Universidad de Valladolid

Facultad de Ciencias

Trabajo Fin de Máster

Máster en Investigación en Matemáticas

El método polinómico en combinatoria y aritmética

Autor: ***David López Soria***

Tutor: ***Antonio Campillo López***

Índice general

Introducción

En el campo de la combinatoria, a menudo aparecen problemas que tratan sobre las propiedades que puede tener un cierto conjunto finito. Puede ser interesante estudiar un conjunto finito de objetos geométricos, sujeto a alguna restricción geométrica de algún tipo. Por ejemplo, fijado un subconjunto finito de un espacio euclídeo, se pueden plantear problemas que se centren en el estudio de las distancias entre los puntos de dicho subconjunto. Desde un punto de vista más aritmético se pueden atacar problemas relacionados con el estudio de un cierto conjunto finito de un cuerpo o anillo, sujeto a alguna restricción que tenga que ver con la estructura algebraica del espacio.

En cualquier caso, vamos a ver en esta memoria como se pueden atacar problemas de tipo combinatorio de manera sistemática, empleando polinomios. Así como en álgebra y geometría algebraica, el empleo de polinomios surge de manera más o menos natural, en combinatoria su utilidad no parece evidente. A primera vista, siempre puede parecer más adaptado emplear un razonamiento combinatorio, cuando uno trata con un problema puramente combinatorio. Desde luego, algunos de los resultados que presentamos en este trabajo, admiten una demostración combinatoria, aunque iremos viendo a lo largo del trabajo por qué son interesantes estas demostraciones polinómicas.

En muchos casos, los métodos polinómicos que veremos dan demostraciones cortas y elementales a problemas complejos de la combinatoria, pero ahí no acaba su utilidad. Los métodos polinómicos son versátiles, ya que en muchas ocasiones, el mismo razonamiento puede ser empleado en diferentes ámbitos, a condición de alterar ligeramente los polinomios con los que se trabaja. En algunos casos, estos procedimientos resuelven problemas para los que no se conoce otra demostración. Por ejemplo, podemos destacar entre estos problemas a la conjetura finita de Kakeya o el problema de las distancias distintas de Erdős (En [GK] encontramos una fantástica referencia al respecto), que solo han podido atacarse de forma efectiva empleando métodos polinómicos. Es claro que el conocimiento y dominio de estos razonamientos

puede resultar muy interesante. En [Tao], Terence Tao realiza un compendio de algunos de estos métodos, que tomaremos como punto de partida para la redacción de este trabajo.

De forma general, la idea de estos métodos se basa en encontrar de alguna forma un polinomio P de grado controlado, que nos permita entender la naturaleza combinatoria del problema al estudiar su conjunto de ceros $V(\mathrm{P})$, gracias a razonamientos geométricos. En muchos casos, los argumentos pasaran por garantizar la existencia de un cierto polinomio interpolador, con las propiedades adecuadas para atacar el problema combinatorio que se plantea.

Aunque algunos de estos procedimientos ya eran conocidos (por ejemplo el método de Stepanov en [Ste]), la extensión y generalización del empleo de métodos polinómicos en combinatoria es reciente.

En el primer capitulo, introduciremos las técnicas y razonamientos fundamentales a la hora de trabajar con polinomios. Por ejemplo, abordaremos resultados de interpolación o de ordenes de anulación de polinomios, que nos permitirán trabajar. Veremos como en algunos casos se deduce de la naturaleza combinatoria del problema, la existencia de un polinomio que se anula en una cierta cantidad de puntos, con una multiplicidad conocida. El objetivo de este primer capítulo, es presentar la prueba de Dvir de la conjetura finita de Kakeya, que trataremos con todo detalle (introduciendo también el problema general en $\mathbb{R}^n$).

El segundo capítulo está dedicado al estudio del llamado Teorema de los ceros combinatorio. Para la prueba de este teorema, nos basaremos en [Alon]. Se trata de un teorema combinatorio, basado en el teorema de los ceros de Hilbert. Veremos de forma clara, como los razonamientos de la geometría algebraica pueden inspirar resultados en combinatoria. Comprobaremos como se puede emplear este teorema para dar nuevas pruebas para algunos teorema clásicos como el de Chevalley y Chevalley-Warning. Además, veremos algunos aplicaciones de los métodos polinómicos en teoría aditiva de números.

A continuación, estudiaremos el llamado Método de Stepanov. Se trata de un procedimiento polinómico que nos va a permitir comprender la estructura combinatoria de las curvas definidas sobre cuerpos finito. Nos centraremos en la prueba de la cota de Hasse para una familia de curvas planas. Es un resultado clásico de teoria de números, que ha inspirado muchos de los otros razonamientos presentados en este trabajo.

Finalmente, en el último capítulo del trabajo estudiaremos las aplicaciones de los métodos polinómicos en $\mathbb{R}^n$. Es especialmente interesante ver como se puede explotar la estructura topológica de $\mathbb{R}^n$, en particular el teorema de Borsuk-Ulam, para complementar los razonamientos polinómicos. De este teorema topológico, se puede deducir el llamado teorema del sándwich de jamón, y su versión polinómica. Gracias a estos resultados, podremos ver como se puede dar una descomposición polinómica del espacio, debida a Guth y Katz ([GK]), que permite atacar problemas combinatorios en $\mathbb{R}^n$. Veremos una prueba del teorema de Szemerédi-Trotter, que servirá para ilustrar como se aplican los razonamientos desarrollados a lo largo del capítulo.

A medida que vaya avanzando el trabajo, se irán precisando las nociones, resultados y bibliografía empleada.

Capítulo 1

Principios generales, la conjetura finita de Kakeya

1 Factorización e interpolación

Comencemos introduciendo los objetos y resultados básicos que vamos a estudiar a lo largo del trabajo. Nos apoyaremos en [Tao] y [DKSS] para introducir estos primeros resultados.

Durante todo el trabajo vamos a estar razonando con cuerpos, ya sean finitos o infinitos. En lo que sigue, cuando nos refiramos a K este será siempre un cuerpo. Desde luego, cuando sea necesario se precisará si es finito o no, así como su característica cuando se requiera.

Definición 1.1. Sea K un cuerpo. Dado un subconjunto S de $\mathrm{K}[\mathrm{X}_1, ..., \mathrm{X}_n]$, denotamos con $V(S)$ al conjunto formado por los puntos de K^n donde se anulan todos los polinomios de S.

Es sencillo comprobar que el conjunto $V(S)$ coincide con el conjunto formado por los puntos donde se anulan todos los polinomios del ideal de $\mathrm{K}[\mathrm{X}_1, ..., \mathrm{X}_n]$ generado por los elementos de S. Del Teorema de la base de Hilbert se deduce que los conjuntos $V(S)$ corresponden al lugar donde se anula una cantidad finita de polinomios.

Los conjuntos $V(S)$ son exactamente los conjuntos cerrados para una topología en K^n, que recibe el nombre de topología de Zariski. De esta forma los conjuntos $V(S)$ reciben el nombre de cerrados algebraicos de K^n. Cuando estos cerrados sean irreducibles (y el cuerpo K algebraicamente cerrado) hablaremos de variedades y cuando se trate de cerrados dados por un único polinomio

hablaremos de hipersuperficies.

La estructura de las hipersuperficies de K es conocida. Lo vemos con la siguiente proposición.

Proposición 1.2. Sea K un cuerpo. Se tiene:

- Si $\mathrm{P} \in \mathrm{K}[\mathrm{X}]$ es un polinomio no nulo de grado menor o igual que $d \geq 0$, entonces el conjunto de $V(\mathrm{P})$ tiene a lo sumo d puntos.
- Si S es un subconjunto de K con cardinal $d \geq 0$, entonces existe un polinomio no trivial $\mathrm{P} \in \mathrm{K}[\mathrm{X}]$ de grado d y tal que $S \subset V(\mathrm{P})$.

Demostración. Si el conjunto $V(\mathrm{P})$ contiene un punto $z \in \mathrm{K}$, entonces $d \geq 1$ y el polinomio P se factoriza como $\mathrm{P}(\mathrm{X}) = (\mathrm{X} - z)\mathrm{Q}(\mathrm{X})$ con Q un polinomio de grado menor o igual que $d-1$. Procediendo por inducción sobre el grado, observamos que necesariamente $V(\mathrm{P})$ tiene cardinal menor o igual que d.

Recíprocamente, si $S \neq \emptyset$ tiene cardinal menor o igual que d, simplemente consideramos el polinomio dado por el producto de los $(\mathrm{X} - z)$ donde $z \in S$. Se construye de esta forma un polinomio no nulo que cumple que $S \subset V(\mathrm{P})$. □

Dado un conjunto S, a lo largo de todo el trabajo, denotaremos con $|S|$ al cardinal de dicho conjunto. Gracias a lo que ocurre en dimensión 1, se pueden deducir propiedades en dimensión mayor. Para el caso en que K es un cuerpo finito, se tiene el siguiente resultado:

Proposición 1.3 (Schwartz-Zippel)**.** Sean K un cuerpo finito y P un polinomio de $\mathrm{K}[\mathrm{X}_1, ..., \mathrm{X}_n]$, para $n \geq 1$. Si P no es el polinomio nulo y tiene grado menor o igual que d, entonces:

$$|V(\mathrm{P})| \leq d\,|\mathrm{K}|^{n-1}\,.$$

Demostración. La prueba se realiza por inducción sobre el número de variables. Para $n = 1$ el resultado se obtiene como consecuencia del primer apartado de la proposición anterior. Supongamos el resultado probado para $n - 1$, veamos que ocurre para $n > 1$.

Para todo $t \in \mathrm{K}$, definimos el polinomio $\mathrm{P}_t \in \mathrm{K}[\mathrm{X}_1, ..., \mathrm{X}_{n-1}]$ dado por:

$$\mathrm{P}_t(\mathrm{X}_1, ..., \mathrm{X}_{n-1}) = \mathrm{P}(\mathrm{X}_1, ..., \mathrm{X}_{n-1}, t).$$

Simplemente estamos evaluando en la última coordenada. Como el polinomio P de partida tiene grado menor o igual que d, los polinomios P_t también

tienen grado menor o igual que d, para todo $t \in \mathrm{K}$. Si el polinomio P_t es idénticamente nulo para $t \in \mathrm{K}$, podemos escribir:

$$\mathrm{P}(\mathrm{X}_1, ..., \mathrm{X}_n) = (\mathrm{X}_n - t)\mathrm{Q}(\mathrm{X}_1, ..., \mathrm{X}_n)$$

donde Q es un polinomio de grado menor o igual que $d-1$. Si tomamos $s \neq t$, gracias a la escritura de P, observamos que Q_s es idénticamente nulo si, y sólo si P_s es idénticamente nulo. Sea N el conjunto formado por los $t \in$K tales que P_t es el polinomio nulo. Necesariamente el cardinal de N es menor o igual que d y podemos reescribir el polinomio de partida:

$$\mathrm{P}(\mathrm{X}_1, ..., \mathrm{X}_n) = (\prod_{t \in N}(\mathrm{X}_n - t))\mathrm{R}(\mathrm{X}_1, ..., \mathrm{X}_n)$$

para algún polinomio $\mathrm{R} \in \mathrm{K}[\mathrm{X}_1, ..., \mathrm{X}_n]$ que tiene grado menor o igual que $d - |N|$, tal que R_s no es idénticamente nulo para todo $s \in \mathrm{K}$ (recordemos que $\mathrm{K}[\mathrm{X}_1, ..., \mathrm{X}_n]$ es un dominio de factorización única). Podemos evaluar P en los puntos de K^n. Vemos que si P se anula en $(a_1, ..., a_n) \in \mathrm{K}^n$ solo hay dos opciones posibles, o bien tenemos $a_n = t \in N$ o bien $s = a_n \notin N$ y entonces se debe dar que R_s se anula en $(a_1, ...a_{n-1})$. Recapitulando:

$$V(\mathrm{P}) \ \subset \ (\mathrm{K}^{n-1} \times N) \ \cup \ \bigcup_{s \in \mathrm{K} \setminus N} (V(\mathrm{R}_s) \times \{s\}).$$

Aplicando la hipótesis de inducción a los polinomios R_s, en $n-1$ variables, podemos deducir que $|V(\mathrm{R}_s) \times \{s\}| \leq (d - |N|)\,|\mathrm{K}|^{n-2}$.

$$\begin{aligned} |V(\mathrm{P})| &\leq |\mathrm{K}|^{n-1}\,|N| \ + \sum_{s \in \mathrm{K} \setminus N} (d - |N|)\,|\mathrm{K}|^{n-2} \\ &\leq |\mathrm{K}|^{n-1}\,|N| \ + |\mathrm{K}|\,(d - |N|)\,|\mathrm{K}|^{n-2} \\ &= d\,|\mathrm{K}|^{n-1} \end{aligned}$$

□

Observamos que para cuerpos finitos, el número de puntos de una hipersuperficie está controlado por el grado del polinomio.

Proposición 1.4. Sea K un cuerpo y sean $n \geq 1$, $d \geq 0$ dos enteros. Si un subconjunto S de K^n tiene cardinal menor que $\binom{d+n}{n}$, entonces existe un polinomio no nulo $\mathrm{P} \in \mathrm{K}[\mathrm{X}_1, ..., \mathrm{X}_n]$ de grado menor o igual que d tal que $S \subset V(\mathrm{P})$.

Demostración. Recordemos primero que el K-espacio vectorial E formado por los polinomios de grado menor o igual que d de $\mathrm{K}[\mathrm{X}_1, ..., \mathrm{X}_n]$ tiene dimensión igual a $\binom{d+n}{n} = \frac{(d+n)...(d+1)}{n!}$. Lo vemos examinando el cardinal de una base. Una base de este espacio vectorial está dada por el conjunto formado por todos los monomios $\mathrm{X}_1^{d_1}\mathrm{X}_2^{d_2}...\mathrm{X}_n^{d_n}$, de grado menor o igual que d. Esta base tiene tantos elementos como el conjuntos de elementos $(d_1, ..., d_n) \in \mathbb{Z}_{\geq 0}^n$ con $d_1 + ... + d_n \leq d$, que tiene cardinal igual a $\binom{d+n}{n}$.

Para probar el resultado simplemente si $|S| < \binom{d+n}{n}$, consideremos la aplicación de evaluación, que es lineal y está dada por:

$$\begin{aligned} \phi : \mathrm{E} &\longrightarrow \mathrm{K}^S \\ \mathrm{P} &\longmapsto (\mathrm{P}(x))_{x \in S} \end{aligned}$$

donde K^S es un K-espacio vectorial que tiene dimensión menor estrictamente que $\binom{d+n}{n}$. Deducimos que ϕ tiene un núcleo no trivial y podemos concluir. □

Como $\binom{d+n}{n} \geq \frac{d^n}{n^n}$ se puede deducir que todo subconjunto finito S de K^n está contenido en una hipersuperficie de grado menor o igual que $n\,|S|^{\frac{1}{n}}$.

Proposición 1.5. Sean K un cuerpo, $n \geq 1$ un entero, $V(\mathrm{P}) \subset \mathrm{K}^n$ una hipersuperficie de grado menor o igual que d y r una recta de K^n. Entonces o r está contenida en $V(\mathrm{P})$, o en caso contrario, $V(\mathrm{P}) \cap r$ tiene cardinal menor o igual que d.

Demostración. Tenemos una recta, que podemos escribir:

$$r_{x,y} = \{x + ty : t \in \mathrm{K}\},$$

donde $x, y \in \mathrm{K}^n$. Para ver el resultado basta aplicar la primera proposición al polinomio $\mathrm{Q}(t) \in \mathrm{K}[t]$, dado por $\mathrm{Q}(t) = \mathrm{P}(x + tv)$, resultado de evaluar P en los puntos de la recta. □

Vemos de esta forma que si una recta contiene $d + 1$ puntos de una hipersuperficie de grado d, entonces la recta está contenida en la hipersuperficie.

Proposición 1.6 (Nykodyn)**.** Sean K un cuerpo finito y $n, d \geq 1$ enteros. Sea S un subconjunto de K^n que cumple que por cada punto $a \in \mathrm{K}^n$, existe una recta r que pasa por a y contiene más de d puntos de S. Entonces $|S| \geq \binom{d+n}{n}$.

Demostración. Por hipótesis tenemos $d < |K|$. Razonemos por reducción al absurdo y supongamos que $|S| < \binom{d+n}{n}$.

Por la proposición 1.4, sabemos que S está contenido en una hipersuperficie $V(\mathrm{P})$, donde P es un polinomio de grado menor o igual que d. Si $a \in \mathrm{K}^n$, por hipótesis existe una recta r que contiene más de d puntos de S, por lo que contiene más de d puntos de $V(\mathrm{P})$. La proposición anterior permite concluir que r está contenida en $V(\mathrm{P})$. En particular, $a \in V(\mathrm{P})$ para cada $a \in \mathrm{K}^n$, por lo que $|V(\mathrm{P})| = |\mathrm{K}^n| = |\mathrm{K}|^n$ y llegamos a contradecir la proposición 1.3. □

2 Multiplicidad

Podemos mejorar los resultados expuestos anteriormente si no nos fijamos únicamente en los puntos donde se anula un polinomio y tenemos en cuenta que un polinomio puede anularse en un punto dado con una cierta multiplicidad. De esta forma pueden afinarse las cotas expuestas.

