

# Puntos enteros en politopos

Alexey Beshenov (cadadr@gmail.com)

Agosto 2016

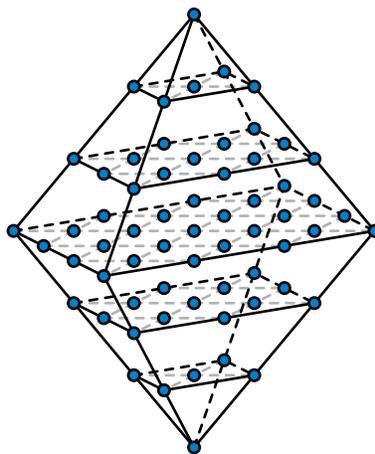
Estos apuntes acompañaron un minicurso impartido en agosto de 2016 en la Universidad de El Salvador.

En teoría, el curso estaría dedicado al “álgebra conmutativa combinatoria”, una rama de matemáticas explorada en los años 70 por Mel Hochster y Richard Stanley. Al principio, mi motivación era entender la relación entre la *teoría de Ehrhart* y los *anillos de Cohen–Macaulay*, explicada, por ejemplo, en el libro “Cohen–Macaulay rings” de Winfried Bruns y Jürgen Herzog. Sin embargo, mis alumnos se dividieron en dos grupos distintos: a unos les interesaba el aspecto geométrico y combinatorio, mientras que los otros seguían mis clases más abstractas dedicadas al álgebra homológica y conmutativa. Por el momento no tengo planeado redactar en  $\text{\LaTeX}$  mis apuntes completos, pero decidí compartir el presente texto, que contiene la parte elemental del curso.

Como fuente principal para la parte I, he usado el libro “Konvexe Mengen” de Kurt Leichtweiß. En la parte II y III, he seguido el libro “Computing the continuous discretely: integer-point enumeration in polyhedra” de Matthias Beck y Sinai Robins (<http://math.sfsu.edu/beck/ccd.html>).

## Agradecimientos

Agradezco al Ministerio de Educación de El Salvador por haber financiado mi estancia y personalmente al Ministro de Educación, el Ing. CARLOS MAURICIO CANJURA LINARES; asimismo al Director de la Escuela de Matemática de la Universidad de El Salvador, el Dr. JOSÉ NERYS FUNES TORRES y al director del Programa “Jovenes Talento”, el Lic. ERNESTO AMÉRICO HIDALGO CASTELLANOS por haberme invitado a dar clases. El Profesor GABRIEL ALEXANDER CHICAS REYES se ocupó de la organización de mis lecciones y me ayudó con la redacción de estos apuntes.



# Índice

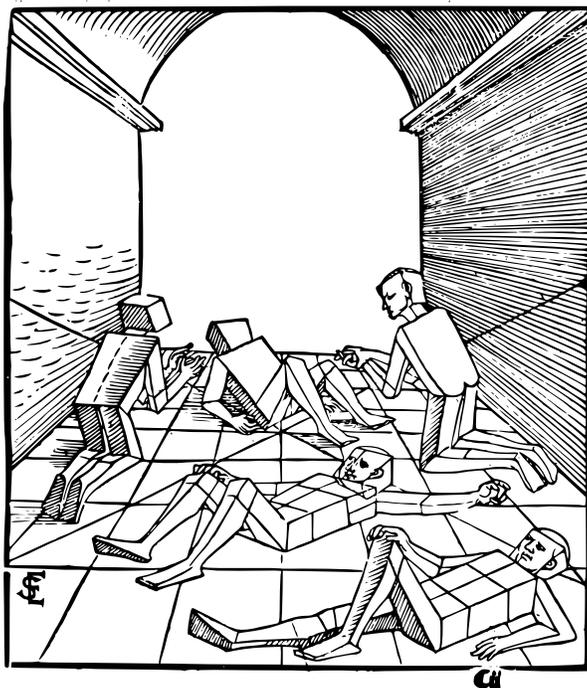
<b>I Poliedros convexos</b>	<b>3</b>
1 Conjuntos convexos .....	4
2 La envolvente convexa .....	12
3 Separación por hiperplanos .....	15
4 Conjuntos polares .....	18
5 Poliedros convexos .....	22
6 Politopos convexos .....	23
7 Subconjuntos extremos y caras de poliedros .....	24
8 Hiperplanos de soporte .....	26
9 Puntos extremos y el teorema de Krein–Milman .....	28
10 Teorema de Minkowski–Weyl.....	29
11 Poliedros en SAGE.....	30
<b>II Polinomios de Ehrhart</b>	<b>35</b>
12 Teorema de Pick.....	35
13 Polinomios de Ehrhart .....	39
14 Polinomios de Ehrhart en SAGE.....	46
<b>III Cuadrados mágicos</b>	<b>47</b>
15 Un poco de historia .....	47
16 Un ejemplo de enumeración: $H_3(m)$ .....	51
17 Politopos de Birkhoff.....	53
18 Teorema de los matrimonios de Hall .....	56
19 Teoría de Ehrhart para los politopos de Birkhoff.....	59
20 Politopos de Birkhoff en SAGE .....	60

# Parte I

## Poliedros convexos

En esta parte voy a explicar qué es un politopo convexo, el objeto principal de nuestro curso. Intuitivamente, todo el mundo conoce de qué se trata: es un subconjunto acotado en  $\mathbb{R}^n$  formado por intersección de un número finito de semiespacios, que es la misma cosa que la *envolvente convexa* de un número finito de puntos (sus vértices). Desde el punto de vista *topológico*, conjuntos convexos son poco interesantes: son espacios contraíbles. Lo que nos interesa es la teoría *combinatoria*.

Aquí voy a recordar qué es la envolvente convexa y también demostrar que las dos definiciones—con semiespacios y la envolvente convexa—son equivalentes. Como siempre sucede en geometría, es algo intuitivamente obvio, pero no es tan fácil de demostrar. En cierto sentido, las dos definiciones son duales, y la dualidad está relacionada con la noción del *conjunto polar*.



Erhard Schön, Cinco figuras en un edificio (1538)

Los objetos geométricos que nos interesan en esta historia son subconjuntos de  $\mathbb{R}^n$ . Voy a denotar los puntos de  $\mathbb{R}^n$  por letras  $\mathbf{x}, \mathbf{y}, \mathbf{z}, \dots$  en negrita y las letras cursivas van a denotar las coordenadas, por ejemplo  $\mathbf{x} = (x_1, \dots, x_n)$ , donde  $x_i \in \mathbb{R}$ . (En la pizarra en lugar de “ $\mathbf{x}$ ” será más fácil escribir “ $\underline{x}$ ”.)

Recordemos que  $\mathbb{R}^n$  tiene estructura de espacio vectorial, afín, euclídeo, métrico, topológico, etc. Voy a denotar todas estas estructuras por el mismo símbolo  $\mathbb{R}^n$ . En particular, para cada punto  $\mathbf{x} \in \mathbb{R}^n$  tiene sentido la multiplicación por  $\lambda \in \mathbb{R}$ :

$$\lambda \mathbf{x} := (\lambda x_1, \dots, \lambda x_n),$$

y para cada pareja de puntos  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$  se tiene su suma

$$\mathbf{x} + \mathbf{y} := (x_1 + y_1, \dots, x_n + y_n).$$

Tenemos el producto escalar (definido positivo, simétrico, bilineal)

$$\langle \mathbf{x}, \mathbf{y} \rangle := x_1 y_1 + \dots + x_n y_n.$$

Voy a usar la notación

$$\|\mathbf{x}\| := \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}.$$

Recordemos la **desigualdad del triángulo**

$$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$$

y también la **desigualdad del triángulo inversa**

$$\|\mathbf{x} - \mathbf{y}\| \geq \left| \|\mathbf{x}\| - \|\mathbf{y}\| \right|.$$

Esto define una estructura de espacio métrico con distancia  $d(\mathbf{x}, \mathbf{y}) := \|\mathbf{x} - \mathbf{y}\|$ . En fin, tenemos la topología estándar sobre  $\mathbb{R}^n$  que es inducida por esta distancia. Es decir, los subconjuntos abiertos son uniones arbitrarias o intersecciones finitas de las bolas abiertas

$$B(\mathbf{x}_0, \epsilon) := \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}_0 - \mathbf{x}\| < \epsilon\}$$

para  $\mathbf{x}_0 \in \mathbb{R}^n$  y  $\epsilon > 0$ .

Las letras  $X, Y$  van a denotar subconjuntos de  $\mathbb{R}^n$ . La letra  $K$  va a denotar un subconjunto convexo (de la palabra "konvex" en alemán) y las letras  $P$  y  $Q$  van a denotar politopos y poliedros convexos.

## 1 Conjuntos convexos

**1.1. Definición.** Un subconjunto  $A \subset \mathbb{R}^n$  se llama un **subespacio afín** si se cumple una de las siguientes condiciones equivalentes:

- 1) para cada dos puntos diferentes  $\mathbf{x}_1, \mathbf{x}_2 \in A$  la recta que pasa por  $\mathbf{x}_1$  y  $\mathbf{x}_2$ :

$$\{\mathbf{x}_1 + \lambda(\mathbf{x}_2 - \mathbf{x}_1) \mid \lambda \in \mathbb{R}\} = \{\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 \mid \lambda_1 + \lambda_2 = 1\}$$

está también contenida en  $A$ ,

- 2)  $A$  contiene todos los puntos de  $\mathbb{R}^n$  **afínmente dependientes** de  $A$ . Se dice que  $\mathbf{x} \in \mathbb{R}^n$  es afínmente dependiente de  $A$  si  $\mathbf{x}$  es una **combinación afín** de ciertos puntos  $\mathbf{x}_1, \dots, \mathbf{x}_k \in A$ :

$$\mathbf{x} = \sum_{1 \leq i \leq k} \lambda_i \mathbf{x}_i \quad \text{para ciertos coeficientes } \lambda_i \in \mathbb{R} \text{ tales que } \sum_i \lambda_i = 1.$$

De hecho, para  $k = 2$  la condición 2) es la misma cosa que 1). Entonces 2) implica 1). Ahora sea  $A$  un subconjunto que satisface la condición 1). La condición 2) se demuestra por inducción sobre  $k$ : el caso  $k = 1$  es trivial, y para  $k \geq 2$ , si tenemos un punto

$$\mathbf{x} = \sum_{1 \leq i \leq k} \lambda_i \mathbf{x}_i \quad \text{donde } \sum_i \lambda_i = 1 \quad \text{y } \mathbf{x}_1, \dots, \mathbf{x}_k \in A,$$

entonces sin pérdida de generalidad  $\lambda_k \neq 1$  y  $\lambda_1 + \dots + \lambda_{k-1} \neq 0$ , y

$$\mathbf{x} = (\lambda_1 + \dots + \lambda_{k-1}) \left( \sum_{1 \leq i \leq k-1} \frac{\lambda_i}{\lambda_1 + \dots + \lambda_{k-1}} \mathbf{x}_i \right) + \lambda_k \mathbf{x}_k,$$

lo cual significa que  $\mathbf{x}$  pertenece a la recta que pasa por los puntos  $\left( \sum_{1 \leq i \leq k-1} \frac{\lambda_i}{\lambda_1 + \dots + \lambda_{k-1}} \mathbf{x}_i \right)$  y  $\mathbf{x}_k$ . ■

**1.2. Definición.** Se dice que un subconjunto  $X \subset \mathbb{R}^n$  es **afínmente independiente** si  $X$  no contiene ningún punto  $\mathbf{x} \in X$  que sea afínmente dependiente de  $X \setminus \{\mathbf{x}\}$ .

Notamos que la condición 2) de arriba no depende del origen de coordenadas  $\mathbf{0} \in \mathbb{R}^n$ , y se ve fácilmente que  $X \subset \mathbb{R}^n$  es afínmente independiente si y solamente si para diferentes puntos  $\mathbf{x}_0, \dots, \mathbf{x}_k \in X$  la relación

$$\sum_{0 \leq i \leq k} \alpha_i \mathbf{x}_i = \mathbf{0} \quad \text{para} \quad \sum_i \alpha_i = 0$$

implica que  $\alpha_0 = \dots = \alpha_k = 0$ . Es decir, los vectores  $\mathbf{x}_1 - \mathbf{x}_0, \dots, \mathbf{x}_k - \mathbf{x}_0$  son linealmente independientes.

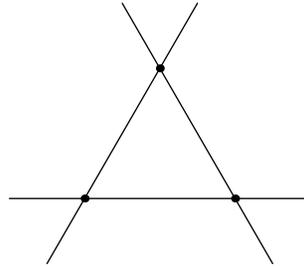
Mediante álgebra lineal básica se ve que cada subespacio afín  $A \subset \mathbb{R}^n$  contiene un subconjunto afínmente independiente máximo  $\{\mathbf{x}_0, \dots, \mathbf{x}_d\}$ , tal que  $A$  es el conjunto de todos los puntos de  $\mathbb{R}^n$  afínmente dependientes de  $\{\mathbf{x}_0, \dots, \mathbf{x}_d\}$ :

$$A = \left\{ \sum_{0 \leq i \leq d} \lambda_i \mathbf{x}_i \mid \sum_i \lambda_i = 1 \right\}.$$

**1.3. Definición.** El número  $d$  de arriba se llama la **dimensión** de  $A$ . Los números  $\lambda_i$  se llaman las **coordenadas baricéntricas** respecto a  $\{\mathbf{x}_0, \dots, \mathbf{x}_d\}$ .

En otras palabras, se puede definir sobre  $A$  una estructura de  $\mathbb{R}$ -espacio vectorial (escogiendo como origen de coordenadas a  $\mathbf{x}_0 \in A$ ), y la dimensión de este espacio vectorial es la dimensión de  $A$ . Los puntos  $\{\mathbf{x}_1, \dots, \mathbf{x}_d\}$  definen una base particular de este espacio vectorial.

**1.4. Ejemplo.** La dimensión del plano  $\mathbb{R}^2$  es 2 porque se puede encontrar 3 puntos  $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2$  afínmente independientes (los vectores  $\mathbf{x}_1 - \mathbf{x}_0$  y  $\mathbf{x}_2 - \mathbf{x}_0$  son linealmente independientes):

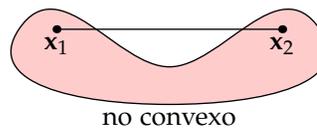
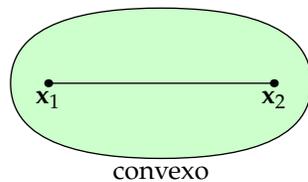


**1.5. Definición.** Un subconjunto  $K \subset \mathbb{R}^n$  se llama **convexo** si se cumple una de las siguientes condiciones equivalentes:

- 1) para cualesquiera dos puntos diferentes  $\mathbf{x}_1, \mathbf{x}_2 \in K$  el segmento de la recta entre  $\mathbf{x}_1$  y  $\mathbf{x}_2$

$$[\mathbf{x}_1, \mathbf{x}_2] := \{\mathbf{x}_1 + \lambda(\mathbf{x}_2 - \mathbf{x}_1) \mid 0 \leq \lambda \leq 1\} = \{\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 \mid \lambda_1, \lambda_2 \geq 0, \lambda_1 + \lambda_2 = 1\}$$

está también contenido en  $K$ ,



2)  $K$  contiene todos los puntos de  $\mathbb{R}^n$  **convexamente dependientes** de  $K$ . Se dice que  $\mathbf{x} \in \mathbb{R}^n$  es convexamente dependiente de  $K$  si  $\mathbf{x}$  es una **combinación convexa** de algunos puntos  $\mathbf{x}_1, \dots, \mathbf{x}_k \in K$ :

$$\mathbf{x} = \sum_{1 \leq i \leq k} \lambda_i \mathbf{x}_i \quad \text{para algunos coeficientes } \lambda_i \geq 0 \text{ tales que } \sum_i \lambda_i = 1.$$

Noten que la única diferencia entre la definición de un subconjunto convexo y un subespacio afín es que para un conjunto convexo se pide que  $\lambda_i$  sean no negativos, y entonces la equivalencia de las condiciones 1) y 2) se demuestra de la misma manera. También notamos que cada subespacio afín de  $\mathbb{R}^n$  es un ejemplo (poco interesante) de un conjunto convexo. El conjunto vacío  $\emptyset$  es también convexo.

**1.6. Observación.** Sean  $K \subset \mathbb{R}^n$  y  $L \subset \mathbb{R}^m$  dos conjuntos convexos y sea  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  una aplicación afín. Entonces  $f(K) \subset \mathbb{R}^m$  y  $f^{-1}(L) \subset \mathbb{R}^n$  son también convexos.

*Demostración.* Recordemos que una **aplicación afín**  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  se caracteriza como una aplicación tal que para cada familia de puntos  $\mathbf{x}_i$ , para  $\sum_i \lambda_i \mathbf{x}_i$  donde  $\sum_i \lambda_i = 1$  se tiene  $f(\sum_i \lambda_i \mathbf{x}_i) = \sum_i \lambda_i f(\mathbf{x}_i)$ .

Si  $f(\mathbf{x}_i) \in f(K)$ , entonces para cualquier combinación convexa tenemos

$$\sum_i \lambda_i f(\mathbf{x}_i) = f\left(\underbrace{\sum_i \lambda_i \mathbf{x}_i}_{\in K}\right) \in f(K).$$

De modo similar, si  $\mathbf{x}_i \in f^{-1}(L)$ , es decir  $f(\mathbf{x}_i) \in L$ , entonces

$$f\left(\sum_i \lambda_i \mathbf{x}_i\right) = \sum_i \lambda_i f(\mathbf{x}_i) \in L,$$

es decir  $\sum_i \lambda_i \mathbf{x}_i \in f^{-1}(L)$ . ■

**1.7. Ejercicio.** ¿Cuáles son los subconjuntos convexos de  $\mathbb{R}^1$ ? Digamos que dos subconjuntos  $X$  y  $Y$  son **diferentes** si uno no se transforma en otro por una homotecia de razón positiva (una aplicación  $x \mapsto \lambda x$  con  $\lambda > 0$ ). ¿Cuáles son los subconjuntos convexos diferentes de  $\mathbb{R}^1$ ?

Aquí hay un par de ejemplos de conjuntos convexos que no son subespacios afines:

**1.8. Ejemplo.** Una bola en  $\mathbb{R}^n$  (cerrada o abierta) es un conjunto convexo. Sin pérdida de generalidad, la bola está centrada en  $\mathbf{0} \in \mathbb{R}^n$ . Sean  $\mathbf{x}_1$  y  $\mathbf{x}_2$  dos puntos en la bola de radio  $r$ , es decir  $\|\mathbf{x}_1\| \leq r$  y  $\|\mathbf{x}_2\| \leq r$ . Sea  $\mathbf{x} = \mathbf{x}_1 + \lambda(\mathbf{x}_2 - \mathbf{x}_1) = (1 - \lambda)\mathbf{x}_1 + \lambda\mathbf{x}_2$  para  $0 \leq \lambda \leq 1$ . Entonces por la desigualdad del triángulo

$$\|\mathbf{x}\| \leq (1 - \lambda)\|\mathbf{x}_1\| + \lambda\|\mathbf{x}_2\| \leq r.$$

Para la bola abierta sustituyan " $\leq$ " por " $<$ ". ▲

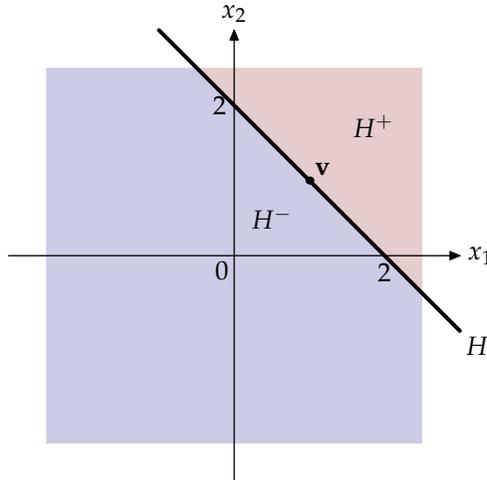
**1.9. Ejemplo.** Sea  $\mathbf{v} \neq \mathbf{0}$  un punto en  $\mathbb{R}^n$  y  $\alpha \in \mathbb{R}$  algún número fijo. Entonces el subconjunto

$$H := \{\mathbf{x} \in \mathbb{R}^n \mid \langle \mathbf{x}, \mathbf{v} \rangle = \alpha\}$$

es un hiperplano que divide  $\mathbb{R}^n$  en dos **semiespacios abiertos**

$$H^+ := \{\mathbf{x} \in \mathbb{R}^n \mid \langle \mathbf{x}, \mathbf{v} \rangle > \alpha\},$$

$$H^- := \{\mathbf{x} \in \mathbb{R}^n \mid \langle \mathbf{x}, \mathbf{v} \rangle < \alpha\}.$$



Ejemplo:  $\mathbf{v} = (1, 1)$ ,  $\alpha = 2$ ,  $H = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 + x_2 = 2\}$ .

$H^+$  y  $H^-$  son convexos. Por ejemplo, si  $\mathbf{x}_1, \mathbf{x}_2 \in H^+$ , entonces para cualquier  $\lambda_1, \lambda_2 \geq 0$  tales que  $\lambda_1 + \lambda_2 = 1$  tenemos

$$\langle \lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2, \mathbf{v} \rangle = \lambda_1 \langle \mathbf{x}_1, \mathbf{v} \rangle + \lambda_2 \langle \mathbf{x}_2, \mathbf{v} \rangle > \alpha.$$

Si en lugar de las condiciones  $>$  y  $<$  se consideran  $\geq$  y  $\leq$ , entonces se tiene **semiespacios cerrados** correspondientes  $\overline{H}^+$  y  $\overline{H}^-$ . ▲

**1.10. Ejemplo.** Sean  $\mathbf{v}_0, \dots, \mathbf{v}_n \in \mathbb{R}^{n+1}$  algunos puntos diferentes tales que  $\{\mathbf{v}_0, \dots, \mathbf{v}_n\}$  es afínmente independiente. Entonces el **símplice con vértices**  $\mathbf{v}_0, \dots, \mathbf{v}_n$  es el subconjunto de  $\mathbb{R}^n$  definido por

$$S := \left\{ \sum_{0 \leq i \leq n} \alpha_i \mathbf{v}_i \mid \sum_i \alpha_i = 1, \alpha_i \geq 0 \right\}.$$

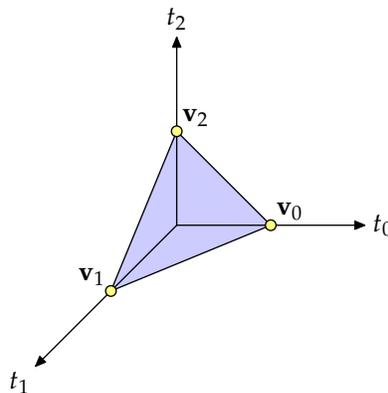
Se ve fácilmente que  $S$  es convexo. De hecho, si  $\mathbf{x}_1 = \sum_{0 \leq i \leq n} \alpha_i^{(1)} \mathbf{v}_i$  y  $\mathbf{x}_2 = \sum_{0 \leq i \leq n} \alpha_i^{(2)} \mathbf{v}_i$  son puntos de  $S$  (es decir,  $\sum_i \alpha_i^{(1)} = \sum_i \alpha_i^{(2)} = 1$ ) y  $\lambda_1, \lambda_2 \geq 0$  son algunos números positivos tales que  $\lambda_1 + \lambda_2 = 1$ , entonces

$$\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 = \sum_{0 \leq i \leq n} (\lambda_1 \alpha_i^{(1)} + \lambda_2 \alpha_i^{(2)}) \mathbf{v}_i,$$

que es también un punto de  $S$  porque  $\sum_i \lambda_1 \alpha_i^{(1)} + \lambda_2 \alpha_i^{(2)} = \lambda_1 \sum_i \alpha_i^{(1)} + \lambda_2 \sum_i \alpha_i^{(2)} = \lambda_1 + \lambda_2 = 1$ .

