



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

Tesis de Licenciatura

La desigualdad de Grothendieck

Gabriel Carvajal

Director: Daniel E. Galicer

Diciembre de 2015

Introducción

Vamos a proponer el siguiente juego. Tenemos una grilla de números:

3	-4	3	4	-2
-7	-1	12	-9	0
2	-11	3	-3	6
1	0	5	5	-8
9	-1	8	-10	3

En este momento la suma de todos los números es 8. Supongamos que vale elegir una columna o una fila e invertir los signos de todos los números que están en ella, y que podemos seguir realizando esta operación tantas veces como queramos. ¿Cuál es el máximo valor que podemos alcanzar sumando todos los números obtenidos? ¿Y qué sucede si empezamos a considerar grillas mucho más grandes?

A lo largo de todo este trabajo, vamos a ocuparnos de hablar de este problema. Consideraremos qué tan difícil es hallar el máximo, qué tan difícil es aproximarlo, qué tan difícil es hallar la secuencia de operaciones que nos llevan a él y qué implicaciones tienen estas respuestas más allá de este juego inocente.

Nuestra historia empieza en 1953, cuando Alexander Grothendieck, poco después de escribir su tesis doctoral, publica en una revista brasileña un artículo en francés bajo el título de “Résumé de la théorie métrique des produits tensoriels topologiques” [Gro56]. En este trabajo, Grothendieck desarrolla una teoría sistemática de normas en productos tensoriales de espacios de Banach. Más precisamente, el problema que le interesaba era considerar qué normas se podían usar sobre un producto tensorial espacios de Banach de manera que se conservara algún tipo de compatibilidad con las normas originales de estos espacios. En un momento clave del artículo, Grothendieck necesita demostrar que dos de las normas “naturales” que definió sobre los productos tensoriales son equivalentes entre sí. El resultado, que el mismo autor llama “Teorema fundamental de la teoría métrica de productos tensoriales”, es básicamente una desigualdad entre dos normas

$$\|z\|_{w_2} \leq \|z\|_{/\pi\setminus} \leq K_G \|z\|_{w_2},$$

donde K_G es una constante, y $z \in E \otimes F$, con E, F espacios de Banach. El teorema demuestra la existencia de esta constante, y muestra además que el menor valor posible K_G para la constante cumple que

$$1,57 \approx \frac{\pi}{2} \leq K_G \leq \sinh\left(\frac{\pi}{2}\right) \approx 2,3.$$

El trabajo de Grothendieck pasa desapercibido muchos años, en los que el autor mismo se dedica a trabajar en otras áreas de la matemática, a las que contribuye enormemente.

Pero 15 años más tarde, Joram Lindenstrauss y Aleksander Pełczyński redescubren este trabajo, y encuentran en él muchos resultados valiosos, incluso algunos que respondían preguntas planteadas por la comunidad luego de su publicación. Entre ambos hacen una gran revisión del paper, y le dan nuevas formulaciones e interpretaciones a su contenido en un trabajo publicado en 1968 [LP68]. En particular, postulan una nueva formulación del Teorema Fundamental que dice que, para cualquier matriz $A \in \mathbb{R}^{n \times n}$ y cualquier espacio de Hilbert \mathcal{H} se cumple la siguiente desigualdad:

Teorema (Desigualdad de Grothendieck).

$$\max_{u_i, v_j \in B_{\mathcal{H}}} \sum_{i,j=1}^n a_{ij} \langle u_i, v_j \rangle \leq K_G \left(\max_{\delta_i, \epsilon_j \in \{-1, 1\}} \sum_{i,j=1}^n a_{ij} \delta_i \epsilon_j \right).$$

Los autores demuestran que el valor de la mejor constante K_G para esta desigualdad es exactamente el mismo que en la formulación original de Grothendieck, a pesar de desconocerlo.

Como el lado derecho de la desigualdad corresponde a multiplicar dos vectores con coeficientes en $\{-1, 1\}$ a izquierda y a derecha de una matriz $A \in \mathbb{R}^{n \times n}$, luego el valor máximo alcanzable es exactamente al resultado del juego que introdujimos al inicio sobre la matriz A . La desigualdad de Grothendieck nos da entonces una primera forma de acotar inferiormente el resultado del juego, al que denotaremos como

$$OPT(A) := \max_{\delta_i, \epsilon_j \in \{-1, 1\}} \sum_{i,j=1}^n a_{ij} \delta_i \epsilon_j.$$

Desde el trabajo de Lindenstrauss y Pełczyński, el resultado, que pasó a ser conocido como la desigualdad de Grothendieck, encuentra numerosas aplicaciones en el análisis y paralelamente su constante asociada K_G va adquiriendo distintas interpretaciones en distintos contextos. Sin embargo, su valor exacto sigue siendo un misterio.

En 1978, Jean-Louis Krivine [Kri78] logra demostrar, con técnicas avanzadas de análisis funcional, que

$$K_G \leq \frac{\pi}{2 \ln(1 + \sqrt{2})} \approx 1,78,$$

y conjetura que este era el valor óptimo. En el año 1991, James Reeds [Ree91] demuestra la primera mejora a la cota inferior original de Grothendieck:

$$1,67 < K_G.$$

Desde entonces, pasan varios años sin ninguna mejora respecto al valor óptimo de la constante. Pero afortunadamente, en la última década se registran varios avances, realizados mayormente por investigadores con una fuerte inclinación por la computación, motivados por la versión de Lindenstrauss-Pełczyński del problema. En esta versión, el problema toma la forma de encontrar la mayor diferencia posible entre dos problemas de optimización. Uno de ellos es $OPT(A)$, equivalente a nuestro juego del inicio y con dominio en $\{-1, 1\}$, y el otro con dominio sobre la bola de un espacio de Hilbert, que corresponde al lado izquierdo de la desigualdad. Además veremos que el problema de optimización con $\{-1, 1\}$, que llamaremos el problema de Grothendieck, y

que es NP-hard, está asociado a un problema de búsqueda de la Cut-Norm de una matriz, con variadas aplicaciones en otros algoritmos.

De este nuevo impulso provienen los resultados principales que contaremos en este trabajo.

Por un lado, hablaremos de la primera mejora a la cota superior de Grothendieck desde 1978, que refuta la conjetura de optimalidad de Krivine. Esta demostración, realizada por Braverman, Makarychev, Makarychev y Naor en el 2013 [BMMN13], viene asociada al mejor algoritmo conocido para hallar soluciones aproximadas del problema de Grothendieck. El algoritmo proviene del hecho de que es mucho más sencillo computacionalmente hallar el máximo de la versión vectorial del problema (la que corresponde al lado izquierdo de la desigualdad), que llamaremos la versión semidefinida del problema:

$$SDP(A) := \max_{u_i, v_j \in B_{\mathcal{H}}} \sum_{i,j=1}^n a_{ij} \langle u_i, v_j \rangle.$$

Una vez que tenemos una solución de esta versión del problema, veremos que tenemos una forma aleatoria de redondear los vectores de la solución a valores en $\{-1, 1\}$ de manera que la solución redondeada alcance en esperanza al menos $\frac{1}{K}$ del valor de la versión semidefinida, para algún K garantizado por el método de redondeo. Es decir, tenemos un método *ALG* tal que en esperanza

$$\frac{OPT(A)}{K} \leq \frac{SDP(A)}{K} \leq ALG(A) \leq OPT(A).$$

Vamos a concentrarnos entonces en encontrar el método de redondeo que garantice el mejor K , lo que nos permitirá demostrar simultáneamente una cota superior para la constante de Grothendieck $K_G \leq K$.

Por otro lado, esto nos pone una restricción sobre las garantías de este tipo de algoritmos de redondeo. Dado que $K_G \leq K$, el mejor algoritmo que podríamos hallar de esta manera no podría garantizar una aproximación mejor que

$$\frac{OPT(A)}{K_G} \leq ALG(A) \leq OPT(A).$$

A priori, podría parecer entonces que no es tan conveniente aproximar $OPT(A)$ usando estos métodos de redondeo. Quizás podríamos conseguir una garantía de aproximación mejor que K_G aproximando de una manera completamente distinta.

Sin embargo, veremos un resultado del año 2009 de Raghavendra y Steurer [RS09] que, bajo cierta hipótesis reciente de complejidad conocida como la Unique Games Conjecture [Kho02], demuestra que el problema de Grothendieck no es aproximable por una mejor constante multiplicativa que K_G . Esto parece sugerir que la constante de Grothendieck tiene una nueva interpretación en la teoría de complejidad computacional, y termina de asegurarnos la estrecha relación entre la desigualdad de Grothendieck y el juego que planteamos al principio.

En el primer capítulo, introduciremos formalmente el problema y la desigualdad de Grothendieck, sus distintas formulaciones y un poco de notación pertinente. Demostraremos ahí las cotas originales para la constante de Grothendieck, y veremos las cotas más modernas de Krivine y de Reeds.

En el segundo capítulo, nos dedicaremos a estudiar la demostración de Krivine, para luego mejorarla usando los argumentos de Braverman, Makarychev, Makarychev y Naor, y hallaremos

la nueva cota superior para la constante de Grothendieck. También veremos la manera de extraer el algoritmo de redondeo prometido de la nueva demostración. Plantearemos algunas preguntas abiertas que podrían llegar a mejorar este algoritmo y la cota superior, quizás hasta obtener una garantía multiplicativa “perfecta” de K_G .

En el tercer capítulo nos ocuparemos del resultado de Raghavendra y Steurer sobre el rol de la constante K_G en la complejidad computacional del problema de Grothendieck. Para esto, introduciremos la Unique Games Conjecture, y veremos por qué es una hipótesis de complejidad bastante utilizada en la actualidad. Luego veremos una aplicación particular de una técnica para trasladar gaps de problemas de optimización y sus relajaciones semidefinidas a gaps en problemas de decisión complejos.

Por último, como un bonus, queremos dejar algunas aplicaciones algorítmicas interesantes de los métodos de redondeo analizados en este trabajo.

Dada la gran cantidad de material disponible sobre el tema, este trabajo no es más que una introducción al tema de la desigualdad de Grothendieck y, en particular, de los temas que nos atraparon en su estudio. Nos gustaría remarcar especialmente la belleza de que un resultado proveniente del contexto más abstracto del análisis funcional haya desembocado en aplicaciones algorítmicas naturales, apreciación que esperamos que sea compartida por el lector.

Índice general

Introducción	III
1. Demostraciones clásicas	1
1.1. Versión matricial	1
1.2. Versión con productos tensoriales	11
1.3. Cota superior de Krivine	15
1.4. Cota inferior de Reeds	19
2. Cotas superiores para la constante de Grothendieck	25
2.1. Generalización de la demostración de Krivine	25
2.2. Uso algorítmico de los Métodos de Redondeo de Krivine	32
2.3. Conjeturas y resultados parciales de optimalidad	36
2.4. La cota de Krivine no es ajustada	44
2.5. Optimalidad de los Métodos de Redondeo de Krivine	58
3. Complejidad computacional y aproximaciones	61
3.1. Complejidad computacional de hallar $\text{OPT}(A)$	61
3.2. La Unique Games Conjecture y técnicas asociadas	67
3.3. Complejidad computacional de aproximar $\text{OPT}(A)$	74
3.4. Aproximación de la constante de Grothendieck	85
Aplicaciones	91
La Cut-Norm de una matriz	91
Usos de la Cut-Norm en algoritmos	95
Bibliografía	103

Capítulo 1

Demostraciones clásicas

En este capítulo vamos a introducir la desigualdad de Grothendieck, contar brevemente el contexto en el que surgió, y qué otras formas fue tomando. Empezaremos enunciando la formulación matricial de la desigualdad que, si bien no es la original, es la que creemos que es más simple y permite el abordaje inmediato del problema. Seguiremos con una primera demostración de la desigualdad con cotas subóptimas, y luego contaremos un poco el contexto original de la desigualdad en productos tensoriales de espacios normados, e introduciendo algunas herramientas que nos ayudarán a ajustar más las cotas. Por último, nos ocuparemos de las mejores cotas clásicas para la constante de Grothendieck: la cota superior demostrada por Krivine en 1978 [Kri78] (y que fue la mejor cota superior conocida hasta el año 2013) y la cota inferior demostrada por Reeds en 1991 [Ree91], aún vigente.

1.1. Versión matricial

Vamos a empezar enunciando la desigualdad de Grothendieck en su forma quizás más elemental, que se debe a Lindenstrauss-Pelczyński [LP68]:

Desigualdad de Grothendieck 1.1.1 (Forma Matricial). *Existe una constante universal $K \in \mathbb{R}_{>0}$ que cumple que, si $A \in \mathbb{R}^{n \times n}$ es una matriz tal que*

$$\left| \sum_{i,j=1}^n a_{ij} \delta_i \epsilon_j \right| \leq 1 \quad \forall \delta_i, \epsilon_j \in \{-1, 1\},$$

entonces para cualquier espacio de Hilbert \mathcal{H} se tiene que

$$\left| \sum_{i,j=1}^n a_{ij} \langle u_i, v_j \rangle \right| \leq K_G \quad \forall u_i, v_j \in \mathcal{H}, \|u_i\| = \|v_j\| = 1.$$

La menor constante K que hace válido el teorema se denomina la **constante de Grothendieck** y la denotaremos como K_G .

Hasta el día de hoy, el valor exacto de la constante de Grothendieck se desconoce. En la demostración original de la desigualdad [Gro56], Grothendieck demostró que la constante K_G cumple

$$1,57 \approx \frac{\pi}{2} \leq K_G \leq \sinh\left(\frac{\pi}{2}\right) \approx 2,3 \tag{1.1}$$

Desde entonces, muchas nuevas demostraciones fueron apareciendo, con cotas peores y mejores a las originales de Grothendieck. La mejor cota inferior conocida hasta hoy es $1,67 \cdots < K_G$, debida a [Ree91] (ver sección (1.4)). Por mucho tiempo, la mejor cota superior conocida fue $K_G \leq \frac{\pi}{2 \ln(1+\sqrt{2})}$, debida a Krivine [Kri78] (ver sección (1.3)), y se conjeturaba que este era el valor óptimo de la constante. Sin embargo en un trabajo muy reciente [BMMN13] se demostró que la cota de Krivine no es ajustada (ver capítulo (2)). Los autores de ese trabajo desistieron de dar una cota superior explícita para la constante, porque su propio método dejó abiertas preguntas que permitirían mejorarla.

Vamos a concentrarnos un poco en esta formulación matricial de la desigualdad. Hagamos algunas observaciones elementales sobre el problema.

Para empezar, notemos que podemos pensar a la desigualdad de Grothendieck como una relación entre dos problemas de búsqueda de valores óptimos para una matriz $A \in \mathbb{R}^{n \times n}$.

Definición 1.1.2. Dada una matriz $A \in \mathbb{R}^{n \times n}$, notamos

$$OPT(A) := \max_{\delta_i, \epsilon_j \in \{-1, 1\}} \sum_{i, j=1}^n \delta_i a_{ij} \epsilon_j. \quad (1.2)$$

Claramente, el valor $OPT(A)$ se alcanza para al menos una elección particular de δ_i, ϵ_j , ¡hay sólo finitas posibilidades! Sin embargo, el total de estas posibilidades es 2^{m+n} , que crece exponencialmente con el tamaño de la matriz. Esta dificultad y sus implicaciones motivará gran parte de la segunda parte de este trabajo. La notación $OPT(A)$ se usará alternativamente para el valor máximo definido arriba como para el problema general de hallarlo. El contexto permitirá siempre salvar este abuso de notación.

Veamos que una definición alternativa del problema $OPT(A)$ es tomar el máximo sobre todos los $\delta_i, \epsilon_j \in [-1, 1]$.

Está claro que

$$\max_{\delta_i, \epsilon_j \in [-1, 1]} \sum_{i, j=1}^n \delta_i a_{ij} \epsilon_j \geq \max_{\delta_i, \epsilon_j \in \{-1, 1\}} \sum_{i, j=1}^n \delta_i a_{ij} \epsilon_j,$$

ya que estamos tomando el máximo sobre un conjunto más grande. Para ver que vale la recíproca, tomemos una solución maximal $\delta_i, \epsilon_j \in [-1, 1]$ (que existe por la continuidad del problema y la compacidad de $[-1, 1]^{m+n}$). Ahora dejemos fijos todos los valores excepto por un δ_{i_0} . El valor alcanzado por la solución se puede reescribir como:

$$\underbrace{\sum_{\substack{i, j=1 \\ i \neq i_0}}^n \delta_i a_{ij} \epsilon_j}_{cte} + \delta_{i_0} \underbrace{\left(\sum_{j=1}^n a_{i_0 j} \epsilon_j \right)}_{cte}.$$

Como esta función es lineal respecto de δ_{i_0} , alcanza un valor maximal en -1 o 1 . Luego podemos reemplazar δ_{i_0} por -1 o 1 , y que el valor de la fórmula no decrezca. Haciendo esto para cada una de las $m+n$ variables, llegamos a un conjunto $\delta_i, \epsilon_j \in \{-1, 1\}$ que alcanza el valor máximo $OPT(A)$.

Definición 1.1.3. Dada una matriz $A \in \mathbb{R}^{n \times n}$, definimos

$$SDP(A) := \sup_{\mathcal{H} \text{ Hilbert}} \sup_{\substack{u_i, v_j \in \mathcal{H} \\ \|u_i\| = \|v_j\| = 1}} \sum_{i,j=1}^n a_{ij} \langle u_i, v_j \rangle. \quad (1.3)$$

Notemos que, sin pérdida de generalidad, podemos considerar que el espacio \mathcal{H} es siempre \mathbb{R}^{2n} . Para ver esto, simplemente tomemos el espacio vectorial generado por los $2n$ vectores u_i, v_j . Al ser un espacio de Hilbert de dimensión $\leq 2n$, resulta ser isométrico a un subespacio de \mathbb{R}^{2n} .

De la misma manera que lo hicimos para $OPT(A)$, podemos ver que la función $\sum_{i,j=1}^n a_{ij} \langle u_i, v_j \rangle$ es afín sobre cada variable u_i o v_j . Luego, el valor $SDP(A)$ no aumenta si permitimos tomar vectores u_i, v_j con norma menor a 1.

Otro punto importante a notar es que cualquier asignación de valores de para el problema $OPT(A)$ constituye una asignación válida para el problema $SDP(A)$. Para ver esto, basta con tomar $\mathcal{H} := \mathbb{R}$ y usar una solución maximal de $OPT(A)$ para el problema $SDP(A)$. Esto nos dice, en particular, que

$$OPT(A) \leq SDP(A). \quad (1.4)$$

Esta última observación podría sugerir que el problema de hallar el valor óptimo o una solución maximal de $SDP(A)$ es más difícil que hacer lo propio para $OPT(A)$ ya que, después de todo, estamos buscando la solución óptima en un conjunto más grande de soluciones válidas. Veremos más adelante que ocurre exactamente lo contrario (el problema $SDP(A)$ está en P, mientras que el problema $OPT(A)$ es NP-hard) y algunas implicaciones de este hecho.

Con las definiciones de los problemas $OPT(A)$ y $SDP(A)$ en mente, podemos llegar a la siguiente formulación de la desigualdad de Grothendieck, quizás su forma más “computacionalmente amistosa”:

Desigualdad de Grothendieck 1.1.4. *Existe una constante universal $K_G \in \mathbb{R}_{>0}$ tal que:*

$$SDP(A) \leq K_G \cdot OPT(A)$$

para todas las matrices A (independientemente de la dimensión).

La desigualdad de Grothendieck vale, con exactamente la misma constante, si permitimos que las matrices A sean rectangulares. Para probar esto, sin pérdida de generalidad, basta notar que, dada $A \in \mathbb{R}^{m \times n}$ con $m > n$, los problemas $OPT(A)$ y $SDP(A)$ alcanzan los mismos valores máximos que $OPT(\tilde{A})$ y $SDP(\tilde{A})$ cuando $\tilde{A} \in \mathbb{R}^{m \times m}$ se define como

$$\tilde{a}_{ij} := \begin{cases} a_{ij} & i \leq n \\ 0 & i > n \end{cases}.$$

Notemos que, para una matriz A dada, el valor de $OPT(A)$ equivale a algunas normas conocidas. Más precisamente:

$$OPT(A) = \max_{\delta_i, \epsilon_j \in [-1, 1]} \sum_{i,j=1}^n \delta_i a_{ij} \epsilon_j = \max_{\delta, \epsilon \in B_{\ell_\infty^n}} \delta A \epsilon^t = \|A\|_{\text{Bil}(\ell_\infty^n, \ell_\infty^n)},$$

donde $Bil(E, F)$ denota el espacio de funciones bilineales en $E \times F$. Como este espacio se identifica isométricamente con $\mathcal{L}(E, F^*)$, tenemos que:

$$OPT(A) = \|A\|_{\mathcal{B}(\ell_\infty^n, \ell_\infty^n)} = \|A\|_{\ell_\infty^n \rightarrow \ell_1^n}.$$

Vamos a contar las demostraciones de las primeras cotas inferior y superior (1.1) que se conocieron para la desigualdad en esta sección. Las demostraciones están basadas en las demostraciones originales de Grothendieck, pero están adaptadas al contexto de la formulación matricial de Lindenstrauss-Pełczyński.

Vamos a empezar por la cota superior, y para esto necesitamos el siguiente lema.

Lema 1.1.5. Sean $u, v \in \mathbb{S}^{k-1}$, y sea G un vector aleatorio en \mathbb{R}^k con distribución Gaussiana estándar. Luego

$$\text{sen} \left(\frac{\pi}{2} \mathbb{E}_G [\text{sgn}(\langle G, u \rangle) \text{sgn}(\langle G, v \rangle)] \right) = \langle u, v \rangle$$

Con este lema, vamos a demostrar que

Teorema 1.1.6.

$$K_G \leq \text{senh} \left(\frac{\pi}{2} \right) \approx 2,3.$$

Notemos que esta será nuestra primera demostración de que K_G existe efectivamente como constante finita, y por lo tanto la primera demostración de la desigualdad de Grothendieck, aunque sea con una constante subóptima.

Demostración. Empecemos con una matriz $A \in \mathbb{R}^{n \times n}$ cualquiera, y sea

$$M := \sup_{\delta_i, \epsilon_j \in \{-1, 1\}} \sum_{i, j=1}^n a_{ij} \delta_i \epsilon_j. \quad (1.5)$$

Entonces tenemos que

$$\left| \sum_{i, j=1}^n a_{ij} \delta_i \epsilon_j \right| \leq M \quad \forall \delta_i, \epsilon_j \in \{-1, 1\}. \quad (1.6)$$

Sea $\{u_1, \dots, u_n, v_1, \dots, v_n\} \subseteq \mathcal{H}$ un conjunto de vectores de norma 1. Sin pérdida de generalidad, podemos suponer que $\mathcal{H} = \mathbb{R}^{2n}$. Queremos ver que

$$\left| \sum_{i, j=1}^n a_{ij} \langle u_i, v_j \rangle \right| \leq K \cdot M,$$

para alguna constante K lo más chica posible.

Apliquemos el Lema (1.1.5) para cada par u_i, v_j , y obtenemos:

$$\begin{aligned}
\sum_{i,j=1}^n a_{ij} \arcsen(\langle u_i, v_j \rangle) &= \sum_{i,j=1}^n a_{ij} \frac{\pi}{2} \mathbb{E}_G [\operatorname{sgn}(\langle u_i, G \rangle) \operatorname{sgn}(\langle v_j, G \rangle)] \\
\sum_{i,j=1}^n a_{ij} \arcsen(\langle u_i, v_j \rangle) &= \frac{\pi}{2} \mathbb{E}_G \left[\sum_{i,j=1}^n a_{ij} \operatorname{sgn}(\langle u_i, G \rangle) \operatorname{sgn}(\langle v_j, G \rangle) \right] \\
\left| \sum_{i,j=1}^n a_{ij} \arcsen(\langle u_i, v_j \rangle) \right| &\leq \frac{\pi}{2} \mathbb{E}_G \left[\left| \sum_{i,j=1}^n a_{ij} \underbrace{\operatorname{sgn}(\langle u_i, G \rangle)}_{\delta_i \in \{-1,1\}} \underbrace{\operatorname{sgn}(\langle v_j, G \rangle)}_{\epsilon_j \in \{-1,1\}} \right| \right] \\
&\leq \frac{\pi}{2} \mathbb{E}_G[M] = M \frac{\pi}{2}.
\end{aligned} \tag{1.7}$$

Nos gustaría tener una forma de aplicar $\operatorname{sen}(-)$ al lado izquierdo de (1.7), para recuperar los productos internos $\langle u_i, v_j \rangle$. Para lograr esto, vamos a usar la serie del seno y acotar separadamente en cada grado.

Definimos, para cada $k \in \mathbb{N}$, $a_{ij}^k = a_{ij} (\arcsen(\langle u_i, v_j \rangle))^k$. Vamos a probar inductivamente que

$$\left| \sum_{i,j=1}^n a_{ij}^k \delta_i \epsilon_j \right| \leq M \left(\frac{\pi}{2} \right)^k \quad \forall \delta_i, \epsilon_j \in \{-1, 1\}. \tag{1.8}$$

El caso $k = 1$ lo acabamos de probar en (1.7). Ahora supongamos que sabemos que (1.8) vale para algún k y demostrémoslo para $k + 1$. Entonces con $\{\delta_i\}_i, \{\epsilon_j\}_j$ fijos tenemos que

$$\begin{aligned}
\left| \sum_{i,j=1}^n a_{ij}^{k+1} \delta_i \epsilon_j \right| &= \left| \sum_{i,j=1}^n a_{ij}^k \arcsen(\langle u_i, v_j \rangle) \delta_i \epsilon_j \right| \\
&= \left| \sum_{i,j=1}^n a_{ij}^k \frac{\pi}{2} \mathbb{E}_G [\operatorname{sgn}(\langle u_i, G \rangle) \operatorname{sgn}(\langle v_j, G \rangle)] \delta_i \epsilon_j \right| \\
&\leq \frac{\pi}{2} \mathbb{E}_G \left[\left| \sum_{i,j=1}^n a_{ij}^k \underbrace{\operatorname{sgn}(\langle u_i, G \rangle)}_{\tilde{\delta}_i \in \{-1,1\}} \underbrace{\operatorname{sgn}(\langle v_j, G \rangle)}_{\tilde{\epsilon}_j \in \{-1,1\}} \right| \right] \\
&\leq \frac{\pi}{2} M \left(\frac{\pi}{2} \right)^k = M \left(\frac{\pi}{2} \right)^{k+1} \quad \text{por hipótesis inductiva.}
\end{aligned}$$

En particular, esto significa que $\forall k \in \mathbb{N}$

$$\left| \sum_{i,j=1}^n a_{ij} \arcsen(\langle u_i, v_j \rangle)^k \right| \leq \left| \sum_{i,j=1}^n a_{ij}^k \underbrace{1}_{\text{"}\delta_i\text{"}} \cdot \underbrace{1}_{\text{"}\epsilon_j\text{"}} \right| \leq M \left(\frac{\pi}{2} \right)^k.$$

Y ahora sí, podemos usar esto para aplicar a (1.7) la serie del seno y “recuperar” los productos internos.

$$\begin{aligned}
\left| \sum_{i,j=1}^n a_{ij} \langle u_i, v_j \rangle \right| &= \left| \sum_{i,j=1}^n a_{ij} \operatorname{sen}(\operatorname{arc\,sen}(\langle u_i, v_j \rangle)) \right| \\
&= \left| \sum_{i,j=1}^n a_{ij} \left(\sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} \operatorname{arc\,sen}(\langle u_i, v_j \rangle)^{2k+1} \right) \right| \\
&\leq \sum_{k=0}^{\infty} \frac{1}{(2k+1)!} \left| \sum_{i,j=1}^n a_{ij} \operatorname{arc\,sen}(\langle u_i, v_j \rangle)^{2k+1} \right| \\
&\leq \sum_{k=0}^{\infty} \frac{1}{(2k+1)!} M \left(\frac{\pi}{2} \right)^k = M \sum_{k=0}^{\infty} \frac{\left(\frac{\pi}{2} \right)^k}{(2k+1)!} = M \operatorname{senh} \left(\frac{\pi}{2} \right).
\end{aligned}$$

□

Ahora vamos a probar la cota inferior de $\frac{\pi}{2}$. Para esto vamos a discutir brevemente un subproblema del problema de Grothendieck. En nuestra formulación computacional (1.1.4), permitíamos que A sea cualquier matriz real. Pero también puede tener interés hallar la constante óptima K_M cuando A pertenece a algún subconjunto más pequeño de matrices $M \subseteq \mathcal{M}(\mathbb{R})$. Como la desigualdad de Grothendieck tiene que valer en cualquier conjunto de matrices, en particular vale para todas las matrices del subconjunto. Esto nos dice que siempre $K_M \leq K_G$, pero la desigualdad podría llegar a ser estricta. En este caso, vamos a restringirnos a que M pertenezca al conjunto de las matrices semidefinidas positivas, y veremos que allí podemos encontrar la constante exacta $K_M = \frac{\pi}{2}$ que separa a los problemas *OPT* y *SDP*.

Para esto vamos a valernos de un nuevo lema ligeramente similar a (1.1.5):

Lema 1.1.7. Sean $u, v \in \mathbb{S}^{k-1}$, y sea G un vector aleatorio en \mathbb{R}^k con distribución Gaussiana estándar. Luego

1. $\mathbb{E}_G [\langle u, G \rangle \langle v, G \rangle] = \langle u, v \rangle$
2. $\sqrt{\frac{\pi}{2}} \mathbb{E}_G [\operatorname{sgn}(\langle u, G \rangle) \langle v, G \rangle] = \langle u, v \rangle$
3. $\frac{\pi}{2} \mathbb{E}_G [\operatorname{sgn}(\langle u, G \rangle) \operatorname{sgn}(\langle v, G \rangle)] = \langle u, v \rangle + \mathbb{E}_G \left[\left(\langle u, G \rangle - \sqrt{\frac{\pi}{2}} \operatorname{sgn}(\langle u, G \rangle) \right) \left(\langle v, G \rangle - \sqrt{\frac{\pi}{2}} \operatorname{sgn}(\langle v, G \rangle) \right) \right]$

Demostración. Veamos primero el ítem (1). Como $\|u\| = \|v\| = 1$, podemos expresar $v = \langle u, v \rangle u + \sqrt{1 - \langle u, v \rangle^2} u^\perp$, con $\|u^\perp\| = 1, u \perp u^\perp$. Luego

$$\begin{aligned}
\mathbb{E}_G [\langle u, G \rangle \langle v, G \rangle] &= \mathbb{E}_G \left[\langle u, G \rangle \left(\langle u, v \rangle \langle u, G \rangle + \sqrt{1 - \langle u, v \rangle^2} \langle u^\perp, G \rangle \right) \right] \\
&= \langle u, v \rangle \underbrace{\mathbb{E}_G [\langle u, G \rangle^2]}_{=1} \sqrt{1 - \langle u, v \rangle^2} \underbrace{\mathbb{E}_G [\langle u^\perp, G \rangle]}_{=0} = \langle u, v \rangle,
\end{aligned}$$

porque $\langle u, G \rangle$ y $\langle u^\perp, G \rangle$ tienen distribución normal estándar, y de hecho son independientes.

Valiéndonos de este hecho, podemos seguir con el ítem (2):

$$\begin{aligned}
\mathbb{E}_G [\operatorname{sgn}(\langle u, G \rangle) \langle v, G \rangle] &= \mathbb{E}_G [\operatorname{sgn}(\langle u, G \rangle) \langle v, G \rangle] \\
&= \mathbb{E}_G \left[\operatorname{sgn}(\langle u, G \rangle) \left(\langle u, v \rangle \langle u, G \rangle + \sqrt{1 - \langle u, v \rangle^2} \langle u^\perp, G \rangle \right) \right] \\
&= \langle u, v \rangle \underbrace{\mathbb{E}_G [|\langle u, G \rangle|]}_{=\sqrt{\frac{2}{\pi}}} + \left(\sqrt{1 - \langle u, v \rangle^2} \right) \underbrace{\mathbb{E}_G [\operatorname{sgn}(\langle u, G \rangle)]}_{=0} \underbrace{\mathbb{E}_G [\langle u^\perp, G \rangle]}_{=0} \\
&= \sqrt{\frac{2}{\pi}} \langle u, v \rangle,
\end{aligned}$$

porque, nuevamente, $\langle u, G \rangle$ y $\langle u^\perp, G \rangle$ tienen distribuciones Gaussianas independientes. Notar que usamos que, en dimensión 1, la esperanza de la norma de una distribución Gaussiana estándar es $\sqrt{\frac{2}{\pi}}$.

Para terminar, demostramos (3) usando (1) y (2):

$$\begin{aligned}
\langle u, v \rangle + \mathbb{E}_G \left[\left(\langle u, G \rangle - \sqrt{\frac{\pi}{2}} \operatorname{sgn}(\langle u, G \rangle) \right) \left(\langle v, G \rangle - \sqrt{\frac{\pi}{2}} \operatorname{sgn}(\langle v, G \rangle) \right) \right] &= \\
= \langle u, v \rangle + \mathbb{E}_G [\langle u, G \rangle \langle v, G \rangle] - \sqrt{\frac{\pi}{2}} \mathbb{E}_G [\langle u, G \rangle \operatorname{sgn}(\langle v, G \rangle)] & \\
- \sqrt{\frac{\pi}{2}} \mathbb{E}_G [\operatorname{sgn}(\langle u, G \rangle) \langle v, G \rangle] + \frac{\pi}{2} \mathbb{E}_G [\operatorname{sgn}(\langle u, G \rangle) \operatorname{sgn}(\langle v, G \rangle)] &= \\
= \langle u, v \rangle + \langle u, v \rangle - \langle u, v \rangle - \langle u, v \rangle + \frac{\pi}{2} \mathbb{E}_G [\operatorname{sgn}(\langle u, G \rangle) \operatorname{sgn}(\langle v, G \rangle)] &= \\
= \frac{\pi}{2} \mathbb{E}_G [\operatorname{sgn}(\langle u, G \rangle) \operatorname{sgn}(\langle v, G \rangle)]. &
\end{aligned}$$

□

Vamos a usar también este lema sobre la distribución uniforme en \mathbb{S}^{k-1} .

Lema 1.1.8. *Si u y v son dos vectores aleatorios independientes en \mathbb{S}^{k-1} distribuidos con densidad uniforme, entonces*

$$\mathbb{E} [\langle u, v \rangle^2] = \frac{1}{k} \quad \text{y} \quad \mathbb{E} [|\langle u, v \rangle|] = \left(\sqrt{\frac{2}{\pi}} + o(1) \right) \frac{1}{\sqrt{k}},$$

donde $o(1)$ representa una expresión en función de k tal que $o(1) \xrightarrow{k \rightarrow \infty} 1$.

Ahora sí, vamos a demostrar que el mayor gap posible entre los valores de $SDP(D)$ y $OPT(D)$ en el caso particular de matrices semidefinidas positivas es exactamente $\frac{\pi}{2}$.

Teorema 1.1.9. *Para toda matriz D semidefinida positiva, se cumple que*

$$SDP(D) \leq \frac{\pi}{2} \cdot OPT(D).$$

Además $\frac{\pi}{2}$ es la menor constante que cumple esta propiedad.

Demostración. Empecemos por notar que, en una matriz semidefinida $D \in \mathbb{R}^{n \times n}$, siempre existe una solución maximal $\{\delta_i\}, \{\epsilon_j\} \subseteq \{-1, 1\}$ de $OPT(D)$ con $\delta_i = \epsilon_i \quad \forall 1 \leq i \leq n$. Para ver esto, tomemos $\{\delta_i\}, \{\epsilon_j\} \subseteq \{-1, 1\}$ una solución maximal de $OPT(D)$ cualquiera. Como D es semidefinida positiva, resulta que existe $C \in \mathbb{R}^{n \times n}$ tal que $D = C^t C$. Luego

$$OPT(D) = |\delta D \epsilon^t| = |\delta C^t C \epsilon^t| = |\langle C \delta^t, C \epsilon^t \rangle| \leq \|C \delta^t\| \|C \epsilon^t\|,$$

por la desigualdad de Cauchy-Schwartz. Pero por otro lado, como multiplicar por δ a izquierda y a derecha es una solución particular de $OPT(D)$,

$$\|C \delta^t\|^2 = |\langle C \delta^t, C \delta^t \rangle| = |\delta C^t C \delta^t| = |\delta D \delta^t| \leq OPT(D).$$

Análogamente, $\|C \epsilon^t\|^2 \leq OPT(D)$. Por lo tanto

$$OPT(D) \leq \|C \delta^t\| \|C \epsilon^t\| \leq OPT(D),$$

y la igualdad sólo se puede cumplir si $\|C \delta^t\| = \|C \epsilon^t\| = \sqrt{OPT(D)}$. Es decir, si $|\delta D \delta^t| = |\epsilon D \epsilon^t| = OPT(D)$, que era lo que queríamos probar.

Ahora veamos que lo mismo ocurre con el problema $SDP(D)$. Para esto, dada una matriz $A \in \mathbb{R}^{n \times n}$, construyamos $A^{(2n)} \in \mathbb{R}^{2n^2 \times 2n^2}$ una matriz de bloques:

$$A^{(2n)} := \underbrace{\begin{bmatrix} A & 0 & \cdots & 0 \\ 0 & A & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & A \end{bmatrix}}_{2n \text{ bloques}}.$$

Con esta notación, está claro que $D^{(2n)}$ es semidefinida positiva, con $D^{(2n)} = \left(C^{(2n)}\right)^t C^{(2n)}$. Dada una evaluación $\{u_i\}, \{v_j\} \subset \mathbb{S}^{2n-1}$ del problema $SDP(D)$, definimos $u, v \in \mathbb{R}^{2n^2}$ como

$$\begin{aligned} u &= (u_1(1), u_2(1), \dots, u_n(1), u_1(2), u_2(2), \dots, u_n(2), \dots, u_1(n), u_2(n), \dots, u_n(n)) \\ v &= (v_1(1), v_2(1), \dots, v_n(1), v_1(2), v_2(2), \dots, v_n(2), \dots, v_1(n), v_2(n), \dots, v_n(n)). \end{aligned}$$

Resulta entonces que

$$\sum_{i,j=1}^n d_{ij} \langle u_i, v_j \rangle = u D^{(2n)} v^t = (C^{(2n)} u^t)^t (C^{(2n)} v^t)$$

En particular, y de forma similar a lo que hicimos anteriormente, si $\{u_i\}, \{v_j\} \subset \mathbb{S}^{2n-1}$ son una evaluación maximal de $SDP(D)$, entonces resulta que

$$SDP(D) = \left| \sum_{i,j=1}^n d_{ij} \langle u_i, v_j \rangle \right| = |u D^{(2n)} v^t| = |u (C^{(2n)})^t C^{(2n)} v| \leq \|(C^{(2n)} u^t)\| \|(C^{(2n)} v^t)\| \leq SDP(D),$$

con $\|(C^{(2n)} u^t)\| = u D^{(2n)} u^t \leq SDP(D)$ y $\|(C^{(2n)} v^t)\| = v D^{(2n)} v^t \leq SDP(D)$. Entonces resulta que:

$$u D^{(2n)} u^t = v D^{(2n)} v^t = SDP(D).$$

Ahora sea $D \in \mathbb{R}^{n \times n}$ una matriz semidefinida cualquiera y apliquemos el ítem (3) de (1.1.7) a una solución óptima del problema $SDP(D)$ (de la que podemos asumir, como acabamos de ver, que $u_i = v_i \quad \forall 1 \leq i \leq n$). Entonces

$$\frac{\pi}{2} \mathbb{E}_G \left[\sum_{i,j=1}^n d_{ij} \operatorname{sgn}(\langle u_i, G \rangle) \operatorname{sgn}(\langle u_j, G \rangle) \right] =$$

$$SDP(D) + \sum_{i,j=1}^n d_{ij} \mathbb{E}_G \left[\underbrace{\left(\langle u_i, G \rangle - \sqrt{\frac{\pi}{2}} \operatorname{sgn}(\langle u_i, G \rangle) \right)}_{=: g_i \in L_2(\Omega)} \underbrace{\left(\langle u_j, G \rangle - \sqrt{\frac{\pi}{2}} \operatorname{sgn}(\langle u_j, G \rangle) \right)}_{=: g_j \in L_2(\Omega)} \right] \leq SDP(D),$$

ya que $\sum_{i,j=1}^n d_{ij} \langle g_i, g_j \rangle \geq 0$, por ser D semidefinida positiva. Para cada G ,

$$\{\operatorname{sgn}(\langle u_i, G \rangle)\}, \{\operatorname{sgn}(\langle u_j, G \rangle)\} \subseteq \{-1, 1\}$$

son una evaluación posible para el problema $OPT(D)$. Como el promedio ponderado por G de esas evaluaciones es $\frac{2}{\pi} SDP(D)$, entonces al menos una de ellas resulta ser $\geq \frac{2}{\pi} SDP(D)$, y por lo tanto, para el caso D semidefinida positiva tenemos que

$$\frac{\pi}{2} OPT(D) \geq SDP(D).$$

Nos falta ver que efectivamente $\pi/2$ es la constante óptima para este problema. Para ver esto, vamos a construir matrices D semidefinidas positivas para las cuales el cociente $\frac{SDP(D)}{OPT(D)}$ se aproxima a $\frac{\pi}{2}$. Vamos a fijar un $d \in \mathbb{R}_{>0}$, y luego tomemos $n \gg k$. Sean $\{v_i\}_{1 \leq i \leq n}$ variables aleatorias independientes elegidas independientemente sobre \mathbb{S}^{k-1} . Definimos luego la matriz de Gram de los vectores $\frac{v_i}{n}$, que es claramente semidefinida positiva:

$$d_{ij} := \frac{\langle v_i, v_j \rangle}{n^2}.$$

Ahora si evaluamos el problema $SDP(D)$ en los vectores v_i resulta:

$$\sum_{i,j=1}^n d_{ij} \langle v_i, v_j \rangle = \sum_{i,j=1}^n \frac{\langle v_i, v_j \rangle}{n^2} \langle v_i, v_j \rangle = \frac{1}{n^2} \sum_{i,j=1}^n (\langle v_i, v_j \rangle)^2.$$

Si tomamos esperanza sobre las variables v_i , vemos que

$$\mathbb{E} \left[\sum_{i,j=1}^n d_{ij} \langle v_i, v_j \rangle \right] = \frac{1}{n^2} \sum_{i,j=1}^n \mathbb{E} [(\langle v_i, v_j \rangle)^2] = \frac{1}{n^2} \sum_{i,j=1}^n \frac{1}{k} = \frac{1}{k}, \quad (1.9)$$

por el Lema (1.1.8). Esto quiere decir que el valor óptimo del problema $SDP(D)$ es al menos $\frac{1}{k}$. Para ser más precisos, dado que no hay una matriz D definida, sino que la estamos definiendo a partir de ciertos vectores aleatorios $\{v_i\}$, lo que podemos decir es que existe al menos una matriz $D \in \mathbb{R}^{n \times n}$ y al menos una asignación para el problema $SDP(D)$ de esa matriz con valor al menos $\frac{1}{k}$.

Ahora vamos a acotar el valor del problema $OPT(D)$ para este tipo de matrices. Sea $\{\delta_i\}_{1 \leq i \leq n}$, y recordemos que ya sabemos que el mayor valor de $OPT(D)$ tiene que alcanzarse con $\delta_i = \epsilon_i$ para todo $i \in \{1, \dots, n\}$. Luego

$$\sum_{i,j=1}^n d_{ij} \delta_i \delta_j = \sum_{i,j=1}^n \frac{\langle v_i, v_j \rangle}{n^2} \delta_i \delta_j = \left\langle \sum_{i=1}^n \frac{\delta_i v_i}{n}, \sum_{j=1}^n \frac{\delta_j v_j}{n} \right\rangle = \left\| \frac{1}{n} \sum_{i=1}^n \delta_i v_i \right\|^2.$$

Es decir que el mayor valor posible del problema $OPT(D)$ está ligado a la mayor norma posible del vector $\sum_{i=1}^n \delta_i v_i$. Supongamos que $\{\delta_i\}$ es la elección óptima de signos para maximizar la norma de $\sum_{i=1}^n \delta_i v_i$, y sea w la dirección de este vector maximal. Entonces está claro que los signos δ_i de este vector tienen que hacer que los vectores v_i “apunten lo más posible” hacia la dirección de w . Es decir

$$\left\langle \sum_{i=1}^n \delta_i v_i, w \right\rangle = \left\| \sum_{i=1}^n \delta_i v_i \right\| \quad \text{y} \quad \delta_i = \text{sgn}(\langle v_i, w \rangle) \quad \forall i, 1 \leq i \leq n.$$

Entonces resulta que, en este tipo de matrices, para algún $w \in \mathbb{S}^{d-1}$

$$OPT(D) = \left\| \frac{1}{n} \sum_{i=1}^n \text{sgn}(\langle v_i, w \rangle) v_i \right\|^2 = \left(\frac{1}{n} \sum_{i=1}^n \text{sgn}(\langle v_i, w \rangle) \langle v_i, w \rangle \right)^2 = \left(\frac{1}{n} \sum_{i=1}^n |\langle v_i, w \rangle| \right)^2.$$

En particular, esto nos dice que

$$OPT(D) \leq \max_{w \in \mathbb{S}^{k-1}} \left(\frac{1}{n} \sum_{i=1}^n |\langle v_i, w \rangle| \right)^2.$$

Como el argumento de la derecha es claramente uniformemente continuo en w , entonces, dado $\epsilon > 0$ podemos tomar un subconjunto $\mathcal{W}_\epsilon \subset \mathbb{S}^{k-1}$ finito tal que

$$OPT(D) \leq \max_{w \in \mathcal{W}} \left(\frac{1}{n} \sum_{i=1}^n |\langle v_i, w \rangle| \right)^2 + \epsilon.$$

Tomando esperanza sobre los v_i usando el Lema (1.1.8), obtenemos

$$\mathbb{E}[OPT(D)] \leq \max_{w \in \mathcal{W}_\epsilon} \left(\frac{1}{n} \sum_{i=1}^n \mathbb{E}[|\langle v_i, w \rangle|] \right)^2 + \epsilon \leq \max_{w \in \mathcal{W}_\epsilon} \frac{1}{n} \sum_{i=1}^n \left(\left(\sqrt{\frac{2}{\pi}} + o(1) \right) \frac{1}{\sqrt{k}} \right)^2 + \epsilon = \left(\frac{2}{\pi} + o(1) \right) \frac{1}{k} + \epsilon. \quad (1.10)$$

Sólo resta poner todos los argumentos de (1.9) y (1.10). Supongamos que tenemos k, ϵ y el conjunto $\mathcal{W}_\epsilon \subset \mathbb{S}^{k-1}$ fijos. Lo que dicen estas dos ecuaciones es que, con n suficientemente grande, alguna de las matrices de Gram D cumple que

$$SDP(D) \geq \frac{1}{k} - \epsilon \quad \text{y} \quad OPT(D) \leq \left(\frac{2}{\pi} + o(1) \right) \frac{1}{k} + \epsilon.$$

Luego esta matriz D tiene un integrality gap $\frac{SDP(D)}{OPT(D)} \geq \frac{\frac{1}{k} - \epsilon}{\left(\frac{2}{\pi} + o(1) \right) \frac{1}{k} + \epsilon} \approx \frac{\pi}{2} - \epsilon'$. Como podemos hacer esto para cualquier $k, \epsilon > 0$, y $o(1) \xrightarrow{k \rightarrow \infty} 1$, entonces existen matrices D semidefinidas positivas con integrality gap arbitrariamente cerca de $\frac{\pi}{2}$. \square

1.2. Versión con productos tensoriales

El contexto original en el que se desarrolla la desigualdad de Grothendieck es el de normas en el producto tensorial de espacios normados. Grothendieck estudió el problema de, dados dos espacios normados $(E, \|\cdot\|_E), (F, \|\cdot\|_F)$, cómo construir una norma $\|\cdot\|_\otimes$ en el producto tensorial $E \otimes F$ de manera que esta nueva norma tenga alguna relación de compatibilidad con las normas originales. En general son interesantes las siguientes propiedades, que pedimos en la definición de una norma tensorial.

