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

Tesis de Licenciatura

Técnicas para contar puntos enteros y racionales en curvas algebraicas y trascendentes

Marcelo E. Paredes

Director: Román Sasyk

Jueves 15 de agosto del 2013

Índice

Introducción

1. Objetivo del trabajo

Un problema central de la teoría de números es el de resolver una ecuación diofántica: dado un polinomio $P \in \mathbb{Z}[x_1, \cdots, x_n]$ no constante, consideramos

$$P(x_1, \cdots, x_n) = 0. \tag{1.1}$$

Una solución entera de (1.1) es un vector $\mathbf{x} \in \mathbb{Z}^n$ tal que $P(\mathbf{x}) = 0$. De manera análoga, definimos una solución racional de (1.1).

Pregunta 0.1. *¿La ecuación* (1.1) *tiene alguna solución entera?*

Pregunta 0.2. *¿La ecuación* (1.1) *tiene alguna solución racional?*

Pregunta 0.3. *¿Existe alguna forma de calcular todas las soluciones enteras de* (1.1)*? (décimo problema de Hilbert)*

Pregunta 0.4. *¿Existe alguna forma de calcular todas las soluciones racionales de* (1.1)*?*

Las preguntas anteriores son naturales, aunque para una ecuación arbitraria como (1.1), son muy complicadas. Por ejemplo, si consideramos $P(x, y, z) = x^n + y^n - z^n$, la ecuación (1.1) se convierte en

$$x^n + y^n = z^n. \tag{1.2}$$

Tenemos soluciones triviales como $(0, 0, 0)$ ó $(x, 0, x)$. ¿Pero hay otras soluciones? En el caso $n = 2$ la respuesta no es difícil; una tal solución se conoce como terna pitagórica y tales ternas admiten una representación sencilla, de la que se deduce que hay (infinitas) soluciones no triviales. Para $n > 2$, esta pregunta es el célebre último teorema de Fermat, cuya demostración es sumamente complicada.

Consideramos ahora $P(x, y, z) = x^2 + y^2 + z^2 + 1$. La ecuación (1.1) se convierte en

$$x^2 + y^2 + z^2 = -1, \tag{1.3}$$

con lo que (1.3) no tiene siquiera soluciones racionales.

Se sabe que no existe un algoritmo para hallar soluciones enteras de $P \in \mathbb{Z}[x_1, \cdots, x_n]$ arbitrario ([**33**]). Esto dice que la Pregunta 0.3 tiene respuesta negativa si se aspira a encontrar una manera general de resolver ecuaciones diofánticas. Sin embargo, permanece abierta la pregunta análoga para soluciones racionales, con lo que la Pregunta 0.4 permanece abierta.

La Pregunta 0.4 es particularmente interesante, pues en algunos casos es posible obtener soluciones racionales a partir del conocimiento de alguna solución. Por ejemplo, si $P \in \mathbb{Q}[x, y]$ define una curva elíptica, el conocimiento de una solución racional permite calcular todas las soluciones racionales de $P(x, y) = 0$.

En esta tesis vamos a estudiar preguntas relacionadas a las anteriores que permiten obtener información de una ecuación diofántica. Estas preguntas son:

Pregunta 0.5. *¿Hay alguna estimación para la cantidad de soluciones enteras de la ecuación* (1.1) *en una región acotada, digamos,* $[0, r]^n$*?*

Pregunta 0.6. *¿Hay alguna estimación para la cantidad de soluciones racionales de altura a lo sumo* H *de la ecuación* (1.1)*?*

La noción de altura que vamos a considerar es la siguiente. Dado $x \in \mathbb{Q}$, si $x = \frac{a}{b}$ con $(a : b) = 1$ y $b > 0$, definimos $H(x) = \max\{|a|, b\}$. Si $P = (x_1, \cdots, x_n) \in \mathbb{Q}^n$, definimos $H(P) := \max_{1 \leq i \leq n}\{H(x_i).\}$

Estas preguntas están relacionadas con el estudio de la distribución que tienen los conjuntos de soluciones enteras y racionales de una ecuación diofántica. Para mostrar por qué responder estas preguntas da información útil, damos algunos ejemplos, extraídos de la exposición de Heath-Brown [**14**, pág. 51-57].

EJEMPLO 0.1. *Sea* $P(x_1, x_2, x_3) = x_1^k + x_2^k + x_3^k - N$. *Tenemos la ecuación*

$$x_1^k + x_2^k + x_3^k = N, x_1, x_2, x_3 \geq 0. \tag{1.4}$$

Para $k \leq 6$ *se sabe que hay infinitos* N *para los que* (1.4) *tiene infinitas soluciones "esencialmente distintas"(no son permutaciones unas de otras). Para* $k \geq 7$ *en cambio se cree que hay a lo sumo una solución, salvo permutaciones. Este problema está relacionado con el famoso problema de Waring, en el que se pregunta cuál es el menor* r *para el que todo* N *suficientemente grande se escribe como suma de* r *potencias* k*-perfectas.*

EJEMPLO 0.2. *En virtud del ejemplo anterior, se puede estudiar la ecuación*

$$x_1^k + \cdots + x_s^k = x_{s+1}^k + \cdots + x_{2s}^k,\ 0 \leq x_1, \cdots, x_{2s} \leq N. \tag{1.5}$$

En este caso, la pregunta es "de cuántas maneras se puede escribir un número como suma de potencias k*-perfectas". Este problema fue estudiado por Hardy-Littlewood, por medio del método circular. Para un tratamiento clásico de este problema, consultar la exposición de Davenport en* [**11**].

EJEMPLO 0.3. *El valor medio de Vinogradov, relacionado con el valor promedio de una suma exponencial, puede ser estimado mediante el sistema de ecuaciones diofánticas*

$$x_1^h + \cdots x_s^h = x_{s+1}^h + \cdots x_{2s}^h,\ \ (1 \leq h \leq k), \tag{1.6}$$

con $0 \leq x_i \leq B$ *para todo* $1 \leq i \leq 2s$. *Se sabe que si* s *es suficientemente grande, el número de soluciones enteras del sistem* (1.6) *es a lo sumo* $cB^{2s-\frac{k(k+1)}{2}}$, *con* $c > 0$ *una constante. Tales cotas tienen numerosas aplicaciones, como dar estimaciones de la región no nula de la función zeta de Riemann. Una heurística por medio del método circular de Hardy-Littlewood lleva a conjeturar que la cota mencionada debería valer para* $s > \frac{k(k+1)}{2}$. *Si esta conjetura fuera cierta, se obtendrían resultados nuevos sobre la función zeta.*

EJEMPLO 0.4. *Se conjetura que todo polinomio irreducible* $f \in \mathbb{Z}[X]$ *que satisface cierta condiciones de congruencia debería asumir infinitos valores que son libres de cuadrados. Esto se sabe para polinomios de grado a lo sumo* 3. *Consideramos entonces* $P(x, y, z) = f(x) - y^2 z$. *Tenemos la ecuación*

$$f(x) = y^2 z, 1 \leq x \leq N,\ y \geq N. \tag{1.7}$$

Tener una buena cota para la cantidad de soluciones enteras de (1.7) *permitiría avanzar en la conjetura mencionada al inicio del ejemplo.*

Existe una pregunta más, que surge de manera natural, que es preguntarse si las estimaciones que se desean se pueden hacer uniformes, esto es, hallar una cota uniforme para la cantidad de soluciones de todas las ecuaciones diofánticas definidas por un polinomio de grado total fijo. Entonces, volvemos a considerar $P \in \mathbb{Z}[x_1, \cdots, x_n]$ no constante, pero ahora de grado total d.

PREGUNTA 0.7. *¿Hay alguna estimación uniforme sobre los polinomios* P *para la cantidad de soluciones enteras de* (1.1) *en una región acotada, digamos,* $[0, r]^n$*?. Con un poco más de precisión, nos preguntamos, por ejemplo, si existe una constante* $c = c(\deg(P))$, *dependiente sólo de* $\deg(P)$, *y una función* f *tal que*

$$\#\{\boldsymbol{x} \in \mathbb{Z}^n : P(\boldsymbol{x}) = 0\} \cap [0, r]^n\} \leq c(\deg(P)) f(r). \tag{1.8}$$

De manera análoga, reformulamos la Pregunta 0.7 pero para puntos racionales de altura a lo sumo H. Esta tesis trata la Pregunta 0.7, principalmente en el caso que P es un polinomio de dos variables. Ya este caso tiene ejemplos interesantes, como son las ecuaciones

$$y^2 = x^3 + Ax + B, \tag{1.9}$$

con $A, B \in \mathbb{Z}$ (curvas elípticas en forma de Weierstrass). También se tienen las ecuaciones

$$y^n = f(x), \tag{1.10}$$

con $f \in \mathbb{Z}[x]$ que no es una potencia n-ésima de algún polinomio entero, y tal que su discriminante es no nulo (curvas hiperelípticas). Un último ejemplo es el de la ecuación de Pell

$$x^2 - ny^2 = 1, \tag{1.11}$$

con $n \in \mathbb{N}$ un número que no es un cuadrado perfecto.

Para estudiar las Preguntas 0.5, 0.6 y 0.7, observamos que el problema que nos interesa, de naturaleza aritmética, puede plantearse de manera geométrica, pues una ecuación $P(x, y) = 0$ en $\mathbb{R}^2$ define (siempre que tenga alguna solución) una curva algebraica plana C. Entonces la Pregunta 0.5 se traduce en estimar la cantidad de puntos enteros de una curva.

PREGUNTA 0.8. *¿Hay alguna estimación para la cantidad de puntos enteros de una curva algebraica plana C en una región acotada, digamos, $[0, r]^2$?*

La ventaja de este planteo es que ahora es posible implementar técnicas geométricas para obtener información sobre C.

La Pregunta 0.8 motiva una pregunta opuesta, que también estudiamos en esta tesis:

PREGUNTA 0.9. *¿Hay alguna estimación para la cantidad de puntos enteros de una curva plana "no muy algebraica"?*

Por ejemplo, una curva plana "fuertemente no algebraica" sería $\{(t, \exp(t)) \in \mathbb{R}^2 : t \in \mathbb{R}\}$, el gráfico de la función exponencial.

Esta tesis estudia las Preguntas 0.5, 0.6, 0.7, 0.8, principalmente siguiendo el trabajo de Bombieri y Pila [**6**] y Helfgott y Venkatesh [**18**]. En referencia a la Pregunta 0.7, Bombieri-Pila desarrollan el método del determinante con el que dan el resultado:

TEOREMA (Bombieri-Pila). *Sea $C \subseteq \mathbb{R}^2$ una curva algebraica irreducible de grado d y R un cuadrado de lado N. Para todo $\varepsilon > 0$, se tiene que existe $c(d, \varepsilon)$ tal que*

$$\#(C \cap R \cap \mathbb{Z}^2) \leq c(d, \varepsilon) N^{\frac{1}{d}+\varepsilon},$$

donde la constante $c(d, \varepsilon)$ es efectiva.

Helfgott-Venkatesh [**18**] obtienen la estimación del teorema anterior, utilizando el método de Bombieri-Pila, pero mediante una estrategia distinta. En el proceso, desarrollan un método de criba en el plano, que les permite obtener el siguiente resultado, de interés propio:

TEOREMA (Helfgott-Venkatesh). *Sea $S \subseteq [N]^2 \subseteq \mathbb{Z}^2$ con $N \geq 1$. Supongamos que existe $\alpha > 0$ tal que*

$$\#\{(x \pmod p, y \pmod p) : (x, y) \in S\} \leq \alpha p,$$

para todo primo p. Entonces, para todo $\varepsilon > 0$ se tiene que ocurre alguna de las proposiciones siguientes:

- *Existe una constante $c = c(\alpha, \varepsilon) > 0$ tal que $\#S \leq c(\alpha, \varepsilon) N^{\varepsilon}$*
- *existe una curva algebraica plana C de grado $O_{\alpha,\varepsilon}(1)$ tal que al menos $(1 - \varepsilon)\#S$ puntos de S se encuentran en C*

Para la Pregunta 0.9, estudiamos el caso que Γ sea el gráfico de una función analítica trascendente. Tenemos el resultado de Bombieri-Pila, también deducido del método del determinante:

TEOREMA (Bombieri-Pila). *Sea $f : [0, 1] \to \mathbb{R}$ una función analítica trascendente. Sea Γ el gráfico de f. Para todo $t \geq 1$ y $\varepsilon > 0$, existe $r(f, \varepsilon)$ tal que*

$$\#((t\Gamma) \cap \mathbb{Z}^2) \leq r(f, \varepsilon) t^{\varepsilon}.$$

Estudiar los puntos de intersección de curvas "no algebraicas", o equivalentemente, estudiar curvas que tienen pocos puntos de intersección con curvas algebraicas racionales, es importante para muchos problemas diofánticos, como la conjetura de Manin-Mumford y la conjetura de André-Oort. Pila en [**40**] da una demostración incondicional de la conjetura de André-Oort para producto de curvas modulares, mostrando que se pueden estudiar las curvas "no algebraicas" mediante el estudio de curvas definidas en alguna estructura O-minimal. Por esta razón, nos van a interesar las estructuras O-minimales, y vamos a estudiar el teorema de Pila-Wilkie, que en última instancia es una generalización adecuada de la Pregunta 0.9:

Teorema (Pila-Wilkie). *Sea $X \subseteq \mathbb{R}^n$ un conjunto definible en $\mathcal{L}$, una estructura O-minimal que contiene la suma y el producto de los números reales. Sea $\varepsilon > 0$. Se tiene la estimación*

$$(X - X^{\mathrm{alg}})(\mathbb{Q}, H) = O_{X,\varepsilon}(T^{\varepsilon}).$$

Para entender el teorema de Pila-Wilkie, introducimos el lenguaje de las estructuras O-minimales y analizamos las propiedades que tienen. Finalizamos explicando cómo se conecta el estudio de curvas "no algebraicas", o en general, el estudio de conjuntos definidos en una estructura O-minimal, con estudiar un problema diofántico. Concretamente, estudiamos la estrategia de Pila-Zannier de [**43**] aplicada al teorema de Mann:

Teorema (Mann, 1965). *Sea $n \geq 1$ un entero positivo. Sea $G(x_1, \cdots, x_l) \in \mathbb{C}[x_1, \cdots, x_l]$ un polinomio no nulo. Entonces el conjunto de puntos torsión de $\{(x_1, \cdots, x_l) \in \mathbb{C}^l : G(x_1, \cdots, x_l) = 0\}$ es una unión finita de coclases de subgrupos de $\mathbb{C}^{\times l}$.*

2. Estructura de la tesis

El trabajo central para esta tesis es el trabajo de Bombieri y Pila [**6**], en el que desarollan el "método del determinante". Por este motivo, hemos decidido aislar la parte "métodica" de [**6**] y los resultados que se obtienen como parte de este método. De esta forma, el capítulo 1 desarrolla la estrategia general del método del determinante, y el capítulo 2 muestra aplicaciones del método del determinante para estimar soluciones enteras de curvas; estas aplicaciones son del trabajo [**6**]. Se presentan dos casos en algún sentido extremos: curvas analíticas trascendentes, y curvas algebraicas. También se esbozan algunas ideas de qué cosas se pueden intentar hacer con el método del determinante en dimensiones superiores, notando también cómo estos primeros intentos tienen limitaciones. Concluímos el capítulo 2 esbozando algunos resultados para dimensiones superiores, de Heath-Brown [**17**], que estudian ecuaciones diofánticas dadas por formas homogéneas.

La idea subyacente del método del determinante es decidir cuándo una familia finita de puntos está contenida en una curva algebraica controlada (grado acotado por algunos parámetros; definida de formas particulares). Esta idea se puede traducir en detectar una noción de estructura algebraica de un conjunto de puntos. Esta interpretación del método es la que se estudia en el capítulo 3, en el que se prueba, siguiendo los argumentos de [**18**], que los conjuntos que están "mal distribuidos" y son "grandes" tienen algún tipo de estructura algebraica: están contenidos de manera controlada en curvas algebraicas. Este resultado en sí es interesante, pues da una suerte de resultado recíproco a la estimación de Lang-Weil. Desde el punto de vista de esta tesis, el teorema de Lang-Weil dice que las curvas (y las hipersuperficies) algebraicas son conjuntos mal distribuidos. Dado que el argumento de Helfgott-Venkatesh se basa esencialmente en desarrollar una criba para los puntos del plano, introducimos en este capítulo las ideas básicas de los métodos de criba.

En el capítulo 4 estudiamos el problema de estimar los puntos racionales "pocos algebraicos" de curvas trascendentes, en un contexto más general, que es el de las estructuras O-minimales. Por esta razón, desarrollamos la teoría de las estructuras O-minimales para poder estudiar en general los puntos racionales "poco algebraicos" de superficies trascendentes. El objetivo principal de este capítulo es esbozar la idea detrás del teorema de Pila-Wilkie [**42**] y sus aplicaciones a la geometría diofántica, concretamente, a la conjetura de Manin-Mumford y a la conjetura de André-Oort. Debido a que estas conjeturas son sumamente profundas, estudiamos casos muy particulares, concretamente, el teorema de Mann, siguiendo la exposición de Scanlon de [**46**].

3. Notación

La notación que empleamos es en general estándar. Para A un conjunto finito, $\#A$ denotará el cardinal del conjunto A. El máximo común divisor y mínimo común múltiplo van a ser denotados $\gcd, \mathrm{lcm}$. Dado $x \in \mathbb{Q}^k$, vamos a denotar $H(x)$ a la función $H : \mathbb{Q}^n \to \mathbb{R}$ que se define como

$$H(\frac{a}{b}) := \max\{|a|, b\} \text{ si } a, b \in \mathbb{Z}, b > 0 \text{ y } \gcd(a, b) = 1.$$
$$H(x_1, \cdots, x_n) = \max_{1 \leq i \leq n}\{H(x_i)\}.$$

Si $X \subseteq \mathbb{R}^k$, vamos a escribir

$$X(\mathbb{Z}) := X \cap \mathbb{Z}^k.$$
$$X(\mathbb{Q}, H) := \{x \in (X \cap \mathbb{Q}^k) : H(x) \leq H\}.$$

De la misma forma, si $X \subseteq \overline{\mathbb{F}_q}^k$ con $\mathbb{F}_q$ un cuerpo de q elementos y $\overline{\mathbb{F}_q}$ su clausura algebraica, vamos a denotar

$$X(\mathbb{F}_{q^n}) := \{x \in (X \cap \mathbb{F}_{q^n}^k)\}.$$

Las dilataciones de un conjunto X son una forma sencilla de perturbar al conjunto. Para $t > 0$, notamos

$$tX := \{t\mathbf{x} : \mathbf{x} \in X\}.$$

En muchos casos vamos a utilizar estimaciones de cotas, con una notación más comprimida. Si $f, g : \mathbb{R} \to \mathbb{R}$ son funciones con g positiva, si se verifica para todo $x \in \mathbb{R}$

$$|f(x)| \leq Cg(x),$$

notamos $f = O(g)$. En general, las funciones pueden usar varios parámetros y las cotas tener dependencias de algunos de los parámetros. Por ejemplo, si

$$|f(x, y, n)| \leq 2^n g(x, y),$$

en este caso la constante C de antes, ahora 2^n, depende de n. Notamos $f = O_n(g)$. Así en general, si la constante que mayora a f contiene parámetros, se los agregamos a la notación como subíndice. También en algunos casos

$$|f(x)| \leq Cg(x),$$

lo vamos a notar $f \ll g$, y como recién, si la constante depende de parámetros a, ε, n o los que fuera, los agregamos como subíndices, por ejemplo $f \ll_{a,\varepsilon,n} g$ denota que existe $C = C(a, \varepsilon, n)$ (una función de estos parámetros) tal que

$$|f(x)| \leq C(a, \epsilon, n)g(x).$$

Usamos la notación $f \asymp g$ para describir dos funciones positivas tales que se comportan asintóticamente iguales, esto es,

$$\lim_{x\to+\infty} \frac{f(x)}{g(x)} = 1.$$

Dada una función $f \in \mathcal{C}^k(I)$ con I un intervalo cerrado y acotado, para cada par N, k de enteros positivos, denotamos

$$||f||_{N,k} := \max_{1\leq\kappa\leq k, x\in I} N^{\kappa-1} \frac{|f^{(\kappa)}(x)}{\kappa!}.$$

Para notar al conjunto de números naturales del 1 al N vamos a usar la notación $[N]$. Dado un conjunto $S \subseteq \mathbb{Z}^d$, para p primo notamos

$$S_p := \{(x_1, \cdots, x_d) \pmod p : (x_1, \cdots, x_d) \in S\}.$$

Estamos denotando $(x_1, \cdots, x_k) \equiv (y_1, \cdots, y_k) \pmod p$ a la relación $x_l \equiv y_l \pmod p$ para todo $1 \leq l \leq k$.

Vamos a denotar 1_A a la función indicadora de A, es decir, la función que en x vale 1 si $x \in A$ y 0 si no. Con la notación $\text{Prob}, \mathbb{E}$ vamos a denotar respectivamente a la probabilidad y la esperanza de una variable aleatoria.

En general, para distinguir entre escalares y vectores, utilizamos x para denotar un escalar y $\mathbf{x}$ para denotar un vector.

4. Agradecimientos

Acaso como una costumbre, tradición incuestionable, parece ser necesaria la sección de Agradecimientos en las tesis de licenciatura. No sé hasta qué punto unas palabras en un papel (o, para ser precisos, en un archivo pdf) pueden transmitir agradecimientos. De cualquier forma, he cedido a escribir esta sección, puesto que sé, por experiencia propia, que es la sección más fácil de leer de una tesis de licenciatura (¡no necesariamente es la sección más fácil de escribir!). En vistas de esto, espero que las personas que prosigo a agradecer, puedan detenerse en esta sección, y, todavía con un poco más de fé, espero puedan percibir los agradecimientos que les debo (y no se detengan a criticar mi estilo complicado de redacción, probablemente poco óptimo y no suficientemente claro).

Les doy las gracias:

- A toda mi familia, por las enseñanzas de vida que continuamente me dan, por su continua atención, por todo.
- A mis amigos de la facu, tanto con los que cursé muchas materias, como con los que usualmente tengo conversaciones, con los que usualmente aprendo cosas nuevas, con los que me usualmente me junto. En especial, a Aye, Dani, Di, Jaz, Juan, Eli, Facu, Fede, Lu, Juli, Kari, Maru, Maxi, Rafa, Rocha, Sofi, Quimey, Yami, Xime.
- A mis jtps y docentes, en especial a Norberto, Pablo, Gabriel, Mariano, Charly, Román, que entre otras cosas, no sólo me enseñaron matemática, si no que también me transmitieron ganas de seguir estudiando matemática.
- A mis jurados Ariel y Teresa, que se leyeron toda la tesis, y me dieron numerosas correcciones importantes, que me orientaron a mejorar y me hicieron volver a leer la tesis de nuevo... lo que fue duro, pero las correcciones lo ameritaron. Muchas gracias.
- A Román, por los numerosos consejos, por las numerosas charlas, por mostrar una perspectiva amplia, quizás, lo menos local posible, de lo que es la matemática.
- A mis amigos del club Touche, con los que practico esgrima. A Guille, gracias por continuar enseñándome la bella disciplina que es la esgrima.

