
&= \frac{m+1}{2m} = \frac{1}{\mathbf{q}(i)}.
\end{aligned}$$

Aplicando el Lema 4.1.3 obtenemos

$$\|a\|_{\mathbf{q}} = \| \prod_{S\in[m]_k} |a|^{\theta} \|_q \leqslant \prod_{S\in[m]_k} \| |a|^{\theta} \|_{\frac{\mathbf{q}_S}{\theta}} = \prod_{S\in[m]_k} \|a\|^{\theta}_{\mathbf{q}_S}.$$

Claramente tenemos que:

$$\|a\|_q = \left(\sum_{\mathbf{i}\in\mathcal{I}(d,n)} |a_{\mathbf{i}}|^{\frac{2m}{m+1}} \right)^{\frac{m+1}{2m}}.$$

Veamos que para cada $S \in [m]_k$, tenemos que

$$\|a\|_{\mathbf{q}_S} \leqslant \left(\sum_{\mathbf{i}\in\mathcal{I}(S,n)} \left(\sum_{\mathbf{j}\in\mathcal{I}(\hat{S},n)} |a_{\mathbf{i}\oplus\mathbf{j}}|^2 \right)^{\frac{1}{2}\times\frac{2k}{k+1}} \right)^{\frac{k+1}{2k}}. \tag{4.5}$$

Supongamos $S = \{j_1, ..., j_k\}$. Para simplificar la notación, para $\mathbf{i} = (i_1, ..., i_m)$, escribimos

$$\alpha_{i_1,\dots,i_{j_k}} = \left(\sum_{i_{j_k+1},\dots,i_m=1} |a_{\mathbf{i}}|^2 \right)^{\frac{1}{2}}.$$

Ahora, como $\frac{2k}{k+1} \leqslant 2$, podemos aplicar la desigualdad de Minkowski

$$\begin{aligned}
&\left(\sum_{i_{j_{k-1}+1},\dots,i_{j_k-1}} \left[\sum_{i_{j_k}} \left(\sum_{i_{j_k+1},\dots,i_m} |a_{\mathbf{i}}|^2 \right)^{\frac{1}{2}\times\frac{2k}{k+1}} \right]^{\frac{k+1}{2k}\times 2} \right)^{\frac{1}{2}} \\
&= \left(\sum_{i_{j_{k-1}+1},\dots,i_{j_k-1}} \left[\sum_{i_{j_k}} \alpha_{i_1,\dots,i_{j_k-1}}^{\frac{2k}{k+1}} \right]^{\frac{k+1}{2k}\times 2} \right)^{\frac{1}{2}} \leqslant \left(\sum_{i_{j_k}} \left[\sum_{i_{j_{k-1}+1},\dots,i_{j_k-1}} \alpha^2_{i_1,\dots,i_{j_k}} \right]^{\frac{1}{2}\times\frac{2k}{k+1}} \right)^{\frac{k+1}{2k}} \\
&= \left[\sum_{i_{j_k}} \left(\sum_{i_{j_{k-1}+1},\dots,i_{j_k-1},\hat{i}_j,i_{j_k+1},\dots,i_m} |a_{\mathbf{i}}|^2 \right)^{\frac{1}{2}\times\frac{2k}{k+1}} \right]^{\frac{k+1}{2k}}.
\end{aligned}$$

Repitiendo este procedimiento, extraemos todos los coeficientes de S a la suma externa y obtenemos (4.5).

4.2. Teoremas de hipercontractividad

Comencemos introduciendo el siguiente concepto que llamaremos estabilidad al ruido, a modo de motivar definiciones y generar algún tipo de intuición sobre los teoremas de hipercontractividad.

4.2.1. Estabilidad al ruido

Supongamos que $f : \{\pm 1\}^n \to \{\pm 1\}$ es una regla de votación para una elección de 2 candidatos y n votantes; esto es, asigna los votos de los votantes al ganador de las elecciones. Quizás la regla de votación más familiar es la función de mayoría:

Definición 4.2.1. *Para n impar, la función mayoría o mayoritaria $Maj_n : \{\pm 1\}^n \to \{\pm 1\}$ está definida por $Maj_n(x) = sign(x_1 + x_2 + ... + x_n)$. Ocasionalmente, para n par podemos decir que f es una función mayoría si $f(x)$ es igual al signo de $x_1 + ... + x_n$ siempre que este número sea distinto de cero.*

Entonces tenemos $f : \{\pm 1\}^n \to \{\pm 1\}$ una regla de votación. Haciendo una suposición de cultura imparcial, es decir los n votantes eligen sus votos de forma independiente y uniforme al azar $x = (x_1, ..., x_n)$. Imaginemos que cuando cada votante va a las urnas hay alguna posibilidad de que su voto sea mal registrado. Específicamente, digamos que cada voto se registra correctamente con probabilidad $\rho \in [0, 1]$ y está confuso, es decir, cambiado a un bit aleatorio, con probabilidad $1 - \rho$. Escribiendo $y = (y_1, ..., y_n)$ para los votos que finalmente se registran, podemos preguntar sobre la probabilidad de que $f(x) = f(y)$, es decir, si los votos se registraron incorrectamente afecta o no el resultado de la elección. Esto tiene que ver con la estabilidad al ruido de f.

Definición 4.2.2. *Sea $\rho \in [0, 1]$. Para $x \in \{\pm 1\}^n$ escribimos $\mathbf{y} \sim N_\rho(x)$ para denotar el vector aleatorio $\mathbf{y}$ el cual se define de la siguiente manera: para cada $i \in [n]$*

$$\mathbf{y}_i = \begin{cases} x_i & \textit{con probabilidad } \rho \\ \\ \textit{uniformemente aleatorio} & \textit{con probabilidad } 1 - \rho, \end{cases} \tag{4.6}$$

extendemos la notación para todo $\rho \in [-1, 1]$ como sigue:

$$\mathbf{y}_i = \begin{cases} x_i & \textit{con probabilidad } \frac{1}{2} + \frac{1}{2}\rho \\ \\ -x_i & \textit{con probabilidad } \frac{1}{2} - \frac{1}{2}\rho. \end{cases} \tag{4.7}$$

Decimos que $\mathbf{y}$ está ρ-correlacionado con x.

Definición 4.2.3. *Si $\mathbf{x} \sim \{\pm 1\}^n$ se toma uniformemente al azar e $\mathbf{y} \sim N_\rho(\mathbf{x})$, decimos que $(\mathbf{x}, \mathbf{y})$ es un par $\rho-$ correlacionado de vectores aleatoreos. Esta definición es simétrica en $\mathbf{x}$ e $\mathbf{y}$; esto es equivalente a decir que independientemente para cada $i \in [n]$, el par de bit aleatorio $(\mathbf{x}_i, \mathbf{y}_i)$ satisface $\mathbb{E}[\mathbf{x}_i] = \mathbf{E}[\mathbf{y}_i] = 0$ y $\mathbb{E}[\mathbf{x}_i\mathbf{y}_i] = \rho$.*

Con esto podemos definir el concepto de estabilidad al ruido, que mide la correlación entre $f(\mathbf{x})$ y $f(\mathbf{y})$ cuando $(\mathbf{x}, \mathbf{y})$ es un par ρ-correlacionado.

Definición 4.2.4. *Para $f : \{\pm 1\}^n \to \mathbb{R}$ y $\rho \in [-1, 1]$, la estabilidad al ruido de f para ρ es*

$$\mathbf{Stab}_\rho[f] = \underset{\substack{(\mathbf{x},\mathbf{y}) \\ \rho-\textit{correlacionado}}}{\mathbb{E}} [f(\mathbf{x})f(\mathbf{y})].$$

Si $f : \{\pm 1\}^n \to \{\pm 1\}$ tenemos que

$$\begin{aligned}\mathbf{Stab}_\rho[f] =& \underset{\substack{(\mathbf{x},\mathbf{y})\\ \rho-\text{correlacionado}}}{\mathbb{P}}[f(\mathbf{x}) = f(\mathbf{y})] - \underset{\substack{(\mathbf{x},\mathbf{y})\\ \rho-\text{correlacionado}}}{\mathbb{P}}[f(\mathbf{x}) \neq f(\mathbf{y})]\\ =&2 \underset{\substack{(\mathbf{x},\mathbf{y})\\ \rho-\text{correlacionado}}}{\mathbb{P}}[f(\mathbf{x}) = f(\mathbf{y})] - 1.\end{aligned}$$

En el escenario de votación descripto anteriormente, la probabilidad de que el registro erróneo de los votos no afecte el resultado de la elección es $\frac{1}{2} + \frac{1}{2}\mathbf{Stab}_\rho[f]$. Cuando ρ está cerca de 1 (es decir, el "ruido" es pequeño) a veces es más natural preguntarse sobre la probabilidad de revertir una pequeña fracción de los votos para que se revierta el resultados de la elección.

A continuación presentaremos uno de los operadores más importante en el análisis de las funciones Booleanas: el operador de ruido, denotado por T_ρ, si bien en otros capítulos ya definimos este operador, aquí lo haremos de manera distinta (ambas definiciones resultan equivalentes).