Definición 1.2.1. Dados E, F espacios vectoriales, no necesariamente normados, podemos identificar $x^* \otimes y^* \in E^* \otimes F^*$ con una función en $(E \otimes F)^*$ dual algebraico de la siguiente manera:

$$(x^* \otimes y^*)(x \otimes y) := x^*(x)y^*(y) \quad \forall x \in E, y \in F.$$

Definición 1.2.2. Si $(E, \|\cdot\|_E), (F, \|\cdot\|_F)$ son dos espacios normados, $\|\cdot\|_\otimes$ una norma en el espacio $E \otimes F$, y $((E \otimes F)^*, \|\cdot\|_{\otimes^*})$ su espacio dual estándar. Decimos que la norma $\|\cdot\|_\otimes$ es razonable si:

1. $\|x \otimes y\|_\otimes = \|x\|_E \|y\|_F \quad \forall x \in E, y \in F.$
2. $x^* \otimes y^* \in (E \otimes F, \|\cdot\|_\otimes)^*$, con $\|x^* \otimes y^*\|_{\otimes^*} = \|x^*\|_{E^*} \|y^*\|_{F^*} \quad \forall x^* \in (E, \|\cdot\|_E)^*, y^* \in (F, \|\cdot\|_F)^*.$

Definición 1.2.3. Una norma tensorial α en la categoría de espacios normados es un método para, dado cualquier par de espacios normados $(E, \|\cdot\|_E), (F, \|\cdot\|_F)$, construir una norma $\alpha(-; E, F)$ en el espacio vectorial $E \otimes F$ que cumple lo siguiente:

1. La norma $\alpha(-; E, F)$ es razonable.
2. Se cumple la propiedad de mapeo tensorial:

$$\|T_1 \otimes T_2\|_\alpha := \|T_1 \otimes T_2 : (E_1 \otimes E_2, \alpha(-; E_1, E_2)) \rightarrow (F_1 \otimes F_2, \alpha(-; F_1, F_2))\| \leq \|T_1\| \|T_2\|.$$

A Grothendieck le interesaba estudiar cuáles son las normas tensoriales “naturales” que se pueden construir en productos tensoriales de espacios normados. En su paper [Gro56], que es básicamente el paper fundacional de la teoría de normas tensoriales, en algún momento quiere demostrar la equivalencia de dos normas tensoriales $\|\cdot\|_\otimes, w_2$, lo que lo lleva a plantear por primera vez lo que hoy conocemos como “Desigualdad de Grothendieck”:

Desigualdad de Grothendieck 1.2.4 (Théorème fondamental de la théorie métrique des produits tensoriels). *Existe una constante $K_G \in \mathbb{R}_{>0}$ tal que*

$$\|\cdot\|_\otimes \leq K_G w_2.$$

El trabajo de Grothendieck pasó en cierta medida desapercibido por la comunidad matemática por bastantes años. En [Pis12], Pisier conjetura que algunos motivos para esto fueron que era un trabajo largo, publicado en francés en una revista brasileña de poca circulación, y que lidiaba con un tema que no estaba en la moda de la época. Pero 15 años más tarde de la publicación del paper de Grothendieck, un trabajo de Lindenstrauss-Pełczyński [LP68] retomó

varios de los caminos abiertos por éste, mostrando cómo incluso mucho de su contenido era aplicable para responder preguntas que habían surgido en la comunidad luego de su publicación. Una de las contribuciones más interesantes de este nuevo trabajo fue la de dar una formulación del Teorema Fundamental de Grothendieck mucho más elemental en términos de matrices.

Por mucho tiempo, la mejor cota superior conocida para la constante Grothendieck fue la hallada por Krivine en [Kri78], y cuya demostración veremos en la próxima sección.

Para esto, vamos a introducir las normas tensoriales w_2 y π .

Definición 1.2.5. Dados dos espacios normados $(E, \|\cdot\|_E), (F, \|\cdot\|_F)$, la norma proyectiva π en $E \otimes F$ se define como

$$\pi(z; E, F) = \inf \left\{ \sum_{k=1}^m \|x_k\|_E \|y_k\|_F \mid z = \sum_{k=1}^m x_k \otimes y_k \right\}$$

Definición 1.2.6. Dados dos espacios normados $(E, \|\cdot\|_E), (F, \|\cdot\|_F)$, la norma w_2 en $E \otimes F$ se define como

$$w_2(z; E, F) = \inf \left\{ \sup_{\substack{x^* \in B_{E^*} \\ y^* \in B_{F^*}}} \sqrt{\sum_{k=1}^m |x^*(x_k)|^2} \sqrt{\sum_{k=1}^m |y^*(y_k)|^2} \mid z = \sum_{i=1}^m x_k \otimes y_k \right\}$$

Cuando los espacios E y F son $\mathcal{C}(S)$ y $\mathcal{C}(T)$ para compactos S y T , la definición se puede simplificar un poco.

Lema 1.2.7. Para el espacio $\mathcal{C}(S) \otimes \mathcal{C}(T)$ vale que

$$w_2(z; \mathcal{C}(S), \mathcal{C}(T)) = \inf \left\{ \left\| \sqrt{\sum_{k=1}^m |x_k|^2} \right\|_{\infty} \left\| \sqrt{\sum_{k=1}^m |y_k|^2} \right\|_{\infty} \mid z = \sum_{k=1}^m x_k \otimes y_k \right\}$$

En el caso de w_2 para $\ell_{\infty}^n \otimes \ell_{\infty}^n$, tenemos una interpretación alternativa muy conveniente de la norma w_2 . Si $S := T := \{1, \dots, n\}$ podemos identificar $\ell_{\infty}^n \otimes \ell_{\infty}^n \sim \mathcal{C}(S) \otimes \mathcal{C}(T)$. A su vez, hay una inclusión inmediata de $\mathcal{C}(S) \otimes \mathcal{C}(T) \subseteq \mathcal{C}(S \times T)$.

Ahora supongamos que tenemos un tensor $z \in \ell_{\infty}^n \otimes \ell_{\infty}^n$, y lo interpretamos como una función $\tilde{z} \in \mathcal{C}(\ell_1^n \times \ell_1^n)$. Vamos a decir que \tilde{z} (o z) tiene una representación como producto interno si existe un Hilbert \mathcal{H} y dos asignaciones $s \mapsto u_s, t \mapsto v_t$, con $u_s, v_t \in \mathcal{H}$, de manera que $\tilde{z}(s, t) = \langle u_s, v_t \rangle_{\mathcal{H}} \quad \forall s, t \in \ell_1^n$.

Dada una representación $\{u_s, v_t\}$, podemos calcular su “tamaño” como $\max_{s \in S} \|u_s\|_{\mathcal{H}} \max_{t \in T} \|v_t\|_{\mathcal{H}}$. Luego para cualquier tensor z , podemos preguntarnos cuál es el menor “tamaño” posible que puede tener una representación como producto interno. Esta métrica sobre z resulta coincidir con la norma w_2 .

Lema 1.2.8.

$$w_2(z; \ell_{\infty}^n, \ell_{\infty}^n) = \inf_{\substack{s \mapsto u_s, t \mapsto v_t \\ \text{representaciones de } z}} \left\{ \max_{\substack{s \in \ell_1^n \\ t \in \ell_1^n}} \|u_s\|_{\mathcal{H}} \|v_t\|_{\mathcal{H}} \right\}.$$

Demostración. La demostración es inmediata con la siguiente correspondencia. Dada una representación $z = \sum_{k=1}^m x_k \otimes y_k$ del tensor z en $\ell_\infty^n \otimes \ell_\infty^n$, consideramos la representación:

$$\begin{array}{ccc} \ell_1^n & \rightarrow & \ell_2^m \\ s & \mapsto & (x_k(s))_k \end{array} \quad \begin{array}{ccc} \ell_1^n & \rightarrow & \ell_2^m \\ t & \mapsto & (y_k(t))_k \end{array}.$$

Por otro lado, partiendo de una representación como producto interno de z

$$s \mapsto u_s \in \ell_2^m, t \mapsto v_t \in \ell_2^m,$$

sin pérdida de generalidad, podemos conseguir una representación tensorial de z definiendo:

$$z = \sum_{k=1}^m x_k \otimes y_k, \quad x_k \in \ell_\infty^n, y_k \in \ell_\infty^n$$

$$\begin{array}{ccc} x_k : \ell_1^n & \rightarrow & \mathbb{R} \\ s & \mapsto & (u_s)_k \end{array} \quad \begin{array}{ccc} y_k : \ell_1^n & \rightarrow & \mathbb{R} \\ t & \mapsto & (v_t)_k \end{array}.$$

Esta claro que estas dos operaciones se corresponden, y que sus “tamaños” coinciden:

$$\left\| \sqrt{\sum_{k=1}^m |x_k|^2} \right\|_\infty \left\| \sqrt{\sum_{k=1}^n |y_k|^2} \right\|_\infty = \max_{\substack{s \in \ell_1^n \\ t \in \ell_1^n}} \|u_s\|_{\mathcal{H}} \|v_t\|_{\mathcal{H}}.$$

Tomando ínfimo a ambos lados, obtenemos el enunciado del lema. □

Si tenemos un tensor z que tiene norma w_2 menor a 1, entonces podemos hallar una representación como producto interno de z donde $\|u_s\| = \|v_t\| = 1 \forall s \in \ell_1^n, t \in \ell_1^n$. Veamos esto. Si la norma w_2 de z es menor a 1, entonces por el lema anterior (1.2.8) existe una representación $u_s, v_t \in \ell_2^m$ con $\|u_s\|, \|v_t\| < 1$. Definamos entonces una nueva representación $\tilde{u}_s, \tilde{v}_t \in \ell_2^{m+2}$:

$$(\tilde{u}_s)_k := \begin{cases} (u_s)_k & k \leq m \\ \sqrt{1 - \|u_s\|_2^2} & k = m + 1 \\ 0 & k = m + 2 \end{cases} \quad (\tilde{v}_t)_k := \begin{cases} (v_t)_k & k \leq m \\ 0 & k = m + 1 \\ \sqrt{1 - \|v_t\|_2^2} & k = m + 2 \end{cases}.$$

Claramente $s \mapsto \tilde{u}_s, t \mapsto \tilde{v}_t$ es una representación de z , y $\|\tilde{u}_s\| = \|\tilde{v}_t\| = 1 \forall s \in \ell_1^n, t \in \ell_1^n$.

Usando Cauchy-Schwarz, esta interpretación permite además ver inmediatamente que

$$\|\tilde{z}\|_\infty^{\mathcal{C}(\ell_1^n \times \ell_1^n)} \leq w_2(z; \ell_\infty^m, \ell_\infty^n).$$

Ahora que ya tenemos alguna familiaridad muy elemental con un par de normas tensoriales, vamos a volver a la formulación original de Grothendieck (1.2.4). Si bien no pretendemos una cobertura profunda de este tema, nos gustaría de todos modos contar un poco el contexto original en que surge este problema. La norma γ_2 a la que se hace referencia en (1.2.4) es el resultado de aplicarle el operador γ_2 , conocido como capsula inyectiva a la norma π . Vamos a dar las definiciones mínimas para seguir.

Definición 1.2.9. Decimos que una norma tensorial α es inyectiva si para todos los espacios normados $A \subset E, B \subset F$, y para todo tensor $z \in A \otimes B \subset E \otimes F$ se cumple que

$$\alpha(z; A, B) = \alpha(z; E, F).$$

Definición 1.2.10. Dada una norma tensorial α , definimos a la **capsula inyectiva** $/\alpha\backslash$ como la única norma tensorial inyectiva con la siguiente propiedad: Si $\beta \leq \alpha$ es una norma tensorial inyectiva, entonces $\beta \leq / \alpha \backslash$.

Notemos que $/\alpha\backslash$ representa entonces un supremo entre las normas inyectivas menores a α . Se puede demostrar que efectivamente existe tal norma tensorial $/\alpha\backslash$, y que, dado un tensor z y un par de espacios normados E, F se puede calcular

$$/\alpha\backslash(z; E, F) = \sup_{\substack{\beta \text{ inyectiva} \\ \beta \leq \alpha}} \beta(z; E, F).$$

Con estas definiciones ya podemos, al menos, darle sentido al enunciado original de Grothendieck. Vamos a ver, a grandes rasgos, cómo este enunciado resulta equivalente a la forma tensorial (1.3.1) que conocemos.

Idea de la demostración de (1.2.4). Para empezar, notemos que en el espacio $\ell_\infty^n \otimes \ell_\infty^n$ se cumple que

$$/\pi\backslash(-; \ell_\infty^n, \ell_\infty^n) = \pi(-; \ell_\infty^n, \ell_\infty^n).$$

Si bien esto no es evidente a priori, se tiene la igualdad entre estas normas, ya que ℓ_∞^n es un espacio inyectivo (consecuencia directa de Hahn-Banach).

Esto ya nos dice que la formulación original de Grothendieck implica nuestra formulación tensorial "finita". Para ir en la otra dirección, vamos a necesitar un poco más de esfuerzo. Comencemos con dos espacios normados E, F , y un tensor $z \in E \otimes F$, para el que queremos probar que $/\pi\backslash(z; E, F) \leq K_G w_2(z; E, F)$. Resulta que las normas $/\pi\backslash$ y w_2 son finitamente generadas, es decir, se pueden calcular de la siguiente manera:

$$/\pi\backslash(z; E, F) = \inf_{\substack{A \subset E, B \subset F \\ \text{dimensión finita}}} / \pi \backslash(z; A, B) \quad w_2(z; E, F) = \inf_{\substack{A \subset E, B \subset F \\ \text{dimensión finita}}} w_2(z; A, B).$$

Por otro lado, cualquier espacio normado E se puede ver isométricamente como un subespacio de $L_\infty(\mathcal{B}_{E^*})$, con la identificación de $x \in E$ con la función $\hat{x} \in X^{**} \subset L_\infty(\mathcal{B}_{E^*})$.

Cualquier espacio $L_\infty(\Omega)$ tiene la propiedad de ser un $\mathcal{L}_\infty^\lambda$ para cualquier $\lambda > 1$. Esto significa que, dado $\lambda > 1$ y un subespacio de dimensión finita $A \subset L_\infty(\Omega)$, existe un subespacio L de dimensión finita, $A \subset L \subset L_\infty(\Omega)$ y un isomorfismo $T : L \rightarrow \ell_\infty^m$ para algún $m \in \mathbb{N}$ tal que $\|T\| \|T^{-1}\| \leq \lambda$. Es decir, cualquier subespacio de dimensión finita A está contenido en otro subespacio de dimensión finita arbitrariamente "parecido" a algún ℓ_∞^m .

Ahora es sólo cuestión de ordenar un poco lo que sabemos. Recordemos que tenemos $z \in E \otimes F$ y queremos probar $/\pi\backslash(z; E, F) \leq K_G w_2(z; E, F)$. Sean $A \subset E, B \subset F$ de dimensión finita. Luego para cada $\lambda > 1$ existen L_A^λ, L_B^λ de dimensión finita tales que $A \subset L_A^\lambda \subset E, B \subset L_B^\lambda \subset F$, y existen isomorfismos $T_A : L_A \rightarrow \ell_\infty^m, T_B : L_B \rightarrow \ell_\infty^m$ con $\|T_A\| \|T_A^{-1}\|, \|T_B\| \|T_B^{-1}\| \leq \lambda$, para algún $m \in \mathbb{N}$.

Luego podemos definir $T := T_A \otimes T_B : L_A \otimes L_B \rightarrow \ell_\infty^m \otimes \ell_\infty^m$, con inversa $T^{-1} := T_A^{-1} \otimes T_B^{-1}$. Por la propiedad de mapeo métrico (1.2.3), se cumple que

$$\|T\|_\alpha \leq \|T_A\| \|T_B\|, \|T^{-1}\|_\alpha \leq \|T_A^{-1}\| \|T_B^{-1}\| \quad (1.11)$$

para cualquier norma tensorial α . Entonces

$$\begin{aligned} / \pi \backslash (z; L_A, L_B) &\leq \|T^{-1}\|_{/\pi \backslash} / \pi \backslash (T(z); \ell_\infty^m, \ell_\infty^m) \stackrel{(1.2)}{=} \|T^{-1}\|_{/\pi \backslash} \pi(T(z); \ell_\infty^m, \ell_\infty^m) \leq \\ &\leq \|T^{-1}\|_{/\pi \backslash} K_G w_2(T(z); \ell_\infty^m, \ell_\infty^m) \leq \|T^{-1}\|_{/\pi \backslash} K_G \|T\|_{w_2} w_2(z; L_A, L_B) \leq \\ &\leq \|T^{-1}\|_{/\pi \backslash} \|T\|_{w_2} \cdot K_G w_2(z; L_A, L_B) \stackrel{(1.11)}{\leq} \|T_A\| \|T_B\| \|T_A^{-1}\| \|T_B^{-1}\| \\ &\leq \lambda^2 \cdot K_G w_2(z; L_A, L_B). \end{aligned}$$

Pero esto vale para cualquier $\lambda > 1$, luego $/ \pi \backslash (z; L_A, L_B) \leq K_G w_2(z; L_A, L_B)$.

Por último, como para cualquier norma tensorial $A \subset L_A, B \subset L_B \Rightarrow \alpha(z; L_A, L_B) \leq \alpha(z; A, B)$, tenemos que

$$\begin{aligned} / \pi \backslash (z; E, F) &= \inf_{\substack{A \subset E, B \subset F \\ \text{dimensión finita}}} / \pi \backslash (z; A, B) = \inf_{L_A, L_B} / \pi \backslash (z; L_A, L_B) \leq \\ &\leq K_G \inf_{L_A, L_B} w_2(z; L_A, L_B) = K_G \inf_{\substack{A \subset E, B \subset F \\ \text{dimensión finita}}} w_2(z; A, B) = K_G w_2(z; E, F). \end{aligned}$$

□

1.3. Cota superior de Krivine

En esta sección vamos a demostrar que la desigualdad de Grothendieck vale usando la constante $\frac{\pi}{\ln(1+\sqrt{2})}$ o, dicho de otra manera, que $K_G \leq \frac{\pi}{\ln(1+\sqrt{2})}$. Esta demostración fue realizada por Krivine alrededor de 1977 [Kri78], y usa algunas herramientas de productos tensoriales. Si bien en el próximo capítulo reduciremos esta demostración a técnicas elementales, nos parece interesante contar el contexto original en que surgen estas ideas.

La demostración original de Krivine tiene que ver con una formulación de la desigualdad de Grothendieck más cercana a la formulación original

Desigualdad de Grothendieck 1.3.1 (Desigualdad entre normas π y w_2). *Para todo $n \in \mathbb{N}$ se tiene que*

$$\pi(z; \ell_\infty^n, \ell_\infty^n) \leq K_G w_2(z; \ell_\infty^n, \ell_\infty^n) \quad \forall z \in \ell_\infty^n \otimes \ell_\infty^n.$$

Antes de dar la demostración, vamos a probar el siguiente lema que resultará fundamental.

Lema 1.3.2. *El espacio $\ell_\infty^n \otimes \ell_\infty^n$ es un álgebra normada con la norma dada por $w_2(-; \ell_\infty^n, \ell_\infty^n)$.*

Demostración. Para empezar vamos a definir el producto que queremos darle a $\ell_\infty^n \otimes \ell_\infty^n$ y bajo el cual será un álgebra. Como podemos interpretar el espacio $\ell_\infty^n \otimes \ell_\infty^n$ como el espacio de formas bilineales de $\ell_1^n \times \ell_1^n$ a escalares, tenemos un producto muy natural simplemente

tomando el producto de funciones punto a punto. Es decir, si tenemos $z, w \in \ell_\infty^n \otimes \ell_\infty^n$, definimos $\widetilde{z \cdot w} \in \text{Bil}(\ell_1^n \times \ell_1^n, \mathbb{R})$ de la siguiente manera:

$$\widetilde{z \cdot w}(s, t) := \widetilde{z}(s, t) \cdot \widetilde{w}(s, t) \quad s, t \in \ell_1^n.$$

Luego $z \cdot w$ se define como el único tensor en $\ell_\infty^n \otimes \ell_\infty^n$ al que le corresponde esa forma bilineal.

Nos faltaría ver, entonces que $w_2(z \cdot w; \ell_\infty^n, \ell_\infty^n) \leq w_2(z; \ell_\infty^n, \ell_\infty^n) \cdot w_2(w; \ell_\infty^n, \ell_\infty^n)$.

Por el Lema (1.2.8), dado $\varepsilon > 0$, podemos tomar representaciones

$$\{u_s^z\}_s, \{v_t^z\}_t \subset \mathcal{H}_z, \quad \{u_s^w\}_s, \{v_t^w\}_t \subset \mathcal{H}_w,$$

con $\mathcal{H}_z, \mathcal{H}_w$ espacios de Hilbert de forma que

$$\begin{aligned} \widetilde{z}(s, t) &= \langle u_s^z, v_t^z \rangle & \|u_s^z\| &= \|v_t^z\| = w_2(z; \ell_\infty^n, \ell_\infty^n) + \varepsilon, \\ \widetilde{w}(s, t) &= \langle u_s^w, v_t^w \rangle & \|u_s^w\| &= \|v_t^w\| = w_2(w; \ell_\infty^n, \ell_\infty^n) + \varepsilon. \end{aligned}$$

Con esto podemos construir una representación como producto interno de $z \cdot w$ en el espacio de Hilbert $\mathcal{H}_z \otimes \mathcal{H}_w$. Esta representación está conformada por los vectores $\{u_s^z \otimes u_s^w\}_s, \{v_t^z \otimes v_t^w\}_t$. Verifiquemos que efectivamente

$$\begin{aligned} \widetilde{z \cdot w}(s, t) &= \widetilde{z}(s, t) \cdot \widetilde{w}(s, t) = \langle u_s^z, v_t^z \rangle \langle u_s^w, v_t^w \rangle = \\ &= \langle u_s^z \otimes u_s^w, v_t^z \otimes v_t^w \rangle. \end{aligned}$$

En cuanto al “tamaño” de la representación, tenemos que

$$\begin{aligned} \|u_s^z \otimes u_s^w\|_{\mathcal{H}_z \otimes \mathcal{H}_w} &= \|u_s^z\|_{\mathcal{H}_z} \|u_s^w\|_{\mathcal{H}_w} = (w_2(z; \ell_\infty^n, \ell_\infty^n) + \varepsilon) (w_2(w; \ell_\infty^n, \ell_\infty^n) + \varepsilon) \\ &= w_2(z; \ell_\infty^n, \ell_\infty^n) \cdot w_2(w; \ell_\infty^n, \ell_\infty^n) + O(\varepsilon). \end{aligned}$$

Y análogamente con los vectores $v_t^z \otimes v_t^w$. Luego, con $\varepsilon \rightarrow 0$, tendremos representaciones de “tamaño” tan cercano a $w_2(z; \ell_\infty^n, \ell_\infty^n) \cdot w_2(w; \ell_\infty^n, \ell_\infty^n)$ como queramos. Nuevamente por el Lema (1.2.8), esto implica que

$$w_2(z \cdot w; \ell_\infty^n, \ell_\infty^n) \leq w_2(z; \ell_\infty^n, \ell_\infty^n) \cdot w_2(w; \ell_\infty^n, \ell_\infty^n).$$

□

Usando este lema, vamos a ver la demostración de Krivine que halla una cota $K_G \leq \frac{\pi}{2 \ln(1+\sqrt{2})} < \sinh\left(\frac{\pi}{2}\right)$ menor a la original de Grothendieck. Luego probaremos la equivalencia de las distintas formulaciones de la desigualdad.

Demostración de la Desigualdad de Grothendieck (1.3.1). Sea $n \in \mathbb{N}$ fijo. Para abreviar, en este contexto y sin ambigüedad, denominaremos $\|z\|_{w_2} := w_2(z; \ell_\infty^n, \ell_\infty^n)$, $\|z\|_\pi := \pi(z; \ell_\infty^n, \ell_\infty^n)$.

Sea $z \in \ell_\infty^n \otimes \ell_\infty^n$, con $\|z\|_{w_2} < 1$, y sea $a := \ln(1 + \sqrt{2})$. Queremos ver que $\|z\|_\pi < \frac{\pi}{2a}$.

Sabemos por (1.3.2) que $(\ell_\infty^n \otimes \ell_\infty^n, \|\cdot\|_{w_2})$ es un álgebra normada. Luego, como $\sin(x) = \sum_{j=0}^{\infty} \frac{(-1)^j}{(2j+1)!} x^{2j+1}$ es una serie con radio de convergencia infinito, entonces podemos definir $\text{sen}(z) := \sum_{j=0}^{\infty} \frac{(-1)^j}{(2j+1)!} z^{2j+1} \quad \forall z \in \ell_\infty^n \otimes \ell_\infty^n$.

Se tiene entonces:

$$\|\operatorname{sen}(az)\|_{w_2} \leq \operatorname{senh}(a\|z\|_{w_2}) < \operatorname{senh}(a) = 1$$

Entonces, como vimos anteriormente, existe una representación $u_s, v_t \in \ell_2^k$ como producto interno de $\operatorname{sen}(az)$, con $\|u_s\| = \|v_t\| = 1 \forall s, t \in \ell_1^n$. Entonces, por el Lema (1.1.5) tenemos que:

$$\operatorname{sen}(a\tilde{z}(s, t)) = \widetilde{\operatorname{sen}(az)}(s, t) = \langle u_s, v_t \rangle = \operatorname{sen} \left(\frac{\pi}{2} \mathbb{E}_G [\operatorname{sgn}(\langle u_s, G \rangle) \operatorname{sgn}(\langle v_t, G \rangle)] \right) \quad \forall s, t \in \ell_1^n. \quad (1.12)$$

Por un lado $a\tilde{z}(s, t) \leq a\|\tilde{z}\|_{\infty}^{C(\ell_1^n \times \ell_1^n)} \leq a\|z\|_{w_2} < a < \frac{\pi}{2}$. Por el otro, $\frac{\pi}{2} \mathbb{E}_G [\operatorname{sgn}(\langle u_s, G \rangle) \operatorname{sgn}(\langle v_t, G \rangle)] \leq \frac{\pi}{2}$. Luego, por inyectividad del seno en $[-\frac{\pi}{2}, \frac{\pi}{2}]$, resulta que en (1.12) los argumentos son iguales:

$$a\tilde{z}(s, t) = \frac{\pi}{2} \mathbb{E}_G [\operatorname{sgn}(\langle u_s, G \rangle) \operatorname{sgn}(\langle v_t, G \rangle)] \quad \forall s, t \in \ell_1^n.$$

Entonces, tenemos que

$$\begin{aligned} az &= \sum_{i,j=1}^n a\tilde{z}(e_i, e_j) e_i^* \otimes e_j^* \\ az &= \sum_{i,j=1}^n \frac{\pi}{2} \mathbb{E}_G [\operatorname{sgn}(\langle u_{e_i}, G \rangle) \operatorname{sgn}(\langle v_{e_j}, G \rangle)] e_i^* \otimes e_j^* \\ az &= \frac{\pi}{2} \mathbb{E}_G \left[\sum_{i,j=1}^n \operatorname{sgn}(\langle u_{e_i}, G \rangle) \operatorname{sgn}(\langle v_{e_j}, G \rangle) e_i^* \otimes e_j^* \right] \\ az &= \frac{\pi}{2} \mathbb{E}_G \left[\left(\sum_{i=1}^n \operatorname{sgn}(\langle u_{e_i}, G \rangle) e_i^* \right) \otimes \left(\sum_{j=1}^n \operatorname{sgn}(\langle v_{e_j}, G \rangle) e_j^* \right) \right] \\ \|az\|_{\pi} &\leq \frac{\pi}{2} \mathbb{E}_G \left[\left\| \left(\sum_{i=1}^n \operatorname{sgn}(\langle u_{e_i}, G \rangle) e_i^* \right) \otimes \left(\sum_{j=1}^n \operatorname{sgn}(\langle v_{e_j}, G \rangle) e_j^* \right) \right\|_{\pi} \right] \\ \|az\|_{\pi} &\leq \frac{\pi}{2} \mathbb{E}_G \left[\underbrace{\left\| \sum_{i=1}^n \operatorname{sgn}(\langle u_{e_i}, G \rangle) e_i^* \right\|_{\infty}}_{=1} \underbrace{\left\| \sum_{j=1}^n \operatorname{sgn}(\langle v_{e_j}, G \rangle) e_j^* \right\|_{\infty}}_{=1} \right] \\ \|z\|_{\pi} &\leq \frac{\pi}{2a} = \frac{\pi}{2 \ln(1 + \sqrt{2})}. \end{aligned}$$

□

Vamos a ver que las formulaciones (1.1.1) y (1.3.1) son equivalentes. Dada una matriz $A \in \mathbb{R}^{n \times n}$, la podemos identificar con una forma bilineal en $\ell_{\infty}^n \times \ell_{\infty}^n$, y por lo tanto con una función del dual $A \in (\ell_{\infty}^n \otimes \ell_{\infty}^n, \pi)^*$. De hecho, es un hecho conocido que para todos E, F normados hay una isometría entre $\mathcal{B}(E, F)$, el espacio de formas bilineales en $E \times F$ acotadas y $(E \otimes_{\pi} F)^* := (E \otimes F, \pi(-; E, F))^*$. Vamos a demostrar esto brevemente.

Lema 1.3.3. *Para cualesquiera espacios normados E y F , $\mathcal{B}(E, F)$ y $(E \otimes_{\pi} F)^*$ son isométricos.*

Demostración. El isomorfismo lineal es obvio. Veamos la igualdad de normas. Sea $B \in (E \otimes_{\pi} F)^*$ y sea $\tilde{B} : E \times F \rightarrow \mathbb{R}$ su forma bilineal asociada. Si z es un tensor en $E \otimes_{\pi} F$ con $\|z\| = 1$ y con una representación $z = \sum_{i=1}^k x_i \otimes y_i$, entonces:

$$|B(z)| = \left| B \left(\sum_{i=1}^k x_i \otimes y_i \right) \right| = \left| \sum_{i=1}^k \tilde{B}(x_i, y_i) \right| \leq \sum_{i=1}^k \|\tilde{B}\| \|x_i\| \|y_i\| \leq \|\tilde{B}\| \left(\sum_{i=1}^k \|x_i\| \|y_i\| \right).$$

Luego tomando ínfimo a la derecha sobre todas las representaciones de z , obtenemos:

$$|B(z)| \leq \|\tilde{B}\| \|z\|_{\pi}.$$

Y ahora tomando supremo sobre todos los $\|z\|_{\pi} = 1$:

$$\|B\|_{\pi} \leq \|\tilde{B}\|.$$

Por otro lado, por definición de $\|\tilde{B}\|$, $\forall \varepsilon > 0$ existen $x_0 \in E, y_0 \in F, \|x_0\| = \|y_0\| = 1$, tales que $\tilde{B}(x_0, y_0) \geq \|\tilde{B}\| - \varepsilon$. Entonces si tomamos $z_0 := x_0 \otimes y_0$, resulta que $\|z_0\|_{\pi} = \|x_0\| \|y_0\| = 1$ y luego

$$\|B\|_{\pi} \geq \sup_{\|z\|=1} |B(z)| \geq |B(z_0)| = |\tilde{B}(x_0, y_0)| \geq \|\tilde{B}\| - \varepsilon.$$

Como esto vale $\forall \varepsilon > 0$, entonces $\|B\|_{\pi} = \|\tilde{B}\|$. \square

Ahora sí, con este lema a mano, vamos a demostrar la equivalencia entre las formulaciones matricial (1.1.1) y tensorial (1.3.1).

Partamos primero de la versión tensorial (1.3.1).

Dada una matriz $A \in \mathbb{R}^{n \times n}$, vamos a escribir el tensor

$$\hat{A} := \sum_{i,j=1}^n a_{ij} e_i \otimes e_j \quad \in (\ell_1^n \otimes \ell_1^n) = (\ell_1^n \otimes \ell_1^n)^*.$$

Sean $\{u_i\}_{i=1}^n, \{v_j\}_{j=1}^n \subset B_{\mathcal{H}}$, con \mathcal{H} Hilbert. Luego el tensor

$$z := \sum_{i,j=1}^n \langle u_i, v_j \rangle e_i^* \otimes e_j^* \quad \in (\ell_{\infty}^n \otimes \ell_{\infty}^n) \quad (1.13)$$

cumple que $\|z\|_{w_2} \leq 1$, ya que los u_i, v_j inducen una representación donde $\|u_s\|, \|v_t\| \leq 1 \forall s, t \in \ell_{\infty}^n$. Análogamente, se puede ver que si $\|z\|_{w_2} \leq 1$, entonces existe una representación de z de tamaño 1.

Si z es un tensor en $\ell_{\infty}^n \otimes \ell_{\infty}^n$, resulta que

$$\hat{A}z = \sum_{i,j=1}^n a_{ij} (e_i \otimes e_j)(f) = \sum_{i,j=1}^n a_{ij} z(e_i, e_j) = \sum_{i,j=1}^n a_{ij} \langle u_{e_i}, v_{e_j} \rangle, \quad (1.14)$$

para cualquier representación $s \mapsto u_s, t \mapsto v_t$ de z como producto interno.

Usando esta expresión, podemos ver que

$$SDP(A) = \max_{\|u_i\|=\|v_j\|=1} \left| \sum_{i,j=1}^n a_{ij} \langle u_i, v_j \rangle \right| = \max_{\|f\|_{w_2} \leq 1} |\hat{A}f| = \|\hat{A}\|_{w_2^*}.$$

Por otro lado, por el Lema (1.3.3), resulta que $\|\hat{A}\|_{\pi^*} = \|A\|_{\mathcal{B}(\ell_\infty^n \otimes \ell_\infty^n)} = \|A\|_{\infty \rightarrow 1} = OPT(A)$.

Con esto, la equivalencia entre las dos formulaciones de Grothendieck queda perfectamente clara. La formulación (1.1.1) asevera que existe una constante K_G tal que, para toda matriz A ,

$$SDP(A) \leq K_G OPT(A).$$

Y acabamos de ver que esto es equivalente a que

$$\|\hat{A}\|_{w_2^*} \leq K_G \|\hat{A}\|_{\pi^*} \quad \forall \hat{A} \in \ell_1^n \otimes \ell_1^n.$$

¡Pero esto es exactamente la desigualdad dual de (1.3.1)!

$$\|z\|_\pi \leq K_G \|z\|_{w_2} \quad \forall z \in \ell_\infty^n \otimes \ell_\infty^n.$$

Como estamos trabajando en espacios de dimensión finita, cada ecuación implica a la otra. Es más, cualquier constante que encontremos que funciona para una de las ecuaciones, inmediatamente funciona para la otra. Luego las constantes óptimas de cada caso son equivalentes, y por lo tanto las dos formulaciones son exactamente equivalentes, como queríamos probar.

1.4. Cota inferior de Reeds

En esta sección vamos a ver la demostración de la mejor cota inferior conocida para el problema de Grothendieck. La idea viene de intentar perturbar un poco las matrices con las que trabajamos en el caso de matrices semidefinidas positivas, para intentar agrandar el gap entre $SDP(A)$ y $OPT(A)$.

Teorema 1.4.1 (Reeds).

$$1,67 < K_G.$$

Vamos a usar la siguiente formulación de la desigualdad de Grothendieck:

Desigualdad de Grothendieck 1.4.2. Para todo $T : L_\infty(X, \mu) \rightarrow L_1(X, \mu)$ operador lineal acotado y para todas funciones $f_1, \dots, f_n \in L_\infty(X, \mu)$ se cumple que

$$\left\| \sqrt{\sum_{i=1}^{\infty} (Tf_i)^2} \right\|_1 \leq K_G \|T\| \left\| \sqrt{\sum_{i=1}^{\infty} f_i^2} \right\|_\infty.$$

La idea va a ser construir operadores T_λ tales que $K_G \geq \frac{1-\lambda}{\|T_\lambda\|}$ y luego encontrar el $\lambda \in \mathbb{R}$ tal que maximiza esta expresión. Definimos el “radio de Grothendieck” para un operador $T \in \mathcal{L}(L_\infty, L_1)$ como

$$r(T) := \sup \left\{ \frac{\left\| \sqrt{\sum_{i=1}^{\infty} (Tf_i)^2} \right\|_1}{\|T\| \left\| \sqrt{\sum_{i=1}^{\infty} f_i^2} \right\|_\infty} \mid f_i \in L_\infty \right\}. \quad (1.15)$$

La idea sobre la que se basa la demostración es que perturbar un poco a un operador con un múltiplo pequeño de la identidad parece hacer crecer su radio de Grothendieck:

$$r(T_\lambda) = r(T - \lambda Id) > r(T). \quad (1.16)$$

El operador inicial T que vamos a usar es uno para el cual $r(T) = \frac{\pi}{2}$, y resulta ser el mismo con el que Grothendieck probó inicialmente que $\frac{\pi}{2} \leq K_G$. Sea (Ω, \mathbb{P}) un espacio de probabilidad, y sea $\{G_n\}$ una secuencia de variables aleatorias independientes Gaussianas estándar. Definimos T como la restricción a $L_\infty(\Omega, \mathbb{P}) \subseteq L_2(\Omega, \mathbb{P})$ de la proyección ortogonal al subespacio

$$G := \overline{\langle G_1, G_2, \dots \rangle}. \quad (1.17)$$

Como $\{G_n\}$ es base ortonormal de ese subespacio, esto quiere decir:

$$Tf := \sum_{n \geq 1}^{\infty} \mathbb{E}[fG_n]G_n. \quad (1.18)$$

Y como habíamos adelantado, definimos $T_\lambda := T - \lambda$, con $\lambda \geq 0$.

Lema 1.4.3.

$$r(T_\lambda) \geq \frac{|1 - \lambda|}{\|T_\lambda\|}.$$

Demostración. Fijamos $n \in \mathbb{N}$, y definimos las variable aleatorias

$$G := (G_1, G_2, \dots, G_n) \quad R := \|G\|_2 \quad X = (X_1, X_2, \dots, X_n) := \frac{G}{R}. \quad (1.19)$$

Luego R^2 tiene distribución χ_n^2 , X se distribuye uniformemente sobre \mathbb{S}^{n-1} . Además cada X_i tiene esperanza $\frac{1}{n}$ y es independiente de R . Calculamos

$$TX_i = \sum_{j=0}^n \underbrace{\mathbb{E}[X_i G_j]}_{0, \text{ si } i \neq j} G_j = \mathbb{E}[X_i G_i] G_i = \mathbb{E}[X_i^2 R] G_i = \mathbb{E}[X_i^2] \mathbb{E}[R] G_i = \frac{\mathbb{E}[R] G_i}{n}.$$

Entonces

$$T_\lambda X_i = \frac{\mathbb{E}[R] G_i}{n} - \lambda X_i = X_i \left(\frac{R \mathbb{E}[R]}{n} - \lambda \right).$$

y por lo tanto

$$\begin{aligned} \sqrt{\sum_{j=0}^n (T_\lambda X_i)^2} &= \sqrt{\sum_{j=0}^n \left(X_i^2 \left(\frac{R \mathbb{E}[R]}{n} - \lambda \right)^2 \right)} \\ &= \sqrt{\left(\sum_{j=0}^n X_i^2 \right) \left(\frac{R \mathbb{E}[R]}{n} - \lambda \right)^2} = \left| \frac{R \mathbb{E}[R]}{n} - \lambda \right|. \end{aligned}$$

Tomando entonces $f_i := X_i$ para $1 \leq i \leq n$, por (1.15) tenemos que:

$$\begin{aligned} r(T_\lambda) &\geq \frac{\left\| \sqrt{\sum_{i=1}^{\infty} (T_\lambda X_i)^2} \right\|_1}{\|T_\lambda\|} \left(\left\| \sqrt{\sum_{i=1}^{\infty} X_i^2} \right\|_\infty \right)^{-1} = \frac{\left\| \frac{R \mathbb{E}[R]}{n} - \lambda \right\|_1}{\|T_\lambda\|} \\ &= \frac{\mathbb{E} \left[\left| \frac{R \mathbb{E}[R]}{n} - \lambda \right| \right]}{\|T_\lambda\|} \geq \frac{\left| \mathbb{E} \left[\frac{R \mathbb{E}[R]}{n} \right] - \mathbb{E}[\lambda] \right|}{\|T_\lambda\|} = \frac{\left| \frac{\mathbb{E}[R]^2}{n} - \lambda \right|}{\|T_\lambda\|} \end{aligned}$$

y como $\frac{\mathbb{E}[R]}{\sqrt{n}} \xrightarrow{n \rightarrow \infty} 1$:

$$r(T_\lambda) \geq \frac{|1 - \lambda|}{\|T_\lambda\|}.$$

□

Vamos a calcular la norma de este T_λ .

Sea K es conjunto de los puntos extremos de la bola unidad de $L_\infty(\Omega, \mathbb{P})$, es decir, $K := \{g \in L_\infty(\Omega, \mathbb{P}) \mid |g| = 1 \text{ c.t.p.}\}$. Como $\|T_\lambda(\cdot)\|_1$ es una función convexa, resulta que $\|T_\lambda\| = \sup\{\|T_\lambda g\|_1 \mid g \in K\}$

Supongamos que g está fijo. Podemos suponer, sin pérdida de generalidad, que la proyección de g sobre el subespacio G definido en (1.17) es exactamente su proyección sobre el primer generador en dirección no negativa. Es decir, que $Tg = \mathbb{E}[gG_1] \geq 0$. Como esto vale para todo $g \in K$:

$$\begin{aligned} \|T_\lambda\| &= \sup\{\|T_\lambda g\|_1 \mid g \in K\} = \sup\{\|T_\lambda g\|_1 \mid g \in K \text{ y } Tg = \mathbb{E}[gG_1]\} \\ &= \sup \bigcup_{\alpha \geq 0} \{\|T_\lambda g\|_1 \mid g \in K \text{ y } Tg = \mathbb{E}[gG_1] = \alpha\}. \end{aligned}$$

Definiendo $K_\alpha := \{g \in K \mid Tg = \mathbb{E}[gG_1] = \alpha\}$:

$$\begin{aligned} \|T_\lambda\| &= \sup \bigcup_{\alpha \geq 0} \{\|T_\lambda g\|_1 \mid g \in K_\alpha\} \\ &= \sup_{\alpha \geq 0} \sup_{g \in K_\alpha} \|T_\lambda g\|_1. \end{aligned} \tag{1.20}$$

Sea $\alpha \in \mathbb{R}_{\geq 0}$ fijo. Dado $g \in K_\alpha$, definimos como $\theta_g(x)$ a la esperanza condicional de g respecto de G_1 : $\theta_g(x) := \mathbb{E}[g \mid G_1 = x]$. Entonces tenemos que $|\theta_g(x)| \leq 1$ y, si $\phi(x) := \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ es la función de densidad de G_1 :

$$\begin{aligned} \alpha &= \mathbb{E}[gG_1] = \mathbb{E}_x [\mathbb{E}[gx \mid G_1 = x]] = \mathbb{E}_x [x\mathbb{E}[g \mid G_1 = x]] \\ &= \mathbb{E}_x [x\theta_g(x)] = \int_{\mathbb{R}} \theta_g(x)x\phi(x)dx. \end{aligned}$$

Por otro lado, dada una función $\theta(x)$, con $|\theta(x)| \leq 1$ c.t.p. y $\alpha = \int_{\mathbb{R}} \theta(x)x\phi(x)dx$, podemos construir $g \in K_\alpha$ tal que $\theta = \theta_g$ de la siguiente manera. Sea U una variable uniforme sobre $[0, 1]$ independiente de todos los G_i con $i \in \mathbb{N}$. Definimos:

$$g(U, Z_1) = \begin{cases} 1 & U \leq \frac{1+\theta(G_1)}{2} \\ -1 & U > \frac{1+\theta(G_1)}{2} \end{cases}.$$

Esto nos dice que tenemos una correspondencia:

$$g \in K_\alpha \quad \longleftrightarrow \quad \theta, \|\theta\|_\infty \leq 1, \int_{\mathbb{R}} \theta_g(x)x\phi(x)dx = \alpha. \tag{1.21}$$

Siguiendo con α fijo, vamos a hallar la forma que toma $\|T_\lambda g\|_1$ con la segunda expresión. Vamos a usar que:

$$\left. \begin{aligned} 1 &= \mathbb{P}(g = 1 \mid G_1 = x) + \mathbb{P}(g = -1 \mid G_1 = x) \\ \theta_g(x) &= \mathbb{P}(g = 1 \mid G_1 = x) - \mathbb{P}(g = -1 \mid G_1 = x) \end{aligned} \right\} \Rightarrow \begin{aligned} \mathbb{P}(g = 1 \mid G_1 = x) &= \frac{1 + \theta_g(x)}{2} \\ \mathbb{P}(g = -1 \mid G_1 = x) &= \frac{1 - \theta_g(x)}{2} \end{aligned} \quad (1.22)$$

$$\begin{aligned} \|T_\lambda g\|_1 &= \mathbb{E}[|Tg - \lambda g|] = \mathbb{E}[|\alpha G_1 - \lambda g|] = \mathbb{E}_x[\mathbb{E}[|\alpha x - \lambda g| \mid G_1 = x]] \\ &= \mathbb{E}_x \left[|\alpha x - \lambda| \cdot \mathbb{P}(g = 1 \mid G_1 = x) + |\alpha x - \lambda(-1)| \cdot \mathbb{P}(g = -1 \mid G_1 = x) \right] \\ &= \mathbb{E}_x \left[|\alpha x - \lambda| \cdot \frac{1 + \theta_g(x)}{2} + |\alpha x + \lambda| \cdot \frac{1 - \theta_g(x)}{2} \right] \\ &= \mathbb{E}_x \left[\frac{|\alpha x - \lambda| + |\alpha x + \lambda|}{2} + \frac{|\alpha x - \lambda| - |\alpha x + \lambda|}{2} \theta_g(x) \right] \\ &= \int_{\mathbb{R}} \left(\frac{|\alpha x - \lambda| + |\alpha x + \lambda|}{2} + \frac{|\alpha x - \lambda| - |\alpha x + \lambda|}{2} \theta_g(x) \right) \phi(x) dx \\ &= \int_{\mathbb{R}} \frac{|\alpha x - \lambda| + |\alpha x + \lambda|}{2} \phi(x) dx + \int_{\mathbb{R}} \frac{|\alpha x - \lambda| - |\alpha x + \lambda|}{2} \theta_g(x) \phi(x) dx. \end{aligned} \quad (1.23)$$

Entonces, queremos maximizar (1.23) restringido a (1.21). Notemos que el primer término de (1.21) no depende de θ , luego podemos desestimarlos para maximizar.

Notemos además que los factores de los integrandos tanto de la función a maximizar como de la restricción son impares. Esto quiere decir que podemos, sin pérdida de generalidad, asumir que θ es impar. Si no lo fuera, podemos tomar su parte impar $\tilde{\theta}(x) = \frac{\theta(x) - \theta(-x)}{2}$, que alcanza los mismos valores para ambas integrales, y por lo tanto resuelve el problema de maximización.