Existe otra razón más por la cuál opté por escribir una sección de agradecimientos, motivados por la siguiente analogía: en una adivinanza cuyo tema es el ajedrez, ¿cuál es la única palabra prohibida? la palabra ajedrez. En cualquier caso, estos agradecimientos están basados en esa única palabra que no dije hasta ahora: “pasión”. Todos los que hicieron posible esta tesis, para bien o para mal, transmitieron esa sensación difícil de describir, pero fácil de sentir.

Y una última razón más para estos agradecimientos es, probablemente, divertirme imaginando las injurias, protestas, los ”inmoral”, ”deshonesto”, ”cruel” y malos calificativos, ¡que podés estar diciendo mientras leíste los agradecimientos y no encontraste tu nombre! Gracias por todo, amiga mía, Morita.

CAPÍTULO 1

El método del determinante

En geometría, el determinante es un objeto algebraico que sirve para modelar el concepto de "integrar" sobre superficies de dimensiones superiores a 2. Esto es debido a que permite dar una forma de "detectar" cuándo un conjunto de vectores es linealmente independiente, lo que se traduce en detectar puntos que se encuentran en posiciones "degeneradas" en el espacio. Concretamente, el determinante en, digamos, $\mathbb{R}^n$, mide si $v_1, \cdots, v_n \in \mathbb{R}^n$ son los vértices de un paralelepípedo de dimensión n o no. Esta forma de "detección" es buena, debido a que el determinante es una función multilineal (alternada). Por "buena" nos estamos refiriendo a que como objeto algebraico es relativamente sencillo.

La idea de usar el determinante como un objeto que detecta "posiciones concretas" de un conjunto es bastante común. El método del determinante de Bombieri-Pila tiene el mismo principio: "detectar cuándo un conjunto de puntos se encuentra contenido en una misma estructura algebraica, sujeta a algunas condiciones". En el caso que la estructura algebraica sea, por ejemplo, una curva algebraica, la pregunta tiene bastante sentido, pues el problema de decidir si un conjunto está contenido en una curva algebraica, digamos, de grado a lo sumo d, es equivalente a preguntarse si cierto sistema de ecuaciones lineales tiene alguna solución (que no sea la nula). En la práctica esto se traduce a considerar una matriz $(a_{ij}(x))_{ij}$ que contiene información acerca de la estructura algebraica de los puntos a estudiar. El método luego busca establecer una cota del estilo

$$|\det(a_{ij}(x))_{i,j}| \leq D. \tag{0.1}$$

Utilizando algún dato adicional, como que los $a_{ij}(x)$ son puntos enteros, se tiene que $\det(a_{ij})_{ij} \in \mathbb{Z}$ con lo que si $D < 1$ resulta $\det(a_{ij}(x))_{ij} = 0$. Teniendo la estimación (0.1) se obtiene:

Teorema (Bombieri-Pila). *Sea $\mathcal{M}$ un conjunto finito de monomios en x, y. Decimos que C es una curva algebraica definida en $\mathcal{M}$ si existe un polinomio $P \in \mathbb{R}[x, y]$ cuyos monomios son elementos de $\mathcal{M}$. Sean*

$$D := \#\mathcal{M}, \quad J = \{j = (j_1, j_2) : x^{j_1} y^{j_2} \in \mathcal{M}\},$$
$$p = \sum_{j=(j_1,j_2)\in J} (j_1 + j_2), \quad q = \sum_{j=(j_1,j_2)\in J} j_2.$$

Sea $f \in \mathcal{C}^{D-1}([0, N])$ y sea Γ el gráfico de f. Se tiene que $\Gamma(\mathbb{Z})$ está contenido en la unión de no mas de

$$\left(D^p ||f||^q_{N,D-1}\right)^{\frac{2}{D(D-1)}} N^{\frac{2p}{D(D-1)}} + 1$$

curvas algebraicas definidas en $\mathcal{M}$.

El Teorema 1 resulta útil, pues en varios casos es posible determinar explicitamente la cantidad de puntos enteros comunes de una curva suave Γ y una curva algebraica, lo que, combinado con el Teorema 1 nos permite obtener una estimación de $\#\Gamma(\mathbb{Z})$.

1. Resultados preliminares

Sean x, x_i, y_{ij} para $i, j = 1, \cdots n$ indeterminadas y sea $V(x_1, \cdots, x_k)$ el determinante de la matriz de Vandermonde en $x_1, \cdots, x_k$, es decir

$$V(x_1, \cdots, x_k) = \prod_{1\leq i<j\leq k} (x_i - x_j)$$

Sea

$$g_{ij}(x) := \frac{-1}{V(x_1,\cdots,x_i)} \det \begin{pmatrix} 1 & x_1 & \cdots & x_1^{i-1} & y_{1j} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 1 & x_i & \cdots & x_i^{i-1} & y_{ij} \\ 1 & x & \cdots & x^{i-1} & 0 \end{pmatrix} \tag{1.1}$$

el polinomio interpolador de grado a lo sumo $i-1$ de los puntos $\{y_{kj}\}$ en los nodos $x_1,\cdots,x_i$, es decir, $g_{ij} \in \mathbb{R}[x]$ es el único polinomio de grado a lo sumo $i-1$ que verifica $g_{ij}(x_l) = y_{lj}$ para todo j. Esto puede corroborarse considerando el determinante

$$\det \begin{pmatrix} 1 & x_1 & \cdots & x_1^{i-1} & y_{1j} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 1 & x_i & \cdots & x_i^{i-1} & y_{ij} \\ 1 & x & \cdots & x^{i-1} & g_{ij}(x) \end{pmatrix}$$

que se anula evaluando en $x_1,\cdots,x_i$ y notando que

$$\det \begin{pmatrix} 1 & x_1 & \cdots & x_1^{i-1} & y_{1j} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 1 & x_i & \cdots & x_i^{i-1} & y_{ij} \\ 1 & x & \cdots & x^{i-1} & g_{ij}(x) \end{pmatrix} = V(x_1,\cdots,x_i)g_{ij}(x) + \begin{pmatrix} 1 & x_1 & \cdots & x_1^{i-1} & y_{1j} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 1 & x_i & \cdots & x_i^{i-1} & y_{ij} \\ 1 & x & \cdots & x^{i-1} & 0 \end{pmatrix}.$$

La expresión (1.1) del polinomio interpolador nos permite obtener:

PROPOSICIÓN 1.1. *Con las notaciones anteriores, se tiene que*

$$\det(y_{ij}) = \frac{V(x_1,\cdots,x_n)}{1!\cdots(n-1)!} \det(g_{ij}^{(i-1)}).$$

DEMOSTRACIÓN. Primero observemos que $g_{ij}(x)$ nos da, desarrollando el determinante por la última fila:

$$g_{ij}(x) = \frac{-1}{V(x_1,\cdots,v_i)} \sum_{1\le j\le i} (-1)^{i+j} C_j x^{j-1},$$

donde los C_j son los cofactores correspondientes, que no dependen de x. Luego, derivando $i-1$ veces respecto de x, sólo queda el coeficiente de x^{i-1}:

$$g_{ij}^{(i-1)}(x) \equiv g_{ij}^{i-1} = \frac{-1}{V(x_1,\cdots,x_i)}(i-1)!(-1)^{i+i}C_i = \frac{-1}{V(x_1,\cdots,x_i)}(i-1)! \det \begin{pmatrix} 1 & x_1 & \cdots & x_1^{i-2} & y_{1j} \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ 1 & x_i & \cdots & x_i^{i-2} & y_{ij} \end{pmatrix}.$$

Si escribimos $\hat{x_k}$ para denotar que la variable x_k es omitida, desarrollando el último determinante por la última columna obtenemos

$$g_{ij}^{(i-1)} = \frac{-1}{V(x_1,\cdots,x_i)}(i-1)! \sum_{1\le k\le i} (-1)^{k-i} V(x_1,\cdots,\hat{x_k},\cdots,x_i) y_{kj}. \tag{1.2}$$

Observemos que la ecuación (1.2) se puede escribir como un producto de matrices $(a\cdot y)_{ij}$, con $y = (y_{ij})$ y $a = (a_{ij})$ triangular inferior

$$a_{ij} = (-1)^{i-j} V(x_1,\cdots,\hat{x_j},\cdots,x_i) 1_{i\ge j},$$

donde $1_{i\ge j}$ denota la función caractéristica del conjunto $\{(i,j) : i \ge j\}$. Lo que acabamos de obtener es la expresión

$$\frac{V(x_1,\cdots,x_n)}{1!\cdots(n-1)!} \det(g_{ij}^{(i-1)}) = \frac{V(x_1,\cdots,x_n)}{1!\cdots(n-1)!} \det(\frac{-1}{V(x_1,\cdots,x_i)}(i-1)!(a\cdot y)_{ij}),$$

con lo que para concluir la proposición nos basta probar que se tiene

$$\frac{V(x_1,\cdots,x_n)}{1!\cdots(n-1)!}\det(\frac{-1}{V(x_1,\cdots,x_i)}(i-1)!(a\cdot y)_{ij})=\det(y_{ij}). \tag{1.3}$$

Utilizando la multilinealidad del determinante, en particular, la propiedad

$$\det(A\cdot B)=\det(A)\det(B),$$

se deduce la igualdad (1.3) y por lo tanto la proposición. □

Es de interés admitir que los coeficientes y_{ij} sean funciones, dependientes de un parámetro. Vamos a suponer entonces que $y_{ij}:=f_j(x_i)$ con $f_j\in\mathcal{C}^{n-1}([a,b])$ y los $x_1,\cdots,x_n$ son n puntos distintos del intervalo $[a,b]$. Escribiendo

$$\det(g_{ij}^{(i-1)})=\sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)\prod_{1\leq k\leq n}g_{k\sigma(k)}^{(k-1)}, \tag{1.4}$$

queremos usar la identidad algebraica probada en la Proposición 1.1 para poder acotar. Dado que los g_{ij} eran los polinomios interpoladores de grado a lo sumo $i-1$ de puntos concretos, vamos a recordar un resultado clásico que involucra dichos polinomios.

Proposición 1.2. *Dado $n>0$ un entero positivo, sean $x_0<x_1<\cdots<x_n$ $n+1$ números reales distintos y sea f una función n veces diferenciable en $[x_0,x_n]$. Sea g el polinomio interpolador de grado a lo sumo n de f en los puntos x_i y llamemos a al coeficiente principal de g. Se tiene que existe $\xi\in(x_0,x_n)$ que verifica $f^{(n)}(\xi)=n!a$*

Demostración. Como g tiene grado a lo sumo n, se tiene que $g^{(n)}(x)=n!a$ para todo x. Sea $h(x):=f(x)-g(x)$. Dado que h se anula en los x_i, luego en $n+1$ puntos distintos, tenemos en virtud del teorema de Rolle que h' se anula en n puntos distintos del intervalo (x_0,x_n). Procediendo así, concluímos que $h^{(n)}$ se anula en algún punto $\xi\in(x_0,x_n)$. Entonces $f^{(n)}(\xi)=g^{(n)}(\xi)=n!a$. □

Dado que g_{ij} es el polinomio interpolador de grado a lo sumo $i-1$ de f_j en los puntos $x_1,\cdots,x_i$, la proposición anterior nos dice que existe $\xi_j\in(0,N)$ tal que

$$g_{ij}^{(i-1)}={f_j}^{(i-1)}(\xi_j).$$

Esto en particular nos permite deducir la siguiente proposición, que es un corolario inmediato de la Proposición 1.1.

Proposición 1.3. *Sean $f_1,\cdots f_n\in\mathcal{C}^{n-1}([a,b])$ y sean $x_1,\cdots,x_n$ puntos distintos de $[a,b]$. Entonces existen puntos intermedios $(\xi_{ij})_{ij}$ tales que*

$$\det(f_j(x_i)_{ij})=\frac{V(x_1,\cdots,x_n)}{1!\cdots(n-1)!}\det(f_j^{(i-1)}(\xi_{ij}))_{ij}.$$

Volviendo a la expresión (1.4), como $\#S_n=n!$, se tiene

$$|\det(g_{ij}^{(i-1)})|\leq\sum_{\sigma\in S_n}\prod_{1\leq k\leq n}\frac{|f_{\sigma(k)}^{(k-1)}(\xi_k)|}{(k-1)!}\leq n!\max_{\sigma\in S_n}\prod_{1\leq k\leq n}\frac{|f_{\sigma(k)}^{(k-1)}(\xi_k)|}{(k-1)!}. \tag{1.5}$$

Para escribir de manera un poco más compacta (1.5), introducimos una norma adecuada sobre $\mathcal{C}^k([a,b])$.

Definición 1.1. *Sea I un intervalo cerrado y acotado. Sean N,k enteros positivos. Dada $f\in\mathcal{C}^k(I)$ definimos*

$$||f||_{N,k}:=\max_{1\leq\kappa\leq k,x\in I}N^{\kappa-1}\frac{|f^{(\kappa)}(x)|}{\kappa!}.$$

Notemos que $||\cdot||_{N,k}$ define una norma en $\mathcal{C}^k(I)$. Ahora, sea $f\in\mathcal{C}^k(I)$ y consideremos $f_t:tI\to\mathbb{R}$ definida por

$$f_t(x):=tf(\frac{x}{t}).$$

Se verifica que $f_t\in\mathcal{C}^k(tI)$. Por otro lado, haciendo el cambio de variables $x\mapsto\frac{x}{t}$ tenemos que $||f_t||_{tN,k}=||f||_{N,k}$. Esto nos dice que la norma es "invariante por dilataciones". Notemos en particular, que para $f(x):=x$, tenemos $f'(x)=1, f^{(\kappa)}(x)=0$ para $\kappa>1$, luego en este caso $||f||_{N,k}=1$ para $I\subseteq[0,N]$.

PROPOSICIÓN 1.4. *Supongamos que $f_1, \cdots, f_m \in \mathcal{C}^k(I)$. Entonces*

$$|| \prod_{1\leq j\leq m} f_j||_{N,k} \leq [(k+1)N]^{m-1} \prod_{1\leq j\leq m} ||f_j||_{N,k}.$$

DEMOSTRACIÓN. Para cada $\kappa \leq k$, tenemos

$$\frac{d^\kappa}{dx^\kappa}\left(\prod_{1\leq j\leq m} f_j\right) = \sum_{i_1+\cdots+i_m=\kappa} \frac{\kappa!}{i_1!\cdots i_m!} \prod_{1\leq j\leq m} f_j^{i_j}.$$

Dado que $i_1+\cdots+i_m=\kappa$, tenemos $\prod_{1\leq j\leq m} N^{i_j} = N^\kappa$. Entonces

$$N^{\kappa-1} = N^{\kappa-1}\cdot N^{m-\kappa}\cdot \prod_{1\leq j\leq m} N^{i_j-1},$$

con lo que, usando la desigualdad triangular:

$$\frac{N^{\kappa-1}}{\kappa!}|\frac{d^\kappa}{dx^\kappa}\left(\prod_{1\leq j\leq m} f_j(x)\right)| \leq N^{\kappa-1} \sum_{i_1+\cdots+i_m=\kappa} N^{m-\kappa} \prod_{1\leq j\leq m} N^{i_j-1}\frac{|f_j{}^{(i_j)}(x)|}{i_j!}.$$

Acotando $N^{i_j-1}\frac{|f^{(i_j)}(x)|}{i_j!} \leq ||f_j||_{N,k}$, y notando que la cantidad de soluciones en enteros positivos de $i_1+\cdots+i_m=\kappa$ está acotada por $(\kappa+1)^{m-1}$, concluímos la proposición. □

De la Proposición 1.4 obtenemos la siguiente estimación:

COROLARIO 1.1. *Sea $f \in \mathcal{C}^k(I)$. Si llamamos $g_{p,q}(x) := x^p f(x)^q$, se tiene*

$$||g_{p,q}||_{N,k} \leq [(k+1)N]^{p+q-1}\,||f||^q{}_{N,k}.$$

Podemos reescribir la cota que obtuvimos antes. Usando la Proposición 1.1 y la definición de las normas $||\cdot||_{N,k}$ y acotando, tenemos:

LEMA 1.1. *Sean $x_1 < \cdots < x_n$ puntos distintos del intervalo $[0,N]$ y para cada n, sea $f_j \in \mathcal{C}^{n-1}([0,N])$. Se tiene que*

$$|\det(f_j(x_i))| \leq |V(x_1,\cdots,x_n)|n!N^{-\frac{n(n-3)}{2}} \prod_{1\leq k\leq n} ||f_k||_{N,N-1}. \tag{1.6}$$

La estimación (1.6) da una cota superior para un determinante que contiene información de una cierta familia de (f_j) con condiciones de regularidad, evaluadas en ciertos puntos. En las aplicaciones, lo que se hace es elegir los puntos x_i y las funciones f_j.

2. Detectando puntos en curvas algebraicas

En lo que sigue, nuestra meta es buscar una forma de detectar cuándo ciertos puntos de una curva están en una curva algebraica; en este caso esa sería la noción de detectar que los puntos poseen algún tipo de estructura algebraica.

Sea $d \geq 1$. Observemos que una curva algebraica plana de grado d definida por la ecuación $F(x,y)=0$ con $\deg(F)=d$ tiene a lo sumo $\frac{1}{2}(d+1)(d+2)$ monomios. En efecto, si homogeneizamos F, sustituyendo $x=\frac{x'}{z}, y=\frac{y'}{z}$, y multiplicamos por z^d obtenemos

$$z^d F(\frac{x'}{z},\frac{y'}{z}) = \sum_{a+b+c=d} e_{a,b,c}x'^a y'^b z^c,$$

con lo que la cantidad total de posibles polinomios F de grado d, que es la máxima cantidad de monomios que puede poseer un polinomio de grado d, coincide con la cantidad de soluciones no negativas de la ecuación $a+b+c=d$, que es $\binom{d+3-1}{d} = \binom{d+2}{2} = D$. Entonces:

PROPOSICIÓN 1.5. *Sea k un cuerpo. Un polinomio $F \in k[x,y]$ de grado d tiene a lo sumo D monomios, con*

$$D = \frac{1}{2}(d+1)(d+2). \tag{2.1}$$

Podemos concluir:

PROPOSICIÓN 1.6. *Sea $D = \frac{1}{2}(d+1)(d+2)$. Dada una familia $\{P_s\}_{1\leq s\leq k}$ de puntos del plano, con $k < D$, existe una curva algebraica plana de grado a lo sumo d que se anula en todos los P_s.*

DEMOSTRACIÓN. Sea $\{P_s\}_s$ una familia de puntos como en la proposición y notemos $P_s = (a_s, b_s)$. Reemplazando en una ecuación general de un polinomio de grado d, tenemos el sistema de ecuaciones

$$\sum_{i+j\leq d} c_{ij} a_s^i b_s^j = 0. \tag{2.2}$$

Dado que hay D incógnitas (una por cada monomio genérico) y $k < D$ ecuaciones, la ecuación (2.2) admite una solución no trivial,. Luego existe una curva algebraica plana de grado a lo sumo d que contiene a los k puntos $\{P_s\}_{1\leq s\leq k}$. □

Ahora estamos en condiciones de probar una generalización de la Proposición 1.6. Fijado $d \geq 1$, vamos a denotar

$$J_d := \{j = (j_1, j_2) : 0 \leq j_1, 0 \leq j_2, j_1 + j_2 \leq d\}.$$

Por la Proposición 1.5, tenemos $\#J_d = D$. Si $P = (x, y)$ entonces

$$P^j = P^{(j_1,j_2)} := x^{j_1} y^{j_2}.$$

LEMA 1.2. *Los puntos $P_1, \cdots, P_t$ pertenecen a alguna curva algebraica plana de grado a lo sumo d si y sólo si*

$$\operatorname{rank}(P_i^j)_{1\leq i\leq t, j\in J_d} < D.$$

DEMOSTRACIÓN. Sea A la matriz dada por

$$A = (P_i^j)_{n+1\leq i\leq n+t, j\in J_d}.$$

Se trata de una matriz de $t \times \#J_d = t \times D$. Si $t < D$, la cota del rango se verifica automáticamente y la Proposición 1.6 nos dice que existe alguna curva algebraica plana de grado a lo sumo d que contiene los puntos. Supongamos entonces $t \geq D$ y que

$$P(x, y) = \sum_{j\in J_d} a_j x^{j_1} y^{j_2},$$

con $a_j \in \mathbb{R}$ definen una curva algebraica plana de grado a lo sumo d que anula a los puntos $P_1, \cdots, P_t$. Entonces tenemos que los a_j verifican

$$\sum_{j\in J_d} a_j P_i^j = 0 \text{ para todo } 1 \leq i \leq t.$$

Esto es lo mismo que decir que para cada conjunto de índices $I \subseteq [t]$ con $\#I = D$, la matriz cuadrada $(P_i^j)_{i\in I, j\in J_d}$ no es inversible, condición que es equivalente a que el rango de dicha matriz sea estrictamente menor a D. Puesto que se trata de una submatriz de A, lo que acabamos de probar es que el rango de A es estrictamente menor a D. Recíprocamente, supongamos que el rango de A es estrictamente menor a D. Sea $\operatorname{rank}(A) = r < D$. Sea

$$A_{RJ} = (P_i^j)_{i\in R, j\in J},$$

un menor principal de tamaño $r \times r$ de rango maximal con $R \subseteq [t], J \subseteq J_d$. En particular, la matriz A_{RJ} es inversible luego su determinante es no nulo. Como $r < D$, tenemos que $J \neq J_d$ luego existe $j^* = (j_1^*, j_2^*) \in J_d \backslash J$. Definamos

$$Q(x, y) := \det \begin{pmatrix} A_{Ij} \\ x^{j_1} y^{j_2} \end{pmatrix}_{j\in J\cup\{j^*\}}.$$

Al tenerse que $j_1 + j_2 \leq d$, resulta f un polinomio de dos variables de grado a lo sumo d y con $\det(A_{RJ})$ de cofactor de $x^{j_1^*} y_2^{j^*}$. Para cada i con $1 \leq i \leq t$ resulta

$$f(P_i) = \det\left[\binom{A_{Ij}}{P_i^j}\right]_{j\in J\cup\{j^*\}} = 0,$$

dado que si $i \in R$, la matriz en la que se evalúa el determinante tiene dos filas iguales, mientras que si $i \notin I$, el determinante está forzado a ser 0 pues en caso contrario encontramos una submatriz cuadrada de A de tamaño $(r+1)\times(r+1)$ e inversible, lo que contradice que r sea el rango de A. □

Nuestro objetivo es detectar si los puntos enteros de una curva Γ en una región como $[0,N]^2$ están en una curva algebraica, digamos, de grado d. Supongamos que $\Gamma(\mathbb{Z})\cap[0,N]^2 = \{P_1,\cdots,P_s\}$ donde los puntos $\{P_i\}_i$ están ordenados por abscisa creciente. Podemos apelar a la Proposición 1.6: si $s < D$ con $D = \frac{1}{2}(d+1)(d+2)$, entonces tenemos que todos los puntos están sobre una curva algebraica plana de grado a lo sumo d. De tenerse $s \geq D$, agrupamos todos los $P_1,\cdots,P_{n_1-1}$ que podemos cubrir por una misma curva algebraica de grado a lo sumo d, es decir, manteniendo el orden de los puntos $\{P_i\}_i$, se tiene que los puntos $P_1,\cdots,P_{n_1-1}$ están contenidos en alguna curva algebraica de grado a lo sumo d, pero los puntos $P_1,\cdots,P_{n_1}$ no. Luego continuamos este procedimiento, agrupando $P_{n_1},\cdots,P_{n_2-1}$ de manera "maximal". Este procedimiento necesariamente termina (hay finitos puntos). La noción de "estructura algebraica" en este caso es, entonces, que los puntos estén contenidos de manera "controlada" en curvas algebraicas de grado "pequeño".

