Universidad Nacional Autónoma de México

Facultad de Ciencias

A 100 AÑOS DEL PROBLEMA DE KAKEYA: ORIGEN Y PERSPECTIVAS

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

MATEMÁTICO

PRESENTA:

JONATHAN LÓPEZ RUIZ

TUTOR:
DR. VINICIO ANTONIO GÓMEZ GUTIÉRREZ

CIUDAD DE MÉXICO, 2017

Hoja de datos del Jurado.

1. Datos del alumno
   López
   Ruiz
   Jonathan
   59706390
   Universidad Nacional Autónoma de México
   Facultad de Ciencias
   Matemáticas
   106005817

2. Datos del tutor
   Dr.
   Vinicio Antonio
   Gómez
   Gutiérrez.

3. Datos del sinodal 1
   Dra.
   Martha
   Takane
   Imay.

4. Datos del sinodal 2
   Dr.
   Hugo Alberto
   Rincón
   Mejía.

5. Datos del sinodal 3
   Dr.
   Valente
   Santiago
   Vargas.

6. Datos del sinodal 4
   M. en C.
   Francisco de Jesús
   Struck
   Chávez.

7. Datos del trabajo escrito.
   A 100 años del problema de Kakeya: Origen y perspectivas
   121 p.
   2017

I

Nunca perseguí la gloria
ni dejar en la memoria
de los hombres mi canción;
yo amo los mundos sutiles,
ingrávidos y gentiles
como pompas de jabón.
Me gusta verlos pintarse
de sol y grana, volar
bajo el cielo azul, temblar
súbitamente y quebrarse.

II

¿Para qué llamar caminos
a los surcos del azar?...
Todo el que camina anda,
como Jesús, sobre el mar.

XXIX

Caminante, son tus huellas
el camino, y nada más;
caminante, no hay camino:
se hace camino al andar.
Al andar se hace camino,
y al volver la vista atrás
se ve la senda que nunca
se ha de volver a pisar.
Caminante, no hay camino,
sino estelas en la mar.

XLIV

Todo pasa y todo queda;
pero lo nuestro es pasar,
pasar haciendo caminos,
caminos sobre la mar.

Antonio Machado

Índice general

Introducción.

En 1917 el matemático japonés Sôichi Kakeya propuso el siguiente problema:

Problema de Kakeya. *¿Cuál es la menor área requerida para rotar* 180° *y de forma continua un segmento de recta de longitud uno en el plano, de modo que vuelva a ocupar su posición original?*

Podría pensarse que este curioso acertijo, conocido como el ***problema de Kakeya***, es meramente una curiosidad matemática, una atractiva pregunta a la cual dedicar unos minutos de ocio. Sin embargo, a partir de 1970 (y, principalmente, después de los años 90) han surgido conexiones de este problema con áreas insospechadas de las matemáticas contemporáneas. Hoy en día la palabra «Kakeya» aparece cada vez con mayor frecuencia en la literatura matemática y, casi de manera constante, los vínculos que se han logrado establecer con el problema de Kakeya parecen ser una fuente inagotable de sorpresas.

El propósito fundamental, pues, de esta tesis es, por una parte, describir de forma detallada la solución del problema de Kakeya y, por otro lado, presentar al lector un enfoque más reciente del mismo. En este sentido, establecido un doble objetivo, el presente trabajo se divide esencialmente en dos partes:

La primera parte está dedicada, naturalmente, al problema de Kakeya. En el capítulo uno se introduce, muy brevemente, la historia del problema así como los personajes involucrados en su solución. En el capítulo dos se expone el primer paso para dar solución al problema: los *conjuntos de Besicovitch*. Un ***conjunto de Besicovitch*** es un subconjunto del plano que contiene un segmento de recta de longitud uno en cada dirección. En 1928 el matemático ruso Abram S. Besicovitch -de él el nombre de dichos conjuntos– demostró que existen conjuntos de Besicovitch ¡de área cero! En el capítulo tres, finalmente, se expone la solución del problema de Kakeya a partir de los conjuntos de Besicovitch.

La segunda parte de esta tesis trata una versión más reciente del problema de Kakeya conocida como el ***problema finito de Kakeya***. En 1999 el matemático Thomas Wolff propuso un análogo para campos finitos del problema de Kakeya: si $\mathbb{F}^n$ es un espacio vectorial sobre un campo finito $\mathbb{F}$, y si definimos un ***conjunto finito de Kakeya*** como un subconjunto de $\mathbb{F}^n$ que contiene una recta en cualquier dirección, entonces Wolff conjeturaba lo siguiente:

Problema finito de Kakeya. *Sea $K \subseteq \mathbb{F}^n$ un conjunto finito de Kakeya. Entonces K tiene cardinalidad a lo menos*

$$|K| \geq C_n |\mathbb{F}|^n,$$

donde $C_n > 0$ es una constante que depende únicamente de n.

En el capítulo cuatro se introduce de forma más detallada la conjetura anterior. Así, el objetivo del resto de la tesis es claro: dar solución al problema finito de Kakeya. Con esto en mente, en el capítulo cinco se desarrolla la teoría necesaria para, finalmente, en el capítulo seis dar completa y cabal solución al problema finito de Kakeya. En el capítulo siete se aborda un conjunto relacionado con los conjuntos finitos de Kakeya: los *conjuntos finitos de Nikodym*. La finalidad de lo anterior es ejemplificar las técnicas introducidas en los capítulos cuatro y cinco resolviendo una conjetura análoga para los *conjuntos de Nikodym*. Por último, se presentan dos apéndices: uno sobre la *medida de Jordan* y otro sobre campos finitos. El primero de ellos formaliza la noción intuitiva de área. Ahora bien, aunque en la mayoría de los resultados no se utilizan propiedades de los campos finitos más allá de su finitud, en el segundo apéndice se exponen las propiedades básicas de los campos finitos para el lector interesado en profundizar en el tema.

Parte I

El problema de Kakeya.

Capítulo 1

Introducción: motivación histórica.

1.1. El problema de Kakeya

En su artículo de 1917 (Kak17) el matemático japonés Sôichi Kakeya, entonces profesor en la Universidad de Sendai, propuso el siguiente interesante problema con aspecto de acertijo:

Problema 1.1.1 (Problema de Kakeya)**.** *En la clase de las figuras convexas del plano en las que puede rotarse* 180° *y de manera continua un segmento de recta de longitud uno (siempre permaneciendo dentro de dicha figura), ¿cuál es la de menor área?* [1]

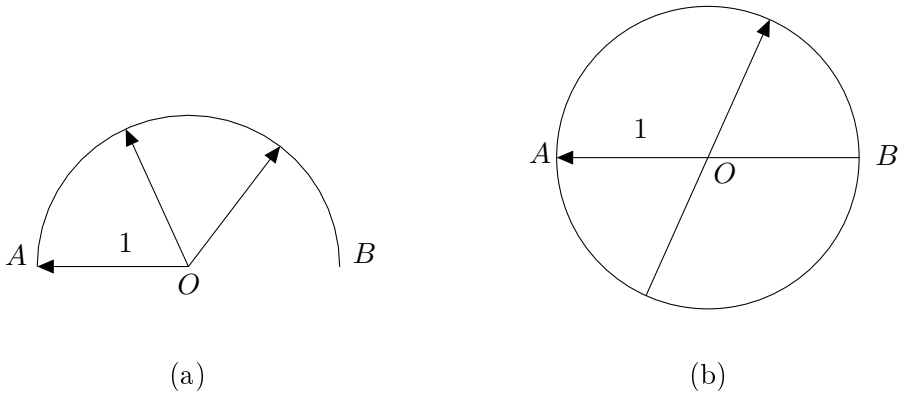
Sería interesante que el amable lector abandonase temporalmente la lectura del presente trabajo y dedicase un momento a buscar su propia solución al problema de Kakeya. Suponiendo que ya se ha dedicado un tiempo a la cuestión, y antes de dar soluciones definitivas, vamos a explorar posibles soluciones que quizá se hayan pensado:

Un primer intento de solución al problema 1.1.1 apunta a rotar el segmento a lo largo de media circunferencia de radio uno como se muestra en la figura 1.1 (a). En efecto, basta tomar el segmento OA como el radio de la semicircunferencia y rotarlo 180° alrededor de O hasta coincidir con el punto B. Después, se puede trasladar horizontalmente hacia la izquierda –esto no modifica el área utilizada– hasta su posición inicial. El área utilizada con este procedimiento es $\frac{\pi}{2} = 1.5707...$.

Puede mejorarse la solución anterior si consideramos un círculo de diámetro uno como en la figura 1.1 (b). Haciendo coincidir el diámetro con el segmento AB y girándolo 180° alrededor del centro de la circunferencia, éste queda invertido ocupando un área igual a $\frac{\pi}{4} = 0.7853...$. Sin embargo, todavía podemos hacerlo mejor. Consideremos un triángulo equilátero de altura uno, llamémoslo ABC. Coloquemos el segmento sobre el lado AB con uno de sus extremos en el vértice B. Giremos 60° alrededor del extremo del segmento que se ha fijado y deslicemos sobre el lado BC. De nuevo, giremos 60° alrededor del vértice C y avancemos a lo largo de AC. Ahora, podemos rotar de nueva cuenta

[1] Algunos autores enuncian el problema de Kakeya considerando una rotación de 360°; ambas formulaciones son equivalentes. En efecto, si se puede lograr la rotación del segmento 180°, repitiendo el procedimiento se obtiene una rotación de 360°. Recíprocamente, si se puede realizar una rotación de 360°, claramente se puede obtener una de 180°. Por simplicidad, en este trabajo se considerará la formulación del problema de Kakeya con 180°.

60° alrededor del vértice A y, finalmente, desplazarnos hacia abajo sobre AB. Como resultado, hemos invertido el segmento dentro del triángulo ABC. Este procedimiento, en efecto, mejora la solución pues el área del triángulo es $\frac{1}{\sqrt{3}} = 0.5773...$. La figura 1.2 muestra lo anterior.



(a) (b)

Figura 1.1: Movimiento de un segmento de longitud uno en: (a) Semicírculo de radio 1 y área $\frac{\pi}{2}$ (b) Círculo de diámetro 1 y área $\frac{\pi}{4}$.

El mismo Kakeya conjeturó que el triángulo equilátero era la figura que resolvía el problema mas no presentó una demostración de dicha afrimación. Por otro lado, según menciona Kakeya en su artículo con M. Fujiwara de 1917 (FK17), a sugerencia del matemático Tadahiko Kubota éstos consideraron el problema sin la restricción de la convexidad para la figura buscada. Naturalmente, esto hizo al problema más atractivo pues permitía considerar un número mayor de figuras. Kakeya y Fujiwara conjeturaron que la deltoide [1] (también llamada tricúspide o hipocicloide de Steiner) inscrita en una circunferencia de diámetro $\frac{3}{2}$ resolvía el problema modificado. (Ver figura 1.3.) Es fácil convencerse de que un segmento de longitud uno puede rotarse dentro de tal figura pues ésta posee la siguiente interesante propiedad: si γ denota a la deltoide, para cualquier punto P en γ, la tangente en P a γ contiene un segmento AB interior a la figura y de longitud uno, independientemente del punto P que se tome. Sorprendentemente, el área de esta figura es $\frac{\pi}{8} = 0.3926...$ (¡incluso mejor que para el caso convexo!), aunque nada daba indicios de que esto no pudiese mejorarse.

El primer avance significativo al problema fue hecho por el matemático húngaro Julius Pál. Huyendo del caos político de posguerra posterior a la Primera guerra mundial, en 1919 se traslada de su natal Hungría a Copenhague, Dinamarca, donde pudo concentrarse en su investigación. Así, en 1921, demostró en (Pál21) que, de hecho, el triángulo equilátero sí es la figura convexa de menor área en la que puede rotarse un segmento

[1] Consideremos un circunferencia de radio R (llamada directriz) y en su interior otra circunferencia (llamada generatriz) de radio menor, $r = \frac{R}{3}$, tangente a la primera. Una *deltoide* es la curva que describe el movimiento de un punto situado sobre la circunferencia generatriz que gira, sin deslizamiento, por el interior de la circunferencia directriz. Es un caso particular de las *hipocicloides*.

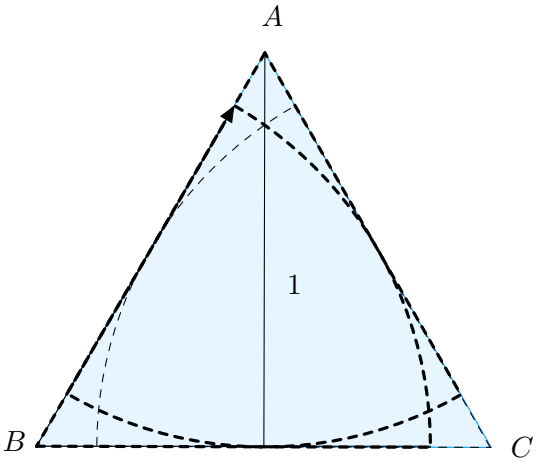


Figura 1.2: Triángulo equilátero de altura uno y área $\frac{1}{\sqrt{3}}$.

de longitud uno. Esto llevó al convencimiento de que el problema sin la restricción de convexidad era de mayor interés. Más tarde, el problema modificado de esta manera es lo que llegaría a conocerse como *El problema de Kakeya*.

Con todo, el problema seguía abierto: ¿en qué figura no convexa del plano, de área muy pequeña, podía invertirse un segmento de longitud uno? Sin lugar a duda quedaba mucho camino por recorrer para dar solución definitiva al problema de Kakeya.

1.2. El problema de Besicovitch

Aquí comienza la parte más interesante en la historia del problema de Kakeya. En 1917 Rusia atravesaba años cruciales; la revolución civil transformaría el país y nada volvería a ser igual. Tras graduarse en 1912 de la Universidad de San Petersburgo, el matemático Abram S. Besicovitch [1] se trasladó a Perm, en los Urales, para trabajar como profesor en la última universidad rusa creada antes de la revolución soviética. Ahí Besicovitch trabajaba en problemas relacionados con el análisis y, en particular, hubo uno en el que centró su atención. Este problema involucraba integrales de Riemann y puede enunciarse como sigue:

Problema 1.2.1 (Problema de Besicovitch)**.** *Dada una función* $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ *Riemann-integrable en una región del plano, ¿es siempre posible encontrar un sistema ortogonal de ejes coordenados* u, v *tal que la función* $g(u) = \int f(u, v)\, dv$ *exista como integral*

[1] Los datos biográficos que se mencionan a continuación sobre Abram S. Besicovitch fueron tomados de (Tay75)

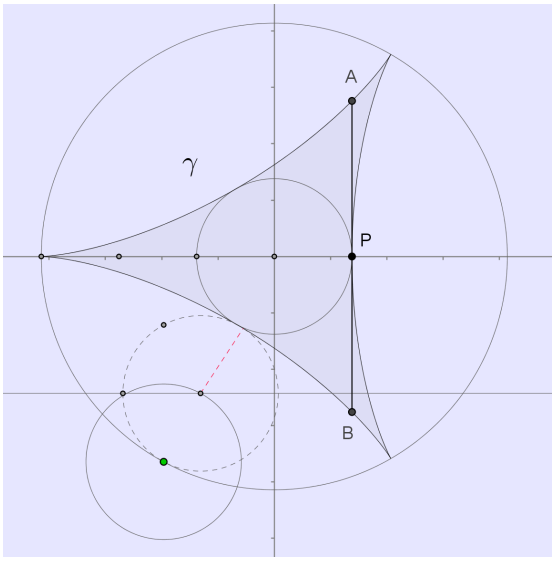


Figura 1.3: Deltoide de área $\frac{\pi}{8}$ inscrita en una circunferencia de diámetro $\frac{3}{2}$.

de Riemann para todo u, que g(u) sea Riemann-integrable y que $\int g(u)du$ *sea igual al valor de la integral sobre la región de* $\mathbb{R}^2$*?*

Besicovitch observó que podía dar respuesta negativa a esta pregunta –acentuando que la integral sobre regiones del plano y la integral iterada son objetos de naturaleza distinta. La idea del contraejemplo de Besicovitch dependía fundamentalmente de la existencia de un conjunto compacto B, *de área cero* (más precisamente, de *medida de Jordan cero* en el plano. Ver el apéndice A), que tuviese un segmento de longitud uno en cualquier dirección. Si dicho conjunto podía construirse, entonces era posible manipularlo de manera conveniente y, de forma muy ingeniosa, definir una función en dicho conjunto que sería integrable en el plano pero no de manera iterada. No abordaremos en el presente trabajo la solución al problema de Besicovitch, para esto nos referimos a (Dun06); lo importante aquí es resaltar que, de hecho, él fue capaz de contruir dicho conjunto (Bes19), llamado *Conjunto de Besicovitch*. Sin embargo, el aislamiento provocado por el conflicto armado entre la Guardia blanca del Zar contra el Ejército rojo bolchevique no permitió la circulación del trabajo de Besicovitch (en 1928 se publicaría de nuevo en *Mathematische Zeitschrift* (Bes28)). Por las mismas circunstancias, el problema propuesto por Kakeya tampoco fue investigado en Rusia en ese periodo. Dado que ambos problemas involucraban un conjunto, con la menor área posible, que tuviese un segmento de longitud uno en cualquier dirección (en adición para el de Kakeya que pedía la condición del movimiento continuo del segmento dentro del conjunto) el vínculo era claro. De hecho, Besicovitch llegaría a considerar ambos problemas como «problemas gemelos» (Bes63).

La oportunidad de recibir la beca Rockefeller daba la oportunidad a Besicovitch de

salir de Rusia hacía el occidente, pero esto se vio truncado al serle negado el permiso de salida por parte de las autoridades rusas. En 1924, finalmente, decide abandonar Rusia por cuenta propia, pasando un año en Copenhague junto a Harald Bohr, importante matemático menos conocido que su hermano, el físico Niels Bohr. De allí Besicovitch pasó a Inglaterra donde estuvo varios meses en Oxford. G. H. Hardy, uno de los más ilustres matemáticos ingleses, le ayudó a ingresar como profesor en la Universidad de Liverpool. Finalmente, en 1927 se establecería en Cambridge donde realizaría el resto de su labor de investigación.

Realmente no se conoce quién hizo la conexión entre le problema de Besicovitch y el de Kakeya, pero lo cierto es que una vez pasados los periplos que Besicovitch tuvo que pasar para abandonar Rusia, éste conoció de algún modo el problema y notó que de manera muy simple podía modificar su construcción de 1919 para dar una notable solución al problema, dando una respuesta totalmente inesperada: no hay una figura de área mínima con las condiciones del problema de Kakeya. En otras palabras, la tarea de rotar continuamente un segmento de longitud uno dentro de un conjunto puede realizarse ¡ocupando un área tan pequeña como se desee!

Capítulo 2

El conjunto de Besicovitch.

El objetivo último de este capítulo es describir la construcción de un *conjunto de Besicovitch* en el plano de *área cero* [1]. Los conceptos desarrollados para esto último serán el primer paso en la solución del problema de Kakeya. Vale la pena precisar entonces a qué nos referiremos a lo largo del presente trabajo con un *conjunto de Besicovitch*:

Definición 2.0.1. *Un conjunto compacto B de $\mathbb{R}^n$ se llamará* **conjunto de Besicovitch** *si contiene un segmento de recta de longitud uno en cada dirección. Más precisamente,*

$$\forall\, e \in S^{n-1}\ \exists\, x_e \in \mathbb{R}^n \text{ tal que } \left\{ x_e + te : t \in \left[-\frac{1}{2}, \frac{1}{2}\right] \right\} \subseteq B$$

donde $S^{n-1} = \{x \in \mathbb{R}^n : \|x\| = 1\}$, la esfera unitaria en $\mathbb{R}^n$.

Nuestro interés, pues, estará en el caso $n = 2$. La laboriosa construcción original de Besicotvich (Bes19) (Bes28) ha sido significativamente simplificada muchas veces (Perron (Per28), van Alphén (vA42), Rademacher (Rad62), Schoenberg (Sch62), el mismo Besicovitch (Bes63), Cunningham (Cun71) y Fisher (Fis73)). La idea central en todas las construcciones es formar un «árbol de Perron-Schoenberg». En este trabajo se presenta una versión detallada que puede consultarse en (Fal86), (dG76) y (Ste16).

2.1. Un pequeño árbol con muchos frutos.

Con el objetivo de exhibir un conjunto de Besicovitch de *área tan pequeña como se desee*, O. Perron e I.J. Schoenberg observaron, de manera independiente, que era mucho más fácil formular primero la manera en que puede ser construido un conjunto, de área tan pequeña como sea posible, con un segmento en toda dirección dentro de un intervalo de 60°; después, sería suficiente tomar copias rotadas apropiadamente de dicho conjunto para obtener uno nuevo con un segmento de recta en cualquier dirección y con las características del área requeridas. Ésta es la idea detrás del «árbol de Perron-Schoenberg». Pero ¿cómo puede construirse un «árbol de Perron-Schoenberg»? En términos generales, la construcción se basa en el siguiente procedimiento:

[1] De forma más precisa se puede hablar de *medida de Jordan cero*, pero, para los objetivos de esta tesis, basta el concepto usual e intuitivo de *área*.

Partamos de un triángulo equilátero, llamemóslo ABC, de altura uno y con base sobre una recta L. Evidentemente, éste contiene un segmento de recta de longitud uno para cada ángulo en el intervalo $[60°, 120°]$ con respecto a L. Podemos entonces construir triángulos más pequeños, digamos $T_1, T_2, ..., T_n$, dividiendo la base de ABC en n intervalos de igual longitud $I_1, I_2, ..., I_n$ y después trasladarlos de forma adecuada para que el área de la unión de los triángulos trasladados $T_1', T_2', ..., T_n'$ sea tan pequeña como queramos. (Ver figuras 2.1 y 2.2.)

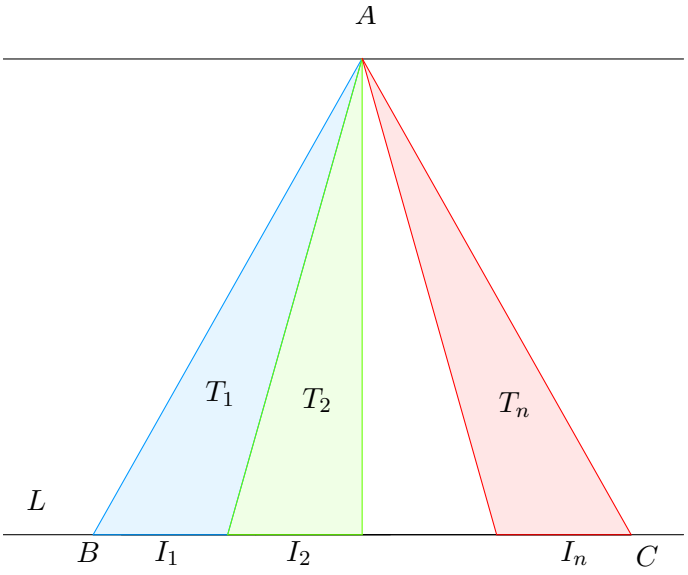


Figura 2.1: División de un triángulo ABC en n subtriángulos.

Los elementos clave de la construcción un «árbol de Perron-Schoenberg» [1] son, por una parte, en cuántos triángulos pequeños dividamos al triángulo original y, por otro lado, la distancia que sean trasladados dichos triángulos. De estos factores dependerá qué tan pequeña sea el área de la figura resultante. Los siguientes dos resultados establecen de manera rigurosa esta idea.

Notación 2.1.1. *En todo lo que sigue, al área de un conjunto $S \subseteq \mathbb{R}^2$ se le denotará por $A(S)$.*

Lema 2.1.2. *Sea T un triángulo con base sobre una recta L. Dividamos la base de T en dos segmentos de igual longitud, y únase el punto de división con el vértice opuesto del triángulo T para formar dos triángulos adyacentes T_1 y T_2 (que llamaremos* triángulos elementales*), con base y altura de longitudes, respectivamente, b y h. (Ver figura 2.3.) Tómese $\frac{1}{2} < \alpha < 1$. Entonces, si movemos a lo largo de L al triángulo T_2 una distancia*

[1] De aquí en adelante por un «árbol de Perron-Schoenberg» nos referiremos a una figura obtenida con el procedimiento descrito.

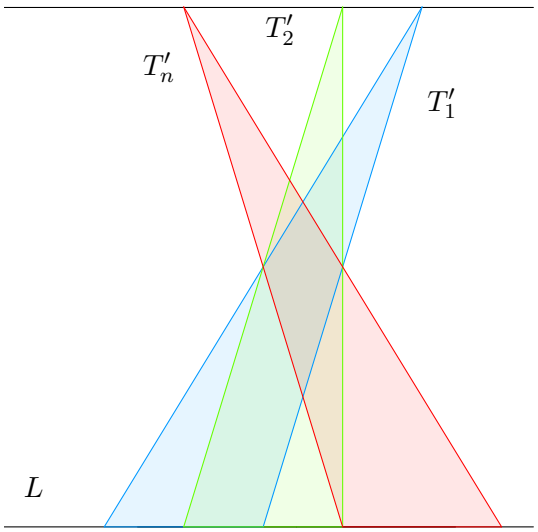


Figura 2.2: Construcción del «árbol de Perron-Schoenberg».

igual a $2(1-\alpha)b$ *hasta sobreponerse con el triángulo* T_1*, la figura resultante* S *satisface las siguientes propiedades:*

1. *S consiste de un triángulo T', semejante a T, y de dos triángulos más pequeños T'_1 y T'_2, que llamaremos* triángulos auxiliares. *Además,* $A(T') = \alpha^2 A(T)$.

2. *El área de S está dada por la siguiente expresión:*

$$A(S) = \left[\alpha^2 + 2(1-\alpha)^2\right] A(T).$$

(Ver figura 2.4.)

Demostración. 1. Basta inspeccionar los ángulos del triángulo T' para concluir que éste es semejante a T. Además, los triángulos auxiliares son resultado inmediato de la construcción de S. Ahora, la longitud de la base de T' es igual a $2b - 2(1-\alpha)b = 2\alpha b$; mientras que la base de T mide $2b$. De estos dos hechos se sigue que la razón de semejanza entre éstos es α, y, por tanto, $A(T') = \alpha^2 A(T)$. De esta última igualdad podemos precisar la altura de T' (algo que usaremos más adelante en la prueba): ésta es igual a αh. Aquí concluye la primera parte de la demostración.

2. Para calcular el área de los triángulos T'_1 y T'_2 vamos a hacer uso de la siguiente construcción: por el punto de intersección de los triángulos auxiliares, tracemos una recta paralela a L; esto define cuatro triángulos: $A_{1,1}, A_{1,2}, A_{2,1}$ y $A_{2,2}$. (Ver figura 2.5.)

Nótese que $A_{1,1}$ es semejante a T_1 y que $A_{2,1}$ es semejante a T_2 –una vez más, es suficiente observar los ángulos para convencerse de esto–, ambos a razón de semejanza

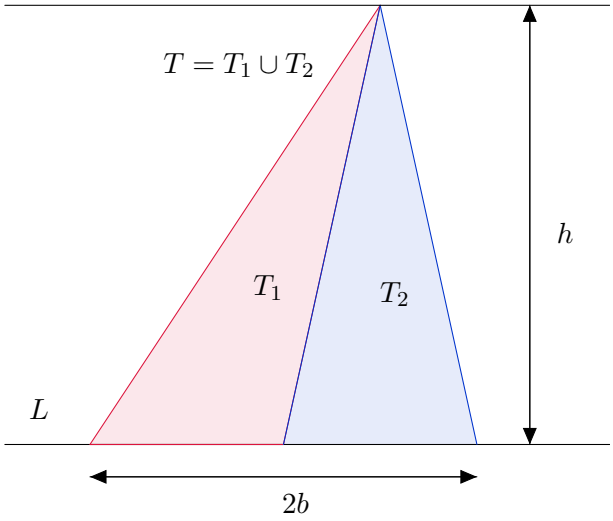


Figura 2.3: División del triángulo T para obtener los dos triángulos elementales T_1 y T_2.

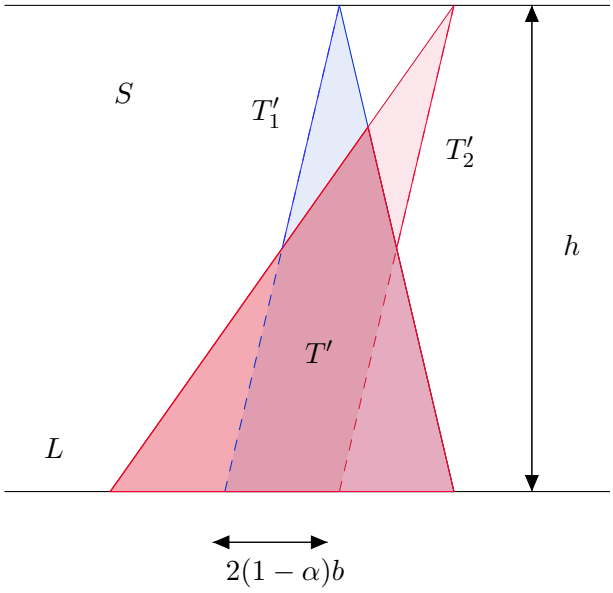


Figura 2.4: Traslación del triángulo T_2 a lo largo de L hasta sobreponerse con T_1.

$1-\alpha$. Más aún, $A_{1,1}$ es congruente con $A_{2,2}$ y $A_{1,2}$ es congruente con $A_{2,1}$. Así, los cuatro triángulos tienen base de longitud $(1-\alpha)b$ y altura $(1-\alpha)h$. Luego,

$$A(A_{1,1}) = A(A_{1,2}) = A(A_{2,1}) = A(A_{2,2}) = \frac{1}{2}(1-\alpha)^2 A(T).$$

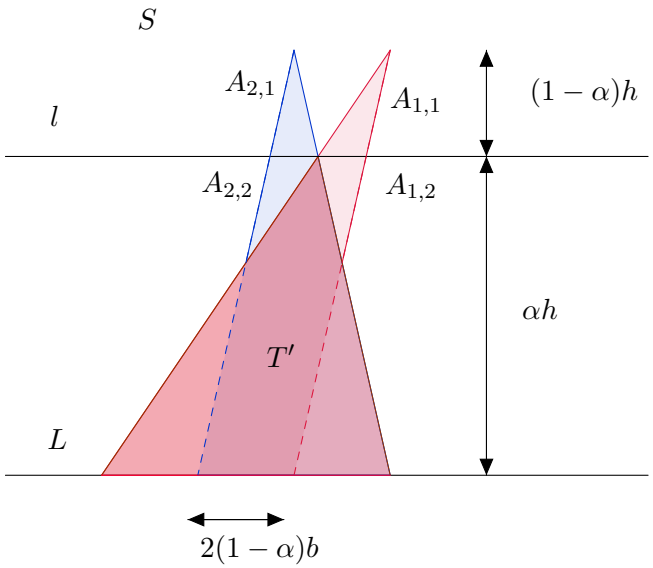


Figura 2.5: La recta l paralela a L. Note que los triángulos $A_{1,1}, A_{1,2}, A_{2,1}$ y $A_{2,2}$ tienen la misma área.

Finalmente, como $A(S) = A(T') + A(T_1') + A(T_2')$, podemos combinar los resultados anteriores para concluir que

$$A(S) = \left[\alpha^2 + 2(1-\alpha)^2\right] A(T).$$

Esto termina la demostración. □

Teorema 2.1.3. *Consideremos un triángulo, T, con base sobre la recta L. Divídase la base de T en 2^n segmentos de igual longitud, y unamos los puntos de división con el vértice opuesto del triángulo T para formar 2^n triángulos elementales $T_1, T_2, ..., T_{2^n}$ (ver figura 2.6). Si n es lo suficientemente grande, existe una traslación, a lo largo de L, de cada T_i $(1 \leq i \leq 2^n)$ de tal forma que el área de la figura (cerrada) resultante S, que es la unión de los triángulos trasladados T_i, sea tan pequeña como se desee.*

Observación 2.1.4. *La traslación de cada triángulo T_i como se describe en el teorema 2.1.3 se aplica tanto al triángulo como a su frontera. Esto es, la imagen de cada T_i bajo la traslación es una figura cerrada. Por lo tanto, la figura S resultante es compacta. Nótese además que, como algunos triángulos comparten frontera, se agregan $2^n - 1$ segmentos de recta en la construcción. Sin embargo, esto no afecta en el área pues hay un número finito de dichos segmentos.*

Demostración. Construyamos la figura S a través de una serie de pasos que involucran repetidas aplicaciones del lema 2.1.2 para un valor fijo de α que será precisado más adelante.

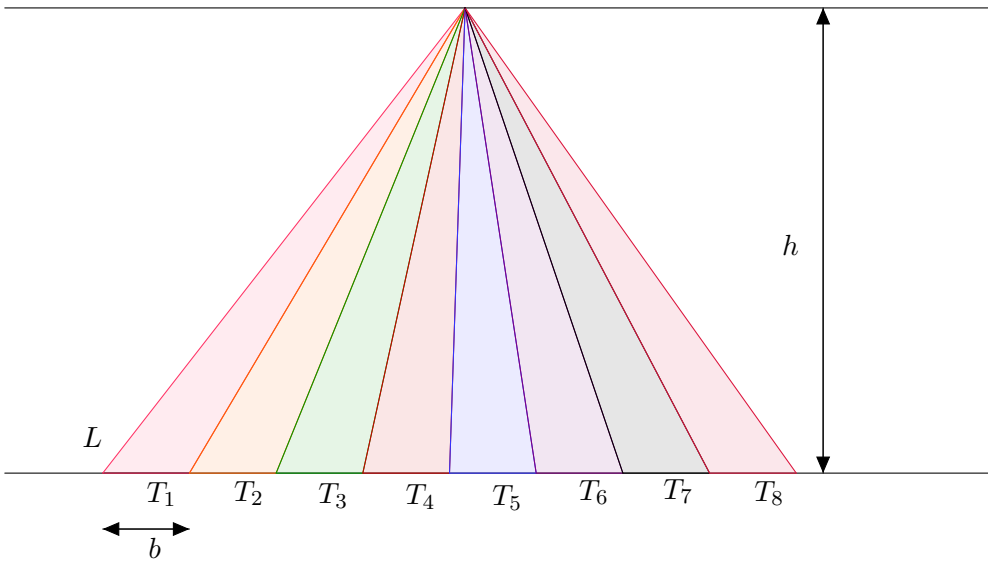


Figura 2.6: La construcción descrita en el teorema 2.1.3 para $n = 3$.

Paso 1. Para este primer paso consideraremos pares de triángulos elementales T_i. Para cada i, $1 \leq i \leq 2^{n-1}$, movamos a lo largo de la recta L al triángulo T_{2i} en la dirección de T_{2i-1} una distancia igual a $2\,(1-\alpha)\,b$ para obtener figuras S_i^1, cada una de éstas formada por el triángulo T_i^1 (semejante a $T_{2i-1} \cup T_{2i}$) y dos triángulos auxiliares A_{2i-1}^1, A_{2i}^1. Más aún, $A\left(T_i^1\right) = \alpha^2 A\left(T_{2i-1} \cup T_{2i}\right)$. Además, según el lema anterior, cada figura S_i cumple que

$$A\left(S_i^1\right) = \left[\alpha^2 + 2(1-\alpha)^2\right] A\left(T_{2i-1} \cup T_{2i}\right).$$

Esta primera aplicación del lema 2.1.2 para todos los triángulos elementales da lugar a una colección $\left\{S_i : 1 \leq i \leq 2^{n-1}\right\}$ de nuevas figuras que satisface lo siguiente:

$$\begin{aligned}\sum_{i=1}^{2^{n-1}} A\left(S_i^1\right) &\leq \left[\alpha^2 + 2(1-\alpha)^2\right] \sum_{i=1}^{2^{n-1}} A\left(T_{2i-1} \cup T_{2i}\right) \\ &= \left[\alpha^2 + 2(1-\alpha)^2\right] A\left(T\right).\end{aligned}$$

La figura 2.7 ilustra esta situación para $n = 3$.

Hagamos una pausa. Es importante observar que, para cada i, $1 \leq i \leq 2^{n-1}$, un lado del triángulo T_{2i-1}^1 es paralelo y de igual longitud al lado opuesto del triángulo T_{2i}^1; así, podemos trasladar las 2^{n-1} figuras S_i^1 de manera que la figura que formen contenga un triángulo D_1, semejante al triángulo inicial T con razón de semejanza α, más los triángulos auxiliares que ya teníamos con un grado de superposición mayor (hecho que

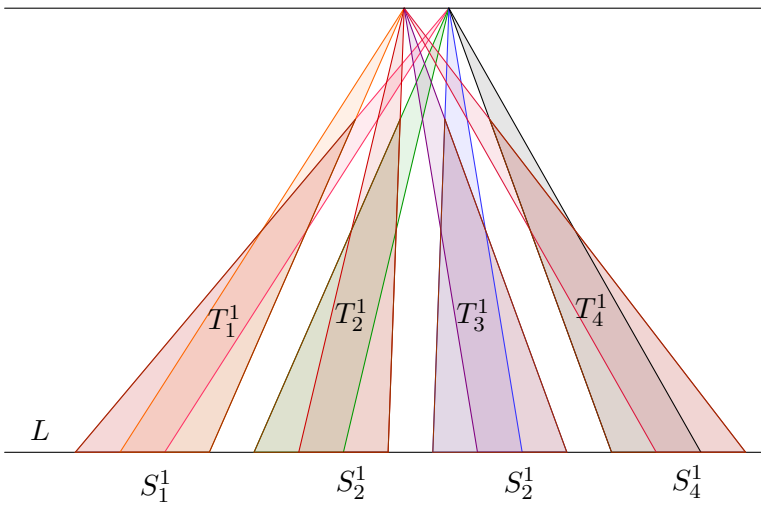


Figura 2.7: La traslación de T_i para formar S_i^2 para $n = 3$.

asegura que el área, en efecto, va disminuyendo cada vez más). Nótese que la longitud de la base del triángulo D_1 es igual a $2^n \alpha b$. En efecto, la base del triángulo original T es $2^n b$; ahora, cada triángulo elemental T_{2i} se desplazó a lo largo de la recta L una distancia $2\,(1-\alpha)\,b$, y como hay 2^{n-1} de dichos triángulos, se sigue que la base del triángulo D_1 es igual a

$$2^n b - \left[2^{n-1} \cdot 2\,(1-\alpha)\,b\right] = 2^n \alpha b.$$

Más aún, por construcción, la base del triángulo D_1 está dividida en 2^{n-1} partes iguales de longitud $2\alpha b$. Esto permite repetir el lema 2.1.2 en el segundo paso. De hecho, un razonamiento análogo entre cada paso es lo que permitirá hacer la construcción de S de manera iterada. En otras palabras, entre cada paso de la construcción de la figura final S se deberá realizar una traslación adecuada de las figuras para poder aplicar el lema 2.1.2. La figura 2.8 muestra esto para $n = 3$

En resumen, la primera aplicación del lema 2.1.2 para todos los triángulos elementales T_i dio lugar a una colección de figuras $\left\{S_i^1 : 1 \leq i \leq 2^{n-1}\right\}$ con las siguientes propiedades:

(a) Cada figura S_i^1 está formada por un triángulo T_i^1 semejante al triángulo $T_{2i-1} \cup T_{2i}$ más dos triángulos auxiliares A_{2i-1}^1, A_{2i}^1.

(b) Trasladando adecuadamente las figuras S_i^1 es posible obtener un triángulo D_1 (que es la unión de los triángulos T_i^1) semejante a T –con razón de semejanza α–, con base de longitud $2^n \alpha b$ y tal que

$$A\,(D_1) = A\left(\bigcup_{i=1}^{2^{n-1}} T_i^1\right) = \alpha^2 A\,(T)\,.$$

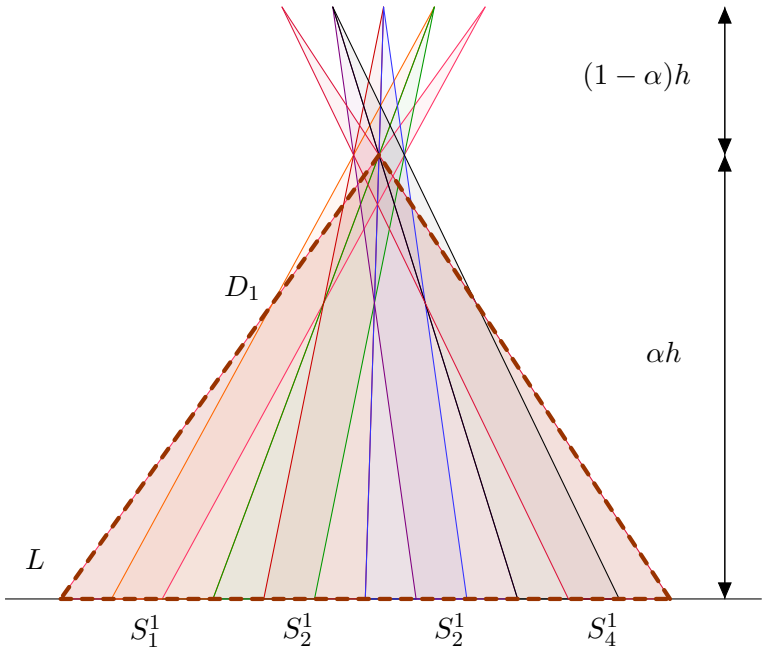


Figura 2.8: El triángulo D_1 que se obtiene en el *paso* 1 aparece con su frontera punteada.

(c) El área total de todos los triángulos auxiliares A_{2i-1}^1, A_{2i}^1 que acompañan a los triángulos T_i^1 no será mayor que $2\left(1-\alpha\right)^2 A\left(T\right)$.

(d)

$$\sum_{i=1}^{2^{n-1}} A\left(S_i^1\right) \leq \left[\alpha^2 + 2(1-\alpha)^2\right] A\left(T\right).$$

Ahora sí, ya podemos seguir la demostración.

Paso 2. En el segundo paso de la construcción centraremos nuestra atención en consecutivos S_i^1. Para cada i, $1 \leq i \leq 2^{n-2}$, traslademos a lo largo de L a S_{2i}^1 en la dirección de S_{2i-1}^1 un distancia $2\left(1-\alpha\right)\left(2\alpha b\right) = 2^2\left(1-\alpha\right)\alpha b$ para obtener 2^{n-2} figuras S_i^2, cada una formada por un triángulo T_i^2 semejante a $T_{2i-1}^1 \cup T_{2i}^1$ más dos triángulos auxiliares A_{2i-1}^2, A_{2i}^2. Además, al igual que en el paso 1, se pueden trasladar convenientemente las figuras S_i^2 para obtener una colección de figuras $\left\{S_i^2 : 1 \leq i \leq 2^{n-2}\right\}$ con las siguientes características:

1. Cada figura S_i^2 contiene un triángulo T_i^2, semejante a $T_{2i-1}^1 \cup T_{2i}^1$. Además, si D_2 es la unión de todos los T_i^2 (trasladados), D_2 es un triángulo semejante al triángulo D_1 del paso anterior y cuya área está dada por $A\left(D_2\right) = \alpha^4 A(T)$.

   (Ver figura 2.9.)

2. Los triángulos T_i^2 están acompañados de triángulos auxiliares cuya área total no es mayor que $2(1-\alpha)^2\alpha^2 A(T)$.