Resumiendo, si definimos $\psi(x) := \frac{|\alpha x - \lambda| - |\alpha x + \lambda|}{2}$, lo que queremos es hallar $\theta \in L_\infty(\mathbb{R}_{\geq 0})$, $|\theta(x)| \leq 1$ tal que maximice:

$$2 \int_0^\infty \theta(x) \psi(x) \phi(x) dx \quad \text{restringido a que} \quad 2 \int_0^\infty \theta(x) x \phi(x) dx = \alpha. \quad (1.24)$$

La idea para resolver este problema de maximización es usar un algoritmo “goloso”. Como $\text{Im}(g) \subseteq [-1, 1]$, y $2\psi\phi \geq 0$, nos gustaría que $\theta(x) = 1$ para “la mayor cantidad” de x posibles, pero esto se ve impedido por la condición de restricción. La idea es entonces hacer que $\theta(x) = 1$ en los x más “efectivos”, que son aquellos donde $\frac{\psi(x)}{x}$ es lo más grande posible. Como resulta que $\frac{\psi(x)}{x}$ es monótona positiva, los valores más efectivos serán los más grandes, y por eso nuestra candidata a maximizar el problema será una función del tipo:

$$\theta_\gamma(x) := \begin{cases} 1 & x \geq \gamma \\ -1 & x < \gamma \end{cases} \quad (1.25)$$

Para comprobarlo, sea $\gamma \geq 0$ tal que

$$\alpha = \int_0^\gamma 2x\phi(x) dx - \int_\gamma^\infty 2x\phi(x) dx = \sqrt{\frac{2}{\pi}} (2e^{-\frac{\gamma^2}{2}} - 1). \quad (1.26)$$

Si existiera otra función $\tilde{\theta}$ que satisfaga la restricción de (1.21), entonces:

$$\underbrace{\left(\frac{\psi(x)}{x} - \frac{\psi(\gamma)}{\gamma}\right)}_{\text{sgn}(\cdot)=\text{sgn}(\theta_\gamma)} \theta_\gamma(x) = \left|\frac{\psi(x)}{x} - \frac{\psi(\gamma)}{\gamma}\right| \geq \left(\frac{\psi(x)}{x} - \frac{\psi(\gamma)}{\gamma}\right) \tilde{\theta}(x).$$

Luego

$$2 \int_0^\infty \left(\frac{\psi(x)}{x} - \frac{\psi(\gamma)}{\gamma}\right) x \theta_\gamma(x) \phi(x) dx \geq 2 \int_0^\infty \left(\frac{\psi(x)}{x} - \frac{\psi(\gamma)}{\gamma}\right) x \tilde{\theta}(x) \phi(x) dx. \quad (1.27)$$

Y como por la restricción:

$$2 \int_0^\infty \underbrace{\frac{\psi(\gamma)}{\gamma}}_{\text{cte}} x \theta_\gamma(x) \phi(x) dx = \frac{\psi(\gamma)}{\gamma} \alpha = 2 \int_0^\infty \underbrace{\frac{\psi(\gamma)}{\gamma}}_{\text{cte}} x \tilde{\theta}(x) \phi(x) dx. \quad (1.28)$$

Luego haciendo (1.27) menos (1.28):

$$\begin{aligned} 2 \int_0^\infty \frac{\psi(x)}{x} x \theta_\gamma(x) \phi(x) dx &\geq 2 \int_0^\infty \frac{\psi(x)}{x} x \tilde{\theta}(x) \phi(x) dx \\ 2 \int_0^\infty \psi(x) \theta_\gamma(x) \phi(x) dx &\geq 2 \int_0^\infty \psi(x) \tilde{\theta}(x) \phi(x) dx. \end{aligned}$$

Y entonces θ_γ es mejor que cualquier otra función que cumple la restricción. Luego, por la equivalencia de (1.21), tenemos una $g_\gamma \leftrightarrow \theta_\gamma$.

Volviendo a (1.20), ya sabemos que, para cada $\alpha \in \mathbb{R}_{\geq 0}$, resulta que el máximo se alcanza siempre en un g_γ apropiado para α como se define en (1.25), según (1.24). Pero por otro lado, (1.24) nos dice que, si $\alpha \geq 0$, entonces $0 \leq \gamma \leq \sqrt{2 \ln(2)}$. Luego, alcanza con tomar el supremo (1.20) sobre todas las g_γ con $0 \leq \gamma \leq \sqrt{2 \ln(2)}$. Como la asignación $\gamma \mapsto \|T_\lambda g_\gamma\|_1$ es continua en $[0, \sqrt{2 \ln(2)}]$, entonces existe $\gamma_0 \in [0, \sqrt{2 \ln(2)}]$ tal que

$$\|T_\lambda\|_{\infty \rightarrow 1} = \|T_\lambda g_{\gamma_0}\|_1. \quad (1.29)$$

Reescribimos entonces (1.23) para una g_γ del estilo de (1.25):

$$\|T_\lambda g_\gamma\|_1 = 2 \int_\gamma^\infty \phi(z) |\alpha z - \lambda| dz + 2 \int_0^\gamma \phi(z) |\alpha z + \lambda| dz,$$

donde, reescribiendo también (1.21), se tiene que

$$\alpha := 2 \int_\gamma^\infty \phi(z) z dz - 2 \int_0^\gamma \phi(z) z dz = \sqrt{\frac{2}{\pi}} \left(2e^{-\frac{\gamma^2}{2}} - 1 \right).$$

Como vimos, queremos maximizar $\frac{|1-\lambda|}{\|T_\lambda g_\gamma\|_1}$ con $\gamma \in [0, \sqrt{2\ln(2)}]$, lo que nos deja en una expresión explícita del problema a maximizar. Luego de algunas reducciones que omitiremos, se llega numéricamente a que los valores óptimos resultan ser:

$$\begin{aligned} \gamma &\approx 0,18009 & \alpha &\approx 0,77222 & \lambda &\approx 0,19748 \\ \frac{|1-\lambda|}{\|T_\lambda g_\gamma\|_1} &\approx 1,67695667\dots \end{aligned}$$

Y esto nos da la cota inferior que buscábamos.

Capítulo 2

Cotas superiores para la constante de Grothendieck

En este capítulo contaremos los avances que se realizaron en los últimos años respecto a la estimación de la constante de Grothendieck. El resultado central del capítulo es la demostración del paper [BMMN13] de que la cota superior de Krivine de $K_G \leq \frac{\pi}{2 \ln(1+\sqrt{2})}$ no es ajustada, lo que representó la primera mejora a las cotas superiores de la constante de Grothendieck en casi 35 años. Veremos que los autores de [BMMN13] no plantean una nueva cota superior explícita para K_G por considerar que el mejor valor que podían alcanzar era subóptimo. Dejan en cambio planteadas algunas preguntas que consideran que pueden ayudar a mejorar su propio método y quizás a encontrar la cota óptima en el futuro. Este resultado fue hallado por investigadores con una fuerte inclinación hacia la computación, que identificaron en la teoría un algoritmo aleatorio para estimar el valor de $OPT(A)$, que es el mejor algoritmo práctico que se conoce hasta el momento para esta tarea. Describiremos este algoritmo en este capítulo, ya que creemos que ayuda a la comprensión de las demostraciones y sus mejoras.

2.1. Generalización de la demostración de Krivine

En esta sección vamos a desarmar la demostración original de Krivine, e identificar los elementos generalizables para mejorar un poco la cota superior de la constante de Grothendieck. Veremos que en el proceso de esta desconstrucción nos encontraremos con una interpretación algorítmica de la demostración. El algoritmo que extraeremos es el mejor algoritmo conocido para el problema de calcular $OPT(A)$ para una matriz A .

Empecemos por mirar la utilización del Lema del seno (1.1.5) en la demostración de Krivine. Las funciones $\text{sgn}(x)$, $\text{sgn}(y)$ que se usan en la demostración nos sirven por dos motivos. Primero, el espacio de llegada es el conjunto $\{-1, 1\}$, que es el conjunto que usamos en las evaluaciones del problema $OPT(A)$.

Vamos a tratar de generalizar entonces, reemplazando el par de funciones $\text{sgn}(x)$, $\text{sgn}(y)$ por un par de funciones f, g medibles impares que caigan en $\{-1, 1\}$. Para mayor generalidad, podemos incluso permitir que el dominio de f y g sea \mathbb{R}^k para algún $k \in \mathbb{N}$, en lugar de simplemente \mathbb{R} . A la hora de redondear, vamos a tener que usar una matriz Gaussiana estándar de dimensión $k \times m$ en lugar de usar un vector n -dimensional.

Supongamos que ya tenemos elegidas funciones $f, g : \mathbb{R}^k \rightarrow \{-1, 1\}$. Por otro lado, el lema del seno nos da una forma de evaluar la esperanza:

$$\operatorname{sen} \left(\frac{\pi}{2} \mathbb{E}_G [\operatorname{sgn}(\langle u, G \rangle) \operatorname{sgn}(\langle v, G \rangle)] \right) = \langle u, v \rangle, \quad u, v \in \mathbb{S}^{m-1}.$$

Vamos a reescribir esto en una forma más conveniente para la generalización. En el fondo, lo que nos interesa es poder evaluar:

$$\mathbb{E}_G [\operatorname{sgn}(Gu) \operatorname{sgn}(Gv)] = \frac{2}{\pi} \operatorname{arc} \operatorname{sen} (\langle u, v \rangle) \quad u, v \in \mathbb{S}^{m-1}.$$

Entonces para nuestras funciones f, g nos gustaría también tener alguna función, que llamaremos $H_{f,g} : \mathbb{R} \rightarrow \mathbb{R}$ que cumpla similarmente:

$$\mathbb{E}_G [f(Gu)g(Gv)] = H_{f,g}(\langle u, v \rangle) \quad u, v \in \mathbb{S}^{m-1}.$$

Entonces, en la demostración original de Krivine podríamos decir que $H_{\operatorname{sgn}, \operatorname{sgn}} = \frac{2}{\pi} \operatorname{arc} \operatorname{sen} (-)$. Vamos a ver pronto que la función $H_{f,g}$, que sólo depende del producto interno de u y v y no de los vectores en sí, efectivamente existe.

Por último, para la demostración de Krivine es fundamental el uso de la constante $\gamma := \frac{2 \ln(1+\sqrt{2})}{\pi}$, es decir, la constante tal que $\operatorname{senh} \left(\frac{\pi}{2} - \right)$ evaluado en γ da 1. Notemos que el seno hiperbólico viene de haber invertido la función $\frac{\pi}{2} \operatorname{arc} \operatorname{sen}$ para obtener la función $\operatorname{sen} \left(\frac{\pi}{2} - \right)$, y luego mirar la serie absoluta de esa función para obtener $\operatorname{senh} \left(\frac{\pi}{2} \right)$. Dada una función definida por una serie de potencias $h(x) := \sum_{k=0}^{\infty} a_k x^k$, vamos a definir su serie absoluta como

$$h^{abs}(x) := \sum_{k=0}^{\infty} |a_k| x^k.$$

Luego podemos reescribir la condición anterior como que $\gamma = \frac{2 \ln(1+\sqrt{2})}{\pi}$ es el número que cumple que

$$\left(\left(\frac{2}{\pi} \operatorname{arc} \operatorname{sen} \right)^{-1} \right)^{abs} (\gamma) = \left(\operatorname{sen} \left(\frac{\pi}{2} - \right) \right)^{abs} (\gamma) = \operatorname{senh} \left(\frac{\pi}{2} \gamma \right) = 1.$$

En nuestro nuevo contexto de funciones f, g más generales, esto equivale a pedir que exista un $\gamma_{f,g} \in \mathbb{R}_{>0}$ tal que

$$\left(H_{f,g}^{-1} \right)^{abs} (\gamma_{f,g}) = 1. \quad (2.1)$$

Resumiendo todo esto un poco, estos son los elementos que nos permitirán generalizar la demostración de Krivine, y sus realizaciones en la demostración original:

$$\begin{array}{ccc} \operatorname{sgn}, \operatorname{sgn} : \mathbb{R} \rightarrow \{-1, 1\} & \longleftarrow & f, g : \mathbb{R}^k \rightarrow \{-1, 1\} \text{ medibles impares} \\ \frac{2}{\pi} \operatorname{arc} \operatorname{sen} : \mathbb{R} \rightarrow \mathbb{R} & \longleftarrow & H_{f,g} : \mathbb{R} \rightarrow \mathbb{R} \\ \mathbb{E}_G [\operatorname{sgn}(Gu) \operatorname{sgn}(Gv)] = \frac{2}{\pi} \operatorname{arc} \operatorname{sen} (\langle u, v \rangle) & \longleftarrow & \mathbb{E}_G [f(Gu)g(Gv)] = H_{f,g}(\langle u, v \rangle) \\ \exists \gamma = \frac{2 \ln(1+\sqrt{2})}{\pi}, \operatorname{senh} \left(\frac{\pi}{2} \gamma \right) = 1 & \longleftarrow & \exists \gamma_{f,g} \in \mathbb{R}_{>0}, \left(H_{f,g}^{-1} \right)^{abs} (\gamma_{f,g}) = 1. \end{array}$$

A los pares de funciones f, g que tienen estas propiedades, los denominaremos Métodos de Redondeo de Krivine (MRK).

Para empezar, vamos a demostrar que cuando f, g son funciones impares medibles, efectivamente existe una función $H_{f,g}$ como la de la arriba que sólo depende del producto interno de los dos vectores u, v . Por comodidad a la hora de hacer las cuentas, vamos a usar una versión levemente modificada de la esperanza que calculábamos antes:

$$\mathbb{E}_G \left[f \left(\frac{1}{\sqrt{2}} Gu \right) g \left(\frac{1}{\sqrt{2}} Gv \right) \right].$$

Notemos que esto no da mayor generalidad, ya que hay una biyección entre las funciones $f \leftrightarrow f \left(\frac{1}{\sqrt{2}} - \right), g \leftrightarrow g \left(\frac{1}{\sqrt{2}} - \right)$, y todas ellas pertenecen al espacio de funciones de \mathbb{R}^k en $\{-1, 1\}$.

Ahora sí, dados $u, v \in \mathbb{S}^{m-1}$, $\|u\| = \|v\| = 1$ definimos:

$$\tilde{H}_{f,g}(u, v) := \mathbb{E}_G \left[f \left(\frac{1}{\sqrt{2}} Gx \right) g \left(\frac{1}{\sqrt{2}} Gy \right) \right]. \quad (2.2)$$

Como $u, v \in \mathbb{S}^{m-1}$, de manera análoga a lo que hicimos en (1.1.7), tenemos que $v = \langle u, v \rangle u + \sqrt{1 - \langle u, v \rangle^2} u^\perp$. Por lo tanto, la distribución de las variables aleatorias (Gu, Gv) es equivalente a la distribución de $(Gu, \langle u, v \rangle Gu + \sqrt{1 - \langle u, v \rangle^2} Gu^\perp)$. Pero Gu, Gu^\perp resultan ser vectores Gausianos de dimensión k estándar, e independientes entre sí. Luego esta distribución equivale a la distribución $(G_1, \langle u, v \rangle G_1 + \sqrt{1 - \langle u, v \rangle^2} G_2)$, donde G_1, G_2 son vectores aleatorios Gausianos estándar e independientes de dimensión k .

Resulta entonces que:

$$\tilde{H}_{f,g}(u, v) = \mathbb{E}_{G_1, G_2} \left[f \left(\frac{1}{\sqrt{2}} G_1 \right) g \left(\frac{\langle u, v \rangle}{\sqrt{2}} G_1 + \frac{\sqrt{1 - \langle u, v \rangle^2}}{\sqrt{2}} G_2 \right) \right].$$

Como la única información de u y v que usa $\tilde{H}_{f,g}$ es el valor del producto interno $\langle u, v \rangle$, podemos definir simplemente:

$$H_{f,g}(t) = \mathbb{E}_{G_1, G_2} \left[f \left(\frac{1}{\sqrt{2}} G_1 \right) g \left(\frac{t}{\sqrt{2}} G_1 + \frac{\sqrt{1 - t^2}}{\sqrt{2}} G_2 \right) \right].$$

De esta manera $\tilde{H}_{f,g}(u, v) = H_{f,g}(\langle u, v \rangle)$.

Tratemos entonces de conseguir una expresión más explícita para $H_{f,g}(t)$:

Definición 2.1.1. Un Método de Redondeo de Krivine de dimensión $k \in \mathbb{N}$ consiste en un par de funciones medibles impares $f, g : \mathbb{R}^k \rightarrow \{-1, 1\}$, tales que la función definida a partir de ellas:

$$H_{f,g}(z) = \mathbb{E}_{G_1, G_2} \left[f \left(\frac{1}{\sqrt{2}} G_1 \right) g \left(\frac{z}{\sqrt{2}} G_1 + \frac{\sqrt{1 - z^2}}{\sqrt{2}} G_2 \right) \right]$$

cumple que $H_{f,g}^{-1}$ es inversible en un entorno del origen, y si $H_{f,g}^{-1}(z) = \sum_{j=0}^{\infty} a_j z^j$ es la expresión como serie de la función inversa en ese entorno, entonces existe $\gamma := \gamma(f, g) \in \mathbb{R}_{>0}$ tal que

$$\sum_{j=0}^{\infty} |a_j| \gamma^j = 1.$$

Vamos a ver como con estos elementos alcanza para probar que $K_G < \frac{1}{\gamma_{f,g}}$ de forma análoga a lo que hicimos en (1.3.1).

Sea $z \in \ell_\infty^n \otimes \ell_\infty^n$, con $\|z\|_{w_2} < 1$, y sea $a := \ln(1 + \sqrt{2})$. Queremos ver que $\|z\|_\pi < \frac{1}{\gamma_{f,g}}$

Se tiene entonces, usando que w_2 es una norma de álgebra:

$$\|H_{f,g}^{-1}(\gamma_{f,g}z)\|_{w_2} \leq \left(H_{f,g}^{-1}\right)^{abs}(\gamma_{f,g}\|z\|_{w_2}) < \left(H_{f,g}^{-1}\right)^{abs}(\gamma_{f,g}) = 1.$$

Entonces, para algún $m \in \mathbb{N}$, existe una representación $\tilde{u}_i, \tilde{v}_j \in \mathbb{S}^{m-1}$ como producto interno de $H_{f,g}^{-1}(\gamma_{f,g}z)$. Entonces, para cada $i, j \in \{1, \dots, n\}$:

$$\gamma_{f,g}z(e_i, e_j) = H_{f,g}\left(H_{f,g}^{-1}(\gamma_{f,g}z(e_i, e_j))\right) = H_{f,g}(\langle \tilde{u}_i, \tilde{v}_j \rangle) = \mathbb{E}_G \left[f\left(\frac{1}{\sqrt{2}}G\tilde{u}_i\right) g\left(\frac{1}{\sqrt{2}}G\tilde{v}_j\right) \right].$$

Entonces, sumando sobre cada $e_i^* \otimes e_j^*$ canónico:

$$\begin{aligned} \gamma_{f,g}z &= \sum_{i,j=1}^n \gamma_{f,g}z(e_i, e_j) e_i^* \otimes e_j^* = \sum_{i,j=1}^n \mathbb{E}_G \left[f\left(\frac{1}{\sqrt{2}}G\tilde{u}_i\right) g\left(\frac{1}{\sqrt{2}}G\tilde{v}_j\right) \right] e_i^* \otimes e_j^* = \\ \gamma_{f,g}z &= \mathbb{E}_G \left[\left(\sum_{i=1}^n f\left(\frac{1}{\sqrt{2}}G\tilde{u}_i\right) e_i^* \right) \otimes \left(\sum_{j=1}^n g\left(\frac{1}{\sqrt{2}}G\tilde{v}_j\right) e_j^* \right) \right]. \end{aligned}$$

Tomando $\|-\|_\pi$ a ambos lados, tenemos que:

$$\begin{aligned} \|\gamma_{f,g}z\|_\pi &\leq \mathbb{E}_G \left[\left\| \sum_{i=1}^n f\left(\frac{1}{\sqrt{2}}G\tilde{u}_i\right) e_i^* \right\| \left\| \sum_{j=1}^n g\left(\frac{1}{\sqrt{2}}G\tilde{v}_j\right) e_j^* \right\| \right] = 1 \\ \|z\|_\pi &\leq \frac{1}{\gamma_{f,g}}. \end{aligned}$$

Entonces el $\gamma_{f,g}$ que le corresponde a un Método de Redondeo de Krivine (f, g) está fuertemente ligado a la constante de Grothendieck. En particular, nos va a interesar hallar, para mejorar las cotas superiores de la constante, los MRK con mayor valor de γ asociado posible. Por este motivo, queremos conocer lo más posible sobre las funciones $H_{f,g}$, sus inversas y las series absolutas de estas. Empecemos por encontrar una expresión más explícita de $H_{f,g}$ para un MRK dado:

$$\begin{aligned} H_{f,g}(t) &= \mathbb{E}_{G_1, G_2} \left[f\left(\frac{1}{\sqrt{2}}G_1\right) g\left(\frac{t}{\sqrt{2}}G_1 + \frac{\sqrt{1-t^2}}{\sqrt{2}}G_2\right) \right] \\ &= \frac{1}{(2\pi)^k} \int_{\mathbb{R}^k} \int_{\mathbb{R}^k} f\left(\frac{\tilde{x}}{\sqrt{2}}\right) g\left(\frac{t}{\sqrt{2}}\tilde{x} + \frac{\sqrt{1-t^2}}{\sqrt{2}}\tilde{y}\right) e^{-\frac{\|\tilde{x}\|^2 + \|\tilde{y}\|^2}{2}} d\tilde{x} d\tilde{y} \end{aligned}$$

Si $|t| < 1$, usando $\tilde{x} := \sqrt{2}x, \tilde{y} := \frac{\sqrt{2}y - \sqrt{2}tx}{\sqrt{1-t^2}}$ con $|J| = \frac{2^k}{(1-t^2)^{k/2}}$:

$$\begin{aligned} &= \frac{1}{(2\pi)^k} \frac{2^k}{(1-t^2)^{k/2}} \int_{\mathbb{R}^k} \int_{\mathbb{R}^k} f(x) g(y) e^{-\frac{\|x\|^2 + \|y\|^2 - 2t\langle x, y \rangle}{1-t^2}} dx dy \\ &= \frac{1}{\pi^k (1-t^2)^{k/2}} \int_{\mathbb{R}^k} \int_{\mathbb{R}^k} f(x) g(y) e^{-\frac{\|x\|^2 + \|y\|^2 - 2t\langle x, y \rangle}{1-t^2}} dx dy. \end{aligned}$$

Notemos que a partir de esta expresión es totalmente plausible evaluar a $H_{f,g}$ en valores complejos. De hecho, se puede demostrar el siguiente hecho:

Lema 2.1.2. *Dado un Método de Redondeo de Krivine (f, g) , la función $H_{f,g}$ asociada tiene una extensión analítica en la franja $\mathcal{S} := \{z \in \mathbb{C} \mid |\operatorname{Re}(z)| < 1\}$. Más aún, todas las $H_{f,g}$ están uniformemente acotadas en \mathcal{S} .*

Como f y g son impares, es sencillo demostrar que $H_{f,g}$ también debe serlo. Para esto usamos el cambio de variables $\tilde{x} := -x$:

$$\begin{aligned} H_{f,g}(-z) &= \frac{1}{\pi^k(1-z^2)^{k/2}} \int_{\mathbb{R}^k} \int_{\mathbb{R}^k} f(x) g(y) e^{-\frac{\|x\|^2 + \|y\|^2 + 2z\langle x, y \rangle}{1-z^2}} dx dy = \\ &= -\frac{1}{\pi^k(1-z^2)^{k/2}} \int_{\mathbb{R}^k} \int_{\mathbb{R}^k} f(\tilde{x}) g(y) e^{-\frac{\|\tilde{x}\|^2 + \|y\|^2 - 2z\langle \tilde{x}, y \rangle}{1-z^2}} d\tilde{x} dy = -H_{f,g}(z). \end{aligned}$$

Esto nos dice entonces que también la inversa $H_{f,g}^{-1}$ debe ser impar, y por lo tanto su expresión como serie de potencias debe ser del tipo:

$$H_{f,g}^{-1}(z) = \sum_{j=0}^{\infty} a_{2j+1} z^{2j+1}.$$

Consideremos el caso particular en que esta serie de potencias es alternada en sus potencias impares. Es decir, el caso en que $\operatorname{sgn}(a_{2j+1}) = (-1)^j$. Este es el caso del MRK original de Krivine de las funciones $(\operatorname{sgn}, \operatorname{sgn})$, ya que en este caso $H_{f,g}^{-1}(z) = \operatorname{sen}\left(\frac{\pi}{2}\right)$, que tiene una serie de potencias alternada. En estos casos particulares, se cumple la siguiente propiedad:

$$\begin{aligned} H_{f,g}^{-1}(i\gamma_{f,g}) &= \sum_{j=0}^{\infty} a_{2j+1} (i\gamma_{f,g})^{2j+1} = \sum_{j=0}^{\infty} (-1)^j |a_{2j+1}| i^{2j} i \gamma_{f,g}^{2j+1} \\ &= i \left(\sum_{j=0}^{\infty} (-1)^j |a_{2j+1}| (-1)^j \gamma_{f,g}^{2j+1} \right) = i \underbrace{\left(\sum_{j=0}^{\infty} |a_{2j+1}| \gamma_{f,g}^{2j+1} \right)}_{=1 \text{ por definición de } \gamma_{f,g}} = i. \end{aligned}$$

Pasando la $H_{f,g}^{-1}$ como inversa, y la i dividiendo al otro lado:

$$\frac{H_{f,g}(i)}{i} = \gamma_{f,g}. \quad (2.3)$$

Es decir, en los casos en que la $H_{f,g}^{-1}$ sea alternada en el sentido que describimos, tenemos una expresión explícita para calcular $\gamma_{f,g}$. Esto es lo que ocurre en la demostración original de Krivine cuando usamos $\gamma_{\operatorname{sgn}, \operatorname{sgn}} := \frac{2}{\pi} \operatorname{arc} \operatorname{sen}(i) = \frac{2 \ln(1+\sqrt{2})}{\pi}$. Dada esta característica especial de este tipo de MRK, nos vemos motivados a darles una denominación especial:

Definición 2.1.3. Un Método de Redondeo de Krivine se dice **alternado** si los coeficientes de la serie de potencias de $H_{f,g}^{-1}$ son alternados en signo:

$$\operatorname{sgn}(a_{2j+1}) = (-1)^j \quad \forall j \in \mathbb{N}_0. \quad (2.4)$$

Y, dada su gran importancia a futuro, enunciamos el hecho que acabamos de demostrar en (2.3):

Lema 2.1.4. *Si (f, g) es un Método de Redondeo de Krivine Alternado, entonces*

$$\gamma_{f,g} = \frac{H_{f,g}(i)}{i}.$$

Por supuesto, ahora tenemos un especial interés en evaluar nuestra función $H_{f,g}$ en $z = i$, y ver qué obtenemos. Podemos hacer esta cuenta incluso sin la hipótesis de que el MRK sea alternado.

$$\begin{aligned} \frac{H_{f,g}(i)}{i} &= \frac{1}{i\pi^k(1-i^2)^{k/2}} \int_{\mathbb{R}^k} \int_{\mathbb{R}^k} f(x)g(y) e^{-\frac{\|x\|^2 + \|y\|^2 - 2i\langle x,y \rangle}{1-i^2}} dx dy \\ &= \frac{1}{i(\sqrt{2\pi})^k} \int_{\mathbb{R}^k} \int_{\mathbb{R}^k} f(x)g(y) e^{-\frac{\|x\|^2 + \|y\|^2 - 2i\langle x,y \rangle}{2}} dx dy \\ &= \frac{1}{i(\sqrt{2\pi})^k} \int_{\mathbb{R}^k} \int_{\mathbb{R}^k} f(x)g(y) e^{-\frac{\|x\|^2 + \|y\|^2}{\sqrt{2}}} (\cos(\langle x,y \rangle) + i \cdot \text{sen}(\langle x,y \rangle)) dx dy \end{aligned}$$

Como f y g son impares y tanto el coseno como la exponencial son pares en x y en y , entonces

$$\frac{1}{i(\sqrt{2\pi})^k} \int_{\mathbb{R}^k} \int_{\mathbb{R}^k} f(x)g(y) e^{-\frac{\|x\|^2 + \|y\|^2}{\sqrt{2}}} \cos(\langle x,y \rangle) dx dy = 0$$

Entonces la ecuación de arriba nos queda:

$$\frac{H_{f,g}(i)}{i} = \frac{1}{(\sqrt{2\pi})^k} \int_{\mathbb{R}^k} \int_{\mathbb{R}^k} f(x)g(y) e^{-\frac{\|x\|^2 + \|y\|^2}{\sqrt{2}}} \text{sen}(\langle x,y \rangle) dx dy.$$

La expresión de la derecha consiste básicamente de una constante $\frac{1}{(\sqrt{2\pi})^k}$, que depende de k , y de una expresión que podemos ver como una forma bilineal $B_K : L^\infty(\mathbb{R}^k) \times L^\infty(\mathbb{R}^k) \rightarrow \mathbb{R}$ evaluada en f y en g :

Definición 2.1.5. Definimos a la forma bilineal de König de dimensión $k \in \mathbb{N}$, $B_K : L^\infty(\mathbb{R}^k) \times L^\infty(\mathbb{R}^k) \rightarrow \mathbb{R}$ como:

$$B_K(f, g) := \int_{\mathbb{R}^k} \int_{\mathbb{R}^k} f(x)g(y) e^{-\frac{\|x\|^2 + \|y\|^2}{2}} \text{sen}(\langle x,y \rangle) dx dy.$$

Con esta notación, resulta que acabamos de probar el siguiente lema:

Lema 2.1.6. *Para cualquier Método de Redondeo de Krivine (f, g) de dimensión k se tiene que*

$$\frac{H_{f,g}(i)}{i} = \frac{B_K(f, g)}{(\sqrt{2\pi})^k}.$$

Juntando los resultados de (2.1.4) y (2.1.6) obtenemos entonces:

$$\gamma_{f,g} \stackrel{f,g \text{ MRKA}}{=} \frac{H_{f,g}(i)}{i} = \frac{B_K(f,g)}{(\sqrt{2\pi})^k},$$

donde la segunda igualdad se cumple para cualquier MRK, pero la primera se cumple si la MRK es alternada. Esto nos dice que parece haber una relación entre los $\gamma_{f,g}$ altos (que son los que nos dan mejores cotas superiores para la constante de Grothendieck) y las funciones f, g que dan valores altos sobre la forma bilineal de König. Vamos entonces a analizar qué tipo de funciones $\|f\|_\infty \|g\|_\infty \leq 1$ dan los valores más altos en la forma bilineal de König, con la esperanza de que esto nos acerque a MRK, tal vez alternados, que nos permitan acotar mejor K_G .

Empecemos por ver que la forma bilineal de König se maximiza en cualquiera de sus dimensiones.

Lema 2.1.7. *Para todo $k \in \mathbb{N}$ existen funciones $f_{max}^{(k)}, g_{max}^{(k)} \in L_\infty(\mathbb{R}^k)$, $\|f_{max}^{(k)}\|_\infty, \|g_{max}^{(k)}\|_\infty \leq 1$ tales que*

$$B_K(f_{max}^{(k)}, g_{max}^{(k)}) = \sup_{\|f\|_\infty, \|g\|_\infty \leq 1} B_K(f, g).$$

Más aún, se cumple que $|f_{max}^{(k)}| = |g_{max}^{(k)}| = 1$ para casi todo punto y

$$f_{max}^{(k)}(x) = \operatorname{sgn} \left(\int_{\mathbb{R}^k} g_{max}^{(k)}(y) e^{-\frac{\|y\|^2}{2}} \operatorname{sen}(\langle x, y \rangle) dy \right).$$

Demostración. Definimos

$$M := \sup_{\|f\|_\infty, \|g\|_\infty \leq 1} B_K(f, g).$$

Esto implica inmediatamente que existen sucesiones de funciones $\{f_n\}, \{g_n\}$, $\|f_n\|_\infty, \|g_n\|_\infty \leq 1$ tales que $B_K(f_n, g_n) \xrightarrow{n \rightarrow \infty} M$.

Como las funciones f_n y g_n son esencialmente acotadas, podemos considerarlas como funciones en $L_2(\mathcal{G})$, el espacio de medida finita dado por la medida de densidad $\mu_{\mathcal{G}}(x) := e^{-\frac{\|x\|^2}{2}}$. De hecho, todas las funciones f_n, g_n tienen que estar en la bola de $L_2(\mathcal{G})$ centrada en cero con radio $(2\pi)^k$, ya que es lo máximo que puede integrar una función esencialmente acotada por 1 en ese espacio.

La bola cerrada es w^* -compacta por Banach-Alaoglu. Luego es w -compacta, porque L_2 es un espacio de Hilbert y por lo tanto reflexivo. Por último, por el Teorema de Eberlein-Smulian en un espacio de Banach un conjunto es w -compacto si y sólo si es secuencialmente w -compacto (es decir, las sucesiones tienen subsucesiones convergentes, no sólo subredes). Entonces, pasando a una subsucesión sin perder generalidad, podemos suponer que existen funciones $f_{max}, g_{max} \in L_2(\mathcal{G})$ tales que $f_n \xrightarrow{w(L_2(\mathcal{G}))} f_{max}, g_n \xrightarrow{w(L_2(\mathcal{G}))} g_{max}$.

Se puede ver fácilmente que $\|f_{max}\|_\infty, \|g_{max}\|_\infty \leq 1$. Si no fuera así, existiría, por ejemplo, un conjunto de medida positiva \mathcal{U} donde $f \geq 1 + \varepsilon$ para $\varepsilon > 0$. Pero entonces, con $\chi_{\mathcal{U}}$ la función característica del conjunto \mathcal{U} , tenemos que

$$\mu_{\mathcal{G}}(\mathcal{U})(1 + \varepsilon) \leq \langle f_{max}, \chi_{\mathcal{U}} \rangle = \lim_{n \rightarrow \infty} \langle f_n, \chi_{\mathcal{U}} \rangle \leq \mu_{\mathcal{G}}(\mathcal{U}).$$

Absurdo, luego $\|f_{max}\|_\infty < 1$, y análogamente para g .

Ahora miremos la diferencia:

$$B_K(f_{max}, g_{max}) - B_K(f_n, g_n) = \underbrace{B_K(f_{max}, g_{max}) - B_K(f_n, g_{max})}_{\xrightarrow{n \rightarrow \infty} 0, \text{ porque } f_n \xrightarrow{w} f_{max}} + \underbrace{B_K(f_n, g_{max}) - B_K(f_n, g_n)}_{\xrightarrow{n \rightarrow \infty} 0, \text{ porque } g_n \xrightarrow{w} g_{max} \text{ y } |f_n| \text{ se mayor por } 1}.$$

Luego $B_K(f_{max}, g_{max}) = M$, que es lo que queríamos probar.

Ahora que ya sabemos de la existencia de al menos un par de funciones $f_{max}^{(k)}, g_{max}^{(k)}$, no necesariamente único, que maximiza la forma bilineal de König de dimensión k , veamos que relación tienen estas dos funciones entre sí. Sabemos que

$$\begin{aligned} B_K(f_{max}^{(k)}, g_{max}^{(k)}) &= \int_{\mathbb{R}^k} \int_{\mathbb{R}^k} f_{max}^{(k)}(x), g_{max}^{(k)}(y) e^{-\frac{\|x\|^2 + \|y\|^2}{2}} \text{sen}(\langle x, y \rangle) dy dx \\ &= \int_{\mathbb{R}^k} \underbrace{e^{-\frac{\|x\|^2}{2}}}_{>0} \underbrace{f_{max}^{(k)}(x)}_{\in [-1,1]} \left(\int_{\mathbb{R}^k} g_{max}^{(k)}(y) e^{-\frac{\|y\|^2}{2}} \text{sen}(\langle x, y \rangle) dy \right) dx \\ &\leq \int_{\mathbb{R}^k} e^{-\frac{\|x\|^2}{2}} \left| \int_{\mathbb{R}^k} g_{max}^{(k)}(y) e^{-\frac{\|y\|^2}{2}} \text{sen}(\langle x, y \rangle) dy \right| dx. \end{aligned}$$

Y este último valor se alcanza exactamente si

$$f_{max}^{(k)}(x) = \text{sgn} \left(\int_{\mathbb{R}^k} g_{max}^{(k)}(y) e^{-\frac{\|y\|^2}{2}} \text{sen}(\langle x, y \rangle) dy \right) \quad \text{en casi todo punto } x \in \mathbb{R}^k.$$

Si hacemos la misma cuenta con $g_{max}^{(k)}$, llegamos equivalentemente a que

$$g_{max}^{(k)}(y) = \text{sgn} \left(\int_{\mathbb{R}^k} f_{max}^{(k)}(x) e^{-\frac{\|x\|^2}{2}} \text{sen}(\langle x, y \rangle) dx \right) \quad \text{en casi todo punto } y \in \mathbb{R}^k.$$

Y además, como tanto $f_{max}^{(k)}$ como $g_{max}^{(k)}$ son en casi todo punto el signo de alguien, resulta que $f_{max}^{(k)}(x), g_{max}^{(k)}(y) \in \{-1, 1\}$ para casi todo punto con $x, y \in \mathbb{R}^k$, que era lo último que nos faltaba probar. \square

Con el objetivo de hallar entonces la cota óptima de la desigualdad de Grothendieck, parece que tendrá sentido seguir analizando la forma bilineal de König y sus maximizadores en cada dimensión. Vamos a analizar lo que se sabe en este sentido en las próximas secciones luego de hacer un breve paréntesis para contar una interpretación algorítmica de los Métodos de Redondeo de Krivine, que nos permitirá alejarnos un poco más del contexto de normas tensoriales.

2.2. Uso algorítmico de los Métodos de Redondeo de Krivine

Vamos a contar en esta sección la forma en que un Método de Redondeo de Krivine como definimos en (2.1.1) puede darnos un algoritmo de redondeo aleatorio polinomial para hallar un valor aproximado de $OPT(A)$, con la garantía de alcanzar, en esperanza, el valor $\gamma_{f,g} OPT(A)$. Esto es particularmente importante, ya que veremos en la próxima sección que el problema de calcular el valor exacto de $OPT(A)$ es NP-hard, y que hay algunos motivos para creer que

el problema de aproximar $OPT(A)$ con garantía mejor que $\frac{OPT(A)}{KG}$ también es NP-hard. Este algoritmo, que es el mejor que se conoce hasta el momento, tiene varias aplicaciones que veremos en la sección 4, y es probable que haya sido una parte importante de las motivaciones para que gente de la computación atacara y avanzara en el problema de la desigualdad de Grothendieck. Si bien este algoritmo está más relacionado con la segunda parte de la tesis, vamos a exponerlo ahora porque creemos que ilustra un poco mejor el proceso por el cual los Métodos de Redondeo de Krivine nos llevan a cotas superiores, y que puede quedar un poco escondido en el uso del álgebra dada por la norma w_2 .

Básicamente, en la demostración que construimos usando un MRK se usaba que

$$\|z\|_{w_2} < 1 \Rightarrow \|H_{f,g}^{-1}(\gamma_{f,g}z)\|_{w_2} < 1, \quad (2.5)$$

y luego se usaban este hecho y las funciones f, g explícitamente para equiparar el tensor $H_{f,g}^{-1}(\gamma_{f,g}z)$ a la esperanza de los tensores dados por una especie de “redondeo”:

$$\gamma_{f,g}z = \mathbb{E}_G \left[\left(\sum_{i=1}^n f\left(\frac{1}{\sqrt{2}}G\tilde{u}_i\right) e_i \right) \otimes \left(\sum_{j=1}^n g\left(\frac{1}{\sqrt{2}}G\tilde{v}_j\right) e_j \right) \right]. \quad (2.6)$$

Vamos a entender un poco esto. En el paso (2.5), lo que se está haciendo es ver que poder representar un tensor como producto interno de cierto tamaño acotado implica que otro tensor en función de este se puede representar como producto interno con tamaño acotado. Más precisamente, habíamos visto que $\|z\|_{w_2} < 1$ implica que el tensor $z \in \ell_\infty^n \otimes \ell_\infty^n$ se puede representar como $z(e_i, e_j) = \langle u_i, v_j \rangle$, con $u_i, v_j \in \mathbb{S}^{2n-1}$.

Aplicarle la función $H_{f,g}^{-1}$ al tensor $\gamma_{f,g}$ quiere decir aplicarle la serie $\sum_{d=1}^{\infty} a_d x^d$ al tensor con el producto definido por el álgebra normada w_2 . Dicho producto no era otra cosa que el producto punto a punto de los tensores de $\ell_\infty^n \otimes \ell_\infty^n$ interpretados como formas bilineales en $\ell_1^n \times \ell_1^n$. Dicho de otra manera, para cada $i, j \in \{1, \dots, n\}$ se cumple que

$$\left(H_{f,g}^{-1}(\gamma_{f,g}z) \right) (e_i, e_j) = H_{f,g}^{-1}(\gamma_{f,g}z(e_i, e_j)).$$

Y entonces el hecho de que $\|H_{f,g}^{-1}(\gamma_{f,g}z)\|_{w_2} < 1$ quiere decir que existe una representación \tilde{u}_i, \tilde{v}_j como producto interno de $H_{f,g}^{-1}(\gamma_{f,g}z)$ tal que

$$H_{f,g}^{-1}(\gamma_{f,g}z)(\langle u_i, v_j \rangle) = \langle \tilde{u}_i, \tilde{v}_j \rangle.$$

Lo interesante de esta conversión de representaciones vectoriales $\langle u_i, v_j \rangle \rightsquigarrow \langle \tilde{u}_i, \tilde{v}_j \rangle$ es que es consistente en el sentido que podemos pensarla como conversiones $u_i \rightsquigarrow \tilde{u}_i, v_j \rightsquigarrow \tilde{v}_j$ que cumplen

$$\begin{pmatrix} \langle u_1, v_1 \rangle & \cdots & \langle u_1, v_n \rangle \\ \vdots & \ddots & \vdots \\ \langle u_n, v_1 \rangle & \cdots & \langle u_n, v_n \rangle \end{pmatrix} \xrightarrow{H_{f,g}^{-1}(\gamma_{f,g}z)} \begin{pmatrix} \langle \tilde{u}_1, \tilde{v}_1 \rangle & \cdots & \langle \tilde{u}_1, \tilde{v}_n \rangle \\ \vdots & \ddots & \vdots \\ \langle \tilde{u}_n, \tilde{v}_1 \rangle & \cdots & \langle \tilde{u}_n, \tilde{v}_n \rangle \end{pmatrix}, \text{ coordenada a coordenada.} \quad (2.7)$$

La existencia de esta representación, que habíamos justificado con el uso de la norma w_2 , también podemos verla describiendo explícitamente la transformación que lleva los vectores u_i, v_j a los vectores \tilde{u}_i, \tilde{v}_j .

Para esto, suponiendo que los vectores u_i, v_j pertenecen a un espacio de Hilbert \mathcal{H} , se construye el espacio

$$\mathcal{H}^\infty := \bigoplus_{d=1}^{\infty} \mathcal{H}^{\otimes d}, \quad (2.8)$$

donde cada $\mathcal{H}^{\otimes d}$ es el espacio $\underbrace{\mathcal{H} \otimes \mathcal{H} \otimes \cdots \otimes \mathcal{H}}_{d \text{ veces}}$ con el producto interno

$$\langle (a_l)_l, (b_l)_l \rangle_{\mathcal{H}^\infty} := \prod_{l=1}^d \langle a_l, b_l \rangle.$$

En este espacio se definen los vectores $\tilde{u}_i, \tilde{v}_j \in \mathcal{H}^\infty$ de la siguiente manera:

$$\tilde{u}_i := \sum_{d=0}^{\infty} \sqrt{|a_d| \gamma^d} x^{\otimes d} \quad (2.9)$$

$$\tilde{v}_j := \sum_{d=0}^{\infty} \operatorname{sgn}(a_d) \sqrt{|a_d| \gamma^d} x^{\otimes d}. \quad (2.10)$$

Veamos que, con esta definición, los vectores \tilde{u}_i, \tilde{v}_j cumplen las propiedades que queremos:

$$\begin{aligned} \|\tilde{u}_i\|_{\mathcal{H}^\infty}^2 &= \left\| \bigoplus_{d=1}^{\infty} \operatorname{sgn}(a_d) \sqrt{|a_d| \gamma^d} u_i^{\otimes d} \right\|_{\mathcal{H}^\infty}^2 = \sum_{d=1}^{\infty} \left\| \sqrt{|a_d| \gamma^d} u_i^{\otimes d} \right\|_{\mathcal{H}^{\otimes d}}^2 \\ &= \sum_{d=1}^{\infty} \left(\sqrt{|a_d| \gamma^d} \|u_i^{\otimes d}\|_{\mathcal{H}^{\otimes d}} \right)^2 = \sum_{d=1}^{\infty} |a_d| \gamma^d \underbrace{\|u_i\|_{\mathcal{H}}^{2d}}_{=1} = \sum_{d=1}^{\infty} |a_d| \gamma^d = 1, \end{aligned}$$

por definición del γ del MRK.

Análogamente tenemos que, para todo j ,

$$\begin{aligned} \|\tilde{v}_j\|_{\mathcal{H}^\infty}^2 &= \left\| \bigoplus_{d=1}^{\infty} \operatorname{sgn}(a_d) \sqrt{|a_d| \gamma^d} v_j^{\otimes d} \right\|_{\mathcal{H}^\infty}^2 = \sum_{d=1}^{\infty} \left\| \operatorname{sgn}(a_d) \sqrt{|a_d| \gamma^d} v_j^{\otimes d} \right\|_{\mathcal{H}^{\otimes d}}^2 \\ &= \sum_{d=1}^{\infty} \left(\sqrt{|a_d| \gamma^d} \|v_j^{\otimes d}\|_{\mathcal{H}^{\otimes d}} \right)^2 = \sum_{d=1}^{\infty} |a_d| \gamma^d \underbrace{\|v_j\|_{\mathcal{H}}^{2d}}_{=1} = \sum_{d=1}^{\infty} |a_d| \gamma^d = 1 \end{aligned}$$

Y por lo tanto, $\|\tilde{u}_i\|_{\mathcal{H}^\infty} = \|\tilde{v}_j\|_{\mathcal{H}^\infty} = 1$, para todo $i, j \in \{1, \dots, n\}$. Para el producto interno entre distintos vectores, sólo nos interesan los productos del tipo $\langle \tilde{u}_i, \tilde{v}_j \rangle$:

$$\begin{aligned} \langle \tilde{u}_i, \tilde{v}_j \rangle &= \left\langle \bigoplus_{d=1}^{\infty} \sqrt{|a_d| \gamma^d} u_i^{\otimes d}, \bigoplus_{d=1}^{\infty} \operatorname{sgn}(a_d) \sqrt{|a_d| \gamma^d} v_j^{\otimes d} \right\rangle_{\mathcal{H}^\infty} \\ &= \sum_{d=1}^{\infty} \sqrt{|a_d| \gamma^d} \cdot \operatorname{sgn}(a_d) \sqrt{|a_d| \gamma^d} \cdot \langle u_i^{\otimes d}, v_j^{\otimes d} \rangle_{\mathcal{H}^{\otimes d}} = \sum_{d=1}^{\infty} a_d \gamma^d \langle u_i, v_j \rangle^d = H_{f,g}^{-1}(\gamma \langle u_i, v_j \rangle), \end{aligned}$$

y esto es exactamente lo que queríamos en (2.5).

Si bien esto nos garantiza, con un método alternativo, la existencia de los vectores \tilde{u}_i, \tilde{v}_j , el proceso para hallarlos no es muy práctico algorítmicamente, ya que se encontrarían en el espacio \mathcal{H}^∞ , que es difícil de describir. En realidad es más sencillo simplemente hallar la matriz $(\langle \tilde{u}_i, \tilde{v}_j \rangle)_{ij}$, aplicando $H_{f,g}^{-1}(\gamma-)$ punto a punto como habíamos hecho en (2.7), y luego planteando un problema de programación semidefinida positiva para hallar en tiempo polinomial vectores $\{\tilde{u}_i\}, \{\tilde{v}_j\}$ tales que sus productos cruzados correspondan efectivamente a los valores de la matriz hallada. Por supuesto, el hecho de que hayamos demostrado la existencia de al menos una realización de vectores $\{\tilde{u}_i\}, \{\tilde{v}_j\}$ que cumplen esta condición es la garantía de que el problema será resuelto por esta técnica.

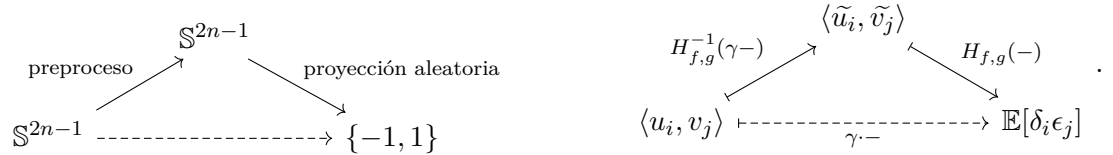
A todo este proceso de pasar de los vectores originales u_i, v_j a los nuevos vectores \tilde{u}_i, \tilde{v}_j lo llamaremos el **preprocesamiento** de los vectores u_i, v_j .