Definición 4.2.5. *Para $\rho \in [-1, 1]$, el operador de ruido con parámetro ρ es el operador lineal T_ρ en funciones $f : \{\pm 1\}^n \to \mathbb{R}$ definido por*

$$T_\rho f(x) = \underset{\mathbf{y}\sim N_\rho(x)}{\mathbb{E}}[f(\mathbf{y})].$$

Proposición 4.2.6. *Para toda $f : \{\pm 1\}^n \to \mathbb{R}$, la expansión de Fourier-Walsh de $T_\rho f$ está dada por*

$$T_\rho f = \sum_{S\subset[n]} \rho^{|S|}\hat{f}(S)\chi_S = \sum_{k=0}^{n} \rho^k f_k.$$

Demostración. Como T_ρ es un operador lineal, es suficiente verificar que $T_\rho\chi_S = \rho^{|S|}\chi_S$:

$$T_\rho\chi_S(x) = \underset{\mathbf{y}\sim N_\rho(x)}{\mathbb{E}}[\mathbf{y}^S] = \prod_{i\in S} \underset{\mathbf{y}\sim N_\rho(x)}{\mathbb{E}}[\mathbf{y_i}] = \prod_{i\in S}(\rho x_i) = \rho^{|S|}\chi_S(x).$$

Aquí usamos el hecho que para $\mathbf{y} \sim N_\rho(x)$ los bits $\mathbf{y_i}$ son independientes y satisfacen $\mathbb{E}[\mathbf{y}_i] = \rho x_i$ □

La conección entre T_ρ y la estabilidad al ruido está dada por la siguiente igualdad

$$\mathbf{Stab}_\rho[f] = \underset{\substack{(\mathbf{x},\mathbf{y})\\ \rho-\text{correlacionado}}}{\mathbb{E}}[f(\mathbf{x})f(\mathbf{y})] = \mathbb{E}_{\mathbf{x}}\left[f(x) \underset{\mathbf{y}\sim N_\rho(\mathbf{x})}{\mathbb{E}}[f(\mathbf{y})]\right]. \tag{4.8}$$

Por lo tanto:

$$\mathbf{Stab}_\rho[f] = \langle f, T_\rho f\rangle.$$

Hechas estas definiciones podemos comenzar a enunciar las desigualdades hipecontractivas. En 1970, Bonami probó el siguiente resultado, normalmente conocido como el Teorema de hipercontractividad.

Teorema 4.2.7. *Sea* $f : \{\pm 1\}^n \to \mathbb{R}$ *y sean* $1 \leqslant p \leqslant q \leqslant \infty$. *Entonces*

$$\|T_\rho f\|_q \leqslant \|f\|_p, \qquad \text{para } 0 \leqslant \rho \leqslant \sqrt{\frac{p-1}{q-1}}. \tag{4.9}$$

Si bien no probaremos el caso general, que se puede encontrar en [O'D14, Capítulo 10], sí demostraremos algunos casos particulares, como por ejemplo Lema de Bonami, el $(2, q)$-Teorema de hipercontractividad y el $(p, 2)$-Teorema de hipercontractividad, con el fin de mostrar con cierto detalle las ideas y estrategias que hay de fondo cuando uno intenta conseguir este tipo de desigualdades en el cubo Booleano.

Así como tales resultado fueron de gran utilidad en esta tesis para poder obtener teoremas significativos, como para acotar subexponencialmente la constante de Bohnenblust-Hille y obtener estimaciones ajustadas para el radio Booleano. Pasa lo mismo en muchos otros problemas donde entra en juego el cubo Booleano, por poner algún ejemplo los resultados de hipercontractividad son muy usados en teoría de informatica cuántica como brevemente comentamos en la Sección 2.3. Podemos encontrar más aplicaciones en el artículo de Montanaro [Mon12] ó en el libro de R. O'Donnell [O'D14, Capítulos 3, 9 y 10].

Lema 4.2.8. *(Lema de Bonami) Sea* $f : \{\pm 1\}^n \to \mathbb{R}$ *de grado* k. *Entonces*

$$\|f\|_4 \leqslant (\sqrt{3})^k \|f\|_2.$$

La idea fundamental de esta afirmación es que si $\mathbf{x} \sim \{\pm 1\}^n$ variable aleatoria uniforme y $f : \{\pm 1\}^n \to \mathbb{R}$ tiene grado bajo entonces la variable aleatoria $f(\mathbf{x})$ es bastante "razonable", es decir se distribuye bien alrededor de su esperanza. El Lema de Bonami se demuestra elementalmente por inducción y es lo suficientemente poderoso para obtener muchas de las conocidas aplicaciones de hipercontractividad, incluido el Teorema KKL (Kahn-Kalai-Linial) el cual afirma que para cada función $f : \{\pm 1\}^n \to \{\pm 1\}$, existe un $j \in [n]$ tal que

$$Inf_j[f] \geqslant Var[f] \cdot \Omega(\frac{\log n}{n}),$$

(ver [O'D14, Pág. 265]) y el Principio de Invarianza (veáse [O'D14, Pág. 355]).

También tenemos, por ejemplo el **(2,q)-Teorema de hipercontractividad**.

Teorema 4.2.9. *Sean* $f : \{\pm 1\}^n \to \mathbb{R}$ *y* $2 \leqslant q \leqslant \infty$. *Entonces*

$$\|T_{\frac{1}{\sqrt{q-1}}} f\|_q \leqslant \|f\|_2.$$

Como consecuencia, si f tiene grado a lo sumo k entonces $\|f\|_q \leqslant (\sqrt{q-1})^k \|f\|_2$. Este teorema cuantifica la medida en que T_ρ es un operador "suavizante"; equivalentemente, esto da aún más control sobre la razonabilidad de los polinomios de bajo grado.

También tenemos el **(p,2)-Teorema de hipercontractividad**

Teorema 4.2.10. *Sean* $f : \{\pm 1\}^n \to \mathbb{R}$ *y* $1 \leqslant p \leqslant 2$. *Entonces*

$$\|T_{\sqrt{p-1}} f\|_2 \leqslant \|f\|_p.$$

Este resultado como veremos luego es equivalente al $(2, q)$-Teorema hipercontractivo en virtud de la desigualdad de Hölder .

4.2.2. Razonabilidad de polinomios de bajo grado

Como habitualmente ocurre en probabilidad, una variable aleatoria a veces puede comportarse de manera bastante “irracional”. Puede que nunca esté cerca de su esperanza. Puede exceder su esperanza casi siempre, o casi nunca. Puede tener $1^{er.}$, $2^{do.}$ y $3^{er.}$ momento finito, pero el $4^{to.}$ momento puede dar infinito.

Este comportamiento irregular puede causar muchos problemas. Sería bueno quedarse con ciertas variables aleaorias que sean más estables en este sentido. Una simple condición para que una variable aleatoria garantice un buen comportamiento, es que el cuarto momento no sea demasiado grande en comparación con su segundo momento.

Definición 4.2.11. *Para un número real $B \geqslant 1$ decimos que la variable aleatoria $\boldsymbol{X}$ es B-razonable, si*

$$\mathbb{E}[\boldsymbol{X}^4] \leqslant B\mathbb{E}[\boldsymbol{X}^2]^2. \tag{4.10}$$

Equivalentemente, si $\|\boldsymbol{X}\|_4 \leqslant B^{\frac{1}{4}}\|\boldsymbol{X}\|_2$.

Cuanto más pequeño B, decimos que más razonable es $\mathbf{X}$. Esta definición es invariante por escalares (es decir, $c\mathbf{X}$ es B-razonable si y sólo si $\mathbf{X}$ lo es, para $c \neq 0$) pero no es invariante por traslación ($c + \mathbf{X}$ y $\mathbf{X}$ pueden no ser razonablemente equivalentes). El último hecho puede resultar molesto algunas veces. Se pueden encontrar algunas condiciones alternativas que también cumplen con la razonabilidad. Por ejemplo, podemos considerar el análogo pero para el $3^{er.}$ momento, es decir que cumpla la condición $\mathbb{E}[|\mathbf{X}|^3] \leqslant B\mathbb{E}[\mathbf{X}^2]^{\frac{3}{2}}$. Estrictamente hablando la condición del $4^{to.}$ momento es más fuerte:

Si $\mathbf{X}$ es B-razonable, entonces

$$\mathbb{E}[|\mathbf{X}|^3] = \mathbb{E}[|\mathbf{X}| \cdot \mathbf{X}^2] \leqslant \sqrt{\mathbb{E}[\mathbf{X}^2]}\sqrt{\mathbb{E}[\mathbf{X}^4]} \leqslant \sqrt{B}\mathbb{E}[\mathbf{X}^2]^{\frac{3}{2}}.$$

Por otro lado, existe una variable aleatoria con $3^{er.}$ momento finito y $4^{to.}$ momento infinito (un ejemplo de una variable aleatoria así es la que tiene distribución t de Student con 4 grados de libertad). Sin embargo, tales variables aleatorias casi nunca surgen para nosotros, y esto dice que la condición de tercer y cuarto momento son igualmente buenos para la razonabilidad.