En efecto,

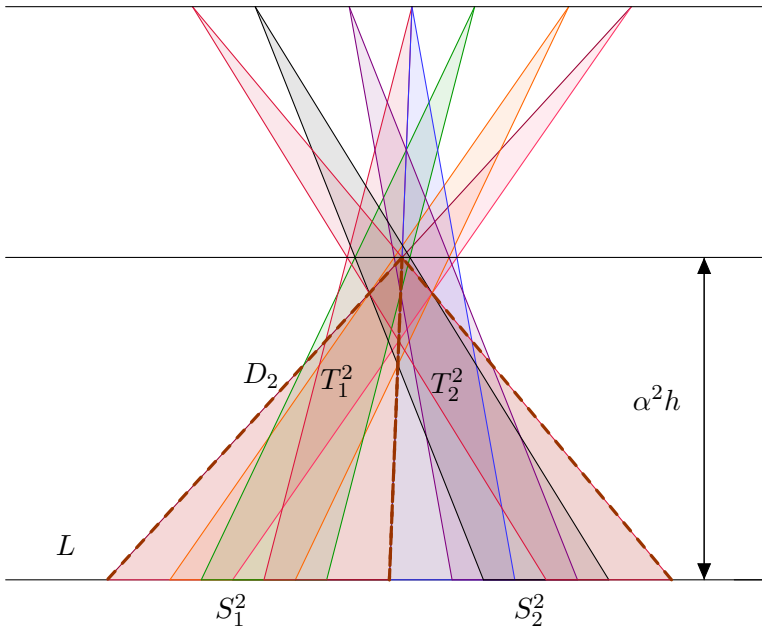


Figura 2.9: La construcción del *paso* 2 para $n = 3$. El triángulo D_2 es la unión de los triángulos T_1^2 y T_2^2, cuya frontera aparece punteada en la imagen.

1. La semejanza de T_i^2 con $T_{2i-1}^1 \cup T_{2i}^1$ es inmediata. Sea D_2 la unión de todos los triángulos (trasladados) T_i^2. Claramente D_2 y D_1 son semejantes. Ahora bien, podemos calcular la longitud de la base del triángulo D_2 restando el total del desplazamiento de las figuras S_{2i}^1 a la longitud de la base del triángulo D_1 del paso anterior:

   La base de D_1 es igual a $2^n\alpha b$, mientras que el desplazamiento total de las figuras S_{2i} es $2^n(1-\alpha)\alpha b$, ya que cada figura S_{2i} se desplazó $2^2(1-\alpha)\alpha b$ y hay 2^{n-2} de ellas. Por lo tanto, la longitud de la base del triángulo D_2 es

   $$2^n\alpha b - 2^n(1-\alpha)\alpha b = 2^n\alpha^2 b.$$

   De aquí se sigue que la razón de semejanza de los triángulos D_2 y D_1 es α. Por lo tanto

   $$A(D_2) = \alpha^2 A(D_1) = \alpha^4 A(T).$$

2. Para esta parte utilizaremos la misma técnica que en el lema 2.1.2 para calcular el área de los triángulos auxiliares. Cada T_i^2 está acompañado de dos triángulos auxiliares A_{2i-1}^2, A_{2i}^2. Por el punto de intersección de éstos, tracemos una recta l paralela a L. Esto da lugar a cuatro triángulos de áreas iguales; basta entonces calcular el área de uno de ellos, digamos, $A_{2i-1,1}^2$. (Ver figura 2.10.)

   Dicho triángulo es semejante a T_{2i}^1, con razón de semejanza $1-\alpha$ pues, mientras que el primero tiene base igual a $2(1-\alpha)\alpha b$, la del segundo mide $2\alpha b$. Luego,

$$A\left(A_{2i-1,1}^2\right) = (1-\alpha)^2 A\left(T_{2i}^1\right) = (1-\alpha)^2\alpha^2 hb.$$

   De aquí se sigue que $A\left(A_{2i-1}^2\right) = 2(1-\alpha)^2\alpha^2 bh$. Ya que puede ocurrir que los triángulos auxiliares se sobrepongan, el área total de éstos está acotada superiormente; es decir,

$$\begin{aligned} A\left(\bigcup_{i=1}^{2^{n-1}} A_{2i-1}^2\right) &\leq \sum_{i=1}^{2^{n-1}} A\left(A_{2i-1}^2\right) \\ &= 2^n(1-\alpha)\alpha^2 bh \\ &= 2(1-\alpha)\alpha^2 A\left(T\right). \end{aligned}$$

Para acotar el área total de las figuras S_i^2, $1 \leq i \leq 2^{n-1}$, tenemos que tener en cuenta el área del triángulo D_2, los triángulos auxiliares del paso 1 y los nuevos triángulos auxiliares obtenidos en este paso que acompañan a cada uno de los triángulos T_i^2. Así, se tiene que

$$\begin{aligned} \sum_{i=1}^{2^{n-2}} A(S_i^2) &\leq \alpha^4 A\left(T\right) + 2(1-\alpha)^2 A\left(T\right) + 2(1-\alpha)^2\alpha^2 A\left(T\right) \\ &= \left[\alpha^4 + 2(1-\alpha)^2\left(1+\alpha^2\right)\right] A\left(T\right). \end{aligned}$$

En resumen, la segunda aplicación del lema 2.1.2 para todas las figuras S_i^1 dio lugar a una colección de figuras $\left\{S_i^2 : 1 \leq i \leq 2^{n-2}\right\}$ con las siguientes propiedades:

(a) Cada figura S_i^2 está formada por un triángulo T_i^2 semejante al triángulo $T_{2i-1}^1 \cup T_{2i}^1$ más dos triángulos auxiliares A_{2i-1}^2, A_{2i}^2.

(b) Trasladando adecuadamente las figuras S_i^2 es posible obtener un triángulo D_2 (que es la unión de los triángulos T_i^2 trasladados) semejante a D_1 –con razón de semejanza α–, con base de longitud $2^n\alpha^2 b$ y tal que $A\left(D_2\right) = \alpha^4 A\left(T\right)$.

(c) El área total de todos los triángulos auxiliares A_{2i-1}^2, A_{2i}^2 que acompañan a los triángulos T_i^2 no será mayor que $2\left(1-\alpha\right)^2\alpha^2 A\left(T\right)$.

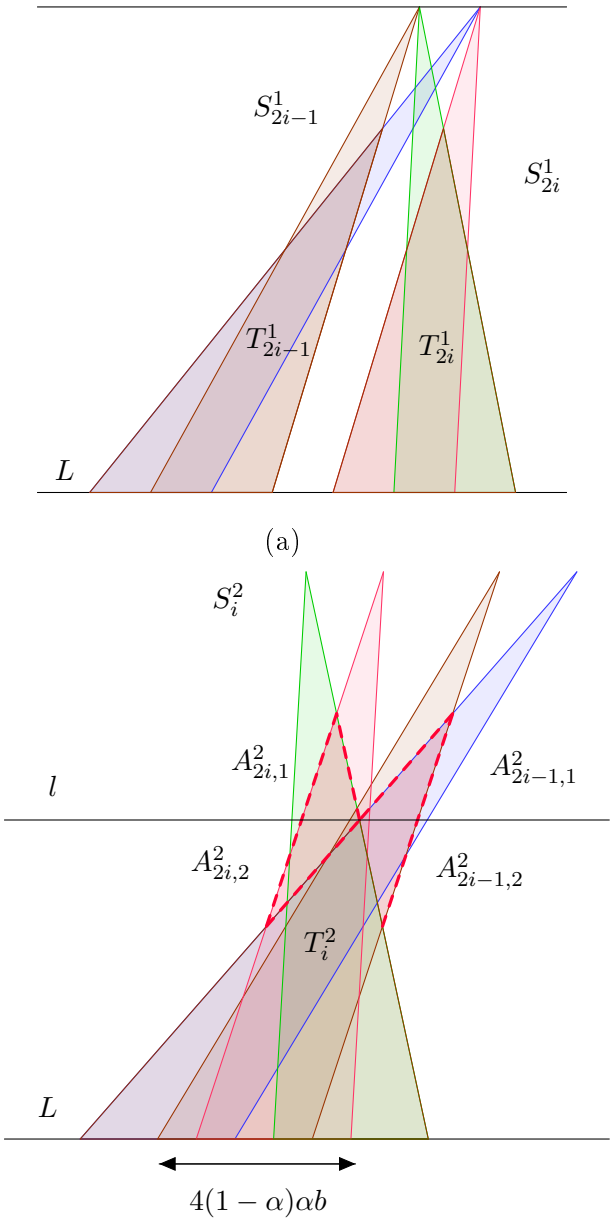


Figura 2.10: (a) A partir de S^1_{i-1} y S^1_i se obtiene la figura S^2_i. (b) Los triángulos auxiliares A^2_{2i-1} y A^2_{2i} aparecen con su frontera punteada.

(d)

$$\sum_{i=1}^{2^{n-2}} A(S^2_i) \leq \left[\alpha^4 + 2(1-\alpha)^2\left(1+\alpha^2\right)\right] A\left(T\right).$$

A este punto de la demostración ya podemos inferir lo que ocurre en cada etapa de la construcción: en el paso k $(2 \leq k \leq n)$ partimos de una colección de figuras $\left\{S_i^{k-1} : 1 \leq i \leq 2^{n-1}\right\}$ tal que

(i) Cada figura S_i^{k-1} contiene un triángulo T_i^{k-1}.

(ii) La unión de todas las figuras S_i^{k-1} es una figura formada por

1. Un triángulo D_{k-1}, que es la unión de todos los triángulos T_i^{k-1}, con base de longitud $2^n\alpha^{k-1}b$ y dividida en 2^{n-k+1} partes iguales de longitud $2^{k-1}\alpha^{k-1}b$. Más aún, $A\left(D_{k-1}\right) = \alpha^{2(k-1)}A\left(T\right)$.
2. Triángulos auxiliares cuya área total no será mayor que

$$\left[2\left(1-\alpha\right)\sum_{j=0}^{k-2}\alpha^{2j}\right]A\left(T\right). \tag{2.1}$$

Como la base del triángulo D_{k-1} está dividida en partes iguales, podemos aplicar el lema 2.1.2 a las figuras S_i^{k-1}. Para cada i $(1 \leq i \leq 2^{n-k})$ trasladamos a lo largo de la recta L a S_{2i}^{k-1} en la dirección de S_{2i-1}^{k-1} una distancia

$$2\left(1-\alpha\right)\left(2^{k-1}\alpha^{k-1}b\right) = 2^k\left(1-\alpha\right)\alpha^{k-1}b$$

para obtener 2^{n-k} figuras S_i^k, cada una de ellas formada por un triángulo T_i^k más dos triángulos auxiliares A_{2i-1}^k, A_{2i}^k.

Análogamente a lo hecho en los pasos 1 y 2 es posible realizar una traslación adecuada de las figuras S_i^k de modo que su unión contenga un triángulo D_k, semejante al triángulo D_{k-1}, con razón de semejanza α, junto con triángulos auxiliares. Más aún,

$$A\left(D_k\right) = \alpha^2 A\left(D_{k-1}\right) = \alpha^{2k}A\left(T\right),$$

y el área total de los triángulos auxiliares A_{2i-1}^k, A_{2i}^k que acompañan a los triángulos T_i^k no será mayor que $2(1-\alpha)^2\alpha^{2k-2}A\left(T\right)$.

En consecuencia, para acotar el área total de las figuras S_i^k, con $1 \leq i \leq 2^{n-k}$, tenemos que considerar el área del triángulo D_k, el área de los triángulos auxiliares de la etapa anteior (ecuación 2.1) y el área de los triángulos auxiliares de este paso. Con esto en cuenta, se sigue que

$$\sum_{i=1}^{2^{n-k}} A\left(S_i^k\right) \leq \left[\alpha^{2k} + 2\left(1-\alpha\right)\sum_{j=0}^{k-1}\alpha^{2j}\right]A\left(T\right).$$

Paso n. En el paso final, terminamos con una figura S (que siguiendo la notación usada antes correspondería a S_1^n), para la cual se cumple que

$$A(S) \leq \left[\alpha^{2n} + 2(1-\alpha)^2\sum_{i=0}^{n-1}\alpha^{2i}\right]A\left(T\right).$$

(Ver figura 2.11.)

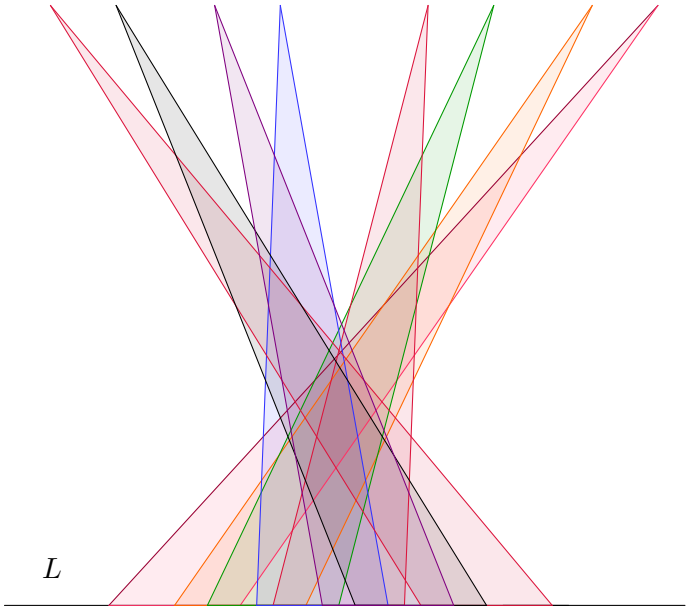


Figura 2.11: La figura que resulta en el último paso para $n = 4$.

Pero

$$\begin{aligned} 2(1-\alpha)^2 \sum_{i=0}^{n-1} \alpha^{2i} &\leq 2(1-\alpha)^2 \sum_{i=0}^{\infty} \alpha^{2i} \\ &= \frac{2(1-\alpha)^2}{1-\alpha^2} \\ &= \frac{2(1-\alpha)}{1+\alpha} \\ &< 2(1-\alpha), \end{aligned}$$

de donde,

$$A(S) < \left[\alpha^{2n} + 2(1-\alpha)\right] A(T).$$

Para finalizar la demostración mostremos que la figura S puede tener área tan pequeña como se desee. Sea $\varepsilon > 0$. Elijamos α, $\frac{1}{2} < \alpha < 1$, tan próximo a uno como sea necesario para que $(1-\alpha) < \dfrac{\varepsilon}{4A(T)}$ y $n \in \mathbb{N}$, suficientemente grande, de forma que $\alpha^{2n} < \dfrac{\varepsilon}{2A(T)}$. Con esta elección se puede garantizar que $A(S) < \varepsilon$. □

Observación 2.1.5. *Nótese que fijando la posición del primer triángulo elemental T_1, y realizando la construcción anterior con respecto a éste, en el paso final de la construcción del teorema 2.1.3, el vértice superior de cada T_i no se moverá a la izquierda del vértice superior de T_1 una distancia mayor que $2^n b$. De hecho, el vértice superior del triángulo T_{2^n}, que por construcción es el triángulo que más se desplaza en las construcción, se moverá una distancia igual a $2^n (1 - \alpha^n) b$ del vértice superior de T_1.*

En este sentido, es posible asegurar que, utilizando la construcción del teorema 2.1.3, se puede obtener una figura de área tan pequeña como se desee trasladando cada triángulo elemental una distancia no muy grande. La idea detrás es relativamente simple: dividir la base del triángulo original en triángulos elementales de bases suficientemente pequeñas y, después, aplicar la construcción anterior en cada uno de ellos. Este será el camino que seguiremos en la prueba del siguiente resultado que formaliza estas ideas.

Teorema 2.1.6. *Con la misma notación que en el teorema 2.1.3, sea $U \subseteq \mathbb{R}^2$ un conjunto abierto que contenga al triángulo T. Entonces, dado $\varepsilon > 0$, la construcción de S puede realizarse de manera que $S \subseteq U$ y $A(S) < \varepsilon$.*

Demostración. Sea $U \subseteq \mathbb{R}^2$ con las características del enunciado anterior y sea $\varepsilon > 0$. Realmente la región de U que nos interesa es aquella que se encuentra muy cerca de T, así, al ser éste último acotado, podemos suponer que U también lo es. (Más tarde esto facilitará recurrir a argumentos relacionados con la compacidad.) Para probar el resultado buscamos trasladar los puntos de T de manera que el área de éste sea muy pequeña, pero sin moverlos demasiado ya que buscamos permanecer dentro de U. En este sentido, primero vamos a encontrar una cota superior, digamos δ, con la propiedad que al trasladar en cualquier dirección a un punto arbitrario de T una distancia menor que δ, éste permanezca en U.

Con el fin de encontrar dicha δ, consideremos la función distancia $d_p : \partial U \longrightarrow \mathbb{R}$, donde ∂U denota la frontera de U p un punto del triángulo T, definida por $d_p(x, p) = \|x - p\|$. Se puede demostrar que d_p es una función continua en la frontera de U. Además, como ∂U es un conjunto compacto, la función distancia alcanza su ínfimo según el teorema del valor extremo. Con esto en cuenta, podemos considerar una nueva función $\psi : T \longrightarrow \mathbb{R}$ (que sí está bien definida por lo anterior) tal que, para cada punto $p \in T$,

$$\psi(p) = inf\{\|x - p\| : x \in \partial U\}.$$

La función ψ definida de este modo es continua y, por la compacidad de T y otra aplicación del teorema del valor extremo, ésta también alcanza su ínfimo; hagamos éste igual a δ. Escojamos ahora $n > \dfrac{b}{\delta}$ donde b es la longitud de la base del triángulo T. Dividamos la base de T en n triángulos elementales $T_1, T_2, ..., T_n$, y apliquemos el teorema 2.1.3 a cada uno de ellos (no al triángulo T, el teorema se aplica a cada T_i).

Consideremos algún T_i. Según el teorema 2.1.3 a dicho triángulo lo podemos dividir en triángulos elementales más pequeños cuya traslación tiene como resultado una figura S_i con $A(S_i) < \dfrac{\varepsilon}{n}$. Más aún, como se mencionó en la observación posterior al teorema 2.1.3, lo anterior puede realizarse moviendo cada triángulo elemental involucrado una

distancia más pequeña que la longitud de su base. En otras palabras, las figuras que conforman a S_i se han movido no más que $\frac{b}{n} < \delta$. Por tanto, los triángulos que forman S_i quedan contenidos en U. Repetir esto para cada T_i nos lleva una colección de figuras $\{S_i : i = 1, 2...n\}$. Sea

$$S = \bigcup_{i=0}^{n} S_i$$

Entonces

$$A(S) = A\left(\bigcup_{i=1}^{n} S_i\right) \leq \sum_{i=1}^{n} A\left(S_i\right) < \varepsilon,$$

y, en la obtención de S a partir de T, ninguna figura se ha movido una longitud mayor a δ, con lo cual $S \subseteq U$. □

2.2. Conjuntos de Besicovitch de área cero.

Para concluir este capítulo demostraremos la existencia de conjuntos de Besicovitch en $\mathbb{R}^2$ de área cero.

Teorema 2.2.1. *Existen conjuntos de Besicovitch en $\mathbb{R}^2$ de área cero.*

Demostración. Primero vamos a construir un conjunto B_0 de área cero que contiene un segmento de longitud uno en toda dirección posible dentro de un sector de 60°. Después, tomando la unión de B_0 con copias congruentes de éste rotadas 60° y 120°, se obtiene el conjunto requerido.

La construcción consistirá de aplicaciones repetidas del teorema 2.1.6. Sea S_1 un triángulo equilátero de altura uno y con base sobre una recta L. Sea, además, U_1 un conjunto abierto que contenga a S_1 y tal que $A\left(\overline{U_1}\right) \leq 2A\left(S_1\right)$, donde $\overline{U_1}$ denota la cerradura de U_1. Como sólo nos interesa lo que ocurre cerca de S_1, podemos suponer que U_1 es acotado (siempre puede tomarse la intersección de U_1 con una bola de radio suficientemente grande que contenga a S_1). Aplicando el teorema 2.1.6 a S_1 podemos obtener una nueva figura cerrada $S_2 \subseteq U_1$ con $A\left(S_2\right) \leq 2^{-2}$. Puesto que S_2 es la unión de un número finito de triángulos, podemos encontrar un conjunto abierto U_2 tal que $S_2 \subseteq U_2 \subseteq U_1$ y de forma que $A(\overline{U_2}) \leq 2A\left(S_2\right)$. De manera similar, es posible aplicar de nuevo el teorema 2.1.6 a cada triángulo elemental de la figura S_2 de tal forma que la figura resultante S_3 esté contenida en U_2 y que $A(S_3) \leq 2^{-3}$. Al igual que antes, $S_3 \subseteq U_3 \subseteq U_2$ para algún conjunto abierto U_3 con $A\left(\overline{U_3}\right) \leq 2A\left(S_3\right)$.

Repitiendo este proceso obtenemos una sucesión de figuras (S_i) así como una sucesión de conjuntos abiertos (U_i) que satisfacen la siguientes propiedades:

1. Cada conjunto U_i es abierto.
2. $S_i \subseteq U_i$.
3. $U_i \subseteq U_{i-1}$.

4. $A\left(\overline{U_i}\right) \leq 2A\left(S_i\right) \leq 2^{-i+1}$.

5. Cada S_i es la unión de un número finito de triángulos elementales con bases sobre L y de alturas de longitud uno. Así cada S_i contiene un segmento de longitud uno para cada ángulo (con respecto a L) en el intervalo $[60°, 120°]$.

Sea

$$B_0 = \bigcap_{i=0}^{\infty} \overline{U_i}.$$

Es importante destacar que B_0 es cerrado y está contenido en U_1, donde este último es acotado. Luego, B_0 es compacto. Veamos que este conjunto tiene las propiedades mencionadas al principio de la prueba.

(i) $A\left(B_0\right) = 0$. En efecto, dado $\varepsilon > 0$ arbitrario, por 4, $A(B_0) \leq 2^{-i+1} < \varepsilon$ para i suficientemente grande. Así, $0 \leq A\left(B_0\right) < \varepsilon$ para todo $\varepsilon > 0$. Por lo tanto, $A\left(B_0\right) = 0$, como se afirmó.

(ii) B_0 contiene un segmento de longitud uno en toda dirección dentro de un sector de 60°. Por construcción, cada S_i, y por tanto cada $\overline{U_i}$, tiene un segmento de longitud uno en cada dirección formando un angulo de al menos 60° con L. Debemos probar que esto es también cierto para B_0. Sea θ un ángulo en el intervalo $[60°, 120°]$. Según 5, para cada i existe $M_i \subseteq S_i$, un segmento de longitud uno en la dirección θ. Para toda i, el segmento M_i tiene extremos de la forma (x_i, y_i), $(x_i + cos(\theta), y_i + sen(\theta))$. Además, por 2, $M_i \subseteq \overline{U_i}$. Notemos que, para una j fija, $M_i \subseteq \overline{U_j}$ para $i \geq j$ pues la sucesión de conjuntos $\left\{\overline{U_i}\right\}_{i\in\mathbb{N}}$ es decreciente (utilizando la propiedad 3) y que $\overline{U_j}$ es un conjunto compacto. En consecuencia, la sucesión de extremos $\{(x_i, y_i)\}_{i\in\mathbb{N}}$ tiene una subsucesión que convergente. Sea (x, y) el punto al que converge dicha subsucesión. Análogamente, la sucesión de extremos $\{(x_i + cos(\theta), y_i + sen(\theta))\}_{i\in\mathbb{N}}$ contiene una subsucesión que converge al punto $(x + cos(\theta), y + sen(\theta))$.

Consideremos el segmento M de longitud uno cuyos extremos son los puntos (x, y) y $(x + cos(\theta), y + sen(\theta))$. Veamos que $M \subseteq B_0$. En efecto, puesto que $M_i \subseteq \overline{U_j}$ si $i \geq j$, por ser cada $\overline{U_j}$ cerrado se cumple que $M \subseteq \overline{U_j}$ para toda j. Por lo tanto

$$M \subseteq \bigcap_{i=0}^{\infty} \overline{U_i} = B_0.$$

□

Capítulo 3

Calles en todas direcciones ocupando área mínima.

El objetivo de este capítulo es describir la solución al problema de Kakeya. Para simplificar un poco, vamos a adoptar primero la siguiente definición.

Definición 3.0.1. *A un segmento de recta de longitud uno en* $\mathbb{R}^2$ *le llamaremos* **aguja**.

En estos términos, el problema de Kakeya puede enunciarse como sigue:

Problema 3.0.2. *Consideremos una aguja en el plano. ¿Cuál es la menor área requerida para rotar continuamente una aguja* 180°*, hasta regresar a su posición original pero invertida?*

En la introducción se mencionó que Besicovich utilizó las ideas que le permitieron construir el conjunto del teorema 2.2.1 para dar solución al problema de Kakeya. Es fácil observar que, aunque dicho conjunto contenga una aguja en cada dirección, las operaciones realizadas para su construcción imposibilitan el movimiento continuo de la aguja dentro de él. De manera sencilla, pero absolutamente eficaz, Julius Pál comunicó a Besicovitch la manera de remendar el problema. En la siguiente sección describiremos la sugerencia de Pál a Besicovitch para, finalmente, en la sección 3.2 presentar la solución al problema 3.0.2.

3.1. Ingeniosas traslaciones.

La sugerencia de Pál a Besicovitch consiste en utilizar las «conexiones de Pál». Claro, esto no dice nada al lector, pero antes de explicar qué son y cómo funcionan dichas conexiones, motivemos la idea de Pál con el siguiente problema:

Supongamos que tenemos una aguja AB sobre una recta l y queremos trasladarla continuamente hasta colocarla sobre el segmento $A'B'$ en otra recta l' (paralela a l), separada de la primera una distancia vertical h. ¿Cuál es la menor área posible que se debe ocupar para completar esta operación? (Ver figura 3.1(a).) Quizás, lo primero que venga a la mente sea hacerlo mediante una traslación en la dirección del vector que une a A con A'. Al realizarlo de este modo, el área que ocupamos es el área del paralelogramo que se muestra en la figura 3.1 (b), y que es igual a $1 \cdot h = h$.

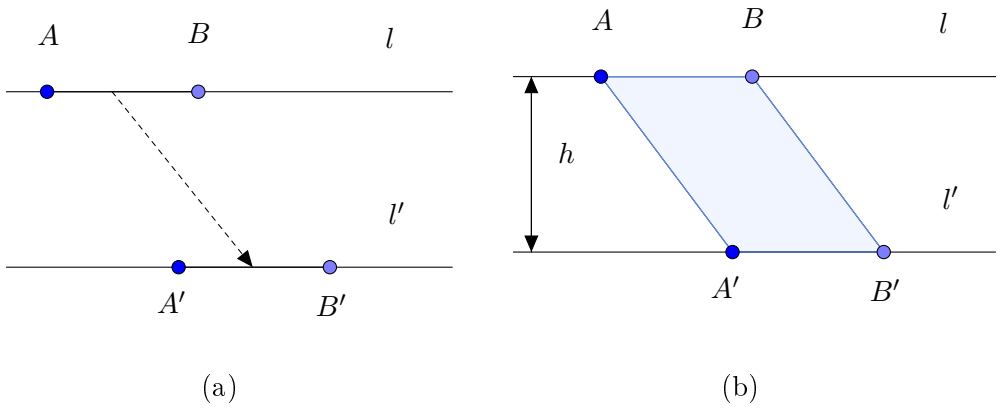


Figura 3.1: Traslación de una aguja de la recta l hasta otra paralela l'.

La solución que buscamos es la idea detrás de las «conexiones de Pál»: primero, tracemos el segmento AB'. Si rotamos AB alrededor del extremo A hasta sobreponerse con AB', podemos deslizar la aguja hasta que B y B' coincidan. Por último, basta rotar AB alrededor de B para colocarla sobre la recta l'. La figura 3.2 muestra lo anterior. El área que abarcamos con esta operación corresponde a los sectores descritos durante las rotaciones de AB. Lo conveniente de esta construcción es que la rotación –y, por tanto, el área de los sectores– puede realizarse girando un ángulo muy pequeño.

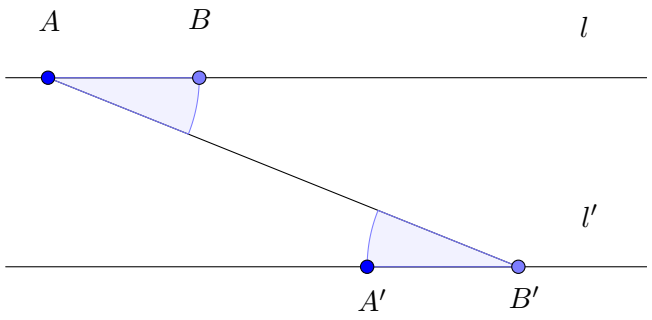


Figura 3.2: Ejemplo de una conexión de Pál.

Si tomamos en cuenta que el objetivo de Besicovitch era remediar el problema del movimiento de la aguja sin agregar demasiada área al conjunto que había logrado construir, la sugerencia de Pál cae como anillo al dedo. El siguiente lema presenta las ideas hasta aquí discutidas de forma rigurosa.

Lema 3.1.1. *Sean L_1 y L_2 rectas paralelas en el plano. Entonces, dado $\varepsilon > 0$, existe un conjunto E, con $A(E) < \varepsilon$, que contiene a dichas rectas y tal que una aguja puede*

moverse continuamente de L_1 a L_2 sin salirse de E.

Demostración. Sean x_1 y x_2 puntos en las rectas L_1 y L_2, respectivamente. Definamos E como el conjunto que contiene a L_1 y L_2, el segmento M que une los puntos x_1 y x_2, y los sectores de círculos con centros en x_i, radio 1 y el menor de los ángulos que forman las rectas L_i con M $(i = 1, 2)$. Es claro que el área total de E puede hacerse tan pequeña como se desee haciendo x_1 y x_2 suficientemente alejados entre sí. Más aún, el segmento de longitud uno puede moverse de L_1 a L_2 rotando a éste en el primer sector tomando un área menor que $\frac{\varepsilon}{2}$ –lo que permite un ángulo aproximadamente igual a ε–, trasladarlo a lo largo de M y, por último, rotarlo en el segundo sector. (Ver figura 3.3.) □

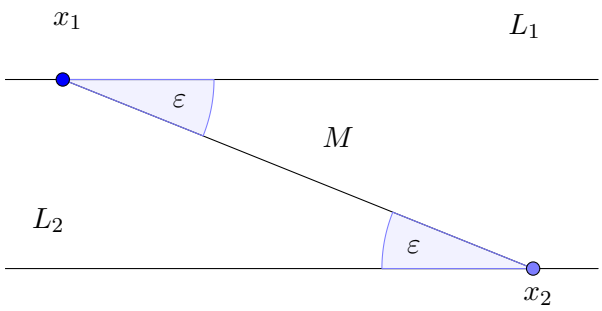


Figura 3.3: Podemos construir conexiones de Pál con un área tan chica como queramos.

3.2. La solución al problema de la aguja.

En esta sección daremos solución al problema de Kakeya construyendo un *conjunto de Kakeya*. La siguiente definición aclara a qué nos referimos con dicho conjunto:

Definición 3.2.1. *Llamaremos* **conjunto de Kakeya** *a cualquier conjunto de $\mathbb{R}^2$ en el que una aguja puede invertirse continuamente hasta su posición original.*

Observación 3.2.2. *Usualmente no se hace distinción en la literatura entre un «conjunto de Besicovitch» y un «conjunto de Kakeya». Aquí optamos por distinguirlos debido a que para el segundo se tiene la condición extra del movimiento para la aguja.*

Los árboles de Perron–Schoenberg junto con las conexiones de Pál son los ingredientes esenciales de la solución para el problema de Kakeya:

Partamos de un triángulo equilátero de altura uno como en la figura 3.4. Formemos ahora un árbol de Perron- Schoenberg como en el teorema 2.1.3 (ver figura 3.5). Como

ya se mencionó, la aguja no puede moverse continuamente dentro de esta figura, sin embargo, sí puede rotarse dentro de los triángulos elementales que forman el árbol. Con esto en cuenta, fijemos nuestra atención en el subtriángulo (T_I) que tiene por lado el lado izquierdo del triángulo original y en aquel (T_D) que tiene el lado derecho. Como cada triángulo elemental tiene un lado paralelo al lado de otro subtriángulo, podemos usar las conexiones de Pál, de área tan pequeña como queramos, para mover la aguja entre dos triángulos con lados paralelos. Así, para completar el movimiento de la aguja en un arco de 60° necesitamos rotarla desde el lado izquierdo de T_I hasta el lado derecho de T_D usando el resto de los triángulos elementales junto con las conexiones para dar saltos de uno a otro.

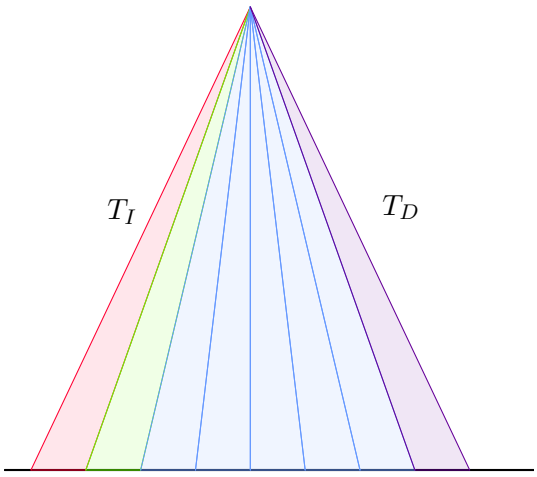


Figura 3.4: Triángulo con subtriángulos distinguidos T_I y T_D.

En resumen, la estrategia para resolver el problema de Kakeya es la siguiente:

1. Coloque la aguja sobre el lado izquierdo del triángulo T_I con uno de sus extremos sobre el vértice superior (ver figura 3.6).

2. Rote la aguja, en el sentido contrario de las manecillas del reloj, alrededor del vértice superior hasta llegar al lado derecho de T_I. Después, localice el triángulo con lado paralelo al lado derecho de T_I. (Ver figura 3.7.)

3. Utilice las conexiones de Pál para llevar la aguja hasta el triángulo con lado paralelo al lado derecho de T_I. (Ver figuras 3.8 3.9 y 3.10.)

4. Continúe hasta llegar al lado derecho del triángulo T_D.

Por último, agregue dos copias del árbol de Perron–Schoenberg rotadas 60° y 120°; realice los mismos pasos anteriores en cada una de ellas para terminar la rotación de

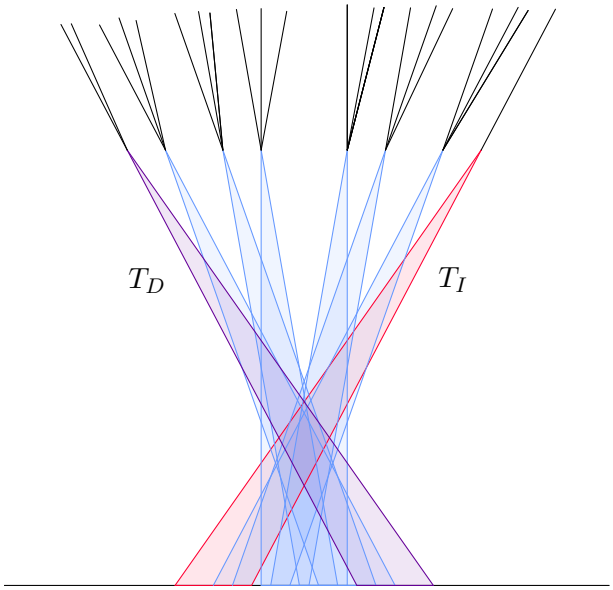


Figura 3.5: Movimiento de la aguja desde T_I hasta T_D dentro de un árbol de Perron-Schoenberg.

180° y sitúe la aguja ya invertida en su posición original. En la figura 3.11 se aprecia el resultado de este paso final.

La meta original era realizar todo lo anterior ocupando un área tan pequeña como se desee. El teorema a continuación establece rigurosamente esto.

Teorema 3.2.3. *Para cada $\varepsilon > 0$, existe un conjunto de Kakeya K en $\mathbb{R}^2$ de área menor que ε.*

Demostración. Sea T un triángulo equilátero de altura uno y con base en una recta L. Utilizando el teorema 2.1.3 podemos dividir a T en $m = 2^k$ triángulos elementales y deslizarlos a lo largo de L para obtener una figura S_k, con $A(S_k) < \frac{\varepsilon}{6}$, que es la unión de los triángulos $T_1', T_2', ..., T_m'$ (es decir, T_i' es el resultado de trasladar el triángulo T_i'). Más aún, S_k contiene una aguja en toda dirección dentro de un intervalo de 60°. Tomando tres copias apropiadamente rotadas de S_k se obtiene un conjunto K_1 con $A(K_1) < \frac{\varepsilon}{2}$ y que contiene una aguja en toda dirección.

Ahora, K_1 es la unión de $3m$ triángulos elementales T_i $(1 \leq i \leq 3m)$; además, para cada i, un lado de T_i' es paralelo al lado opuesto de T_{i+1}', por lo cual, usando el lema 3.1.1, podemos agregar un conjunto K_2 a K_1, que es la unión de $3m - 1$ conexiones de Pál entre los los triángulos T_i' y T_{i+1}, de área menor que $\frac{\varepsilon}{2}$. Lo anterior da como resultado un conjunto $K = K_1 \cup K_2$, de área menor que ε, dentro del cual es posible rotar continuamente una aguja 180°. □

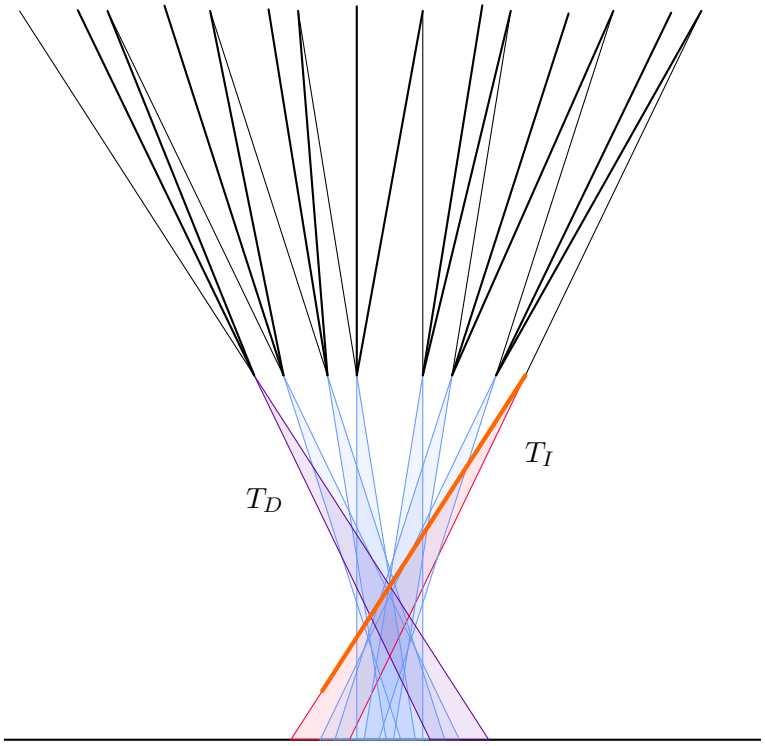


Figura 3.6: Aguja sobre el lado izquierdo del triángulo T_I.

¡Problema resuelto!

John Edensor Littlewood, en su famoso texto *A Mathematician's Miscellany* (Lit53), menciona que la solución que Besicovitch dio al problema de Kakeya establece el primero de dos resultados contraintuitivos en el trabajo de Besicovitch. El segundo se dio a conocer en 1947 cuando Besicovitch dio solución al *problema de Crum* (ver (Bes47)). Este problema pregunta por el máximo número de poliedros convexos no sobrepuestos para los que, por pares, comparten al menos una parte de sus caras. La respuesta en dimensión dos es 4 (apelando al teorema de los 4 colores) y se pensaba que para el caso de dimensión tres el número estaba entre 10 o 12. De hecho, para dimensión tres, Besicovitch demostró que ¡hay un número infinito de dichos poliedros! Así, para una pregunta bastante sencilla Besicovitch dio la sorprendente respuesta de ser cero y para otra razonable pregunta dio la igualmente sorprendente respuesta de ser infinito. Cabe mencionar que en 1958 la Asociación Americana de Matemáticas produjo un filme sobre el problema de Kakeya donde el mismo Besicovitch expone su solución al problema (ver (oA62)). Cualquier oportunidad de ver este trabajo no debe ser desaprovechada. También pueden ser de interés (Num15) y (Mat15a).

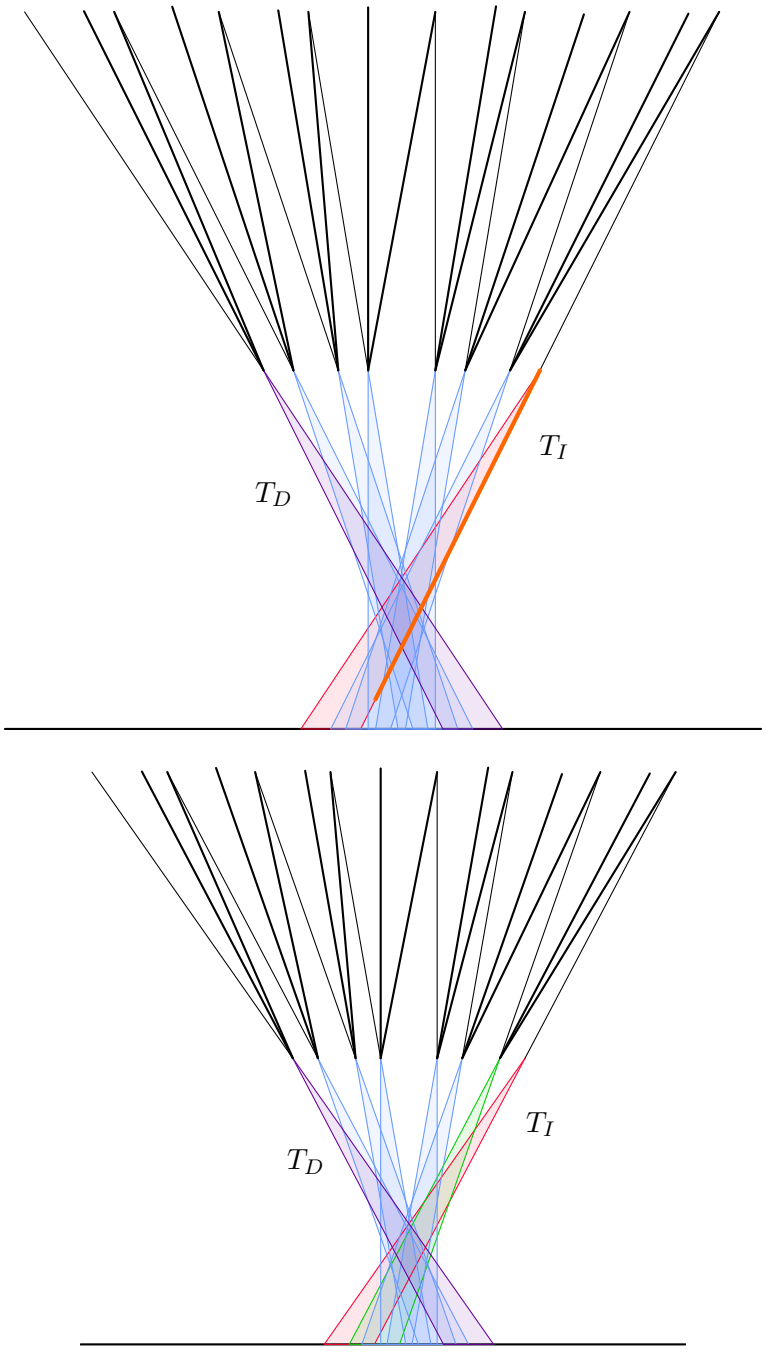


Figura 3.7: Aguja sobre el lado derecho del triángulo T_I. El triángulo de color verde tiene un lado paralelo al lado derecho del triángulo T_I (ver figura 3.4.