A priori, para saber la cantidad de curvas empleadas en este procedimiento, requerimos conocer $s = \#\Gamma(\mathbb{Z})$, que es exactamente lo que queremos acotar. El método del determinante de Bombieri-Pila permite invertir este proceso.

Formalicemos los pasos que mencionamos. Dada Γ una curva (el gráfico de una función suave, una curva regular), sean $P_1,\cdots,P_s$ los puntos de coordenadas enteras de $\Gamma\cap[0,N]^2$, ordenados por orden creciente de abcisa. Luego $P_t = (x_t, y_t)$ con $x_t, y_t \in \mathbb{Z}$. Sea $d \geq 1$ un entero positivo.

Definimos la secuencia finita de enteros $(n_l)_l$ de la siguiente forma:

(1) $n_0 = 1$
(2) Supongamos n_{l-1} ya definido. Entonces n_l es el único entero tal que los puntos P_i para $n_{l-1} \leq i < n_l$ se encuentran en alguna curva algebraica plana de grado a lo sumo d pero los puntos P_i para $n_{l-1} \leq i \leq n_l$ no, si tal entero n_l existe. Caso contrario, la secuencia termina con n_{l-1}

Por la Propiedad 1.6, tenemos que $D-1$ puntos se encuentran siempre en alguna curva algebraica plana de grado a lo sumo d. Luego se tiene $n_l - n_{l-1} \geq D-1$. El Lema 1.2 nos permite obtener

Corolario 1.2.

- $\operatorname{rank}(P_i^j)_{n_{l-1}\leq i<n_l, j\in J_d} = D-1$.
- $\operatorname{rank}(P_i^j)_{n_{l-1}\leq i\leq n_l, j\in J_d} = D$.

Veamos ahora cómo combinaremos el Lema 1.2 y el Lema 1.1 para acotar $s = \#\Gamma\cap[0,N]^2$. Para ello, notemos lo siguiente. Si conseguimos $A > 0$ que cumpla

$$|x_{n_{k+1}} - x_{n_k}| \geq A,$$

para todo k, con $(n_k)_k$ la secuencia que armamos cuando cubrimos los P_k, entonces tenemos que los puntos $x_{n_1}, x_{n_2},\cdots,x_{n_l}$ están distribuídos en $[0,N]$ separados por distancias de al menos A. Esto implica que $n_l \leq \frac{N}{A}$ y como los puntos de abcisa entre x_{n_k} y $x_{n_{k+1}}$ están todos contenidos en una curva algebraica plana de grado a lo sumo d, resulta que los $P_1,\cdots,P_s$ están contenidos en a lo sumo $\frac{N}{A}$ curvas algebraicas planas de grado a lo sumo d. Esto quiere decir que conseguir una cota suficiente buena de A implica conseguir una cota buena de s. Ese es el objetivo del próximo lema.

Sea $f:[0,N]\to\mathbb{R}$ una función suave. Consideremos Γ el gráfico de f. Notemos que $\Gamma(\mathbb{Z})$ es finito. Lo que vamos a hacer es cubrir a $\Gamma(\mathbb{Z})$ por curvas algebraicas planas de grado a lo sumo d. Para estimar la cantidad de curvas del cubrimiento, usamos el Lema 1.1, para dar una cota inferior a $|x_{n_{l+1}} - x_{n_l}|$.

Lema 1.3. *Sea $d \geq 1$ un entero positivo y sea $D = \frac{1}{2}(d+1)(d+2)$. Sea Γ el gráfico de $f \in \mathcal{C}^{D-1}([0,N])$ y escribamos $\Gamma(\mathbb{Z}) = \{P_1,\cdots,P_s\}$. Consideremos la secuencia finita de enteros positivos (n_l) introducida en el párrafo anterior para Γ. Si $P_k = (x_k, f(x_k))$ se tiene que*

$$|x_{n_{l+1}} - x_{n_l}| \geq \left(D^2||f||_{N,D-1}\right)^{-\frac{4}{3(d+3)}} N^{1-\frac{8}{3(d+3)}}.$$

Demostración. El Corolario 1.2 dice que la matriz $(P_i^j)_{n_l\leq i\leq n_{l+1}, j\in J_d}$ tiene rango D. Entonces existe un subconjunto $I \subseteq \{n_l,\cdots,n_{l+1}\}$ de cardinal D tal que la matriz $(P_i^j)_{i\in I, j\in J_d}$ de tamaño $D\times D$ es inversible. Como los P_i tenían coordenadas enteras, si $j = (j_1, j_2)$, los valores $P_i^j = x_i^{j_1}\cdot f(x_i)^{j_2}$ son enteros. Esto nos permite concluir que $(P_i^j)_{i\in I, j\in J_d}$ es

una matriz inversible de coordenadas enteras, con lo que $\Delta = \det((P_i^j)_{i\in I, j\in J_d})$ es un entero no nulo. Tenemos en particular $|\Delta| \geq 1$. Ahora usamos el Lema 1.1 con la matriz (P_i^j), obteniendo:

$$|\Delta| = |\det(P_i^j)_{i\in I, j\in J_d}| \leq |V(x_{n_l}, x_{n_l+1}, \cdots, x_{n_{l+1}})| N^{\frac{-D(D-3)}{2}} D! \prod_{j\in J_d} ||f_j||_{N,D-1}. \tag{2.3}$$

donde

$$f_{(j_1,j_2)}(x) := x^{j_1} f(x)^{j_2}.$$

Dado que $V(x_{n_l}, \cdots, x_{n_{l+1}})$ consiste de todos los productos $x_i - x_j$ con $i < j, i, j \in I$, podemos acotar cada $|x_i - x_j|$ por $|x_{n_{l+1}} - x_{n_l}|$. Dado que hay $\frac{D(D-1)}{2}$ de tales posibles productos, concluímos que

$$|V(x_{n_l}, \cdots, x_{n_{l+1}})| \leq |x_{n_{l+1}} - x_{n_l}|^{\frac{D(D-1)}{2}}. \tag{2.4}$$

Ahora, el Corolario 1.1 nos da una cota para las normas de las funciones f_j:

$$||f_j||_{N,D-1} \leq (DN)^{j_1+j_2-1} ||f||_{N,D-1}^{j_2}. \tag{2.5}$$

Usamos la expresión (2.5) para acotar el producto de las normas $||f_j||_{N,D-1}$ en la expresión (2.3). Nos queda

$$\prod_{j\in J_d} (DN)^{j_1+j_2-1} ||f||_{N,D-1}^{j_2}. \tag{2.6}$$

Contando bien las potencias de cada factor que aparece en el producto, combinando las desigualdades (2.4), (2.6) y usando la cota $D! \leq D^D$ en 2.3, llegamos a

$$1 \leq |\Delta| \leq \left(\frac{|x_{n_{l+1}} - x_{n_l}|}{N}\right)^{\frac{D(D-1)}{2}} (DN)^{\frac{dD}{3}} ||f||_{N,D-1}^{\frac{dD}{6}}.$$

Despejando y usando que

$$\frac{d}{D-1} = \frac{2d}{d^2+3d+1} \leq \frac{2d}{d^2+3d} \leq \frac{2}{d+3},$$

obtenemos la cota del lema. □

Podemos entonces concluir:

LEMA 1.4. *Sea $d \geq 1$ un entero positivo, $D = \frac{1}{2}(d+1)(d+2)$ y $f \in \mathcal{C}^{D-1}([0,N])$. Entonces los puntos de coordenadas enteras de Γ dada por el gráfico de f están contenidos en la unión de no más de*

$$3\left(||f||_{N,D-1}^{\frac{1}{2}} N\right)^{\frac{8}{3(d+3)}} + 1$$

curvas algebraicas de grado a lo sumo d. Si $||f||_{N,D-1}^{\frac{1}{2}} N \geq 1$, entonces tenemos que la cantidad anterior no excede

$$4\left(||f||_{N,D-1}^{\frac{1}{2}} N\right)^{\frac{8}{3(d+3)}}.$$

3. Una generalización

Nuestro objetivo ahora es refinar el Lema 1.4, algo que conseguiremos obteniendo mayor control de las curvas algebraicas con las que cubrimos una curva. Esto lo vamos a conseguir restringiendo los monomios que aparecen en las ecuaciones que definen a las curvas

Sea $\mathcal{M}$ un conjunto finito de monomios en las indeterminadas x, y y sea C una curva algebraica plana definida por $G(x,y) = 0$ con $G \in \mathbb{R}[x,y]$.

DEFINICIÓN 1.2. *Decimos que C está definida en $\mathcal{M}$ si los monomios que aparecen en G pertenecen a $\mathcal{M}$.*

OBSERVACIÓN 1.1. *Observemos que tomando $\mathcal{M} = \{x^{j_1} y^{j_2} : j_1 + j_2 \leq d\}$, decir que C está definida en $\mathcal{M}$ es equivalente a decir que C tiene grado a lo sumo d. Estas son las curvas que detectamos en la sección anterior.*

Usando argumentos análogos, podemos deducir el siguiente lema:

LEMA 1.5. *Sea $\mathcal{M}$ un conjunto finito de monomios en x, y. Definamos $D := \#\mathcal{M}$ y $J_{\mathcal{M}} := \{j = (j_1, j_2) : x^{j_1}y^{j_2} \in M\}$. Los puntos $P_1, \cdots, P_t$ pertenecen a alguna curva algebraica plana definida en $\mathcal{M}$ si y sólo si*

$$\text{rank}(P_i^j)_{1\leq i\leq t, j\in J_{\mathcal{M}}} < D.$$

Sean ahora $P_1, \cdots, P_s$ los puntos de coordenadas enteras de $\Gamma \cap [0, N]^2$ con Γ una curva (suave, algebraica). Escribamos $P_t = (x_t, y_t)$ con $x_t, y_t \in \mathbb{Z}$. Definimos una secuencia finita $(n_l)_l$, que generaliza la secuencia de la sección anterior, como sigue:

(1) $n_0 = 1$
(2) Supongamos n_{l-1} ya definido. Entonces n_l es el único entero tal que los puntos P_i para $n_{l-1} \leq i < n_l$ se encuentran en alguna curva algebraica plana definida en $\mathcal{M}$ pero los puntos P_i para $n_{l-1} \leq i \leq n_l$ no, si tal entero n_l existe. Caso contrario, la secuencia termina con n_{l-1}

Igual que antes, deducimos:

COROLARIO 1.3. *Sea $\mathcal{M}$ un conjunto finito de monomios en x, y. Sean $D := \#\mathcal{M}$ y $J_{\mathcal{M}} := \{j = (j_1, j_2) : x^{j_1}y^{j_2} \in M\}$. Para la sucesión $(n_l)_l$ del párrafo anterior se tiene:*

- $n_{l+1} - n_l \geq D - 1$.
- $\text{rank}(P_i^j)_{n_{l-1}\leq i<n_l, j\in J_d} = D - 1$.
- $\text{rank}(P_i^j)_{n_{l-1}\leq i\leq n_l, j\in J_d} = D$.

También obtenemos una cota inferior para $|x_{n_{l+1}} - x_{n_l}|$, siguiendo los mismos pasos que antes.

LEMA 1.6. *Sea $\mathcal{M}$ un conjunto finito de monomios en x, y. Sean $D := \#\mathcal{M}$ y $J := \{j = (j_1, j_2) : x^{j_1}y^{j2} \in \mathcal{M}\}$. Definimos $p := \sum_{j\in J}(j_1 + j_2)$, $q = \sum_{j\in J} j_2$. Sea Γ el gráfico de $f \in \mathcal{C}^{D-1}([0, N])$ y escribamos $\Gamma(\mathbb{Z}) = \{P_1, \cdots, P_s\}$. Consideremos la secuencia finita de enteros positivos (n_l) introducida en el párrafo anterior para Γ. Si $P_k = (x_k, f(x_k))$ se tiene que*

$$|x_{n_{l+1}} - x_{n_l}| \geq \left(D^p||f||^q_{N,D-1}\right)^{-\frac{2}{D(D-1)}} N^{1-\frac{2p}{D(D-1)}}.$$

Finalmente de este lema deducimos la versión generalizada del Lema 1.4:

LEMA 1.7. *Sea $\mathcal{M}$ un conjunto finito de monomios en x, y. Sean D, J, p, q como en el lema anterior. Sea $f \in \mathcal{C}^{D-1}([0, N])$ y Γ el gráfico de f. Entonces los puntos de coordenadas enteras de Γ están contenidos en la unión de no más de*

$$\left(D^p||f||^q_{N,D-1}\right)^{\frac{2}{D(D-1)}} N^{\frac{2p}{D(D-1)}} + 1$$

curvas algebraicas definidas en $\mathcal{M}$. Si $||f||^{\frac{q}{p}}_{N,D-1}N \geq 1$, entonces tenemos que la cantidad anterior no excede

$$2\left(D^p||f||^q_{N,D-1}\right)^{\frac{2}{D(D-1)}} N^{\frac{2p}{D(D-1)}}$$

.

□

4. El problema en dimensiones superiores

El objetivo de esta sección es ver cómo las ideas anteriores se pueden aplicar a superficies regulares en vez de curvas regulares. Vamos a dar resultados probados en [**39**][pág. 210-213].

Las dificultades que surgen al buscar estructura algebraica en conjuntos de dimensiones superiores se deben a que la geometría misma del espacio ya de por sí es más rica. Por ejemplo, en $\mathbb{R}^2$ nuestra noción de ''estructura algebraica" de un conjunto estaba basada en que el conjunto estuviera contenido de manera controlada en curvas algebraicas planas. En $\mathbb{R}^3$, en un principio tenemos hipersuperficies algebraicas, con lo que podemos aspirar a estudiar si un conjunto está contenido en hipersuperficies algebraicas, de manera controlada, pero no es esperable distinguir si estas hipersuperficies son curvas o superficies algebraicas. Incluso el concepto clásico de superficie suave es más delicado.

En lo que resta del capítulo, vamos a estudiar conjuntos $\Omega \subseteq \mathbb{R}^n$ que admiten una parametrización de la forma $\phi : [0, N]^k \to \mathbb{R}^n$, es decir, $\phi([0, N]^k) = \Omega$. Vamos a usar la siguiente notación. Si $P_i = (x_{1i}, \cdots, x_{ki})$ entonces vamos a notar

$$P_i^j := \prod_{l=1}^{k} x_{li}^{j_l}.$$

Dados k, d enteros positivos, vamos a definir

$$J_{k,d} = \{\mathbf{j} \in \mathbb{N}^k : |\mathbf{j}| \leq d\}, \;\; \#J_k(d) = D_k(d),$$

Para $\mathbf{a} = (a_1, \cdots, a_k) \in \mathbb{N}^k$, definimos

$$|\mathbf{a}| := \sum_{i=1}^{k} a_i.$$

LEMA 1.8. *Los puntos $P_{n+1}, \cdots, P_{n+t} \in \mathbb{R}^k$ pertenecen a una misma hipersuperficie algebraica de grado a lo sumo d si y sólo si*

$$\operatorname{rank}(P_i^j)_{n+1 \leq n+t, j \in J_k(d)} < D_k(d),$$

La demostración de este lema es exactamente la misma que la del Lema 1.2. Ahora obtenemos un resultado como en el Lema 1.1.

LEMA 1.9. *Sean k, n, d enteros positivos y $D = D_n(d)$ del lema. Existe un entero positivo $b = b(k, n, d)$ y una constante positiva $B = B(k, n, d)$ tal que si $\phi_1, \cdots, \phi_D : \mathbb{R}^k \to \mathbb{R}$ son de clase $\mathcal{C}^{b+1}([0, N]^k)$ entonces dado $U \subseteq \mathbb{R}^k$ de radio $r \leq 1$ y dados $z^{(1)}, \cdots, z^{(D)} \in [0, 1]^k \cap U$, se tiene*

$$|\det(\phi_i(z^{(j)})_{1 \leq i,j \leq D})| = O_{N,\phi_1,\cdots,\phi_D,d}(r^B).$$

Las constantes del Lema 1.9 son efectivas; la dependencia en las funciones ϕ_i aparece en término de las normas $||\cdot||_{N,k}$ aunque, en este caso, son muchas más derivadas las que hay que considerar y por ello consideramos la dependencia dentro de la constante. De nuevo, la demostración es bastante similar al Lema 1.2. Consultar los detalles en [**39**].

Dado $t > 0$, consideramos la dilatación

$$\Omega(\mathbb{Z}, t) := \{\mathbf{a} \in \Omega : t\mathbf{a} \in t\Omega(\mathbb{Z})\}.$$

Es de interés estudiar las dilataciones pues permiten entender cómo un tipo sencillo de perturbación afecta a la estructura algebraica de una superficie. Se tiene:

TEOREMA 1.1. *Sean k, n, d enteros positivos. Entonces existe un entero positivo $b = b(k, n, d)$ y una constante positiva $\varepsilon = \varepsilon(k, n, d)$ tal que si $\phi : [0, N]^k \to \mathbb{R}^n$ es de clase $\mathcal{C}^{b+1}([0, N]^k)$ y $t \geq 1$, entonces se tiene que $\Omega(\mathbb{Z}, t)$ está contenido en la unión de a lo sumo*

$$O_{\phi,d,N}(t^{\varepsilon})$$

hipersuperficies algebraicas de grado a lo sumo d.

DEMOSTRACIÓN. Al igual que en el Lema 1.1, si escribimos $\phi = (\phi_1, \cdots, \phi_n)$, consideramos

$$\phi_{\mathbf{a}} = \prod_{i=1}^{n} \phi_i^{a_i},$$

con $|\mathbf{a}| \leq d$. Sean $U \subseteq \mathbb{R}^n$ un disco de radio $r \leq 1$ a elegir y $P_1, \cdots, P_D$ puntos de $[0, 1]^k \cap U$ tales que $\phi(P_i) \in \Omega(\mathbb{Z}, t)$ para todo i. A continuación, usamos el Lema 1.9 y obtenemos

$$|\det(\phi_{\mathbf{a}}(P_j)_{|\mathbf{a}| \leq d, 1 \leq j \leq D})| = O_{N,\phi,d}(r^B).$$

Como $\phi(Q_j) \in \Omega(t, \mathbb{Z})$, se tiene que $t\phi(Q_j) \in \mathbb{Z}$. Entonces, multiplicando por t en cada fila de la matriz $\phi_{\mathbf{a}}(P_j)_{|\mathbf{a}| \leq d, 1 \leq j \leq D}$, se tiene que existe un $V = V(k, n, d)$ tal que

$$t^V \det(\phi_{\mathbf{a}}(P_j)_{|\mathbf{a}| \leq d, 1 \leq j \leq D}) \in \mathbb{Z}. \tag{4.1}$$

Puesto que r podemos elegirlo muy chico, podemos considerarlo de manera que los $\phi_i(Q_j)$ sean muy pequeños, de manera que el producto (4.1) tenga módulo menor a 1, luego se verifica

$$\det(\phi_{\mathbf{a}}(P_j)_{|\mathbf{a}| \leq d, 1 \leq j \leq D}) = 0,$$

para cualquier elección de los $P_j \in [0,1]^k \cap U$, con $\phi(P_j) \in \Omega(\mathbb{Z}, t)$. Por ejemplo, si la constante del Lema 1.9 es c, entonces alcanza con tomar $r < (ct^V)^{-\frac{1}{B}}$. Por el Lema 1.8, todos estos puntos se encuentran en una sola hipersuperficie algebraica de grado a lo sumo d. Como $[0,1]^k$ es compacto, si cubrimos con discos cerrados de radio suficientemente pequeño, por ejemplo $r < \frac{1}{2}(ct^V)^{-\frac{1}{B}}$ podemos extraer un subcubrimiento finito y, por cada disco que nos quedó, sabemos que los puntos están contenidos en una sola hipersuperficie algebraica de grado a lo sumo d. Notemos que la cantidad de discos del subcubrimiento es del orden de $O_{N,\phi,d}(t^{\frac{kV}{B}})$ donde la constante involucrada contiene a la constante c. Llamando $\varepsilon = \frac{kV}{B}$, esto nos da la estimación que queremos. □

Utilizando un razonamiento similar a la demostración del Teorema 1.1, podemos dar una estimación para los puntos racionales de Ω de altura acotada. Recordemos que para nosotros, si $\mathbf{a} = (\frac{a_1}{b_1}, \cdots, \frac{a_n}{b_n})$, con $\gcd(a_i, b_i) = 1$ y $b_i > 0$ para todo i, entonces $H(a) := \max_{1 \leq i \leq n}\{|a_i|, b_i\}$

Teorema 1.2. *Sean k, n, d enteros positivos. Entonces existe un entero positivo $b = b(k,n,d)$ y una constante positiva $\varepsilon' = \varepsilon'(k,n,d)$ tal que si $\phi : [0,N]^k \to \mathbb{R}^n$ es de clase $\mathcal{C}^{b+1}([0,1])$ y $H \geq 1$, entonces se tiene que $\Omega(\mathbb{Q}, H)$ está contenido en la unión de a lo sumo*

$$O_{\phi,d,N}(H^{\varepsilon})$$

hipersuperficies algebraicas de grado a lo sumo d. Además, para k, n fijos con $k < n$ se tiene $\varepsilon'(d) \to 0$ cuando $d \to +\infty$.