Ejemplos 4.2.12. En lo siguiente daremos ejemplos básicos de variables aleatorias que cumplen la definición de ser B-razonable:

Si $\mathbf{x} \sim \{-1, 1\}$ es una variable uniforme (es decir es una variable aleatoria que elige con misma probabilidad a 1 que a -1) entonces $\mathbf{x}$ es 1-razonable, ya que calculando los momentos pares obtenemos:

$$\mathbb{E}(|\mathbf{x}^k|) = \mathbb{P}(\mathbf{x} = 1) + \mathbb{P}(\mathbf{x} = -1) = 1,$$

para todo $k \in \mathbb{N}$ par.

Para calcular los momentos de una variable aleatoria X, una estrategía es conocer la función generadora de momentos, que se define como

$$M_{\mathbf{X}}(t) = \mathbb{E}(e^{t\mathbf{X}}),$$

siempre que esta esperanza exista. La función generadora de momentos se llama así porque, si existe en un entorno de $t = 0$, permite generar los momentos de la distribución de probabilidad:

$$\mathbb{E}\left(\mathbf{X}^n\right) = M_{\mathbf{X}}^{(n)}(0) = \frac{d^n M_{\mathbf{X}}}{dt}(0). \tag{4.11}$$

Sea $\mathbf{g} \sim N(0,1)$ una variable aleatoria normal estandar, sabemos que la función generadora de momentos es

$$M_{\mathbf{g}}(t) = e^{\frac{t^2}{2}}.$$

Luego de un simple cálculo usando (4.11), conseguimos que $\mathbb{E}[\mathbf{g}^4] = 3$ y $\mathbb{E}[\mathbf{g}^2] = 1$, entonces $\mathbf{g}$ es 3-razonable.

Si $\mathbf{u} \sim [-1,1]$ es uniforme, entonces se puede calcular que es $\frac{9}{5}$-razonable siguiendo la misma idea que antes, sabiendo que

$$M_{\mathbf{u}}(t) = E(e^{tx}) = \frac{e^t - e^{-t}}{2t}.$$

En todos estos ejemplos B es una constante "pequeña", y pensamos a estas variables aleatorias simplemente como "razonables". Un ejemplo de alguna que no sea "razonable" en este sentido, está basado en una variable aleatoria de Bernoulli $\mathbf{y}$ altamente sesgada, es decir que cumpla $\mathbb{P}(\mathbf{y} = 1) = 2^{-n}$ y $\mathbb{P}(\mathbf{y} = 0) = 1 - 2^{-n}$, donde n es grande. En este caso tenemos que

$$M_{\mathbf{y}(t)} = \frac{1}{2^n}e^t.$$

Entonces podemos afirmar que esta variable aleatoria $\mathbf{y}$ no es B-razonable a menos que $B \geqslant 2^n$.

Veamos los beneficios de trabajar con variables razonables. Primero, tienen límite de cola ligeramente mejores a las que se obtiene de la desigualdad de Chebyshev.

Proposición 4.2.13. *Sea una variable aleatoria* $\mathbf{X} \not\sim 0$, *B-razonable. Entonces*

$$\mathbb{P}(|\mathbf{X}| \geqslant t\|\mathbf{X}\|_2) \leqslant \frac{B}{t^4} \qquad \textit{para todo } t > 0.$$

Demostración. Esta demostración es inmediata por la desigualdad de Markov:

$$\mathbb{P}(|\mathbf{X}| \geqslant t\|\mathbf{X}\|_2) = \mathbb{P}(\mathbf{X^4} \geqslant t^4\|\mathbf{X}\|_2^4) \leqslant \frac{\mathbb{E}[\mathbf{X}^4]}{t^4\mathbb{E}[\mathbf{X}^2]^2} \leqslant \frac{B}{t^4}.$$

□

Más interesante aún, también satisfacen los límites de anticoncentración; por ejemplo, se puede limitar la probabilidad de que estén cerca de 0.

Proposición 4.2.14. *Sea una variable aleatoria* $\mathbf{X} \not\sim 0$, *B-razonable. Entonces*

$$\mathbb{P}(|\mathbf{X}| > t\|\mathbf{X}\|_2) \geqslant \frac{(1-t^2)^2}{2} \qquad \textit{para todo } t \in [0,1].$$

Antes de demostrar la proposición, enunciemos el teorema de Paley-Zygmund (también conocido como el método de segundo momento).

Teorema 4.2.15. *(Paley-Zygmund) Sea $Z > 0$ una varible aleatoria con varianza finita, y sea $0 \leqslant \theta \leqslant 1$ se tiene*

$$\mathbb{P}(Z > \theta\mathbb{E}[Z]) \geqslant (1-\theta)^2 \frac{\mathbb{E}[Z]^2}{\mathbb{E}[Z^2]}.$$

Demostración. (del teorema de Paley-Zygmund.) Comenzamos observando que

$$\mathbb{E}[Z] = \mathbb{E}[Z\mathbf{1}_{\{Z \leqslant \theta\mathbb{E}[Z]\}}] + \mathbb{E}[Z\mathbf{1}_{\{Z > \theta\mathbb{E}[Z]\}}].$$

El primer término es a lo sumo $\theta\mathbb{E}[Z]$, mientras que el segundo término es como mucho $\mathbb{E}[Z^2]^{\frac{1}{2}}\mathbb{P}(Z > \theta\mathbb{E}[Z])^{\frac{1}{2}}$ por la desigualdad de Cauchy-Schwarz . Luego la conclusión deseada se sigue despejando. □

La demostración de la Proposición 4.2.14 es una simple aplicación del método de segundo momento.

Demostración. Aplicando la Desigualdad de Paley-Zygmund para $\mathbf{X}^2$, tenemos que

$$\mathbb{P}(|\mathbf{X}| \geqslant t\|\mathbf{X}\|_2|) = \mathbb{P}(\mathbf{X}^2 \geqslant t^2\mathbb{E}[\mathbf{X}^2]) \geqslant \frac{(1-t^2)^2\mathbb{E}[\mathbf{X}^2]^2}{\mathbb{E}[\mathbf{X}^4]} \geqslant \frac{(1-t^2)^2}{B},$$

como se quería. □

Para una variable aleatoria discreta $\mathbf{X}$, una simple condición que genera la razonabilidad es que $\mathbf{X}$ tome cada uno de sus valores con probabilidad no despreciable.

Proposición 4.2.16. *Sea $\mathbf{X}$ una variable aleatoria discreta con función de probabilidad π. Escribimos*

$$\lambda = \operatorname{mín}(\pi) = \operatorname*{mín}_{x \in Ran(\mathbf{X})} \{\mathbb{P}(\mathbf{X} = x)\}.$$

Entonces $\mathbf{X}$ es $\frac{1}{\lambda}$-razonable.

Demostración. Sea $M = \|\mathbf{X}\|_\infty$. Como $\mathbb{P}(|\mathbf{X}| = M) \geqslant \lambda$ tenemos que

$$\mathbb{E}[\mathbf{X}^2] \geqslant \lambda M^2,$$

entonces

$$M^2 \leqslant \frac{\mathbb{E}[\mathbf{X}^2]}{\lambda}.$$

Por otro lado

$$\mathbb{E}[\mathbf{X}^4] = \mathbb{E}[\mathbf{X}^2\mathbf{X}^2] \leqslant M^2\mathbb{E}[\mathbf{X}^2].$$

Por lo tanto, $\mathbb{E}[\mathbf{X}^4] \leqslant \frac{1}{\lambda}\mathbb{E}[\mathbf{X}^2]^2$ que es la condición de $\frac{1}{\lambda}$-razonabilidad deseada. □

La recíproca para la proposición anterior es falsa. Por ejemplo, si

$$\mathbf{X} = \frac{1}{\sqrt{n}}\mathbf{x}_1 + ... + \frac{1}{\sqrt{n}}\mathbf{x}_n,$$

donde $\mathbf{x}_i \sim \{-1,1\}^n$, es decir $\mathbf{x}_i(x) = x_i$ con $x \in \{\pm 1\}^n$ independientes entre sí, entonces por el teorema central del límite $\mathbf{X}$ tiende a una variable aleatoria Gaussiana estandar y es 3-razonable. Por otro lado, el λ (mínimo de la función de probabilidad) para la variable aleatoria $\mathbf{X}$ es un minúsculo 2^{-n}, pues como $X(1,...,1) = \sqrt{n}$ se tiene

$$\sqrt{n} \in ran(X) \ \text{ y } \ \mathbb{P}(X = \sqrt{n}) = \frac{1}{2^n}.$$

Esta discusión plantea la cuestión de cómo se podría intentar construir una variable aleatoria uniformemente independiente a ± 1 bits que sea no razonable. Por la Proposición 4.2.16, al menos debe usar muchos de ellos. Por ejemplo, para construir la variable aleatoria no razonable, como en el Ejemplo 4.2.12 se requiere grado n: $\mathbf{y} = \frac{\mathbf{(1+x_1)(1+x_2)...(1+x_n)}}{\mathbf{2^n}}$.