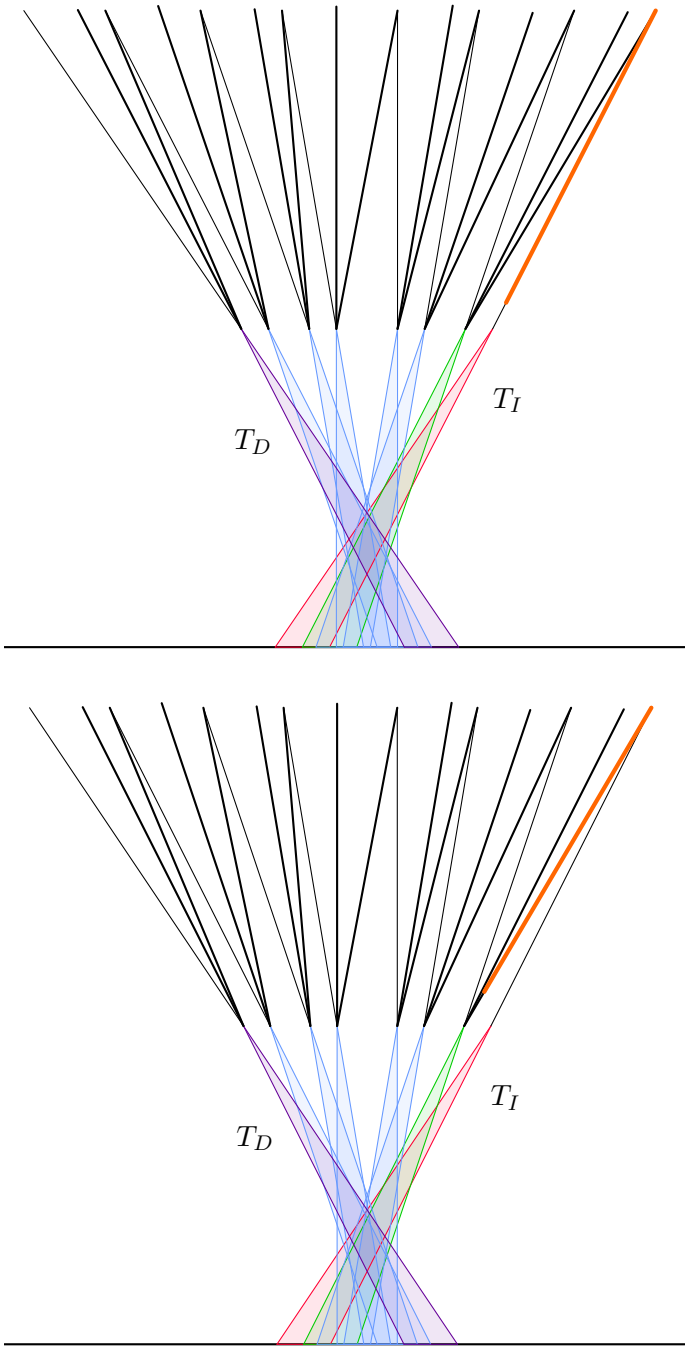


Figura 3.8: Uso de las conexiones de Pál para mover la aguja.

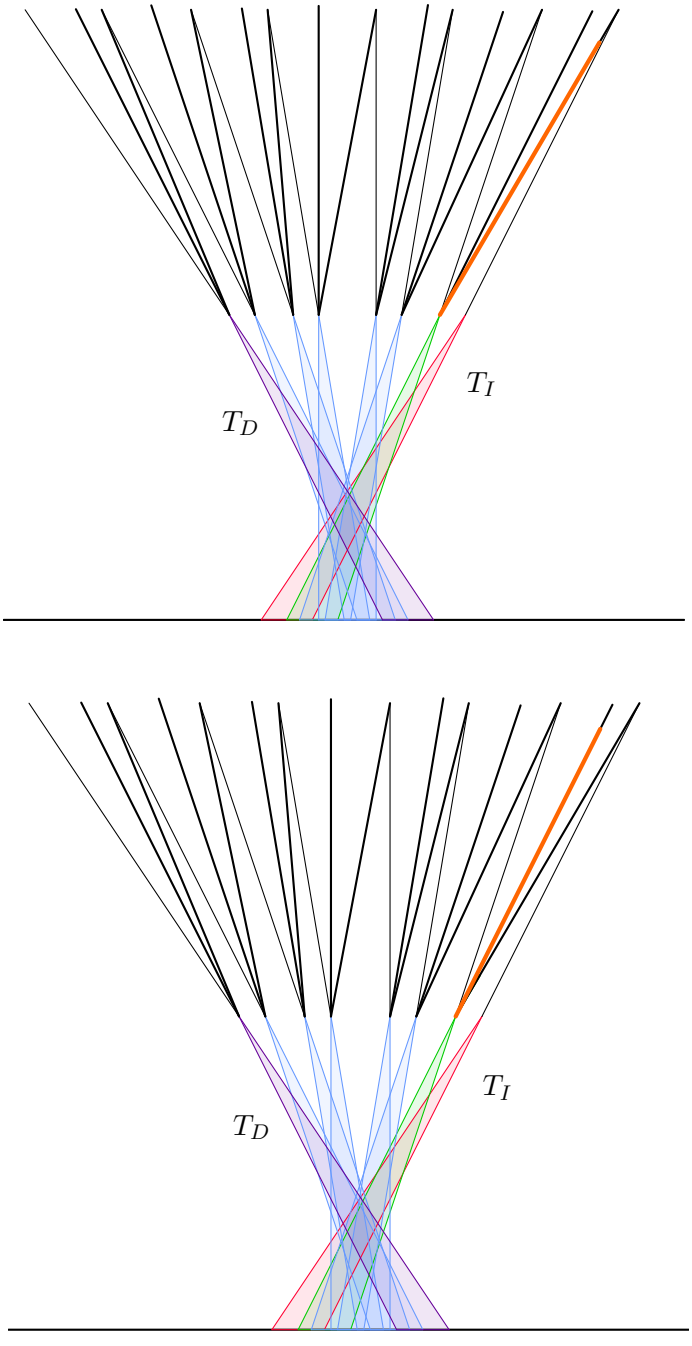


Figura 3.9: Uso de las conexiones de Pál para mover la aguja.

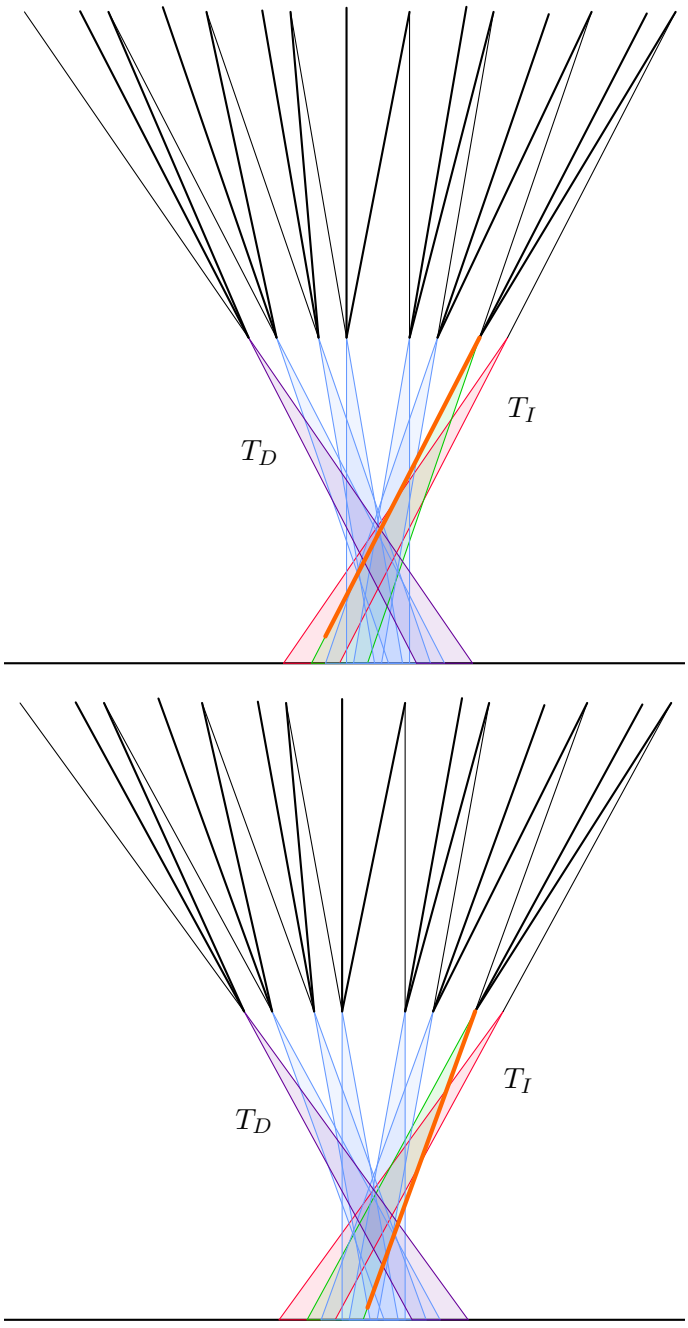


Figura 3.10: Uso de las conexiones de Pál para mover la aguja.

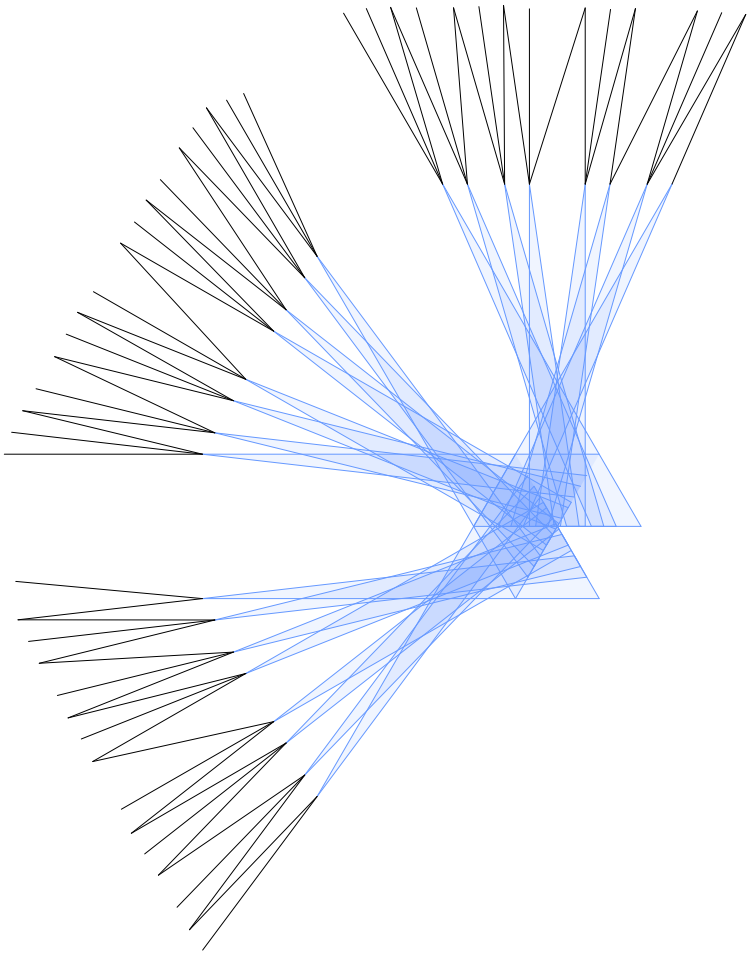


Figura 3.11: Ejemplo de un conjunto de Kakeya.

Parte II

El problema finito de Kakeya.

Capítulo 4

Introducción.

Por el sencillo planteamiento del problema de Kakeya, podría pensarse que éste es únicamente un poco más que una curiosidad matemática. Esto, sin embargo, no podría estar más alejado de la realidad. En recientes años, la palabra «Kakeya» aparece cada vez con mayor frecuencia en la literatura matemática, y los conceptos relacionados con el problema original han llevado a establecer nuevos vínculos con áreas insospechadas de las matemáticas, entre las que destacan, por mencionar algunas, análisis armónico, trasformada de Fourier, combinatoria, teoría de números y hasta ecuaciones de onda. (Excelentes referencias donde se explican las conexiones mencionadas son (Wol99), (Bou00) y (Tao01).) En este sentido, lo que se ha llegado a conocer como la *conjetura de Kakeya* ha jugado un papel fundamental para establecer nuevos horizontes alrededor del problema de Kakeya. Para poder enunciar dicha conjetura, recordemos que un conjunto de Besicovith en $\mathbb{R}^n$ es un subconjunto compacto $B \subseteq \mathbb{R}^n$ que contiene un segmento de recta de longitud uno en cualquier dirección [1]. En estos términos la *conjetura de Kakeya* establece lo siguiente:

Conjetura 4.0.1 (Conjetura de Kakeya)**.** *Un conjunto de Besicovitch en $\mathbb{R}^n$ tiene dimensión* [2] *igual a n.*

Para $n = 1$ la conjetura es inmediata, y para $n = 2$ fue completamente resuelta por Roy O. Davies (Dav71). Sin embargo, permanece abierta en los demás casos y parece volverse cada vez menos asequible conforme la dimensión aumenta. Hoy en día se considera uno de los mayores problemas abiertos en la *teoría geométrica de la medida*.

Dadas las numerosas conexiones de la conjetura 4.0.1 con otras ramas de las matemáticas, así como los pocos avances significativos en su solución, se han considerado diversos análogos a ésta con la esperanza de obtener nuevas ideas que permitan establecerla de manera completa. Como consecuencia natural, se ha desarrollado una maqui-

[1] Al igual que en el caso del plano, se sabe que existen conjuntos de Besicovitch en $\mathbb{R}^n$ de *medida de Lebesgue cero* (ver (Wol99)). El lector puede pensar el concepto de *medida de Lebesgue* como una generalización a mayores dimensiones del concepto de área y volumen.

[2] De forma más precisa, se pregunta por la *dimensión Hausdorff* o *Minkowski*. No entraremos en detalles sobre estos conceptos ya que no es el tema central del presente trabajo. Para una completa exposición sobre las definiciones y demás propiedades sugerimos al lector consultar (Fal86), (Mat15b) y (SS09).

naria de nuevas técnicas que, en muchas ocasiones de forma inesperada, han ayudado a resolver problemas aparentemente pertenecientes a un contexto completamente ajeno.

En este trabajo estaremos interesados en un análogo particular a la conjetura de Kakeya, conocido como ***el problema finito de Kakeya***, propuesto por Thomas Wolff en 1999 (Wol99). El enunciado es un «modelo a escala» de la conjetura 4.0.1, y es extremadamente conveniente pues evita todas las tecnicidades involucradas en el concepto de *dimensión* (*dimensión Hausdorff* o *Minkowski*). Antes de entrar con todo detalle al *problema finito de Kakeya*, primero establezcamos los conceptos necesarios.

Notación 4.0.2. *En todo lo que sigue, $\mathbb{F}$ denotará un campo (y a menos que se haga alguna precisión, como ser finito o infinito, se pensará en un campo arbitrario). Además, dado un campo $\mathbb{F}$, denotaremos por $\mathbb{F}^n$ al espacio vectorial sobre $\mathbb{F}$ (también llamado $\mathbb{F}$–espacio vectorial) de dimensión n.*

Definición 4.0.3. *Sea $\mathbb{F}$ un campo. Para cada par de vectores $x, y \in \mathbb{F}^n$, con $x \neq 0$, definimos la recta $l(y;x) \subseteq \mathbb{F}^n$ que pasa por y en la dirección de x como*

$$l(y;x) = \{y + tx : t \in \mathbb{F}\}.$$

Con esto en cuenta, pasamos ahora a definir un *conjunto finito de Kakeya*.

Definición 4.0.4. *Sea $\mathbb{F}$ un campo finito. Un conjunto $K \subseteq \mathbb{F}^n$ es un* **conjunto finito de Kakeya** *si, para cada $x \neq 0$ en $\mathbb{F}^n$, existe $y = y_x \in \mathbb{F}^n$ tal que $l(y;x) \subseteq K$.*

Para todo lo que sigue, nos referiremos a un conjunto finito de Kakeya simplemente como conjunto de Kakeya. Además, la gran ventaja de considerar estos conjuntos de Kakeya es que la única noción de *tamaño* a considerar es la de cardinalidad. Como estamos trabajando en campos finitos al considerar conjuntos de Kakeya, la cardinalidad de estos conjuntos será siempre finita. La cardinalidad de un conjunto será denotada por $|\cdot|$.

Podemos entonces enunciar el *problema finito de Kakeya* en este contexto.

Conjetura 4.0.5 (Problema finito de Kakeya)**.** *Sea $\mathbb{F}$ un campo finito. Si $K \subseteq \mathbb{F}^n$ es un conjunto de Kakeya, entonces*

$$|K| \geq C_n |\mathbb{F}|^n,$$

donde $C_n > 0$ depende sólo de n, pero no de la cardinalidad de $\mathbb{F}$. [1]

Para $n = 1$ la solución es inmediata (el único conjunto de Kakeya es $\mathbb{F}$), así que el problema se vuelve interesante para $n \geq 2$. El mismo Wolff estableció en (Wol99) una cota de la forma $C_n|\mathbb{F}|^{\frac{n+2}{2}}$ (obsérvese que esto resuelve la conjetura 4.0.5 para $n = 2$). Subsecuentemente esta cota fue mejorada en (Rog01), (BKT04), (MT$^+$04) y (Tao05) tanto para el caso general n como para pequeños valores de éste ($n = 3, 4$). Pese a

[1] En la conjetura 4.0.5 se debe pensar a n como fijo. Además, algunos autores precisan que se deben considerar valores muy grandes para la cardinalidad del campo $\mathbb{F}$.

que importantes matemáticos trabajaron en el problema finito de Kakeya, por ejemplo Terence Tao, ganador de la medalla Fields, el avance era similar al de la conjetura 4.0.1 original (aunque hay que decirlo, el problema llevaba relativamente poco de haberse propuesto y recibió mucha menos atención que su análogo en el caso euclidiano). Hasta antes del 2009 la mejor cota para la conjetura 4.0.5 era de la forma $C_n|\mathbb{F}|^{\frac{4n}{7}}$ ((Rog01) y (MT$^+$04), basados en resultados de (KT99)). Sin embargo, en ese mismo año, Zeev Dvir sorprendió a la comunidad matemática tras resolver en (Dvi09) el problema finito de Kakeya utilizando un breve (¡tan sólo una página de extensión!) y hermoso argumento que explota el comportamiento de ciertos polinomios en los conjuntos de Kakeya; una prueba digna de «El libro» (AZ14).

Las técnicas introducidas por Dvir en (Dvi09) forman parte de lo que ahora se conoce como ***el método polinomial***, el cual busca imponer una estructura algebraica a problemas geométricos utilizando polinomios. En los últimos años ha sido de gran utilidad para la solución de problemas en combinatoria (ver (Tao13), (Dvi12), (UoTv16), (oTv16) y (Fra13)); además ha tenido una influencia importante en las ciencias de la computación (ver (Juk11) y (Gut16)). Sin embargo, como menciona Terence Tao en (Tao09), es una pena que el *método polinomial* sea fundamentalmente dependiente de la estructura algebraica de los campos finitos, ya que dificulta una aplicación directa para resolver la conjetura de Kakeya en el caso euclidiano. Sin embargo, el *método polinomial* es interesante por sí mismo y, en la medida de lo posible, en el presente trabajo se buscará resaltar lo atractivo de esta nueva técnica. Así, el objetivo está marcado: lo que nos ocupará de aquí en adelante será presentar la solución de Dvir al problema finito de Kakeya con todo detalle, preservando el espíritu pionero –salvo por algunas observaciones sugeridas a Dvir por Terence Tao y Noga Alon– de la prueba original de Dvir para apreciar el *método polinomial* en todo su esplendor.

Capítulo 5

El método polinomial.

A grandes rasgos, hay dos ingredientes esenciales en la demostración de Dvir de la conjetura finita de Kakeya. El primero de ellos es la generalización a polinomios en varias variables de la siguiente proposición:

Proposición 5.0.1. *Sea $\mathbb{F}$ un campo y $d \geq 1$ un entero. Denotemos por $\mathbb{F}[x]$ al conjunto de los polinomios en x con coeficientes en $\mathbb{F}$.*

(i) Si $p(x) \in \mathbb{F}[x]$ es un polinomio no cero de grado a lo más d, entonces

$$|\{x \in \mathbb{F} : p(x) = 0\}| \leq d.$$

(ii) Dado un conjunto $A \subseteq \mathbb{F}$ de cardinalidad a lo más d, existe un polinomio diferente de cero $p(x) \in \mathbb{F}[x]$ de grado a lo más d tal que $p(a) = 0$, para cada $a \in A$.

Si ya se cuenta con cierta experiencia estudiando polinomios en x con coeficientes en $\mathbb{R}$ o $\mathbb{C}$, la proposición anterior debe ser un tanto familiar. Sin embargo, las ideas desarrolladas por Dvir exigen introducirnos en un contexto más general, contemplando un campo arbitrario $\mathbb{F}$. De hecho, cuando nos adentremos en su prueba de la conjetura finita de Kakeya, será evidente la dependencia absoluta de un tipo particular de campos: los campos finitos. [1]

Por ahora no se mencionará nada del segundo ingrediente, pues es más atractivo abordarlo una vez que se haya generalizado la proposición 5.0.1. En este sentido, la tarea que nos ocupará en este capítulo será, primero, probar en la sección 5.1 la proposición 5.0.1; después, en la sección 5.2, veremos su generalización a polinomios en varias variables. Finalmente se abordará el segundo ingrediente de la demostración de Dvir.

5.1. Polinomios en una variable.

Retomando la posible familiaridad que pudiese existir con los polinomios en x con coeficientes en $\mathbb{R}$ o $\mathbb{C}$, también deben conocerse conceptos como el *grado de un polinomio*, *polinomio cero*, así como las operaciones de *suma* y *multiplicación* entre polinomios. Así,

[1] Ver apéndice B.

para empezar, vamos a recuperar todas estas nociones considerando *polinomios en x con coeficientes en un campo arbitrario.*

Primero vamos a definir lo que es un *polinomio en x con coeficientes en un campo.*

Definición 5.1.1. *Sea $\mathbb{F}$ un campo. Un* ***polinomio*** *$p(x)$* ***con coeficientes en*** *$\mathbb{F}$ es una suma formal infinita*

$$\sum_{i\in\mathbb{N}} a_i x^i$$

donde $a_i \in \mathbb{F}$ y $a_i = 0$ para todos, salvo un número finito de valores de i.

Llamaremos a los valores a_i los **coeficientes** de $p(x)$. Además, diremos que a_i es el **coeficiente** de x^i en $p(x)$, para cada $i = 0, 1, ...$ Si para alguna $i > 0$ ocurre que $a_i \neq 0$, el más grande de dichos valores de i es el **grado** de $p(x)$, el cual se denotará por $\deg(p)$. De ser el caso que no existiese dicha $i > 0$, si además $a_0 \neq 0$, entonces $p(x)$ es de **grado cero**. Por último, si $a_i = 0$ para cada valor de i, llamaremos a $p(x)$ el **polinomio cero**, denotado por 0, y definimos su grado como -1[1].

Por supuesto, siguiendo estrictamente a la definición 5.1.1, para $\mathbb{F} = \mathbb{R}$, ¡$2 + x^2$ no sería un polinomio! Siempre tendríamos que escribir $2 + 0x + 1x^2 + 0x^3 + ...$ En otras palabras, en la práctica, estamos acostumbrados a identificar un polinomio con expresiones finitas de la forma

$$\sum_{i=0}^{n} a_i x^i = a_n x^n + ... + a_1 x + a_0.$$

Por lo tanto, acordemos que si en $p(x) = a_0 + a_1 x + ... + a_n x^n + ...$ ocurre que $a_i = 0$ para toda $i > n$, entonces podemos denotar $p(x)$ por $a_n x^n + ... + a_1 x + a_0$. Esto es una convención práctica, y no una buena manera de definir lo que es un polinomio, ya que, ciertamente, si aceptáramos la definición de un polinomio como una suma formal finita, los polinomios $0 + a_1 x$ y $0 + a_1 x + 0x^2$ con coeficientes en algún campo $\mathbb{F}$ serían distintos como sumas formales, pero queremos entenderlos como el mismo polinomio tal como lo hacemos en la práctica. Siguiendo la definición 5.1.1, ya no existe este problema.

Otros acuerdos que facilitarán la notación son los siguientes: si alguna $a_i = 1$, podemos quitarla de la suma formal, así que identificaremos, por ejemplo, $2 + x$ con el polinomio $2 + 1x$ con coeficientes en $\mathbb{R}$. Por último, convengamos que es posible omitir de la suma formal cualquier término $0x_i$ o $a_0 = 0$, si $a_0 = 0$ pero no todas las $a_i = 0$. De este modo, por ejemplo, $0, 5, x, 7 + x^3$ son todos polinomios con coeficientes en $\mathbb{R}$.

Antes de continuar, vale la pena dar algunos ejemplos sobre el grado de un polinomio. Consideremos los polinomios $7x^5 + 3x^2 + 8$, -9 y 0 con coeficientes en $\mathbb{R}$. Entonces, $\deg(7x^5 + 3x^2 + 8) = 5$, $\deg(-9) = 0$ y $\deg(0) = -1$. Dicho sea de paso, cuando un elemento del campo $\mathbb{F}$ se considera como un polinomio, a éste se le llama **polinomio constante**.

[1] Algunos autores definen el grado del polinomio cero como $-\infty$; otros, por el contrario, adoptan no definirlo. Nosotros seguiremos la definición que aparece en (AZ14).

La suma y multiplicación de polinomios con coeficientes en un campo $\mathbb{F}$ están definidas de la manera usual. Si

$$f(x) = a_0 + a_1x + ...a_nx^n + ...$$

y

$$g(x) = b_0 + b_1x + ... + b_nx^n + ...$$

entonces, para el polinomio suma, se tiene

$$f(x) + g(x) = c_0 + c_1x + ... + c_nx^n + ...,$$

donde $c_n = a_n + b_n$, y, para el polinomio multiplicación, tenemos

$$f(x)g(x) = d_0 + d_1x + ... + d_nx^n + ...,$$

donde $d_n = \sum_{i=0}^{n} a_ib_{n-i}$, para cada $n \in \mathbb{N}$. De nuevo, c_i y d_i ambas son cero para todos, excepto un número finito de valores de i, así que las definiciones anteriores tienen sentido.

Con lo anterior en consideración, tenemos el siguiente teorema:

Teorema 5.1.2. *El conjunto $\mathbb{F}[x]$ de todos los polinomios en x con coeficientes en un campo $\mathbb{F}$, es un anillo bajo la suma y multiplicación polinomial.*

Así, $\mathbb{Q}[x]$ es el anillo de los polinomios en x con coeficientes racionales; $\mathbb{Z}_p[x]$ (p primo) el de los polinomios con coeficientes en los enteros módulo p, etc.

Ejemplo 5.1.3. En $\mathbb{Z}_3[x]$ tenemos

$$(x^3 + 2x + 1) + (x + 1) = x^3 + (2 + 1)x + (1 + 1) = x^3 + 2.$$

Todavía en $\mathbb{Z}_3[x]$ se tiene que

$$(x + 2)(2x^2 + x + 1) = 2x^3 + (1 + 1)x^2 + (2 + 1)x + 2 = 2x^3 + 2x^2 + 2.$$

Otro importante aspecto sobre polinomios, usado quizá de forma muy ingenua en la práctica, es la «evaluación de un polinomio». La siguiente definición precisa a qué nos referimos con lo anterior.

Definición 5.1.4. *Sea $\mathbb{F}$ un campo. Para cualquier $p(x) \in \mathbb{F}[x]$, podemos asociar al polinomio $p(x)$ una función $p : \mathbb{F} \longrightarrow \mathbb{F}$ como sigue: si*

$$p(x) = a_nx^n + a_{n-1}x^{n-1} + ... + a_1x + a_0,$$

entonces, para cada $r \in \mathbb{F}$, se define $p(r) \in \mathbb{F}$ como

$$p(r) = a_nr^n + a_{n-1}r^{n-1} + ... + a_1r + a_0.$$

La función p es llamada ***función polinomial inducida por*** *$p(x)$.*

Por ejemplo, si consideramos el polinomio $p(x) = x^2 + 1$ con coeficientes en $\mathbb{R}$, podemos *evaluar* $p(x)$ en, digamos, $\sqrt{2}$ para obtener $p\left(\sqrt{2}\right) = 3$. No entraremos en la formalización de este hecho, pero sí nos valdremos de él en lo sucesivo. Una excelente referencia donde se explica de manera rigurosa y detallada el trasfondo de la «evaluación de un polinomio» es el texto de Fraleigh (Fra03).

Observación 5.1.5. *Estudiando polinomios en x con coeficientes en $\mathbb{R}$ ($\mathbb{C}$), es raro hacer la distinción entre el polinomio como elemento de $\mathbb{R}[x]$ ($\mathbb{C}[x]$) o como función de $\mathbb{R}$ en $\mathbb{R}$* (de $\mathbb{C}$ en $\mathbb{C}$). *Mas al considerar campos arbitrarios, puede suceder que dos polinomios sean diferentes incluso si definen la misma función polinomial.*

Consideremos, por ejemplo, los polinomios con coeficientes en $\mathbb{Z}_3$ $f(x) = x^4 + x + 1$ y $g(x) = x^3 + x^2 + 1$. Ciertamente son distintos polinomios (note que tienen grado distinto), sin embargo, $f(0) = 1 = g(0)$, $f(1) = 0 = g(1)$ y $f(2) = 1 = g(2)$; esto es, como funciones de $\mathbb{Z}_3$ en $\mathbb{Z}_3$, f y g son iguales. Más aún, si $\mathbb{F}$ es un campo con q elementos, existen sólo q^q funciones distintas de $\mathbb{F}$ en $\mathbb{F}$, mientras que hay un número infinito de distintos polinomios con coeficientes en $\mathbb{F}$.

Esta ambigüedad, por otro lado, no ocurre al estudiar polinomios con coeficientes en un campo infinito. Es por esta razón que al estudiar polinomios en $\mathbb{R}[x]$ o $\mathbb{C}[x]$ no actuamos cautelosamente en este sentido. Más tarde se dará una prueba de este hecho. (Ver teorema 5.1.13.)

Evaluar un polinomio en un elemento distinguido está íntimamente relacionado con el concepto clásico de «resolver una ecuación polinomial». En lugar de hablar de resolver una ecuación polinomial, nos referiremos a encontrar las *raíces* o *ceros* de un polinomio.

Definición 5.1.6. *Sea $\mathbb{F}$ un campo y $p(x) \in \mathbb{F}[x]$ un polinomio. Un elemento $a \in \mathbb{F}$ es una **raíz** (a veces también llamado **cero**) **de** $p(x)$ si $p(a) = 0$ (pensando a $p(x)$ como función de $\mathbb{F}$ en $\mathbb{F}$, esto significa que $a \longmapsto 0$ bajo p).*

Para aclarar a qué nos referíamos en el párrafo anterior a la definición 5.1.6, veamos un ejemplo.

Ejemplo 5.1.7. En el contexto de la definición 5.1.6 podemos llevar el problema clásico de hallar todas las soluciones reales de la ecuación polinomial $x^2 + x - 2 = 0$ a *encontrar todos los ceros de* $p(x) = x^2 + x - 2$ en $\mathbb{R}$. Ambos problemas, en efecto, tienen la misma respuesta pues

$$\{a \in \mathbb{R} : p(a) = 0\} = \left\{x \in \mathbb{R} : x^2 + x - 2 = 0\right\} = \{-2, 1\}.$$

Con lo dicho hasta aquí sobre polinomios en x con coeficientes en un campo es suficiente para comenzar a probar la proposición 5.0.1.

En todos los resultados siguientes $\mathbb{F}$ denotará un campo arbitrario. Comenzamos con el siguiente lema sobre el grado del producto de dos polinomios.

Lema 5.1.8. *Si $f(x), g(x) \in \mathbb{F}[x]$ son dos polinomios distintos de cero, entonces*

$$\deg(fg) = \deg(f) + \deg(g).$$

Demostración. Escribamos

$$f(x) = a_n x^n + ... + a_1 x + a_0$$

y

$$g(x) = b_m x^m + ... + b_1 x + b_0,$$

donde $n = \deg(f)$ y $m = \deg(g)$; en particular, $a_n \neq 0$ y $b_m \neq 0$. Así,

$$f(x)g(x) = a_0 b_0 + (a_0 b_1 + a_1 b_0)\, x + ... + a_n b_m x^{n+m}$$

donde $a_n b_m \neq 0$ es el coeficiente de la potencia más grande de x que puede aparecer en $p(x)g(x)$. Por lo tanto

$$\deg(fg) = n + m = \deg(f) + \deg(g).$$

□

El siguiente teorema es una herramienta imprescindible para probar el primer punto de la proposición 5.0.1.

Teorema 5.1.9 (Algoritmo de la división para $\mathbb{F}[x]$)**.** *Sean*

$$f(x) = a_n x^n + a_{n-a} x^{n-1} + ... + a_0$$

y

$$g(x) = b_m x^m + b_{m-1} x^{m-1} + ... + b_0$$

dos polinomios de $\mathbb{F}[x]$*, con* $b_m \neq 0$ *y* $m > 0$*. Entonces, existen polinomios únicos* $q(x)$ *y* $r(x)$ *en* $\mathbb{F}[x]$ *tales que* $f(x) = g(x)q(x) + r(x)$ *y* $\deg(r) < \deg(g)$*.*

Demostración. Primero veamos la parte de la unicidad del teorema. Supongamos que el polinomio $f(x)$ puede escribirse como $f(x) = g(x)q_1(x) + r_1(x)$ y también en la forma $f(x) = g(x)q_2(x) + r_2(x)$, con $\deg(r_1) < \deg(g)$ y $\deg(r_2) < \deg(g)$. Igualando estas dos expresiones para $f(x)$, se tiene que $g(x)q_1(x) + r_1(x) = g(x)q_2(x) + r_2(x)$, o bien, $[q_1(x) - q_2(x)]\, g(x) = r_2(x) - r_1(x)$.

Ahora bien, observemos que el lado derecho de la ecuación anterior tiene grado menor que $\deg(g)$, así que el lado derecho debe cumplir lo mismo. Esto implica que $q_1(x) - q_2(x)$ debe ser igual al polinomio cero pues, de no ser así, por el lema 5.1.8 se cumpliría que $\deg(g) \leq \deg[(q_1 - q_2)\, g]$. Luego, $q_1(x) = q_2(x)$. Finalmente,

$$r_1(x) = f(x) - q_1(x)g(x) = f(x) - q_2(x)g(x) = r_2(x).$$

Por lo tanto, los polinomios $q(x)$ y $r(x)$ son únicos.

Ahora sí, pasemos a la parte de la existencia de los polinomios $q(x)$ y $r(x)$. Se distinguen dos casos inmediatos: si $f(x)$ es igual al polinomio cero, o bien, $\deg(f) < \deg(g)$. En cualquier de ellos basta tomar $q(x) = 0$ y $r(x) = f(x)$. Así, podemos suponer que el polinomio $f(x)$ es distinto del polinomio cero y que $\deg(f) \geq \deg(g)$.

Vamos a utilizar inducción matemática sobre $n = \deg(f)$. El caso base es $n = 0$. En esta situación tanto $f(x)$ como $g(x)$ son polinomios constantes (diferentes de cero), y es suficiente tomar $q(x) = \dfrac{f(x)}{g(x)}$ y $r(x) = 0$ para obtener el teorema si $\deg(f) = 0$.

Supongamos ahora que el resultado es verdadero para cualquier polinomio de grado menor que n (no sólo para aquellos de grado $n-1$). Sean

$$f(x) = a_n x^n + a_{n-a} x^{n-1} + ... + a_0$$

y

$$g(x) = b_m x^m + b_{m-1} x^{m-1} + ... + b_0,$$

con a_n, b_m distintos de cero y donde $n \geq m$. Notemos que el polinomio $\dfrac{a_n}{b_m} x^{n-m} g(x)$ comparte con $f(x)$ el término $a_n x^n$, por lo cual

$$h(x) = f(x) - \frac{a_n}{b_m} x^{n-m} g(x) \tag{5.1}$$

es un polinomio de grado menor que n (o posiblemente igual a cero) al ser el lado derecho de la ecuación 5.1 igual a $\left(a_{n-1} - \dfrac{a_n}{b_m} b_{n-1}\right) x^{n-1}$ más términos de grado menor. Se distinguen, entonces, dos casos:

I Si el polinomio $h(x) = 0$. En este caso se tiene que

$$f(x) = \frac{a_n}{b_m} x^{n-m} g(x).$$

Por lo tanto, basta tomar $q(x) = \dfrac{a_n}{b_m} x^{n-m}$ y $r(x) = 0$ para obtener el resultado en este caso.

II Si el polinomio $h(x)$ es diferente del polinomio cero. Como $\deg(h) \geq 0$, podemos aplicar la hipótesis de inducción a los polinomios $h(x)$ y $g(x)$ para obtener $q_1(x)$ y $r(x)$ en $\mathbb{F}[x]$ ocurriendo que $h(x) = g(x)q_1(x) + r(x)$, con $\deg(r) < \deg(g)$. Luego,

$$f(x) = g(x) \left[\frac{a_n}{b_m} x^{n-m} + q_1(x)\right] + r(x). \tag{5.2}$$

Finalmente, basta tomar $q(x) = \dfrac{a_n}{b_m} x^{n-m} + q_1(x)$ en la ecuación 5.2 para concluir el resultado para este caso.

En consecuencia, por el principio de inducción matemática, se concluye que el teorema es válido para cada $n \in \mathbb{N}$. □

Los siguientes dos corolarios son consecuencia inmediata del teorema 5.1.9:

Corolario 5.1.10. *Un elemento $a \in \mathbb{F}$ es un cero de $p(x) \in \mathbb{F}[x]$ si, y sólo si, $x - a$ es factor*[1] *de $p(x)$.*

Demostración. Supóngase primero que para $a \in \mathbb{F}$ se tiene que $p(a) = 0$. Entonces por el teorema 5.1.9, existen polinomios $q(x)$ y $r(x)$ en $\mathbb{F}[x]$ tales que

$$f(x) = (x - a)q(x) + r(x)$$

donde el grado de $r(x)$ es menor que 1 (el grado de $x - a$). Luego, $r(x)$ debe ser un polinomio constante, digamos $r(x) = c$ para algún $c \in \mathbb{F}$. De donde,

$$p(x) = (x - a)q(x) + c.$$

Con lo anterior en cuenta se sigue que

$$0 = p(a) = 0q(a) + c,$$

es decir, $c = 0$. Por lo tanto $p(x) = (x - a)q(x)$, de modo que $x - a$ es un factor de $p(x) \in \mathbb{F}$.

Recíprocamente, si para $a \in \mathbb{F}$, $x - a$ es un factor de $p(x) \in \mathbb{F}[x]$, entonces, para algún polinomio $q(x) \in \mathbb{F}[x]$, $p(x) = (x - a)q(x)$. Así, $p(a) = 0q(a) = 0$. □

Corolario 5.1.11. *Un polinomio distinto de cero $p(x) \in \mathbb{F}[x]$ de grado d tiene a lo más d raíces en $\mathbb{F}$.*

Demostración. Vamos a realizar la prueba utilizando inducción matemática sobre d, el número de raíces de $p(x)$. Para $d = 0$, el polinomio $p(x) \in \mathbb{F}[x]$ es constante, digamos $p(x) = c$ para alguna $c \neq 0$ en $\mathbb{F}$. Luego, $p(x)$ no tiene raíces en $\mathbb{F}$, quedando así demostrado el caso base.

Supongamos ahora que el resultado es verdadero para cualquier polinomio diferente de cero en $\mathbb{F}[x]$ de grado $d - 1$. Sea $p(x) \in \mathbb{F}[x]$ no cero y de grado d. Si $p(x)$ no tiene raíces, se concluye la prueba. Por otro lado, si $p(x)$ tiene una raíz $a \in \mathbb{F}$, entonces, por el corolario anterior, $p(x) = (x - a)q(x)$ para algún polinomio $q(x) \in \mathbb{F}[x]$ de grado $d - 1$. Por la hipótesis de inducción, $q(x)$ tiene a lo más $d - 1$ raíces en $\mathbb{F}$. Ahora bien, cualquier raíz $b \neq a$ en $\mathbb{F}$ del polinomio $p(x)$ satisface que $0 = p(b) = (b-a)q(b)$, y, como $\mathbb{F}$ no tiene divisores de cero [2], b es también una raíz de $q(x)$ en $\mathbb{F}$. En consecuencia, el número total de raíces de $p(x)$ en $\mathbb{F}$ es a lo más $1 + (d - 1) = d$.

Por lo tanto, por el principio de inducción matemática, se concluye que el resultado es verdadero para cada $d \in \mathbb{N}$. □

Con el corolario 5.1.11 queda demostrado el primer punto de la proposición 5.0.1. Este resultado será muy útil en su forma contrapositiva para concluir que un polinomio determinado es, en realidad, el polinomio cero, si éste tiene demasiadas raíces. Vale la pena entonces enunciar el corolario 5.1.11 en estos términos.

[1] Si $f(x), g(x) \in \mathbb{F}[x]$, se dice que $g(x)$ es un factor de $f(x)$ si $f(x) = g(x)q(x)$, para algún polinomio $q(x) \in \mathbb{F}[x]$

[2] Recordemos que si $a \neq 0$ y $b \neq 0$ son dos elementos en un anillo R tales que $ab = 0$, decimos que a y b son divisores de cero.

Corolario 5.1.12. *Sea $p(x) \in \mathbb{F}[x]$ un polinomio de grado d. Si el número de raíces de $p(x)$ en $\mathbb{F}$ es (estrictamente) mayor que d, entonces $p(x)$ es el polinomio cero.*

Obsérvese también que, para el caso de campos finitos, no podemos usarlo de esta manera con polinomios de grado $|\mathbb{F}|$, o mayor, ya que no hay $|\mathbb{F}| + 1$ elementos que puedan funcionar como raíces.

Con ayuda del corolario 5.1.11, podemos demostrar que no hay ambigüedad al considerar funciones polinomiales cuando se trabaja en campos infinitos. Más precisamente:

Teorema 5.1.13. *Sea $\mathbb{F}$ un campo infinito y $p_1(x), p_2(x) \in \mathbb{F}[x]$. Entonces $p_1(x)$ y $p_2(x)$ inducen la misma función de $\mathbb{F}$ en $\mathbb{F}$ si, y sólo si, $p_1(x) = p_2(x)$.*

Demostración. Supongamos primero que $p_1(x) = p_2(x)$. Veamos que ambos polinomios definen la misma función polinomial. Siguiendo la definición 5.1.4, tanto $p_1(x)$ como $p_2(x)$ tienen asociados, respectivamente, las funciones p_1 y p_2 de $\mathbb{F}$ en $\mathbb{F}$. Así, sólo resta probar que $p_1(r) = p_2(r)$, para cada $r \in \mathbb{F}$ (es decir, tienen la misma *regla de correspondencia*). Pero esto se sigue inmediatamente de la hipótesis.

Recíprocamente, sean, respectivamente, p_1 y p_2 las funciones de $\mathbb{F}$ en $\mathbb{F}$ inducidas por $p_1(x)$ y $p_2(x)$, y supongamos que éstas coinciden. Entonces $p_1(r) = p_2(r)$, para cada $r \in \mathbb{F}$, por lo cual, $p_1(r) - p_2(r) = 0$, para toda $r \in \mathbb{F}$. Esto quiere decir que cada $r \in \mathbb{F}$ es una raíz del polinomio $p_1(x) - p_2(x)$, y como $\mathbb{F}$ es infinito, quiere decir que hay un número infinito de ellas. De aquí se puede concluir que $p_1(x) - p_2(x)$ es el polinomio cero, pues de no ser el caso se estaría contradiciendo el corolario 5.1.11. Luego, $p_1(x) - p_2(x) = 0$, de donde, $p_1(x) = p_2(x)$. □

Para abordar el punto (ii) de la proposición 5.0.1, primero revisemos la siguiente definición que nos ayudará a enunciarlo de una manera más breve.

Definición 5.1.14. *Sea $\mathbb{F}$ un campo y $p(x) \in \mathbb{F}[x]$ un polinomio. Se dice que $p(x)$ se **anula** en un conjunto $A \subseteq \mathbb{F}$ si $p(a) = 0$ para cada $a \in A$.*

Lema 5.1.15. *Sea $d \geq 1$ un número entero. Para cualquier conjunto $A \subseteq \mathbb{F}$ de cardinalidad $|A| \leq d$, existe un polinomio diferente de cero $p \in \mathbb{F}[x]$ de grado a lo más d que se anula en A.*

Demostración. Es suficiente tomar el polinomio $p(x) = \prod_{a \in A} (x - a)$. □

Alternativamente, se presenta a continuación otro argumento para demostrar el lema 5.1.15 usando las técnicas introducidas por Dvir. Como se mencionó en la introducción, gran parte de sus ideas están relacionadas con conceptos clásicos de álgebra lineal.