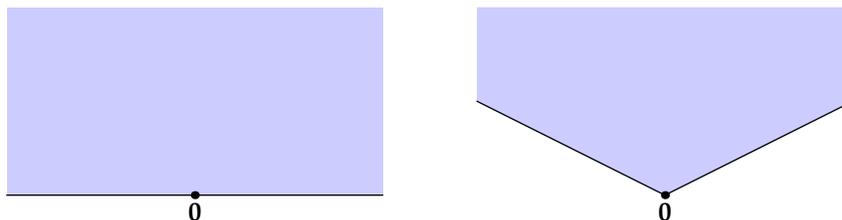
Normalmente en  $\mathbb{R}^{n+1}$  se consideran los puntos  $\mathbf{v}_0 = (1, 0, \dots, 0)$ ,  $\mathbf{v}_1 = (0, 1, \dots, 0)$ ,  $\dots$ ,  $\mathbf{v}_n = (0, 0, \dots, 1)$ , y el conjunto correspondiente se llama el  **$n$ -símplice estándar**:

$$\Delta^n := \left\{ (\alpha_0, \dots, \alpha_n) \in \mathbb{R}^{n+1} \mid \sum_i \alpha_i = 1, \alpha_i \geq 0 \right\}.$$





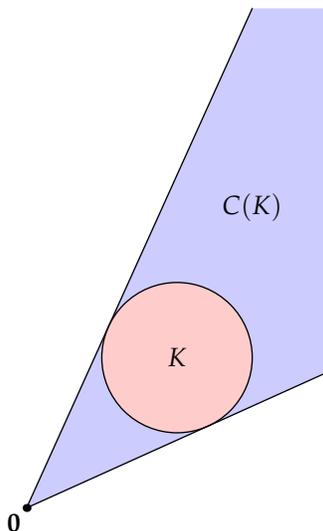
**1.11. Ejemplo.** Un subconjunto  $C \subset \mathbb{R}^n$  se llama un **cono** de vértice  $\mathbf{0}$  si  $\mathbf{0} \in C$  y para cada punto  $\mathbf{x} \in C$  y cada  $\lambda > 0$  también  $\lambda \mathbf{x} \in C$ . Notemos que  $C$  es convexo si y solamente si para cada  $\mathbf{x}, \mathbf{y} \in C$  tenemos  $\mathbf{x} + \mathbf{y} \in C$ . En otras palabras, un cono convexo  $C$  es un subconjunto de  $\mathbb{R}^n$  tal que  $\mathbf{0} \in C$  y  $C$  es cerrado bajo combinaciones lineales positivas. En general, un cono de vértice  $\mathbf{a}$  se obtiene por la traslación paralela de un cono de vértice  $\mathbf{0}$ . Un hiperplano es un caso particular de un cono.



Ahora si  $X \subset \mathbb{R}^n$  es un subconjunto, entonces el conjunto

$$C(X) := \{\lambda \mathbf{x} \mid \mathbf{x} \in X, \lambda \geq 0\}$$

es el mínimo cono que contiene a  $X$ . Si  $K \subset \mathbb{R}^n$  es convexo, entonces  $C(K)$  es un cono convexo.



Las siguientes propiedades son fáciles de demostrar y las dejo como un ejercicio:

- 1.12. Ejercicio.**
- 1) Si  $K \subset \mathbb{R}^n$  y  $L \subset \mathbb{R}^m$  son conjuntos convexos, entonces  $K \times L \subset \mathbb{R}^{n+m}$  es también convexo.
  - 2) Sea  $\{K_\alpha\}$  una familia de conjuntos convexos en  $\mathbb{R}^n$  (indexada por algún parámetro real  $\alpha$ ). Entonces su intersección  $\bigcap_\alpha K_\alpha$  es también convexa.
  - 3) La unión de conjuntos convexos  $\bigcup_\alpha K_\alpha$  casi nunca es convexa. Sin embargo, si la familia  $\{K_\alpha\}_\alpha$  satisface  $\alpha \leq \beta \Leftrightarrow K_\alpha \subseteq K_\beta$ , entonces su unión es también convexa.
  - 4) Para dos subconjuntos  $X, Y \subseteq \mathbb{R}^n$  su **suma de Minkowski** es el subconjunto

$$X + Y := \{\mathbf{x} + \mathbf{y} \mid \mathbf{x} \in X, \mathbf{y} \in Y\}.$$

Demuestre que si  $X$  y  $Y$  son convexos, entonces  $X + Y$  es también convexo.

**HERMANN MINKOWSKI** (1864–1909) fue un matemático alemán de origen judío, conocido por sus contribuciones en geometría, teoría de números y física matemática.



Nació en el Imperio ruso, en el territorio actual de Lituania. Su familia se mudó a Königsberg en el 1872. Minkowski recibió su doctorado en Königsberg bajo la dirección de Ferdinand von Lindemann. Estudió formas cuadráticas y estableció la “geometría de los números”. Minkowski trabajó en universidades de Bonn, Gotinga, Königsberg y Zürich. En Zürich fue uno de los profesores de Albert Einstein. Murió repentinamente a los 44 años de ruptura del apéndice.

Varios conceptos y resultados en matemáticas tienen su nombre, entre ellos la suma de Minkowski que hemos definido arriba y los teoremas sobre conjuntos convexos y retículos:

- Sea  $\Lambda \subset \mathbb{R}^n$  un retículo (por ejemplo,  $\Lambda = \mathbb{Z}^n$ ). Sea  $K \subset \mathbb{R}^n$  un conjunto convexo, simétrico respecto a  $\mathbf{0}$  (es decir, si  $\mathbf{x} \in K$ , entonces  $-\mathbf{x} \in K$ ). Si  $\text{vol } K > 2^n \cdot \det \Lambda$ , entonces el **teorema de Minkowski** dice que  $K$  contiene un punto  $\mathbf{x} \in \Lambda$ ,  $\mathbf{x} \neq \mathbf{0}$ .
- Sea  $\Lambda \subset \mathbb{R}^n$  un retículo y sea  $K \subset \mathbb{R}^n$  un conjunto convexo, simétrico respecto a  $\mathbf{0}$ , tal que  $\text{vol } K < \infty$ . Entonces el **segundo teorema de Minkowski** da una estimación de  $\text{vol } K$ :

$$\frac{2^n}{n!} \det \Lambda \leq \lambda_1 \cdots \lambda_n \text{vol } K \leq 2^n \det \Lambda,$$

donde

$$\lambda_i := \inf\{\lambda > 0 \mid \lambda K \text{ contiene } i \text{ vectores linealmente independientes de } \Lambda\}.$$

- **El teorema de Hasse–Minkowski:** si  $F(X_1, \dots, X_n) \in \mathbb{Q}[X_1, \dots, X_n]$  es una forma cuadrática (un polinomio homogéneo de grado 2 en  $n$  variables), entonces la ecuación  $F(\mathbf{X}) = 0$  tiene una solución no trivial  $\mathbf{x} \in \mathbb{Q}^n$ ,  $\mathbf{x} \neq \mathbf{0}$  si y solamente si  $\bar{F}(\mathbf{X}) = 0$  tiene solución no trivial en cada completación  $\mathbb{Q}_p$  (incluso  $\mathbb{Q}_\infty = \mathbb{R}$ ). La demostración del teorema de Hasse–Minkowski usa el teorema de Minkowski de arriba.

**1.13. Observación.** Si  $K \subset \mathbb{R}^n$  es un conjunto convexo, entonces su *interior* y *clausura topológica*

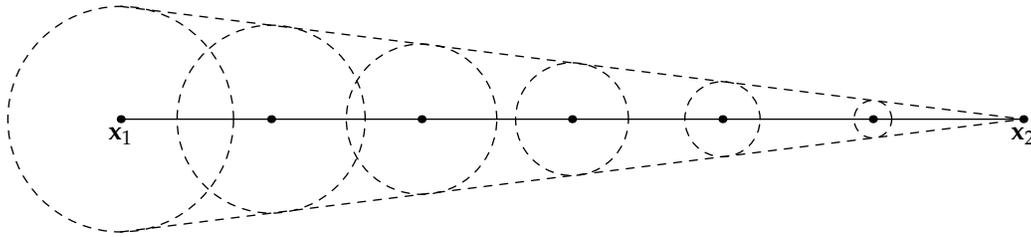
$$\text{int } K := \bigcup_{\substack{F \subset K \\ F \text{ abierto en } \mathbb{R}^n}} F = \{ \mathbf{x} \in K \mid B(\mathbf{x}, \epsilon) \subset K \text{ para algún } \epsilon > 0 \},$$

$$\bar{K} := \bigcap_{\substack{F \supset K \\ F \text{ cerrado en } \mathbb{R}^n}} F = \{ \mathbf{x} \in \mathbb{R}^n \mid B(\mathbf{x}, \epsilon) \cap K \neq \emptyset \text{ para cualquier } \epsilon > 0 \}.$$

son también convexos.

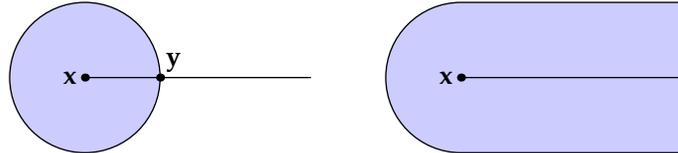
*Demostración.* Vamos a demostrar la parte sobre  $\text{int } K$ . Sea  $\mathbf{x}_1 \in \text{int } K$ . Entonces  $B(\mathbf{x}_1, \epsilon) \subset K$  para algún  $\epsilon > 0$ . Luego, para cualquier punto  $\mathbf{x}_2 \in K$  por la convexidad de  $K$  tenemos

$$\lambda_1 B(\mathbf{x}_1, \epsilon) + \lambda_2 \mathbf{x}_2 = B(\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2, \lambda_1 \epsilon) \subset K \quad \text{para } \lambda_1 > 0, \lambda_2 \geq 0, \lambda_1 + \lambda_2 = 1.$$



En otras palabras, para cualquier  $\mathbf{x}_1 \in \text{int } K$  y  $\mathbf{x}_2 \in K$  tenemos  $[\mathbf{x}_1, \mathbf{x}_2) \subset \text{int } K$ . En particular, si  $\mathbf{x}_1, \mathbf{x}_2 \in \text{int } K$ , entonces  $[\mathbf{x}_1, \mathbf{x}_2] \subset \text{int } K$ . ■

El argumento de arriba demuestra que los conjuntos convexos tienen la siguiente propiedad especial: si  $R$  es un rayo con origen en  $\mathbf{x} \in \text{int } K$ , entonces  $R$  interseca la frontera  $\text{fr } K$  como mucho en un punto. Si  $K \cap \text{fr } K = \{ \mathbf{y} \}$ , entonces  $[\mathbf{x}, \mathbf{y}] \subset \text{int } K$  y  $R \setminus [\mathbf{x}, \mathbf{y}] \subset \mathbb{R}^n \setminus \bar{K}$ . Si  $K$  no es acotado, puede ser que  $R \cap \text{fr } K = \emptyset$ , y en este caso  $R \subset \text{int } K$ .



**1.14. Ejercicio.** Termine la demostración del caso de  $\bar{K}$ .

Noten que en general el interior  $\text{int } K$  de un conjunto convexo  $K \subset \mathbb{R}^n$  puede ser vacío. Por ejemplo, si

$$S = \left\{ \sum_{0 \leq i \leq n} \alpha_i \mathbf{v}_i \mid \sum_{0 \leq i \leq n} \alpha_i = 1, \alpha_i \geq 0 \right\}$$

es un símplice en  $\mathbb{R}^{n+1}$ , entonces  $S$  no contiene ningún subconjunto que sea abierto en  $\mathbb{R}^{n+1}$ , simplemente porque  $S$  pertenece al hiperplano

$$H = \left\{ \sum_{0 \leq i \leq n} \alpha_i \mathbf{v}_i \mid \sum_{0 \leq i \leq n} \alpha_i = 1 \right\}.$$

Pero  $H$  es un espacio afín, de hecho el espacio afín mínimo que contiene a  $S$ , y el interior de  $S$  como subespacio topológico de  $H$  no es vacío, es dado por

$$\text{intrel } S = \left\{ \sum_{0 \leq i \leq n} \alpha_i \mathbf{v}_i \mid \sum_i \alpha_i = 1, \alpha_i > 0 \right\}.$$

Esto recibe el nombre de **interior relativo** (a  $H$ ). La **frontera relativa** de  $S$  está dada por

$$S \setminus \text{intrel } S = \bigcup_{0 \leq i \leq n} S_i,$$

donde  $S_i$  es el simplejo con vértices  $\{\mathbf{v}_0, \dots, \mathbf{v}_n\} \setminus \{\mathbf{v}_i\}$ , que se obtiene poniendo  $\alpha_i = 0$ .

**1.15. Definición.** Si  $K \subset \mathbb{R}^n$  es un conjunto convexo, entonces la **envolvente afín** de  $K$  es el conjunto

$$\text{af } K := \bigcap_{\substack{A \supseteq K \\ A \text{ es un subespacio afín de } \mathbb{R}^n}} A.$$

La **dimensión** de  $K$  es la dimensión del espacio afín  $\text{af } K$ .

Se ve que

$$\text{af } K = \left\{ \sum_{0 \leq i \leq m} \lambda_i \mathbf{x}_i \mid \sum_{0 \leq i \leq m} \lambda_i = 1 \right\},$$

donde  $\{\mathbf{x}_0, \dots, \mathbf{x}_m\} \subset K$  es un conjunto máximo de puntos afinmente independientes y  $m = \dim K$ . La envolvente afín es interesante por la siguiente razón:

**1.16. Observación.** Sea  $K$  un conjunto convexo no vacío. Entonces el interior de  $K$  como subconjunto de  $\text{af } K$  no es vacío. Este se llama el **interior relativo** y se denota por  $\text{intrel } K$ .

En particular, si  $\text{int } K = \emptyset$ , entonces  $\text{af } K \subsetneq \mathbb{R}^n$  y  $K$  pertenece a algún hiperplano en  $\mathbb{R}^n$ .

*Demostración.* Si  $K$  no es vacío, entonces

$$\text{af } K = \left\{ \sum_{0 \leq i \leq m} \lambda_i \mathbf{x}_i \mid \sum_{0 \leq i \leq m} \lambda_i = 1 \right\}$$

contiene el simplejo

$$S = \left\{ \sum_{0 \leq i \leq m} \lambda_i \mathbf{x}_i \mid \sum_{0 \leq i \leq m} \lambda_i = 1, \lambda_i \geq 0 \right\},$$

con vértices en ciertos puntos  $\mathbf{x}_0, \dots, \mathbf{x}_m \in K$ , y luego su **baricentro**  $\frac{1}{m+1} \sum_{0 \leq i \leq m} \mathbf{x}_i$  pertenece a  $\text{intrel } S \subseteq \text{intrel } K$ . ■

Los conjuntos convexos tienen buenas propiedades topológicas. Por ejemplo, para cualquier espacio topológico  $X$  tenemos

$$\begin{aligned} \overline{\text{int } X} &\subseteq \overline{X}, \\ \text{int } X &\subseteq \text{int } \overline{X}. \end{aligned}$$

Pero en general  $\overline{\text{int } X} \neq \overline{X}$  (por ejemplo, si  $X$  es un punto) y  $\text{int } X \neq \text{int } \overline{X}$  (por ejemplo, si  $X$  es una bola en  $\mathbb{R}^n$  sin un punto interior). Sin embargo, si  $X$  es convexo y  $\text{int } X \neq \emptyset$ , entonces hay inclusiones inversas.

**1.17. Proposición.** Sea  $K \subseteq \mathbb{R}^n$  un conjunto convexo. Entonces

$$\begin{aligned} \overline{\text{int } K} &= \overline{K}, \quad \text{si } \text{int } K \neq \emptyset, \\ \text{int } K &= \text{int } \overline{K}, \\ \overline{\text{intrel } K} &= \overline{K}, \quad \text{si } \text{intrel } K \neq \emptyset, \\ \text{intrel } K &= \text{intrel } \overline{K}, \\ \text{fr } K &= \text{fr}(\text{int } K) = \text{fr}(\overline{K}), \quad \text{si } \text{int } K \neq \emptyset. \end{aligned}$$

Aquí  $\text{fr } K := \overline{K} \setminus \text{int } K$  es la **frontera** de  $K$ .

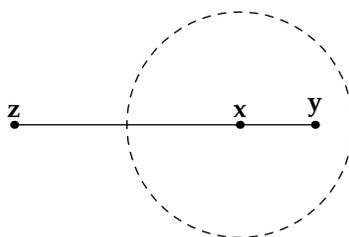
*Demostración.* La última fórmula es una consecuencia formal de las dos primeras. Las fórmulas con  $\text{intrel}$  son también consecuencias de las fórmulas con  $\text{int}$ .

Para la primera fórmula, si  $x \in \bar{K}$ , entonces para cualquier  $\epsilon > 0$  existe  $y \in K$  tal que  $\|x - y\| < \epsilon$ . Por nuestra hipótesis, existe un punto  $z \in \text{int} K$ , y por lo tanto  $[z, y) \subset \text{int} K$ , como hemos visto en 1.13. Entonces para cualquier  $\epsilon > 0$  existe  $y' \in \text{int} K$  tal que  $\|x - y'\| < \epsilon$ . Hemos demostrado que  $\bar{K} \subseteq \overline{\text{int} K}$ .

Para la segunda fórmula, tenemos que demostrar que  $\text{int} \bar{K} \subseteq \text{int} K$ . Primero supongamos que  $\text{int} K = \emptyset$ . Tenemos dos posibilidades:

- $K = \emptyset$ , y en este caso obviamente  $\text{int} K = \text{int} \bar{K}$ ,
- $K \neq \emptyset$ , y en este caso  $\text{intrel} K \neq \emptyset$  y por lo tanto  $\text{af} K \subsetneq \mathbb{R}^n$  y  $\text{int} K = \text{int} \bar{K} = \emptyset$ .

Entonces, el caso  $\text{int} K = \emptyset$  es trivial y podemos suponer que  $\text{int} K \neq \emptyset$ . Escojamos un punto  $z \in \text{int} K$ . Para un punto  $x \in \text{int} \bar{K}$  sea  $R$  el rayo que empieza en  $z$  y pasa por  $x$ . La condición  $x \in \text{int} \bar{K}$  quiere decir que  $B(x, \epsilon) \subset \bar{K}$  para algún  $\epsilon > 0$ , y por lo tanto existe  $y \in B(x, \epsilon)$  tal que  $x \in (z, y)$ . Pero  $[z, y) \subset \text{int} K$ , como hemos observado antes (el rayo puede intersectar la frontera solo una vez, y esta intersección no puede estar antes de  $y$  porque  $y \in \bar{K}$ ).



■

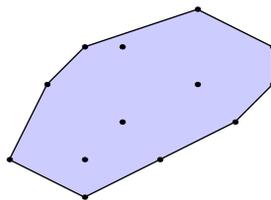
## 2 La envolvente convexa

**2.1. Definición.** Si  $X \subset \mathbb{R}^n$  es cualquier conjunto, entonces la **envolvente convexa** de  $X$  es el conjunto

$$\text{conv} X := \bigcap_{\substack{K \supseteq X \\ K \text{ convexo}}} K.$$

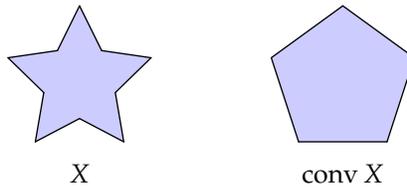
La intersección de conjuntos convexos es convexa, de donde  $\text{conv} X$  es un conjunto convexo, específicamente el conjunto convexo mínimo que contiene a  $X$ . Tenemos  $\text{conv} X = X$  si y solamente si  $X$  es convexo.

**2.2. Ejemplo.** He aquí la envolvente convexa de un conjunto finito de puntos en  $\mathbb{R}^2$ :



▲

**2.3. Ejemplo.** He aquí un subconjunto no convexo de  $\mathbb{R}^2$  y su envolvente convexa:



**2.4. Ejercicio.** 1) Demuestre que si  $X$  es un conjunto abierto, entonces  $\text{conv } X$  es también abierto. (Indicación:  $\text{conv } X$  es convexo, de donde  $\text{int } \text{conv } X$  es también convexo.)

2) Encuentre algún ejemplo de un conjunto cerrado  $X$  tal que  $\text{conv } X$  no es cerrado.

**2.5. Ejercicio.** Si  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  es una aplicación afín, entonces para cualquier subconjunto  $X \subset \mathbb{R}^n$  tenemos

$$f(\text{conv } X) = \text{conv } f(X)$$

(utilice el hecho que la imagen y la imagen inversa de un conjunto convexo, respecto a una aplicación afín, es también convexa).

**2.6. Ejercicio.** 1) Tenemos la siguiente descripción de la envolvente convexa de  $X \subset \mathbb{R}^n$ :

$$\text{conv } X = \left\{ \sum_{1 \leq i \leq k} \lambda_i \mathbf{x}_i \mid \sum_i \lambda_i = 1, \lambda_i \geq 0, \mathbf{x}_1, \dots, \mathbf{x}_k \in X, k = 1, 2, \dots \right\}.$$

(denotemos el conjunto a la derecha por  $K$ ; note que  $K \subseteq \text{conv } X$  y  $X \subset K$ , por lo que es suficiente de demostrar que  $K$  es convexo).

2) En particular, para un conjunto finito  $X = \{\mathbf{v}_0, \dots, \mathbf{v}_m\}$  y afínmente independiente, su envolvente convexa es el  $m$ -símplice con vértices  $\mathbf{v}_0, \dots, \mathbf{v}_m$ .

3) Si  $K_1$  y  $K_2$  son dos subconjuntos convexos no vacíos en  $\mathbb{R}^n$ , entonces

$$\text{conv}(K_1 \cup K_2) = \bigcup_{\substack{\mathbf{x}_1 \in K_1 \\ \mathbf{x}_2 \in K_2}} [\mathbf{x}_1, \mathbf{x}_2].$$

4) Si  $C_1$  y  $C_2$  son dos conos convexos, entonces su suma de Minkowski  $C_1 + C_2$  coincide con  $\text{conv}(C_1 \cup C_2)$ .

**2.7. Proposición (Teorema de Carathéodory).** En la fórmula de arriba para  $\text{conv } X$  es suficiente considerar  $k = 1, 2, \dots, n + 1$  y suponer que los puntos  $\mathbf{x}_i$  son afínmente independientes. En otras palabras, para  $X \subset \mathbb{R}^n$  la envolvente convexa  $\text{conv } X$  es la unión de todos  $m$ -símplices con vértices en  $X$  para  $m \leq n$ .

