



Universidad de Valladolid

Facultad de Ciencias

TRABAJO FIN DE GRADO

Grado en Matemáticas

**Sobre la conjetura de la sensibilidad
y su resolución vía teoría de grafos**

Autora: Sara Asensio Ferrero

Tutor: Philippe Gimenez

*A Philippe Gimenez, por haber confiado tanto en mí desde el principio y porque con la luz que desprende ha sido capaz de contagiarme su entusiasmo
Gracias por haber sido tan buen profesor y tan buena persona*

A Ana Núñez, por su disposición, por su trabajo y porque su carácter y su valía la han convertido para mí en un modelo a seguir

A mis padres, mis abuelos, mi hermano y demás familia, por la vida que me han regalado, por la educación que me han dado y por haberme impulsado con tanto amor para cumplir mis objetivos

A mis amigos, por haberse quedado a mi lado hasta el final de este trayecto y porque ojalá este viaje juntos siga mucho más allá

A todos los que alguna vez me han demostrado que confiaban en mí

Índice general

1. Introducción	1
2. Funciones booleanas	1
2.1. Definición y primeras propiedades	1
2.2. Medidas de complejidad	4
2.2.1. El grado	5
2.2.2. El grado aproximado	9
2.2.3. La sensibilidad	18
2.2.4. La sensibilidad por bloques	19
2.2.5. Otras medidas de complejidad interesantes	20
2.3. Primeras relaciones entre medidas	22
3. La conjetura de la sensibilidad	31
3.1. Enunciado y consecuencias	31
3.2. Intervención de la teoría de grafos	31
3.2.1. El teorema de equivalencia	32
3.2.2. Sobre los subgrafos inducidos de un cubo	39
3.3. Demostración de la conjetura	48
4. Nuevas líneas de investigación	55
A. La función de Rubinstein	59
A.1. Sensibilidad y sensibilidad por bloques	60
Bibliografía	63

Capítulo 1

Introducción

La teoría de la complejidad computacional es una de las grandes áreas que podemos encontrar dentro de las ciencias de la computación. Por tanto, podemos enmarcarla dentro de aquellas ciencias formales que estudian las bases teóricas de la computación, así como su implementación en sistemas computacionales. Es dentro de este estudio donde aparecen las matemáticas, y uno de los objetivos de este trabajo es la exposición del enfoque matemático dentro de este campo.

El principal objetivo de la teoría de la complejidad computacional es estudiar los recursos requeridos por los sistemas computacionales para resolver un problema, con el fin de clasificar estos problemas computacionales de acuerdo con su dificultad intrínseca. Para medir esta dificultad se van a definir las llamadas *medidas de complejidad*. Establecer relaciones entre estas medidas es otro de los grandes objetivos dentro de este campo de estudio, en tanto que facilita la clasificación que queremos llevar a cabo.

En este trabajo nos vamos a centrar en el caso particular de las funciones booleanas. Expondremos algunas de sus medidas de complejidad más importantes, y comprobaremos que todas ellas son equivalentes o, en un sentido que precisaremos más adelante, que están polinómicamente relacionadas.

Si bien es cierto que desde 1994 se conoce gran parte de las relaciones entre las medidas de complejidad de funciones booleanas que vamos a estudiar, hasta 2019 no se logró probar la equivalencia de todas ellas. Esto es lo que se conoce como la *conjetura de la sensibilidad*, que es el resultado principal de este trabajo y cuyo planteamiento veremos que surge de forma natural.

La conjetura de la sensibilidad fue demostrada en 2019 por Hao Huang ([7]), un matemático que actualmente trabaja como profesor asociado en el departamento de matemáticas de la Universidad Nacional de Singapur, y que ha reconocido haber tardado aproximadamente siete años en demostrar este resultado.

Lo más fascinante de su demostración es el hecho de que Huang se apoya en dos resultados previos que pueden resultar sorprendentes por su naturaleza. Uno de ellos es un resultado de teoría de grafos de naturaleza pura-

mente combinatoria, y fue demostrado en 1988 por Chung, Füredi, Graham y Seymour ([3]). El otro resultado establece una equivalencia entre dos afirmaciones: una se encuadra, al igual que el resultado que acabamos de mencionar, dentro de la teoría de grafos; mientras que la otra hace referencia a las medidas de complejidad, en términos de las cuales se había establecido inicialmente la conjetura de la sensibilidad. Gracias a este resultado, que fue demostrado en 1992 por Gotsman y Linial ([6]), se aprecia una conexión muy potente y bonita entre estas dos áreas de trabajo dentro de las matemáticas. Huang centró sus esfuerzos en resolver un problema de teoría de grafos y, gracias al teorema de equivalencia, pudo concluir inmediatamente la demostración de la conjetura de la sensibilidad.

El nombre de este resultado se debe a que la sensibilidad de una función booleana es una de sus medidas de complejidad, y la mencionada conjetura precisamente buscaba probar que la sensibilidad era equivalente a todas las otras medidas que se conocían hasta el momento. En los últimos años se han desarrollado otras medidas de complejidad que no vamos a tratar en este trabajo, como por ejemplo la complejidad cuántica relacionada con los ordenadores cuánticos. También se ha probado que estas medidas están relacionadas polinómicamente con las más “clásicas”, incluso antes de probarlo para la sensibilidad.

Además, la demostración de la conjetura ha abierto nuevas líneas de investigación sobre las que también hablaremos, centradas sobre todo en tratar de extender los resultados que damos en este trabajo para funciones booleanas y grafos de hipercubos n -dimensionales a otros tipos de grafos. Esto nos va a permitir reflejar la contemporaneidad del tema tratado, teniendo en cuenta que los últimos artículos relacionados con este tema han sido publicados en 2022, el mismo año de la elaboración de este trabajo.

Capítulo 2

Funciones booleanas

2.1. Definición y primeras propiedades

Comenzamos este capítulo introduciendo las funciones booleanas, que van a ser el objeto de estudio principal en este trabajo.

Definición 2.1.1. Una *función booleana* $f(\mathbf{x}) = f(x_1, \dots, x_n)$ en n variables es una aplicación $f : \{0, 1\}^n \rightarrow \{0, 1\}$.

Como en el conjunto $\{0, 1\}^n$ hay 2^n elementos, y cada uno de ellos puede tomar dos valores posibles a través de una función booleana, de la definición anterior se deduce inmediatamente que el número total de funciones booleanas es 2^{2^n} .

Aunque la definición de *función booleana* es muy simple, vamos a ilustrar el concepto a través de algunos ejemplos sobre los que volveremos más adelante:

Ejemplo 2.1.2. La función de paridad asigna a cada $\mathbf{x} = (x_1, \dots, x_n) \in \{0, 1\}^n$ el valor 1 si \mathbf{x} tiene un número impar de entradas no nulas, y 0 si el número de entradas no nulas de \mathbf{x} es par. Utilizando el lenguaje de congruencias y denotando por \oplus_n a la función de paridad, tenemos que

$$\oplus_n(\mathbf{x}) = 1 \Leftrightarrow x_1 + \dots + x_n \equiv 1 \pmod{2}.$$

En resumen, la función de paridad consiste en asignar a cada \mathbf{x} el valor correspondiente a la suma módulo 2 de sus componentes.

Ejemplo 2.1.3. La función de mayoría asigna a cada $\mathbf{x} = (x_1, \dots, x_n) \in \{0, 1\}^n$ el valor 1 cuando el número de entradas no nulas de \mathbf{x} es al menos la mitad en el caso de n par, o al menos el número entero más próximo por exceso a la mitad en el caso de n impar. Se asignará el valor 0 en caso contrario. Utilizando la notación habitual para la función techo, que precisamente

asigna a cada número real el entero más próximo por exceso, y denotando a la función de mayoría por Maj_n , tenemos que:

$$\text{Maj}_n(\mathbf{x}) = 1 \Leftrightarrow x_1 + x_2 + \dots + x_n \geq \left\lceil \frac{n}{2} \right\rceil.$$

Esta función de mayoría no es más que un caso particular de las funciones umbral que presentamos en el siguiente ejemplo.

Ejemplo 2.1.4. La función umbral Th_k^n asigna a cada $\mathbf{x} = (x_1, \dots, x_n) \in \{0, 1\}^n$ el valor 1 si y solamente si el número de entradas no nulas de \mathbf{x} es al menos k . Es decir,

$$\text{Th}_k^n(\mathbf{x}) = 1 \Leftrightarrow x_1 + x_2 + \dots + x_n \geq k.$$

Utilizando esta notación, la función de mayoría descrita en el ejemplo anterior se podría representar mediante $\text{Th}_{\lceil \frac{n}{2} \rceil}^n$.

Además de estas funciones booleanas en n variables, vamos a describir a continuación algunas de las operaciones booleanas básicas más importantes en informática, indicando entre paréntesis el nombre habitual que reciben dentro de esta disciplina:

- Negación (NOT): $\neg x = 1 - x$.
- Conjunción (AND): $x \wedge y = x \cdot y$.
- Disyunción (OR): $x \vee y = 1 - (1 - x) \cdot (1 - y)$.
- Paridad (XOR): $x \oplus y = x(1 - y) + y(1 - x) = (x + y) \bmod 2$.

Tradicionalmente, las funciones booleanas se han definido como aplicaciones $f : \{F, T\}^n \rightarrow \{F, T\}$, donde la F hace referencia al valor “falso” y la T se refiere al valor “verdadero” (ambas iniciales procedentes de las palabras inglesas *false* y *true* con el mismo significado). Sin embargo, actualmente es más común representar “verdadero” a través del número 1 y “falso” a través del 0, tal y como hemos hecho en nuestra primera definición, con el fin de llevar a cabo un manejo numérico, que resulta más cómodo.

Definición 2.1.5. Se dice que un polinomio real multivariante $p : \mathbb{R}^n \rightarrow \mathbb{R}$ representa a una función booleana f si para cada $\mathbf{x} \in \{0, 1\}^n$ se tiene que $f(\mathbf{x}) = p(\mathbf{x})$.

Por ejemplo, evaluando en todas las entradas posibles es fácil ver que podemos representar la función de conjunción de tres variables, que se denota por AND_3 , a través del polinomio $p(x_1, x_2, x_3) = x_1 x_2 x_3$; y la función de disyunción de tres variables, a la que denotamos por OR_3 , a través de $q(x_1, x_2, x_3) = x_1 + x_2 + x_3 - x_1 x_2 - x_1 x_3 - x_2 x_3 + x_1 x_2 x_3$.

El primer resultado que vamos a probar afirma que cada función booleana se puede representar de forma única a través de un polinomio. Si recordamos que un polinomio multilineal es un polinomio en varias variables donde el exponente de cada variable puede tomar solo los valores 0 o 1, tenemos el siguiente resultado:

Proposición 2.1.6. *Para toda función booleana $f : \{0, 1\}^n \rightarrow \{0, 1\}$ existe un único polinomio multilineal que la representa.*

Notación 2.1.7. A lo largo de todo el trabajo, denotaremos por $[n]$ al conjunto $\{1, \dots, n\}$.

Demostración. En primer lugar, si existe una representación polinomial de una función booleana f , esta siempre se va a poder tomar multilineal. Esto se debe a que para cada $x \in \{0, 1\}$ y $k \neq 0$ siempre tenemos que $x^k = x$. De esta manera, ninguna de las variables del mencionado polinomio necesitará un exponente mayor que 1.

Pasamos a probar ahora la existencia de tal representación, y lo hacemos por inducción en el número de variables n .

- Si $n = 1$, solo hay 4 funciones booleanas posibles f , y es muy fácil determinar un polinomio real en una variable que represente a cada una de ellas.
 - Si $f(0) = 0$ y $f(1) = 0$, basta considerar $p(x) = 0$.
 - Si $f(0) = 0$ y $f(1) = 1$, basta considerar $p(x) = x$.
 - Si $f(0) = 1$ y $f(1) = 0$, basta considerar $p(x) = 1 - x$.
 - Si $f(0) = 1$ y $f(1) = 1$, basta considerar $p(x) = 1$.
- Supongamos ahora que el resultado es cierto para funciones booleanas en $n - 1$ variables. Si consideramos las dos funciones booleanas en $n - 1$ variables dadas por $f_1(x_1, \dots, x_{n-1}) = f(x_1, \dots, x_{n-1}, 0)$ y $f_2(x_1, \dots, x_{n-1}) = f(x_1, \dots, x_{n-1}, 1)$, aplicando la hipótesis de inducción es posible considerar sus representaciones polinomiales multilineales Q y R , respectivamente. Considerando ahora el polinomio

$$P(x_1, \dots, x_n) = x_n R(x_1, \dots, x_{n-1}) + (1 - x_n) Q(x_1, \dots, x_{n-1}),$$

es claro que P es multilineal y, por la forma en la que se ha definido, que representa a f .

Por último, vamos a demostrar que esta representación es, de hecho, única. Para ello, supongamos que P_1 y P_2 son dos representaciones multilineales de f . Entonces se verifica que $(P_1 - P_2)(\mathbf{x}) = 0$ para cada $\mathbf{x} \in \{0, 1\}^n$.

Razonamos por reducción al absurdo, suponiendo que $P_1 - P_2$ no es idénticamente nulo y que, por tanto, contiene a algún monomio distinto de 0. Sea

ahora $S \subseteq [n]$ un conjunto minimal de índices tal que el monomio $\prod_{i \in S} x_i$ aparece en $P_1 - P_2$ con coeficiente no nulo pero los monomios de la forma $\prod_{i \in S'} x_i$, con $S' \subsetneq S$, tienen todos coeficiente nulo. Si denotamos por χ_S al elemento de $\{0, 1\}^n$ cuyas entradas toman valor distinto de 0 solo cuando corresponden a aquellos índices $i \in S$, tenemos que $(P_1 - P_2)(\chi_S) \neq 0$, lo que contradice el hecho de que $(P_1 - P_2)(\mathbf{x}) = 0$ para cada $\mathbf{x} \in \{0, 1\}^n$. Así, podemos concluir que $P_1 - P_2$ es idénticamente nulo y queda por tanto probada la unicidad de nuestra representación. \square

La elección de representar “verdadero” a través de un 1 y “falso” mediante un 0 es algo arbitrario. Otra elección conveniente con la que trabajaremos en algún punto de este trabajo (siempre con previo aviso) consiste en representar “verdadero” a través del valor -1 y “falso” mediante un 1. Las representaciones polinomiales de una función booleana f que resultan al utilizar diferentes codificaciones para los valores “verdadero” y “falso” pueden ser muy distintas; por ejemplo, una puede ser muy dispersa (con pocos coeficientes no nulos), mientras que la otra puede ser densa. Sin embargo, el estudio de las funciones booleanas ha permitido probar que el grado de esta representación polinomial multilineal permanece invariante en las distintas representaciones. Para ello, basta darse cuenta de que toda función booleana $f : \{0, 1\}^n \rightarrow \{0, 1\}$ se puede transformar en una función $\tilde{f} : \{-1, 1\}^n \rightarrow \{-1, 1\}$ dada por

$$\tilde{f}(x_1, \dots, x_n) = 1 - 2 \cdot f\left(\frac{1 - x_1}{2}, \dots, \frac{1 - x_n}{2}\right),$$

que toma el valor 1 donde f tomaba el valor 0 y toma el valor -1 donde f tomaba el valor 1, y cuya representación polinomial tiene el mismo grado que la de f .

Este carácter invariante del grado de la representación polinomial de una función booleana f va a permitirnos definir en la Subsección 2.2.1 el grado de una función booleana como una medida de complejidad.

El uso de esta representación polinomial de las funciones booleanas para estudiar sus propiedades computacionales se inició en el libro *Perceptrons*, de Minsky y Papert ([12]). Sin embargo, a partir de ese momento se incrementó considerablemente el número de estudios que utilizan estas representaciones polinomiales para estudiar algunas de las medidas de complejidad de las funciones booleanas.

2.2. Medidas de complejidad

El propósito de esta sección es presentar algunas de las medidas de complejidad más interesantes de las funciones booleanas, para poder demostrar posteriormente varias de las relaciones entre ellas. En este proceso va a ser

posible contemplar cómo surge de manera natural el planteamiento de la conjetura de la sensibilidad, como veremos al final del capítulo.

2.2.1. El grado

Como hemos visto en la Proposición 2.1.6, toda función booleana $f : \{0, 1\}^n \rightarrow \{0, 1\}$ se puede representar a través de un único polinomio multilinear. Esto nos lleva a definir nuestra primera medida de complejidad de una función booleana como el parámetro más básico de su representación polinomial: el grado.

Definición 2.2.1. Para una función booleana f se define el *grado* de f , que representaremos por $\deg(f)$ en virtud de la palabra inglesa para referirse al grado (*degree*), como el grado del único polinomio real multilinear que representa a f .

A continuación, vamos a presentar algunos resultados necesarios para la demostración de un teorema enunciado y probado por primera vez por Nisan y Szegedy en [14] en 1994, y que responde a la pregunta de cuál es el menor grado posible para una función booleana que depende de n variables.

En nuestro camino hacia la prueba de este teorema, conviene ver una función booleana como una aplicación $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$, utilizando el valor -1 para representar “verdadero” y 1 para representar “falso”, dado que es así como aparece en varios de los resultados más clásicos que vamos a utilizar.

Notación 2.2.2. Para cada subconjunto $S \subseteq [n]$, vamos a denotar $X_S = \prod_{i \in S} x_i$.

Los primeros lemas que vamos a utilizar, así como la primera definición, aparecen en [10] y son los que exponemos a continuación. No obstante, al igual que hacen Nisan y Szegedy en [14], no entraremos en los detalles de las demostraciones debido a que estas corresponden a áreas de las matemáticas en las que no estamos trabajando.

Lema 2.2.3 (Igualdad de Parseval). *Si representamos una función booleana f a través del polinomio multilinear $P = \sum_S \alpha_S X_S$, entonces*

$$\sum_S \alpha_S^2 = 1.$$

Definición 2.2.4. Para cada función booleana f y cada variable x_i se define la *influencia* de x_i sobre f , que denotaremos por $\text{Inf}_i(f)$, como la siguiente probabilidad:

$$\text{Prob}[f(x_1, \dots, x_{i-1}, -1, x_{i+1}, \dots, x_n) \neq f(x_1, \dots, x_{i-1}, 1, x_{i+1}, \dots, x_n)],$$

donde $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n$ se toman aleatoriamente en $\{-1, 1\}$.

Lema 2.2.5. *Para cada función booleana f en n variables, si representamos f a través del polinomio multilinear $P = \sum_S \alpha_S X_S$, entonces*

$$\sum_{i=1}^n \text{Inf}_i(f) = \sum_S |S| \alpha_S^2 .$$

A partir de los dos lemas anteriores, podemos deducir fácilmente el siguiente resultado:

Corolario 2.2.6. *Para toda función booleana f ,*

$$\sum_{i=1}^n \text{Inf}_i(f) \leq \deg(f).$$

Demostración. En virtud del Lema 2.2.5, $\sum_{i=1}^n \text{Inf}_i(f) = \sum_S |S| \alpha_S^2$. Ahora bien, como evidentemente $|S| \leq \deg(f)$ para todos aquellos $S \subseteq [n]$ tales que $\alpha_S \neq 0$, se tiene que

$$\sum_S |S| \alpha_S^2 \leq \deg(f) \sum_S \alpha_S^2$$

y, aplicando la igualdad de Parseval, se obtiene el resultado. \square

A continuación vamos a enunciar y demostrar un resultado conocido como lema de Schwartz, que será útil en la demostración del teorema que perseguimos y que proporciona una cota superior para el número de ceros de cualquier polinomio multilinear en $\{-1, 1\}^n$ en términos probabilísticos.

Lema 2.2.7 (Schwartz). *Sea $p(x_1, \dots, x_n)$ un polinomio multilinear no nulo de grado d . Si tomamos x_1, \dots, x_n de forma independiente y aleatoria en $\{-1, 1\}$, entonces se verifica la siguiente desigualdad:*

$$\text{Prob}[p(x_1, \dots, x_n) \neq 0] \geq 2^{-d} .$$

Demostración. Razonamos por inducción sobre n .

- Para $n = 1$, tenemos que p es un polinomio lineal en una variable, por lo que tiene grado 0 o 1, y no nulo. En consecuencia, p solo puede tener a lo sumo un cero y es claro que

$$\text{Prob}[p(x) \neq 0] \geq \frac{1}{2} .$$

- Para el paso inductivo, vamos a escribir p de la siguiente manera:

$$p(x_1, \dots, x_n) = x_n \cdot g(x_1, \dots, x_{n-1}) + h(x_1, \dots, x_{n-1}).$$

Esto siempre es posible, sin más que tomar x_n como factor común de todos aquellos monomios en los que aparezca.