Demostración. Vamos a utilizar el Teorema 1.9 para ver que, usando el Lema 1.8, tenemos puntos de un disco dentro de una hipersuperficie algebraica de grado a lo sumo d. Luego cubrimos por discos de radio pequeño, extraemos un subcubrimiento y concluímos una cota. Sea entonces $U \subseteq \mathbb{R}^k$ disco de radio $r \leq 1$ y consideremos $P_1, \cdots, P_{D_n(d)}$ puntos tales que para todo j se tiene $P_j \in [0,1]^k \cap U$ y $\phi(P_j) \in \Omega(\mathbb{Q}, H)$. Sean $b_1, \cdots, b_{D_n(d)}$ racionales con $|b_j| \leq H$ para todo j tales que $b_j \phi(Q_j) \in b_j \Omega(\mathbb{Z})$. Entonces

$$\left[\prod_{j=1}^{D_n(d)} {b_j}^d\right] \det(\phi_{\mathbf{a}}(P_j)_{|\mathbf{a}| \leq d, 1 \leq j \leq D}) \in \mathbb{Z}. \tag{4.2}$$

Como $\prod_{j=1}^{D_n(d)} b_j^d \leq H^{dD_n(d)}$, considerando U de radio $r < (cH^{-dD_n(d)})^{-\frac{1}{B}}$ donde c es la cota del Teorema 1.9, tenemos que el determinante (4.2) tiene módulo menor a 1. Concluímos

$$\det(\phi_{\mathbf{a}}(P_j)_{|\mathbf{a}| \leq d, 1 \leq j \leq D}) = 0,$$

de lo que deducimos, mediante el Lema 1.8, que los puntos $P_1, \cdots, P_{D_n(d)}$ están contenidos en una sola hipersuperficie de grado a lo sumo d. Tomando $\varepsilon'(k,n,d) = \frac{kdD_n(d)}{B(k,n,d)}$, un análisis cuidadoso de quién es la cota $B(k,n,d)$ (ver [**39**]) muestra que $\varepsilon'(k,n,d) \to 0$ cuando $d \to +\infty$. Concluímos la demostración mediante un argumento similar al del Teorema 1.1. □

Observación 1.2. *Los métodos anteriores no permiten distinguir entre hipersuperficies algebraicas irreducibles; el cubrimiento que aseguran los resultados es de hipersuperficies en general.*

Observación 1.3. *Es posible dar alguna cota explícita para la cantidad de discos que cubren el $(0,1)^k$ (aunque dar una cota buena del cubrimiento no es un problema sencillo). También es posible observar que si cambiamos $[0,1]^k$ por un dominio $J \subseteq \mathbb{R}^k$ convexo compacto, exactamente los mismos resultados se mantienen. La convexidad es usada en el Lema 1.9; consultar los detalles en* [**39**].

5. Esquema general del método del determinante

Los resultados de las secciones anteriores se suelen utilizar siguiendo una estrategia similar, conocida como "método del determinante de Bombieri-Pila", debido a que fue introducido por Bombieri y Pila en [**6**], e involucra la estimación de un determinante. Podemos resumir la estrategia del método en los siguientes pasos:

- Consideramos un conjunto finito $\{P_i\}_{1 \leq i \leq n}$ de $\mathbb{R}^m$, con $m \geq 2$. Consideramos una matriz adecuada $(a_{ij})_{ij}$ tal que $\det(a_{ij})_{ij} = 0$ si y sólo si los puntos $\{P_i\}_i$ están contenidos en una hipersuperficie algebraica particular (como en el Lema 1.2, Lema 1.5, Lema 1.8).
- Obtenemos una cota superior e inferior para el determinante $\det(a_{ij})_{ij}$ (como en el Lema 1.1, Lema 1.9).

Usualmente, el método del determinante se aplica para estimar la cantidad de puntos enteros, o puntos racionales de altura acotada, de un conjunto, como el gráfico de una función suave. En estos casos, el método se aplica de la siguiente manera. Sea $X \subseteq \mathbb{R}^m$, con $m \geq 2$.

- Cubrimos el conjunto $X(\mathbb{Z})$ o $X(\mathbb{Q}, H)$ con hipersuperficies algebraicas.
- Por medio de la estrategia del método del determinante esbozada en el párrafo anterior, obtenemos una cota para la cantidad de hipersuperficies algebraicas que cubren $X(\mathbb{Z})$ o $X(\mathbb{Q}, H)$ (como en el Lema 1.4, Lema 1.7, Lema 1.2).
- Estimamos $\#(X \cap S)$, con S una hipersuperficie algebraica del cubrimiento del primer paso.
- Combinando las estimaciones de los dos pasos anteriores, obtenemos una cota superior para $X(\mathbb{Z}), X(\mathbb{Q}, H)$.

El tercer paso de la estrategia recién esbozada es delicado, debido a que podría ocurrir que $\#(X \cap S) = +\infty$ para alguna hipersuperficie algebraica. Esta problemática, sin embargo, se puede "solucionar" cubriendo al conjunto X a estudiar con hipersuperficies algebraicas particulares; en el caso que $X \subseteq \mathbb{R}^2$, por ejemplo podemos restringirnos a curvas algebraicas definidas en un conjunto de monomios $\mathcal{M}$. De cualquier forma, aún cuando $\#(X \cap S)$ es finito, es difícil obtener cotas superiores para $\#(X \cap S)$ suficientemente uniformes, por ejemplo, en el grado de la hipersuperficie S.

El método del determinante es utilizado en varios trabajos. Además del trabajo original [**6**], es empleado en [**38**], [**39**], [**41**], [**18**], y es generalizado por Heath-Brown en [**17**]. En los próximos capítulos vamos a usar el método del determinante de manera clave.

CAPÍTULO 2

Puntos enteros y racionales en curvas y superficies

Consideremos una curva plana rectificable Γ simple parametrizada por $\sigma : [0, N] \to \mathbb{R}^2$, es decir, σ parametriza una curva de $\mathbb{R}^2$ a la que se le puede calcular su longitud aproximando con poligonales, y que no se corta a sí misma. Notemos que Γ contiene finitos puntos de coordenadas enteras. Sean $t_0 < t_1 < \cdots < t_k$ los tiempos en los que σ toma valores enteros, es decir, $\Gamma(\mathbb{Z}) = \{\sigma(t_i)\}_{1 \leq i \leq k}$. Tenemos que para $1 \leq i \leq k$, $\sigma([t_{i-1}, t_i])$ es una curva simple, debido a que Γ era simple. Poniendo $\Gamma_i = \sigma([t_{i-1}, t_i])$, se tiene

$$\bigcup_{1 \leq i \leq k+1} \Gamma_i \subseteq \Gamma,$$

de donde deducimos, usando que Γ es simple,

$$\text{long}(\Gamma) \geq \sum_{1 \leq i \leq k} \text{long}(\Gamma_i),$$

con $\text{long}(\Gamma), \text{long}(\Gamma_i)$ las longitudes de las curvas respectivas. Como los puntos $\sigma(t_i)_{1 \leq i \leq k}$ tienen coordenadas enteras, tenemos que la longitud del segmento que une dos puntos consecutivos $\sigma(t_{i-1}), \sigma(t_i)$ es al menos 1, con $1 \leq i \leq k$. Esto nos dice que la longitud de cada tramo Γ_i es al menos 1, con $1 \leq i \leq k$. Concluímos entonces

$$\text{long}(\Gamma) \geq \sum_{1 \leq i \leq k} \Gamma_i \geq k.$$

Si llamamos $\text{long}(\Gamma) = l$, lo que obtuvimos recién no es otra cosa que la desigualdad $k \leq l$. Dado que había $k + 1$ puntos de coordenadas enteras, acabamos de probar:

PROPOSICIÓN. *Sea Γ una curva rectificable simple de longitud l. Entonces*

$$\#\Gamma(\mathbb{Z}) \leq l + 1.$$

Sin hipótesis adicionales, la cota anterior es óptima. En efecto, si consideramos $\sigma(t) := (t, 0)$ con $t \in [0, N]$, resulta que $\Gamma = \sigma([0, N])$ tiene longitud N y posee exactamente $N + 1$ puntos de coordenadas enteras. Sin embargo, es posible mejorar la cota mediante hipótesis adicionales de regularidad. Jarnik (1925) en [**20**] dio un resultado, asumiendo que la curva venía dada por el gráfico de una función $f : [0, N] \to \mathbb{R}$ continua y estrictamente convexa, es decir, para todo $t \in [0, 1]$ y todo par $x, y \in [0, N]$ se tiene

$$f(tx + (1 - t)y) < t(fx) + (1 - t)f(y).$$

La cota de Jarnik queda enunciada en la siguiente proposición.

TEOREMA (Jarnik). *Sea Γ una curva de longitud l que es el gráfico de una función $f : [0, N] \to \mathbb{R}$ continua y estrictamente convexa. Entonces se tiene que*

$$\#\Gamma(\mathbb{Z}) \leq 3(4\pi)^{-\frac{1}{3}} l^{\frac{2}{3}} + O(l^{\frac{1}{3}}). \tag{0.1}$$

Podemos decir a secas, que si $l \geq 1$, entonces $\#\Gamma(\mathbb{Z}) = O(l^{\frac{2}{3}})$, siendo la primera hipótesis para nada restrictiva a nuestro interés debido a que una curva de longitud $0 < l < 1$ no tiene puntos de coordenadas enteras. Swinnerton-Dyer (1972) [**52**] mejora el exponente $\frac{2}{3}$ en (0.1) asumiendo que la curva convexa cumple condiciones adicionales de suavidad.

Si bien resulta intuitivo que la longitud de una curva Γ determina una cota general para $\#\Gamma(\mathbb{Z})$, para poder obtener una cota más precisa, fue necesario agregar una hipótesis acerca de la convexidad de la curva. Si Γ viene dada por el gráfico de una función $f : [0, N] \to \mathbb{R}^2$, para determinar la convexidad de f se puede recurrir a estudiar el signo de f'', siempre que exista. Si $f'' > 0$ tenemos que f es estrictamente convexa; si $f'' < 0$, tomando $g(x) := -f(x)$ podemos aplicar el teorema 2 a g. No obstante, puede ocurrir que $f''(x) = 0$ para posiblemente muchos puntos. Estas observaciones nos dicen que para una f bastante complicada, se requeriría conocer demasiada información para aplicar el Teorema 2. Esencialmente, el problema

que surge aquí es el de obtener una cota uniforme (que no dependa de cada tramo). En esta dirección, Schmidt (1985) en [**48**] obtiene una cota uniforme:

TEOREMA (Schmidt). *Sea $f \in \mathcal{C}^3([0,N])$ con $|f(x)| \leq N$ y $f'''(x) \neq 0$ para $x \in [0,N]$. Denotemos Γ al gráfico de f. Entonces para todo $\varepsilon > 0$, se tiene la estimación*

$$\#\Gamma(\mathbb{Z}) = O_\epsilon(N^{\frac{3}{5}+\epsilon}). \tag{0.2}$$

Las curvas algebraicas planas vienen dadas por una ecuación $f(x,y) = 0$ con $f \in \mathbb{R}[x,y]$. Si f es irreducible, posee finitos puntos singulares. Esto en un principio nos dice que si Γ es una curva algebraica plana geométricamente irreducible, se podría descomponer en tramos que son gráficos de funciones suaves, estudiar la convexidad de estos tramos, y luego tratar de utilizar el Teorema 0.1 o el Teorema 0.2, sobre cada porción $[0,N]^2$ del plano. De nuevo, calcular los puntos singulares de una curva involucra calcular las soluciones de un sistema de ecuaciones de dos polinomios de dos variables. Adicionalmente, hace falta tener un buen estudio de cada parametrización de los tramos. Si bien teóricamente parece funcionar este procedimiento, requiere demasiada información de Γ, y con tanta información no haría falta recurrir a las estimaciones del teorema 0.1 y el teorema 0.2. Además, no parece esperable obtener una cota razonablemente uniforme para las curvas algebraicas de, digamos, grado d (ver pregunta 0.7), pues deberíamos poder acotar uniformemente las derivadas de las parametrizaciones de los tramos. Bombieri y Pila (1989) estudian estas problemáticas y otros relacionadas en [**6**]. Con respecto al resultado de Schmidt, obtienen:

TEOREMA (Bombieri-Pila). *Sean $d \geq 4$ y $N \geq 1$ enteros positivos. Existe $c = c(d)$ tal que para toda curva Γ que es el gráfico de una función $f \in \mathcal{C}^D(I)$ con $D := \frac{1}{2}(d+1)(d+2)$ e I es un intervalo de longitud N, $|f'| \leq 1$ y $f^{(D)}$ con a lo sumo m ceros, se tiene la estimación:*

$$\#\Gamma(\mathbb{Z}) \leq (m+1)c(d)N^{\frac{1}{2}+\frac{3}{d+3}}.$$

Para el problema de curvas algebraicas, Bombieri y Pila obtienen

TEOREMA (Bombieri-Pila). *Sea C una curva algebraica geométricamente irreducible de grado $d \geq 2$ y sea $N \geq \exp(d^6)$. Entonces*

$$\#(\Gamma(\mathbb{Z}) \cap [0,N]^2) \leq N^{\frac{1}{d}} \exp(12\sqrt{d\log(N)\log\log(N)}).$$

Lo novedoso[1] del trabajo es, justamente, la obtención de cotas uniformes en el grado de las curvas algebraicas.

Bombieri y Pila en [**6**] también dan una estimación para los puntos enteros en las dilataciones de curvas trascendentes:

TEOREMA. *Sea $f : I \to \mathbb{R}$ una función analítica trascendente, con I un intervalo cerrado y acotado. Sea Γ el gráfico de f. Sea $\varepsilon > 0$. Entonces para todo $t \geq 1$ se tiene*

$$\#(t\Gamma)(\mathbb{Z}) = O_{f,\epsilon}(t^\varepsilon).$$

Este resultado es de interés, debido a que da una estimación para los puntos racionales de una curva analítica trascendente. Concretamente, tomando $t = N$ un entero positivo, el teorema anterior da una cota para los puntos racionales de denominador N de una curva trascendente Γ.

En este capítulo, vamos a aplicar el método del determinante para probar las estimaciones de Bombieri-Pila para curvas analíticas trascendentes y curvas algebraicas planas, siguiendo esencialmente los pasos del trabajo original [**6**]. El caso de curvas analíticas trascendentes es tratado en la sección 1. El caso de curvas algebraicas, requiere un poco más de cuidado, y es desarrollado en las secciones 2, 3 y 4. En la sección 5 explicamos brevemente el problema de contar puntos enteros y racionales en hipersuperficies algebraicas y curvas algebraicas del plano proyectivo, principalmente mencionando los resultados de Heath-Brown de [**17**].

[1] En verdad, por métodos de naturaleza distinta a los empleados en [**6**], ya se conocían estimaciones uniformes en el grado de las curvas algebraicas; por ejemplo, en [**5**], se emplea la técnica de la criba grande para acotar los puntos enteros de las curvas algebraicas de grado a lo sumo d, obteniendo una cota uniforme en el grado de las curvas.

1. Puntos enteros en curvas analíticas

El ejemplo opuesto a una curva algebraica podría decirse que es una curva trascendente (por ejemplo el gráfico de $f(x) := \exp(x)$). Es de esperar que una curva así tenga pocos puntos enteros. En esta sección vamos a estudiar los puntos enteros de curvas trascendentes analíticas. Bajo estas condiciones, el uso del Lema 1.4 nos permite concluir una estimación de los puntos enteros de Γ. Primero, establecemos un lema sencillo.

LEMA 2.1. *Sea $f : [0, N] \to \mathbb{R}$ una función analítica trascendente. Sea Γ el gráfico de f. Si γ es una curva algebraica plana real* [2]*, entonces $\#(\Gamma \cap \gamma) < \infty$.*

DEMOSTRACIÓN. Sea $P \in \mathbb{R}[x, y]$ que define a la curva algebraica plana γ. Como f es trascendente, se tiene que $g(x) := P(x, f(x))$ no es idénticamente nula en $[0, N]$. Supongamos que $\#(\Gamma \cap \gamma) = \infty$,. Luego existe $(x_n, y_n)_{n\in\mathbb{N}}$ una sucesión en $\Gamma \cap \gamma$ que tiene límite $(x, y) \in \Gamma \cap \gamma$ y (x, y) es un punto de acumulación de $\Gamma \cap \gamma$. Como $(x_n, y_n) \in \gamma$, se tiene que $P(x_n, y_n) = 0$. Adicionalmente $(x_n, y_n) \in \Gamma$ con lo que $y_n = f(x_n)$. Esto nos dice que $P(x_n, f(x_n)) = 0$ para todo n. De manera similar se tiene que $P(x, f(x)) = 0$. Dado que $g(x) = P(x, f(x))$, tenemos que g es analítica en $[0, N]$ y se anula en un conjunto con un punto de acumulación. Este hecho implica que g es idéntícamente nula en $[0, N]$, lo que contradice que f sea trascendente. □

OBSERVACIÓN 2.1. *Fijado $d \in \mathbb{N}$ y Γ el gráfico de $f : [0, N] \to \mathbb{R}$ una función analítica trascendente, se tiene que existe $\rho(\Gamma, d)$ tal que $\#(\Gamma \cap \gamma) \leq \rho(f, d)$ para toda γ una curva algebraica plana real de grado d, es decir, existe una cota uniforme para la intersección de Γ con cualquier curva algebraica de grado d.*

Ahora podemos estudiar las dilataciones de una curva analítica no algebraica:

TEOREMA 2.1. *Sea $f : I \to \mathbb{R}$ una función analítica trascendente, con I un intervalo cerrado y acotado. Sea Γ el gráfico de f. Sea $\varepsilon > 0$. Entonces para todo $t \geq 1$ se tiene*

$$\#(t\Gamma)(\mathbb{Z}) = O_{f,\epsilon}(t^{\varepsilon}).$$

DEMOSTRACIÓN. sea $I \subseteq [0, N]$. Por el lema 2.1, tenemos que para toda γ una curva algebraica real, se tiene $\#(\Gamma \cap \gamma) < \infty$. Por la observación 2.1 existe un número $\rho(f, d)$ tal que $\#(\Gamma \cap \gamma) \leq \rho(f, d)$ para toda γ una curva algebraica plana real de grado d.

Del Lema 1.4, sabemos una cota para la cantidad de curvas algebraicas de grado a lo sumo d con las que podemos cubrir $\Gamma(\mathbb{Z})$. Como cada una de estas curvas del cubrimiento tiene a lo sumo $\rho(f, d)$ puntos de Γ, concluímos

$$\#\Gamma(\mathbb{Z}) \leq \rho(f, d) \left[3(||f||_{N,D-1}^{\frac{1}{2}} N)^{\frac{8}{3(d+3)}} + 1\right].$$

Ahora, sea $t \geq 1$. Como $t\Gamma = \{(tx, tf(x)) : x \in I\}$, haciendo el cambio de variable $tx = y$, obtenemos que $t\Gamma$ es el gráfico de la función $g(y) := tf(\frac{y}{t})$ para $y \in tI$. Dado que la norma $|| \cdot ||_{N,D-1}$ es invariante por dilataciones (ver comentarios posteriores a la definición 1.1), resulta $||g||_{tI,D-1} = ||f||_{I,D-1}$ lo que da, usando la cota anterior con el gráfico de g:

$$\#(t\Gamma)(\mathbb{Z}) \leq \rho(f, d) \left[3||f||_{N,D-1}^{\frac{2}{3(d+3)}} t^{\frac{8}{3(d+3)}} N^{\frac{8}{3(d+3)}} + 1\right].$$

Multiplicando y dividiendo por t^{ε} y usando el hecho que $t^{-\varepsilon} \leq 1$ para $t \geq 1$ y $\varepsilon > 0$, concluímos el teorema. □

OBSERVACIÓN 2.2. *Notemos que el Teorema 2.1 nos da, para $t = N$, una cota para la cantidad de puntos racionales $(\frac{a}{N}, \frac{b}{N}) \in \Gamma$, ya que, si consideramos (Nc_1, Nc_2) con $(c_1, c_2) \in \Gamma(\mathbb{Z})$, entonces $(\frac{c_1}{N}, \frac{c_2}{N})$ es un punto racional de Γ, de denominador N. Recíprocamente, todo punto racional de Γ de la forma $c = (\frac{a}{N}, \frac{b}{N})$ verifica $Nc \in N\Gamma(\mathbb{Z})$. Luego, el Teorema 2.1 es equivalente a*

$$\#\{(\tfrac{a}{N}, \tfrac{b}{N}) \in \Gamma : a, b \in \mathbb{Z}\} = O_{f,\varepsilon}(N^{\varepsilon}).$$

OBSERVACIÓN 2.3. *Los coeficientes $\rho(f, d)$ que aparecieron en la demostración, no fueron construídos de manera explícita. En este sentido, el Teorema 2.1 no da una cota efectiva.*

A continuación, damos tres ejemplos que estudian distintos aspectos del Teorema 2.1. El primer ejemplo es de [**6**], mientras que el segundo y tercer ejemplo son de [**39**].

[2] Esto quiere decir que la ecuación $f(x, y) = 0$ que la define viene dada por un polinomio con coeficientes en $\mathbb{R}$

EJEMPLO 2.1 (Bombieri-Pila). *Destaquemos que los $\rho(f,d)$, para una f dada, pueden crecer arbitrariamente, con lo que no se puede obtener una cota independiente de d. Por ejemplo, sea $(d_n)_n$ una sucesión estrictamente creciente de enteros positivos. Definamos la expresión formal*

$$f(x) := \sum_{i=0}^{+\infty} 2^{-d_i} \prod_{k=1}^{d_i} (x - 2^{-k}) = \sum_{i=0}^{+\infty} f_i(x).$$

donde

$$f_i(x) := 2^{-d_i} \prod_{k=1}^{d_i} (x - 2^{-k}).$$

Veamos que esta expresión formal tiene sentido analítico en el disco unitario. Para ello, primero notemos que si $|z| \leq 1$, entonces $|z - 2^{-k}|$ alcanza el valor máximo con $z = -\frac{1}{2}$. Luego, la función $f_i(x)$ de la suma verifica la cota

$$|f_i(x)| \leq 2^{-d_i} \prod_{k=1}^{d_i} (\frac{1}{2} + 2^{-k}) = 2^{-2d_i} \prod_{k=1}^{d_i} (1 + 2^{-(k-1)}).$$

Para acotar el producto, tomamos logaritmo y usamos $\log(x) \leq x - 1$.

$$\sum_{k=1}^{d_i} \log(1 + 2^{-k+1}) \leq \sum_{k=1}^{d_i} 2^{-k+1} = 2[1 - 2^{-d_i}].$$

Podemos entonces acotar $|f_i(x)|$ como

$$|f_i(x)| \leq 2^{-2d_i}[e^2 e^{2^{-d_i+1}}] \leq A 2^{-d_i},$$

donde A es una constante positiva. Se concluye que las f_i convergen uniformemente para $|z| \leq 1$ y por lo tanto f define una función analítica en el disco. Ahora, consideremos γ_i la curva algebraica plana definida por la ecuación

$$g_i(x,y) := \sum_{j=0}^{i} f_j(x) - y.$$

Resulta que g_i tiene grado d_i y, como

$$f(2^{-j}) = \sum_{j=0}^{i} f_i(x) \text{ si } 1 \leq j \leq d_{i+1},$$

tenemos que $g_i(2^{-j}, f(2^{-j})) = 0$ para tales j. Luego $\rho(f,d_i) \geq \#(\Gamma \cap \gamma_i) \geq d_{i+1}$. Como los d_i pueden crecer arbitrariamente, los $\rho(f,d_i)$ también.