Ahora para llegar a una elección de signos aleatoria δ_i, ϵ_j , solo resta elegir una matriz aleatoria Gaussiana $G \in \mathbb{R}^{k \times m}$, y realizar este redondeo para todos los $i, j \in \{1, \dots, n\}$, siempre con la misma matriz G :

$$\delta_i := f\left(\frac{1}{\sqrt{2}}G\tilde{u}_i\right) \quad \epsilon_j := f\left(\frac{1}{\sqrt{2}}G\tilde{v}_j\right).$$

A esta segunda parte del proceso de elección de signos la denominamos **proyección aleatoria** de los vectores preprocesados.

Veamos que con esta evaluación de signos nos aseguramos que, en esperanza, el valor del problema $OPT(A)$ sea al menos $\frac{SDP(A)}{\gamma}$, como muestra el siguiente diagrama:



Sea $A \in \mathbb{R}^{n \times n}$ una matriz, y sean $\{u_i\}, \{v_j\} \subset \mathbb{S}^{2n-1}$ una evaluación del problema $SDP(A)$ definido como en (1.1.3). Hacemos el preproceso y la proyección aleatoria sobre estos vectores que acabamos de describir para obtener signos δ_i, ϵ_j . Estos signos entonces cumplirán

$$\begin{aligned} \mathbb{E} \left[\sum_{i,j=1}^n a_{ij} \delta_i \epsilon_j \right] &= \sum_{i,j=1}^n a_{ij} \mathbb{E} [\delta_i \epsilon_j] = \sum_{i,j=1}^n a_{ij} \mathbb{E}_G \left[f\left(\frac{1}{\sqrt{2}}G\tilde{u}_i\right) g\left(\frac{1}{\sqrt{2}}G\tilde{v}_j\right) \right] = \\ &= \sum_{i,j=1}^n a_{ij} H_{f,g}(\langle \tilde{u}_i, \tilde{v}_j \rangle) = \sum_{i,j=1}^n a_{ij} H_{f,g}\left(H_{f,g}^{-1}(\gamma\langle u_i, v_j \rangle)\right) = \\ &= \sum_{i,j=1}^n a_{ij} \gamma \langle u_i, v_j \rangle = \gamma \sum_{i,j=1}^n a_{ij} \langle u_i, v_j \rangle. \end{aligned}$$

Esto nos dice que, dada una evaluación del problema $SDP(A)$, tenemos un algoritmo aleatorio polinomial que nos permite hallar una evaluación del problema $OPT(A)$ que alcanza, en esperanza, una fracción γ del valor de la evaluación original.

Partiendo de una matriz A , es posible hallar el valor óptimo del problema $SDP(A)$ en tiempo polinomial mediante un problema de programación semidefinida positiva, como discutiremos en

la próxima sección. Supongamos que partimos de esa solución efectivamente de vectores u_i, v_j que alcanzan este valor máximo. Por lo que acabamos de ver, esto implica que:

$$SDP(A) = \sum_{i,j=1}^n a_{ij} \langle u_i, v_j \rangle = \frac{1}{\gamma} \mathbb{E} \left[\sum_{i,j=1}^n a_{ij} \delta_i \epsilon_j \right] \leq \frac{1}{\gamma} OPT(A).$$

Esto nos da una demostración levemente alternativa de que $K_G \leq \frac{1}{\gamma_{f,g}}$ para cualquier MRK (f, g) , pero nos dice además que la evaluación de signos aleatoria que encontramos es, en esperanza, al menos $\gamma_{f,g} SDP(A) \geq \gamma_{f,g} OPT(A)$. Es decir, tenemos un algoritmo polinomial aleatorio que, en esperanza, queda a una fracción $\gamma_{f,g}$ del valor máximo del problema $OPT(A)$. Este algoritmo es el mejor algoritmo de aproximación del problema $OPT(A)$ que se conoce hasta el momento, aunque no se conoce el mejor valor posible de $\gamma_{f,g}$ y, por lo tanto, no se conocen los Métodos de Redondeo de Krivine que mejor lo explotan.

2.3. Conjeturas y resultados parciales de optimalidad

Ahora que ya pasamos por la interpretación algorítmica de los Métodos de Redondeo de Krivine, volvamos al marco de la forma bilineal de König (2.1.5), que sospechamos que está fuertemente vinculada con el valor óptimo que podemos hallar para $\gamma_{f,g}$. Vamos a hacer algunas observaciones.

Primero, se puede ver inmediatamente desde la fórmula que la forma bilineal de König es estable por rotaciones, debido a que las funciones $e^{-\frac{\|x\|^2 + \|y\|^2}{2}}$ y $\text{sen}(\langle x, y \rangle)$ solo dependen de los valores $\|x\|, \|y\|$ y $\langle x, y \rangle$, que no cambian si los vectores $x, y \in \mathbb{R}^k$ rotan juntos. Esto quiere decir que si $\mathcal{O} \in \mathbb{R}^{k \times k}$ es una matriz ortogonal, entonces:

$$B_K(f, g) = B_K(f(\mathcal{O}-), g(\mathcal{O}-)) \quad \forall f, g \in L_\infty(\mathbb{R}^k). \quad (2.11)$$

Entonces no vamos a esperar que los MRK $(f_{max}^{(k)}, g_{max}^{(k)})$ sean únicos, sino que cualquiera de sus rotaciones sincronizadas es válida como maximal de la forma bilineal de König.

Un resultado muy interesante es que, si la forma bilineal de König tiene al menos un par maximizador $(f_{max}^{(k)}, g_{max}^{(k)})$ que es MRK alternado, entonces resulta que no hay ningún MRK (f, g) , ni siquiera entre los no alternados, que tenga un $\gamma_{f,g}$ mayor a $\gamma_{f_{max}^{(k)}, g_{max}^{(k)}} = \frac{B_K(f_{max}^{(k)}, g_{max}^{(k)})}{(\sqrt{2\pi})^k}$. Notemos que, si la MRK (f, g) no es alternada, la conclusión de que

$$B_K(f_{max}^{(k)}, g_{max}^{(k)}) \geq B_K(f, g) \quad \Rightarrow \quad \gamma_{f_{max}^{(k)}, g_{max}^{(k)}} \geq \gamma_{f,g}$$

no es trivial, ya que a priori no se puede establecer la conexión entre $\gamma_{f,g}$ y $B_K(f, g)$. Vamos entonces a enunciar y probar este hecho:

Lema 2.3.1. *Si para $k \in \mathbb{N}$ la forma bilineal de König de dimensión k se maximiza en un Método de Redondeo de Krivine **alternado** (f_M, g_M) , entonces*

$$\gamma_{f_M, g_M} \geq \gamma_{f,g}, \quad \forall (f, g) \text{ MRK.}$$

Demostración. Supongamos que existiera un Método de Redondeo de Krivine (f, g) tal que $\gamma_{f,g} > \gamma_{f_M, g_M}$. Sea $R > 0$ el radio de convergencia de la función $H_{f,g}^{-1}$. Dado que, por definición, la serie absoluta de $H_{f,g}^{-1}$ converge en $\gamma_{f,g}$ (y de hecho da 1), entonces $\gamma_{f,g}$ tiene que ser menor o igual al radio de convergencia de la serie. Es decir

$$R \geq \gamma_{f,g} > \gamma_{f_M, g_M}.$$

Sea $\alpha := \frac{H_{f,g}(i)}{i} = \frac{B_K(f,g)}{(\sqrt{2\pi})^k}$, y notemos que podría ser negativo. Como $B_K(f_M, g_M) \geq B_K(f, g)$, entonces

$$|\alpha| = \left| \frac{H_{f,g}(i)}{i} \right| = \frac{|B_K(f,g)|}{(\sqrt{2\pi})^k} \leq \frac{B_K(f_M, g_M)}{(\sqrt{2\pi})^k} = \frac{H_{f,g}(i)}{i} = \gamma_{f_M, g_M} < R. \quad (2.12)$$

Si (f, g) fuera MRK alternado ya estaría, porque en ese caso $\gamma_{f,g} = |\alpha| \leq \gamma_{f_M, g_M}$. Pero si no, esto igual nos dice que la serie absoluta de $H_{f,g}^{-1}$ converge en $|\alpha|$. Luego, por definición de α :

$$1 = \frac{H_{f,g}^{-1}(i\alpha)}{i}.$$

Tomando módulo a ambos lados, y usando que vale pasar a la serie absoluta nos queda:

$$1 = |H_{f,g}^{-1}(\alpha)| \leq \left(H_{f,g}^{-1} \right)^{abs} (|\alpha|).$$

Pero como, por definición de $\gamma_{f,g}$, $\left(H_{f,g}^{-1} \right)^{abs} (\gamma_{f,g}) = 1$, entonces

$$\left(H_{f,g}^{-1} \right)^{abs} (\gamma_{f,g}) = 1 \leq \left(H_{f,g}^{-1} \right)^{abs} (|\alpha|).$$

Pero como $\left(H_{f,g}^{-1} \right)^{abs}$ es una serie de potencias de términos no negativos, esto implica que $\gamma_{f,g} \leq |\alpha|$, y luego por (2.12) tenemos $\gamma_{f,g} \leq |\alpha| \leq \gamma_{f_M, g_M}$. ¡Absurdo! \square

En el caso $k = 1$, se puede demostrar que la forma bilineal de König es maximizada por el Método de Redondeo de Krivine alternado $(\text{sgn}(x), \text{sgn}(y))$. Con el resultado anterior. Este hecho fue descubierto primero por König y Tomczak-Jaegermann, y es citado en [Kön01], pero esta demostración nunca fue publicada. La única demostración publicada es la de [BMMN13], que vamos a omitir en este trabajo por ser muy técnica. La idea es básicamente proceder en pasos incrementales, demostrando que las funciones que maximizan B_K en dimensión 1 tienen que ser constantes en un pequeño intervalo. Este intervalo luego se va extendiendo gracias a acotaciones más finas, hasta que finalmente se ve que las funciones maximizadoras tienen que ser constantes en $(0, \infty)$. Este teorema dice básicamente que no tiene sentido considerar otros MRK de dimensión 1, ya que (sgn, sgn) es el mejor de todos en el sentido de tener el γ más grande posible. Sin embargo, queda abierta la pregunta de si aporta algo considerar MRK de dimensión mayor a 1.

König creía que no. Más precisamente, König conjetura en [Kön01], para cada $k \in \mathbb{N}$, las funciones $f_0(x) := \text{sgn}(x_1)$, $g_0(y) := \text{sgn}(y_1)$ (o equivalentemente, la partición del espacio R^k dada por cualquier hiperplano que pase por el 0), son las maximizadoras de la forma bilineal B_K de cada dimensión. Con esta conjetura en mente, König prueba en ese mismo paper uno de los teoremas más motivadores de esta teoría. Ya sabemos que cada MRK (f, g) nos permite demostrar que $K_G \leq \frac{1}{\gamma_{f,g}}$. El teorema de König nos permite ir en el otro sentido:

Teorema 2.3.2. *Si para una sucesión $(k_q)_q \subseteq \mathbb{N}$ la forma bilineal de König de dimensión k_q se maximiza en un Método de Redondeo de Krivine alternado $(f_{\max}^{(k_q)}, g_{\max}^{(k_q)})$ entonces vale que*

$$\frac{1}{\lim_{k \rightarrow \infty} \gamma_{f_{\max}^{(k_q)}, g_{\max}^{(k_q)}}} = K_G.$$

Esto quiere decir que, si existieran estos MRKA de los que habla el teorema, podríamos usarlos para calcular el valor exacto de K_G . En particular, la conjetura de König dice que esa situación ocurre para todas las dimensiones.

Definición 2.3.3. Para cada $k \in \mathbb{N}$, definimos a las funciones $f_0, g_0 : \mathbb{R}^k \rightarrow \{-1, 1\}$ como

$$f_0(x) := \operatorname{sgn}(x_1), g_0(y) := \operatorname{sgn}(y_1).$$

Conjetura 2.3.4. *Para todo $k \in \mathbb{N}$ y cualesquiera $f, g : \mathbb{R}^k \rightarrow \{-1, 1\}$ se cumple que*

$$B_K(f, g) \leq B_K(f_0, g_0).$$

Usando el Lema (1.1.5) es sencillo ver que los Métodos de Redondeo de Krivine f_0, g_0 para todas las dimensiones $k \in \mathbb{N}$ coinciden en los parámetros más importantes:

$$\begin{aligned} H_{f_0, g_0}(z) &= \frac{2}{\pi} \operatorname{arc} \operatorname{sen}(z), \\ H_{f_0, g_0}^{-1}(z) &= \operatorname{sen}\left(\frac{\pi}{2} - \right), \\ \gamma_{f_0, g_0} &= \frac{2 \ln(1 + \sqrt{2})}{\pi}. \end{aligned} \tag{2.13}$$

Notemos que ya habíamos visto esto en el caso $k = 1$, donde $f_0 \equiv g_0 \equiv \operatorname{sgn}$.

Entonces, si la conjetura (2.3.4) fuera cierta, eso implicaría que

$$K_G = \frac{\pi}{2 \ln(1 + \sqrt{2})}.$$

La conjetura de König fue refutada en [BMMN13], como veremos pronto. Sin embargo, el Teorema sigue siendo válido, y podría llegar a tener la clave para hallar el valor exacto de K_G . Vamos a demostrar entonces el teorema de König, y luego hablaremos de cómo refutar su conjetura.

Partimos de la siguiente versión de la desigualdad de Grothendieck:

Desigualdad de Grothendieck 2.3.5. *Sean (Ω, μ) un espacio de medida, \mathcal{H} un espacio de Hilbert, $K \in L_1((\Omega, \mu)^2)$ un núcleo para $\Omega \times \Omega$ y $f, g \in L_\infty((\Omega, \mu); \mathcal{H})$. Si definimos el operador $T_K : L_\infty(\Omega, \mu) \rightarrow L_1(\Omega, \mu)$:*

$$T_K f(y) := \int_{\Omega} K(x, y) f(x) d\mu(x),$$

entonces vale que

$$\int_{\Omega} \int_{\Omega} K(x, y) \langle f(x), g(y) \rangle d\mu(x) d\mu(y) \leq K_G \|T_K\|_{\infty \rightarrow 1} \|f\|_{\infty} \|g\|_{\infty}.$$

Demostración. Esta forma de escribir la desigualdad de Grothendieck no es más que una formulación continua de la forma matricial (1.1.1). Para pasar de esta formulación a la formulación matricial alcanza con tomar $\Omega := \{1, \dots, n\}$ y μ la medida uniforme sobre Ω con densidad $\frac{1}{n}$ en cada punto.

Por otro lado, para llegar desde la formulación matricial a la formulación continua podemos considerar aproximaciones del núcleo K en L_1 con funciones elementales. \square

Pasando de términos:

$$\frac{\int_{\Omega} \int_{\Omega} K(x, y) \langle f(x), g(y) \rangle d\mu(x) d\mu(y)}{\|T_K\|_{\infty \rightarrow 1} \|f\|_{\infty} \|g\|_{\infty}} \leq K_G.$$

La constante K_G se puede definir como la menor constante que realiza esa desigualdad $\forall H, K, f, g$.

En nuestro caso particular, nos alcanza con considerar $\Omega = \mathbb{R}^k$ con la medida de Lebesgue, $\mathcal{H} = \mathbb{R}$, $f(x) = g(x) = \frac{x}{\|x\|}$. La desigualdad se convierte entonces en:

$$\frac{\int_{\mathbb{R}^k} \int_{\mathbb{R}^k} K(x, y) \left\langle \frac{x}{\|x\|}, \frac{y}{\|y\|} \right\rangle dx dy}{\|T_K\|_{\infty \rightarrow 1}} \leq K_G. \quad (2.14)$$

König propuso que el núcleo $K(x, y) := e^{-\frac{\|x\|^2 + \|y\|^2}{2}} \text{sen}(\langle x, y \rangle)$ hace que el valor del lado izquierdo tienda a K_G cuando $k \rightarrow \infty$. Esta conjetura sería cierta en el caso de que existan funciones $f_{max}^{(k)}, g_{max}^{(k)}$ que formen un Método de Redondeo de Krivine Alternado para cada $k \in \mathbb{N}$.

El motivo de König para realizar esta conjetura probablemente es que König creía que, $\forall k \in \mathbb{N}$, las funciones $f_{max}^{(k)}(x) = g_{max}^{(k)}(x) = \text{sgn}(x_1)$ maximizaban la forma bilineal $B_K(f, g)$. La evidencia para creer en esto era que se había demostrado que el par $(\text{sgn}(x), \text{sgn}(y))$ maximizaba B_K para $k = 1$. Como adelantamos, esta demostrada falsa for [BMMN13], cuando hallaron un par de funciones $f, g : \mathbb{R}^2 \rightarrow \{-1, 1\}$ tales que $B_K(f, g) > B_K(\text{sgn}(x_1), \text{sgn}(y_1))$, y por lo tanto, el par $(\text{sgn}(x_1), \text{sgn}(y_1))$ no puede ser maximal para ningún $k \geq 2$.

Sin embargo, la demostración del paper de König sigue funcionando bajo la suposición de que $f_{max}^{(k)}, g_{max}^{(k)}$ forman un Método de Redondeo de Krivine Alternado para infinitos $k \in \mathbb{N}$. Esto viene del hecho de que, en tal caso, con k fijo,

$$\begin{aligned} \|T_K\|_{L_{\infty}(\mathbb{R}^k) \rightarrow L_1(\mathbb{R}^k)} &= \sup_{\|f\|_{\infty}=1} \|T_K f\|_{L_1(\mathbb{R}^k)} = \sup_{\|f\|_{\infty}=1} \sup_{\|g\|_{\infty}=1} \langle T_K f, g \rangle = \\ &= \sup_{\|f\|_{\infty}, \|g\|_{\infty}=1} \int_{\mathbb{R}^k} K(x, y) f(x) g(y) dx \\ &= B_K(f_{max}^{(k)}, g_{max}^{(k)}) = \gamma_{f_{max}^{(k)}, g_{max}^{(k)}} \cdot (\sqrt{2}\pi)^k. \end{aligned}$$

Nos queda entonces la tarea de estimar el numerador de (2.14). Enunciamos entonces el siguiente lema:

Lema 2.3.6.

$$\int_{\mathbb{R}^k} \int_{\mathbb{R}^k} K(x, y) \left\langle \frac{x}{\|x\|}, \frac{y}{\|y\|} \right\rangle dx dy = (\sqrt{2}\pi)^k \cdot r_k, \text{ con } r_k \xrightarrow{k \rightarrow \infty} 1$$

Demostración. Empezamos expresando la integral en coordenadas polares en \mathbb{R}^k . Representamos a $x \in \mathbb{R}^k$ como el par $\{s, u\}$, donde $s = \|x\| \in \mathbb{R}_{\geq 0}$ y $u = \frac{x}{\|x\|} \in \mathbb{S}^{k-1}$. Análogamente, representamos y como el par $\{t, v\}$. El Jacobiano correspondiente a esta transformación es $|J| = (st)^{k-1}$.

$$\begin{aligned} \int_{\mathbb{R}^k} \int_{\mathbb{R}^k} K(x, y) \left\langle \frac{x}{\|x\|}, \frac{y}{\|y\|} \right\rangle dx dy &= \\ &= \int_0^\infty \int_{\mathbb{S}^{k-1}} \int_0^\infty \int_{\mathbb{S}^{k-1}} K(s \cdot u, t \cdot v) \langle u, v \rangle dv dt du ds = \end{aligned}$$

Reordenando un poco las integrales y desarrollando el núcleo K:

$$= \int_0^\infty \int_0^\infty (st)^{k-1} e^{-\frac{s^2+t^2}{2}} \int_{\mathbb{S}^{k-1}} \int_{\mathbb{S}^{k-1}} \text{sen}(st \langle u, v \rangle) \langle u, v \rangle dv du dt ds. \quad (2.15)$$

Vamos a concentrarnos en la integral más interior, dejando u fijo e integrando sobre v :

$$\int_{\mathbb{S}^{k-1}} \text{sen}(st \langle u, v \rangle) \langle u, v \rangle dv.$$

Como v recorre todo \mathbb{S}^{k-1} con densidad uniforme, el valor $\langle u, v \rangle$ tiene la misma distribución para todo $u \in \mathbb{S}^{k-1}$. Sin pérdida de generalidad, suponemos entonces que $u = e_1 \in \mathbb{R}^k$. Luego si \mathbb{S}^{k-2} representa la circunferencia unidad en las últimas $k-1$ coordenadas de \mathbb{R}^k :

$$\begin{aligned} \int_{\mathbb{S}^{k-1}} \text{sen}(st \langle u, v \rangle) \langle u, v \rangle dv &= \int_{-1}^1 \int_{p \cdot e_1 + (\sqrt{1-p^2})^{n-3} \mathbb{S}^{k-2}} \text{sen}(st \langle e_1, (p, \tilde{p}) \rangle) \langle e_1, (p, \tilde{p}) \rangle d\tilde{p} dp \\ &= \int_{-1}^1 \int_{p \cdot e_1 + \sqrt{1-p^2} \mathbb{S}^{k-2}} \text{sen}(st \cdot p) p d\tilde{p} dp \end{aligned}$$

Por ser una traslación, claramente $|p \cdot e_1 + \sqrt{1-p^2} \cdot \mathbb{S}^{k-2}| = |\sqrt{1-p^2} \cdot \mathbb{S}^{k-2}|$:

$$\begin{aligned} &= \int_{-1}^1 \int_{\sqrt{1-p^2} \mathbb{S}^{k-2}} \text{sen}(st \cdot p) p d\tilde{p} dp \\ &= \int_{-1}^1 \left| \sqrt{1-p^2} \mathbb{S}^{k-2} \right| \text{sen}(st \cdot p) p dp \\ &= |\mathbb{S}^{k-2}| \int_{-1}^1 \text{sen}(st \cdot p) p \left(\sqrt{1-p^2} \right)^{k-3} dp \\ &= 2|\mathbb{S}^{k-2}| \int_0^1 \text{sen}(st \cdot p) p \left(\sqrt{1-p^2} \right)^{k-3} dp. \end{aligned}$$

Usando el cambio de variables $p = \cos \theta$,

$$\begin{aligned} \int_{\mathbb{S}^{k-1}} \text{sen}(st \langle u, v \rangle) \langle u, v \rangle dv &= 2|\mathbb{S}^{k-2}| \int_{\frac{\pi}{2}}^0 \text{sen}(st \cos \theta) \cos \theta (\text{sen } \theta)^{k-3} (-\text{sen } \theta) d\theta \\ &= 2|\mathbb{S}^{k-2}| \int_0^{\frac{\pi}{2}} \text{sen}(st \cos \theta) \cos \theta (\text{sen } \theta)^{k-2} d\theta. \end{aligned}$$

Integramos por partes con $u := \text{sen}(st \cos \theta)$, $v' := \cos \theta (\text{sen } \theta)^{k-2}$:

$$= 2|\mathbb{S}^{k-2}| \frac{st}{n-1} \int_0^{\frac{\pi}{2}} \cos(st \cos \theta) (\text{sen } \theta)^k d\theta$$

Esta integral la sacamos de la tabla (8.411/4) de [Zwi14], y usamos que $\Gamma(\frac{1}{2}) = \sqrt{\pi}$:

$$\int_{\mathbb{S}^{k-1}} \text{sen}(st \langle u, v \rangle) \langle u, v \rangle dv = |\mathbb{S}^{k-2}| \sqrt{\pi} \frac{st}{k-1} \Gamma\left(\frac{k+1}{2}\right) \left(\frac{2}{st}\right)^{k/2} J_{\frac{k}{2}}(st),$$

donde $J_{\frac{k}{2}}$ es la función de Bessel de orden $\frac{k}{2}$.

Habíamos visto que este integrando no dependía de qué valor explícito $u \in \mathbb{S}^{k-1}$ tomaba. Luego, al integrarlo con u recorriendo uniformemente ese dominio, sólo se multiplica el valor por la medida de \mathbb{S}^{k-1} .

Resulta entonces que

$$(2.15) = \int_0^\infty \int_0^\infty (st)^{k-1} e^{-\frac{s^2+t^2}{2}} |\mathbb{S}^{k-1}| |\mathbb{S}^{k-2}| \sqrt{\pi} \frac{st}{k-1} \Gamma\left(\frac{k+1}{2}\right) \left(\frac{2}{st}\right)^{k/2} J_{\frac{k}{2}}(st) dt ds$$

Reordenando y usando que $|\mathbb{S}^{k-1}| |\mathbb{S}^{k-2}| \Gamma(\frac{k+1}{2}) = \frac{2\pi^{k-\frac{1}{2}}}{\Gamma(\frac{k}{2})}$:

$$= \frac{2^{\frac{k}{2}+1} \pi^k}{\Gamma\left(\frac{k}{2}\right)} \int_0^\infty \int_0^\infty (st)^{\frac{k}{2}} e^{-\frac{s^2+t^2}{2}} J_{\frac{k}{2}}(st) dt ds. \quad (2.16)$$

Llegado este paso, la integral se puede calcular usando las propiedades (6.631/1), (7.525/1) y (9.131/1) de [Zwi14], que son propiedades sobre integrales con funciones hipergeométricas ${}_1F_1$ y ${}_2F_1$. Usando estas propiedades se obtiene que

$$(2.16) = 2^{\frac{k}{2}} \pi^k \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right) \Gamma\left(\frac{k}{2} + 1\right)} {}_2F_1\left(\frac{1}{2}, \frac{1}{2}; \frac{k}{2} + 1; -1\right)$$

Como se cumple que

$$(2.15) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right) \Gamma\left(\frac{k}{2} + 1\right)} {}_2F_1\left(\frac{1}{2}, \frac{1}{2}; \frac{k}{2} + 1; -1\right) \xrightarrow{k \rightarrow \infty} 1,$$

la demostración está terminada. □

Ahora sí, vamos a terminar la demostración de que si infinitas particiones maximales son Métodos de Redondeo de Krivine Alternados entonces $K_G = \lim_{q \rightarrow \infty} \frac{1}{\gamma_{k_q}}$.

Demostración del Teorema (2.3.2). Para cada $k_q \in \mathbb{N}$, notamos $\gamma_{k_q} := \gamma_{f_{max}^{(k_q)}, g_{max}^{(k_q)}}$. Usando el Lema (2.3.6) en (2.14), tenemos que:

$$\begin{aligned} \frac{\int_{\mathbb{R}^{k_q}} \int_{\mathbb{R}^{k_q}} K(x, y) \langle \frac{x}{\|x\|}, \frac{y}{\|y\|} \rangle dx dy}{\|TK\|_{\infty \rightarrow 1}} &\leq K_G \\ \frac{(\sqrt{2}\pi)^{k_q} \cdot r_{k_q}}{(\sqrt{2}\pi)^{k_q} \cdot \gamma(f_{max}^{(k_q)}, g_{max}^{(k_q)})} &\leq K_G \\ \frac{r_{k_q}}{\gamma_{k_q}} &\leq K_G \end{aligned}$$

Y haciendo $q \rightarrow \infty$:

$$\frac{1}{\lim_{q \rightarrow \infty} \gamma_{k_q}} \leq K_G.$$

Como estamos bajo el supuesto de que cada par $(f_{max}^{(k_q)}, g_{max}^{(k_q)})$ es un Método de Redondeo de Krivine Alternado, resulta que cada uno de estos pares da una cota superior para K_G . Luego:

$$\begin{aligned} \frac{1}{\lim_{q \rightarrow \infty} \gamma_{f_{max}^{(k_q)}, g_{max}^{(k_q)}}} &\leq K_G \leq \frac{1}{\gamma_{f_{max}^{(k_q)}, g_{max}^{(k_q)}}} \quad \forall q \in \mathbb{N} \\ \frac{1}{\lim_{q \rightarrow \infty} \gamma_{f_{max}^{(k_q)}, g_{max}^{(k_q)}}} &\leq K_G \leq \frac{1}{\lim_{q \rightarrow \infty} \gamma_{f_{max}^{(k_q)}, g_{max}^{(k_q)}}}. \end{aligned}$$

Y por lo tanto

$$K_G = \frac{1}{\lim_{q \rightarrow \infty} \gamma_{f_{max}^{(k_q)}, g_{max}^{(k_q)}}}.$$

□

El Teorema de König nos alienta a encontrar las particiones maximales para la forma bilineal de König, y habíamos mencionado que König conjeturó que en todas las dimensiones las particiones por hiperplanos son maximales. Vamos a ver ahora un primer argumento por el cuál esto parece falso.

Supongamos que tenemos una partición maximal $f_{max}^{(k)}, g_{max}^{(k)} : \mathbb{R}^k \rightarrow \{-1, 1\}$. Habíamos visto en (2.1.7) que cualesquiera funciones que maximicen B_K tienen que tener la forma

$$f_{max}^{(k)}(x) = \operatorname{sgn} \left(\int_{\mathbb{R}^k} g_{max}^{(k)}(y) e^{-\frac{\|y\|^2}{2}} \operatorname{sen}(\langle x, y \rangle) dy \right), \quad g_{max}^{(k)}(y) = \operatorname{sgn} \left(\int_{\mathbb{R}^k} f_{max}^{(k)}(x) e^{-\frac{\|x\|^2}{2}} \operatorname{sen}(\langle x, y \rangle) dx \right).$$

Es decir que las funciones $f_{max}^{(k)}$ y $g_{max}^{(k)}$ tienen que ser, de alguna forma, “conjugadas” para un operador $\sigma : L_\infty(\mathbb{R}^k) \rightarrow L_\infty(\mathbb{R}^k)$ que podemos definir como

$$\sigma(f)(x) := \operatorname{sgn} \left(\int_{\mathbb{R}^k} g_{max}^{(k)}(y) e^{-\frac{\|y\|^2}{2}} \operatorname{sen}(\langle x, y \rangle) dy \right).$$

Con esta definición, la propiedad (2.3) se expresa simplemente como

$$\sigma(f_{max}^{(k)}) = g_{max}^{(k)}, \quad \sigma(g_{max}^{(k)}) = f_{max}^{(k)},$$

y de ahí se deduce inmediatamente que $f_{max}^{(k)}$ y $g_{max}^{(k)}$ tienen que ser puntos fijos del operador σ^2 .

Por un lado, las funciones f_0, g_0 conjeturadas por König cumplen efectivamente con $\sigma(f_0) = g_0, \sigma(g_0) = f_0$, y esto implica, como dijimos, que ambas son puntos fijos de σ^2 . ¿Pero qué sucede si le aplicamos el operador σ^2 a otras funciones?

Para $k = 2$, según [BMMN13], resulta que comenzando desde muchas otras funciones y aplicando numéricamente el operador σ^2 las funciones parecen “converger” siempre a alguna de dos particiones específicas (excepto por alguna rotación, que no importa, como explicamos en (2.11)). Estas dos particiones parecen ser, naturalmente, “conjugadas” en el sentido de σ .

A modo de ilustración, esta es una de las dos particiones, ilustrada entre $[-20, 20] \times [-20, 20]$, que los autores llaman la Partición del Tigre. La partición “conjugada” tiene un dibujo estéticamente similar.

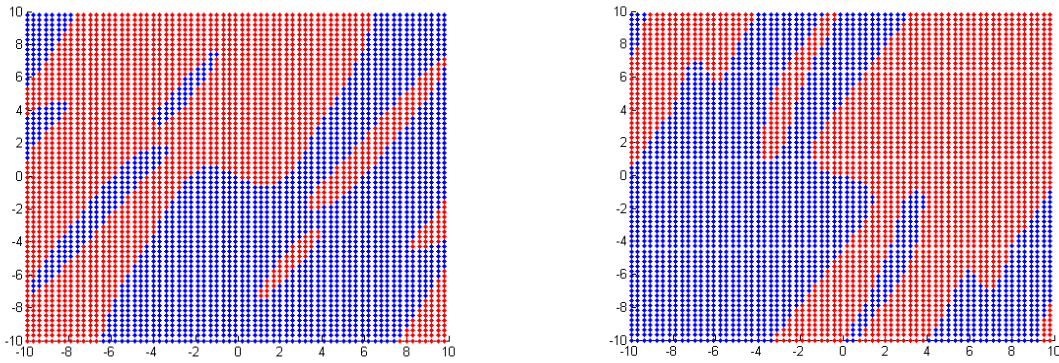


Casi todo se desconoce sobre esta Partición del Tigre y su conjugada. Estas son algunas de las preguntas que consideramos más interesantes:

- ¿Hay alguna fórmula cerrada que defina a la Partición del Tigre y/o a su conjugada?
- ¿Las particiones de \mathbb{R}^2 convergen en algún sentido formal a la Partición del Tigre?

- ¿La Partición del Tigre y su conjugada forman un Método de Redondeo de Krivine?
- En caso afirmativo, ¿es este un Método de Redondeo de Krivine Alternado?
- ¿Cuál es la respuesta a todas estas preguntas en dimensiones más altas?

Con un algoritmo muy básico en Matlab, logramos reproducir parte del experimento de [BMMN13]. Incluso con una resolución de gráfico mucho menor y con todos los errores numéricos que esto implica, en nuestras funciones finales también llega a verse claramente una curva principal como la que domina el dibujo anterior. Las siguiente es la partición que obtuvimos y su conjugada (recordar que las rotaciones son irrelevantes por (2.11)).



Veremos en la próxima sección que (f_0, g_0) no son particiones maximales para $k = 2$, y de esta manera refutaremos la conjetura de König. Con esto en mente, y viendo que parece muy creíble que la partición del Tigre y su conjugada sean los únicos otros puntos fijos de σ^2 para $k = 2$, pareciera que la Partición del Tigre y su conjugada sí maximizan B_K . Lamentablemente no podemos probarlo, y la falta de una fórmula para estas particiones imposibilitan trabajar con ellas para acotar K_G . Por estos motivos, en la próxima sección, la demostración de que (f_0, g_0) no es óptima en $k = 2$ la realizaremos con una partición que sabemos subóptima (no es punto fijo de σ^2), pero con la que sí podemos hacer algunas cuentas. Incluso con esta partición explícita, no se conoce una demostración de que el Método de Redondeo de Krivine que induce sea alternado, aunque se sospecha que sí lo es. Esta incerteza nos obligará a tomar un desvío para acotar γ en este caso particular, que ocupará gran parte de la próxima sección, y nos llevará a usar una versión un poco más relajada de los MRK.

2.4. La cota de Krivine no es ajustada

El objetivo de esta sección es probar el resultado principal de [BMMN13], es decir, que la constante de Grothendieck es estrictamente menor que el valor $\frac{\pi}{2\ln(1+\sqrt{2})}$ conjeturado por Krivine. Como la conjetura de König (2.3.4) implicaba la conjetura de Krivine, este resultado prueba también que la conjetura de König es falsa. De hecho, la demostración se basará en describir una familia de MRK (f_η, g_η) para algún η en un entorno positivo del 0 apropiado, tales que cualquiera de estos pares de funciones cumple que $B_K(f_\eta, g_\eta)$. Esto ya refuta la conjetura de König. Si bien

se sospecha que los MRK de esta familia son alternados, no se ha logrado demostrarlo. Por este motivo, se usará una técnica alternativa para llegar desde el contraejemplo de König a la demostración de que $K_G < \frac{\pi}{2 \ln(1+\sqrt{2})}$. Este método consistirá en combinar adecuadamente los MRK dados por (f_η, g_η) y (f_0, g_0) de forma que se conserven las propiedades buenas de cada uno. Más precisamente, el primer MRK nos da una evaluación de $B_K(f_\eta, g_\eta)$ mayor a las que teníamos, mientras que el segundo MRK, al ser alternado, permite trasladar el valor de $B_K(f_0, g_0)$ al γ del método.

Vamos a empezar entonces, por la familia de funciones que nos da el contraejemplo a la conjetura de König (2.3.4).

Teorema 2.4.1. *En $k = 2$, existen $f_K, g_K : \mathbb{R}^2 \rightarrow \{-1, 1\}$ tales que:*

$$B_K(f_K, g_K) > B_K(f_0, g_0) = \ln(1 + \sqrt{2})$$

La idea de los autores fue usar un par de particiones que pudieran ser definidas a partir de dos funciones impares $\alpha, \beta : \mathbb{R} \rightarrow \mathbb{R}$ y una constante $\eta \in \mathbb{R}$ de la siguiente manera:

$$f_\eta(x_1, x_2) := \begin{cases} 1 & x_2 \geq \eta\alpha(x_1), \\ -1 & x_2 < \eta\alpha(x_1), \end{cases} \quad y \quad g_\eta(x_1, x_2) := \begin{cases} 1 & x_2 \geq \eta\beta(x_1), \\ -1 & x_2 < \eta\beta(x_1), \end{cases} \quad (2.17)$$

Notemos que con esta definición, f_0 y g_0 corresponden efectivamente a $\text{sgn}(x_1)$ y $\text{sgn}(y_1)$, y que el hecho de que α y β sean impares implica que también f_0 y g_0 lo son.

Definamos entonces $\phi(\eta) := B_K(f_\eta, g_\eta)$. La idea va a ser ver si, para alguna elección apropiada de α y β , la función ϕ no es máximo local en $\eta = 0$. Esto probaría que f_0 y g_0 no son máximos absolutos de la forma bilineal de König.

Para hacer esto, vamos a mirar las derivadas de ϕ en $\eta = 0$. Como $f_{-\eta} = -f_\eta$ y $g_{-\eta} = -g_\eta$, entonces ϕ_η es par. Luego $\phi'(0) = 0$, y de hecho todas las derivadas impares de ϕ dan 0 en $\eta = 0$. Como esto no contradice que $\eta = 0$ sea máximo local, tenemos que mirar al menos la derivada de segundo orden, que nos gustaría que fuera positiva. Y a esta altura ya es necesario hacer algunas cuentas para avanzar. Para empezar, escribamos explícitamente ϕ :

$$\phi(\eta) = \int_{\mathbb{R}} \int_{\mathbb{R}} \int_{\mathbb{R}} \int_{\mathbb{R}} f_\eta(x) g_\eta(y) e^{-\frac{x_1^2 + x_2^2 + y_1^2 + y_2^2}{2}} \text{sen}(x_1 y_1 + x_2 y_2) dx_2 dy_2 dx_1 dy_1. \quad (2.18)$$

Miremos un poco la integral sobre y_1 e y_2 , sacando constantes:

$$\int_{\mathbb{R}^2} g_\eta(y) e^{-\frac{y_2^2}{2}} \text{sen}(x_1 y_1 + x_2 y_2) dx_2 = \int_{\mathbb{R}} g_\eta(y) e^{-\frac{y_2^2}{2}} \text{sen}(x_1 y_1 + x_2 y_2) dy. \quad (2.19)$$

Como $\text{sen}(\langle x, - \rangle)$ y g_η son impares, y la exponencial $e^{-\frac{y_2^2}{2}}$ es par, resulta que el integrando da lo mismo en (y_1, y_2) que en $(-y_1, -y_2)$ para cualquier $y_1, y_2 \in \mathbb{R}$. Como

$$y_2 > \eta\beta(y_1) \Leftrightarrow -y_2 < \eta\beta(-y_1),$$

podemos pensar a la integral (2.19) como dos veces la integral sobre los pares (y_1, y_2) tales que $y_2 > \eta\beta(y_1)$. Es decir que

$$\begin{aligned} \int_{\mathbb{R}^2} g_\eta(y) e^{-\frac{y_2^2}{2}} \operatorname{sen}(x_1 y_1 + x_2 y_2) dy_1 dy_2 &= 2 \int_{\mathbb{R}} \int_{\eta\beta(y_1)}^{\infty} g_\eta(y) e^{-\frac{y_2^2}{2}} \operatorname{sen}(x_1 y_1 + x_2 y_2) dy_1 dy_2 = \\ &= 2 \int_{\mathbb{R}} \int_{\eta\beta(y_1)}^{\infty} e^{-\frac{y_2^2}{2}} \operatorname{sen}(x_1 y_1 + x_2 y_2) dy_1 dy_2, \end{aligned}$$

ya que $g_\eta \equiv 1$ en el dominio $\{y_2 > \eta\beta(y_1)\}$.

Haciendo lo mismo para x , podemos expresar ϕ de la siguiente manera:

$$\phi(\eta) = 4 \int_{\mathbb{R}} \int_{\mathbb{R}} \int_{\eta\alpha(x_1)}^{\infty} \int_{\eta\beta(y_1)}^{\infty} e^{-\frac{x_1^2 + y_2^2 + y_1^2 + y_2^2}{2}} \operatorname{sen}(x_1 y_1 + x_2 y_2) dx_2 dy_2 dx_1 dy_1.$$

Para derivar esto en η , vamos a usar la función auxiliar $F : \mathbb{R}^2 \rightarrow \mathbb{R}$:

$$F(a, b) = 4 \int_{\mathbb{R}} \int_{\mathbb{R}} \int_{a\alpha(x_1)}^{\infty} \int_{b\beta(y_1)}^{\infty} e^{-\frac{x_1^2 + x_2^2 + y_1^2 + y_2^2}{2}} \operatorname{sen}(x_1 y_1 + x_2 y_2) dx_2 dy_2 dx_1 dy_1.$$

Esto nos permite expresar, usando $\operatorname{diag} : \mathbb{R} \rightarrow \mathbb{R}^2$, $\operatorname{diag}(\eta) := (\eta, \eta)$,

$$\phi(\eta) = F \circ \operatorname{diag}(\eta).$$

Veamos las derivadas:

$$\begin{aligned} \phi(\eta) &= (F \circ \operatorname{diag})(\eta) \\ \phi'(\eta) &= \nabla F(\eta, \eta) \cdot D(\operatorname{diag})(\eta) = \nabla F(\eta, \eta) \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \left(\frac{\partial}{\partial a} + \frac{\partial}{\partial b} \right) F(\eta, \eta) \\ \phi''(\eta) &= \left(\frac{\partial^2}{\partial a^2} + \frac{\partial^2}{\partial a \partial b} + \frac{\partial^2}{\partial b \partial a} + \frac{\partial^2}{\partial b^2} \right) F(\eta, \eta). \end{aligned} \tag{2.20}$$

Vamos a calcular estas derivadas parciales de F :

$$\begin{aligned} \frac{\partial}{\partial a} F(a, b) &= -4 \int_{\mathbb{R}} \int_{\mathbb{R}} \int_{b\beta(y_1)}^{\infty} \alpha(x_1) e^{-\frac{x_1^2 + (a\alpha(x_1))^2 + y_1^2 + y_2^2}{2}} \operatorname{sen}(x_1 y_1 + a\alpha(x_1) y_2) dy_2 dx_1 dy_1 \\ \frac{\partial}{\partial b} F(a, b) &= -4 \int_{\mathbb{R}} \int_{\mathbb{R}} \int_{a\alpha(x_1)}^{\infty} \beta(y_1) e^{-\frac{x_1^2 + x_2^2 + y_1^2 + (b\beta(y_1))^2}{2}} \operatorname{sen}(x_1 y_1 + x_2 b\beta(y_1)) dy_2 dx_1 dy_1. \end{aligned}$$

Para las derivadas segundas, podemos ahorrar tiempo y espacio eliminando términos enteros,

por saber que vamos a evaluarla en $(0, 0)$:

$$\begin{aligned} \frac{\partial^2}{\partial a^2} F(0, 0) &= -4 \int_{\mathbb{R}} \int_{\mathbb{R}} \int_0^{\infty} \alpha(x_1)^2 y_2 e^{-\frac{x_1^2 + y_1^2 + y_2^2}{2}} \cos(x_1 y_1) dy_2 dx_1 dy_1 \\ &= -4 \int_{\mathbb{R}} \int_{\mathbb{R}} \alpha(x_1)^2 e^{-\frac{x_1^2 + y_1^2}{2}} \cos(x_1 y_1) dx_1 dy_1 \end{aligned} \quad (2.21)$$

$$\begin{aligned} \frac{\partial^2}{\partial b^2} F(0, 0) &= -4 \int_{\mathbb{R}} \int_{\mathbb{R}} \int_0^{\infty} x_2 \beta(y_1)^2 e^{-\frac{x_1^2 + x_2^2 + y_1^2}{2}} \cos(x_1 y_1) dx_2 dx_1 dy_1 \\ &= -4 \int_{\mathbb{R}} \int_{\mathbb{R}} \beta(y_1)^2 e^{-\frac{x_1^2 + y_1^2}{2}} \cos(x_1 y_1) dx_1 dy_1 \end{aligned} \quad (2.22)$$

$$\frac{\partial^2}{\partial a \partial b} F(a, b) = 4 \int_{\mathbb{R}} \int_{\mathbb{R}} \alpha(x_1) \beta(y_1) e^{-\frac{x_1^2 + y_1^2}{2}} \operatorname{sen}(x_1 y_1) dx_1 dy_1, \quad (2.23)$$

donde en (2.21) y (2.23) usamos que

$$\int_0^{\infty} x e^{-\frac{x^2}{2}} dx = 1.$$

El tipo de integrales que se le están aplicando a las funciones α y β nos sugiere tratar de usar polinomios de Hermite, que se llevan bien con las funciones exponenciales.