Distinguiamos ahora tres casos:

1. $h + g$ es idénticamente nula. En este caso $h = -g$ y podemos escribir $p = (x_n - 1)g$. Como el grado de p es d por hipótesis, de la expresión anterior se deduce que $\deg(g) = d - 1$. Esto se debe a que si el grado de g ya fuese d , al multiplicar el monomio de grado d de g (en el que no aparece x_n por ser g un polinomio en x_1, \dots, x_{n-1}) por x_n obtendríamos un monomio de p de grado $d+1$, lo que es absurdo; y si el grado de g fuese estrictamente menor que $d - 1$, entonces p tendría grado estrictamente menor que d , que también es absurdo. Además, g es un polinomio multilineal en $n - 1$ variables, de manera que podremos aplicarle nuestra hipótesis de inducción. Aplicando el teorema de la probabilidad total y teniendo presente que en este caso $p = (x_n - 1)g$ tenemos que

$$\begin{aligned}
\text{Prob}[p(\mathbf{x}) \neq 0] &= \text{Prob}[x_n = 1] \cdot \text{Prob}[p(\mathbf{x}) \neq 0 \mid x_n = 1] + \\
&+ \text{Prob}[x_n = -1] \cdot \text{Prob}[p(\mathbf{x}) \neq 0 \mid x_n = -1] = \\
&= \frac{1}{2} \text{Prob}[(-2) \cdot g(x_1, \dots, x_{n-1}) \neq 0] = \\
&= \frac{1}{2} \text{Prob}[g(x_1, \dots, x_{n-1}) \neq 0] \stackrel{(*)}{\geq} \\
&\stackrel{(*)}{\geq} \frac{1}{2} \cdot 2^{-(d-1)} = \frac{1}{2} \cdot 2^{-d+1} = 2^{-d},
\end{aligned}$$

donde en la desigualdad (*) se ha utilizado la hipótesis de inducción y previamente se ha utilizado el hecho de que x_1, \dots, x_n se toman de forma independiente, que implica que

$$\text{Prob}[g(x_1, \dots, x_{n-1}) \neq 0 \mid x_n = -1] = \text{Prob}[g(x_1, \dots, x_{n-1}) \neq 0].$$

2. $h - g$ es idénticamente nula. En este caso $h = g$ y podemos escribir $p = (x_n + 1)g$. Nuevamente tenemos que $\deg(g) = d - 1$ y que podemos aplicar a g la hipótesis de inducción, de manera que gracias al teorema de la probabilidad total

$$\begin{aligned}
\text{Prob}[p(\mathbf{x}) \neq 0] &= \text{Prob}[x_n = 1] \cdot \text{Prob}[p(\mathbf{x}) \neq 0 \mid x_n = 1] + \\
&+ \text{Prob}[x_n = -1] \cdot \text{Prob}[p(\mathbf{x}) \neq 0 \mid x_n = -1] = \\
&= \frac{1}{2} \text{Prob}[2 \cdot g(x_1, \dots, x_{n-1}) \neq 0] = \\
&= \frac{1}{2} \text{Prob}[g(x_1, \dots, x_{n-1}) \neq 0] \stackrel{(*)}{\geq} \\
&\stackrel{(*)}{\geq} \frac{1}{2} \cdot 2^{-(d-1)} = \frac{1}{2} \cdot 2^{-d+1} = 2^{-d},
\end{aligned}$$

donde en la desigualdad (*) se ha utilizado la hipótesis de inducción y previamente se ha utilizado el hecho de que x_1, \dots, x_n se toman de forma independiente.

3. Ni $h + g$ ni $h - g$ son idénticamente nulas, con sus grados acotados ambos por d .

Podemos aplicar la hipótesis de inducción a $h + g$ para aquellos (x_1, \dots, x_n) tales que $x_n = 1$ y a $h - g$ para aquellos (x_1, \dots, x_n) tales que $x_n = -1$, y lo hacemos de la siguiente manera, aplicando el teorema de la probabilidad total:

$$\begin{aligned}
\text{Prob}[p(\mathbf{x}) \neq 0] &= \text{Prob}[x_n = 1] \cdot \text{Prob}[p(\mathbf{x}) \neq 0 \mid x_n = 1] + \\
&+ \text{Prob}[x_n = -1] \cdot \text{Prob}[p(\mathbf{x}) \neq 0 \mid x_n = -1] = \\
&= \frac{1}{2} \text{Prob}[(h + g)(x_1, \dots, x_{n-1}) \neq 0 \mid x_n = 1] + \\
&+ \frac{1}{2} \text{Prob}[(h - g)(x_1, \dots, x_{n-1}) \neq 0 \mid x_n = -1] = \\
&= \frac{1}{2} \text{Prob}[(h + g)(x_1, \dots, x_{n-1}) \neq 0] + \\
&+ \frac{1}{2} \text{Prob}[(h - g)(x_1, \dots, x_{n-1}) \neq 0] \geq \\
&\stackrel{(*)}{\geq} \frac{1}{2} \cdot 2^{-d} + \frac{1}{2} \cdot 2^{-d} = 2^{-d},
\end{aligned}$$

donde en la desigualdad (*) se ha utilizado la hipótesis de inducción, y previamente se ha utilizado el hecho de que x_1, \dots, x_n se toman de forma independiente.

Concluimos así la prueba del lema de Schwartz. \square

Después de los resultados que acabamos de introducir, nos encontramos en condiciones de enunciar y demostrar el siguiente teorema, que, como habíamos dicho, nos va a permitir dar una cota inferior para el grado de una función booleana que depende de n variables.

Teorema 2.2.8. *Sea f una función booleana que depende de n variables. Entonces se verifica que*

$$\deg(f) + \log_2(\deg(f)) \geq \log_2(n).$$

Demostración. Para cada $i \in \{1, \dots, n\}$, comenzamos definiendo una función f^i en $n - 1$ variables de la forma siguiente:

$$\begin{aligned}
f^i(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) &= \\
&= f(x_1, \dots, x_{i-1}, -1, x_{i+1}, \dots, x_n) - f(x_1, \dots, x_{i-1}, 1, x_{i+1}, \dots, x_n).
\end{aligned}$$

Utilizando esta notación y el concepto de influencia de una variable sobre una función booleana introducido en la Definición 2.2.4, es claro que

$$\text{Inf}_i(f) = \text{Prob}[f^i(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) \neq 0],$$

donde $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n$ se toman aleatoriamente en $\{-1, 1\}$.

Como f depende de todas sus variables por hipótesis, para cada i se tiene que la función f^i no es idénticamente nula. En consecuencia, podemos recurrir a la Proposición 2.1.6 para afirmar que para cada i existe un único polinomio multilinear p_i de grado $d_i := \deg(f^i) \leq \deg(f) = d$ que representa a f^i y por tanto coincide con f^i en cada $\mathbf{x} \in \{-1, 1\}^{n-1}$.

El hecho de que f^i no sea idénticamente nula implica que p_i tampoco lo es. Así, es posible aplicar el lema de Schwartz para obtener que

$$\begin{aligned} \text{Inf}_i(f) &= \text{Prob}[f^i(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) \neq 0] = \\ &= \text{Prob}[p_i(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) \neq 0] \geq 2^{-d_i} \geq 2^{-d}. \end{aligned}$$

Por otra parte, en el Corolario 2.2.6 habíamos probado que

$$\sum_{i=1}^n \text{Inf}_i(f) \leq \deg(f) = d.$$

Si combinamos ahora estas dos desigualdades, obtenemos lo siguiente:

$$d \geq \sum_{i=1}^n \text{Inf}_i(f) \geq \sum_{i=1}^n 2^{-d} = n \cdot 2^{-d} = \frac{n}{2^d},$$

lo que implica que $d \cdot 2^d \geq n$.

Como el logaritmo en base 2 es una función monótona creciente (de hecho, es estrictamente creciente), podemos aplicarla a ambos lados de la desigualdad anterior conservando el sentido de la misma. De esta manera, lo que se obtiene es que

$$\log_2(d) + d = \log_2(d) + \log_2(2^d) = \log_2(d \cdot 2^d) \geq \log_2(n),$$

que es lo que queríamos probar. \square

Las técnicas empleadas dentro de la teoría de la complejidad computacional son suficientemente potentes para dar buenas cotas incluso para el menor grado necesario para “aproximar” funciones booleanas en la norma infinito, en un sentido que precisaremos en la siguiente subsección.

2.2.2. El grado aproximado

En esta sección volvemos a considerar las funciones booleanas como aplicaciones $\{0, 1\}^n \rightarrow \{0, 1\}$.

Definición 2.2.9. Sea f una función booleana, y sea p un polinomio real. Decimos que p *aproxima* f si para cada $\mathbf{x} \in \{0, 1\}^n$ se verifica que

$$|p(\mathbf{x}) - f(\mathbf{x})| \leq \frac{1}{3}.$$

El *grado aproximado* de f , que denotaremos por $\widetilde{\deg}(f)$, se define como el mínimo grado de p cuando p varía en el conjunto de los polinomios que aproximan f ; es decir,

$$\widetilde{\deg}(f) := \min \{ \deg(p), \text{ siendo } p \text{ un polinomio que aproxima } f \}.$$

La constante $\frac{1}{3}$ en la definición anterior es arbitraria y puede sustituirse por cualquier constante ϵ tal que $0 < \epsilon < \frac{1}{2}$ sin que esto afecte a nuestros resultados en nada más que un factor constante.

Tras la introducción de este concepto, nuestro objetivo pasa a ser la prueba de un resultado que nos permitirá obtener cotas inferiores tanto para el grado de una función booleana f que cumple ciertas condiciones como para su grado aproximado.

El primer paso consiste en introducir el método de simetrización, que fue usado por primera vez por Minsky y Papert en [12].

Definición 2.2.10. Si $p : \mathbb{R}^n \rightarrow \mathbb{R}$ es un polinomio en varias variables, entonces la *simetrización* de p se define como

$$p^{sym}(x_1, \dots, x_n) = \frac{\sum_{\pi \in S_n} p(x_{\pi(1)}, \dots, x_{\pi(n)})}{n!}.$$

Un hecho importante y que vamos a justificar en el siguiente lema es que si solo consideramos entradas $\mathbf{x} \in \{0, 1\}^n$, entonces resulta que p^{sym} depende solo de $x_1 + \dots + x_n$. En consecuencia, lo podemos representar como un polinomio en una variable, siendo esta variable precisamente $x_1 + \dots + x_n$. Para demostrarlo, seguiremos la prueba proporcionada en [11], que hace a su vez referencia a la prueba proporcionada originalmente por Minsky y Papert en [12].

Lema 2.2.11. Si $p : \mathbb{R}^n \rightarrow \mathbb{R}$ es un polinomio en varias variables, entonces existe un polinomio $\tilde{p} : \mathbb{R} \rightarrow \mathbb{R}$ de grado máximo n tal que, para cada $(x_1, \dots, x_n) \in \{0, 1\}^n$, se tiene que

$$p^{sym}(x_1, \dots, x_n) = \tilde{p}(x_1 + \dots + x_n).$$

Además, $\deg(\tilde{p}) \leq \deg(p)$.

Demostración. Para cada $k \in \{1, \dots, n\}$, definimos

$$P_k(\mathbf{x}) := \sum_{S \in \binom{[n]}{k}} \prod_{i \in S} x_i,$$

donde estamos denotando por $\binom{[n]}{k}$ al conjunto de subconjuntos de $[n]$ de cardinal k ; es decir, P_k no es más que la suma de todos los monomios de grado k con coeficiente 1.

Sea d el grado de p .

Razonamos por inducción sobre d para probar que existen constantes $c_i \in \mathbb{R}$, con i comprendido entre 1 y d , tales que

$$p^{sym}(\mathbf{x}) = c_0 + c_1 P_1(\mathbf{x}) + \dots + c_d P_d(\mathbf{x}).$$

- Si $d = 0$, entonces p es una constante y coincide con su simetrizado, con lo cual el resultado es cierto.
- Supongamos que el resultado es cierto para cualquier valor menor o igual que $d - 1$; es decir, nuestra hipótesis de inducción es que si q es un polinomio de grado a lo sumo $d - 1$, entonces existen $a_i \in \mathbb{R}$, con i comprendido entre 1 y $d - 1$, tales que $q^{sym}(\mathbf{x}) = a_0 + a_1 P_1(\mathbf{x}) + \dots + a_{d-1} P_{d-1}(\mathbf{x})$ (notemos que si el grado de q fuese estrictamente menor que $d - 1$, bastaría tomar $a_{d-1} = 0$ y el resultado seguiría siendo cierto).

En este paso vamos a hacer uso de un resultado muy sencillo: el simetrizado de una suma de polinomios coincide con la suma de los simetrizados. En efecto,

$$\begin{aligned} (p + q)^{sym}(\mathbf{x}) &= \frac{\sum_{\pi \in S_n} (p + q)(x_{\pi(1)}, \dots, x_{\pi(n)})}{n!} = \\ &= \frac{\sum_{\pi \in S_n} (p(x_{\pi(1)}, \dots, x_{\pi(n)}) + q(x_{\pi(1)}, \dots, x_{\pi(n)}))}{n!} = \\ &= p^{sym}(x_1, \dots, x_n) + q^{sym}(x_1, \dots, x_n). \end{aligned}$$

Si tomamos ahora un polinomio p de grado d , entonces podemos escribir $p(\mathbf{x}) = q(\mathbf{x}) + c \cdot h(\mathbf{x})$, donde q es un polinomio de grado menor o igual que $d - 1$, $c \in \mathbb{R}$ y h es una suma de monomios de grado d . A partir de lo que acabamos de ver, tenemos que $p^{sym}(\mathbf{x}) = q^{sym}(\mathbf{x}) + c \cdot h^{sym}(\mathbf{x})$, y aplicando la hipótesis de inducción obtenemos la igualdad

$$p^{sym}(\mathbf{x}) = c_0 + c_1 P_1(\mathbf{x}) + \dots + c_{d-1} P_{d-1}(\mathbf{x}) + c \cdot h^{sym}(\mathbf{x})$$

para ciertos $c_i, c \in \mathbb{R}$.

Ahora bien, recordemos que h es una suma de monomios de grado d y por lo tanto su simetrizado será la suma de los simetrizados de los monomios. Si probamos que el simetrizado de un monomio de grado d es el producto de una constante real por $P_d(\mathbf{x})$, habremos terminado.

Sin pérdida de generalidad, vamos a considerar el monomio de grado d que resulta de multiplicar las d primeras variables, $x_1 \cdot \dots \cdot x_d$. Tenemos que su simetrizado es

$$(x_1 \cdot \dots \cdot x_d)^{sym} = \frac{\sum_{\pi \in S_n} x_{\pi(1)} \cdot \dots \cdot x_{\pi(d)}}{n!}.$$

A partir de esta expresión es evidente que en el simetrizado del monomio que estamos tratando van a aparecer todos los monomios posibles

de grado d . Para que el resultado fuera el producto de un número real por $P_d(\mathbf{x})$ tendría que cumplirse que cada monomio de grado d aparezca en la expresión anterior el mismo número de veces, y esto efectivamente es cierto: el número de veces que aparece cada monomio $x_{i_1} \cdot \dots \cdot x_{i_d}$ coincide con el número de permutaciones de $[n]$ que envían el conjunto $\{1, \dots, d\}$ en el conjunto $\{i_1, \dots, i_d\}$, que son $d! \cdot (n - d)!$.

Queda probado entonces que

$$p^{sym}(\mathbf{x}) = c_0 + c_1 P_1(\mathbf{x}) + \dots + c_d P_d(\mathbf{x})$$

para algunos $c_i \in \mathbb{R}$, donde i está comprendido entre 1 y d .

Veamos ahora que para todo $k \in \{1, \dots, n\}$ se cumple que $P_k(\mathbf{x}) = \binom{|\mathbf{x}|}{k}$, donde nos referimos al módulo de un vector como la suma de sus componentes, y lo vamos a hacer razonando nuevamente por inducción, esta vez sobre k .

- Si $k = 1$, es claro:

$$P_1(\mathbf{x}) = x_1 + \dots + x_n = |\mathbf{x}| = \binom{|\mathbf{x}|}{1}.$$

- Supongamos ahora que el resultado es cierto para $k - 1$; es decir, supongamos que $P_{k-1}(\mathbf{x}) = \binom{|\mathbf{x}|}{k-1}$. Entonces

$$\begin{aligned} \binom{|\mathbf{x}|}{k} &= \frac{1}{k!} |\mathbf{x}| (|\mathbf{x}| - 1)(|\mathbf{x}| - 2) \dots (|\mathbf{x}| - (k - 1)) = \\ &= \frac{1}{k} (|\mathbf{x}| - (k - 1)) \left[\frac{1}{(k - 1)!} |\mathbf{x}| (|\mathbf{x}| - 1) \dots (|\mathbf{x}| - (k - 2)) \right] = \\ &= \frac{1}{k} \cdot (|\mathbf{x}| - (k - 1)) \cdot P_{k-1}(\mathbf{x}) = \\ &= \frac{1}{k} \cdot (x_1 + \dots + x_n - (k - 1)) \cdot P_{k-1}(\mathbf{x}) = \\ &= \frac{1}{k} \cdot [(x_1 + \dots + x_n) \cdot P_{k-1}(\mathbf{x}) - (k - 1) \cdot P_{k-1}(\mathbf{x})]. \end{aligned}$$

Queremos probar que el último miembro de la igualdad anterior coincide con $P_k(\mathbf{x})$, y lo vamos a hacer probando que coincide cada término de ambas expresiones.

Basta probar de hecho que los monomios en los que aparece x_1 coinciden en ambas expresiones (análogamente, coincidirán todos). Empezamos considerando la expresión que acabamos de obtener,

$$\frac{1}{k} \cdot [(x_1 + \dots + x_n) \cdot P_{k-1}(\mathbf{x}) - (k - 1) \cdot P_{k-1}(\mathbf{x})].$$

Si desarrollamos, $(x_1 + \dots + x_n) \cdot P_{k-1}(\mathbf{x}) = x_1 \cdot P_{k-1}(\mathbf{x}) + \dots + x_n \cdot P_{k-1}(\mathbf{x})$, suma esta última a la que vamos a denotar por $(*)$.

Cuando hacemos el producto $x_1 \cdot P_{k-1}(\mathbf{x})$ nos encontramos lo siguiente:

- Los monomios de P_{k-1} en los que no aparecía x_1 pasan a ser monomios de grado k en los que sí que aparece x_1 : serán de la forma $x_1 \cdot (x_{i_1} \dots x_{i_{k-1}})$. Cada uno de estos monomios aparecerá en (*) un total de k veces, una por cada x_{i_j} que multiplica a P_{k-1} . Además, ninguno de estos monomios aparece en el sustraendo $(k-1) \cdot P_{k-1}(\mathbf{x})$ por tener grado k . En consecuencia, al dividir por k obtendremos la suma de todos los monomios de grado k en los que aparece x_1 con coeficiente 1.
- Los monomios en los que ya aparecía x_1 permanecen intactos porque para $x_1 \in \{0, 1\}$ se cumple que $x_1^2 = x_1$. Así pues, serán de la forma $x_1 \cdot (x_{i_1} \dots x_{i_{k-2}})$. Cada uno de estos monomios aparecerá en (*) un total de $k-1$ veces, una por cada x_{i_j} que multiplica a P_{k-1} . En consecuencia, todos estos monomios desaparecerán al restar $(k-1) \cdot P_{k-1}$ y no quedará ningún monomio de grado $k-1$.

Esto nos permite concluir que efectivamente

$$\binom{|\mathbf{x}|}{k} = \frac{1}{k} \cdot [(x_1 + \dots + x_n) \cdot P_{k-1}(\mathbf{x}) - (k-1) \cdot P_{k-1}(\mathbf{x})] = P_k(\mathbf{x}).$$

El polinomio \tilde{p} en una variable de grado menor o igual que $d = \deg(p)$ definido como

$$\tilde{p}(t) := c_0 + c_1 \cdot \binom{t}{1} + \dots + c_d \cdot \binom{t}{d}$$

es el que buscamos, ya que efectivamente

$$\begin{aligned} \tilde{p}(x_1 + \dots + x_n) &= c_0 + c_1 \cdot \binom{x_1 + \dots + x_n}{1} + \dots + c_d \cdot \binom{x_1 + \dots + x_n}{d} = \\ &= c_0 + c_1 \cdot P_1(\mathbf{x}) + \dots + c_d \cdot P_d(\mathbf{x}) = \\ &= p^{sym}(\mathbf{x}), \end{aligned}$$

lo que concluye la demostración. \square

Pasamos ahora a introducir un resultado que será imprescindible en la demostración del lema que perseguimos. Se trata de un teorema de teoría de la aproximación que podemos encontrar en [4] y [15], y cuya demostración hemos seguido en [14]. Sin embargo, antes de conocer su enunciado y su demostración vamos a enunciar un teorema clásico de Markov del que veremos que es consecuencia directa. No entraremos en los detalles de la demostración de este resultado clásico porque no se ajusta al contenido de este trabajo.

Teorema 2.2.12 (Markov, [2]). *Sea $p : \mathbb{R} \rightarrow \mathbb{R}$ un polinomio de grado d en una variable tal que para cada número real $x \in [a_1, a_2]$ se verifica que $b_1 \leq p(x) \leq b_2$. Entonces la derivada de p satisface que $|p'(x)| \leq d^2 \cdot \frac{b_2 - b_1}{a_2 - a_1}$ para cada $x \in [a_1, a_2]$.*

Teorema 2.2.13 (Ehlich, Zeller [4]; Rivlin, Cheney [15]). *Sea p un polinomio con las siguientes propiedades:*

1. *Para cada entero $i \in [0, n]$ se tiene que $b_1 \leq p(i) \leq b_2$.*
2. *Para algún número real $x \in [0, n]$ la derivada del polinomio p satisface que $|p'(x)| \geq c \geq 0$.*

Entonces

$$\deg(p) \geq \sqrt{\frac{cn}{c + b_2 - b_1}} .$$

Demostración. Sea $c' = \max_{0 \leq x \leq n} |p'(x)| \geq c$. Se cumple que para cada número real $x \in [0, n]$,

$$b_1 - \frac{c'}{2} \leq p(x) \leq b_2 + \frac{c'}{2} .$$

Vamos a justificar solo la primera de las desigualdades, ya que el razonamiento para la segunda es idéntico. Para ello vamos a analizar cuál es el mínimo valor que puede tomar p en un número real $x \in [0, n]$ bajo las condiciones del enunciado.

En virtud de la primera de las propiedades, el mínimo valor que puede tomar p en un entero $i \in [0, n]$ es b_1 . Para que p pudiese alcanzar el valor mínimo tendría que tomar el valor b_1 en un entero i , tener pendiente negativa y de módulo c' en ese entero i , y tomar el valor b_1 de nuevo en el entero $i + 1$. De esta manera, p alcanzaría su valor mínimo posible en el número real $i + \frac{1}{2}$ y este sería $b_1 - \frac{c'}{2}$, lo que concluye la justificación de la primera desigualdad.

Utilizando la desigualdad de Markov introducida en el Teorema 2.2.12, tenemos que

$$|p'(x)| \leq \deg(p)^2 \cdot \frac{b_2 + \frac{c'}{2} - (b_1 - \frac{c'}{2})}{n - 0} = \deg(p)^2 \cdot \frac{c' + b_2 - b_1}{n}$$

para cada $x \in [0, n]$, luego

$$c' \leq \deg(p)^2 \cdot \frac{c' + b_2 - b_1}{n}$$

y es posible comprobar que se verifica que

$$\deg(p)^2 \geq \frac{c'n}{c' + b_2 - b_1} \geq \frac{cn}{c + b_2 - b_1} ,$$

con lo que tomando raíces cuadradas a ambos lados habremos terminado.