EJEMPLO 2.2 (Pila). *En general, no se puede mejorar la cota del Teorema 2.1. Sea $\varepsilon : [1,\infty)$ una función estrictamente decreciente tal que $\varepsilon(t) \to 0$ cuando $t \to +\infty$. Definimos una sucesión $\{N_j\}_{j\in\mathbb{N}}$ de enteros positivos de manera inductiva, de la siguiente manera: sea $N_0 = 1$. Suponiendo definidos $N_0, \cdots, N_{k-1}$, sea N_k definido de manera que $N_k \geq k, N_{k-1} \mid N_k$, y $\varepsilon(N_k^{N_{k-1}} 2^{k-1}) \leq \frac{1}{2N_{k-1}}$. Para cada $k \in \mathbb{N}$, sea $t_k := N_k^{N_{k-1}} 2^{k-1}, X_k := \{\frac{i}{N_k} : i \in \mathbb{Z}, 0 \leq i \leq N_k\}$. Definamos*

$$f(x) := \sum_{k=0}^{+\infty} 2^{-k} \prod_{z\in X_k} (x - z).$$

Al igual que en el ejemplo 2.1, la función f es analítica en $[0,1]$. Sea Γ el gráfico de f. Si $x \in X_k$, entonces $N_k x \in \mathbb{Z}$ y $t_k f(x) \in \mathbb{Z}$, con lo que

$$\#(t_k \Gamma(\mathbb{Z})) \geq N_k \geq \exp(\frac{\log(t_k)}{2N_{k-1}}) = t_k^{\frac{1}{2N_{k-1}}} \geq t_k^{\varepsilon(t_k)}.$$

EJEMPLO 2.3 (Pila). *Este ejemplo muestra que es clave la compacidad del intervalo I en el Teorema 2.1. Sea, como en el Ejemplo 2.2, $\varepsilon : [1,+\infty) \to \mathbb{R}$ una función estrictamente decreciente con $\varepsilon(t) \to 0$ cuando $t \to +\infty$. Definimos la secuencia $\{t_j : j \in \mathbb{N}\}$ de enteros positivos, inductivamente, de la siguiente manera. Sea $t_0 = 1$. Supongamos que $t_0, t_1, \cdots, t_{k-1}$ están definidos, luego sea t_k definido de manera que $t_{k-1} \mid t_k$ y $\varepsilon(t_k) \leq 2^{-k-1}$. Construímos una función $f : (0,1]$ analítica trascendente de manera que en el intervalo $(2^{-k-1,2^{-k}}]$, si $t_k x \in \mathbb{Z}$, entonces $t_k f(x) \in \mathbb{Z}$ (ver el ejemplo de [**56**]). Sea Γ el gráfico de f. Entonces*

$$\#t_k\Gamma(\mathbb{Z}) \geq \sum_{j=0}^{k} \frac{t_j}{2^{j+1}} \geq \frac{t_k}{2^{k+1}} \geq \varepsilon(t_k)t_k,$$

con lo que para k suficientemente grande, se concluye $\#t_k\Gamma(\mathbb{Z}) \geq t_k^{\varepsilon(t_k)}$.

Destaquemos que el Teorema 2.1 tiene una versión para los puntos racionales de altura acotada, probada por Pila en [**41**].

TEOREMA 2.2 (Pila). *Sea $f : I \to \mathbb{R}$ una función analítica trascendente, con I un intervalo cerrado y acotado. Sea Γ el gráfico de f. Sea $\varepsilon > 0$. Entonces para todo $N \geq 1$ entero positivo se tiene*

$$\#((t\Gamma)(\mathbb{Q}), N) = O_{f,\epsilon}(N^{\varepsilon}).$$

No vamos a probar el Teorema 2.2; lo mencionamos porque nos va a interesar una generalización del mismo en el capítulo 4. No obstante, las técnicas empleadas en la demostración del Teorema 2.2 son similares a las del trabajo [**6**].

2. Puntos enteros en curvas algebraicas

Sea C una curva algebraica plana. Podemos particionar a la curva algebraica C en tramos suficientemente suaves, en los que se pueda aplicar el Lema 1.7 y luego tratar de convertir esta información algebraica en información aritmética. Por medio de alguna hipótesis de irreducibilidad, las intersecciones de γ con algunas curvas algebraicas van a ser finitas y acotadas uniformemente. Luego, un tramo de γ va a tener pocas intersecciones con algunas curvas algebraicas. Si adicionalmente estas curvas algebraicas cubren los puntos enteros de γ en el tramo suave, podríamos esperar que, usando los métodos anteriores, tengamos una cota para los puntos enteros del tramo. Conociendo la cantidad de tramos suaves en los que podemos particionar la curva γ, tenemos una estimación para $\#(\gamma(\mathbb{Z}) \cap [0, N]^2)$. Estos pasos se realizan con suma facilidad con las técnicas que desarrollamos en el capítulo 1.

En lo que sigue, vamos a requerir una estimación para los puntos de intersección entre dos curvas algebraicas sin factores comunes. Para ello, vamos a recurrir al teorema de Bézout, que recordamos:

TEOREMA 2.3 (Teorema de Bézout). *Sean C, C' dos curvas de grado d y d' respectivamente, que no poseen factores comunes. Entonces, C y C' tienen a lo sumo dd' puntos en común.*

Sea $\mathcal{D}$ un entero positivo. Consideremos $\mathcal{M}$ un conjunto de monomios de modo que las curvas definidas en $\mathcal{M}$ no sean factores de C. Consideramos $F \in \mathbb{R}[x, y]$ de grado $d \geq 2$.

Sea $j_F = (j_1^F, j_2^F)$ el índice del monomio de grado d de mayor grado en la indeterminada y que aparece en F. Por ejemplo, si $F(x, y) = xy^3 + x^2y^2 + x^3y + x + y$, se tiene $j_F = (1, 3)$. Para conseguir que F no aparezca como componente en una curva G, lo que para nosotros será equivalente a la condición $F(x, y) \nmid G(x, y)$, ya que asumiremos F geométricamente irreducible, podemos pedir que los monomios de G no sean divisibles por $(x, y)^{j_F}$. Por ello, definamos

$$\mathcal{M}_F(\mathcal{D}) = \{x^{j_1}y^{j_2} : d \leq j_1 + j_2 \leq \mathcal{D}, x^{j_1^F}y^{j_2^F} \nmid x^{j_1}y^{j_2}\}. \tag{2.1}$$

Notemos que $d \leq \mathcal{D}$. Tenemos:

PROPOSICIÓN 2.1. *Sea $G(x, y) \in \mathbb{R}[x, y]$ tal que G está definido en $\mathcal{M}_F(\mathcal{D})$. Entonces $F(x, y) \nmid G(x, y)$.*
En particular, por el teorema de Bézout, las curvas que definen F y G se intersecan en a lo sumo $d\mathcal{D}$ puntos.

DEMOSTRACIÓN. Supongamos que $G(x, y) = H(x, y)F(x, y)$ para algún $H \in \mathbb{R}[x, y]$. Usando una notación análoga, definimos j_H. Tenemos que $(x, y)^{j_H}(x, y)^{j_F}$ es un monomio que aparece en $G(x, y)$. Como G está definido en $\mathcal{M}_F(\mathcal{D})$, debe tenerse $(x, y)^{j_H}(x, y)^{j_F} \in \mathcal{M}_F(\mathcal{D})$ luego $(x, y)^{j_F} \nmid (x, y)^{j_H}(x, y)^{j_F}$, lo que es una contradicción. □

En la demostración no usamos que $d \leq j_1 + j_2$; esto se debe a que esta condición es de naturaleza técnica. En lo que sigue, vamos a suponer $\mathcal{D} \geq 2d$.

Sea $h \geq d$. La cantidad de monomios de grado h no divisibles por un monomio fijo de grado d es exactamente d. En efecto, si el monomio fijo es $(x, y)^{(i,j)} = x^iy^j$ con $i + j = d$, para que $(x, y)^{(i,j)} \nmid (x,)^{(i',j')}$ con $i' + j' = h$, debe ocurrir que $i > i'$ o $j > j'$. Si $i > i'$, hay i posibilidades para i', a saber $0, 1, \cdots, i-1$. Como $i' + j' = h$, i' determina j' luego hay i posibilidades. De manera análoga, hay j posibilidades para $j > j'$. Finalmente, si $i > i', j > j'$, entonces $d = i + j > i' + j' = h \geq d$ lo que es absurdo. Entonces el principio de inclusión-exclusión nos dice que hay $i + j = d$ monomios de grado h no divisibles por $(x, y)^{(i,j)}$.

Sea $f \in \mathcal{C}^\infty(I)$ con I un intervalo cerrado y acotado contenido en $[0, N]$ tal que $F(x, f(x)) = 0$ y $|f'(x)| \leq 1$, es decir, consideramos una función muy suave que está contenida en la curva algebraica definida por F, con velocidad acotada. Sea Γ el gráfico de f. El corolario en la Proposición 2.1 nos dice que toda curva algebraica plana definida en $\mathcal{M}_F(\mathcal{D})$ tiene a lo sumo $d\mathcal{D}$ intersecciones con la curva definida por $F(x, y) = 0$. Luego toda curva algebraica definida en $\mathcal{M}_F(\mathcal{D})$ se interseca con el gráfico de f en a lo sumo $d\mathcal{D}$ puntos. Puesto que el Lema generalizado 1.7 da una cota superior para la cantidad de curvas algebraicas definidas en $\mathcal{M}$ que cubren $\Gamma(\mathbb{Z})$, podemos concluir el siguiente resultado:

TEOREMA 2.4. *Sea Γ el gráfico de $f \in \mathcal{C}^\infty(I)$ con I intervalo cerrado y acotado, contenido en $[0, N]$. Supongamos que $|f'(x)| \leq 1$ para $x \in I$ y que está contenida en una curva algebraica definida por la ecuación $F(x, y) = 0$. Entonces se tiene la estimación*

$$\#\Gamma(\mathbb{Z}) \leq d\mathcal{D}[(D^p||f||^q_{N,D-1})^{\frac{2}{D(D-1)}} N^{\frac{2p}{D(D-1)}} + 1].$$

OBSERVACIÓN 2.4. *Lo que dice el Teorema 2.4 es que, siguiendo la notación del Teorema 2.1, si $\rho(f, d)$ es una cota superior para $\#(\Gamma \cap \gamma)$ con γ una curva algebraica definida en $\mathcal{M}$, podemos tomar $\rho(f, d) \leq d\mathcal{D}$ (ver el Teorema 2.1 y la Observación 2.3).*

OBSERVACIÓN 2.5. *En caso que $|f'(x)|$ no está acotada por 1 para todo x, si la derivada no se anula, podemos parametrizar por medio de $x = g(y)$ con $g = f^{-1}$ y $|g'(y)| \leq 1$ en un intervalo adecuado y aplicar ahora el Teorema 2.4.*

Con lo que hicimos hasta ahora detectamos curvas algebraicas que no tienen demasiadas intersecciones con la curva algebraica original, hecho reflejado en el Teorema 2.4. Sin embargo, las cotas que obtenemos son locales: dependen de la parametrización de la curva algebraica en una región. Esto nos dice que, para Γ una curva algebraica irreducible de grado $d \geq 2$, el Teorema 2.4 nos permite obtener una estimación de $\#(\Gamma(\mathbb{Z}) \cap [0, N]^2)$, **pero a priori esta estimación depende de** Γ (las cotas dependen de las parametrizaciones de los tramos suaves de Γ). El objetivo de la próxima sección es probar que, si una curva que es el gráfico de una función f muy suave, está contenida en una curva algebraica irreducible, entonces es posible obtener cotas independientes de la parametrización. Esto será posible debido a que vamos a poder particionar suficientemente el intervalo I para luego controlar los tamaños de las derivadas $|f^i(x)|$. Por otro lado, también resultará clave que si $|f^i(x)|$ es grande, entonces el intervalo en el que esto ocurre es muy chico. Lo que está ocurriendo aquí es que, al estar contenida f en una curva algebraica, hay una estructura adicional que le impone restricciones adicionales a las parametrizaciones.

3. Control de las oscilaciones de una curva parametrizada

Comenzamos con un lema:

LEMA 2.2. *Supongamos que $G(x, y) \in \mathbb{R}[x, y]$ es un polinomio geométricamente irreducible y de grado b. Sea $g \in \mathcal{C}^\infty(I)$ tal que $G(x, g) = 0$. Supongamos que g no es un polinomio de grado menor o igual a k. Entonces para cada $k \geq 1$ y $c \in \mathbb{R}$ la ecuación*

$$g^{(k)}(x) = c \tag{3.1}$$

tiene a lo sumo $b(b-1)(2k-1)$ soluciones reales.

DEMOSTRACIÓN. Primero notemos que la hipótesis de que g no es un polinomio de grado menor o igual a k es necesaria, pues en caso contrario, la ecuación del lema tiene infinitas soluciones. Lo que vamos a hacer es mostrar que las soluciones de la ecuación (3.1) están contenidas en la intersección de dos curvas algebraicas de grado que es posible acotar superiormente. Utilizando el teorema de Bézout, obtenemos la cota del lema.

Como $G(x, g) = 0$, si derivamos, tenemos

$$G_x + G_y g' = 0. \tag{3.2}$$

Volviendo a derivar, tenemos

$$G_{xx} + G_{yx}g' + G_{xy}g' + G_{yy}(g')^2 + G_y g'' = G_{xx} + 2G_{xy}g' + G_{yy}(g')^2 + G_y g'' = 0. \tag{3.3}$$

Multiplicando por G_y^2 en la ecuación (3.3) y luego sustituyendo (3.2), obtenemos

$$G_{xx} - 2G_x G_y G_{xy} + G_x^2 G_{yy} + G_y^3 g'' = 0,$$

lo que puede escribirse de manera más compacta como

$$H_2(x,g) + G_y^3 g'' = 0,$$

con $P_2(x,y) = G_{xx}(x,y) - 2G_x(x,y)G_y(x,y)G_{xy}(x,y) + G_x^2(x,y)G_{yy}(x,y)$.

En general, por inducción, tenemos que para cada k, existen $H_k \in \mathbb{R}[x,y]$ y $a_k \in \mathbb{N}$ tales que

$$H_k + G_y^{a_k} g^{(k)} = 0. \tag{3.4}$$

Derivando la ecuación (3.4), obtenemos:

$$(H_k)_x + (H_k)_y g' + a_k G_y^{a_k-1}\left[G_{yx} + G_{yy}g'\right] g^{(k)} + G_y^{a_k} g^{(k+1)} = 0. \tag{3.5}$$

Multiplicando G_y^2 en la ecuación (3.5) y luego sustituyendo la ecuación (3.4), tenemos

$$G_y(H_{kx}G_y - H_{ky}G_x) - a_k H_k(G_{yx}G_y - G_{yy}G_x) + G_y^{a_k+2} g^{(k+1)} = 0.$$

Poniendo $h_k = \deg(H_k)$, la última expresión da una recursión para a_k, h_k

$$\begin{cases} a_{k+1} = a_k + 2 \\ a_1 = 1 \end{cases} \qquad \begin{cases} h_{k+1} \leq h_k + 2b - 3 \\ h_1 = b-1 \end{cases}$$

cuyas soluciones son $a_k = 2k-1$ y $h_k \leq (2k-1)(b-1) - k < (2k-1)(b-1)$. En virtud del razonamiento que realizamos, si x verifica $g^{(k)}(x) = c$ entonces en particular existe y (por ejemplo $y = g(x)$) tal que

$$\begin{cases} G(x,y) = 0 \\ H_k + G_y^{a_k} c = 0 \end{cases},$$

es decir, x es la coordenada de alguna de los puntos en la intersección de las curvas algebraicas $G(x,y) = 0$ y $H_k + G_y^{a_k} c = 0$. Como g no es un polinomio de grado menor o igual a k, $g^{(k)}$ no es constante. La irreducibilidad de G y el teorema de Bézout nos dicen que $G(x,y) = 0$ y $H_k + G_y^{a_k} c = 0$ tienen a lo sumo $b(b-1)(2k-1)$ puntos en común, lo que da la cota del lema. □

El siguiente lema pone en manifiesto que las curvas suaves contenidas en una curva algebraica poseen cierta "uniformidad" en sus derivadas.

LEMA 2.3. *Sean $g \in \mathcal{C}^\infty(I)$ y $G(x,y) \in \mathbb{R}[x,y]$ un polinomio irreducible de grado $b \geq 2$, satisfaciendo $G(x,g) = 0$. Sean $(A_l)_l$ números reales positivos para $l = 1, \cdots, k$. Entonces podemos dividir el intervalo I en a lo sumo $2b^2k^2$ subintervalos I_ν tales que para cada ν y l, tenemos que se cumple alguna de las siguientes condiciones:*

(1) $|g^{(l)}(x)| \leq A_l,\ \forall x \in I_\nu$.
(2) $|g^{(l)}(x)| \geq A_l,\ \forall x \in I_\nu$.

DEMOSTRACIÓN. Vamos a separar dos casos:

Caso 1: Si g no es un polinomio de grado menor o igual a b, consideramos las ecuaciones

$$g^{(i)}(x) = \pm A_i,\ i = 1, \cdots, k. \tag{3.6}$$

Por el Lema 2.2, para cada i, hay a lo sumo $2b(b-1)(2i-1)$ soluciones a la ecuación (3.6), siendo el factor 2 debido a las dos posibilidades de signo por cada A_i. Luego hay a lo sumo

$$\sum_{i=1}^{k} 2b(b-1)(2i-1) = 2b(b-1)k^2$$

valores de x que verifican para algún i la condición (3.6). Estos puntos determinan una división del intervalo I en a lo sumo $2b(b-1)k^2 + 1 \leq 2b^2k^2$ intervalos.

Caso 2: Si g es un polinomio de grado a lo sumo b, consideramos las ecuaciones

$$g^{(i)(x)} = \pm A_i,\ i = 1, \cdots, b-1. \tag{3.7}$$

Para cada i, el grado de $g^{(i)}$ es $\deg(g) - i \leq b - i$, luego cada ecuación anterior tiene a lo sumo $b - i$ posibles soluciones. De nuevo, por los dos signos posibles de cada A_i, tenemos

$$\sum_{i=1}^{b-1} 2(b-i) = 2b(b-1) - b(b-1) = b(b-1) \leq 2b(b-1)k^2$$

posibles soluciones a todas las ecuaciones (3.7). Estos puntos determinan una división del intervalo I en a lo sumo $2b^2k^2$ intervalos.

En cualquier caso, notamos que para cada uno de los intervalos $I_\nu = [a_\nu, b_\nu]$ se cumple la condición del Lema 2.3, pues en caso contrario existen x, y puntos interiores del intervalo tales que $g^{(l)}(x) > A_l$ ó $g^{(l)}(x) < -A_l$, $-A_l < g^{(l)}(y) < A_l$. Entonces, la continuidad de $g^{(l)}$ nos permite aplicar el teorema del valor medio, que nos dice que existe z un punto interior a I_ν con $g^{(l)}(z) = \pm A_l$, lo que contradice la construcción de los intervalos I_ν que contenían de manera exhaustiva los puntos que verifican las ecuaciones (3.6) y (3.7). □

Lo que probaremos a continuación es que si una curva algebraica es localmente el gráfico de una función $f : I \to \mathbb{R}$ muy suave y la función f tiene norma $||\cdot||_{N,k}$ muy grande en un intervalo, entonces este intervalo tiene que ser pequeño.

LEMA 2.4. *Supongamos que $I = [a, b], g \in \mathcal{C}^k([a, b])$ y para algún A, N reales positivos, se tiene*

$$|g^{(i)}(x)| \leq i! A^{\frac{i}{k}} N^{1-i},\ \forall\ x \in I, i = 0, \cdots, k-1,$$
$$|g^{(k)}(x)| \geq k! A N^{1-k},\ \forall x \in I.$$

Entonces $|I| \leq 2A^{-\frac{1}{k}} N$.

DEMOSTRACIÓN. Usando el desarrollo de Taylor de g centrado en a, tenemos que existe $\xi \in (a, b)$ tal que

$$g(b) = g(a) + \sum_{i=1}^{k-1} \frac{g^{(i)}(a)}{i!}(b-a)^i + \frac{g^{(k)}(\xi)}{k!}(b-a)^k.$$

Despejando y usando las hipótesis del lema, y notando que $|b - a| = |I|$ tenemos:

$$|I|^k A N^{1-k} \leq |\frac{g^{(k)}(\xi)}{k!}||I|^k \leq |g(b) - g(a)| + \sum_{i=1}^{k-1} |\frac{g^{(i)}(a)}{i!}||I|^i \leq 2N + \sum_{i=1}^{k-1} |I|^i A^{\frac{i}{k}} N^{1-i}.$$

Sea $\lambda = \left(\frac{|I|}{N}\right) A^{\frac{1}{k}}$. Simplificando una N en cada término, obtenemos la desigualdad

$$\lambda^k \leq \sum_{i=1}^{k-1} \lambda^i + 2. \tag{3.8}$$

Lo que queremos probar es $\lambda \leq 2$. Dado que $\lambda^k - 1 = (\lambda - 1)\sum_{i=0}^{k-1} \lambda^i$, (3.8) nos da

$$(\lambda - 1)\sum_{i=0}^{k-1} \lambda^i \leq \sum_{i=0}^{k-1} \lambda^i.$$

Como $\lambda > 0$, simplificamos sin cambiar el sentido de la desigualdad, quedando

$$\lambda - 1 \leq 1,$$

lo que da $\lambda \leq 2$. □

Ahora estamos en condiciones de probar el siguiente teorema, que esencialmente refleja que es posible obtener una cota uniforme para la cantidad de puntos enteros del gráfico de una curva suave siempre que ésta se encuentre contenida en una curva algebraica.