*Demostración.* Está claro que cada símplice con vértices en  $X$  pertenece a  $\text{conv } X$ . Lo que tenemos que demostrar es que cada punto  $\mathbf{x} \in \text{conv } X$  pertenece a un  $m$ -símplice para  $m \leq n$ . En el ejercicio de arriba hemos visto que

$$\mathbf{x} = \sum_{1 \leq i \leq k} \lambda_i \mathbf{x}_i$$

para algunos  $\mathbf{x}_i \in X$  y  $\lambda_i \geq 0, \sum_i \lambda_i = 1$ . Sin pérdida de generalidad,  $\lambda_i > 0$ . Si los puntos  $\mathbf{x}_i$  son afínmente independientes, entonces  $k \leq n + 1$ . Supongamos que  $\mathbf{x}_i$  son afínmente dependientes, es decir que existe alguna relación

$$\sum_{1 \leq i \leq k} \alpha_i \mathbf{x}_i = 0,$$

donde  $\sum_i \alpha_i = 1$ . Sin pérdida de generalidad,  $\alpha_k > 0$ . También, cambiando la numeración de los términos con  $\alpha_i \neq 0$ , podemos suponer que

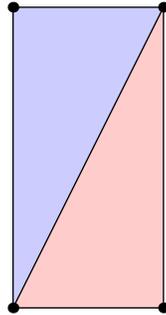
$$\lambda_i / \alpha_i \geq \lambda_k / \alpha_k.$$

Entonces

$$\mathbf{x} = \sum_{1 \leq i \leq k} \lambda_i \mathbf{x}_i = \sum_{1 \leq i \leq k} \lambda_i \mathbf{x}_i - \frac{\lambda_k}{\alpha_k} \underbrace{\sum_{1 \leq i \leq k} \alpha_i \mathbf{x}_i}_{=0} = \sum_{1 \leq i \leq k-1} \underbrace{\left( \lambda_i - \frac{\lambda_k}{\alpha_k} \alpha_i \right)}_{\geq 0} \mathbf{x}_i.$$

Entonces, usando la relación  $\sum_{1 \leq i \leq k} \alpha_i \mathbf{x}_i = 0$ , hemos escrito  $\mathbf{x}$  como una combinación convexa de  $k - 1$  puntos  $\mathbf{x}_1, \dots, \mathbf{x}_{k-1}$ . Si estos puntos son afínmente independientes, la demostración está terminada. Sino, podemos repetir el proceso y reducir el número de términos. ■

**2.8. Ejemplo.** El rectángulo es un conjunto convexo que es la unión de dos triángulos.



Cada punto de  $\mathbb{R}^2$  pertenece a un triángulo, cada punto de  $\mathbb{R}^n$  pertenece a un  $n$ -símplice; en general, cada punto de un conjunto convexo  $K$  pertenece a un símplice con vértices en  $K$ . El teorema de Carathéodory dice que para formar  $\text{conv } X$ , hay que tomar todos los símplices con vértices en  $X$ . ▲

**2.9. Ejemplo.** Sea  $C \subset \mathbb{R}^n$  un cono de vértice  $\mathbf{0}$  (no necesariamente convexo), entonces cada punto de  $\text{conv } C$  puede ser escrito como  $\mathbf{x}_1 + \dots + \mathbf{x}_n$  para algunos  $\mathbf{x}_i \in C$ . El hecho que  $n$  términos sean suficientes es álgebra lineal. ▲

**CONSTANTIN CARATHÉODORY** (1873–1950) fue un matemático de origen griego que trabajó en el análisis real, cálculo de variaciones y teoría de la medida.



Nació en Berlín, siendo su padre un embajador del Imperio otomano. Estudió matemáticas en la universidad de Berlín y la universidad de Gotinga bajo la dirección de Minkowski. Luego trabajó en varias universidades alemanas y en 1920 fue invitado a fundar una universidad en Esmirna tras la invasión griega de esta ciudad durante la partición del Imperio otomano en la guerra greco-turca de 1919–1922. Sin embargo, en 1922 la ciudad fue destrozada por un incendio y los turcos recuperaron el control. Carathéodory fue rescatado en el último momento por un barco de guerra. La universidad de Esmirna nunca admitió estudiantes; Carathéodory salvó la biblioteca universitaria y trabajó hasta 1924 en Atenas. Desde 1924 hasta su jubilación en 1938 fue profesor de matemáticas en la Universidad de Múnich. Aunque perteneció a la Academia de Ciencias de Baviera durante la época de Hitler, no tomaba directamente parte de las actividades del partido nazi.

**2.10. Proposición.** Si  $X \subset \mathbb{R}^n$  es un conjunto compacto, entonces  $\text{conv } X$  es también compacto.

Un subconjunto de  $\mathbb{R}^n$  es compacto si y solamente si es cerrado y acotado. Ya hemos mencionado que para  $X$  cerrado,  $\text{conv } X$  no es necesariamente cerrado. Esta proposición dice que si  $X$  es cerrado y acotado, entonces  $\text{conv } X$  es también cerrado y acotado.

*Demostración.* Consideramos el símplice

$$\Delta := \{(\lambda_1, \dots, \lambda_{n+1}) \in \mathbb{R}^{n+1} \mid \sum_i \lambda_i = 1, \lambda_i \geq 0\}.$$

Es un subconjunto compacto de  $\mathbb{R}^{n+1}$ . Ahora el teorema de Carathéodory nos dice que  $\text{conv } X$  es exactamente la imagen de la aplicación continua

$$\underbrace{X \times \dots \times X}_{n+1} \times \Delta \rightarrow \mathbb{R}^n,$$

$$(\mathbf{x}_1, \dots, \mathbf{x}_{n+1}, \lambda_1, \dots, \lambda_{n+1}) \mapsto \sum_{1 \leq i \leq n+1} \lambda_i \mathbf{x}_i.$$

Si  $X$  es compacto, entonces  $X \times \dots \times X \times \Delta$  es también compacto (como producto de conjuntos compactos). La imagen de un conjunto compacto por una aplicación continua es también compacta. ■

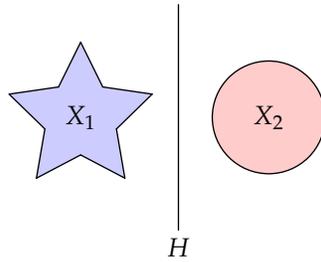
**2.11. Corolario.** Si  $X$  es un conjunto acotado, entonces  $\text{conv } X$  es también acotado.

*Demostración.* Recuerden que  $X$  es acotado si y solamente si  $X \subseteq Y$  donde  $Y$  es compacto. En este caso  $\text{conv } X \subseteq \text{conv } Y$ , donde  $\text{conv } Y$  es compacto por la proposición precedente. ■

### 3 Separación por hiperplanos

**3.1. Definición.** Se dice que dos conjuntos no vacíos  $X_1, X_2 \subseteq \mathbb{R}^n$  están **separados por el hiperplano**  $H$  si  $X_1 \subseteq \overline{H^+}$  y  $X_2 \subseteq \overline{H^-}$  para los dos semiespacios cerrados definidos por  $H$ .

Si  $X_1 \subseteq H^+$  y  $X_2 \subseteq H^-$  para los semiespacios abiertos correspondientes, entonces se dice que  $X_1$  y  $X_2$  están **estrictamente separados**.

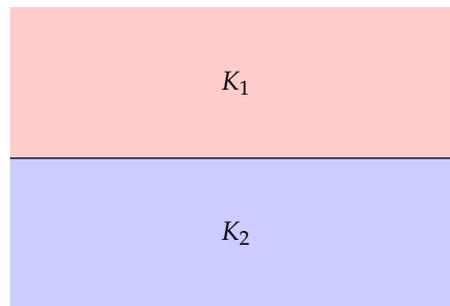


Es claro que hay parejas de conjuntos  $X_1, X_2 \subseteq \mathbb{R}^n$  que no pueden ser separadas por un hiperplano. Vamos a analizar el caso de la separación estricta. He aquí un criterio suficiente:

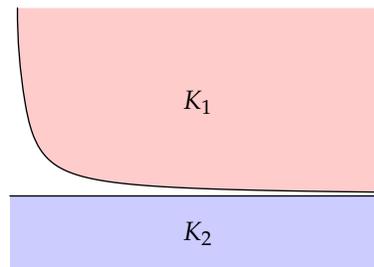
**3.2. Teorema.** Sean  $K_1, K_2 \subseteq \mathbb{R}^n$  dos conjuntos no vacíos, disjuntos, convexos, cerrados. También supongamos que  $K_2$  es acotado. Entonces existe un hiperplano que separa  $K_1$  y  $K_2$  estrictamente.

**3.3. Ejemplo.** La condición  $K_1 \cap K_2 = \emptyset$  es obviamente necesaria. También sin la hipótesis que  $K_1$  y  $K_2$  son convexos, se puede construir muchos contraejemplos.

Es necesario que ambos conjuntos sean cerrados, porque en el caso contrario tenemos el contraejemplo estúpido donde  $K_1$  y  $K_2$  son dos semiespacios que corresponden al mismo hiperplano,  $K_1$  abierto,  $K_2$  cerrado y  $K_1 \cup K_2 = \mathbb{R}^n$ .



En fin, la hipótesis que uno de los conjuntos es compacto es necesaria porque, por ejemplo, los conjuntos  $K_1 := \{x > 0, y \geq 1/x\}$  y  $K_2 := \{y \leq 0\}$  en  $\mathbb{R}^2$  no admiten separación estricta: son disjuntos, pero la distancia entre  $K_1$  y  $K_2$  es arbitrariamente pequeña.



El punto clave de la demostración del teorema 3.2 es el siguiente

**3.4. Lema.** Bajo las hipótesis de 3.2, existen puntos  $\mathbf{x}_0 \in K_1$  y  $\mathbf{y}_0 \in K_2$  tales que la distancia  $\|\mathbf{x}_0 - \mathbf{y}_0\|$  es mínima.

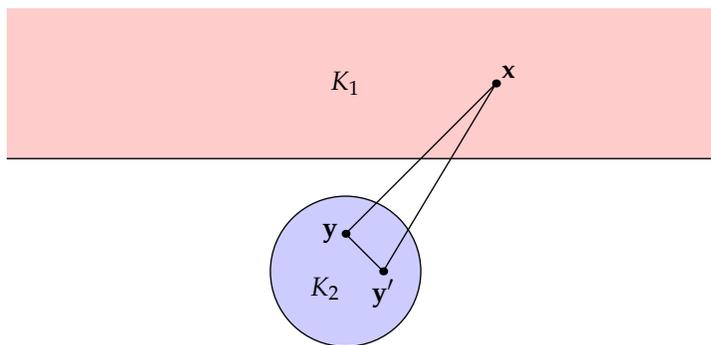
*Demostración.* Consideremos la distancia entre los puntos de nuestros conjuntos:

$$K_1 \times K_2 \rightarrow \mathbb{R}_{>0},$$

$$(\mathbf{x}, \mathbf{y}) \mapsto \|\mathbf{x} - \mathbf{y}\|.$$

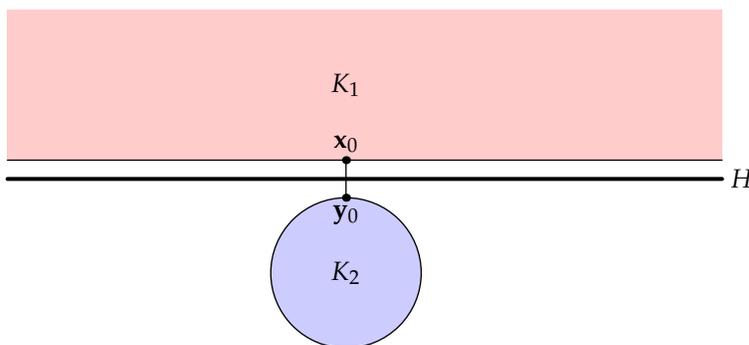
Es una función continua, por lo que si  $K_1$  y  $K_2$  son compactos, la función alcanza su valor mínimo en ciertos puntos  $\mathbf{x}' \in K_1$  y  $\mathbf{y}' \in K_2$ . Según nuestras hipótesis, solamente  $K_2$  es compacto, pero esto no es un problema. Fijemos cualquier punto  $\mathbf{y}' \in K_2$ . Luego para dos puntos  $\mathbf{x} \in K_1$  y  $\mathbf{y} \in K_2$  tenemos la desigualdad inversa del triángulo que dice que cada lado de un triángulo es más largo que la diferencia de los otros dos lados:

$$\|\mathbf{x} - \mathbf{y}\| \geq \|\mathbf{x} - \mathbf{y}'\| - \underbrace{\|\mathbf{y} - \mathbf{y}'\|}_{\leq \text{diam } K_2}.$$



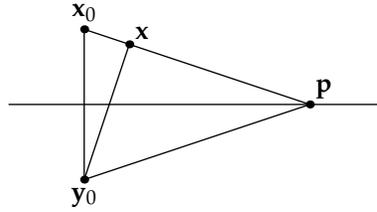
Podemos escoger otro punto  $\mathbf{x}' \in K_1$  y reemplazar  $K_1$  por el conjunto  $K'_1 := K_1 \cap B(\mathbf{y}', r)$ , donde  $\bar{B}(\mathbf{y}', r)$  es la bola cerrada de radio  $r := \|\mathbf{x}' - \mathbf{y}'\| + \text{diam}(K_2)$ . Ahora si  $\mathbf{x}$  no está en esta bola, entonces  $\|\mathbf{x} - \mathbf{y}\| > \|\mathbf{x}' - \mathbf{y}'\|$ . Esto quiere decir que la distancia mínima, si existe, se alcanza sobre  $K'_1 \times K_2$ . Pero ahora  $K'_1$  es compacto, y el mínimo existe. ■

*Demostración de 3.2.* Sean  $\mathbf{x}_0 \in K_1$  y  $\mathbf{y}_0 \in K_2$  dos puntos tales que la distancia  $\|\mathbf{x}_0 - \mathbf{y}_0\|$  es mínima. Afirmamos que el hiperplano que es perpendicular al segmento  $[\mathbf{x}_0, \mathbf{y}_0]$  y pasa por su mitad separa a  $K_1$  y  $K_2$ :

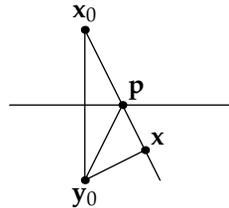


De hecho, si  $H$  no los separa, entonces  $H$  tiene un punto de intersección con  $K_1$  o  $K_2$ . Por ejemplo, supongamos que  $\mathbf{p} \in H \cap K_1$ . Consideremos el triángulo isósceles  $\mathbf{x}_0\mathbf{y}_0\mathbf{p}$ . Sea  $\mathbf{y}_0\mathbf{x}$  el perpendicular desde  $\mathbf{y}_0$  al lado  $\mathbf{x}_0\mathbf{p}$ . Tenemos dos posibilidades:

- $x \in [x_0, p]$ , y en este caso, por convexidad de  $K_1$  (¡es el punto donde tenemos que usar esta hipótesis!) también  $x \in K_1$  y  $\|x - y_0\| < \|x_0 - y_0\|$ . Pero esto contradice el hecho que  $\|x_0 - y_0\|$  es la distancia mínima entre los puntos de  $K_1$  y  $K_2$ .



- $x \notin [x_0, p]$ . En este caso vemos que  $\|p - y_0\| < \|x_0 - y_0\|$ , contradicción.



Así el hiperplano  $H$  no puede tener intersecciones con  $K_1$ . De modo similar se ve que  $H$  tampoco puede tener intersecciones con  $K_2$ . ■

Un caso particular del teorema es cuando  $K_2$  es un punto:

**3.5. Corolario.** Cada espacio convexo y cerrado  $K \subsetneq \mathbb{R}^n$  es una intersección de semiespacios cerrados.

*Demostración.* Para cada  $x \notin K$  sea  $\bar{H}_x^+$  el semiespacio cerrado tal que  $x \notin \bar{H}_x^+$ . Entonces  $K = \bigcap_{x \notin K} \bar{H}_x^+$ . ■

**3.6. Corolario.** Si  $X \subset \mathbb{R}^n$  es compacto, entonces la envolvente convexa de  $X$  es la intersección de todos los semiespacios cerrados que contienen  $X$ :

$$\text{conv } X = \bigcap_{\bar{H}^+ \supset X} \bar{H}^+$$

*Demostración.* Notemos que  $x \notin \text{conv } X$  si y solamente si  $x$  puede ser estrictamente separado de  $X$ . ■

## 4 Conjuntos polares

**4.1. Definición.** Para un subconjunto  $X \subseteq \mathbb{R}^n$  el **conjunto polar** correspondiente es el subconjunto  $X^\circ \subseteq \mathbb{R}^n$  definido por

$$X^\circ := \{x \in \mathbb{R}^n \mid \langle x, y \rangle \leq 1 \text{ para cada } y \in X\}.$$

**4.2. Ejemplo.** Si  $X = \{y\}$  es un punto, entonces el conjunto polar es simplemente

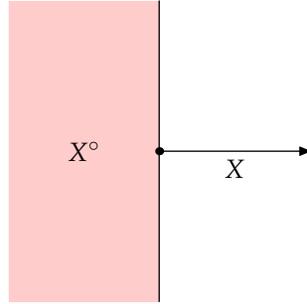
$$\{y\}^\circ = \{x \in \mathbb{R}^n \mid \langle x, y \rangle \leq 1\}.$$

Si  $y = 0$ , este coincide con  $\mathbb{R}^n$  y si  $y \neq 0$ , es un semiespacio cerrado. ▲

**4.3. Ejemplo.** Sea  $y \neq 0$  y  $X := \{\lambda y \mid \lambda \geq 0\}$  el rayo correspondiente con origen en  $0$ . Entonces

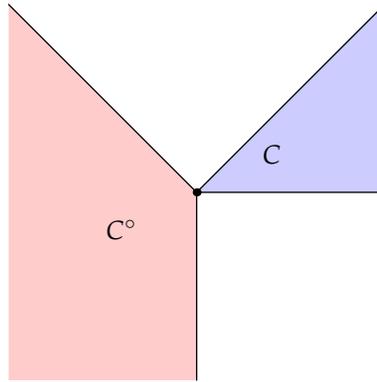
$$X^\circ = \{x \in \mathbb{R}^n \mid \langle x, y \rangle \leq 0\}$$

es el semiespacio que pasa por  $0$  y que es perpendicular a  $y$ .



En general, si C es un cono convexo de vértice  $\mathbf{0}$ , entonces

$$C^\circ = \{\mathbf{x} \in \mathbb{R}^n \mid \langle \mathbf{x}, \mathbf{y} \rangle \leq 0 \text{ para cada } \mathbf{y} \in C\}.$$



**4.4. Ejercicio.** Sea  $B(\mathbf{0}, r)$  la bola cerrada centrada en  $\mathbf{0} \in \mathbb{R}^n$  de radio  $r$ . Entonces  $B(\mathbf{0}, r)^\circ = B(\mathbf{0}, 1/r)$ .

De la definición de  $X^\circ$  está clara la siguiente

**4.5. Observación.**

- 1)  $X^\circ = (X \cup \{\mathbf{0}\})^\circ$ .
- 2) Si  $X \subseteq Y$ , entonces  $X^\circ \supseteq Y^\circ$ .
- 3) Para uniones de conjuntos tenemos  $(\bigcup_i X_i)^\circ = \bigcap_i X_i^\circ$ .

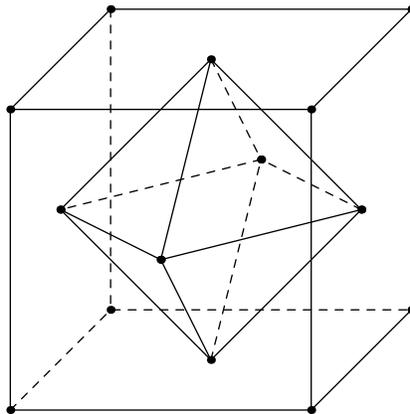
**4.6. Ejemplo.** Sea X el cubo en  $\mathbb{R}^3$  definido por

$$\{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid |x_i| \leq 1\}.$$

Notemos que es la misma cosa que la envolvente convexa de los 8 puntos  $(\pm 1, \pm 1, \pm 1)$ , y por lo tanto

$$\begin{aligned} X^\circ &= (\text{conv}\{(\pm 1, \pm 1, \pm 1)\})^\circ \\ &= \{(\pm 1, \pm 1, \pm 1)\}^\circ \\ &= \left(\bigcup\{(\pm 1, \pm 1, \pm 1)\}\right)^\circ \\ &= \bigcap\{(\pm 1, \pm 1, \pm 1)\}^\circ \\ &= \bigcap\{\pm x_1 \pm x_2 \pm x_3 \leq 1\} \\ &= \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid |x_1| + |x_2| + |x_3| \leq 1\}. \end{aligned}$$

Este conjunto es conocido como **octaedro**. Es la envolvente convexa de los 6 puntos  $(\pm 1, 0, 0)$ ,  $(0, \pm 1, 0)$ ,  $(0, 0, \pm 1)$ .



▲

**4.7. Ejercicio.** Sea  $X$  el cuadrado con vértices  $(1, 1)$ ,  $(-1, 1)$ ,  $(1, -1)$ ,  $(-1, -1)$ . Calcule  $X^\circ$ .

**4.8. Observación.** Si  $\mathbf{0} \in \text{int } X$ , entonces  $X^\circ$  es un conjunto acotado. Si  $X$  es acotado, entonces  $\mathbf{0} \in \text{int}(X^\circ)$ .

*Demostración.* Si  $\mathbf{0} \in \text{int } X$ , entonces  $X$  contiene una bola centrada en  $\mathbf{0}$ . Pero  $B \subseteq X$  implica  $X^\circ \subseteq B^\circ$ . De la misma manera, si  $X$  es acotado, entonces  $X \subseteq B$  para una bola centrada en  $\mathbf{0}$  lo suficientemente grande. ■

**4.9. Ejercicio.** Para cualquier  $X \subseteq \mathbb{R}^n$  el conjunto polar  $X^\circ$  es convexo y cerrado. Siempre tenemos  $\mathbf{0} \in X^\circ$ .

**4.10. Observación.** Para cualquier conjunto  $X \subseteq \mathbb{R}^n$  tenemos

$$X^\circ = (\text{conv } X)^\circ.$$

*Demostración.*  $X \subseteq \text{conv } X$ , entonces  $X^\circ \supseteq (\text{conv } X)^\circ$ . Tenemos que demostrar que también  $X^\circ \subseteq (\text{conv } X)^\circ$ . Si  $\mathbf{x} \in X^\circ$ , entonces  $\langle \mathbf{x}, \mathbf{y} \rangle \leq 1$  para cada  $\mathbf{y} \in X$ . Como hemos visto,  $\text{conv } X$  consiste en las combinaciones convexas  $\sum_i \lambda_i \mathbf{y}_i$  donde  $\mathbf{y}_i \in X$ ,  $\sum_i \lambda_i = 1$ ,  $\lambda_i \geq 0$ . Luego

$$\left\langle \mathbf{x}, \sum_i \lambda_i \mathbf{y}_i \right\rangle = \sum_i \lambda_i \underbrace{\langle \mathbf{x}, \mathbf{y}_i \rangle}_{\leq 1} \leq 1.$$

■

**4.11. Observación.** Para cualquier conjunto  $X \subseteq \mathbb{R}^n$  tenemos

$$X^\circ = (\overline{X})^\circ.$$

*Demostración.*  $X \subseteq \overline{X}$ , de donde  $(\overline{X})^\circ \subseteq X^\circ$ . Tenemos que demostrar que para cada punto  $\mathbf{x} \in X^\circ$  tenemos también  $\mathbf{x} \in (\overline{X})^\circ$ . Recordemos la caracterización analítica de subconjuntos cerrados de espacios métricos completos:  $X$  es cerrado si y solamente si para cada sucesión convergente de puntos  $\mathbf{y}_i \in X$  tenemos  $\lim \mathbf{y}_i \in X$ . Entonces la clausura topológica  $\overline{X}$  es precisamente el conjunto de todos los  $\lim \mathbf{y}_i$  para cada sucesión convergente  $\{\mathbf{y}_i \in X\}$ . Para  $\mathbf{x} \in X^\circ$

$$\langle \mathbf{x}, \lim \mathbf{y}_i \rangle = \lim \langle \mathbf{x}, \mathbf{y}_i \rangle \leq 1,$$

y por lo tanto  $\mathbf{x} \in (\overline{X})^\circ$ . ■

Entonces, sabemos que  $X^\circ$  es un conjunto especial: es convexo, cerrado y contiene el punto  $\mathbf{0}$ . Ahora vamos a demostrar que si  $X$  también tiene estas propiedades, entonces  $X$  es el conjunto polar para  $X^\circ$ , es decir  $(X^\circ)^\circ = X$ .

**4.12. Proposición.** Si  $K \subseteq \mathbb{R}^n$  es un conjunto convexo, cerrado, tal que  $\mathbf{0} \in K$ , entonces  $K^{\circ\circ} = K$ .