La primera desigualdad de la expresión anterior se obtiene despejando. Para probar la segunda desigualdad, calcularemos la diferencia de los dos

miembros y analizaremos su signo:

$$\begin{aligned}
\frac{c'n}{c' + b_2 - b_1} - \frac{cn}{c + b_2 - b_1} &= \frac{cc'n + b_2c'n - b_1c'n - c'cn - b_2cn + b_1cn}{(c' + b_2 - b_1)(c + b_2 - b_1)} = \\
&= \frac{b_2n(c' - c) + b_1n(c - c')}{(c' + b_2 - b_1)(c + b_2 - b_1)} = \\
&= \frac{n(c' - c)(b_2 - b_1)}{(c' + b_2 - b_1)(c + b_2 - b_1)} \geq 0
\end{aligned}$$

gracias a que n , $c' - c$, $b_2 - b_1$, c' y c son mayores o iguales que 0. Así, queda probada la desigualdad y por tanto el teorema. \square

Los dos teoremas anteriores son similares, aunque en el segundo no conocemos información sobre el valor $p(x)$ para cada x real en el intervalo indicado sino solo para cada entero. Así, el segundo teorema puede verse como una generalización del teorema de Markov, aunque hemos visto que se puede demostrar a partir de él de una forma simple.

Ahora ya podemos enunciar y demostrar el resultado principal de esta subsección, en el que aparece un concepto cuya definición introduciremos previamente.

Definición 2.2.14. Dado un vector \mathbf{x} , su *peso de Hamming* se define como el número de entradas no nulas de \mathbf{x} .

Este concepto aparece de forma natural en teoría de códigos al hablar de la distancia mínima de un código lineal, que es uno de sus parámetros fundamentales, y va a aparecer entre las hipótesis de nuestro siguiente resultado:

Lema 2.2.15. *Sea f una función booleana tal que $f(0, 0, \dots, 0) = 0$ y que, para cada vector booleano \mathbf{x} de peso de Hamming 1, verifique que $f(\mathbf{x}) = 1$. Entonces se verifican las siguientes desigualdades:*

$$\deg(f) \geq \sqrt{\frac{n}{2}}, \quad \widetilde{\deg}(f) \geq \sqrt{\frac{n}{6}}.$$

Demostración. Empezamos probando la cota para $\widetilde{\deg}(f)$.

Sea p un polinomio que aproxima f , de manera que se verifica que $|p(\mathbf{x}) - f(\mathbf{x})| \leq \frac{1}{3}$ para cada $\mathbf{x} \in \{0, 1\}^n$ o, equivalentemente, $-\frac{1}{3} \leq p(\mathbf{x}) - f(\mathbf{x}) \leq \frac{1}{3}$ para cada $\mathbf{x} \in \{0, 1\}^n$.

Denotaremos por \tilde{p} al polinomio en una variable que representa la simetrización de p , que verifica las siguientes propiedades:

1. Como ya habíamos probado en el Lema 2.2.11, $\deg(\tilde{p}) \leq \deg(p)$.
2. Como $|p(\mathbf{x}) - f(\mathbf{x})| \leq \frac{1}{3}$ para cada $\mathbf{x} \in \{0, 1\}^n$, para cada vector booleano \mathbf{x} se verifica que $p(\mathbf{x})$ está como mucho a distancia $\frac{1}{3}$ de un

valor booleano, que es 0 o 1. Esto implica que, en particular, $-\frac{1}{3} \leq p(\mathbf{x}) \leq \frac{4}{3}$ para cada $\mathbf{x} \in \{0, 1\}^n$. En consecuencia, teniendo en cuenta que para cada entero i se cumple que $\tilde{p}(i) = p^{sym}(x_1, \dots, x_n)$, con $x_1 + \dots + x_n = i$, llegamos a que

$$\begin{aligned} -\frac{1}{3} &= \frac{1}{n!} \cdot \left(-\frac{1}{3}\right) \cdot n! \leq \tilde{p}(i) = p^{sym}(x_1, \dots, x_n) = \\ &= \frac{1}{n!} \sum_{\pi \in S_n} p(x_{\pi(1)}, \dots, x_{\pi(n)}) \leq \frac{1}{n!} \cdot \frac{4}{3} \cdot n! = \frac{4}{3}; \end{aligned}$$

es decir, $-\frac{1}{3} \leq \tilde{p}(i) \leq \frac{4}{3}$ para cada entero $i \in [0, n]$.

3. $\tilde{p}(0) \leq \frac{1}{3}$ ya que

$$\tilde{p}(0) = p^{sym}(0, \dots, 0) = \frac{n! \cdot p(\mathbf{0})}{n!} = p(\mathbf{0})$$

y, como $f(0, 0, \dots, 0) = 0$,

$$p(\mathbf{0}) = p(\mathbf{0}) - f(\mathbf{0}) \leq \frac{1}{3}.$$

4. $\tilde{p}(1) \geq \frac{2}{3}$ ya que $\tilde{p}(1) = p^{sym}(\mathbf{x}_{H_1})$, donde \mathbf{x}_{H_1} es un vector booleano de peso de Hamming 1, en el que por tanto f va a tomar el valor 1. Así pues, $|p(\mathbf{x}_{H_1}) - f(\mathbf{x}_{H_1})| = |p(\mathbf{x}_{H_1}) - 1| \leq \frac{1}{3}$ y esto implica que $-\frac{1}{3} \leq p(\mathbf{x}_{H_1}) - 1 \leq \frac{1}{3}$ y por tanto $0 < \frac{2}{3} \leq p(\mathbf{x}_{H_1}) \leq \frac{4}{3}$. En consecuencia,

$$\tilde{p}(1) = p^{sym}(\mathbf{x}_{H_1}) = \frac{\sum_{\pi \in S_n} p(\mathbf{x}_{H_1}^\pi)}{n!} \geq \frac{\frac{2}{3}n!}{n!} = \frac{2}{3},$$

donde hemos denotado por $\mathbf{x}_{H_1}^\pi$ al resultado de permutar las entradas de \mathbf{x}_{H_1} de acuerdo con la permutación π , que vuelve a ser de nuevo un vector booleano con peso de Hamming 1.

Como \tilde{p} es un polinomio real en una variable, tenemos que es una función continua en $[0, 1]$ y derivable en $(0, 1)$. Esto nos permite aplicar el teorema de los incrementos finitos de Lagrange para concluir que existe al menos un punto $z \in (0, 1)$ tal que

$$\tilde{p}'(z) = \frac{\tilde{p}(1) - \tilde{p}(0)}{1 - 0} = \tilde{p}(1) - \tilde{p}(0),$$

y combinando las propiedades 3 y 4 anteriores tenemos que

$$\tilde{p}'(z) = \tilde{p}(1) - \tilde{p}(0) \geq \frac{2}{3} - \frac{1}{3} = \frac{1}{3}.$$

Tomando ahora $b_1 = -\frac{1}{3}$, $b_2 = \frac{4}{3}$ y $c = \frac{1}{3}$, podemos aplicar el Teorema 2.2.13 para obtener

$$\deg(\tilde{p}) \geq \sqrt{\frac{\frac{1}{3}n}{\frac{1}{3} + \frac{4}{3} + \frac{1}{3}}} = \sqrt{\frac{n}{6}},$$

y como $\deg(p) \geq \deg(\tilde{p})$, también $\deg(p) \geq \sqrt{\frac{n}{6}}$. Esto nos permite concluir que

$$\widetilde{\deg}(f) = \min_{p \text{ aproxima } f} \deg(p) \geq \sqrt{\frac{n}{6}}.$$

El razonamiento que nos permite obtener la cota para $\deg(f)$ sigue exactamente las mismas líneas. Sin embargo, vemos que la cota que se obtiene es mejor.

Sea ahora p un polinomio que representa a f . Las propiedades que verifica \tilde{p} en este caso son:

1. $\deg(\tilde{p}) \leq \deg(p)$.
2. Como p representa a f , para cada vector booleano \mathbf{x} se cumple que $p(\mathbf{x})$ coincide con el valor booleano $f(\mathbf{x})$. En consecuencia, como $\tilde{p}(i) = p^{sym}(x_1, \dots, x_n)$, con i entero entre 0 y n y $x_1 + \dots + x_n = i$, tenemos que

$$\begin{aligned} 0 \leq \tilde{p}(i) = p^{sym}(x_1, \dots, x_n) &= \frac{1}{n!} \sum_{\pi \in S_n} p(x_{\pi(1)}, \dots, x_{\pi(n)}) = \\ &= \frac{1}{n!} \sum_{\pi \in S_n} f(x_{\pi(1)}, \dots, x_{\pi(n)}) \leq \frac{n! \cdot 1}{n!} = 1, \end{aligned}$$

luego $0 \leq \tilde{p}(i) \leq 1$ para cada entero $i \in [0, n]$.

3. Se verifica que

$$\tilde{p}(0) = p^{sym}(\mathbf{0}) = \frac{\sum_{\pi \in S_n} p(\mathbf{0})}{n!} = 0,$$

donde la última igualdad se debe a que $p(\mathbf{0}) = f(\mathbf{0}) = 0$.

4. Con las notaciones anteriores,

$$\tilde{p}(1) = p^{sym}(\mathbf{x}_{H_1}) = \frac{\sum_{\pi \in S_n} p(\mathbf{x}_{H_1}^\pi)}{n!} = 1,$$

donde la última igualdad se debe a que $p(\mathbf{x}_{H_1}^\pi) = f(\mathbf{x}_{H_1}^\pi) = 1$ para toda permutación $\pi \in S_n$.

Aplicando de nuevo el teorema de los incrementos finitos de Lagrange, tenemos que existe $z \in (0, 1)$ tal que $\tilde{p}'(z) = \tilde{p}(1) - \tilde{p}(0) = 1$. Así pues, podemos aplicar el Teorema 2.2.13 tomando $c = 1$, $b_1 = 0$ y $b_2 = 1$ para concluir que

$$\deg(f) = \deg(p) \geq \deg(\tilde{p}) \geq \sqrt{\frac{1 \cdot n}{1 + 1}} = \sqrt{\frac{n}{2}}.$$

□

En el momento en que se llevó a cabo la anterior demostración ya se conocía un ejemplo con polinomios de Chebyshev que permitía demostrar que la cota para $\deg(f)$ sí que se alcanza salvo producto por una constante; es decir, lo que se alcanza es el orden de la misma. Sin embargo, todavía no se conoce si la cota para $\deg(f)$ se alcanza o no.

Hemos encontrado por tanto cotas inferiores para el grado y el grado aproximado de funciones booleanas que pertenecen a un tipo muy concreto, pero nos gustaría extender estas cotas al conjunto de todas las funciones booleanas. Para ello, tendremos que introducir primero otras medidas de complejidad de funciones booleanas, y es esto a lo que dedicamos las siguientes subsecciones.

2.2.3. La sensibilidad

Llegamos en este punto a la medida de complejidad que da nombre a la conjetura que constituye el resultado principal de este trabajo.

Definición 2.2.16. Sea $f : \{0, 1\}^n \rightarrow \{0, 1\}$ una función booleana y sea $\mathbf{x} = (x_1, \dots, x_n) \in \{0, 1\}^n$.

Denotamos por $\mathbf{x}^{\{i\}}$ al vector booleano obtenido a partir de \mathbf{x} cambiando su i -ésima entrada.

Se define la *sensibilidad local* de f en \mathbf{x} , que denotaremos por $s(f, \mathbf{x})$, como el número de índices i tales que $f(\mathbf{x}) \neq f(\mathbf{x}^{\{i\}})$.

A partir de este concepto, la *sensibilidad* de f , que denotaremos por $s(f)$, se define como el máximo de las sensibilidades locales; es decir,

$$s(f) = \max_{\mathbf{x} \in \{0, 1\}^n} s(f, \mathbf{x}).$$

En base a la definición anterior, tenemos que la sensibilidad de una función booleana mide su comportamiento local con respecto a la distancia de Hamming, siendo la distancia de Hamming entre dos vectores el número de entradas en las que ambos se diferencian (o, utilizando la terminología que habíamos introducido antes, el peso de Hamming de su diferencia). Además, esta sensibilidad se puede entender como el análogo en el caso discreto de la regularidad de una función continua. Informalmente, podemos entenderla

también como una medida de la inestabilidad de una función booleana frente a pequeñas perturbaciones en la entrada.

Presentamos a continuación algunos ejemplos:

Ejemplo 2.2.17. La sensibilidad de la función booleana $\text{AND}_n : \{0, 1\}^n \rightarrow \{0, 1\}$, que toma el valor 1 si y solo si todas las entradas del vector booleano que estamos tratando toman también el valor 1, es n .

Si consideramos $\mathbf{1} = (1, 1, \dots, 1)$, para cualquier entero $i \in [1, n]$ resulta que $1 = \text{AND}_n(\mathbf{1}) \neq \text{AND}_n(\mathbf{1}^{\{i\}}) = 0$.

Sin embargo, $s(\text{AND}_n, \mathbf{0}) = 0$ porque $\text{AND}_n(\mathbf{0}) = \text{AND}_n(\mathbf{0}^{\{i\}}) = 0$ para cada entero $i \in [1, n]$.

Ejemplo 2.2.18. La sensibilidad de la función booleana $\text{OR}_n : \{0, 1\}^n \rightarrow \{0, 1\}$, que toma el valor 0 si y solo si todas las entradas del vector booleano que estamos tratando toman el valor 0, es también n .

De forma análoga a lo que sucedía en el ejemplo anterior, si consideramos $\mathbf{0} = (0, 0, \dots, 0)$ resulta que $0 = \text{OR}_n(\mathbf{0}) \neq \text{OR}_n(\mathbf{0}^{\{i\}}) = 1$ para cualquier entero $i \in [1, n]$.

Sin embargo, $s(\text{OR}_n, \mathbf{1}) = 0$ porque $\text{OR}_n(\mathbf{1}) = \text{OR}_n(\mathbf{1}^{\{i\}}) = 1$ para cada entero $i \in [1, n]$.

Ejemplo 2.2.19. Si consideramos la función de paridad introducida en el Ejemplo 2.1.2, es claro que su sensibilidad también es n .

Esto se debe a que modificar una entrada cualquiera en un vector booleano de n componentes altera la suma módulo 2 de estas componentes.

2.2.4. La sensibilidad por bloques

Definición 2.2.20. Sea $f : \{0, 1\}^n \rightarrow \{0, 1\}$ una función booleana y sea $\mathbf{x} = (x_1, \dots, x_n) \in \{0, 1\}^n$.

Para cada subconjunto $S \subseteq [n]$, denotamos por \mathbf{x}^S al vector booleano obtenido a partir de \mathbf{x} cambiando todas las entradas x_i tales que $i \in S$.

Se define la *sensibilidad local por bloques* de f en \mathbf{x} , que denotaremos por $bs(f, \mathbf{x})$, como el máximo número k de subconjuntos B_1, \dots, B_k de $[n]$ disjuntos tales que para cada B_i se verifica que $f(\mathbf{x}) \neq f(\mathbf{x}^{B_i})$.

A partir de este concepto, la *sensibilidad por bloques* de f , que denotaremos por $bs(f)$, se define como el máximo de las sensibilidades locales por bloques; es decir,

$$bs(f) = \max_{\mathbf{x} \in \{0, 1\}^n} bs(f, \mathbf{x}).$$

Ejemplo 2.2.21. Si consideramos una vez más la función AND_n , es claro que $bs(\text{AND}_n, \mathbf{1}) = n$ porque cambiar una entrada cualquiera del vector $\mathbf{1}$ cambia el valor que le asigna nuestra función de conjunción. Así, $\{1\}, \{2\}, \dots, \{n\}$ son n bloques disjuntos y sensibles. Como el máximo valor que puede tomar

la sensibilidad por bloques es precisamente n , esto nos permite concluir que $bs(\text{AND}_n) = n$.

Sin embargo, $bs(\text{AND}_n, \mathbf{0}) = 1$ porque el único bloque sensible para este vector es $\{1, \dots, n\}$. Esto se debe a que, partiendo del vector $(0, \dots, 0)$, si queremos que su valor a través de AND_n pase a ser 1 tenemos que cambiar cada una de sus entradas.

Aunque volveremos sobre ello en la sección dedicada a estudiar las primeras relaciones entre medidas de complejidad, es claro a partir de las definiciones que el valor de la sensibilidad de una función booleana es siempre menor o igual que el de la sensibilidad por bloques. Podemos preguntarnos si existe alguna función booleana para la que la desigualdad anterior sea estricta, y la respuesta a esta pregunta, que es afirmativa, fue proporcionada por primera vez por Rubinstein en [16]. Entraremos en los detalles de la misma en el Apéndice A.

2.2.5. Otras medidas de complejidad interesantes

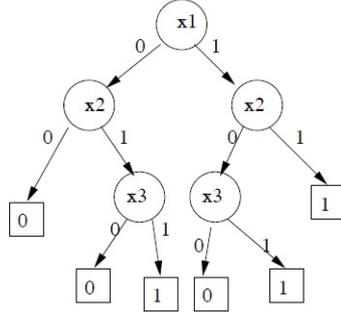
Dedicamos esta sección a introducir otras medidas de complejidad de funciones booleanas interesantes. Todas ellas aparecerán en la siguiente sección y nos permitirán plantear la conjetura de la sensibilidad, así como apreciar la relevancia que tiene.

Empezamos hablando de árboles de decisión.

Un árbol de decisión es uno de los modelos computacionales más sencillos, y se utiliza para estudiar el número de bits de la entrada que se necesita conocer para hallar el valor de una función en esta entrada. Un árbol de decisión para una función booleana es un árbol en el cual cada vértice representa a una de las variables y de él salen dos aristas, una que corresponde al valor 0 de esa variable, y la otra correspondiente al valor 1.

Dado un valor de entrada $\mathbf{x} = (x_1, \dots, x_n)$, el cálculo se efectúa de la manera siguiente en cada nodo. En primer lugar, analizamos el valor que toma la variable x_i correspondiente a ese nodo. Si $x_i = 1$, continuamos hacia el nodo al que nos conduce la arista correspondiente al valor 1, y tomamos la arista correspondiente al valor 0 en caso contrario. De esta manera, cada entrada \mathbf{x} determina un camino dentro del árbol y el valor de la hoja al final de ese camino es precisamente $f(\mathbf{x})$.

Un ejemplo de árbol de decisión para calcular los distintos valores que toma la función de mayoría en tres variables, y que podemos encontrar en las notas [1] correspondientes a un curso de la Universidad de Princeton, es el siguiente:



El coste de un árbol t para un valor de entrada \mathbf{x} , que denotaremos por $\text{cost}(t, \mathbf{x})$, no es más que el menor número de bits de \mathbf{x} cuyo valor necesita conocer el árbol antes de proporcionarnos el valor de la función en esa entrada.

Definición 2.2.22. La *profundidad* de un árbol de decisión t se define como la longitud del camino más largo de t desde la raíz hasta alguna de las hojas.

Si utilizamos la nomenclatura que ha aparecido en el párrafo anterior, tenemos que la profundidad de t se puede expresar como $\max_{\mathbf{x} \in \{0,1\}^n} \text{cost}(t, \mathbf{x})$.

Antes de pasar a la siguiente definición, notemos que los distintos órdenes en los que se pueden pedir los valores de las entradas de \mathbf{x} determinan distintos árboles de decisión para una misma función f .

Definición 2.2.23. La *complejidad por árbol de decisión* de f , que denotaremos por $D(f)$, se define como la menor profundidad de todos los árboles de decisión posibles para calcular los valores de f . Es decir, si denotamos por T el conjunto de los árboles de decisión que permiten hallar los valores de f , se tiene que

$$D(f) = \min_{t \in T} \max_{\mathbf{x} \in \{0,1\}^n} \text{cost}(t, \mathbf{x}).$$

La complejidad por árbol de decisión es conocida también como *complejidad por consulta determinista* ya que es equivalente al menor número de consultas que necesita un algoritmo determinista con el objetivo de hallar distintos valores de f .

Pasamos ahora a hablar de una nueva medida de complejidad, que será la última que introduzcamos en el trabajo: la complejidad por certificado.

Definición 2.2.24. Un *certificado* de una función booleana f para una entrada $\mathbf{x} \in \{0,1\}^n$ es un subconjunto $S \subseteq [n]$ tal que f es constante en todas las entradas que coinciden con \mathbf{x} en los bits cuyos índices pertenecen a S . Es decir, S es un certificado de f para \mathbf{x} si $\mathbf{y}|_S = \mathbf{x}|_S$ implica que $f(\mathbf{y}) = f(\mathbf{x})$.

En particular, si $f(\mathbf{x}) = 0$, un 0-certificado para \mathbf{x} es un conjunto S tal que $\mathbf{y}|_S = \mathbf{x}|_S$ implica que $f(\mathbf{y}) = 0$. Y si $f(\mathbf{x}) = 1$, un 1-certificado para \mathbf{x} es un conjunto S tal que $\mathbf{y}|_S = \mathbf{x}|_S$ implica que $f(\mathbf{y}) = 1$.

La *complejidad por certificado* de f en \mathbf{x} , que denotaremos por $C(f, \mathbf{x})$, se define como el tamaño del menor certificado de f para \mathbf{x} .

A partir de estos conceptos, vamos a definir la *complejidad por 1-certificado* de f , que denotaremos por $C_1(f)$, como $C_1(f) = \max_{\mathbf{x} \text{ tal que } f(\mathbf{x})=1} C(f, \mathbf{x})$. Análogamente, definimos la *complejidad por 0-certificado* de f , $C_0(f)$, como $C_0(f) = \max_{\mathbf{x} \text{ tal que } f(\mathbf{x})=0} C(f, \mathbf{x})$.

Por último, se define la *complejidad por certificado* de f , que representaremos por $C(f)$, como $C(f) = \max\{C_0(f), C_1(f)\} = \max_{\mathbf{x} \in \{0,1\}^n} C(f, \mathbf{x})$.