TEOREMA 2.5. *Sean d, N enteros positivo. Definimos $G(d) = G(d, N)$ al máximo número de puntos enteros que puede contener el gráfico de una función $g \in \mathcal{C}^\infty(I)$ con I un intervalo cerrado y acotado de longitud a lo sumo N, con $|g'(x)| \leq 1$, que verifica que existe un polinomio $P(x, y) \in \mathbb{R}[x, y]$ no nulo, geométricamente irreducible, tal que $P(x, g(x)) = 0$ para todo $x \in I$[3]. Se tiene que para $N \geq \exp(d^6)$, se verifica*

$$G(N) \leq N^{\frac{1}{d}} \exp(11\sqrt{d \log(N) \log\log(N)}).$$

DEMOSTRACIÓN. Sea $g \in \mathcal{C}^\infty(I)$ con I un intervalo cerrado y acotado de longitud a lo sumo N y $|g'(x)| \leq 1$ para todo $x \in I$, tal que existe $G(x, y) \in \mathbb{R}[x, y]$ no nulo, que verifica $G(x, g(x)) = 0$ para todo $x \in I$, y tal que si Γ es el gráfico de g, se tiene $\#\Gamma(\mathbb{Z}) = G(N)$. Supongamos primero que $|g(x)| \leq N$. Usando el Lema 2.3, vamos a particionar el intervalo I donde está definida g de manera que, o bien en cada subintervalo la norma de g es pequeña, o bien la norma es grande. Usando el Lema 2.4, vamos a tener que los intervalos donde la norma es grande son pequeños. Luego usamos el método del determinante con una familia de monomios $\mathcal{M}$ adecuada, aplicado al gráfico de g restringido a cada subintervalo.

Sea $\mathcal{D} \geq 2d$ que elegiremos posteriormente. Consideremos el conjunto de monomios $\mathcal{M}_G(\mathcal{D})$ (ver expresión (2.1)) y sea $D = \#\mathcal{M}_F(\mathcal{D})$. Sea $A \geq 1$. Utilizando el Lema 2.3, podemos dividir el intervalo I en a lo sumo $2d^2(D-1)^2 \leq 2d^2D^2$ subintervalos I_ν tales que para cada I_ν y cada $l = 1, \cdots, D-1$ se tienen alguna de las condiciones

(1) $|g^{(l)}(x)| \leq l! A^{\frac{l}{D-1}} N^{1-l}$, $\forall x \in I_\nu$.
(2) $|g^{(l)}(x)| \geq l! A^{\frac{l}{D-1}} N^{1-l}$, $\forall x \in I_\nu$.

Si la condición (1) se cumple siempre, debido a que $|g(x)| \leq N$, se tiene $||g||_{N,D-1} \leq A$. Aplicando el Teorema 2.4, concluímos

$$\Gamma(\mathbb{Z}) \leq d\mathcal{D}\left[\left(D^p ||g||_{N,D-1}^q\right)^{\frac{2}{D(D-1)}} N^{\frac{2p}{D(D-1)}} + 1\right], \tag{3.9}$$

donde, si $J = \{j = (j_1, j_2) : x^{j_1} y^{j_2} \in \mathcal{M}_F(\mathcal{D})\}$, se tienen $p = \sum_{j=(j_1,j_2)\in J}(j_1 + j_2), q = \sum_{j=(j_1,j_2)\in J} j_2$.

Si la condición (1) no se cumple siempre, existe $k < D-1$ tal que

$$|g^{(l)}(x)| \leq l! A^{\frac{l}{D-1}} N^{1-l}, \forall x \in I_\nu, \forall l < k,$$
$$|g^{(k)}(x)| \geq k! A^{\frac{k}{D-1}} N^{1-k}, \forall x \in I_\nu.$$

En este caso, usando el Lema 2.4 con $A^{\frac{k}{D-1}}$, obtenemos $|I_\nu| \leq 2A^{-\frac{1}{D-1}} N$.

Denotemos Γ_ν al gráfico de $g|_{I_\nu}$.

- Para los intervalos I_ν en los que $||g|_{I_\nu}||_{N,D-1} \leq A$, usamos la cota (3.9) para obtener:
$$\Gamma_\nu(\mathbb{Z}) \leq d\mathcal{D}\left[\left(D^p ||g||_{N,D-1}^q\right)^{\frac{2}{D(D-1)}} N^{\frac{2p}{D(D-1)}} + 1\right].$$
- Para los intervalos I_ν en los que $||g|_{I_\nu}||_{N,D-1} > A$, como $|I_\nu| \leq 2A^{-\frac{1}{D-1}} N$, acotamos $\#\Gamma_\nu(\mathbb{Z})$ por $G(2A^{-\frac{1}{D-1}} N)$.

Usando que $\Gamma = \bigcup_\nu \Gamma_\nu = (\bigcup_{\nu:||g|_{I_\nu}||_{N,D-1}\leq A} \Gamma_\nu) \cup (\bigcup_{\nu:||g|_{I_\nu}||_{N,D-1}>A} \Gamma_\nu)$, obtenemos:

$$G(N) \leq \sum_{\nu:||g|_{I_\nu}||_{N,D-1}\leq A} \#\Gamma_\nu(\mathbb{Z}) + \sum_{\nu:||g|_{I_\nu}||_{N,D-1}>A} \#\Gamma_\nu(\mathbb{Z}). \tag{3.10}$$

Cada suma la acotamos respectivamente:

$$\sum_{\nu:||g|_{I_\nu}||_{N,D-1}\leq A} d\delta\left[\left(D^p ||g|_{I_\nu}||_{N,D-1}^q\right)^{\frac{2}{D(D-1)}} N^{\frac{2p}{D(D-1)}} + 1\right] + \sum_{\nu:||g|_{I_\nu}||_{N,D-1}>A} G(2A^{-\frac{1}{D-1}} N). \tag{3.11}$$

La última expresión la podemos acotar por

$$2d^2D^2 2d\mathcal{D}(D^p A^q)^{\frac{2}{D(D-1)}} N^{\frac{2p}{D(D-1)}} + 2d^2D^2 G(2A^{-\frac{1}{D-1}} N), \tag{3.12}$$

con lo que podemos concluir la desigualdad

[3]La razón por la cual $G(N)$ es finito se debe a que $|g'(x)| \leq 1$, por lo tanto Γ tiene longitud acotada (uniformemente).

$$G(N) \leq 2d^2D^2 2d\mathcal{D}(D^pA^q)^{\frac{2}{D(D-1)}} N^{\frac{2p}{D(D-1)}} + 2d^2D^2 G(2A^{-\frac{1}{D-1}}N). \tag{3.13}$$

Recordemos que vamos a elegir $\mathcal{D} \geq 2d$. Luego, notemos que los parámetros $D, p, \mathcal{D}$ verifican:

$$D = d(\mathcal{D} - d + 1) \leq d\mathcal{D},$$
$$\frac{1}{d} \leq \frac{2p}{\mathcal{D}(\mathcal{D}-1)} \leq \frac{1}{d} + \frac{4}{\mathcal{D}}.$$

Si escribimos

$$H = 4d^5\delta^3(D^pA^q)^{\frac{2}{D(D-1)}},$$
$$K = 2d^4\delta^2,$$
$$\alpha = \frac{2p}{D(D-1)},$$
$$\lambda = 2A^{-\frac{1}{D-1}},$$

la expresión (3.13) permite concluir la desigualdad

$$G(N) \leq HN^\alpha + KG(\lambda N). \tag{3.14}$$

Si $\lambda N < 1$, tenemos que $G(\lambda N)$ mide el máximo de puntos de coordenadas enteras en un cuadrado de lado a lo sumo λN de los gráficos de funciones $g \in \mathcal{C}^\infty(I)$, con I un intervalo cerrado y acotado de longitud a lo sumo λN, donde el gráfico de g está contenido en una curva algebraica plana. Dado que estos puntos son de la forma $(t, g(t))$ con $t \in I$ un invervalo cerrado de longitud a lo sumo λN, los puntos de coordenadas enteras tienen $t \in \mathbb{Z}$ y como un intervalo de longitud $\lambda N < 1$ tiene a lo sumo un entero, deducimos $G(\lambda N) < 1$. Entonces

$$G(N) \leq HN^\alpha + KG(\lambda N) \leq HN^\alpha + K.$$

En general, si $\lambda^{n-1}N \geq 1$, iteramos la cota (3.14), con $G(\lambda N)$ en lugar de $G(N)$, para obtener

$$G(N) \leq HN^\alpha(1 + K\lambda^\alpha + \cdots + (K\lambda^\alpha)^{n-1}) + K^nG(\lambda^n N). \tag{3.15}$$

Sea λ tal que $K\lambda^\alpha = \frac{1}{2}$, es decir,

$$\lambda = \left(\frac{1}{2K}\right)^{\frac{1}{\alpha}} = \left(4d^4\delta^2\right)^{-\frac{D(D-1)}{2p}}.$$

Sea

$$A = \left(\frac{2}{\lambda}\right)^{D-1} = 2^{D-1}(4d^4\delta^2)^{\frac{D(D-1)^2}{2p}}.$$

Puesto que $A > 1$, resulta una elección válida de parámetro para aplicar el Lema 2.3, con lo que la estimación (3.15) permanece válida. Concluímos:

$$G(N) \leq HN^\alpha \sum_{i=0}^{n-1} \left(\frac{1}{2}\right)^i + K^nG(\lambda^nN) \leq 2HN^\alpha + K^n\lambda^{\alpha n}\lambda^{-\alpha n}G(\lambda^nN) \leq 2HN^\alpha + 2^{-n}\lambda^{-\alpha n}G(\lambda^nN) \leq 2HN^\alpha + 2^{-n}\lambda^{-\alpha}N^\alpha G(\lambda^nN). \tag{3.16}$$

En la última desigualdad de la expresión (3.16), usamos que $\lambda^{-n\alpha} \leq \lambda^{-\alpha}N^\alpha$, lo que se deduce de $\lambda^{n-1}N \geq 1$. Sea n tal que

$$\frac{\lambda}{N} \leq \lambda^n < \frac{1}{N}.$$

Entonces $\lambda^nN \geq 1$, con lo que la estimación (3.16) permanece válida. Además, $G(\lambda^nN) \leq 1$ por las mismas razones de antes. Concluímos:

$$G(N) \leq 2HN^\alpha + 2^{-n}\lambda^{-\alpha}N^\alpha \leq 2(H+K)N^\alpha.$$

Elijamos ahora $\mathcal{D}$, que necesitamos que verifique $\mathcal{D} \geq 2d$ para que todas las estimaciones que hicimos permanezcan válidas. Como $p \leq q$, tenemos

$$H + K \leq 5d^5\mathcal{D}^3(DA)^{\frac{2p}{D(D-1)}} \leq 5d^5\mathcal{D}^3(D2^{D-1})^{\frac{2p}{D(D-1)}}(4d^4\mathcal{D}^2)^{D-1} \leq 5d\mathcal{D}d^2\mathcal{D}^2(16d^4\mathcal{D}^2)^D \leq (d^4\mathcal{D}^5)^{d\mathcal{D}} \leq \frac{1}{2}\mathcal{D}^{9d\mathcal{D}}. \tag{3.17}$$

Dado que $N^\alpha \leq N^{\frac{1}{d}+\frac{4}{\delta}}$, de (3.17) deducimos:

$$G(N) \leq 2\frac{1}{2}\mathcal{D}^{9d\mathcal{D}}N^{\frac{1}{d}}N^{\frac{4}{\mathcal{D}}} = N^{\frac{1}{d}}\exp(\frac{4}{\mathcal{D}}\log(N) + 9d\mathcal{D}\log(\mathcal{D})).$$

Si

$$\mathcal{D} = \sqrt{\frac{4\log(N)}{d\log\log(N)}},$$

se tiene $\mathcal{D} \geq 2d$, siempre que $N \geq \exp(d^6)$. Podemos concluir de (3.17) la cota superior:

$$G(N) \leq N^{\frac{1}{d}}\exp(11\sqrt{d\log(N)\log\log(N)})$$

que es exactamente lo que queríamos probar.

Recordamos ahora que al principio supusimos $|g(x)| \leq N$. En caso de que esto no se verifique, para cada uno de los intervalos I_ν realizamos una traslación de g por un entero, es decir, consideramos $h(x) = g|_{I_\nu}(x) + r_\nu$ con $r_\nu \in \mathbb{Z}$ y usando el hecho que $|g'(x)| \leq 1$, podemos elegir los r_ν para que $|h(x)| \leq N$. Esta h va a tener un gráfico con la misma cantidad de puntos de coordenadas enteras que g, es igual de suave, con lo que procedemos como antes pero ahora con $\mathcal{M}_{G_\nu}(\delta)$ con G_ν es la traslación correspondiente a G en I_ν. Luego podemos repetir el razonamiento que hicimos para g con h. □

Corolario 2.1. *Sea $f \in \mathcal{C}^\infty(I)$ con I un intervalo cerrado contenido en $[0, N]$, y supongamos que $F(x, f(x)) = 0$, donde $F(x, y) \in \mathbb{R}[x, y]$ es un polinomio geométricamente irreducible de grado $d \geq 2$. Supongamos que $|f'(x)| \leq 1$. Si $N \geq \exp(d^6)$ entonces el número de puntos de coordenadas enteras en el gráfico de f es a lo sumo*

$$N^{\frac{1}{d}}\exp(11\sqrt{d\log(N)\log\log(N)}).$$

Demostración. Sea Γ el gráfico de f. Por el Teorema 2.5, se tiene que

$$\Gamma(\mathbb{Z}) \leq G(N) \leq N^{\frac{1}{d}}\exp(11\sqrt{d\log(N)\log\log(N)}).$$

□

4. Una cota uniforme

Nuestra meta en esta sección es utilizar el Corolario 2.1 en el caso de una curva algebraica. Esto lo vamos a conseguir particionando a la curva en tramos muy regulares adecuados. Expliquemos esto en detalle.

Supongamos que $F(x, y) = 0$ define una curva algebraica geométricamente irreducible C, con $F \in \mathbb{R}[x, y]$. Sea $N \geq 1$. Si consideramos $(x, y) \in C \cap [0, N]^2$, puede ocurrir que la curva tenga o no puntos singulares. Si (x, y) no es un punto singular de C, por el teorema de la función implícita tenemos que C es localmente el gráfico de una función (respecto a alguno de los ejes), muy suave. Si C no posee puntos singulares en $[0, N]^2$, debido a la compacidad de $C \cap [0, N]^2$, podemos deducir que C es unión finita de gráficos de funciones muy suaves. Sin embargo, notemos que podemos decir más. Si $\frac{\partial F}{\partial x}$ no se anula en todo $C \cap [0, N]^2$, el teorema de la función implícita y la compacidad del $[0, N]$ nos permiten deducir que $C \cap [0, N]^2$ es el gráfico de una función muy suave.

Si ahora $(x, y) \in C \cap [0, N]^2$ es un punto singular, consideremos $[x, N] \times [y, N]$ y asumamos que en este rectángulo sólo (x, y) es punto singular de C. Perturbando $\varepsilon > 0$ muy pequeño, $C \cap [x + \varepsilon, N] \times [y + \varepsilon, N]$ resulta el gráfico de una función muy suave por lo que ya dijimos. Supongamos entonces que queremos estudiar los puntos enteros de C en $[0, N]^2$. Si (x, y), el punto singular, no es un punto entero, considerando las cuatro porciones de rectángulo $[0, x - \varepsilon] \times [y + \varepsilon, N], [0, x - \varepsilon] \times [0, y - \varepsilon], [x + \varepsilon, y + \varepsilon], [x + \varepsilon, y - \varepsilon]$, en cada una C es el gráfico de una función suave muy regular y a efecto práctico de estudiar los puntos enteros de C en $[0, N]^2$, podemos suponer que C no tiene un punto singular (tomando ε muy pequeño). En caso que (x, y) sea un punto entero, el análisis anterior nos hace perder un punto entero, con lo que debemos agregarlo de alguna manera (por ejemplo, luego de estudiar fuera de los puntos singulares, los agregamos). Esto no es demasiado problemático, pues la estimación que tenemos para los puntos singulares depende del grado de la curva C.

PROPOSICIÓN 2.2. *Sea $f \in \mathbb{R}[x,y]$ irreducible. Si $\frac{\partial f}{\partial x}$ no es el polinomio nulo, entonces el sistema de ecuaciones*

$$\begin{cases} f(x,y) = 0 \\ \dfrac{\partial f}{\partial x} = 0 \end{cases},$$

tiene a lo sumo $\deg(f)[\deg(f)-1]$ soluciones. En particular, concluímos que un polinomio $f \in k[x,y]$ irreducible tiene a lo sumo $\deg(f)[\deg(f)-1]$ puntos singulares.

DEMOSTRACIÓN. Las curvas C, C', definidas por $f(x,y) = 0$ y $\frac{\partial f}{\partial x} = 0$ respectivamente, no tienen componentes en común. Por el teorema de Bézout, $C \cap C'$ consiste de a lo sumo $\deg(f)[\deg(f)-1]$ puntos.

Dado que un punto singular P de f verifica $\frac{\partial f}{\partial x}(P) = \frac{\partial f}{\partial y}(P) = 0$, concluímos que si $f \in \mathbb{R}[x,y]$ es irreducible, tiene a lo sumo $\deg(f)[\deg(f)-1]$ puntos singulares. □

OBSERVACIÓN 2.6. *La Proposición 2.2 sigue siendo válida para $f \in k[x,y]$ irreducible con k un cuerpo de característica positiva, o más en general, con k un cuerpo perfecto.*

TEOREMA 2.6. *Sea C una curva algebraica geométricamente irreducible de grado $d \geq 2$ y sea $N \geq \exp(d^6)$. Entonces*

$$\#(\Gamma(\mathbb{Z}) \cap [0,N]^2) \leq N^{\frac{1}{d}} \exp(12\sqrt{d\log(N)\log\log(N)}).$$

DEMOSTRACIÓN. Sea $S = [0,N] \times [0,N]$. En vista del análisis que hicimos al comienzo de la sección, vamos a suponer que no hay puntos singulares en S y luego los agregamos al final. Por la Proposición 2.2, ya sabemos que una curva algebraica C de grado $d \geq 2$ definida por $F(x,y) = 0$ tiene $d(d-1)$ puntos donde $\frac{\partial F}{\partial x} = 0$ y $d(d-1)$ puntos donde $\frac{\partial F}{\partial y} = 0$. Luego, hay a lo sumo $2d(d-1)$ puntos donde el vector tangente es horizontal o vertical. Esto implica, debido a que F es una función suave, que la pendiente del vector tangente a C es ± 1 en a lo sumo $2d(d-1)+1$ puntos. Separamos ahora la curva C en tramos suaves que son gráficos de funciones f con $|f'(x)| \leq 1$ (o $|f'(y)| \leq 1$). No es difícil de comprobar que $C \cap S$ es la unión de a lo sumo $4d^2$ tramos Γ_j, que son gráficos de funciones $f_j \in \mathcal{C}^\infty$ con pendiente acotada por 1 con respecto a uno de los ejes. Sea $G(N)$ el máximo número de puntos enteros que puede contener el gráfico de una función $g \in \mathcal{C}^\infty(I)$ con I un intervalo cerrado y acotado de longitud a lo sumo N, con $|g'(x)| \leq 1$, que verifica que existe un polinomio $P(x,y) \in \mathbb{R}[x,y]$ no nulo, geométricamente irreducible, tal que $P(x,g(x)) = 0$ para todo $x \in I$ (ver Teorema 2.5). Entonces $\#\Gamma_j(\mathbb{Z}) \leq G(N)$, por lo tanto, por el Teorema 2.5, para $N \geq \exp(d^6)$:

$$4d^2 G(N) \leq 4d^2 N^{\frac{1}{d}} \exp(11\sqrt{d\log(N)\log\log(N)}). \tag{4.1}$$

Agregando los $d(d-1) \leq d^2$ puntos singulares a la desigualdad (4.1), obtenemos

$$d^2(4N^{\frac{1}{d}} \exp(11\sqrt{d\log(N)\log\log(N)}) + 1) \leq 8d^2 N^{\frac{1}{d}} \exp(11\sqrt{d\log(N)\log\log(N)}).$$

Como $N \geq \exp(d^6)$, resulta:

$$\exp(\sqrt{d\log(N)\log\log(N)}) \geq \exp(\sqrt{dd^6 6\log(d)}) = \exp(\sqrt{6d^7\log(d)}) \geq \sqrt{6d^7\log(d)} \geq 8d^2,$$

lo que termina de probar el teorema. □

OBSERVACIÓN 2.7. *En la demostración del Teorema 2.6, sólo usamos que $S = [0,N]^2$ es un cuadrado de lado a lo sumo N, con lo que la estimación del Teorema 2.6 permanece válida para todo cuadrado S de longitud N. Por otro lado, al usar el Teorema 2.5, estamos usando que $[0,N]^2$ es un cuadrado. Los métodos que desarrollamos en el primer y segundo capítulo no alcanzan para obtener una estimación en una región como $[0,N] \times [0,M]$ con $N \neq M$.*

Notemos que

$$\exp(12\sqrt{d\log(N)\log\log(N)}) = \exp(\log(N))^{12\sqrt{d}\frac{\log\log(N)}{\sqrt{\log(N)}}} = N^{12\sqrt{d}\frac{\log\log(N)}{\sqrt{\log(N)}}}.$$

Como $\frac{\log\log(N)}{\sqrt{\log(N)}} = O_\varepsilon(\varepsilon)$, concluímos que

$$\exp(12\sqrt{d\log(N)\log\log(N)}) = O_{d,\varepsilon}(N^{\frac{1}{\varepsilon}}).$$

Deducimos:

COROLARIO 2.2. *Sea C una curva algebraica plan geométricamente irreducible de grado $d \geq 2$. Para todo $\varepsilon > 0$, se verifica*

$$\#(\Gamma(\mathbb{Z}) \cap [0,N]^2) = O_{d,\varepsilon}(N^{\frac{1}{d}+\varepsilon}).$$

Observación 2.8. *Sea C la curva algebraica plana definida por la ecuación $f(x,y) = x^d - y = 0$. El polinomio $f(x,y) = x^d - y$ es irreducible sobre $\mathbb{C}[x,y]$, luego C es geométricamente irreducible. No es difícil de ver que*

$$\#(C(\mathbb{Z}) \cap [0,N]^2) = \left\lfloor N^{\frac{1}{d}} \right\rfloor \asymp N^{\frac{1}{d}}.$$

Esto muestra que el exponente en el Corolario 2.2 es óptimo.