*Demostración.* Está claro que  $K \subseteq K^{\circ\circ}$ , porque para cada  $\mathbf{x} \in K$  y cada  $\mathbf{y} \in K^\circ$  tenemos  $\langle \mathbf{x}, \mathbf{y} \rangle \leq 1$ :

$$K^\circ := \{\mathbf{x} \in \mathbb{R}^n \mid \langle \mathbf{x}, \mathbf{y} \rangle \leq 1 \text{ para cada } \mathbf{y} \in K\},$$

$$K^{\circ\circ} := \{\mathbf{x} \in \mathbb{R}^n \mid \langle \mathbf{x}, \mathbf{y} \rangle \leq 1 \text{ para cada } \mathbf{y} \in K^\circ\}.$$

La inclusión no trivial es  $K^{\circ\circ} \subseteq K$ . Vamos a demostrar que si  $\mathbf{p} \notin K$ , entonces  $\mathbf{p} \notin K^{\circ\circ}$ . Para cada  $\mathbf{p} \notin K$  existe un hiperplano

$$H = \{\mathbf{x} \in \mathbb{R}^n \mid \langle \mathbf{x}, \mathbf{v} \rangle = \alpha\}, \quad \mathbf{v} \neq \mathbf{0}, \quad \alpha \in \mathbb{R}$$

tal que  $H$  separa  $K$  y  $\mathbf{p}$  estrictamente (la separación es posible gracias a la hipótesis de que  $K$  es convexo y cerrado). Notemos que  $\alpha \neq 0$  porque el hiperplano  $H$  no pasa por  $\mathbf{0}$ : de hecho,  $\mathbf{0} \in K$  por nuestra hipótesis y  $H \cap K = \emptyset$ , entonces  $\mathbf{0} \notin H$ . Sin pérdida de generalidad, podemos suponer que  $\alpha = 1$  (si no, reemplazamos  $\mathbf{v}$  por  $1/\alpha \cdot \mathbf{v}$ ). Entonces  $K$  pertenece a uno de los dos semiespacios abiertos

$$H^- := \{\mathbf{x} \in \mathbb{R}^n \mid \langle \mathbf{x}, \mathbf{v} \rangle < 1\},$$

$$H^+ := \{\mathbf{x} \in \mathbb{R}^n \mid \langle \mathbf{x}, \mathbf{v} \rangle > 1\}.$$

Para ver a cuál, es suficiente examinar el punto  $\mathbf{0} \in K$ : tenemos  $\langle \mathbf{0}, \mathbf{v} \rangle = 0 < 1$ , y entonces  $K \subset H^-$ , es decir

$$\langle \mathbf{x}, \mathbf{v} \rangle < 1 \text{ para todo } \mathbf{x} \in K,$$

y  $\mathbf{v} \in K^\circ$ . El punto  $\mathbf{p}$  pertenece al otro semiespacio  $H^+$ , de donde

$$\langle \mathbf{p}, \mathbf{v} \rangle > 1,$$

lo que quiere decir que  $\mathbf{p} \notin K^{\circ\circ}$ . ■

**4.13. Corolario.** Para cualquier conjunto  $X \subseteq \mathbb{R}^n$  tenemos

$$X^{\circ\circ} = \overline{\text{conv}(X \cup \{\mathbf{0}\})}.$$

*Demostración.* De hecho, el conjunto  $Y := \overline{\text{conv}(X \cup \{\mathbf{0}\})}$  satisface las hipótesis de la proposición de arriba y por lo tanto  $Y^{\circ\circ} = Y$ . Pero  $Y^\circ = X^\circ$  porque, como hemos observado, tomando la clausura topológica o envolvente convexa, o añadiendo  $\mathbf{0}$ , el conjunto polar no se cambia. ■

**4.14. Corolario.** Sean  $K_1, K_2 \subseteq \mathbb{R}^n$  conjuntos convexos, cerrados, que contienen  $\mathbf{0}$ .

- 1) Si  $K_1^\circ = K_2^\circ$ , entonces  $K_1 = K_2$ .
- 2) Si  $K_1^\circ \subseteq K_2^\circ$ , entonces  $K_1 \supseteq K_2$ .
- 3)  $(K_1 \cap K_2)^\circ = \overline{\text{conv}(K_1^\circ \cup K_2^\circ)}$ .

*Demostración.* Las partes 1) y 2) son consecuencias inmediatas de la fórmula  $K^{\circ\circ} = K$ .

En 3), notemos que los conjuntos  $K_1 \cap K_2$ ,  $(K_1 \cap K_2)^\circ$ ,  $\overline{\text{conv}(K_1^\circ \cup K_2^\circ)}$  son también convexos, cerrados y contienen  $\mathbf{0}$ , y

$$(K_1 \cap K_2)^{\circ\circ} = K_1 \cap K_2 = K_1^{\circ\circ} \cap K_2^{\circ\circ} = (K_1^\circ \cup K_2^\circ)^\circ = \overline{\text{conv}(K_1^\circ \cup K_2^\circ)},$$

y por la parte 1) la igualdad  $(K_1 \cap K_2)^{\circ\circ} = \overline{\text{conv}(K_1^\circ \cup K_2^\circ)}$  implica  $(K_1 \cap K_2)^\circ = \overline{\text{conv}(K_1^\circ \cup K_2^\circ)}$ . ■

## 5 Poliedros convexos

Nuestro objetivo es estudiar politopos, pero por el momento es conveniente definir un objeto un poco más general:

**5.1. Definición.** Se dice que  $X \subseteq \mathbb{R}^n$  es un **poliedro convexo** si  $X$  es una intersección finita de semiespacios cerrados.

$\mathbb{R}^n$  y  $\emptyset$  son también considerados como poliedros convexos (intersección de una familia vacía y la intersección vacía). En otras palabras, un poliedro convexo es la solución de un sistema de desigualdades lineales:

$$\{\mathbf{x} \in \mathbb{R}^n \mid f_1(\mathbf{x}) \geq 0, \dots, f_k(\mathbf{x}) \geq 0\},$$

donde  $f_i: \mathbb{R}^n \rightarrow \mathbb{R}$  son aplicaciones afines, es decir de la forma

$$f_i(\mathbf{x}) = f_i(x_1, \dots, x_n) = a_1 x_1 + \dots + a_n x_n + a_0$$

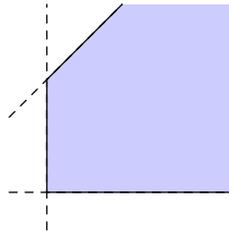
para algunos  $a_0, \dots, a_n \in \mathbb{R}$ .

**5.2. Observación.**

- 1) Cada poliedro convexo es convexo y cerrado (siendo una intersección finita de semiespacios que son convexos y cerrados).
- 2) Si  $P$  y  $Q$  son poliedros convexos, entonces  $P \cap Q$  es un poliedro convexo.
- 3) Si  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  es una aplicación afín y  $P \subset \mathbb{R}^n$  es un poliedro convexo, entonces  $f(P) \subset \mathbb{R}^m$  es también un poliedro convexo.

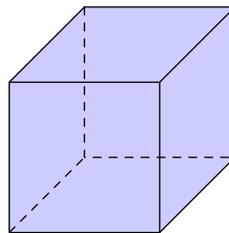
**5.3. Ejemplo.** Cada subespacio afín  $A \subset \mathbb{R}^n$  es un poliedro convexo (poco interesante). Si  $X \subset A$ , entonces  $X$  es un poliedro convexo respecto a  $A$  si y solamente si  $X$  es un poliedro convexo respecto a  $\mathbb{R}^n$ . ▲

**5.4. Ejemplo.** He aquí un poliedro convexo en  $\mathbb{R}^2$  que no es acotado:



▲

**5.5. Ejemplo.** La intersección de semiespacios  $x_1 \leq 1, x_1 \geq -1, x_2 \leq 1, x_2 \geq -1, x_3 \leq 1, x_3 \geq -1$  es un cubo en  $\mathbb{R}^3$ . Es un poliedro convexo acotado.



▲

## 6 Polítopos convexos

**6.1. Definición.** Un **politopo convexo**  $P \subset \mathbb{R}^n$  es la envolvente convexa de un conjunto finito de puntos  $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{R}^n$ .

En particular, cada politopo convexo es compacto (cerrado y acotado).

La palabra **poliedro** viene del griego clásico πολυεδρον (poliedron), de la raíz πολυς (polis), “muchas” y ἔδρα (edra), “base”, “asiento”, “cara”. La palabra **politopo** viene también de πολυς y τόπος (topos), “lugar”, “región”, “posición” (la misma raíz aparece en las palabras como “topología” y “topos”). Entonces, si las traducimos del griego, las palabras “poliedro” y “politopo” quieren decir la misma cosa. Pero para nosotros son dos objetos diferentes: un poliedro es la intersección de semiespacios cerrados (que puede ser o no acotada) y un politopo es la envolvente convexa de puntos (que es siempre acotada). Más abajo vamos a ver la relación entre las dos cosas.

Aunque construido con las raíces griegas, el termino politopo (“polytop” en alemán) fue introducido por el matemático alemán **REINHOLD HOPPE** (1816–1900) que estudió politopos regulares en dimensiones superiores. En inglés la palabra “polytope” fue introducida por **ALICIA BOOLE STOTT** (1860–1940), hija de **GEORGE BOOLE**. Curiosamente, los politopos habían sido estudiados antes por el matemático suizo **LUDWIG SCHLÄFLI** (1814–1895), pero su trabajo en geometría no era muy conocido. Alicia Boole Stott descubrió que en la dimensión 4 hay exactamente 6 politopos regulares (que se llaman pentácoron, octácoron, hexadécácoron, icositetrácoron, hecatonicosácoron, hexacosícoron) y después encontró los trabajos de Schläfli, quien había publicado el mismo resultado en el 1852.

**6.2. Observación.** Si  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  es una transformación afín y  $P \subset \mathbb{R}^n$  es un politopo convexo, entonces  $f(P) \subset \mathbb{R}^m$  es también un politopo convexo.

*Demostración.* Si  $P = \text{conv}\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ , entonces por el ejercicio 2.5  $f(P) = \text{conv}\{f(\mathbf{x}_1), \dots, f(\mathbf{x}_k)\}$ . ■

**6.3. Observación.** Si  $P$  y  $Q$  son politopos convexos, entonces  $\text{conv}(P \cup Q)$  es también un politopo convexo.

*Demostración.* Si  $P = \text{conv}\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$  y  $Q = \text{conv}\{\mathbf{y}_1, \dots, \mathbf{y}_\ell\}$ , entonces

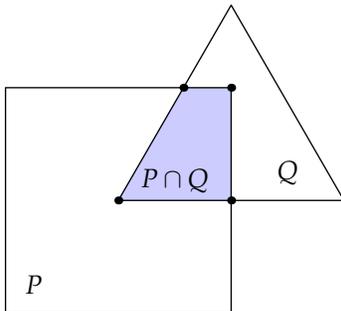
$$\text{conv}(P \cup Q) = \text{conv}(\text{conv}\{\mathbf{x}_1, \dots, \mathbf{x}_k\} \cup \text{conv}\{\mathbf{y}_1, \dots, \mathbf{y}_\ell\}) = \text{conv}\{\mathbf{x}_1, \dots, \mathbf{x}_k, \mathbf{y}_1, \dots, \mathbf{y}_\ell\}.$$

■

Intuitivamente, la siguiente observación también está clara:

**6.4. Observación.** Si  $P$  y  $Q$  son dos politopos convexos, entonces  $P \cap Q$  es también un politopo convexo.

Pero ¿cuál es la demostración? Lo que sabemos es que el conjunto  $P \cap Q$  es también convexo y compacto. No está claro cuáles son los vértices cuya envolvente convexa sea  $P \cap Q$  porque la intersección produce nuevos vértices:



El hecho de que  $P \cap Q$  sea un politopo será obvio una vez establecida la siguiente caracterización:

$$\text{politopos convexos} = \text{poliedros convexos acotados.}$$

Esto se llama el **teorema de Minkowski–Weyl** y lo vamos a demostrar más adelante en §10.

## 7 Subconjuntos extremos y caras de poliedros

**7.1. Definición.** Sea  $K \subseteq \mathbb{R}^n$  un conjunto convexo cerrado. Un subconjunto  $F \subseteq K$  se llama un **subconjunto extremo** si

- 1)  $F$  es convexo,
- 2) para cada  $x_1, x_2 \in K$  si  $(x_1, x_2) \cap F \neq \emptyset$ , entonces  $x_1, x_2 \in F$ .

**7.2. Ejemplo.**  $K$  y  $\emptyset$  son subconjuntos extremos de  $K$  (triviales). ▲

**7.3. Ejemplo.** Notemos que para un punto  $x \in K$  las siguientes condiciones son equivalentes:

- 1)  $\{x\} \subset K$  es un subconjunto extremo,
- 2)  $K$  no contiene segmentos abiertos  $(x_1, x_2)$  que pasan por  $x$ ,
- 3)  $K \setminus \{x\}$  es también convexo.

En este caso se dice que  $x$  es un **punto extremo** de  $K$ . ▲

**7.4. Definición.** Un subconjunto extremo de un poliedro se llama **cara**. Las caras de dimensión 0 se llaman **vértices** y las caras de dimensión 1 se llaman **aristas**.

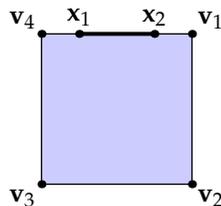
**7.5. Ejercicio.** Sea  $K \subseteq \mathbb{R}^n$  un conjunto convexo cerrado.

- 1) Si  $F_1 \subseteq K$  y  $F_2 \subseteq K$  son dos subconjuntos extremos, entonces  $F_1 \cap F_2$  es un subconjunto extremo. En particular, la intersección de caras de un poliedro es también una cara.
- 2) Sea  $F \subseteq K$  un subconjunto extremo de  $K$ . Entonces los subconjuntos extremos de  $F$  son los subconjuntos extremos de  $K$  contenidos en  $F$ .

**7.6. Ejemplo.** Sea  $K \subset \mathbb{R}^2$  el cuadrado definido por

$$\begin{aligned} K &:= \text{conv}\{\mathbf{v}_1 := (1, 1), \mathbf{v}_2 := (1, -1), \mathbf{v}_3 := (-1, -1), \mathbf{v}_4 := (-1, 1)\} \\ &= \{(x_1, x_2) \in \mathbb{R}^2 \mid |x_1| \leq 1, |x_2| \leq 1\}. \end{aligned}$$

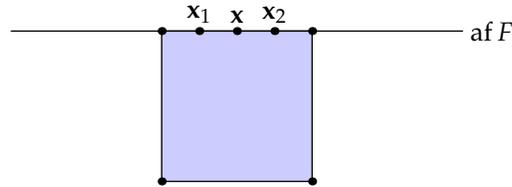
- Cada vértice  $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$  forma un conjunto extremo, porque  $K \setminus \{\mathbf{v}_i\}$  es convexo (y son los únicos puntos que se puede quitar sin perder la convexidad).
- Cada segmento  $[\mathbf{v}_i, \mathbf{v}_j]$  para  $i \neq j$  es también un subconjunto extremo. Tales segmentos son aristas de  $K$ . Las intersecciones de aristas son los vértices.





**7.7. Observación.** Sea  $K$  un conjunto convexo cerrado y  $F \subseteq K$  un subconjunto extremo. Entonces  $F = K \cap \text{af } F$ .  
 En particular, cada cara  $F$  de un poliedro  $P$  es también un poliedro, siendo la intersección de  $P$  con un subespacio afín de  $\mathbb{R}^n$ .

*Demostración.* Obviamente  $F \subseteq K \cap \text{af } F$ , porque  $F \subseteq K$  y  $F \subseteq \text{af } F$ . El interior relativo de  $F$ , que es el interior de  $F$  como un subespacio de  $\text{af } F$ , no es vacío. Escojamos un punto  $x \in \text{intrel } F$ . Para cualquier punto  $x_1 \in K \cap \text{af } F$ , si  $x_1 = x$ , entonces la demostración está terminada. Si  $x \neq x_1$ , entonces existe otro punto  $x_2 \in F$ ,  $x_2 \neq x_1$ ,  $x_2 \neq x$  tal que  $x \in (x_1, x_2)$  (porque  $x$  está en el interior), y por la extremidad de  $F$  concluimos que  $x_1 \in F$ .



**7.8. Observación.** Sea  $K$  un conjunto convexo cerrado. Entonces  $K$  es la unión disjunta de  $\text{intrel } F$  para cada subconjunto extremo  $F \subseteq K$ . En particular, todos los subconjuntos extremos propios  $F \subsetneq K$  están contenidos en  $K \setminus \text{intrel } K$ .

**7.9. Ejemplo.** Aquí está la descomposición en  $\text{intrel } F$  para el cuadrado.



Primero demostremos el siguiente

**7.10. Lema.** Sea  $K$  un conjunto convexo cerrado. Si  $F, F' \subset K$  son dos subconjuntos extremos tales que  $F' \cap \text{intrel } F \neq \emptyset$ , entonces  $F \subseteq F'$ .

*Demostración.* Fijemos un punto  $x \in F' \cap \text{intrel } F$ . Entonces

$$F = K \cap \text{af } F = \bigcup_{\ell \text{ recta en } \text{af } F, x \in \ell} (K \cap \ell) = \bigcup_{\substack{x_1, x_2 \in K \\ x_1 \neq x, x_2 \neq x \\ x \in (x_1, x_2)}} [x_1, x_2] \subseteq F',$$

porque  $F'$  es extremo y  $x \in (x_1, x_2) \cap F'$  implica  $[x_1, x_2] \subset F'$ .



*Demostración de 7.8.* Antes de todo, veamos la unión es disjunta, porque el lema de arriba implica que si  $F$  y  $F'$  son dos subconjuntos extremos tales que  $\text{intrel } F \cap \text{intrel } F' \neq \emptyset$ , entonces  $F \subseteq F'$  y  $F' \subseteq F$ .

Ahora tenemos que ver por qué la unión de  $\text{intrel } F$  es todo  $K$ . Usamos inducción sobre la dimensión. Si  $K$  es de dimensión 0 (un punto), es obvio. Luego, cada punto  $x \in K$  pertenece o bien a  $\text{intrel } K$ , o bien a un subconjunto extremo de dimensión menor.



**7.11. Proposición.** Sea  $P \subseteq \mathbb{R}^n$  un poliedro convexo, es decir una intersección finita

$$(*) \quad P = \overline{H}_1^+ \cap \cdots \cap \overline{H}_k^+,$$

donde  $\overline{H}_i^+$  son semiespacios cerrados que corresponden a hiperplanos  $H_i := \text{fr } \overline{H}_i^+$ . Supongamos que  $\text{int } P \neq \emptyset$  (es decir,  $\dim P = n$ ). Entonces cada cara propia  $F \subsetneq P$  de dimensión máxima  $n - 1$ , es de la forma  $H_i \cap P$  para algún  $1 \leq i \leq k$ . En particular,  $P$  tiene un número finito de caras de dimensión  $n - 1$ .

(En general, no todo subconjunto  $H_i \cap P$  es una cara de dimensión máxima  $n - 1$ , pero es verdad si la presentación (\*) es mínima, es decir, si quitamos algún  $\overline{H}_i^+$ , la intersección ya no coincide con  $P$ .)

*Demostración.* Sea  $F \subsetneq P$  una cara de dimensión  $n - 1$ . Entonces  $\mathbf{x} \in \text{intrel } F$  implica  $\mathbf{x} \in \text{fr } P$ . Notemos que

$$\text{fr } P = P \setminus \text{int } P = \bigcap_i \overline{H}_i^+ \setminus \bigcap_i \text{int } \overline{H}_i^+.$$

Entonces,  $\mathbf{x} \in H_i$  para algún  $i$ . Luego  $F = \text{af } F \cap P = H_i \cap P$ . (De hecho, si  $\text{af } F \neq H_i$ , pero  $\mathbf{x} \in F$  y  $\mathbf{x} \in H_i$ ; entonces el hiperplano  $H_i$  divide  $\text{af } F$  en dos partes, lo cual es imposible.) ■

**7.12. Proposición.** Sea  $P$  un politopo de dimensión  $n$ . Si  $F \subsetneq P$  es una cara de  $P$  de dimensión  $k < n$ , entonces  $F \subsetneq F'$  para alguna cara  $F'$  de dimensión  $k + 1$ .

*Demostración.* Es suficiente de demostrar que para cada cara  $F$  existe una cara  $F'$  de dimensión  $n - 1$  tal que  $F \subset F'$ . Esto nos da la proposición para  $k = n - 2$ , y luego para  $k < n - 2$  podemos usar inducción sobre  $k$ .

Sin pérdida de generalidad,  $\text{int } P \neq \emptyset$ . Entonces por la proposición de arriba, si  $\mathbf{x} \in \text{intrel } F$ , existe una cara  $F'$  de dimensión  $n - 1$  tal que  $\mathbf{x} \in F'$ . Pero esto implica  $F \subset F'$  por 7.10. ■

**7.13. Proposición.** Sea  $P = \bigcap_{1 \leq i \leq k} \overline{H}_i^+$  un poliedro convexo. Entonces para cada propia cara  $F \subsetneq P$  tenemos

$$\text{af } F = H_{i_1} \cap \cdots \cap H_{i_\ell}$$

para algunos índices  $1 \leq i_1 < \cdots < i_\ell \leq k$ . En particular, el número de caras es finito.

*Demostración.* Sea  $\mathbf{x} \in \text{intrel } F$ . Sean  $i_1, \dots, i_\ell$  los índices tales que  $\mathbf{x} \in H_{i_j}$ . Entonces  $\text{af } F \subset H_{i_j}$  para cada  $j$  por 7.10.

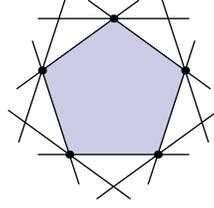
Ahora tenemos que ver  $H_{i_1} \cap \cdots \cap H_{i_\ell} \subset \text{af } F$ . Supongamos que es falso, es decir existe una bola  $B(\mathbf{x}, \epsilon)$  tal que  $B(\mathbf{x}, \epsilon) \cap H_{i_1} \cap \cdots \cap H_{i_\ell} \subset P$ . Pero esto implica que  $F$  no es una cara, porque existen puntos  $\mathbf{x}_1, \mathbf{x}_2 \in B(\mathbf{x}, \epsilon) \cap H_{i_1} \cap \cdots \cap H_{i_\ell}$  tales que  $\mathbf{x}_1, \mathbf{x}_2 \notin \text{af } F$  y  $\mathbf{x} \in (\mathbf{x}_1, \mathbf{x}_2)$ . ■

## 8 Hiperplanos de soporte

**8.1. Definición.** Sea  $X \subseteq \mathbb{R}^n$  un conjunto. Se dice que un hiperplano  $H \subseteq \mathbb{R}^n$  es un **hiperplano de soporte** para  $X$  si  $H \cap \overline{X} \neq \emptyset$  y  $X$  está en el semiespacio cerrado  $\overline{H}^+$  o  $\overline{H}^-$ .

**8.2. Proposición.** Si  $K \subseteq \mathbb{R}^n$  es un conjunto convexo y  $\mathbf{x} \in \text{fr } K := \overline{K} \setminus \text{int } K$ , entonces existe un hiperplano de soporte  $H \ni \mathbf{x}$ .