2.3. Primeras relaciones entre medidas

Después de haber introducido algunas de las medidas de complejidad de funciones booleanas más interesantes, vamos a dedicar esta sección a exponer las relaciones entre ellas que ya se conocían en el momento en que se enunció la conjetura de la sensibilidad.

El primer resultado que voy a enunciar y demostrar con detalle nos proporciona cotas inferiores para el grado y el grado aproximado de una función booleana en términos de su sensibilidad por bloques.

Lema 2.3.1. *Para cada función booleana f se satisfacen las siguientes desigualdades:*

$$\deg(f) \geq \sqrt{\frac{bs(f)}{2}}, \quad \widetilde{\deg}(f) \geq \sqrt{\frac{bs(f)}{6}}.$$

Demostración. Sea f una función booleana y sean \mathbf{x} y S_1, \dots, S_t la entrada y los conjuntos que nos permiten alcanzar el valor de la sensibilidad por bloques de f definida en 2.2.20. En particular, se tiene que S_1, \dots, S_t son conjuntos disjuntos y para cada i entre 1 y t se cumple que $f(\mathbf{x}) \neq f(\mathbf{x}^{S_i})$, de acuerdo con las notaciones que habíamos introducido antes.

Podemos suponer sin pérdida de generalidad que $f(\mathbf{x}) = 0$. Esto se debe a que si $f(\mathbf{x}) = 1$ basta considerar la función booleana $g = 1 - f$, que verifica $g(\mathbf{x}) = 0$ y $\deg(g) = \deg(f)$.

Denotando por \oplus al operador que suma bits o vectores de bits módulo 2, definimos una función $f'(y_1, \dots, y_t)$ como sigue:

$$f'(y_1, \dots, y_t) = f(\mathbf{x} \oplus y_1 S_1 \oplus \dots \oplus y_t S_t);$$

es decir, el j -ésimo bit proporcionado a f es $x_j \oplus y_i$ si j pertenece a S_i , y es x_j si j no pertenece a ninguno de los conjuntos S_i .

Tenemos que se cumple lo siguiente:

1. $\deg(f') \leq \deg(f)$, ya que los bits x_j actúan como constantes en la definición de f' y $t \leq n$.
2. f' satisface las hipótesis del Lema 2.2.15:

- $f'(0, \dots, 0) = f(\mathbf{x} \oplus 0S_1 \oplus \dots \oplus 0S_t) = f(\mathbf{x})$, y estamos trabajando con la suposición de que $f(\mathbf{x}) = 0$ sin pérdida de generalidad.
- Sea \mathbf{y}_{H_1} un vector booleano con peso de Hamming 1; es decir, \mathbf{y}_{H_1} tiene todas las entradas nulas salvo una, cuyo valor es 1.

Queremos probar que $f'(\mathbf{y}_{H_1}) = 1$.

Supongamos que la entrada no nula de \mathbf{y}_{H_1} es la i -ésima. Entonces $f'(\mathbf{y}_{H_1}) = f(\mathbf{x} \oplus 1S_i)$, donde es claro que el vector $\mathbf{x} \oplus 1S_i$ coincide, utilizando la notación anterior, con \mathbf{x}^{S_i} . Así pues,

$$f'(\mathbf{y}_{H_1}) = f(\mathbf{x}^{S_i}) \neq f(\mathbf{x}) = 0,$$

y como $f(\mathbf{x}^{S_i}) \in \{0, 1\}$ podemos concluir que $f'(\mathbf{y}_{H_1}) = 1$.

Aplicando el Lema 2.2.15 a f' , y teniendo en cuenta que $t = bs(f)$, obtenemos que

$$\deg(f') \geq \sqrt{\frac{bs(f)}{2}} \quad \text{y} \quad \widetilde{\deg}(f') \geq \sqrt{\frac{bs(f)}{6}}.$$

Teniendo por último en cuenta la primera observación de las anteriores, en virtud de la cual $\deg(f) \geq \deg(f')$, concluimos la prueba. \square

En cuanto a las dos medidas de complejidad introducidas en la Subsección 2.2.5, la complejidad por árbol de decisión $D(f)$ y la complejidad por certificado $C(f)$, podemos empezar diciendo que $C(f) \leq D(f)$. Esto se debe a que el camino que se sigue en un árbol de decisión, desde la raíz hasta una hoja, para hallar el valor de f en \mathbf{x} determina inmediatamente un certificado para \mathbf{x} , pero puede ser que existan certificados de f para \mathbf{x} de menor tamaño.

También es posible probar, aunque no entraremos en los detalles de la demostración por el trasfondo que conllevan, que se verifica el siguiente resultado. Para encontrar los detalles de la prueba, se puede consultar [11].

Proposición 2.3.2. *Para toda función booleana f se cumple que*

$$D(f) \leq C_0(f) \cdot C_1(f) \leq C(f)^2.$$

Asimismo, es posible hallar alguna relación polinómica que nos permita comparar la complejidad por certificado tanto con la sensibilidad como con la sensibilidad por bloques de una función booleana f , y a esto dedicamos los dos siguientes resultados:

Proposición 2.3.3. *Para toda función booleana f ,*

$$s(f) \leq C(f).$$

Demostración. Si $f(\mathbf{x}^{\{i\}}) \neq f(\mathbf{x})$ para algún entero $i \in [n]$, entonces i debe estar presente en cualquier certificado de f para \mathbf{x} .

En efecto, supongamos que existe un certificado S de f para \mathbf{x} al que i no pertenece. Entonces, si consideramos \mathbf{x} y $\mathbf{x}^{\{i\}}$, tenemos que ambas coinciden en las entradas cuyos índices pertenecen a S (de hecho, coinciden en todas las entradas salvo la i -ésima) y sin embargo los valores que toma f en \mathbf{x} y $\mathbf{x}^{\{i\}}$ son distintos, lo cual es absurdo.

Ahora bien, el conjunto de índices i para los cuales $f(\mathbf{x}) \neq f(\mathbf{x}^{\{i\}})$ tiene cardinal $s(f, \mathbf{x})$, y cualquier certificado de f para \mathbf{x} va a contener a dicho conjunto. Esto implica que $s(f, \mathbf{x}) \leq C(f, \mathbf{x})$ y de aquí se sigue inmediatamente la desigualdad del enunciado, sin más que tomar el máximo cuando $\mathbf{x} \in \{0, 1\}^n$ a ambos lados. \square

Teorema 2.3.4. *Para toda función booleana f ,*

$$bs(f) \geq \sqrt{C(f)} .$$

Demostración. Sea $\mathbf{x} \in \{0, 1\}^n$ tal que la complejidad por certificado de f se alcanza en \mathbf{x} ; es decir, $C(f) = \max_{\mathbf{y} \in \{0, 1\}^n} C(f, \mathbf{y}) = C(f, \mathbf{x})$. En particular, como $C(f, \mathbf{x})$ es el tamaño del menor certificado de f para \mathbf{x} , se tiene que cualquier certificado de f para \mathbf{x} tiene tamaño al menos $C(f)$.

En primer lugar, vamos a demostrar que si B es un bloque sensible y minimal de f en alguna entrada \mathbf{x} , entonces $|B| \leq s(f)$. El hecho de que B sea un bloque sensible quiere decir que $f(\mathbf{x}^B) \neq f(\mathbf{x})$, y como además B es minimal se cumple que $f(\mathbf{x}^C) = f(\mathbf{x})$ para cualquier subconjunto propio C de B .

Si $i \in B$, entonces $(\mathbf{x}^B)^{\{i\}}$ es el resultado de cambiar el valor de $|B| - 1$ bits del vector booleano \mathbf{x} (los $|B|$ que habíamos cambiado para obtener \mathbf{x}^B menos el i -ésimo bit, que es devuelto a su valor original). En consecuencia, de la minimalidad de $|B|$ se sigue que $f((\mathbf{x}^B)^{\{i\}}) = f(\mathbf{x})$. Como $f(\mathbf{x}^B) \neq f(\mathbf{x})$, llegamos a que $f((\mathbf{x}^B)^{\{i\}}) \neq f(\mathbf{x}^B)$ y en consecuencia $|B| \leq s(f, \mathbf{x}^B) \leq s(f)$.

Sean ahora B_1, \dots, B_m m subconjuntos disjuntos de $[n]$ como los que aparecían en la definición de la sensibilidad por bloques de f , de manera que $m = bs(f, \mathbf{x})$. Para cada $i \in [m]$, sustituimos B_i por un subconjunto $\tilde{B}_i \subseteq B_i$ que sea un bloque sensible minimal. Veamos que $S := \cup_{i=1}^m \tilde{B}_i$ es un certificado de f para \mathbf{x} ; es decir, que si $\mathbf{y}|_S = \mathbf{x}|_S$ entonces $f(\mathbf{y}) = f(\mathbf{x})$.

Para probar esto razonamos por reducción al absurdo, suponiendo que existe un vector booleano \mathbf{y} tal que $\mathbf{y}|_S = \mathbf{x}|_S$ y sin embargo $f(\mathbf{y}) \neq f(\mathbf{x})$. Definimos el conjunto $B^* = \{i \in [n] \text{ tales que } x_i \neq y_i\}$. Si cambiamos en \mathbf{x} los bits cuyos índices pertenecen a B^* , lo que hacemos es cambiar en \mathbf{x} todos aquellos bits que no coinciden con los de \mathbf{y} , obteniendo precisamente \mathbf{y} como resultado; es decir, $\mathbf{x}^{B^*} = \mathbf{y}$. En consecuencia, $f(\mathbf{x}^{B^*}) = f(\mathbf{y}) \neq f(\mathbf{x})$, lo que implica que B^* es un bloque sensible de f para \mathbf{x} . También, por la forma en la que se ha definido, $\emptyset = B^* \cap S = \cup_{i=1}^m (\tilde{B}_i \cap B^*)$. Esto nos conduce a que

$\tilde{B}_1, \dots, \tilde{B}_m, B^*$ son bloques sensibles disjuntos de f para \mathbf{x} , en contradicción con el hecho de que $m = bs(f, \mathbf{x})$.

Tenemos entonces que $S = \cup_{i=1}^m \tilde{B}_i$ es un certificado de f para \mathbf{x} , y gracias a que la unión anterior es disjunta y estamos suponiendo que el tamaño mínimo de cualquier certificado de f para \mathbf{x} es $C(f)$, tenemos que $|S| = \sum_{i=1}^m |\tilde{B}_i| \geq C(f)$.

Ahora, si $m = bs(f, \mathbf{x}) \geq \sqrt{C(f)}$, hemos terminado.

En caso contrario, ha de existir $i \in [m]$ tal que $|\tilde{B}_i| \geq \sqrt{C(f)}$. Si no fuera así, obtendríamos que $\sum_{i=1}^m |\tilde{B}_i| < m\sqrt{C(f)} < C(f)$, en contradicción con el párrafo anterior. Ahora bien, notemos que a partir de las definiciones de sensibilidad y sensibilidad por bloques de una función booleana f se deduce inmediatamente que $s(f) \leq bs(f)$. Esto se debe a que todos los índices para los cuales f es sensible dan lugar a bloques unipuntuales y también sensibles. Aplicando esto y lo que habíamos probado al comenzar la demostración:

$$bs(f) \geq s(f) \geq |\tilde{B}_i| \geq \sqrt{C(f)},$$

lo que nos permite concluir la prueba del teorema. \square

Dentro de la demostración anterior hemos mencionado y demostrado un resultado que, aunque sencillo, tiene una gran importancia dentro de este trabajo. Por este motivo, lo enunciamos a continuación:

Proposición 2.3.5. *Para toda función booleana f se verifica que*

$$s(f) \leq bs(f).$$

Más adelante volveremos sobre este resultado. De momento, combinándolo con el Lema 2.3.1 podemos obtener inmediatamente cotas inferiores para el grado y el grado aproximado de una función booleana en términos de su sensibilidad:

Proposición 2.3.6. *Para cada función booleana f se satisfacen las siguientes desigualdades:*

$$\deg(f) \geq \sqrt{\frac{s(f)}{2}}, \quad \widetilde{\deg}(f) \geq \sqrt{\frac{s(f)}{6}}.$$

Es posible también establecer una relación entre la complejidad por árbol de decisión de una función booleana f y su grado de la siguiente manera:

Proposición 2.3.7. *Para toda función booleana f se verifica que*

$$\deg(f) \leq D(f).$$

Demostración. Consideramos un árbol de decisión T que nos permita hallar los valores de f de manera que la profundidad de T sea $D(f)$. Sea L una hoja con valor de salida 1 y sean x_1, \dots, x_r las entradas que intervienen en el camino desde la raíz hasta la hoja L , con valores b_1, \dots, b_r en cada una.

Definimos el polinomio $P_L(\mathbf{x}) = \prod_{i:b_i=1} x_i \cdot \prod_{i:b_i=0} (1-x_i)$, que tiene grado menor o igual que la profundidad del árbol; es decir, $\deg(P_L) \leq D(f)$. Más allá de eso, $P_L(\mathbf{x}) = 1$ si la entrada \mathbf{x} nos conduce a la hoja L a lo largo del árbol y $P_L(\mathbf{x}) = 0$ en otro caso.

Por último, sea $P = \sum_{L \in \mathcal{A}} P_L$ la suma de todos los polinomios P_L como el definido antes, donde L varía en el conjunto \mathcal{A} de las hojas que proporcionan valor de salida 1. Entonces

$$\deg(P) = \max_{L \in \mathcal{A}} (\deg(P_L)) \leq D(f),$$

y $P(\mathbf{x}) = 1$ si y solo si la entrada \mathbf{x} nos permite alcanzar una hoja con valor de salida 1, lo que implica que P representa a f y por tanto $\deg(f) = \deg(P)$. \square

Notemos ahora que, por la forma en la que se han definido, todo polinomio que representa a una función booleana f es también un polinomio que aproxima f . En efecto, la distancia entre una función booleana y el polinomio multilineal que la representa es igual a 0, que es menor que $\frac{1}{3}$, cuando evaluamos ambos en el mismo vector booleano. Teniendo en cuenta que el grado aproximado de una función booleana es el mínimo de los grados de todos los polinomios que aproximan f , esto implica directamente el siguiente resultado:

Proposición 2.3.8. *Para toda función booleana f se verifica que*

$$\widetilde{\deg}(f) \leq \deg(f).$$

A continuación vamos a enunciar un lema que será útil en la demostración del siguiente teorema y que muestra que la complejidad por árbol de decisión se puede relacionar polinómicamente con la sensibilidad por bloques de la siguiente manera:

Lema 2.3.9. *Para toda función booleana f se verifica que*

$$bs(f) \leq D(f) \leq bs^4(f).$$

Demostración. En el caso de este lema, vamos a probar solo la primera desigualdad. Esto se debe a que la demostración de la segunda desigualdad nos llevaría a introducir nuevos resultados auxiliares y no queremos extendernos demasiado en este primer capítulo. En cualquier caso, es posible encontrar la demostración completa de este resultado en [13].

Para demostrar que $bs(f) \leq D(f)$, sean \mathbf{x} y B_1, \dots, B_t la entrada y los conjuntos que nos permiten alcanzar el valor de la sensibilidad por bloques de f , de manera que $t = bs(f)$.

Si probamos que cualquier árbol de decisión tiene que pedir el valor de al menos una entrada de cada uno de los bloques anteriores para conocer el valor que toma f en \mathbf{x} , ya podremos concluir.

Y esto es efectivamente cierto: para comprobarlo, supongamos que no lo fuera y que existen un valor j entre 1 y t y un árbol de decisión T que no pide el valor de ninguna entrada x_i de \mathbf{x} con $i \in B_j$. Así, cambiando en \mathbf{x} las entradas x_i cuyo índice pertenece a ese conjunto B_j , tendríamos que el valor de la función en \mathbf{x} cambiaría (por ser B_j un bloque sensible), y sin embargo el árbol de decisión T seguiría proporcionando el mismo valor, lo cual es absurdo. \square

A partir de estos resultados, y probando una cota superior para la complejidad por árbol de decisión de f en términos de su grado aproximado, ha sido posible observar, quizás sorprendentemente, que calcular una aproximación no es mucho más fácil que hallar una representación exacta de una función booleana. Esto es lo que se recoge en el siguiente resultado, que nos proporciona además la relación polinómica existente entre la complejidad por árbol de decisión de una función booleana y su grado aproximado:

Teorema 2.3.10. *Para toda función booleana f se verifica que*

$$\widetilde{\deg}(f) \leq \deg(f) \leq D(f) \leq 1296 \cdot \widetilde{\deg}(f)^8.$$

Demostración. La primera desigualdad y la segunda han sido probadas en la Proposición 2.3.8 y la Proposición 2.3.7, respectivamente.

Para probar la última de las desigualdades vamos a utilizar, por un lado, que $D(f) \leq bs^4(f)$, como ya habíamos comentado en el Lema 2.3.9. Por otro lado, del Lema 2.3.1 se deduce fácilmente (sin más que elevar al cuadrado los dos miembros de la segunda desigualdad y despejar) que $bs(f) \leq 6 \cdot \widetilde{\deg}(f)^2$. Así pues,

$$D(f) \leq bs^4(f) \leq 6^4 \cdot \widetilde{\deg}(f)^8 = 1296 \cdot \widetilde{\deg}(f)^8.$$

\square

De forma análoga, utilizando la Proposición 2.3.7, se ha conseguido probar también el siguiente teorema:

Teorema 2.3.11. *Para toda función booleana f se tiene que*

$$\deg(f) \leq D(f) \leq 16 \deg(f)^8.$$

Demostración. La primera desigualdad ha sido probada en la Proposición 2.3.7. Por tanto, basta con probar la segunda.

Para ello vamos a utilizar, por un lado, que $D(f) \leq bs^4(f)$, como ya habíamos comentado en el Lema 2.3.9. Por otro lado, del Lema 2.3.1 se deduce fácilmente (sin más que elevar al cuadrado los dos miembros de la primera desigualdad y despejar) que $bs(f) \leq 2 \cdot \deg(f)^2$. Así pues,

$$D(f) \leq bs(f)^4 \leq 2^4 \cdot \deg(f)^8 = 16 \deg(f)^8$$

y concluimos. \square

Después de haber probado todos estos resultados, vamos a introducir una última noción: se dice que dos medidas de complejidad de funciones booleanas s_1 y s_2 son equivalentes o están polinómicamente relacionadas si existen constantes $C_1, C_2 > 0$ tales que para toda función booleana f se cumple que

$$s_2(f)^{C_1} \leq s_1(f) \leq s_2(f)^{C_2},$$

salvo producto de algún término de la cadena de desigualdades anterior por una constante positiva.

Veamos qué podemos decir de las medidas de complejidad con las que estamos trabajando:

- Del Teorema 2.3.11 se deduce que el grado y la complejidad por árbol de decisión de una función booleana están polinómicamente relacionados.
- Del Teorema 2.3.10 se deduce que el grado aproximado y la complejidad por árbol de decisión de una función booleana son equivalentes.
- Del Lema 2.3.9 se deduce que la sensibilidad por bloques y la complejidad por árbol de decisión de una función booleana son equivalentes.
- De la Proposición 2.3.2 y el razonamiento del párrafo anterior a ella se sigue que

$$C(f) \leq D(f) \leq C(f)^2,$$

que implica que la complejidad por certificado y la complejidad por árbol de decisión de una función booleana son equivalentes.

- En el Teorema 2.3.4 veíamos que $\sqrt{C(f)} \leq bs(f)$. Como también hemos visto que $bs(f) \leq D(f) \leq C(f)^2$, resulta que

$$\sqrt{C(f)} \leq bs(f) \leq C(f)^2,$$

y esto implica que la sensibilidad por bloques y la complejidad por certificado de una función booleana están polinómicamente relacionadas.

- De las cadenas de desigualdades $C(f) \leq D(f) \leq C(f)^2$ y $\widetilde{\deg}(f) \leq D(f) \leq 1296 \cdot \widetilde{\deg}(f)^8$ se sigue que

$$\sqrt{\widetilde{\deg}(f)} \leq \sqrt{D(f)} \leq C(f) \leq D(f) \leq 1296 \cdot \widetilde{\deg}(f)^8,$$

con lo que

$$\sqrt{\widetilde{\deg}(f)} \leq C(f) \leq 1296 \cdot \widetilde{\deg}(f)^8$$

y esto implica que el grado aproximado y la complejidad por certificado de una función booleana están polinómicamente relacionadas.

- Análogamente, de las cadenas de desigualdades $C(f) \leq D(f) \leq C(f)^2$ y $\deg(f) \leq D(f) \leq 16 \cdot \deg(f)^8$ se sigue que

$$\sqrt{\deg(f)} \leq \sqrt{D(f)} \leq C(f) \leq D(f) \leq 16 \cdot \deg(f)^8,$$

con lo que

$$\sqrt{\deg(f)} \leq C(f) \leq 16 \cdot \deg(f)^8$$

y esto implica que el grado y la complejidad por certificado de una función booleana están polinómicamente relacionadas.

- En el Lema 2.3.1 probamos que $\sqrt{\frac{bs(f)}{6}} \leq \widetilde{\deg}(f)$. Como también hemos visto que $\widetilde{\deg}(f) \leq \deg(f) \leq D(f) \leq bs(f)^4$, tenemos que

$$\sqrt{\frac{bs(f)}{6}} \leq \widetilde{\deg}(f) \leq bs(f)^4$$

y por tanto el grado aproximado y la sensibilidad por bloques de una función booleana están polinómicamente relacionados.

- También en el Lema 2.3.1 habíamos probado que $\sqrt{\frac{bs(f)}{2}} \leq \deg(f)$. Haciendo uso de que $\deg(f) \leq D(f) \leq bs(f)^4$, tenemos que

$$\sqrt{\frac{bs(f)}{2}} \leq \deg(f) \leq bs(f)^4,$$

con lo cual el grado y la sensibilidad por bloques de una función booleana son equivalentes.

- Sabemos que $\widetilde{\deg}(f) \leq \deg(f)$. Además, combinando los resultados 2.3.7 y 2.3.10 resulta que $\deg(f) \leq D(f) \leq 1296 \cdot \widetilde{\deg}(f)^8$. En definitiva,

$$\widetilde{\deg}(f) \leq \deg(f) \leq 1296 \cdot \widetilde{\deg}(f)^8$$

y esto implica que el grado y el grado aproximado de una función booleana están polinómicamente relacionados.