Sea K un cuerpo y P un polinomio de K[X]. Dado un punto $z \in \mathrm{K}$, decimos que P se anula con multiplicidad m en z si el polinomio P es divisible por $(\mathrm{X} - z)^m$. Es útil expresar la multiplicidad en términos de la derivada de Hasse.

Definición 1.7. Para todo $\mathrm{P} = \sum_{i=0}^{d} a_i \mathrm{X}^i \in \mathrm{K}[\mathrm{X}]$, definimos la r-ésima derivada de Hasse de P:

$$\mathrm{D}^r(\mathrm{P})(\mathrm{X}) = \mathrm{D}^r\left(\sum_{i=0}^{d} a_i \mathrm{X}^i\right) = \sum_{i=0}^{d} \binom{i}{r} a_i \mathrm{X}^{i-r}.$$

La derivada de Hasse tiene muchas propiedades en común con la derivada usual. Se deduce inmediatamente de la definición que D^r es una aplicación K-lineal de K[X] en K[X]. Además, podemos ver que la derivada de Hasse satisface la regla de Leibniz, por lo que se trata realmente de una derivación de K[X]. Lo comprobamos:

Queremos ver que se cumple $\mathrm{D}^r(\mathrm{PQ}) = \sum_{i=0}^{r} \mathrm{D}^i(\mathrm{P})\mathrm{D}^{r-i}(\mathrm{Q})$, para cada par de polinomios $\mathrm{P}, \mathrm{Q} \in \mathrm{K}[\mathrm{X}]$.

Si fijamos $\mathrm{P} \in \mathrm{K}[\mathrm{X}]$, obtenemos una aplicación K-lineal de K[X] en K[X] que envía Q en $\mathrm{D}^r(\mathrm{PQ})$. Deducimos que para probar la fórmula, basta considerar el caso $\mathrm{P} = \mathrm{X}^n$ para n natural. De forma análoga, razonando para Q

deducimos que basta considerar el caso $\mathrm{P} = \mathrm{X}^m$ para m natural. Tenemos:

$$\mathrm{D}^r(\mathrm{X}^n\mathrm{X}^m) = \mathrm{D}^r(\mathrm{X}^{n+m}) = \binom{n+m}{r}\mathrm{X}^{n+m-r}$$

y

$$\mathrm{D}^i(\mathrm{X}^n)\mathrm{D}^{r-i}(\mathrm{X}^m) = \binom{n}{i}\mathrm{X}^{n-i}\binom{m}{r-i}\mathrm{X}^{m-(r-i)} = \binom{n}{i}\binom{m}{r-i}\mathrm{X}^{n+m-r}.$$

Vemos que basta probar que $\sum_{i=0}^{r}\binom{n}{i}\binom{m}{r-i} = \binom{n+m}{k}$, pero esto es simplemente la identidad de Vandermonde para los coeficientes binomiales.

En general, no se tienen todas las propiedades conocidas de la derivada usual, para la derivada de Hasse. Por ejemplo, es sencillo probar utilizando argumentos similares a los expuestos, que $\mathrm{D}^i\mathrm{D}^j = \binom{i+j}{j}\mathrm{D}^{i+j}$, por lo que en general $\mathrm{D}^i\mathrm{D}^j \neq \mathrm{D}^{i+j}$.

Lo que nos va a permitir trabajar con la multiplicidad es la existencia de una fórmula de Taylor dada gracias a la derivada de Hasse. Para cada $\mathrm{P} \in \mathrm{K}[\mathrm{X}]$ de grado $d(\mathrm{P})$, se tiene:

$$\mathrm{P}(\mathrm{X}) = \sum_{i=0}^{d(\mathrm{P})}(\mathrm{D}^i(\mathrm{P}))(z)(\mathrm{X}-z)^i,$$

donde $z \in \mathrm{K}$. Para ver que la fórmula es correcta, razonamos como antes y observamos que por linealidad basta probarla para $\mathrm{P} = \mathrm{X}^n$ donde n es natural. Entonces se tiene:

$$\begin{aligned}\sum_{i=0}^{n}(\mathrm{D}^i(\mathrm{X}^n))(z)(\mathrm{X}-z)^i &= \sum_{i=0}^{n}\left(\binom{n}{i}\mathrm{X}^{n-i}\right)(z)(\mathrm{X}-z)^i \\ &= \sum_{i=0}^{n}\binom{n}{i}z^{n-i}(\mathrm{X}-z)^i \\ &= ((\mathrm{X}-z)+z)^n = \mathrm{X}^n.\end{aligned}$$

Por construcción observamos que $\mathrm{P} \in \mathrm{K}[\mathrm{X}]$ se anula en un punto $z \in \mathrm{K}$ con multiplicidad m si, y sólo si todas las derivadas $\mathrm{D}^0\mathrm{P}$, ...,$\mathrm{D}^{m-1}\mathrm{P}$ se anulan en z.

Es importante hacer notar que el empleo de la derivada de Hasse solventa los problemas que plantea la derivada usual, cuando tratamos con cuerpos de característica positiva. En particular nos permite definir una fórmula de

Taylor para estos cuerpos.

Esta construcción se generaliza al caso de varias variables. Para cada polinomio $\mathrm{P} \in \mathrm{K}[\mathrm{X}_1, ..., \mathrm{X}_n]$, podemos definir su derivada de Hasse multidimensional $\mathrm{D}^r(\mathrm{P}) = \mathrm{D}^{r_1,...,r_n}(\mathrm{P})$ con $r = (r_1, ..., r_n) \in \mathbb{Z}^n_{\geq 0}$, dada por:

$$\begin{aligned}\mathrm{D}^{r_1,...,r_n}(\mathrm{P}) &= \mathrm{D}^{r_1,...,r_n}\left(\sum_{i_1,...,i_n} a_{i_1,...,i_n}\mathrm{X}_1^{i_1}...\mathrm{X}_n^{i_n}\right) \\ &= \sum_{i_1,...,i_n} \mathrm{D}^{r_1,...,r_n} a_{i_1,...,i_n}\binom{i_1}{r_1}...\binom{i_n}{r_n}\mathrm{X}_1^{i_1}...\mathrm{X}_n^{i_n}.\end{aligned}$$

Se tiene la fórmula de Taylor correspondiente:

$$\mathrm{P}(\mathrm{X}_1, ..., \mathrm{X}_n) = \sum_{i_1,...,i_n} \mathrm{D}^{i_1,...,i_n}\mathrm{P}(z_1, ..., z_n)(\mathrm{X} - z_1)^{i_1}...(\mathrm{X} - z_n)^{i_n}.$$

Diremos que $\mathrm{P} \in \mathrm{K}[\mathrm{X}_1, ..., \mathrm{X}_n]$ se anula con multiplicidad m en $z \in \mathrm{K}^n$ si todas las derivadas $\mathrm{D}^{r_1,...,r_n}\mathrm{P}$ con $r_1 + ... + r_n < m$ se anulan en z. Llamaremos orden o multiplicidad de P en z al mayor m para el que se cumple esta condición de anulación para todas las derivadas. Lo denotaremos con $\mathrm{ord}_z(\mathrm{P})$. Emplearemos la convención $\mathrm{ord}_z(\mathrm{P}) = \infty$ cuando P sea el polinomio nulo. Por construcción tenemos $\mathrm{ord}_z(\mathrm{P}) > 0$ si, y sólo si $z \in V(\mathrm{P})$. Se comprueba fácilmente que el orden cumple que para cada par de polinomios $\mathrm{P}, \mathrm{Q} \in \mathrm{K}[\mathrm{X}_1, ..., \mathrm{X}_n]$ y cada punto $z \in \mathrm{K}^n$, $\mathrm{ord}_z(\mathrm{PQ}) = \mathrm{ord}_z(\mathrm{P}) + \mathrm{ord}_z(\mathrm{Q})$.

Estamos ahora en condiciones de mejorar los resultados presentados con anterioridad.

Proposición 1.8. Sea K un cuerpo. Se tiene:

- Si $\mathrm{P} \in \mathrm{K}[\mathrm{X}]$ es un polinomio no trivial de grado menor o igual que $d \geq 0$, entonces $\sum_{z\in\mathrm{K}} \mathrm{ord}_z(\mathrm{P}) \leq d$.
- Si $\{a_z\}_{z\in\mathrm{K}}$ son números naturales que cumplen $\sum_{z\in\mathrm{K}} a_z \leq d$, entonces existe un polinomio no trivial $\mathrm{P} \in \mathrm{K}[\mathrm{X}]$ de grado d y tal que $\mathrm{ord}_z(\mathrm{P}) \geq a_z$ para cada $z \in \mathrm{K}$.

Demostración. Para probar el primer punto, simplemente repetimos el argumento inductivo empleado para probar el primer punto de la proposición 1.2. donde permitimos que se repitan los puntos que consideramos.

El segundo punto se prueba de forma análoga, repitiendo el argumento empleado para probar el segundo punto de la proposición 1.2., construyendo el polinomio $\mathrm{P}(\mathrm{X}) = \prod_{z\in\mathrm{K}}(\mathrm{X}-z)^{a_z}$. Como la suma de los a_z es menor o igual que d, realmente obtenemos un polinomio puesto que el producto es finito, y además tiene el grado necesario y cumple la propiedad que buscábamos.

□

Proposición 1.9 (Schwartz-Zippel con multiplicidad)**.** Sea K un cuerpo finito y P un polinomio de $\mathrm{K}[\mathrm{X}_1, ..., \mathrm{X}_n]$ para $n \geq 1$. Si P no es el polinomio nulo y tiene grado menor o igual que $d \geq 0$, entonces:

$$\sum_{t\in\mathrm{K}^n} \mathrm{ord}_t(\mathrm{P}) \leq d\,|\mathrm{K}|^{n-1}.$$

Demostración. De nuevo la prueba consiste en adaptar la prueba realizada para la proposición 1.3. empleando un razonamiento inductivo sobre el número de variables. Para $n = 1$ el resultado se deduce del primer punto de la proposición anterior. Supongamos el resultado probado para $n-1$, veamos que ocurre para $n > 1$. Factorizamos en P todos los términos $(\mathrm{X}_n - z)$ que aparecen en P, obteniendo una expresión de P de la forma:

$$\mathrm{P}(\mathrm{X}_1, ..., \mathrm{X}_n) = \left(\prod_{z\in\mathrm{K}}(\mathrm{X}_n - z)^{a_z}\right)\mathrm{Q}(\mathrm{X}_1, ..., \mathrm{X}_n)$$

donde los $\{a_z\}_{z\in\mathrm{K}}$ cumplen por construcción $\sum_{z\in\mathrm{K}} a_z \leq d$. El polinomio Q es no nulo por serlo P, tiene grado menor o igual $d - \sum_{z\in\mathrm{K}} a_z$ y cumple que todos los polinomios $\mathrm{Q}_z = \mathrm{Q}(\mathrm{X}_1, ..., \mathrm{X}_{n-1}, z) \in \mathrm{K}[\mathrm{X}_1, ..., \mathrm{X}_{n-1}]$ para $z \in \mathrm{K}$ son no nulos, ya que si alguno no lo fuera se podría factorizar un $\mathrm{X}_n - z$. Por construcción se tiene para cada $t = (t_1, ..., t_n) \in \mathrm{K}^n$:

$$\mathrm{ord}_t(\mathrm{P}) = a_{t_n} + \mathrm{ord}_t(\mathrm{Q}).$$

Si sumamos todas estas cantidades para cada $t \in \mathrm{K}^n$ obtenemos:

$$\sum_{t\in\mathrm{K}^n} \mathrm{ord}_t(\mathrm{P}) = |\mathrm{K}|^{n-1}\sum_{z\in\mathrm{K}} a_z + \sum_{t\in\mathrm{K}^n} \mathrm{ord}_t(\mathrm{Q})$$

Además, por construcción de los polinomios $\mathrm{Q}_z \in \mathrm{K}[\mathrm{X}_1, ..., \mathrm{X}_{n-1}]$ para $z \in \mathrm{K}$ se cumple para cada $(t_1, ..., t_n) \in \mathrm{K}^n$:

$$\mathrm{ord}_{(t_1,...,t_n)}(\mathrm{Q}) \leq \mathrm{ord}_{(t_1,...,t_{n-1})}(\mathrm{Q}_{t_n}).$$

Aplicando la hipótesis de inducción a los polinomios Q_z, se deduce:

$$\sum_{(t_1,...,t_{n-1})\in\mathrm{K}^{n-1}} \mathrm{ord}_{(t_1,...,t_{n-1})}(\mathrm{Q}_{t_n}) \leq \left(d - \sum_{z\in\mathrm{K}} a_z\right)|\mathrm{K}|^{n-2}.$$

Sumando ahora en t_n tenemos que:

$$\sum_{t\in \mathrm{K}^n} \mathrm{ord}_t(\mathrm{Q}) \leq \left(d - \sum_{z\in \mathrm{K}} a_z \right) |\mathrm{K}|^{n-1}$$

que permite concluir.

□

Proposición 1.10. Sea K un cuerpo y sean $n \geq 1$, $d \geq 0$ dos enteros. Si $\{a_z\}_{z\in \mathrm{K}^n}$ son números naturales tales que $\sum_{z\in \mathrm{K}^n} \binom{a_z+n-1}{n} < \binom{d+n}{n}$, entonces existe un polinomio no nulo $\mathrm{P} \in \mathrm{K}[\mathrm{X}_1, ..., \mathrm{X}_n]$ de grado menor o igual que d tal que $\mathrm{ord}_z(\mathrm{P}) \geq a_z$ para cada $z \in \mathrm{K}^n$.

Demostración. Para realizar esta prueba simplemente hay que modificar el argumento utilizado en la prueba de la proposición 1.4. Sea E el K-espacio vectorial formado por los polinomios de $\mathrm{K}[\mathrm{X}_1, ..., \mathrm{X}_n]$ de grado menor o igual que d, que como vimos tiene dimensión $\binom{d+n}{n}$. Consideramos la aplicación:

$$\mathrm{P} \mapsto (\mathrm{D}^{r_1,...,r_n}\mathrm{P}(z))_{z\in \mathrm{K}^n \text{ y } r_1+...+r_n<a_z}$$

El espacio de llegada tiene dimensión igual a $\sum_{z\in \mathrm{K}^n} \binom{a_z+n-1}{n}$ que es estrictamente menor que $\binom{d+n}{n}$. Como en la proposición 1.4. se deduce que el núcleo de la aplicación es no trivial y con ello se obtiene la existencia del polinomio deseado.

□

Vemos que si tomamos todos los a_z iguales a un cierto entero m, para todo subconjunto S de K^n se puede encontrar una hipersuperficie de grado menor o igual que d que se anula con multiplicidad m en todo punto de S siempre que se cumpla la cota $\binom{m+n-1}{n} |S| < \binom{d+n}{n}$ dada por la proposición, para este caso concreto. De las cotas $\binom{m+n-1}{n} < \frac{(m+n)^n}{n!}$ y $\binom{d+n}{n} \geq \frac{d^n}{n!}$ deducimos que se puede tomar $d = (m+n)\,|S|^{\frac{1}{n}}$.

Finalmente nos queda la mejora de la proposición 1.5.

Proposición 1.11. Sean K un cuerpo, $n \geq 1$ un entero, $V(\mathrm{P}) \subset \mathrm{K}^n$ una hipersuperficie de grado menor o igual que d y $r_{x,y} = \{x + ty : t \in \mathrm{K}\}$ una recta dada por $x \in \mathrm{K}^n, y \in \mathrm{K}^n \setminus \{0\}$ distintos. Entonces o $r_{x,y}$ está contenida en $V(\mathrm{P})$, o se cumple que $\sum_{z\in \mathrm{K}} \mathrm{ord}_z \mathrm{P}(x + Ty) \leq d$ con $\mathrm{P}(x + Ty) \in \mathrm{K}[\mathrm{T}]$.

Demostración. Se deduce directamente del primer apartado de la proposición 1.8.

□

3 Un primer ejemplo: La conjetura finita de Kakeya

Vamos a ver una primera aplicación de las proposiciones desarrolladas anteriormente. Diremos que un subconjunto S de $\mathbb{R}^n$ es de Kakeya cuando contiene un segmento unidad para cada dirección. Por ejemplo, es sencillo comprobar que los conjuntos conexos y acotados, del plano que tienen por frontera una circunferencia o un triángulo de tamaño suficientemente grande, son conjuntos de Kakeya. La conjetura de Kakeya trata sobre las dimensiones de tales conjuntos, y para ello necesitamos introducir el concepto de dimensión de Hausdorff.