Probemos primero el siguiente lema:

Lema 5.1.16. *Sean $Poly_d(\mathbb{F})$ y $\mathbb{F}^A$, respectivamente, el $\mathbb{F}$-espacio vectorial de los polinomios de grado a lo más d con coeficientes en $\mathbb{F}$ y el de las sucesiones finitas $(y_a)_{a \in A}$. Entonces, la función $T : Poly_d(\mathbb{F}) \longrightarrow \mathbb{F}^A$ definida por $T(p) = (p(a))_{a \in A}$, para cada $p(x) \in Poly_d(\mathbb{F})$, es una trasformación lineal.*

Demostración. En efecto,

$$\begin{aligned} T(cp_1 + p_2) &= \left((cp_1 + p_2)(a)\right)_{a\in A} \\ &= \left(cp_1(a)\right)_{a\in A} + \left(p_2(a)\right)_{a\in A} \\ &= c\left(p_1(a)\right)_{a\in A} + \left(p_2(a)\right)_{a\in A} \\ &= c\,T(p_1) + T(p_2), \end{aligned}$$

para cualesquiera $p_1(x), p_2(x) \in Poly_d(\mathbb{F})$ y $c \in \mathbb{F}$. Por lo tanto, T es una trasformación lineal. □

Corolario 5.1.17. *Sea $d \geq 1$ un número entero. Dado un conjunto $A \subseteq \mathbb{F}$ de cardinalidad $|A| \leq d$, existe un polinomio no cero $p(x) \in \mathbb{F}[x]$ de grado a lo más d que se anula en A.*

Demostración. Sean $Poly_d(\mathbb{F})$, $\mathbb{F}^A$ y la transformación lineal T como en el lema anterior. Una base para el espacio $Poly_d(\mathbb{F})$ está dada por el conjunto $\{1, x, ..., x^d\}$. Luego, $dim\left[Poly_d(\mathbb{F})\right] = d + 1$. Por otro lado, la dimensión del espacio $\mathbb{F}^A$ es igual a $|A| \leq d$. Puesto que el rango de T, $R(T)$, es un subespacio de $\mathbb{F}^A$, $dim\left[R(T)\right] \leq dim\left(\mathbb{F}^A\right)$. De aquí se sigue que

$$\begin{aligned} 0 &< (d+1) - |A| \\ &= dim\left[Poly_d(\mathbb{F})\right] - dim\left(\mathbb{F}^A\right) \\ &\leq dim\left[Poly_d(\mathbb{F})\right] - dim\left[R(T)\right]. \end{aligned}$$

En otras palabras, el núcleo de T es no trivial, por lo cual existe $p(x) \in Poly_d(\mathbb{F})$ que se anula en A. □

5.2. Polinomios en varias variables.

En esta sección nos ocuparemos de la generalización de la proposición 5.0.1 para el caso de polinomios en varias variables. Para comenzar, mencionaremos, muy brevemente, algunos aspectos básicos sobre dichos polinomios.

Sea $\mathbb{F}$ un campo. Un **polinomio en las variables $x_1, \ldots, x_n$ con coeficientes en $\mathbb{F}$** se define como una expresión formal $p(x_1, \ldots, x_n)$ de la forma

$$p(x_1, \ldots, x_n) = \sum_{i_1,\ldots,i_n \geq 0} c_{i_1,\ldots,i_n} x_1^{i_1} \cdots x_n^{i_n},$$

donde los coeficientes $c_{i_1,\ldots,i_n}$ están en $\mathbb{F}$, y sólo un número finito de ellos son diferentes de cero. Por supuesto, los mismos acuerdos para facilitar la notación mencionados para polinomios en una variable se utilizarán en este caso. Así, por ejemplo,

$$p(x_1, x_2) = x_1^2 + x_2^2 - x_1 x_2 + 1$$

y

$$q(x_1, x_2, \cdots, x_n) = x_1 \ldots x_n + x_1 + \ldots + x_n$$

son polinomios, respectivamente, en dos y n variables con coeficientes en algún campo. En completa analogía con el anillo $\mathbb{F}[x]$, denotaremos por $\mathbb{F}[x_1, x_2, \ldots, x_n]$ al anillo de los polinomios en $x_1, x_2, \ldots, x_n$ con coeficientes en $\mathbb{F}$.

En n variables, un **monomio** es un polinomio de la forma $x_1^{i_1} \cdots x_n^{i_n}$, donde la suma $i_1 + \ldots + i_n$ se llama el **grado** del monomio. Así, el único monomio de grado cero es la constante 1. (Obsérvese que todo polinomio es una combinación lineal de monomios con coeficientes diferentes de cero tomados del campo $\mathbb{F}$.) El **grado** de un polinomio $p(x) \in \mathbb{F}[x_1, x_2, \ldots, x_n]$[1], $\deg(p)$, se define como el máximo de los grados de sus monomios para los que sus coeficientes sean distintos de cero. También diremos que $p(x) \in \mathbb{F}[x_1, x_2, \ldots, x_n]$ es **homogéneo** si todos sus monomios con coeficientes no cero tienen el mismo grado. Por ejemplo, el monomio $x_1 x_2^3 x_3$ tiene grado 5, mientras que el polinomio $p(x_1, x_2, x_3) = 2x_1^4 x_2^2 + 3x_1 x_2^3 x_3$ tiene grado 6 ya que $x_1^4 x_2^2$ es el monomio de mayor grado. Además, el polinomio $p(x_1, x_2) = x_1^2 + x_1 x_2$ es homogéneo de grado 2. Finalmente, un polinomio $p(x) \in \mathbb{F}[x_1, x_2, \ldots, x_n]$ se dice ser el **polinomio cero**, denotado por 0, si todos sus coeficientes son iguales a cero. Se define el grado del polinomio cero $0 \in \mathbb{F}[x_1, x_2, \ldots, x_n]$ como -1.

En algunas ocasiones estaremos interesados en considerar todas las variables, salvo una, de un polinomio como «fijas» y considerar al polinomio como uno en una variable con sus coeficientes siendo ahora polinomios en el resto de las variables. Por ejemplo, si pensamos al polinomio $p(x_1, x_2) = x_1^3 x_2 + x_1^3 + x_1 x_2^3$ como un polinomio (sólo) en la variable x_1, entonces el coeficiente de x_1^3 y de x_1 en $p(x_1, x_2)$ son, respectivamente, $x_2 + 1$ y x_2^3. También, interpretando las indeterminadas $x_1, \ldots, x_n$ como variables, será útil interpretar un polinomio $p(x_1, \ldots, x_n) \in \mathbb{F}[x_1, x_2, \ldots, x_n]$ como una función de $\mathbb{F}^n$ en $\mathbb{F}$.

Definición 5.2.1. *Sea $\mathbb{F}$ un campo. Para cualquier $p(x_1, \ldots, x_n) \in \mathbb{F}[x_1, x_2, \ldots, x_n]$, podemos asociar al polinomio $p(x_1, \ldots, x_n)$ una función $p : \mathbb{F}^n \longrightarrow \mathbb{F}$ como sigue: si*

$$p(x_1, \ldots, x_n) = \sum_{i_1, \ldots, i_n \geq 0} c_{i_1, \ldots, i_n} x_1^{i_1} \cdots x_n^{i_n},$$

entonces, para cada $(r_1, \ldots, r_n) \in \mathbb{F}^n$, se define $p(r_1, \ldots, r_n) \in \mathbb{F}$ como

$$p(r_1, \ldots, r_n) = \sum_{i_1, \ldots, i_n \geq 0} c_{i_1, \ldots, i_n} r_1^{i_1} \cdots r_n^{i_n}.$$

*La función p es llamada **función polinomial inducida por** $p(x_1, \ldots, x_n)$.*

Por supuesto, al igual que en el caso de una variable, debemos ser cuidadosos al considerar a dicho polinomio como función o como elemento de $\mathbb{F}[x_1, x_2, \ldots, x_n]$, pues, trabajando con campos finitos, puede ocurrir que dos polinomios induzcan la misma función. En este sentido, es importante tener en mente el siguiente ejemplo:

[1] Para facilitar la notación, cuando sea conveniente escribiremos $p(x)$ en lugar de $p(x_1, ..., x_n)$ para denotar a un polinomio en $\mathbb{F}[x_1, ..., x_n]$.

Ejemplo 5.2.2. Consideremos el polinomio $p(x_1, x_2) = x_1^2 + x_1$ con coeficientes en $\mathbb{Z}_2$. Este polinomio satisface que $p(x_1, x_2) = 0$ para cada $(x_1, x_2) \in \mathbb{Z}_2^2$; así, éste induce la misma función que el polinomio $0 \in \mathbb{Z}[x_1, x_2]$. Sin embargo, claramente, $p(x_1, x_2)$ no es el polinomio cero.

Con esto en cuenta, para un campo $\mathbb{F}$, no se debe caer en la trampa de pensar que el polinomio $0 \in \mathbb{F}[x_1, x_2, \ldots, x_n]$ es el único que, considerado como función, *evalúa* a cada elemento de $\mathbb{F}^n$ en el cero de $\mathbb{F}$.

Por último, los siguientes conceptos serán de gran utilidad para todo lo sucesivo:

Se dice que un vector $(a_1, \ldots, a_n) \in \mathbb{F}^n$ es un **cero** o **raíz** del polinomio $p(x_1, \ldots, x_n)$ en $\mathbb{F}[x_1, x_2, \ldots, x_n]$ si $p(a_1, \ldots, a_n) = 0$. En el corolario 5.1.11 probamos que, para el caso de un polinomio $p(x) \in \mathbb{F}[x]$ diferente de cero, éste no podía tener más de $\deg(p)$ raíces en $\mathbb{F}$. Por supuesto, en general, este no es el caso para polinomios en varias variables: por ejemplo, el polinomio $p(x_1, x_2) = x_1 x_2$ con coeficientes en $\mathbb{R}$ se anula en todas las parejas $(x_1, 0), (0, x_2)$, con $x_1, x_2 \in \mathbb{R}$. Por lo tanto $p(x_1, x_2)$ tiene un número infinito de raíces. En consecuencia, si de alguna forma queremos generalizar el teorema del factor para polinomios en varias variables, conservando la relación entre el grado y el número de ceros de un polinomio, un punto central será considerar polinomios con coeficientes en un campo finito. Más adelante veremos en el teorema 5.2.4 que éste es el caso. Finalmente, diremos que un polinomio $p(x) \in \mathbb{F}[x_1, x_2, \ldots, x_n]$ **se anula** es un conjunto $A \subseteq \mathbb{F}^n$ si $p(a) = 0$ para cada $a \in A$.

Ahora sí estamos listos para demostrar las generalizaciones de los puntos (i) y (ii) en la proposición 5.0.1.

5.2.1. Lema de DeMillo-Lipton-Zippel-Schwartz.

Primero vamos a probar la generalización del corolario 5.1.11 analizando el comportamiento de los polinomios en rectas. Para esto, necesitamos el siguiente lema previo:

Lema 5.2.3. *Sea $\mathbb{F}$ un campo finito, con $|\mathbb{F}| = q$, y sea $p(x) \in \mathbb{F}[x_1, x_2, \ldots, x_n]$ un polinomio de grado $d < q$. Si $p(x) = 0$ para cada $x \in \mathbb{F}^n$, entonces, $p(x)$ es el polinomio cero.*

Demostración. La prueba se realizará por inducción matemática sobre n, el número de variables de $p(x)$. Caso base: $n = 1$. Sea $p(x) \in \mathbb{F}[x]$. Si $p(x)$ no es el polinomio cero, entonces, sabemos que éste tiene a lo más $d < q$ raíces en $\mathbb{F}$. Así, existe $a \in \mathbb{F}$ tal que $p(a) \neq 0$. Con esto queda demostrado el resultado para $n = 1$. Analicemos ahora el caso $n \geq 2$:

Hipótesis de inducción: para cualquier polinomio $f(x)$ en $\mathbb{F}[x_1, x_2, \ldots, x_{n-1}]$ de grado menor que q se cumple que, si $f(x) = 0$ para cada $x \in \mathbb{F}^{n-1}$, entonces $f(x)$ es el polinomio cero.

Sea $p(x_1, \ldots, x_n) \in \mathbb{F}[x_1, x_2, \ldots, x_n]$ un polinomio de grado $d < q$. Supongamos, además, que $p(x_1, \ldots, x_n) = 0$ para todo $(x_1, \ldots, x_n) \in \mathbb{F}^n$. Así, tenemos que mostrar

que $p(x_1, \ldots, x_n)$ es el polinomio cero. Para esto, vamos a descomponer a $p(x_1, \ldots, x_n)$ en sumandos de acuerdo a las potencias de la variable x_n como sigue:

$$p(x_1, \ldots, x_n) = \sum_{i=0}^{k} f_i(x_1, \ldots, x_{n-1}) x_n^i,$$

donde $f_i \in \mathbb{F}[x_1, \ldots, x_{n-1}]$ para $0 \leq i \leq k \leq d$, con k la mayor potencia a la cual aparece la variable x_n en $p(x)$. Ahora bien, sea $a = (a_1, \ldots, a_{n-1}) \in \mathbb{F}^{n-1}$, y definamos $p_a(x_n) \in \mathbb{F}[x_n]$ como $p_a(x_n) = f(a_1, \ldots, a_{n-1}, x_n)$. Sabemos que $p_a(x_n) = 0$ para cada $x_n \in \mathbb{F}$; además, $\deg(p_a) \leq k \leq d < q$, por lo que volvemos al caso $n = 1$. Luego, $p_a(x_n)$ es el polinomio cero, es decir, $f_i(a_1, \ldots, a_{n-1}) = 0$, para toda $0 \leq i \leq k$. Pero la elección de $a \in \mathbb{F}^{n-1}$ fue arbitraria, por lo cual, en realidad, $f_i(x) = 0$ para cada $x \in \mathbb{F}^{n-1}$. También se cumple que $\deg(f_i) \leq d < q$, así que, utilizando la hipótesis de inducción, para cada $0 \leq i \leq k$, $f_i(x)$ debe ser el polinomio cero. En otras palabras, todos los coeficientes de $f_i(x)$, y por lo tanto de $p(x) \in \mathbb{F}[x_1, x_2, \ldots, x_n]$, son cero. En consecuencia, $p(x)$ es el polinomio cero.

Por lo tanto, por el principio de inducción matemática, se concluye que el resultado es verdadero para cada $n \in \mathbb{N}$. □

Teorema 5.2.4. *Sea $\mathbb{F}$ un campo finito con q elementos. Todo polinomio diferente de cero $p(x)$ en $\mathbb{F}[x_1, x_2, \ldots, x_n]$ de grado d tiene a lo más dq^{n-1} raíces en $\mathbb{F}^n$.*

Así, por ejemplo, este teorema asegura que el polinomio $p(x_1, x_2, x_3) = x_3^2 + x_1x_2 + x_3$ tiene a lo más 8 raíces en $\mathbb{Z}_2^3$.

La prueba a continuación se debe a Zeev Dvir (Dvi09) y a Dana Moshkovitz (Mos10).

Demostración. El caso $n = 1$ del teorema está considerado en el corolario 5.1.11, así que podemos suponer que $n \geq 2$. Más aún, sin perdida de generalidad, podemos suponer que $1 \leq d < q$ [1]. La demostración se hará reduciendo al caso $n = 1$. Escribamos $p(x) = f(x) + g(x)$, donde $f(x)$ es un polinomio, no cero, homogéneo de grado d y $g(x)$ tiene sólo monomios de grado estrictamente menores que d. Por el lema 5.2.3, $f(w) \neq 0$ para algún vector w en $\mathbb{F}^n$. Más aún, como $f(x)$ es homogéneo de grado $d \geq 1$, $w \neq 0$ (pues $f(0) = 0$ al ser $f(x)$ homogéneo). Ahora, a cada vector $u \in \mathbb{F}^n$ asociémosle la recta $l_u = \{u + tw : t \in \mathbb{F}\}$ por u en la dirección w. Así, $l_u \cap l_v = \emptyset$ siempre que $v \notin l_u$. En efecto, si $y \in l_u \cap l_v$, entonces existen $t_1, t_2 \in \mathbb{F}$ tales que $y = u + t_1w$ y $y = v + t_2w$. De donde, $u + t_1w = v + t_2w$, o bien, $v = u + (t_1 - t_2)w$. Luego, $v \in l_u$, lo cual es una contradicción a la suposición inicial.

Obsérvese que al ser $w \neq 0$, cada recta l_u contiene $|l_u| = q$ puntos. De este modo, el espacio $\mathbb{F}^n$ puede partirse en $\dfrac{q^n}{q} = q^{n-1}$ rectas. Queda, por tanto, demostrar que el número de raíces del polinomio $p(x)$ en cada cada recta l_u es a lo más d.

[1] En efecto, si $d \geq q$ entonces $dq^{n-1} \geq q^n$, lo cual da lugar a una cota superior inmediata para el número de raíces posibles.

Para probar esto, nótese que, para cada $u \in \mathbb{F}^n$, la función $p_u(t) = p(u + tw)$ es un polinomio en t de grado a lo más d. Más aún, este polinomio es diferente de cero ya que el coeficiente del término t^d en $p_u(t)$ es $g(w) \neq 0$. Por lo tanto, utilizando el teorema del factor, $p_u(t)$ tiene a lo más d raíces. En otras palabras, el polinomio $p(x)$ se puede anular en a lo más d puntos de la recta l_u. Como sólo hay q^{n-1} rectas en la partición de $\mathbb{F}^n$, el número total de raíces de $p(x)$ no puede exceder dq^{n-1}, como buscábamos probar. □

Ejemplo 5.2.5. Consideremos el espacio $\mathbb{Z}_3^2$ y sea $p(x_1, x_2) = x_1^2 + x_2 + 1 \in \mathbb{Z}_3[x_1, x_2]$. En la notación del teorema anterior, $p(x_1, x_2) = f(x_1, x_2) + g(x_1, x_2)$, donde el polinomio $f(x_1, x_2) = x_1^2$ es homogéneo de grado 2 y $g(x_1, x_2) = x_2 + 1$. Además, $p(x)$ tiene a lo más 6 raíces en $\mathbb{Z}_3^2$ según el teorema 5.2.4. Por otro lado, $(1, 1) \in \mathbb{Z}_3^2$ es tal que $f(1, 1) \neq 0$. Así, las rectas

$$\begin{aligned} l_{(0,0)} &= \{t(1,1) : t \in \mathbb{F}\} = \{(0,0), (1,1), (2,2)\}, \\ l_{(0,1)} &= \{(0,1) + t(1,1) : t \in \mathbb{F}\} = \{(0,1), (1,2), (2,0)\}, \\ l_{(0,2)} &= \{(0,2) + t(1,1) : t \in \mathbb{F}\} = \{(0,2), (1,0), (2,1)\}, \end{aligned}$$

forman una partición del espacio $\mathbb{Z}_3^2$. En la recta $l_{(0,0)}$, $(1, 1)$ es una raíz de $p(x)$; en la recta $l_{(0,1)}$, $p(x)$ no tiene raíces; y en la recta $l_{(0,2)}$, $(0, 2)$ y $(2, 1)$ son raíces de $p(x)$. En consecuencia, el polinomio $p(x)$ tiene exactamente 3 raíces en $\mathbb{Z}_3^2$.

Un resultado más general fue demostrado independientemente por Richard A. DeMillo y Richard J. Lipton (DL78), Richard Zippel (Zip79) y Jacob T. Schwartz (Sch80).

Teorema 5.2.6 (DeMillo-Lipton-Zippel-Schwartz)**.** *Sea $\mathbb{F}$ un campo y d un entero positivo. Para cada conjunto finito $A \subseteq \mathbb{F}$ tal que $|A| \geq d$, se cumple que cualquier polinomio no cero $p(x) \in \mathbb{F}[x_1, x_2, \ldots, x_n]$ de grado d tiene a lo más $d|A|^{n-1}$ raíces en A^n.*

Demostración. Vamos a demostrar el resultado utilizando inducción matemática sobre n, el número de variables de $p(x)$.

Para $n = 1$, por el corolario 5.1.11, el polinomio no cero en una variable $p(x)$ tiene a lo más d raíces en $\mathbb{F}$. Ahora analicemos el caso $n \geq 2$. Supongamos que el resultado es verdadero para cualquier polinomio de $n-1$ variables, y consideremos $p(x)$ un polinomio diferente de cero en $\mathbb{F}[x_1, x_2, \ldots, x_n]$. Para contar el número de raíces de $p(x)$ en A^n, primero vamos descomponerlo en sumandos de acuerdo a las potencias de la variable x_n como sigue:

$$p(x) = \sum_{i=0}^{k} g_i(x_1, x_2, \ldots, x_{n-1}) x_n^i,$$

donde $g_i(x_1, x_2, \ldots, x_{n-1}) \in \mathbb{F}[x_1, x_2, \ldots, x_{n-1}]$ para $0 \leq i \leq k \leq d$. Además, como $p(x)$ es diferente cero, existe un índice i de forma que $g_i(x_1, x_2, \ldots, x_{n-1})$ sea no cero; sea k el mayor de dichos índices. Con esto en cuenta, escribamos cada $v \in A^n$ en la forma $v = (a, b)$ con $a \in A^{n-1}$, $b \in A$, y estimemos el número de raíces $p(a, b) = 0$. Se pueden distinguir los siguientes dos casos:

Caso 1. Raíces (a, b) tal que $g_k(a) = 0$.

Como $g_k(x_1, x_2, \ldots, x_{n-1}) \neq 0$ y $\deg(g_k) \leq d - k$ –pues el grado del polinomio $g_i(x_1, x_2, \ldots, x_{n-1}) x_n^i$ es a lo más d–, el polinomio $g_k(x_1, x_2, \ldots, x_{n-1})$ tiene a lo más $(d-k)|A|^{n-2}$ raíces en A^{n-1}, por hipótesis de inducción. Ahora bien, para cada $a \in A^{n-1}$ hay a lo más $|A|$ diferentes posibles elecciones para el elemento b. Esto quiere decir que el número de raíces de $p(x)$ en A^n que caen en este primer caso son a lo más $(d-k)|A|^{n-1}$.

Caso 2. Raíces (a, b) con $g_k(a) \neq 0$.

Observemos que $p(a, x_n) \in \mathbb{F}[x_n]$ es un polinomio distinto de cero, de una sola variable, x_n, y de grado k. Así, para cada elemento a, hay a lo más k posibles elecciones de b en A de modo que $p(a, b) = 0$. Ya que existen a lo más $|A|^{n-1}$ posibles elecciones para el elemento a, tenemos a lo más $k|A|^{n-1}$ raíces de $p(x)$ en este caso.

Sumando el número máximo de raíces posibles en ambos casos, se sigue que hay a lo más

$$(d-k)|A|^{n-1} + k|A|^{n-1} = d|A|^{n-1}$$

raíces de $p(x)$ en A^n. Por lo tanto, por el principio de inducción matemática, se concluye que el resultado es verdadero para cada $n \in \mathbb{N}$. □

Observación 5.2.7. *Aunque la demostración, utilizando inducción matemática, del teorema 5.2.6 parezca menos intuitiva (en el sentido geométrico) que la prueba del teorema 5.2.4, la primera tiene la ventaja de no requerir el lema 5.2.3. De hecho, este último resultado puede deducirse del teorema 5.2.6.*

Con esto queda demostrada la generalización de la parte (i) en la proposición 5.0.1. Por otro lado, como ya se mencionó, no hay un análogo de ese resultado que permita acotar el número de raíces de un polinomio en varias variables para el caso de campos infinitos. Sin embargo, hay un resultado que sí dice «algo» en este sentido.

Lema 5.2.8. *Sea $\mathbb{F}$ un campo infinito y sea $p(x) \in \mathbb{F}[x_1, x_2, \ldots, x_n]$ un polinomio. Si $p(x) = 0$ para cada $x \in \mathbb{F}^n$, entonces, $p(x)$ es el polinomio cero.*

Observación 5.2.9. *Del lema 5.2.8 se sigue que, si $p(x) \in \mathbb{F}[x_1, x_2, \ldots, x_n]$ es diferente de cero, entonces existe $a \in \mathbb{F}^n$ tal que $p(a) \neq 0$; un resultado bastante intuitivo.*

Demostración. La prueba se realizará por inducción matemática sobre n, el número de variables de $p(x)$. Caso base: $n = 1$. Sea $p(x) \in \mathbb{F}[x]$. Si $p(x)$ no es el polinomio cero, entonces, sabemos que éste tiene a lo más $\deg(p)$ raíces en $\mathbb{F}$. Como $\mathbb{F}$ es infinito, claramente existe $a \in \mathbb{F}$ tal que $p(a) \neq 0$. Con esto queda demostrado el resultado para $n = 1$. Analicemos ahora el caso $n \geq 2$:

Hipótesis de inducción: para cualquier polinomio $f(x)$ en $\mathbb{F}[x_1, x_2, \ldots, x_{n-1}]$ se cumple que, si $f(x) = 0$ para cada $x \in \mathbb{F}^{n-1}$, entonces $f(x)$ es el polinomio cero.

Sea $p(x_1, \ldots, x_n) \in \mathbb{F}[x_1, x_2, \ldots, x_n]$ un polinomio tal que $p(x_1, \ldots, x_n) = 0$ para todo $(x_1, \ldots, x_n)$ en $\mathbb{F}^n$. Así, tenemos que mostrar que $p(x_1, \ldots, x_n)$ es el polinomio cero. Para esto, vamos a descomponer a $p(x_1, \ldots, x_n)$ en sumandos de acuerdo a las

potencias de la variable x_n como sigue:

$$p(x_1, \ldots, x_n) = \sum_{i=0}^{k} f_i(x_1, \ldots, x_{n-1}) x_n^i,$$

donde $f_i \in \mathbb{F}[x_1, x_2, \ldots, x_{n-1}]$ para $0 \leq i \leq k \leq \deg(p)$, con k la mayor potencia a la cual aparece la variable x_n en $p(x)$. Ahora bien, sea $a = (a_1, \ldots, a_{n-1}) \in \mathbb{F}^{n-1}$, y definamos $p_a(x_n) \in \mathbb{F}[x_n]$ como $p_a(x_n) = f(a_1, \ldots, a_{n-1}, x_n)$. Sabemos que $p_a(x_n) = 0$ para cada $x_n \in \mathbb{F}$, por lo que volvemos al caso $n = 1$. Luego, $p_a(x_n)$ es el polinomio cero, es decir, $f_i(a_1, \ldots, a_{n-1}) = 0$, para toda $0 \leq i \leq k$. Pero la elección de $a \in \mathbb{F}^{n-1}$ fue arbitraria, por lo cual, en realidad, $f_i(x) = 0$ para cada $x \in \mathbb{F}^{n-1}$. De donde, utilizando la hipótesis de inducción, para cada $0 \leq i \leq k$, $f_i(x)$ debe ser el polinomio cero. En otras palabras, todos los coeficientes de $f_i(x)$, y por lo tanto de $p(x) \in \mathbb{F}[x_1, x_2, \ldots, x_n]$, son cero. En consecuencia, $p(x)$ es el polinomio cero.

Por lo tanto, por el principio de inducción matemática, se concluye que el resultado es verdadero para cada $n \in \mathbb{N}$. □

Con ayuda del lema 5.2.8 podemos establecer un análogo al teorema 5.1.13 para polinomios en varias variables.

Teorema 5.2.10. *Sea $\mathbb{F}$ un campo infinito y $p_1(x), p_2(x) \in \mathbb{F}[x_1, x_2, \ldots, x_n]$. Entonces $p_1(x)$ y $p_2(x)$ inducen la misma función de $\mathbb{F}^n$ en $\mathbb{F}$ si, y sólo si, $p_1(x) = p_2(x)$.*

Demostración. Supongamos primero que $p_1(x) = p_2(x)$. Veamos que ambos polinomios definen la misma función polinomial. Siguiendo la definición 5.2.1, tanto $p_1(x)$ como $p_2(x)$ tienen asociados, respectivamente, las funciones p_1 y p_2 de $\mathbb{F}^n$ en $\mathbb{F}$. Así, sólo resta probar que $p_1(r) = p_2(r)$, para cada $r \in \mathbb{F}^n$ (en otras palabras, tienen la misma *regla de correspondencia*). Pero esto es una consecuencia inmediata de la igualdad de los polinomios $p_1(x)$ y $p_2(x)$.

Recíprocamente, sean, respectivamente, p_1 y p_1 las funciones de $\mathbb{F}^n$ en $\mathbb{F}$ inducidas por $p_1(x)$ y $p_2(x)$, y supongamos que éstas son iguales. Entonces $p_1(r) = p_2(r)$, para cada $r \in \mathbb{F}^n$, por lo que $p_1(r) - p_2(r) = 0$, para toda $r \in \mathbb{F}^n$. Esto quiere decir que el polinomio $p_1(x) - p_2(x)$ se anula en todo $\mathbb{F}^n$. Utilizando el lema 5.2.8 se sigue que $p_1(x) - p_2(x)$ es el polinomio cero. Por lo tanto, $p_1(x) - p_2(x) = 0$, de donde, $p_1(x) = p_2(x)$. □

Pasemos ahora a establecer la generalización de la parte (ii) del teorema 5.0.1. Para esto, nos valdremos del siguiente lema que versa sobre el espacio de los polinomios de grado acotado.

Lema 5.2.11. *Sea $\mathbb{F}$ un campo. El $\mathbb{F}$-espacio vectorial $Poly_d(\mathbb{F}^n)$ de polinomios en $\mathbb{F}[x_1, x_2, \ldots, x_n]$ de grado a lo más d tiene dimensión $\binom{d+n}{d}$.*

Demostración. Una base para el espacio $Poly_d(\mathbb{F}^n)$ está dada por el conjunto de los monomios $x^{i_1} x^{i_2} \ldots x^{i_n}$ con $i_1 + i_2 + \ldots + i_n \leq d$:

$$\left\{1, x_1, \ldots, x_n, x_1^2, x_1 x_2, \ldots, x_1^3, \ldots, x_n^d\right\}.$$

Así, necesitamos contar el número total de monomios en la base. En este sentido, lo anterior es equivalente a contar el número de sucesiones $(i_1, i_2, \ldots, i_n)$ de enteros no negativos restringidos a la condición $i_1 + i_2 + \ldots + i_n \leq d$. Consideremos pues la función f definida por

$$f(i_1, i_2, \ldots, i_n) = (i_1 + 1, i_1 + i_2 + 2, \ldots, i_1 + i_2 + \ldots + i_n + n).$$

Es decir, la imagen de cada sucesión $(i_1, i_2, \ldots, i_n)$ es la sucesión estrictamente creciente

$$i_1 + 1 < i_1 + i_2 + 2 < \ldots < i_1 + i_2 + \ldots + i_n + n,$$

lo cual determina el subconjunto $\{i_1 + 1, i_1 + i_2 + 2, \ldots, i_1 + i_2 + \ldots + i_n + n\}$ de n elementos (el hecho que la sucesión sea estrictamente creciente asegura esto) del conjunto $\{1, 2, \ldots, n + d\}$.

Veamos que la función f es inyectiva. Sean $a = (a_1, a_2, \ldots, a_n)$ y $b = (b_1, b_2, \ldots, b_n)$ dos sucesiones de números enteros no negativos tales que $a_1 + a_2 + \ldots + a_n \leq d$ y $b_1 + b_2 + \ldots + b_n \leq d$, y supongamos que $f(a) = f(b)$. Debemos mostrar entonces que $a = b$, es decir, tenemos que probar que $a_i = b_i$, para cada $i = 1, 2, \ldots, n$. En efecto, como

$$(a_1+1, a_1+a_2+2, \ldots, a_1+a_2+\ldots+a_n+n) = (b_1+1, b_1+b_2+2, \ldots, b_1+b_2+\ldots+b_n+n),$$

entonces

$$\begin{aligned} a_1 + 1 &= b_1 + 1 \\ a_1 + a_2 + 2 &= b_1 + b_2 + 2 \\ &\vdots \\ a_1 + a_2 + \ldots + a_n + n &= b_1 + b_2 + \ldots + b_n + n. \end{aligned}$$

De la primera ecuación se obtiene que $a_1 = b_1$ y, sustituyendo en la segunda ecuación, resulta que $a_2 = b_2$. Continuando con este proceso se obtiene de la última igualdad que $a_n = b_n$, y, en consecuencia, $a = b$.

Así, f es una función biyectiva en su imagen, por lo cual el número de monomios es igual al número de subconjuntos de n elementos escogidos de uno con $n + d$, es decir, igual a $\binom{n+d}{n}$.

Por lo tanto,

$$dim\left[Poly_d\left(\mathbb{F}^n\right)\right] = \binom{n+d}{n}.$$

□

Ejemplo 5.2.12. Consideremos el espacio $Poly_3\left(\mathbb{F}^2\right)$. Una base para éste está dada por:

$$\left\{1, x_1, x_2, x_1^2, x_1x_2, x_2^2, x_1^3, x_1^2x_2, x_1x_2^2, x_2^3\right\},$$

cuya cardinalidad es $\binom{2+3}{3} = 10$. Así, cualquier polinomio $p(x_1, x_2) \in Poly_3(\mathbb{F}^2)$ se puede expresar como

$$\begin{aligned} p(x_1, x_2) = a_{0,0} + a_{1,0}x_1 + a_{0,1}x_2 + a_{2,0}x_1^2 + a_{1,1}x_1x_2 + \\ + a_{0,2}x_2^2 + a_{3,0}x_1^3 + a_{2,1}x_1^2x_2 + a_{1,2}x_1x_2^2 + a_{0,3}x_2^3. \end{aligned}$$

Lema 5.2.13. *Sea $\mathbb{F}$ un campo y $d \geq 0$ un número entero. Dado un conjunto $A \subseteq \mathbb{F}^n$ de cardinalidad $|A| < \binom{n+d}{d}$, existe un polinomio no cero $p(x) \in \mathbb{F}[x_1, x_2, \ldots, x_n]$ de grado a lo más d que se anula en A.*

Antes de probar el lema 5.2.13, veamos un ejemplo para apreciar cómo funciona. Consideremos el espacio $\mathbb{Z}_2^5$. Según el lema mencionado, dado el conjunto

$$A = \{(0,0,1,1,1), (0,0,1,1,0), (0,1,1,0,1), (1,0,0,1,1), (0,0,1,0,0)\}$$

existe un polinomio no nulo $p(x_1, \ldots, x_5) \in \mathbb{Z}_2[x_1, \ldots, x_5]$ de grado uno que se anula en A ya que

$$|A| = 5 < \binom{5+1}{1} = 6.$$

Por ejemplo, $p(x_1, \ldots, x_5) = x_1 + x_3 + 1$.

Ahora daremos la demostración del lema 5.2.13.

Demostración. Sea $Poly_d(\mathbb{F}^n) \subseteq \mathbb{F}[x_1, x_2, \ldots, x_n]$ el $\mathbb{F}$-espacio vectorial de los polinomios de grado a lo más d. Por el lema 5.2.11, sabemos que $dim[Poly_d(\mathbb{F}^n)] = \binom{n+d}{d}$. Consideremos también al espacio vectorial $\mathbb{F}^A$ de todas las sucesiones finitas $(y_a)_{a \in A}$. Éste tiene dimensión igual a $|A|$, la cual por hipótesis es menor que $dim[Poly_d(\mathbb{F}^n)]$.

Ahora bien, consideremos la transformación lineal $T : Poly_d(\mathbb{F}^n) \longrightarrow \mathbb{F}^A$ definida por $T(p) = (p(a))_{a \in A}$ (ver lema 5.1.16). Como el rango de T, $R(T)$, es un subespacio vectorial de $\mathbb{F}^A$, se sigue que $dim[R(T)] \leq dim(\mathbb{F}^A)$. En consecuencia,

$$\begin{aligned} 0 &< \binom{n+d}{d} - |A| \\ &= dim[Poly_d(\mathbb{F}^n)] - dim(\mathbb{F}^A) \\ &\leq dim[Poly_d(\mathbb{F}^n)] - dim[R(T)]. \end{aligned}$$

Esto es, la transformación T tiene núcleo no trivial. Por lo tanto, existe un polinomio no cero $p(x) \in Poly_d(\mathbb{F}^n)$ que se anula en el conjunto A. □

Para finalizar esta sección, pasemos al segundo ingrediente de la prueba de Dvir. Éste es una perspicaz observación sobre el comportamiento de ciertos polinomios que se anulan en un conjunto de Kakeya.

Teorema 5.2.14. *Sea $\mathbb{F}$ un campo, con $|\mathbb{F}| = q$, y $p(x) \in \mathbb{F}[x_1, x_2, \ldots, x_n]$ un polinomio de grado a lo más $q-1$. Si $p(x)$ se anula en un conjunto de Kakeya $K \subseteq \mathbb{F}^n$, entonces $p(x)$ es el polinomio cero.*

Demostración. Supongamos, por el contrario, que $p(x)$ no es el polinomio cero. Escribamos a éste de la siguiente manera:

$$p(x) = p_0(x) + p_1(x) + \ldots + p_d(x), \tag{5.3}$$

donde $p_i(x) \in \mathbb{F}[x_1, x_2, \ldots, x_n]$ es homogéneo de grado i, para cada $0 \leq i \leq d$. Podemos notar que $p_d(x)$ no es el polinomio cero pues d es justo el grado de $p(x)$. Además, como sabemos que este último se anula en el conjunto no vacío K, se sigue que $d > 0$.

Sea $v \in \mathbb{F}^n \setminus \{0\}$ una dirección arbitraria. Como K es un conjunto de Kakeya, sabemos que éste contiene a la recta $\{u + tv : t \in \mathbb{F}\}$, para algún $u = u_v \in \mathbb{F}^n$. Luego, como el polinomio $p(x)$ se anula en todo K, en particular, se tiene que $p(u + tv) = 0$ para cada $t \in \mathbb{F}$.

Observemos que el lado izquierdo de la ecuación $p(u + tv) = 0$ es un polinomio en la variable t de grado a lo más $q - 1$ –ya que $\deg(p) \leq q - 1$ – que se anula en todo el campo $\mathbb{F}$. Esto implica que $p(u + tv)$ es exactamente el polinomio cero pues, de no serlo, por el teorema del factor (corolario 5.1.11), éste tendría a lo más $q - 1$ raíces en $\mathbb{F}$, pero hemos dicho ya que se anula en $\mathbb{F}$, el cual contiene q elementos.

Ahora, fijemos nuestra atención en el coeficiente del monomio t^d en la ecuación 5.3. Dicho coeficiente es exactamente $p_d(v)$. Lo anterior junto con el hecho que $p(u+tv)$ sea el polinomio cero permite concluir que $p_d(v) = 0$. Pero al ser la elección del vector v arbitraria, resulta que $p_d(v)$ para cada $v \in \mathbb{F}^n \setminus \{0\}$; además, $p_d(0) = 0$ ya que $d > 0$ y $p_d(x)$ es homogéneo de grado d. Así, el polinomio $p_d(x)$ se anula en todo el espacio $\mathbb{F}^n$.

En consecuencia de lo anterior se sigue que $p_d(x)$ debe ser el polinomio cero. Si esto no fuese de esta manera, como $\deg(p_d) \leq q - 1$, según el lema 5.2.4 $p_d(x)$ tendría a lo más

$$dq^{n-1} \leq (q-1)q^{n-1} < q^n$$

raíces en $\mathbb{F}^n$.

Sin embargo, establecimos al principio de la prueba que el polinomio $p_d(x)$ era distinto de cero ya que d es el grado del polinomio $p(x)$. Así hemos llegado a una contradicción y el resultado queda demostrado. □

En pocas palabras, lo que el teorema 5.2.14 asegura es que un polinomio de grado suficientemente pequeño que se anule en un conjunto de Kakeya, debe ser el polinomio cero. Combinando el lema 5.2.13 y el teorema 5.2.14 se obtiene la solución a la conjetura de Kakeya, la cual se abordará en el siguiente capítulo.

Capítulo 6

Solución al problema finito de Kakeya.

El objetivo de este capítulo es presentar la solución de Dvir al problema finito de Kakeya. Sin embargo, vamos a comenzar con el caso particular $n = 2$. (Este caso ya se había establecido desde la publicación del artículo de Wolff (Wol99).) Analógamente a lo que ocurre con la conjetura de Kakeya, es interesante que por muchos años fuera el único avance contundente en la conjetura finita de Kakeya. La prueba que se presenta aquí, y que se debe a Keith M. Rogers (Rog01), es un argumento relativamente sencillo de conteo y, aunque la prueba es notable por sí misma, es de nuestro interés exponerla al lector para hacer clara la diferencia conceptual de la prueba de Dvir con la de Rogers.

La demostración de Rogers toma como base la siguiente observación.

Observación 6.0.1. *Sea $\mathbb{F}$ un campo, con $|\mathbb{F}| = q$.*

(i) *Hay $\dfrac{q^n - 1}{q - 1}$ distintas direcciones para orientar una recta en $\mathbb{F}^n$.*

(ii) *Cualesquiera dos rectas distintas en $\mathbb{F}^n$ se intersecan en a lo más un punto.*

Estas propiedades se establecen, respectivamente, en los lemas 6.0.2 y 6.0.3.

Lema 6.0.2. *Sea $\mathbb{F}$ un campo finito. Dados dos vectores $x_1, x_2 \in \mathbb{F}^n \setminus \{0\}$ y $y \in \mathbb{F}^n$, entonces, $l(y; x_1) = l(y; x_2)$ si, y sólo si, existe $\alpha \neq 0$ en $\mathbb{F}$ tal que $x_1 = \alpha x_2$.*

Demostración. Supongamos primero que existe $\alpha \neq 0$ en $\mathbb{F}$ tal que $x_1 = \alpha x_2$.

1. Si $y + tx_1 \in l(y; x_1)$, con $t \in \mathbb{F}$, entonces $y + tx_1 = y + (t\alpha)\, x_2$, donde $t\alpha \in \mathbb{F}$. Luego, $y + tx_1 \in l(y; x_2)$.

2. Si $y + tx_2 \in l(y; x_2)$, con $t \in \mathbb{F}$, entonces $y + tx_2 = y + \left(\dfrac{t}{\alpha}\right) x_1$, donde $\dfrac{t}{\alpha} \in \mathbb{F}$ (note que se puede dividir entre α pues éste es diferente de cero). Luego, $y + tx_2 \in l(y; x_1)$.

De 1 y 2, se sigue que $l(y; x_1) = l(y; x_2)$.

Recíprocamente, si $l(y; x_1) = l(y; x_2)$, entonces, para $1 = t_1 \in \mathbb{F}$, existe $t_2 \in \mathbb{F}$ tal que $y + x_1 = y + t_2 x_2$. De donde, $x_1 = \alpha x_2$, con $\alpha = t_2$ (note que $\alpha = t_2 \neq 0$, pues, si lo fuera, $x_1 = 0$ contradiciendo que $x_1 \in \mathbb{F}^n \setminus \{0\}$). □

De aquí en adelante se dirá que dos vectores $x_1, x_2 \in \mathbb{F}^n \setminus \{0\}$ apuntan en direcciones distintas si $x_1 \neq \alpha x_2$ para todo $\alpha \neq 0$ en $\mathbb{F}$. Ahora bien, si $|\mathbb{F}| = q$, entonces hay $q^n - 1$ vectores no cero en $\mathbb{F}^n$; además, según el lema anterior, para un vector no cero dado hay otros $q-1$ que apuntan en la misma dirección. Por lo tanto, el número de posibles direcciones es $\dfrac{q^n - 1}{q-1}$.