Observación 2.9. *Una curva algebraica C definida por $G(x,y) = ax + by + c = 0$ verifica la condición $y = -\frac{1}{b}(ax+c)$ en caso de tener grado positivo en y, con lo que el valor de x determina unívocamente el valor de y para que esté en γ; se tiene la situación análoga si $a \neq 0$ con x. Entonces $\#(\Gamma(\mathbb{Z}) \cap [0,N]^2) \leq N$ y esta estimación se alcanza con el caso sencillo $G(x,y) = x - y$.*

Observación 2.10. *Una posible forma de interpretar el Teorema 2.6 es la siguiente. Dada una curva algebraica C de grado d, contamos los puntos enteros que posee en una región $S = [0,N]^2$. Si estos puntos son muchos, digamos al menos $N^{\frac{1}{d}+\varepsilon}$, por lo anterior debe tenerse que C no es geométricamente irreducible. Luego el Teorema 2.6 nos permite obtener información geométrica (ser irreducible) a partir de información aritmética (la densidad de puntos enteros en regiones rectangulares).*

Observación 2.11. *La cota del Corolario 2.2, por lo que vimos, tiene exponente óptimo. Sin embargo, no se sabe si la cota del corolario es óptima, es decir, para toda curva algebraica C, geométricamente irreducible, de grado d, no se sabe si se tiene la estimación*

$$\#(C(\mathbb{Z}) \cap [0,N]^2) = O_d(N^{\frac{1}{d}}).$$

5. Puntos enteros en superficies

Veamos qué tipo de estimación podemos conseguir con el Teorema 2.6 en un caso concreto. Consideremos la hipersuperficie algebraica de grado 3 dada por la ecuación

$$x^3 + y^3 = z^3. \tag{5.1}$$

Ya sabemos que no tiene soluciones enteras no triviales (es un caso del último teorema de Fermat), esto es, en $[0,N]^3$ seguro son solución $(x,0,x), (0,x,x)$ pero no hay más. Fijemos entonces $z \in [0,N]$; tenemos ahora una curva algebraica plana C_z de grado 3, a saber, el conjunto de (x,y) con $x^3 + y^3 = z^3$. Si $z = 0$, el $(0,0,0)$ es la única solución en $[0,N]^3$ con lo que podemos suponer $z \in (0,N]$. Puesto que $(y-z) \mid y^3 - z^3$ pero $(y-z) \nmid x^3, (y-z)^2 \nmid y^3 - z^3$, por el criterio de Eisenstein el polinomio $x^3 + y^3 = z^3$ es irreducible para $z \in (0,N]$.

Particionamos ahora la curva C_z en tramos suaves; no es difícil de ver que en este caso, es posible hacerlo en 4 tramos. Por el Teorema 2.6, tenemos que C_z tiene a lo sumo

$$N^{\frac{1}{3}} \exp(12\sqrt{3\log(N)\log\log(N)}),$$

siempre que $N \geq \exp(3^6)$. Dado que hay a lo sumo N posibilidades para z, la cota anterior dice que hay a lo sumo

$$N^{1+\frac{1}{3}} \exp(12\sqrt{3\log(N)\log\log(N)}) = O(N^{1+\frac{1}{3}}\log(N)).$$

Observemos que esta cota no usa nada especial de la ecuación (5.1), salvo que es de grado 3 y que, fijado z, el polinomio obtenido es geométricamente irreducible (salvo para $z = 0$). Dicho de otra forma, podemos aplicar esta idea para una hipersuperficie algebraica geométricamente irreducible de grado $d \geq 2$ tal que si está definida por la ecuación $F(x_1, \cdots, x_n) = 0$ con $F \in \mathbb{R}[x_1, \cdots, x_n]$ irreducible sobre $\mathbb{C}[x_1, \cdots, x_n]$, y este polinomio permanece irreducible fijada, digamos, la variable x_n, salvo para finitos x_n, entonces tomando N suficientemente grande, la cantidad de puntos enteros $\#S(\mathbb{Z})$ en el n-cubo $[0,N]^n$ es del orden $O(N^{n-2+\frac{1}{d}+\varepsilon})$ para todo $\varepsilon > 0$. Adicionalmente, la cota es óptima pues la hipersuperficie algebraica definida por $F(x_1, \cdots, x_n) = x_1 - x_n^d = 0$ tiene exactamente

$$N^{n-2}\left\lfloor N^{\frac{1}{d}} \right\rfloor$$

puntos enteros en el n-cubo $[0,N]^n$.

Esto muestra por un lado que las técnicas que introdujimos dan estimaciones de un orden que no se puede mejorar sin hipótesis adicionales a la irreducibilidad. Por otro lado, tenemos las limitaciones de los resultados que podemos aspirar obtener: si deseamos estimar la cantidad de soluciones de una ecuación como (5.1) para ver que "no hay muchas soluciones",

el Teorema 2.6 da una cota "extremal" (considera el caso con más puntos enteros posibles) con lo que no nos permite obtener una cota del orden correcto para comparar, salvo en dimensiones bajas (en el caso del plano).

Muchos problemas diofánticos vienen descritos por hipersuperficies algebraicas. Esto hace que las limitaciones de las técnicas traten de ser mejoradas y extendidas. La ecuación (5.1) tiene una propiedad más a destacar, que es que viene dada por uns forma homogénea, luego define una hipersuperficie algebraica en el plano proyectivo. El problema de estudiar las ecuaciones diofánticas dadas por formas homogéneas es importante; escapa el contenido de la tesis, pero explicamos brevemente resultados en esa dirección, siguiendo el trabajo [**17**]. Sea $F \in \mathbb{Z}[x_1, \cdots, x_s]$ una forma homogénea de grado $k \geq 2$. Sea

$$F(x_1, \cdots, x_s) = 0. \tag{5.2}$$

Nos preguntamos ahora por cotas para la cantidad de soluciones de la ecuación (5.2), de ser posible cotas uniformes en el grado de F. En esta dirección, Heath-Brown en [**17**] estudia el problema de estimar la cantidad de puntos racionales de cierta altura B, de la hipersuperficie de $\mathbb{P}^{n-1}(\mathbb{R})$ definida por una forma $F \in \mathbb{Z}[x_1, \cdots, x_n], n \geq 3$, geométricamente irreducible, de grado d. Obtiene resultados en dimensiones superiores a [**6**], pero siguiendo pasos similares. Para ver el tipo de resultados que prueba, introducimos un poco de notación.

Sea F como antes. Consideremos los puntos de altura a lo sumo B, en la hipersuperficie algebraica proyectiva que define F. Tales puntos admiten una representación particular. En efecto, podemos elegir los representantes de la forma $(x_1, \cdots, x_n) \in \mathbb{Z}^n$ tales que

- los x_i no todos nulos.
- $\gcd(x_1, \cdots, x_n) = 1$.
- si i es el menor entero positivo tal que $x_i \neq 0$, entonces $x_i > 0$.

Vamos a definir Z_n como el conjunto de los representantes recién indicados. Definimos, para $B > 0$, la cantidad

$$N(F, B) := \#\{\mathbf{x} \in Z_n : F(\mathbf{x}) = 0\} \cap [-B, B]^n.$$

O sea, $N(F, B)$ es la cantidad de representantes contenidos en la hipersuperficie definida por los ceros de F que están en el cubo $[-B, B]^n$. Si

$$F(x_1, \cdots, x_n) = \sum_{r_1, \cdots, r_n} a_{r_1 \cdots r_n} x_{i_1}^{r_1} \cdots x_{i_n}^{r_n},$$

definimos

$$||F|| := \max_{r_1, \cdots, r_n} |a_{r_1 \cdots r_n}|.$$

Podemos entonces enunciar los siguientes resultados de[**17**] [teorema 3, teorema 5], que corresponden al estudio de las curvas algebraicas del plano proyectivo $\mathbb{P}^3(\mathbb{R})$:

Teorema 2.7. *[Heath-Brown] Sea $\varepsilon > 0$. Si $F(x_1, x_2, x_3) \in \mathbb{Q}[x_1, x_2, x_3]$ es una forma homogénea de grado d, irreducible en $\mathbb{Q}[x_1, x_2, x_3]$, entonces para todo $B \geq 1$ se tiene*

$$N(F, B) = O_\varepsilon(B^{\frac{2}{d}} + \varepsilon).$$

Teorema 2.8. *[Heath-Brown] Sea C una curva irreducible en $\mathbb{P}^3(\mathbb{R})$, de grado d, no necesariamente definida sobre $\mathbb{Q}$. Entonces dado $\varepsilon > 0$, se tiene para C la estimación*

$$Z_4 \cap [-B, B]^n = O_\varepsilon(B^{\frac{2}{d} + \varepsilon}).$$

En [**17**] también se estudian las hipersuperficies de $\mathbb{P}^3(\mathbb{R})$. En [**17**] [teorema 9], se prueba:

Teorema 2.9. *[Heath-Brown] Sea $F \in \overline{\mathbb{Q}}[x_1, x_2, x_3, x_4]$ una forma geométricamente irreducible de grado $d \geq 2$. Entonces se tiene la estimación*

$$N(B, F) = O_\varepsilon(B^{2+\varepsilon}).$$

Consideremos ahora $F(x_1, x_2, x_3, x_4) = x_1^d + x_2^d - x_3^d - x_4^d$. Tenemos "muchas soluciones triviales", como las de la forma (a, b, a, b). Esto hace que la cantidad $N(B, F)$ medida sin cuidado, pueda ser grande. Resulta entonces natural eliminar la cantidad $N(B, F)$ sin estas soluciones triviales, considerando $N_1(B, F)$ a la cantidad $N(B, F)$ a la que se le quitaron los puntos racionales contenidos en rectas de la hipersuperficie $F(x_1, x_2, x_3, x_4) = 0$. Se tiene la estimación de [**17**] [teorema 7]:

TEOREMA 2.10 (Heath-Brown). *Sea* $F \in \overline{\mathbb{Q}}[x_1, x_2, x_3, x_4]$ *una forma geométricamente irreducible de grado* $d \geq 2$. *Entonces se tiene la estimación*

$$N_1(B, F) = O_\epsilon(B^{1+\frac{3}{\sqrt{d}}+\epsilon}).$$

Para dimensiones superiores, un análogo al Teorema 2.9 no se conoce. Se conjetura en [**17**][conjetura 2]:

CONJETURA 2.1. *Sea* $F \in \overline{\mathbb{Q}}[x_1, \cdots, x_n]$ *una forma geométricamente irreducible de grado* $d \geq 3$ *con* $n \geq 5$. *Entonces se tiene la estimación*

$$N(B, F) = O_\varepsilon(B^{n-2+\varepsilon}).$$

Las técnicas empleadas para probar el Teorema 2.8 y el Teorema 2.9 son de naturaleza geométrica, aunque requieren más elementos de geometría algebraica que los empleados en esta tesis, además de técnicas del análisis p-ádico.

CAPÍTULO 3

Los conjuntos algebraicos como conjuntos mal distribuidos

El problema de resolver una ecuación diofántica tiene sentido una vez que se sabe que hay soluciones. Para dar un ejemplo bien clásico, relacionado con el problema de Waring, consideramos para cada N la ecuación diofántica

$$x_1^3 + x_2^3 + x_3^3 = N.$$

Antes de intentar acotar las soluciones o hacer un estudio, podemos observar que para muchos casos de N, no hay soluciones. Para comprobarlo, miramos módulo 9 la ecuación anterior. Dado que $x^3 \equiv \pm 1 \pmod 9$, deducimos que si tomamos $N \equiv 4 \pmod 9$, la ecuación diofántica no tiene solución. Concluímos que para algunos N, no hay soluciones.

Restrigiéndonos a m primo, podemos pensar que para detectar soluciones de una ecuación

$$P(x_1, \cdots, x_n) = 0, \tag{0.3}$$

con $P \in \mathbb{Z}[x_1, \cdots, x_n]$, un primer intento es estudiar la ecuación anterior en $\mathbb{Z}/p\mathbb{Z}$, que es un cuerpo finito de p elementos, para todo primo p. Por lo que dijimos si la ecuación tiene solución, tiene que tener solución sobre $\mathbb{Z}/p\mathbb{Z}$ para todo primo p. Surgen entonces (al menos) dos interrogantes

PREGUNTA. *Si la ecuación* (0.3) *tiene solución sobre todo* $\mathbb{Z}/p\mathbb{Z}$*, ¿hay alguna solución entera?*

PREGUNTA. *De tener una estimación para la cantidad de soluciones sobre* $\mathbb{Z}/p\mathbb{Z}$ *de la ecuación* (0.3)*, ¿podemos decir algo de las soluciones enteras de la ecuación?*

La primera pregunta tiene una respuesta negativa. Más aún, Selmer en [**49**] mostró que la ecuación

$$3x^3 + 4y^3 + 5z^3 = 0,$$

tiene soluciones reales y en $\frac{\mathbb{Z}}{p\mathbb{Z}}$[1], pero no tiene soluciones enteras. Para nuestro caso, bastará un ejemplo bastante más sencillo. Consideremos la ecuación

$$x^2 + y^2 + z^2 + w^2 = -1.$$

Tal ecuación no puede tener soluciones enteras, pues la suma de cuadrados es no negativa. Sin embargo, al mirar módulo p, resulta

$$x^2 + y^2 + z^2 + w^2 \equiv p - 1 \pmod p.$$

Como $p - 1$ es positivo, acudimos al siguiente resultado clásico de Lagrange (ver [**37**][capítulo 7, pág. 281]):

TEOREMA 3.1 (Lagrange). *Sea N un entero no negativo. Entonces existen $x, y, z, w \in \mathbb{Z}$ tales que*

$$x^2 + y^2 + z^2 + w^2 = N.$$

Como $p - 1$ es un entero no negativo, por el Teorema 3.1, existen $x, y, z, w \in \mathbb{Z}$ tales que la suma de sus cuadrados es $p - 1$. Reduciendo módulo p, tenemos que la ecuación tiene soluciones módulo p.

En cuanto a la segunda pregunta, Lang y Weil [**25**] dan una estimación para la cantidad de soluciones módulo p de una ecuación diofántica:

TEOREMA (Cota de Lang-Weil). *Sea $S \subseteq \overline{\mathbb{Q}}^n$ una hipersuperficie algebraica, geométricamente irreducible, definida por los ceros de un polinomio $P(x_1, \cdots, x_n) \in \mathbb{Z}[x_1, \cdots, x_n]$ de grado d. Se tiene la estimación*

$$\#V(\mathbb{Z}/p\mathbb{Z}) = \left(1 + O_d(p^{-\frac{1}{2}})\right) p^{n-1}.$$

Sea $Q(x, y) \in \mathbb{Z}[x, y]$ geométricamente irreducible. Entonces, tenemos una estimación para la cantidad de soluciones módulo p de la ecuación $Q(x, y) = 0$. Helfgott y Venkatesh [**18**] obtienen información acerca de las soluciones enteras de $Q(x, y) = 0$ en un cuadrado $[0, N]^2$, con N un entero positivo, por medio de la siguiente estrategia:

[1] En verdad, lo que Selmer probó es que hay soluciones en los racionales p-ádicos $\mathbb{Q}_p$, para cada primo p.

- Para cada primo p, reducimos módulo p al polinomio Q. El polinomio reducido $\overline{Q}(x,y) \in \mathbb{Z}/p\mathbb{Z}[x,y]$ puede no ser irreducible; factorizamos en irreducibles $\overline{Q} = \overline{Q}_1 \cdots \overline{Q_{e_p}}$. Se verifica que los primos p para los que Q es reducible, que llamaremos $\mathcal{P}$, son finitos.
- Usamos la estimación de Lang-Weil para cada polinomio irreducible que apareció en el paso anterior.
- Sea C la curva algebraica geométricamente irreducible definida por la ecuación $Q(x,y) = 0$. Si $R = [0,N]^2$, consideramos $C(\mathbb{Z}) \cap R \subseteq [N]^2$. Sea $p \in \mathcal{P}$, luego $\overline{Q} = \overline{Q}_1 \cdots \overline{Q_{e_p}}$. Notemos que siempre $e_p \leq d$. Cada factor Q_i define una curva algebraica geométricamente irreducible $C_{i,p}$. Dado $x \in C(\mathbb{Z}) \cap R$, debe tenerse que existe un primo $p \in \mathcal{P}$ tal que $x \pmod p$ pertenece a sólo una curva $C_{i,p}$. De esta forma, podemos particionar $C(\mathbb{Z}) \cap R$ en conjuntos $S_1, \cdots, S_k$.
- Por la construcción de los S_k, estos conjuntos están "mal distribuidos", es decir, ocupan pocas clases residuales módulo p para todo primo p suficientemente grande (esencialmente por la estimación de Lang-Weil). Helfgott y Venkatesh estudian los conjuntos mal distribuidos por medio de una adaptación del método del determinante de Bombieri-Pila indexMétodo del determinante. Por medio de este estudio, consiguen obtener una cota para cada $\#S_i$ y, por lo tanto, una cota para $\#(C(\mathbb{Z}) \cap R)$. Más aún, la estimación que obtienen es la misma de Bombieri-Pila (ver Teorema 2.2).

Podemos concluir entonces que es posible extraer una estimación de la cantidad de soluciones enteras de $Q(x,y) = 0$ a partir de tener una estimación de las soluciones enteras de $Q(x,y) \equiv 0 \pmod p$ para todo primo p suficientemente grande.

El trabajo [**18**] es importante porque, entre otras cosas, da una interpretación "local" del método del determinante de Bombieri-Pila, desarrollado en los primeros dos capítulos. Esta "interpretación local" se basa en el estudio de los conjuntos "mal distribuidos". Un conjunto aleatorio $S \subseteq \mathbb{Z}$ puede pensarse como un conjunto que, reducido módulo p, con p primo, cubre muchas clases residuales módulo p, para todos los primos. Entonces resulta razonable pensar que el conjunto S está mal distribuido si ocupa pocas clases residuales módulo p para muchos primos p.

Para los conjuntos $S \subseteq \mathbb{Z}^2$, Helfgott y Venkatesh adaptan la definición de "mal distribuido" de los conjuntos de $\mathbb{Z}$, del párrafo anterior. Obtienen la siguiente caracterización, que probaremos en la sección 4 de este capítulo:

Teorema 1 (Helfgott-Venkatesh). *Sea $S \subseteq [N]^2 \subseteq \mathbb{Z}^2$ con $N \geq 1$. Supongamos que existe $\alpha > 0$ tal que $\#S_p \leq \alpha p$ para todo primo p. Entonces, para todo $\varepsilon > 0$ se tiene que ocurre alguna de las proposiciones siguientes:*

- *$\#S = O_{\alpha,\varepsilon}(N^\varepsilon)$,*
- *existe una curva algebraica plana C de grado $O_{\alpha,\varepsilon}(1)$ tal que al menos $(1-\varepsilon)\#S$ puntos de S se encuentran en C.*

En la estrategia de la demostración del Teorema 2.2 de Helfgott y Venkatesh se utiliza una adaptación del método del determinante de Bombieri-Pila, basada en una técnica analítica conocida como criba más grande de Gallagher. Mencionamos a continuación esta adaptación

Teorema 2. *Sea $S \subseteq [N]^2, N \geq 1$. Supongamos que existe $\alpha > 0$ tal que $\#S_p \leq \alpha p$ para todo primo $p > c$.*
Sea $\mathcal{M} \subseteq \mathbb{Z}[x,y]$ un conjunto finito de polinomios que verifica:

- *ser linealmente independiente,*
- *separar puntos,*
- *incluye al polinomio $P(x,y) = 1$.*

Sea $D := \#\mathcal{M}$ y $d_{\mathcal{M}} := \sum_{P \in \mathcal{M}} \deg(P)$. Decimos que C es una $\mathcal{M}$-curva si es una curva algebraica plana definida por la ecuación $g(x,y) = 0$ con g un polinomio que es combinación $\mathbb{Q}$-lineal de los polinomios de $\mathcal{M}$. Entonces, para todo $\delta \in (0,1)$, se cumple algunas de las siguientes condiciones:

(1) *existe una $\mathcal{M}$-curva que contiene al menos $\delta\#S$ puntos de S,*
(2) *se tiene $\#S \ll_{\alpha,\delta,\mathcal{M}} (N^{\frac{2\alpha d_{\mathcal{M}}}{D(D+1)} + O_{\alpha,\mathcal{M}}(\delta)})$.*

El teorema anterior es el resultado central de [**18**]. De hecho, la caracterización del Teorema 1 es una aplicación de este teorema.

En este capítulo vamos a probar el Teorema 1 y el Teorema 2. Vamos a dedicar las primeras dos secciones a motivar la caracterización de los conjuntos "mal distribuidos" del Teorema 1. En la sección 3, explicamos cómo se adapta el método del determinante en el contexto de la criba más grande de Gallagher, para luego probar el Teorema 2 y damos como aplicación el Teorema 1. Finalizamos el capítulo en la sección 4, explicando qué se puede decir del estudio de conjuntos mal distribuidos en dimensiones superiores (conjuntos contenidos en $\mathbb{Z}^n$ con $n > 2$) y en la recta (conjuntos contenidos en $\mathbb{Z}$).

1. Conjuntos mal distribuidos

Consideremos un conjunto $S \subseteq \mathbb{R}^n$. A veces no tenemos suficiente control del conjunto, sin embargo tenemos una idea de qué tan bien distribuido se encuentra. Por ejemplo, si consideramos $S = \{x_n := \{nx\} : n \in \mathbb{Z}\}$, donde $\{a\}$ es la mantisa de a, es decir, $\{a\} = a - \lfloor a \rfloor$, es sabido que si $x \notin \mathbb{Q}$, S es un conjunto denso del $[0,1]$. Sin embargo, es más que denso, está uniformemente distribuida[2] en el $[0,1]$, lo que significa que tiene densidad total en cada subintervalo del $[0,1]$. Concretamente, si $I = (a,b) \subseteq [0,1]$ lo que se tiene es

$$\lim_{N\to+\infty} \frac{\#\{x_n \in S \cap I : 1 \leq n \leq N\}}{N} = |I| = (b-a).$$

En general, tenemos la siguiente definición.