Sea S un subconjunto no vacío del espacio métrico $(\mathbb{R}^n, d)$. Denotaremos con $\delta(S)$ al diámetro del subconjunto S, que corresponde a la cantidad $\sup\{d(x,y) : x, y \in S\}$. Para $p \geq 0$ y $s \geq 0$, definimos:

$$\mathrm{H}_{p,s}(S) = \inf\left\{\sum_{i=1}^{\infty} \delta(A_i)^p : S \subset \bigcup_{i=1}^{\infty} A_i, \delta(A_i) < s\right\}.$$

con el convenio $\inf(\emptyset) = \infty$. Si no podemos recubrir S por una cantidad numerable de conjuntos de diámetro menor que s, asignamos el valor ∞ a $H_{p,s}(S)$. Precisamos además el límite que nos permitirá definir la dimensión de Hausdorff:

$$\mathrm{H}_p(S) = \lim_{s\to 0} \mathrm{H}_{p,s}(S) = \sup_{s>0} \mathrm{H}_{p,s}(S).$$

Observamos que $\mathrm{H}_p(S) \in [0, \infty]$ y $\mathrm{H}_p(\emptyset) = 0$. Es sencillo comprobar que para todo $S \subset \mathbb{R}^n$ y todo $p > n$, $\mathrm{H}_p(S) = 0$.

Una medida exterior en X es una aplicación μ del conjunto de partes de X en el intervalo $[0, \infty]$ que verifica:

- $\mu(\emptyset) = 0$.
- $\mu(A) \leq \mu(B)$ si $A \subset B$.
- $\mu(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} \mu(A_i)$ para cada familia numerable $\{A_i\}_{i=1}^{\infty}$.

Veamos que $\mathrm{H}_{p,s}$ y H_p son medidas exteriores. Sea $S \subset \bigcup_{i=1}^{\infty} A_i$. Sin pérdida de generalidad, podemos suponer que para todo i, $\mathrm{H}_{p,s}(A_i) < \infty$. Para todo $\epsilon > 0$, existen conjuntos B_i^j tales que $A_i \subset \bigcup_{j=1}^{\infty} B_i^j$, $\delta(B_i^j) < s$ y además:

$$\mathrm{H}_{p,s}(A_i) + \frac{\epsilon}{2^i} \geq \sum_{j=1}^{\infty} \delta(B_i^j)^p.$$

Como $S \subset \bigcup_{i=1}^{\infty}(\bigcup_{j=1}^{\infty} B_i^j)$, tenemos:

$$\sum_{i=1}^{\infty} \mathrm{H}_{p,s}(A_i) + \epsilon \geq \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \delta(B_i^j)^p \geq \mathrm{H}_{p,s}(S).$$

Haciendo $\epsilon \to 0$, comprobamos que $\mathrm{H}_{p,s}(S) \leq \sum_{i=1}^{\infty} \mathrm{H}_{p,s}(A_i)$ y concluimos que $\mathrm{H}_{p,s}$ es medida exterior. Que H_p es medida exterior se deduce de que $\mathrm{H}_{p,s}$ lo es, ya que tenemos:

$$\mathrm{H}_{p,s}(S) \leq \sum_{i=1}^{\infty} \mathrm{H}_{p,s}(A_i) \leq \sum_{i=1}^{\infty} \mathrm{H}_p(A_i).$$

Tomando el límite para $s \to 0$ se deduce que $\mathrm{H}_p(S) < \sum_{i=1}^{\infty} \mathrm{H}_p(A_i)$ y por lo tanto H_p es medida exterior.

Veamos ahora que H_p es una medida exterior métrica, es decir, veremos que dados $A, B \subset \mathbb{R}^n$ con $d(A, B) = d_0 > 0$, tenemos:

$$\mathrm{H}_p(A \cup B) = \mathrm{H}_p(A) + \mathrm{H}_p(B).$$

De las dos desigualdades a probar, $\mathrm{H}_p(A \cup B) \leq \mathrm{H}_p(A) + \mathrm{H}_p(B)$ se deriva inmediatamente por construcción. La desigualdad inversa se cumple también cuando $\mathrm{H}_p(A \cup B) = \infty$. En el caso finito, consideramos conjuntos $\{E_i\}_{i=1}^{\infty}$ tales que $A \cup B \subset \bigcup_{i=1}^{\infty} E_i$ y $\delta(E_i) < s < \frac{d_0}{4}$. Por construcción, cada E_i interseca únicamente a uno de los dos conjuntos A o B. Podemos entonces separar los conjuntos E_i en dos familias $\{E_i^A\}_{i=1}^{\infty}$ y $\{E_i^B\}_{i=1}^{\infty}$, de forma que $A \subset \bigcup_{i=1}^{\infty} E_i^A$ y $B \subset \bigcup_{i=1}^{\infty} E_i^B$. Entonces:

$$\sum_{i=1}^{\infty} \delta(E_i)^p = \sum_{i=1}^{\infty} \delta(E_i^A)^p + \sum_{i=1}^{\infty} \delta(E_i^B)^p \geq \mathrm{H}_{p,s}(A) + \mathrm{H}_{p,s}(B).$$

Se deduce que para todo s tal que $0 < s < \frac{d_0}{4}$, se tiene:

$$\mathrm{H}_{p,s}(A \cup B) \geq \mathrm{H}_{p,s}(A) + \mathrm{H}_{p,s}(B).$$

Basta tomar limites para obtener la desigualdad deseada, y por lo tanto concluimos que H_p es una medida exterior métrica. Como toda medida exterior métrica define una medida de Borel, observamos que H_p define una medida de Borel a la que llamamos medida p-dimensional de Hausdorff.

Comprobamos ahora que dados $S \subset \mathbb{R}^n$ y $0 \leq p < q$, si $\mathrm{H}_p(S) < \infty$, entonces $\mathrm{H}_q(S) = 0$. Como $\mathrm{H}_p(S) < \infty$, existen $\{A_i\}_{i=1}^{\infty}$ con $\delta(A_i) < s$ de forma que $S \subset \bigcup_{i=1}^{\infty} A_i$ y cumple:

$$\sum_{i=1}^{\infty} \delta(A_i)^p \leq \mathrm{H}_p(S) + 1$$

Además, puesto que $\delta(A_i)^p = \delta(A_i)^q \delta(A_i)^{p-q} \geq s^{p-q}\delta(A_i)^q$, deducimos que $\mathrm{H}_{q,s}(S) \leq \sum_{i=1}^{\infty} \delta(A_i)^q \leq s^{q-p}(\mathrm{H}_p(S) + 1)$. Como el último término tiende a 0 cuando s tiende a 0, se concluye.

Finalmente, definimos la dimensión de Hausdorff de $S \subset \mathbb{R}^n$:

$$\dim_{\mathrm{H}}(S) = \sup\{p \geq 0 : \mathrm{H}_p(S) = \infty\} = \inf\{p > 0 : \mathrm{H}_p(S) = 0\}.$$

En $\mathbb{R}^n$, se conoce como construir conjuntos de Kakeya de medida tan pequeña como se quiera. Los llamados conjuntos de Besicovitch, son conjuntos de Kakeya que tienen medida nula, pero todos ellos tienen dimensión de Hausdorff igual a n. En [Fal] podemos encontrar todos los detalles de su construcción y propiedades.

La conjetura de Kakeya en $\mathbb{R}^n$ puede enunciarse: Si S es un conjunto de Kakeya, entonces $\dim_{\mathrm{H}}(S) = n$.

Para $n = 1$ y $n = 2$ se ha probado positivamente la conjetura, aunque para $n > 2$ el problema sigue abierto Se trata de uno de los problemas centrales abiertos del análisis matemático, sobre el que han contribuido numerosos autores, entre ellos J.Bourgain y T. Tao. Para n=2 fue probado por R. Davies y una prueba de 1971 puede encontrarse en [Dav].

Nos interesa estudiar en este trabajo la llamada conjetura de Kakeya para cuerpos finitos. Se trata de una simplificación del problema de $\mathbb{R}^n$, propuesta por Wolff en 1999. El enunciado es más sencillo y evita todos los problemas técnicos ligados al manejo de la dimensión de Hausdorff.

Dado K un cuerpo finito, diremos que un subconjunto de $S \subset \mathrm{K}^n$ es de Kakeya si contiene una recta en cada dirección, es decir, si para cada elemento no nulo $x \in \mathrm{K}^n$ existe un $y \in \mathrm{K}^n$ tal que $x + ty \in S$ para todo $t \in \mathrm{K}$.

Teorema 1.12 (Conjetura finita de Kakeya)**.** Sean K un cuerpo finito, un entero $n \geq 1$ y un conjunto de Kakeya $E \subset \mathrm{K}^n$. Entonces $|E| \geq \binom{|\mathrm{K}|+n-1}{n}$, y en particular $|E| \geq \frac{1}{n!}|\mathrm{K}|^n$.

Demostración. Razonemos por reducción al absurdo. Supongamos que se tiene $|E| < \binom{|\mathrm{K}|+n-1}{n}$. Gracias a la proposición 1.4. sabemos que existe un polinomio no nulo $\mathrm{P} \in \mathrm{K}[\mathrm{X}_1, ..., \mathrm{X}_n]$ de grado d menor o igual que $|\mathrm{K}| - 1$, tal que $E \subset V(\mathrm{P})$. Denotamos con P_l a la componente homogénea de grado l del polinomio P. Es importante hacer notar que como $E \neq \emptyset$, forzosamente tenemos $d \geq 1$.

Sea $y = (y_1, ..., y_n) \in \mathrm{K}^n \setminus \{0\}$. Por hipótesis existe un elemento $x = (x_1, ..., x_n) \in E$ tal que $x + ty \in E$ para cada $t \in \mathrm{K}$. Como $E \subset V(\mathrm{P})$, queda claro que se tiene $\mathrm{P}(x + ty) = 0$, para todo $t \in \mathrm{K}$. Definimos el polinomio $\mathrm{Q} \in \mathrm{K}[\mathrm{T}]$ dado por:

$$\mathrm{Q}(\mathrm{T}) = \mathrm{P}(x + \mathrm{T}y) = \mathrm{P}(x_1 + \mathrm{T}y_1, ..., x_n + \mathrm{T}y_n).$$

Por construcción el polinomio Q tiene grado d menor o igual $|\mathrm{K}| - 1$ y se anula en todo punto de K. Aplicando la proposición 1.3. para el caso $n = 1$, obtenemos que Q es el polinomio nulo. En particular, el coeficiente correspondiente al término de grado d de Q es $\mathrm{P}_d(y) = \mathrm{P}_d(y_1, ..., y_n)$ y debe ser cero.

Deducimos que para cada $y \in \mathrm{K}^n \setminus \{0\}$, $P_d(y) = 0$. Como por construcción P_d es un polinomio homogéneo de grado $d \geq 1$, tenemos también $\mathrm{P}_d(0) = 0$. De esta forma vemos que P_d se anula en todo K^n, es decir, se anula en $|\mathrm{K}|^n$ puntos. Puesto que $d < |\mathrm{K}| - 1$, aplicando la proposición 1.3. deducimos que P_d es el polinomio nulo, por lo que P debe ser el polinomio nulo y llegamos a contradicción completando la prueba.

□

La demostración es puramente polinómica. Es importante mencionar que se trata de una prueba que no proviene de la reutilización de argumentos relacionados con la conjetura en $\mathbb{R}^n$. Es más, no se conoce una demostración de este resultado que no sea polinómica.

La cota que acabamos de dar para el cardinal de los conjuntos de Kakeya, se puede mejorar gracias a los resultados relacionados con multiplicidades.

Teorema 1.13. Sean K un cuerpo finito, un entero $n \geq 1$ y un conjunto de Kakeya $E \subset \mathrm{K}^n$. Entonces $|E| \geq 2^{-n} |K|^n$.

Demostración. Sean $1 \leq l \leq m$ enteros. La proposición 1.10 garantiza que existe un polinomio no nulo $\mathrm{P} \in \mathrm{K}[\mathrm{X}_1, ..., \mathrm{X}_n]$ de d menor o igual que tal que $(m + n) |E|^{\frac{1}{n}}$, que se anula con multiplicidad mayor o igual que m en cada punto de E. Además, si tomamos $i = (i_1, ..., i_n)$ una n-upla de enteros no negativos tales que $i_1 + ... + i_n \leq l$, entonces $\mathrm{D}^{i_1,...,i_n}(\mathrm{P})$ se anula con

multiplicidad mayor o igual que $m-(i_1+...+i_n)$ en todos los puntos de la recta $r_{x,y}=\{x+ty : t\in \mathrm{K}\}$, para cada recta $r_{x,y}$ asociada al conjunto de Kakeya E. Es claro que el polinomio $\mathrm{D}^{i_1,...,i_n}(\mathrm{P})$ tiene grado menor o igual que $(m+n)\,|E|^{\frac{1}{n}}-(i_1+...+i_n)$.

Aplicando la proposición 1.11 deducimos que para cada una de estas rectas se cumple que $r_{x,y}$ está contenida en $V(\mathrm{D}^{i_1,...,i_n}(\mathrm{P}))$ o bien se debe satisfacer que:

$$|\mathrm{K}|\,(m-(i_1+...+i_n))\leq (m+n)\,|E|^{\frac{1}{n}}-(i_1+...+i_n).$$

Podemos tomar l y m de forma que $|\mathrm{K}|\,(m-l)>(m+n)\,|E|^{\frac{1}{n}}-l$. De esta forma todas las rectas $r_{x,y}$ estarán contenidas en $V(\mathrm{D}^{i_1,...,i_n}(\mathrm{P}))$, para cada $i=(i_1,...,i_n)$ con $i_1+...+i_n\leq l$. Razonando ahora como en la demostración anterior, se llega a contradecir la proposición 1.9, siempre que se de $l\,|\mathrm{K}|>(m+n)\,|E|^{\frac{1}{n}}$. Si $|E|^{\frac{1}{n}}<\frac{1}{2}\,|\mathrm{K}|$, entonces tomando l suficientemente grande y eligiendo $m=2l$, se verifican las dos condiciones que llevan a contradicción. Podemos concluir que $|E|^{\frac{1}{n}}\geq\frac{1}{2}\,|\mathrm{K}|$.

□

Es interesante observar que estas pruebas están muy ligadas a las propiedades algebraicas de los cuerpos finitos. Las técnicas empleadas tienen difícil extensión al caso $\mathbb{R}^n$.

Veamos a continuación como se pueden dar conjuntos de Kakeya E de cardinal $|E|\leq 2^{-(n+1)}\,|\mathrm{K}|^n+O(|\mathrm{K}|^{n-1})$, donde empleamos la notación de Landau. Sera necesario explorar dos construcciones según sea la característica del cuerpo par o impar. En [DKSS] podemos encontrar más detalles sobre estas construcciones

Proposición 1.14. Sea K un cuerpo finito de característica impar y $n\geq 1$ un entero. Sea E' el conjunto dado por:

$$E'=\{(x_1,...,x_{n-1},y)\in \mathrm{K}^n : \forall i, x_i+y^2 \text{ es un cuadrado en K}\}.$$

El conjunto $E=E'\cup(\mathrm{K}^{n-1}\times\{0\})$ es de Kakeya y tiene cardinal menor o igual que $2^{-(n+1)}\,|\mathrm{K}|^n+O(|\mathrm{K}|^{n-1})$.

Demostración. Veamos primero que el conjunto así definido es de Kakeya. Sea $v=(v_1,...,v_n)\in \mathrm{K}^n\setminus\{0\}$. Si $v_n=0$, tomamos $u=(0,...,0)\in \mathrm{K}^n$ y para todo $t\in \mathrm{K}$, $u+tv\in \mathrm{K}^{n-1}\times\{0\}\subset E$.

Si $v_n \neq 0$, tomamos $u = \left(\left(\frac{v_1}{2v_n}\right)^2, ..., \left(\frac{v_{n-1}}{2v_n}\right)^2, 0\right)$. Para cada $t \in \mathrm{K}$, el punto $u + tv$ tiene coordenadas $(x_1, ..., x_{n-1}, y)$ donde $x_i = \left(\frac{v_1}{2v_n}\right)^2 + tv_i$, $y = tv_n$ para $1 \leq i \leq n-1$. Por construcción, para cada i tenemos:

$$\left(\frac{v_1}{2v_n} + tv_n\right)^2 = x_i + y^2$$

Concluimos que $x + ty \in E' \subset E$, y por tanto E es un conjunto de Kakeya. El conjunto E' tiene cardinal igual a $|\mathrm{K}| \left(\frac{|\mathrm{K}|+1}{2}\right)^{n-1} = \frac{|\mathrm{K}|^n}{2^{n-1}} + O(|K|^{n-1})$, ya que tenemos $\frac{|\mathrm{K}|+1}{2}$ elecciones posibles para cada $x_i + y^2$ y $|\mathrm{K}|$ elecciones posibles para y. Deducimos que el cardinal de E es $|\mathrm{K}| = |\mathrm{E'}| + |\mathrm{K}|^{n-1} = \frac{|\mathrm{K}|^n}{2^{n-1}} + O(|\mathrm{K}|^{n-1})$ y podemos concluir.