Enunciemos a continuación el Lema de Bonami:

Lema 4.2.17. *(Lema de Bonami) Para cada $k \in \mathbb{N}$, sea $f : \{\pm 1\}^n \to \mathbb{R}$ con grado k, y sean $\mathbf{x}_1, ..., \mathbf{x}_n$ variables aleatorias independientes, uniformes a ± 1, entonces la variable aleatoria $f(\mathbf{x})$ es 9^k-razonable, es decir*

$$\mathbb{E}[f^4] \leqslant 9^k \mathbb{E}[f^2]^2 \iff \|f\|_4 \leqslant \sqrt{3}^k \|f\|_2.$$

En otras palabras, los polinomios de bajo grado con ± 1 bits uniformemente independientes son razonables. También debemos tener en cuenta que el nombre "Lema de Bonami" no es estandar, sin embargo, el resultado fue primero demostrado por Bonami y a menudo se utiliza como lema, por lo que el nombre encaja.

Demostración. Asumimos $k \geqslant 1$ de lo contrario f debe ser constante y es claramente trivial. La demostración es por inducción en n. Si $n = 0$, entonces f debe ser una constante y es claro que la afirmación es cierta. Para $n \geqslant 1$ consideramos la descomposición $f(x) = x_n D_n f(x) + E_n f(x)$ donde $deg(D_n f) \leqslant k-1$, $deg(E_n f) \leqslant k$, tanto el polinomio $D_n f(x)$ como $E_n f(x)$ no depende de x_n. Para abreviar escribimos $f := f(x)$, $d := D_n f(x)$ y $e := E_n f(x)$. Entonces

$$\begin{aligned}
\mathbb{E}[f^4] &= \mathbb{E}[(\mathbf{x}_n d + e)^4] \\
&= \mathbb{E}[\mathbf{x}_n^4 d^4] + 4\mathbb{E}[\mathbf{x}_n^3 d^3 e] + 6\mathbb{E}[\mathbf{x}_n^2 d^2 e^2] + 4\mathbb{E}[\mathbf{x}_n d^3 e^3] + \mathbb{E}[e^4] \\
&= \mathbb{E}[\mathbf{x}_n^4]\mathbb{E}[d^4] + 4\mathbb{E}[\mathbf{x}_n^3]\mathbb{E}[d^3 e] + 6\mathbb{E}[\mathbf{x}_n^2]\mathbb{E}[d^2 e^2] + 4\mathbb{E}[\mathbf{x}_n]\mathbb{E}[d^3 e^3] + \mathbb{E}[e^4].
\end{aligned}$$

En el último paso usamos que $\mathbf{x}_n$ es independiente de d y e, ya que $D_n f$ y $E_n f$ no dependen de x_n. Ahora como $\mathbb{E}[\mathbf{x}_n] = \mathbb{E}[\mathbf{x}_n^3] = 0$ y $\mathbb{E}[\mathbf{x}_n^2] = \mathbb{E}[\mathbf{x}_n^4] = 1$, se deduce:

$$\mathbb{E}[f^4] = \mathbb{E}[d^4] + 6\mathbb{E}[d^2 e^2] + \mathbb{E}[e^4]. \tag{4.12}$$

Con una secuencia similar de pasos se puede probar que

$$\mathbb{E}[f^2] = \mathbb{E}[d^2] + \mathbb{E}[e^2]. \tag{4.13}$$

Para acotar por arriba (4.12) recordemos que $d = D_n f(x)$ donde $D_n f(x)$ es un polinomio tetraedral de grado a lo sumo $k-1$ dependiendo de $n-1$ variables. Por lo tanto, podemos aplicar hipótesis inductiva $\mathbb{E}[d^4] \leqslant 9^{k-1}\mathbb{E}[d^2]^2$. Similarmente, $\mathbb{E}[e^4] \leqslant 9^k \mathbb{E}[e^2]^2$ ya que el grado de $E_n f$ es menor o igual que k.

Para acotar $\mathbb{E}[d^2 e^2]$ aplicamos Cauchy-Schwarz obteniendo $\sqrt{\mathbb{E}[d^4]}\sqrt{\mathbb{E}[e^4]}$ y usando inducción nuevamente tenemos

$$\begin{aligned}\mathbb{E}[f^4] &\leqslant 9^{k-1}\mathbb{E}[d^2]^2 + 6\sqrt{9^{k-1}\mathbb{E}[d^2]^2}\sqrt{9^k\mathbb{E}[e^2]^2} + 9^k\mathbb{E}[e^2]^2 \\ &\leqslant 9^k\left(\mathbb{E}[d^2]^2 + 2\mathbb{E}[d^2]\mathbb{E}[d^2] + \mathbb{E}[e^2]^2\right) = 9^k\left(\mathbb{E}[d^2] + \mathbb{E}[e^2]]\right)^2,\end{aligned}$$

donde usamos $9^{k-1}\mathbb{E}[d^2]^2 \leqslant 9^k\mathbb{E}[d^2]^2$. Luego por (4.13), la demostración está completa. □

Una consecuencia inmediata del Lema de Bonami es que para toda $f : \{\pm 1\}^n \to \mathbb{R}$ k-homogénea, con $k \in \mathbb{N}$, se tiene

$$\|T_{\frac{1}{\sqrt{3}}} f\|_4 = \frac{1}{\sqrt{3}^k}\|f\|_4 \leqslant \|f\|_2. \tag{4.14}$$

Este es un caso especial del $(2,4)$ -Terorema de hipercontractividad (cuyo nombre se explicará en breve), el cual dice que el supuesto de homogeneidad de grado k no es necesario:

Teorema 4.2.18. $(2,4)$-*Teorema de hipercontractividad. Sea* $f : \{\pm 1\}^n \to \mathbb{R}$*, entonces*

$$\|T_{\frac{1}{\sqrt{3}}} f\|_4 \leqslant \|f\|_2.$$

La demostración es simplemente repetir la técnica de inducción utilizada para el Lema de Bonami.

Demostración. Probaremos $\mathbb{E}[T_{\frac{1}{\sqrt{3}}} f(\mathbf{x})^4] \leqslant \mathbb{E}[f(\mathbf{x})^2]^2$ usando inducción de la misma manera que antes. Mantenemos las notaciones $d := D_n f(x)$, $e := E_n f(x)$ y agregamos $T := T_{\frac{1}{\sqrt{3}}}$, tenemos

$$Tf = \mathbf{x}_n \cdot \frac{1}{\sqrt{3}} Td + Te.$$

Por cálculos similares a los hechos en la prueba del Lema de Bonami obtenemos

$$\begin{aligned}\mathbb{E}[(Tf)^4] &= \left(\frac{1}{\sqrt{3}}\right)^4 \mathbb{E}[(Td)^4] + 6\left(\frac{1}{\sqrt{3}}\right)^2 \mathbb{E}[(Td)^2(Te)^2] + \mathbb{E}[(Te)^4] \\ &\leqslant \mathbb{E}[(Td)^4] + 2\mathbb{E}[(Td)^2(Te)^2] + \mathbb{E}[(Te)^4] \\ &\leqslant \mathbb{E}[(Td)^4] + 2\sqrt{\mathbb{E}[(Td)^4]}\sqrt{\mathbb{E}(Te)^4]} + \mathbb{E}[(Te)^4] \\ &\leqslant \mathbb{E}[d^2]^2 + 2\mathbb{E}[d^2]\mathbb{E}[e^2] + \mathbb{E}[e^2]^2 \\ &= (\mathbb{E}[d^2] + \mathbb{E}[e^2])^2 = \mathbb{E}[f^2]^2.\end{aligned}$$

Notar que aplicamos Cauchy-Schwarz en la segunda desigualdad, en la tercera inducción, y la desigualdad final es un simple cálculo análogo a los hechos en el lema anterior. □

El nombre "hipercontractividad" en este teorema describe que no sólo $T_{\frac{1}{\sqrt{3}}}$ es una "contracción" en $L^2(\{\pm 1\}^n)$ (en el sentido $\|T_{\frac{1}{\sqrt{3}}} f\|_2 \leqslant \|f\|_2$ para todo f) incluso es una contracción cuando se ve como un operador de $L^2(\{\pm 1\}^n)$ a $L^4(\{\pm 1\}^n)$. Pensamos en los teoremas de hipercontractividad como cuantificadores de en qué medida el operador T_ρ es "razonable". Mediante un truco simple del análisis, usando el hecho de que $T_{\frac{1}{\sqrt{3}}}$ es un operador autoadjunto podemos "voltear la norma 2" usando la desigualdad de Hölder:

Teorema 4.2.19. $(4/3, 2)$-*Teorema de hipercontractividad. Sean* $f : \{\pm 1\}^n \to \mathbb{R}$, *entonces*

$$\|T_{\frac{1}{\sqrt{3}}} f\|_2 \leqslant \|f\|_{4/3};$$

i.e.,

$$\mathbf{Stab}_{\frac{1}{3}}[f] \leqslant \|f\|^2_{\frac{4}{3}}.$$

Demostración. Escribimos $T = T_{\frac{1}{\sqrt{3}}}$ para abreviar notación, tenemos

$$\|Tf\|_2^2 = \langle Tf, Tf\rangle = \langle f, TTf\rangle \leqslant \|f\|_{4/3}\|TTf\|_4 \leqslant \|f\|_{4/3}\|Tf\|_2,$$

por la desigualdad de Hölder y el $(2, 4)$-Teorema de hipercontractividad. Dividiendo por $\|Tf\|_2$ (que podemos suponer que no es cero) completamos la demostración. □

4.2.3. $(2, q)-$ y $(p, 2)-$hipercontractividad para un bit

Aunque se pueden obtener mucho resultados al estudiar la norma 4 de estas variables aleatorias, también es natural considerar otras normas. Por ejemplo, obtendríamos versiones mejoradas de los resultados de concentración 4.2.13 y anticoncentración 4.2.14, si pudieramos vincular las normas más altas de una variable aleatoria en términos de su norma 2. Como veremos, también podemos obtener "desigualdades de nivel k" más fuertes al limitar la $(2+\epsilon)$-norma de una función Booleana para pequeños $\epsilon > 0$. Comenzamos con la 4-norma debido a la simplicidad de las pruebas del Lema de Bonami y el $(2, 4)$- Teorema de hipercontractividad. Para generalizar estos resultados a otras normas, es un poco más específico trabajar con estos últimos. En parte, esto se debe a que es "formalmente más fuerte" como veremos luego. Pero la razón principal es que la versión de hipercontractividad alivia el problema poco formal de ser "B-razonable", que no es invariante por traslaciones. Por lo tanto, en lugar de generalizar la condición $\|\rho\mathbf{X}\|_4 \leqslant \|\mathbf{X}\|_2$ ($\mathbf{X}$ es "ρ^{-4}-razonable") generalizaremos la siguiente condición $\|a + \rho b\mathbf{X}\|_4 \leqslant \|a + b\mathbf{X}\|_2$.

Definición 4.2.20. *Sea* $1 \leqslant p \leqslant q \leqslant \infty$ *y sea* $0 \leqslant \rho < 1$. *Decimos que una variable aleatoria real* $\mathbf{X}$ *(con* $\|\mathbf{X}\|_q \leqslant \infty$*) es* (p, q, ρ)*-hipercontractiva si*

$$\|a + \rho b\mathbf{X}\|_q \leqslant \|a + b\mathbf{X}\|_p, \quad \textit{para todos } a, b \in \mathbb{R}.$$

Observar que por homogeneidad, es suficiente chequear la condición para $a = 1$, $b \in \mathbb{R}$ ó para $a \in \mathbb{R}$, $b = 1$. Además es cierto que si $\mathbf{X}$ es (p, q, ρ)-hipercontractiva entonces es $(p, q, \tilde{\rho})$-hipercontractiva para $\tilde{\rho} < \rho$.

También se puede demostrar que si $\mathbf{X}$ es hipercontractiva, entonces $\mathbb{E}[\mathbf{X}]$ debe ser 0 (considerar la expansión Taylor para $\|1 + \rho\epsilon\mathbf{X}\|_r$ alrededor de $\epsilon = 0$, notando que $\rho < 1$ por definición). Por lo tanto, la hipercontractividad, como la razonabilidad, no es una noción invariante por traslaciones. Sin embargo, el hecho de que la definición implica la traslación por un arbitrario facilita enormemente las pruebas por inducción. Por ejemplo, una propiedad que obtenemos de la definición (podemos encontrar los pasos de la demostración en [O'D14, pág 312]) es la siguiente:

Proposición 4.2.21. *Sean $\mathbf{X}$ e $\mathbf{Y}$ variables aleatorias independientes (p, q, ρ)-hipercontractivas. Entonces $\mathbf{X} + \mathbf{Y}$ también es (p, q, ρ)-hipercontractiva.*

El caso $n = 1$ del $(2, 4)$- Teorema de hipercontractividad 4.2.18 dice que $\mathbf{x}$ un bit ± 1 uniformemente aleatorio es $(2, 4, 1/\sqrt{3})$-hipercontractivo; el $(4/3, 2)$-Teorema de hipercontractividad 4.2.19 dice que $\mathbf{x}$ es también $(4/3, 2, 1/\sqrt{3})$-hipercontractivo. En lo siguiente desarrollaremos las generalizaciones de estos hechos a $(2, q, \rho)$ - y $(p, 2, \rho)$-hipercontractividad para otros valores de p y q. Observemos que hasta el momento nos centramos principalmente en los casos de $p = 2$ ó $q = 2$. El estudio de la hipercontractividad para $p, q \neq 2$ y para variables aleatorias que no sean uniformes ± 1 no se demostrará, está detallado [O'D14, Capítulo 10].

Ahora consideramos la hipercontractividad de un bit $\mathbf{x}$ uniformemente aleatorio ± 1. Sabemos que $\mathbf{x}$ es $(2, q, 1/\sqrt{3})$-hipercontractivo para $q = 4$, ¿qué ocurre con otros valores de q? Las cosas mejoran cuando q es un entero par porque entonces no es necesario tomar el valor absoluto al calcular $\|a + \rho b\mathbf{X}\|_q$. Por esto intentamos con $q = 6$.

Proposición 4.2.22. *Para $\mathbf{x}$ un bit uniforme ± 1, tenemos $\|a + \rho b\mathbf{x}\|_6 \leqslant \|a + b\mathbf{x}\|_2$ para todo $a, b \in \mathbb{R}$ si (y sólo si) $\rho \leqslant 1/\sqrt{5}$. Es decir, $\mathbf{x}$ es $(2, 6, 1/\sqrt{5})$- hipercontractivo.*

Demostración. Elevando la desigualdad a la potencia sexta, tenemos que mostrar

$$\mathbb{E}[(a + \rho b\mathbf{x})^6] \leqslant \mathbb{E}[(a + b\mathbf{x})^2]^3. \tag{4.15}$$

Este resultado es trivial cuando $a = 0$; de lo contrario suponemos que $a = 1$ por homogeneidad. Expandimos ambas cantidades dentro de las esperanzas y utilizamos el hecho de que $\mathbb{E}[\mathbf{x}^k]$ es 0 cuando k es impar y 1 cuando k es par. Por lo tanto (4.15) es equivalente a

$$1 + 15\rho^2 b^2 + 15\rho^4 b^4 + \rho^6 b^6 \leqslant (1 + b^2)^3 = 1 + 3b^2 + 3b^4 + b^6. \tag{4.16}$$

Comparando los dos lados término por término, vemos que el coeficiente de b^2 es el factor limitante: para que (4.16) se mantenga para todo $\rho \in \mathbb{R}$ es suficiente que $15\rho^2 \leqslant 3$; es decir, $\rho \leqslant 1/\sqrt{5}$. Al considerar $b \to 0$ también es fácil ver que esta condición es necesaria. $\square$

Si se repite este análisis para el caso $q = 8$, se encontrará que nuevamente el factor limitante es el coeficiente de b^2, y que $\mathbf{x}$ es $(2, 8, \rho)$-hipercontractiva si (y sólo si) $\binom{8}{2}\rho^2 \leqslant \binom{4}{1}$; i.e., $\rho \leqslant 1/\sqrt{7}$. En base a esto, es natural proponer lo siguiente:

Teorema 4.2.23. *Sea* $\mathbf{x}$ *un bit uniforme* ± 1 *y sea* $q \in (2, \infty]$. *Entonces*

$$\|a + \rho b\mathbf{x}\|_q \leqslant \|a + b\mathbf{x}\|_2,$$

para todo $a, b \in \mathbb{R}$, *asumiendo* $\rho \leqslant 1/\sqrt{q-1}$.

Afirmaciones equivalentes son $\|a + (1/\sqrt{q-1})b\mathbf{x}\|_q^2 \leqslant a^2 + b^2$, $\mathbf{x}$ *es* $(2, q, 1/\sqrt{q-1})$-*hipercontractiva, y* $\|T_{1/\sqrt{q-1}}\|_q \leqslant \|f\|_2$ *para toda función* $f : \{\pm 1\} \to \mathbb{R}$.

Para un número entero par q no es difícil probar el Teorema 4.2.23 tal como lo hicimos para $q = 6$. Desafortunadamente, obtener el teorema para todos los $q > 2$ reales requiere de más trucos. Una idea natural es intentar avanzar como en la Proposición 4.2.22, utilizando las expansiones en serie para $(1 + \rho b\mathbf{x})^q$ y $(1 + b^2)^{q/2}$ proporcionadas por el teorema del binomio generalizado. Sin embargo, incluso cuando $|b| < 1$ (para que la convergencia no sea un problema) hay una dificultad porque los coeficientes en la expansión de $(1 + b^2)^{q/2}$ son a veces negativo. Afortunadamente, este problema de coeficientes negativos en la expansión de la serie desaparece si se intenta probar la afirmación análoga para el Teorema de $(p, 2, \rho)$-hipercontractivo. Por lo tanto, la hábil prueba del Teorema 4.2.23 procede probando primero dicha afirmación, luego "volteando la norma 2".