Acabamos de probar que todas las medidas de complejidad introducidas en este trabajo, salvo la sensibilidad, están polinómicamente relacionadas. Sin embargo, todavía no hemos encontrado ninguna cota inferior para la

sensibilidad de una función booleana en términos de alguna de las otras medidas.

Ante esto, Nisan y Szegedy se plantearon en el artículo [14] si existiría una relación polinómica entre la sensibilidad y la sensibilidad por bloques de una función booleana, para las que ya sabemos que $s(f) \leq bs(f)$. La existencia de esta relación es lo que se conoce como conjetura de la sensibilidad y tardó aproximadamente 30 años en ser probada. De hecho, el resultado que Hao Huang probó en 2019 en [7] es que $s(f) \geq \sqrt{\deg(f)}$, y esto, además de permitirnos concluir que efectivamente la conjetura de la sensibilidad se verifica (como veremos en el siguiente capítulo), nos permite deducir junto a lo que acabamos de probar que todas las medidas de complejidad de funciones booleanas que aparecen en este trabajo están polinómicamente relacionadas.

Capítulo 3

La conjetura de la sensibilidad

3.1. Enunciado y consecuencias

Volviendo sobre el último párrafo del capítulo anterior, vamos a escribir el enunciado original de la conjetura de la sensibilidad, tal y como lo plantearon Nisan y Szegedy en 1994.

Conjetura 3.1.1 (Conjetura de la sensibilidad). *Existe una constante $C > 0$ tal que para toda función booleana f*

$$bs(f) \leq s(f)^C.$$

Antes de conocer si esta conjetura era cierta o no, ya se podía apreciar que su veracidad tendría una gran importancia. Como ya hemos comentado en el capítulo anterior, la prueba de esta conjetura nos va a permitir concluir que todas las medidas de complejidad de funciones booleanas que aparecen en este trabajo están polinómicamente relacionadas. Así, desde el punto de vista computacional, la conjetura de la sensibilidad implicaría que las funciones booleanas “suaves” (de baja sensibilidad) serían fáciles de calcular utilizando algunos de los modelos más simples, como por ejemplo los árboles de decisión. Algebraicamente, esta conjetura afirma que las funciones booleanas con estas propiedades tienen un grado bajo como polinomios reales. Por último, tal y como vamos a ver en la siguiente sección, Gotsman y Linial observaron en [6] que la prueba de esta conjetura era equivalente a resolver un problema de naturaleza puramente combinatoria dentro de la teoría de grafos.

Vamos a dedicar la siguiente sección precisamente a esto: veamos el papel que juega la teoría de grafos dentro de un problema enunciado originalmente en el marco de la teoría de la complejidad computacional.

3.2. Intervención de la teoría de grafos

A pesar de los esfuerzos realizados durante cerca de 30 años en el contexto de la teoría de la complejidad computacional para probar que la sensibilidad

y la sensibilidad por bloques de una función booleana están polinómicamente relacionadas, lo único que se había conseguido hasta 2019 eran cotas exponenciales y ejemplos que hacían sospechar que la relación entre ambas medidas tenía que ser de orden al menos cuadrático.

Vamos a mostrar a continuación, para hacernos una idea, cuáles eran las cotas superiores que se habían obtenido hasta ese momento para la sensibilidad por bloques de una función booleana en términos de su sensibilidad, todas ellas de tipo exponencial:

- $bs(f) = O(s(f) \cdot 4^{s(f)})$ (Simon, 1983).
- $bs(f) \leq \frac{e}{\sqrt{2\pi}} \cdot e^{s(f)} \cdot \sqrt{s(f)}$ (Kenyon y Kutin, 2004).
- $bs(f) \leq 2^{s(f)-1} \cdot s(f)$ (Ambainis, Bavarian, Gao, Mao, Sun y Zuo, 2013).

A pesar de que solo se obtenían cotas de tipo exponencial, algunos autores consiguieron construir ejemplos de funciones booleanas particulares para las cuales la relación entre su sensibilidad y su sensibilidad por bloques era de orden cuadrático:

- $bs(f) = \frac{1}{2}s(f)^2$ (Rubinstein, 1995).
- $bs(f) = \frac{1}{2}s(f)^2 + s(f)$ (Virza, 2011).
- $bs(f) = \frac{2}{3}s(f)^2 - \frac{1}{2}s(f)$ (Ambainis y Sun, 2011).

Esto hacía sospechar que esas cotas exponenciales obtenidas se podían mejorar hasta obtener una cota de tipo polinómico.

Es en este momento cuando entra en juego la teoría de grafos. Gotsman y Linial enuncian y demuestran un teorema en 1992 que establece una equivalencia entre dos afirmaciones: una de ellas se encuadra en el marco de la teoría de la complejidad computacional, a la que pertenece originalmente la conjetura de la sensibilidad, mientras que la otra es de naturaleza combinatoria y pertenece a la teoría de grafos. Veámoslo.

3.2.1. El teorema de equivalencia

Antes de presentar el teorema de equivalencia, conviene introducir un nuevo concepto asociado a las funciones booleanas.

Si la representación polinomial de una función booleana viene dada de la siguiente manera,

$$p(\mathbf{x}) = \sum_{I \subseteq [n]} \left[\alpha_I \prod_{i \in I} x_i \right],$$

diremos que el coeficiente α_I es la transformada de Fourier de f en I y lo denotaremos por $\hat{f}(I)$. Este concepto está claramente relacionado con el que

recibe el mismo nombre dentro de la rama del análisis matemático, sin más que observar que las funciones de la forma $\prod_{i \in I} x_i$, con $I \subseteq [n]$, forman una base ortonormal del espacio vectorial de todas las funciones $f : \{-1, 1\}^n \rightarrow \mathbb{R}$ dotado del producto interno

$$\langle f, g \rangle = \frac{1}{2^n} \sum_{\mathbf{x} \in \{-1, 1\}^n} f(\mathbf{x})g(\mathbf{x}).$$

Se pueden encontrar estos detalles en [9]. En ocasiones, y por cuestión de comodidad, denotaremos al polinomio anterior con el mismo nombre de la función booleana a la que representa.

En cuanto a la notación que vamos a utilizar dentro de este capítulo, vamos a considerar nuevamente que las funciones booleanas están definidas en $\{-1, 1\}^n$ y tienen su llegada en el conjunto $\{-1, 1\}$, en lugar de ser funciones de $\{0, 1\}^n$ en $\{0, 1\}$. Como ya hemos comentado previamente, esto no supone ningún problema. Q_n va a denotar el grafo del cubo n -dimensional $C^n = \{-1, 1\}^n$, en el que dos vértices son adyacentes si y solo si difieren exactamente en una única de sus componentes. Para un subgrafo inducido G de Q_n , vamos a denotar por $\Delta(G)$ a su grado máximo; es decir, $\Delta(G) = \max_{\mathbf{x} \in V(G)} \deg_G(\mathbf{x})$, donde $\deg_G(\mathbf{x})$ hace referencia al grado de \mathbf{x} visto como vértice del subgrafo inducido G . Finalmente, vamos a denotar por $\Gamma(G)$ a la siguiente cantidad:

$$\Gamma(G) = \max\{\Delta(G), \Delta(Q_n - G)\}.$$

Después de haber introducido las herramientas necesarias, ya nos encontramos en condiciones de enunciar y demostrar el teorema de equivalencia.

Teorema 3.2.1. *Dada una función estrictamente creciente $h : \mathbb{N} \rightarrow \mathbb{R}$, las siguientes afirmaciones son equivalentes:*

1. *Para todo subgrafo inducido G de Q_n tal que $|V(G)| \neq 2^{n-1}$, $\Gamma(G) \geq h(n)$.*
2. *Para toda función booleana f , $s(f) \geq h(\deg(f))$.*

Nota 3.2.2. Notemos que la primera de las afirmaciones hace referencia simplemente a subgrafos inducidos del cubo n -dimensional que no tienen exactamente la mitad de los vértices.

Demostración. En primer lugar, y con el fin de aproximar su contenido al de la segunda afirmación, vamos a transformar la primera de las afirmaciones en un aserto sobre funciones booleanas. Para ello, vamos a asociar a cada subgrafo inducido G de Q_n una función booleana g tal que $g(\mathbf{x}) = 1$ si y solo si $\mathbf{x} \in V(G)$.

Si recordamos la definición de sensibilidad local, teníamos que $s(g, \mathbf{x})$ era el número de coordenadas i de \mathbf{x} tales que $g(\mathbf{x}) \neq g(\mathbf{x}^{\{i\}})$. Traduciendo esto

al caso de los grafos que estamos considerando, como los vectores que difieren de \mathbf{x} en una única coordenada son precisamente los correspondientes a los vértices adyacentes a él, resulta que $s(g, \mathbf{x})$ es el número de vértices \mathbf{y} de Q_n adyacentes a \mathbf{x} tales que $g(\mathbf{x}) \neq g(\mathbf{y})$.

Sea ahora $\mathbf{x} \in V(G)$. Para este \mathbf{x} se verifica que $\deg_G(\mathbf{x})$ coincide con el número de vértices de Q_n que son adyacentes a \mathbf{x} menos el número de vértices \mathbf{y} de Q_n que son adyacentes a \mathbf{x} pero no pertenecen al subgrafo G . Estos vértices \mathbf{y} son precisamente aquellos para los cuales $g(\mathbf{y}) = -1 \neq 1 = g(\mathbf{x})$ o, en otras palabras, son aquellos vectores que corresponden a las coordenadas de sensibilidad de la función g en \mathbf{x} . Con la notación que venimos utilizando, y teniendo en cuenta que todo vértice de Q_n tiene grado n allí porque es adyacente a un vértice en cada dirección, esto se traduce en la siguiente expresión:

$$\deg_G(\mathbf{x}) = n - s(g, \mathbf{x}) \quad \forall \mathbf{x} \in V(G).$$

Un razonamiento idéntico nos permite probar que lo mismo ocurre en $Q_n - G$ para cada $\mathbf{x} \notin V(G)$:

$$\deg_{Q_n - G}(\mathbf{x}) = n - s(g, \mathbf{x}) \quad \forall \mathbf{x} \notin V(G).$$

Vamos a denotar por $E(g)$ al valor medio de la función g en C^n ; es decir,

$$E(g) = \frac{1}{2^n} \sum_{\mathbf{x} \in C^n} g(\mathbf{x}).$$

Veamos que se cumple lo siguiente:

$$E(g) = 0 \Leftrightarrow |V(G)| = 2^{n-1}.$$

Para probarlo, vamos a demostrar la doble implicación:

\Rightarrow Vamos a probar el contrarrecíproco. Para ello, supongamos que G es un subgrafo inducido de Q_n tal que $|V(G)| \neq 2^{n-1}$. Entonces, teniendo en cuenta que la función booleana g verifica que $g(\mathbf{x}) = 1$ si y solo si $\mathbf{x} \in V(G)$, se verifica que $E(g) \neq 0$. Esto se debe a que, en caso contrario,

$$\frac{1}{2^n} \sum_{\mathbf{x} \in C^n} g(\mathbf{x}) = \frac{1}{2^n} \cdot \left[\sum_{\mathbf{x} \in V(G)} g(\mathbf{x}) + \sum_{\mathbf{x} \notin V(G)} g(\mathbf{x}) \right] = 0$$

implica que

$$|V(G)| = \sum_{\mathbf{x} \in V(G)} g(\mathbf{x}) = - \sum_{\mathbf{x} \notin V(G)} g(\mathbf{x}) = -(-1) \cdot |V(Q_n) - V(G)|$$

y de aquí se sigue que $|V(G)| = |V(Q_n) - V(G)|$.

Como $|V(G)| + |V(Q_n) - V(G)| = |V(Q_n)| = 2^n$, esto nos permite concluir, junto con lo anterior, que $|V(G)| = 2^{n-1}$, en contra de lo que estábamos suponiendo.

⇐ También en este caso vamos a probar el contrarrecíproco.

Para ello, supongamos que la función booleana g verifica que $E(g) \neq 0$. Entonces, teniendo en cuenta que el subgrafo inducido G de Q_n cumple que $V(G) = \{\mathbf{x} \in C^n / g(\mathbf{x}) = 1\}$, se verifica que $|V(G)| \neq 2^{n-1}$. Esto se debe a que, en caso contrario,

$$\begin{aligned} E(g) &= \frac{1}{2^n} \sum_{\mathbf{x} \in C^n} g(\mathbf{x}) = \frac{1}{2^n} \left[\sum_{\mathbf{x} \in V(G)} g(\mathbf{x}) + \sum_{\mathbf{x} \notin V(G)} g(\mathbf{x}) \right] = \\ &= \frac{1}{2^n} \cdot [2^{n-1} \cdot 1 + 2^{n-1} \cdot (-1)] = 0, \end{aligned}$$

en contra de la hipótesis que estábamos asumiendo.

Ahora estamos en condiciones de demostrar que la primera afirmación es equivalente a:

(1') Para toda función booleana g , $E(g) \neq 0$ implica que existe $\mathbf{x} \in C^n$ tal que $s(g, \mathbf{x}) \leq n - h(n)$.

En efecto, vamos a demostrar la doble implicación:

(1) \Rightarrow (1') Sea g una función booleana con $E(g) \neq 0$. Por lo que acabamos de probar, esto implica que $|V(G)| \neq 2^{n-1}$, siendo G el subgrafo inducido de Q_n cuyos vértices son aquellos $\mathbf{x} \in C^n$ tales que $g(\mathbf{x}) = 1$. Dado que se verifica (1), se tiene que $\Gamma(G) \geq h(n)$. En consecuencia, de la definición de $\Gamma(G)$ se deduce que existe $\mathbf{x} \in C^n$ tal que $\deg(\mathbf{x})$, viendo \mathbf{x} como vértice de uno de los subgrafos inducidos G o $Q_n - G$, es mayor o igual que $h(n)$. Pero también hemos visto antes que el valor de este grado es precisamente $n - s(g, \mathbf{x})$, luego tenemos que existe $\mathbf{x} \in C^n$ tal que $n - s(g, \mathbf{x}) \geq h(n)$.

A partir de aquí concluimos que existe $\mathbf{x} \in C^n$ tal que $s(g, \mathbf{x}) \leq n - h(n)$, que es equivalente a decir que se verifica (1').

(1') \Rightarrow (1) Sea G un subgrafo inducido de Q_n tal que $|V(G)| \neq 2^{n-1}$. Consideramos la función booleana g que toma el valor 1 exactamente en aquellos \mathbf{x} pertenecientes a $V(G)$. Entonces, por lo que hemos visto antes, se verifica que $E(g) \neq 0$. Como estamos suponiendo que se verifica (1'), esto implica que existe $\mathbf{x} \in C^n$ tal que $s(g, \mathbf{x}) \leq n - h(n)$. En consecuencia, si denotamos por d el grado de \mathbf{x} visto como vértice de G o de $Q_n - G$, tenemos que $d = n - s(g, \mathbf{x}) \geq h(n)$.

A partir de aquí concluimos que $\Gamma(G) \geq h(n)$, que es equivalente a decir que se verifica (1).

Lo que hacemos a continuación es probar que la segunda afirmación es equivalente a:

(2') Para toda función booleana f , $s(f) < h(n)$ implica que $\deg(f) < n$.

Nuevamente, vamos a probar la doble implicación:

(2) \Rightarrow (2') Sea f una función booleana con $s(f) < h(n)$. Como estamos suponiendo que se verifica (2), se cumple que $\deg(f) \leq h^{-1}(s(f))$. Entonces

$$\deg(f) \leq h^{-1}(s(f)) < h^{-1}(h(n)) = n .$$

(2') \Rightarrow (2) Razonamos por reducción al absurdo, suponiendo que existe una función booleana f tal que $s(f) < h(\deg(f))$.

Como estamos suponiendo que se verifica (2'), esto implica que $\deg(f) < \deg(f)$, lo cual es absurdo.

En resumen, para demostrar el teorema basta probar que son equivalentes las siguientes afirmaciones:

(1') Para toda función booleana g , $E(g) \neq 0$ implica que existe $\mathbf{x} \in C^n$ tal que $s(g, \mathbf{x}) \leq n - h(n)$.

(2') Para toda función booleana f , $s(f) < h(n)$ implica que $\deg(f) < n$.

Para probar esto, comenzamos definiendo la función booleana $g(\mathbf{x}) = f(\mathbf{x}) \cdot p(\mathbf{x})$, donde $p(\mathbf{x})$ es la función de paridad de \mathbf{x} (en notación multiplicativa porque estamos considerando funciones booleanas definidas en $\{-1, 1\}^n$ con llegada en el conjunto $\{-1, 1\}$); es decir, $p(\mathbf{x}) = \prod_{i=1}^n x_i$.

Se cumple que $s(g, \mathbf{x}) = n - s(f, \mathbf{x})$ para cada $\mathbf{x} \in C^n$. En efecto, para cada vector \mathbf{y} que difiere de \mathbf{x} en una única coordenada se tiene que $p(\mathbf{x}) \neq p(\mathbf{y})$. Así pues, como las funciones f , g y p toman solo los valores $+1$ o -1 y $g(\mathbf{x}) = f(\mathbf{x}) \cdot p(\mathbf{x})$, si tuviéramos $f(\mathbf{x}) \neq f(\mathbf{y})$ se cumpliría que $g(\mathbf{x}) = g(\mathbf{y})$.

En consecuencia, los vectores que difieren de \mathbf{x} en una única coordenada y proporcionan valores distintos de g se obtienen a partir del conjunto de todos los vectores que difieren de \mathbf{x} en una única coordenada, que tiene cardinal n , eliminando todos aquellos vectores que proporcionan distintos valores de f . Esto se traduce en la expresión deseada:

$$s(g, \mathbf{x}) = n - s(f, \mathbf{x}).$$

También es posible demostrar que se cumple que $\hat{g}(I) = \hat{f}([n] - I)$ para todo $I \subseteq [n]$.

En efecto, si tenemos $I \subseteq [n]$ y $\hat{f}(I) \cdot \prod_{i \in I} x_i$ es un monomio de f , al multiplicar por $p(\mathbf{x}) = \prod_{i=1}^n x_i$ para obtener la función g resulta que las x_i con $i \in I$ se multiplican por sí mismas, obteniendo de esta manera $x_i^2 = 1$. Por otra parte, las x_i con $i \in [n] - I$ pasan a aparecer en este monomio en el que antes de multiplicar por $p(\mathbf{x})$ no aparecían.

En consecuencia, tras multiplicar el monomio $\hat{f}(I) \cdot \prod_{i \in I} x_i$ por $p(\mathbf{x})$ obtenemos el siguiente monomio de g : $\hat{f}(I) \cdot \prod_{i \in [n]-I} x_i$. De aquí se sigue inmediatamente que $\hat{g}([n] - I) = \hat{f}(I)$ y en consecuencia $\hat{g}(I) = \hat{f}([n] - I)$.

De esta propiedad se deduce también que $\hat{f}([n]) = \hat{g}(\emptyset)$, donde se puede probar que este último valor coincide con $E(g)$. Veámoslo:

$$\begin{aligned} E(g) &= \frac{1}{2^n} \cdot \sum_{\mathbf{x} \in C^n} g(\mathbf{x}) = \frac{1}{2^n} \cdot \sum_{\mathbf{x} \in C^n} \left(\sum_{I \subseteq [n]} \left[\hat{g}(I) \prod_{i \in I} x_i \right] \right) = \\ &= \frac{1}{2^n} \left[\sum_{\mathbf{x} \in C^n} \hat{g}(\emptyset) + \sum_{\mathbf{x} \in C^n} \left(\sum_{I \subseteq [n], I \neq \emptyset} \left[\hat{g}(I) \prod_{i \in I} x_i \right] \right) \right] = \\ &= \frac{1}{2^n} \left[2^n \cdot \hat{g}(\emptyset) + \sum_{I \subseteq [n], I \neq \emptyset} \left(\hat{g}(I) \sum_{\mathbf{x} \in C^n} \prod_{i \in I} x_i \right) \right]. \end{aligned}$$

Así, si probamos que $\sum_{\mathbf{x} \in C^n} \prod_{i \in I} x_i = 0$ para todo $I \subseteq [n]$, $I \neq \emptyset$, entonces podremos concluir que $E(g) = \hat{g}(\emptyset)$.

Sea I un subconjunto no vacío de $[n]$ de cardinal k . Entonces

$$\begin{aligned} \sum_{\mathbf{x} \in C^n} \prod_{i \in I} x_i &= \sum_{\substack{\mathbf{x} \in C^n \\ \#\{i \in I / x_i = -1\} \text{ par}}} \prod_{i \in I} x_i + \sum_{\substack{\mathbf{x} \in C^n \\ \#\{i \in I / x_i = -1\} \text{ impar}}} \prod_{i \in I} x_i = \\ &= \sum_{\substack{\mathbf{x} \in C^n \\ \#\{i \in I / x_i = -1\} \text{ par}}} 1 - \sum_{\substack{\mathbf{x} \in C^n \\ \#\{i \in I / x_i = -1\} \text{ impar}}} 1. \end{aligned}$$

La pregunta que nos hacemos ahora es: ¿cuántos elementos \mathbf{x} de C^n verifican que el cardinal del conjunto $\{i \in I / x_i = -1\}$ es par? La forma de obtener estos elementos es escoger un subconjunto de cardinal par de I en el que colocar los -1 's (como el cardinal de I es k , hay 2^{k-1} posibilidades), y colocar 1 's en el resto de coordenadas cuyos índices pertenecen a I . Para las $n - k$ coordenadas restantes (aquellas cuyos índices no corresponden a elementos de I) se pueden asignar los valores 1 o -1 indistintamente. Así, resulta que hay $2^{k-1} \cdot 2^{n-k}$ elementos \mathbf{x} de C^n que verifican que el cardinal del conjunto $\{i \in I / x_i = -1\}$ es par. Con un razonamiento idéntico, resulta que hay $2^{k-1} \cdot 2^{n-k}$ elementos \mathbf{x} de C^n que verifican que el cardinal del conjunto $\{i \in I / x_i = -1\}$ es impar. En consecuencia,

$$\sum_{\mathbf{x} \in C^n} \prod_{i \in I} x_i = 2^{n-1} \cdot 2^{n-k} - 2^{n-1} \cdot 2^{n-k} = 0,$$

que es lo que queríamos probar.