□

Proposición 1.15. Sea K un cuerpo finito de característica 2 y $n \geq 1$ un entero. El conjunto dado por:

$$E = \{(x_1, ..., x_{n-1}, y) \in \mathrm{K}^n : \forall i, \text{ existe } a_i \in \mathrm{K} \text{ tal que } x_i = a_i^2 + a_i y\}.$$

es de Kakeya y tiene cardinal menor o igual que $2^{-(n+1)} |\mathrm{K}|^n + O(|\mathrm{K}|^{n-1})$.

Demostración. Sea $v = (v_1, ..., v_n) \in \mathrm{K}^n \setminus \{0\}$. Si $v_n = 0$, tomamos de nuevo $u = (0, ..., 0) \in \mathrm{K}^n$ y para cada $t \in \mathrm{K}$ tenemos:

$$u + tv = (tv_1, ..., tv_n - 1, 0) = (a_1^2 + ya_1, ..., a_{n-1}^2 + ya_{n-1}, y)$$

con $y = 0$, $a_i = (tv_i)^{\frac{|\mathrm{K}|}{2}}$ y concluimos que $u + tv \in E$ para todo $t \in \mathrm{K}$.

Si $v_n \neq 0$, tomamos $u = \left(\left(\frac{v_1}{v_n}\right)^2, ..., \left(\frac{v_{n-1}}{v_n}\right)^2, 0\right)$. Para cada $t \in \mathrm{K}$, el punto $u + tv$ tiene coordenadas $(x_1, ..., x_{n-1}, y)$, donde $x_i = \left(\frac{v_i}{v_n}\right)^2 + tv_i$, $y = tv_n$. Para $a_i = \frac{v_i}{v_n}$, se tiene:

$$a_i^2 + a_i y = \left(\frac{v_i}{v_n}\right)^2 + tv_i = x_i.$$

Concluimos que efectivamente E es un conjunto de Kakeya. Verificamos finalmente que el conjunto tiene el cardinal deseado. Es claro que el cardinal del conjunto de puntos de la forma $(x_i, ..., x_{n-1}, 0)$ es $|\mathrm{K}|^{n-1}$. Veamos que

pasa cuando $y \neq 0$. Para ello necesitamos conocer el cardinal de los conjuntos $\Lambda(y) = \{a^2 + ay : a \in \mathrm{K}\}$. Para cada $a \in \mathrm{K}$, se verifica:

$$a^2 + ay = (a+y)^2 + y(a+y) \text{ cuando } y \neq 0.$$

Si consideramos la aplicación $s \mapsto s^2 + sy$, observamos que cada imagen tiene exactamente dos antecedentes que son s y $s+y$, por lo que deducimos que cada $\Lambda(y)$ tiene cardinal $\frac{|\mathrm{K}|}{2}$. El cardinal del conjunto formado por los puntos $(x_1, ..., x_{n-1}, y)$ con $y \neq 0$ es $\left(\frac{|\mathrm{K}|}{2}\right)^{n-1}$, y por lo tanto obtenemos:

$$|E| = (q-1)\left(\frac{|\mathrm{K}|}{2}\right)^{n-1} + |\mathrm{K}|^{n-1} = \frac{|\mathrm{K}|^n}{2^{n-1}} + O(|\mathrm{K}|^{n-1})$$

□

Capítulo 2

Teorema de los ceros combinatorio

En geometría algebraica, el Teorema de los ceros de Hilbert es un resultado fundamental que nos permite relacionar los ideales de $\mathrm{K}[\mathrm{X}_1, ..., \mathrm{X}_n]$ y los cerrados de K^n, cuando K es un cuerpo algebraicamente cerrado. El teorema nos dice que si un polinomio $\mathrm{P} \in \mathrm{K}[\mathrm{X}_1, ..., \mathrm{X}_n]$ se anula en $V(\mathrm{I})$, para un ideal I del mismo anillo de polinomios, entonces se tiene un número natural r, tal que $\mathrm{P}^r \in I$. De esta forma, los polinomios que se anulan en toda la variedad definida por un ideal I pertenecen a su radical, que denotamos con $\sqrt{\mathrm{I}}$.

Dado un subconjunto $S \subset \mathrm{K}^n$, podemos considerar el conjunto formado por los polinomios de $\mathrm{K}[\mathrm{X}_1, ..., \mathrm{X}_n]$, que se anulan en todos los puntos de S:

$$I(S) = \{\mathrm{P} \in \mathrm{K}[\mathrm{X}_1, ..., \mathrm{X}_n] : \mathrm{P}(s) = 0 \text{ para cada } s \in S\}.$$

Con esta notación, el Teorema de los ceros de Hilbert nos indica que se tiene $I(V(\mathrm{I})) \subset \sqrt{I}$, siempre que el cuerpo K sea algebraicamente cerrado. Además, como por construcción siempre se tiene la contención contraria $\sqrt{I} \subset I(V(\mathrm{I}))$ y realmente se obtiene una igualdad.

Sea $V(\mathrm{I})$ un cerrado algebraico de K^n. Las propiedades fundamentales de los cerrados algebraicos nos dicen que siempre tenemos $V(\mathrm{I}) = V(\sqrt{\mathrm{I}})$. De esta forma, el Teorema de los ceros de Hilbert da una correspondencia biunívoca entre cerrados algebraicos de K^n e ideales radicales (que coinciden con su radical) del anillo de polinomios $\mathrm{K}[\mathrm{X}_1, ..., \mathrm{X}_n]$, en el caso algebraicamente cerrado. Esta correspondencia nos dice además, que los ideales primos del anillo de polinomios son exactamente los ideales que dan lugar a los cerrados irreducibles, es decir, a las variedades algebraicas de K^n. Adicionalmente los ideales maximales del anillo $\mathrm{K}[\mathrm{X}_1, ..., \mathrm{X}_n]$ son exactamente los de la

forma $(X_1 - a_1, ..., X_n - a_n)$ donde $a_1, ..., a_n \in K$ y corresponden a los puntos $(a_1, ..., a_n) \in K^n$.

Veremos como se puede explotar la idea del Teorema de los ceros de Hilbert, para dar un teorema combinatorio. Para la prueba del teorema de los ceros combinatorio nos apoyaremos en [Alon]. Probaremos primero un lema que nos permitira a continuación realizar las demostraciones.

Lema 2.1. Sean K un cuerpo y P un polinomio de $K[X_1, ..., X_n]$. Se puede ver P como un polinomio en una sola variable X_i para $i = 1, ..., n$ y en tal caso, denotamos con d_i a su grado. Sean $S_1, ...S_n$ subconjuntos de K de forma que $|S_i| \geq d_i + 1$. Con estas condiciones, si el polinomio P se anula en todos los puntos de $S_1 \times ... \times S_n$, entonces P debe ser el polinomio nulo.

Demostración. Razonamos por inducción sobre el número de variables. Para una sola variable, el resultado simplemente es la proposición 1.2. Suponemos que el lema es correcto para $n - 1$ variables, veamos que se cumple también para n.

Para $i = 1, ...n$, sean S_i subconjuntos de K que satisfacen las hipótesis del lema. Escribimos el polinomio $P \in K[X_1, ..., X_n]$ como un polinomio en la variable X_n:

$$P(X_1, ..., X_n) = \sum_{j=0}^{d_n} P_j(X_1, ..., X_{n-1}) X_n^j.$$

Los polinomios P_j son polinomios en las primeras $n-1$ variables. Para cada variable X_i con $i = 1, .., n-1$, por construcción se cumple que el grado de cada P_j con $j = 0, ..., d_n$, visto como polinomio en la variable X_i es menor o igual que d_i.

Supongamos que el polinomio P se anula en todos los puntos de $S_1 \times ... \times S_n$. Dado un elemento $(s_1, ..., s_{n-1}) \in S_1 \times ... \times S_{n-1}$, podemos considerar el polinomio Q de $K[X_n]$ que se obtiene al evaluar P en dicho elemento, manteniendo la última variable. Como Q tiene grado d_n y se anula para todo $s_n \in S_n$, deducimos que Q es identicamente nulo. De esta forma, para cada $j = 0, ..., d_n$, obtenemos que $P_j(s_1, ..., s_{n-1}) = 0$ para todo elemento $(s_1, ..., s_{n-1}) \in S_1 \times ... \times S_{n-1}$. Aplicando la hipótesis de inducción concluimos que cada P_j es identicamente nulo y por lo tanto, P también debe ser identicamente nulo.

□

Los dos resultados que vienen a continuación son lo que llamaremos Teorema de los ceros combinatorio.

Teorema 2.2. Sean K un cuerpo, $\mathrm{P} \in \mathrm{K}[\mathrm{X}_1, ..., \mathrm{X}_n]$ y $S_1, ..., S_n$ subconjuntos finitos de K. Definimos los polinomios $\mathrm{Q}_i = \prod_{s \in S_i} (\mathrm{X}_i - s)$ para $i \in \{1, ..., n\}$ y consideramos el ideal I de $\mathrm{K}[\mathrm{X}_1, ..., \mathrm{X}_n]$, generado por los polinomios Q_i. Si P se anula en todo $V(\mathrm{I})$, entonces existen polinomios $\mathrm{R}_1, ..., \mathrm{R}_n \in \mathrm{K}[\mathrm{X}_1, ..., \mathrm{X}_n]$, donde el grado de cada R_i es menor o igual que la diferencia de los grados de P y Q_i y que cumplen:

$$\mathrm{P} = \sum_{i=1}^{n} \mathrm{R}_i \mathrm{Q}_i.$$

Además, si todos los coeficientes de los polinomios $\mathrm{P}, \mathrm{Q}_1, ..., \mathrm{Q}_n$ pertenecen a un subanillo A de K, entonces se cumple que cada $\mathrm{R}_i \in \mathrm{A}[\mathrm{X}_1, ..., \mathrm{X}_n]$.

Demostración. Definimos $d_i = |S_i| - 1$ para $i = 1, ..., n$. Por hipótesis, el polinomio P se anula en todos los elementos de $S_1 \times ... \times S_n$.

Para $i = 1, ..., n$, se tiene:

$$\mathrm{Q}_i = \prod_{s \in S_i} (\mathrm{X}_i - s) = \mathrm{X}_i^{d_i+1} - \sum_{j=0}^{d_i} \alpha_{ij} \mathrm{X}_i^j.$$

Los coeficientes α_{ij} son elementos de K. Por construcción, los polinomios Q_i se anulan en todos los elementos de S_i, luego todos los elementos $s \in S_i$ cumplen la misma combinación lineal:

$$s^{d_i+1} = \sum_{j=0}^{d_i} \alpha_{ij} \mathrm{s}_i^j.$$

Construimos ahora un polinomio $\tilde{\mathrm{P}}$ a partir del polinomio P. Podemos escribir P como combinación lineal con coeficientes en K de monomios donde aparecen las variables $\mathrm{X}_1, ..., \mathrm{X}_n$ elevadas a unas ciertas potencias. Para cada $i = 1, ..., n$, cuando en la escritura de P aparezca la variable X_i con una potencia menor o igual que d_i no realizaremos ningún cambio. Sin embargo, cuando aparezcan potencias de X_i mayores que d_i, expresaremos esas potencias como combinaciones lineales con coeficientes en K de potencias de X_i menores o iguales que d_i, utilizando las relaciones lineales presentadas anteriormente que satisfacen los elementos de S_i. Construimos el polinomo $\tilde{\mathrm{P}}$ iterando el proceso hasta que no aparezca ningun X_i elevado a un exponente mayor que d_i. Es claro que por la naturaleza del problema, este proceso es

finito.

Lo que estamos haciendo para construir el polinomio $\tilde{\mathrm{P}}$ es restarle a P productos de la forma $\mathrm{R}_i\mathrm{Q}_i$, para unos ciertos polinomios $\mathrm{R}_i \in \mathrm{K}[\mathrm{X}_1, ..., \mathrm{X}_n]$. Por construcción, el grado de los polinomios R_i es menor o igual que la diferencia de los grados de P y Q_i. Además, los coeficientes de los R_i se encuentran en el menor anillo que contenga a los coeficientes de P y $\mathrm{Q}_1, ..., \mathrm{Q}_n$. Tenemos:

$$\tilde{\mathrm{P}} = \mathrm{P} - \mathrm{R}_1\mathrm{Q}_1 - ... - \mathrm{R}_n\mathrm{Q}_n.$$

De esta forma, vemos que $\tilde{\mathrm{P}}$ toma los mismos valores que P al evaluar en los puntos de $S_1 \times ... \times S_n$. Como por hipótesis P se anula en estos puntos, deducimos que $\tilde{\mathrm{P}}$ también lo hace. Estamos entonces en condición de aplicar a $\tilde{\mathrm{P}}$ el lema 2.1, y concluimos que $\tilde{\mathrm{P}}$ debe ser el polinomio nulo. De esta forma se completa la prueba del teorema obteniendo la expresión de P como combinación de los Q_i.

□

Teorema 2.3. Sean K un cuerpo y P un polinomio no nulo de $\mathrm{K}[\mathrm{X}_1, ..., \mathrm{X}_n]$. Suponemos que el grado del polinomio P, que denotamos con d, se puede escribir $d = d_1 + ... + d_n$ donde cada d_i es un entero no negativo. Además, suponemos que el coeficiente del monomio $\mathrm{X}_1^{d_1}...\mathrm{X}_n^{d_n}$ de P es no nulo. Entonces si $S_1, ..., S_n$ son subconjuntos de K con $|S_i| > d_i$ para $i = 1, ..., n$, existen $s_1 \in S_1, ..., s_n \in S_n$ tales que:

$$\mathrm{P}(s_1, ..., s_n) \neq 0$$

Demostración. Podemos suponer sin pérdida de generalidad que $|S_i| = d_i + 1$ para $i = 1, ..., n$. Razonaremos por reducción al absurdo. Supongamos que en las condiciones del teorema, se tiene $\mathrm{P}(s_1, ..., s_n) = 0$ para cada $(s_1, ..., s_n) \in S_1 \times ... \times S_n$. Definimos los polinomios $\mathrm{Q}_i = \prod_{s \in S_i}(\mathrm{X}_i - s)$ para $i = 1, ..., n$. Podemos aplicar el teorema 2.2, y deducimos que existen polinomios $\mathrm{R}_1, ..., \mathrm{R}_n \in \mathrm{K}[\mathrm{X}_1, ..., \mathrm{X}_n]$ con:

$$\mathrm{P} = \sum_{i=1}^{n} \mathrm{R}_i\mathrm{Q}_i.$$

Además, los grados de cada uno de los polinomios R_i son menores o iguales que la diferencia de los grados de P y Q_i. Por hipótesis sabemos que en la escritura de P como suma de monomios, el coeficiente que acompaña al monomio $\mathrm{X}_1^{d_1}...\mathrm{X}_n^{d_n}$ es no nulo. Como es lógico, esto es valido también para el polinomio dado por la suma de los $\mathrm{R}_i\mathrm{Q}_i$.

El grado de cada $\mathrm{R}_i\mathrm{Q}_i$ es a lo sumo igual a d, el grado de P. Si podemos encontrar algún monomio de grado d en $\mathrm{R}_i\mathrm{Q}_i$ igual de ese grado, por construcción de los polinomios Q_i debe ser divisible por $\mathrm{X}_i^{d_i+1}$. De esta forma, observamos que el coeficiente que acompaña al monomio $\mathrm{X}_1^{d_1}...\mathrm{X}_n^{d_n}$ de $\sum_{i=1}^{n}\mathrm{R}_i\mathrm{Q}_i$ debe ser nulo, y por tanto llegamos a contradicción. Concluimos que P no puede anularse en todo $S_1 \times ... \times S_n$.

□

De forma general, cuando buscamos aplicar el teorema de los ceros combinatorio, necesitamos construir un polinomio para el que podamos extraer información sobre alguno de sus coeficientes. La construcción de estos polinomios puede realizarse mediante los métodos expuestos en este trabajo, o depender totalmente del problema en cuestión. Trateremos a continuación, algunos ejemplos que ilustran la importancia del teorema expuesto.

Podemos obtener de forma inmediata un resultado clásico, el teorema de Cauchy-Davenport, como corolario del teorema de los ceros combinatorio. Se trata de un teorema con númerosas generalizaciones y que puede probarse de forma puramente combinatoria, pero es intersante ver como se puede deducir muy facilmente del teorema de los ceros combinatorio.

Proposición 2.4 (Cauchy-Davenport)**.** Sea p un número primo y sean A, B subconjuntos no vacios de $\mathbb{F}_p$, el cuerpo finito con p elementos. Podemos considerar el conjunto $A+B$, formado por los elementos de $\mathbb{F}_p$ que son suma un elemento de A y uno de B. Con estas condiciones, se tiene:

$$|A+B| \geq \min\{|A|+|B|-1, p\}$$

Demostración. Si se cumple que $|A|+|B| > p$, el resultado es trivial. En este caso, se tiene que para cada $x \in \mathbb{F}_p$, el conjunto A y el conjunto formado por los elementos $x-b$ con $b \in B$ se cortan. De esta forma, todo elemento de $\mathbb{F}_p$ puede verse como suma de un elemento de A y uno de B. Se deduce que $A+B = \mathbb{F}_p$.