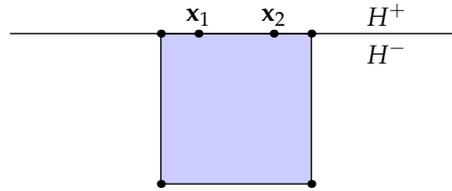
He aquí un ejemplo en  $\mathbb{R}^2$  (noten que en ciertos puntos de  $\text{fr } K$  hay varios hiperplanos de soporte):



*Demostración.* Si  $\text{int } K = \emptyset$ , entonces  $K$  pertenece a algún hiperplano en  $\mathbb{R}^n$  que será un hiperplano de soporte. Así podemos suponer que  $\text{int } K \neq \emptyset$  y, como hemos visto antes (1.17), en este caso  $\text{int } K = \text{int } \bar{K}$  y  $\text{fr } K = \bar{K}$ , y por lo tanto sin pérdida de generalidad podemos reemplazar  $K$  por  $\bar{K}$ . También podemos suponer que  $\mathbf{x} = \mathbf{0}$ . Lo que sabemos es cómo se puede separar un punto  $\mathbf{y} \notin K$  por un hiperplano. El hecho que  $\mathbf{x}$  esté en la frontera significa que existe una sucesión  $\mathbf{x}_i \in \mathbb{R}^n \setminus K$  tal que  $\lim \mathbf{x}_i = \mathbf{x}$ . Para cada  $i$  sea  $H_i = \{\mathbf{y} \in \mathbb{R}^n \mid \langle \mathbf{y}, \mathbf{v}_i \rangle = \alpha_i\}$  un hiperplano que separa  $\mathbf{x}_i$  de  $K$ . Aquí  $\mathbf{v}_i$  es el vector normal a  $H_i$  (tal que  $\|\mathbf{v}_i\| = 1$ ) y  $\alpha_i > 0$  son parámetros tales que  $\lim \alpha_i = \lim \text{dist}(\mathbf{0}, H_i) = 0$ . Reemplazando  $\{\mathbf{v}_i\}$  por una subsucesión convergente, podemos tomar  $\mathbf{v}_0 := \lim \mathbf{v}_i$ . Consideremos el hiperplano  $H := \{\mathbf{y} \in \mathbb{R}^n \mid \langle \mathbf{y}, \mathbf{v}_0 \rangle = 0\}$ . Para cada  $i$  tenemos  $K \subset H_i^+$ , es decir  $\langle \mathbf{y}, \mathbf{v}_i \rangle > \alpha_i$  para cada  $\mathbf{y} \in K$ , entonces  $\langle \mathbf{y}, \mathbf{v}_0 \rangle \geq 0$  y  $K \subseteq \bar{H}^+$ . ■

**8.3. Observación.** Si  $H$  es un hiperplano de soporte para  $K$ , entonces  $K \cap H$  es un subconjunto extremo.

*Demostración.* Supongamos que para  $\mathbf{x}_1, \mathbf{x}_2 \in K$  se tiene  $(\mathbf{x}_1, \mathbf{x}_2) \cap (K \cap H) \neq \emptyset$ . Supongamos por absurdo que  $[\mathbf{x}_1, \mathbf{x}_2] \not\subset H$ . Entonces los puntos  $\mathbf{x}_1$  y  $\mathbf{x}_2$  están en diferentes semiespacios  $\bar{H}^+$  y  $\bar{H}^-$ . Pero esto contradice el hecho de que todo  $K$  pertenece a uno de estos semiespacios (por la definición de hiperplanos de soporte).



**8.4. Proposición.** Cada cara de un poliedro convexo  $P = \bigcap_{1 \leq i \leq k} H_i^+$  es la intersección de  $P$  con algún hiperplano de soporte.

*Demostración.* Podemos suponer que  $\text{int } P \neq \emptyset$ . Sea  $F$  una cara tal que

$$\text{af } F = H_{i_1} \cap \cdots \cap H_{i_\ell},$$

donde  $H_i^+ = \{\mathbf{x} \in \mathbb{R}^n \mid f_i(\mathbf{x}) \geq 0\}$  son algunos semiespacios cerrados definidos por aplicaciones afines  $f_i: \mathbb{R}^n \rightarrow \mathbb{R}$ . Entonces

$$\text{af } F = H := \{\mathbf{x} \in \mathbb{R}^n \mid f_{i_1}(\mathbf{x}) + \cdots + f_{i_\ell}(\mathbf{x}) = 0\}$$

es un hiperplano de soporte tal que  $H \cap P = F$ . ■

## 9 Puntos extremos y el teorema de Krein–Milman

Sea  $K$  un conjunto convexo cerrado. Les recuerdo que  $x \in K$  se llama un **punto extremo** si  $\{x\} \subset K$  es un subconjunto extremo, que es la misma cosa que  $K \setminus \{x\}$  sea convexo.

**9.1. Ejemplo.** Aquí está un plano de estadio de fútbol con sus puntos extremos representados por las líneas gruesas.



El siguiente hecho demuestra el significado de los puntos extremos:

**9.2. Teorema (Krein, Milman).** Sea  $K \subset \mathbb{R}^n$  un conjunto convexo y compacto. Sea

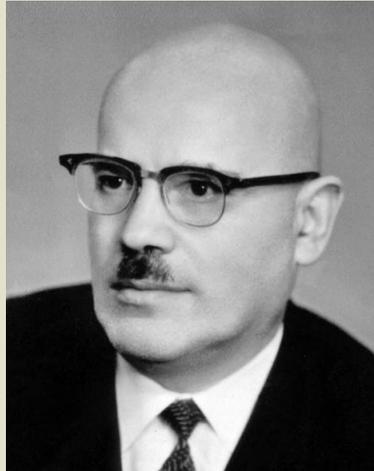
$$\text{ex}(K) := \{x \in K \mid \{x\} \subset K \text{ es extremo}\} = \{x \in K \mid K \setminus \{x\} \text{ es convexo}\}.$$

el conjunto de sus puntos extremos. Entonces

$$K = \text{conv ex } K.$$

*Demostración.* Obviamente,  $\text{conv ex } K \subseteq K$  y tenemos que demostrar que  $K \subseteq \text{conv ex } K$ . Usamos inducción sobre la dimensión. En dimensiones 0 y 1, cuando  $K$  es un punto o un segmento cerrado, está claro que  $K = \text{conv ex } K$ . Para cualquier punto  $x \in K$  sea  $\ell$  una recta que pasa por  $x$ . Tenemos  $[x_1, x_2] := \ell \cap K$  (porque  $K$  es compacto por nuestra hipótesis) para algunos puntos  $x_1, x_2$  en la frontera  $\text{fr } K$ . Por convexidad de  $\text{conv ex } K$ , es suficiente demostrar que  $x_1, x_2 \in \text{conv ex } K$ . Para  $i = 1, 2$  sea  $H_i$  un hiperplano de soporte que pasa por  $x_i$ . El subconjunto  $H_i \cap K$  tiene dimensión menor y podemos aplicar la hipótesis de inducción a  $H_i \cap K$  como un subconjunto del espacio afín  $H_i$ . Entonces  $H_i \cap K = \text{conv ex}(H_i \cap K)$ . Pero  $\text{ex}(H_i \cap K) = H_i \cap \text{ex } K \subseteq \text{ex } K$ , entonces  $H_i \cap K \subseteq \text{conv ex } K$ . ■

La proposición de arriba (para un subconjunto convexo compacto de  $\mathbb{R}^n$ ) se atribuye a Minkowski, pero en su mayor generalidad se conoce como el **teorema de Krein–Milman** y fue demostrada por los matemáticos soviéticos **MARK KREIN** (1907–1989) y su estudiante **DAVID MILMAN** (1912–1982) que trabajaban en el análisis funcional. Krein era estudiante de **NIKOLÁI CHEBOTARIOV** (1894–1947), otro matemático soviético famoso por sus resultados en teoría de números. Krein y Milman trabajaban en Odessa (Ucrania). Durante la guerra la universidad fue evacuada, y después de su regreso en 1944 Krein fue despedido por acusaciones del “nacionalismo judío” y tuvo posiciones en el Instituto de ingeniería marítima y el Instituto de ingeniería civil de Odessa. No obstante la persecución política y imposibilidad de atender los congresos en el extranjero, Krein logró reconocimiento internacional y publicó más de 270 artículos.



Mark Krein

Milman también fue discriminado por sus orígenes judíos, fue despedido y tuvo que trabajar en el Instituto de telecomunicaciones de Odessa. En 1973, después de su jubilación, se mudó a Israel, como muchos otros judíos soviéticos.

## 10 Teorema de Minkowski–Weyl

**10.1. Teorema (Minkowski, Weyl).** Para un conjunto  $P \subset \mathbb{R}^n$  las siguientes afirmaciones son equivalentes:

- 1)  $P$  es un poliedro (una intersección de un número finito de semiespacios cerrados) y  $P$  es acotado,
- 2)  $P$  es un politopo convexo (la envolvente convexa de un conjunto finito  $\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ ).

*Demostración.* Si  $P$  es un poliedro convexo acotado, entonces  $P = \text{conv} P$  por el teorema de Krein–Milman. El conjunto de puntos extremos  $\text{ex} P$  está formado por los vértices de  $P$  y es un conjunto finito.

Ahora sea  $P = \text{conv}\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$  un politopo convexo. Sin pérdida de generalidad, podemos considerar  $P$  como un politopo en  $\text{af} P$  y suponer que  $\text{int} P \neq \emptyset$ . También podemos suponer que  $\mathbf{0} \in \text{int} P$  aplicando una traslación. Como hemos observado en 4.8,  $\mathbf{0} \in \text{int} P$  implica que el conjunto polar  $P^\circ$  es acotado y además se ve que es un poliedro:

$$\begin{aligned}
 P^\circ &= (\text{conv}\{\mathbf{x}_1, \dots, \mathbf{x}_k\})^\circ \\
 &= \{\mathbf{x}_1, \dots, \mathbf{x}_k\}^\circ \\
 &= \left( \bigcup_{1 \leq i \leq k} \{\mathbf{x}_i\} \right)^\circ \\
 &= \bigcap_{1 \leq i \leq k} \{\mathbf{x}_i\}^\circ
 \end{aligned}$$

donde  $\{\mathbf{x}_i\}^\circ$  son semiespacios cerrados. Entonces,  $P^\circ$  es también un poliedro convexo acotado. El mismo razonamiento demuestra que  $P^{\circ\circ}$  es un politopo convexo acotado, pero  $P^{\circ\circ} = P$ . ■

El teorema anterior fue demostrado por Minkowski (Minkowski, H. Allgemeine Lehrsätze über die konvexe Polyeder, Nachr. Ges. Wiss., Göttingen, 1897, 198–219) y otro matemático alemán **HERMANN WEYL** (1885–1955) (Weyl, H. Elementare Theorie der konvexen Polyeder, Comment. Math. Helvetici, 1935, 7). Minkowski y Weyl representaban la escuela de matemáticas de la Universidad de Gotinga, donde también trabajaba Hilbert, que era el director de tesis de Weyl. Al principio de la guerra en 1933 Weyl huyó a los Estados Unidos, donde tuvo plaza en el Instituto de Estudios Avanzados.



Weyl hizo varias contribuciones en geometría (entre otras cosas, dio definiciones rigurosas de las superficies de Riemann y variedades topológicas, estudió representaciones de grupos de Lie compactos, etc.), lógica y física matemática.

No confundir con el matemático francés **ANDRÉ WEIL** (1906–1998). El apellido de Weyl se pronuncia [vail] y el apellido de Weil se pronuncia [veyl].

## 11 Poliedros en SAGE

Para hacer cálculos con poliedros existen varios programas, y aquí voy a explicar cómo usar SAGE. Pueden descargar este programa y obtener documentación completa en el sitio <http://sagemath.org/>. Para probar los ejemplos de abajo sin instalar el programa (que pesa alrededor de 3 gigabytes), pueden también usar el sitio <https://cloud.sagemath.com/>.

Para definir un politopo  $P = \text{conv}\{\mathbf{v}_0, \dots, \mathbf{v}_k\}$ , se puede especificar una lista de puntos. Por ejemplo, si  $\mathbf{v}_0 = (1, 0, 0)$ ,  $\mathbf{v}_1 = (0, 1, 0)$ ,  $\mathbf{v}_2 = (0, 0, 1)$ , entonces escribimos

```
sage: P = Polyhedron(vertices = [[1,0,0],[0,1,0],[0,0,1]])
sage: P
A 2-dimensional polyhedron in ZZ^3 defined as the convex hull of 3 vertices
```

La lista de vértices de  $P$  se obtiene por `P.vertices()`:

```
sage: P.vertices()
(A vertex at (0, 0, 1), A vertex at (0, 1, 0), A vertex at (1, 0, 0))
```

Para obtener las desigualdades y ecuaciones que definen a  $P$ , escribimos `P.Hrepresentation()`:

```
sage: P.Hrepresentation()
(An equation (1, 1, 1) x - 1 == 0,
```

An inequality  $(0, -1, -1) x + 1 \geq 0$ ,  
 An inequality  $(0, 1, 0) x + 0 \geq 0$ ,  
 An inequality  $(0, 0, 1) x + 0 \geq 0$

Esto quiere decir que  $P = \text{conv}\{\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2\}$  es la misma cosa que el conjunto en  $\mathbb{R}^3$  definido por

$$\begin{aligned} x_1 + x_2 + x_3 - 1 &= 0, \\ -x_2 - x_3 + 1 &\geq 0, \\ x_2 &\geq 0, \\ x_3 &\geq 0. \end{aligned}$$

Para obtener las caras de dimensión  $d$ , escribimos `P.faces(d)`:

```
sage: P.faces(0)
(<0>, <1>, <2>)
sage: P.faces(1)
(<0,1>, <0,2>, <1,2>)
sage: P.faces(2)
(<0,1,2>,)
sage: P.faces(3)
()
```

Por ejemplo, `P.faces(2)` nos da `(<0,1,2>,)`, lo cual significa que hay una cara de dimensión 2, a saber  $\text{conv}\{\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2\}$  (los números 0,1,2 corresponden a los elementos de la lista `P.vertices()`).

Algunos politopos básicos ya están definidos por SAGE. Por ejemplo,

- `polytopes.simplex(d)` es el símplice de dimensión  $d$  en  $\mathbb{R}^{d+1}$  que tiene como vértices los  $d+1$  puntos  $\mathbf{v}_i := (0, \dots, 1, \dots, 0)$ . Por ejemplo, en nuestros ejemplos de arriba,  $P = \text{polytopes.simplex}(2)$ .
- `polytopes.hypercube(d)` es el hipercubo de dimensión  $d$  en  $\mathbb{R}^d$  que tiene como vértices los  $2^d$  puntos con coordenadas  $\pm 1$ .
- `polytopes.cube()` es la misma cosa que `polytopes.hypercube(3)`.
- `polytopes.tetrahedron()` es el tetraedro en  $\mathbb{R}^3$  de vértices  $(0,0,0)$ ,  $(0,1,1)$ ,  $(1,0,1)$ ,  $(1,1,0)$ .
- `polytopes.octahedron()` es el octaedro en  $\mathbb{R}^3$  de vértices  $(\pm 1,0,0)$ ,  $(0,\pm 1,0)$ ,  $(0,0,\pm 1)$ .
- `polytopes.icosahedron()` es el icosaedro en  $\mathbb{R}^3$  de vértices:

```
sage: I = polytopes.icosahedron()
sage: I
A 3-dimensional polyhedron in
(Number Field in sqrt5 with defining polynomial x^2 - 5)^3
defined as the convex hull of 12 vertices
sage: I.vertices()
(A vertex at (0, 1/2, 1/4*sqrt5 + 1/4),
 A vertex at (0, -1/2, 1/4*sqrt5 + 1/4),
 A vertex at (1/2, 1/4*sqrt5 + 1/4, 0),
 A vertex at (1/2, -1/4*sqrt5 - 1/4, 0),
 A vertex at (1/4*sqrt5 + 1/4, 0, 1/2),
 A vertex at (1/4*sqrt5 + 1/4, 0, -1/2),
 A vertex at (-1/2, 1/4*sqrt5 + 1/4, 0),
```

```

A vertex at (-1/2, -1/4*sqrt5 - 1/4, 0),
A vertex at (-1/4*sqrt5 - 1/4, 0, 1/2),
A vertex at (0, 1/2, -1/4*sqrt5 - 1/4),
A vertex at (0, -1/2, -1/4*sqrt5 - 1/4),
A vertex at (-1/4*sqrt5 - 1/4, 0, -1/2))

```

- `polytopes.dodecahedron()` es el dodecaedro en  $\mathbb{R}^3$  con 20 vértices.
- Etc. — vean la lista completa en <http://doc.sagemath.org/html/en/reference/geometry/sage/geometry/polyhedron/library.html>

Para calcular el politopo polar, se usa `P.polar()`. Por ejemplo, podemos comprobar que el politopo polar del cubo es el octaedro y el politopo polar del icosaedro es el dodecaedro:

```

sage: polytopes.cube().polar() == polytopes.octahedron()
True
sage: polytopes.icosahedron().polar() == polytopes.dodecahedron()
True

```

SAGE trata de modo diferente los politopos definidos sobre anillos diferentes. Por ejemplo, el tetraedro en SAGE es definido como la envoltura convexa de los puntos  $(0,0,0)$ ,  $(0,1,1)$ ,  $(1,0,1)$ ,  $(1,1,0)$ , y entonces es definido sobre  $\mathbb{Z}$ . SAGE calcula el politopo polar solo cuando  $\mathbf{0} \in \text{int } P$  (porque si  $\mathbf{0} \notin \text{int } P$ , entonces  $P^\circ$  no es un conjunto acotado), lo cual no es el caso para el tetraedro:

```

sage: polytopes.tetrahedron().polar()
-----
. . . . .
ValueError: The polytope must have the IP property.

```

Pero podemos definir el mismo politopo sobre  $\mathbb{Q}$  escribiendo

```

Polyhedron(vertices = [[0,0,1],[0,1,0],[1,0,0],[0,0,0]], base_ring=QQ)

```

En este caso SAGE va a trasladar  $P$  de modo que  $\mathbf{0} \in \text{int } P$  y después calcular  $P^\circ$  (que es también un tetraedro):

```

sage: P = Polyhedron(vertices = [[0,0,1],[0,1,0],[1,0,0],[0,0,0]], base_ring=QQ)
sage: Q = P.polar()
sage: Q.vertices()
(A vertex at (4, 4, 4),
 A vertex at (0, 0, -4),
 A vertex at (-4, 0, 0),
 A vertex at (0, -4, 0))

```

Para calcular el volumen, hay que escribir `P.volume()`:

```

sage: polytopes.cube().volume()
8
sage: polytopes.tetrahedron().volume()
1/3
sage: polytopes.octahedron().volume()
4/3
sage: polytopes.icosahedron().volume()
5/12*sqrt5 + 5/4

```

```
sage: polytopes.dodecahedron().volume()
-176*sqrt(5) + 400
```

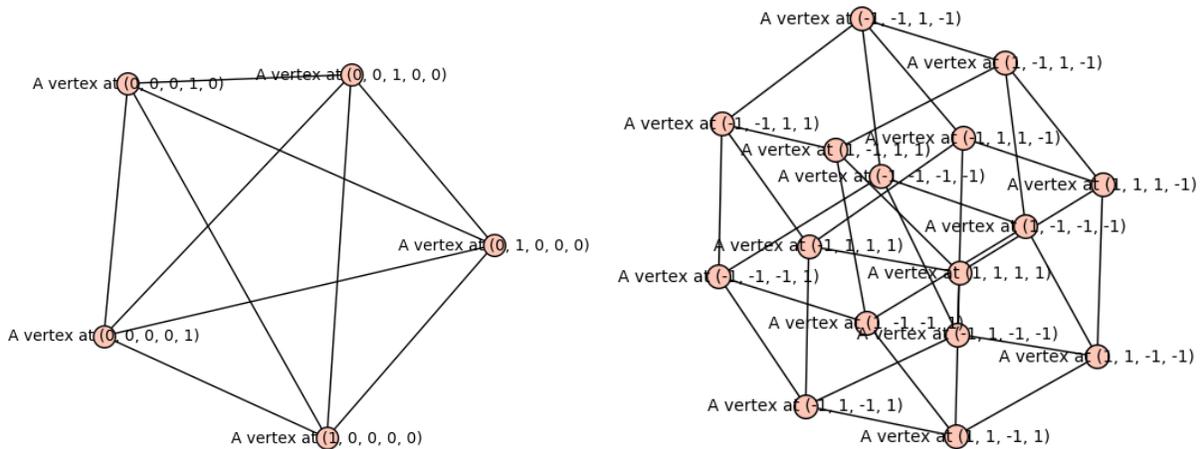
Noten que el volumen se calcula como el volumen en el espacio ambiente. Por ejemplo, el  $d$ -símplice tiene volumen 0 porque es considerado como un subconjunto en  $\mathbb{R}^{d+1}$ . Pero `P.affine_hull()` nos da el mismo politopo pero considerado como un subconjunto en su envolvente afín. Así `polytopes.simplex(d).volume() = 0`, pero `polytopes.simplex(d).affine_hull().volume() = 1/d!`:

```
sage: polytopes.simplex(2).volume()
0
sage: polytopes.simplex(2).affine_hull().volume()
1/2
sage: polytopes.simplex(3).affine_hull().volume()
1/6
sage: polytopes.simplex(4).affine_hull().volume()
1/24
```

Podemos visualizar el grafo de vértices de un politopo, o el mismo politopo, si tiene dimensión  $\leq 4$ . Por ejemplo, el código

```
sage: polytopes.simplex(4).vertex_graph().graphplot().show()
sage: polytopes.hypercube(4).vertex_graph().graphplot().show()
```

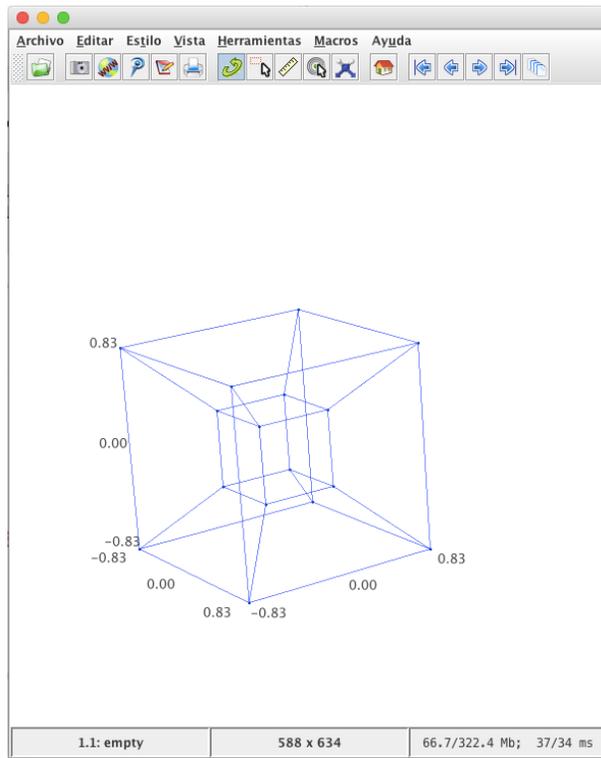
produce las siguientes imágenes:



Para ver mejor la estructura del 4-cubo, podemos también escribir

```
sage: polytopes.hypercube(4).show()
```

El resultado es un modelo interactivo:



**11.1. Ejercicio.** Descargue SAGE, busque la documentación sobre polítopos y pruebe algunos ejemplos.

# Parte II

## Polinomios de Ehrhart

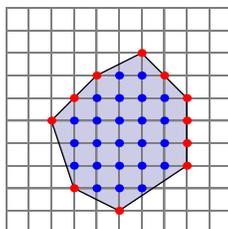
En esta parte vamos a ver algunos ejemplos del siguiente fenómeno sorprendente: si  $P \subset \mathbb{R}^d$  es un politopo convexo de dimensión  $d$  con vértices en  $\mathbb{Z}^d$  y  $mP$  son sus dilataciones por un factor  $m = 0, 1, 2, \dots$ , entonces el número de puntos enteros en  $mP$ , es decir  $\#(mP \cap \mathbb{Z}^d)$ , es un polinomio  $\ell_P(m)$  de grado  $d$  en el parámetro  $m$ . Este polinomio se llama el **polinomio de Ehrhart** de  $P$ . Luego, la evaluación de este polinomio en  $-m$  calcula el número de puntos que están en el interior de  $P$ . Estas propiedades generalizan el famoso teorema de Pick sobre polígonos convexos en  $\mathbb{R}^2$ .