Para terminar la demostración, vamos a probar que las afirmaciones (1') y (2') son equivalentes a través de la doble implicación.

Recordemos cuáles eran estas afirmaciones:

- (1') Para toda función booleana g , $E(g) \neq 0$ implica que existe $\mathbf{x} \in C^n$ tal que $s(g, \mathbf{x}) \leq n - h(n)$.
- (2') Para toda función booleana f , $s(f) < h(n)$ implica que $\deg(f) < n$.

Veamos que son equivalentes:

- (1') \Rightarrow (2') Para probar esta implicación, comencemos observando que siempre se tiene que $\deg(f) \leq n$, luego decir que $\deg(f) < n$ equivale a decir que $\deg(f) \neq n$.

Supongamos que se verifica (1') y veamos que entonces se verifica (2') o, equivalentemente, que para toda función booleana f , $\deg(f) = n$ implica que $s(f) \geq h(n)$.

Si $\deg(f) = n$, entonces $\hat{f}([n]) \neq 0$. Como hemos probado que $\hat{f}([n]) = \hat{g}(\emptyset) = E(g)$, con $g = f \cdot p$, esto implica que $E(g) \neq 0$; y como estamos suponiendo que se verifica (1') se tiene que existe $\mathbf{x} \in C^n$ tal que $s(g, \mathbf{x}) \leq n - h(n)$.

En consecuencia, para este \mathbf{x} se verifica que $h(n) \leq n - s(g, \mathbf{x}) = s(f, \mathbf{x})$ y esto ya nos permite concluir que

$$s(f) = \max_{\mathbf{x} \in C^n} s(f, \mathbf{x}) \geq h(n),$$

que es lo que queríamos probar.

- (2') \Rightarrow (1') Para probar esta implicación, supongamos que se verifica (2') y veamos que para toda función booleana g , si se cumple que $s(g, \mathbf{x}) > n - h(n)$ para cada $\mathbf{x} \in C^n$, entonces $E(g) = 0$.

De nuestra hipótesis se deduce que $h(n) > n - s(g, \mathbf{x}) = s(f, \mathbf{x})$ para todo $\mathbf{x} \in C^n$, donde $g = f \cdot p$. En consecuencia, como C^n tiene una cantidad finita de elementos, se tiene que $s(f) = \max_{\mathbf{x} \in C^n} s(f, \mathbf{x}) < h(n)$. Como estamos suponiendo que se verifica (2'), esto implica que $\deg(f) < n$. Por lo tanto, $\hat{f}([n])$, que ya hemos visto previamente que coincide con $E(g)$, es igual a 0, que es exactamente lo que queríamos probar.

Concluimos de esta manera la demostración. □

Si sustituimos $h(\mathbf{x})$ por $\sqrt{\mathbf{x}}$, que es estrictamente creciente, en el enunciado del teorema de equivalencia, resulta que las cotas $\Gamma(G) \geq \sqrt{n}$ y $\deg(f) \leq s(f)^2$ son equivalentes.

Vamos a hablar a continuación de un resultado enunciado y demostrado por Chung, Füredi, Graham y Seymour en [3], donde construyeron un subgrafo inducido G de Q_n con $2^{n-1} + 1$ vértices tal que $\Delta(G) < \sqrt{n} + 1$. Al mismo tiempo, proporcionaron una cota inferior para $\Delta(G)$, siendo G

un subgrafo inducido de Q_n cualquiera con más de la mitad de los vértices: $\Delta(G) > \frac{1}{2} \log_2 n - \frac{1}{2} \log_2 \log_2 n + \frac{1}{2}$. En virtud del teorema de equivalencia, esto permite obtener una cota superior de tipo exponencial para el grado de una función booleana en términos de su sensibilidad, pero no es esto lo que nos interesa en la conjetura de la sensibilidad. Lo que sí que podemos ver fácilmente es que una prueba de que $\Delta(G) \geq \sqrt{n}$ implicaría inmediatamente que nuestra conjetura es cierta, y esto es precisamente lo que probó Hao Huang en [7].

Dejamos esa prueba para un poco más adelante. Primero vamos a mostrar el trabajo de Chung, Füredi, Graham y Seymour.

3.2.2. Sobre los subgrafos inducidos de un cubo

Antes de enunciar el resultado principal de esta subsección, vamos a introducir la notación con la que vamos a trabajar.

Notación 3.2.3. Recordemos que estamos denotando por Q_n el grafo del cubo n -dimensional; es decir, el conjunto de vértices de Q_n consiste en todos los vectores de $\{-1, 1\}^n$. Además, decimos que dos vectores de este conjunto son adyacentes si y solo si difieren exactamente en una componente.

Para un grafo $G = (V, E)$, estamos utilizando $\Delta(G)$ para denotar su grado máximo; es decir, $\Delta(G) = \max_{\mathbf{x} \in V(G)} \deg_G(\mathbf{x})$.

Vamos a definir ahora el grado medio de un grafo G como

$$\overline{\deg}(G) = \frac{1}{|V(G)|} \sum_{\mathbf{x} \in V(G)} \deg_G(\mathbf{x}).$$

Además, diremos que $G \in Q_n(N)$ si G es un subgrafo inducido de Q_n con N vértices; es decir, $|V(G)| = N$, $V(G) \subseteq \{-1, 1\}^n$ y $E(G) = E(Q_n) \cap (V(G) \times V(G))$.

Sabemos que Q_n es un grafo bipartito y que por lo tanto existe un subgrafo G de Q_n con exactamente la mitad de los vértices y sin ninguna arista. El principal resultado de esta sección, que está recogido en [3], muestra que incluso aunque el grado medio de un grafo $G \in Q_n(2^{n-1} + 1)$ puede ser muy pequeño, este debe tener un grado máximo elevado.

Teorema 3.2.4. *Sea G un subgrafo inducido de Q_n con al menos $2^{n-1} + 1$ vértices. Entonces existe un vértice v en G tal que*

$$\deg_G(v) > \frac{1}{2} \log_2 n - \frac{1}{2} \log_2 \log_2 n + \frac{1}{2}.$$

Por otra parte, existe un subgrafo $G \in Q_n(2^{n-1} + 1)$ con $\Delta(G) < \sqrt{n} + 1$.

Nota 3.2.5. Aunque el enunciado consta de dos partes claramente diferenciadas, solo vamos a entrar en los detalles de la demostración correspondiente

a la segunda parte. Esta es la que requiere un desarrollo más largo, y también la única que vamos a utilizar posteriormente. El resultado de Huang, que es precisamente donde vamos a utilizar la segunda de las afirmaciones del enunciado, consigue mejorar la cota que proporciona la primera parte de este teorema. Así pues, nos limitaremos a demostrar esa mejor cota en la siguiente subsección.

Demostración. Como acabamos de comentar en la Nota 3.2.5, nos vamos a limitar a detallar la prueba de la cota superior, correspondiente a la segunda parte del enunciado.

De la misma manera que hemos hecho previamente, vamos a denotar por $[n]$ al conjunto $\{1, \dots, n\}$.

Ya sabemos que el conjunto $2^{[n]}$ de aplicaciones de $\{1, \dots, n\}$ en $\{-1, 1\}$ se identifica con $\mathcal{P}([n])$, el conjunto de todos los subconjuntos de $[n]$, a través de la aplicación que envía a cada $\phi \in 2^{[n]}$ en $\{i \in [n] / \phi(i) = 1\} \in \mathcal{P}([n])$.

Del mismo modo, la aplicación $\{-1, 1\}^n \rightarrow 2^{[n]} \simeq \mathcal{P}([n])$ que envía a cada \mathbf{x} en $\{i \in [n] / x_i = 1\}$ es una biyección. Precisamente esto nos permite trabajar directamente con familias de subconjuntos finitos de $[n]$ en lugar de hacerlo con los conjuntos de vértices de determinados subgrafos de Q_n .

Afirmamos ahora que existe una partición de $[n]$ de la forma $F_1 \cup \dots \cup F_k$ tal que $|k - \sqrt{n}| < 1$ y $||F_i| - \sqrt{n}| < 1$ para todo $i \in \{1, \dots, k\}$.

Si n es un cuadrado, está claro. En este caso basta tomar $k = \sqrt{n} = |F_i|$ para todo i .

En el que caso en el que n no es un cuadrado, tenemos que

$$|k - \sqrt{n}| < 1 \Leftrightarrow -1 < k - \sqrt{n} < 1 \Leftrightarrow \sqrt{n} - 1 < k < \sqrt{n} + 1 \quad \text{y}$$

$$||F_i| - \sqrt{n}| < 1 \Leftrightarrow -1 < |F_i| - \sqrt{n} < 1 \Leftrightarrow \sqrt{n} - 1 < |F_i| < \sqrt{n} + 1.$$

Además, como k y $|F_i|$ son números naturales para cada i , deducimos que $k, |F_i| \in \{\lfloor \sqrt{n} \rfloor, \lceil \sqrt{n} \rceil\}$ para todo i .

En consecuencia, lo que buscamos es escribir $[n]$ como una unión de x conjuntos de cardinal $\lfloor \sqrt{n} \rfloor$ e y conjuntos de cardinal $\lceil \sqrt{n} \rceil$, de manera que también se satisfaga que $x + y \in \{\lfloor \sqrt{n} \rfloor, \lceil \sqrt{n} \rceil\}$. Nos encontramos entonces con el problema de teoría de números consistente en resolver la ecuación diofántica $n = x \cdot \lfloor \sqrt{n} \rfloor + y \cdot \lceil \sqrt{n} \rceil$ con las restricciones $x, y \geq 0$ y $x + y \in \{\lfloor \sqrt{n} \rfloor, \lceil \sqrt{n} \rceil\}$. No incluimos aquí los detalles de la prueba de que este problema tiene solución en nuestro caso porque su contenido no se corresponde con el del trabajo que estamos desarrollando.

Definimos ahora la familia \mathbf{X} como sigue: \mathbf{X} está formada por los conjuntos pares (es decir, los subconjuntos de $[n]$ de cardinal par) que contienen a algún F_i , con $1 \leq i \leq k$, y los conjuntos impares que no contienen a ningún F_i .

Esta familia \mathbf{X} cumple las dos propiedades siguientes:

Afirmación 3.2.6. $|\mathbf{X}| = 2^{n-1} \pm 1$ dependiendo de si $n + k$ es impar o par, respectivamente.

Afirmación 3.2.7. Para los subgrafos inducidos por \mathbf{X} y $2^{[n]} - \mathbf{X}$ se tiene que $\Delta \leq k$.

Sin embargo, es posible generalizar la construcción anterior y probar estas dos afirmaciones en un contexto más general. Es precisamente esto lo que vamos a hacer.

Sea $\mathbf{F} \subseteq 2^{[n]}$ una colección de conjuntos finitos.

Definimos ahora

$$\begin{aligned} \mathbf{X}(\mathbf{F}) = & \{S \subseteq [n] \text{ tales que } |S| \text{ es par y existe } F \in \mathbf{F} \text{ con } F \subseteq S\} \cup \\ & \cup \{S \subseteq [n] \text{ tales que } |S| \text{ es impar y } F - S \neq \emptyset \text{ para todo } F \in \mathbf{F}\}. \end{aligned}$$

Sea $G(\mathbf{F})$ el subgrafo inducido de Q_n con vértices $\mathbf{X}(\mathbf{F})$ (identificando, como ya hemos dicho al principio, subconjuntos de $[n]$ con elementos de $\{-1, 1\}^n$), y sea $G'(\mathbf{F})$ el subgrafo inducido de Q_n con vértices $2^{[n]} - \mathbf{X}(\mathbf{F})$. Estos subgrafos inducidos de Q_n van a ser los que nos permitan probar la cota superior que figura en el enunciado, generalizando al mismo tiempo lo que recoge el enunciado de la Afirmación 3.2.7.

Definimos el rango de \mathbf{F} , que se denotará por $r(\mathbf{F})$, como

$$r(\mathbf{F}) = \max_{F \in \mathbf{F}} |F|.$$

Vamos a denotar por $t(\mathbf{F})$ el máximo valor de t para el cual podemos encontrar $F_1, \dots, F_t \in \mathbf{F}$ y $x_i \in F_i$, con $1 \leq i \leq t$, tales que para $i \neq j$ se tiene que $x_i \notin F_j$. Se dice que una subfamilia de \mathbf{F} de esta forma es disjuntamente representable, de manera que $t(\mathbf{F})$ no es más que el tamaño máximo de una subfamilia de \mathbf{F} disjuntamente representable.

Veamos ahora la siguiente proposición:

Proposición 3.2.8.

$$\Delta(G(\mathbf{F})) \leq \max\{r(\mathbf{F}), t(\mathbf{F})\} \quad y$$

$$\Delta(G'(\mathbf{F})) \leq \max\{r(\mathbf{F}), t(\mathbf{F})\}.$$

Demostración. Si (S, S') es una arista de Q_n con $S, S' \in \mathbf{X}(\mathbf{F})$ y S es par, entonces $S' \subsetneq S$. En efecto, si S es par, como S' es adyacente a S resulta que S' ha de ser impar. S y S' son conjuntos que corresponden a elementos de $\{-1, 1\}^n$ que difieren en exactamente una entrada, de manera que o bien $S \subsetneq S'$ o bien $S' \subsetneq S$ (tienen todos los elementos iguales salvo uno, que corresponde a la entrada en la que difieren los vectores asociados, y este elemento solo aparecerá en uno de los dos conjuntos). Pero si $S \subsetneq S'$, como $S \in \mathbf{X}(\mathbf{F})$ es par, existe $F \in \mathbf{F}$ tal que $F \subseteq S \subsetneq S'$. Entonces $F - S' = \emptyset$, con

$F \in \mathbf{F}$, lo cual va en contra de que S' es un conjunto impar que pertenece a $\mathbf{X}(\mathbf{F})$. Queda probado así que, efectivamente, $S' \subsetneq S$.

Además de esto, si $F \subseteq S$, con $F \in \mathbf{F}$, entonces $S - S' \in F$. Efectivamente, por lo que hemos comentado en el párrafo anterior, ya sabemos que $|S - S'| = 1$. Así pues,

$$S - S' \in F \Leftrightarrow (S - S') \cap F \neq \emptyset \Leftrightarrow F - ((S - S') \cap F) \neq F.$$

Como $F - ((S - S') \cap F) = (F - (S - S')) \cup \emptyset = F - (S - S')$, llegamos a que $S - S' \in F \Leftrightarrow F - (F - (S - S')) = S - S' \neq \emptyset$, y esta última condición sabemos que sí que es cierta.

A partir de esto, podemos probar que $\deg_{G(\mathbf{F})}(S) \leq r(\mathbf{F})$. Veámoslo:

El grado de S como vértice de $G(\mathbf{F})$ es el número de conjuntos S' tales que $(S, S') \in (\mathbf{X}(\mathbf{F}) \times \mathbf{X}(\mathbf{F}))$ es una arista de Q_n , y habrá tantos conjuntos S' de esta forma como elementos $S - S'$ con las condiciones pedidas.

Como $S \in \mathbf{X}(\mathbf{F})$ es par, sabemos que existe $F_0 \in \mathbf{F}$ tal que $F_0 \subseteq S$, y en este caso ya hemos visto que $S - S' \in F_0$. En consecuencia, habrá a lo sumo tantos elementos $S - S'$ como elementos de F_0 . Esto implica que

$$\deg_{G(\mathbf{F})}(S) \leq |F_0| \leq r(\mathbf{F}) = \max_{F \in \mathbf{F}} |F|.$$

Por otra parte, si S es impar, entonces $S \subsetneq S'$. En efecto, tal y como habíamos razonado al principio de la demostración, en este caso se tiene que S' es par. De igual modo, volvemos a tener que o bien $S \subsetneq S'$ o bien $S' \subsetneq S$. Si se tuviera que $S' \subsetneq S$, entonces, por ser $S' \in \mathbf{X}(\mathbf{F})$ par, existiría $F' \in \mathbf{F}$ con $F' \subseteq S' \subsetneq S$. Esto implicaría que $F' - S = \emptyset$, con $F' \in \mathbf{F}$, lo cual va en contra de que S es un conjunto impar que pertenece a $\mathbf{X}(\mathbf{F})$.

Como $S \subsetneq S'$, con S impar y S' par, tenemos que existe $F \in \mathbf{F}$ tal que $F \subseteq S'$ pero $F \not\subseteq S$. Por lo tanto, si $S \subsetneq S'_1, S'_2, \dots, S'_m$, siendo los S'_i los conjuntos adyacentes a S en el sentido de los vértices asociados, entonces F_1, \dots, F_m (con $F_i \subseteq S'_i$ pero $F_i \not\subseteq S$) son disjuntamente representables. Veámoslo:

Por cuestiones de simetría, para ver que F_1, \dots, F_m son disjuntamente representables basta probar que existe $x_1 \in F_1$ tal que $x_1 \notin F_2$.

Como $F_1 \not\subseteq S$, existe $x_1 \in F_1$ tal que $x_1 \notin S$. Dado que $F_1 \subseteq S'_1$, tenemos que $x_1 \in S'_1 - S$, que ya hemos comentado que es un conjunto unipuntual por la adyacencia. Como los conjuntos unipuntuales $S'_1 - S, S'_2 - S, \dots, S'_m - S$ son todos distintos si lo son los S'_i , tenemos que $x_1 \notin S'_2 - S$. Como $x_1 \notin S$ y $S'_2 = (S'_2 - S) \cup (S'_2 \cap S)$, resulta que $x_1 \notin S'_2$. De la contención $F_2 \subseteq S'_2$ se sigue que $x_1 \notin F_2$, que es lo que queríamos probar.

De todo esto se deduce que $m \leq t(\mathbf{F})$. Entonces el grado de S como vértice de $G(\mathbf{F})$, que es el número de conjuntos $S' \in \mathbf{X}(\mathbf{F})$ tales que (S, S') es una arista de Q_n , coincide con el número de conjuntos $S' \in \mathbf{X}(\mathbf{F})$ tales que son “adyacentes” a S (en el sentido de los vértices asociados) y $S \subsetneq S'$.

A su vez, con la notación que estamos utilizando, este número es menor o igual que m , que acabamos de probar que es menor o igual que $t(\mathbf{F})$.

Tenemos entonces que $\Delta(G(\mathbf{F})) \leq \max\{r(\mathbf{F}), t(\mathbf{F})\}$.

La prueba de que $\Delta(G'(\mathbf{F})) \leq \max\{r(\mathbf{F}), t(\mathbf{F})\}$ es análoga. Después de haber redactado el caso anterior con todo detalle, veamos este caso brevemente:

- Si (S, S') es una arista de Q_n con $S, S' \in 2^{[n]} - \mathbf{X}(\mathbf{F})$ y S es par, entonces el hecho de que $S \notin \mathbf{X}(\mathbf{F})$ y S sea par implica que no existe ningún conjunto $F \in \mathbf{F}$ tal que $F \subseteq S$. Del mismo modo, el hecho de que $S' \notin \mathbf{X}(\mathbf{F})$ y S' sea impar implica que existe un conjunto $F \in \mathbf{F}$ tal que $F - S' = \emptyset$, que implica a su vez que $F \subseteq S'$.

De aquí se deduce que la contención que se da entre S y S' es $S \subsetneq S'$ y por tanto existe $F \in \mathbf{F}$ tal que $F \subseteq S'$ y $F \not\subseteq S$. Esto nos permite concluir como en el caso que hemos desarrollado previamente con todo detalle cuando S era impar.

- Si (S, S') es una arista de Q_n con $S, S' \in 2^{[n]} - \mathbf{X}(\mathbf{F})$ y S es impar, entonces el hecho de que $S \notin \mathbf{X}(\mathbf{F})$ y S sea impar implica que existe un conjunto $F \in \mathbf{F}$ tal que $F - S = \emptyset$, que implica a su vez que $F \subseteq S$. Del mismo modo, el hecho de que $S' \notin \mathbf{X}(\mathbf{F})$ y S' sea par implica que no existe ningún conjunto $F \in \mathbf{F}$ tal que $F \subseteq S'$.

De aquí se deduce inmediatamente que la contención que se da entre S y S' es $S' \subsetneq S$ y por tanto podemos concluir como en el caso que hemos desarrollado previamente con todo detalle cuando S era par.

De aquí se sigue análogamente que $\Delta(G'(\mathbf{F})) \leq \max\{r(\mathbf{F}), t(\mathbf{F})\}$. □

Pasamos a calcular ahora el cardinal de $\mathbf{X}(\mathbf{F})$.

Sea $F \subseteq [n]$. El número de conjuntos pares que contienen a F es

$$\begin{cases} 0 & \text{si } |F| = n, \text{ con } n \text{ impar} \\ 1 & \text{si } |F| = n, \text{ con } n \text{ par} \\ 2^{n-|F|-1} & \text{si } |F| < n. \end{cases}$$

Los dos primeros casos son evidentes, pero el último no lo es tanto. Por este motivo, vamos a desarrollarlo detalladamente, distinguiendo los casos con $|F|$ par y con $|F|$ impar.

- Supongamos que $|F| = k$ es impar. Entonces los conjuntos de cardinal par que contienen a F se corresponden con la siguiente unión:

$$\begin{aligned} & \{\text{conjuntos de cardinal } k + 1 \text{ que contienen a } F\} \cup \\ & \cup \{\text{conjuntos de cardinal } k + 3 \text{ que contienen a } F\} \cup \dots \end{aligned}$$

$\dots \cup$ {conjuntos de cardinal $(n - 1)$ o n (el que sea par)
que contienen a F }.

Ahora bien, todo subconjunto de $[n]$ de cardinal i que contiene a F se puede identificar, fijando los elementos de F , con un subconjunto de $[n] - F$ de cardinal $i - |F|$.