Supongamos que $|A|+|B| \leq p$, y que el resultado es falso, es decir, $|A+B| \leq |A|+|B|-2$. Razonamos por reducción al absurdo. Sea C un subconjunto de $\mathbb{F}_p$, de forma que $A+B \subset C$ y tal que $|C| = |A|+|B|-2$. Definimos el polinomio $\mathrm{P} \in \mathbb{F}_p[\mathrm{X},\mathrm{Y}]$, de grado $|A|+|B|-2$, que nos permite aplicar el teorema de los ceros combinatorio:

$$\mathrm{P}(\mathrm{X},\mathrm{Y}) = \prod_{c\in C}(\mathrm{X}+\mathrm{Y}-c).$$

Por construcción del polinomio y del conjunto C, siempre se cumple que $\mathrm{P}(a,b)=0$ para cada $(a,b)\in A\times B$. El coeficiente que acompaña al monomio $\mathrm{X}^{|A|-1}\mathrm{Y}^{|B|-1}$ es $\binom{|A|+|B|-2}{|B|-1}\neq 0$. Aplicando el teorema 2.3, deducimos que debe existir un elemento $(a,b)\in A\times B$ tal que $\mathrm{P}(a,b)\neq 0$ y llegamos a contradicción.

□

Aunque existan otras formas de probar este resultado, esta es interesante ya que se puede emplear el mismo razonamiento para probar resultados parecidos, cuando imponemos condiciones adicionales de tipo algebraico. Ilustramos este punto probando el siguiente resultado.

Proposición 2.5. Sea p un número primo y sean A,B subconjuntos no vacios de $\mathbb{F}_p$, de forma que $|A|\neq|B|$. Consideramos el conjunto:

$$\widetilde{A+B}=\{a+b : a\in A, b\in B, a\neq b\}.$$

Se cumple entonces que $\left|\widetilde{A+B}\right|\geq \min\{|A|+|B|-2,p\}$.

Demostración. El razonamiento es análogo al empleado para demostrar la proposición anterior. Como en la demostración anterior, el caso $|A|+|B|>p+1$ es trivial. El $|A|=1$ o $|B|=1$ es sencillo también. Podemos suponer que se cumple $|A|+|B|\leq p+1$ y $|A|,|B|\geq 2$.

De nuevo razonamos por reducción al absurdo. Suponemos que $\widetilde{A+B}\subset C$ para un subconjunto C de $\mathbb{F}_p$ que cumple $|C|=|A|+|B|-3$. De nuevo, definimos un polinomio P que nos permite aplicar el teorema de los ceros combinatorio:

$$\mathrm{P}(\mathrm{X},\mathrm{Y})=(\mathrm{X}-\mathrm{Y})\prod_{c\in C}(\mathrm{X}+\mathrm{Y}-c).$$

El polinomio P tiene grado $|A|+|B|-2$, se anula en todos los elementos de $A\times B$. De nuevo, si observamos el coeficiente que acompaña al monomio $\mathrm{X}^{|A|-1}\mathrm{Y}^{|B|-1}$, se trata de $\binom{|A|+|B|-3}{|A|-2}-\binom{|A|+|B|-2}{|A|-1}\neq 0$. La aplicación del teorema 2.3, permite llegar a contradicción y concluir.

□

De esta forma queda patente la robustez del teorema de los ceros combinatorio. En este caso por ejemplo, el teorema nos permite tratar con condiciones polinómicas siguiendo los mismos razonamientos para las demostraciones.

Otro teorema interesante que tiene que ver con la combinatoria y los polinomios sobre cuerpos finitos es el teorema de Chevalley-Warning, que enunciamos a continuación. Una referencia interesante para esta parte del capítulo

corresponde a [Sch]. Para realizar la demostración introducimos primero dos lemas previos.

Lema 2.6. Sea $\mathbb{F}_q$ un cuerpo finito. Sea un entero N que cumple $0 \leq N < q-1$. Se verifica que:

$$\sum_{x \in \mathbb{F}_q} x^N = 0.$$

Demostración. Si $N = 0$, tenemos:

$$\sum_{x \in \mathbb{F}_q} x^0 = \sum_{x \in \mathbb{F}_q} 1 = q = 0.$$

Supongamos entonces $0 < N < q-1$. Sea μ un generador del grupo cíclico $\mathbb{F}_q^\star$, formado por las unidades del cuerpo $\mathbb{F}_q$. Como μ tiene orden $q-1$, es claro que $\mu^N \neq 1$. Tenemos entonces:

$$\sum_{x \in \mathbb{F}_q} x^N = \sum_{x \in \mathbb{F}_q} (\mu x)^N = \mu^N \sum_{x \in \mathbb{F}_q} x^N.$$

Esto permite concluir.

□

Lema 2.7. Sea P un polinomio de $\mathbb{F}_q[\mathrm{X}_1, ..., \mathrm{X}_n]$, de grado $d < n(q-1)$. Se cumple que:

$$\sum_{(x_1,...,x_n) \in \mathbb{F}_q^n} \mathrm{P}(x_1, ..., x_n) = 0.$$

Demostración. Basta probar el resultado para un polinomio formado por un solo monomio $\mathrm{X}_1^{d_1}, ..., \mathrm{X}_n^{d_n}$, ya que el resultado se extiende por linealidad. En ese caso, tenemos:

$$\sum_{(x_1,...,x_n) \in \mathbb{F}_q^n} \mathrm{P}(x_1, ..., x_n) = \sum_{(x_1,...,x_n) \in \mathbb{F}_q^n} x_1^{d_1} ... x_n^{d_n} = \prod_{i=1}^{n} \Big(\sum_{x_i \in \mathbb{F}_q} x_i^{d_i} \Big).$$

Por hipótesis, el grado de P está acotado y se tiene $d_1 + ... + d_n = d < n(q-1)$. Existe entonces un j tal que $nd_j < d < n(q-1)$ y por lo tanto $d_j < q-1$. Aplicando el lema 2.6 obtenemos:

$$\sum_{x_j \in \mathbb{F}_q} x_j^{d_j} = 0.$$

Esto permite concluir que se satisface la relación esperada.

□

Proposición 2.8 (Chevalley-Warning)**.** Sea $\mathbb{F}_q$ un cuerpo finito de característica p. Sean $\mathrm{P}_1, ..., \mathrm{P}_m$ polinomios de $\mathbb{F}_q[\mathrm{X}_1, ..., \mathrm{X}_n]$, de grados $d_1, ..., d_m$ respectivamente. Supongamos que $d = d_1 + ... + d_m < n$. Podemos considerar el conjunto de ceros asociado a este conjunto de polinomios:

$$V(\mathrm{P}_1, ..., \mathrm{P}_m) = \{(x_1, ..., x_n) \in \mathbb{F}_q^n : \mathrm{P}_1(x_1, ..., x_n) = ... = \mathrm{P}_m(x_1, ..., x_n) = 0\}.$$

Denotando N al cardinal de este conjunto, se debe cumplir que $N \equiv 0 \ (\mathrm{mod} p)$.

Demostración. Construimos un polinomio $\mathrm{Q} \in \mathbb{F}_q[\mathrm{X}_1, ..., \mathrm{X}_n]$ a partir de los datos del problema:

$$\mathrm{Q}(\mathrm{X}_1, ..., \mathrm{X}_n) = \prod_{i=1}^{m} (1 - \mathrm{P}_i^{q-1}(\mathrm{X}_1, ..., \mathrm{X}_n)).$$

Por construcción, el polinomio Q tiene grado $d(q-1) < n(q-1)$. Podemos aplicarle el lema anterior, y deducimos que:

$$\sum_{(x_1,...,x_n) \in \mathbb{F}_q^n} \mathrm{Q}(x_1, ..., x_n) = 0.$$

Para cada $((x_1, ..., x_n) \in \mathbb{F}_q^n)$ y cada $i = 1, ..., m$, los valores $\mathrm{P}_i^{q-1}(x_1, ..., x_n)$ solo pueden ser 1, salvo si $(x_1, ..., x_n)$ es un cero del polinomio, en cuyo caso el valor es 0. Observamos que Q se anula en todos los elementos de $\mathbb{F}_q^n$, salvo en aquellos que son ceros comunes de los polinomios $\mathrm{P}_1, ..., \mathrm{P}_m$, donde Q toma el valor 1. De esta forma, tenemos que:

$$N = \sum_{(x_1,...,x_n) \in \mathbb{F}_q^n} \mathrm{Q}(x_1, ..., x_n) = 0.$$

Y esto permite concluir.

□

Gracias a este teorema, es inmediato y sencillo deducir el teorema de Chevalley, que enunciamos a continuación. Para poner de manifiesto la importancia del teorema de los ceros combinatorio, vamos a realizar una demostración alternativa del teorema de Chevalley. De esta forma, veremos como el teorema de los ceros combinatorio nos permite en muchas ocasiones simplificar pruebas y obtener resultados a veces conocidos de una forma distinta.

Proposición 2.9 (Chevalley)**.** Sea $\mathbb{F}_q$ un cuerpo finito y sean $\mathrm{P}_1, ..., \mathrm{P}_m$ polinomios de $\mathbb{F}_q[\mathrm{X}_1, ..., \mathrm{X}_n]$, de grados $d_1, ..., d_m$. Si $n > d_1 + ... + d_m$ y todos los polinomios P_i tienen un cero común $(y_1, ..., y_n)$, entonces deben de tener otro cero común distinto.

Demostración. Razonamos por reducción al absurdo. Supongamos que solo hay un solo cero común. Definimos un polinomio $Q \in \mathbb{F}_q[X_1, ..., X_n]$:

$$Q(X_1, ..., X_n) = \prod_{i=1}^{m}(1 - P_i^{q-1}(X_1, ..., X_n)) - \lambda \prod_{j=1}^{n} \prod_{x \in \mathbb{F}_q : x \neq y_j} (X_j - x).$$

De esta forma. Q se escribe como suma de dos términos. Cuando evaluamos en el cero común, el término de la izquierda vale 1 y el de la derecha, por como está construido, no se anula El elemento $\lambda \in \mathbb{F}_q$, se toma de forma que $Q(y_1, ..., y_n) = 0$. De esta forma elegimos λ para que se produzca una cancelación y su valor está determinado por el valor que toma el término de la izquierda al evaluar en el cero común. Queda claro además que λ es no nulo.

Si evaluamos Q en cualquier elemento $(x_1, ..., x_n)$ de $\mathbb{F}_q^n$ que no sea el cero común, debemos tener un j tal que $P_j(x_1, ..., x_n) \neq 0$. De esa forma vemos que el término del lado derecho se anula en todo $\mathbb{F}_q^n$, salvo en $(y_1, ..., y_n)$. Además, por construcción el término del lado izquierdo se anula en los mismos puntos. Deducimos que Q se anula en todo $\mathbb{F}_q^n$.

Nos fijamos ahora en el coeficiente que acompaña al monomio $X_1^{q-1}...X_n^{q-1}$ del polinomio Q. Vemos que se trata de $-\lambda$, ya que el término del lado izquierdo tiene grado estrictamente menor que $(p-1)n$. Aplicando el teorema 2.3 llegamos a contradicción y podemos concluir.

□

Capítulo 3

Método de Stepanov

En este capítulo, vamos a centrarnos en ver como se pueden emplear métodos polinómicos, a la hora de estudiar propiedades combinatorias de algunos objetos de la geometría sobre cuerpos finitos. Se conoce como método de Stepanov al procedimiento polinómico que vamos a exponer. Se pueden encontrar referencias interesantes para este capítulo, como son [Ste] y [IK].

Sea p un número primo y q una potencia de p. Consideramos $\mathbb{F}_q$, un cuerpo con q elementos, de característica p y fijamos una clausura algebraica de este cuerpo que denotamos con $\overline{\mathbb{F}_q}$. Podemos definir ahora el objeto que nos va a servir para ilustrar el funcionamiento del método de Stepanov. Vamos a estudiar un ejemplo suficientemente complejo para poder apreciar la eficacia del método y suficientemente sencillo para que la construcción sea fácilmente comprensible. Consideramos la siguiente curva algebraica plana:

$$X(\overline{\mathbb{F}_q}) = \{(x, y) \in \overline{\mathbb{F}_q}^2 : y^2 = \mathrm{P}(x)\}$$

de forma que P es un polinomio del anillo $\mathbb{F}_q[\mathrm{X}]$, que no es un cuadrado en $\overline{\mathbb{F}_q}[\mathrm{X}]$. Denotaremos con d al grado del polinomio P. Como tenemos $\mathbb{F}_q \subset \overline{\mathbb{F}_q}$, podemos estudiar cuantos puntos de la curva $X(\overline{\mathbb{F}_q})$ tienen coordenadas $x, y \in \mathbb{F}_q$. Sea N al número de puntos de $X(\overline{\mathbb{F}_q})$ que tienen coordenadas en $\mathbb{F}_q$. El método de Stepanov nos va a permitir dar una aproximación para N. El objetivo es probar el siguiente teorema, mediante el empleo de razonamientos polinómicos.

Teorema 3.1 (Hasse)**.** Con estas condiciones, si $d \geq 3$ y $q > 4d^2$ entonces se tiene la siguiente cota:

$$|N - q| < 8d\sqrt{q}$$

Gracias al método de Stepanov, vamos a dar una demostración elemental de este teorema que se fundamentará en la construcción de un polinomio con

unas propiedades de anulación adaptadas al problema.

Observamos que si consideramos el caso $p = 2$, la aplicación que envía cada $x \in \mathbb{F}_q$ en x^2 es un automorfismo de $\mathbb{F}_q$. En este caso, tenemos simplemente $N = q$, ya que en $X(\mathbb{F}_q)$ hay tantos puntos como elementos $(x, \mathrm{P}(x))$ con $x \in \mathbb{F}_q$. Podemos suponer por lo tanto que p es un primo impar.

Se pueden distinguir dos tipos de puntos entre los elementos (x, y) de $X(\mathbb{F}_q)$. Si consideramos los puntos tales que $y = 0$, simplemente estamos contando los puntos $(x, 0)$ tales que x es un cero de P y denotamos con N' a esta cantidad. De forma similar, si (x, y) es un punto de $X(\mathbb{F}_q)$ con $y \neq 0$, entonces $\mathrm{P}(x)$ es un cuadrado en $\mathbb{F}_q$. Esto último se cumple si, y sólo si, $\mathrm{P}^{\frac{q-1}{2}}(x) = y^{q-1} = 1$. Recíprocamente, dado $x \in \mathbb{F}_q$ con $\mathrm{P}^{\frac{q-1}{2}}(x) = 1$ existen únicamente dos elementos distintos $y, -y \in \mathbb{F}_q$ que cumplen que $y^2 = \mathrm{P}(x)$. Denotando con N'' al número de elementos $x \in \mathbb{F}_q$ tales que $\mathrm{P}^{\frac{q-1}{2}}(x) = 1$, tenemos:

$$N = N' + 2N''.$$

Para realizar la prueba, nos va a interesar estudiar el cardinal de los subconjuntos de $\mathbb{F}_q$, dados por $S(\lambda) = \{x \in \mathbb{F}_q : \mathrm{P}(x) = 0 \text{ o } \mathrm{P}^{\frac{q-1}{2}}(x) = \lambda\}$ para $\lambda \in \mathbb{F}_q$.

La idea de la demostración consiste en construir un polinomio, para el que podamos controlar el grado y que tenga ceros de un orden fijo en cada uno de los puntos $S(\lambda)$. Vamos a presentar con detalle en la siguiente proposición, cuales son las propiedades del polinomio que vamos a emplear para la prueba del teorema.

Proposición 3.2. Si $q > 8d$, sean l es un entero que satisface $d < l \leq \frac{q}{8}$ y $\lambda \in \mathbb{F}_q$. Entonces existe un polinomio no nulo $\mathrm{Q(X)} \in \mathbb{F}_q[\mathrm{X}]$ de grado $\deg(\mathrm{Q})$ menor o igual que $\frac{q-1}{2}l + 2dl(l-1) + dq$, que se anula con orden mayor o igual que l en cada punto de $S(\lambda)$.

Veamos ahora como construir el polinomio para probar la proposición y que posteriormente utilizaremos para probar el teorema 3.1. Queremos encontrar el polinomio buscado de la forma:

$$\mathrm{Q(X)} = \mathrm{P}^l(\mathrm{X}) \sum_{0 \leq j < J} (\mathrm{R}_j(\mathrm{X}) + \mathrm{S}_j(\mathrm{X})\mathrm{P}^{\frac{q-1}{2}}(\mathrm{X}))\mathrm{X}^{jq}$$

donde $\mathrm{R}_j, \mathrm{S}_j$ son polinomios de $\mathbb{F}_q[\mathrm{X}]$, de grados $\deg(\mathrm{R}_j)$, $\deg(\mathrm{S}_j)$ respectivamente, menores estrictamente que $\frac{q-1}{2} - d$ (esta cantidad es positiva por construcción). La cantidad J (no necesariamente entera) la tomaremos más adelante, de forma que cumpla las propiedades especificadas en la proposición.