Definición 2.4.2. Para cada $n \in \mathbb{N}$ se define el polinomio de Hermite de grado n como

$$H_n(x) = (-1)^n e^{x^2} \frac{\partial^n}{\partial x^n} e^{-x^2}.$$

Los polinomios de Hermite cumplen que son una conjunto ortonormal en $L_2(\mathcal{G})$, el espacio de Hilbert con el producto interno dado por

$$\langle f, g \rangle_{L_2(\mathcal{G})} := \int_{\mathbb{R}} f(x) g(x) e^{-x^2} dx. \quad (2.24)$$

Vamos a usar las siguientes propiedades de los polinomios de Hermite, ambas consecuencia de que $e^{-\frac{x^2}{2}} H_n(x)$ es autovector de autovalor i^n de la Transformada de Fourier (ver 6.1.15 en [AAR99]):

$$\int_{\mathbb{R}} H_{2n}(y) e^{-\frac{y^2}{2}} \cos(xy) dy = (-1)^n \sqrt{2\pi} e^{-\frac{x^2}{2}} H_{2n}(x) \quad \forall n \in \mathbb{N}_0 \quad (2.25)$$

$$\int_{\mathbb{R}} H_{2n+1}(y) e^{-\frac{y^2}{2}} \operatorname{sen}(xy) dy = (-1)^n \sqrt{2\pi} e^{-\frac{x^2}{2}} H_{2n+1}(x) \quad \forall n \in \mathbb{N}_0. \quad (2.26)$$

Vamos a expresar entonces α y β como series de polinomios de Hermite. Para que α y β sean impares en el desarrollo como serie sólo incluimos los polinomios de grado impar.

$$\alpha(x) = \sum_{k=0}^{\infty} \alpha_k H_{2k+1}(x) \quad \text{y} \quad \beta(y) = \sum_{k=0}^{\infty} \beta_k H_{2k+1}(y). \quad (2.27)$$

Volvamos a las derivadas parciales de F :

$$\begin{aligned} \frac{\partial^2}{\partial a^2} F(0,0) &\stackrel{(2.21)}{=} -4 \int_{\mathbb{R}} \int_{\mathbb{R}} \alpha(x_1)^2 e^{-\frac{x_1^2+y_1^2}{2}} \cos(x_1 y_1) dx_1 dy_1 = \\ &= -4 \int_{\mathbb{R}} \int_{\mathbb{R}} \alpha(x_1)^2 e^{-\frac{x_1^2}{2}} \left(e^{-\frac{y_1^2}{2}} \cos(x_1 y_1) dy_1 \right) dx_1 = \\ &\stackrel{(2.25)}{=} -4 \int_{\mathbb{R}} \int_{\mathbb{R}} \alpha(x_1)^2 e^{-\frac{x_1^2}{2}} \sqrt{2\pi} e^{-\frac{x_1^2}{2}} dx_1 = \\ &= -4\sqrt{2\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} \alpha(x_1)^2 e^{-x_1^2} dx_1 = -4\sqrt{2\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} \left(\sum_{k=0}^{\infty} \alpha_k H_{2k+1}(x_1) \right)^2 e^{-x_1^2} dx_1 = \\ &= -4\sqrt{2\pi} \sum_{k_1, k_2=0}^{\infty} \int_{\mathbb{R}} \int_{\mathbb{R}} \alpha_{k_1} \alpha_{k_2} H_{2k_1+1}(x_1) H_{2k_2+1}(x_1) e^{-x_1^2} dx_1 \\ &= -4\sqrt{2\pi} \sum_{k=0}^{\infty} \alpha_k^2. \end{aligned} \quad (2.28)$$

Análogamente, tenemos

$$\frac{\partial^2}{\partial b^2} F(0,0) = -4\sqrt{2\pi} \sum_{k=0}^{\infty} \beta_k^2. \quad (2.29)$$

Por último, miramos la derivada cruzada:

$$\begin{aligned}
\frac{\partial^2}{\partial a \partial b} F(a, b) &\stackrel{(2.23)}{=} 4 \int_{\mathbb{R}} \int_{\mathbb{R}} \alpha(x_1) \beta(y_1) e^{-\frac{x_1^2 + y_1^2}{2}} \operatorname{sen}(x_1 y_1) dx_1 dy_1 = \\
&= 4 \int_{\mathbb{R}} \alpha(x_1) e^{-\frac{x_1^2}{2}} \left(\int_{\mathbb{R}} \beta(y_1) e^{-\frac{y_1^2}{2}} \operatorname{sen}(x_1 y_1) dy_1 \right) dx_1 = \\
&= 4 \int_{\mathbb{R}} \alpha(x_1) e^{-\frac{x_1^2}{2}} \left(\int_{\mathbb{R}} \sum_{k=0}^{\infty} \beta_k H_{2k+1}(y_1) e^{-\frac{y_1^2}{2}} \operatorname{sen}(x_1 y_1) dy_1 \right) dx_1 = \\
&= 4 \int_{\mathbb{R}} \alpha(x_1) e^{-\frac{x_1^2}{2}} \left(\sum_{k=0}^{\infty} \beta_k \int_{\mathbb{R}} H_{2k+1}(y_1) e^{-\frac{y_1^2}{2}} \operatorname{sen}(x_1 y_1) dy_1 \right) dx_1 = \\
&\stackrel{(2.26)}{=} 4 \int_{\mathbb{R}} \alpha(x_1) e^{-\frac{x_1^2}{2}} \left(\sum_{k=0}^{\infty} \beta_k (-1)^k \sqrt{2\pi} H_{2k+1}(x_1) e^{-\frac{x_1^2}{2}} \right) dx_1 = \\
&= 4 \int_{\mathbb{R}} \left(\sum_{k_1=0}^{\infty} \alpha_{k_1} H_{2k_1+1}(x_1) \right) e^{-\frac{x_1^2}{2}} \left(\sum_{k_2=0}^{\infty} \beta_{k_2} (-1)^{k_2} \sqrt{2\pi} H_{2k_2+1}(x_1) e^{-\frac{x_1^2}{2}} \right) dx_1 = \\
&= 4\sqrt{2\pi} \sum_{k_1, k_2=0}^{\infty} \int_{\mathbb{R}} (-1)^k \alpha_{k_1} \beta_{k_2} H_{2k_1+1}(x_1) H_{2k_2+1}(x_1) e^{-x_1^2} dx_1 = \\
&= 4\sqrt{2\pi} \sum_{k=0}^{\infty} (-1)^k \alpha_k \beta_k. \tag{2.30}
\end{aligned}$$

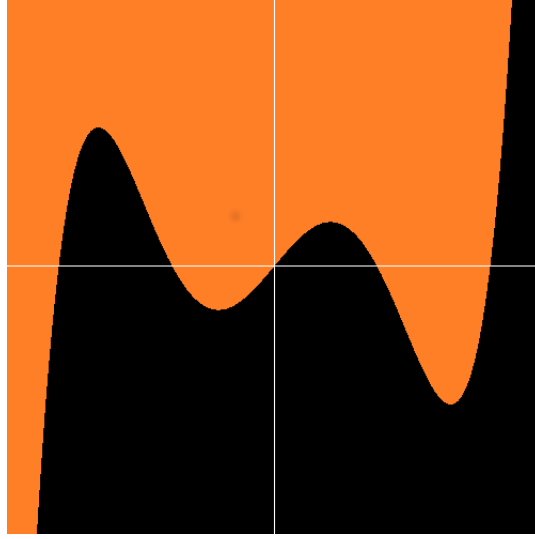
Insertando lo hallado en (2.28), (2.30) y (2.29) en (2.20) obtenemos:

$$\begin{aligned}
\phi''(0) &= -4\sqrt{2\pi} \sum_{k=0}^{\infty} \alpha_k^2 + 4\sqrt{2\pi} \sum_{k=0}^{\infty} (-1)^k \alpha_k \beta_k + 4\sqrt{2\pi} \sum_{k=0}^{\infty} (-1)^k \alpha_k \beta_k - 4\sqrt{2\pi} \sum_{k=0}^{\infty} \beta_k^2 \\
&= -4\sqrt{2\pi} \sum_{k=0}^{\infty} (\alpha_k - (-1)^k \beta_k)^2 \leq 0 \tag{2.31}
\end{aligned}$$

Esto implica que, para tener alguna oportunidad de que ϕ tenga un máximo local en $\eta = 0$, es necesario que al menos $\phi''(0) = 0$, para tener la esperanza de que las derivadas de órdenes mayores tengan el signo buscado. Esto implica que $\alpha_k = (-1)^k \beta_k \quad \forall k \in \mathbb{N}_0$

Dentro de estas restricciones, los casos más simples son $\alpha = \beta = h_1$, $\alpha = -\beta = h_3$ y $\alpha = \beta = h_5$. El primer caso no funciona, porque nos devuelve siempre $f_\eta = g_\eta = f_0 = g_0$, entonces ϕ es constante. Los autores de [BMMN13] eligen el tercer caso sobre el segundo por tener la simplicidad extra de que $\alpha \equiv \beta$, y con estas funciones logran probar que $\phi''''(0) > 0$, y el teorema (2.4.1).

Ilustramos la partición $f_1 \equiv g_1$ dada por h_5 .



Las particiones $f_\eta \equiv g_\eta$ cuando $\eta \in (0, 1)$ son distorsiones de esta partición, así que está claro que ninguna de estas particiones se parecerá a la Partición del Tigre que describimos antes. De hecho, ninguna partición que sea del tipo (2.17), podría coincidir con la Partición del Tigre, ya que ésta no corresponde a separar el plano por la imagen de la curva de una función. Esto sugiere que un análisis más ajustado de la constante de Grothendieck requeriría algunos métodos nuevos. Sin embargo, las particiones correspondientes a h_5 sí nos alcanzarán para demostrar que la cota superior de Krivine no es óptima.

Lema 2.4.3. *Si para cada $\eta \in \mathbb{R}_{>0}$ definimos las funciones $f_\eta, g_\eta : \mathbb{R}^2 \rightarrow \mathbb{R}$ de la siguiente manera:*

$$f_\eta(x_1, x_2) := \begin{cases} 1 & x_2 \geq \eta H_5(x_1), \\ -1 & x_2 < \eta H_5(x_1), \end{cases} \quad y \quad g_\eta(x_1, x_2) := \begin{cases} 1 & x_2 \geq \eta H_5(x_1), \\ -1 & x_2 < \eta H_5(x_1), \end{cases},$$

entonces existe η_0 tal que

$$B_K(f_\eta, g_\eta) > B_K(f_0, g_0) \quad \forall \eta \in (0, \eta_0).$$

La demostración es, como dijimos, una vez visto que $\phi''(0) = 0$, demostrar que $\phi''''(0) > 0$ para este caso particular, con cuentas similares a las que venimos haciendo. Dado que la idea principal y la motivación del uso de estas funciones está clara, vamos a omitir las cuentas que demuestran que $\phi''''(0) > 0$, ya que creemos que no agregan demasiado a este trabajo.

Como corolario, sabemos que existen al menos un par de funciones medibles impares $f_K, g_K : \mathbb{R}^2 \rightarrow \mathbb{R}$ tales que $B_K(f_K, g_K) > B_K(f_0, g_0)$, ya que basta tomar $f_K := f_\eta, g_K := g_\eta$, para cualquier $\eta \in (0, \eta_0)$ como en (2.4.3).

Vamos a pasar entonces a demostrar usando estas funciones, que $K_G < \frac{\pi}{2 \ln(1+\sqrt{2})}$. Si f_K y g_K formaran una Método de Redondeo de Krivine Alternado de dimensión 2 como en (2.1.3), entonces el resultado sería inmediato. Por (2.3) y (2.13) tendríamos que

$$\gamma_{f_K, g_K} = \frac{H_{f_K, g_K}(i)}{i} = \frac{B_K(f_K, g_K)}{(\sqrt{2}\pi)^2} > \frac{B_K(f_0, g_0)}{(\sqrt{2}\pi)^2} = \gamma_{f_0, g_0} = \frac{2 \ln(1 + \sqrt{2})}{\pi}.$$

Luego, por lo visto en este capítulo, la existencia de un Método de Redondeo de Krivine (f_K, g_K) implica que

$$K_G \leq \frac{1}{\gamma_{f_K, g_K}} < \frac{\pi}{2 \ln(1 + \sqrt{2})}.$$

En [BMMN13], los autores conjeturan que las f_K y g_K que hallamos en son un Método de Redondeo de Krivine Alternado de dimensión 2, pero no pueden probarlo. Por eso es necesaria esta demostración alternativa que construye un nuevo método de redondeo haciendo una combinación lineal con el MRK original (f_0, g_0) .

A grandes rasgos, (f_0, g_0) es un Método de Redondeo de Krivine Alternado, pero su valor de $\frac{H_{f_0, g_0}(i)}{i} \leq \frac{2 \ln(1 + \sqrt{2})}{\pi}$ (de hecho, vale la igualdad por (2.13)). Por otro lado, tenemos el par (f_K, g_K) que no es, hasta donde sabemos, un Método de Redondeo de Krivine Alternado, pero que sí cumple $\frac{H_{f_K, g_K}(i)}{i} > \frac{2 \ln(1 + \sqrt{2})}{\pi}$. La idea es entonces hacer un nuevo método de redondeo a partir de estos dos pares de funciones, de forma tal que este método conserve la propiedad buena que queremos de cada ingrediente. Más adelante veremos un resultado sobre métodos de redondeo similares a estos, que en efecto generalizan a los Métodos de Redondeo de Krivine como los conocíamos hasta el momento.

Teorema 2.4.4. *Sean $f_K, g_K : \mathbb{R}^2 \rightarrow \{-1, 1\}$ dos funciones medibles impares, tales que $\frac{B_K(f_K, g_K)}{(\sqrt{2}\pi)^2} > \frac{2 \ln(1 + \sqrt{2})}{2}$. Entonces $K_G < \frac{2}{2 \ln(1 + \sqrt{2})}$. Además, existe un método de redondeo aleatorio tal que las evaluaciones de signos que encuentra para cualquier problema $OPT(A)$ son, en esperanza, estrictamente mayores a $\frac{2 \ln(1 + \sqrt{2})}{2} SDP(A)$.*

Para probar esto, vamos a construir el método de redondeo que prometimos, basándonos en la interpretación algorítmica de los Métodos de Redondeo de Krivine que vimos anteriormente. Empezamos definiendo cómo va a funcionar el método de “proyección aleatoria”.

Dado $p \in (0, 1)$, vamos a definir un método de proyección aleatoria μ_p que usa el par (f_0, g_0) con probabilidad $1 - p$, y el par (f_K, g_K) con probabilidad p . Podemos modelar esto como:

$$\begin{aligned} \delta_i &:= \Lambda \cdot f_0 \left(\frac{1}{\sqrt{2}} G u_i \right) + (1 - \Lambda) \cdot f_K \left(\frac{1}{\sqrt{2}} G u_i \right), \\ \epsilon_j &:= \Lambda \cdot g_0 \left(\frac{1}{\sqrt{2}} G v_j \right) + (1 - \Lambda) \cdot g_K \left(\frac{1}{\sqrt{2}} G v_j \right), \end{aligned}$$

donde Λ es una variable aleatoria de Bernoulli que toma el valor 0 con probabilidad p y el valor 1 con probabilidad $1 - p$, independiente de la variable aleatoria Gaussiana $G \in \mathbb{R}^{2 \times n}$ usual.

Nos gustaría ahora construir un “preproceso” que corresponda a esta proyección aleatoria. Para esto querríamos, emulando lo que ocurre con los Métodos de Redondeo de Krivine normales, una función del estilo

$$\begin{aligned} \widetilde{H}_{\mu_p}(u, v) = \\ \mathbb{E}_{G, \Lambda} \left[\left(\Lambda \cdot f_0 \left(\frac{1}{\sqrt{2}} G u \right) + (1 - \Lambda) \cdot f_K \left(\frac{1}{\sqrt{2}} G u \right) \right) \left(\Lambda \cdot g_0 \left(\frac{1}{\sqrt{2}} G v \right) + (1 - \Lambda) \cdot g_K \left(\frac{1}{\sqrt{2}} G v \right) \right) \right], \end{aligned} \tag{2.32}$$

para cualesquiera $u, v \in \mathbb{S}^{n-1}$, de forma que \widetilde{H}_{μ_p} sólo depende del producto interno $\langle u, v \rangle$. Y en efecto, distribuyendo la esperanza por linealidad e independencia y usando que

$$\mathbb{E}[\Lambda(1 - \Lambda)] = \mathbb{E}[(1 - \Lambda)\Lambda] = 0 \quad \mathbb{E}[(1 - \Lambda)^2] = p \quad \mathbb{E}[\Lambda^2] = 1 - p,$$

el lado derecho de (2.32) se simplifica a

$$\begin{aligned} \widetilde{H}_{\mu_p}(u, v) &= (1 - p)\mathbb{E}_G \left[f_0 \left(\frac{1}{\sqrt{2}}Gu \right) g_0 \left(\frac{1}{\sqrt{2}}Gv \right) \right] + p\mathbb{E}_G \left[f_K \left(\frac{1}{\sqrt{2}}Gu \right) g_K \left(\frac{1}{\sqrt{2}}Gv \right) \right] = \\ &= (1 - p)H_{f_0, g_0}(\langle u, v \rangle) + pH_{f_K, g_K}(\langle u, v \rangle). \end{aligned}$$

Es decir que efectivamente $\widetilde{H}_{\mu_p}(u, v)$ depende sólo de $\langle u, v \rangle$, y entonces podemos definir

$$H_{\mu_p}(z) := (1 - p)H_{f_0, g_0}(z) + pH_{f_K, g_K}(z).$$

De nuevo, emulando a los Métodos de Redondeo de Krivine normales, lo que queremos es conseguir un $\gamma_{\mu_p} \in \mathbb{R}_{>0}$ tal que

$$\left(H_{\mu_p}^{-1} \right)^{abs}(\gamma_{\mu_p}) = 1. \quad (2.33)$$

De esta manera, para cualquier conjunto de vectores $\{u_i\}_i \{v_j\}_j \subseteq \mathbb{S}^{n-1}$, otro conjunto de vectores “preprocesados” $\{\tilde{u}_i\}_i \{\tilde{v}_j\}_j \subseteq \mathbb{S}^{n-1}$ tales que

$$\langle \tilde{u}_i, \tilde{v}_j \rangle = H_{\mu_p}^{-1}(\gamma_{\mu_p} \langle u_i, v_j \rangle). \quad (2.34)$$

Con este preproceso, seguido por la proyección aleatoria μ_p que definimos, conseguimos un proceso de redondeo aleatorio de vectores que cumple que, en esperanza,

$$\mathbb{E}[\delta_i, \epsilon_j] = H_{\mu_p}(\langle \tilde{u}_i, \tilde{v}_j \rangle) = H_{\mu_p} \left(H_{\mu_p}^{-1}(\gamma_{\mu_p} \langle u_i, v_j \rangle) \right) = \gamma_{\mu_p} \langle u_i, v_j \rangle.$$

Y por lo tanto, usando este redondeo sobre la evaluación óptima del problema $SDP(A)$ para cada matriz A , obtenemos que

$$K_G \leq \frac{1}{\gamma_{\mu_p}}.$$

Entonces, nuestro objetivo es probar que, para algún $p \in (0, 1)$, existe un γ_{μ_p} como en (2.33), y que

$$\gamma_{\mu_p} > \frac{2 \ln(1 + \sqrt{2})}{\pi}. \quad (2.35)$$

Para simplificar un poco las cuentas, vamos a definir las funciones auxiliares

$$H_p := \frac{\pi}{2} H_{\mu_p} \quad H_0 := \frac{\pi}{2} H_{f_0, g_0} \quad H_K := \frac{\pi}{2} H_{f_K, g_K}, \quad (2.36)$$

y por lo tanto vale que

$$H_p^{-1}(-) = H_{\mu_p}^{-1} \left(\frac{2}{\pi} - \right) \quad H_0^{-1}(-) = H_{f_0, g_0}^{-1} \left(\frac{2}{\pi} - \right) \quad H_K^{-1}(-) = H_{f_K, g_K}^{-1} \left(\frac{2}{\pi} - \right).$$

La ventaja es que ahora $H_0(z) = \arcsen z$, $H_0^{-1}(z) = \sen z$, y no estorba la constante de $\frac{\pi}{2}$.

Con esta definición, vale que

$$H_p(z) = (1 - p)H_0(z) + pH_K(z).$$

Además encontrar el γ_{μ_p} que buscábamos en (2.33) y (2.35) equivale a encontrar $\gamma \in \mathbb{R}_{>0}$ tal que

$$\left(H_p^{-1}\right)^{abs}(\gamma) = 1 \quad \text{con } \gamma = \frac{\pi}{2}\gamma_{\mu_p} > \frac{\pi}{2} \frac{2 \ln(1 + \sqrt{2})}{\pi} = \ln(1 + \sqrt{2}). \quad (2.37)$$

Vamos a necesitar algunos lemas para la demostración. Empezamos recordando el Teorema de Rouché de análisis complejo.

Teorema 2.4.5. *Sean $K \subset \mathcal{U} \subseteq \mathbb{C}$, K un conjunto acotado con frontera ∂K que es un lazo simple homotópico a un punto, \mathcal{U} abierto, y sean $f, g : \mathcal{U} \rightarrow \mathbb{C}$ dos funciones holomorfas. Si se cumple la desigualdad estricta*

$$|g(z)| < |f(z)| \quad \forall z \in \partial K,$$

entonces f y $f + g$ tienen la misma cantidad de ceros en K .

Lo primero va a ser demostrar que la función H_p^{-1} es analítica en algún entorno lo suficientemente bueno como para trabajar bien.

Lema 2.4.6. *Para todo $0 < r < 1$, existen $p_r \in (0, 1)$ y $\Omega_r \subseteq \mathcal{S}$ tal que $\forall p \in (0, p_r)$ se cumple que $H_p : \Omega_r \rightarrow r\mathbb{D}$ es analítica con inversa analítica. Además existe $C_r \in (0, \infty)$ tal que:*

$$|H_p^{-1}(z) - \text{sen } z - p(z - H_K(\text{sen } z)) \cos z| \leq C_r p^2 \quad \forall z \in r\mathbb{D}, p \in (0, p_r).$$

Demostración. La idea va a ser valernos del Teorema de Rouché, y de una acotación uniforme sobre H_η para encontrar un entorno donde H_p y H_0 sean “parecidos” y entonces H_p “herede” la inyectividad de H_0 .

Empecemos por notar que, como H_0 coincide con la función $\text{arc sen}(z)$ en el intervalo $(-1, 1)$, por unicidad de la extensión analítica resulta que $H_0(z) = \text{arc sen}(z)$ en la región \mathcal{S} . Es fácil que ver $\text{arc sen}(z)$ es inyectiva en \mathcal{S} , y que $|\text{Re}(\text{sen}(z))| < 1 \quad \forall z \in \mathbb{D}$. Esto nos dice que, $\forall r \in (0, 1)$, existe $\Gamma_r \subset \mathcal{S}$ tal que $H_0 = \text{arc sen} : \Gamma_r \rightarrow r\mathbb{D}$ es analítica.

Definimos los conjuntos $E_n := \left\{ z \in \mathbb{C} \mid |\text{Re}(z)| < 1 - \frac{1}{n} \text{ y } |\text{Im}(z)| < n \right\}$. Como se tiene que $\bigcup_{n=1}^{\infty} E_n = \mathcal{S}$, entonces existe algún n_r tal que $r\overline{\mathbb{D}} \subseteq H_0(E_{n_r})$.

Sea $\zeta \in r\overline{\mathbb{D}}$. Entonces existe $z \in E_{n_r}$ tal que $H_0(z) = \zeta$, y como arc sen es inyectiva, entonces $H_0(z) \neq \zeta \quad \forall z \in \partial E_{n_r+1}$. Luego, usando que $r\overline{\mathbb{D}}$ y ∂E_{n_r+1} son compactos, tiene sentido definir

$$m := \min_{\substack{\zeta \in r\overline{\mathbb{D}} \\ z \in \partial E_{n_r+1}}} |H_0(z) - \zeta| > 0. \quad (2.38)$$

Por otro lado, con el Lema (2.1.2), existe algún $M > 0$ tal que $|H_\eta(z)| < M \quad \forall \eta \geq 0, z \in \partial E_{n_r+1}$.

Con estas cotas a mano, vamos a definir $p_r := \frac{m}{2M}$, y fijamos $\zeta \in r\overline{\mathbb{D}}$. Si $p \in (0, p_r)$, entonces tenemos que

$$\underbrace{|p(H_\eta(z) - H_0(z))|}_{=:g(z)} \leq \frac{m}{2M} (|H_\eta(z)| + |H_0(z)|) < \frac{m}{2M} 2M \leq m \leq \underbrace{|H_0(z) - \zeta|}_{=:f(z)} \quad \forall z \in \partial E_{n_r+1}.$$

En particular con f y g definidas de esta manera $|g(z)| < |f(z)| \quad \forall z \in \partial E_{n_r+1}$. Entonces, por el teorema de Rouché (2.4.5), resulta que f y $f+g$ tienen la misma cantidad de ceros en E_{n_r+1} . Esto quiere decir que $p(H_\eta(z) - H_0(z))$ y $p(H_\eta(z) - H_0(z)) + H_0(z) - \zeta = pH_\eta(z) + (1-p)H_0(z) - \zeta = H_p - \zeta$ tienen la misma cantidad de ceros en E_{n_r+1} . Pero por inyectividad de H_0 en \mathcal{S} , hay un solo $z \in E_{n_r+1}$ tal que $H_0(z) = \zeta$. Luego hay un solo $z \in E_{n_r+1}$ tal que $H_p(z) = \zeta$. Como esto vale $\forall \zeta \in r\mathbb{D}$, esto quiere decir que H_p^{-1} está bien definida en $r\mathbb{D}$ y es analítica en ese dominio. Basta con definir entonces, $\Omega_r := H_p(E_{n_r+1})$.

Empecemos por ver que efectivamente la expresión restada corresponde al polinomio de Taylor de H_p^{-1} . Sea $z \in r\mathbb{D}$ fijo. Entonces

$$z = H_K(H_p^{-1}(z)) = (1-p)H_0(H_p^{-1}(z)) + pH_\eta(H_p^{-1}(z)).$$

Entonces podemos derivar en p a cada lado:

$$\begin{aligned} 0 &= H_K(H_p^{-1}(z)) - H_0(H_p^{-1}(z)) + \left((1-p) \frac{\partial}{\partial z} H_0(H_p^{-1}(z)) + p \frac{\partial}{\partial z} H_\eta(H_p^{-1}(z)) \right) \frac{\partial}{\partial p} H_p^{-1}(z) \\ 0 &= H_K(H_p^{-1}(z)) - H_0(H_p^{-1}(z)) + \frac{\partial}{\partial z} H_p(H_p^{-1}(z)) \frac{\partial}{\partial p} H_p^{-1}(z). \end{aligned}$$

Y por Teorema de la Función Inversa:

$$0 = H_K(H_p^{-1}(z)) - H_0(H_p^{-1}(z)) + \frac{1}{\frac{\partial}{\partial z} H_p^{-1}(z)} \frac{\partial}{\partial p} H_p^{-1}(z).$$

Entonces con un pasaje de términos, obtenemos

$$\frac{\partial}{\partial p} H_p^{-1}(z) = \left(H_0(H_p^{-1}(z)) - H_K(H_p^{-1}(z)) \right) \frac{\partial}{\partial z} H_p^{-1}(z). \quad (2.39)$$

Evaluemos entonces la función $H_p^{-1}(z)$ y su derivada en $p=0$:

$$\begin{aligned} H_p^{-1}(z) \Big|_{p=0} &= H_0^{-1}(z) = \text{sen } z. \\ \frac{\partial}{\partial p} H_p^{-1}(z) \Big|_{p=0} &= \left(H_0(H_0^{-1}(z)) - H_K(H_0^{-1}(z)) \right) \frac{\partial}{\partial z} H_0^{-1}(z) = \\ &= (z - H_K(\text{sen}(z))) \frac{\partial}{\partial z} \text{sen}(z) = (z - H_K(\text{sen } z)) \cos z. \end{aligned}$$

Si miramos el polinomio de Taylor de $H_p^{-1}(z)$ centrado en 0, resulta que

$$\begin{aligned} \left| H_p^{-1}(z) - H_p^{-1}(z) \Big|_{p=0} - p \frac{\partial}{\partial p} H_p^{-1}(z) \Big|_{p=0} \right| &\leq \frac{\frac{\partial^2}{\partial p^2} H_p^{-1}(z) \Big|_{p=p_\xi} p^2}{2!}, \quad p_\xi \in (0, p). \\ \left| H_p^{-1}(z) - \text{sen } z - p(z - H_K(\text{sen } z)) \cos z \right| &\leq \frac{\frac{\partial^2}{\partial p^2} H_p^{-1}(z) \Big|_{p=p_\xi} p^2}{2!}, \quad p_\xi \in (0, p). \end{aligned} \quad (2.40)$$

La derivada segunda $\frac{\partial^2}{\partial p^2} H_p^{-1}(z)$ la podemos hallar a partir de (2.39), derivando una vez más en p y volviendo a reemplazar $\frac{\partial}{\partial p} H_p^{-1}(z)$ usando (2.39) cada vez que aparezca en el desarrollo. Haciendo esto, la expresión de la derivada segunda queda en función de H_0, H_K , sus derivadas primeras (que están todas acotadas uniformemente por (2.1.2)), y las derivadas en z de $H_p^{-1}(z)$

(que podemos acotar en función de r usando la fórmula de Cauchy, porque H_p^{-1} es analítica y $H_p(\Gamma_r) \subseteq r\mathbb{D}$ para cualquier $p \in (0, p_r)$). Luego la derivada segunda puede acotarse en función de r para cualquier $p \in (0, r)$:

$$\left| \frac{\partial^2}{\partial p^2} H_p^{-1}(z) \right| \leq C_r, \quad \forall p \in (0, r).$$

Usando este último resultado en (2.40), completamos la demostración. \square

Por el Lema (2.4.6) sabemos que $\forall r \in (0, 1)$ existe $p_r \in (0, 1)$ y un abierto acotado $\Omega_r \subseteq \overline{\Omega_r} \subseteq \mathcal{S}$, tal que para todo $p \in (0, p_r)$ la función $H_p : \Omega_r \rightarrow r\mathbb{D}$ es analítica con inversa analítica y

$$|H_p^{-1}(z) - \text{sen } z - p(z - H_K(\text{sen } z)) \cos z| \leq C_r p^2 \quad \forall z \in r\mathbb{D}. \quad (2.41)$$

Escribamos la serie correspondiente a $H_p^{-1}(z)$ centrada en 0 (notar que como H_0 y H_K son impares, H_p y H_p^{-1} lo son):

$$H_p^{-1}(z) =: \sum_{k=0}^{\infty} a_{2k+1}(p) z^{2k+1}.$$

Nuestro objetivo es probar que, para alguna elección correcta de p , la existencia de un γ como en (2.37) con $\gamma > \ln(1 + \sqrt{2})$.

Consideramos ahora también las expansiones como series de:

$$\text{sen } z = \sum_{k=0}^{\infty} b_{2k+1} z^{2k+1} = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} z^{2k+1}.$$

y, motivados por (2.41)

$$\phi(z) = (z - H_K(\text{sen } z)) \cos z = \sum_{k=0}^{\infty} c_{2k+1} z^{2k+1}.$$

Vamos a fijar $r \in (\frac{9}{10}, 1)$. Por la Fórmula Integral de Cauchy, para cada $k \in \mathbb{N}_0$ se tiene:

$$|a_{2k+1}(p) - b_{2k+1} - p c_{2k+1}| = \left| \frac{1}{2\pi i r} \oint_{r\partial\mathbb{D}} \frac{H_p^{-1}(z) - \text{sen } z - p\phi(z)}{z^{2k+2}} dz \right| \leq \frac{C_r p^2}{r^{2k+2}}. \quad (2.42)$$

Además, por (2.1.2), resulta que H_K , y por lo tanto ϕ es analítica en \mathcal{S} . En particular, su serie converge absolutamente para cualquier $|z| < 1$. Usamos esto y (2.42) para acotar inferiormente la serie absoluta de H_p en $\frac{9}{10}$:

$$\begin{aligned} \sum_{k=0}^{\infty} |a_{2k+1}(p)| \left(\frac{9}{10}\right)^{2k+1} &\geq \underbrace{\sum_{k=0}^{\infty} \frac{\left(\frac{9}{10}\right)^{2k+1}}{(2k+1)!}}_{\text{sinh}(9/10)} - p \underbrace{\sum_{k=0}^{\infty} |c_{2k+1}| \left(\frac{9}{10}\right)^{2k+1}}_{< \infty, \text{ no depende de } p} - p^2 \underbrace{\frac{C_r}{r} \sum_{k=0}^{\infty} \left(\frac{9}{10r}\right)^{2k+1}}_{< \infty, \text{ no depende de } p} \\ &\geq \text{sinh}\left(\frac{9}{10}\right) - O(p) - O(p^2), \end{aligned}$$

y acotando por abajo el \sinh , y con $p \rightarrow 0$ tenemos:

$$> 1,02 - O(p).$$

Esto quiere decir que para p suficientemente pequeños, $\sum_{k=0}^{\infty} |a_{2k+1}(p)| \left(\frac{9}{10}\right)^{2k+1} > 1$. Por Bolzano, para cada p existe $\gamma = \gamma(p) \in (0, \frac{9}{10})$ tal que

$$\sum_{k=0}^{\infty} |a_{2k+1}(p)| (\gamma)^{2k+1} = 1.$$

Queremos probar que $\gamma > \ln(1 + \sqrt{2})$. Supongamos entonces que

$$\gamma \leq \ln(1 + \sqrt{2}). \quad (2.43)$$

Luego, como $\frac{9}{10} < r$:

$$\frac{\gamma}{r} \leq \frac{10 \ln(1 + \sqrt{2})}{9} < \frac{49}{50}. \quad (2.44)$$

Tenemos que $\text{sen}(\frac{9}{10}\mathbb{D}) \subseteq \text{sen}(\mathbb{D}) \subseteq \mathcal{S}$, donde H_K es analítica. Luego $H_K \circ \text{sen}$ es analítica en $\frac{9}{10}\mathbb{D}$ por ser composición de analíticas. Claramente entonces $\phi = (id - H_K \circ \text{sen}) \cdot \cos$ es analítica en $\frac{9}{10}\mathbb{D}$. En particular, ϕ tiene radio de convergencia al menos $\frac{9}{10}$. Como $0 < \gamma < \frac{9}{10}$, esto implica que la serie de ϕ converge absolutamente en γ :

$$\sum_{k=0}^{\infty} |c_{2k+1}| \gamma^{2k+1} < \infty,$$

y por lo tanto, $\forall \varepsilon > 0$, existe $n \in \mathbb{N}$ tal que la cola de la serie es arbitrariamente chica:

$$\sum_{k=n+1}^{\infty} |c_{2k+1}| \gamma^{2k+1} < \frac{\varepsilon}{2}. \quad (2.45)$$

Además, una vez elegidos ε y n , existe $p(\varepsilon) > 0$ tal que $\forall p \in (0, p(\varepsilon))$ se tiene que:

$$p|c_{2k+1}| < \frac{1}{2}|b_{2k+1}| \quad \forall k \in \{1, \dots, n\}.$$

Esto implica que $p|c_{2k+1}|$ no es lo suficientemente grande como para cambiar el signo de b_{2k+1} , y luego:

$$\text{sgn}(b_{2k+1} + pc_{2k+1}) = \text{sgn}(b_{2k+1}) = (-1)^k. \quad (2.46)$$

Ahora tratemos de acotar:

$$\begin{aligned} \left| 1 - \frac{H_p^{-1}(i\gamma)}{i} \right| &= \left| \sum_{k=0}^{\infty} |a_{2k+1}(p)| \gamma^{2k+1} - \frac{1}{i} \sum_{k=0}^{\infty} a_{2k+1}(p) (i\gamma)^{2k+1} \right| \\ &= \left| \sum_{k=0}^{\infty} \left(|a_{2k+1}(p)| - (-1)^k a_{2k+1}(p) \right) \gamma^{2k+1} \right| \end{aligned}$$

A esta altura, el módulo es supérfluo, porque $|a_{2k+1}(p)| \geq (-1)^k a_{2k+1}(p)$ y $\gamma > 0$:

$$= \sum_{k=0}^{\infty} \left(|a_{2k+1}(p)| - (-1)^k a_{2k+1}(p) \right) \gamma^{2k+1}.$$

Usando dos veces (2.42), y que $|A - B| \leq C \Rightarrow \pm A \leq C \pm B$, tenemos:

$$\begin{aligned} &\leq \sum_{k=0}^{\infty} \left(\frac{C_r p^2}{r^{2k+2}} + |b_{2k+1} + pc_{2k+1}| + \frac{C_r p^2}{r^{2k+2}} - (-1)^k (b_{2k+1} + pc_{2k+1}) \right) \gamma^{2k+1} \\ &\leq \underbrace{\sum_{k=0}^{\infty} \left(|b_{2k+1} + pc_{2k+1}| - (-1)^k (b_{2k+1} + pc_{2k+1}) \right) \gamma^{2k+1}}_{(1)} + 2 \underbrace{\sum_{k=0}^{\infty} \frac{C_r p^2}{r^{2k+2}} \gamma^{2k+1}}_{(2)}. \end{aligned}$$

Para acotar (1), notamos que por (2.46), los primeros n términos de la suma se anulan:

$$\begin{aligned} &\sum_{k=0}^{\infty} \left(|b_{2k+1} + pc_{2k+1}| - (-1)^k (b_{2k+1} + pc_{2k+1}) \right) \gamma^{2k+1} \leq \\ &\leq \sum_{k=n+1}^{\infty} \left(|b_{2k+1} + pc_{2k+1}| - |b_{2k+1}| - (-1)^k pc_{2k+1} \right) \gamma^{2k+1} = \\ &\leq 2p \sum_{k=0}^{\infty} c_{2k+1} \gamma^{2k+1} \stackrel{(2.45)}{<} p\varepsilon. \end{aligned}$$

Para acotar (2), simplemente hacemos:

$$2 \sum_{k=0}^{\infty} \frac{C_r p^2}{r^{2k+2}} \gamma^{2k+1} \leq p^2 \frac{2C_r}{r} \sum_{k=0}^{\infty} \left(\frac{r}{\gamma} \right)^{2k+1} \stackrel{(2.44)}{<} p^2 \underbrace{\frac{2C_r}{r} \sum_{k=0}^{\infty} \left(\frac{r}{\gamma} \right)^{2k+1}}_{=: C'_r} \leq p^2 C'_r,$$

donde C'_r depende solamente de r .

Sea entonces $\beta := H_p^{-1}(i\gamma) - i$. Acabamos de probar que:

$$|\beta| = \left| 1 - \frac{H_p^{-1}(i\gamma)}{i} \right| \leq p^2 C'_r + p\varepsilon. \quad (2.47)$$

Por último, como H_0 y H_K son analíticas en S , entonces ambas son Lipschitz en $i + \frac{1}{2}$. Como H_p es un promedio ponderado de ambas, entonces H_p también es Lipschitz con L la constante de Lipschitz más grande entre ambas, independientemente de p . Luego

$$\begin{aligned} \ln(1 + \sqrt{2}) &\stackrel{(2.43)}{\geq} \gamma = \frac{H_p(\beta + i)}{i} \geq \frac{H_p(i)}{i} - L|\beta| \stackrel{(2.47)}{\geq} \frac{(1-p)H_0(i) + pH_K(i)}{i} - Lp(pC'_r + \varepsilon) \\ &\stackrel{(2.36)(2.13)}{=} (1-p) \ln(1 + \sqrt{2}) + p \frac{H_K(i)}{i} - Lp(pC'_r + \varepsilon). \end{aligned}$$

Reordenando, obtenemos:

$$\frac{H_K(i)}{i} \leq \frac{p}{p} \ln(1 + \sqrt{2}) + pLC'_r + L\varepsilon \leq \ln(1 + \sqrt{2}) + pLC'_r + L\varepsilon.$$

Esto vale para cualquier $\varepsilon > 0$ y $p \in (0, p(\varepsilon))$. Pero esto claramente contradice (2.43). Luego $\gamma > \ln(1 + \sqrt{2})$.

2.5. Optimalidad de los Métodos de Redondeo de Krivine

En esta sección vamos a contar brevemente un resultado de [NR14] sobre una generalización de los Métodos de Redondeo de Krivine.

Recordemos que un Método de Redondeo de Krivine f, g viene asociado a una constante $\gamma_{f,g} \in \mathbb{R}_{>0}$, y que siempre $K_G \leq \frac{1}{\gamma_{f,g}}$. Sin embargo, no sabemos si los $\gamma_{f,g}$ asociados a MRK son ajustados, en el sentido de que

$$K_G = \inf_{f,g \text{ MRK}} \frac{1}{\gamma_{f,g}}. \quad (2.48)$$

Esto nos interesa por dos motivos. Primero, si esto fuera cierto, estudiar los Métodos de Redondeo de Krivine nos podría acercar al valor exacto de la constante de Grothendieck. Pero además, como cada MRK nos induce un algoritmo aleatorio de redondeo de soluciones, los MRK con $\frac{1}{\gamma_{f,g}}$ cercano a K_G nos darían algoritmos aleatorios que consiguen evaluaciones del problema $OPT(A)$ que en esperanza alcanzan un valor arbitrariamente cercano a $\frac{OPT(A)}{K_G}$ (y veremos en el próximo capítulo motivos para creer que esta aproximación sería óptima).

Hasta ahora el único resultado en este sentido que tenemos es el de (2.3.2), que dice que (2.48) es cierto si, para infinitos $k \in \mathbb{N}$, la forma bilineal de König se maximiza en un Método de Redondeo de Krivine Alternado. Sin embargo, aun no sabemos si esto es cierto para ningún $k > 1$.

Vamos a contar ahora un resultado en el espíritu de (2.48), pero para una definición un poco más relajada de los Métodos de Redondeo de Krivine. Introducimos entonces los Métodos de Redondeo de Krivine por Medida.

Definición 2.5.1. Un Método de Redondeo de Krivine por Medida (MRKM) de dimensión k consiste en una medida de probabilidad Boreliana μ en $\{-1, 1\}^{\mathbb{R}^k} \times \{-1, 1\}^{\mathbb{R}^k}$. Se pide además que el soporte de esta medida esté contenido en

$$\{(f, g) \mid f, g : \mathbb{R}^k \rightarrow \{-1, 1\} \text{ son medibles Lebesgue}\}$$

La idea es que un Método de Redondeo de Krivine Distribuido representa una familia de pares de particiones de \mathbb{R}^k , de forma que a cada partición se le asigna un “peso” (densidad de la medida). Queremos usar esta medida para la “proyección aleatoria” de manera similar a lo que venimos haciendo con los Métodos de Redondeo de Krivine. Dada una configuración de vectores preprocesada $\{\tilde{u}_i\}_i \{\tilde{v}_j\}_j \subset \mathbb{S}^{2n-1}$, elegimos una partición (f, g) al azar usando la medida μ y elegimos $G \in \mathbb{R}^{k \times 2n}$ una matriz Gaussiana estándar. Luego tomamos:

$$\delta_i := f(G\tilde{u}_i), \quad \epsilon_j := g(G\tilde{v}_j).$$

Con esta definición, la esperanza del producto de los δ_i, ϵ_j es

$$\mathbb{E}_G \left[\int f(G\tilde{u}_i) g(G\tilde{v}_j) d\mu(f, g) \right].$$

Esta última expresión depende sólo de $\langle \tilde{u}_i, \tilde{v}_j \rangle$, por la usual consideración de que la distribución de que, para dos vectores $u, v \in \mathcal{B}_{\mathcal{H}}$ son equivalentes las distribuciones

$$(Gu, Gv) \sim (G_1, \langle u, v \rangle G_1 + \sqrt{1 - \langle u, v \rangle^2} G_2),$$

donde $G_1, G_2 \in \mathbb{R}^k$ son vectores Gaussianos estándar.

Nos interesarán los casos en que podamos realizar el “preproceso” de una configuración de vectores $\{u_i\}_i, \{v_j\}_j \subset \mathbb{S}^{2n-1}$. Para esto planteamos la siguiente definición.

Definición 2.5.2. Decimos que un Método de Redondeo de Krivine por Medida tiene calidad $K \in \mathbb{R}_{>0}$ si $\forall \{u_i\}_i, \{v_j\}_j \in \mathbb{S}^{2n-1}$, existen $\{\tilde{u}_i\}_i, \{\tilde{v}_j\}_j \in \mathbb{S}^{2n-1}$ tales que

$$\mathbb{E}_G \left[\int f(G\tilde{u}_i) g(G\tilde{v}_j) d\mu(f, g) \right] = \frac{\langle u, v \rangle}{K}.$$

Notemos que un Método de Redondeo de Krivine a secas (f, g) es un Método de Redondeo de Krivine por Medida, donde la medida es atómica sobre una única partición $\{\tilde{f}, \tilde{g}\}$, definiendo $\tilde{f}(u) := f\left(\frac{1}{\sqrt{2}}u\right)$, $\tilde{g}(v) := g\left(\frac{1}{\sqrt{2}}v\right)$. Nuestros resultados de este capítulo nos dicen que un Método de Redondeo de Krivine (f, g) tiene calidad $K = \frac{1}{\gamma_{f,g}}$.

Un ejemplo más interesante es el método de redondeo μ_p que usamos para demostrar que $K_G > \frac{\pi}{2\ln(1+\sqrt{2})}$ en la sección anterior. El método redondeo que habíamos definido es exactamente el Método de Redondeo de Krivine por Medida que le corresponde a una medida atómica μ_p que tiene densidad $1-p$ en la partición (f_0, g_0) y densidad p en (f_K, g_K) . Lo que demostramos en el Lema (2.35) equivale a decir que el MRKM μ_p tiene calidad estrictamente mayor a $\frac{\pi}{2\ln(1+\sqrt{2})}$.

El siguiente teorema de [NR14] dice que para las MRKM sí podemos garantizar que las calidades aproximan la constante de Grothendieck.

Teorema 2.5.3 (Naor-Regev). *Existe para cada $k \in \mathbb{N}$ un Método de Redondeo de Krivine por Medida μ_k de dimensión k , de forma que las respectivas calidades K_k cumplen:*

$$K_k \xrightarrow{k \rightarrow \infty} K_G$$

Vamos a dar una idea básica de la demostración, que parte desde esta formulación equivalente de la desigualdad de Grothendieck:

Desigualdad de Grothendieck 2.5.4. *Sea $k \in \mathbb{N}$ fijo. Existe una medida de probabilidad Boreliana ν_k en $\{-1, 1\}^{\mathbb{S}^{k-1}} \times \{-1, 1\}^{\mathbb{S}^{k-1}}$ tal que*

$$\int_{\{-1, 1\}^{\mathbb{S}^{k-1}} \times \{-1, 1\}^{\mathbb{S}^{k-1}}} f(u)g(v)\nu_k(f, g) = \frac{1}{K_G} \langle u, v \rangle.$$

Usando esta formulación equivalente, la medida μ_k para el teorema (2.5.3) se define como una radialización de la medida ν_k en el siguiente sentido. Para cada par $(f, g) \in \{-1, 1\}^{\mathbb{S}^{k-1}} \times \{-1, 1\}^{\mathbb{S}^{k-1}}$, definimos el par $\{\tilde{f}, \tilde{g}\} \in \{-1, 1\}^{\mathbb{R}^k} \times \{-1, 1\}^{\mathbb{R}^k}$ de forma que

$$\tilde{f}(u) := f\left(\frac{u}{\|u\|}\right) \quad \tilde{g}(u) := g\left(\frac{u}{\|u\|}\right) \quad \forall u \in \mathbb{R}^k.$$

Se puede demostrar que estas medidas cumplen la propiedad enunciada en el teorema.

Lamentablemente, las medidas que se usan en la demostración vienen de resultados de existencia, y no nos permiten hacer cuentas que acoten la constante de Grothendieck, o nos den un algoritmo de redondeo explícito. No conocemos ningún resultado que mejore esta situación, ni que permita restringirnos a considerar algún conjunto más pequeño de medidas, por ejemplo, medidas con densidad concentrada en finitos puntos.

Capítulo 3

Complejidad computacional y aproximaciones

En este capítulo vamos a hablar sobre la complejidad de hallar el valor óptimo del problema $OPT(A)$ para matrices $A \in \mathbb{R}^{n \times n}$. El primer resultado, que ya mencionamos varias veces, es que este problema es NP-hard. Vamos a demostrar esto reduciendo el problema NP-hard MAX-CUT al problema $OPT(A)$. También veremos que el problema de hallar $SDP(A)$ está se puede resolver en tiempo polinomial, esbozando un algoritmo para su resolución.

Luego, dada la dificultad de hallar una solución exacta, nos abocaremos a la tarea de aproximar el valor de $OPT(A)$. Ahí encontraremos el resultado principal del capítulo, que dice que, asumiendo una hipótesis de complejidad conocida como Unique Games Conjecture, el valor de $OPT(A)$ no se puede aproximar polinomialmente con una garantía mejor que $\frac{OPT(A)}{K_G}$. Para llegar a este resultado, dedicaremos un tiempo a introducir la Unique Games Conjecture y explicar por qué es una hipótesis que parece razonable.

Finalmente, usaremos algunas de estas técnicas para dar, para $\varepsilon > 0$ fijo, un algoritmo aleatorio polinomial para hallar una solución de $OPT(A)$ mejor o igual que $\frac{OPT(A)}{K_G + \varepsilon}$, y otro para conseguir una aproximación de K_G con error menor a ε , pero en tiempo exponencial en ε .

3.1. Complejidad computacional de hallar $OPT(A)$

A pesar de todos estos avances en los últimos años, la constante de Grothendieck es aún desconocida, al punto de que no se sabe si su segundo dígito es 6 o 7. Dado que el problema tiene una formulación más “discreta” en [LP68], descripta en (1.1.1), cabría esperar que métodos computacionales directos hayan encontrado a esta altura mejores aproximaciones que los métodos analíticos del capítulo anterior.

Pero consideremos, por ejemplo, el problema de hallar cotas inferiores para K_G . Esto consiste, en general, en conseguir una matriz A , y usar el valor de $\frac{SDP(A)}{OPT(A)}$ como cota inferior de K_G , como lo definimos en (1.1.4). El valor de $SDP(A)$ se puede acotar desde abajo con sólo hallar una realización de vectores u_i, v_j que alcancen un valor alto, y esto resulta relativamente fácil computacionalmente (veremos de hecho que es “fácil” hallar la solución vectorial óptima). La dificultad de esta forma de encarar el problema es que acotar desde arriba el valor $OPT(A)$ de forma justa requiere efectivamente el cálculo de este valor. Esto significa, en una implementación

inocente al menos, hallar la asignación de valores δ_i, ϵ_j que maximiza el problema. Lamentablemente, es bastante sencillo probar que el problema de hallar $OPT(A)$ para matrices arbitrarias es NP-hard. Es decir, tenemos una gran diferencia de dificultad entre hallar el máximo del problema de optimización $OPT(A)$ (NP-hard) y hallar el máximo de $SDP(A)$, la relajación semidefinida positiva del mismo problema (¡resoluble en tiempo polinomial!).

Empecemos por demostrar que el problema de optimización de hallar $OPT(A)$ para una matriz $A \in n \times n$ es NP-hard. Para esto, vamos a reducir el problema de optimización MAX-CUT, uno de los problemas NP-hard más famosos, al problema de Grothendieck.