Así pues, identificamos los conjuntos de cardinal par que contienen a F con la unión

{subconjuntos de cardinal 1 de un conjunto de cardinal $n - |F|$ } \cup
 \cup {subconjuntos de cardinal 3 de un conjunto de
cardinal $n - |F|$ } $\cup \dots$
 $\dots \cup$ {subconjuntos de cardinal $(n - k - 1)$ o $(n - k)$ (el que sea impar)
de un conjunto de cardinal $n - |F|$ }.

Pero esta unión no es más que el conjunto de todos los subconjuntos de cardinal impar de un conjunto de cardinal $n - |F|$, de los cuales sabemos que hay exactamente $2^{n-|F|-1}$, que es lo que queríamos probar.

- Supongamos que $|F| = k$ es par. Entonces, razonando de forma idéntica a como acabamos de hacerlo, es posible observar que los conjuntos de cardinal par que contienen a F se corresponden con el conjunto de todos los subconjuntos de cardinal par de un conjunto de cardinal $n - |F|$, de los cuales sabemos también que hay exactamente $2^{n-|F|-1}$.

Razonando de la misma manera, el número de conjuntos impares que contienen a $F \subseteq [n]$ es

$$\begin{cases} 1 & \text{si } |F| = n, \text{ con } n \text{ impar} \\ 0 & \text{si } |F| = n, \text{ con } n \text{ par} \\ 2^{n-|F|-1} & \text{si } |F| < n. \end{cases}$$

Sea $\mathbf{F} = \{F_1, \dots, F_N\}$.

- Comenzamos calculando en primer lugar el cardinal de la primera parte de $\mathbf{X}(\mathbf{F})$, que recordamos que es

$$(*) = \{S \subseteq [n] \text{ tales que } |S| \text{ es par y existe } F \in \mathbf{F} \text{ con } F \subseteq S\} .$$

Como

$$\begin{aligned} (*) &= \{S \subseteq [n] \text{ tales que } |S| \text{ es par y } F_1 \subseteq S\} \cup \dots \\ &\dots \cup \{S \subseteq [n] \text{ tales que } |S| \text{ es par y } F_N \subseteq S\}, \end{aligned}$$

aplicando el principio de inclusión-exclusión resulta que

$$\begin{aligned}
|(*)| &= \sum_{i=1}^N |\{S \subseteq [n] \text{ tales que } |S| \text{ es par y } F_i \subseteq S\}| - \\
&- \sum_{\{i,j\} \subseteq [N]} |\{S \subseteq [n] \text{ tales que } |S| \text{ es par y } F_i \cup F_j \subseteq S\}| + \dots = \\
&= \sum_{i=1}^N (2^{n-|F_i|-1})^* - \sum_{\{i,j\} \subseteq [N]} (2^{n-|F_i \cup F_j|-1})^* + \dots,
\end{aligned}$$

donde

$$(2^A)^* = \begin{cases} 2^A & \text{si } A \geq 0 \\ 1 - \epsilon, \text{ con } n \equiv \epsilon \pmod{2} & \text{si } A = -1. \end{cases}$$

- Razonamos de forma análoga para calcular el cardinal de la segunda parte de $\mathbf{X}(\mathbf{F})$, que recordamos que es

$$(**) = \{S \subseteq [n] \text{ tales que } |S| \text{ es impar y } F - S \neq \emptyset \text{ para todo } F \in \mathbf{F}\}.$$

Como

$$(**) = \{S \subseteq [n] \text{ tales que } |S| \text{ es impar}\} -$$

$$- \{S \subseteq [n] \text{ tales que } |S| \text{ es impar y existe } F \in \mathbf{F} \text{ tal que } F \subseteq S\},$$

aplicando el principio de inclusión-exclusión al segundo de los conjuntos anteriores de la misma manera en la que lo hacíamos previamente resulta que

$$|(**)| = 2^{n-1} - \sum_{i=1}^N (2^{n-|F_i|-1})^{**} + \sum_{\{i,j\} \subseteq [N]} (2^{n-|F_i \cup F_j|-1})^{**} - \dots,$$

donde

$$(2^A)^{**} = \begin{cases} 2^A & \text{si } A \geq 0 \\ \epsilon, \text{ con } n \equiv \epsilon \pmod{2} & \text{si } A = -1. \end{cases}$$

Ahora, basta darse cuenta de que

$$(2^A)^* - (2^A)^{**} = \begin{cases} 0 & \text{si } A \geq 0 \\ 1 - 2\epsilon, \text{ con } n \equiv \epsilon \pmod{2} & \text{si } A = -1 \end{cases}$$

para deducir que

$$\begin{aligned}
|\mathbf{X}(\mathbf{F})| &= |(*)| + |(**)| = \\
&= 2^{n-1} + (1 - 2\epsilon) \cdot \left[\sum_{\substack{F_i \in \mathbf{F} \\ |F_i|=n}} 1 - \sum_{\substack{F_i, F_j \in \mathbf{F} \\ |F_i \cup F_j|=n}} 1 + \sum_{\substack{F_i, F_j, F_k \in \mathbf{F} \\ |F_i \cup F_j \cup F_k|=n}} 1 - \dots \right].
\end{aligned}$$

Denotamos por $f(\mathbf{F})$ a la expresión entre corchetes del último miembro de la igualdad anterior, de manera que $|\mathbf{X}(\mathbf{F})| = 2^{n-1} + (1 - 2\epsilon) \cdot f(\mathbf{F})$.

Si \mathbf{F} es una partición de $[n]$ en k conjuntos no vacíos, entonces se tiene que $|F_1 \cup \dots \cup F_k| = n$, pero el cardinal de la unión de menos de k conjuntos de \mathbf{F} es estrictamente menor que n . Así pues, en este caso

$$\begin{aligned} f(\mathbf{F}) &= \sum_{\substack{F_i \in \mathbf{F} \\ |F_i|=n}} 1 - \sum_{\substack{F_i, F_j \in \mathbf{F} \\ |F_i \cup F_j|=n}} 1 + \sum_{\substack{F_i, F_j, F_k \in \mathbf{F} \\ |F_i \cup F_j \cup F_k|=n}} 1 - \dots = \\ &= (-1)^{1+1} \sum_{\substack{F_i \in \mathbf{F} \\ |F_i|=n}} 1 + (-1)^{2+1} \sum_{\substack{F_i, F_j \in \mathbf{F} \\ |F_i \cup F_j|=n}} 1 + \dots = \\ &= (-1)^{k+1} \sum_{\substack{F_{i_1}, \dots, F_{i_k} \in \mathbf{F} \\ |F_{i_1} \cup \dots \cup F_{i_k}|=n}} 1 = (-1)^{k+1} \cdot 1 = (-1)^{k+1}. \end{aligned}$$

En consecuencia, dado que si $n \equiv \epsilon \pmod{2}$,

$$1 - 2\epsilon = \begin{cases} 1 & \text{si } n \text{ es par} \\ -1 & \text{si } n \text{ es impar} \end{cases},$$

vemos que $(1 - 2\epsilon) = (-1)^n$. Juntando todo lo que hemos visto hasta este momento:

$$|\mathbf{X}(\mathbf{F})| = 2^{n-1} + (1 - 2\epsilon) \cdot f(\mathbf{F}) = 2^{n-1} + (-1)^n \cdot (-1)^{k+1} = 2^{n-1} + (-1)^{n+k+1}$$

y queda probada la Afirmación 3.2.6 en este contexto más general.

En general, no es posible calcular explícitamente el valor de $f(\mathbf{F})$ porque esto tiende a complicarse. Sin embargo, sí que es posible probar algunas de sus propiedades. Por ejemplo:

1. Si $F_0 = [n] \in \mathbf{F}$, entonces $f(\mathbf{F}) = f(\mathbf{F} - F_0)$.

Para probar esto, basta ver que la aportación de F_0 a la expresión que permite calcular $f(\mathbf{F})$ es nula. En cada sumatorio, como $|F_0| = n$, el resto de conjuntos que intervienen a la vez que F_0 pueden ser cualesquiera. Es decir, si tenemos m conjuntos de \mathbf{F} , uno de los cuales es F_0 , su unión siempre va a tener cardinal n . En consecuencia, en el sumatorio correspondiente a tomar m conjuntos de \mathbf{F} cuya unión tenga cardinal n va a haber tantos sumandos en los que intervenga F_0 como formas de escoger otros $m - 1$ conjuntos de $\mathbf{F} - F_0$. Si llamamos k a $|\mathbf{F}|$, entonces la aportación de F_0 al cálculo de $f(\mathbf{F})$ es

$$\binom{k-1}{0} - \binom{k-1}{1} + \binom{k-1}{2} - \binom{k-1}{3} + \dots + (-1)^{k-1} \cdot \binom{k-1}{k-1}.$$

Reagrupando los términos de la expresión anterior, tenemos que este valor coincide con

$$\left[\binom{k-1}{0} + \binom{k-1}{2} + \dots \right] - \left[\binom{k-1}{1} + \binom{k-1}{3} + \dots \right],$$

que a su vez coincide con el número de subconjuntos de cardinal par de un conjunto de cardinal $k-1$ menos el número de subconjuntos de cardinal impar de un conjunto de cardinal $k-1$. Sabemos que todo conjunto tiene el mismo número de subconjuntos de cardinal par que de cardinal impar, luego el valor anterior es 0 y esto nos permite concluir que $f(\mathbf{F}) = f(\mathbf{F} - F_0)$.

2. Si $F_0 = \emptyset \in \mathbf{F}$, entonces $f(\mathbf{F}) = 0$.

Recordemos que

$$f(\mathbf{F}) = \sum_{\substack{F_i \in \mathbf{F} \\ |F_i|=n}} 1 - \sum_{\substack{F_i, F_j \in \mathbf{F} \\ |F_i \cup F_j|=n}} 1 + \dots + (-1)^{k+1} \sum_{\substack{F_{i_1}, \dots, F_{i_k} \in \mathbf{F} \\ |F_{i_1} \cup \dots \cup F_{i_k}|=n}} 1.$$

Ahora, es claro que podemos dividir cada sumatorio en dos: uno en el que uno de los F_i 's que tomamos es F_0 , que no aporta nada al cardinal de la unión de los F_i 's, y otro en el que todos los F_i 's considerados son no vacíos. Así, tenemos que

$$\begin{aligned} f(\mathbf{F}) = & \sum_{\substack{F_i \in \mathbf{F} \\ |F_i|=n \\ i \neq 0}} 1 - \sum_{\substack{F_i \in \mathbf{F} \\ |F_i|=n \\ i \neq 0}} 1 - \sum_{\substack{F_i, F_j \in \mathbf{F} \\ |F_i \cup F_j|=n \\ i, j \neq 0}} 1 + \sum_{\substack{F_i, F_j \in \mathbf{F} \\ |F_i \cup F_j|=n \\ i, j \neq 0}} 1 + \sum_{\substack{F_i, F_j, F_k \in \mathbf{F} \\ |F_i \cup F_j \cup F_k|=n \\ i, j, k \neq 0}} 1 + \dots \\ & \dots + (-1)^{k+1} \sum_{\substack{F_{i_1}, \dots, F_{i_{k-1}} \in \mathbf{F} \\ |F_{i_1} \cup \dots \cup F_{i_{k-1}}|=n \\ i_1, \dots, i_{k-1} \neq 0}} 1 + (-1)^{k+1} \sum_{\substack{F_{i_1}, \dots, F_{i_k} \in \mathbf{F} \\ |F_{i_1} \cup \dots \cup F_{i_k}|=n \\ i_1, \dots, i_k \neq 0}} 1. \end{aligned}$$

En la expresión anterior, vemos que todos los sumatorios salvo el último se cancelan. Pero el último vale 0 porque si $|\mathbf{F}| = k$, entonces al tomar k elementos de \mathbf{F} los estamos tomando todos, no podemos excluir a F_0 . Así pues, acabamos de probar que en este caso $f(\mathbf{F}) = 0$.

Para concluir la prueba de la cota superior recordemos que, como habíamos comentado al principio de la demostración, podemos tomar una partición de $[n]$ de la forma $F_1 \cup \dots \cup F_k$ tal que $\sqrt{n} - 1 < k < \sqrt{n} + 1$ y $\sqrt{n} - 1 < |F_i| < \sqrt{n} + 1$ para cada $i \in \{1, \dots, k\}$.

Consideramos $\mathbf{F} = \{F_1, \dots, F_k\}$ y construimos $\mathbf{X}(\mathbf{F})$ de la forma que hemos descrito antes. En estas condiciones, es obvio a partir de las definiciones de $r(\mathbf{F})$ y $t(\mathbf{F})$ que $r(\mathbf{F}) < \sqrt{n} + 1$ y $t(\mathbf{F}) \leq k < \sqrt{n} + 1$. En consecuencia, $\max\{r(\mathbf{F}), t(\mathbf{F})\} < \sqrt{n} + 1$.

Si $n + k$ es impar, entonces $|\mathbf{X}(\mathbf{F})| = 2^{n-1} + 1$. Así pues, si consideramos $G(\mathbf{F})$ como antes, es decir, $G(\mathbf{F})$ es el subgrafo inducido de Q_n con vértices $\mathbf{X}(\mathbf{F})$, tenemos que el número de vértices de $G(\mathbf{F})$ es precisamente $2^{n-1} + 1$. Además, de la Proposición 3.2.8 se sigue que $\Delta(G(\mathbf{F})) \leq \max\{r(\mathbf{F}), t(\mathbf{F})\}$, y acabamos de probar que este valor es estrictamente menor que $\sqrt{n} + 1$. De esta manera, el subgrafo inducido $G(\mathbf{F})$ de Q_n nos permite concluir la prueba de la cota superior en este caso.

Para terminar, si $n + k$ es par, entonces $|\mathbf{X}(\mathbf{F})| = 2^{n-1} - 1$. Así pues, si consideramos $G'(\mathbf{F})$ como antes, es decir, $G'(\mathbf{F})$ es el subgrafo inducido de Q_n con vértices $2^{[n]} - \mathbf{X}(\mathbf{F})$, tenemos que el número de vértices de $G'(\mathbf{F})$ es precisamente $2^{n-1} + 1$. Además, de la Proposición 3.2.8 se sigue que $\Delta(G'(\mathbf{F})) \leq \max\{r(\mathbf{F}), t(\mathbf{F})\}$, y acabamos de probar que este valor es estrictamente menor que $\sqrt{n} + 1$. Así pues, el subgrafo inducido $G'(\mathbf{F})$ de Q_n nos permite concluir la prueba de la cota superior en el caso en el que $n + k$ es par.

Queda probado de esta manera lo que queríamos: existe un subgrafo inducido G de Q_n con $2^{n-1} + 1$ vértices y tal que $\Delta(G) < \sqrt{n} + 1$. \square

Ahora sí, nos encontramos en condiciones de estudiar el resultado en torno al cual gira este trabajo.

3.3. Demostración de la conjetura

Vamos a dedicar esta sección a exponer detalladamente el trabajo de Hao Huang recogido en [7], donde demostró la conjetura de la sensibilidad casi 30 años después de que esta fuera enunciada.

Como acabamos de ver en la sección anterior, Chung, Füredi, Graham y Seymour probaron en [3] que si G es un subgrafo inducido de Q_n con más de la mitad de los vértices, entonces el grado máximo de G , al que denotábamos por $\Delta(G)$, es al menos $\frac{1}{2} \log_2 n - \frac{1}{2} \log_2 \log_2 n + \frac{1}{2}$. Además de esto, construyeron un subgrafo inducido de Q_n con $2^{n-1} + 1$ vértices cuyo grado máximo está acotado superior y estrictamente por $\sqrt{n} + 1$.

El resultado principal de esta sección, que nos permitirá obtener la conjetura de la sensibilidad como una consecuencia, mejora la cota de tipo logarítmico que acabamos de comentar, proporcionando de hecho la mejor cota posible (es decir, una cota que se alcanza en algún caso). Veámoslo:

Teorema 3.3.1. *Para todo entero $n \geq 1$, sea H un subgrafo inducido cualquiera de Q_n con $2^{n-1} + 1$ vértices. Entonces*

$$\Delta(H) \geq \sqrt{n}.$$

Además, esta cota es la mejor posible.

Antes de pasar a estudiar la demostración de este importante teorema, conviene señalar una observación e introducir la notación necesaria.

Nota 3.3.2. Como Q_n es un grafo bipartito, sabemos que existe un subgrafo de Q_n con exactamente la mitad de los vértices que tiene grado 0 porque está formado por vértices aislados, sin aristas.

El teorema muestra que el grado máximo de cualquier subgrafo con simplemente un vértice más aumenta repentinamente hasta \sqrt{n} .

Notación 3.3.3. Para un grafo no dirigido G , vamos a denotar por $\lambda_1(G)$ al mayor autovalor de su matriz de adyacencia, siendo la matriz de adyacencia de un grafo aquella matriz simétrica cuyas filas y columnas están indexadas en el conjunto de vértices del grafo, de manera que todos los valores de la diagonal son 0 y la entrada de la fila u y la columna v vale 1 si u y v son vértices adyacentes y 0 en caso contrario.

En cuanto a la nomenclatura utilizada: dada una matriz A , una submatriz principal de A se obtiene eliminando en A el mismo conjunto de filas y de columnas. Por ejemplo, si eliminamos la segunda y la tercera filas de A , debemos eliminar también la segunda y la tercera columnas.

Ahora sí, veamos la demostración del teorema:

Demostración. Dentro de la demostración, vamos a necesitar una serie de lemas.

El primero de ellos es un resultado clásico, consecuencia de otro resultado conocido como el principio minimax de Courant-Fischer-Weyl. Aunque no vamos a entrar en su demostración, precisamente por tratarse de un resultado clásico que no ha sido desarrollado explícitamente para nuestro propósito, escribimos a continuación su enunciado:

Lema 3.3.4 (Teorema del entrelazamiento de Cauchy). *Sea A una matriz simétrica $n \times n$ y sea B una submatriz principal $m \times m$ de A para algún $m < n$. Si los autovalores de A son $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$, y los autovalores de B son $\mu_1 \geq \mu_2 \geq \dots \geq \mu_m$, entonces para todo i con $1 \leq i \leq m$ se cumple que*

$$\lambda_i \geq \mu_i \geq \lambda_{i+n-m}.$$

Necesitamos otros dos lemas, que enunciamos y demostramos a continuación:

Lema 3.3.5. *Definimos una sucesión de matrices cuadradas simétricas de la forma recursiva siguiente:*

$$A_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad A_n = \begin{pmatrix} A_{n-1} & I \\ I & -A_{n-1} \end{pmatrix}.$$

Entonces A_n es una matriz $2^n \times 2^n$ cuyos autovalores son \sqrt{n} , de multiplicidad 2^{n-1} , y $-\sqrt{n}$, de multiplicidad 2^{n-1} .

Demostración. Para demostrar el lema, empezamos probando por inducción que $A_n^2 = nI$.

- Si $n = 1$,

$$A_1^2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I.$$

- Supongamos que el resultado es cierto para $n - 1$, es decir, que $A_{n-1}^2 = (n - 1)I$, y veamos que entonces también lo es para n .

$$\begin{aligned} A_n^2 &= A_n \cdot A_n = \begin{pmatrix} A_{n-1} & I \\ I & -A_{n-1} \end{pmatrix} \cdot \begin{pmatrix} A_{n-1} & I \\ I & -A_{n-1} \end{pmatrix} = \\ &= \begin{pmatrix} A_{n-1}^2 + I & 0 \\ 0 & I + A_{n-1}^2 \end{pmatrix} = \begin{pmatrix} (n-1)I + I & 0 \\ 0 & I + (n-1)I \end{pmatrix} = nI. \end{aligned}$$

Tenemos entonces que $A_n^2 = nI$, de manera que el único autovalor de A_n^2 es n . Ahora bien, es sencillo probar que los autovalores de una matriz B cualquiera se pueden obtener a partir de las raíces cuadradas de los autovalores de B^2 . En efecto, la ecuación que resolvemos para calcular los autovalores de la matriz B^2 es $\det(B^2 - \lambda I) = 0$, siendo λ nuestra incógnita.

Como

$$(B - \sqrt{\lambda}I) \cdot (B + \sqrt{\lambda}I) = B^2 + B\sqrt{\lambda}I - \sqrt{\lambda}IB - \lambda I = B^2 - \lambda I$$

y el determinante de un producto de matrices es el producto de los determinantes, resulta que $\det(B^2 - \lambda I) = \det(B - \sqrt{\lambda}I) \cdot \det(B + \sqrt{\lambda}I)$ y de aquí se deduce lo que queríamos.

Así pues, en nuestro caso concreto, los posibles autovalores de A_n son \sqrt{n} y $-\sqrt{n}$.

Algo que ya sabemos es que la traza de una matriz es la suma de sus autovalores. En consecuencia, como la traza de A_n es igual a 0 y los posibles autovalores son \sqrt{n} y $-\sqrt{n}$, tenemos que la mitad de los autovalores corresponden al primer valor y la otra mitad corresponden al segundo; es decir, \sqrt{n} tiene multiplicidad $\frac{2^n}{2} = 2^{n-1}$, y lo mismo ocurre con $-\sqrt{n}$. \square

Lema 3.3.6. *Supongamos que H es un grafo no dirigido con m vértices y A es una matriz simétrica cuyas entradas pertenecen al conjunto $\{-1, 0, 1\}$ y cuyas filas y columnas están indexadas por $V(H)$, y siempre que u y v no son adyacentes en H , $A_{u,v} = 0$ (a veces se dice que A es una matriz de adyacencia signada o con signo del grafo H). Entonces $\Delta(H) \geq \lambda_1 := \lambda_1(A)$ (el mayor autovalor de A).*

Demostración. Supongamos que \mathbf{v} es el autovector correspondiente al autovalor λ_1 . Entonces $A\mathbf{v} = \lambda_1\mathbf{v}$. Si λ_1 fuese igual a 0, entonces el resultado es trivial. Por lo tanto, vamos a trabajar en la demostración con $\lambda_1 \neq 0$, y también con $\mathbf{v} \neq \mathbf{0}$.