### 12 Teorema de Pick

Sea  $P \subset \mathbb{R}^2$  un **polígono convexo entero**, es decir la envolvente convexa de algunos puntos de  $\mathbb{Z}^2 \subset \mathbb{R}^2$ . Entonces el área de  $P$  puede ser calculada contando los puntos enteros en  $P$ . A saber,

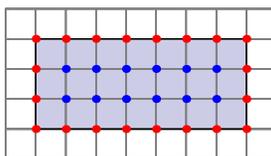
**12.1. Teorema (Teorema de Pick).** *Sea  $P$  un polígono convexo entero. Sea  $B$  el número de puntos enteros en la frontera de  $P$  y  $I$  el número de los puntos enteros en el interior de  $P$ . Entonces el área de  $P$  viene dada por*

$$A = I + \frac{1}{2} B - 1.$$

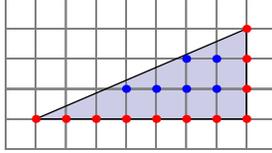


Hay muchas demostraciones de este teorema. Voy a explicar el argumento elemental del libro de Beck y Robbins. Antes de todo, examinemos ciertos casos básicos:

**12.2. Ejercicio.** *Sea  $P$  un rectángulo con lados paralelos a los ejes. Demuestre directamente la identidad  $A = I + \frac{1}{2} B - 1$ .*



**12.3. Ejercicio.** *Sea  $P$  un triángulo rectángulo con dos lados paralelos a los ejes. Demuestre la identidad  $A = I + \frac{1}{2} B - 1$ . Para simplificar su vida, puede suponer que la hipotenusa no contiene puntos enteros (excepto los dos vértices).*

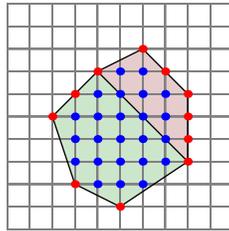


**12.4. Ejercicio.** A propósito, ¿cuándo es que el rectángulo del ejercicio precedente no contiene puntos enteros en la hipotenusa? Cuente el número de puntos enteros en el segmento de la línea entre  $(a,b)$  y  $(c,d)$  para algunos  $a, b, c, d \in \mathbb{Z}$ .

Ahora la observación clave es que se puede cortar el polígono en polígonos más pequeños trazando algunas de sus diagonales, de modo que la identidad  $A = I + \frac{1}{2} B - 1$  para cada parte implica la misma identidad para  $P$ :

**12.5. Observación.** Supongamos que  $P$  está dividido por algunas de sus diagonales en los polígonos  $P_1, \dots, P_r$ . Entonces  $A_k = I_k + \frac{1}{2} B_k - 1$  para cada  $1, \dots, r$  implica  $A = I + \frac{1}{2} B - 1$ .

*Demostración.* Por inducción, es suficiente de examinar el caso  $r = 2$ :

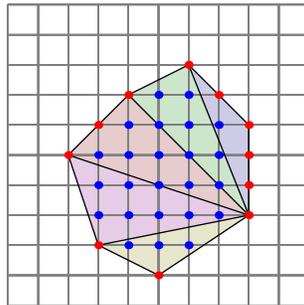


Supongamos que  $A_1 = I_1 + \frac{1}{2} B_1 - 1$  y  $A_2 = I_2 + \frac{1}{2} B_2 - 1$ . Como en el dibujo de arriba, algunos puntos enteros pueden estar en la línea que divide  $P$ , y por lo menos los dos vértices son entre ellos. Sea  $L$  el número de estos puntos. Se ve que  $I = I_1 + I_2 + L - 2$  y  $B = B_1 + B_2 - 2L + 2$ , y por lo tanto

$$I + \frac{1}{2} B - 1 = (I_1 + I_2 + L - 2) + \frac{1}{2}(B_1 + B_2 - 2L + 2) - 1 = (I_1 + \frac{1}{2} B_1 - 1) + (I_2 + \frac{1}{2} B_2 - 1) = A_1 + A_2 = A.$$

■

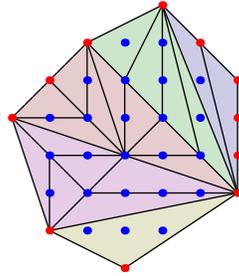
*Demostración del teorema de Pick.* En particular, se puede cortar nuestro polígono  $P$  en triángulos  $P_1, \dots, P_r$ — es suficiente escoger cualquier vértice y considerar las líneas que parten del mismo:



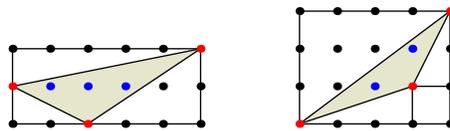
Refinando la triangulación, podemos suponer que cada uno de los lados de estos triángulos son

- o bien paralelos a los ejes,
- o bien no tienen puntos enteros, excepto los dos vértices.

Aquí está tal triangulación para nuestro ejemplo:

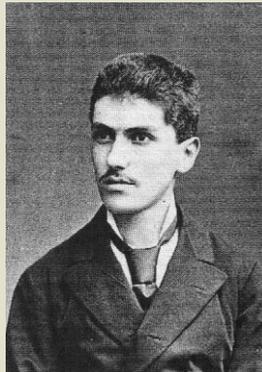


Entonces es suficiente demostrar el teorema para los triángulos que satisfacen la condición de arriba. Pero tales triángulos pueden ser encajados en un rectángulo entero con lados paralelos a los ejes, añadiendo triángulos rectángulos (sin puntos enteros en la hipotenusa) y rectángulos. Por ejemplo,



Por la aditividad de la identidad de Pick, hemos reducido la demostración a los ejercicios [12.2](#) y [12.3](#). ■

**GEORG ALEXANDER PICK** (1859–1942) fue un matemático austríaco de origen judío.



Después de obtener su doctorado en la universidad de Viena, en 1881 tuvo una posición de profesor en la Universidad alemana de Praga.

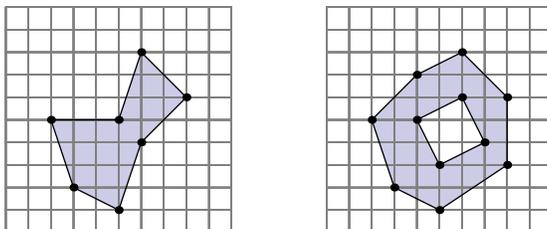
Su famoso teorema apareció en su artículo “Geometrisches zur Zahlenlehre” (1899) y fue popularizado posteriormente por Hugo Steinhaus al incluirlo en la edición de 1969 de su libro “Mathematical Snapshots”.

Pick fue el director del comité que nombró a Albert Einstein profesor de física matemática en 1911. También fue Pick que introdujo a Einstein al trabajo de los matemáticos italianos Gregorio Ricci-Curbastro y Tullio Levi-Civita en geometría diferencial, que ayudó a Einstein en 1915 a formular su teoría general de la relatividad.

Después de su jubilación en 1927 Pick volvió a Viena. En 1938 tuvo que regresar a Praga cuando los nazis invadieron Austria. En 1939 llegaron a Checoslovaquia y Pick fue enviado al campo de concentración Theresienstadt el 13 de julio de 1942, donde murió dos semanas después a los 82 años.

**12.6. Ejercicio.** Supongamos que el polígono es

- a) simplemente conexo (no tiene agujeros), pero no necesariamente convexo,
- b) no simplemente conexo.



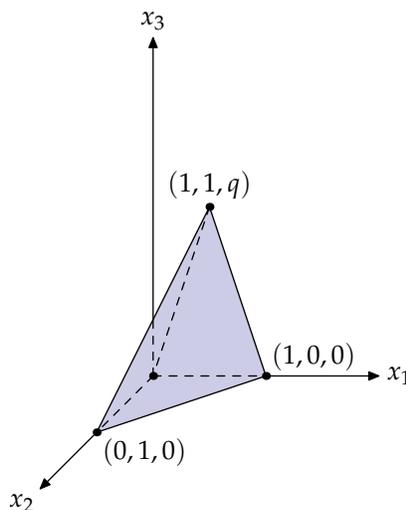
¿Se cumple todavía la identidad de Pick  $A = I + \frac{1}{2}B - 1$ ? Si no, ¿es posible corregirla?

Una pregunta natural es si el teorema de Pick se cumple en dimensiones superiores: sea  $P \subset \mathbb{R}^d$  un poliedro convexo entero, sea  $\text{vol } P := \int_P dx$  su volumen,  $I := \overset{\circ}{P} \cap \mathbb{Z}^d$  el número de puntos en su interior y  $B := \#(P \cap \mathbb{Z}^d) - I$  el número de puntos en su frontera. ¿Existe alguna relación entre  $\text{vol } P$ ,  $I$  y  $B$ , similar a la identidad  $A = I + \frac{1}{2}B - 1$ ? La respuesta es **no**. Aquí está la diferencia clave:

- En  $\mathbb{R}^2$  todo triángulo de vértices enteros que no contiene puntos enteros en su interior necesariamente tiene área  $1/2$  (este hecho es equivalente al teorema de Pick y se usa en algunas demostraciones).
- En  $\mathbb{R}^n$  para  $n > 2$  un símplice de vértices enteros y sin puntos enteros en su interior puede tener volumen arbitrariamente grande.

Aquí está un ejemplo del artículo J.E. Reeve, "On the volume of lattice polyhedra", Proc. London Math. Soc. (3) 7 (1957), 378–395:

**12.7. Ejercicio.** Para  $q = 1, 2, 3, \dots$  sea  $T_q$  el símplice con vértices  $(0, 0, 0)$ ,  $(1, 0, 0)$ ,  $(0, 1, 0)$ ,  $(1, 1, q)$ .



Demuestre que para cada  $q$  tenemos  $\#(T_q \cap \mathbb{Z}^3) = 4$ ; los únicos puntos enteros en  $T_q$  son sus vértices, así que  $I = 0$  y  $B = 4$ . Sin embargo, observe que  $\text{vol } T_q \rightarrow \infty$  para  $q \rightarrow \infty$ .

Entonces en dimensiones superiores no existe un análogo simple del teorema de Pick. La generalización correcta es dada por los **polinomios de Ehrhart**. A saber,  $\text{vol } P$  es un coeficiente de cierto polinomio de grado  $d = \dim P$  que está relacionado con cálculo de puntos enteros en  $P$ .

## 13 Polinomios de Ehrhart

Para obtener la generalización correcta de la identidad  $A = I + \frac{1}{2}B - 1$ , notamos que el número de puntos enteros en un polígono  $P$  puede ser escrito como

$$\#(P \cap \mathbb{Z}^2) = I + B = A + \frac{1}{2}B + 1.$$

De modo similar, para  $m = 1, 2, 3, \dots$  sea  $mP$  el polígono  $P$  dilatado por un factor  $m$ :

$$mP := \{(mx, my) \mid (x, y) \in P\}.$$

Consideremos el número de puntos enteros en  $mP$ :

$$\ell_P(m) := \#(mP \cap \mathbb{Z}^2).$$

El área de  $mP$  es igual a  $Am^2$ , donde  $A$  es el área de  $P$ . El número de puntos enteros en la frontera de  $mP$  es  $Bm$  (véase 12.4). Entonces tenemos

$$\ell_P(m) = Am^2 + \frac{1}{2}Bm + 1.$$

Luego, si contamos el número de puntos en el interior de  $mP$ , tenemos

$$\ell_P(m) := \#(m\overset{\circ}{P} \cap \mathbb{Z}^2) = \ell_P(m) - Bm = Am^2 - \frac{1}{2}Bm + 1 = \ell_P(-m),$$

donde la expresión  $\ell_P(-m)$  significa que el polinomio  $\ell_P(m) = Am^2 + \frac{1}{2}Bm + 1$  se evalúa formalmente en  $-m$ .

En general, tenemos el siguiente

**13.1. Teorema.** Si  $P \subset \mathbb{R}^d$  es un politopo convexo entero (la envolvente convexa de algunos puntos en  $\mathbb{Z}^d$ ) de dimensión  $d$ , entonces el número de puntos enteros

$$\ell_P(m) := \#(mP \cap \mathbb{Z}^d)$$

es un polinomio en  $m$  de grado  $d$  llamado el **polinomio de Ehrhart** de  $P$ .

Los polinomios de Ehrhart son muy especiales: para  $m \in \mathbb{Z}$  tenemos  $\ell_P(m) \in \mathbb{Z}$ . Aquí está una caracterización de tales polinomios:

**13.2. Proposición.** Sea  $p(X) \in \mathbb{Q}[X]$  un polinomio de grado  $d$  con coeficientes en  $\mathbb{Q}$ . Entonces las siguientes condiciones son equivalentes:

- 1)  $p(m) \in \mathbb{Z}$  para cada  $m \in \mathbb{Z}$ .
- 2)  $p(X) = \sum_{0 \leq i \leq d} a_i \binom{X}{i}$  para algunos  $a_0, \dots, a_d \in \mathbb{Z}$ . Aquí  $\binom{X}{i} := \frac{1}{i!} X \cdot (X-1) \cdot (X-2) \cdots (X-i+1)$ .

*Demostración.* Está claro que 2) implica 1) porque  $\binom{m}{i} \in \mathbb{N}$  para cada  $m \in \mathbb{N}$  (por ejemplo por la definición combinatoria de los coeficientes binomiales), y para valores negativos tenemos la identidad

$$\binom{-m}{i} = (-1)^i \binom{m+i-1}{i}.$$

Para ver que 1) implica 2), antes que nada se ve que  $\binom{X}{i}$  forman una base de  $\mathbb{Q}[X]$  sobre  $\mathbb{Q}$ , por lo que cada polinomio con coeficientes en  $\mathbb{Q}$  puede ser escrito como

$$p(X) = \sum_{0 \leq i \leq d} a_i \binom{X}{i} \quad \text{para algunos } a_0, \dots, a_d \in \mathbb{Q}.$$

Tenemos que ver que si  $p$  tiene valores enteros sobre  $\mathbb{Z}$ , entonces  $a_i \in \mathbb{Z}$ . Esto es básicamente la **fórmula de diferencias divididas de Newton** que dice que si  $p$  es un polinomio de grado  $d$ , entonces

$$p(X) = \sum_{0 \leq i \leq d} (\Delta^i p)(0) \binom{X}{i},$$

donde  $\Delta^i p$  es el polinomio (que también tiene valores enteros en  $\mathbb{Z}$ ) definido por

$$\begin{aligned} (\Delta p)(X) &:= p(X+1) - p(X), \\ \Delta^i p &:= \Delta(\Delta^{i-1} p) \quad \text{para } i > 1. \end{aligned}$$

En particular,

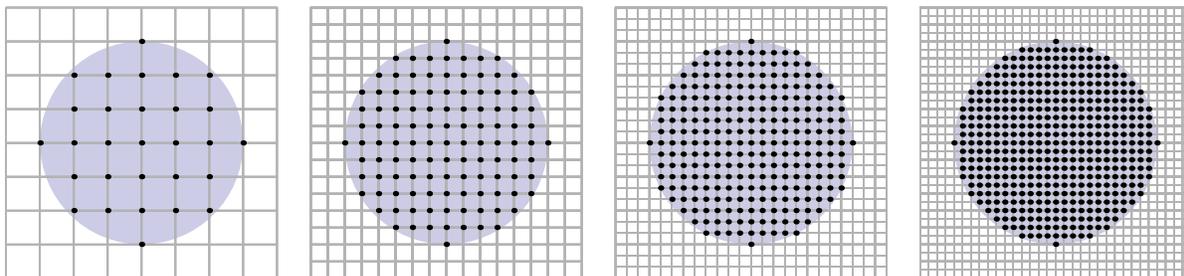
$$\begin{aligned} (\Delta^1 p)(0) &= p(1) - p(0), \\ (\Delta^2 p)(0) &= p(2) - 2p(1) + p(0), \\ (\Delta^3 p)(0) &= p(3) - 3p(2) + 3p(1) - p(0), \\ (\Delta^4 p)(0) &= p(4) - 4p(3) + 6p(2) - 4p(1) + p(0), \\ &\vdots \\ (\Delta^i p)(0) &= p(i) - \binom{i}{1} p(i-1) + \binom{i}{2} p(i-2) + \dots + (-1)^{i-1} p(1) + (-1)^i p(0). \end{aligned}$$

Entonces  $a_i = (\Delta^i p)(0) \in \mathbb{Z}$ . ■

Supongamos que ya sabemos que  $\ell_P(m)$  es un polinomio en  $m$  de grado  $d$ . Tenemos un corolario muy interesante que generaliza al teorema de Pick:

**13.3. Proposición.** Si  $\ell_P(m) = a_d m^d + \dots + a_1 m + a_0$  es el polinomio de Ehrhart de un politopo convexo entero de dimensión  $d$ , entonces  $a_d = \text{vol } P := \int_P dx$ .

*Demostración.* Para calcular aproximadamente el volumen de  $P$ , podemos considerar el subconjunto  $(\frac{1}{m}\mathbb{Z})^d \subset \mathbb{R}^d$  y contar cuántos de sus puntos pertenecen a  $P$ . Es la misma cosa que cortar a  $P$  en pequeños cubitos, cada uno de volumen  $\frac{1}{m^d}$ . Cuando  $m$  tiende a 0, el volumen total de los cubitos en  $P$  tiende al volumen de  $P$ .



En vez de considerar el conjunto  $(\frac{1}{m}\mathbb{Z})^d$ , podemos considerar la intersección del politopo dilatado  $mP$  con  $\mathbb{Z}^d$ :

$$\text{vol } P = \lim_{m \rightarrow \infty} \frac{1}{m^d} \cdot \# \left( P \cap \left( \frac{1}{m}\mathbb{Z} \right)^d \right) = \lim_{m \rightarrow \infty} \frac{1}{m^d} \cdot \# (mP \cap \mathbb{Z}^d) = \lim_{m \rightarrow \infty} \frac{1}{m^d} \cdot \ell_P(m).$$

Finalmente, ya que  $\ell_P(m)$  es un polinomio de grado  $d$ , se sigue que este límite es precisamente el coeficiente de  $m^d$ . ■

(Por supuesto, este argumento elemental funciona porque  $P$  es un subconjunto convexo compacto de  $\mathbb{R}^d$ ; en general la teoría de la medida es más sofisticada que cortes de figuras en cubitos.)

Notamos también que en el polinomio de Ehrhart  $\ell_P(m) = a_d m^d + \dots + a_1 m + a_0$  siempre tenemos  $a_0 = \ell_P(0) = 1$ .

Gracias a las **fórmulas de interpolación**, para calcular el polinomio  $\ell_P(m)$  para un politopo  $P \subset \mathbb{R}^d$  de dimensión  $d$  es suficiente de saber sus  $d + 1$  valores. Por ejemplo, podemos calcular directamente el número de puntos enteros en  $mP \cap \mathbb{Z}^d$  para  $m = 1, \dots, d + 1$ . Por lo menos para pequeños valores de  $d$  esto es un método útil para calcular  $a_d = \text{vol } P$ . O viceversa, si podemos calcular  $\text{vol } P$ , se ahorra un poco de trabajo para calcular  $\ell_P(m)$  porque ya conocemos dos coeficientes  $a_d = \text{vol } P$  y  $a_0 = 1$ .

**13.4. Ejercicio.** Sea  $p(X) = a_d X^d + a_{d-1} X^{d-1} + \dots + a_1 X + a_0 \in \mathbb{Q}[X]$  un polinomio de grado  $d$ . Consideremos  $d + 1$  números  $x_1, x_2, \dots, x_{d+1} \in \mathbb{Q}$ . Observemos que tenemos

$$(*) \quad \begin{pmatrix} p(x_1) \\ p(x_2) \\ \vdots \\ p(x_{d+1}) \end{pmatrix} = \begin{pmatrix} x_1^d & x_1^{d-1} & \dots & x_1 & 1 \\ x_2^d & x_2^{d-1} & \dots & x_2 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_{d+1}^d & x_{d+1}^{d-1} & \dots & x_{d+1} & 1 \end{pmatrix} \cdot \begin{pmatrix} a_d \\ a_{d-1} \\ \vdots \\ a_0 \end{pmatrix}.$$

La matriz con potencias de  $x_1, \dots, x_d$  se llama la **matriz de Vandermonde**. Calcule su determinante

$$\begin{vmatrix} x_1^d & x_1^{d-1} & \dots & x_1 & 1 \\ x_2^d & x_2^{d-1} & \dots & x_2 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_{d+1}^d & x_{d+1}^{d-1} & \dots & x_{d+1} & 1 \end{vmatrix} = \prod_{1 \leq j < k \leq d+1} (x_j - x_k).$$

En consecuencia, si tomamos diferentes  $x_1, x_2, \dots, x_{d+1} \in \mathbb{Q}$ , la matriz es invertible y la ecuación (\*) permite calcular los coeficientes  $a_d, a_{d-1}, \dots, a_1, a_0$  de  $p$  a partir de sus valores  $p(x_1), p(x_2), \dots, p(x_{d+1})$ .

(Note que para polinomios sobre cuerpos finitos, esto no funciona; los coeficientes de un polinomio  $p \in \mathbb{F}_q[X]$  no están determinados por los valores de  $p$ .)

Intuitivamente está claro que el volumen de un politopo  $P$  puede ser calculado contando los puntos enteros en  $mP$  para  $m \rightarrow \infty$ . El resultado notable es que  $\ell_P(m)$  es un polinomio de grado  $d$ . Otra cosa sorprendente es que el polinomio

$$\ell_P(-m) = (-1)^d a_d m^d + (-1)^{d-1} a_{d-1} m^{d-1} + \dots - a_1 m + a_0$$

también tiene significado combinatorio: a saber, calcula los puntos enteros en el interior de  $P$ :

**13.5. Teorema (Reciprocidad de Ehrhart–Macdonald).** Sea  $P \subset \mathbb{R}^d$  un politopo convexo entero de dimensión  $d$ . Sea  $\ell_P(m) := \#(mP \cap \mathbb{Z}^d)$  su polinomio de Ehrhart. Entonces el número de puntos enteros en el interior de  $mP$  está dado por

$$\ell_{\overset{\circ}{P}}(m) := \#(m\overset{\circ}{P} \cap \mathbb{Z}^d) = (-1)^d \ell_P(-m),$$

donde " $\ell_P(-m)$ " es el polinomio  $\ell_P(m)$  evaluado en  $-m$  de modo formal.

Para una demostración de 13.1 y 13.5 con métodos del álgebra homológica y conmutativa, véase el libro “Cohen–Macaulay rings” de Winfried Bruns y Jürgen Herzog; otra demostración, en términos elementales, aparece en “Computing the continuous discretely” de Matthias Beck y Sinai Robins

**13.6. Ejemplo.** Sea  $P$  el octaedro

$$P := \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid |x_1| + |x_2| + |x_3| \leq 1\} = \text{conv}\{(\pm 1, 0, 0), (0, \pm 1, 0), (0, 0, \pm 1)\}.$$

Su polinomio de Ehrhart  $\ell_P(m)$  es de grado 3:

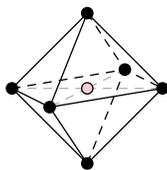
$$\ell_P(m) = a_3 m^3 + a_2 m^2 + a_1 m + a_0.$$

Ya sabemos que  $a_3 = \text{vol } P = \frac{4}{3}$  —el octaedro consiste de dos pirámides, cada una de volumen

$$\frac{1}{3} \times \text{área de la base} \times \text{altura} = \frac{1}{3} \cdot 2 \cdot 1$$

y también sabemos que  $a_0 = 1$ . Los únicos coeficientes que no conocemos son  $a_2$  y  $a_1$ , así que necesitamos solo dos valores de  $\ell_P(m)$ . El octaedro tiene 7 puntos enteros: son sus vértices y el punto  $(0, 0, 0)$ , que está en el interior, de donde

$$\ell_P(1) = 7, \quad \ell_P(-1) = -\ell_{\hat{P}}(1) = -1.$$



Tenemos un sistema de ecuaciones

$$\begin{aligned} \frac{4}{3} + a_2 + a_1 + 1 &= 7, \\ -\frac{4}{3} + a_2 - a_1 + 1 &= -1 \end{aligned}$$

cuya solución es  $a_2 = 2$  y  $a_1 = \frac{8}{3}$ , y entonces

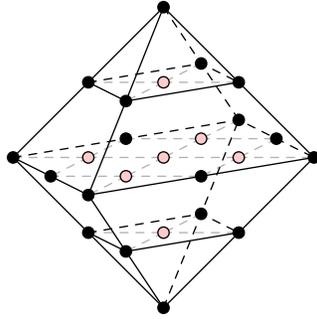
$$\begin{aligned} \ell_P(m) &= \frac{4}{3} m^3 + 2 m^2 + \frac{8}{3} m + 1 = 8 \binom{m}{3} + 12 \binom{m}{2} + 6 \binom{m}{1} + 1, \\ \ell_{\hat{P}}(m) &= \frac{4}{3} m^3 - 2 m^2 + \frac{8}{3} m - 1 = 8 \binom{m}{3} + 4 \binom{m}{2} + 2 \binom{m}{1} - 1. \end{aligned}$$

Usando estas fórmulas, calculamos que por ejemplo  $\ell_P(2) = 25$  y  $\ell_P(-2) = -7$ . Podemos comprobar estos resultados examinando  $2P$ . Es más fácil enumerar los puntos que están en el interior; son 7:

$$2\hat{P} \cap \mathbb{Z}^3 = \{(n_1, n_2, n_3) \in \mathbb{Z}^3 \mid |n_1| + |n_2| + |n_3| < 2\} = \{(\pm 1, 0, 0), (0, \pm 1, 0), (0, 0, \pm 1), (0, 0, 0)\}.$$

Y además en la frontera hay 18 puntos más:

$$\begin{aligned} 2(P \setminus \hat{P}) \cap \mathbb{Z}^3 &= \{(n_1, n_2, n_3) \in \mathbb{Z}^3 \mid |n_1| + |n_2| + |n_3| = 2\} \\ &= \{(\pm 2, 0, 0), (\pm 1, \pm 1, 0), (\pm 1, 0, \pm 1), (0, \pm 2, 0), (0, \pm 1, \pm 1), (0, 0, \pm 2)\}. \end{aligned}$$



Los valores de  $\ell_p(m)$  ([OEIS A001845](#)) son conocidos como **números octaédricos centrados** o **números octaédricos de Haüy**, porque aparecieron en el trabajo del mineralogista francés RENÉ JUST HAÜY (1743–1822), el fundador de la cristalografía; son la misma cosa que los **números de Delannoy**  $D(3, m)$  (HENRI-AUGUSTE DELANNOY (1833–1915), militar francés y matemático aficionado).