Lema 6.0.3. *Sea $\mathbb{F}$ un campo finito. Dos rectas distintas $l(y_1; x_1)$ y $l(y_2; x_2)$ en $\mathbb{F}^n$ son ajenas o se intersecan en un punto.*

Demostración. Supongamos que $a, b \in l(y_1; x_1) \cap l(y_2; x_2)$. Entonces

$$y_1 + t'_a x_1 = a = y_2 + t''_a x_2 \tag{6.1}$$

y

$$y_1 + t'_b x_1 = b = y_2 + t''_b x_2, \tag{6.2}$$

con $t'_a, t''_a, t'_b, t''_b \in \mathbb{F}$, y donde $t'_a \neq t'_b, t''_a \neq t''_b$ (si ocurriese la igualdad, en ambos casos, se llegaría a que $a = b$). De 6.1 y 6.2 se obtiene, respectivamente, que

$$y_1 - y_2 = t''_a x_2 - t'_a x_1$$

y

$$y_1 - y_2 = t''_b x_2 - t'_b x_1.$$

De donde, $t''_a x_2 - t'_a x_1 = t''_b x_2 - t'_b x_1$, o bien, $(t'_b - t'_a) x_1 = (t''_b - t''_a) x_2$. En consecuencia,

$$x_1 = \frac{t''_b - t''_a}{t'_b - t'_a} x_2,$$

con $\dfrac{t''_b - t''_a}{t'_b - t'_a} \in \mathbb{F}$ diferente de cero. Por lo tanto, por el lema 6.0.2 , $l(y_1; x_1) = l(y_2; x_2)$; lo cual contradice nuestra suposición inicial. Se concluye entonces que dos rectas distintas en $\mathbb{F}^n$, si se cortan, lo hacen en tan sólo un punto. □

Con esto ya podemos establecer el problema finito de Kakeya para $n = 2$.

Teorema 6.0.4. *Sea $\mathbb{F}$ un campo finito con q elementos. Si K es un conjunto de Kakeya en $\mathbb{F}^2$, entonces*

$$|K| \geq \frac{q^2}{2}.$$

Demostración. Por (i) de la observación 6.0.1 sabemos que hay $\dfrac{q^2 - 1}{q-1} = q + 1$ direcciones para orientar una recta en $\mathbb{F}^2$. Como K debe contener al menos una recta en cada dirección, éste debe contener al menos $q + 1$ rectas distintas; sean $l_1, l_2, ..., l_{q+1}$ dichas

rectas. Además, recordemos que cada una de estas rectas tiene q elementos y que dos rectas distintas se cortan a lo más en un punto. Por lo tanto,

$$\begin{aligned}|K| &\geq |l_1| + |l_2 - l_1| + |l_3 - (l_1 \cup l_2)\,| + ... + |l_{q+1} - (l_1 \cup l_2 \cup ... \cup l_q)\,| \\ &\geq q + (q-1) + ... + 1 + 0 \\ &= \frac{q(q+1)}{2} \\ &\geq \frac{q^2}{2}\end{aligned}$$

□

Ejemplo 6.0.5. Consideremos el espacio vectorial $\mathbb{Z}_3^2$. Por (ii) de la observación 6.0.1 sabemos que hay cuatro posibles direcciones para orientar una recta: tomemos, por ejemplo, $(1,0)$, $(1,1)$, $(1,2)$, $(0,1)$. Utilizando la misma notación que en el teorema 6.0.4, tomemos las rectas l_1, l_2, l_3 y l_4 como sigue:

$$\begin{aligned}l_1 &= \{(1,0) + t(1,1) : t \in \mathbb{Z}_3\} = \{(1,0),(2,1),(0,2)\}, \\ l_2 &= \{(1,0) + t(1,2) : t \in \mathbb{Z}_3\} = \{(1,0),(2,2),(0,1)\}, \\ l_3 &= \{(1,0) + t(1,0) : t \in \mathbb{Z}_3\} = \{(1,0),(2,0),(0,0)\}, \\ l_4 &= \{(2,0) + t(0,1) : t \in \mathbb{Z}_3\} = \{(2,0),(2,1),(2,2)\}.\end{aligned}$$

Ciertamente, cada una de las rectas anteriores interseca a todas las que le preceden. Ahora, consideremos K un conjunto de Kakeya en $\mathbb{Z}_3^2$. Para ilustrar la estrategia que seguimos en la prueba del teorema anterior para acotar la cardinalidad de K, pensemos que los únicos elementos de K son aquellos elementos en $l_1 \cup l_2 \cup l_3 \cup l_4$. Entonces

$$7 = |K| \geq 3 + 2 + 1 + 0 = 6 \geq \frac{3^2}{2}.$$

Pasemos ahora al caso general n. Combinando el lema 5.2.13 y el teorema 5.2.14 se obtiene de manera inmediata la solución a la conjetura finita de Kakeya.

Teorema 6.0.6 (Zeev Dvir)**.** *Sea $\mathbb{F}$ un campo finito. Si $K \subseteq \mathbb{F}^n$ es un conjunto de Kakeya, entonces*

$$|K| \geq \binom{|\mathbb{F}| + n - 1}{n} \geq \frac{|\mathbb{F}|^n}{n!}.$$

Demostración. Supongamos que $|\mathbb{F}| = q$. La segunda desigualdad se sigue de la definición del coeficiente binomial:

$$\binom{q+n-1}{n} = \frac{1}{n!}\, q(q+1)...(q+n-1) \geq \frac{q^n}{n!}.$$

Vamos a concentrarnos, pues, en la primera desigualdad. Si

$$|K| < \binom{q+n-1}{n} = \binom{q+n-1}{q-1},$$

entonces, por el lema 5.2.13 existe un polinomio no cero $p(x) \in \mathbb{F}[x_1, x_2, ..., x_n]$ de grado $d \leq q-1$ que se anula en el conjunto K. Pero esto contradice al teorema 5.2.14. Por lo tanto,

$$|K| \geq \binom{q+n-1}{n} \geq \frac{q^n}{n!},$$

como buscábamos probar. □

Obsérvese que, en realidad, hemos probado que todo conjunto de Kakeya K en $\mathbb{F}^n$ satisface que

$$|K| \geq \binom{n+q-1}{n} = \frac{1}{n!}q^n + \mathcal{O}_n\left(q^{n-1}\right), \tag{6.3}$$

donde $\mathcal{O}_n\left(q^{n-1}\right)$ denota términos que son potencias de q con exponente a lo más $n-1$. Para $n = 2$ la cota inferior que se obtiene a partir de la ecuación 6.3 corrobora información que se conocía previo al artículo de Dvir. Por ejemplo, Terence Tao y Gerd Mockenhaupt probaron en (MT+04) que existen conjuntos de Kakeya en $\mathbb{F}^2$ de cardinalidad $\frac{|\mathbb{F}|^2}{2} + \frac{|\mathbb{F}|}{2}$. Para dimensiones más grandes, sin embargo, la constante $\frac{1}{n!}$ fue mejorada significativamente a $\frac{1}{2^n}$ en (DKSS09). De manera más precisa, Dvir et al. establecieron el siguiente resultado:

Teorema 6.0.7. *Sea $\mathbb{F}$ un campo finito, con $|\mathbb{F}| = q$. Para cualquier conjunto de Kakeya K en $\mathbb{F}^n$, se cumple que $|K| \geq \frac{1}{2^n}q^n$.*

Las ideas desarrolladas en (DKSS09) para llegar al teorema anterior son, en esencia, semejantes a las que presentamos aquí, mas ellos utilizan argumentos más sofisticados sobre polinomios con *ceros de mayores multiplicidades*. Así, posterior a la prueba de Dvir, el estudio de los conjuntos de Kakeya está enfocado en calcular valores exactos para la constante C_n del problema finito de Kakeya. El caso $n = 2$ está completamente resuelto pues tenemos la cota $\frac{q^2}{2}$ y en la siguiente sección veremos una construcción del mismo tamaño. Para $n = 3$, el resultado más reciente se puede encontrar en (LSW16). En este artículo se prueba el siguiente teorema:

Teorema 6.0.8. *Sea $\mathbb{F}$ un campo finito con q elementos. Existe una constante $C > 0$ tal que, si $q > C$ y $K \subseteq \mathbb{F}^3$ es un conjunto de Kakeya, entonces se cumple que*

$$|K| \geq 0.2107q^3.$$

Queda, por tanto, mucho por estudiar sobre los conjuntos finitos de Kakeya.

Para concluir este capítulo, continuamos nuestra descripción de los conjuntos de Kakeya proporcionando una cota superior para éstos. Contrastando los teoremas 6.0.7 y 6.0.11 podemos observar que hay un factor de 2 entre la cota inferior y superior de los conjuntos de Kakeya, y un interesante problema abierto es tratar de reducir dicho factor.

Antes de presentar con todo detalle el teorema 6.0.11, enunciaremos la siguiente definición y un lema que será fundamental para establecer la cota superior buscada.

Definición 6.0.9. *Sea $\mathbb{F}$ un campo. Si existe un entero positivo m tal que $ma = 0$ para todo $a \in \mathbb{F}$, entonces al menor de dichos enteros positivos m se le llamará la* ***característica del campo*** *$\mathbb{F}$, y se dirá que $\mathbb{F}$ tiene* ***característica positiva****. Si no existen dichos enteros positivos, entonces se dirá que $\mathbb{F}$ es de* ***característica cero****.*

Lema 6.0.10. *Sea $\mathbb{F}$ un campo finito con q elementos.*

(i) Si $\mathbb{F}$ es de característica impar, el número de cuadrados [1] en $\mathbb{F}$ es $\frac{1}{2}(q+1)$.

(ii) Si Si $\mathbb{F}$ es de característica par, el número de cuadrados en $\mathbb{F}$ es q.

Demostración. Ver lema B.1.54 en el apéndice B. □

Pasemos ahora al resultado central de esta sección: el conjunto de Kakeya $K \subseteq \mathbb{F}^n$ más pequeño que se conoce tiene cardinalidad

$$|K| \leq \frac{q^n}{2^{n-1}} + \mathcal{O}_n\left(q^{n-1}\right).$$

La construcción se debe a Dvir (DKSS09) (para el caso en que $\mathbb{F}$ es un campo de característica par) y a Shubhangi Saraf y Madhu Sudan (SS08) (para el caso en que $\mathbb{F}$ es de característica par), y ambas son una generalización de la construcción hecha por Mockenhaupt y Tao para el caso $n = 2$ que mencionamos antes.

Teorema 6.0.11. *Sea $\mathbb{F}$ un campo finito de cardinalidad q. Para cada $n \geq 2$ existe un conjunto de Kakeya K de cardinalidad*

$$|K| \leq \frac{q^n}{2^{n-1}} + \mathcal{O}_n\left(q^{n-1}\right).$$

Demostración. Para realizar la demostración, se considerarán dos casos según la característica de $\mathbb{F}$ sea par o impar.

Característica impar. Consideremos el conjunto

$$K_n = \left\{(\alpha_1, ..., \alpha_{n-1}, \beta) \in \mathbb{F}^n : \alpha_i, \beta \in \mathbb{F} \text{ y } \alpha_i + \beta^2 \text{ es un cuadrado en } \mathbb{F}\right\},$$

y sea $K = K_n \cup \left(\mathbb{F}^{n-1} \times \{0\}\right)$. Se afirma que (i) K es un conjunto de Kakeya y (ii) K tiene cardinalidad como se indica en el enunciado de este teorema.

(i) Por definición de conjunto de Kakeya, tenemos que mostrar que, para cualquier $v = (v_1, ..., v_n) \neq 0$ en $\mathbb{F}^n$, existe $a = a_v \in \mathbb{F}^n$ tal que la recta

$$l(a; v) = \{a + tv : t \in \mathbb{F}\} \subseteq K.$$

Caso 1. Si $v_n = 0$, para $a = (0, ..., 0) \in \mathbb{F}^n$ se tiene que $a + tv \in \mathbb{F}^{n-1} \times \{0\} \subseteq K$, para todo $t \in \mathbb{F}$. Luego, $l(a; v) \subseteq K$.

[1] Sea $\mathbb{F}$ un campo. Diremos que $a \in \mathbb{F}$ es un cuadrado en $\mathbb{F}$, si existe $b \in \mathbb{F}$ tal que $a = b^2$.

Caso 2. Si $v_n \neq 0$, hagamos

$$a = \left(\left(\frac{v_1}{2v_n} \right)^2, ..., \left(\frac{v_{n-1}}{2v_n} \right)^2, 0 \right).$$

Entonces, para cada $t \in \mathbb{F}$, el punto $a+tv$ tiene coordenadas $(\alpha_1, ..., \alpha_{n-1}, \beta)$, donde $\alpha_i = \left(\frac{v_i}{2v_n} \right)^2 + tv_i$ y $\beta = tv_n$. De aquí tenemos que

$$\alpha_i + \beta^2 = \left(\frac{v_i}{2v_n} + tv_n \right)^2,$$

el cual es un cuadrado para cada $i = 1, ..., n-1$.
Por lo tanto, $a + tv \in K_n \subseteq K$, de donde, $l(a; v) \subseteq K$.

Por los casos 1 y 2 de (i) se concluye que K es, en efecto, un conjunto de Kakeya.

(ii) Veamos ahora que la cardinalidad de K satisface la cota superior indicada en el enunciado del teorema. Primero, note que la cardinalidad de K_n es exactamente

$$|K_n| = q \left(\frac{q+1}{2} \right)^{n-1}.$$

Esto es así pues hay q posibilidades para seleccionar a $\beta \in \mathbb{F}$ y $\frac{q+1}{2}$ elecciones para cada $\alpha_i + \beta^2$, con $i = 1, ..., n-1$. Así, la cardinalidad del conjunto K es a lo más

$$|K| \leq |K_n| + q^{n-1} = q \left(\frac{q+1}{2} \right)^{n-1} + q^{n-1} = \frac{q^n}{2^{n-1}} + \mathcal{O}_n\left(q^{n-1}\right).$$

Característica par. Sea

$$K = \left\{ (\alpha_1, ..., \alpha_{n-1}, \beta) \in \mathbb{F}^n : \alpha_i, \beta \in \mathbb{F} \text{ y } \exists \gamma_i \in \mathbb{F} \text{ tal que } \alpha_i = \gamma_i^2 + \gamma_i \beta \right\}.$$

Nótese que K así definido ya contiene al conjunto $\mathbb{F}^{n-1} \times \{0\}$ que se unió a K_n en la construcción anterior. Esto es así ya que al ser el campo $\mathbb{F}$ de característica par, cada elemento de $\mathbb{F}$ es un cuadrado.

Nuevamente veamos que (i) K es un conjunto de Kakeya y (ii) K tiene la cardinalidad indicada.

(i) Una vez más, por la definición de conjunto de Kakeya, debemos probar que, para cada $v = (v_1, ..., v_n) \neq 0$ en $\mathbb{F}^n$, existe $a = a_v \in \mathbb{F}^n$ tal que la recta

$$l(a; v) = \{a + tv : t \in \mathbb{F}\} \subseteq K.$$

Caso 1. Si $v_n = 0$, al igual en la construcción anterior, tomemos $a = (0, ..., 0) \in \mathbb{F}^n$. Entonces, para cada $t \in \mathbb{F}$,

$$a + tv = (tv_1, ..., tv_{n-1}, 0) \in \mathbb{F}^{n-1} \times \{0\} \subseteq K.$$

Por lo tanto, $l(a; v) \subseteq K$.

Caso 2. Si $v_n \neq 0$, consideremos

$$a = \left(\left(\frac{v_1}{v_n} \right)^2, ..., \left(\frac{v_{n-1}}{v_n} \right)^2, 0 \right).$$

Así, para cada $t \in \mathbb{F}$, el punto $a + tv$ tiene coordenadas $(\alpha_1, ..., \alpha_{n-1}, \beta)$, con $\alpha_i = \left(\frac{v_i}{v_n} \right)^2 + tv_i$ y $\beta = tv_n$. Con esto en cuenta, haciendo $\gamma_i = \frac{v_i}{v_n}$, para cada $i = 1, ..., n-1$, se obtiene que

$$\gamma_i^2 + \gamma_i \beta = \left(\frac{v_i}{v_n} \right)^2 + tv_i = \alpha_i.$$

Por lo tanto, $a + tv \in K$, lo cual permite concluir que $l(a; v) \subseteq K$.

Por los casos 1 y 2 de (i) se concluye que K es, ciertamente, un conjunto de Kakeya.

(ii) El número de elementos de la forma $(\alpha_1, ..., \alpha_{n-1}, 0) \in K$ es q^{n-1} (una vez más estamos utilizando el hecho que, al ser $\mathbb{F}$ un campo de característica par, todos sus elementos son cuadrados). Así, queda por estimar el número de puntos de la forma $(\alpha_1, ..., \alpha_{n-1}, \beta) \in K$ para $\beta \neq 0$ fijo. Se afirma primero que

$$| \{ \gamma^2 + \beta\gamma : \gamma \in \mathbb{F} \} | = \frac{q}{2}.$$

Para probar la afirmación, hagamos $A = \{ \gamma^2 + \beta\gamma : \gamma \in \mathbb{F} \}$ y consideremos la función $f : \mathbb{F} \longrightarrow A$ tal que $f(\gamma) = \gamma^2 + \gamma\beta$. Nótese que para esta función se cumple que, para toda $\gamma \in \mathbb{F}$, $f(\gamma) = f(\tau)$ con $\tau = \gamma + \beta \neq \gamma$. En efecto,

$$\begin{aligned} f(\tau) &= \tau^2 + \tau\beta \\ &= (\gamma + \beta)^2 + (\gamma + \beta)\beta \\ &= \gamma^2 + \gamma\beta + 2(\gamma\beta + \beta^2) \\ &= \gamma^2 + \gamma\beta \\ &= f(\gamma) \end{aligned}$$

donde $2(\gamma\beta + \beta^2) = 0$ ya que el campo $\mathbb{F}$ es de característica par. Así, la función f «envía» dos elementos distintos de $\mathbb{F}$ a uno mismo en A, esto es, la función f es 2 a 1 sobre su imagen. Luego, el conjunto A tiene la cardinalidad que se afirmó. En

consecuencia, para $\beta \neq 0$ fijo, el número de puntos de la forma $(\gamma_1, ..., \gamma_{n-1}, \beta) \in K$ es $\left(\frac{q}{2}\right)^{n-1}$.

Por lo tanto, el conjunto K tiene cardinalidad

$$|K| = (q-1)\left(\frac{q}{2}\right)^{n-1} + q^{n-1} = \frac{q^n}{2^{n-1}} + \mathcal{O}_n\left(q^{n-1}\right) \tag{6.4}$$

(el factor $q-1$ después de la primera igualdad en 6.4 corresponde a las posibles elecciones de $\beta \neq 0$ en $\mathbb{F}$).

□

Capítulo 7

El problema finito de Nikodym.

7.1. Introducción.

Como se mencionó en la introducción, se sabe que el artículo de Zeev Dvir (Dvi09) significó el principio de una nueva técnica, conocida como ***el método polinomial***, que en los últimos años ha permitido la solución de problemas combinatorios utilizando polinomios en varias variables. A grandes rasgos, el *método polinomial* impone un estructura algebraica a problemas geométricos–combinatorios (es decir, problemas que sólo involucran puntos, rectas e *incidencias*) y se puede resumir, *grosso modo*, como sigue:

El problema que se busca resolver es entender ciertas propiedades de un subconjunto A de un espacio vectorial finito. Para esto, se encuentra un polinomio, de grado «conveniente», que se anule en el conjunto A; después se usa dicho polinomio para estudiar, ya con un nuevo enfoque establecido –a saber, el enfoque polinomial–, el problema original. Lo verdaderamente atractivo, la sorpresa, de esta nueva estrategia –como se pudo apreciar en la solución del problema finito de Kakeya– es que al estudiar conjuntos finitos, la transición del problema original al estudio de polinomios requiere únicamente argumentos elementales de álgebra lineal.

En lugar de abordar *el método polinomial* desde un punto de vista teórico, en este capítulo se presentará, a través de una aplicación íntimamente relacionada a los conjuntos de Kakeya, la nueva técnica en plena acción. En este sentido, en la práctica se pueden distinguir los siguientes tres pasos a seguir para utilizar *el método polinomial*:

(a) Dado un conjunto $A \subseteq \mathbb{F}^n$, con $\mathbb{F}$ un campo finito, el problema a resolver es acotar la cardinalidad de A.

(b) Encontrar (vía el lema 5.2.13) un polinomio $p(x) \in \mathbb{F}[x_1, x_2, ..., x_n]$, diferente de cero, de grado «suficientemente pequeño».

(c) Utilizar las propiedades inherentes al conjunto A junto con el polinomio $p(x)$ para concluir que este último se anula en «demasiados» puntos de $\mathbb{F}^n$. Concluir una contradicción (apelando al teorema 5.2.4).

Intuitivamente, el *método polinomial* es una técnica eficaz, pues, como ya vimos en el corolario 5.1.11 y en el teorema 5.2.4, el número de raíces de un polinomio está

estrechamente relacionado con el grado del mismo.

7.1.1. El problema finito de Nikodym.

El problema que abordaremos para el ilustrar *el método polinomial* será el ***problema finito de Nikodym***. Comenzamos definiendo *un conjunto finito de Nikodym.*

Definición 7.1.1. *Sea $\mathbb{F}$ un campo finito con q elementos. Un conjunto $N \subseteq \mathbb{F}^n$ será llamado **conjunto finito de Nikodym** si, para cada $x \in \mathbb{F}^n \setminus N$, existe $y \neq 0$ en $\mathbb{F}^n$ tal que la recta $l(x;y)$ que pasa por x en la dirección de y satisface que*

$$l^*(x;y) = l(x;y) \setminus \{x\} \subseteq N.$$

Ejemplo 7.1.2.

(i) Claramente, dado un campo finito $\mathbb{F}$, el $\mathbb{F}$-espacio vectorial $\mathbb{F}^n$ es un conjunto finito de Nikodym.

(ii) Sea
$$N = \{(1,1),(2,2),(1,2),(2,0),(1,0),(2,1)\} \subseteq \mathbb{Z}_3^2.$$
Veamos que N es un conjunto finito de Nikodym. En efecto;
$$\mathbb{Z}_3^2 \setminus N = \{(0,0),(0,1),(0,2)\}.$$
Ahora, el vector $(1,1) \in \mathbb{Z}_3^2$ satisface que
$$\begin{aligned} l^*((0,0);(1,1)) &= \{(0,0)+t(1,1) : t \in \mathbb{F}\} \setminus \{(0,0)\} = \{(1,1),(2,2)\} \subseteq N, \\ l^*((0,1);(1,1)) &= \{(0,1)+t(1,1) : t \in \mathbb{F}\} \setminus \{(0,1)\} = \{(1,2),(2,0)\} \subseteq N, \\ l^*((0,2);(1,1)) &= \{(0,2)+t(1,1) : t \in \mathbb{F}\} \setminus \{(0,2)\} = \{(1,0),(2,1)\} \subseteq N. \end{aligned}$$
Por lo tanto, $N \subseteq \mathbb{Z}_3^2$ es un conjunto finito de Nikodym.

De forma similar a los conjuntos finitos de Kakeya, la pregunta que nos ocupará en lo que sigue es: ¿qué tan pequeño (en el sentido de la cardinalidad del conjunto) puede ser un conjunto de Nikodym? Esta pregunta lleva a la siguiente conjetura.

Conjetura 7.1.3 (Problema finito de Nikodym)**.** *Sea $\mathbb{F}$ un campo finito. Si $N \subseteq \mathbb{F}^n$ es un conjunto finito de Nikodym, entonces*

$$|N| \geq C_n |\mathbb{F}|^n,$$

donde $C_n > 0$ depende sólo de n, pero no de la caridalidad de $\mathbb{F}$. [1]

[1] Al igual que en la conjetura finita de Kakeya, en el problema finito de Nikodym se debe pensar a n como fijo y se deben considerar valores muy grandes para la cardinalidad del campo $\mathbb{F}$.

Al igual que en el caso del problema finito de Kakeya, el correspondiente para el conjunto de Nikodym también es un «modelo a escala» de un problema en el espacio euclidiano. A continuación se presenta una breve introducción histórica a dicho problema.

En 1927 el matemático polaco Otto M. Nikodym, estudiando la estructura geométrica de *los conjuntos medibles en el plano*, mostró en (Nik27) cómo construir en $\mathbb{R}^2$ un conjunto N_0 dentro del cuadrado de lado uno $Q = [0,1]^2$ (que llamaremos cuadrado unitario), tal que N_0 tiene área $A(N_0) = 1$ y satisface la siguiente propiedad: para cada $x \in N_0$, existe una recta $l(x) \subseteq \mathbb{R}^2$ que pasa por x para la cual $l(x) \cap N_0 = \{x\}$ (cuando un conjunto cumple esta propiedad usualmente se le llama *linealmente accesible*).

La construcción original de Nikodym es complicada, pero fue significativamente simplificada por Davies (Dav52) quién, además, probó que era posible construir el conjunto N_0 de forma que exista un número no numerable de rectas por cada punto $x \in N_0$ que sólo intersequen a N_0 en x. De manera sorpresiva la construcción del conjunto N_0 puede realizarse a partir de los árboles de Perron–Schoenberg, hecho que establece un punto de contacto con los conjuntos de Kakeya en $\mathbb{R}^2$. Para los detalles de esta construcción nos referiremos al texto de M. de Guzmán (dG76).

Con lo anterior en cuenta, llamaremos *conjuntos de Nikodym* a los complementos de conjuntos como N_0. Más precisamente, diremos que un conjunto $N \subseteq [0,1]^n$ es un ***conjunto de Nikodym***, si tiene *medida de Lebesgue cero* y para cada $x \in [0,1]^n$ existe una recta $l(x) \subseteq \mathbb{R}^n$ que pasa por x tal que $l(x) \cap [0,1]^n \setminus \{x\}$ está contenido en N. Al igual que en el caso de los conjuntos de Besicovitch, se tiene la siguiente conjetura sobre los conjuntos de Nikodym:

Conjetura 7.1.4 (Conjetura de Nikodym)**.** *Todo conjunto de Nikodym en $\mathbb{R}^n$ tiene dimensión* [1] *igual a n.*

La conjetura de Kakeya y la correspondiente para los conjuntos de Nikodym están estrechamente relacionadas: la conjetura de Kakeya implica la conjetura de Nikodym (ver (Mat15b)). Esta relación entre los conjuntos de Kakeya y de Nikodym hacen presentir la posibilidad de una relación profunda entre ambos conjuntos. En efecto, dichos conjuntos se han estudiado paralelamente, y así como hay un vínculo entre ambas conjeturas, se ha logrado relacionar a los dos conjuntos a partir de formulaciones distintas de la conjetura de Kakeya que no contemplaremos aquí (una vez más, sacrificio hecho para evitar la parte técnica de estos conceptos), pero que pueden consultarse en (Mat15b), (Car92) y (T$^+$99) para el lector interesado. En ese sentido, para el contexto que nos aqueja, es decir, el contexto de los campos finitos, ya no es ninguna sorpresa que también haya una relación entre los conjuntos finitos de Kakeya y de Nikodym. Más adelante estudiaremos a fondo esta relación (ver teoremas 7.2.8 y 7.2.9). Nuestra primera tarea, sin embargo, será resolver el problema finito de Nikodym.

[1] *Dimensión Hausdorff* o *Minkowski*

7.2. Solución al problema finito de Nikodym.

En todo lo que sigue, por brevedad, llamaremos a un conjunto finito de Nikodym simplemente como *conjunto de Nikodym*. Para comenzar, conviene presentar los siguientes dos lemas.

Lema 7.2.1 (Lema de anulación). *Sea $\mathbb{F}$ un campo. Si un polinomio $p(x)$ de grado d en $\mathbb{F}[x_1, x_2, ..., x_n]$ se anula en $d+1$ puntos de la recta $l(y;x) = \{y + tx : t \in \mathbb{F}\}$, entonces, en realidad, $p(x)$ se anula en toda la recta $l(y;x)$.*

Observación 7.2.2. *Claramente, estamos presuponiendo en el enunciado del resultado anterior que, en el caso de ser $\mathbb{F}$ un campo finito, $d < |\mathbb{F}|$ ya que de otro modo no se cumplirían las hipótesis del lema. Es decir, si el grado del polinomio $p(x)$ fuese igual a $|\mathbb{F}|$, o mayor, no habría $d+1$ puntos en los cuales $p(x)$ pueda anularse.*

Demostración. Definamos el polinomio $q(t) \in \mathbb{F}[t]$ como $q(t) = p(y + tx)$, es decir, $q(t)$ es el polinomio $p(x)$ restringido al a recta $l(y;x)$. Así, $q(t)$ es un polinomio en una variable de grado menor o igual que d. Ahora bien, puesto que el polinomio $p(x)$ se anula en $d+1$ puntos de la recta $l(y;x)$, se sigue que $q(t)$ se anula en $d+1$ valores de $t \in \mathbb{F}$ (en otras palabras, $q(t)$ tiene $d+1$ raíces en $\mathbb{F}$). De aquí se puede concluir que $q(t)$ es el polinomio cero. En efecto, si esto no fuese así, por el corolario 5.1.11) $q(t)$ tendría a lo más d raíces en $\mathbb{F}$; lo cual es una contradicción. Por lo tanto $q(t)$ es el polinomio cero, y, en consecuencia, $p(x)$ se anula en toda la recta $l(y;x)$. □

El siguiente lema es una ligera modificación al lema 5.2.13.

Lema 7.2.3. *Sea $\mathbb{F}$ un campo. Para cualquier conjunto finito $A \subseteq \mathbb{F}^n$, existe un polinomio diferente de cero $p(x)$ en $\mathbb{F}[x_1, x_2, ..., x_n]$ que se anula en A y de grado a lo más $n|A|^{1/n}$.*

Demostración. Supongamos que $n \geq 2$ [1], y definamos $d \geq 0$ como el mayor entero tal que $d \leq n|A|^{1/n}$. Por como se ha elegido a d, se sigue que $d+1$ satisface que

$$d + 1 > n|A|^{1/n},$$

de donde,

$$\left(\frac{d+1}{n}\right)^n > |A|.$$

Ahora bien,

$$\binom{n+d}{d} = \binom{n+d}{n} = \frac{(d+1)(d+2)...(d+n)}{1 \cdot 2 \ ... \ n} > \frac{(d+1)^n}{n^n},$$

por lo cual $|A| < \binom{n+d}{d}$. Luego, por el lema 5.2.13 existe un polinomio diferente de cero $p(x) \in \mathbb{F}[x_1, x_2, ..., x_n]$ de grado a lo más d y que se anula en el conjunto A. □

[1] El caso $n = 1$ ya quedó establecido en el corolario 5.1.11

Ahora sí estamos listos para dar solución al problema finito de Nikodym. Se presentan aquí dos soluciones distintas: la primera de ellas puede consultarse en el texto de Larry Guth (Gut16), mientras que la segunda surge como una consecuencia de un resultado de Liangpan Li sobre cotas inferiores que satisfacen los conjuntos de Nikodym (ver (Li08)). Comencemos por la solución de Larry Guth.

Teorema 7.2.4. *Sea $\mathbb{F}$ un campo, con $|\mathbb{F}| = q$. Si $N \subseteq \mathbb{F}^n$ es un conjunto de Nikodym, entonces*

$$|N| \geq \frac{1}{(10n)^n} q^n.$$

Demostración. Sea $N \subseteq \mathbb{F}^n$ un conjunto de Nikodym. Supongamos, por el contrario, que

$$|N| < \frac{1}{(10n)^n} q^n.$$

Entonces, por el lema 7.2.3, existe un polinomio no cero $p(x) \in \mathbb{F}[x_1, x_2, ...x_n]$ que se anula en N y tal que

$$\deg(p) \leq n|N|^{1/n} \leq n \left[\frac{q^n}{(10n)^n} \right]^{\frac{1}{n}} = \frac{q}{10} < q - 1.$$

Utilizando las propiedades del conjunto N veamos que, de hecho, el polinomio $p(x)$ se anula en todo el espacio $\mathbb{F}^n$. Sea x_0 un punto arbitrario de $\mathbb{F}^n \setminus N$. Por la definición de conjunto de Nikodym, existe $y \in \mathbb{F}^n$ diferente de cero tal que $l^*(x_0; y) \subseteq N$. Ahora bien, el polinomio $p(x)$ se anula en N, así que $p(x)$ se anula en $q - 1$ puntos de la recta $l(x_0; y)$. Puesto que $\deg(p) < q - 1$, el lema de anulación implica que $p(x)$, en realidad, se anula en toda la recta $l(x_0; y)$. En particular, $p(x_0) = 0$, y, como la elección de $x_0 \in \mathbb{F}^n \setminus N$ fue arbitraria, se sigue que $p(x)$ se anula completamente en $\mathbb{F}^n \setminus N$. Lo anterior, junto con el hecho que $p(x)$ se anula en el conjunto N, permite concluir que $p(x)$ se anula todo $\mathbb{F}^n$, como se afirmó.

Tenemos, entonces, que $p(x) = 0$ para cada $x \in \mathbb{F}^n$ y que $\deg(p) < q$, así, según el lema 5.2.3, $p(x)$ debe ser el polinomio cero. Sin embargo, se estableció al principio de la prueba que $p(x)$ era un polinomio no cero. Por tanto hemos llegado a una contradicción y el resultado queda demostrado. □

Pasemos ahora a la segunda solución del problema finito de Nikodym. Para esto, precisemos primero la cota inferior para los conjuntos de Nikodym establecida por Liangpan Li.

Teorema 7.2.5. *Sea $\mathbb{F}$ un campo finito con q elementos. Cualquier conjunto de Nikodym $N \subseteq \mathbb{F}^n$ satisface que*

$$|N| \geq \binom{n+q-2}{n}.$$

Demostración. Sea $N \subseteq \mathbb{F}^n$ un conjunto de Nikodym. Supongamos, por el contrario, que

$$|N| < \binom{n+q-2}{n}.$$

Entonces, por el lema 5.2.13, existe un polinomio no cero $p(x) \in \mathbb{F}[x_1, x_2, ...x_n]$ que se anula en N y de grado a lo más $q-2 < q-1$.

Al igual que en el teorema 7.2.4, veamos que el polinomio $p(x)$ se anula en todo el espacio $\mathbb{F}^n$. Sea x_0 un punto cualquiera de $\mathbb{F}^n \setminus N$. Por la definición de conjunto de Nikodym, existe $y \neq 0$ en $\mathbb{F}^n$ tal que $l^*(x_0; y) \subseteq N$. Ahora bien, el polinomio $p(x)$ se anula en N, así que $p(x)$ se anula en $q-1$ puntos de la recta $l(x_0; y)$. Puesto que $\deg(p) < q-1$, el lema de anulación implica que $p(x)$, en realidad, se anula en toda la recta $l(x_0; y)$. En particular, $p(x_0) = 0$, y, al ser $x_0 \in \mathbb{F}^n \setminus N$ arbitrario, resulta que $p(x)$ se anula completamente en $\mathbb{F}^n$.

Así, $p(x) = 0$ para cada $x \in \mathbb{F}^n$ y $\deg(p) < q$, por lo cual, según el lema 5.2.3, $p(x)$ debe ser el polinomio cero. Sin embargo, al principio de la prueba se dijo que $p(x)$ era un polinomio no cero. Por tanto hemos llegado a una contradicción y el resultado queda demostrado. □

Observación 7.2.6. *Alternativamente, tanto la prueba del teorema anterior como la del teorema 7.2.4 se pueden concluir a partir del teorema 5.2.4 como sigue:*

Después de probar que $p(x)$ se anula en todo $\mathbb{F}^n$, se puede decir lo siguiente: al ser $p(x)$ un polinomio diferente de cero y de grado $\deg(p) \leq q-2$, por el teorema 5.2.4, éste tiene a lo más

$$(q-2)q^{n-1} < q \cdot q^{n-1} = q^n$$

raíces en $\mathbb{F}^n$. Lo cual contradice que $p(x) = 0$ para cada $x \in \mathbb{F}^n$.

Como consecuencia del teorema 7.2.5 se tiene el siguiente resultado:

Corolario 7.2.7. *Sea $\mathbb{F}$ un campo, con $|\mathbb{F}| = q$. Si $N \subseteq \mathbb{F}^n$ es un conjunto de Nikodym, entonces*

$$|N| \geq \frac{1}{2n!} q^n.$$

Demostración. Sea $N \subseteq \mathbb{F}^n$ un conjunto de Nikodym. Entonces, por el teorema 7.2.5 sabemos que

$$|N| \geq \binom{n+q-2}{n}.$$

Ahora bien, como $q \geq 2$, entonces:

(a)
$$\binom{n+q-2}{n} = \frac{(n+q-2)...q(q-1)}{n!} \geq \frac{q^{n-1}(q-1)}{n!}.$$

(b) Como $q^{n-1} \leq \dfrac{q^n}{2}$, se sigue que

$$\frac{q^n}{n!} - \frac{q^{n-1}}{n!} = \frac{q^n}{2n!} + \left(\frac{q^n}{2} - q^{n-1}\right)\left(\frac{1}{n!}\right) \geq \frac{1}{2n!}q^n.$$

Juntando (a) y (b) se concluye el resultado. □

Una vez que se dio solución al problema finito de Nikodym, podemos pasar a la siguiente pregunta: ¿cómo se relacionan los conjuntos de Kakeya con los conjuntos de Nikodym? Una primera respuesta se obtiene a partir del siguiente teorema.

Teorema 7.2.8. *Sea $\mathbb{F}$ un campo finito de cardinalidad q. Si $K \subseteq \mathbb{F}^n$ es un conjunto de Kakeya, entonces el conjunto*

$$N = \{tk : t \in \mathbb{F}, k \in K\}$$

es un conjunto de Nikodym. Así, $|N| \leq q|K|$.

Demostración. Sea $K \subseteq \mathbb{F}^n$ un conjunto de Kakeya. Por definición de conjunto de Nikodym, necesitamos demostrar que, para cada $x \in \mathbb{F}^n \setminus N$, existe $y \neq 0$ en $\mathbb{F}^n$ tal que $l^*(x;y) \subseteq N$.

Sea $x \in \mathbb{F}^n \setminus N$. Obsérvese que $0 \in N$, así que $x \neq 0$. Luego, como K es un conjunto de Kakeya, existe $y_x \in \mathbb{F}^n$ tal que $y_x + sx \in K$, para todo $s \in \mathbb{F}$. Se distinguen los siguientes dos casos sobre $y_x \in \mathbb{F}^n$:

(i) Si $y_x \neq 0$. En este caso, haciendo $y = y_x$ se sigue que

$$l^*(x;y) = \left\{x + s^{-1}y : s \in \mathbb{F} \setminus \{0\}\right\} = \left\{s^{-1}(y + sx) : s \in \mathbb{F} \setminus \{0\}\right\} \subseteq N.$$

(ii) Si $y_x = 0$. En esta situación se tiene que $sx \in K$, para toda $s \in \mathbb{F}$. Ahora bien, si $s_0 \in \mathbb{F}$ es no cero, $s_0 x \neq 0$. Luego, haciendo $y = s_0 x$ se obtiene que

$$l^*(x;y) = \{x + r(s_0 x) : r \in \mathbb{F} \setminus \{0\}\} = \left\{r^{-1}[(r + s_0)x] : r \in \mathbb{F} \setminus \{0\}\right\} \subseteq N.$$

De los casos (i) y (ii) se concluye que N es un conjunto de Nikodym. □

Una segunda conexión entre los conjuntos de Kakeya y de Nikodym puede extraerse de la primera versión de Dvir para solución al problema finito de Kakeya. Dvir menciona en (Dvi09) que el uso del *método polinomial* lo llevó en primera instancia a una cota inferior de la forma $C_n q^{n-1}$ (posteriormente, con algunas observaciones hechas por Terence Tao y Noga Alon, pudo finalmente establecer la cota requerida); dicha cota puede obtenerse tomando el conjunto de Nikodym que se consideró en el teorema anterior. La prueba que se presenta a continuación contiene ligeras modificaciones de los argumentos originales de Dvir, pero el espíritu de su prueba original permanece a través del uso del *método polinomial*.

Teorema 7.2.9. *Sea $\mathbb{F}$ un campo finito, con $|\mathbb{F}| = q$. Si $K \subseteq \mathbb{F}^n$ es un conjunto de Kakeya, entonces*

$$|K| \geq \left(\frac{1}{10n}\right)^n q^{n-1}.$$

Demostración. Supongamos, por el contrario, que

$$|K| < \left(\frac{1}{10n}\right)^n q^{n-1}.$$

Sea N un conjunto de Nikodym como en el teorema 7.2.8. Así, el conjunto N satisface que

$$|N| \leq q|K| \leq \left(\frac{1}{10n}\right)^n q^n.$$

Ahora bien, por el lema 7.2.3, podemos encontrar un polinomio diferente de cero $p(x) \in \mathbb{F}[x_1, x_2, ..., x_n]$ que se anula en N y de grado

$$\deg(p) \leq n|N|^{1/n} \leq n\left[\frac{q^n}{(10n)^n}\right]^{\frac{1}{n}} = \frac{q}{10} < q - 1.$$

Sea x_0 un punto cualquiera de $\mathbb{F}^n \setminus N$. Por la definición de conjunto de Nikodym, existe $y \neq 0$ en $\mathbb{F}^n$ tal que $l^*(x_0; y) \subseteq N$. Ahora bien, el polinomio $p(x)$ se anula en N, así que $p(x)$ se anula en $q-1$ puntos de la recta $l(x_0; y)$. Puesto que $\deg(p) < q - 1$, a partir del lema de anulación 7.2.1 se sigue que $p(x)$ se anula en toda la recta $l(x_0; y)$. En particular, $p(x_0) = 0$, y, al ser $x_0 \in \mathbb{F}^n \setminus N$ arbitrario, resulta que $p(x)$ se anula completamente en $\mathbb{F}^n \setminus N$. Lo anterior, junto con el hecho que el polinomio $p(x)$ se anule en N, permite concluir que $p(x)$ se anula en todo $\mathbb{F}^n$.

Por otro lado, al ser $p(x)$ un polinomio diferente de cero y de grado $\deg(p) \leq q - 2$, por el teorema 5.2.4, éste tiene a lo más

$$(q-2)q^{n-1} < q \cdot q^{n-1} = q^n$$

raíces en $\mathbb{F}^n$. Lo cual contradice que $p(x) = 0$ para cada $x \in \mathbb{F}^n$. Por lo tanto,

$$|K| \geq \left(\frac{1}{10n}\right)^n q^{n-1}.$$

□

Una vez resuelta la conjetura finita de Nikodym, el estudio de estos conjuntos se ha centrado en tratar de establecer mejores cotas inferiores para éstos. En ese sentido, para concluir este capítulo, se presentarán en la siguiente sección algunos resultados para el caso $n = 2$ y, muy brevemente, se pasará revista de lo que se sabe para mayores dimensiones.

7.3. Conjuntos de Nikodym en dos dimensiones.

En contraste con la cota que se obtiene en el teorema 6.0.6 para los conjuntos de Kakeya, la correspondiente cota para los conjuntos de Nikodym en el teorema 7.2.5 no se compagina con las estimaciones conocidas del tamaño de dichos conjuntos. Para $n = 2$, por ejemplo, el teorema 7.2.5 arroja una cota de la forma

$$\binom{q+2-2}{2} = \frac{q(q-1)}{2} = \frac{q^2}{2} + \mathcal{O}_n(q).$$

Como veremos en el siguiente teorema (ver (Li08)), esta cota se puede mejorar a $\frac{2q^2}{3} + \mathcal{O}_n(q)$, que es mayor que $q^2/2 + \mathcal{O}_n(q)$ para $q \geq 14$.

Teorema 7.3.1. *Sea* F *un campo, con* $|\mathbb{F}| = q$. *Cualquier conjunto de Nikodym* $N \in \mathbb{F}^2$ *satisface que*

$$|N| \geq \frac{2}{3}q^2 + \mathcal{O}_n(q).$$

Demostración. Sea s el máximo entero menor o igual que $\frac{q}{3}$. La prueba se hará por casos.