Supongamos sin pérdida de generalidad que v_1 es la coordenada de \mathbf{v} con mayor valor absoluto. Entonces

$$|\lambda_1 v_1| = |(A\mathbf{v})_1| = \left| \sum_{j=1}^m A_{1,j} v_j \right| \leq \sum_{j=1}^m |A_{1,j}| |v_j| \leq |v_1| \sum_{j=1}^m |A_{1,j}|.$$

Además, por la forma en la que se ha definido la matriz A , se tiene que la suma de los valores absolutos de las entradas de la primera fila es igual al grado del vértice de H correspondiente a esa fila y, en consecuencia, menor o igual que el grado máximo de H . Por lo tanto,

$$|\lambda_1| |v_1| = |\lambda_1 v_1| \leq |v_1| \Delta(H),$$

y como $v_1 \neq 0$ obtenemos que efectivamente $|\lambda_1| \leq \Delta(H)$. \square

Volvamos, ahora sí, a la demostración del teorema.

Sea A_n la n -ésima de las matrices de la sucesión definida en el Lema 3.3.5, matrices todas ellas cuyas entradas pertenecen al conjunto $\{-1, 0, 1\}$. Gracias a la construcción recursiva de estas matrices, es claro que si cambiamos en A_n aquellas entradas que toman el valor -1 por 1 obtenemos exactamente la matriz de adyacencia del cubo n -dimensional Q_n . En efecto, podemos ver los bloques superior izquierdo e inferior derecho de esta matriz como las matrices de adyacencia de los dos subcubos $(n-1)$ -dimensionales de Q_n , y los dos bloques formados por la matriz identidad como los correspondientes a las aristas que conectan estos dos subcubos. En consecuencia, la matriz A_n y el grafo Q_n satisfacen las hipótesis del Lema 3.3.6. Pero también un subgrafo inducido H de Q_n con $2^{n-1} + 1$ vértices y la submatriz principal A_H de A_n inducida de forma natural por H las satisfacen. Como consecuencia, $\Delta(H) \geq \lambda_1(A_H)$.

Por otra parte, gracias al Lema 3.3.5, sabemos que los autovalores de A_n son \sqrt{n} y $-\sqrt{n}$, ambos con multiplicidad 2^{n-1} .

Notemos que, como A_H es la submatriz de A_n inducida de forma natural por el subgrafo H de Q_n , que tiene $2^{n-1} + 1$ vértices, resulta que A_H es una submatriz de dimensiones $(2^{n-1} + 1) \times (2^{n-1} + 1)$ de la matriz A_n , que tiene dimensiones $2^n \times 2^n$. En consecuencia, aplicando el Teorema del entrelazamiento de Cauchy y teniendo en cuenta que $2^n - (2^{n-1} + 1) = 2^{n-1} - 1$, tenemos que

$$\lambda_1(A_H) \geq \lambda_{1+2^{n-1}-1}(A_n) = \lambda_{2^{n-1}}(A_n) = \sqrt{n}.$$

De esta manera, combinando las dos desigualdades que hemos obtenido gracias a los lemas previos, podemos concluir que

$$\Delta(H) \geq \lambda_1(A_H) \geq \sqrt{n},$$

que es lo que queríamos probar.

Finalmente, para el subgrafo inducido de Q_n con $2^{n-1} + 1$ vértices construido por Chung, Füredi, Graham y Seymour cuya construcción hemos presentado detalladamente en la subsección anterior, la cota se alcanza. Efectivamente, el grado máximo de ese grafo está acotado superior y estrictamente por $\sqrt{n} + 1$ o, lo que es lo mismo, es menor o igual que \sqrt{n} . Como acabamos de probar que este grado máximo es también mayor o igual que \sqrt{n} , ya podemos concluir que se da la igualdad y por lo tanto la cota se alcanza. \square

Volvamos, para terminar, al problema de la conjetura de la sensibilidad.

Acabamos de probar que para todo subgrafo inducido H de Q_n con $2^{n-1} + 1$ vértices se verifica que $\Delta(H) \geq \sqrt{n}$. Así pues, si consideramos cualquier subgrafo inducido G de Q_n con $|V(G)| \neq 2^{n-1}$, o bien G o bien $Q_n - G$ será un grafo con al menos $2^{n-1} + 1$ vértices. Como la función Δ es monótona (en el sentido de que $\Delta(G_1) \leq \Delta(G_2)$ si G_1 es un subgrafo de G_2), el teorema que acabamos de probar nos permite afirmar que $\Gamma(G) = \max\{\Delta(G), \Delta(Q_n - G)\} \geq \sqrt{n}$, y esto no es más que la primera de las afirmaciones del enunciado del Teorema de equivalencia 3.2.1 tomando como h la función raíz cuadrada, que es estrictamente creciente. En virtud de este teorema, lo que acabamos de probar es equivalente a que toda función booleana f verifica que $s(f) \geq \sqrt{\deg(f)}$, obteniendo de esta manera una cota inferior para la sensibilidad de una función booleana en términos de otra de las medidas de complejidad que hemos considerado a lo largo del trabajo.

Para poder relacionar la sensibilidad y la sensibilidad por bloques de una función booleana, que es como se había enunciado originalmente la conjetura de la sensibilidad, basta recordar que, como vimos en el primer capítulo, el grado de una función booleana y su sensibilidad por bloques están polinómicamente relacionados. En particular, habíamos probado que

$$\sqrt{\frac{bs(f)}{2}} \leq \deg(f) \leq bs(f)^4.$$

Combinando esto con la desigualdad que acabamos de obtener, resulta que

$$bs(f) \leq 2 \cdot \deg(f)^2 \leq 2 \cdot s(f)^4,$$

lo que concluye definitivamente la prueba de la conjetura de la sensibilidad.

Aunque en el primer capítulo nos centramos en demostrar la desigualdad $bs(f) \leq 2 \cdot \deg(f)^2$, que es la que probaron Nisan y Szegedy en su trabajo publicado en 1994, esta desigualdad fue mejorada por Tal en un trabajo publicado en 2013, en el que demostró que $bs(f) \leq \deg(f)^2$.

Así pues, la relación entre la sensibilidad y la sensibilidad por bloques de una función booleana f se puede mejorar también para obtener

$$bs(f) \leq s(f)^4.$$

Aunque esta es la mejor cota inferior para la sensibilidad de una función booleana en términos de su sensibilidad por bloques que se conoce hasta el

momento, la función de Rubinstein a la que hemos dedicado el Apéndice A, junto con otras funciones booleanas concretas que se han podido construir, sugiere que la separación real entre estas dos medidas de complejidad es de orden al menos cuadrático y no cuártico.

Capítulo 4

Nuevas líneas de investigación

La demostración de la conjetura de la sensibilidad ha abierto nuevas líneas de investigación, de las cuales vamos a hablar brevemente en este capítulo.

En primer lugar, y como ya hemos comentado al final del capítulo anterior, construyendo funciones booleanas específicas se ha obtenido una separación de, al menos, orden cuadrático entre la sensibilidad y la sensibilidad por bloques de ciertas funciones booleanas. Varios autores sugieren que esta es la mayor separación que puede haber entre ambas medidas de complejidad. No obstante, la cota superior que hemos obtenido para la sensibilidad por bloques en términos de la sensibilidad de una función booleana es de orden cuatro y procede de una relación ajustada entre la sensibilidad y el grado que se obtiene combinando el resultado de Huang con el teorema de equivalencia. Además, todavía no se ha conseguido probar que la separación de orden cuadrático sea cierta en general. Actualmente, el autor de la demostración de la conjetura de la sensibilidad cree que esta separación de orden cuadrático ha de ser cierta en general (es decir, que la sensibilidad por bloques de una función booleana está acotada, salvo constantes, por el cuadrado de su sensibilidad) y que este resultado podría probarse utilizando técnicas del álgebra lineal, aunque en estos momentos es un problema abierto.

Otra de las preguntas que Huang plantea tras su demostración es la siguiente: si G es un grafo con “muchas” simetrías y denotamos por $\alpha(G)$ su número de independencia, ¿qué podemos decir acerca del mínimo del conjunto de grados máximos de subgrafos inducidos de G con $\alpha(G) + 1$ vértices? Si denotamos por $f(G)$ a este valor, ¿en qué condiciones podremos proporcionar una cota inferior para $f(G)$ que se alcance? Antes de comentar brevemente los resultados obtenidos en relación a estas dos preguntas, conviene precisar el significado de algunos de los términos que aparecen en ellas.

Dado un grafo G , se dice que G es *simétrico* si, dados dos pares (u_1, u_2) y (v_1, v_2) de vértices adyacentes, existe un automorfismo $f : V(G) \rightarrow V(G)$ tal que $f(u_1) = v_1$ y $f(u_2) = v_2$. Sin embargo, los grafos con “muchas” simetrías

a los que se refiere Huang en su pregunta no han de ajustarse exactamente a esta definición, sino que la idea que busca consiste en generalizar este concepto, que sí que es preciso en el caso del hipercubo n -dimensional con el que hemos estado trabajando, a otros tipos de grafos con propiedades interesantes para nuestro fin.

Además, se llama *conjunto independiente de G* a un conjunto de vértices de G tales que no existen entre ellos dos vértices adyacentes, y *número de independencia de G* al cardinal del mayor conjunto independiente. Es fácil apreciar que esto busca generalizar la situación a la que Huang se enfrenta en su resultado, con el grafo simétrico Q_n y sus subgrafos inducidos con $2^{n-1} + 1$ vértices, valor que coincide en este caso con $\alpha(Q_n) + 1$.

Gran parte de los trabajos que se han publicado en relación con estas dos preguntas consideran a los grafos de Cayley como grafos con suficiente simetría que permitan generalizar los resultados sobre hipercubos n -dimensionales que hemos probado a lo largo del trabajo. Dado un grupo G y un subconjunto $S \subseteq G$, se define *el grafo de Cayley de G con respecto a S* , que denotaremos por $Cay(G, S) = (V, E)$, como el grafo cuyos vértices son elementos de G y

$$\{x, y\} \in E \Leftrightarrow xy^{-1} \in S \text{ o } yx^{-1} \in S.$$

Aunque existen artículos previos en los que se prueba que el valor que hemos denotado por $f(G)$ no puede ser acotado inferiormente por una función del grado para grafos de Cayley en general, vamos a destacar uno de los más recientes: ha sido publicado en 2022 y sus autores son Ignacio García-Marco y Kolja Knauer ([5]), que abordan distintos problemas en relación con este tema. En primer lugar, proporcionan 3 familias de grafos de Cayley que tienen grado no acotado pero tales que $f(G) = 1$ para todos sus miembros G . Sin entrar en detalles, estas familias son: grafos de Cayley del grupo diédrico; grafos de Cayley del grupo simétrico S_n con respecto a todas las trasposiciones que contienen al 1; y grafos de Cayley d -regulares con n vértices que tienen un subgrafo inducido de grado máximo 1 y con $\frac{d}{2d-1}n$ vértices.

Además, en relación con la pregunta de si se puede dar una cota inferior para $f(G)$ que se alcance, los autores enuncian y demuestran como consecuencia del resultado de Huang que si un grafo de Cayley bipartito G tiene como subgrafo inducido a un hipercubo de dimensión máxima $\kappa(G)$, entonces $f(G) \geq \sqrt{\kappa(G)}$. A continuación, obtienen infinitas familias de grafos de Cayley con el valor κ no acotado para las cuales la cota inferior anterior se alcanza. Para ello, estudian los grupos de Coxeter, sobre los cuales simplemente vamos a dar aquí la definición. Un *sistema finito de Coxeter* es un par (W, S) , donde W es un grupo con generadores $S = \{a_1, \dots, a_n\}$ y presentación $W = \langle a_1, \dots, a_n \mid (a_i a_j)^{m_{ij}} = 1 \rangle$, donde $m_{ij} > 1$ y $m_{ii} = 2$. Es precisamente W quien recibe el nombre de *grupo de Coxeter*. Dentro de este estudio, los autores conjeturan que para todo grafo de Cayley G de un grupo

de Coxeter se verifica que $f(G) = \lceil \sqrt{\kappa(G)} \rceil$.

Más allá de esto, y volviendo al problema del hipercubo, Huang planteó una nueva pregunta. Si $g(n, k)$ es el mínimo valor t tal que todo subgrafo inducido H de Q_n con t vértices tiene grado máximo al menos k , en este trabajo hemos demostrado que $g(n, \sqrt{n}) = 2^{n-1} + 1$. Huang sugiere que podría resultar interesante hallar el valor de $g(n, k)$ para distintos valores de k .

Después de todo, podemos concluir que el tema tratado es de gran actualidad y que se espera que en los próximos años se convierta en un área de trabajo muy prolífica.

Apéndice A

La función de Rubinstein

Como hemos visto en el Capítulo 2, siempre se verifica que la sensibilidad de una función booleana es menor o igual que su sensibilidad por bloques. Además, probamos que para la función de conjunción AND se cumple que ambos valores coinciden.

Vamos a dedicar este apéndice a la construcción de una función booleana para la cual se verifica la desigualdad estricta. Esta función fue construida por primera vez por David Rubinstein en [16], y por este motivo recibe el nombre de *función de Rubinstein*.

La forma en que vamos a construir esta función no es exactamente la forma original en la que lo hizo Rubinstein, sino que se trata de una adaptación más sencilla que Hao Huang presentó en una charla para el *Simons Institute for the Theory of Computing* de la Universidad de California ([8]).

El conjunto de partida de nuestra función booleana, a la que vamos a denotar por f , va a ser $\{0, 1\}^{n^2}$. Así, nos va a resultar más cómodo representar cada vector booleano como una matriz de la siguiente manera:

$$\mathbf{x} = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nn} \end{pmatrix},$$

donde cada x_{ij} puede tomar solamente los valores 0 o 1. Así, si denotamos por \bigvee a la función OR de disyunción con la que ya hemos trabajado antes, vamos a definir la función booleana $f : \{0, 1\}^{n^2} \rightarrow \{0, 1\}$ como

$$f \left(\begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nn} \end{pmatrix} \right) = \bigvee_{i=1}^n g(x_{i1}, \dots, x_{in}),$$

donde $g : \{0, 1\}^n \rightarrow \{0, 1\}$ es una función booleana que toma el valor 1 en $\mathbf{x} = (x_1, \dots, x_n)$ si y solamente si existe un único valor j entre 1 y $n - 1$ tal que $x_j = x_{j+1} = 1$ y $x_k = 0$ para todas las demás coordenadas.

A.1. Sensibilidad y sensibilidad por bloques

Empecemos probando que la sensibilidad por bloques de f es al menos del orden de n^2 . Para ello, vamos a prestar atención al valor de la sensibilidad local por bloques de f en $\mathbf{0} \in \{0,1\}^{n^2}$, teniendo en cuenta que $f(\mathbf{0})$ es claramente 0. En primer lugar, notemos que si en una fila cambiamos el valor de dos coordenadas nulas consecutivas por 1, el valor que toma g en esa fila pasa también a ser 1 y lo mismo ocurre con el valor de f . Así pues, vamos a distinguir ahora dos casos:

- Si n es par, en cada fila es posible tomar un máximo de $\frac{n}{2}$ bloques sensibles y disjuntos. Si nos fijamos en la fila i -ésima, se trata de los bloques correspondientes a los índices (i, j) e $(i, j + 1)$, con j impar.

Como hay un total de n filas, resulta que la sensibilidad local por bloques de f en $\mathbf{0}$ en el caso de n par es $n \cdot \frac{n}{2} = \frac{n^2}{2}$.

- Si n es impar, en cada fila es posible tomar un máximo de $\frac{n-1}{2}$ bloques sensibles y disjuntos. Por ejemplo, si consideramos la fila i -ésima, podemos considerar los bloques correspondientes a los índices (i, j) e $(i, j + 1)$, con j impar, de manera que la última entrada de cada fila no pertenezca a ningún bloque.

Como hay un total de n filas, resulta que la sensibilidad local por bloques de f en $\mathbf{0}$ en el caso de n impar es $n \cdot \frac{n-1}{2} = \frac{n^2-n}{2}$.

Entonces $bs(f) \geq bs(f, \mathbf{0}) \sim n^2$.

Veamos ahora que $s(f) = O(n)$.

Nota A.1.1. Escribimos $f = O(g)$ y leemos “ f es O grande de g cuando n tiende a ∞ ” si existe una constante $K > 0$ tal que para todo n a partir de un cierto n_0 se cumple que $|f(n)| \leq K \cdot |g(n)|$.

Esta notación recibe el nombre de O grande de Landau en honor a Edmund Landau, matemático alemán que la introdujo.

Vamos a considerar el valor de la sensibilidad local de f en dos casos distintos.

- En primer lugar, consideramos aquellos \mathbf{x} tales que $f(\mathbf{x}) = 0$. En este caso, tenemos que el valor de la función g en cada fila de nuestra matriz booleana es 0.

Si la fila en cuestión tiene todas sus coordenadas nulas, no existe ninguna entrada sensible en ella, ya que cambiar cualquiera de las entradas por 1 no cambia del mismo modo el valor que toma g en la fila.

Sin embargo, si la fila sí tiene coordenadas no nulas, entonces tendrá como mucho 2 coordenadas sensibles. Esto corresponderá al caso en

el que la fila sea de la forma $(0 \dots 0 1 0 \dots 0)$, con una única entrada no nula, y las coordenadas sensibles serán las que se encuentran a la izquierda y a la derecha de esa coordenada no nula.

Así pues, en cada fila tendremos a lo sumo 2 coordenadas sensibles. Como hay un total de n filas, esto nos permite concluir que $s(f, \mathbf{x}) \leq 2n$.

- Si consideramos ahora los \mathbf{x} tales que $f(\mathbf{x}) = 1$, es posible distinguir de nuevo dos posibilidades:
 - Si existen al menos dos filas en las que g toma el valor 1, entonces la sensibilidad local de f en \mathbf{x} es 0, ya que modificando una entrada de una fila solo podríamos cambiar el valor que toma g en esa fila, mientras que en la otra seguiría tomando el valor 1 y en consecuencia f también.
 - Si g toma el valor 1 en una única fila, pongamos la i -ésima, cambiar una entrada de cualquier otra fila no afecta al valor de f . Así pues, todas las coordenadas sensibles se encuentran en la fila i -ésima, que será de la forma $(0 \dots 0 1 1 0 \dots 0)$ porque g toma en ella el valor 1. En consecuencia, todas las entradas de esta fila corresponden a coordenadas sensibles y resulta que $s(f, \mathbf{x}) = n$.

Hemos probado entonces que $s(f, \mathbf{x}) \leq 2n$ para cada $\mathbf{x} \in \{0, 1\}^{n^2}$ y por tanto $s(f) \leq 2n$, de manera que, como habíamos anunciado, $s(f) = O(n)$.

Así pues, tenemos que para la función de Rubinstein f que hemos definido se cumple que $bs(f)$ es al menos del orden de n^2 , mientras que $s(f) = O(n)$. Esto nos muestra que en este caso se va a verificar la desigualdad estricta $s(f) < bs(f)$ y sugiere que la separación entre ambos valores podría ser de, al menos, orden cuadrático.

Bibliografía

- [1] S. ARORA Y J. KRAMER, *Lecture 17: Decision Tree Complexity*, Computer Science 522: Computational Complexity Theory, Princeton University, 2003.
<https://users.cs.duke.edu/~reif/courses/complectures/Arora/lec17.pdf> (13/06/2022)
- [2] E. W. CHENEY, *Introduction to approximation theory*. McGraw-Hill Book Co., 1966.
- [3] F. R. K. CHUNG, Z. FÜREDI, R. L. GRAHAM Y P. SEYMOUR, *On induced subgraphs of the cube*, Journal of Combinatorial Theory, Series A **49**, pp. 180-187, 1988.
- [4] H. EHLICH Y K. ZELLER, *Schwankung von Polynomen zwischen Gitterpunkten*, Mathematische Zeitschrift **86**, pp. 41-44, 1964.
- [5] I. GARCÍA-MARCO Y K. KNAUER, *On sensitivity in bipartite Cayley graphs*, Journal of Combinatorial Theory, Series B **154**, pp. 211-238, 2022.
- [6] C. GOTSMAN Y N. LINIAL, *The equivalence of two problems on the cube*, Journal of Combinatorial Theory, Series A **61**, pp. 142-146, 1992.
- [7] H. HUANG, *Induced subgraphs of hypercubes and a proof of the sensitivity conjecture*, Annals of Mathematics **190**(3), pp. 949-955, 2019.
- [8] H. HUANG, *Sensitivity Conjecture and Its Applications*, Simons Institute, YouTube, 22/07/2020.
<https://www.youtube.com/watch?v=EJoe4qH6kLs&t=3107s>
(13/06/2022)
- [9] S. JUKNA, *Analysis of Boolean Functions*, Boolean Function Complexity, Algorithms and Combinatorics **27**, pp. 55-77, Springer, Berlin, Heidelberg, 2012.
- [10] J. KAHN, G. KALAI Y N. LINIAL, *The influence of variables on Boolean functions*, [Proceedings 1988] 29th Annual Symposium on Foundations of Computer Science, pp. 68-80, 1988.

- [11] S. KOPPARTY, E. CAI Y Z. LANGLEY, *Lecture 3: Boolean Function Complexity Measures*, Advanced Topics in Algorithms and Complexity, Rutgers University, 2020.
<https://sites.math.rutgers.edu/~sk1233/courses/topics-S20/lec3.pdf> (13/06/2022)
- [12] M. MINSKY Y S. PAPERT, *Perceptrons*, expanded edition: An introduction to Computational Geometry. The MIT press, Cambridge, 1988. La primera edición se publicó en 1969.
- [13] N. NISAN, *CREW PRAM's and decision trees*, STOC '89: Proceedings of the twenty-first annual ACM symposium on Theory of computing, pp. 327-335, 1989.
- [14] N. NISAN Y M. SZEGEDY, *On the degree of Boolean functions as real polynomials*, Computational complexity **4**, pp. 301-313, 1994.
- [15] T. J. RIVLIN Y E. W. CHENEY, *A comparison of Uniform Approximations on an interval and a finite subset thereof*, SIAM Journal on Numerical Analysis **3**(2), pp. 311-320, 1966.
- [16] D. RUBINSTEIN, *Sensitivity vs. block sensitivity of Boolean functions*, Combinatorica **15**, pp. 297-299, 1995.