Definición 3.1.1. El problema de MAX-CUT es un problema de optimización combinatoria donde cada instancia es un grafo G . Una solución de la instancia G consiste en “pintar” cada nodo del grafo de uno de dos colores disponibles. La función de costo de la solución es la cantidad de aristas cuyos extremos están pintados de distintos colores sobre la cantidad total de aristas del grafo, y el objetivo del problema es maximizar esta función de costo. Denotaremos a este valor máximo como $OPT_{MC}(G)$.

El nombre del problema viene de separar los nodos del grafo G en dos grupos y “cortar” todas las aristas que cruzan de un grupo a otro. Con esta interpretación, lo que se está buscando es cortar la mayor cantidad de aristas posible, es decir, el MAX-CUT. Como dijimos, MAX-CUT es uno de los problemas NP-hard más famosos, y de hecho su versión como problema de decisión y con pesos en las aristas es uno de los 21 problemas de Karp, que son algunos de los primeros problemas interesantes para los que se demostró complejidad NP-hard.

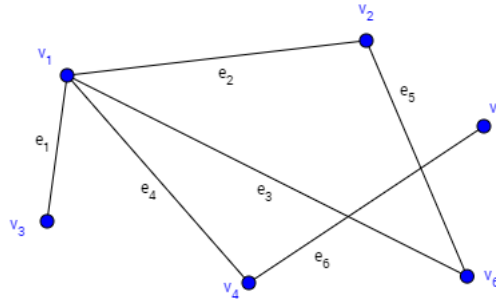
Vamos entonces a contar la reducción de MAX-CUT al problema de Grothendieck. El argumento es una adaptación de una demostración de [AN06].

Teorema 3.1.2. *El problema de hallar el valor óptimo $OPT(A)$ del problema de Grothendieck para una matriz $A \in \mathbb{R}^{m \times n}$ es NP-hard.*

Demostración. Empezamos con una instancia del problema MAX-CUT, es decir, un grafo G , que tiene un conjunto de nodos $V := \{v_1, \dots, v_n\}$ y un conjunto de aristas $E := \{e_1, \dots, e_m\} \subseteq V \times V$, y vamos a asumir que cada arista e_j viene dada por un par ordenado de nodos (e_j^1, e_j^2) . Vamos a construir entonces una matriz $A \in \mathbb{R}^{m \times n}$ de la siguiente manera:

$$a_{ij} := \begin{cases} 1/2, & \text{si } v_i = e_j^1 \\ -1/2, & \text{si } v_i = e_j^2 \\ 0, & \text{si no.} \end{cases}$$

En palabras, en la fila de una arista todos los coeficientes son cero, excepto por los nodos que corresponden a los extremos de la arista, que tienen los valores $\frac{1}{2}$ y $-\frac{1}{2}$, en algún orden. Por ejemplo, al siguiente grafo



le puede corresponder la siguiente matriz

$$\begin{array}{r}
 \\
 \\
 e_1 \rightarrow \\
 e_2 \rightarrow \\
 e_3 \rightarrow \\
 e_4 \rightarrow \\
 e_5 \rightarrow \\
 e_6 \rightarrow
 \end{array}
 \begin{pmatrix}
 v_1 & v_2 & v_3 & v_4 & v_5 & v_6 \\
 \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\
 1/2 & 0 & -1/2 & 0 & 0 & 0 \\
 -1/2 & 1/2 & 0 & 0 & 0 & 0 \\
 1/2 & 0 & 0 & 0 & 0 & -1/2 \\
 -1/2 & 0 & 0 & 1/2 & 0 & 0 \\
 0 & -1/2 & 0 & 0 & 0 & 1/2 \\
 0 & 0 & 0 & -1/2 & 1/2 & 0
 \end{pmatrix}.$$

La construcción de esta matriz claramente se puede realizar en tiempo polinomial en el tamaño del grafo G . Ahora supongamos que estamos buscando el valor óptimo $OPT(A)$ sobre la matriz A . Esto lo podemos pensar como encontrar signos para multiplicar cada columna y cada fila, de manera que se maximice la suma de los coeficientes de la matriz. Asumamos que dejamos fijos los signos que se van a asignar a las columnas de la matriz. Esto correspondería a asignarle un -1 o un 1 a cada nodo del grafo original G , y lo podemos pensar como que efectivamente está “pintando” el grafo con dos colores.

Con los signos de las columnas fijos, ¿cuál es el mayor valor de suma de coeficientes que podemos alcanzar cambiando los signos de las filas? Miremos, por ejemplo, la fila e_5 . Hay dos posibilidades. Si los extremos de e_5 tienen asignados signos iguales a sus columnas, entonces la fila va a contribuir 0 a la suma total. Por otro lado, si los signos de v_2 y v_6 son distintos, entonces podemos ajustar el signo de la fila para que contribuyan 1 a la suma total de coeficientes.

Haciendo esto fila por fila, está claro que, con los signos de las columnas fijos, la máxima suma de coeficientes posibles es exactamente sumar 1 por cada arista que tenga sus extremos en columnas de distinto signo. ¡Pero esas son exactamente las aristas del “corte” generado por la asignación de $+$ o $-$ a los nodos/columnas! Es decir, para cada asignación de $+$ o $-$ a las columnas, el valor máximo es exactamente el valor del corte que separa los nodos de acuerdo a qué signo les tocó en sus respectivas columnas.

Luego, el valor del mejor corte para el grafo G tiene que coincidir con el mejor valor posible para $OPT(A)$. Como la construcción de la matriz se puede realizar en tiempo polinomial, si tuviéramos un algoritmo que pudiera hallar en tiempo polinomial el valor $OPT(A)$ para cualquier matriz A , lo podríamos usar para hallar en tiempo polinomial el valor óptimo $OPT_{MC}(G)$ para cualquier grafo G . Por lo tanto el problema de Grothendieck es NP-hard. \square

Con esto demostramos que el problema de hallar $OPT(A)$ es NP-hard. En cambio, el problema de hallar $SDP(A)$ para cualquier matriz está es resoluble en tiempo polinomial. Esto se debe a que el problema se puede resolver con una estrategia conocida como Programación Semi-definida (de donde viene la notación $SDP(A)$) que describiremos brevemente. La manera más sencilla de explicar por qué se puede demostrar que $SDP(A)$ se resuelve en tiempo polinomial y por qué la demostración no funciona para $OPT(A)$ es intentar reducir el problema $OPT(A)$ a un problema de programación lineal. Vamos a ver que la reducción no funciona para $OPT(A)$, pero sí para $SDP(A)$. Como el problema de programación lineal se puede resolver en tiempo polinomial, esto implica que nuestro problema $SDP(A)$ también.

Básicamente, nos gustaría tratar de escribir las condiciones del problema $OPT(A)$ como las condiciones de un problema de programación lineal, que consiste, a grandes rasgos, en maximizar una función lineal en una región del espacio compacta, no degenerada, delimitada por desigualdades lineales. Para describir $OPT(A)$, definimos variables b_{ij} que van a cumplir la función de $\delta_i \epsilon_j$. Lo que nos interesa entonces es, dada una matriz $A \in \mathbb{R}^{n \times n}$, maximizar el programa lineal

$$\begin{cases} \text{máx} \sum_{i,j=1}^n a_{ij} b_{ij} \\ -1 \leq b_{ij} \leq 1 \end{cases} . \quad (3.1)$$

Este programa lineal se puede maximizar sin problemas. Pero sin embargo, la solución no nos sirve para $OPT(A)$ porque no tenemos ninguna garantía de que las variables b_{ij} correspondan efectivamente a variables del estilo

$$\begin{cases} b_{ij} = \delta_i \epsilon_j \\ -1 \leq \delta_i \leq 1 \\ -1 \leq \epsilon_j \leq 1 \end{cases} . \quad (3.2)$$

Dicho de otra manera, el programa lineal (3.1) nos devolverá siempre $b_{ij} = \text{sgn}(a_{ij})$, y el valor máximo de la función será $\sum_{i,j=1}^n |a_{ij}|$. Necesitamos entonces tratar de agregar la condición (3.2). El problema es que la condición de $b_{ij} = \delta_i \epsilon_j$ no es lineal, y tampoco hay forma de expresarla con varias condiciones lineales.

Sin embargo, se puede relajar un poco el problema para expresar una condición similar. Pensemos a los δ_i, ϵ_j como una sola lista de variables x_k , con $1 \leq k \leq 2n$:

$$x_k = \begin{cases} \delta_k & 1 \leq k \leq n \\ \epsilon_k & n+1 \leq k \leq 2n \end{cases} .$$

Usemos ahora variables y_{kl} , con $1 \leq k, l \leq 2n$. Entonces las condiciones (3.1) y (3.2) se expresan como

$$\begin{cases} \text{máx} \sum_{k=1}^n \sum_{l=n+1}^{2n} a_{kl} y_{kl} \\ y_{kl} = x_k x_l, & 1 \leq k \leq n; n+1 \leq l \leq 2n \\ -1 \leq y_{kl} \leq 1, & 1 \leq k \leq n; n+1 \leq l \leq 2n \\ -1 \leq x_k \leq 1, & 1 \leq k \leq 2n \end{cases} .$$

Este es el programa que nos gustaría resolver. Notemos que de las variables y_{kl} sólo estamos usando aquellas donde $1 \leq k \leq n$ y $n+1 \leq l \leq 2n$. No perdemos nada pidiendo consistencia y

usando la misma condición para todos los y_{kl} :

$$\begin{cases} \text{máx} \sum_{k=1}^n \sum_{l=n+1}^{2n} a_{kl} y_{kl} \\ y_{kl} = x_k x_l, \quad 1 \leq k, l \leq 2n \\ -1 \leq y_{kl} \leq 1, \quad 1 \leq k, l \leq 2n \\ -1 \leq x_k \leq 1, \quad 1 \leq k \leq 2n \end{cases} .$$

Seguimos teniendo el problema de la condición $y_{kl} = x_k x_l$ no lineal. Pero notemos que esta condición implica que y_{kl} es una matriz semidefinida positiva (es la matriz de Gram de los x_k). ¡Y resulta que la propiedad de ser una matriz semidefinida positiva sí es expresable linealmente! Equivale simplemente a decir que

$$qYq^t = \sum_{k=1}^{2n} v_k v_l y_{kl} \leq 0, \quad \forall q \in \mathbb{R}^{2n}. \quad (3.3)$$

Llegado este punto, no necesitamos las variables x_k en la formulación. Podemos escribir simplemente:

$$\begin{cases} \text{máx} \sum_{k=1}^{2n} \sum_{l=1}^{2n} a_{kl} y_{kl} \\ \sum_{k=1}^{2n} \sum_{l=1}^{2n} q_k q_l y_{kl} \leq 0, \quad \forall q \in \mathbb{R}^{2n} \\ -1 \leq y_{kl} \leq 1, \quad 1 \leq k, l \leq 2n \end{cases} .$$

Las restricciones con $q \in \mathbb{R}^{2n}$ son en realidad un número infinito, y a priori no se podrían usar para definir el programa. La manera de arreglar esto es no considerarlas parte formal del programa, sino agregar esas restricciones a medida que las soluciones halladas no satisfacen con la propiedad de que Y es una matriz semidefinida positiva. Más precisamente, si Y no es una matriz semidefinida positiva, se puede hallar en tiempo polinomial un vector $c \in \mathbb{R}^{2n}$ tal que $cYc^t < 0$. Luego se agrega en tiempo real de ejecución la restricción $cYc^t \leq 0$ al problema y se sigue ejecutando para encontrar una mejor solución. Esto funciona bien, por ejemplo, con el método de la elipsoide [GLS81] para resolver programas lineales. En este método, podemos incorporar la verificación de la existencia de q en el algoritmo que verifica si un punto pertenece al conjunto de soluciones válidas, y en caso contrario devuelve un hiperplano separador.

Ahora bien, ¿podemos garantizar que la solución hallada y_{kl} para el problema

$$\text{máx} \sum_{k=1}^{2n} \sum_{l=1}^{2n} a_{kl} y_{kl}$$

sea una solución que provenga de $y_{kl} = x_k x_l$, con $x_k \in \mathbb{R}$? Claramente no, pero por un resultado clásico de Álgebra Lineal, el hecho de que Y sea semidefinida positiva nos permite dar una descomposición $Y = CC^t$, con $C \in \mathbb{R}^{2n \times 2n}$ y que se puede hallar en tiempo polinomial usando la descomposición de Cholesky, por ejemplo. Luego las filas de C son vectores $v_k \in \mathbb{R}^{2n}$ tales que $y_{kl} = \langle v_k, v_l \rangle, \forall 1 \leq k, l \leq 2n$.

Es decir, nuestra relajación del problema de pedir la condición (3.2) al problema de pedir (3.3), nos da exactamente la formulación del problema $SDP(A)$. Esto ilustra cómo hallar la solución $SDP(A)$ en tiempo polinomial, y a la vez ilumina por qué este mismo método no nos sirve para $OPT(A)$.

Debido al Teorema (3.1.2) nuestras esperanzas de encontrar un algoritmo rápido para hallar $OPT(A)$ son inmediatamente destruidas. ¿Pero podríamos encontrar algoritmos eficientes que aproximen $OPT(A)$, al menos dentro de un factor razonable? Es decir, ¿podríamos encontrar algún algoritmo polinomial ALG , quizás aleatorio, que reciba cómo argumento una matriz A y nos devuelva

$$\frac{OPT(A)}{K} \leq ALG(A), \quad \text{para alguna constante } K \in \mathbb{R}_{>0} \quad (3.4)$$

Por supuesto, este algoritmo no puede ser polinomial independientemente de K , ya que esto nos daría una solución polinomial al problema de calcular $OPT(A)$. Pero sí cabría la posibilidad de que, con $\varepsilon > 0$, tengamos un algoritmo polinomial en la dimensión del problema que obtenga siempre

$$\frac{OPT(A)}{1 + \varepsilon} \leq ALG(A),$$

con la reserva de que el tiempo de este algoritmo crezca mucho al achicar ε y por esto no nos sirva para aproximaciones arbitrariamente cercanas de tiempo polinomial.

Esta posibilidad también se ve desterrada al considerar la reducción que hicimos de MAX-CUT a nuestro problema, que conservaba exactamente el valor de la solución óptima. Es un hecho conocido que el problema de MAX-CUT tiene un “inapproximation gap” de $\frac{17}{16}$. Es decir, que existe un $K \in \mathbb{R}_{>0}$ tal que, si asumimos $P \neq NP$, no existe ningún algoritmo polinomial ALG_{MC} que, para cualquier instancia G del problema MAX-CUT obtenga:

$$\frac{16}{17} OPT_{MC}(G) < ALG_{MC}(G).$$

Pero de la reducción que usamos para (3.1.2), sale inmediatamente que cualquier garantía de aproximación para el problema de Grothendieck sería inmediatamente trasladable al problema de MAX-CUT. Esto nos dice entonces que no existe bajo $P \neq NP$, al igual que con el problema de MAX-CUT, ningún algoritmo polinomial ALG que cumple

$$\frac{16}{17} OPT(A) \leq ALG(A) \quad \forall A \in \mathcal{M}(\mathbb{R}). \quad (3.5)$$

Pero nos queda ahora una nueva pregunta, ¿cuál es el menor K que cumple esta condición?

Un resultado muy fuerte de Raghavendra y Steurer en [RS09] demuestra que, bajo cierta hipótesis de complejidad conocida como la Unique Games Conjecture (UGC, ver (3.2.6)), la mejor constante es $K := K_G$, ¡la constante de Grothendieck!

Teorema 3.1.3. *Supongamos que la Unique Games Conjecture es cierta y sea $K \in \mathbb{R}$ tal que $K > K_G$. Entonces el problema de hallar, para cualquier matriz $A \in \mathbb{R}^{n \times n}$, un valor $ALG(A)$ alcanzable en el problema $OPT(A)$ tal que*

$$\frac{OPT(A)}{K} \leq ALG(A) \leq OPT(A)$$

es NP-hard. En particular, si $P \neq NP$, no existe ningún algoritmo de tiempo polinomial ALG que para cualquier $A \in \mathbb{R}^{n \times n}$ obtenga una evaluación válida del problema $OPT(A)$ que garantice

$$\frac{OPT(A)}{K} \leq ALG(A) \leq OPT(A).$$

Notar que, sorprendentemente, el teorema funciona sin conocer el valor exacto de la constante de Grothendieck K_G . Es decir, no conocemos el valor exacto de la constante de Grothendieck definida según las desigualdades de los primeros capítulos. Pero sea cuál sea este valor, coincide exactamente con la mejor constante de aproximación que se puede conseguir, bajo UGC, para el problema $OPT(A)$. Este teorema está entonces, para el problema de optimización de Grothendieck, vinculando el valor del gap de integralidad de una relajación semidefinida positiva al valor del gap de inaproximación del problema.

El Teorema (3.1.3) está basado en un trabajo anterior de Raghavendra [Rag08] que consigue un resultado similar, bajo la misma hipótesis de complejidad UGC, para toda una familia de problemas conocidos como Constraint Satisfiability Problems (CSP). El resultado dice que para todos estos problemas el gap de inaproximación es exactamente el gap de integralidad, es decir, el gap entre la solución óptima del problema y la solución óptima de una relajación de programación semidefinida del problema (SDP). Además, se obtiene un algoritmo aleatorio polinomial que alcanza esta mejor aproximación salvo por un factor aditivo arbitrariamente chico. Aún más, para una subclase de estos problemas conocida como 2-CSP, donde las restricciones son fórmulas que siempre se aplican sobre exactamente dos variables, [Rag08] consigue un resultado extra. Dado $\eta > 0$ y sin la hipótesis de UGC, se describe un algoritmo de redondeo de ciertas relajaciones semidefinidas positivas naturales que obtiene una evaluación para el problema que está a lo sumo a un error aditivo de η del valor cuya existencia garantiza el gap de integralidad K_{int} del problema. Es decir, si

$$K_{int} := \sup_{I \in \mathcal{I}} \frac{SDP(I)}{OPT(I)},$$

entonces el algoritmo ALG garantiza

$$\frac{SDP(I)}{K_{int}} - \eta \leq ALG(I).$$

El problema de Grothendieck es un caso particular de un CSP y también resulta ser un 2-CSP ya que las “restricciones” del problema las podemos pensar como restricciones del tipo $\delta_i \epsilon_j = 1$. Sin embargo, el resultado de [Rag08] no se puede aplicar directamente en este caso por dos motivos. Segundo, la reducción general del paper no conserva la propiedad de “bipartición” del problema de Grothendieck (las variables $\{u_i\}$ y $\{v_j\}$ se relacionan de grupo a grupo, pero no directamente entre ellas). Y además el resultado de aproximar bien excepto un factor aditivo no es muy práctico en este caso, porque el valor óptimo del problema de Grothendieck normalizado puede ser muy pequeño incluso para instancias muy grandes, y por este motivo el error aditivo supera ampliamente en magnitud a la aproximación alcanzada.

3.2. La Unique Games Conjecture y técnicas asociadas

La Unique Games Conjecture es una conjetura introducida por Subhash Khot en su paper [Kho02]. Desde la demostración del teorema de PCP en [ALM⁺98], se generó un gran avance en el estudio de la complejidad de aproximación de problemas de optimización. Los resultados de esta época probaron que un problema de decisión basado en un problema de optimización conocido como Label Cover es NP-hard. Subsecuentes avances lograron reducir este problema a problemas de decisión de otros problemas de optimización (como por ejemplo, MAX-3SAT), demostrando a su vez que estos problemas de decisión son NP-hard.

Lo interesante es que la dificultad de un problema de decisión apropiado para un problema de optimización implica la dificultad de aproximar el problema de optimización con cierta constante. Vamos a explicar más en detalle esta relación.

Supongamos que tenemos un problema de optimización combinatoria \mathcal{I} . Es decir, a grandes rasgos, un problema donde hay que buscar entre muchas soluciones posibles la solución que maximiza cierta función de costo. Denotaremos una instancia del problema \mathcal{I} como I , y denotaremos al valor máximo que puede alcanzar una solución de I como $OPT(I)$. Supongamos que tenemos un algoritmo ALG que recibe una instancia I y nos devuelve el costo $ALG(I)$ de una solución válida para la instancia I . Entonces, la inexistencia de un algoritmo ALG que pueda distinguir velozmente entre las situaciones $OPT(I) \leq s$ y $OPT(I) \geq c$ implica la inexistencia de un algoritmo que garantice siempre $\frac{s}{c}OPT(I) < ALG(I)$. Vamos a definir estas situaciones con algo más de precisión:

Definición 3.2.1. Dado un problema de optimización \mathcal{I} , vamos a definir el problema de decisión $\mathcal{I}_{c,s}^{gap}$. Las instancias de $\mathcal{I}_{c,s}^{gap}$, son las instancias de \mathcal{I} para las cuales o bien $OPT(I) \leq s$, o bien $OPT(I) \geq c$. El objetivo del problema de decisión es, dada una instancia de $\mathcal{I}_{c,s}^{gap}$, decidir a cuál de los dos casos corresponde. El parámetro c se denomina Completeness y el parámetro s se denomina Soundness.

Lema 3.2.2. Sea \mathcal{I} un problema de optimización. Luego si el problema de decisión $\mathcal{I}_{c,s}^{gap}$ es NP-hard, entonces el problema de hallar un algoritmo ALG para \mathcal{I} tal que $\frac{s}{c}OPT(I) < ALG(I)$ es NP-hard. En otras palabras, bajo la hipótesis de $P \neq NP$, no existe ningún algoritmo polinomial ALG que pueda garantizar $\frac{s}{c}OPT(I) < ALG(I)$.

Demostración. La prueba es bastante sencilla. Vamos a mostrar que la existencia de este algoritmo ALG nos permitiría discernir entre los dos tipos de instancia del problema $\mathcal{I}_{c,s}^{gap}$.

Supongamos que I es una instancia de $\mathcal{I}_{c,s}^{gap}$. Si $ALG(I) > s$, en particular $OPT(I) \geq ALG(I) > s$, entonces $OPT(I) \not\leq s$, y por lo tanto la instancia tiene que ser del otro tipo. Luego $OPT(I) \geq c$.

Por otro lado, si $ALG(I) \leq s$, entonces por la hipótesis $OPT(I) < \frac{c}{s}ALG(I) \leq \frac{c}{s}s = c$. Luego I tiene que ser una instancia del tipo $OPT(I) \leq s$.

Es decir que correr el algoritmo ALG y ver si el resultado es mayor que s permite discernir a qué caso corresponde $I \in \mathcal{I}_{c,s}^{gap}$. \square

Resultados relacionados al Teorema PCP demostraron que, para el problema de Label Cover \mathcal{LC} (3.2.3), con $\varepsilon > 0$ tal que el problema de decisión $\mathcal{LC}_{1,\varepsilon}^{gap}$ es NP-hard. Desde entonces, se logró en reiteradas oportunidades para algún problema de optimización (llamémosle \mathcal{I}), reducir polinomialmente el problema $\mathcal{LC}_{1,\varepsilon}^{gap}$ a $\mathcal{I}_{c,s}^{gap}$ para algunos $c, s \in \mathbb{R}$. De esta manera, se prueba que el problema $\mathcal{I}_{c,s}^{gap}$ es NP-hard, y por lo tanto que aproximar el problema \mathcal{I} con mejor constante que $\frac{s}{c}$ es NP-hard. Por ejemplo, de esta manera se logró probar que aproximar MAX-3SAT con una constante mejor que $\frac{7}{8}$ es NP-hard, lo que coincide con la aproximación alcanzada por los mejores algoritmos.

Sin embargo, para muchos problemas de aproximación NP muy importantes, seguía sin poder demostrarse una constante de aproximación óptima. Con esto queremos decir que, a partir de la NP-hardness de $\mathcal{LC}_{1,\varepsilon}^{gap}$, se lograban demostrar resultados de inaproximación para alguna constante K_2 , pero el mejor algoritmo conocido apenas llegaba a garantizar un valor de $\frac{OPT(I)}{K_1}$ para alguna constante K_1 , con $K_1 > K_2$.

En este contexto es que Khot introduce la hipótesis de que una variante del problema de decisión de Label Cover, que denominó Unique Label Cover, es NP-hard. Sorprendentemente incluso para su autor, la hipótesis resultó muy fructífera. En muchos problemas en los que el progreso estaba trabado, se logró demostrar que hallar mejores aproximaciones para el problema que las conocidas implica resolver el problema de Unique Label Cover, y por lo tanto, es al menos igual de difícil. Luego, bajo la hipótesis de que resolver Unique Label Cover es NP-hard, es NP-hard aproximar en tiempo polinomial estos problemas mejor que los algoritmos conocidos.

El impacto de este resultado fue aumentado por el hecho de que muchos de estos algoritmos no se creían óptimos en ese momento, sino que activamente se estaban buscando mejores opciones. Algunos ejemplos de avances en este sentido son [KR08], [KKMO07], [KD06], [Rag08]. Hay varios surveys que recopilan esta información, por ejemplo [Kho]. En particular, [Rag08] sienta las bases y las ideas del trabajo siguiente [RS09], del que nos vamos a ocupar en este capítulo.

Para dar un ejemplo concreto en un problema que ya conocemos, bajo la hipótesis de $P \neq NP$ la mejor cota de inaproximación conocida para el problema de MAX-CUT es la de $\frac{17}{16}$. Sin embargo, el mejor algoritmo conocido para el problema [GW95] apenas alcanza una garantía en esperanza de $\alpha_{MC} OPT_{MC}(G)$, donde $\frac{1}{\alpha_{MC}} \approx 1,13 > \frac{17}{16}$. Por lo tanto, o bien había que encontrar un algoritmo con mejor garantía, o bien demostrar que ese algoritmo era óptimo (lo que se creía falso). Sin embargo, bajo la hipótesis de UGC, se puede probar que el problema de MAX-CUT tiene un gap de inaproximación de exactamente $\frac{1}{\alpha_{MC}}$.

Un hecho muy interesante es que el algoritmo óptimo para MAX-CUT funciona pasando por una relajación semidefinida del problema análoga a la que describimos para el problema de Grothendieck y que nos llevaba de $OPT(A)$ a $SDP(A)$. De hecho, MAX-CUT es uno de los problemas que originaron estas técnicas, y para los que se probó la equivalencia, bajo UGC, de la equivalencia entre el gap de inaproximación y el gap de integralidad entre la relajación semidefinida y el problema original.

Vamos a definir entonces el problema de Unique Label Cover, y la Unique Games Conjecture.

Empecemos por una formulación del problema de Label Cover. Básicamente el problema nos da dos grupos de labels M y N que podemos pensar como “colores” con los que queremos pintar los nodos de un grafo bipartito. Una instancia de Label Cover consistirá en un grafo bipartito, y ciertas restricciones para cada arista sobre cómo se pueden pintar sus extremos. El objetivo es pintar los nodos de una manera que satisfaga la mayor cantidad de restricciones posibles.

Definición 3.2.3. Sean M, N conjuntos de labels finitos, con $\#M \geq \#N$ (y en general $\#M \gg \#N$). Una instancia I del problema de Label Cover $\mathcal{LC}(M, N)$ consiste en un grafo bipartito $G(V, W)$ completo, donde cada arista (u, v) tiene asociada una función $\pi_{u,v} : M \rightarrow N$.

Una asignación de labels a la instancia I , consiste en dar labels $L(u) \in M, L(v) \in N$ para cada $u \in U, v \in V$. Decimos que la asignación L satisface a la arista (u, v) si $\pi_{u,v}(L(u)) = L(v)$ (es decir, si la asignación de labels es “consistente” con la restricción impuesta sobre esa arista).

Definimos como el valor de una asignación L a la fracción del total de aristas que son satisfechas por esta asignación. El valor $OPT(I)$ corresponde entonces al mejor valor que puede alcanzar una asignación de labels L sobre la instancia I .

Teorema 3.2.4 (Consecuencia del Teorema PCP y el Teorema de Repetición Paralela de Raz). *Para cualquier $\delta > 0$ suficientemente chico, existe una constante $k := k(\delta)$ tal que para el problema de Label Cover con conjuntos de labels M, N con $\#M, \#N \leq k$, es NP-hard distinguir*

entre los casos $OPT(I) = 1$ y $OPT(I) \leq \delta$. En otras palabras, para δ suficientemente chico, existe k tal que el problema $\mathcal{LC}(M, N)_{1, \delta}^{gap}$ con $\#M, \#N \leq k$ es NP-hard.

La Unique Games Conjecture surge debido a las dificultades que traía el hecho de que $\#M \gg \#N$ en algunas de las reducciones que se buscaban realizar. Para salvar este problema, Khot propone un tipo especial de problema de Label Cover llamado Unique Label Cover Game (o Unique Game, para acortar):

Definición 3.2.5. Una instancia de Unique Game $\mathcal{UG}(M)$ es una instancia de Label Cover en el que las dimensiones de los conjuntos de labels coinciden (podemos suponer que hay un sólo conjunto de labels $M = N$), y la restricción $\pi_{u,v}$ de cada arista es una biyección.

Por supuesto, los Unique Games son instancias particulares del problema de Label Cover, y el problema es decidir si son instancias suficientemente “flexibles” como para capturar o conservar la NP-hardness del problema de gap de Label Cover.

Notemos que la introducción de esta restricción de biyección implica que el problema de decisión de determinar si todas las restricciones son simultáneamente satisfacibles está en P. Esto se debe a que, si buscamos un label satisfacible, decidir el label que le corresponde a un nodo inmediatamente determina el label de cada uno de sus vecinos. Así sucesivamente, la asignación queda determinada en toda la componente conexa del grafo. Luego, basta con intentar la viabilidad de cada label en un sólo nodo para determinar si la instancia es 100% satisfacible.

Esto quiere decir que el problema $\mathcal{UG}(M)_{1, \delta}^{gap}$ está en P $\forall \delta < 1$. La solución para conseguir complejidad NP-hard en este tipo de instancias es relajar la Completeness del problema de gap. En tal caso, al ir “tanteando” asignaciones de labels, no poder cumplir la restricción de una arista no implica el descarte del proceso parcial, y esto podría potencialmente aumentar mucho la complejidad de la decisión. Esto es precisamente lo que postula la Unique Games Conjecture:

Definición 3.2.6 (Unique Games Conjecture). Para cualesquiera constantes suficientemente chicas $\varepsilon, \delta > 0$ existe una constante $k := k(\varepsilon, \delta)$ tal que, para instancias de Unique Game con conjunto de label de tamaño k , es NP-hard determinar si $OPT(I) \geq 1 - \varepsilon$ o $OPT(I) \leq \delta$. Es decir, para M con $\#M = k$, el problema $\mathcal{UG}(M)_{1-\varepsilon, \delta}^{gap}$ es NP-hard.

La Unique Games Conjecture fue introducida por Khot en 2002, y sigue siendo un problema abierto al día de hoy. De hecho, dada la cantidad de resultados óptimos que implicaría una respuesta positiva, es considerada como uno de los problemas abiertos más importantes de la teoría de complejidad. Y quizás, el hecho de ser bastante similar al problema de Label Cover, cuya complejidad fue demostrada después de un largo proceso, alienta a ilusionarse con la posibilidad de una demostración más cercana que la de $P \neq NP$. De todas maneras, y aunque el resultado principal que contaremos en este capítulo es dependientes de la veracidad de la Unique Games Conjecture, cabe destacar que su postulación llevó a la creación de nuevas técnicas y algunos resultados que seguirían siendo válidos incluso si ésta fuera demostrada como falsa.

Una gran parte de las reducciones de Unique Games a otros problemas, en especial problemas de 2-CSP como es el caso del problema de Grothendieck, se basa en la utilización de Long Codes, desarrollado por Bellare y Håstad. Esto permite, generalmente mediante la utilización de algún test de dictadura apropiado, reducir una instancia de Unique Game a una instancia del 2-CSP, donde las variables a evaluar corresponden a bits de los Long Codes de las asignaciones del Unique Game. Para explicar esto en más detalle, vamos a definir estos términos:

Definición 3.2.7. Definimos como el cubo de Hamming \mathbb{H}^k al cubo k -dimensional dado por $\{-1, 1\}^k$. Vamos a asociarle a \mathbb{H}^k la probabilidad uniforme, con densidad $\frac{1}{2^k}$ en cada punto, y que usaremos en general silenciosamente.

Vamos a darle muy especial atención a las funciones $f : \mathbb{H}^k \rightarrow \{-1, 1\}$. Mucho del vocabulario que usaremos para este análisis viene de la economía y teoría de decisiones. El razonamiento es que una función f de este tipo equivale a un método de tomar decisiones entre dos opciones $\{-1, 1\}$ de acuerdo a las preferencias de k individuos. De esta manera, $f(x)$ representa la decisión tomada cuando el individuo i tiene la preferencia x_i , para cada $1 \leq i \leq k$. En este contexto, queda claro la siguiente definición, que en esta interpretación equivale a que un individuo toma la decisión unilateralmente:

Definición 3.2.8. Decimos que una función $f : \mathbb{H}^k \rightarrow \{-1, 1\}$ es una dictadura si existe $i_0 \in \{1, \dots, k\}$ tal que $f(x) = x_{i_0} \quad \forall x \in \mathbb{H}^k$. Para cada $i_0 \in \{1, \dots, k\}$, denotaremos $d_{i_0} \in L_2(\mathbb{H}^k)$,

$$d_{i_0}(x) := x_{i_0},$$

a la dictadura sobre la i_0 -ésima coordenada.

De la misma manera, tiene sentido cuantificar la influencia de la preferencia de un individuo sobre la decisión final. Esto se puede modelar con la pregunta “¿Qué tan probable es que un cambio de preferencia del individuo i_0 cambie la decisión final?”. Definimos pues:

Definición 3.2.9. Definimos la “influencia” de la coordenada i_0 en $f : \mathbb{H}^k \rightarrow \{-1, 1\}$ como:

$$\text{Inf}_{i_0}(f) = \mathbb{P}[f(x_1, x_2, \dots, \mathbf{x}_{i_0}, \cdot, x_k) \neq f(x_1, x_2, \dots, -\mathbf{x}_{i_0}, \cdot, x_k)],$$

donde la probabilidad se toma uniformemente sobre todos los $x \in \mathbb{H}^k$.

Notar que una dictadura como definimos en (3.2.8) tendrá influencia 1 en su coordenada dictadora e influencia 0 en todo el resto. Por este motivo, tiene sentido considerar como “opuestas” de las dictaduras a aquellas funciones en las que la influencia de una ninguna variable es muy grande. Con este objetivo, definimos:

Definición 3.2.10. Definimos la máxima influencia de $f : \mathbb{H}^k \rightarrow \{-1, 1\}$ como

$$\text{MaxInf } f := \max_{i \in \{1, \dots, k\}} \text{Inf}_i f.$$

Todas estas definiciones son estándar dentro de la teoría de complejidad de aproximación. Vamos a usar una definición extra, forzada por la particularidad de que el problema de Grothendieck es bipartito (las variables se pueden separar en dos grupos, las i y las j , de forma que las valuaciones del problema usan siempre una variable de cada grupo).

Definición 3.2.11. Definimos a la máxima influencia común de un par de funciones $f, g : \mathbb{H}^k \rightarrow \{-1, 1\}$ como

$$\text{MaxComInf}(f, g) := \max_{i \in \{1, \dots, k\}} \{\min \{ \text{Inf}_i f, \text{Inf}_i g \}\}.$$

Es decir, $\text{MaxComInf}(f, g)$ es la máxima influencia que f y g tienen simultáneamente en una misma variable.

El método usual para reducir instancias de Unique Game a otros problemas es codificar las asignaciones de labels del problema con funciones del cubo de Hamming. Más específicamente, si L es una asignación de una instancia I de $\mathcal{UG}(M)$ con $\#M = k$, vamos a asignarle a cada nodo w del grafo de I una dictadura $f : \mathbb{H}^k \rightarrow \{-1, 1\}$ con $f(x) := f_{L(w)}$. Luego, para pasar a la situación de variables booleanas, vamos a codificar cada una de estas funciones en un Long Code.

Definición 3.2.12. Supongamos que los vértices del cubo de Hamming \mathbb{H}^k están ordenados de alguna manera $\{h_1, h_2, \dots, h_{2^k}\}$. A cada función $f : \mathbb{H}^k \rightarrow \{-1, 1\}$ le asignamos un código “Long Code” de 2^k bits, dónde el bit i es el valor de $f(h_i)$. Dicho de otra manera, el Long Code de f son las coordenadas de f en la base de $L_2(\mathbb{H}^k)$ dada por $\{\chi_{\{h_1\}}, \chi_{\{h_2\}}, \dots, \chi_{\{h_{2^k}\}}\}$, con $\chi_{\{h\}}(k) = 1$ para $h = k$ y $\chi_{\{h\}}(k) = 0$ para $h \neq k$.

Claramente tenemos una biyección entre los Long Codes y las funciones $f : \mathbb{H}^k \rightarrow \{-1, 1\}$. Lo que nos va a permitir sacar ventaja de esta forma de codificar los datos es el hecho de que, bajo UGC, es difícil distinguir entre que una función sea una dictadura o tenga baja máxima influencia mirando pocos bits de su Long Code.

Para explotar esta diferencia entre dictaduras y funciones de baja máxima influencia vamos a usar el “operador de ruido” u “operador de Bonami-Beckner”. La idea es, dado un punto $x \in \mathbb{H}^k$, estudiar lo que le pasa al valor de una de nuestras funciones en un punto si cambiamos cada bit x_i de x con probabilidad $r \in [0, 1]$. Si hacemos esto para cada $x \in \mathbb{H}^k$, obtenemos una función que denotaremos $T_{1-2r}f$ (el motivo de esta notación quedará claro en un par de párrafos). Es decir:

Definición 3.2.13. Para $x \in \mathbb{H}^k$ fijo $y \sim_r x$ denota que la variable $y \in \mathbb{H}^k$ tiene la distribución resultante de cambiar el signo de cada coordenada de x independientemente con probabilidad r . El “operador de ruido” u “operador de Bonami-Beckner” $T_{1-2r} : L_2(\mathbb{H}^k) \rightarrow L_2(\mathbb{H}^k)$ se define como:

$$T_{1-2r}f(x) = \mathbb{E}_{y \sim_r x}[f(y)] \quad \forall x \in \mathbb{H}^k.$$

Para entender mejor qué es lo que hace efectivamente el operador de ruido, vamos a considerar la descomposición de las funciones $L_2(\mathbb{H}^k)$ en la siguiente base ortonormal de polinomios:

$$Pol_{\mathbb{H}^k} := \left\{ \chi_S(x) := \prod_{i \in S} x_i \mid S \subseteq \{1, \dots, k\} \right\} \subseteq L_2(\mathbb{H}^k). \quad (3.6)$$

Vamos a ver qué ocurre al aplicar el operador de ruido sobre esta base:

$$T_{1-2r}\chi_S(x) = \mathbb{E}_{y \sim_r x}[\chi_S(y)] = \mathbb{E}_{y \sim_r x} \left[\prod_{i \in S} y_i \right],$$

y por independencia de las coordenadas de y :

$$\begin{aligned} &= \prod_{i \in S} \mathbb{E}_{y_i \sim_r x_i} [y_i] = \prod_{i \in S} (x_i \cdot \mathbb{P}(y_i = x_i) + (-x_i) \cdot \mathbb{P}(y_i = -x_i)) \\ &= \prod_{i \in S} (x_i \cdot (1 - r) + (-x_i) \cdot r) = \prod_{i \in S} (x_i \cdot (1 - 2r)) = (1 - 2r)^{\#S} \chi_S(x). \end{aligned}$$

Esto ilumina nuestra notación para el operador, y nos permite dar la siguiente definición alternativa.

Definición 3.2.14. Para cada $d \in \{1, \dots, k\}$, definimos el proyector $P_d : L_2(\mathbb{H}^k) \rightarrow L_2(\mathbb{H}^k)$ como el proyector ortogonal sobre el espacio

$$Pol_{\mathbb{H}^k}^{(d)} := \langle \chi_S \rangle_{\substack{S \subseteq \{1, \dots, k\} \\ \#S=d}}.$$

Definición 3.2.15. Una definición alternativa, dado un $\rho \in [0, 1]$, del operador de ruido $T_\rho : L_2(\mathbb{H}^k) \rightarrow L_2(\mathbb{H}^k)$ es:

$$T_\rho f := \sum_{d=0}^k \rho^d P_d f.$$

Resulta claro en esta definición, además, que el operador de ruido es autoadjunto y contractivo en $L_2(\mathbb{H}^k)$, ya que es diagonalizable en una base ortonormal, y todos sus autovalores son menores o iguales a 1.

Una de las herramientas más útiles para trabajar con funciones en $L_2(\mathbb{H}^k)$ es compararlas con funciones en $L_2(\mathcal{G}^k)$, donde \mathcal{G}^k denota al espacio \mathbb{R}^k con la medida Gaussiana estándar k -dimensional.

En particular, se puede demostrar que una función $f \in L_2(\mathbb{H}^k)$ tiene mucho en común con una extensión natural a $L_2(\mathcal{G}^k)$ que denotaremos la extensión multilineal.

Definición 3.2.16. Dada una versión $f \in L_2(\mathbb{H}^k)$, consideramos su serie única de Fourier sobre los polinomios definidos en (3.6)

$$f = \sum_{S \subseteq \{1, \dots, k\}} f_S \chi_S,$$

luego definimos $\tilde{f} \in L_2(\mathcal{G}^k)$ la extensión multilineal de f como

$$\tilde{f} = \sum_{S \subseteq \{1, \dots, k\}} f_S \widetilde{\chi}_S,$$

donde $\widetilde{\chi}_S$ es el mismo polinomio multilineal $\prod_{i \in S} x_i$, pero interpretado como polinomio con dominio en \mathbb{R}^k .

El operador de ruido tal como lo definimos en (3.2.15) tiene sentido para las funciones en $L_2(\mathcal{G}^k)$, con la diferencia de que ahora en lugar de ser suryectivo, su imagen es un subespacio propio (el subespacio de los polinomios multilineales de grado menor o igual a k). Es trivial ver que

$$\widetilde{T_\rho f} = T_\rho \tilde{f} \quad \forall f \in L_2(\mathbb{H}^k), \rho \in (0, 1).$$

El siguiente lema expone otra de las formas en que las funciones de $L_2(\mathbb{H}^k)$ se relacionan con sus extensiones multilineales.

Lema 3.2.17. Sean $u, v \in \mathbb{S}^{m-1}$, y $f, g \in L_2(\mathbb{H}^k)$. Entonces

$$\mathbb{E}_G \left[\tilde{f}(Gu) \tilde{g}(Gv) \right] = \langle f, T_{\langle u, v \rangle} g \rangle,$$

donde $G \in \mathbb{R}^{k \times m}$ es una matriz Gaussiana estándar.

Demostración. Por linealidad, alcanza para probarlo para f y g funciones de la base $Pol_{\mathbb{H}^k}$. Como las variables tanto en \mathbb{H}^k como en \mathcal{G}^k son independientes entre sí, entonces se ve claro que, si $S_1, S_2 \subseteq \{1, \dots, k\}$ con $S_1 \neq S_2$, entonces

$$\mathbb{E}_G [\chi_{S_1}(Gu) \chi_{S_2}(Gv)] = 0 = \langle \chi_{S_1}, T_{\langle u, v \rangle} \chi_{S_2} \rangle,$$

porque al menos una de las variables queda “suelta”, y todas tienen esperanza cero. Por último, si $S \subseteq \{1, \dots, k\}$, tenemos

$$\mathbb{E}_G [\chi_S(Gu) \chi_S(Gv)] = \prod_{i \in S} (Gu)_i (Gv)_i = \langle u, v \rangle^{\#S} = \langle u, v \rangle^{\#S} \mathbb{E}_{\mathbb{H}^k} [\chi_S^2(x)] = \langle \chi_S, T_{\langle u, v \rangle} \chi_S \rangle,$$

con lo que el lema queda probado. \square

3.3. Complejidad computacional de aproximar $\text{OPT}(\mathbf{A})$

En las reducciones UGC se suele utilizar como paso intermedio un test de dictaduras. Esto consiste en algún “evaluador” de bits de Long Code con alguna lógica similar a la del problema al que queremos reducir el Unique Game, de manera que el “evaluador” alcance un valor “alto” en dictaduras y un valor “bajo” en funciones con baja máxima influencia. En nuestro contexto del problema de Grothendieck, tiene sentido la idea de tomar dos Long Codes y multiplicarlos, con algunos pesos apropiados que corresponderán a los a_{ij} . Para garantizar cierta consistencia, vamos a aplicarles alguna cantidad de ruido variable a las funciones.

Definición 3.3.1. En nuestro contexto, un test de dictadura será un operador $\mathcal{D} : L_2(\mathbb{H}^k) \rightarrow L_2(\mathbb{H}^k)$ de la forma

$$\mathcal{D} := \sum_{d=0}^k \lambda_d P_d$$

donde $\lambda_d \in \mathbb{R}$, y $|\lambda_1| \geq |\lambda_d| \quad \forall 0 \leq d \leq k$.

La condición de $|\lambda_1| \geq |\lambda_d|$ nos va a permitir asegurar que si $|\lambda_1| = 1$, entonces \mathcal{D} es un operador contractivo: $\|\mathcal{D}f\|_2 \leq \|f\|_2$. En general, si $|\lambda_1| \geq |\lambda_d|$, podremos normalizarlo dividiendo por una constante para obtener un operador contractivo.

La forma en que vamos a usar los test de dictadura es, dados $f, g \in \mathbb{H}^k \rightarrow \{-1, 1\}$, evaluar

$$\langle T_{1-\eta} f, \mathcal{D} T_{1-\eta} g \rangle$$

Notemos que esto equivale a una evaluación del problema de Grothendieck donde la matriz es la matriz asociada al operador $T_{1-\eta} \mathcal{D} T_{1-\eta}$ en la base $Pol_{\mathbb{H}^k}$, y los valores de δ_i, ϵ_j corresponden a los bits de los Long Codes de f y g respectivamente. Con esto en mente, vamos a dar las definiciones asociadas a este “evaluador”.

Definición 3.3.2. Dado un test de dictadura $\mathcal{D} \in L_2(\mathbb{H}^k)$, definimos los siguientes parámetros:

$$\text{Completeness}_\eta(\mathcal{D}) := \inf_{i \in \{1, \dots, k\}} \langle T_{1-\eta} d_{x_i}, \mathcal{D} T_{1-\eta} d_{x_i} \rangle = \lambda_1 (1 - \eta)^2$$

$$\text{Soundness}_{\eta, \tau}(\mathcal{D}) := \sup_{\substack{f, g \in L_2(\mathbb{H}^k) \\ \|f\|_\infty, \|g\|_\infty \leq 1 \\ \text{MaxComInf}(T_{1-\eta} f, T_{1-\eta} g) \leq \tau}} \langle T_{1-\eta} f, \mathcal{D} T_{1-\eta} g \rangle$$

$$\text{Completeness}(\mathcal{D}) := \text{Completeness}_0(\mathcal{D}) = \lambda_1.$$

Veamos que dada una matriz $A \in \mathbb{R}^{n \times n}$, podemos construir un test de dictadura $\mathcal{D}_\eta(A)$ con Completeness $\approx SDP(A)$ y Soundness $\approx OPT(A)$. Para empezar, hallamos una solución exacta (a efectos prácticos) del problema $SDP(A)$ con el método descrito en la sección 1. Sea $\{u_i\}_n, \{v_i\}_n \in \mathbb{R}^{2n}$ tal solución.