I. Si $|\mathbb{F}^2 \setminus N| \leq s(q-1) + 2q$. Entonces,

$$\begin{aligned} |N| &\geq q^2 - s(q-1) - 2q \\ &\geq q^2 - \frac{q}{3}(q-1) - 2q \\ &= \frac{2}{3}q^2 - \frac{5}{3}q. \end{aligned}$$

II. Si $|\mathbb{F}^2 \setminus N| \geq s(q-1) + 2q$. Como N es un conjunto de Nikodym, para cada $x \in \mathbb{F}^2 \setminus N$ existe $y_x \neq 0$ en $\mathbb{F}^2$ tal que $l^*(x; y_x) \subseteq N$. Ahora bien, por (i) de la observación 6.0.1 sabemos que hay $q+1$ posibles direcciones para orientar una recta en $\mathbb{F}^2$, por lo cual, podemos partir el conjunto

$$\left\{ l(x; y_x) \subseteq \mathbb{F}^2 : x \in \mathbb{F}^2 \setminus N \right\}$$

en $q+1$ de acuerdo a las direcciones de cada recta. Sea $\{G_i : 0 \leq i \leq q\}$ dicha partición. Sin perdida de generalidad, podemos asumir que

$$|G_0| \geq |G_1| \geq |G_2| \geq \ldots \geq |G_q|.$$

Luego, como cada recta en $\mathbb{F}^2$ tiene q puntos, se sigue que

$$\begin{aligned} q + q + |G_2|(q-1) &\geq \sum_{i=0}^{q} |G_i| \\ &\geq |\mathbb{F}^2 \setminus N| \\ &\geq s(q-1) + 2q. \end{aligned}$$

De donde, $2q+|G_2|(q-1) \geq s(q-1)+2q$, y, en consecuencia, $|G_2| \geq s$. Puesto que $|G_0| \geq |G_1| \geq |G_2|$, podemos tomar s rectas paralelas de cada conjunto G_0, G_1 y G_2. Sea W_0, W_1 y W_2, respectivamente, los conjuntos formados por las S rectas paralelas tomadas de G_0, G_1 y G_2. Es decir, $W_i \subseteq G_i$ y $|W_i| = s$ $(i = 0, 1, 2)$.

Cada recta en W_0 tiene $q-1$ puntos en N, y todas las rectas ahí son paralelas (i.e., no tienen puntos en común). Por lo tanto,

$$|N| \geq |N \cap W_o| = s(q-1).$$

Ahora centremos nuestra atención en las rectas de W_1:

(i) Cada recta tiene $q-1$ puntos en N.

(ii) Dos rectas diferentes en W_1 no tienen puntos en común.

(iii) Cada recta en W_0 interseca a todas las rectas de W_1 (según el punto (ii) de la observación 6.0.1).

En resumen, cada recta en W_1 tiene $q-1-s$ puntos en N que no pertenecen a W_0. De aquí se obtiene que

$$|N| \geq s(q-1) + s(q-1-s).$$

Haciendo un análisis análogo para las rectas en W_2, se puede concluir que

$$\begin{aligned} |N| &\geq s(q-1) + s(q-1-s) + s(q-1-2s) \\ &= 3s(q-1-s) \\ &\geq 3\left(\frac{q-3}{3}\right)\left(q-1-\frac{q}{3}\right) \\ &= \frac{2}{3}q^2 - 3q + 3. \end{aligned}$$

Juntando los casos I y II se obtiene el resultado. □

En (FLS10), Chunrong Feng, Liangpan Li y Jian Shen mejoraron la cota del teorema 7.3.1 a $q^2-q^{3/2}-q$, y, recientemente, en (LSW16) se logró un pequeño (pero significativo) avance al establecer la cota $q^2 - q^{3/2} - 1$, que es la mejor que se conoce hasta ahora. Para $n = 3$, se conoce la cota $(0.38 - o(1))\, q^3$ [1]. Por último, para el caso general n se sabe la cota $(1 - o(1))\, q^n$, que marca una clara distinción con las cotas que se conocen para los conjuntos de Kakeya. Las demostraciones tanto para el caso $n = 3$ como para el caso general n se pueden consultar en (LSW16).

[1] $o(1)$ representa una función de q que *tiende* a cero cuando q *tiende a infinito*.

Apéndice A

Medida de Jordan.

Salta a la vista lo árido de este apéndice, pues se trató de ser lo más concreto posible. Aquí se presenta el concepto preciso de *área* que se adopta para este trabajo. Dicho concepto de *área* recibe el nombre de *medida de Jordan*. Cabe destacar que esta no es la única manera de definir el concepto de *área* de manera rigurosa (una definición extremadamente importante, que se utiliza para conjuntos más generales, es la *medida de Lebesgue*), sin embargo, la *medida de Jordan* tiene la ventaja de ser más intuitiva y bastante adecuada para los fines de esta tesis. Para el lector interesado en profundizar en el concepto abstracto de *medida*, se recomienda consultar (Tao11) y (Coh80).

El tratamiento del concepto de *área* que se muestra aquí sigue el texto de Richard Courant (CJ12), por lo que las pruebas que aquí se omitan pueden consultarse ahí. Se hace notar que, aunque nos enfoquemos aquí en el caso del plano, las ideas se pueden generalizar de manera natural a dimensiones superiores con cambios en la terminología: el remplazo del término *área* por *volumen*, *cuadrado* por *cubo* y así sucesivamente.

A.1. Áreas en el plano.

Para definir el *área* –y por tanto llegar a un valor $A(S)$ determinado de modo único– para un conjunto acotado $S \subseteq \mathbb{R}^2$ se usan subdivisiones sucesivas del plano en cuadrados de lado $1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, ...$ (se puede demostrar que cualquier otra manera de dividir el plano en cuadrados o rectángulos conducirá a la misma *área*) por medio de rectas paralelas equidistantes de los ejes coordenados. Los cuadrados congruentes proporcionan un modo sencillo de cubrir el plano sin espacios vacíos o traslapes [1].

Como primer paso, utilizaremos la rejilla asociada al plano coordenado, proporcionada por las rectas $x = 0, \pm 1, \pm 2, \pm 3, ...$ y $y = 0, \pm 1, \pm 2, \pm 3, ...$, la cual divide al plano completo en cuadrados cerrados de lado 1. Denotemos por $A_0^+(S)$ a la suma de las *áreas* de todos los cuadrados que tienen puntos en común con el conjunto S y por $A_0^-(S)$ a la suma de las *áreas* de aquellos cuadrados contenidos completamente en S. A continuación, dividamos cada cuadrado en cuatro cuadrados iguales de lado $\frac{1}{2}$ y denotemos por $A_1^+(S)$ a la suma de las *áreas* de aquellos subcuadrados que tienen puntos en común

[1] Diremos que los conjuntos $A_1, A_2, ..., A_n$ *no se traslapan* si para cada $i \in \{1, 2, ..., n\}$ todo punto interior de A_i es exterior a A_j para toda $j \neq i$.

con S y por $A_1^-(S)$ a la suma de las *áreas* de aquéllos contenidos completamente en S. Como cada cuadrado unitario completamente contenido en S da lugar a cuatro subcuadrados completamente contenidos en S, se tiene que $A_0^-(S) \leq A_1^-(S)$ y, de modo semejante, $A_0^+(S) \geq A_1^+(S)$.

La n-ésima subdivisión del plano, con $n \in \mathbb{N}$, se obtiene de las rectas

$$x = \frac{i}{2^n},\ y = \frac{k}{2^n},$$

donde $i, k \in \mathbb{Z}$. Entonces se divide al plano en los cuadrados cerrados R_{ik}^n de lado 2^{-n} dados por

$$R_{ik}^n = \left\{(x,y) \in \mathbb{R}^2 : \frac{i}{2^n} \leq x \leq \frac{i+1}{2^n},\ \frac{k}{2^n} \leq y \leq \frac{k+1}{2^n}\right\}.$$

La idea es formar aproximaciones desde abajo y desde arriba para el *área* del conjunto S, formando la suma $A_n^-(S)$ de las *áreas* de todos los cuadrados R_{ik}^n que están completamente contenidos en S y la suma $A_n^+(S)$ de las *áreas* de todos los cuadrados R_{ik}^n que tienen puntos en común con S. En este momento, seguramente, el lector ya haya observado que, para cada $n \in \mathbb{N}$, se está suponiendo conocida el *área* de cada cuadrado R_{ik}^n en las definiciones de las sumas A_n^+ y A_n^-. Y es que sí la conocemos, es decir, si queremos que el concepto de *área* que vamos a desarrollar se corresponda con el concepto intuitivo de *área*, se espera recuperar que el *área* de un cuadrado de lado l sea l^2. Puesto que necesitamos partir de algún lugar nuestro tratamiento del concepto de *área*, se definirá el *área* de un cuadrado R_{ik}^n de lado 2^{-n} como 2^{-2n}. Más adelante veremos que la *medida de Jordan* –lo que llamaremos *área*– de cada cuadrado R_{ik}^n es igual a su *área*.

Con lo anterior en cuenta, se tiene [1] que

$$A_n^-(S) = \sum_{\substack{i,k \\ R_{ik}^n \subset S}} 2^{-2n},\ A_n^+(S) = \sum_{\substack{i,k \\ R_{ik}^n \cap S \neq \emptyset}} 2^{-2n}.$$

(Ver figura A.1.)

Es claro, por la definición, que

$$0 \leq A_n^-(S) \leq A_n^+(S).$$

Ahora bien, al pasar de la n-ésima subdivisión hacia la $(n+1)$-ésima, se divide cada cuadrado R_{ik}^n en cuatro cuadrados R_{rs}^{n+1}. Si el cuadrado R_{ik}^n está contenido en el conjunto S, deben estarlo sus partes R_{rs}^{n+1}. Si, además, una parte R_{rs}^{n+1} contiene un punto de S, entonces lo mismo se cumple para el cuadrado completo R_{ik}^n. Se concluye [2] que sumas sucesivas satisfacen la desigualdad

$$A_n^-(S) \leq A_{n+1}^-(S) \leq A_{n+1}^+(S) \leq A_n^+(S). \tag{A.1}$$

[1] Si ningún cuadrado R_{ik}^n está completamente contenido en S, se hace $A_n^-(S) = 0$.

[2] Aquí se está usando implícitamente el hecho que la suma de las áreas de los cuatro cuadrados R_{rs}^{n+1} que constituyen a R_{ik}^n es igual al área de R_{ik}^n, lo cual, en este contexto se deduce de la identidad $4 \cdot 2^{-2(n+1)} = 2^{-2n}$.

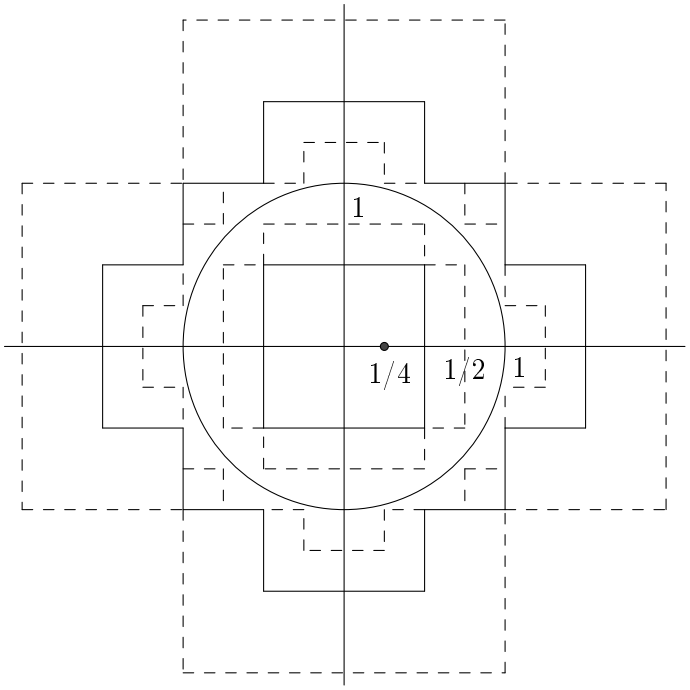


Figura A.1: Aproximaciones interior y exterior para el área del disco unitario $D : x^2+y^2 \leq 1$, para $n = 0, 1, 2$, donde $A_0^-(D) = 0, A_1^-(D) = 1, A_2^-(D) = 2$ y $A_2^+(D) = 4\frac{1}{4}, A_1^+(D) = 6, A_0^+(D) = 12$.

Por lo tanto, de A.1 se sigue que la sucesión de sumas $\{A_n^-(S)\}_{n\in\mathbb{N}}$ es no decreciente, con la cota superior A_1^+; así, converge hacia un límite, digamos,

$$A^-(S) = \lim_{n\to\infty} A_n^-(S).$$

Análogamente, la sucesión de sumas $\{A_n^+(S)\}_{n\in\mathbb{N}}$, es no creciente, con la cota inferior A_1^- y, por tanto, convergente:

$$A^+(S) = \lim_{n\to\infty} A_n^+(S).$$

Por A.1 se tiene que, para cada $n \in \mathbb{N}$

$$0 \leq A_n^-(S) \leq A^-(S) \leq A^+(S) \leq A_n^+(S).$$

A los números $A^-(S)$ y $A^+(S)$ se les denomina, respectivamente, ***área interior*** y ***área exterior*** del conjunto S.

Obsérvese que el área interior $A^-(S) = 0$ si, y sólo si, S no tiene puntos interiores, ya que un conjunto sin puntos interiores no contiene cuadrado R_{ik}^n alguno, de modo que $A_n^-(S) = 0$ para todo $n \in \mathbb{N}$ y, en consecuencia, $A^-(S) = 0$. Un conjunto con puntos interiores contiene algún cuadrado R_{ik}^n para n suficientemente grande, de modo que $A_n(S) > 0$ para n grande y, así, $A^-(S) > 0$.

A.2. Conjuntos medibles según Jordan.

Definición A.2.1. *Sea $S \subseteq \mathbb{R}^2$ un conjunto acotado. Se dice que S es* **medible según Jordan** [1] *si el área interior y el área exterior de S coinciden.*

Se denotará al valor común por A y se le dará el nombre de **área** *o* **medida de Jordan** *de S:*

$$A^-(S) = A^+(S) = A(S).$$

Al comenzar, definimos el *área* de los cuadrados R_{ik}^n de las definiciones de la manera usual en que se está acostumbrado a calcular el área de un cuadrado. Veamos que, de hecho, nuestra nueva noción de área, es decir, la medida de Jordan, y la definida para dichos cuadrados coincide. Esto es, cada cuadrado R_{ik}^n tiene área –medida de Jordan– 2^{-2n} en el sentido de la definición general pues, si $S = R_{ik}^n$ y $m > n$, $m, n \in \mathbb{N}$, entonces

$$A_m^-(S) = \left(2^{m-n}\right)^2 2^{-2m} = 2^{-2n}, \tag{A.2}$$

y

$$A_m^+(S) = \left[\left(2^{m-n}\right)^2 + 4\left(2^{m-n}\right) + 4\right] 2^{-2m} = 2^{-2n} + 2^{2-m-n} + 2^{2-2m}. \tag{A.3}$$

De las ecuaciones A.2 y A.3, se sigue que

$$\lim_{m\to\infty} A_m^-(S) = \lim_{m\to\infty} A_m^+(S) = 2^{-2n}.$$

Por lo tanto, $A(S) = 2^{-2n}$.

De forma más general, cualquier rectángulo S con lados paralelos a los ejes coordenados:

$$S = \left\{(x,y) \in \mathbb{R}^2 :\ a \le x \le b,\ c \le y \le d\right\}$$

tiene área igual a $(b-a)(d-c)$, como es de esperar con base en la geometría clásica. En efecto, dado un número entero positivo n, se pueden hallar números enteros α, β, γ y δ tales que

$$\alpha 2^{-n} < a \le (\alpha+1)2^{-n},\ \beta 2^{-n} \le b < (\beta+1)2^{-n}$$
$$\gamma 2^{-n} < c \le (\gamma+1)2^{-n},\ \delta 2^{-n} \le d < (\delta+1)2^{-n}.$$

Así,

$$A_n^-(S) = (\beta-\alpha-1)(\delta-\gamma-1)2^{-2n} \ge \left(b-a-2^{1-n}\right)\left(d-c-2^{1-n}\right)$$

[1] En lugar de usar la frase «el conjunto S es medible según Jordan, simplemente se dirá «S tiene área».

y

$$A_n^+(S) = (\beta - \alpha + 1)(\delta - \gamma + 1)\, 2^{-2n} \leq \left(b - a + 2^{1-n}\right)\left(d - c + 2^{1-n}\right).$$

Luego, cuando $n \longrightarrow \infty$,

$$A(S) = \lim_{n\to\infty} A_n^-(S) = \lim_{n\to\infty} A_n^+(S) = (b-a)(d-c).$$

El siguiente resultado establece condiciones necesarias y suficientes para que un conjunto acotado del plano tenga área.

Teorema A.2.2. *Sea $S \subseteq \mathbb{R}^2$ un conjunto acotado. Entonces S tiene área si, y sólo si, la frontera de S, ∂S, tiene área cero.*

Ejemplo A.2.3. Un ejemplo de un conjunto que no tenga área –en el sentido de la medida de Jordan– es el conjunto $S = [0,1]^2 \cap \mathbb{Q}^2$, es decir, el conjunto de los puntos (x, y), donde x y y son números racionales entre 0 y 1. En efecto, como $\partial S = [0,1]^2$, entonces $A(\partial S) = 1$. A partir del teorema A.2.2 se concluye que el conjunto S no es medible según Jordan.

Para concluir este apéndice, se resumen en el siguiente teorema las propiedades elementales de la medida de Jordan.

Teorema A.2.4.

(i) Cualquier subconjunto de un conjunto de área cero, tiene área cero.

(ii) La unión de cualquier número finito de conjuntos de área cero tiene área cero. En particular, cualquier conjunto finito de puntos tiene área cero.

(iii) Sea $S \subseteq \mathbb{R}^2$. Entonces S tiene área cero si para cada $\varepsilon > 0$ existen subconjuntos acotados $S_1, S_2, ..., S_n$ de $\mathbb{R}^2$ tales que

$$S \subseteq \bigcup_{i=1}^{n} S_i \quad y \quad \sum_{i=1}^{n} A^+(S_i) < \varepsilon.$$

(iv) Supóngase que la frontera de un conjunto acotado $S \subseteq \mathbb{R}^2$ está contenida en un número finito de arcos, cada uno de los cuales está dado por una ecuación $y = f(x)$ o $x = g(y)$, con la función f o g definida y continua en un intervalo cerrado acotado $I \subseteq \mathbb{R}$. Entonces S tiene área.

(v) Sean S y T dos subconjuntos medibles según Jordan de $\mathbb{R}^2$.

(a) $S \cup T$ y $S \cap T$ son medibles según Jordan. Además, para el caso de la unión, se cumple que

$$A(S \cup T) \leq A(S) + A(T).$$

Más aún, si S y T no se traslapan, entonces

$$A(S \cup T) = A(S) + A(T).$$

(b) Si, además, $S \subseteq T$*, entonces* $T \setminus S$ *tiene área y*

$$A(T \setminus S) = A(T) - A(S).$$

Apéndice B

Campos finitos.

Al igual que en el apéndice anterior, el carácter árido de este apéndice se debe a la brevedad. Aquí se presentan algunos resultados básicos sobre *campos finitos* con el objetivo de brindar al lector cierta familiaridad con este tipo de campos. Se tomó como base el texto de R. Lidl y H. Niederreiter (LN97). Para las pruebas que aquí se omiten, el lector puede consular el mismo (LN97) así como (Fra03), (Rot10) y (Nic12).

B.1. Estructuras algebraicas.

B.1.1. Grupos.

En el conjunto de los enteros $\mathbb{Z}$ las operaciones de suma y multiplicación son bien conocidas. Se puede generalizar la noción de *operación* a conjuntos arbitrarios.

Definición B.1.1. *Sea A un conjunto, y denotemos por $A \times A$ al conjunto de todas las parejas (a_1, a_2), con $a_i \in A$ $(i = 1, 2)$. Una función $\bullet : A \times A \longrightarrow A$ se llamará una* **operación binaria en el conjunto** *A.*

Por supuesto, bajo esta definición, estamos asegurando que (i) se asigna exactamente un elemento a cada pareja posible de elementos en el conjunto A, y (ii) la imagen de $(a_1, a_2) \in A \times A$ esté en en conjunto A; a esto último se le conoce como *la propiedad de cerradura* de la operación $\bullet$. Por convención, a la imagen del elemento $(a_1, a_2) \in A \times A$ bajo la operación $\bullet$ se le suele denotar por $a_1 \bullet a_2$.

Ejemplo B.1.2.

1. En el conjunto de los enteros positivos $\mathbb{Z}^+$, para cada $a, b \in \mathbb{Z}^+$, defínase la operación $\bullet$ por $a \bullet b$ que es igual al mínimo entre a y b, o al valor común si $a = b$. Así, $5 \bullet 23 = 5; 34 \bullet 10 = 10$ y $4 \bullet 4 = 4$.
2. No hay una operación en $\mathbb{Q}$ tal que $(a, b) \longmapsto \dfrac{a}{b}$.
3. Sea A el conjunto de todas las funciones de $\mathbb{R}$ en $\mathbb{R}$. Defínase $\bullet$ como la suma usual de dos funciones, esto es, $f \bullet g = h$ donde $h(x) = f(x) + g(x)$, para $f, g \in A$ y $x \in \mathbb{R}$. Esta manera de definir $\bullet$ satisface la definición B.1.1 y nos da una operación binaria en A.

Con el concepto de operación binaria en un conjunto, podemos introducir las siguientes *estructuras algebraicas*[1], que se suponen familiares al lector.

Definición B.1.3. *Un* grupo $(G, \bullet)$ *es un conjunto* G*, junto con una operación binaria* $\bullet$ *definida en* G *tal que se cumplen las siguientes propiedades:*

*G*1. $\bullet$ *es* asociativa*, es decir, para cualesquiera* $a, b, c \in G$*,*

$$a \bullet (b \bullet c) = (a \bullet b) \bullet c.$$

*G*2. *Existe un elemento* $e \in G$ *tal que,*

$$a \bullet e = e \bullet a = a$$

para todo $a \in G$*. El elemento* $e \in G$ *se llama* ***elemento identidad*** *para* $\bullet$ *en* G*.*

*G*3. *Para cada* $a \in G$ *existe un elemento* $a^{-1} \in G$ *tal que*

$$a \bullet a^{-1} = a^{-1} \bullet a = e.$$

Al elemento $a^{-1} \in G$ *se le llama el inverso de* $a \in G$ *respecto a* $\bullet$*.*

Si además el grupo satisface

*G*4. *Para cualesquiera* $a, b \in G$

$$a \bullet b = b \bullet a,$$

el grupo se llamará ***abeliano*** *o* ***conmutativo****.*

Por brevedad, escribiremos simplemente G para denotar al grupo $(G, \bullet)$. Como es de esperar, tanto el elemento identidad $e \in G$ como el inverso a^{-1} de un elemento $a \in G$ son únicos. Más aún, $(a \bullet b)^{-1} = b^{-1} \bullet a^{-1}$, para $a, b \in G$. También para facilitar la notación se escribirá ab, en lugar de $a \bullet b$, para denotar el producto de los elementos $a, b \in G$. En algunas ocasiones es útil escribir, respectivamente, $a + b$ y $-a$ en vez de $a \bullet b$ y a^{-1}, y, en este caso, se suele decir que se está tomando un *grupo aditivo* (pero, preferentemente, se apelará a esta notación cuando se traten grupos abelianos), en lugar de uno *multiplicativo*).

La propiedad asociativa posibilita escribir expresiones de la forma $a_1 a_2 ... a_n$, con $a_i \in G$ $(1 \leq i \leq n)$, sin caer en ambigüedades pues, sin importar cómo se coloquen los paréntesis, dicha expresión siempre representa un mismo elemento en G. Para denotar que se ha aplicado n veces, $n \in \mathbb{N}$, la operación $\bullet$ a un elemento $a \in G$ con él mismo, se escribirá

$$a^n = aa...a \quad \text{(n factores a)}.$$

[1]Por *estructura algebraica* nos referimos a un conjunto con una o varias operaciones binarias definidas en él.

Cuando G es un grupo multiplicativo, la expresión a^n se llamará la n–ésima potencia de $a \in G$. Por otro lado, si G se toma como grupo aditivo, se escribirá

$$na = a + a + ... + a \quad \text{(n sumandos a)}.$$

Con la notación descrita antes, se tiene las siguientes propiedades elementales:

Proposición B.1.4. *Sea G un grupo.*

(i) Si G es un grupo aditivo, entonces para cualquier $a \in G$ se cumplen las siguientes propiedades:

1. $(-n)a = n(-a)$.
2. $na + ma = (m+n)a$.
3. $m(na) = (mn)a$.

(ii) Si G es un grupo multiplicativo, todo $a \in G$ satisface que

1. $a^{-n} = \left(a^{-1}\right)^n$.
2. $a^n a^m = a^{n+m}$.
3. $(a^n)^m = a^{nm}$.

Para finalizar las convenciones, adoptaremos también que, para $n = 0 \in \mathbb{Z}$, $a^0 = e$, en la notación multiplicativa, y que $0a = 0$ (aquí el 0 del lado derecho representa al elemento identidad del grupo), en la notación aditiva.

Ejemplo B.1.5.

1. *El conjunto $\mathbb{Z}$ de los números enteros forma un grupo con la operación suma (de hecho, es un grupo abeliano).*
2. *El conjunto $GL_n(\mathbb{R})$ de matrices invertibles de $n \times n$ con entradas reales, junto con la multiplicación de matrices forma un grupo. (Para $n > 1$ este grupo no es abeliano.)*

Definición B.1.6. *Un grupo multiplicativo G es* **cíclico** *si existe un elemento $a \in G$ tal que, para cada $b \in G$, $b = a^i$, para algún número entero i. El elemento $a \in G$ con esa característica se llama* **generador** *del grupo cíclico, y escribimos $G = \langle a \rangle$.*

Se sigue de la definición que todo grupo cíclico es conmutativo. Además, nótese que puede haber más de un elemento generador del grupo, por ejemplo, en el grupo aditivo $\mathbb{Z}$ los generadores son 1 y -1.

Definición B.1.7. *Dado un conjunto A, un subconjunto R de $A \times A$ es una* **relación de equivalencia** *en A si éste satisface lo siguiente:*

- *Para cada $a \in A$, $(a, a) \in R$ (reflexividad)*

- *Si* $(a, b) \in R$*, entonces* $(b, a) \in R$ *(simetría)*
- *Si* $(a, b) \in R$ *y* $(b, c) \in R$*, entonces* $(a, c) \in R$ *(transitividad)*

Recordemos que una relación de equivalencia R en un conjunto A induce una *partición* en A –esto es, A se puede expresar como la unión de subconjuntos no vacíos y mutuamente ajenos. Dado un elemento fijo $a \in A$, podemos tomar el conjunto de todos los elementos equivalentes a él bajo la relación R, para obtener la clase de equivalencia del elemento $a \in A$, denotada por

$$[a] = \{b \in A : (a, b) \in R\}.$$

La colección de todas las diferentes clases de equivalencia forman una partición del conjunto A.

Definición B.1.8. *Sean* $a, b \in \mathbb{Z}$ *y* n *un entero positivo. Decimos que* a *es* **congruente** *con* b **módulo** n*, lo escribimos como* $a \equiv b \pmod{n}$*, si* $a = b + kn$*, para algún entero* k *(i.e., la diferencia* $a - b$ *es un múltiplo de* n*).*

Como se sabe, la relación de «congruencia módulo n» es una relación de equivalencia en el conjunto $\mathbb{Z}$ de los números enteros. Así, podemos tomar la partición del conjunto $\mathbb{Z}$ inducida por las clases de equivalencia, las cuales son:

$$\begin{aligned}
[0] &= \{..., -2n, -n, 0, n, 2n, ...\}, \\
[1] &= \{..., -2n+1, -n+1, 1, n+1, 2n+1, ...\}, \\
&\vdots \\
[n-1] &= \{..., -n-1, -1, n-1, 2n-1, 3n-1, ...\}.
\end{aligned}$$

Podemos definir en el conjunto $\{[0], [1], ..., [n-1]\}$ de clases de equivalencia la operación $[a] + [b] = [a + b]$, donde a y b son, respectivamente, elementos arbitrarios en los conjuntos $[a]$ y $[b]$, y $a + b$ es la suma usual de números enteros. Con esto en cuenta tenemos el siguiente teorema.

Teorema B.1.9. *Sea* $n \in \mathbb{N}$*. El conjunto* $\{[0], [1], ..., [n-1]\}$ *de clases de equivalencia módulo* n *junto con la operación* $[a]+[b] = [a+b]$ *forma un grupo. (El elemento identidad es* $[0]$ *y el inverso de* $[a]$ *es* $[-a]$*.)*

A este grupo se le llama ***grupo de enteros módulo*** n*, y se denota por* $\mathbb{Z}_n$*.*

$\mathbb{Z}_n$ es, en realidad, un grupo cíclico con el elemento [1] como generador. Más aún, este grupo tiene *orden* n según la siguiente definición.

Definición B.1.10. *Un grupo se dice* **finito (infinito)** *si tiene un número finito (infinito) de elementos. El número de elementos en un grupo finito se llama el* **orden** *del grupo.*

Notación B.1.11. *El orden de un grupo finito* G *se denotará por* $|G|$*.*

Definición B.1.12. *Un subconjunto H de un grupo G es un* **subgrupo** *de G si él mismo es un grupo con respecto a la operación de G. (Los subgrupos de G diferentes de los subgrupos* triviales *G y $\{e\}$ se llaman subgrupos* no triviales.*)*

Definición B.1.13. *El subgrupo de G que consiste de todas las potencias de un elemento $a \in G$, se llama el subgrupo* **generado** *por el elemento a y se denota por $\langle a \rangle$. Este grupo es necesariamente cíclico. Si $\langle a \rangle$ es finito, entonces se dirá que el* **orden** *del elemento a es el entero más pequeño i tal que $a^i = e$. En caso contrario se dirá que a es de* **orden infinito**.

Para $n \in \mathbb{N}$, el subgrupo $\langle n \rangle$ del grupo aditivo $\mathbb{Z}$ de los números enteros está estrechamente ligado con la relación de congruencia módulo n, pues $a \equiv b \ (mod\ n)$ si, y sólo si, $a - b \in \langle n \rangle$. Así, el subgrupo $\langle n \rangle$ define un relación de equivalencia. Esta situación se generaliza con el siguiente teorema.

Teorema B.1.14. *Sea H un subgrupo de un grupo G. La relación R_H en G definida por $(a, b) \in R_H$ si, y sólo si, $a = bh$ para algún $h \in H$ (y cuando el grupo es aditivo, $a = b + h$ para algún $h \in H$), es una relación de equivalencia.*

Las clases de equivalencia bajo la relación R_H se llaman ***clases laterales izquierdas*** de G módulo H y se denotan por $aH = \{ah : h \in H\}$ (o $a + H = \{a + h : h \in H\}$ cuando el grupo es aditivo), donde a es un elemento fijo de G. De manera análoga se pueden obtener las ***clases laterales derechas*** de G módulo H, las cuales son de la forma $Ha = \{ha : h \in H\}$. Claramente, si el grupo es abeliano las clases laterales izquierdas y derechas módulo H coinciden.

Ejemplo B.1.15. *Sea $(\mathbb{Z}_{12}, +)$ y H el subgrupo $\{[0], [3], [6], [9]\}$. Las distintas clases laterales izquierdas módulo H están dadas por:*

$$\begin{aligned} [0] + H &= \{[0], [3], [6], [9]\}, \\ [1] + H &= \{[1], [4], [7], [10]\}, \\ [2] + H &= \{[2], [5], [8], [11]\}. \end{aligned}$$

Teorema B.1.16. *Si H es un subgrupo finito de un grupo G, toda clase lateral (izquierda o derecha) de G módulo H tiene el mismo número de elementos que H.*

Definición B.1.17. *El número de clases laterales se llama el* **índice** *de H en G.*

Teorema B.1.18. *Sea G un grupo finito y sea H un subgrupo de G. El orden de G es igual al orden de H por el índice de H en G. En particular, tanto el orden de H como el orden de cualquier elemento $a \in G$ dividen al orden de G.*

El siguiente teorema resume las propiedades principales de los grupos cíclicos. En dicho resultado se apela a la función ϕ de Euler, la cual está definida como sigue: para cada $n \in \mathbb{N}$,

$$\phi(n) = |\{k \in \mathbb{N} : 1 \leq k \leq n \text{ y } k \text{ es primo relativo con } n\}|.$$

Si $n \in \mathbb{N}$ tiene descomposición en números primos de la forma $p_1^{a_1} p_2^{a_2} ... a_r^{a_r}$, entonces

$$\phi(n) = n\left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right)...\left(1 - \frac{1}{p_r}\right).$$

Por ejemplo, $\phi(30) = 2 \cdot 3 \cdot 5 \cdot \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{4}{5} = 8$.

Teorema B.1.19.

(a) Todo subgrupo de un grupo cíclico es también un grupo cíclico.

En los siguientes inciso se habla de un grupo cíclico $\langle a \rangle$ de orden m.

(b) En un grupo cíclico finito $\langle a \rangle$ de orden m el elemento a^k genera un subgrupo de orden $\frac{m}{mcd(k,m)}$, donde $mcd(k,m)$ denota al máximo común divisor *de k y m*

(c) Para cualquier divisor positivo d de m, $\langle a \rangle$ contiene exactamente un subgrupo de índice d. Para cualquier divisor positivo f de m, $\langle a \rangle$ contiene precisamente un subgrupo de orden f.

(d) Sea k un divisor positivo de m. Entonces, $\langle a \rangle$ contiene $\phi(k)$ elementos de orden k.

(e) El grupo $\langle a \rangle$ contiene $\phi(m)$ generadores, a saber, las potencias a^r para las cuales $mcd(r,m) = 1$.

Definición B.1.20. *Sean G y H dos grupos. Una función $f : G \longrightarrow H$ es un* **homomorfismo** *de G en H si f preserva la operación de G. Esto es, si $\bullet$ y $*$ son las operaciones de G y H respectivamente, entonces, f preserva la operación de G si para cualesquiera $a, b \in G$, $f(a \bullet b) = f(a) * f(b)$. Si, además, es una función biyectiva el homomorfismo se llama* **isomorfismo** *y se dice que G y H son isomorfos, lo cual se denota como $G \cong H$. Por último, un isomorfismo de G en sí mismo se llama* **automorfismo**.

Observación B.1.21. *Si $f : G \longrightarrow H$ es un homomorfismo de un grupo G en un grupo H, y e es el elemento identidad en G, entonces del hecho que $ee = e$, se sigue que $f(e)f(e) = f(e)$, de donde $e' = f(e)$, con e' el elemento identidad en H. Además, como*

$$e = f(e) = f\left(aa^{-1}\right) = f(a)f\left(a^{-1}\right)$$

se obtiene que $f\left(a^{-1}\right) = (f(a)^{-1}$, para cada $a \in G$. de la ecuación $aa^{-1} = e$ se obtiene que $f\left(a^{-1}\right) = (f(a))^{-1}$, para todo $a \in G$.

Definición B.1.22. *El* **núcleo** *de un homomorfismo $f : G \longrightarrow H$ de un grupo G en un grupo H es el conjunto (de hecho, es un subgrupo de G)*

$$ker\,(f) = \{a \in G : f(a) = e\},$$

donde e denota el elemento identidad de H.

La **imagen** *de f es el conjunto (es, en realidad, un subgrupo de H)*

$$im\,(f) = \{f(a) : a \in G\}\,.$$

Ejemplo B.1.23. Consideremos la función f del grupo aditivo $\mathbb{Z}$ de los números enteros al grupo $\mathbb{Z}_n$ definida como $f(a) = [a]$, para cada $a \in \mathbb{Z}$. Entonces,

$$f(a+b) = [a+b] = [a] + [b] = f(a) + f(b),$$

para $a, b \in \mathbb{Z}$. Luego, f es un homomorfismo. Más aún, $ker\,(f)$ está formado por todos los elementos $a \in \mathbb{Z}$ tal que $[a] = [0]$. Ya que lo anterior se cumple para los múltiplos a de n, se sigue que $ker\,(f) = \langle n \rangle$, el subgrupo de $\mathbb{Z}$ generado por n.

Definición B.1.24. *Un subgrupo H de un grupo G se llama subgrupo* **normal** *de G si $aha^{-1} \in H$ para todo $a \in G$ y $h \in H$.*

Nótese que todo subgrupo de un grupo abeliano es normal. El siguiente teorema brinda algunas caracterizaciones alternativas de los subgrupos normales.

Teorema B.1.25. *Sea H un subgrupo de un grupo G.*

(i) Dado $a \in G$, definamos el conjunto aHa^{-1} como $aHa^{-1} = \left\{aha^{-1} : h \in H\right\}$. Entonces, H es un subgrupo normal de G si, y sólo si, $H = aHa^{-1}$, para todo $a \in G$.

(ii) H es un subgrupo normal de G si, y sólo si, $aH = Ha$, para cada $a \in G$. En otras palabras, H es un subgrupo normal de G si, y sólo si, las clases laterales izquierdas y derechas de G módulo H coinciden.

El siguiente teorema asegura que el conjunto de clases laterales (izquierdas) forman un grupo.

Teorema B.1.26. *Si H es un subgrupo normal de un grupo G, entonces el conjunto de clases laterales (izquierdas) de G módulo H forma un grupo con la operación $(aH)(bH) = (ab)H$.*

Definición B.1.27. *Dado un subgrupo normal H de un grupo G, el grupo formado por las clases laterales (izquierdas) bajo la operación descrita en el teorema B.1.26 es llamado* **grupo cociente** *(o* **grupo factor***) de G módulo H, y se denota por G/H.*

Si G/H es finito, su orden es igual al índice de H en G. Por lo tanto, según el teorema B.1.18, para un grupo finito G, se cumple que

$$|G/H| = \frac{|G|}{|H|}.$$

El siguiente teorema es un resultado fundamental en la *teoría de grupos*.

Teorema B.1.28 (Primer teorema de isomorfismo)**.** *Sea $f : G \longrightarrow H$ un homomorfismo de grupos. Entonces, el núcleo de f es un subgrupo normal de G y*

$$G/ker\,(f) \cong im\,(f)$$

(bajo el isomorfismo $aker\,(f) \longmapsto f(a)$, con $a \in G$).

Recíprocamente, si H es un subgrupo normal de G, la función $\psi : G \longrightarrow G/H$ definida por $\psi(a) = aH$, para $a \in G$, es un homomorfismo de G en G/H tal que $ker\,(\psi) = H$.

Ejemplo B.1.29. Con la notación del teorema anterior, tomemos $G = \mathbb{Z}$, $H = \mathbb{Z}_n$ y el homomorfismo $f : \mathbb{Z} \longrightarrow \mathbb{Z}_n$ definido por $f(a) = [a]$. Entonces, $ker\,(f) = \langle n \rangle$ y $im\,(f) = \mathbb{Z}_n$, de donde, por el primer teorema de isomorfismo, $\langle n \rangle$ es un subgrupo normal de $\mathbb{Z}$ y $\mathbb{Z}/\langle n \rangle \cong \mathbb{Z}_n$.

B.1.2. Anillos y campos.

En la mayoría de los conjuntos numéricos usados en la aritmética elemental hay dos operaciones binarias distintas: suma y multiplicación. Ejemplos de lo anterior son los números enteros, racionales y reales. A continuación se introduce el concepto de *anillo*; estructura algebraica bien conocida que comparte muchas de las propiedades de los conjuntos numéricos mencionados.

Definición B.1.30. *Un **anillo** $(R, +, \cdot)$ es un conjunto R, junto con dos operaciones binarias $+$ y $\cdot$, que llamaremos suma y multiplicación, definidas en R tales que se satisfacen las siguientes propiedades:*

*R*1. *$(R, +)$ es un grupo abeliano con elemento identidad $0 \in R$.*

*R*2. *La multiplicación es asociativa.*

*R*3. *Para cualesquiera $a, b, c \in R$ se cumple*

*(i) la **ley distributiva izquierda**, es decir*

$$a \cdot (b + c) = a \cdot b + a \cdot c.$$

*(ii) la **ley distributiva derecha**, es decir*

$$(a + b) \cdot c = a \cdot c + b \cdot c.$$

Al igual que en el caso de los grupos, denotaremos por R al anillo $(R, +, \cdot)$. Además, usaremos el símbolo 0 (llamado *elemento cero*, o simplemente *cero*) para referirnos al elemento identidad del grupo abeliano R con respecto a la suma, y al inverso (aditivo) del elemento $a \in R$ se denotará por $-a$; también abreviaremos $a + (-b)$ por $a - b$. En vez de escribir $a \cdot b$, usualmente escribiremos ab para el producto de a y b. Como consecuencia de la definición de anillo tenemos la siguiente propiedad: $a0 = 0a = 0$, para cada $a \in R$. Más aún, de lo anterior se deduce que $(-a)b = a(-b) = -ab$, para cualesquiera $a, b \in R$.

La definición que sigue contempla anillos con algunas propiedades específicas.

Definición B.1.31. *Sea R un anillo.*

(i) *R es un* **anillo con uno** *o* **unitario** *si existe un elemento $1 \in R$ tal que $1 \neq 0$ y $1 \cdot x = x \cdot 1 = x$ para cada $x \in R$.*

(ii) *R es un* **anillo conmutativo** *si la operación $\cdot$ es conmutativa, es decir, si para cada $x, y \in R$, se cumple que $x \cdot y = y \cdot x$.*

(iii) *R es un* **dominio entero** *si es un anillo conmutativo y con identidad $1 \neq 0$ para el cual, si $ab = 0$, entonces, $a = 0$ o $b = 0$, con $a, b \in R$.*

(iv) *R es un* **anillo de división** *si el conjunto de los elementos en R diferentes de $0 \in R$ forman un grupo con la operación $\cdot$.*

(iv) *Un anillo de división conmutativo es un* **campo**.

Ejemplo B.1.32.

1. Sea $(R, +)$ cualquier grupo abeliano. Defina $a \cdot b = 0$ para todo $a, b \in R$. Entonces R es un anillo.
2. Los números enteros forman un dominio entero, pero no un campo.
3. Los números enteros pares forman un anillo conmutativo sin uno.
4. Las funciones de $\mathbb{R}$ en $\mathbb{R}$ forman un anillo conmutativo con uno bajo las definiciones de $f + g$ y $f \cdot g$ dados por $(f + g)(x) = f(x) + g(x)$ y $(f \cdot g)(x) = f(x) \cdot g(x)$ para cada $x \in \mathbb{R}$.
5. Las matrices de 2×2 con entradas en los números reales forman un anillo no conmutativo con uno respecto a las operaciones de suma y multiplicación de matrices.

Puesto que nuestro estudio está enfocado en campos, haremos algunas precisiones sobre la definición de campo. En primer lugar, un campo es un conjunto $\mathbb{F}$ con dos operaciones binarias, llamadas suma y multiplicación, y con dos elementos distinguidos 0 y 1, con $1 \neq 0$. Más aún, $\mathbb{F}$ es un grupo abeliano con respecto a la suma, teniendo como elemento identidad al $0 \in \mathbb{F}$. Los elementos diferentes de cero forman un grupo abeliano con respecto a la multiplicación (a este grupo lo llamaremos *grupo multiplicativo*, y se denotará por $\mathbb{F}^*$) con elemento identidad $1 \in \mathbb{F}$ (al que llamaremos *uno*). Por último, las operación de suma y multiplicación están relacionadas por las leyes de distributividad. El elemento $0 \in \mathbb{F}$ es llamado *elemento cero* (o simplemente *cero*) y el elemento $1 \in \mathbb{F}$ será llamado *identidad multiplicativa* o, simplemente, *uno*.

La propiedad *(iii)* que aparece en la definición B.1.31 –a saber, si $ab = 0$, entonces $a = 0$ o $b = 0$– se expresa diciendo que, en el anillo R, no hay *divisores de cero distintos de* 0. En particular, un campo no tiene divisores propios de cero pues, si $ab = 0$ y $a \neq 0$, multiplicando por a^{-1} se obtiene que $b = 0$.

Como vimos antes un campo es, en particular, un dominio entero. El recíproco en general no es cierto (ver punto 2 en el ejemplo B.1.32) salvo que el dominio entero sea finito.