Podemos ver que el grado del polinomio Q construido de esta manera está acotado. Por construcción:

$$\begin{aligned}\deg(\mathrm{Q}) &\leq ld + \frac{q-1}{2}d + \frac{q-1}{2} - d + Jq \\ &\leq \frac{5q-12}{8}d + \frac{q-1}{2} + Jq\end{aligned}$$

Como $d \geq 3$ y $q > 4d^2$, la cantidad $\frac{5q-12}{8}d + \frac{q-1}{2}$ es menor o igual que qd y deducimos que:

$$\deg(\mathrm{Q}) \leq (d+J)q.$$

De esta forma, observamos que la elección de J nos permite controlar el grado del polinomio Q. Gracias al lema que enunciamos a continuación, veremos cuando el polinomio Q que construimos es no nulo.

Lema 3.3. $\mathrm{Q} \in \mathbb{F}_q[\mathrm{X}]$ es el polinomio nulo si, y sólo si, para todo j los polinomios $\mathrm{R}_j, \mathrm{S}_j$ son idénticamente nulos.

Demostración. Es claro que si para cada j, los polinomios $\mathrm{R}_j, \mathrm{S}_j$ son nulos, entonces Q debe ser necesariamente nulo. Veamos la otra implicación.

Podemos suponer sin pérdida de generalidad que el polinomio P no se anula en $0 \in \mathbb{F}_q$. En caso contrario, simplemente realizaríamos un cambio lineal de coordenadas. Razonemos por reducción al absurdo, y supongamos que Q es el polinomio nulo, sin que todos los $\mathrm{R}_j, \mathrm{S}_j$ sean idénticamente nulos. Sea k el menor de los índices tales que unos de los dos polinomios $\mathrm{R}_k, \mathrm{S}_k$ es no nulo. Por construcción, podemos dividir Q por $\mathrm{P}^l\mathrm{X}^{kq}$ y obtenemos:

$$\sum_{k \leq j < J} (\mathrm{R}_j(\mathrm{X}) + \mathrm{S}_j(\mathrm{X})\mathrm{P}^{\frac{q-1}{2}}(\mathrm{X}))\mathrm{X}^{(j-k)q} = 0.$$

Podemos reescribir esta igualdad, $\mathrm{T}_1(\mathrm{X}) + \mathrm{T}_2(\mathrm{X})\mathrm{P}^{\frac{q-1}{2}}(\mathrm{X}) = 0$, definiendo:

$$\mathrm{T}_1(\mathrm{X}) = \sum_{k \leq j < J} \mathrm{R}_j(\mathrm{X})\mathrm{X}^{(j-k)q}, \quad \mathrm{T}_2(\mathrm{X}) = \sum_{k \leq j < J} \mathrm{S}_j(\mathrm{X})\mathrm{X}^{(j-k)q}.$$

De la igualdad, tomando cuadrados y multiplicando por el polinomio P, resulta que $\mathrm{T}_1^2\mathrm{P} = \mathrm{T}_2^2\mathrm{P}^q$. Como $\mathrm{P} \in \mathbb{F}_q[\mathrm{X}]$, se comprueba que:

$$\mathrm{P}^q(\mathrm{X}) = \mathrm{P}(\mathrm{X}^q) \equiv \mathrm{P}(0) \pmod{\mathrm{X}^q}.$$

Deducimos entonces que $\mathrm{R}_j^2\mathrm{P} \equiv \mathrm{S}_j^2\mathrm{P}(0) \pmod{\mathrm{X}^q}$. Los grados de los polinomios R_j y S_j cumplen las siguientes desigualdades:

$$2\deg(\mathrm{R}_j) + d \leq 2(\frac{q-1}{2} - d) + d < q.$$

$$2\deg(\mathrm{S}_j) < 2(\frac{q-1}{2} - d) < q.$$

Se debe dar la igualdad $\mathrm{R}_j^2\mathrm{P} = \mathrm{S}_j^2\mathrm{P}(0)$ y llegamos a una contradicción. Esto va en contra de la definición del polinomio P, que no es un cuadrado en $\overline{\mathbb{F}_q}[\mathrm{X}]$ y podemos concluir.

□

Para garantizar que el polinomio Q se anule con el orden necesario, vamos a tener que examinar sus derivadas de Hasse. Para ello, enunciamos el lema que viene a continuación.

Lema 3.4. Sea k un entero natural menor o igual que l. Entonces existen polinomios $\mathrm{R}_j^{(k)}, \mathrm{S}_j^{(k)} \in \mathbb{F}_q[\mathrm{X}]$ de grados menores o iguales que $\frac{q-1}{2} - d + k(d-1)$ y tales que:

$$\mathrm{D}^k\mathrm{Q}(\mathrm{X}) = \mathrm{P}^{l-k}(\mathrm{X}) \sum_{0 \leq j < J} (\mathrm{R}_j^{(k)}(\mathrm{X}) + \mathrm{S}_j^{(k)}(\mathrm{X})\mathrm{P}^{\frac{q-1}{2}}(\mathrm{X}))\mathrm{X}^{jq}.$$

Demostración. Damos una escritura del polinomio Q que nos permita trabajar con las derivadas de Hasse. Tenemos $\mathrm{Q}(\mathrm{X}) = \mathrm{H}(\mathrm{X}, \mathrm{X}^q)$, donde H el polinomio de $\mathbb{F}_q[\mathrm{X}, \mathrm{Y}]$, dado por:

$$\mathrm{H}(\mathrm{X}, \mathrm{Y}) = \mathrm{P}^l(\mathrm{X}) \sum_{0 \leq j < J} (\mathrm{R}_j(\mathrm{X}) + \mathrm{S}_j(\mathrm{X})\mathrm{P}^{\frac{q-1}{2}}(\mathrm{X}))\mathrm{Y}^j.$$

De forma general, si $r < q$, la r-ésima derivada de Hasse de un polinomio $\mathrm{Q} \in \mathbb{F}_q[\mathrm{X}]$ de la forma de $\mathrm{Q}(\mathrm{X}) = \mathrm{H}(\mathrm{X}, \mathrm{X}^q)$, como el indicado, viene dada por la fórmula:

$$\mathrm{D}^r\mathrm{Q}(\mathrm{X}) = \mathrm{D}^{r,0}\mathrm{H}(\mathrm{X}, \mathrm{X}^q).$$

Esto se prueba fácilmente, viendo que la fórmula funciona para monomios y extendiendo por linealidad a polinomios cualesquiera. Sea un monomio $\mathrm{H}(\mathrm{X}, \mathrm{Y}) = \mathrm{X}^n\mathrm{Y}^m$, utilizando las propiedades básicas de la derivada de Hasse que vimos en el primer capítulo, veamos que se cumple la fórmula deseada:

$$\mathrm{D}^r(\mathrm{X}^{n+mq}) = \sum_{i=0}^{r} \mathrm{D}^{r-i}(\mathrm{X}^n)\mathrm{D}^i(\mathrm{X}^{mq})$$

Basta observar que $\mathrm{D}^i(\mathrm{X}^{mq}) = 0$ para $0 < i < q$, y esto ocurre ya que por definición:

$$\mathrm{D}^i(\mathrm{X}^{mq}) = \binom{mq}{i}\mathrm{X}^{mq-i} = \frac{mq}{i}\binom{mq-1}{i-1}\mathrm{X}^{mq-i} = 0.$$

Volviendo al caso que estamos tratando en esta demostración, podemos calcular la derivada de Hasse utilizando lo que acabamos de ver:

$$\mathrm{D}^k\mathrm{Q}(\mathrm{X}) = \mathrm{D}^{k,0}\mathrm{H}(\mathrm{X}, \mathrm{X}^q) = \sum_{0\leq j<J} (\mathrm{D}^k(\mathrm{P}^l(\mathrm{X})\mathrm{R}_j(\mathrm{X}))+\mathrm{D}^k(\mathrm{P}^{l+\frac{q-1}{2}}(\mathrm{X})\mathrm{S}_j(\mathrm{X})))\mathrm{X}^{jq}.$$

Vamos a examinar los dos términos que aparecen en los sumandos, para ver que efectivamente se tiene la fórmula enunciada en el lema:

$$\mathrm{D}^k(\mathrm{P}^l(\mathrm{X})\mathrm{R}_j(\mathrm{X})) = \sum_{i=0}^{k} \mathrm{D}^i\mathrm{P}^l(\mathrm{X})\ \mathrm{D}^{k-i}\mathrm{R}_j(\mathrm{X}) = \mathrm{P}^{l-k}(\mathrm{X})\mathrm{R}_j^{(k)}(\mathrm{X}).$$

Lo único que hemos hecho es factorizar el polinomio P^{l-k} que aparece en todos los sumandos. El polinomio $\mathrm{R}_j^{(k)}$ no es mas que el resultado de esta factorización. Además, por construcción el grado de $\mathrm{R}_j^{(k)}$, denotado con $\deg(\mathrm{R}_j^{(k)})$, está acotado:

$$\deg(\mathrm{R}_j^{(k)}) \leq \deg(\mathrm{R}_j) + kd - k.$$

Esta cota se obtiene observando que el grado de $\mathrm{P}^{l-k}(\mathrm{X})\mathrm{R}_j^{(k)}(\mathrm{X})$ está acotado por el de $\mathrm{P}^l(\mathrm{X})\mathrm{R}_j(\mathrm{X})$ menos k, por la derivación. Eliminando la contribución de $\mathrm{P}^{l-k}(\mathrm{X})$, obtenemos la cota. Utilizando la cota que ya teníamos para $\deg(\mathrm{R}_j)$, obtenemos:

$$\deg(\mathrm{R}_j^{(k)}) \leq \frac{q-1}{2} - d + kd - k.$$

Se razona de forma similar para el término correspondiente a S_j que aparece en los sumandos y se puede concluir.

□

Queremos que el polinomio Q tenga ceros de orden mayor o igual que l en todos los puntos de $S(\lambda)$. Sea $x \in S(\lambda) \subset \mathbb{F}_q$, es claro que si $\mathrm{P}(x) = 0$ entonces por definición de Q se cumple lo que buscamos. Supongamos que P no se anula en $x \in S(\lambda)$. Aplicando el lema 3.4, podemos evaluar una derivada de Hasse de orden k en el punto x:

$$\mathrm{D}^k\mathrm{Q}(x) = \mathrm{P}^{l-k}(x) \sum_{0\leq j<J} (\mathrm{R}_j^{(k)}(x) + \mathrm{S}_j^{(k)}(x)\mathrm{P}^{\frac{q-1}{2}}(x))x^{jq}.$$

El punto x por definición cumple $\mathrm{P}^{\frac{q-1}{2}}(x) = \lambda$. Además como trabajamos en característica p, x coincide con x^q y se verifica que:

$$\mathrm{D}^k\mathrm{Q}(x) = \mathrm{P}^{l-k}(x) \sum_{0\leq j<J} (\mathrm{R}_j^{(k)}(x) + \lambda\mathrm{S}_j^{(k)}(x))x^j = \mathrm{P}^{l-k}(x)\sigma^{(k)}(x).$$

Donde $\sigma^{(k)}(x)$ es el polinomio de $\mathbb{F}_q[\mathrm{X}]$ dado por:

$$\sigma^{(k)}(\mathrm{X}) = \sum_{0 \leq j < J} (\mathrm{R}_j^{(k)}(\mathrm{X}) + \lambda \mathrm{S}_j^{(k)}(\mathrm{X}))\mathrm{X}^j.$$

Observamos que si conseguimos que $\sigma^{(k)}(x)$ sea 0 para cada $x \in S(\lambda)$ y cada $k < l$, obtendremos que Q tiene un cero de orden mayor o igual que l en cada punto $x \in S(\lambda)$. Llegados a este punto, podemos probar la proposición 3.2. Podemos estudiar el sistema de ecuaciones dado por:

$$\sigma^{(k)} = 0 \text{ para cada } k < l.$$

Los coeficientes del polinomio $\sigma^{(k)}$ son combinaciones lineales de los coeficientes de los polinomios $\mathrm{R}_j^{(k)}, \mathrm{S}_j^{(k)}$.Se trata de un sistema de ecuaciones lineales homogéneo, donde las incógnitas son los coeficientes de $\mathrm{R}_j^{(k)}$ y $\mathrm{S}_j^{(k)}$. Hay tantas ecuaciones como coeficientes tienen todos los polinomios $\sigma^{(k)}$, para $k < l$. Observamos que el grado de $\sigma^{(k)}$, denotado con $\deg(\sigma^{(k)})$ está acotado:

$$\deg(\sigma^{(k)}) < \frac{q-1}{2} - d + k(d-1) + J$$

Recordando que la suma de los primeros m enteros naturales vale $\frac{m(m-1)}{2}$, vemos que el número de ecuaciones es estrictamente menor que:

$$\mu_1 = l(\frac{q-1}{2} - d + J) + \frac{l(l-1)(d-1)}{2}$$

Por un razonamiento de la misma naturaleza, podemos obtener que al menos hay $\mu_2 = (q-1-2d)J$ incógnitas. Tomando J suficientemente grande, podemos hacer que $\mu_2 > \mu_1$. En ese caso, el sistema tiene una solución no trivial, por lo que gracias al lema 3.3, podemos garantizar que este proceso esta construyendo un polinomio no nulo.

Tomando $J = \frac{l}{q}(\frac{q-1}{2} + 2d(l-1))$, se cumple $\mu_2 > \mu_1$. Además, se verifica la condición sobre el grado de la proposición 3.2:

$$\begin{aligned}
\deg(\mathrm{Q}) &\leq (J+d)q \\
&\leq (\frac{l}{q}(\frac{q-1}{2} + 2dl(l-1)) + d)q \\
&\leq \frac{l(q-1)}{2} + 2dl(l-1) + dq
\end{aligned}$$

De esta forma queda probada la proposición y podemos probar ahora el teorema 3.1.

Nos situamos en las condiciones del teorema. Recordemos que teníamos $q > 4d^2$ y como $d > 2$, se cumple $q > 8d$. Sea $\lambda \in \mathbb{F}_q$. Podemos aplicar en estas condiciones la proposición 3.2. Para cada entero l con $d < l \leq \frac{q}{8}$, sabemos que existe un polinomio auxiliar Q que se anula con orden mayor o igual que l en todos los puntos de $S(\lambda)$, para $\lambda \in \mathbb{F}_q$. De esta forma, podemos asegurar que:

$$l\,|S(\lambda)| \leq deg(\mathrm{Q}) \leq \frac{l(q-1)}{2} + 2dl(l-1) + dq.$$

Dividiendo por $l \neq 0$, deducimos:

$$|S(\lambda)| \leq \frac{(q-1)}{2} + 2d(l-1) + \frac{dq}{l}.$$

Tomamos $l = 1 + \left\lfloor \frac{\sqrt{q}}{2} \right\rfloor$, donde $\lfloor x \rfloor = \max\{k \in \mathbb{Z} : q \leq \frac{\sqrt{q}}{2}\}$. Esta elección cumple las cotas $d < l < \frac{q}{8}$ necesarias para aplicar la proposición, ya que por ejemplo para la primera cota, sabemos que $q > 4d^2$ y por tanto $\frac{\sqrt{q}}{2} > d$ y $l > d$. Llevando este valor de l, a la cota obtenida para $|S(\lambda)|$:

$$\begin{aligned} |S(\lambda)| &\leq \frac{q-1}{2} + 2d(1 - \frac{\sqrt{q}}{2} - 1) + \frac{2dq}{\sqrt{q}} \\ &\leq \frac{q-1}{2} + d\sqrt{q} + 2d\sqrt{q} \\ &< \frac{q-1}{2} + 4d\sqrt{q} \end{aligned}$$

Si nos fijamos en el caso $\lambda = 1$, obtenemos la cota:

$$N' + N'' = |S(1)| < \frac{q-1}{2} + 4d\sqrt{q}.$$

De esta cota, podemos deducir para N:

$$N = N' + 2N'' < 2(N' + N'') < q + 8d\sqrt{q}.$$

Para probar el teorema 3.1 solo nos falta probar la desigualdad contraria. En este punto, es útil expresar el polinomio $\mathrm{X}^q - \mathrm{X}$ como producto de tres polinomios, es decir, $\mathrm{X}^q - \mathrm{X}(\mathrm{X}^{\frac{q-1}{2}} - 1)(\mathrm{X}^{\frac{q-1}{2}} + 1)$ (recordando que $\frac{q-1}{2}$ es un entero). Si cambiamos X por P(X), se verifica:

$$\mathrm{P}^q(\mathrm{X}) - \mathrm{P}(\mathrm{X}) = \mathrm{P}(\mathrm{X})(\mathrm{P}^{\frac{q-1}{2}}(\mathrm{X}) - 1)(\mathrm{P}^{\frac{q-1}{2}}(\mathrm{X}) + 1).$$

Evaluando en $x \in \mathbb{F}_q$, como cada elemento $y \in \mathbb{F}_q$ cumple $y^q - y = 0$, tenemos:

$$\mathrm{P}^q(x) - \mathrm{P}(x) = \mathrm{P}(x)(\mathrm{P}^{\frac{q-1}{2}}(x) - 1)(\mathrm{P}^{\frac{q-1}{2}}(x) + 1) = 0.$$

Como $\mathbb{F}_q$ tiene q elementos, sabemos entonces que hay exactamente q elementos que satisfacen la ecuación anterior. Además, por integridad, el producto de los tres polinomios se anula, únicamente cuando uno de los factores se anula. Si denotamos con N''' al cardinal del conjunto $\{x \in \mathbb{F}_q : \mathrm{P}^{\frac{q-1}{2}}(\mathrm{x}) = -1\}$, deducimos que:

$$N' + N'' + N''' = q.$$

Sea $\lambda = -1$. Por definición del conjunto $S(-1)$ y aplicando la cota obtenida para $S(\lambda)$, se tiene:

$$N' + N''' = |S(-1)| < \frac{q}{2} + 4d\sqrt{q}.$$

Utilizando esta cota y la igualdad anterior, se deduce que:

$$N'' = q - N' - N''' > \frac{q}{2} - 4d\sqrt{q}.$$

De esta última cota, obtenemos:

$$N = N' + 2N'' \geq 2N'' > q - 8d\sqrt{q}.$$

Finalmente, gracias a las dos cotas que satisface N, concluimos la demostración del teorema. Es interesante observar que pese a ser elemental, la demostración del teorema no es inmediata y no deja de ser bastante técnica. El método de Stepanov puede emplearse para atacar el mismo problema, planteado para curvas más generales. En tales casos, se requieren resultados adicionales de geometría algebraica más avanzados, como el teorema de Riemann-Roch. Es interesante recalcar que filosofía de las pruebas sigue siendo la misma. En casos mas generales, lo que se construye no es un polinomio pero si una función racional, es decir, un cociente de polinomios y lo que se acota no es el grado, sino el número de polos de la función. Como una referencia para esto último se tiene [Bom]. Destacar que las cotas obtenidas son del mismo orden que la que hemos probado en este capítulo.