Definición 3.3.3. Sea $A \in \mathbb{R}^{n \times n}$, con $\{u_i\}_n, \{v_i\}_n \in \mathbb{R}^{2n}$ una solución de $SDP(A)$. Para $0 \leq d \leq k$ definimos:

$$\lambda_d := \sum_{i,j=1}^n a_{ij} \langle u_i, v_j \rangle^d.$$

Luego definimos los tests de dictadura asociados

$$\mathcal{D}(A) := \sum_{d=0}^k \lambda_d P_d, \quad \mathcal{D}_\eta(A) := T_{1-\eta} \mathcal{D}(A) T_{1-\eta} = \sum_{d=0}^k (1-\eta)^{2d} \lambda_d P_d.$$

Por supuesto, queremos verificar que efectivamente $|\lambda_1| \geq |\lambda_d| \quad \forall 1 \leq d \leq k$, tal como pedimos en (3.3.1). Para esto, notemos simplemente que, si $\{u_i\}, \{v_j\} \subseteq B_{\mathcal{H}}$ Hilbert, entonces podemos definir $\{u_i^d\}, \{v_j^d\} \subseteq B_{\mathcal{H}^{\otimes d}}$ Hilbert de la siguiente manera:

$$u_i^d := \underbrace{u_i \otimes u_i \otimes \cdots \otimes u_i}_{d \text{ veces}}, \quad v_j^d := \underbrace{v_j \otimes v_j \otimes \cdots \otimes v_j}_{d \text{ veces}}.$$

Claramente, con esta definición $\langle u_i^d, v_j^d \rangle = \langle u_i, v_j \rangle^d$ y por lo tanto, al ser $\{u_i^d\}, \{v_j^d\}$ una solución particular del problema semidefinido para cualquier d , $1 \leq d \leq k$, se tiene:

$$|\lambda_1| = \left| \sum_{i,j=1}^n a_{ij} \langle u_i, v_j \rangle \right| = SDP(A) \geq \left| \sum_{i,j=1}^n a_{ij} \langle u_i^d, v_j^d \rangle \right| = \left| \sum_{i,j=1}^n a_{ij} \langle u_i, v_j \rangle^d \right| = |\lambda_d|.$$

Aprovechamos para probar un lema que usaremos un poco más adelante:

Lema 3.3.4.

$$\|\mathcal{D}(A)\|_{L_2(\mathbb{H}^k) \rightarrow L_2(\mathbb{H}^k)} \leq SDP(A) \leq 2 \cdot OPT(A).$$

Demostración. La igualdad de Parseval nos dice que

$$\|f\|_2^2 = \sum_{S \subseteq \{1, \dots, k\}} f_{\chi_S}^2.$$

Luego, si $\|f\|_2^2 = 1$, usando que $|\lambda_1| \geq \lambda_d \quad \forall d \in \{1, \dots, k\}$:

$$\begin{aligned} \|\mathcal{D}(A)f\|_2^2 &= \sum_{S \subseteq \{1, \dots, k\}} (\mathcal{D}(A)f)_{\chi_S}^2 = \sum_{S \subseteq \{1, \dots, k\}} \lambda_{\#S}^2 f_{\chi_S}^2 \\ &\leq \lambda_1^2 \underbrace{\sum_{S \subseteq \{1, \dots, k\}} f_{\chi_S}^2}_{=1} = \lambda_1^2 = SDP(A)^2. \end{aligned}$$

Entonces, usando que $K_G \leq 2$:

$$\|\mathcal{D}(A)\|_2 = \sup_{\|f\|_2=1} \|\mathcal{D}(A)f\|_2 \leq SDP(A) \leq 2 \cdot OPT(A).$$

□

Teorema 3.3.5. Dada $A \in \mathbb{R}^{n \times n}$, se tiene que, para τ y η suficientemente pequeños,

$$\begin{aligned} \text{Completeness}(\mathcal{D}(A)) &= \text{SDP}(A) \\ \text{Soundness}_{\eta, \tau}(\mathcal{D}(A)) &\leq \text{OPT}(A)(1 + \eta). \end{aligned}$$

La parte de Completeness es inmediata por la definición de λ_1 para el test $\mathcal{D}(A)$

$$\text{Completeness}(\mathcal{D}(A)) = \lambda_1 = \text{SDP}(A). \quad (3.7)$$

La parte del Soundness es considerablemente más complicada. Vamos a definir, para cada par de funciones $f, g \in L_2(\mathbb{H}^k)$ y cada $\eta > 0$, un método de redondeo aleatorio $R_{\eta, f, g}$ que recibe una evaluación vectorial para el problema $\text{SDP}(A)$ y nos devuelve una evaluación de signos, de forma similar a lo que hacíamos en la Sección 2.2. Luego veremos que dadas dos funciones $f, g \in L_2(\mathbb{H}^k)$ con bajas influencias comunes, es posible encontrar un par de funciones f', g' tal que la esperanza del valor de la evaluación de signos del redondeo $R_{\eta, f', g'}$ sea casi $\langle T_{1-\eta}f, \mathcal{D}T_{1-\eta}g \rangle$. Pero la evaluación de signos no puede tener un valor más grande que $\text{OPT}(A)$, por definición. Por lo tanto, el máximo valor que puede alcanzar $\langle T_{1-\eta}f, \mathcal{D}T_{1-\eta}g \rangle$ sobre funciones de baja influencia común, es decir, el $\text{Soundness}_{\eta, \tau}$, está prácticamente acotado por $\text{OPT}(A)$.

Definamos entonces cómo funciona el proceso de redondeo $R_{\eta, f, g}$.

Definición 3.3.6. Dadas dos funciones $f, g \in L_2(\mathbb{H}^k)$ y $0 < \eta < 1$, y dados vectores $\{u_i\}, \{v_j\} \in \mathbb{S}^{m-1}$, vamos a elegir aleatoriamente signos δ_i, ϵ_j de la siguiente manera:

- Se computan \tilde{f}, \tilde{g} , las extensiones multilineales de f y g .
- Se genera una matriz Gaussiana aleatoria $G \in \mathbb{R}^{k \times m}$.
- Se asignan $\delta_i := \text{trunc } T_{1-\eta} \tilde{f}(Gu_i)$, $\epsilon_j := \text{trunc } T_{1-\eta} \tilde{g}(Gv_j)$.

Definición 3.3.7. Dadas dos funciones $f, g \in L_2(\mathbb{H}^k)$ y $0 < \eta < 1$, y dados vectores $\{u_i\}, \{v_j\} \in \mathbb{S}^{m-1}$ que son una solución óptima del problema $\text{SDP}(A)$ para una matriz A , vamos a definir

$$R_{\eta, f, g}(A) := \mathbb{E}_G \left[\sum_{i, j=1}^n a_{ij} \delta_i \epsilon_j \right],$$

que es el valor esperado del proceso de redondeo $R_{\eta, f, g}$ sobre la mejor solución vectorial posible.

Para demostrar que un valor de Soundness alto implica un valor de redondeo alto usaremos el siguiente lema. Nos dice que a partir de funciones f, g con máxima influencia común chica, se pueden conseguir funciones f', g' que tienen todas las influencias chicas, pero que alcanzan un valor bastante parecido con el operador $\langle T_{1-\eta}(-), \mathcal{D}T_{1-\eta}(-) \rangle$ al alcanzado por f, g .

Lema 3.3.8. Sea \mathcal{D} un test de dictadura en $L_2(\mathbb{H}^k)$, y $f, g \in L_2(\mathbb{H}^k)$, con

$$\|f\|_\infty, \|g\|_\infty \leq 1, \quad \text{MaxComInf}(T_{1-\eta}f, T_{1-\eta}g) \leq \tau.$$

Entonces para cada $\tau' > 0$ existen funciones $f', g' \in L_2(\mathbb{H}^k)$, con

$$\|f'\|_\infty, \|g'\|_\infty \leq 1, \quad \text{MaxInf}T_{1-\eta}f', \text{MaxInf}T_{1-\eta}g' \leq \tau',$$

tales que

$$\langle T_{1-\eta}f', \mathcal{D}T_{1-\eta}g' \rangle \geq \langle T_{1-\eta}f, \mathcal{D}T_{1-\eta}g \rangle - 2\|\mathcal{D}\| \sqrt{\frac{\tau}{\tau'\eta}}.$$

Además, se cumple que $\text{MaxComInf}(T_{1-\eta}f', T_{1-\eta}g') \leq \tau$.

Ahora sí, usando este lema, vamos a tomar $\tau' := \sqrt{\tau}$. Luego existen f', g' como en (3.3.8) tales que

$$\begin{aligned} \langle T_{1-\eta}f', \mathcal{D}(A)T_{1-\eta}g' \rangle &\geq \langle T_{1-\eta}f, \mathcal{D}(A)T_{1-\eta}g \rangle - 2\|\mathcal{D}(A)\| \sqrt{\frac{\tau}{\sqrt{\tau}\eta}} \\ &= \langle T_{1-\eta}f, \mathcal{D}(A)T_{1-\eta}g \rangle - 2\|\mathcal{D}(A)\| \frac{\sqrt[4]{\tau}}{\eta}. \end{aligned}$$

Y ahora usando (3.3.4), resulta que $\|\mathcal{D}(A)\| \leq 2OPT(A)$

$$\langle T_{1-\eta}f', \mathcal{D}(A)T_{1-\eta}g' \rangle \geq \langle T_{1-\eta}f, \mathcal{D}(A)T_{1-\eta}g \rangle - 4 \cdot OPT(A) \frac{\sqrt[4]{\tau}}{\eta}. \quad (3.8)$$

Supongamos que $\{u_i\}, \{v_j\} \in \mathbb{R}^{2n}$ es la solución óptima de $SDP(A)$ sobre la que construimos $\mathcal{D}(A)$ como en (3.3.3). Esto nos dice que

$$\begin{aligned} \langle T_{1-\eta}f', \mathcal{D}(A)T_{1-\eta}g' \rangle &= \sum_{d=0}^k \langle T_{1-\eta}f', \lambda_d P_d T_{1-\eta}g' \rangle \\ &= \sum_{d=0}^k \left\langle T_{1-\eta}f', \left(\sum_{i,j} a_{ij} \langle u_i, v_j \rangle^d \right) P_d T_{1-\eta}g' \right\rangle \\ &= \sum_{i,j} a_{ij} \left\langle T_{1-\eta}f', \sum_{d=0}^k \langle u_i, v_j \rangle^d P_d T_{1-\eta}g' \right\rangle \\ &= \sum_{i,j} a_{ij} \langle T_{1-\eta}f', T_{\langle u_i, v_j \rangle} T_{1-\eta}g' \rangle, \end{aligned}$$

que por (3.2.17), resulta ser

$$= \sum_{i,j} a_{ij} \mathbb{E}_G \left[T_{1-\eta} \tilde{f}'(Gu_i) T_{1-\eta} \tilde{g}'(Gv_j) \right]. \quad (3.9)$$

Notemos que $T_{1-\eta} \tilde{f}'(-\cdot u_i)$ y $T_{1-\eta} \tilde{g}'(-\cdot v_j)$ son vectores en el espacio de Hilbert probabilístico $L_2(\mathcal{G}^{k \times d})$, donde el producto interno es la esperanza sobre la distribución Gaussiana del producto de funciones. Sacando la esperanza hacia afuera, nos encantaría a esta altura acotar esta última expresión en función de $OPT(A)$, acotando en cada valor que toma G . El problema es que $T_{1-\eta} \tilde{f}'$ y $T_{1-\eta} \tilde{g}'$ no están acotados. Acá es donde surge la idea de aproximar estas funciones con funciones truncadas, y acotar un poco el impacto que esto tiene sobre la esperanza.

Usamos el siguiente lema de invarianza de [MOO10]:

Lema 3.3.9. *Existe una constante universal C tal que $\forall \eta \in (0, 1)$ se cumple que, dada $f \in L_2(\mathbb{H}^k)$, $\|f\|_\infty \leq 1$, $\text{Inf}_i(T_{1-\eta}f) \leq \tau \quad \forall i \in \{1, \dots, k\}$,*

$$\|T_{1-\eta} \tilde{f} - \text{trunc } T_{1-\eta} \tilde{f}\| \leq \tau^C \eta,$$

donde la función $\tilde{f} \in L_2(\mathcal{G}^k)$ representa la extensión única multilineal de f a \mathbb{R}^k como en 3.2.16.

Como comentario al margen, este lema, que a priori parece bastante inocente, es realmente muy profundo, y es de hecho el motivo por el que estamos usando los operadores de ruido $T_{1-\eta}$ en nuestra reducción. Básicamente, los autores de [MOO10] encuentran toda una serie de cotas para la diferencia entre una generalización de los polinomios sobre espacios de probabilidad abstractos que cumplen cierta condición de hipercontractividad. El origen de estos teoremas tiene que ver con la demostración de una conjetura que dice que la función de mayoría del cubo de Hamming es la función más estable de baja influencia para el operador de ruido. Como consecuencia de esa conjetura fue que se logró demostrar que la mejor constante de inaproximación del problema de MAX-CUT bajo la hipótesis de UGC coincide con la garantía del mejor algoritmo conocido. Más aún, estos resultados encontraron algunas aplicaciones muy interesantes en teorías de elección social, donde las funciones del cubo de Hamming modelan cómo tomar una decisión de acuerdo a las preferencias de k personas. En el Teorema (3.4.4) de la última sección, la demostración, de la que sólo contamos la idea, usa fuertemente los resultados de este paper en un contexto mucho más general que el del lema que estamos citando en este momento.

Acotemos, por último, cuán lejos está el valor

$$R_{\eta, f', g'}(A) = \sum_{i,j}^n a_{ij} \mathbb{E}_G \left[\text{trunc } T_{1-\eta} \tilde{f}'(Gu_i) \text{ trunc } T_{1-\eta} \tilde{g}'(Gv_j) \right].$$

del valor de (3.9). Para esto, definimos los vectores en $L_2(\mathcal{G}^{k \times m})$:

$$\begin{aligned} u'_i &= T_{1-\eta} \tilde{f}'(Gu_i) & u''_i &= \text{trunc } T_{1-\eta} \tilde{f}'(Gu_i) \\ v'_i &= T_{1-\eta} \tilde{g}'(Gv_i) & v''_i &= \text{trunc } T_{1-\eta} \tilde{g}'(Gv_i). \end{aligned}$$

Con esta notación tenemos que:

$$\sum_{i,j}^n a_{ij} \langle u''_i, v''_j \rangle = \sum_{i,j}^n a_{ij} \langle u'_i, v'_j \rangle - \sum_{i,j}^n a_{ij} \langle u'_i - u''_i, v'_j \rangle - \sum_{i,j}^n a_{ij} \langle u''_i, v'_i - v''_i \rangle. \quad (3.10)$$

La desigualdad de Grothendieck nos permite acotar por arriba cada uno de los términos usando, por ejemplo, que $K_G \leq 2$:

$$\sum_{i,j}^n a_{ij} \langle \vec{w}_i, \vec{z}_j \rangle \leq SDP(A) \max_{1 \leq i \leq n} \|\vec{z}_j\| \max_{1 \leq j \leq n} \|\vec{w}_i\| \leq 2OPT(A) \max_{1 \leq i \leq n} \|\vec{w}_i\| \max_{1 \leq j \leq n} \|\vec{z}_j\|. \quad (3.11)$$

Ahora sólo nos resta acotar la norma de los vectores que aparecen. Por (3.9), tenemos que

$$\|u'_i - u''_i\|, \|v'_j - v''_j\| \leq \tau^{\frac{C\eta}{2}}. \quad (3.12)$$

Por otro lado, $\|f'\|_2 \leq \|f'\|_\infty \leq 1$ y $\|g'\|_2 \leq \|g'\|_\infty \leq 1$ porque \mathbb{H}^k es un espacio de probabilidad. Tomar la extensión multilineal de una función del cubo de Hamming conserva su norma. La rotación sobre el dominio de las \tilde{f}' , \tilde{g}' que hacen los u_i y v_j no cambia la norma porque son vectores de norma 1, y la medida del espacio es invariante por rotaciones. Por último, el operador de ruido es contractivo porque $1 - \eta \leq 1$. Luego

$$\|u'_i\| \leq 1 \quad \|v'_j\| \leq 1.$$

Finalmente, también podemos acotar las normas de los vectores u_i'' y v_j'' porque trunc es claramente un operador contractivo:

$$\|u_i''\| \leq \|u_i'\| \leq 1 \quad \|v_j''\| \leq \|v_j'\| \leq 1. \quad (3.13)$$

Insertando lo hallado en (3.11), (3.12) y (3.13) en la ecuación (3.10), obtenemos:

$$\begin{aligned} R_{\eta,f,g}(A) &= \sum_{i,j} a_{ij} \langle u_i'', v_j'' \rangle = \\ &= \sum_{i,j} a_{ij} \langle u_i', v_j' \rangle - \sum_{i,j} a_{ij} \langle u_i' - u_i'', v_j' \rangle - \sum_{i,j} a_{ij} \langle u_i'', v_j' - v_j'' \rangle \\ &\geq \sum_{i,j} a_{ij} \langle u_i', v_j' \rangle - SDP(A) \max_{1 \leq i \leq n} \|u_i' - u_i''\| \max_{1 \leq j \leq n} \|v_j'\| \\ &\quad - SDP(A) \max_{1 \leq i \leq n} \|u_i''\| \max_{1 \leq j \leq n} \|v_j' - v_j''\| \\ &\geq \sum_{i,j} a_{ij} \langle u_i', v_j' \rangle - 2 \cdot OPT(A) \tau^{\frac{C\eta}{2}} - 2 \cdot OPT(A) \tau^{\frac{C\eta}{2}} \\ &= \sum_{i,j} a_{ij} \langle u_i', v_j' \rangle - 4 \cdot OPT(A) \tau^{\frac{C\eta}{2}}, \end{aligned}$$

y entonces por (3.8)

$$\geq \langle f, \mathcal{D}(A)g \rangle - 4 \cdot OPT(A) \tau^{\frac{C\eta}{2}} - 4 \cdot OPT(A) \frac{\sqrt[4]{\tau}}{\eta}.$$

Reagrupando, lo que obtuvimos es que para cualquier $f, g \in L_2(\mathbb{H}^k)$, con

$$\|f\|_{\infty}, \|g\|_{\infty} \leq 1, MaxComInf(T_{1-\eta}f, T_{1-\eta}g) \leq \tau, \quad (3.14)$$

existen $f', g' \in L_2(\mathbb{H}^k)$ que cumplen las mismas propiedades (3.14) y tales que

$$\langle f, \mathcal{D}(A)g \rangle \leq R_{\eta,f',g'}(A) + \left(4 \cdot \tau^{\frac{C\eta}{2}} + 4 \cdot \frac{\sqrt[4]{\tau}}{\eta} \right) OPT(A).$$

Entonces podemos tomar supremo sobre todas las f, g que cumplen (3.14) a ambos lados, y la desigualdad se mantiene.

$$\begin{aligned} \sup_{\substack{f,g \in L_2(\mathbb{H}^k) \\ \|f\|_{\infty}, \|g\|_{\infty} \leq 1 \\ MaxComInf(T_{1-\eta}f, T_{1-\eta}g) \leq \tau}} \langle f, \mathcal{D}(A)g \rangle &\leq \sup_{\substack{f,g \in L_2(\mathbb{H}^k) \\ MaxComInf(T_{1-\eta}f, T_{1-\eta}g) \leq \tau}} R_{\eta,f,g}(A) + O\left(\tau^{\frac{C\eta}{2}}\right) OPT(A) \\ Soundness_{\eta,\tau}(A) &\leq \sup_{\substack{f,g \in L_2(\mathbb{H}^k) \\ MaxComInf(T_{1-\eta}f, T_{1-\eta}g) \leq \tau}} R_{\eta,f,g}(A) + O\left(\tau^{\frac{C\eta}{2}}\right) OPT(A). \end{aligned} \quad (3.15)$$

Y como $R_{\eta,f,g}(A) \leq OPT(A)$ por ser la esperanza de soluciones del problema $OPT(A)$:

$$Soundness_{\eta,\tau}(A) \leq OPT(A) \left(1 + O\left(\tau^{\frac{C\eta}{2}}\right) \right). \quad (3.16)$$

Para $\eta > 0$ fijo, podemos encontrar τ tal que la cota sea tan ajustada como queremos. Luego, usando los resultados (3.7) y (3.16), podemos demostrar el Teorema 3.3.5 cuando η y τ sean lo suficientemente chicos.

Vamos a ocuparnos ahora de la reducción principal de este paper, que dice que no puede aproximarse el valor de $OPT(A)$ mejor que el gap de inaproximación de cualquier test de dictadura. Para esto, vamos a asociar los labels asignados a un grafo G con funciones $f \in L_2(\mathbb{H}^k)^{|U|+|V|}$, donde f es como un conjunto de $|U| + |V|$ funciones en $L_2(\mathbb{H}^k)$ indexadas por los nodos U y V .

Dada una biyección $\pi : \{1, \dots, k\} \rightarrow \{1, \dots, k\}$, definimos el operador $(- \circ \pi)$ en $L_2(\mathbb{H}^k)$ tal que

$$(f \circ \pi)(x_1, x_2, \dots, x_n) := f(x_{\pi(1)}, x_{\pi(2)}, \dots, x_{\pi(k)}),$$

para cada punto $(x_1, x_2, \dots, x_n) \in \mathbb{H}^k$. El interés de este operador es que nos permite convertir las restricciones de las aristas de la instancia de Unique Games en operadores sobre las funciones del cubo de Hamming.

Supongamos que tenemos un test de dictadura \mathcal{D} fijo. Para cada instancia G de Unique Games, con condiciones (u, v, π) , vamos a definir un operador $G \otimes_\eta \mathcal{D}$ en $L_2(\mathbb{H}^k)$, mediante el Teorema de Representación de Riesz, de la siguiente manera.

Definición 3.3.10. Definimos a $G \otimes_\eta \mathcal{D} : L_2(\mathbb{H}^k)^{|U|+|V|} \rightarrow L_2(\mathbb{H}^k)^{|U|+|V|}$ como el único operador tal que:

$$\langle f, (G \otimes_\eta \mathcal{D})g \rangle = \mathbb{E}_{(u,v,\pi) \in G} [\langle T_{1-\eta}(f_u \circ \pi), \mathcal{D}T_{1-\eta}g_v \rangle] \quad \forall f, g \in L_2(\mathbb{H}^k)^{|U|+|V|}.$$

Ahora consideremos una base de \mathbb{H}^k , dada por las funciones canónicas $\delta_x \in L_2(\mathbb{H}^k)$, $x \in \mathbb{H}^k$:

$$\delta_{x_0}(x) := \begin{cases} 1 & \text{si } x = x_0 \\ 0 & \text{si } x \neq x_0. \end{cases}$$

Notemos que si expresamos una función $f : \mathbb{H}^k \rightarrow \{-1, 1\}$ en la base $\{\delta_x\}_{x \in \mathbb{H}^k}$, lo que obtenemos es el Long Code de la función f . Podemos definir una base de $L_2(\mathbb{H}^k)^{|U|+|V|}$, usando una base $\{\delta_x\}_{x \in \mathbb{H}^k}$ para cada $u \in U$ y para cada $v \in V$:

$$(\delta_{w_0, x_0})_w(x) := \begin{cases} 1 & \text{si } w = w_0 \text{ y } x = x_0 \\ 0 & \text{si } w \neq w_0 \text{ o } x \neq x_0. \end{cases}$$

Expresemos entonces, en la base $E := \{\delta_{w,x} \mid w \in U \cup V, x \in \mathbb{H}^k\}$, la forma bilineal generada por $G \otimes_\eta \mathcal{D}$:

$$\begin{pmatrix} -f_{u_1} & \cdots & -g_{v_{|V|}} \end{pmatrix} \begin{pmatrix} [G \otimes_\eta \mathcal{D}]_E \end{pmatrix} \begin{pmatrix} f_{u_1} \\ \vdots \\ g_{v_{|V|}} \end{pmatrix}. \quad (3.17)$$

Ahora, está claro que encontrar los Long Codes que maximizan el producto (3.17) equivale a encontrar funciones $f, g \in L_2(\mathbb{H}^k)^{|U|+|V|}$, $\|f\|_\infty, \|g\|_\infty \leq 1$.

Vamos a probar el siguiente teorema, con el que demostramos que la reducción de problema de gap de Unique Game al problema de Grothendieck funciona.

Teorema 3.3.11. *Sea G una instancia de $\mathcal{UG}(M)$, con $\#M = k$, y sea $Val_{LC}(G)$ su valor óptimo, que corresponde a la máxima proporción de aristas simultáneamente satisfacibles. Entonces, para τ, η y ε suficientemente chicos, se cumple que*

$$\text{Si } Val_{LC}(G) > 1 - \varepsilon, \quad \Rightarrow \quad OPT(G \otimes_\eta \mathcal{D}) \geq Completeness(\mathcal{D})(1 - O(\varepsilon + \eta)).$$

$$\text{Si } Val_{LC}(G) < (\tau\eta)^3, \quad \Rightarrow \quad OPT(G \otimes_\eta \mathcal{D}) \leq Soundness_{\eta,\tau}(\mathcal{D}) + \tau\eta Completeness(\mathcal{D}).$$

Demostración. Vamos a demostrar este teorema en dos partes.

Primero veremos que si $Val_{LC}(G) > 1 - \varepsilon$, entonces podemos asignar un Long Code a f y a g tal que $\langle f, (G \otimes_\eta \mathcal{D})g \rangle \geq Completeness(\mathcal{D})(1 - 2\varepsilon - 2\eta) = Completeness(\mathcal{D})(1 - O(\varepsilon + \eta))$. En cambio para probar la segunda parte procederemos por el contrareciproco. Partiendo de una asignación de $\{-1, 1\}$ a f y g tal que $\langle f, (G \otimes_\eta \mathcal{D})g \rangle > Soundness_{\eta,\tau}(\mathcal{D}) + \tau\varepsilon Completeness(\mathcal{D})$, deduciremos la existencia de una asignación de labels en el problema de Label Cover G tal que $Val_{LC}(G) < (\tau\eta)^3$.

Para simplificar las cuentas, vamos a asumir, sin pérdida de generalidad, que en el test de dictadura \mathcal{D} , se tiene $\lambda_1 = 1$, y consiguientemente $\lambda_d \in [-1, 1], \forall 0 \leq d \leq k$. En caso, contrario, se puede reescalar \mathcal{D} , lo que reescala también linealmente los valores $Completeness(\mathcal{D})$ y $Soundness_{\eta,\tau}(\mathcal{D})$.

Comencemos con la primera parte. Supongamos entonces que $Val_{LC}(G) > 1 - \varepsilon$. Esto quiere decir que existe al menos una asignación de labels a U y V tal que satisface al menos $1 - \varepsilon$ de las aristas del Label Cover G . Elijamos una de estas asignaciones. Supongamos que $k_0 \in \{1, \dots, k\}$ es el label que se le asigna al nodo $u \in U$. Luego, en la función $f := (f_w)_{w \in U \cup V}$ vamos a definir $f_{w_0}(x) := d_{k_0}(x) = x_{k_0}$. Es decir, le asignamos a f en la coordenada w_0 la función dictadura de x_{k_0} . Hacemos el proceso análogo para g con los nodos de V . Vamos a ver que, con estas asignaciones, $\langle f, (G \otimes_\eta \mathcal{D})g \rangle > Completeness(\mathcal{D})(1 - 2\varepsilon - 2\eta)$.

Como $\langle f, (G \otimes_\eta \mathcal{D})g \rangle = \mathbb{E}_{(u,v,\pi) \in G} [\langle T_{1-\eta}(f_u \circ \pi), \mathcal{D}T_{1-\eta}g_v \rangle]$, vamos a intentar calcular el valor de $\langle T_{1-\eta}(f_u \circ \pi), \mathcal{D}T_{1-\eta}g_v \rangle$ para un (u, v, π) fijo. Supongamos que (u, v, π) es una de las aristas satisfechas por la asignación elegida. Luego hay dos labels k_u y k_v tales que $\pi(k_u) = k_v$ y, como $f_u = d_{k_u}$ y $g_v = d_{k_v}$, resulta que $f_u \circ \pi_{u,v} = d_{\pi_{u,v}(k_u)} = d_{k_v} = g_v$. Entonces en este caso tendremos que:

$$\begin{aligned} \langle T_{1-\eta}(f_u \circ \pi_{u,v}), \mathcal{D}T_{1-\eta}g_v \rangle &= \langle T_{1-\eta}d_{k_u}, \mathcal{D}T_{1-\eta}d_{k_v} \rangle = \langle (1-\eta)d_{k_u}, \lambda_1(1-\eta)d_{k_v} \rangle \\ &= \lambda_1(1-\eta)^2 = Completeness(\mathcal{D})(1-\eta)^2. \end{aligned}$$

Por otro lado, si la asignación no satisface la arista (u, v, π) , como $\|\mathcal{D}\|, \|T_{1-\eta}\| \leq 1$, entonces podemos acotar:

$$|\langle T_{1-\eta}(f_u \circ \pi_{u,v}), \mathcal{D}T_{1-\eta}g_v \rangle| \leq 1.$$

Lo que nos dice que para el caso general:

$$|\langle T_{1-\eta}(f_u \circ \pi_{u,v}), \mathcal{D}T_{1-\eta}g_v \rangle| \leq Completeness(\mathcal{D}). \quad (3.18)$$

Luego, considerando que la asignación que elegimos satisface al menos $1 - \varepsilon$ de las aristas, tenemos:

$$\begin{aligned} \mathbb{E}_{(u,v,\pi) \in G} [\langle T_{1-\eta}(f_u \circ \pi), \mathcal{D}T_{1-\eta}g_v \rangle] &\geq (1 - \varepsilon) \cdot (1 - \eta)^2 \text{Completeness}(\mathcal{D}) - \varepsilon \cdot \text{Completeness}(\mathcal{D}) \\ &= \text{Completeness}(\mathcal{D})((1 - \varepsilon)(1 - \eta)^2 - \varepsilon) \\ &\geq \text{Completeness}(\mathcal{D})((1 - \varepsilon)(1 - 2\eta + \eta^2) - \varepsilon) \\ &\geq \text{Completeness}(\mathcal{D})(1 - 2\varepsilon - 2\eta). \end{aligned}$$

Ahora para la segunda parte, supongamos que $\langle f, (G \otimes_{\eta} \mathcal{D})g \rangle > \text{Soundness}_{\eta,\tau}(\mathcal{D}) + \tau\varepsilon$ para algún par f, g . La idea para construir una asignación de labels a partir de este hecho es mostrar que, a grandes rasgos, esto sólo puede ocurrir si hay variables en las f_u, g_v con mucha influencia que se corresponden a través de $\pi_{u,v}$, y por lo tanto se “encuentran” en el producto interno que usamos. Por otro lado, la cantidad de variables con mucha influencia para cada f_u o g_v está acotada por la norma de estas funciones en $L_2(\mathbb{H}^k)$, y por lo tanto estas variables que se están correspondiendo son relativamente pocas. En particular, es posible realizar una elección más o menos consistente de una variable por cada función, y esto equivale a realizar una asignación así de consistente para el problema original de Label Cover G .

Vamos a los detalles: Para $u \in U$ y $v \in V$, definimos los “conjuntos de labels tentativos” análogamente en cada caso:

$$J_u := \{i \mid \text{Inf}_i T_{1-\eta}(f_u) > \tau\} \quad J_v := \{j \mid \text{Inf}_j T_{1-\eta}(g_v) > \tau\}.$$

Usaremos el siguiente lema para acotar la cantidad de “labels tentativos” en cada uno de estos conjuntos:

Lema 3.3.12. *Sea $f \in L_2(\mathbb{H}^k)$ y $\eta \in [0, 1]$. Se tiene que:*

$$\sum_{i=1}^k \text{Inf}_i T_{1-\eta} f \leq \frac{\|f\|_{\infty}^2}{\eta}.$$

Usando el lema, resulta que:

$$\#J_u \leq \frac{1}{\tau} \sum_{i=1}^k \text{Inf}_i T_{1-\eta} f \leq \frac{1}{\tau} \frac{\|f\|_{\infty}^2}{\eta} \leq \frac{1}{\tau\eta}.$$

Lo mismo vale para $\#J_v$.

Supongamos que para algún $(u, v, \pi_{u,v}) \in G$ fijo se tiene que

$$\langle T_{1-\eta}(f_u \circ \pi_{u,v}), \mathcal{D}T_{1-\eta}g_v \rangle > \text{Soundness}_{\eta,\tau}(\mathcal{D}).$$

Entonces por la definición de $\text{Soundness}_{\eta,\tau}$ de un test de dictadura, resulta que no puede ser que $\text{MaxComInf}(T_{1-\eta}f_u \circ \pi_{u,v}, T_{1-\eta}g_v) \leq \tau$. Esto quiere decir que existe $k_v \in \{1, \dots, k\}$ tal que

$$\text{Inf}_{k_v} T_{1-\eta}(f_u \circ \pi_{u,v}), \text{Inf}_{k_v} T_{1-\eta}g_v > \tau.$$

Pero por la definición de $f_u \circ \pi_{u,v}$,

$$\text{Inf}_{k_v} T_{1-\eta}(f_u \circ \pi_{u,v}) = \text{Inf}_{k_u} T_{1-\eta}(f_u), \text{ con algún } k_u \in \{1, \dots, k\} \text{ tal que } \pi_{u,v}(k_u) = k_v.$$

En particular, todo esto nos dice que, en el caso de que

$$\langle T_{1-\eta}(f_u \circ \pi_{u,v}), \mathcal{D}T_{1-\eta}g_v \rangle > Soundness_{\eta,\tau}(\mathcal{D}),$$

resulta que

$$\exists k_u \in J_u, k_v \in J_v \text{ tales que } \pi_{u,v}(k_u) = k_v. \quad (3.19)$$

Por otro lado, por (3.18), tenemos una cota superior para todos estos productos.

Vamos a ver en cuántas aristas $(u, v, \pi_{u,v})$ podemos garantizar que existen conjuntos J_u, J_v apareables en el sentido (3.19). Para esto, vamos a ver cuántos pares podemos garantizar que tienen valor $\langle T_{1-\eta}(f_u \circ \pi_{u,v}), \mathcal{D}T_{1-\eta}g_v \rangle > Soundness_{\eta,\tau}(\mathcal{D})$. Supongamos que

$$\mathbb{P}_{(u,v,\pi) \in G} \{ \langle T_{1-\eta}(f_u \circ \pi_{u,v}), \mathcal{D}T_{1-\eta}g_v \rangle > Soundness_{\eta,\tau}(\mathcal{D}) \} < \delta.$$

Luego

$$\begin{aligned} & \mathbb{E}_{(u,v,\pi) \in G} [\langle T_{1-\eta}(f_u \circ \pi), \mathcal{D}T_{1-\eta}g_v \rangle] \leq \\ & \leq \underbrace{\mathbb{P}_{(u,v,\pi) \in G} \{ \langle T_{1-\eta}(f_u \circ \pi_{u,v}), \mathcal{D}T_{1-\eta}g_v \rangle > Soundness_{\eta,\tau}(\mathcal{D}) \}}_{< \delta} \cdot 1 \\ & + \underbrace{\mathbb{P}_{(u,v,\pi) \in G} \{ \langle T_{1-\eta}(f_u \circ \pi_{u,v}), \mathcal{D}T_{1-\eta}g_v \rangle \leq Soundness_{\eta,\tau}(\mathcal{D}) \}}_{\leq 1} \cdot Soundness_{\eta,\tau}(\mathcal{D}) \\ & \leq \delta + Soundness_{\eta,\tau}(\mathcal{D}). \end{aligned}$$

Esto quiere decir que si el valor alcanzado por un f y g es mayor a $\delta + Soundness_{\eta,\tau}(\mathcal{D})$, luego la asignación que se derivará de ellos tendrá al menos δ de las aristas que cumplen la condición (3.19), a las que llamaremos “buenas”.

Elijamos entonces $\delta := \tau\eta$. En ese caso, tenemos al menos $\tau\eta$ de aristas buenas. Para cada nodo $w \in U \cup V$, elijamos un label al azar $k_w \in J_w$ (si $J_w = \emptyset$, elijamos un label $k_w \in \{1, \dots, k\}$ cualquiera). Supongamos que $(u, v, \pi_{u,v})$ es una arista buena. Acotemos desde abajo la probabilidad de satisfacer esa arista con nuestra asignación:

$$\mathbb{P}_{\{k_w | k_w \in J_w\}} \{ \pi_{u,v}(k_u) = k_v \} \geq \frac{1}{\#J_u} \frac{1}{\#J_v} \geq (\tau\eta)^2.$$

Notemos $Val_{LC}(G)(\{k_w\})$ al valor alcanzado en el problema de Label Cover G por un labeling $\{k_w\}_{w \in U \cup V}$. Como sabemos que por lo menos $\tau\eta$ de las aristas son buenas, resulta que:

$$\begin{aligned} \mathbb{E}_{\{k_w | k_w \in J_w\}} [Val_{LC}(G)(\{k_w\})] &= \mathbb{E}_{\{k_w | k_w \in J_w\}} \left[\mathbb{E}_{(u,v,\pi) \in G} \left[\mathbb{1}_{\{ \pi_{u,v}(k_u) = k_v \}} \right] \right] = \\ &= \mathbb{E}_{(u,v,\pi) \in G} \left[\mathbb{E}_{\{k_w | k_w \in J_w\}} \left[\mathbb{1}_{\{ \pi_{u,v}(k_u) = k_v \}} \right] \right] = \\ &= \mathbb{E}_{(u,v,\pi) \in G} \left[\mathbb{P}_{\{k_w | k_w \in J_w\}} \{ \pi_{u,v}(k_u) = k_v \} \right] \geq \\ &\geq \tau\eta \cdot (\tau\eta)^2 = (\tau\eta)^3. \end{aligned}$$

Es decir, el valor en esperanza de nuestra asignación del Label Cover G será $\geq (\tau\eta)^3$. Esto significa que dentro de nuestras asignaciones posibles hay al menos una que tiene valor $\geq (\tau\eta)^3$. Por lo tanto, el Label Cover G tiene una asignación que satisface $\geq (\tau\eta)^3$ de las aristas, y por lo tanto $Val_{LC}(G) \geq (\tau\eta)^3$. El contrareciproco nos dice, como queríamos, que si $Val_{LC}(G) < (\tau\eta)^3$, entonces $OPT(G \otimes_{\eta} \mathcal{D}) \leq Soundness_{\eta,\tau}(\mathcal{D}) + \tau\eta$. \square

Notemos, por último, que esta reducción se puede realizar en tiempo polinómico. La matriz de $G \otimes_{\eta} \mathcal{D}$ tiene dimensión $(|U| + |V|)2^k \times (|U| + |V|)2^k$, pero k es una constante (la cantidad de labels). Lo que crece con el tamaño del grafo es $|U| + |V|$, mientras que el tamaño de la matriz sólo crece cuadráticamente en relación a esta variable. Como las cuentas que hacemos para generar la matriz toman tiempo polinomial para cada coeficiente, entonces la reducción a la matriz es efectivamente polinomial.

Ahora sí, finalmente estamos en condiciones de probar el Teorema 3.1.3, el Teorema central del capítulo.

Demostración del Teorema 3.1.3. Veamos lo que nos dice el Teorema 3.3.11.

Para cualquier $\mathcal{D} : L_2(\mathbb{H}^k) \rightarrow L_2(\mathbb{H}^k)$ test de dictaduras, tenemos una reducción polinomial del problema $\mathcal{UG}_{1-\varepsilon,(\tau\eta)^3}^{gap}(M)$, con $\#M = k$, a un problema de gap del problema de Grothendieck $OPT(A)$. Más precisamente, la reducción es al problema de decidir para una matriz $A \in \mathbb{R}^{n \times n}$ entre estas dos posibilidades:

$$\begin{aligned} OPT(A) &\geq Completeness(\mathcal{D})(1 - O(\varepsilon + \eta)) \\ OPT(A) &\leq Soundness_{\eta,\tau}(\mathcal{D}) + \tau\eta Completeness(\mathcal{D}). \end{aligned} \quad (3.20)$$

Como podemos tomar ε, τ y η arbitrariamente chicos, si asumimos la UGC, entonces los podemos tomar tan chicos que el problema $\mathcal{UG}_{1-\varepsilon,(\tau\eta)^3}^{gap}(M)$ sea NP-hard para algún tamaño suficientemente grande de M .

Luego, para ε, τ y η suficientemente chicos y k suficientemente grande, y para cualquier $\mathcal{D} : L_2(\mathbb{H}^k) \rightarrow L_2(\mathbb{H}^k)$ test de dictaduras, distinguir entre los casos (3.20) para toda matriz A es NP-hard.

Por el Teorema 3.3.5, para cualquier matriz Q y cualquier $k \in \mathbb{N}$, existe un test de dictaduras $\mathcal{D}(Q)$ tal que

$$\begin{aligned} Completeness(\mathcal{D}(Q)) &= SDP(Q) \\ Soundness_{\eta,\tau}(\mathcal{D}(Q)) &\leq OPT(Q)(1 + \eta). \end{aligned}$$

Usando el test de dictadura $\mathcal{D}(Q)$ en (3.20), obtenemos que es NP-hard distinguir entre

$$\begin{aligned} OPT(A) &\geq SDP(Q)(1 - O(\varepsilon + \eta)) \\ OPT(A) &\leq OPT(Q)(1 + \eta) + \tau\eta SDP(Q). \end{aligned}$$

Y por lo tanto, por el Lema 3.2.2, el problema $OPT(A)$ no se puede aproximar con una constante mejor que

$$\frac{OPT(Q)(1 + \eta) + \tau\eta SDP(Q)}{SDP(Q)(1 - O(\varepsilon + \eta))} OPT(A) = \left(\frac{OPT(Q)}{SDP(Q)} (1 + \eta) + \tau\eta \right) \frac{1}{1 - O(\varepsilon + \eta)} OPT(A).$$

Como esto vale para toda matriz Q , en particular vale para matrices tales que $\frac{OPT(Q)}{SDP(Q)}$ está arbitrariamente cerca de $\frac{1}{K_G}$. Como además vale para η, τ y ε suficientemente chicos, entonces la expresión puede acercarse tanto a $\frac{1}{K_G}OPT(A)$ como queramos.

Luego aproximar $OPT(A)$ con una constante mejor que K_G es NP-hard, como queríamos demostrar. \square

3.4. Aproximación de la constante de Grothendieck

Nuestra demostración de la sección anterior de que K_G es el gap de integralidad del problema de Grothendieck se basa en una reducción que traslada gaps. Más precisamente, partimos de una matriz A con gap de integralidad $\frac{SDP(A)}{OPT(A)}$, y logramos con nuestra reducción construir un test de dictadura $\mathcal{D}(A)$ tal que básicamente

$$\frac{SDP(A)}{OPT(A)} \approx \frac{Completeness(\mathcal{D}(A))}{Soundness_{\eta,\tau}(\mathcal{D}(A))}.$$

En esta sección vamos a proceder al revés, vamos a contar que dado un test de dictadura \mathcal{D} , como los definimos en (3.3.1), podemos construir una instancia del problema de Grothendieck $\mathcal{G}(\mathcal{D})$ tal que

$$\frac{SDP(\mathcal{G}(\mathcal{D}))}{OPT(\mathcal{G}(\mathcal{D}))} \approx \frac{Completeness(\mathcal{D})}{Soundness_{\eta,\tau}(\mathcal{D})}.$$

La pequeña diferencia es el espacio donde va a vivir el operador $\mathcal{G}(\mathcal{D})$. En la sección anterior, trabajamos siempre con matrices $A \in \mathbb{R}^{n \times n}$. En cambio, ahora vamos a pasarnos al contexto continuo que habíamos descrito brevemente en (2.3.5). A tal efecto, definimos, dado un test de dictadura \mathcal{D} , un operador en el espacio de $L_2(\mathcal{G}^k)$. La idea va a ser “copiar” el funcionamiento del test \mathcal{D} en el espacio $L_2(\mathbb{H}^k)$ al espacio $L_2(\mathcal{G}^k)$. Para esto vamos a necesitar el operador de ruido Gaussiano.

El espacio $L_2(\mathcal{G}^k)$ tiene una base ortonormal dada por los productos de polinomios de Hermite en cada variable. Es decir que

$$\left\{ H_\alpha(x) := \prod_{i=1}^k H_{\alpha_i}(x_i) \mid \alpha \in \mathbb{N}_0^k \right\}$$

es una base ortonormal de $L_2(\mathcal{G}^k)$. Usando esta base, podemos definir el siguiente operador.

Definición 3.4.1. Dado $d \in \mathbb{N}_0$, se define la proyección $Q_d : L_2(\mathcal{G}^k) \rightarrow L_2(\mathcal{G}^k)$ como la proyección al subespacio

$$H^{(d)} := \left\{ H_\alpha(x) := \prod_{i=1}^k H_{\alpha_i}(x_i) \mid \alpha \in \mathbb{N}_0^k, \sum_{i=1}^k \alpha_i = d \right\}.$$

Y luego usamos estos operadores para definir el operador de ruido Gaussiano, de manera prácticamente análoga a (3.2.15).

Definición 3.4.2. Sea $\rho \in (0, 1)$. Definimos el operador de ruido Gaussiano U_ρ para el espacio $L_2(\mathcal{G}^k)$ como

$$U_\rho := \sum_{d=0}^k \rho^d Q_d.$$

Definición 3.4.3. Dado un test de dictadura $\mathcal{D} : L_2(\mathbb{H}^k) \rightarrow L_2(\mathbb{H}^k)$ de la forma

$$\mathcal{D} = \sum_{d=0}^k \lambda_d P_d,$$

definimos el operador $\mathcal{G}(\mathcal{D}) : L_2(\mathcal{G}^k) \rightarrow L_2(\mathcal{G}^k)$ como

$$\mathcal{G}(\mathcal{D}) = \sum_{d=0}^k \lambda_d Q_d,$$

y el operador $\mathcal{G}_\eta(\mathcal{D}) : L_2(\mathcal{G}^k) \rightarrow L_2(\mathcal{G}^k)$ como

$$\mathcal{G}_\eta(\mathcal{D}) = U_{1-\eta} \mathcal{G}(\mathcal{D}) U_{1-\eta} = \sum_{d=0}^k \lambda_d (1-\eta)^2 Q_d.$$

Con estos operadores $\mathcal{G}(\mathcal{D})$ o $\mathcal{G}_\eta(\mathcal{D})$ en el espacio de $L_2(\mathcal{G}^k)$, hallar una solución al problema $SDP(\mathcal{G}(\mathcal{D}))$ corresponde a dar dos funciones $f, g : \mathcal{G}^k \rightarrow \mathcal{H}$ para algún Hilbert tal que las imágenes de f, g sean siempre vectores unitarios. Por otro lado, construir una solución al problema $OPT(\mathcal{G}(\mathcal{D}))$ corresponde a dar dos funciones $f, g \in L_2(\mathcal{G}^k)$ con $\|f\|_\infty, \|g\|_\infty \leq 1$. La versión de la desigualdad de Grothendieck (2.3.5) nos dice efectivamente que el gap entre estos dos valores es a lo sumo K_G :

$$\frac{SDP(\mathcal{G}(\mathcal{D}))}{OPT(\mathcal{G}(\mathcal{D}))} \leq K_G.$$

Ahora sí, vamos a usar el siguiente teorema, que prueba cómo se traslada el gap del inaproximación del test de dictadura al gap de integralidad del operador que creamos.

Teorema 3.4.4. Para todo $\eta > 0$, existen k, τ tales que para cualquier test de dictadura $\mathcal{D} \in L_2(\mathbb{H}^k)$ se cumple que

$$\begin{aligned} SDP(\mathcal{G}_\eta(\mathcal{D})) &\geq \text{Completeness}(\mathcal{D})(1 - 5\eta), \\ OPT(\mathcal{G}_\eta(\mathcal{D})) &\leq \text{Soundness}_{\eta, \tau}(\mathcal{D})(1 + \eta) + \eta \text{Completeness}(\mathcal{D}). \end{aligned}$$

Idea de la demostración. Para la cota de $SDP(\mathcal{G}_\eta(\mathcal{D})) \geq \text{Completeness}(\mathcal{D})(1 - 5\eta)$, el objetivo es encontrar funciones f, g que den un valor cercano a $\text{Completeness}(\mathcal{D}) = \lambda_1$ en la evaluación del problema $SDP(\mathcal{G}_\eta(\mathcal{D}))$. Para esto, notemos que esto se lograría si pudiéramos usar funciones $f, g : \mathbb{R}^k \rightarrow \mathcal{H}$ donde cada coordenada sea una función lineal (ya que estas son las funciones que el operador $\mathcal{G}_\eta(\mathcal{D})$ multiplica por $\lambda_1(1 - \eta)^2$).