$m:$	0	1	2	3	4	5	6	7	8	9	10
$\ell_p(m):$	1	7	25	63	129	231	377	575	833	1159	1561

Por definición,  $D(p, q)$  es el número de caminos desde el punto  $(0, 0)$  al punto  $(p, q)$ , donde cada paso es en la dirección norte  $((i, j) \rightarrow (i, j + 1))$ , este  $((i, j) \rightarrow (i + 1, j))$ , o noreste  $((i, j) \rightarrow (i + 1, j + 1))$ . En particular, de la definición es obvio que

$$D(p, q) = D(q, p).$$

Por ejemplo, para  $(p, q) = (1, 1)$  hay 3 caminos:



y para  $(p, q) = (3, 1)$  hay 7 caminos:





En el triángulo de abajo están marcados los valores  $D(3, m) = \ell_P(m)$ :

				1							
				1	1						
				1	3	1					
			1	5	5	1					
			1	7	13	7	1				
			1	9	25	25	9	1			
			1	11	41	63	41	11	1		
			1	13	61	129	129	61	13	1	
			1	15	85	231	321	231	85	15	1



**13.7. Ejercicio.** Volvamos al tetraedro  $T_q$  de 12.7. Usando interpolación, demuestre que

$$\ell_{T_q}(m) = \frac{q}{6} m^3 + m^2 + \left(2 - \frac{q}{6}\right) m + 1$$

(recuerde que  $a_3 = \text{vol } T_q$  y  $a_0 = 1$ ).

**13.8. Ejercicio.** La fórmula de Pick para  $P \subset \mathbb{R}^2$  puede ser escrita como

$$A = \frac{1}{2}(\ell_P(1) + \ell_{\hat{P}}(1) - 2).$$

Demuestre que en la dimensión  $d = 3$  la fórmula análoga es

$$\text{vol } P = \frac{1}{6}(\ell_P(2) - 3 \ell_P(1) - \ell_{\hat{P}}(1) + 3)$$

(de nuevo, use interpolación).

EUGÈNE EHRHART (1906–2000) fue un matemático francés que trabajó de profesor en varios *lycées* (colegios secundarios franceses) y hacía investigación como pasatiempo. No escribió su tesis de doctorado "Sur un problème de géométrie diophantienne linéaire" sino hasta 1966, instado por sus colegas. Se jubiló en 1971 pero siguió publicando resultados. Incluimos a continuación una página dedicada a Ehrhart: <http://icps.u-strasbg.fr/~clauss/Ehrhart.html>



El resultado de que  $\ell_P(m)$  es un polinomio proviene del artículo "Sur les polyèdres rationnels homothétiques à  $n$  dimensions", C.R. Acad. Sci. Paris 254 (1962), 616–618.

## 14 Polinomios de Ehrhart en SAGE

Para hacer cálculos con puntos enteros en politopos y polinomios de Ehrhart, SAGE utiliza la biblioteca LattE\* ("Lattice point Enumeration", <https://www.math.ucdavis.edu/~latte/>). Para instalarla, hay que ejecutar la comanda `sage -i latte_int`.

```
sage: S = polytopes.simplex(4)
sage: p = S.ehrhart_polynomial()
sage: p
1/24*t^4 + 5/12*t^3 + 35/24*t^2 + 25/12*t + 1
sage: S.affine_hull().volume()
1/24
sage: len((2*S).integral_points())
15
sage: p(2)
15

sage: polytopes.tetrahedron().ehrhart_polynomial()
1/3*t^3 + t^2 + 5/3*t + 1
```

Por ejemplo, podemos calcular el polinomio de Ehrhart del octaedro:

```
sage: P = polytopes.octahedron()
sage: P
A 3-dimensional polyhedron in ZZ^3 defined as the convex hull of 6 vertices
sage: P.vertices()
(A vertex at (-1, 0, 0),
 A vertex at (0, -1, 0),
 A vertex at (0, 0, -1),
 A vertex at (0, 0, 1),
 A vertex at (0, 1, 0),
 A vertex at (1, 0, 0))
sage: P.ehrhart_polynomial()
4/3*t^3 + 2*t^2 + 8/3*t + 1
```

El resultado coincide con nuestro cálculo en [13.6](#).

---

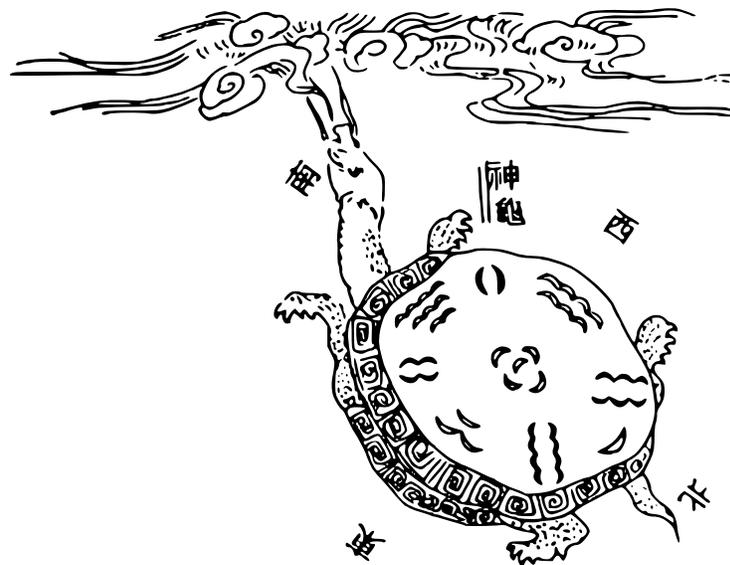
\*¡café con leche!

# Parte III

## Cuadrados mágicos

### 15 Un poco de historia

El primer cuadrado mágico en la historia es el cuadrado **Lo Shu** que fue descubierto por un matemático desconocido chino, probablemente alrededor del siglo I a. C. Es un cuadrado de  $3 \times 3$  que según la leyenda apareció en el caparazón de una tortuga que surgió de un río.



Es el siguiente cuadrado:

4	9	2
3	5	7
8	1	6

Es mágico porque

$$\begin{aligned} 15 &= 4 + 9 + 2 = 3 + 5 + 7 = 8 + 1 + 6 \quad (\text{sumas de filas}) \\ &= 4 + 3 + 8 = 9 + 5 + 1 = 2 + 7 + 6 \quad (\text{sumas de columnas}) \\ &= 4 + 5 + 6 = 8 + 5 + 2 \quad (\text{sumas de diagonales}). \end{aligned}$$

Para mayor información sobre el cuadrado **Lo Shu** y su relación con la filosofía y religión china antigua, pueden leer el artículo

Schuyler Cammann, *The Magic Square of Three in Old Chinese Philosophy and Religion*. History of Religions, Vol. 1, No. 1 (Summer, 1961), pp. 37–80. <http://www.jstor.org/stable/1061970>

El filósofo, alquimista, médico y nigromante alemán **Enrique Cornelio Agripa** (1486–1535) publicó 1531 su libro “De Occulta Philosophia libri III” en el cual los siete planetas conocidos en aquel tiempo (Saturno, Júpiter, Marte, el sol, Venus, Mercurio, la luna) fueron relacionados con ciertos cuadrados mágicos (**LIBER SECVNDVS, p. CXLIX–CLIII**).

4	9	2
3	5	7
8	1	6

4	14	15	1
9	7	6	12
5	11	10	8
16	2	3	13

11	24	7	20	3
4	12	25	8	16
17	5	13	21	9
10	18	1	14	22
23	6	19	2	15

Saturno                      Júpiter                      Marte

El pintor renacentista alemán **Alberto Durero** (1471–1528) hizo en el 1514 el grabado alegórico muy famoso “Melencolía I” donde entre otras cosas se puede observar el politopo convexo conocido como “poliedro de Durero” (recuerden que estamos interesados en politopos) y el siguiente cuadrado mágico:

16	3	2	13
5	10	11	8
9	6	7	12
4	15	14	1

En este cuadrado las sumas de filas y columnas son iguales a 34. También las sumas diagonales son 34 y los cuatro cuadrados de  $2 \times 2$  en los rincones y uno en el centro, a saber 

16	3
5	10

, 

2	13
11	8

, 

9	6
4	15

, 

7	12
14	1

, 

10	11
6	7

 consisten de números con suma 34. Aquí nos va a interesar solamente la condición para sumas de filas y columnas.

Otro cuadrado mágico famoso fue diseñado por Euler, que publicó en 1759 un artículo con el elaborado título “**Solution d’une question curieuse que ne paroît soumise a aucune analyse**” (“Solución de una cuestión curiosa que al parecer no se había sometido a ningún análisis”) dedicado al problema de visitar todo el tablero de ajedrez con un caballo. A saber, el cuadrado de arriba tiene la siguiente propiedad: si uno sigue los números 1, 2, 3, . . . , entonces la sucesión corresponde a movimientos de caballo (verifiquenlo ustedes mismos).

41	59	44	9	40	21	46	7
61	10	41	58	45	8	39	20
12	45	60	55	22	57	6	47
53	62	11	30	25	28	19	38
32	13	54	27	56	23	48	5
63	52	31	24	29	26	37	18
14	33	2	51	16	35	4	49
1	64	15	34	3	50	17	36

Parece que la gente estaba más interesada en cuadrados que son mágicos en un sentido muy fuerte: en los ejemplos de arriba los cuadrados contienen todos los números de 1 a  $n^2$  sin repeticiones y no solamente filas y columnas, sino también ambas diagonales tienen la misma suma —que debe ser igual a  $n(n^2 + 1)/2$  (¿por qué?). Pero no vamos a estudiar tales cuadrados mágicos, porque su número exacto todavía no es conocido:

- Salvo rotaciones y reflexiones, el cuadrado Lo Shu es el único cuadrado de  $3 \times 3$  con esta propiedad (ejercicio).
- El matemático francés Bernard Frénicle de Bessy (alrededor de 1605–1675) descubrió que, salvo rotaciones y reflexiones, hay 880 tales cuadrados  $4 \times 4$  (su libro “Des quarrez ou tables magiques” fue publicado póstumamente en el 1693).
- Richard Schroepel calculó en 1973 que para  $n = 5$  hay 275305224 tales cuadrados.
- Para  $n = 6$  el número exacto no es conocido, pero hay un cálculo estimado que dice que es  $\approx 1,7745 \cdot 10^{19}$ .

Entonces para nosotros un cuadrado mágico va a ser una cosa más simple:

**15.1. Definición.** Un **cuadrado mágico** es una tabla de  $n \times n$  llenada con números naturales arbitrarios (posiblemente 0) tal que las sumas de filas y columnas son iguales al mismo número  $m$ , que se llama la **constante mágica**.

Este objeto es más interesante porque se puede preguntar cuántos cuadrados mágicos de  $n \times n$  de constante mágica  $m$  hay. Por ejemplo, si  $n = 3$  y  $m = 2$ , entonces hay precisamente 21 cuadrados mágicos:

0 0 2	0 0 2	0 0 2	0 1 1	0 1 1	0 1 1	0 1 1
0 2 0	1 1 0	2 0 0	0 1 1	1 0 1	1 1 0	2 0 0
2 0 0	1 1 0	0 2 0	2 0 0	1 1 0	1 0 1	0 1 1
0 2 0	0 2 0	0 2 0	1 0 1	1 0 1	1 0 1	1 0 1
0 0 2	1 0 1	2 0 0	0 1 1	0 2 0	1 0 1	1 1 0
2 0 0	1 0 1	0 0 2	1 1 0	1 0 1	0 2 0	0 1 1
1 1 0	1 1 0	1 1 0	1 1 0	2 0 0	2 0 0	2 0 0
0 0 2	0 1 1	1 0 1	1 1 0	0 0 2	0 1 1	0 2 0
1 1 0	1 0 1	0 1 1	0 0 2	0 2 0	0 1 1	0 0 2



Alberto Durero, Melancholía I (1514)

## 16 Un ejemplo de enumeración: $H_3(m)$

**16.1. Notación.**  $H_n(m)$  es el número de cuadrados mágicos de  $n \times n$  con sumas iguales a  $m$ .

**16.2. Ejemplo.** Aquí están algunos valores de  $H_n(m)$  para valores pequeños de  $n$  y  $m$ , obtenidos con ayuda de una computadora:

$m =$	1	2	3	4	5	6	7
$H_2(m) =$	2	3	4	5	6	7	8
$H_3(m) =$	6	21	55	120	231	406	666
$H_4(m) =$	24	282	2008	10147	40176	132724	381424
$H_5(m) =$	120	6210	153040	2224955	22069251	164176640	976395820
$H_6(m) =$	720	202410	20933840	1047649905	30767936616	602351808741	8575979362560

Por ejemplo, hay 666 cuadrados mágicos de  $3 \times 3$  con sumas iguales a 7 (¡la marca de la Bestia!).

Es fácil de explicar algunos valores. Para  $n = 2$ , una vez está escogido un número  $x \in \{0, 1, \dots, m\}$ , las otras tres entradas del cuadrado están definidas por las condiciones mágicas y se tiene el cuadrado

$x$	$m - x$
$m - x$	$x$

Entonces  $H_2(m) = m + 1$ .

También se ve de la table que  $H_n(1) = n!$  —es porque para  $m = 1$  cada cuadrado mágico debe tener precisamente un coeficiente 1 en cada fila y cada columna, y los otros coeficientes deben ser ceros. Esto se llama **matriz de permutación**: tales matrices actúan sobre  $\mathbb{K}^n$  permutando los vectores de una base, y dan una representación lineal fiel  $S_n \hookrightarrow \text{GL}_n(\mathbb{K})$ . Aquí por ejemplo están los cuadrados mágicos para  $n = 3$  y  $m = 1$ :

0	0	1	0	0	1	0	1	0	0	1	0	1	0	0	1	0	0
0	1	0	1	0	0	0	0	1	1	0	0	0	0	1	0	1	0
1	0	0	0	1	0	1	0	0	0	0	1	0	1	0	0	0	1

▲

El primer cálculo no trivial es  $H_3(m)$ , y ¡no es tan fácil! El matemático británico PERCY ALEXANDER MACMAHON publicó en su libro “Combinatory Analysis” (1915) la fórmula

$$H_3(m) = \binom{m+4}{4} + \binom{m+3}{4} + \binom{m+2}{4},$$

que es un polinomio misterioso de grado 4, que es igual a

$$3 \binom{m}{4} + 9 \binom{m}{3} + 10 \binom{m}{2} + 5m + 1 = \frac{1}{8}m^4 + \frac{3}{4}m^3 + \frac{15}{8}m^2 + \frac{9}{4}m + 1.$$

Hay una demostración bastante sencilla publicada en 1997 por Miklos Bona (<http://www.jstor.org/stable/2691261>) que usa el siguiente razonamiento. Noten que podemos escoger libremente cuatro elementos de la tabla:

$a$	$d$	?
?	$b$	?
?	?	$c$

Y el resto está determinado automáticamente:

$a$	$d$	$m - (a + d)$
$m + c - (a + b + d)$	$b$	$a + d - c$
$b + d - c$	$m - (b + d)$	$c$

Las elecciones de  $a, b, c, d$  no son totalmente arbitrarias: hay que verificar que en la tabla obtenida las entradas no son negativas. Entonces  $M_3(m)$  es el número de las soluciones no negativas del sistema de desigualdades

$$m \geq a + d, \quad m + c \geq a + b + d, \quad a + d \geq c, \quad b + d \geq c, \quad m \geq b + d.$$

Hay tres casos que van a corresponder exactamente a los tres coeficientes binomiales de la fórmula  $H_3(m) = \binom{m+4}{4} + \binom{m+3}{4} + \binom{m+2}{4}$ :

- Supongamos que  $a \leq b$  y  $a \leq c$ . En este caso la desigualdad  $a + d \geq c$ , implica  $b + d \geq c$  y  $m \geq b + d$  implica  $m \geq a + d$ . Luego, tenemos  $m + c \geq b + d + c \geq a + b + d$ . Entonces nuestro sistema de desigualdades se simplifica a

$$a \leq b, \quad a \leq c, \quad a + d \geq c, \quad m \geq b + d.$$

Observamos que esto equivale a la cadena de desigualdades

$$0 \leq a \leq 2a + d - c \leq a + b + d - c \leq b + d \leq m$$

—de hecho, la primera desigualdad es equivalente a  $a + d \geq c$ , la segunda es equivalente a  $a \leq b$  y la tercera es equivalente a  $a \leq c$ .

Una selección de tres números consecutivos  $a, 2a + d - c, a + b + d - c, b + d$ , nos daría valores de  $a, b, c, d$  y el cuadrado mágico correspondiente. Entonces tenemos que escoger 4 elementos con posibles repeticiones entre  $m + 1$  elementos, y en total hay  $\binom{m+4}{4}$  posibilidades.

- Supongamos que  $b < a$  y  $b \geq c$ . Notamos que este caso es simétrico: hemos cambiado  $a$  y  $b$ , y la única diferencia es que el caso  $a = b$  está excluido (porque lo hemos contado antes). Las desigualdades son equivalentes a la cadena de desigualdades

$$0 \leq b \leq 2b + d - c \leq a + b + d - c - 1 \leq a + d - 1 \leq m - 1.$$

Es la misma cadena de desigualdades, solo con  $a$  remplazado por  $b$ , y hemos añadido  $-1$  para no contar el caso  $a = b$ . Entonces estamos contando combinaciones de 4 elementos entre  $m$  con posibles repeticiones, y el resultado es  $\binom{m+3}{4}$ .

- Por fin supongamos que  $c < a$  y  $c < b$ . Las desigualdades  $a + d \geq c$  y  $b + d \geq c$  se vuelven redundantes. Además, la desigualdad  $m + c \geq a + b + d$  implica  $m \geq a + d$  y  $m \geq b + d$ . Entonces nuestro sistema de desigualdades se reduce a

$$c < a, \quad c < b, \quad m + c \geq a + b + d.$$

Se ve que esto es equivalente a la cadena de desigualdades

$$0 \leq c \leq b - 1 \leq b + d - 1 \leq a + b + d - c - 2 \leq m - 2.$$

Contamos el número de combinaciones de 4 elementos entre  $m - 1$  con posibles repeticiones, y la respuesta es  $\binom{m+2}{4}$ .

Los tres casos juntos nos dan

$$M_3(m) = \binom{m+4}{4} + \binom{m+3}{4} + \binom{m+2}{4}.$$

La demostración de arriba explica los coeficientes binomiales de la fórmula, pero es muy tediosa. Si quieren por ejemplo calcular  $M_4(m)$ , será irrazonable decir “bueno, empecemos por nueve parámetros y calculemos las soluciones de las desigualdades correspondientes”... Debe de existir algún modo más listo de comprender  $H_n(m)$ .

## 17 Polítopos de Birkhoff

Notamos que cada cuadrado mágico corresponde a un punto entero del conjunto

$$\left\{ \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nn} \end{pmatrix} \in \mathbb{R}^{n^2} \mid x_{ij} \geq 0, \sum_i x_{ij} = m \text{ para cada } j, \sum_j x_{ij} = m \text{ para cada } i \right\}.$$

Este es cierto politopo convexo dilatado por  $m$ .

**17.1. Definición.** El politopo convexo

$$\mathcal{B}_n := \left\{ \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nn} \end{pmatrix} \in \mathbb{R}^{n^2} \mid x_{ij} \geq 0, \sum_i x_{ij} = 1 \text{ for all } j, \sum_j x_{ij} = 1 \text{ para cada } i \right\}.$$

recibe el nombre de **politopo de Birkhoff** y una matriz que satisface  $\sum_i x_{ij} = 1$  para cada  $j$  y  $\sum_j x_{ij} = 1$  para cada  $i$  se llama **doblemente estocástica**.

Notamos que

$$H_n(m) = \#(m\mathcal{B}_n \cap \mathbb{Z}^{n^2}) = \ell_{\mathcal{B}_n}(m).$$

**17.2. Observación.**  $\dim \mathcal{B}_n = (n-1)^2$ .

*Demostración.* En una matriz doblemente estocástica se pueden escoger elementos  $x_{i1}, \dots, x_{i,n-1}$  en cada fila  $i = 1, \dots, n-1$ , y luego la última columna  $x_{i,n} = 1 - (x_{i1} + \dots + x_{i,n-1})$  está determinada automáticamente. De la misma manera, la última fila está dada por  $x_{n,j} = 1 - (x_{1,j} + \dots + x_{n-1,j})$  para cada  $j = 1, \dots, n$ . ■

$\mathcal{B}_n$  es un politopo convexo porque es una intersección de un número finito de semiespacios afines y es acotado. Recordamos que también hay la descripción dual de polítopos convexos:  $\mathcal{B}_n$  debe ser una combinación convexa de un número finito de puntos. Resulta que estos puntos precisamente corresponden a las matrices de permutación; en particular son puntos enteros, y entonces podemos aplicar la teoría de Ehrhart a  $\mathcal{B}_n$ .

**17.3. Teorema (Birkhoff–von Neumann).**  $\mathcal{B}_n$  es la envolvente convexa de las matrices de permutación:

$$\mathcal{B}_n = \left\{ \sum_{\sigma \in S_n} \lambda_\sigma P_\sigma \mid \lambda_\sigma \geq 0 \text{ y } \sum_{\sigma \in S_n} \lambda_\sigma = 1 \right\},$$

donde  $P_\sigma$  es la matriz de  $n \times n$  que representa a la permutación  $\sigma$ .

**GARRETT BIRKHOFF** (1911–1996) fue un matemático estadounidense. Obtuvo su licenciatura en Harvard en 1932, donde trabajó como profesor por el resto de su carrera. Estuvo entre los primeros algebraistas modernos en los Estados Unidos y publicó varios libros importantes, notablemente “A Survey of Modern Algebra” (1941) y “Algebra” (1967) con Saunders Mac Lane. Sus contribuciones en matemáticas puras están relacionadas con el “álgebra universal”, que ya es una rama bastante exótica del álgebra. Durante la segunda guerra mundial empezó a trabajar en matemáticas aplicadas, radares, balística, fluidodinámica, etc.