Teorema 4.2.24. *Sea* $\mathbf{x}$ *un bit uniforme* ± 1 *y sea* $1 \leqslant p < 2$. *Entonces*

$$\|a + \rho b\mathbf{x}\|_2 \leqslant \|a + b\mathbf{x}\|_p,$$

para todo $a, b \in \mathbb{R}$ *asumiendo que* $0 \leqslant \rho \leqslant \sqrt{p-1}$. *Es decir,* $\mathbf{x}$ *es* $(p, 2, \sqrt{p-1})$-*hipercontractiva.*

Demostración. Como ya comentamos podemos asumir que $a = 1$ y $\rho = \sqrt{p-1}$. También podemos suponer sin pérdida de generalidad que $1 + b\mathbf{x} \geqslant 0$ para $\mathbf{x} \in \{\pm 1\}$ (ver [O'D14, pág 63, 2.34]); i.e., que $|b| \leqslant 1$. Entonces es suficiente probar el resultado para todo $|b| < 1$ porque el caso $|b| = 1$ se sigue por continuidad. Escribiendo $b = \epsilon$, necesitamos probar

$$\|1 + \sqrt{p-1} \cdot \epsilon\mathbf{x}\|_2^p \leqslant \|1 + \epsilon\mathbf{x}\|_p^p, \tag{4.17}$$

$$\iff \mathbb{E}[(1 + \sqrt{p-1} \cdot \epsilon\mathbf{x})^2]^{\frac{p}{2}} \leqslant \mathbb{E}[(1 + \epsilon\mathbf{x})^p]. \tag{4.18}$$

Aquí pudimos sacar el valor absoluto en el lado derecho porque $|\epsilon| < 1$. El lado izquierdo de (4.17) es

$$(1 + (p-1)\epsilon^2)^{p/2} \leqslant 1 + \frac{p(p-1)}{2}\epsilon^2, \tag{4.19}$$

donde usamos la desigualdad $(1 + t)^\theta \leqslant 1 + \theta t$ para $t \geqslant 0$ y $0 \leqslant \theta \leqslant 1$. En cuanto al lado derecho de (4.17), ya que $|\epsilon\mathbf{x}| < 1$ podemos usar la generalización del teorema binomial para mostrar que es igual a

$$\begin{aligned}
&\mathbb{E}\left[1 + p\epsilon\mathbf{x} + \frac{p(p-1)}{2!}\epsilon^2\mathbf{x}^2 + \frac{p(p-1)(p-2)}{3!}\epsilon^3\mathbf{x}^3 + \frac{p(p-1)(p-2)(p-3)}{4!}\epsilon^4\mathbf{x}^4 + \cdots\right] \\
&= 1 + p\epsilon\mathbb{E}[\mathbf{x}] + \frac{p(p-1)}{2!}\epsilon^2\mathbb{E}[\mathbf{x}]^2 + \frac{p(p-1)(p-2)}{3!}\epsilon^3\mathbb{E}[\mathbf{x}]^3 + \frac{p(p-1)(p-2)(p-3)}{4!}\epsilon^4\mathbb{E}[\mathbf{x}]^4 + \cdots \\
&= 1 + \frac{p(p-1)}{2!}\epsilon^2 + \frac{p(p-1)(p-2)(p-3)}{4!}\epsilon^4 + \frac{p(p-1)(p-2)(p-3)(p-4)(p-5)}{6!}\epsilon^6 + \cdots.
\end{aligned}$$

Por (4.19), para verificar (4.17) es suficiente notar que cada "término cuadrático" de la expresión anterior,

$$\frac{p(p-1)(p-2)(p-3)\cdots(p-(2k-1))}{(2k)!}\epsilon^{2k},$$

es no negativo. Esto se desprende de que $1 \leqslant p \leqslant 2$ y el numerador tiene dos factores positivos y un número par de factores negativos. □

Para deducir el Teorema 4.2.23 del Teorema 4.2.24, nuevamente necesitamos voltear la norma 2 usando el hecho que T_ρ es autoadjunto. Esto se logra tomando $\Omega = \{\pm 1\}$, $\pi = \pi_{1/2}$ $q = 2$, $T = T_{\sqrt{p-1}}$ y $C = 1$ en la siguiente proposición general:

Proposición 4.2.25. *Sea T un operador autoadjunto en $L^2(\Omega, \pi)$, sean $1 \leqslant p, q \leqslant \infty$ y sean p', q' los índices conjugados de Hölder. Asumimos $\|Tf\|_q \leqslant C\|f\|_p$ para todo f. Entonces $\|Tg\|_{p'} \leqslant C\|g\|_{q'}$ para toda g.*

Demostración. Se sigue por

$$\|Tg\|_{p'} = \sup_{\|f\|_p=1} \langle f, Tg\rangle = \sup_{\|f\|_p=1} \langle Tf, g\rangle \leqslant \sup_{\|f\|_p=1} \|Tf\|_q\|g\|_{q'} \leqslant C\|g\|_{q'},$$

donde la primera igualdad es el caso extremo de la desigualdad de Hölder, la segunda igualdad se cumple porque T es autoadjunta, la posterior desigualdad es de Hölder y la desigualdad final utiliza la hipótesis $\|Tf\|_q \leqslant C\|f\|_p$. □

4.2.4. Hipercontractividad para dos funciones e inducción

Hasta el momento, hemos establecido que si $\mathbf{x}$ es un bit uniforme ± 1, entonces es $(2, q, 1/\sqrt{q-1})$-hipecontractivo y $(p, 2, \sqrt{q-1})$-hipecontractivo. Lo siguiente será por medio de una inducción muy simple transformar estos hechos en los $(2, q)$- y $(p, 2)$-Teoremas de hipercontractividad establecidos en el comienzo del esta sección.

Dicho de otro modo, hemos conseguido que si $f : \{\pm 1\} \to \mathbb{R}$ entonces para todo $p \leqslant 2 \leqslant q$,

$$\|T_{\sqrt{p-1}}f\|_2 \leqslant \|f\|_p \qquad \|T_{\frac{1}{\sqrt{q-1}}}f\|_q \leqslant \|f\|_2$$

Nos gustaría extender estos hechos al caso general $f : \{\pm 1\}^n \to \mathbb{R}$ es decir, establecer los $(p, 2)$ y $(2, q)$- Teoremas de hipercontractividad. Un enfoque natural es la inducción. En el análisis de las funciones Booleanas, hay dos métodos clásicos para probar afirmaciones sobre $f : \{\pm 1\}^n \to \mathbb{R}$ por inducción en n. Un método, que podría ser llamado "inducción por derivadas", utiliza la descomposición $f(x) = x_n D_n f(x) + E_n f(x)$. Vimos este enfoque en la prueba inductiva del Lema de Bonami. El otro método, que podría llamarse "inducción por restricciones", se consigue por las subfunciones $f_{\pm 1}$ obtenidas fijando la n-ésima coordenada de f en ± 1. En ambos métodos reducimos inductivamente de una función f a dos funciones: $D_n f$ y $E_n f$, ó f_{-1} y f_{+1}. Debido a esto, al intentar probar una afirmación por inducción en n, a menudo es útil intentar un hecho generalizado sobre dos funciones. Entonces, para facilitar la inducción, busquemos una versión de dos funciones de las afirmaciones de hipercontractividad que hemos probado hasta ahora. La afirmación más natural que hemos visto es la reformulación de la estabilidad al ruido del $(4/3, 2)$-Teorema de hipercontractividad, $\mathbf{Stab}_{1/3}[f] \leqslant \|f\|_{4/3}^2$. Al menos en el caso $n = 1$, el Teorema 4.2.24 generaliza esto a $\mathbf{Stab}_{p-1}[f] \leqslant \|f\|_p^2$ para $1 \leqslant p \leqslant 2$. Es decir,

$$\mathbf{Stab}_\rho[f] = \underset{\substack{(\mathbf{x},\mathbf{y}) \\ \rho-\text{correlacionado}}}{\mathbb{E}} [f(\mathbf{x})f(\mathbf{y})] \leqslant \|f\|_{1+\rho}^2,$$

para $0 \leqslant \rho \leqslant 1$. Observando esta situación, podríamos hacer una generalización (correcta) para dos funciones $f, g : \{\pm 1\}^n \to \mathbb{R}$

$$\underset{\substack{(\mathbf{x},\mathbf{y}) \\ \rho-\text{correlacionado}}}{\mathbb{E}} [f(\mathbf{x})g(\mathbf{y})] \leqslant \|f\|_{1+\rho}\|g\|_{1+\rho}. \tag{4.20}$$