Teorema B.1.33. *Todo dominio entero finito es un campo.*

Observación B.1.34. *Sea R un dominio entero finito. Para demostrar que R es un campo necesitamos probar que R es un anillo de división conmutativo. R ya es conmutativo por ser un dominio entero, así que sólo resta probar que es un anillo de división, es decir, que los elementos diferentes de 0 forman un grupo bajo la operación $\cdot$. La asociatividad y la existencia del elemento identidad $1 \in R$ se garantiza por ser R dominio entero. Así, sólo resta probar que cada elemento diferente de cero tiene inverso.*

Demostración. Con la notación de la observación anterior, supongamos que los elementos de R son $a_1, a_2, ..., a_n$. Para cada $a \in R$ fijo, con $a \neq 0$, consideremos los productos $aa_1, aa_2, ..., aa_n$. Nótese que $aa_i \neq aa_j$ para cada i, j, con $i \neq j$. En efecto, si $aa_i = aa_j$ para algún $i \neq j$, entonces $a(a_i - a_j) = 0$, de donde, $a_i = a_j$ pues $a \neq 0$. Así que $a_i = a_j$, lo cual es una contradicción. Ahora, como hay n de dichos productos, cada elemento de R es de la forma aa_i para algún i $(i = 1, 2, ..., n)$. En particular, $1 = aa_i$, $1 \leq i \leq n$. Puesto que R es conmutativo, también se cumple que $1 = a_i a$, por lo cual a_i es el inverso (multiplicativo) de a. Por lo tanto, los elementos diferentes de 0 de R forman un grupo, y, en consecuencia, R es un campo. □

Definición B.1.35. *Un subconjunto S de un anillo R es un* **subanillo** *de R si éste es cerrado bajo las operaciones $+$ y $\cdot$, y, además, S es un anillo bajo dichas operaciones.*

Definición B.1.36. *Un subconjunto J de un anillo R es llamado un* **ideal** *si éste es un subanillo de R y, para cada $a \in J$ y $r \in R$, se tiene que $ar \in J$ y $ra \in J$.*

Definición B.1.37. *Un dominio entero en el que todo ideal es principal se llama* **dominio de ideales principales**.

Ejemplo B.1.38.

1. *Sea $R = \mathbb{Q}$, el conjunto de los números racionales. Entonces el conjunto de los números enteros $\mathbb{Z}$ es un subanillo de $\mathbb{Q}$ pero no es un ideal ya que, por ejemplo, $1 \in \mathbb{Z}$ y $\frac{1}{4} \in \mathbb{Q}$, pero $\frac{1}{4} \cdot 1 \notin \mathbb{Z}$.*
2. *Sea R un anillo conmutativo y $a \in R$. Entonces el conjunto $J = \{ra : r \in R\}$ es un ideal.*

Definición B.1.39. *Sea R un anillo conmutativo. Entonces el ideal más pequeño (en el sentido de la contención) que contiene al elemento $a \in \mathbb{R}$ es $(a) = \{ra : r \in R\}$. Se dice, además, que (a) es el ideal* **principal** *generado por $a \in R$.*

Por definición, los ideales son subgrupos normales del grupo aditivo de un anillo, por lo cual, un ideal J de un anillo R genera una partición en clases laterales ajenas, llamadas *clases residuales* módulo J. La clase residual de un elemento $a \in R$ se denotará por $[a] = a + J$, ya que esta última consiste de todos los elementos en R de la forma $a + j$ para algún $j \in J$. En completa analogía con la relación de congruencia para los enteros, diremos que $a, b \in R$ son *congruentes* módulo J, lo cual escribiremos $a \equiv b \ (mod\ J)$,

si $a - b \in J$ (en otras palabras, si éstos están en la misma clase residual módulo J). Se pueden verificar las siguientes propiedades que cumple la relación de congruencia definida antes:

(a) $a \equiv b \ (mod\ J)$ implica $a + r \equiv b + r \ (mod\ J)$, $ra \equiv rb \ (mod\ J)$, para cualquier $r \in R$, y $na \equiv nb \ (mod\ J)$ para cualquier $n \in \mathbb{Z}$.

(b) Si, además, $r \equiv s \ (mod\ J)$ entonces $a + r \equiv b + s \ (mod\ J)$ y $ar \equiv bs \ (mod\ J)$.

Con esto en cuenta, se puede concluir que el conjunto de clases residuales de un anillo R módulo un ideal J forma un anillo con respecto a las operaciones

$$(a + J) + (b + J) = (a + b) + J, \tag{B.1}$$

$$(a + J)(b + J) = ab + J. \tag{B.2}$$

Definición B.1.40. *El anillo de clases residuales de un anillo R módulo el ideal J bajo las operaciones B.1 y B.2 se llama* **anillo de clases residuales** (o **anillo de factores**) *de R módulo J, y se denota por R/J.*

Ejemplo B.1.41 (El anillo de clases residuales $\mathbb{Z}/(n)$)**.** *Al igual que en el caso de grupos, denotamos la clase lateral o residual de $a \in \mathbb{Z}$ módulo un entero positivo n por $[a]$, así como por $a + (n)$, donde (n) es el ideal principal generado por n. Así, los elementos de $\mathbb{Z}/(n)$ son*

$$[0] = 0 + (n), [1] = 1 + (n), ..., [n - 1] = n - 1 + (n).$$

Con el trabajo desarrollado hasta aquí ya podemos dar un primer resultado sobre la estructura de los *campos finitos* –es decir, sobre campos que contienen sólo un número finito de elementos.

Teorema B.1.42. *$\mathbb{Z}/(p)$, el anillo de clases residuales de enteros módulo el ideal principal generado por un número primo p, es un campo.*

Ejemplo B.1.43. *Sea $p = 3$. Entonces $\mathbb{Z}/(p)$ está formado por los elementos* [0], [1] *y* [2]*. Las operaciones en este campo se describen en las tablas:*

$+$	[0]	[1]	[2]
[0]	[0]	[1]	[2]
[1]	[1]	[2]	[0]
[2]	[2]	[0]	[1]

$\cdot$	[0]	[1]	[2]
[0]	[0]	[0]	[0]
[1]	[0]	[1]	[2]
[2]	[0]	[2]	[1]

Observación B.1.44. *Debe haber cautela y no suponer de antemano que al formar anillos de clases residuales se van preservar las propiedades del anillo original. Por ejemplo, la propiedad de no tener divisores de cero del anillo $\mathbb{Z}$ no se preserva en $\mathbb{Z}/(n)$, con n un número compuesto.*

Se puede extender de manera natural el concepto de homomorfismo de grupos a anillos de la siguiente manera:

Definición B.1.45. *Sean R y S dos anillos. Una función $\varphi : R \longrightarrow S$ es un* **homomorfismo** *si para cualesquiera a, b in$\mathbb{R}$ se cumple que:*

$$\varphi(a+b) = \varphi(a) + \varphi(b) \quad y \quad \varphi(ab) = \varphi(a)\varphi(b).$$

Esto es, un homorfismo para anillos φ preserva las operaciones $+$ y $\cdot$ del anillo R, así como induce un homomorfismo de grupos entre el grupo aditivo de R y el correspondiente de S. Otros conceptos, tales como *núcleo*, *imagen* o *isomorfismo*, se definen de manera análoga para el caso de homomorfismos entre anillos. En este sentido, también tenemos un teorema de isomorfismo análogo al teorema correspondiente para grupos.

Teorema B.1.46 (Teorema de homomorfismo para anillos)**.** *Sean R y S dos anillos, y consideremos el homorfismo de anillos $\varphi : R \longrightarrow S$. Entonces $ker(\varphi)$ es un ideal y el anillo $im(R)$ es isomorfo al anillo de clases residuales $R/ker(\varphi)$ (bajo la función $\varphi(r) \mapsto r + ker(\varphi)$, para $r \in R$).*

Recíprocamente, si J es un ideal del anillo R, entonces la función $\psi : R \longrightarrow R/J$ dada por $\psi(r) = r + J$, para $r \in R$, es un homomorfismo de R en R/J con $ker(\psi) = J$.

Se pueden utilizar funciones entre conjuntos para conferir a un conjunto sin estructura aquélla que ya posea una estructura algebraica conocida. Para aclarar a que nos referimos con lo anterior, consideremos un anillo R y sea φ una función biyectiva de R en un conjunto S; entonces en términos de φ se puede definir una estructura de anillo al conjunto S que convierta a φ a un isomorfismo de anillos. El punto clave es el siguiente: sean $s_1, s_2 \in S$ y sean $r_1, r_2 \in R$ los elementos determinados de manera única (en otras palabras, determinados por la función biyectiva φ) por $\varphi(r_1) = s_1$ y $\varphi(r_2) = s_2$. Entonces definiendo $s_1 + s_2$ como $\varphi(r_1 + r_2)$ y $s_1 s_2$ por $\varphi(r_1 r_2)$ la función φ cumplirá todas aquellas propiedades que lo hagan un isomorfismo. La estructura así dada al conjunto S puede llamarse estructura de anillo *inducida por* φ. En caso que R tenga además otras propiedades (por mencionar algunas, que sea dominio entero o campo) entonces S «heredará» dichas propiedades. La utilidad de todo lo anterior, el punto clave, es que se puede dar una representación más «manejable» para los campos finitos $\mathbb{Z}/(p)$.

Definición B.1.47. *Sea p un número primo. Si $\mathbb{F}_p$ denota al conjunto $\{0, 1, ..., p-1\}$ de números enteros y $\varphi : \mathbb{Z}/(p) \longrightarrow \mathbb{F}_p$ es la función definida por $\varphi(a) = [a]$, para $a = 0, 1, ..., p-1$, entonces, $\mathbb{F}_p$, provisto con la estructura inducida por φ, es un campo finito, llamado el* **campo de Galois de orden** *p.*

Observación B.1.48. *Por lo discutido antes, la función $\varphi : \mathbb{Z}/(p) \longrightarrow \mathbb{F}_p$ es un isomorfismo con*

$$\varphi([a+b]) = \varphi([a]) + \varphi([b]),$$
$$\varphi([ab]) = \varphi([a])\,\varphi([b]).$$

El campo finito $\mathbb{F}_p$ tiene elemento cero 0, identidad 1 y su estructura es la estructura de $\mathbb{Z}/(p)$. Por lo tanto los elementos de $\mathbb{F}_p$ se pueden tomar como los enteros módulo p.

Ejemplo B.1.49. $\mathbb{Z}/(2)$ *es isomorfo a* $\mathbb{F}_2 = \{0, 1\}$. *Las tablas para las dos operaciones* $+$ *y* $\cdot$ *en* $\mathbb{F}_2$ *se muestran a continuación:*

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

Tomando a $\mathbb{Z}$ como grupo aditivo, un elemento diferente de cero $b \in \mathbb{Z}$ tiene orden infinito, esto es, si $nb = 0$, entonces $n = 0$. Por otro lado, para el anillo $\mathbb{F}_p$ (p un número primo), un elemento no cero $b \in \mathbb{F}_p$ satisface que $pb = 0$. En otras palabras, viendo a $\mathbb{F}_p$ como grupo aditivo, $b \in \mathbb{F}$ tiene orden p. Conviene generalizar esta propiedad para entender más a fondo la estructura de los campos finitos.

Definición B.1.50. *Sea* R *un anillo. Si existe* $n \in \mathbb{Z}$, *con* $n > 0$, *tal que* $nr = 0$ *para todo* $r \in R$, *entonces, al menor de dichos enteros positivos* n *se le llamará la* **característica** *del anillo, y se dirá que* R *tiene característica (positiva)* n. *Si no existe tal entero positivo* n, *se dirá que* R *tiene* **característica cero**.

Ejemplo B.1.51.

1. El anillo $\mathbb{Z}_n$ de los enteros módulo n tiene característica n.
2. $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ y $\mathbb{C}$ tienen, todos, característica cero.

Teorema B.1.52. *Sea* $R \neq \{0\}$ *un anillo unitario y sin divisores de cero. Si* R *tiene característica positiva, entonces ésta debe ser un número primo.*

Demostración. Nótese primero que, como el anillo R tiene elementos diferentes del cero, R tiene característica $n \geq 2$. Supongamos que n no es un número primo. Entonces existen $a, b \in \mathbb{Z}$, con $1 < a < n$ y $1 < b < n$, tal que $n = ab$. Entonces, por definición de característica, $0 = n \cdot 1 = (a \cdot b)1 = (a \cdot 1)(b \cdot 1)$, y esto implica que o bien $a \cdot 1 = 0$ o $b \cdot 1 = 0$ pues R no tiene divisores de cero. De aquí se sigue que $ar = (a \cdot 1)r = 0$ para todo $r \in R$ o $br = (b \cdot 1)r = 0$ para todo $r \in R$. Pero, en cualquier caso, esto contradice la definición de característica del anillo R. □

Usaremos principalmente el concepto de característica para un campo. En ese sentido, tenemos el siguiente resultado como consecuencia del teorema B.1.52.

Corolario B.1.53. *Todo campo finito tiene característica igual a un número primo.*

Demostración. Sea $\mathbb{F}$ un campo. Por el teorema B.1.52 es suficiente mostrar que que $\mathbb{F}$ tiene característica positiva. Consideremos los múltiplos $1, 2 \cdot 1, 3 \cdot 1, \ldots$ de la identidad $1 \in \mathbb{F}$. Ya que $\mathbb{F}$ es finito, debe haber sólo un número finito de elementos distintos en $\mathbb{F}$, por lo cual existen números enteros k, m, con $1 \leq k < m$, tal que $k \cdot 1 = m \cdot 1$, o bien, $(m - k)1 = 0$. Así, $(m - k)r = [(m - k)1]\, r = 0$ para todo $r \in \mathbb{F}$. Por lo tanto, $\mathbb{F}$ tiene característica positiva. □

Lema B.1.54. *Sea* $\mathbb{F}$ *un campo finito con* q *elementos.*

(i) Si $\mathbb{F}$ es de característica impar (en otras palabras, si la característica es distinta de 2 según el corolario B.1.53), el número de cuadrados [1] *en $\mathbb{F}$ es $\frac{1}{2}(q+1)$.*

(ii) Si $\mathbb{F}$ es de característica par, el número de cuadrados en $\mathbb{F}$ es q.

Demostración. Sea $\mathbb{F}^*$ el grupo multiplicativo de los elementos no cero del campo. La función $\varphi : \mathbb{F}^* \longrightarrow \mathbb{F}^*$ dada por $\varphi(x) = x^2$ es un homomorfismo de grupos.

(i) Si la característica es impar, $ker\,(\varphi) = \{-1, 1\}$ ya que la ecuación $x^2 - 1 = 0$ tiene dos soluciones: -1 y 1. Luego, por el teorema B.1.28 sabemos que

$$im\,(\varphi) \cong \mathbb{F}^*/ker\,(\varphi)\,.$$

Luego,

$$|im\,(\varphi)\,| = \frac{|\mathbb{F}^*|}{|ker\,(\varphi)\,|} = \frac{q-1}{2}.$$

Estos son los cuadrados de los elementos distintos del cero. En $\mathbb{F}$ hay otro cuadrado, a saber, el cero. Por lo tanto, el número de cuadrados es $\frac{q-1}{2} + 1 = \frac{q+1}{2}$.

(ii) Si la característica de $\mathbb{F}$ es par, es decir, si la característica de $\mathbb{F}$ es 2 (ver corolario B.1.53) entonces, como $x^2 - 1 = (x-1)(x+1) = (x+1)^2$, la ecuación $x^2 - 1 = 0$ tiene sólo una solución. Por lo tanto, siguiendo la misma idea que en el punto anterior, se sigue que

$$|im\,(\varphi)\,| = \frac{|\mathbb{F}^*|}{|ker\,(\varphi)\,|} = q - 1.$$

Con esto ya tenemos el número de cuadrados que corresponden a los elementos distintos del cero. Contando al cero se concluye que el número de cuadrados es igual a q.

□

B.1.3. Polinomios.

Sea R un anillo arbitrario. Un **polinomio** $p(x)$ sobre R es una expresión de la forma

$$p(x) = \sum_{i=0}^{n} a_i x^i = a_0 + a_1 x + ... + a_n x^n,$$

donde n es un número entero no negativo, los *coeficientes* a_i $(0 \leq i \leq n)$ son elementos de R y x es un símbolo que no pertenece a R, llamado una *indeterminada* sobre R. Se aceptará que, cuando $a_i = 0$, el término $a_i x^i$ sea omitido en la expresión de $p(x)$. En particular, el polinomio $p(x)$ puede escribirse, cuando sea conveniente, en la forma

$$p(x) = a_0 + a_1 x + ... + a_n x^n + 0x^{n+1} + ... + 0x^{n+h},$$

[1] Sea $\mathbb{F}$ un campo. Diremos que $a \in \mathbb{F}$ es un cuadrado en $\mathbb{F}$, si existe $b \in \mathbb{F}$ tal que $a = b^2$.

con h un número entero positivo. Así, se puede suponer que, dados dos polinomios $p(x)$ y $g(x)$ sobre R, éstos involucran las mismas potencias de x.

Los polinomios

$$f(x) = \sum_{i=0}^{n} a_i x^i \text{ y } g(x) = \sum_{i=0}^{n} b_i x^i,$$

son iguales si, y sólo si, $a_i = b_i$ para cada $0 \leq i \leq n$. Se define la *suma* de $f(x)$ y $g(x)$ como

$$f(x) + g(x) = \sum_{i=0}^{n} (a_i + b_i) x^i.$$

Para definir el *producto* de dos polinomios sobre R, supongamos que

$$f(x) = \sum_{i=0}^{n} a_i x^i \text{ y } g(x) = \sum_{j=0}^{n} b_j x^j,$$

y defínase

$$f(x)g(x) = \sum_{k=0}^{n+m} c_k x^k,$$

donde

$$c_k = \sum_{\substack{i+j=k \\ 0 \leq i \leq n,\ 0 \leq j \leq n}} a_i b_j.$$

Con estas operaciones los polinomios forman un anillo.

Definición B.1.55. *El anillo, $R[x]$, formado por los polinomios sobre R junto con las operaciones de suma y multiplicación de polinomios se llama el* **anillo de polinomios sobre** R.

El elemento cero de $R[x]$ es el polinomio cuyos coeficientes son todos 0; este polinomio se llama *polinomio cero* y se denota por 0 (se hará explícito en el contexto cuándo el símbolo 0 denota al elemento cero de R o al polinomio cero).

Definición B.1.56. *Sea $p(x) = a_0 + a_1 x + ... + a_n x^n \in R[x]$ un polinomio no cero (así que podemos suponer que $a_n \neq 0$).*

- *El coeficiente a_n es llamado el* **coeficiente principal** *de $p(x)$ y a a_0 se le llama el* **coeficiente constante**.
- *A $n \in \mathbb{N}$ se le llama el* **grado** *de $p(x)$ y se denota por $n = \deg(p)$. Por convención, $\deg(0) = -1$. Además, a los polinomios de grado $d \leq 0$ los llamaremos* **polinomios constantes**.
- *Si el anillo R tiene identidad 1 y el coeficiente principal de $p(x)$ es 1, a $p(x)$ se le llama* **polinomio mónico**.

Sea $\mathbb{F}$ un campo (no necesariamente finito). De aquí en adelante consideraremos polinomios sobre $\mathbb{F}$. Es útil, además, presentar el siguiente concepto: se dice que un polinomio $g(x) \in \mathbb{F}[x]$ *divide* al polinomio $f(x) \in \mathbb{F}[x]$ si existe $h(x) \in \mathbb{F}[x]$ tal que $f(x) = g(x)h(x)$. Además diremos que $g(x)$ es un *divisor* de $f(x)$, o que que $f(x)$ es un *múltiplo* de $g(x)$, o que $f(x)$ es *divisible* entre $g(x)$.

Teorema B.1.57 (Algoritmo de la división para $\mathbb{F}[x]$)**.** *Sea $g(x)$ un polinomio diferente de cero en $\mathbb{F}[x]$. Entonces, para cualquier $f(x) \in \mathbb{F}[x]$ existen únicos polinomios $q(x), r(x) \in \mathbb{F}[x]$ tales que*

$$f(x) = q(x)g(x) + r(x),$$

donde $\deg(r) < \deg(q)$.

Usando el algoritmo de la división se puede demostrar que todo ideal de $\mathbb{F}[x]$ es principal.

Teorema B.1.58. *$\mathbb{F}[x]$ es un dominio de ideales principales. De hecho, para cada ideal $J \neq (0)$ existe un único polinomio mónico $g(x) \in \mathbb{F}[x]$ tal que $J = (g(x))$.*

A continuación se introduce un tipo importante de polinomios.

Definición B.1.59. *Un polinomio $p(x) \in \mathbb{F}[x]$ se dice* **irreducible sobre** $\mathbb{F}$ *(o* **irreducible en** $\mathbb{F}[x]$*, o* **primo en** $\mathbb{F}[x]$*) si $p(x)$ es no constante y cada vez que $p(x) = a(x)b(x)$, con $a(x), b(x) \in \mathbb{F}[x]$, entonces, o $a(x)$ es constante o lo es $b(x)$. En caso contrario se dice que $p(x)$ es* **reducible sobre** $\mathbb{F}$.

Observación B.1.60. *El campo en el que se esté considerando al anillo de polinomios es de suma importancia en la definición anterior. Por ejemplo, el polinomio $x^2 + 1$ es irreducible en $\mathbb{R}[x]$ pero reducible en $\mathbb{C}[x]$, con factorización $(x + i)(x - i)$.*

Para entender la estructura de $\mathbb{F}[x]$, conviene estudiar a los polinomios irreducibles. En este sentido, la importancia de los últimos radica en que todo polinomio en $\mathbb{F}[x]$ pueden expresarse como producto de polinomios irreducibles de una manera (esencialmente) única, tal como se indica en el siguiente teorema.

Teorema B.1.61 (Teorema de factorización única en $\mathbb{F}[x]$)**.** *Cualquier polinomio $p(x)$ en $\mathbb{F}[x]$ no constante puede expresarse de la forma*

$$p(x) = a\,[p_1(x)]^{e_1}\,[p_2(x)]^{e_2} \ldots [p_k(x)]^{e_k}\,, \tag{B.3}$$

donde $a \in \mathbb{F}$, $p_1(x), p_2(x), ..., p_k(x)$ son distintos polinomios mónicos e irreducibles en $\mathbb{F}[x]$ y $e_1, e_2, ..., e_k$ son enteros positivos. Más aún, la factorización B.3 es única salvo el orden de los factores; además a B.3 se le llamará la factorización canónica *de $p(x)$ en $\mathbb{F}[x]$.*

Una pregunta central sobre polinomios en $\mathbb{F}[x]$ es saber cuándo un polinomio dado es irreducible. Para nuestros propósitos (construir ejemplos de campos finitos) estaremos interesados en encontrar polinomios irreducibles sobre $\mathbb{F}_p$. Un método para determinar

los polinomios irreducibles sobre $\mathbb{F}_p$ de grado d consiste en calcular primero los polinomios reducibles sobre $\mathbb{F}_p$, y después descartarlos del conjunto de polinomios mónicos en $\mathbb{F}[x]$ de grado d. Cuando d o p no son muy grandes, este métodos será suficiente para nosotros. Para métodos más sofisticados para encontrar polinomios irreducibles pueden consultarse el capítulo 3 (secciones 2 y 3) de (LN97). Para aclarar ideas, veamos el siguiente ejemplo.

Problema B.1.62. *Encontrar todos los polinomios irreducibles sobre $\mathbb{F}_2$ de grado 3.*

Solución. Primero, nótese que todos los polinomios diferentes de cero en $\mathbb{F}_2[x]$ deben ser mónicos. Ahora bien, todos los polinomios de grado 3 se pueden expresar en la forma $x^3+a_1x^2+a_2x+a_3$, donde cada coeficiente $a_i \in \{0,1\}$. Por lo tanto hay $2^3 = 8$ de dichos polinomios. Por último observe que un polinomio de grado 3 en $\mathbb{F}_2[x]$ es reducible si, y sólo si, éste tiene un divisor de grado 1. Por lo tanto, basta calcular todos los productos $(x + a_0)(x^2 + b_1x + b_0)$ para obtener los polinomios irreducibles. Hay 6 polinomios de grado 3 que son reducibles, restando 2 irreducibles, a saber: x^3+x+1 y x^3+x^2+1. □

Teorema B.1.63. *Sea $p(x) \in \mathbb{F}[x]$ un polinomio. El anillo $\mathbb{F}[x]/(p(x))$ de clases residuales es un campo si, y sólo si, $p(x)$ es irreducible sobre $\mathbb{F}$.*

Para construir ejemplos de campos finitos estaremos interesados en la estructura del anillo de clases residuales $\mathbb{F}[x]/(p(x))$, con $p(x) \in \mathbb{F}[x]$ un polinomio no cero arbitrario. En este sentido, convendrá tener en mente lo siguiente:

i $\mathbb{F}[x]/(p(x))$ está formado por las clases residuales $g(x) + (f(x))$, con $g(x) \in \mathbb{F}[x]$.

ii Sean $g(x), h(x) \in \mathbb{F}[x]$. Dos clases residuales $g(x) + (f(x))$ y $h(x) + (f(x))$ son equivalentes si, y sólo si, $g(x) - h(x) \in (f(x))$, esto es, si, y sólo si $g(x) - h(x)$ es divisible entre $f(x)$. Lo anterior es equivalente a decir que $g(x)$ y $h(x)$ tienen el mismo residuo al dividirlos entre $f(x)$.

iii Según el algoritmo de la división para $\mathbb{F}[x]$ (ver teorema B.1.57) cada clase residual $g(x)+(f(x))$ contiene un único elementos $r(x) \in \mathbb{F}[x]$ tal que $\deg(r) < \deg(f)$ –a saber, el residuo que se obtiene al dividir $g(x)$ entre $f(x)$. El proceso de pasar de $g(x)$ a $r(x)$ se le suele llamar *reducción módulo* $f(x)$.

iv A partir del punto iii se puede concluir que las distintas clases residuales contenidas en $\mathbb{F}[x]/(f(x))$ son precisamente las clases $r(x)+(f(x))$, donde $r(x)$ corre en todos los polinomios en $\mathbb{F}[x]$ con $\deg(r) < \deg(f)$. En particular, si $\mathbb{F} = \mathbb{F}_p$ (p un número primo) y $\deg(f) = n$, entonces el número de elementos en $\mathbb{F}_p/(f(x))$ es igual al número de polinomios en $\mathbb{F}_p[x]$ de grado menor que n, a saber, p^n.

Ejemplo B.1.64.

1. Sea $p(x) = x \in \mathbb{F}[x]$ Los $p^n = 2^1 = 2$ polinomios en $\mathbb{F}_2[x]$ de grado menor que 1 determinan todas las clases residuales en $\mathbb{F}_2[x]/(x)$. Por lo tanto,

$$\mathbb{F}_2[x]/(x) = \{[0] = 0 + (x), [1] = 1 + (x)\}$$

es isomorfo a $\mathbb{F}_2$.

2. Sea $f(x) = x^2 + x + 1 \in \mathbb{F}_2[x]$. Entonces $\mathbb{F}_2[x]/(p(x))$ es un campo finito (ya que $p(x)$ es irreducible y por el teorema B.1.63) con $p^n = 2^2 = 4$ elementos:

$$\mathbb{F}_2[x]/(p(x)) = \{0 + (p(x)), 1 + (p(x)), x + (p(x)), x + 1 + (p(x))\}.$$

Las tablas para las operaciones de este campo se muestran a continuación (recordar que la característica es 2). Al realizar operaciones en $\mathbb{F}_2[x]/(p(x))$ observe que, como se reemplaza cada ocurrencia de $p(x)$ por 0, el representante de cada clase residual tiene grado menor que 2.

$+$	$[0]$	$[1]$	$[x]$	$[x+1]$
$[0]$	$[0]$	$[1]$	$[x]$	$[x+1]$
$[1]$	$[1]$	$[0]$	$[x+1]$	$[x]$
$[x]$	$[x]$	$[x+1]$	$[0]$	$[1]$
$[x+1]$	$[x+1]$	$[x]$	$[1]$	$[0]$

$\cdot$	$[0]$	$[1]$	$[x]$	$[x+1]$
$[0]$	$[0]$	$[0]$	$[0]$	$[0]$
$[1]$	$[0]$	$[1]$	$[x]$	$[x+1]$
$[x]$	$[0]$	$[x]$	$[x+1]$	$[1]$
$[x+1]$	$[0]$	$[x+1]$	$[1]$	$[x]$

Note además que, en la tabla de multiplicación,

$$[x + (p(x))][x + (p(x))] = x^2 + (p(x)) = p(x) - x - 1 = x + 1 + (p(x)),$$

$$[x + (p(x))][x + 1 + (p(x))] = x^2 + x + (p(x)) = p(x) - 1 + (p(x)) = 1 + (p(x)),$$

$$[x + 1 + (p(x))][x + 1 + (p(x))] = x^2 + 1 + (p(x)) = p(x) - x + (p(x)) = x + (p(x)).$$

Este es nuestro primer ejemplo de un campo finito cuyo número de elementos ¡no es un número primo!

3. Sea $p(x) = x^2 + 2 \in \mathbb{F}_3[x]$. Se puede observar que $\mathbb{F}_3[x]/(p(x))$ es un anillo con 9 elementos que no es un dominio entero, mucho menos un campo (observe que $p(x)$ no es irreducible). Sus elementos son:

$$\mathbb{F}_3[x]/(p(x)) = \{[0], [1], [2], [x], [x+1], [x+2], [2x], [2x+1], [2x+2]\}$$

Para ver que no es un dominio entero basta notar que

$$[x+1][x+2] = [x+1][x-1] = [x^2 - 1] = [x^2 + 2] = [0],$$

pero, claramente, ni $[x+1]$ ni $[x+2]$ son iguales a $[0]$.

Si $\mathbb{F}$ es un campo arbitrario y $p(x) \in \mathbb{F}[x]$, podemos «reemplazar» la indeterminada x en $p(x)$ por un elemento fijo de $\mathbb{F}$ para obtener otro elemento de $\mathbb{F}$. Más precisamente, si $p(x) = a_0 + a_1x + ... + a_nx^n \in \mathbb{F}[x]$ y $b \in \mathbb{F}$, al *reemplazar* x por b se obtiene $p(b) = a_0 + a_1b + ... + a_nb^n \in \mathbb{F}$.

Definición B.1.65. *Un elemento $a \in \mathbb{F}$ se llama* **raíz** (*o* **cero**) *del polinomio $p(x)$ en $\mathbb{F}[x]$ si $p(a) = 0$.*

Ejemplo B.1.66. El polinomio $x^2 - 2$ no tiene raíces en $\mathbb{Q}$. En $\mathbb{R}$, por otro lado, $\sqrt{2}$ y $-\sqrt{2}$ son raíces de dicho polinomio.

Las raíces de un polinomio están ligadas al concepto de divisibilidad.

Teorema B.1.67. *Un elemento $b \in \mathbb{F}$ es un raíz de $p(x) \in \mathbb{F}[x]$ si, y sólo si, $x - b$ divide a $p(x)$.*

Definición B.1.68. *Sea $b \in \mathbb{F}$ una raíz de un polinomio $p(x) \in \mathbb{F}[x]$. Si k es un entero positivo tal que $p(x)$ es divisible entre $(x - b)^k$, pero no entre $(x - b)^{k+1}$, entonces a k se le llama la* **multiplicidad** *de b. Si $k = 1$, b se llama* **raíz simple** (*o* **cero simple**) *de $p(x)$, y si $k \geq 2$ se llama* **raíz múltiple** (*o* **cero múltiple**) *de $p(x)$.*

Teorema B.1.69. *Un elemento $a \in \mathbb{F}$ es una raíz de multiplicidad de un polinomio $p(x) \in \mathbb{F}[x]$ si, y sólo si, $a \in \mathbb{F}$ es una raíz tanto de $p(x)$ como de su* derivada *$p'(x)$.*

Demostración. Supongamos primero que $a \in \mathbb{F}$ es un raíz múltiple de $p(x)$. Entonces $p(x) = (x - a)^k q(x)$, con $k \geq 2$, para algún $q(x) \in \mathbb{F}[x]$. Claramente $p(a) = 0$, así que sólo resta comprobar que $a \in \mathbb{F}$ es raíz de $p'(x)$. Tomando la derivada de $p(x)$ se sigue que

$$p'(x) = (x - a)^{k-1}q(x) + (x - a)^k q'(x).$$

De aquí se puede concluir que $p'(a) = 0$ (ya que $k - 1 \geq 1$).

Recíprocamente, supongamos que $a \in \mathbb{F}$ es una raíz tanto de $p(x)$ como de $p'(x)$. Como $p(a) = 0$, existe un polinomio $q_1(x) \in \mathbb{F}[x]$ tal que $p(x) = (x - a)q_1(x)$. Tomando la derivada de $p(x)$ se obtiene que $p'(x) = q_1(x) + (x - a)q_1'(x)$. De aquí se sigue que $p'(a) = q_1(a)$, pero, como $a \in \mathbb{F}$ es raíz de $p'(x)$, resulta que $q_1(a) = 0$. Luego, $q_1(x) = (x - a)q_0(x)$ para algún polinomio $q_0(x) \in \mathbb{F}[x]$. Con lo anterior en cuenta, se concluye que $p(x) = (x - a)^2 q_0(x)$, esto es, la multiplicidad de $a \in \mathbb{F}$ es al menos 2, por lo que es una raíz múltiple. □

Ejemplo B.1.70. Consideremos el polinomio $p(x) = x^3 - 7x^2 + 16x - 12 \in \mathbb{R}[x]$. Dicho polinomio puede expresarse como $p(x) = (x-2)^2(x-3)$, así que 2 y 3 son raíces, respectivamente, de multiplicidades 2 y 1. Ahora bien, $p'(x) = 3x^2 - 14x + 16$ que puede factorizarse como $p'(x) = (x - 2)(3x - 8)$, así que se verifica que 2 es también una raíz de $p'(x)$.

Concluimos esta sección con el siguiente resultado.

Teorema B.1.71. *Sea $\mathbb{F}$ un campo. Si $p(x) \in \mathbb{F}[x]$ es un polinomio no cero de grado n, entonces $p(x)$ tiene a lo más n raíces en $\mathbb{F}$.*

B.1.4. Extensiones de campo.

Definición B.1.72. *Sea $\mathbb{F}$ un campo. Un subconjunto $\mathbb{K}$ de $\mathbb{F}$ se llamará un* **subcampo** *de $\mathbb{F}$ si él mismo es un campo bajo las operaciones definidas en $\mathbb{F}$. En este contexto, $\mathbb{F}$ se dirá ser una* **extensión (de campo)** *de $\mathbb{K}$. Si además $\mathbb{K} \neq \mathbb{F}$, diremos que $\mathbb{K}$ es un* **subcampo propio** *de $\mathbb{F}$.*

Ejemplo B.1.73. El campo $\mathbb{Q}$ de los números racionales es un subcampo propio del campo $\mathbb{R}$ de los números reales, quien a su vez es un subcampo propio del campo de los números complejos $\mathbb{C}$.

Observación B.1.74. *Si $\mathbb{K}$ es un subcampo del campo finito $\mathbb{Z}_p$ (p un número primo) entonces $\mathbb{K}$ debe contener a los elementos 0 y 1 en $\mathbb{F}$ y, por lo tanto, a todos los elementos de $\mathbb{Z}_p$ por la cerradura de $\mathbb{K}$ bajo la suma. Así, el campo $\mathbb{Z}_p$ no contiene subcampos propios.*

Con la observación anterior en cuenta, podemos considerar la siguiente definición:

Definición B.1.75. *Un campo que no contenga subcampos propios se llamará* **campo primo**.

Ejemplo B.1.76. Por el argumento de la observación anterior, cualquier campo de orden p, con p un número primo, es un campo primo. Otro ejemplo de campo primo es el campo $\mathbb{Q}$ de los números racionales.

La intersección de cualquier colección no vacía de subcampos de un campo $\mathbb{F}$ dado es también un subcampo de $\mathbb{F}$. Si se toma la intersección de *todos* los subcampos de $\mathbb{F}$, se obtiene el ***subcampo primo*** de $\mathbb{F}$. Claramente éste es un campo primo.

Teorema B.1.77. *Sea $\mathbb{F}$ un campo. El subcampo primo de $\mathbb{F}$ es isomorfo a $\mathbb{F}_p$ (con p un número primo) o a $\mathbb{Q}$, según la característica de $\mathbb{F}$ sea p o 0.*

Definición B.1.78. *Sea $\mathbb{F}$ un campo y $\mathbb{K}$ un subcampo de $\mathbb{F}$. Para cualquier subconjunto M de $\mathbb{F}$ se define el campo $\mathbb{K}(M)$ como la intersección de todos los subcampos de $\mathbb{F}$ que contengan tanto a M como a $\mathbb{K}$ (es decir, el subcampo más pequeño de $\mathbb{F}$ que contenga tanto a $\mathbb{K}$ como a M). A dicho campo se le llama la extensión (de campo) de $\mathbb{K}$ obtenida* agregando *los elementos en M.*

- *Si M es finito, $M = \{\alpha_1, \alpha_2, ..., \alpha_n\}$, se escribirá $\mathbb{K}(\alpha_1, \alpha_2, ..., \alpha_n)$.*
- *Si $M = (\alpha)$, entonces $\mathbb{L} = \mathbb{K}(\alpha)$ se dice ser una* **extensión simple** *de $\mathbb{K}$.*

Observación B.1.79. *Sea $\mathbb{F}$ un campo, $\mathbb{K}$ un subcampo de $\mathbb{F}$ y $u, v \in \mathbb{F}$. Ya que $\mathbb{K}(u)$ es una extensión de $\mathbb{K}$, podemos considerar la extensión $\mathbb{K}(u)(v)$ de $\mathbb{K}(u)$.*
Veamos que $\mathbb{K}(u, v) = \mathbb{K}(u)(v)$.

(i) Es fácil observar que $\mathbb{K}(u)(v)$ contiene a $\mathbb{K}$, u y v. De donde, al ser $\mathbb{K}(u, v)$ el mínimo subcampo con esa propiedad, se obtiene que $\mathbb{K}(u, v) \subseteq \mathbb{K}(u)(v)$.

(ii) Primero, obsérvese que $\mathbb{K}(u,v)$ *contiene a* $\mathbb{K}(u)$*. En efecto,* $\mathbb{K}(u,v)$ *contiene a* $\mathbb{K}$ *y a* u *por lo cual contiene a* $\mathbb{K}(u)$*, ya que éste es el mínimo subcampo que contiene a* $\mathbb{K}$ *y* u*. También se cumple que* $\mathbb{K}(u,v)$ *contiene a* v*, lo cual implica que* $\mathbb{K}(u)(v) \subseteq \mathbb{K}(u,v)$ *ya que el primero es el mínimo subcampo con esa propiedad.*

De (i) y (ii) se sigue que $\mathbb{K}(u,v) = \mathbb{K}(u)(v)$*. Un razonamiento análogo junto con el principio de inducción matemática permite concluir que*

$$\begin{aligned}
\mathbb{K}(u_1,u_2) &= \mathbb{K}(u_1)(u_2) \\
\mathbb{K}(u_1,u_2,u_3) &= \mathbb{K}(u_1,u_2)(u_3) \\
&\vdots \\
\mathbb{K}(u_1,...,u_{n-1},u_n) &= \mathbb{K}(u_1,...,u_{n-1})(u_n).
\end{aligned}$$

Definición B.1.80. *Sea* $\mathbb{K}$ *un subcampo de un campo* $\mathbb{F}$ *y* $\alpha \in \mathbb{F}$*. Si* α *satisface una ecuación polinomial (no trivial) con coeficientes en* $\mathbb{K}$*, esto es, si*

$$a_n\alpha^n + ... + a_1\alpha + a_0 = 0,$$

con $a_i \in \mathbb{K}$ $(i = 0,1,...,n)$ *y no todos cero, entonces* α *se dice ser algebraico sobre* $\mathbb{K}$*.*

Una extensión $\mathbb{L}$ *de* $\mathbb{K}$ *se llama* **algebraica** *sobre* $\mathbb{K}$ (*o* **extensión algebraica** *de* $\mathbb{K}$) *si todo elemento de* $\mathbb{L}$ *es algebraico sobre* $\mathbb{K}$*.*

Ejemplo B.1.81.

1. $\sqrt[3]{3} \in \mathbb{R}$ es algebraico sobre $\mathbb{Q}$ ya que es raíz del polinomio $x^3 - 3 \in \mathbb{Q}[x]$.
2. $i \in \mathbb{C}$ es algebraico sobre $\mathbb{R}$ ya que es raíz del polinomio $x^2 + 1 \in \mathbb{R}[x]$
3. $\pi \in \mathbb{R}$ no es algebraico sobre $\mathbb{Q}$. Un elemento que no es algebraico sobre un campo $\mathbb{F}$ se dice ***trascendental*** sobre $\mathbb{F}$.

Dado α en un campo $\mathbb{F}$ y algebraico sobre algún subcampo $\mathbb{K}$ de $\mathbb{F}$, el conjunto $J = \{f(x) \in \mathbb{K}[x] : f(\alpha) = 0\}$ es un ideal de $\mathbb{K}[x]$ (además $J \neq (0)$ pues α es algebraico sobre $\mathbb{K}$). Así, por el teorema B.1.58 existe un único polinomio $g(x) \in \mathbb{K}[x]$ tal que $J = (g(x))$.

Definición B.1.82. *Sea* $\mathbb{K}$ *un subcampo de un campo* $\mathbb{F}$*. Si* α *es algebraico sobre* $\mathbb{K}$*, al único polinomio mónico* $g(x) \in \mathbb{K}[x]$ *que genera al ideal* $J = \{f(x) \in \mathbb{K}[x] : f(\alpha) = 0\}$ *de* $\mathbb{K}[x]$ *se le llama* **polinomio mínimo** (*o* **polinomio irreducible**) *de* α *sobre* $\mathbb{K}$*. Nos referiremos al grado de* $g(x)$ *como* **el grado** *de* α *sobre* $\mathbb{K}$*.*

El siguiente teorema resume las propiedades principales del polinomio mínimo de la definición anterior. En particular, la tercera de ellas será la más útil para nosotros.

Teorema B.1.83. *Sea* $\mathbb{F}$ *un campo y* $\mathbb{K}$ *un subcampo de* $\mathbb{F}$*. Si* α *es algebraico sobre* $\mathbb{K}$*, entonces el polinomio mínimo de* α *sobre* $\mathbb{K}$ *tiene las siguientes propiedades:*

(i) $g(x)$ *es irreducible en* $\mathbb{K}[x]$.

(ii) Dado un polinomio $f(x) \in \mathbb{K}[x]$, $f(\alpha) = 0$ *si, y sólo si,* $g(x)$ *divide a* $f(x)$.

(ii) $g(x)$ *es el polinomio mónico de menor grado que tiene a* α *como raíz.*

Ejemplo B.1.84.

1. $\sqrt[3]{3} \in \mathbb{R}$ es algebraico sobre $\mathbb{Q}$ ya que es raíz del polinomio $x^3 - 3 \in \mathbb{Q}[x]$. Además, puesto que $x^3 - 3$ es irreducible sobre $\mathbb{Q}$, éste es el polinomio mínimo de $\sqrt[3]{3}$ sobre $\mathbb{Q}$; el grado de $\sqrt[3]{3}$ sobre $\mathbb{Q}$ es 3.

2. $i = \sqrt{-1} \in \mathbb{C}$ es algebraico sobre el subcampo $\mathbb{R}$ de $\mathbb{C}$ ya que es raíz del polinomio $x^2 + 1 \in \mathbb{R}[x]$. Como $x^2 + 1$ es irreducible sobre $\mathbb{R}$, éste es el polinomio mínimo de i sobre $\mathbb{R}$ y, por tanto, el grado de i sobre $\mathbb{R}$ es 2.

Extensiones de campo como espacios vectoriales.