Es interesante destacar también que todas estas cotas se deducen de las conjeturas de Weil, que tratan el caso más general de las funciones zeta de Weil, asociadas a variedades algebraicas proyectivas no singulares definidas sobre cuerpos finitos.

Desde luego, las pruebas para el caso general requieren muchos más resultados de geometría algebraica, pero son asimismo interesantes. Entre las referencias sobre estas funciones zeta, sus propiedades y los resultados que se derivan de ellas, se encuentran la del apéndice C del libro de R. Hartshorne [Hart] y la de la laudatio de N. Katz [Katz] en el ICM de Helsinki

Capítulo 4

Teorema de Szemerédi-Trotter

En este capítulo, vamos a ilustrar como se pueden emplear técnicas polinomiales para probar resultados combinatorios en $\mathbb{R}^n$. Para ello vamos a probar el teorema de Szmeredi-Trotter en $\mathbb{R}^2$. Se trata de un teorema que tiene que ver con la combinatoria resultante de trabajar con un conjunto finito de puntos y rectas del plano. Será de especial interés el resaltar la importancia de los razonamientos topológicos en $\mathbb{R}^n$ y como estos nos permiten afinar las cotas que se pueden obtener de forma general, en el caso de que el cuerpo base no sea necesariamente $\mathbb{R}$. Para la realización de la prueba del teorema de Szmeredi-Trotter, nos apoyaremos en el artículo [GK]. La realización de esta prueba requiere del empleo del teorema de Borsuk-Ulam y de uno de sus corolarios, el teorema del sándwich de jamón, que veremos en la primera parte del capítulo. Para esta primera parte, una referencia interesante es [Mat].

Durante todo el capítulo, el marco general de trabajo será $\mathbb{R}^n$, donde n es un entero natural.

Los razonamientos que vamos a emplear en este capítulo se basan en el teorema de Borsuk-Ulam. Se trata de un teorema topológico con aplicaciones significativas en combinatoria. Enunciamos a continuación una de las versiones del teorema de Borsuk-Ulam. Se trata de un teorema clásico que puede encontrarse en muchos textos de topología algebraica. Podemos encontrar una prueba y una exposición de algunas de sus aplicaciones en combinatoria en [Mat]. Denotamos con S^n a la esfera unidad de $\mathbb{R}^{n+1}$, con respecto de la norma euclídea.

Teorema 4.1 (Borsuk-Ulam)**.** Para cada aplicación continua $f : S^n \to \mathbb{R}^n$, existe un punto $x \in S^n$ tal que $f(x) = f(-x)$.

Es interesante observar que el teorema sigue siendo cierto cuando consideramos una aplicación continua $f : S^m \to \mathbb{R}^n$ con $m \geq n$, ya que siempre podemos restringir f a una esfera de dimensión inferior.

Vamos a probar ahora el llamado teorema del sándwich de jamón, que se obtiene como consecuencia del teorema de Borsuk-Ulam. De nuevo se trata de un teorema del que podemos encontrar varias versiones. Probaremos primero la variante del teorema para medidas de Borel. De este resultado se deduce de forma inmediata la forma más clásica del teorema, que se aplica a conjuntos abiertos y acotados.

Teorema 4.2. Sean $\mu_1, ..., \mu_n$ medidas de Borel finitas en $\mathbb{R}^n$ tales que cada hiperplano tiene medida 0, para cada medida. Cada hiperplano H divide $\mathbb{R}^n$ en dos componentes conexas denotadas con $\mathrm{H}^1, \mathrm{H}^2$, que llamamos semiespacios. Entonces, existe un hiperplano H de $\mathbb{R}^n$ que cumple:

$$\mu_i(\mathrm{H}^1) = \mu_i(\mathrm{H}^2) = \frac{\mu_i(\mathbb{R}^n)}{2}, \text{ para cada } i = 1, ..., n.$$

Demostración. Sea $u = (u_0, ..., u_n)$ un punto de $S^n \subset \mathbb{R}^{n+1}$. Si al menos una de las últimas n componentes de u es no nula, se puede definir un hiperplano de $\mathbb{R}^n$ gracias a u. Este hiperplano nos da dos semiespacios, que denotamos $\mathrm{H}^1(u), \mathrm{H}^2(u)$. Explícitamente, tenemos por ejemplo:

$$\mathrm{H}^1(u) = \{(x_1, ..., x_n) \in \mathbb{R}^n : u_1x_1 + ... + u_nx_n \leq u_0\}.$$

El otro semiespacio de $\mathbb{R}^n$ se define logicamente cambiando de sentido la desigualdad. De esta forma, podemos asignar a cada punto de $u \in S^n$ un semiespacio de $\mathbb{R}^n$. Siempre que tenemos un punto $u \in S^n$, podemos considerar el punto antipodal asociado, $-u \in S^n$. Por construcción, al asignar un semiespacio a un punto $u \in S^n$, estamos asignando el semiespacio complementario al punto antipodal $-u \in S^n$. En el caso extremo donde consideramos un punto $u \in S^n$, donde las últimas n coordenadas son nulas, se satisface $u_0 = \pm 1$. Además, en tal caso se tiene:

$$\mathrm{H}^1(1, 0, ...0) = \mathbb{R}^n \text{ y } \mathrm{H}^1(1, 0, ...0) = \emptyset.$$

Podemos definir unas función $f : S^n \to \mathbb{R}^n$, dando valor a sus componentes denotadas con f_i. Para $i = 1, ..., n$, definimos $f_i(u) = \mu_i(\mathrm{H}^1(u))$. Veamos que f es una aplicación continua, comprobando que sus componentes son continuas.
Sea $\{u_m\}_{m=0}^{\infty}$ una sucesión de puntos de S^n que converge hacia $u \in S^n$. Queremos ver que la sucesión $\{f_i(u_m)\}_{m=0}^{\infty}$ converge hacia $f_i(u)$, para cada

$i = 1, ..., n$. Consideramos las funciones características de los conjuntos $\mathrm{H}^1(u)$ y $\mathrm{H}^1(u_m)$, que denotamos con χ_u y χ_{u_m} respectivamente. Es sencillo comprobar que si x no es un punto de la frontera de $\mathrm{H}^1(u)$, entonces para m suficientemente grande, $x \in \mathrm{H}^1(u_m)$ si, y sólo si $x \in \mathrm{H}^1(u)$. De esta forma tenemos $\chi_{u_m}(x) \to \chi_u(x)$, para cada x que no esté en la frontera de $\mathrm{H}^1(u)$. Como por construcción, la frontera de $\mathrm{H}^1(u)$ es un conjunto de medida nula (para cada μ_i), sabemos que χ_{u_m} converge hacia χ_u en casi todo punto. La aplicación del teorema de la convergencia dominada de Lebesgue, permite concluir que $f_i(u_m)$ converge hacia $f_i(u)$. Por tanto, las f_i son aplicaciones continuas y f también es una aplicación continua. La aplicación del teorema 4.1 permite concluir que existe un punto $u \in S^n$ tal que $f(u) = f(-u)$. Como el punto u no puede ser ni $(1, 0, ..., 0)$ ni $(-1, 0, ..., 0)$, el hiperplano que buscamos es la frontera de $\mathrm{H}^1(u)$.

□

Al trabajar en $\mathbb{R}^n$, podemos considerar la medida de Lebesgue, que denotamos con λ. Recordemos que los hiperplanos de $\mathbb{R}^n$ son conjuntos de medida nula.

Teorema 4.3 (del sándwich de jamón)**.** Sean $U_1, ..., U_n$ subconjuntos abiertos y acotados de $\mathbb{R}^n$. Existe un hiperplano H de $\mathbb{R}^n$ que corta a todos los U_i y divide cada U_i en dos subconjuntos, ambos con la misma medida de Lebesgue.

Demostración. Recordemos primero que los conjuntos U_i tienen medida de Lebesgue finita. Se obtiene el resultado de forma inmediata al aplicar el teorema 4.2 a las medidas de Borel dadas por:

$$\mu_i(X) = \lambda(X \cap U_i), \text{ para cada } X \in \mathbb{R}^n \text{ de Borel y cada } i = 1, ..., n.$$

□

Existen más variantes del teorema que acabamos de enunciar. Vemos ahora otra variante interesante del teorema del sándwich de jamón, donde reemplazamos el hiperplano por una hipersuperficie. Se trata de una generalización del teorema anterior, donde solo estamos considerando los ceros de polinomios de grado 1.

Teorema 4.4. Sea d un entero mayor o igual que 0. Sean $U_1, ..., U_m$ subconjuntos abiertos y acotados de $\mathbb{R}^n$, donde m es un entero estrictamente menor que $\binom{d+n}{n}$. Entonces, existe un polinomio $\mathrm{P} \in \mathbb{R}[\mathrm{X}_1, ..., \mathrm{X}_n]$ de grado menor o igual que d y tal que $V(\mathrm{P})$ divide cada U_i en dos subconjuntos, ambos con la misma medida de Lebesgue.

Demostración. La demostración de este resultado es análoga a la realizada para el teorema 4.2. En ese caso identificábamos, de forma implícita, $\mathbb{R}^{n+1}$ con el espacio vectorial real formado por los polinomios con coeficientes reales en n variables, de grado menor o igual que 1. Así podíamos identificar los puntos de S^n con polinomios, que nos definen hiperplanos de $\mathbb{R}^n$. Definíamos entonces una aplicación continua. Utilizando el teorema de Borsuk-Ulam, podíamos finalmente quedarnos con un polinomio en concreto y por tanto con un hiperplano.

Para este resultado se procede de la misma forma. El espacio vectorial real V formado por los polinomios con coeficientes reales en n variables y de grado menor o igual que d tiene dimensión $s = \binom{d+n}{n}$. Podemos identificar, de forma no única, V con $\mathbb{R}^s$. De esta forma, se pueden ver de nuevo los puntos de $S^{s-1} \subset \mathbb{R}^s$ como polinomios no nulos de grado menor o igual que d. Para cada polinomio $\mathrm{P} \in S^{s-1}$, se definen de nuevo dos semiespacios $\mathrm{H}^1(\mathrm{P}), \mathrm{H}^2(\mathrm{P})$, tales que:

$$\mathrm{H}^1(\mathrm{P}) = \{(x_1, ..., x_n) \in \mathbb{R}^n : \mathrm{P}(x_1, ..., x_n) \leq 0\}.$$
$$\mathrm{H}^2(\mathrm{P}) = \{(x_1, ..., x_n) \in \mathbb{R}^n : \mathrm{P}(x_1, ..., x_n) \geq 0\}.$$

Recordemos que $s - 1 \geq m$. La construcción de la función $f : S^{s-1} \to \mathbb{R}^m$ así como la comprobación de la continuidad se realizan de la misma forma que en la demostración. La frontera que se obtiene tras aplicar el teorema de Borsuk-Ulam nos proporciona una hipersuperficie con las propiedades requeridas.

□

En combinatoria es interesante trabajar con conjuntos finitos. Veamos ahora una última variante del teorema del sándwich de jamón para conjuntos finitos, que se obtiene como corolario del último teorema enunciado.

Corolario 4.5. Sea d un entero mayor o igual que 0. Sean $S_1, ..., S_m$ subconjuntos finitos de $\mathbb{R}^n$, donde m es un entero estrictamente menor que $\binom{d+n}{n}$. Entonces, existe un polinomio $\mathrm{P} \in \mathbb{R}[\mathrm{X}_1, ..., \mathrm{X}_n]$ de grado menor o igual que d y tal que, para cada $i = 1, ..., m$, las intersecciones de S_i con las regiones $\{x \in \mathbb{R}^n : \mathrm{P}(x) > 0\}$ y $\{x \in \mathbb{R}^n : \mathrm{P}(x) < 0\}$ tienen cardinal menor o igual que $\frac{|S_i|}{2}$.

Demostración. La prueba consiste en cambiar los puntos por bolas abiertas de radio fijo, aplicar el teorema del sándwich de jamón a las uniones de estas bolas y hacer tender los radios a 0.

Sean $\varepsilon > 0$ y $s = \binom{d+n}{n}$. Para cada S_i con $i = 1, .., n$, consideramos en cada punto $x \in S_i$ la bola abierta de centro x y radio ε. Llamamos $U_{i,\varepsilon}$ a la unión

de estas bolas abiertas. Tenemos de esta forma m abiertos acotados a los que podemos aplicar el teorema 4.4. Deducimos que existe $\mathrm{P}_\varepsilon \in \mathbb{R}[\mathrm{X}_1, ..., \mathrm{X}_n]$, de grado menor o igual que d y tal que $V(\mathrm{P}_\varepsilon)$ divide cada S_i en dos conjuntos con la misma medida de Lebesgue.

De nuevo como en las pruebas realizadas de las distintas variantes del teorema del sándwich de jamón, podemos ver los polinomios P_ε como elementos de S^{s-1}. Si tomamos una sucesión $\{\varepsilon_t\}_{t=0}^{\infty}$ que tiende a 0, tenemos que $\mathrm{P}_{\varepsilon_t}$ converge hacia un polinomio $\mathrm{P} \in S^{s-1}$. Los coeficientes de $\mathrm{P}_{\varepsilon_t}$ convergen hacia los coeficientes de P y es sencillo comprobar que $\mathrm{P}_{\varepsilon_t}$ converge uniformemente a P en los compactos. El polinomio P tiene el grado adecuado y solo falta comprobar que cumple la condición de los cardinales.

Razonamos por reducción al absurdo. Supongamos que existe j tal que por ejemplo $\mathrm{P} > 0$ en un número de puntos de S_j mayor estrictamente que $\frac{|S_j|}{2}$. La otra suposición $\mathrm{P} < 0$, es similar. Sea $S_j^+ \subset S_j$ el conjunto formado por los puntos de S_j donde $\mathrm{P} > 0$. Sea $\delta > 0$. Tomando δ suficientemente pequeño, podemos suponer que $\mathrm{P} > \delta$ en todas las bolas abiertas de radio δ, centradas en los puntos de S_j^+. Además, podemos tomar δ suficientemente pequeño para que todas las bolas abiertas de radio δ centradas en los puntos de S_j sean disjuntas. Como los $\mathrm{P}_{\varepsilon_t}$ convergen de forma uniforme hacia P en los compactos, podemos encontrar N suficientemente grande tal que $\mathrm{P}_{\varepsilon_N} > 0$ en toda bola de radio δ centrada en cada punto de S_j^+. Tomando N suficientemente grande también podemos garantizar que $\varepsilon_N < \delta$. De esta forma obtenemos que $\mathrm{P}_{\varepsilon_N} > 0$ en más de la mitad de U_{j,ε_N} (en términos de la medida) y llegamos a contradicción.

□

Vamos a ver con la siguiente proposición como la aplicación sucesiva de este corolario nos permite estudiar conjuntos finitos de puntos en $\mathbb{R}^n$, gracias al empleo de polinomios de grado controlado. Dado un subconjunto finito de $\mathbb{R}^n$, el teorema del sándwich de jamón nos va a permitir construir un polinomio P, cuyos ceros descomponen $\mathbb{R}^n$ en una serie de regiones abiertas, que llamaremos células. Estas células no serán otra cosa que las componentes conexas de $\mathbb{R}^n \setminus V(\mathrm{P})$. Además, podremos controlar cuantos elementos de S se encuentran a lo sumo en cada célula. Destacar finalmente que, por construcción, las fronteras de las células estarán contenidas en $V(\mathrm{P})$. Esta construcción que se debe a Guth y Katz, puede consultarse en [GK].