Obviamente, no se puede tener una función $f, g : \mathbb{R}^k \rightarrow \mathcal{H}$ lineal cuya imagen sean todos vectores unitarios. Pero sí podemos usar las funciones $f : \mathbb{R}^k \rightarrow \mathbb{R}^k$ del tipo $f(x) = g(x) := \frac{x}{\|x\|}$ (como hicimos en la demostración de la cota de Reeds (1.4)), que a medida que $k \rightarrow \infty$ se

parece mucho a la función lineal $\tilde{f}(x) = \tilde{g}(x) = \frac{x}{\sqrt{k}}$. Acotando la diferencia entre las soluciones del problema $SDP(\mathcal{G}_\eta(\mathcal{D}))$ para f, g y \tilde{f}, \tilde{g} por $\|f - \tilde{f}\|$ y $\|g - \tilde{g}\|$ se puede demostrar que, para $k > \frac{1}{\eta}$ suficientemente grande, se cumple que

$$SDP(\mathcal{G}_\eta(\mathcal{D})) \geq \text{Completeness}(\mathcal{D})(1 - 5\eta).$$

La demostración de que $OPT(\mathcal{G}_\eta(\mathcal{D})) \leq \text{Soundness}_{\eta, \tau}(\mathcal{D})(1 + \eta) + \eta \text{Completeness}(\mathcal{D})$ es, nuevamente, mucho más delicada.

Supongamos que esto fuera falso. Luego existiría un par de funciones $f, g \in L_2(\mathcal{G}^k)$, con $\|f\|_\infty, \|g\|_\infty \leq 1$ tales que

$$\langle f, \mathcal{G}_\eta(\mathcal{D})g \rangle > \text{Soundness}_{\eta, \tau}(\mathcal{D})(1 + \eta) + \eta \text{Completeness}(\mathcal{D}).$$

La idea es ver que entonces se pueden construir dos funciones $f'', g'' \in L_2(\mathbb{H}^k)$ de influencia máxima menor a τ tales que

$$\langle T_{1-\eta}f'', \mathcal{D}T_{1-\eta}g'' \rangle > \text{Soundness}_{\eta, \tau}(\mathcal{D}).$$

Esto contradice la definición de Soundness para un test de dictadura (3.3.2).

Vamos a describir brevemente la construcción de las funciones f'', g'' . El primer paso es tratar de bajarle la influencia a las funciones f y g . Notemos que las funciones f y g están en $L_2(\mathcal{G}^k)$, mientras que la definición de influencia la habíamos dado para funciones en $L_2(\mathbb{H}^k)$. Sin embargo, hay una extensión natural de la definición que se lleva bien con los teoría de invarianza estilo [MOO10]. El método para bajar la influencia de la función es “inventar” coordenadas nuevas para distribuir la influencia de cada coordenada. Se definen entonces $f', g' : L_2(\mathcal{G}^{kR})$ de la siguiente manera:

$$\begin{aligned} f'(x) &:= f \left(\frac{1}{\sqrt{R}} \sum_{i=1}^R x_i, \frac{1}{\sqrt{R}} \sum_{i=R+1}^{2R} x_i, \dots, \frac{1}{\sqrt{R}} \sum_{i=(k-1)R+1}^{kR} x_i \right) \\ g'(x) &:= f \left(\frac{1}{\sqrt{R}} \sum_{i=1}^R x_i, \frac{1}{\sqrt{R}} \sum_{i=R+1}^{2R} x_i, \dots, \frac{1}{\sqrt{R}} \sum_{i=(k-1)R+1}^{kR} x_i \right). \end{aligned}$$

Por último, la idea es construir dos funciones $f'', g'' \in L_2(\mathbb{H}^k)$ que mantengan las propiedades de las funciones f', g' . Para esto se usa una variación de los lemas de invarianza de [MOO10], con técnicas muy profundas, para pasar de una “realización” de las funciones f', g' en el espacio de los polinomios de Hermite multivariados a una “realización” en el espacio de polinomios multilineales en \mathbb{H}^k . Si k es suficientemente grande, se obtienen así funciones f'', g'' con

$$\|f''\|_\infty, \|g''\|_\infty \leq 1 \quad \text{MaxInf}(T_{1-\eta}f''), \text{MaxInf}(T_{1-\eta}g'') \leq \tau,$$

y que cumplen

$$\begin{aligned} \langle T_{1-\eta}f'', \mathcal{D}T_{1-\eta}g'' \rangle &\geq \langle f', \mathcal{G}(\mathcal{D}g') - \eta\|\mathcal{D}\| \rangle = \langle f, \mathcal{G}(\mathcal{D}g) - \eta\|\mathcal{D}\| \rangle \\ &= \text{Soundness}_{\eta, \tau}(\mathcal{D})(1 + \eta) + \underbrace{\eta \text{Completeness}(\mathcal{D})}_{=\lambda_1} - \underbrace{\eta\|\mathcal{D}\|}_{=\lambda_1} \\ &= \text{Soundness}_{\eta, \tau}(\mathcal{D})(1 + \eta). \end{aligned}$$

Y esto era lo que queríamos ver. □

Vamos a ver cómo usar este resultado para conseguir, para cualquier $\varepsilon > 0$ un algoritmo aleatorio para el problema $OPT(A)$ que alcance en esperanza un valor de

$$\left(\frac{1}{K_G} - \varepsilon\right) SDP(A).$$

Empecemos por el siguiente corolario del Teorema (3.4.4).

Corolario 3.4.5. *Para cualquier $\varepsilon > 0$ y cualquier $\eta > 0$ suficientemente chico, existen k, τ tales que para cualquier test de dictadura $\mathcal{D} \in L_2(\mathbb{H}^k)$ se cumple*

$$\frac{Soundness_{\eta, \tau}(\mathcal{D})}{Completeness(\mathcal{D})} \geq \frac{1}{K_G} - \varepsilon.$$

Demostración. Es sólo cuestión de dividir las expresiones de (3.4.4), recordar que

$$\frac{OPT(\mathcal{G}_\eta(\mathcal{D}))}{SDP(\mathcal{G}_\eta(\mathcal{D}))} \leq \frac{1}{K_G},$$

y tomar $\eta > 0$ lo suficientemente chico para que se cumpla la desigualdad. \square

Ahora, usando esto, vamos a describir nuestro nuevo algoritmo para aproximar el problema $OPT(A)$ en una matriz $A \in \mathbb{R}^{n \times n}$.

- Comenzamos, como siempre, hallando una solución $\{u_i\}_i, \{v_j\}_j \subset \mathbb{S}^{2n-1}$ para el problema $SDP(A)$ con programación semidefinida.
- Elegimos η y τ suficientemente chicos y k suficientemente grande (a especificar más tarde).
- Para cada función $f, g \in L_2(\mathbb{H}^k)$, $\|f\|, \|g\| \leq 1$ corremos el algoritmo de redondeo $R_{\eta, f, g}(A)$, y nos quedamos con la mejor solución de $OPT(A)$ entre todas.

Con este algoritmo tenemos

$$\begin{aligned} \sup_{\substack{f, g \in L_2(\mathbb{H}^k) \\ \|f\|, \|g\| \leq 1}} R_{\eta, f, g}(A) &\geq \sup_{\substack{f, g \in L_2(\mathbb{H}^k) \\ MaxComInf(T_{1-\eta}f, T_{1-\eta}g) \leq \tau}} R_{\eta, f, g}(A) \\ &\stackrel{(3.15)}{\geq} Soundness_{\eta, \tau}(\mathcal{D}(A)) - O\left(\tau^{\frac{c_\eta}{2}}\right) SDP(A) \\ &\stackrel{(3.4.5)}{\geq} \left(\frac{1}{K_G} - \varepsilon\right) Completeness(\mathcal{D}(A)) - O\left(\tau^{\frac{c_\eta}{2}}\right) SDP(A) \\ &= \left(\frac{1}{K_G} - \varepsilon\right) SDP(A) - O\left(\tau^{\frac{c_\eta}{2}}\right) SDP(A) \\ &= \left(\left(\frac{1}{K_G} - \varepsilon\right) - O\left(\tau^{\frac{c_\eta}{2}}\right)\right) SDP(A). \end{aligned}$$

Con elecciones adecuadas de ε, η, τ y k , en ese orden, podemos conseguir una solución que alcance un valor de al menos

$$\left(\frac{1}{K_G} - \varepsilon'\right) SDP(A).$$

Por supuesto, no podemos usar todas las funciones $f, g \in L_2(\mathbb{H}^k)$, $\|f\|, \|g\| \leq 1$ tal como sugerimos. Pero se puede arreglar definiendo de antemano una red de funciones suficientemente densa en la bola unitaria de $L_2(\mathbb{H}^k)$, ya que se puede demostrar que si los pares (f, g) y (f', g') están suficientemente cerca en norma, entonces también están cerca (uniformemente) los valores de $R_{\eta, f, g}(A)$ y $R_{\eta, f', g'}(A)$.

Una vez definidas todas las constantes y construídas la red de funciones, este algoritmo corre polinomialmente en el tamaño de la matriz A . Lamentablemente el costo de conseguir error ε con este algoritmo agrega una constante de $e^{o(\frac{1}{\varepsilon^3})}$, lo que lo hace impráctico para aplicaciones reales.

Sin embargo, una consecuencia interesante de este algoritmo es que nos da un algoritmo para calcular la constante K_G , aunque con los mismos problemas de lentitud a la hora de conseguir un error pequeño. Vamos a describir este algoritmo.

Para esto, notemos que el hecho de haber demostrado que el gap del problema de Grothendieck es exactamente igual al gap de inaproximación de los test de dictadura (3.3.1), nos permite calcular la constante de Grothendieck iterando sobre tests de dictadura en lugar de iterar sobre las matrices. En particular, en nuestro último algoritmo vimos que para garantizar una aproximación $(\frac{1}{K_G} - \varepsilon')$, se puede fijar k , que tiene que ser suficientemente alto, pero puede quedar fijo.

Luego, con ese k elegido, tenemos

$$\begin{aligned}
\left(\frac{1}{K_G} - \varepsilon'\right) &\leq \inf_{\substack{\mathcal{D} \text{ test de dictadura} \\ \text{Completeness}(\mathcal{D})=1}} \frac{\text{Soundness}_{\eta, \tau}(\mathcal{D})}{\text{Completeness}(\mathcal{D})} && \left(\leq \frac{1}{K_G}\right) \\
&\leq \inf_{\substack{\mathcal{D} \text{ test de dictadura} \\ \lambda_1=1}} \frac{1}{\lambda_1} \sup_{\substack{f, g \in L_2(\mathbb{H}^k) \\ \|f\|_\infty, \|g\|_\infty \leq 1 \\ \text{MaxComInf}(T_{1-\eta}f, T_{1-\eta}g) \leq \tau}} \langle T_{1-\eta}f, \mathcal{D}T_{1-\eta}g \rangle \\
&\leq \inf_{\substack{\lambda_1=1 \\ |\lambda_d| \leq 1 \forall d}} \sup_{\substack{f, g \in L_2(\mathbb{H}^k) \\ \|f\|_\infty, \|g\|_\infty \leq 1 \\ \text{MaxComInf}(T_{1-\eta}f, T_{1-\eta}g) \leq \tau}} \left\langle f, \sum_{d=0}^k \lambda_d (1-\eta)^{2d} P_d g \right\rangle \\
&\leq \inf_{\substack{\lambda_1=1 \\ |\lambda_d| \leq 1 \forall d}} \sup_{\substack{f, g \in L_2(\mathbb{H}^k) \\ \|f\|_\infty, \|g\|_\infty \leq 1 \\ \text{MaxComInf}(T_{1-\eta}f, T_{1-\eta}g) \leq \tau}} \sum_{d=0}^k \lambda_d (1-\eta)^{2d} \langle f, P_d g \rangle. \\
&\leq \sup_{\substack{f, g \in L_2(\mathbb{H}^k) \\ \|f\|_\infty, \|g\|_\infty \leq 1 \\ \text{MaxComInf}(T_{1-\eta}f, T_{1-\eta}g) \leq \tau}} \inf_{\substack{\lambda_1=1 \\ |\lambda_d| \leq 1 \forall d}} \sum_{d=0}^k \lambda_d (1-\eta)^{2d} \langle f, P_d g \rangle.
\end{aligned}$$

Nuevamente podemos definir una red de funciones

$$f, g \in L_2(\mathbb{H}^k), \quad \|f\|_\infty, \|g\|_\infty \leq 1, \quad \text{MaxComInf}(T_{1-\eta}f, T_{1-\eta}g) \leq \tau,$$

que sea suficientemente densa, para aproximar el supremo tanto como deseemos evaluando sólo en finitas funciones. Para cada una de estas funciones de la red, lo que queremos es hallar el valor mínimo del programa lineal:

$$\begin{cases} \sum_{d=0}^k \lambda_d (1 - \eta)^{2d} \langle f, P_d g \rangle \\ \lambda_1 = 1 \\ |\lambda_d| \leq 1 \quad \forall 0 \leq d \leq k \end{cases} .$$

Para nuestra aproximación de $\frac{1}{K_G}$ basta entonces con tomar el máximo sobre el resultado de todos los programas lineales para todas las f, g de la red.

Aplicaciones

La Cut-Norm de una matriz

Vamos a definir la Cut-Norm de una matriz $A \in \mathbb{R}^{m \times n}$.

Definición 4.1.6. Dada una matriz $A \in \mathbb{R}^{m \times n}$, definimos la Cut-Norm de la matriz A como

$$\|A\|_c := \max_{\substack{S \subseteq \{1, \dots, m\} \\ T \subseteq \{1, \dots, n\}}} \left| \sum_{\substack{i \in S \\ j \in T}} a_{ij} \right|.$$

Es decir, $\|A\|_c$ es el máximo valor en módulo de suma de coeficientes alcanzado por un menor de una matriz A .

En general, nos interesará tanto el problema de hallar el valor $\|A\|_c$ como el de hallar los conjuntos $S \subseteq \{1, \dots, m\}, T \subseteq \{1, \dots, n\}$ que realizan ese máximo. Este es un problema natural que surge en distintos contextos algorítmicos, que analizaremos brevemente en la próxima sección.

Veamos la relación de este problema con el problema de hallar $OPT(A)$ que estuvimos estudiando en todo este trabajo.

Lema 4.1.7. Sea $A \in \mathbb{R}^{m \times n}$ una matriz. Luego

$$\|A\|_c \leq OPT(A) \leq 4\|A\|_c.$$

Demostración. Recordemos que el problema de hallar $OPT(A)$ para una matriz A equivale a hallar signos para multiplicar las filas y las columnas de A de forma que se maximice la suma de sus coeficientes. Cada asignación de signos a las filas y las columnas de A nos induce una bipartición de las filas y las columnas. Supongamos que $\{\delta_i\}, \{\epsilon_j\} \subseteq \{-1, 1\}$ es una solución maximal del problema $OPT(A)$. Denotemos:

$$\begin{aligned} S_+ &:= \{i \in \{1, \dots, m\} \mid \delta_i = 1\} \\ S_- &:= \{i \in \{1, \dots, m\} \mid \delta_i = -1\} \\ T_+ &:= \{j \in \{1, \dots, n\} \mid \epsilon_j = 1\} \\ T_- &:= \{j \in \{1, \dots, n\} \mid \epsilon_j = -1\}. \end{aligned}$$

Luego tenemos que

$$\begin{aligned}
OPT(A) &= \sum_{i,j=1}^{m,n} a_{ij} \delta_i \epsilon_j = \\
&= \sum_{\substack{i \in S_+ \\ j \in T_+}} a_{ij} - \sum_{\substack{i \in S_- \\ j \in T_+}} a_{ij} - \sum_{\substack{i \in S_+ \\ j \in T_-}} a_{ij} + \sum_{\substack{i \in S_- \\ j \in T_-}} a_{ij} \leq \\
&\leq 4\|A\|_c.
\end{aligned}$$

Por otro lado, supongamos que tenemos conjuntos $S_0 \subseteq \{1, \dots, m\}$, $T_0 \subseteq \{1, \dots, n\}$ tales que la Cut-Norm de la matriz A se maximiza, sacando el módulo sin pérdida de generalidad, en S_0 y T_0 .

$$\|A\|_c = \sum_{\substack{i \in S \\ j \in T}} a_{ij}.$$

Sean S_0^c y T_0^c los complementos de S_0 y T_0 en $\{1, \dots, m\}$ y $\{1, \dots, n\}$, respectivamente. Vamos a usar estos conjuntos para conseguir asignaciones de signos para el problema $OPT(A)$. Para $i \in S_0$, $j \in T_0$ y $k \in \{1, 2, 3, 4\}$ vamos a usar $\delta_i^k = \epsilon_j^k = 1$, mientras que para $i \in S_0^c$ y $j \in T_0^c$ vamos a usar las 4 combinaciones posibles:

$$\begin{aligned}
\delta_i^1 &= 1, \epsilon_j^1 = 1 & \delta_i^2 &= -1, \epsilon_j^2 = 1 \\
\delta_i^3 &= 1, \epsilon_j^3 = -1 & \delta_i^4 &= -1, \epsilon_j^4 = -1.
\end{aligned}$$

Para cualquier $k \in \{1, 2, 3, 4\}$ tenemos que

$$OPT(A) \geq \sum_{i,j=1}^{m,n} a_{ij} \delta_i^k \epsilon_j^k.$$

Luego sumandos sobre los 4 conjuntos de signos:

$$\begin{aligned}
OPT(A) &\geq \frac{1}{4} \sum_{k=1}^4 \sum_{i,j=1}^{m,n} a_{ij} \delta_i^k \epsilon_j^k = \frac{1}{4} \sum_{k=1}^4 \left(\sum_{\substack{i \in S_0 \\ j \in T_0}} a_{ij} + \sum_{\substack{i \in S_0^c \\ j \in T_0}} a_{ij} \delta_i^k \epsilon_j^k + \sum_{\substack{i \in S_0 \\ j \in T_0^c}} a_{ij} \delta_i^k \epsilon_j^k + \sum_{\substack{i \in S_0^c \\ j \in T_0^c}} a_{ij} \delta_i^k \epsilon_j^k \right) = \\
&= \sum_{\substack{i \in S_0 \\ j \in T_0}} a_{ij} + \frac{1}{4} \sum_{k=1}^4 \left(\sum_{\substack{i \in S_0^c \\ j \in T_0}} a_{ij} \delta_i^k \epsilon_j^k \right) + \frac{1}{4} \sum_{k=1}^4 \left(\sum_{\substack{i \in S_0 \\ j \in T_0^c}} a_{ij} \delta_i^k \epsilon_j^k \right) + \frac{1}{4} \sum_{k=1}^4 \left(\sum_{\substack{i \in S_0^c \\ j \in T_0^c}} a_{ij} \delta_i^k \epsilon_j^k \right) = \\
&= \sum_{\substack{i \in S_0 \\ j \in T_0}} a_{ij} + \frac{1}{4} \sum_{\substack{i \in S_0^c \\ j \in T_0}} a_{ij} \left(\underbrace{\sum_{k=1}^4 \delta_i^k \epsilon_j^k}_{=0} \right) + \frac{1}{4} \sum_{\substack{i \in S_0 \\ j \in T_0^c}} a_{ij} \left(\underbrace{\sum_{k=1}^4 \delta_i^k \epsilon_j^k}_{=0} \right) + \frac{1}{4} \sum_{\substack{i \in S_0^c \\ j \in T_0^c}} a_{ij} \left(\underbrace{\sum_{k=1}^4 \delta_i^k \epsilon_j^k}_{=0} \right) = \\
&= \sum_{\substack{i \in S_0 \\ j \in T_0}} a_{ij} = \|A\|_c.
\end{aligned}$$

□

Es más, podemos ver que, en un caso particular, la desigualdad superior es ajustada.

Lema 4.1.8. *Sea $A \in \mathbb{R}^{m \times n}$ una matriz tal que la suma de los coeficientes de cada fila y la suma de los coeficientes de cada columna da cero. Luego*

$$\|A\|_c = \frac{1}{4}OPT(A).$$

Demostración. Ya tenemos del paso anterior la desigualdad $\|A\|_c \geq \frac{1}{4}OPT(A)$. Veamos la desigualdad inversa.

Sean S_0 y T_0 conjuntos en donde se alcanza la Cut-Norm de A , con signo positivo, sin pérdida de generalidad. Definamos entonces un conjunto de signos

$$\delta_i := \begin{cases} 1 & i \in S_0 \\ -1 & i \notin S_0 \end{cases} \quad \epsilon_j := \begin{cases} 1 & j \in T_0 \\ -1 & j \notin T_0. \end{cases}$$

Luego tenemos que

$$\begin{aligned} \|A\|_c &= \sum_{i,j=1}^n a_{ij} \frac{1 + \delta_i}{2} \frac{1 + \epsilon_j}{2} = \\ &= \frac{1}{4} \left(\sum_{i,j=1}^{m,n} a_{ij} \delta_i \epsilon_j + \sum_{i,j=1}^{m,n} a_{ij} \epsilon_j + \sum_{i,j=1}^{m,n} a_{ij} \delta_i + \sum_{i,j=1}^{m,n} a_{ij} \right) \\ &= \frac{1}{4} \left(\sum_{i,j=1}^{m,n} a_{ij} \delta_i \epsilon_j + \sum_{j=1}^n \underbrace{\left(\sum_{i=1}^m a_{ij} \right)}_{=0} \epsilon_j + \sum_{i=1}^m \underbrace{\left(\sum_{j=1}^n a_{ij} \right)}_{=0} \delta_i + \sum_{j=1}^n \underbrace{\left(\sum_{i=1}^m a_{ij} \right)}_{=0} \right) \\ &= \frac{1}{4} \sum_{i,j=1}^{m,n} a_{ij} \delta_i \epsilon_j \leq \frac{1}{4}OPT(A). \end{aligned}$$

□

Esto tiene un par de aplicaciones útiles.

Para empezar, recordemos la reducción del problema de MAX-CUT al problema de Grothendieck que hicimos en el capítulo 3. Resulta que la matriz que usábamos en la reducción funciona, con una pequeña modificación, para reducir el problema de MAX-CUT al problema de hallar la Cut-Norm.

Para ver esto, tomemos un grafo G instancia del problema de MAX-CUT, con nodos $V := \{v_1, \dots, v_n\}$ y aristas $E := \{e_1, \dots, e_m\} \subseteq V \times V$.

Construimos la matriz $A \in \mathbb{R}^{2m \times n}$, ahora con dos filas por cada arista, de la siguiente manera:

$$a_{ij} := \begin{cases} 1, & \text{si } v_i = e_{2j-1}^1 \\ -1, & \text{si } v_i = e_{2j-1}^2 \\ -1, & \text{si } v_i = e_{2j}^1 \\ +1, & \text{si } v_i = e_{2j}^2 \\ 0, & \text{si no.} \end{cases}$$

$\tilde{A} \in \mathbb{R}^{(m+1) \times (n+1)}$, definida como

$$\tilde{A} := \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & -\sum_{j=1}^n a_{1j} \\ a_{21} & a_{12} & \cdots & a_{2n} & -\sum_{j=1}^n a_{2j} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & -\sum_{j=1}^n a_{mj} \\ -\sum_{i=1}^m a_{i1} & -\sum_{i=1}^m a_{i2} & \cdots & -\sum_{i=1}^m a_{in} & \sum_{i=1}^m \sum_{j=1}^n a_{ij} \end{pmatrix}.$$

Claramente cada fila y columna de \tilde{A} suma cero. Si miramos la demostración del Lema 4.1.8, resulta que, en este tipo de matrices, cualquier elección de signos para el problema de Grothendieck alcanza el mismo valor que $\frac{1}{4}$ del valor del corte de la matriz inducida por las columnas y filas con signos positivos. Luego hay una correspondencia exacta entre soluciones para el problema de Grothendieck y cortes posibles para la matriz \tilde{A} , con la salvedad de que al pasar al problema de Cut-Norm el valor se divide por 4.

Vamos a ver ahora que $\|\tilde{A}\|_c = \|A\|_c$. La parte de $\|\tilde{A}\|_c \geq \|A\|_c$ es obvia porque la matriz \tilde{A} contiene una copia de la matriz A . Para ir en la otra dirección, supongamos que la Cut-Norm de la matriz \tilde{A} se alcanza en dos conjuntos $S_0 \subseteq \{1, \dots, m+1\}$, $T_0 \subseteq \{1, \dots, n+1\}$. Si $m+1 \in S$, entonces podemos cambiar S por su complemento, ya que

$$\sum_{\substack{i \in S_0 \\ j \in T_0}} a_{ij} = \sum_{j \in T_0} \left(\sum_{i \in S_0} a_{ij} \right) = \sum_{j \in T_0} \left(\sum_{i \in S_0^c} a_{ij} \right) = \sum_{\substack{i \in S_0^c \\ j \in T_0}} a_{ij},$$

ya que la suma de coeficientes sobre cada columna es cero. Análogamente, si $n+1 \in T_0$, podemos cambiar T_0 por su complemento y que el valor del corte se mantenga. Luego el Cut-Norm de \tilde{A} se alcanza para algunos $S_0 \subseteq \{1, \dots, m\}$, $T_0 \subseteq \{1, \dots, n\}$. Como este corte también se puede alcanzar en la matriz A , resulta que $\|\tilde{A}\|_c \leq \|A\|_c$.

Esto nos permite usar cualquier algoritmo de aproximación de $OPT(A)$ para calcular $\|A\|_c$. Simplemente calculamos la matriz \tilde{A} , conseguimos una solución aproximada para $OPT(\tilde{A})$, la convertimos en una solución con la misma garantía para $OPT(A)$, y por último la convertimos en una solución con la misma garantía para $\|A\|_c$.

En particular, con los algoritmos del capítulo 2, esto nos permite conseguir un algoritmo polinomial aleatorio para calcular $\|A\|_c$ tal que la solución hallada tiene un valor, en esperanza, de al menos $\frac{2 \ln(1+\sqrt{2})}{\pi} \|A\|_c \approx 0,63 \|A\|_c$. Por supuesto, si tuviéramos mejores Métodos de Redondeo de Krivine obtendríamos mejores resultados, pero ya esta garantía nos servirá para muchas aplicaciones.

Usos de la Cut-Norm en algoritmos

Vamos a contar una aplicación del problema de Cut-Norm, y en particular del algoritmo basado en los Métodos de Redondeo de Krivine a la construcción de particiones de Szemerédi. Vamos a enunciar el Lema de Regularidad de Szemerédi, demostrado por primera vez en [Sze75]. Este lema surgió como una herramienta teórica para demostrar que subconjuntos suficientemente densos de los números naturales contienen progresiones aritméticas suficientemente grandes, pero desde entonces fue utilizado en muchos otros contextos, incluyendo la teoría de grafos misma.

Vamos a dar algunas definiciones necesarias.

Definición 4.2.9. Sean $X, Y \subset V$ dos subconjuntos de nodos disjuntos no vacíos de un grafo $G = (V, E)$. Definimos como $e(X, Y)$ a la cantidad de aristas con un extremo en X y el otro extremo en Y . Luego definimos la densidad de aristas entre X e Y como

$$d(X, Y) := \frac{e(X, Y)}{|X||Y|},$$

es decir, la cantidad de aristas que hay entre X e Y dividida por la cantidad de aristas que “podría haber” entre X e Y .

Y definimos también una noción de “regularidad”.

Definición 4.2.10. Sean $X, Y \subset V$ dos subconjuntos de nodos disjuntos no vacíos de un grafo $G = (V, E)$. Decimos que (X, Y) es ε -regular si para cada $S \subseteq X, T \subseteq Y$ tales que $|S| \geq \varepsilon|X|$ y $|T| \geq \varepsilon|Y|$ se cumple que

$$|d(S, T) - d(X, Y)| \leq \varepsilon.$$

En palabras, la definición de ε -regular está midiendo qué tan lejos están las aristas entre X e Y de ser un grafo aleatorio. Para ilustrar esto, supongamos que empezamos sin aristas entre X e Y , pero ya sabemos de antemano la densidad ρ de aristas que queremos que haya entre ambos. Para cada par de nodos $x \in X, y \in Y$ vamos a dibujar una arista entre x e y con probabilidad ρ independientemente. Luego efectivamente la densidad de aristas entre X e Y va a parecerse a ρ . Pero también va a parecerse mucho a ρ la densidad de aristas entre conjuntos $S \subseteq X, T \subseteq Y$ suficientemente grandes. Luego, es muy razonable esperar que un grafo aleatorio en este sentido cumpla la Definición 4.2.10, y que ésta nos de alguna noción de cuánto un grafo se parece a un grafo aleatorio.

Ahora sí, con estas definiciones, vamos a enunciar el Lema de Regularidad de Szemerédi, que dice que podemos partir cualquier grafo en una cantidad acotada de conjuntos razonablemente regulares dos a dos.

Lema 4.2.11 (Lema de Regularidad de Szemerédi). *Para todos $\eta \in (0, 1)$ y $k \in \mathbb{N}$, existe $K \in \mathbb{N}$ tal que para cualquier grafo $G = (V, E)$ con $|V| > K$ existe una partición de V en m conjuntos $V = S_1 \sqcup S_2 \sqcup \dots \sqcup S_m$ tales que*

- $k \leq m \leq K$
- $|S_i| - |S_j| \leq 1, \quad \forall i, j \in \{1, \dots, m\}$
- Al menos $(1 - \eta)^{\binom{m}{2}}$ de los pares (S_1, S_2) son ε -regulares.

Es un hecho conocido que, dado $\varepsilon > 0$ el problema de decidir si un grafo bipartito $G = (V, E)$ (con $V = X \sqcup Y, |X| = |Y| = n$) es ε -regular es co-NP-completo (ver [ADL⁺94]). Es decir, el complemento del problema, que equivale a decidir si el grafo G no es ε -regular es NP-completo (la pertenencia a NP es sencilla, porque un contraejemplo de ε -regularidad es evaluable en tiempo polinomial).

Sin embargo, existe un algoritmo polinomial que permite hallar una partición de Szemerédi como en el Lema 4.2.11. La forma del algoritmo para eludir la NP-hardness de decidir ε -regularidad es definir un $\varepsilon' < \varepsilon$ y decidir en cambio entre los casos:

1. El par (X, Y) es ε -regular ($\Rightarrow (X, Y)$ es ε' -regular).
2. El par (X, Y) no es ε' -regular ($\Rightarrow (X, Y)$ no es ε -regular).

En el caso de que el par (X, Y) no pertenezca a ninguno de los dos casos, el algoritmo puede responder cualquier cosa.

El algoritmo original de [ADL⁺94] lograba esto con $\varepsilon' := \frac{\varepsilon^4}{16}$. En cambio, gracias al algoritmo basado en los Métodos de Redondeo de Krivine, [AN06] logran mejorar este paso a $\varepsilon' := \frac{\varepsilon^3}{2}$. Gracias a esta mejora en ese paso del algoritmo de particiones de Szemerédi, se logra mejorar el valor del K asegurado (y si bien así se logra la mejor cota conocida, el K sigue teniendo tamaño monstruoso en relación a ε).

Vamos a contar brevemente cómo funciona este paso del algoritmo.

Sea $G = (V, E)$ con $V = X \sqcup Y, X = \{v_1, \dots, v_n\}, Y = \{w_1, \dots, w_n\}$ un grafo bipartito. Sea $d := \frac{e(X, Y)}{|X| \cdot |Y|}$. Construimos la siguiente matriz $A \in \mathbb{R}^{n \times n}$:

$$a_{ij} := \begin{cases} 1 - d & \text{si } (v_i, w_j) \in E \\ -d & \text{si } (v_i, w_j) \notin E \end{cases}$$

Luego, dados $S \subseteq X, T \subseteq Y$, tenemos que

$$\left| \sum_{\substack{i \in S \\ j \in T}} a_{ij} \right| = \left| e(S, T) - |S| \cdot |T| \cdot d \right| = |S| \cdot |T| \cdot \left| \frac{e(S, T)}{|S| \cdot |T|} - \frac{e(X, Y)}{|X| \cdot |Y|} \right|. \quad (4.21)$$

Si (X, Y) no es ε -regular, entonces existen $|S| \geq \varepsilon|X|, |T| \geq \varepsilon|Y|$ tales que $\left| \frac{e(S, T)}{|S| \cdot |T|} - \frac{e(X, Y)}{|X| \cdot |Y|} \right|$, y por lo tanto

$$\|A\|_c \geq \left| \sum_{\substack{i \in S \\ j \in T}} a_{ij} \right| \geq \varepsilon^3 n^2.$$

Luego nuestro algoritmo tiene que poder hallar una solución subóptima S_0, T_0 de al menos $\frac{1}{2}$ de ese valor:

$$\left| \sum_{\substack{i \in S_0 \\ j \in T_0}} a_{ij} \right| \geq \frac{1}{2} \|A\|_c \geq \frac{1}{2} \varepsilon^3 n^2.$$

Pero además como $|S_0|, |T_0| \leq n$ y $\left| \frac{e(S_0, T_0)}{|S_0| \cdot |T_0|} - \frac{e(X, Y)}{|X| \cdot |Y|} \right| \leq 1$, a partir de (4.21) tenemos que

$$\min \left\{ n|S_0|, n|T_0|, n^2 \left| \frac{e(S_0, T_0)}{|S_0| \cdot |T_0|} - \frac{e(X, Y)}{|X| \cdot |Y|} \right| \right\} \geq \left| \sum_{\substack{i \in S_0 \\ j \in T_0}} a_{ij} \right| \geq \frac{1}{2} \varepsilon^3 n^2.$$

Esto nos garantiza que siempre que (X, Y) sea ε -regular podremos encontrar S_0, T_0 contraejemplos de $\frac{1}{2}\varepsilon^3$ -regularidad:

$$|S_0| \geq \frac{1}{2} \varepsilon^3 n^2 \quad |T_0| \geq \frac{1}{2} \varepsilon^3 n^2 \quad \left| \frac{e(S_0, T_0)}{|S_0| \cdot |T_0|} - \frac{e(X, Y)}{|X| \cdot |Y|} \right| \geq \frac{1}{2} \varepsilon^3. \quad (4.22)$$

Luego nuestro algoritmo será aproximar lo mejor posible la Cut-Norm de A , y responder que es ϵ -regular (aunque podría ser falso) si encontramos conjuntos S_0, T_0 como en (4.22). El hecho de que el algoritmo pueda devolver conjuntos explícitos S_0, T_0 también es una mejora al algoritmo original para hallar particiones.

Vamos a ver ahora una aplicación del algoritmo a otro problema, el de la cut-decomposition de una matriz. Para esto introducimos la siguiente definición.

Definición 4.2.12. Decimos que una matriz $D \in \mathbb{R}^{m \times n}$ es una matriz de corte si existen conjuntos $S \subseteq \{1, \dots, m\}$, $T \subseteq \{1, \dots, n\}$ y $d \in \mathbb{R}$ tales que

$$d_{ij} := \begin{cases} d & \text{si } i \in S \text{ y } j \in T \\ 0 & \text{si no,} \end{cases} \quad \forall i \in \{1, \dots, m\}, j \in \{1, \dots, n\}.$$

En tal caso, denotaremos a la matriz D como $CUT(S, T, d)$.

El objetivo es encontrar, para una matriz $A \in \mathbb{R}^{m \times n}$ tal que $|a_{ij}| \leq 1$ para todo $1 \leq i \leq m, 1 \leq j \leq n$, una descomposición en matrices de corte $D_k, 1 \leq k \leq K$, de manera que lo que “sobre” sea pequeño:

$$\left\| A - \sum_{k=1}^K D_k \right\|_c \leq \epsilon mn.$$

Vamos a mostrar como hallar esta descomposición de forma inductiva. Comenzamos con la matriz $A_0 := A$, y vamos definiendo matrices de corte D_1, \dots, D_r , de forma tal que

$$A_r = (a_{ij}^{(r)}) := A - \sum_{k=1}^r D_k.$$

Si en algún momento logramos que $\|A_r\|_c \leq \epsilon mn$, significa que encontramos la descomposición buscada. Supongamos que estamos en el paso r con la matriz A_r y las matrices de corte D_1, \dots, D_r , y que $\|A_r\|_c > \epsilon mn$. Un algoritmo con Métodos de Redondeo de Krivine nos permite encontrar, para alguna constante $\gamma > 0$, subconjuntos $S_{r+1} \subseteq \{1, \dots, m\}$, $T_{r+1} \subseteq \{1, \dots, n\}$ tales que

$$\left| \sum_{\substack{i \in S_{r+1} \\ j \in T_{r+1}}} a_{ij}^{(r)} \right| \geq \gamma \|A_r\|_c \geq \gamma \epsilon mn.$$

Entonces podemos elegir como $(r+1)$ -ésima matriz de corte a la matriz promediada de los conjuntos S_{r+1} y T_{r+1} :

$$D_{r+1} := C(S_{r+1}, T_{r+1}, d_{r+1}), \quad \text{donde } d_{r+1} := \frac{\sum_{\substack{i \in S_{r+1} \\ j \in T_{r+1}}} a_{ij}^{(r)}}{|S| \cdot |T|}.$$

Luego definimos $A_{r+1} := A_r - D_{r+1}$, como veníamos haciendo.

El objetivo es ver que este proceso eventualmente termina. Para esto, vamos a demostrar por inducción que

$$0 \leq \sum_{i=1}^m \sum_{j=1}^n (a_{ij}^{(r)})^2 \leq mn - r\gamma^2 \epsilon^2 mn.$$

El caso $r = 0$ es trivial porque habíamos pedido $|a_{ij}| \leq 1$ para todos los i, j . Y si vale para r , entonces

$$\begin{aligned}
 \sum_{i=1}^m \sum_{j=1}^n (a_{ij}^{(r+1)})^2 &= \sum_{\substack{i \in S_{r+1} \\ j \in T_{r+1}}} (a_{ij}^{(r+1)} - d_{r+1})^2 + \sum_{\substack{i \in S_{r+1} \\ j \in T_{r+1}^c}} (a_{ij}^{(r+1)})^2 + \sum_{\substack{i \in S_{r+1}^c \\ j \in T_{r+1}}} (a_{ij}^{(r+1)})^2 = \\
 &= \sum_{\substack{i \in S_{r+1} \\ j \in T_{r+1}}} (a_{ij}^{(r)} - d_{r+1})^2 + \sum_{\substack{i \in S_{r+1} \\ j \in T_{r+1}^c}} (a_{ij}^{(r)})^2 + \sum_{\substack{i \in S_{r+1}^c \\ j \in T_{r+1}}} (a_{ij}^{(r)})^2 = \\
 &= \sum_{\substack{i \in S_{r+1} \\ j \in T_{r+1}}} (a_{ij}^{(r)})^2 - 2 \sum_{\substack{i \in S_{r+1} \\ j \in T_{r+1}}} a_{ij}^{(r)} d_{r+1} + \sum_{\substack{i \in S_{r+1} \\ j \in T_{r+1}}} d_{r+1}^2 + \sum_{\substack{i \in S_{r+1} \\ j \in T_{r+1}^c}} (a_{ij}^{(r)})^2 + \sum_{\substack{i \in S_{r+1}^c \\ j \in T_{r+1}}} (a_{ij}^{(r)})^2 = \\
 &= \sum_{\substack{i \in S_{r+1} \\ j \in T_{r+1}}} (a_{ij}^{(r)})^2 - d_{r+1} \left(\sum_{\substack{i \in S_{r+1} \\ j \in T_{r+1}}} d_{r+1} - 2 \sum_{\substack{i \in S_{r+1} \\ j \in T_{r+1}}} a_{ij}^{(r)} \right) + \sum_{\substack{i \in S_{r+1} \\ j \in T_{r+1}^c}} (a_{ij}^{(r)})^2 + \sum_{\substack{i \in S_{r+1}^c \\ j \in T_{r+1}}} (a_{ij}^{(r)})^2 = \\
 &= \sum_{\substack{i \in S_{r+1} \\ j \in T_{r+1}}} (a_{ij}^{(r)})^2 - d_{r+1} \sum_{\substack{i \in S_{r+1} \\ j \in T_{r+1}}} a_{ij}^{(r)} + \sum_{\substack{i \in S_{r+1} \\ j \in T_{r+1}^c}} (a_{ij}^{(r)})^2 + \sum_{\substack{i \in S_{r+1}^c \\ j \in T_{r+1}}} (a_{ij}^{(r)})^2 = \\
 &= \sum_{\substack{i \in S_{r+1} \\ j \in T_{r+1}}} (a_{ij}^{(r)})^2 - \frac{1}{|S_{r+1}||T_{r+1}|} \left(\sum_{\substack{i \in S_{r+1} \\ j \in T_{r+1}}} a_{ij}^{(r)} \right)^2 + \sum_{\substack{i \in S_{r+1} \\ j \in T_{r+1}^c}} (a_{ij}^{(r)})^2 + \sum_{\substack{i \in S_{r+1}^c \\ j \in T_{r+1}}} (a_{ij}^{(r)})^2 = \\
 &= \sum_{\substack{i \in S_{r+1} \\ j \in T_{r+1}}} (a_{ij}^{(r)})^2 - \frac{1}{|S_{r+1}||T_{r+1}|} \left(\sum_{\substack{i \in S_{r+1} \\ j \in T_{r+1}}} a_{ij}^{(r)} \right)^2 + \sum_{\substack{i \in S_{r+1} \\ j \in T_{r+1}^c}} (a_{ij}^{(r)})^2 + \sum_{\substack{i \in S_{r+1}^c \\ j \in T_{r+1}}} (a_{ij}^{(r)})^2 = \\
 &= \sum_{i=1}^m \sum_{j=1}^n (a_{ij}^{(r)})^2 - \frac{1}{|S_{r+1}||T_{r+1}|} \left(\sum_{\substack{i \in S_{r+1} \\ j \in T_{r+1}}} a_{ij}^{(r)} \right)^2 \leq \sum_{i=1}^m \sum_{j=1}^n (a_{ij}^{(r)})^2 - \frac{(\gamma \epsilon mn)^2}{mn}.
 \end{aligned}$$

Esto completa el paso inductivo. Esto nos dice que no podemos realizar más de $\frac{1}{\gamma^2 \epsilon^2}$ pasos de este proceso, y que por lo tanto para algún $r \leq \frac{1}{\gamma^2 \epsilon^2} + 1$ se tiene que $\|A_r\|_c \leq \epsilon mn$, y así obtenemos una cut-decomposition válida.

El problema de hallar una cut-decomposition tiene variadas aplicaciones en algoritmos de optimización. Un ejemplo es el Problema de Asignación Cuadrática Métrico, estudiado con estas técnicas en [FK99].

Bibliografía

- [AAR99] George E Andrews, Richard Askey, and Ranjan Roy. *Special functions*, volume 71. Cambridge university press, 1999.
- [ADL⁺94] Noga Alon, Richard A Duke, Hanno Lefmann, Vojtech Rodl, and Raphael Yuster. The algorithmic aspects of the regularity lemma. *Journal of Algorithms*, 16(1):80–109, 1994.
- [ALM⁺98] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. Proof verification and the hardness of approximation problems. *Journal of the ACM (JACM)*, 45(3):501–555, 1998.
- [AN06] Noga Alon and Assaf Naor. Approximating the cut-norm via Grothendieck’s inequality. *SIAM Journal on Computing*, 35(4):787–803, 2006.
- [BMMN13] Mark Braverman, Konstantin Makarychev, Yury Makarychev, and Assaf Naor. The Grothendieck constant is strictly smaller than Krivine’s bound. In *Forum of Mathematics, Pi*, volume 1, page e4. Cambridge Univ Press, 2013.
- [DF92] Andreas Defant and Klaus Floret. *Tensor norms and operator ideals*, volume 176. Elsevier, 1992.
- [DGFS08] Joseph Diestel, Alexandre Grothendieck, Jan H Fourie, and Johan Swart. *The metric theory of tensor products: Grothendieck’s résumé revisited*. American Mathematical Soc., 2008.
- [DW08] Ronald De Wolf. A brief introduction to Fourier Analysis on the boolean cube. *Theory of Computing, Graduate Surveys*, 1:1–20, 2008.
- [FK99] Alan Frieze and Ravi Kannan. Quick approximation to matrices and applications. *Combinatorica*, 19(2):175–220, 1999.
- [GLS81] Martin Grötschel, László Lovász, and Alexander Schrijver. The ellipsoid method and its consequences in combinatorial optimization. *Combinatorica*, 1(2):169–197, 1981.
- [GO13] Venkatesan Guruswami and Ryan O’Donnell. The PCP Theorem and Hardness of Approximation, notas del curso, 2013. Disponible en <http://courses.cs.washington.edu/courses/cse533/05au/>.

- [Gro56] Alexander Grothendieck. Résumé de la théorie métrique des produits tensoriels topologiques. *Bol. Soc. Matem. Sao Paulo*, 8:1–79, 1956.
- [GW95] Michel X Goemans and David P Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM (JACM)*, 42(6):1115–1145, 1995.
- [KD06] Subhash Khot and Ryan O Donnell. SDP gaps and UGC-hardness for MAXCUT-GAIN. In *Foundations of Computer Science, 2006. FOCS'06. 47th Annual IEEE Symposium on*, pages 217–226. IEEE, 2006.
- [Kho] Subhash Khot. On the unique games conjecture (invited survey). In *2010 IEEE 25th Annual Conference on Computational Complexity*.
- [Kho02] Subhash Khot. On the power of unique 2-prover 1-round games. In *Proceedings of the thirty-fourth annual ACM symposium on Theory of computing*, pages 767–775. ACM, 2002.
- [Kho10] Subhash Khot. Inapproximability of NP-complete problems, discrete fourier analysis, and geometry. In *International Congress of Mathematics*, volume 5, 2010.
- [KKMO07] Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O’Donnell. Optimal inapproximability results for MAX-CUT and other 2-variable CSPs? *SIAM Journal on Computing*, 37(1):319–357, 2007.
- [KN12] Subhash Khot and Assaf Naor. Grothendieck-type inequalities in combinatorial optimization. *Communications on Pure and Applied Mathematics*, 65(7):992–1035, 2012.
- [Kön01] Hermann König. On an extremal problem originating in questions of unconditional convergence. In *Recent Progress in Multivariate Approximation*, pages 185–192. Springer, 2001.
- [KR08] Subhash Khot and Oded Regev. Vertex cover might be hard to approximate to within $2-\varepsilon$. *Journal of Computer and System Sciences*, 74(3):335–349, 2008.
- [Kri78] Jean-Louis Krivine. Constantes de Grothendieck et fonctions de type positif sur les spheres. *Séminaire Analyse fonctionnelle*, pages 1–17, 1978.
- [LP68] Joram Lindenstrauss and A Pełczyński. Absolutely summing operators in \mathcal{L}_p -spaces and their applications. *Studia Mathematica*, 3(29):275–326, 1968.
- [MOO10] Elchanan Mossel, Ryan O’Donnell, and Krzysztof Oleszkiewicz. Noise stability of functions with low influences: Invariance and optimality. *Annals of Mathematics*, 171(1):295–341, 2010.
- [NR14] Assaf Naor and Oded Regev. Krivine schemes are optimal. *Proceedings of the American Mathematical Society*, 142(12):4315–4320, 2014.

- [O'D13] Ryan O'Donnell. A Theorist's Toolkit, notas del curso, 2013. Disponible en <http://www.cs.cmu.edu/~Eodonnell/toolkit13/>.
- [Pis12] Gilles Pisier. Grothendieck's theorem, past and present. *Bulletin of the American Mathematical Society*, 49(2):237–323, 2012.
- [Rag08] Prasad Raghavendra. Optimal algorithms and inapproximability results for every CSP? In *Proceedings of the fortieth annual ACM symposium on Theory of computing*, pages 245–254. ACM, 2008.
- [Ree91] James A Reeds. A new lower bound on the real Grothendieck constant. *Manuscript*, 1991. Disponible en <https://web.archive.org/web/20140107030238/http://www.dtc.umn.edu/~Ereedsj/bound2.dvi>.
- [RS09] Prasad Raghavendra and David Steurer. Towards computing the Grothendieck constant. In *Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 525–534. Society for Industrial and Applied Mathematics, 2009.
- [Sze75] Endre Szemerédi. Regular partitions of graphs. Technical report, DTIC Document, 1975.
- [Zwi14] Daniel Zwillinger. *Table of integrals, series, and products*. Elsevier, 2014.