Para ver que podemos esperar de esta afirmación, separemos la correlación de (4.20) en dos partes; es decir, escribir

$$\rho = \sqrt{rs}, \qquad 0 \leqslant r, s \leqslant 1,$$

y usar

$$\underset{\substack{(\mathbf{x},\mathbf{y}) \\ \sqrt{rs}-\text{correlacionado}}}{\mathbb{E}} [f(\mathbf{x})g(\mathbf{y})] = \mathbb{E}[T_{\sqrt{r}}f \cdot T_{\sqrt{s}}g].$$

Entonces Cauchy-Scwarz implica

$$\underset{\substack{(\mathbf{x},\mathbf{y}) \\ \rho-\text{correlacionado}}}{\mathbb{E}} [f(\mathbf{x})g(\mathbf{y})] = \mathbb{E}[T_{\sqrt{r}}f \cdot T_{\sqrt{s}}g] \leqslant \|T_{\sqrt{r}}f\|_2\|T_{\sqrt{s}}g\|_2 \leqslant \|f\|_{1+r}\|g\|_{1+s}, \tag{4.21}$$

donde en el último paso se utilizó el $(p, 2)-$Teorema de hipercontractividad, que hasta ahora sólo hemos demostrado el caso $n = 1$ en 4.2.24. La desigualdad (4.21), es precisamente la versión deseada de dos funciones de los $(2, p)$ y $(p, 2)$-Teoremas hipercontractivos.

Teorema 4.2.26. *Teorema débil de hipercontractividad para dos funciones.*
Sean $f, g : \{\pm 1\}^n \to \mathbb{R}$, sea $0 \leqslant r, s \leqslant 1$, y asumimos $0 \leqslant \rho \leqslant \sqrt{rs} \leqslant 1$. Entonces

$$\underset{\substack{(\mathbf{x},\mathbf{y}) \\ \rho-\textit{correlacionado}}}{\mathbb{E}} [f(\mathbf{x})g(\mathbf{y})] \leqslant \|f\|_{1+r}\|g\|_{1+s}.$$

Llamamos a esto el Teorema de hipercontractividad para dos funciones "débil" porque la hipótesis $r, s \leqslant 1$ no es realmente necesaria; ver [O'D14, Capítulo 10.1]. Como se mencionó, hasta ahora hemos establecido este teorema en el caso $n = 1$. Sin embargo, lo interesante de la hipercontractividad en este formato es que se extiende al caso general n por una inducción sencilla. La forma de inducción que usaremos es "inducción por restricciones" (vale aclarar que también es posible, pero un poco más complicado, extender el $(2, q)$-Teorema de hipercontractividad de $n = 1$ al caso general mediante "inducción por derivadas"), escribiremos la inducción con notación más general:

Teorema 4.2.27. *(Inducción del Teorema de hipercontractividad para dos funciones) Sea* $0 \leqslant \rho \leqslant 1$ *y asumimos que*

$$\mathop{\mathbb{E}}_{\substack{(\mathbf{x},\mathbf{y})\\ \rho-\text{correlacionado}}}[f(\mathbf{x})g(\mathbf{y})] \leqslant \|f\|_p\|g\|_q,$$

para todas $f, g \in L^2(\Omega,\pi)$. *Entonces la desigualdad es cierta para todas* $f, g \in L^2(\Omega^n, \pi^{\otimes n})$.

Demostración. La demostración es por inducción en n, el caso $n = 1$ se obtiene por suposición. Para $n > 1$, sean $f, g \in L^2(\Omega^n, \pi^{\otimes n})$ y sea $(\mathbf{x},\mathbf{y})$ denota un par ρ-correlacionado sobre $\pi^{\otimes n}$. Usemos la notación $\mathbf{x} = (\mathbf{x}', x)$ donde $\mathbf{x}' = (x_1, ..., x_{n-1})$ y similar notación para $\mathbf{y}$. Notar que $(\mathbf{x}', \mathbf{y}')$ y $(\mathbf{x}_n, \mathbf{y}_n)$ son pares ρ-correlacionados (de longuitud $n-1$ y 1, respectivamente). Escribimos $f_{x_n} = f_{[n-1]}|x_n$ para la restricción de f fijando el valor de la última coordenada para el valor de x_n, y símilarmente para g. Tenemos

$$\mathop{\mathbb{E}}_{\substack{(\mathbf{x},\mathbf{y})\\ \rho-\text{correlacionado}}}[f(\mathbf{x})g(\mathbf{y})] = \mathop{\mathbb{E}}_{(\mathbf{x}_n,\mathbf{y}_n)}\mathop{\mathbb{E}}_{(\mathbf{x}',\mathbf{y}')}[f_{\mathbf{x}_n}(\mathbf{x}')g_{\mathbf{y}_n}(\mathbf{y}')] \leqslant \mathop{\mathbb{E}}_{(\mathbf{x}_n,\mathbf{y}_n)}[\|f_{\mathbf{x}_n}\|_p\|g_{\mathbf{y}_n}\|_q]$$

por inducción. Si escribimos $F \in L^2(\Omega, \pi)$ para la función $x_n \to \|f_{x_n}\|_p$ y similarmente escribimos $G(y_n) = \|g_{y_n}\|_q$, entonces podemos continuar lo anterior como

$$\mathop{\mathbb{E}}_{(\mathbf{x}_n,\mathbf{y}_n)}[\|f_{\mathbf{x}_n}\|_p\|g_{\mathbf{y}_n}\|_q] = \mathop{\mathbb{E}}_{(\mathbf{x}_n,\mathbf{y}_n)}[F(\mathbf{x}_n)G(\mathbf{y}_n)] \leqslant \|F\|_{p,\mathbf{x}_n}\|G\|_{q,\mathbf{x}_n},$$

donde usamos el caso base de la inducción. Finalmente,

$$\|F\|_{p,\mathbf{x}_n} = \mathop{\mathbb{E}}_{\mathbf{x}_n}[|F(\mathbf{x}_n)|^p]^{1/p} = \mathop{\mathbb{E}}_{\mathbf{x}_n}[\|f_{\mathbf{x}_n}\|_p^p]^{1/p} = \left(\mathop{\mathbb{E}}_{\mathbf{x}_n}\mathop{\mathbb{E}}_{\mathbf{x}'}|f_{\mathbf{x}_n}(\mathbf{x}')|^p\right)^{1/p} = \|f\|_p$$

por definición, y similarmente para $\|G\|_{q,\mathbf{x}_n}$. Por lo tanto tenemos $\mathbb{E}[f(\mathbf{x})g(\mathbf{y})] \leqslant \|f\|_p\|g\|_q$, completando la inducción. $\square$

De manera más general, si asumimos que la desigualdad es cierta sobre cada (Ω_i, π_i) con $i \in [n]$ entonces también la afirmación se mantiene en $(\Omega_1 \times \cdots \times \Omega_n, \pi_1 \otimes \cdots \otimes \pi_n)$; el único cambio necesario para la prueba es de notación.

En este punto, hemos establecido completamente el Teorema débil de hipercontractividad de dos funciones. Al tomar $g = f$ y $r = s = \rho$ en el teorema, obtenemos el $(p, 2)-$ Teorema de hipercontractividad completa enunciado al comienzo del la sección. Finalmente, aplicando la Proposición 4.2.25 también obtenemos el $(2, q)$-Teorema de hipercontractividad para toda $f : \{\pm 1\}^n \to \mathbb{R}$.

4.2.5. Aplicaciones de hipercontractividad

Con el $(2, q)$- y $(p, 2)$-Teorema de hipercontractividad en mano, revisemos algunas aplicaciones que vimos. Comenzamos deduciendo una generalización del Lema de Bonami:

Teorema 4.2.28. *Sea $f : \{\pm 1\}^n \to \mathbb{R}$ tiene grado a lo sumo k. Entonces*

$$\|f\|_q \leqslant \sqrt{q-1}^k \|f\|_2 \text{ para todo } q \geqslant 2.$$

Demostración. Tenemos

$$\|f\|_q^2 = \|T_{1/\sqrt{q-1}} T_{\sqrt{q-1}} f\|_q^2 \leqslant \|T_{\sqrt{q-1}} f\|_2^2,$$

usando el $(2, q)$-Teorema de hipercontractividad. En lo que sigue extendemos la definición de T_ρ para $\rho > 1$ via $T_\rho f = \sum_j \rho^j f^{=j}$, entonces tenemos

$$\|T_{\sqrt{q-1}} f\|_2^2 = \sum_{j=0}^{k} (q-1)^j \mathbb{W}^j[f] \leqslant (q-1)^k \sum_{j=0}^{k} \mathbb{W}^j[f] = (q-1)^k \|f\|_2^2,$$

donde $\mathbb{W}^k[f] = \sum_{S\subset[n],|S|=k} \hat{f}(S)^2$ para todo $k \in [n]$ y el resultado se deduce. □