Proposición 4.6 (Descomposición celular)**.** Sean $r \geq 1$ y S un subconjunto finito de $\mathbb{R}^n$. Entonces, existe un polinomio $\mathrm{P} \in \mathbb{R}[\mathrm{X}_1, ..., \mathrm{X}_n]$, de grado

$d = O(r)$ tal que $V(\mathrm{P})$ descompone $\mathbb{R}^n$ en células como las descritas anteriormente. Además, cada célula contiene a lo sumo $|S|\, r^{-n}$ puntos de S.

Demostración. La idea de la demostración es aplicar de forma inductiva el corolario 4.5 un número finito de veces para construir el polinomio con las propiedades deseadas. Vamos a detallar la demostración:

Como en un principio solo tenemos un único conjunto finito $S \in \mathbb{R}^n$, aplicando el corolario 4.5 podemos garantizar que existe un polinomio $\mathrm{P}_1 \in \mathbb{R}[\mathrm{X}_1, ..., \mathrm{X}_n]$ de grado $d_1 = O(2^{\frac{1}{n}})$, de forma que $S^0 = S \cap \{x \in \mathbb{R}^n : \mathrm{P}_1(x) > 0\}$ y $S^1 = S \cap \{x \in \mathbb{R}^n : \mathrm{P}_1(x) < 0\}$ tienen a lo sumo $\frac{|S|}{2}$ puntos de S. Recordemos que los puntos de S pueden encontrarse en $V(\mathrm{P}_1)$. Los conjuntos S^0 y S^1 están contenidos en diferentes componentes conexas de $\mathbb{R}^n \backslash V(\mathrm{P}_1)$.

Aplicando de nuevo el corolario, esta vez a los dos conjuntos que hemos obtenido en el primer paso, construimos un polinomio P_1 de grado $d_2 = O(2^{\frac{2}{n}})$. Este polinomio nos proporciona $2^2 = 4$ células, que intersecamos con el conjunto S:

$$\begin{cases} S^{(0,0)} = S \cap \{x \in \mathbb{R}^n : \mathrm{P}_1(x) > 0, \mathrm{P}_2(x) > 0\} \\ S^{(1,0)} = S \cap \{x \in \mathbb{R}^n : \mathrm{P}_1(x) < 0, \mathrm{P}_2(x) > 0\} \\ S^{(0,1)} = S \cap \{x \in \mathbb{R}^n : \mathrm{P}_1(x) > 0, \mathrm{P}_2(x) < 0\} \\ S^{(1,1)} = S \cap \{x \in \mathbb{R}^n : \mathrm{P}_1(x) < 0, \mathrm{P}_2(x) < 0\} \end{cases}$$

Estos 4 conjuntos finitos son a los que se aplica el corolario en la siguiente iteración. El conjunto de ceros que sirve de frontera para las células, en este caso, no es otro que $V(\mathrm{P}_1\mathrm{P}_2)$. Por construcción, estas células contienen a lo sumo $|S|\, 2^{-2}$ puntos de S.

En el paso k de la inducción, construimos gracias al corolario un polinomio P_k de grado $d_k = O(2^{\frac{k}{n}})$. Este polinomio nos proporciona de nuevo una descomposición de $\mathbb{R}^n$, con 2^k células. Los conjuntos finitos que obtenemos vienen dados por S^z, con $z = (z_1, ..., z_k) \in \{0, 1\}^k$. S^z no es más que la intersección de S con la célula dada por los polinomios $\mathrm{P}_1, ..., \mathrm{P}_k$. Los signos de las desigualdades vienen dados por z, cuando en la posición s aparece 0, entonces consideramos la desigualdad $\mathrm{P}_s(x) > 0$, para $x \in \mathbb{R}^n$. En el caso de tener un 1, obviamente consideramos la desigualdad contraria. Recopilando esas k desigualdades se construyen las células. Por construcción, cada célula contiene a lo sumo $|S|\, 2^{-k}$ puntos de S.

Dado $r \geq 1$, podemos fijar k de forma que $2^k \geq r^n > 2^{k-1}$. Tomando $\mathrm{P} = \mathrm{P}_1 ... \mathrm{P}_k$, tenemos un polinomio con las propiedades necesarias. Por ejemplo, cada célula tiene a lo sumo $|S|\, 2^{-k} \leq |S|\, r^{-n}$ puntos de S. Solamente falta por comprobar que el grado de P es el adecuado. Si d es el grado de P, se verifica:

$$d = \sum_{i=1}^{k} d_i \lesssim \sum_{i=1}^{k} 2^{\frac{i}{n}} \leq 2^{\frac{k+1}{n}} < (4r^n)^{\frac{1}{n}} \lesssim r.$$

Esto permite concluir. En la última expresión, por comodidad, hemos utilizado otra notación clásica $\lesssim$ que es equivalente a la que veníamos usando $O()$. Se tiene $A \lesssim B$ si, y sólo si, existe una constante $C > 0$ tal que $A \leq CB$, es decir, $A = O(B)$. Utilizaremos en lo que sigue esta expresión para aligerar la notación cuando trabajemos con desigualdades.

□

Es importante destacar que la proposición no garantiza que los puntos de S se encuentran en el complementario de $\mathbb{R}^n \setminus V(\mathrm{P})$, aunque esta situación pueda darse. Se pueden dar casos extremos como que $S \subset V(\mathrm{P})$. De forma genérica, tendremos unos puntos de S en $V(\mathrm{P})$, pudiendo controlar el grado de P y unos puntos distribuidos en las diferentes regiones abiertas, donde podemos controlar el número de puntos que tenemos.

Las descomposiciones celulares descritas en la proposición anterior nos van a permitir probar el teorema de Szemerédi-Trotter

Consideramos en $\mathbb{R}^2$ un conjunto finito de puntos P y un conjunto finito de rectas R. Definimos el conjunto siguiente:

$$I(P, R) = \{(p, r) \in P \times R : p \in r\}.$$

En la siguiente proposición, damos unas primeras cotas para $|I(P, R)|$, que se obtienen de forma sencilla simplemente explotando la naturaleza combinatoria del problema. El teorema de Szemerédi-Trotter nos va a dar mejores cotas, pero para ello hay que pasar por las descomposiciones celulares que hemos introducido en este capítulo.

Proposición 4.7. En $\mathbb{R}^2$, sean P un conjunto finito de puntos y R un conjunto finito de rectas. Se verifican las siguientes cotas para $|I(P, R)|$:

- $|I(P, R)| \leq |P|^2 + |R|$.
- $|I(P, R)| \leq |P| + |R|^2$.
- $|I(P, R)| \leq |P|\, |R|^{\frac{1}{2}} + |R|$.

- $|I(P,R)| \leq |R|\,|P|^{\frac{1}{2}} + |P|$.

Demostración. Es claro que los dos primeros resultados son duales. Probamos solamente el primero, el segundo se obtiene de forma análoga. Fijamos un punto $p \in P$. Sea R_p el conjunto formado por las rectas de R que únicamente contienen al punto p, de entre los puntos de P. Dado otro punto p', a lo sumo hay una recta de R que pasa por p y p'. Deducimos que $|I(\{x\},R)| \leq |P|+|R_p|$, y por lo tanto teniendo en cuenta todos los puntos de P, se verifica que:

$$|I(P,R)| \leq |P|^2 + \sum_{p\in P} |R_p| \leq |P|^2 + |R|\,.$$

Veamos finalmente la tercera desigualdad. Dada una recta $r \in R$, denotamos $N(r)$ al número de puntos p de P tales que $p \in r$. Se verifica entonces que $|I(P,R)| = \sum_{r\in R} N(r)$. De la desigualdad de Cauchy-Schwarz se deduce que $|I(P,R)|^2 \leq |R| \sum_{r\in R} N(r)^2$. Se puede comprobar mediante una sencilla recurrencia que el término de la derecha de la última desigualdad cuenta el número de ternas$(p,p',r) \in P \times P \times R$ con $p,p' \in r$. Para esto último no hay que olvidar las ternas correspondientes a un mismo punto p contado dos veces. Teniendo en cuenta que dos puntos definen una única recta, podemos acotar el término de la derecha por $n^2 + |I(P,R)|$. Se satisface:

$$\frac{|I(P,R)|^2}{|R|} \leq |P|^2 + |I(P,R)|\,.$$

A partir de esta desigualdad y completando cuadrados se verifica que:

$$\left(|I(P,R)| - \frac{|R|}{2}\right)^2 \leq |P|^2\,|R| + \frac{|R|^2}{4} \leq \left(|P|\,|R|^{\frac{1}{2}} + \frac{|R|}{2}\right)^2.$$

Finalmente se obtiene $|I(P,R)| \leq |P|\,|R|^{\frac{1}{2}} + |R|$. La cuarta desigualdad se obtiene de forma análoga, simplemente en vez de tener en cuenta que por dos puntos distintos pasa a lo sumo una recta, hay que tener en cuenta que dos rectas distintas se cortan en a lo sumo un solo punto.

□

Es interesante observar que esta ultima cota es válida si en vez de $\mathbb{R}$, consideramos un cuerpo finito. La demostración sigue siendo valida y además la cota es ajustada. Sea $\mathbb{F}$ un cuerpo finito. Tomamos P el conjunto formado por todos los puntos de $\mathbb{F}^2$ y R el conjunto formado por todas las rectas de $\mathbb{F}^2$. Se comprueba en este caso que $|I(P,R)| = |\mathbb{F}|^3 + |\mathbb{F}|^2$, ya que se tiene $|R| = |\mathbb{F}|^2 + |\mathbb{F}|$.

Podemos abordar ahora la prueba del teorema de Szemerédi-Trotter, en su versión plana. Este teorema nos permite dar una cota para la cantidad $|I(P,R)|$.

Teorema 4.8 (Szemerédi-Trotter)**.** En $\mathbb{R}^2$, sean P un conjunto finito de puntos y R un conjunto finito de rectas. Tenemos:

$$|I(P,R)| = O(|P|^{\frac{2}{3}}\,|R|^{\frac{2}{3}} + |P| + |R|).$$

Demostración. Aplicamos ahora la proposición 4.6, para un parámetro r al que daremos valor más adelante. Existe entonces un polinomio Q que nos permite descomponer el plano de la siguiente forma:

$$\mathbb{R}^2 = V(\mathrm{Q}) \cup C_1 \cup ... \cup C_N.$$

Obviamente estamos denotando con C_i a las células de la descomposición. Además, recordamos que Q tiene a lo sumo grado $\deg(\mathrm{Q}) = O(r)$ y la descomposición cumple que $|P \cap C_i| \leq |P|\, r^{-n}$. Definimos P_i como el conjunto formado por los puntos de P que se encuentran en C_i, para $i = 1, ..., N$. De forma similar para cada $i = 1, ..., N$, definimos $R_i = \{r \in R : r \cap C_i \neq \emptyset\}$, que es el conjunto formado por las rectas de R, que pasan por la célula C_i. No debemos olvidarnos de los puntos de P que se pueden encontrar en el conjunto de ceros de Q. Denotamos con P_0 al conjunto formado por los puntos $p \in P$ tales que $p \in V(\mathrm{Q})$. De forma similar, definimos $R_0 = \{r \in R : r \cap V(\mathrm{Q}) \neq \emptyset\}$. Teniendo todo esto en cuenta, se verifica que:

$$|I(P,R)| = |I(P_0,R_0)| + \sum_{i=1}^{N} |I(P_i,R_i)|\,.$$

Podemos acotar ahora las cantidades que aparecen en la suma. Aplicando las desigualdades de la proposición 4.7, sabemos que:

$$|I(P_i,R_i)| \leq |P_i|\,|R_i|^{\frac{1}{2}} + |R_i| \leq \frac{|P_i|}{r^2}\,|R_i|^{\frac{1}{2}} + |R_i|\,.$$

Por construcción del conjunto R_i cuando $i \geq 1$, sabemos que las rectas de R_i no están contenidas en $V(\mathrm{Q})$. Tenemos:

$$\sum_{i=1}^{N} |R_i| \leq \deg(\mathrm{Q})\,|R| \lesssim r\,|R|\,.$$

Recordemos que en la demostración de la proposición 4.6, veíamos que con la notación de esta demostración, $N \lesssim r^2$. Aplicando ahora la desigualdad

de Cauchy-Schwarz, se deduce:

$$\left(\sum_{i=1}^{N} |R_i|^{\frac{1}{2}}\right)^2 \lesssim \left(\sum_{i=1}^{N} |R_i|\right) r^2 \lesssim r^3 |R| .$$

Obtenemos que $\sum_{i=1}^{N} |R_i|^{\frac{1}{2}} \lesssim r^{\frac{3}{2}} |R|^{\frac{1}{2}}$. Llevamos esto a la cota de $|I(P_i, R_i)|$ y concluimos:

$$\sum_{i=1}^{N} |I(P_i, R_i)| \lesssim \frac{|P|\,|R|}{r^{\frac{1}{2}}} + r\,|R| .$$

Acotamos ahora el término $|I(P_0, R_0)|$. Las rectas de R_0 que no están contenidas en $V(\mathrm{Q})$, cortan a la hipersuperficie en a lo sumo $\deg(\mathrm{Q})$ puntos. Estas rectas no pueden contribuir a $|I(P_0, R_0)|$ más que $\deg(\mathrm{Q})\,|R| \lesssim r\,|R|$. Razonando de manera similar, solo puede haber como mucho $\deg(\mathrm{Q})$ rectas de R_0 contenidas en la hipersuperficie. De esto y la cuarta desigualdad de la proposición 4.7 se deduce que estas rectas contenidas en la hipersuperficie solo contribuyen a lo sumo a $|I(P_0, R_0)|$ con $\deg(\mathrm{Q})\,|P| + |P| \lesssim r\,|P| + |P|$. Recopilando todo, tenemos la cota para $|I(P_0, R_0)| \lesssim r\,|R| + r\,|P|^{\frac{1}{2}} + |P|$. Llevando todo a la cota de $|I(P, R)|$, tenemos:

$$|I(P, R)| \lesssim \frac{|P|\,|R|}{r^{\frac{1}{2}}} + r\,|R| + r\,|P|^{\frac{1}{2}} + |P| .$$

Tanto si $|P|^{\frac{1}{2}} \geq |R|$ como si $|P|^2 \leq |R|$, la proposición 4.7 permite concluir. En el resto de casos, basta tomar $r = |P|^{\frac{2}{3}}\,|R|^{\frac{-1}{3}}$ para concluir.

□

En [GK], podemos encontrar más aplicaciones de este método de subdivisión polinómica para la resolución de problemas combinatorios en $\mathbb{R}^n$. Entre otros resultados, se pueden encontrar algunos resultados parecidos a este último teorema para dimensiones mayores que 2. En todo caso, los razonamientos siguen la misma estrategia que la que hemos expuesto aquí.

Bibliografía

[Alon] N. Alon, *Combinatorial Nullstellensatz,* Combinatorics, Probability and Computing 8: 7-29, 1999.

[Bom] E. Bombieri, *Counting points on curves over finite fields,* In Séminaire N. Bourbaki, number exp n430, 1972-1973.

[Dav] R.O. Davies, *Some remarks on the Kakeya problem,* Mathematical Proceedings of the Cambridge Philosophical Society, 69 (03):417-421, 1971.

[DKSS] Z.Dvir, S. Kopparty, S. Saraf and M. Sudan, *Extensions to the method of multiplicities, with applications to kakeya sets and mergers,* Proceedings of the 50th Annual IEEE Symposium on Foundations of Computer Science (FOCS), pages 181-190. IEEE Computer Society, Washington DC, 2009.

[Fal] K.J. Falconer, *The Geometry of Fractal Sets,* Cambridge University Press, 1985.

[GK] L. Guth and N.H.Katz, *On the Erdős distinct distances problem in the plane,* Ann. of Math. (2), 181(1):155-190, 2015.

[Hart] R. Hartshorne, *Algebraic Geometry,* GTM Springer, 1977.

[IK] H. Iwaniec and E.Kowalski, *Analytic Number Theory,* AMS Colloquium Publications, Volume 53, 2004.

[Katz] N.Katz, *The work of Pierre Deligne,* Proc. Intern. Congress of Mathematicians, 47-52, 1978.

[Mat] J. Matoušek, *Using the Borsuk-Ulam Theorem,* Lectures on Topological Methods in Combinatorics and Geometry, Springer, Heidelber, 2003.

[Sch] W.M. Schmidt, *Equations over Finite Fields, an Elementary Approach,* Lecture Notes in Mathematics, Vol. 536, Springer, Berlin, 1976.

[Ste] S.A. Stepanov, *Arithmetic of Algebraic Curves,* , Monographs in Contemporary Math., Plenum Publishing, New York, 1994.

[Tao] T. Tao, *Algebraic combinatorial geometry: the polynomial method in arithmetic combinatorics, incidence combinatorics, and number theory,* EMS Surv. Math. Sci. 1, 1-46, 2014.