Sea $\mathbb{L}$ una extensión de campo de $\mathbb{K}$. Un enfoque útil para estudiar a $\mathbb{L}$ es considerar a éste como *espacio vectorial* sobre $\mathbb{K}$. Para convencerse de que, en efecto, $\mathbb{L}$ es un $\mathbb{K}$–espacio vectorial basta observar que los elementos de $\mathbb{L}$ (que son los «vectores») forman un grupo abeliano bajo la suma. Más aún, cada vector $v \in \mathbb{L}$ puede multiplicarse por un «escalar» $r \in \mathbb{K}$ de modo que $rv \in \mathbb{L}$ (aquí rv denota simplemente el producto entre los elementos r y v del campo $\mathbb{L}$). Es claro que también se cumplen las propiedades para el *producto por escalares.*

Ejemplo B.1.85. *Sea* $\mathbb{L} = \mathbb{C}$ *y* $\mathbb{K} = \mathbb{R}$. *Fácilmente se puede comprobar que* $\mathbb{C}$ *es un* $\mathbb{R}$*–espacio vectorial. Además, como* $\mathbb{C} = \{a + bi : a, b \in \mathbb{R}\}$ *es claro que una base para este espacio es* $\{1, i\}$.

Definición B.1.86. *Sea* $\mathbb{L}$ *una extensión de campo de* $\mathbb{K}$. *Si* $\mathbb{L}$, *visto como* $\mathbb{K}$*–espacio vectorial, es dimensionalmente finito, entonces* $\mathbb{L}$ *se dice una* **extensión finita** *de* $\mathbb{K}$. *La dimensión del espacio vectorial* $\mathbb{L}$ *sobre* $\mathbb{K}$ *se llama el* **grado** *de* $\mathbb{L}$ *sobre* $\mathbb{K}$ *y se denota por* $[\mathbb{L} : \mathbb{K}]$.

Ejemplo B.1.87. Del ejemplo B.1.85 se sigue que $\mathbb{C}$ es una extensión finita de $\mathbb{R}$ de grado 2.

Teorema B.1.88. *Si* $\mathbb{L}$ *es una extensión finita de* $\mathbb{K}$, *y* $\mathbb{M}$ *es una extensión finita de* $\mathbb{L}$, *entonces* $\mathbb{M}$ *es una extensión finita de* $\mathbb{K}$ *y*

$$[\mathbb{M} : \mathbb{K}] = [\mathbb{M} : \mathbb{L}]\,[\mathbb{L} : \mathbb{K}]\,. \tag{B.4}$$

Observación B.1.89. *La fórmula B.4 se extiende de manera natural como sigue: para los campos*

$$\mathbb{K} \subseteq \mathbb{F}_1 \subseteq \ldots \subseteq \mathbb{F}_{n-1} \subseteq \mathbb{F}_n,$$

se cumple que

$$[\mathbb{F}_n : \mathbb{K}] = [\mathbb{F}_n : \mathbb{F}_{n-1}] \ldots [\mathbb{F}_1 : \mathbb{K}]\,.$$

Teorema B.1.90. *Toda extensión finita de* $\mathbb{K}$ *es algebraica sobre* $\mathbb{K}$.

Demostración. Sea $\mathbb{L}$ una extensión finita de $\mathbb{K}$ y hagamos $[\mathbb{L}:\mathbb{K}]=n$. Para $\alpha\in\mathbb{L}$, los $n+1$ elementos $1,\alpha,...,\alpha^n$ deben ser linealmente dependientes, por lo cual existe una combinación lineal $a_n\alpha^n+...+a_1\alpha+a_0=0$, con $a_i\in\mathbb{K}$ $(i=0,1,...,n)$ y donde no todos los escalares a_i son cero. Lo anterior implica que $\alpha\in\mathbb{L}$ es algebraico sobre $\mathbb{K}$. □

Sea $\mathbb{F}$ una extensión de $\mathbb{K}$ y $\alpha\in\mathbb{F}$ algebraico sobre $\mathbb{K}$. El siguiente teorema asegura que la extensión simple $\mathbb{K}(\alpha)$ de $\mathbb{K}$ obtenida al agregar el elemento α es una extensión finita (y, por tanto, una extensión algebraica) de $\mathbb{K}$.

Teorema B.1.91. *Sea* $\alpha\in\mathbb{F}$ *un elemento algebraico de grado* n *sobre* $\mathbb{K}$ *y sea* $g(x)$ *el polinomio mínimo de* α *sobre* $\mathbb{K}$. *Entonces:*

(a) $\mathbb{K}(\alpha)$ *es isomorfo a* $\mathbb{K}[x]/(g(x))$.

(b) $[\mathbb{K}(\alpha):\mathbb{K}]=n$ *y* $\{1,\alpha,...,\alpha^{n-1}\}$ *es una base del* $\mathbb{K}$*–espacio vectorial* $\mathbb{K}(\alpha)$.

(c) Cada $\beta\in\mathbb{K}(\alpha)$ *es algebraico sobre* $\mathbb{K}$ *y su grado sobre* $\mathbb{K}$ *divide a* n.

Observación B.1.92. *El teorema B.1.91 nos dice que los elementos de la extensión finita* $\mathbb{K}(\alpha)$ *de* $\mathbb{K}$ *se pueden expresar como polinomios en* α. *Además, cada* $\beta\in\mathbb{K}(\alpha)$ *se puede expresar de manera como*

$$\beta=a_{n-1}\alpha^{n-1}+...+a_1\alpha+a_0,$$

con $a_i\in\mathbb{K}$ $(i=0,1,...,n-1)$.

Ejemplo B.1.93.

1. Consideremos la extensión simple $\mathbb{R}(i)$ de $\mathbb{R}$. En el ejemplo B.1.84 vimos que $i\in\mathbb{C}$ tiene a x^2+1 como polinomio mínimo sobre $\mathbb{R}$. Así, $\mathbb{R}(i)\cong\mathbb{R}[x]/(x^2+1)$ y $\{1,i\}$ es una base para el $\mathbb{R}$–espacio vectorial $\mathbb{R}(i)$. Por lo tanto $\mathbb{C}=\mathbb{R}(i)$.

2. Tomemos la extensión simple $\mathbb{Q}\left(\sqrt[3]{3}\right)$ de $\mathbb{Q}$. Ya vimos en el ejemplo B.1.84 que el polinomio x^3-3 es el polinomio mínimo de $\sqrt[3]{3}$ sobre $\mathbb{Q}$. En consecuencia, $\mathbb{Q}\left(\sqrt[3]{3}\right)\cong\mathbb{Q}[x]/(x^3-3)$ y $\left\{1,\sqrt[3]{3},\left(\sqrt[3]{3}\right)^2\right\}$ es una base para el $\mathbb{Q}$–espacio vectorial $\mathbb{Q}\left(\sqrt[3]{3}\right)$. Además,

$$\mathbb{Q}\left(\sqrt[3]{3}\right)=\left\{a+b\sqrt[3]{3}+c\left(\sqrt[3]{3}\right)^2:a,b,c\in\mathbb{Q}\right\}.$$

Obsérvese que el teorema B.1.91 supone que $\mathbb{K}$ y α son, respectivamente, un subcampo y un elemento de un campo más grande $\mathbb{F}$. En lo que sigue estaremos interesados en construir extensiones algebraicas sin considerar de antemano un campo más grande. En ese sentido, el siguiente teorema es uno de los resultados más importantes en *teoría de campos*; éste dice que, dado un polinomio no constante sobre un campo, existe una extensión de campo en la cual el polinomio referido tiene una raíz.

Teorema B.1.94. *Sea $p(x) \in \mathbb{K}[x]$ un polinomio irreducible sobre un campo $\mathbb{K}$. Entonces, existe una extensión algebraica simple $\mathbb{K}(\alpha)$ de $\mathbb{K}$, donde α es una raíz de $p(x)$.*

Demostración. Sea $\mathbb{L} = \mathbb{K}[x]/(p(x))$ el anillo de clases residuales, el cual es un campo según el teorema B.1.63. Sus elementos son clases residuales $[f(x)] = f(x) + (p(x))$, con $f(x) \in \mathbb{K}[x]$. Para cualquier $a \in \mathbb{K}$, podemos pensar a éste como un polinomio constante en $\mathbb{K}[x]$ y considerar la clase residual $[a] \in \mathbb{L}$.

Con ayuda de la clase residual $[a]$ podemos identificar de manera natural a $\mathbb{K}$ con un subcampo de $\mathbb{L}$ a través del homomorfismo (de anillos) $\varphi : \mathbb{K} \longrightarrow \mathbb{L}$ definido como $\varphi(a) = [a]$, para cada $a \in \mathbb{K}$. Este homomorfismo es inyectivo ya que, si $\varphi(a) = \varphi(b)$, es decir, si $[a] = [b]$, con $a, b \in \mathbb{K}$, entonces $[a - b] = [0]$. Luego, $a - b \in (p(x))$, de donde, $a - b$ debe ser un múltiplo del polinomio $p(x)$; éste último, recordemos, es de grado positivo. Así, necesariamente se debe cumplir que $a - b = 0$ y, por tanto, $a = b$. Como toda función es suprayectiva sobre su imagen, podemos asegurar que φ es un isomorfismo sobre su imagen e identificar a $\mathbb{K}$ con $\mathbb{K}' = im(\varphi)$. En otras palabras, podemos pensar a $\mathbb{L}$ como una extensión de $\mathbb{K}$.

Ahora bien, para cada $f(x) = a_0 + a_1 x + ... + a_n x^m \in \mathbb{K}[x]$, se tiene que

$$\begin{aligned} [f(x)] &= [a_0 + a_1 x + ... + a_m x^m] \\ &= [a_0] + [a_1][x] + ... + [a_m][x]^m \\ &= a_0 + a_1[x] + ... + a_m[x]^m, \end{aligned} \tag{B.5}$$

haciendo la identificación de $[a_i]$ con a_i, para cada $i = 0, 1, ..., m$. Por lo tanto, todo elemento de $\mathbb{L}$ puede expresarse como un polinomio en $[x]$ con coeficientes en $\mathbb{K}$. Como $\mathbb{L}$ contiene a $\mathbb{K}$ (estrictamente, contiene a la «copia» $\mathbb{K}'$ de $\mathbb{K}$) y a $[x]$, entonces $\mathbb{K}([x]) \subseteq \mathbb{L}$ (pues $\mathbb{K}([x])$ es el mínimo subcampo con esta propiedad). Para la otra contención basta observar que cualquier campo que contenga a $\mathbb{K}$ y a $[x]$ debe contener las expresiones de la forma B.5. Por lo tanto, $\mathbb{L}$ es la extensión simple $\mathbb{K}([x])$ de $\mathbb{K}$ obtenida al agregar $[x]$.

Por último, veamos que, en realidad, $\mathbb{L}$ es una extensión algebraica simple de $\mathbb{K}$. En efecto, si $p(x) = b_0 + b_1 x + ...b_n x^n$, entonces

$$p([x]) = b_0 + b_1[x] + ... + b_n[x]^n = [b_0 + b_1 x + ... + b_n x^n] = [p(x)] = [0].$$

Es decir, $[x]$ es una raíz de $p(x)$. En consecuencia, haciendo $\alpha = [x]$ se cumple el resultado.

□

Ejemplo B.1.95. Como ejemplo del proceso formal de agregar una raíz mencionado en el teorema anterior, consideremos el campo $\mathbb{F}_3$ y al polinomio $p(x) = x^2 + x + 2 \in \mathbb{F}_3[x]$, el cual es irreducible sobre $\mathbb{F}_3$ (esto se puede comprobar evaluando a $p(x)$ en cada elemento de $\mathbb{F}_3$). Sea α una «raíz» de $p(x)$ en el sentido que α es la clase residual $[x] = x + (p(x))$

en $\mathbb{L} = \mathbb{F}_3/(p(x))$. Explícitamente se tiene que

$$\begin{aligned} p(\alpha) &= p([x]) \\ &= p(x + (p(x))) \\ &= [x + (p(x))]^2 + [x + (p(x))] + [2 + (p(x))] \\ &= \left(x^2 + x + 2\right) + (p(x)) \\ &= p(x) + (p(x)) \\ &= 0 + (p(x)) \\ &= [0]. \end{aligned}$$

La otra raíz de $p(x)$ en $\mathbb{L}$ es $2\alpha + 2$ ya que

$$\begin{aligned} p(2\alpha + 2) &= (2\alpha + 2)^2 + (2\alpha + 2) + 2 \\ &= \alpha^2 + \alpha + 2 \\ &= 0. \end{aligned}$$

De acuerdo al punto *(b) d*el teorema B.1.91 se sigue que la extensión finita $\mathbb{L} = \mathbb{F}_3(\alpha)$ es igual al conjunto

$$\{0, 1, 2, \alpha, \alpha + 1, \alpha + 2, 2\alpha, 2\alpha + 1, 2\alpha + 2\}.$$

Por último, las operaciones del campo $\mathbb{L}$ se muestran en las tablas B.1 y B.2.

$+$	0	1	2	α	$\alpha+1$	$\alpha+2$	2α	$2\alpha+1$	$2\alpha+2$
0	0	1	2	α	$\alpha+1$	$\alpha+2$	2α	$2\alpha+1$	$2\alpha+2$
1	1	2	0	$\alpha+1$	$\alpha+2$	α	$2\alpha+1$	$2\alpha+2$	2α
2	2	0	1	$\alpha+2$	α	$\alpha+1$	$2\alpha+1$	2α	$2\alpha+1$
α	α	$\alpha+1$	$\alpha+2$	2α	$2\alpha+1$	$2\alpha+2$	0	1	2
$\alpha+1$	$\alpha+1$	$\alpha+2$	α	$2\alpha+1$	$2\alpha+2$	2α	1	2	0
$\alpha+2$	$\alpha+2$	α	$\alpha+1$	$2\alpha+2$	2α	$2\alpha+1$	2	0	1
2α	2α	$2\alpha+1$	$2\alpha+2$	0	1	2	α	$\alpha+1$	$\alpha+2$
$2\alpha+1$	$2\alpha+1$	$2\alpha+2$	$2\alpha+1$	1	2	0	$\alpha+1$	$\alpha+2$	α
$2\alpha+2$	$2\alpha+2$	2α	$2\alpha+1$	2	0	1	$\alpha+2$	α	$\alpha+1$

Tabla B.1: Tabla de suma para el campo $\mathbb{F}_3(\alpha)$.

Ejemplo B.1.96. Consideremos el polinomio $p(x) = x^2 + x + 1 \in \mathbb{F}_2[x]$, el cual es irreducible sobre $\mathbb{F}_2$. Sea α la raíz $[x] = x + (p(x))$ de $p(x)$. Entonces la extensión simple $\mathbb{L} = \mathbb{F}_2(\alpha)$ consiste de cuatro elementos: $0, 1, \alpha, \alpha + 1$ (la otra raíz de $p(x)$ es $\alpha + 1$). El modo de operar los elementos del campo $\mathbb{F}_2(\alpha)$ se basa en las tablas B.3 y B.4. (Comparar con el punto 2 de ejemplo B.1.64.)

Observación B.1.97. *Note que tanto en el ejemplo B.1.95 como en el B.1.96 se obtiene la misma extensión de campo al agregar la otra raíz del polinomio $p(x)$. Esto se debe al siguiente teorema.*

$+$	0	1	2	α	$\alpha+1$	$\alpha+2$	2α	$2\alpha+1$	$2\alpha+2$
0	0	0	0	0	0	0	0	0	0
1	0	1	2	α	$\alpha+1$	$\alpha+2$	2α	$2\alpha+1$	$2\alpha+2$
2	0	2	1	2α	$2\alpha+2$	$2\alpha+1$	α	$\alpha+2$	$\alpha+1$
α	0	α	2α	$2\alpha+1$	1	$\alpha+1$	$\alpha+2$	$2\alpha+2$	2
$\alpha+1$	0	$\alpha+1$	$2\alpha+2$	1	$\alpha+2$	2α	2	α	$2\alpha+1$
$\alpha+2$	0	$\alpha+2$	$2\alpha+1$	$\alpha+1$	2α	2	$2\alpha+2$	1	α
2α	0	2α	α	$\alpha+2$	2	$2\alpha+2$	$2\alpha+1$	$\alpha+1$	1
$2\alpha+1$	0	$2\alpha+1$	$\alpha+2$	$2\alpha+2$	α	1	$\alpha+1$	2	2α
$2\alpha+2$	0	$2\alpha+2$	$\alpha+1$	2	$2\alpha+1$	α	1	2α	$\alpha+2$

Tabla B.2: Tabla de multiplicación para el campo $\mathbb{F}_3(\alpha)$.

$+$	0	1	α	$\alpha+1$
0	0	1	α	$\alpha+1$
1	1	0	$\alpha+1$	α
α	α	$\alpha+1$	0	1
$\alpha+1$	$\alpha+1$	α	1	0

Tabla B.3: Tabla de suma para el campo $\mathbb{F}_2(\alpha)$.

$\cdot$	0	1	α	$\alpha+1$
0	0	0	0	0
1	0	1	α	$\alpha+1$
α	0	α	$\alpha+1$	1
$\alpha+1$	0	$\alpha+1$	1	α

Tabla B.4: Tabla de multiplicación para el campo $\mathbb{F}_2(\alpha)$.

Teorema B.1.98. *Sea $p(x) \in \mathbb{K}[x]$ un polinomio irreducible sobre el campo $\mathbb{K}$. Si α y β son raíces de $p(x)$, entonces $\mathbb{K}(\alpha)$ y $\mathbb{K}(\beta)$ son isomorfos.*

Demostración. Se sabe por *(a)* del teorema B.1.91 que tanto $\mathbb{K}(\alpha)$ como $\mathbb{K}(\beta)$ son isomorfos a $\mathbb{K}[x]/(p(x))$ ya que el polinomio irreducible $p(x)$ es el polinomio mínimo de α y β. Por lo tanto,

$$\mathbb{K}(\alpha) \cong \mathbb{K}[x]/(p(x)) \cong \mathbb{K}(\beta),$$

de donde, $\mathbb{K}(\alpha) \cong \mathbb{K}(\beta)$. □

Nuestro siguiente objetivo será encontrar una extensión de campo que contenga todas las raíces de un polinomio dado.

Definición B.1.99. *Sea $p(x) \in \mathbb{K}[x]$ un polinomio de grado positivo y $\mathbb{F}$ una extensión de campo de $\mathbb{K}$. Se dice que $p(x)$ se* **descompone** *en $\mathbb{F}$ si $p(x)$ puede expresarse como*

producto de factores lineales en $\mathbb{F}[x]$ *–esto es, si existen* $\alpha_1, \alpha_2, ..., \alpha_n \in \mathbb{F}$ *tales que*

$$p(x) = a\,(x-\alpha_1)\,(x-\alpha_2)\,...\,(x-\alpha_n)\,,$$

donde a *es el coeficiente principal de* $p(x)$*. Si* $p(x)$ *se descompone en* $\mathbb{F}$ *y si, además,* $\mathbb{F} = \mathbb{K}(\alpha_1, \alpha_2, ..., \alpha_n)$*, entones* $\mathbb{F}$ *se dice ser un* **campo de descomposición** *de* $p(x)$ *sobre* $\mathbb{K}$*.*

Por lo tanto, un campo de descomposición $\mathbb{F}$ de un polinomio $p(x)$ sobre $\mathbb{K}$ es una extensión de campo que contiene todas las raíces de $p(x)$, y es el «más pequeño» en el sentido que ningún otro subcampo contiene todas las raíces de $p(x)$. El siguiente teorema responde a las pregunta ¿es siempre posible encontrar un campo de descomposición? De ser así, ¿cuántos campos de descomposición existen?

Teorema B.1.100 (Existencia y unicidad de campos de descomposición)**.** *Sea* $\mathbb{K}$ *un campo y* $p(x)$ *un polinomio no constante en* $\mathbb{K}[x]$*.*

(a) Existe un campo de descomposición de $p(x)$ *sobre* $\mathbb{K}$*.*

(b) Cualesquiera dos campos de descomposición de $p(x)$ *sobre* $\mathbb{K}$ *son isomorfos.*

Observación B.1.101. *El teorema anterior permite hablar de* el *campo de descomposición de* $p(x)$ *sobre* $\mathbb{K}$*. Éste se obtiene a partir de* $\mathbb{K}$ *al agregar un número finito de elementos algebraicos sobre* $\mathbb{K}$ *y, por tanto, utilizando el teorema B.1.88 y el punto (b) del teorema B.1.91 se puede mostrar que es una extensión finita. En efecto, sea* $\mathbb{E} = \mathbb{K}(u_1, ..., u_{n-1}, u_n)$*, con* u_i *algebraico sobre* $\mathbb{K}$ *para todo* $i = 1, ..., n-1, n$*. Vamos a probar utilizando inducción sobre* n*, el número de elementos algebraicos sobre* $\mathbb{K}$*, que* $\mathbb{E}$ *es una extensión finita de* $\mathbb{K}$*.*

Si $n = 1$*,* $[\mathbb{E} : \mathbb{K}]$ *es finito por el punto (b) del teorema B.1.91. Veamos el caso* $n \geq 2$*. Hipótesis de inducción: supongamos que* $\mathbb{K}(u_1, ..., u_{n-1})$*, con* $u_1, ..., u_{n-1}$ *algebraicos sobre* $\mathbb{K}$*, es una extensión finita de* $\mathbb{K}$*. Hagamos* $\mathbb{L} = \mathbb{K}(u_1, ..., u_{n-1})$*. Por el inciso (b) del teorema B.1.91, se sabe que* $\mathbb{E} = \mathbb{L}(u_n)$ *es una extensión finita al ser* u_n *algebraico sobre* $\mathbb{K}$ *(y, por tanto, sobre* $\mathbb{L}$*). Así, tanto* $[\mathbb{E} : \mathbb{L}]$ *como* $[\mathbb{L} : \mathbb{K}]$ *son finitos. Luego,* $[\mathbb{E} : \mathbb{K}]$ *es finito por el teorema B.1.88. Por lo tanto, por el principio de inducción matemática, se concluye el resultado.*

Problema B.1.102. *Encontrar el campo de descomposición de* $p(x) = x^2 + 2 \in \mathbb{Q}[x]$ *sobre* $\mathbb{Q}$*.*

Solución. El polinomio $p(x)$ se descompone sobre $\mathbb{C}$ pues $p(x) = \left(x - i\sqrt{2}\right)\left(x + i\sqrt{2}\right)$. Sin embargo, $\mathbb{C}$ no es el campo de descomposición de $p(x)$; es suficiente agregar una raíz de $p(x)$ a $\mathbb{Q}$. Claramente, el campo $\mathbb{K} = \mathbb{Q}\left(i\sqrt{2}\right)$ contiene las dos raíces de $p(x)$, y no hay ningún subcampo más pequeño con esta propiedad. Por lo tanto $\mathbb{K}$ es el campo de descomposición de $p(x)$. □

Los campos de descomposición serán fundamentales para dar la caracterización de los campos finitos en la siguiente sección.

B.2. Caracterización de los campos finitos.

En la sección pasada abordamos los ejemplos mas familiares de campos finitos: para cada número primo p, el anillo de clases residuales $\mathbb{Z}/(p)$ es un campo finito con p elementos (ver teorema B.1.42) que puede ser identificado con el campo de Galois $\mathbb{F}_p$ de orden p (ver definición B.1.47). El campo $\mathbb{F}_p$ juega un rol fundamental en la *teoría de campos* ya que, según el teorema B.1.77, todo campo de característica p debe contener una copia isomorfa de $\mathbb{F}_p$ y, en consecuencia, puede pensarse como una extensión de $\mathbb{F}_p$. Esta observación junto con el hecho que todo campo tiene por característica a un número primo (ver corolario B.1.53), es fundamental para clasificar a los campos finitos.

Para comenzar esta sección, enunciamos el siguiente lema que establece una condición necesaria sobre el número de elementos de un campo finito.

Lema B.2.1. *Sea $\mathbb{F}$ un campo finito con un subcampo $\mathbb{K}$ de orden q. Entonces $|\mathbb{F}| = q^m$, donde $m = [\mathbb{F} : \mathbb{K}]$.*

Demostración. $\mathbb{F}$ es un espacio vectorial sobre $\mathbb{K}$ y, como $\mathbb{F}$ es finito, éste es, de hecho, un $\mathbb{K}$–espacio vectorial dimensionalmente finito. Si $m = [\mathbb{F} : \mathbb{K}]$, una base para $\mathbb{F}$ debe consistir en m elementos, digamos $v_1, v_2, ..., v_m$. Ahora bien, cada elementos de $\mathbb{F}$ puede expresarse de manera única como combinación lineal de la base, esto es, puede expresarse como

$$a_1v_1 + a_2v_2 + ... + a_mv_m,$$

donde $a_i \in \mathbb{K}$ para todo $i = 1, 2, ..., m$. Puesto que cada a_i puede tomar q valores, se sigue que $\mathbb{F}$ tiene exactamente q^m elementos. □

Teorema B.2.2. *Sea $\mathbb{F}$ un campo finito. Entonces, $|\mathbb{F}| = p^n$ (p un número primo), donde p es la característica de $\mathbb{F}$ y n es la dimensión de $\mathbb{F}$ sobre su subcampo primo.*

Demostración. Como $\mathbb{F}$ es finito, su característica es un número primo, digamos p, según el corolario B.1.53. En consecuencia, por el teorema B.1.77, el subcampo primo $\mathbb{K}$ de $\mathbb{F}$ es isomorfo a $\mathbb{F}_p$, y, por tanto, éste contiene p elementos. La conclusión del teorema se sigue del lema B.2.1. □

Así, todos los campos finitos deben tener orden igual a una potencia de un primo –no hay un campo finito con 6 elementos, por ejemplo. La siguiente pregunta interesante que surge en nuestro estudio de campos finitos es la siguiente: dado un número primo p, ¿existe siempre un campo finito de orden p^n, para cada potencia p^n? Los siguientes dos lemas permiten responder la pregunta anterior.

Lema B.2.3. *Sea $\mathbb{F}$ un campo con q elementos. Entonces cualquier elemento $a \in \mathbb{F}$ satisface la ecuación $a^q = a$.*

Demostración. Claramente la ecuación $a^q = a$ se satisface para $a = 0$. Por otro lado, los elementos diferentes de cero forman un grupo de orden $q - 1$ bajo la multiplicación. Usando el hecho que $a^{|G|} = 1$ para cualquier elemento a en un grupo finito G, se sigue que $a^{q-1} = 1$ para todo $a \neq 0$ en $\mathbb{F}$. Luego, $a^q = a$. □

Lema B.2.4. *Si $\mathbb{F}$ es un campo finito con q elementos y $\mathbb{K}$ es un subcampo de $\mathbb{F}$, entonces el polinomio $x^q - x \in \mathbb{K}[x]$ se puede factorizar en $\mathbb{F}[x]$ como*

$$x^q - x = \prod_{a \in \mathbb{F}} (x - a), \tag{B.6}$$

y, por tanto, $\mathbb{F}$ es el campo de descomposición de $x^q - x$ sobre $\mathbb{K}$.

Demostración. Como el polinomio $x^q - x$ tiene grado q, éste tiene a lo más q raíces en $\mathbb{F}$. Por el lema B.2.3 todos los elementos de $\mathbb{F}$ son raíces de dicho polinomio, y hay q de ellas. Por lo tanto, el polinomio $x^q - x$ se factoriza como en la ecuaciónB.6, y no se puede descomponer en un subcampo más pequeño. □

Teorema B.2.5 (Existencia y unicidad de campos finitos)**.** *Para cada número primo p y cada entero positivo n existe un campo finito con p^n elementos. Además, cualquier campo finito de orden $q = p^n$ es isomorfo al campo de descomposición de $x^q - x$ sobre $\mathbb{F}_p$.*

Como consecuencia de la unicidad del teorema B.2.5 podemos hablar de *el* campo finito con q elementos (o sea de *el* campo finito de orden q).

Notación B.2.6. *A un campo finito con q elementos se le suele denotar por $\mathbb{F}_q$. Otra notación usual para el campo finito con q elementos es $GF(q)$, donde «GF» hace referencia a «Campo de Galois». Esta notación es en honor a Évariste Galois $(1811-1832)$ quien, en 1830, fue la primera persona en estudiar con seriedad las propiedades generales de los campos finitos.*

Dado un primo p y un entero positivo n, ya vimos que es siempre posible considerar un campo finito de orden p^n. Sin embargo, ¿cómo se puede construir dicho campo? Bueno, partiendo del campo primo $\mathbb{F}_p$ se puede obtener el campo de orden p^n agregando raíces de polinomios: si $p(x) \in \mathbb{F}_p[x]$ es un polinomio irreducible [1] sobre $\mathbb{F}_p$ de grado n, entonces agregando una raíz de $p(x)$ a $\mathbb{F}_p$ (en otras palabras, tomando la extensión de campo $\mathbb{F}_p(\alpha)$ de $\mathbb{F}_p$, con α una raíz de $p(x)$) se obtiene el campo finito con p^n elementos.

Ejemplo B.2.7.

(i) En el ejemplo B.1.95, se construyó un campo $\mathbb{L} = \mathbb{F}_3(\alpha)$ de 9 elementos, donde α es una raíz del polinomio $x^2 + x + 2 \in \mathbb{F}_3[x]$. Por el teorema B.2.5, $\mathbb{L}$ es el *campo finito con nueve elementos, es decir, $\mathbb{F}_9$.*

(ii) En el ejemplo B.1.96, se construyó un campo $\mathbb{L} = \mathbb{F}_2(\alpha)$ de 4 elementos, donde α es una raíz del polinomio $x^2 + x + 1 \in \mathbb{F}_2[x]$. Por el teorema B.2.5, $\mathbb{L}$ es el *campo finito con cuatro elementos, es decir, $\mathbb{F}_4$.*

[1] En principio no es claro si para cada entero positivo n existe un polinomio irreducible en $\mathbb{F}_p[x]$ de grado n. Puede consultarse en (LN97) que lo anterior sí se cumple.

(iii) Consideremos el polinomio $p(x) = x^3 + x + 1 \in \mathbb{F}_2$*, el cual es irreducible sobre* $\mathbb{F}_2$ *ya que no tiene raíces en* $\mathbb{F}_2$*. Sea* α *la raíz* $[x] = x + (p(x))$ *de* $p(x)$*; entonces la extensión simple* $\mathbb{L} = \mathbb{F}_2(\alpha)$ *consiste de ocho elementos:*

$$0, 1, \alpha, \alpha+1, \alpha^2, \alpha^2+1, \alpha^2+\alpha, \alpha^2+\alpha+1.$$

Por el teorema B.2.5, $\mathbb{L}$ *es* el *campo finito con ocho elementos, es decir,* $\mathbb{F}_8$*. El modo de operar los elementos del campo* $\mathbb{F}_8$ *se basa en las tablas B.5 y B.6.*

$+$	0	1	α	$\alpha+1$	α^2	α^2+1	$\alpha^2+\alpha$	$\alpha^2+\alpha+1$
0	0	1	α	$\alpha+1$	α^2	α^2+1	$\alpha^2+\alpha$	$\alpha^2+\alpha+1$
1	1	0	$\alpha+1$	α	α^2+1	α^2	$\alpha^2+\alpha+1$	$\alpha^2+\alpha$
α	α	$\alpha+1$	0	1	$\alpha^2+\alpha$	$\alpha^2+\alpha+1$	α^2	α^2+1
$\alpha+1$	$\alpha+1$	α	1	0	$\alpha^2+\alpha+1$	$\alpha^2+\alpha$	α^2+1	α
α^2	α^2	α^2+1	$\alpha^2+\alpha$	$\alpha^2+\alpha+1$	0	1	α	$\alpha+1$
α^2+1	α^2+1	α^2	$\alpha^2+\alpha+1$	$\alpha^2+\alpha$	1	0	$\alpha+1$	α
$\alpha^2+\alpha$	$\alpha^2+\alpha$	$\alpha^2+\alpha+1$	α^2	α^2+1	α	$\alpha+1$	0	1
$\alpha^2+\alpha+1$	$\alpha+\alpha+1$	$\alpha^2+\alpha$	α^2+1	α^2	$\alpha+1$	α	1	0

Tabla B.5: Tabla de suma para el campo $\mathbb{F}_8$.

$\cdot$	0	1	α	$\alpha+1$	α^2	α^2+1	$\alpha^2+\alpha$	$\alpha^2+\alpha+1$
0	0	0	0	0	0	0	0	0
1	0	1	α	$\alpha+1$	α^2	α^2+1	$\alpha^2+\alpha$	$\alpha^2+\alpha+1$
α	0	α	α^2	$\alpha^2+\alpha$	$\alpha+1$	1	$\alpha^2+\alpha+1$	α^2+1
$\alpha+1$	0	$\alpha+1$	$\alpha^2+\alpha$	α^2+1	$\alpha^2+\alpha+1$	α^2	1	α
α^2	0	α^2	$\alpha+1$	$\alpha^2+\alpha+1$	$\alpha^2+\alpha$	α	α^2+1	1
α^2+1	0	α^2+1	1	α^2	α	$\alpha^2+\alpha+1$	$\alpha+1$	$\alpha^2+\alpha$
$\alpha^2+\alpha$	0	$\alpha^2+\alpha$	$\alpha^2+\alpha+1$	1	α^2+1	$\alpha+1$	α	α^2
$\alpha^2+\alpha+1$	0	$\alpha^2+\alpha+1$	α^2+1	α	1	$\alpha^2+\alpha$	α^2	$\alpha+1$

Tabla B.6: Tabla de multiplicación para el campo $\mathbb{F}_8$.

Bibliografía

[AZ14] Martin Aigner and G. M. Ziegler. *Proofs from the Book*. Springer, 5 edition, 2014. 43, 46

[Bes19] Abram Samoilovitch Besicovitch. Sur deux questions de l'intégrabilité des fonctions. *J. Soc. Phys.–Math.(Perm)*, 2:105–123, 1919. 8, 11

[Bes28] A.S. Besicovitch. On Kakeya's problem and a similar one. *Mathematische Zeitschrift*, 27(1):312–320, 1928. 8, 11

[Bes47] A. S. Besicovitch. On Crum's problem. *J. London Math. Soc.*, 22:285–287, 1947. 32

[Bes63] A.S. Besicovitch. The Kakeya problem. *The American Mathematical Monthly*, 70(7):697–706, 1963. 8, 11

[BKT04] Jean Bourgain, Nets Katz, and Terence Tao. A sum-product estimate in finite fields, and applications. *Geometric & Functional Analysis GAFA*, 14(1):27–57, 2004. 42

[Bou00] J. Bourgain. Harmonic analysis and combinatorics: how much may they contribute to each other. *Mathematics: Frontiers and Perspectives (V. Arnold. et al., eds.)*, pages 13–32, 2000. 41

[Car92] Anthony Carbery. Restriction implies Bochner–Riesz for paraboloids. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 111, pages 525–529. Cambridge Univ Press, 1992. 73

[CJ12] Richard Courant and Fritz John. *Introduction to calculus and analysis II*. Springer Science & Business Media, 2012. 81

[Coh80] Donald L Cohn. *Measure theory*, volume 165. Springer, 1980. 81

[Cun71] F. Cunningham. The Kakeya problem for simply connected and for star-shaped sets. *The American Mathematical Monthly*, 78(2):114–129, 1971. 11

[Dav52] R.O. Davies. On accessibility of plane sets and differentiation of functions of two real variables. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 48, pages 215–232. Cambridge Univ Press, 1952. 73

[Dav71] Roy O. Davies. Some remarks on the Kakeya problem. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 69, pages 417–421. Cambridge Univ. Press, 1971. 41

[dG76] Miguel de Guzmán. Differentiation of integrals in $\mathbb{R}^n$. In *Measure theory*, pages 181–185. Springer, 1976. 11, 73

[DKSS09] Zeev Dvir, Swastik Kopparty, Shubhangi Saraf, and Madhu Sudan. Extensions to the method of multiplicities, with applications to Kakeya sets and mergers. In *Foundations of Computer Science, 2009. FOCS'09. 50th Annual IEEE Symposium on*, pages 181–190. IEEE, 2009. 66, 67

[DL78] Richard A. DeMillo and Richard J. Lipton. A probabilistic remark on algebraic program testing. *Information Processing Letters*, 7(4):193–195, 1978. 57

[Dun06] David L. Duncan. Constructions regarding integration in the plane and the rotation of segments. Undergraduate Senior Thesis, University of Washington, 2006. Disponible en `http://http://ms.mcmaster.ca/~duncand/PlaneIntegration(SeniorThesis).pdf`. 8

[Dvi09] Zeev Dvir. On the size of Kakeya sets in finite fields. *Journal of the American Mathematical Society*, 22(4):1093–1097, 2009. 43, 56, 71, 77

[Dvi12] Zeev Dvir. Incidence theorems and their applications. *arXiv preprint arXiv:1208.5073*, 2012. 43

[Fal86] Kenneth J. Falconer. *The geometry of fractal sets*, volume 85. Cambridge university press, 1986. 11, 41

[Fis73] B. Fisher. On a problem of Besicovitch. *The American Mathematical Monthly*, 80(7):785–787, 1973. 11

[FK17] Matsusaburô Fujiwara and Sōichi Kakeya. On some problems of maxima and minima for the curve of constant breadth and the in-revolvable curve of the equilateral triangle. *Tohoku Mathematical Journal, First Series*, 11:92–110, 1917. 6

[FLS10] Chunrong Feng, Liangpan Li, and Jian Shen. Some inequalities in functional analysis, combinatorics, and probability theory. *The electronic journal of combinatorics*, 17(1):R58, 2010. 80

[Fra03] John B. Fraleigh. *A first course in abstract algebra*. Pearson Education India, 2003. 48, 87

[Fra13] Kevin Scott Fray. *Polynomial Methods in Combinatorial Geometry*. PhD thesis, The University of Melbourne, 2013. 43

[Gut16] Larry Guth. *Polynomial Methods in Combinatorics*, volume 64. American Mathematical Soc., 2016. 43, 75

[Juk11] Stasys Jukna. *Extremal combinatorics: with applications in computer science*. Springer Science & Business Media, 2011. 43

[Kak17] Sōichi Kakeya. Some problems on maxima and minima regarding ovals. *Tôhoku Science Reports*, 6:71–88, 1917. 5

[KT99] Nets Hawk Katz and Terence Tao. Bounds on arithmetic projections, and applications to the Kakeya conjecture. *Mathematical Research Letters*, 6(5/6):625–630, 1999. 43

[Li08] Liangpan Li. On the size of Nikodym sets in finite fields. *arXiv preprint arXiv:0803.3525*, 2008. 75, 79

[Lit53] John Edensor Littlewood. *A mathematicians's miscellany*. Methuen, 1953. 32

[LN97] Rudolf Lidl and Harald Niederreiter. *Finite fields*, volume 20. Cambridge university press, 1997. 87, 103, 115

[LSW16] Ben Lund, Shubhangi Saraf, and Charles Wolf. Finite field Kakeya and Nikodym sets in three dimensions. *arXiv preprint arXiv:1609.01048*, 2016. 66, 80

[Mat15a] Mathologer. The Kakeya needle problem (the squeegee approach). https://www.youtube.com/watch?v=IM-n9c-ARHU, Oct. 2015. [Accesado 4-Ene-2017]. 32

[Mat15b] Pertti Mattila. *Fourier analysis and Hausdorff dimension*, volume 150. Cambridge University Press, 2015. 41, 73

[Mos10] Dana Moshkovitz. An alternative proof of the Schwartz–Zippel lemma. In *Electronic Colloquium on Computational Complexity (ECCC)*, volume 17, page 34, 2010. 56

[MT+04] Gerd Mockenhaupt, Terence Tao, et al. Restriction and Kakeya phenomena for finite fields. *Duke Mathematical Journal*, 121(1):35–74, 2004. 42, 43, 66

[Nic12] W. Keith Nicholson. *Introduction to abstract algebra*. John Wiley & Sons, 2012. 87

[Nik27] Otton Nikodym. Sur la mesure des ensembles plans dont tous les points sont rectilinéairement accessibles. *Fundamenta Mathematicae*, 10(1):116–168, 1927. 73

[Num15] Numberphile. Kakeya's needle problem. https://www.youtube.com/watch?v=j-dce6QmVAQ, Oct. 2015. [Accesado 4-Ene-2017]. 32

[oA62] Mathematical Association of America. The Kakeya problem. http://av-cah.lib.utexas.edu/index.php/Math:E_math_01262, 1962. [Accesado 4-Ene-2017]. 32

[oTv16] Massachusetts Institute of Technology[videosfromIAS]. Unexpected applications of polynomials in combinatorics. https://www.youtube.com/watch?v=SmL0Oz2mFnw&t=1809s, Ago. 2016. [Accesado 5-Ene-2017]. 43

[Pál21] Julius Pál. Ein minimumproblem für ovale. *Mathematische Annalen*, 83(3):311–319, 1921. 6

[Per28] Oskar Perron. Über einen Satz von Besicovitch. *Mathematische Zeitschrift*, 28(1):383–386, 1928. 11

[Rad62] H.A. Rademacher. On a theorem from Besicovitch. *Studies in Mathematical Analysis and Related Topics: Essays in Honor of George Pòlya*, pages 294–296, 1962. 11

[Rog01] Keith McKenzie Rogers. The finite field Kakeya problem. *The American Mathematical Monthly*, 108(8):756–759, 2001. 42, 43, 63

[Rot10] Joseph J. Rotman. *Advanced modern algebra*, volume 114. American Mathematical Soc., 2010. 87

[Sch62] Isaac J. Schoenberg. On the Besicovitch–Perron solution of the Kakeya problem. *Studies in Mathematical Analysis and Related Topics, Pólya*, 30:359–363, 1962. 11

[Sch80] Jacob T. Schwartz. Fast probabilistic algorithms for verification of polynomial identities. *Journal of the ACM (JACM)*, 27(4):701–717, 1980. 57

[SS08] Shubhangi Saraf and Madhu Sudan. An improved lower bound on the size of Kakeya sets over finite fields. *Analysis & PDE*, 1(3):375–379, 2008. 67

[SS09] Elias M. Stein and Rami Shakarchi. *Real analysis: measure theory, integration, and Hilbert spaces*. Princeton University Press, 2009. 41

[Ste16] Elias M. Stein. *Harmonic Analysis (PMS-43): Real-Variable Methods, Orthogonality, and Oscillatory Integrals.(PMS-43)*, volume 43. Princeton University Press, 2016. 11

[T+99] Terence Tao et al. The Bochner–Riesz conjecture implies the restriction conjecture. *Duke mathematical journal*, 96(2):363–376, 1999. 73

[Tao01] Terence Tao. From rotating needles to stability of waves: Emerging connections between combinatorics, analysis and PDE. *Notices of the AMS*, 48(3), 2001. 41

[Tao05] Terence Tao. A new bound for finite field Besicovitch sets in four dimensions. *Pacific journal of mathematics*, 222(2):337–363, 2005. 42

[Tao09] Terence Tao. *Poincaré's legacies: pages from year two of a mathematical blog.* American Mathematical Soc., 2009. 43

[Tao11] Terence Tao. *An introduction to measure theory*, volume 126. American Mathematical Soc., 2011. 81

[Tao13] Terence Tao. Algebraic combinatorial geometry: the polynomial method in arithmetic combinatorics, incidence combinatorics, and number theory. *arXiv preprint arXiv:1310.6482*, 2013. 43

[Tay75] S.J. Taylor. Abram Samoilovitch Besicovitch. *Bull. London Math. Soc*, 7:191–210, 1975. 7

[UoTv16] School of Mathematics University of Toronto[videosfromIAS]. The polynomial method and applications from finite field Kakeya to distinct distances. https://www.youtube.com/watch?v=Zk077i-dokk&t=973s, Ago. 2016. [Accesado 5-Ene-2017]. 43

[vA42] H.J. van Alphen. Uitbreiding van een stelling von Besicovitsch. *Mathematica (Zutphen) B*, 10:144–157, 1942. 11

[Wol99] Thomas Wolff. Recent work connected with the Kakeya problem. *Prospects in mathematics (Princeton, NJ, 1996)*, 2:129–162, 1999. 41, 42, 63

[Zip79] Richard Zippel. *Probabilistic algorithms for sparse polynomials.* Springer, 1979. 57