Considerando un truco similar al que usamos en la prueba del $(4/3, 2)$-Teorema de hipercontractividad, se puede deducir que $\|f\|_2 \leqslant (1/\sqrt{p-1})^k \|f\|_p$ cuando f tiene grado k y para cualquier $1 < p \leqslant 2$. Sin embargo, un truco diferente produce un resultado estrictamente mejor, incluyendo el caso $p = 1$:

Teorema 4.2.29. *Sea $f : \{\pm 1\}^n \to \mathbb{R}$ tiene grado a lo sumo k. Entonces $\|f\|_2 \leqslant e^k \|f\|_1$. Más generalmente, para $1 \leqslant p \leqslant 2$ se tiene que $\|f\|_2 \leqslant (e^{\frac{2}{p}-1})^k \|f\|_p$.*

Demostración. Probaremos la afirmación para la norma 1. Para $\epsilon > 0$, sea $0 < \theta < 1$ la solución de $\frac{1}{2} = \frac{\theta}{1} + \frac{1-\theta}{2+\epsilon}$. Aplicando la versión general de la desigualdad de Hölder y el Teorema 4.2.28, obtenemos

$$\|f\|_2 \leqslant \|f\|_{2+\epsilon}^{1-\theta} \|f\|_1^\theta \leqslant \sqrt{1+\epsilon}^{k(1-\theta)} \|f\|_2^{1-\theta} \|f\|_1^\theta.$$

Dividiendo por $\|f\|_2^{1-\theta}$(podemos asumir que es no nulo) y luego de elevar el resultado a la potencia de $1/\theta$ conseguimos

$$\|f\|_2 \leqslant \left((1+\epsilon)^{\frac{1-\theta}{2\theta}}\right)^k \|f\|_1 = \left((1+\epsilon)^{\frac{1}{\epsilon}+\frac{1}{2}}\right)^k \|f\|_1.$$

El resultado sigue tomando límite cuando $\epsilon \to 0$. El caso general es similar sólo que debemos tomar a θ como solución de $\frac{1-\theta}{2\theta} = (\frac{2}{p} - 1)\frac{1}{\epsilon} + (\frac{1}{p} - \frac{1}{2})$. □

Bibliografía

[AA14] Scott Aaronson and Andris Ambainis. The need for structure in quantum speedups. *Theory OF Computing*, 10(6):133–166, 2014.

[Bac12] Arturs Backurs. Influences in low-degree polynomials, `http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.304.5820&rep=rep1&type=pdf`. 2012.

[BE95] Peter Borwein and Tamás Erdélyi. *Polynomials and polynomial inequalities*, volume 161. Springer Science & Business Media, 1995.

[BH31] Henri Frédéric Bohnenblust and Einar Hille. On the absolute convergence of Dirichlet series. *Annals of Mathematics*, pages 600–622, 1931.

[BK97] Harold Boas and Dmitry Khavinson. Bohr's power series theorem in several variables. *Proceedings of the American Mathematical Society*, 125(10):2975–2979, 1997.

[BKS88] Yu A Brudnyi, SG Krein, and EM Semenov. Interpolation of linear operators. *Journal of Soviet Mathematics*, 42(6):2009–2113, 1988.

[Ble01] Ron Blei. *Analysis in integer and fractional dimensions*, volume 71. Cambridge University Press, 2001.

[Bon68] Aline Bonami. Ensembles $\Lambda(p)$ dans le dual de D^∞. In *Annales de l'institut Fourier*, volume 18, pages 193–204, 1968.

[Bon70] Aline Bonami. Étude des coefficients de fourier des fonctions de $L^p(G)$. In *Annales de l'institut Fourier*, volume 20, pages 335–402, 1970.

[CJRMF$^+$15] JR Campos, P Jiménez-Rodríguez, Gustavo A Muñoz-Fernández, D Pellegrino, and Juan B Seoane-Sepúlveda. On the real polynomial Bohnenblust–Hille inequality. *Linear Algebra and its Applications*, 465:391–400, 2015.

[DF02] Dimiter Dryanov and Richard Fournier. Bound preserving operators over classes of polynomials. *East journal on approximations*, 8(3):327–354, 2002.

[DFOC+11] Andreas Defant, Leonhard Frerick, Joaquim Ortega-Cerdà, Myriam Ounaïes, and Kristian Seip. The Bohnenblust-Hille inequality for homogeneous polynomials is hypercontractive. *Annals of mathematics*, pages 485–497, 2011.

[DM16] Andreas Defant and Mieczysław Mastyło. Bohnenblust–Hille inequalities for Lorentz spaces via interpolation. *Analysis & PDE*, 9(5):1235–1258, 2016.

[DMP18] Andreas Defant, Mieczysław Mastyło, and Antonio Pérez. Bohr's phenomenon for functions on the boolean cube. *Journal of Functional Analysis*, 275(11):3115–3147, 2018.

[DMP19] Andreas Defant, Mieczysław Mastyło, and Antonio Pérez. On the Fourier spectrum of functions on boolean cubes. *Mathematische Annalen*, 374(1-2):653–680, 2019.

[DT89] Seán Dineen and Richard Timoney. Absolute bases, tensor products and a theorem of Bohr. *Studia Mathematica*, 94(3):227–234, 1989.

[DW08] Ronald De Wolf. A brief introduction to Fourier analysis on the Boolean cube. *Theory of Computing*, pages 1–20, 2008.

[FHH+14] Yuval Filmus, Hamed Hatami, Steven Heilman, Elchanan Mossel, Ryan O'Donnell, Sushant Sachdeva, Andrew Wan, and Karl Wimmer. Real analysis in computer science: A collection of open problems. 2014.

[Fou08] Richard Fournier. Asymptotics of the Bohr radius for polynomials of fixed degree. *Journal of mathematical analysis and applications*, 338(2):1100–1107, 2008.

[FR99] Lance Fortnow and John Rogers. Complexity limitations on quantum computation. *Journal of Computer and System Sciences*, 59(2):240–252, 1999.

[Gua05] Zdeňka Guadarrama. Bohr's radius for polynomials in one complex variable. *Computational Methods and Function Theory*, 5(1):143–151, 2005.

[Har72] Lawrence A Harris. Bounds on the derivatives of holomorphic functions of vectors. In *Proc. Colloq. Analysis, Rio de Janeiro*, volume 145, page 163, 1972.

[IDO06] G. Kindler I. Dinur, E. Friedgut and R. O'Donnell. On the Fourier tails of bounded functions over the discrete cube. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, pages 437–446, 2006.

[Kie69] Konrad Kiener. *Uber Produkte von quadratisch integrierbaren Funktionen endlicher Vielfalt*. PhD thesis, PhD thesis, Dissertation, Universität Innsbruck, 1969.

[KK93] Csam Kahane and Jean-Pierre Kahane. *Some random series of functions*, volume 5. Cambridge University Press, 1993.

[Kli95] Maciej Klimek. Metrics associated with extremal plurisubharmonic functions. *Proceedings of the American Mathematical Society*, 123(9):2763–2770, 1995.

[Kwa87] Stanislaw Kwapien. Decoupling inequalities for polynomial chaos. *The Annals of Probability*, pages 1062–1071, 1987.

[Lit30] John E Littlewood. On bounded bilinear forms in an infinite number of variables. *The Quarterly Journal of Mathematics*, (1):164–174, 1930.

[LPV03] László Lovász, József Pelikán, and Katalin Vesztergombi. Discrete mathematics, undergraduate texts in mathematics, 2003.

[McK89] Brendan D McKay. On Littlewood's estimate for the binomial distribution. *Advances in Applied Probability*, 21(2):475–478, 1989.

[Mon12] Ashley Montanaro. Some applications of hypercontractive inequalities in quantum information theory. *Journal of Mathematical Physics*, 53(12):122206, 2012.

[Nat64] Isidor Pavlovich Natanson. *Constructive function theory*, volume 1. Ungar, 1964.

[O'D08] Ryan O'Donnell. Some topics in analysis of Boolean functions. In *Proceedings of the fortieth annual ACM symposium on Theory of computing*, pages 569–578, 2008.

[O'D14] Ryan O'Donnell. *Analysis of Boolean functions*. Cambridge University Press, 2014.

[OZ16] Ryan O'Donnell and Yu Zhao. Polynomial bounds for decoupling, with applications. In *31st Conference on Computational Complexity (CCC 2016)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2016.

[Pal32] Rymond Paley. A remarkable series of orthogonal functions (i). *Proceedings of the London Mathematical Society*, 2(1):241–264, 1932.

[PSS14] F. Bayart, D. Pellegrino and J.B. Seoane-Sepúlveda. The Bohr radius of the n-dimensional polydisk is equivalent to $\sqrt{(\log n)/n}$. *Advances in mathematics*, 264:726–746, 2014.

[Sch69] Michel Schreiber. Fermeture en probabilité de certains sous-espaces d'un espace L^2. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 14(1):36–48, 1969.

[Vis46] C Visser. A generalization of Tchebychef's inequality to polynomials in more than one variable. *Indag. Math*, 8:310–311, 1946.