El teorema de Birkhoff–von Neumann proviene del artículo

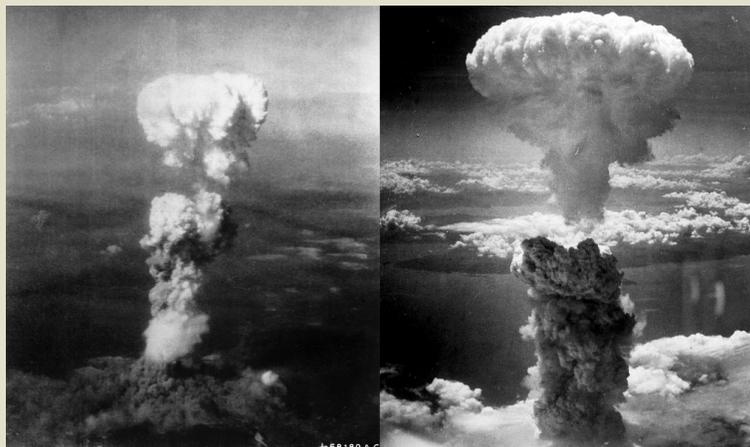
G. Birkhoff, “Tres observaciones sobre el álgebra lineal” (1946), Univ. Nac. Tucumán. Revista A., 5: 147–151, MR 0020547.

(Publicado en español en una revista de la Universidad Nacional de Tucumán en Argentina.)



Garrett Birkhoff

**JOHN VON NEUMANN** (1903–1957) fue un matemático estadounidense de origen judío-húngaro. Empezó su carrera en matemáticas de joven y publicó numerosos artículos. En 1933 fue invitado al *Institute for Advanced Study* en Princeton, donde trabajó por el resto de su vida. Hizo contribuciones en lógica, probabilidad, teoría de la medida, mecánica cuántica, informática, etc. Participó en el Proyecto Manhattan y fue uno de los científicos responsables por los bombardeos atómicos de Hiroshima y Nagasaki.



Nubes de hongo sobre Hiroshima y Nagasaki

**17.4. Ejemplo.** Probablemente el único caso fácil de visualizar es el de  $\mathcal{B}_2$ : es un subconjunto convexo en  $\mathbb{R}^4$ , pero de dimensión 1. Cada matriz doblemente estocástica de  $2 \times 2$  es de la forma

$$\begin{pmatrix} \lambda & 1-\lambda \\ 1-\lambda & \lambda \end{pmatrix} = \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + (1-\lambda) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{para } \lambda \in [0,1],$$

y en consecuencia, el conjunto  $\mathcal{B}_2$  es el segmento de recta entre  $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$  y  $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ . ▲

Una dirección del teorema de Birkhoff–von Neumann es fácil: cada matriz de permutación es doblemente estocástica y tenemos la siguiente

**17.5. Observación.** Si  $A_1, \dots, A_r$  son matrices doblemente estocásticas, entonces la combinación convexa

$$\lambda_1 A_1 + \dots + \lambda_r A_r \quad \text{donde } \lambda_i \geq 0 \text{ y } \sum_{1 \leq i \leq r} \lambda_i = 1$$

es también una matriz doblemente estocástica.

**17.6. Ejemplo.** En particular, la suma de dos cuadrados mágicos es también un cuadrado mágico. ▲

La parte más difícil de ver es que cada matriz  $A \in \mathcal{B}_n$  puede ser escrita como una combinación convexa  $\sum_{\sigma \in S_n} \lambda_\sigma P_\sigma$  para algunos coeficientes  $\lambda_\sigma \geq 0$ ,  $\sum_{\sigma} \lambda_\sigma = 1$ . El punto clave es el siguiente

**17.7. Lema.** Sea  $A = [x_{ij}]$  una matriz doblemente estocástica. Entonces existe una matriz de permutación  $P = [p_{ij}]$  tal que  $p_{ij} = 1 \Rightarrow x_{ij} \neq 0$ .

Primero veamos por qué este lema es suficiente para demostrar el teorema de Birkhoff–von Neumann. Si tenemos tal matriz de permutación, consideramos

$$\lambda := \min_{1 \leq i, j \leq n} \{x_{ij} \mid p_{ij} \neq 0\}.$$

Notamos que este es un número positivo. Sean  $p, q$  índices tales que  $\lambda = x_{pq}$ . Si  $\lambda = 1$ , entonces  $A$  es una matriz de permutación y la afirmación del teorema es trivial. Si  $\lambda \neq 1$ , consideramos la matriz doblemente estocástica

$$A' := \frac{1}{1-\lambda} (A - \lambda P).$$

Tenemos la combinación convexa

$$A = (1-\lambda) A' + \lambda P,$$

donde  $P$  es una matriz de permutación y  $A'$  es una matriz doblemente estocástica y observamos que  $x'_{pq} = 0$  porque  $x_{pq} = \lambda$ . Luego,  $A'$  tiene un coeficiente cero más que  $A$ . Esto significa que podemos aplicar el mismo procedimiento a  $A'$  y el proceso se termina eventualmente porque en cada paso la matriz  $A'$  va a tener más ceros. ■

**17.8. Ejemplo.** Consideremos la matriz  $\begin{pmatrix} 1/5 & 4/5 & 0 \\ 2/5 & 1/5 & 2/5 \\ 2/5 & 0 & 3/5 \end{pmatrix}$ . Según la demostración de arriba, para es-

cribirla como una combinación convexa de matrices de permutación, tenemos que escoger una matriz de permutación cuyos coeficientes no ceros corresponden a coeficientes no nulos en nuestra matriz. El lema (que todavía no hemos demostrado) nos dice que esto es siempre posible. Tomemos por ejemplo

$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ . Tenemos la combinación convexa

$$\begin{pmatrix} 1/5 & 4/5 & 0 \\ 2/5 & 1/5 & 2/5 \\ 2/5 & 0 & 3/5 \end{pmatrix} = \frac{4}{5} \begin{pmatrix} 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 \end{pmatrix} + \frac{1}{5} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

De la misma manera, para la primera matriz de la suma se obtiene una descomposición convexa

$$\begin{pmatrix} 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Por tanto,

$$\begin{pmatrix} 1/5 & 4/5 & 0 \\ 2/5 & 1/5 & 2/5 \\ 2/5 & 0 & 3/5 \end{pmatrix} = \frac{2}{5} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} + \frac{2}{5} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \frac{1}{5} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

▲

**17.9. Ejemplo.** De lo anterior se desprende que cada cuadrado mágico es una suma de matrices de permutación. Por ejemplo,

$$\begin{array}{|c|c|c|} \hline 1 & 4 & 0 \\ \hline 2 & 1 & 2 \\ \hline 2 & 0 & 3 \\ \hline \end{array} = 2 \cdot \begin{array}{|c|c|c|} \hline 0 & 1 & 0 \\ \hline 0 & 0 & 1 \\ \hline 1 & 0 & 0 \\ \hline \end{array} + 2 \cdot \begin{array}{|c|c|c|} \hline 0 & 1 & 0 \\ \hline 1 & 0 & 0 \\ \hline 0 & 0 & 1 \\ \hline \end{array} + \begin{array}{|c|c|c|} \hline 1 & 0 & 0 \\ \hline 0 & 1 & 0 \\ \hline 0 & 0 & 1 \\ \hline \end{array}$$

Esta expresión no es única. Por ejemplo, el cuadrado Lo Shu tiene dos descomposiciones diferentes:

$$\begin{array}{|c|c|c|} \hline 4 & 9 & 2 \\ \hline 3 & 5 & 7 \\ \hline 8 & 1 & 6 \\ \hline \end{array} = \begin{array}{|c|c|c|} \hline 0 & 0 & 1 \\ \hline 0 & 1 & 0 \\ \hline 1 & 0 & 0 \\ \hline \end{array} + \begin{array}{|c|c|c|} \hline 0 & 0 & 1 \\ \hline 1 & 0 & 0 \\ \hline 0 & 1 & 0 \\ \hline \end{array} + 7 \cdot \begin{array}{|c|c|c|} \hline 0 & 1 & 0 \\ \hline 0 & 0 & 1 \\ \hline 1 & 0 & 0 \\ \hline \end{array} + 2 \cdot \begin{array}{|c|c|c|} \hline 0 & 1 & 0 \\ \hline 1 & 0 & 0 \\ \hline 0 & 0 & 1 \\ \hline \end{array} + 4 \cdot \begin{array}{|c|c|c|} \hline 1 & 0 & 0 \\ \hline 0 & 1 & 0 \\ \hline 0 & 0 & 1 \\ \hline \end{array}$$

$$= 2 \cdot \begin{array}{|c|c|c|} \hline 0 & 0 & 1 \\ \hline 0 & 1 & 0 \\ \hline 1 & 0 & 0 \\ \hline \end{array} + 6 \cdot \begin{array}{|c|c|c|} \hline 0 & 1 & 0 \\ \hline 0 & 0 & 1 \\ \hline 1 & 0 & 0 \\ \hline \end{array} + 3 \cdot \begin{array}{|c|c|c|} \hline 0 & 1 & 0 \\ \hline 1 & 0 & 0 \\ \hline 0 & 0 & 1 \\ \hline \end{array} + \begin{array}{|c|c|c|} \hline 1 & 0 & 0 \\ \hline 0 & 0 & 1 \\ \hline 0 & 1 & 0 \\ \hline \end{array} + 3 \cdot \begin{array}{|c|c|c|} \hline 1 & 0 & 0 \\ \hline 0 & 1 & 0 \\ \hline 0 & 0 & 1 \\ \hline \end{array}$$

▲

**17.10. Ejemplo.** Otra aplicación del teorema de Birkhoff–von Neumann: *el producto de dos matrices doblemente estocásticas es también una matriz doblemente estocástica*. Es una consecuencia fácil del hecho de que el producto de dos matrices de permutación  $P_\sigma \cdot P_\tau$  es también una matriz de permutación  $P_{\sigma \cdot \tau}$ .

En particular, el producto de dos cuadrados mágicos (calculado como el producto de matrices) es también mágico. La constante mágica correspondiente es el producto de las constantes mágicas. Por ejemplo,

$$\begin{array}{|c|c|c|} \hline 1 & 4 & 0 \\ \hline 3 & 1 & 1 \\ \hline 1 & 0 & 4 \\ \hline \end{array} \cdot \begin{array}{|c|c|c|} \hline 0 & 1 & 2 \\ \hline 1 & 2 & 0 \\ \hline 2 & 0 & 1 \\ \hline \end{array} = \begin{array}{|c|c|c|} \hline 4 & 9 & 2 \\ \hline 3 & 5 & 7 \\ \hline 8 & 1 & 6 \\ \hline \end{array}$$

▲

## 18 Teorema de los matrimonios de Hall

Volvamos ahora al lema sobre matrices de permutación. Hemos afirmado que para cada matriz doblemente estocástica  $A = [x_{ij}]$  existe una matriz de permutación  $P = [p_{ij}]$  tal que  $p_{ij} = 1 \Rightarrow x_{ij} \neq 0$ . Ahora vamos a demostrarlo usando un famoso resultado combinatorio.

Recordemos que un **grafo bipartito**  $G = (V, E)$  es un grafo tal que su conjunto de vértices es una unión disjunta  $V = X \sqcup Y$  tal que todas las aristas son de la forma  $(x, y)$  donde  $x \in X$  e  $y \in Y$ .

Para un subconjunto  $W \subset V$  su **entorno**  $N_G(W)$  es el conjunto de los vértices adyacentes a los vértices de  $W$ :

$$N_G(W) := \{v \in V \mid (u, v) \in E \text{ para algún } u \in W\}.$$

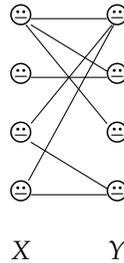
**18.1. Teorema (Teorema de Hall del matrimonio).** Sea  $G = (X \sqcup Y, E)$  un grafo bipartito con  $|X| = |Y|$ . Supongamos que

$$|N_G(W)| \geq |W| \quad \text{para cada } W \subset X.$$

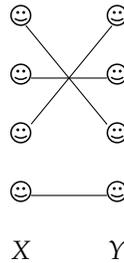
Entonces existe un **emparejamiento perfecto**, es decir una biyección  $\phi: X \rightarrow Y$  tal que  $(x, \phi(x)) \in E$  para cada  $x \in X$ .

Antes que nada, notamos que la condición  $|N_G(W)| \geq |W|$  es necesaria para tener un emparejamiento perfecto.

**18.2. Ejemplo.** Consideramos dos grupos de personas  $X$  e  $Y$ . Ponemos una arista entre  $x \in X$  e  $y \in Y$  si  $x$  e  $y$  se gustan y quieren casarse.



El teorema de Hall dice que si a cada grupo de personas  $W \subset X$  le gustan por lo menos  $|W|$  personas del grupo  $Y$ , entonces todos del grupo  $X$  pueden casarse. En nuestro caso tenemos la siguiente solución:



*Demostración del teorema de los matrimonios de Hall.* Por inducción sobre  $n = |X| = |Y|$ . Para  $n = 1$  la afirmación es trivial.

Supongamos primero que para cada  $W \subset X$  tenemos  $|N_G(W)| \geq |W| + 1$ . Entonces podemos simplemente escoger un vértice  $x \in X$ . Por nuestra hipótesis,  $N_G(\{x\}) \neq \emptyset$ , y entonces podemos escoger  $y \in N_G(\{x\})$ . Si quitamos los vértices  $x$  y  $y$ , el grafo restante  $G'$  satisface la condición  $|N_{G'}(W)| \geq |W|$ , y así inductivamente se obtiene un emparejamiento perfecto.

Ahora supongamos que existe algún  $W \subset X$  tal que  $|N_G(W)| = |W|$ . Consideramos los siguientes grafos bipartitos más pequeños:

$$H := W \sqcup N_G(W) \quad \text{y} \quad K := (X \setminus W) \sqcup (Y \setminus N_G(W)).$$

Noten que para cada  $Z \subset W$  tenemos  $N_H(Z) = N_G(Z)$ , de donde el grafo  $H$  satisface la condición  $|N_H(Z)| \geq |Z|$ , y por la hipótesis de inducción tiene un emparejamiento perfecto.

Para el grafo  $K$ , sea  $Z \subset X \setminus W$ . Tenemos

$$N_G(Z \sqcup W) = N_K(Z) \sqcup N_G(W),$$

y por lo tanto

$$|Z| + |W| = |Z \sqcup W| \leq |N_G(Z \sqcup W)| = |N_K(Z)| + |N_G(W)| \geq |N_K(Z)| + |W|.$$

Entonces tenemos  $|N_K(Z)| \geq |Z|$ , y por la hipótesis de inducción existe un emparejamiento perfecto para  $K$ . ■

En nuestro caso, para una matriz doblemente estocástica  $A = [x_{ij}]$  consideramos el grafo bipartito cuyos vértices representan las filas y columnas de la matriz:

$$V := R \sqcup C := \{r_1, \dots, r_n\} \sqcup \{c_1, \dots, c_n\}.$$

Las aristas corresponden a coeficientes no nulos en la matriz:

$$E := \{(r_i, c_j) \mid x_{ij} \neq 0\}.$$

Este grafo satisface la condición  $|N(I)| \geq |I|$  para cada  $I \subset R$ , es decir, si escogemos  $\ell$  filas  $I = \{r_{i_1}, \dots, r_{i_\ell}\}$ , entonces hay por lo menos  $\ell$  columnas diferentes con coeficientes no nulos en las intersecciones con  $I$ . Notamos que

$$\sum_{\substack{i \in I \\ j \in N(I)}} x_{ij} = \sum_{i \in I} \sum_{j \in N(\{i\})} x_{ij} = \sum_{i \in I} \underbrace{\sum_{1 \leq j \leq n} x_{ij}}_{=1} = |I|,$$

porque la matriz es doblemente estocástica. De la misma manera, para cada  $J \subset C$  tenemos

$$|J| = \sum_{\substack{j \in J \\ i \in N(J)}} x_{ij}$$

Para  $I \subset R$  notamos que  $I \subseteq N(N(I))$ , y por lo tanto

$$|N(I)| = \sum_{\substack{j \in N(I) \\ i \in N(N(I))}} x_{ij} \geq \sum_{j \in N(I)} x_{ij} = |I|.$$

Entonces por el teorema de los matrimonios de Hall, existe un emparejamiento perfecto  $M \subset E$ . Consideramos la matriz  $P = [p_{ij}]$  definida por

$$p_{ij} := \begin{cases} 1, & (r_i, c_j) \in M, \\ 0, & (r_i, c_j) \notin M. \end{cases}$$

Observamos que  $M$  es un emparejamiento perfecto, de donde cada fila y cada columna de  $P$  tiene exactamente un coeficiente 1; es decir,  $P$  es una matriz de permutación. Si  $x_{ij} = 0$ , entonces por la definición del grafo,  $(r_i, c_j) \notin E$  y entonces  $p_{ij} = 0$ , como deseamos. ■

El teorema lleva el nombre del matemático inglés **PHILIP HALL** (1904–1982), conocido por sus contribuciones a la teoría de grupos finitos. El teorema de Hall aparece en su artículo de 1935 “**On Representatives of Subsets**”. El nombre “teorema de los matrimonios” aparece más tarde, por ejemplo en

Paul R. Halmos and Herbert E. Vaughan, *The Marriage Problem*, American Journal of Mathematics, Vol. 72, No. 1 (Jan., 1950), pp. 214-215, <http://www.jstor.org/stable/2372148>

## 19 Teoría de Ehrhart para los politopos de Birkhoff

Hemos demostrado que  $\mathcal{B}_n$  es un politopo convexo entero de dimensión  $(n-1)^2$ , y entonces el teorema de Ehrhart implica

**19.1. Teorema (Ehrhart, Stanley).**  $H_n(m)$  es un polinomio en  $m$  de grado  $(n-1)^2$ .

**19.2. Ejemplo.** Aquí están los primeros polinomios  $H_n(m)$  (los escribimos en la base de coeficientes binomiales  $\binom{m}{i}$  para enfatizar que  $H_n(m) \in \mathbb{N}$  para  $m \in \mathbb{N}$ ):

$$H_0(m) = 1,$$

$$H_1(m) = 1,$$

$$H_2(m) = m + 1,$$

$$H_3(m) = \binom{m+4}{4} + \binom{m+3}{4} + \binom{m+2}{4} = 3 \binom{m}{4} + 9 \binom{m}{3} + 10 \binom{m}{2} + 5m + 1,$$

$$H_4(m) = 352 \binom{m}{9} + 2112 \binom{m}{8} + 5432 \binom{m}{7} + 7800 \binom{m}{6} + 6820 \binom{m}{5} + 3712 \binom{m}{4} + 1233 \binom{m}{3} + 235 \binom{m}{2} + 23m + 1,$$

$$H_5(m) = 4718075 \binom{m}{16} + 47180750 \binom{m}{15} + 215622400 \binom{m}{14} + 595950225 \binom{m}{13} + 1110779300 \binom{m}{12} + 1473782200 \binom{m}{11} + 1431876375 \binom{m}{10} + 1031899475 \binom{m}{9} + 552509140 \binom{m}{8} + 217967015 \binom{m}{7} + 62167090 \binom{m}{6} + 12413375 \binom{m}{5} + 1649576 \binom{m}{4} + 134769 \binom{m}{3} + 5971 \binom{m}{2} + 119m + 1.$$

▲

Recordemos que también tenemos la fórmula de reciprocidad

$$\ell_{\hat{\mathcal{B}}_n}(m) = (-1)^{(n-1)^2} \ell_{\mathcal{B}_n}(-m).$$

¿Cuál es su significado en nuestro caso? La función  $\ell_{\hat{\mathcal{B}}_n}(m)$  cuenta los cuadrados mágicos de  $n \times n$  de suma mágica  $m$  y tal que todas las entradas son estrictamente positivas. En particular,  $\ell_{\hat{\mathcal{B}}_n}(m) = 0$  para  $m = 1, 2, \dots, n-1$ . Si  $m \geq n$ , notamos que contar tales mágicos es la misma cosa que contar los cuadrados mágicos con suma  $m-n$  y luego añadir 1 a cada entrada. Entonces la reciprocidad nos dice que

**19.3. Teorema.** 1)  $H_n(-1) = H_n(-2) = \dots = H_n(-n+1) = 0$ .

2)  $H_n(m-n) = (-1)^{(n-1)^2} H_n(-m)$  para  $m \geq n$ .

**19.4. Ejemplo.** Para valores pequeños de  $n$  esto reduce el trabajo con interpolación polinomial para calcular  $H_n(m)$ . Por ejemplo, para  $n = 3$  sabemos que

$$H_3(-3) = H_3(0) = 1, \quad H_3(-2) = 0, \quad H_3(-1) = 0, \quad H_3(0) = 1, \quad H_3(1) = 3! = 6,$$

y estos cinco valores definen de modo único el polinomio de grado 4

$$H_3(m) = \frac{1}{8} m^4 + \frac{3}{4} m^3 + \frac{15}{8} m^2 + \frac{9}{4} m + 1$$

que coincide con el que hemos visto arriba. Para  $n = 4$  el polinomio  $H_4(m)$  es de grado 9, así que debemos obtener 10 de sus valores. Tenemos

$$\begin{aligned}
 H_4(-7) &= -H_4(3) = -2008, \\
 H_4(-6) &= -H_4(2) = -282, \\
 H_4(-5) &= -H_4(1) = -24, \\
 H_4(-4) &= -H_4(0) = -1, \\
 H_4(-3) &= 0, \\
 H_4(-2) &= 0, \\
 H_4(-1) &= 0, \\
 H_4(0) &= 1, \\
 H_4(1) &= 24, \\
 H_4(2) &= 282, \\
 H_4(3) &= 2008.
 \end{aligned}$$

Los únicos valores no obvios son  $H_4(2) = 282$  y  $H_4(3) = 2008$ , y para encontrarlos pueden ejecutar un pequeño programa—¡escribanlo ustedes mismos! Luego, la interpolación polinomial nos da

$$\begin{aligned}
 H_4(m) &= \frac{11}{11340} m^9 + \frac{11}{630} m^8 + \frac{19}{135} m^7 + \frac{2}{3} m^6 + \frac{1109}{540} m^5 + \frac{43}{10} m^4 + \frac{35117}{5670} m^3 + \frac{379}{63} m^2 + \frac{65}{18} m + 1 \\
 &= 352 \binom{m}{9} + 2112 \binom{m}{8} + 5432 \binom{m}{7} + 7800 \binom{m}{6} + 6820 \binom{m}{5} + 3712 \binom{m}{4} + 1233 \binom{m}{3} + 235 \binom{m}{2} + 23m + 1,
 \end{aligned}$$

y solo evaluando este polinomio podemos obtener el maravilloso hecho que hay 5045326 cuadrados mágicos de  $4 \times 4$  con suma 10 y 1159382525319251 cuadrados mágicos con suma 100. ▲

## 20 Polítopos de Birkhoff en SAGE

SAGE puede construir los polítopos de Birkhoff y calcular sus polinomios de Ehrhart:

```

sage: polytopes.Birkhoff_polytope(3)
A 4-dimensional polyhedron in ZZ^9 defined as the convex hull of 6 vertices
sage: polytopes.Birkhoff_polytope(3).ehrhart_polynomial()
1/8*t^4 + 3/4*t^3 + 15/8*t^2 + 9/4*t + 1

```

Es muy difícil hacer cálculos con  $\mathcal{B}_n$ : su dimensión  $(n - 1)^2$  y número de vértices  $n!$  crecen muy rápido. Por ejemplo, Matthias Beck y Dennis Pixton calcularon el volumen de  $\mathcal{B}_{10}$  usando una red de computadoras: el cálculo empezó en el marzo 2002 y terminó en el mayo 2003. Una computadora normal con una CPU de 1 GHz tomaría casi 17 años para ejecutar el mismo programa. Y se trata solo del coeficiente mayor del polinomio de Ehrhart  $H_{10}$ . Beck y Pixton calcularon  $H_n$  hasta  $n = 9$ ; para mayor información, véase <http://www.math.binghamton.edu/dennis/Birkhoff/>. Con mi computadora portátil y SAGE tuve paciencia solo hasta  $n = 5$ .