Definición 3.1. *Sea $x_n)_n$ una sucesión de números reales. Decimos que $(x_n)_n$ está uniformemente distribuida módulo 1, si para todo intervalo $I = (a,b) \subseteq [0,1]$ se verifica*

$$\lim_{N\to+\infty} \frac{\#\{x_n : \{x_n\} \in I, 1 \leq n \leq N\}}{N} = |I| = (b-a).$$

No es difícil de convencerse que una sucesión $(x_n)_n$ uniformemente distribuida módulo 1 tiene su sucesión asociada $(\{x_n\})_n$ densa en el $[0,1]$. Por otro lado, puede corroborarse que la sucesión $\{\sin(n)\}_{n\in\mathbb{N}}$ es densa en el $[-1,1]$ pero no está uniformemente distribuida módulo 1. Esto quiere decir que "estar uniformemente distribuido" es algo más fuerte que "ser denso".

Las sucesiones uniformemente distribuidas módulo 1 verifican muy buenas propiedades. Por ejemplo, si $f : [0,1] \to \mathbb{R}$ es una función continua, si se consideran los promedios $A_n := \frac{1}{n}\sum_{i=1}^{n} f(\{x_i\})$, con $(x_i)_i$ una sucesión uniformemente distribuida módulo 1, no es difícil de ver que A_n converge a la integral $\int_0^1 f(t)dt$. Además, las sucesiones uniformemente distribuidas dan una buena noción de conjuntos aleatorios: dada $(x_n)_n$ sucesión uniformemente distribuida módulo 1, si elegimos un intervalo del $[0,1]$, la distribución de la sucesión $(\{x_n\})_n$ es uniforme en el intervalo.

Resulta de interés extender la noción de uniforme distribución módulo 1, por ejemplo, para sucesiones de elementos de un grupo localmente compacto y Hausdorff. En nuestro caso veamos cómo podemos definir la noción de uniforme distribución de un conjunto de enteros. Sea $(x_n)_n$ una sucesión de enteros. De manera análoga a una sucesión uniformemente distribuida módulo 1, podríamos decir que la sucesión $(x_n)_n$ está uniformemente distribuida si para cada entero positivo m, la distribución de $(x_n \pmod m)_n$ es uniforme en $\mathbb{Z}/m\mathbb{Z}$. Esta es la definición clásica de una sucesión de enteros uniformemente distribuida módulo $\mathbb{Z}$.

Definición 3.2. *Sea $(x_n)_n$ una sucesión de enteros. Decimos que una sucesión $(x_n)_x$ está uniformemente distribuida módulo m si, para todo $0 \leq j \leq m-1$, se verifica la condición*

$$\lim_{N\to+\infty} \frac{\#\{x_n \in S : x_n \equiv j(m), 1 \leq n \leq N\}}{N} = \frac{1}{m}.$$

Decimos que el conjunto $S \in \mathbb{Z}$ está uniformemente distribuido módulo $\mathbb{Z}$ si está uniformemente distribuido módulo m para todo $m \geq 2$.

En particular, si una sucesión está uniformemente distribuida módulo $\mathbb{Z}$, está uniformemente distribuida módulo p^n para todo primo p y todo entero positivo n.

Para muchas aplicaciones, concretamente, cuando se estudia la distribución de un conjunto $\Omega \subseteq \mathbb{Z}$, resulta natural estudiar Ω_p, el conjunto de restos módulo p de Ω, o Ω_{p^n}, el conjunto de restos módulo p^n , con p un primo y n un entero positivo. Posteriormente veremos que este es el caso de la criba grande y la criba más grande de Gallagher. Resulta entonces natural preguntarse si una sucesión $(x_n)_n$ uniformemente distribuida módulo p^n, para todo p primo y n entero positivo, resulta una sucesión uniformemente distribuida módulo $\mathbb{Z}$. Esta pregunta tiene respuesta negativa.

Ejemplo 3.1 (Niederreiter-Kuipers). *Sea*

$$a_n := \begin{cases} n \text{ si } n \equiv 0,1,2,5 \pmod 6 \\ n-2 \text{ si } n \equiv 3 \pmod 6 \\ n+2 \text{ si } n \equiv 4 \pmod 6 \end{cases}.$$

No es difícil de comprobar que $(a_n)_n$ está uniformemente distribuida módulo m, con $m \geq 2$ y $6 \nmid m$, luego está uniformemente distribuida módulo p^n con p primo y n entero positivo. Pero $(a_n)_n$ no está uniformemente distribuida módulo 6, luego no está uniformemente distribuida módulo $\mathbb{Z}$.

[2]Este resultado es conocido como el lema de Kronecker. Puede encontrarse una demostración en [**23**]

Una forma de estudiar las sucesiones uniformemente distribuidas módulo $\mathbb{Z}$ es estudiar qué sucesiones están "mal distribuidas".

PREGUNTA 3.1. *¿Cómo es una sucesión $(x_n)_n \subseteq \mathbb{Z}$ que no está uniformemente distribuido módulo $\mathbb{Z}$?*

La Pregunta "inversa" 3.1 sigue siendo bastante complicada, porque hay muchas razones por las cuales una sucesión puede estar mal distribuida. Sin embargo, para muchas aplicaciones en teoría de números, la "mala distribución" de una sucesión es de una naturaleza particular.

EJEMPLO 3.2. *Sea $S = (n^2)_n$ el conjunto de los cuadrados perfectos. Sea $p > 2$ un primo. Sea $S_p := \{n^2 \pmod p : n \in \mathbb{N}\}$, es decir, clases residuales módulo p de los elementos de S. Por la ley de reciprocidad cuadrática, $\#S_p = \frac{p+1}{2}$, es decir, S esencialmente ocupa la mitad de las clases residuales módulo p. Esto implica que la sucesión de los cuadrados perfectos no está uniformemente distribuida módulo p y por lo tanto, no está uniformemente distribuida módulo $\mathbb{Z}$.*

El Ejemplo 3.2 muestra una razón típica de mala distribución, debida a que una sucesión ocupa pocas clases residuales módulo p para muchos primos p. Esto motiva la siguiente definición.

DEFINICIÓN 3.3. *Sea $S \subseteq \mathbb{Z}$ no vacío. Decimos que S está mal distribuido si existen $\alpha \in (0,1)$ y $c > 0$ tales que, para todo $p > c$ primo, el conjunto $S_p := \{x \pmod p : x \in S\}$ tiene a lo sumo αp elementos.*

La Definición 3.3 asegura la no uniforme distribución módulo $\mathbb{Z}$ de una sucesión x_n de enteros. Además, la definición 3.3 resulta más conveniente que la Definición 3.2 en algunas aplicaciones, debido a que se refiere a la mala distribución de un conjunto y no a la mala distribución de una sucesión[3]. Remarquemos que el conjunto de los cuadrados perfectos resulta ser un conjunto mal distribuido con la Definición 3.3

Ahora podemos plantear otra pregunta, en la dirección de la Pregunta 3.1.

PREGUNTA 3.2. *¿Cómo es un subconjunto de enteros S mal distribuido?*

Un resultado relacionado a la pregunta anterior es el siguiente.

TEOREMA (Gallagher). *Sea $S \subseteq [N]$ un subconjunto tal que existen $\alpha \in (0,1)$ y $c > 0$ tal que S ocupa a lo sumo αp clases residuales módulo p para todo primo $p > c$. Entonces $\#S \ll_\alpha N^\alpha$.*

El teorema anterior dice que un conjunto finito de enteros positivos, mal distribuido, no puede ser muy grande. Más aún, dado que los conjuntos de enteros que se estudian de manera natural, son infinitos, el conjunto S que uno suele considerar se obtiene de truncar un conjunto infinito mal distribuido, es decir, si $S \subseteq \mathbb{Z}$ está mal distribuido, podemos considerar $S_N = S \cap [N]$, luego el teorema anterior nos dice que, para cada N, S_N no puede ser muy grande, es decir, el tamaño de S_N no crece muy rápido cuando $N \to +\infty$. Sin embargo, es esperable que la naturaleza de la mala distribución pueda ser debida a otras razones, como que el conjunto tenga alguna estructura algebraica, como en el caso de los cuadrados perfectos que el conjunto en cuestión es la imagen del polinomio $P(x) = x^2$.

La Pregunta 3.2 es bastante complicada (ver [**18**][sección 4.2]) y escapa los objetivos de la tesis. Nosotros vamos a estudiar la pregunta análoga a 3.2 para conjuntos mal distribuidos del plano. Para ello, adaptamos la Definición 3.3.

DEFINICIÓN 3.4. *Sea $S \subseteq \mathbb{Z}^2$ un conjunto no vacío. Decimos que S está mal distribuido si existen $\alpha > 0$ y $c > 0$ tales que para todo $p > c$ primo, se verifica que S ocupa a lo sumo αp clases residuales módulo p, es decir, si $S_p := \{(x \pmod p, y \pmod p) : (x,y) \in S\}$, se tiene $\#S_p \leq \alpha p$.*

Notemos que en este caso, ocupar pocas clases residuales se refiere a ocupar una cantidad de un orden menor que p^2 clases residuales. En la definición estamos pidiendo que la cantidad de clases residuales sea $O(p)$. Pedimos esta estimación, y no otras como $O(p\log(p))$, debido a que en la práctica, los conjuntos mal distribuidos del plano típicos verifican que la cantidad de clases residuales es $O(p)$. Esto puede verse en el siguiente ejemplo.

EJEMPLO 3.3. *Sea C una curva algebraica plana, definida por un polinomio $P(x,y) \in \mathbb{Z}[x,y]$, geométricamente irreducible, de grado d. Las estimaciones de Lang-Weil* [**25**] *permiten deducir que la cantidad de clases módulo p ocupadas por $\overline{P}(x,y) \in \mathbb{Z}/p\mathbb{Z}[x,y]$, la reducción módulo p del polinomio P, es del orden de $p + O_d(p^{\frac{1}{2}})$. Por lo tanto, $C(\mathbb{Z}) \subseteq \mathbb{Z}^2$, el conjunto de puntos enteros de C, es un conjunto mal distribuido del plano.*

Ahora consideramos la pregunta análoga a la Pregunta 3.2:

[3]Puede ocurrir que, dado un conjunto $S \subseteq \mathbb{Z}$, admita dos numeraciones $(x_n)_n, (y_n)_n$ tales que $(x_n)_n$ sea uniformemente distribuida módulo $\mathbb{Z}$ pero $(y_n)_n$ no.

PREGUNTA 3.3. *¿Cómo es un subconjunto S mal distribuido del plano?*

Como mencionamos en la introducción a este capítulo, Helfgott y Venkatesh obtienen una caracterización de los subconjuntos del plano mal distribuidos, que da una respuesta satisfactoria a la Pregunta 3.3.

TEOREMA (Helfgott-Venkatesh). *Sea $S \subseteq [N]^2 \subseteq \mathbb{Z}^2$ con $N \geq 1$. Supongamos que existe $\alpha > 0$ tal que $\#S_p \leq \alpha p$ para todo primo p. Entonces, para todo $\varepsilon > 0$ se tiene que ocurre alguna de las proposiciones siguientes:*

- *$\#S = O_{\alpha,\epsilon}(N^{\varepsilon})$,*
- *existe una curva algebraica plana C de grado $O_{\alpha,\varepsilon}(1)$ tal que al menos $(1-\varepsilon)\#S$ puntos de S se encuentran en C.*

Podemos pensar que un "conjunto aleatorio" S verifica que para cada primo p, se ocupan con igual probabilidad todas las clases residuales módulo p. Entonces los conjuntos mal distribuidos son en algún sentido opuestos a los conjuntos aleatorios. El teorema anterior lo que dice es que si $S \subseteq \mathbb{Z}^2$ no es un conjunto aleatorio, si consideramos $S_N = S \cap [N]^2$, la "no aleatoriedad" de S es debida a que las truncaciones S_N crecen en tamaño lentamente cuando $N \to +\infty$, o a bien crecen manteniendo algún tipo de estructura algebraica cuando $N \to +\infty$.

Veamos con varios ejemplos cómo se manifiesta la mala distribución de un conjunto del plano. En lo que sigue, dado $S \subseteq \mathbb{Z}^2$, denotamos $S_p := \{(x \pmod p, y \pmod p) : (x, y) \in S\}$.

EJEMPLO 3.4. *Sea $S \subseteq \mathbb{Z}^2$ un singletón, por ejemplo $S = \{(0,0)\}$. El conjunto S ocupa una sola clase residual módulo p para todo primo p, con lo que $\#S_p \leq \frac{1}{2}p$ para todo primo positivo. El conjunto S está mal distribuido, pero podemos atribuir la mala distribución al tamaño del conjunto, que es el menor posible.*

EJEMPLO 3.5. *Sea $S = \{(n,n) : n \in [N]\}$, es decir, la diagonal de $[N]^2$. Para todo primo $p \leq N$ se tiene que S ocupa sólo las diagonales de las clases residuales módulo p, es decir, p clases residuales. Ahora, si $q > N$ es primo, tenemos que S ocupa de las q^2 clases residuales, sólo N, que es mucho más chico que q para q grande. Dado que $\#S = N = \frac{N}{q}q$, en este caso, si $\alpha = \frac{N}{q}$ se tiene $\#S_p \leq \alpha p$ para todo p primo. El conjunto S está mal distribuido y tiene N elementos en $[N]^2$. Podemos atribuir la mala distribución de S a su estructura algebraica, ya que S es el conjunto de puntos enteros de la curva algebraica definida por $P(x,y) = x - y = 0$.*

EJEMPLO 3.6. *Consideramos ahora un análogo en el plano del Ejemplo 3.2. Sea $S := \{(a^2, a) : a^2 \leq N, a \in \mathbb{Z}\} \subseteq \mathbb{Z}^2$. Por la ley de reciprocidad cuadrática, el conjunto ocupa $\frac{1}{2}(p+1)$ clases residuales módulo p para todo primo $p > 2$, por lo tanto se trata de un conjunto mal del plano distribuido. S posee $O(N^{\frac{1}{2}})$ elementos. Podemos atribuir la mala distribución de S a su estructura algebraica, ya que S es el conjunto de puntos enteros de la curva algebraica definida por $Q(x,y) = x - y^2$.*

Un poco más en general, si k es un entero positivo fijo, sea $N \geq k$. Consideramos $S := \{(a^2, b) : a^2 \leq N, a \in [N], b \in [k]\} \subseteq [N]^2$. Para todo primo $p > 2$, el conjunto S ocupa $\frac{1}{2}(p+1)k$ clases residuales módulo p, con lo que S es un conjunto mal distribuido. Se tiene que S posee $O(N^{\frac{1}{2}})$. La mala distribución de S en este caso se puede atribuir a que S posee estructura algebraica, al estar contenido en k curvas algebraicas de grado pequeño.

EJEMPLO 3.7. *Recordamos el Ejemplo 3.3. Las puntos enteros de una curva algebraica conforman un conjunto mal distribuido. Su mala distribución se puede atribuir a que están contenidos en una curva algebraica.*

Es importante notar que puede pensarse que todo conjunto finito del plano tiene algún tipo de estructura algebraica. Si $S \subseteq [N]^2$, el polinomio

$$P(x,y) = \prod_{a:(a,y)\in S \text{ para algún } y} (x-a),$$

define una curva algebraica cuyos puntos enteros contienen a S. La caracterización de Helfgott y Venkatesh pone en manifiesto que, además de estar contenido en una curva algebraica, el grado de la curva debe ser pequeño.

2. Puntos racionales en curvas sobre cuerpos finitos

El problema de contar puntos en curvas o superficies sobre cuerpos finitos es bastante profundo; Weil formuló varias conjeturas sobre el tema, que sirvieron para el desarrollo de la geometría algebraica. Podríamos dedicar toda la tesis a este tema, para abordarlo de manera superficial; por esta razón, vamos a contar brevemente lo que dicen las conjeturas de Weil y explicar qué relación tienen con el problema de interés de la tesis, que son los conjuntos mal distribuidos.

Sea V un conjunto algebraico proyectivo sobre $\mathbb{F}_q$, definida por el sistema de ecuaciones

$$\begin{cases} f_1(x_1, \cdots, x_n) = 0 \\ \vdots \\ f_k(x_1, \cdots, x_n) = 0, \end{cases}$$

con $f_i \in \mathbb{F}_q[x_1, \cdots, x_n]$ formas homogéneas. Consideremos $\mathbb{F}_{q^r}$ la extensión de $\mathbb{F}_q$ de grado r. Para calcular $\#V(\mathbb{F}_{q^r})$ se considera la función zeta asociada a V, que es la función generatriz exponencial

$$Z(V, \mathbb{F}_q; x) := \exp\left(\sum_{r=1}^{+\infty} (\#V(\mathbb{F}_{q^r}))\frac{x^r}{r}\right).$$

Estamos usando que la exponencial de una serie $F(x) \in \mathbb{Q}[[x]]$ que no posee término constante la definimos como

$$\exp(F(x)) := \sum_{k=0}^{+\infty} \frac{F(x)^k}{k!}.$$

Conocer la función generatriz determina las cantidades $\#V(\mathbb{F}_{q^r})$, pues se tiene

$$\#V(\mathbb{F}_{q^r}) := \frac{1}{(n-1)!}\left(\frac{d}{dx}\right)^n |_{x=0}(\log(Z(V, \mathbb{F}_q; x))).$$

Calculemos la función $Z(V, \mathbb{F}_q, x)$ para V particulares.

Ejemplo 3.8. *Consideremos $V = \mathbb{A}^n$ el plano afín asociado a una clausura algebraica de $\mathbb{F}_q$. Se tiene $\#V(\mathbb{F}_{q^r}) = q^{rn}$. Luego*

$$\log(Z(\mathbb{A}^n, \mathbb{F}_q; x)) = \sum_{r=1}^{+\infty} q^{rn}\frac{x^r}{r} = -\log(1 - q^n x).$$

Exponenciando, obtenemos

$$Z(\mathbb{A}^n, \mathbb{F}_q; x) = \frac{1}{1 - q^n x}.$$

Ejemplo 3.9. *Consideremos $V = \mathbb{P}^n$ el plano proyectivo asociado a una clausura algebraica de $\mathbb{F}_q$. Puesto que un elemento de $V(\mathbb{F}_{q^r})$ está representado por una coordenada homogénea $(x_0 : \cdots : x_n)$ con todos los x_i no nulos, y dos coordenadas homogéneas son equivalentes si y sólo si una es un múltiplo de la otra por un escalar no nulo de $\mathbb{F}_{q^r}$, se tiene*

$$\#\mathbb{P}^n(\mathbb{F}_{q^r}) = \frac{q^{r(n+1)} - 1}{q^r - 1} = \sum_{i=0}^{n} q^{ri},$$

luego

$$\log(Z(\mathbb{P}^n, \mathbb{F}_q; x)) = \sum_{r=1}^{+\infty}\left(\sum_{i=0}^{n} q^{ri}\right)\frac{x^r}{r} = \sum_{i=0}^{n} -\log(1 - q^i x).$$

Exponenciando, llegamos a

$$Z(\mathbb{P}^n, \mathbb{F}_q; x) = \prod_{i=0}^{n} \frac{1}{1 - q^i x}. \tag{2.1}$$

Observación 3.1. *Notemos que la función zeta en los dos ejemplos anteriores tiene coeficientes racionales, luego $Z(\mathbb{A}^n, \mathbb{F}_q; x), Z(\mathbb{P}^n, \mathbb{F}_q; x) \in \mathbb{Q}(x)$. Adicionalmente, la expresión (2.1) se parece mucho a la fórmula del producto para los primos de la función zeta usual*

$$\zeta(z) := \sum_{n=1}^{+\infty} \frac{1}{n^z}.$$

Observación 3.2. *Sea $V \subseteq \mathbb{A}^n$ (respectivamente $V \subseteq \mathbb{P}^n$). Resulta $\#V(\mathbb{F}_{q^r}) \leq q^{rn}$ (respectivamente $\#V(\mathbb{F}_{q^r}) \leq \sum_{i=0}^{r} q^{rn}$). Luego la serie generatriz $Z(X, \mathbb{F}_q, x)$ converge para todo x con $|x| < q^{-n}$.*

Observación 3.3. *Sea $\overline{\mathbb{F}_q}$ una clausura algebraica de $\mathbb{F}_q$, y consideremos el morfismo de Frobenius $F_q : x \in \overline{\mathbb{F}_q} \mapsto x^q \in \overline{\mathbb{F}_q}$. Sea $V \subseteq \mathbb{A}^n$ o $\mathbb{P}^n$ un conjunto algebraico. Tenemos que F_q actúa sobre $V(\overline{\mathbb{F}_q})$ (el conjunto de puntos de V con coordenadas en $\overline{\mathbb{F}_q}$) mediante*

$$F_q \cdot (t_1, \cdots, t_n) = (t_1^q, \cdots, t_n^q).$$

Vamos a decir que $\mathcal{P}$ es un primo de V si es una órbita de la acción de F_q sobre $V(\overline{\mathbb{F}_q})$. Definimos el grado del primo $\mathcal{P}$ como el cardinal de la órbita a la que está asociado. Denotemos al grado de $\mathcal{P}$ como $\deg(\mathcal{P})$ *y* $n\mathcal{P} := q^{\deg P}$. *Si*

$$B_r := \#\{\mathcal{P} \text{ primo de } X : \deg(\mathcal{P}) = r\},$$

de la descomposición de órbitas por la acción de F_q tenemos la relación

$$\#V(\mathbb{F}_{q^r}) = \sum_{i|r} iB_i.$$

Entonces, se tiene, dentro del dominio de convergencia:

$$\sum_{r=1}^{+\infty} \#V(\mathbb{F}_{q^r}) \frac{x^r}{r} = \sum_{r=1}^{+\infty} \frac{x^r}{r} \sum_{i|r} iB_i = \sum_{r=1}^{+\infty} \sum_{i=1}^{+\infty} iB_i 1_{i|r}(i) \frac{x^r}{r}.$$

Intercambiando el orden de sumación:

$$\sum_{r=1}^{+\infty} \sum_{i=1}^{+\infty} iB_i 1_{i|r}(i) \frac{x^r}{r} = \sum_{i=1}^{+\infty} iB_i \sum_{l=1}^{+\infty} \frac{x^{li}}{li} = \sum_{i=1}^{+\infty} B_i \log(\frac{1}{1-x^i}) = \log\left[\prod_{i=1}^{+\infty} \left(1-x^i\right)^{-B_i}\right].$$

Exponenciando, obtenemos la expresión para la función zeta

$$Z(V, \mathbb{F}_q; x) = \prod_{i=1}^{+\infty} \left(1-x^i\right)^{-B_i} = \prod_{\mathcal{P} \text{ primo de } V} \left(1 - x^{\deg(\mathcal{P})}\right)^{-1}. \tag{2.2}$$

Poniendo $x = q^{-s}$ obtenemos la expresión

$$Z(V, \mathbb{F}_q; q^{-s}) = \prod_{\mathcal{P}} \left(1 - n\mathcal{P}^{-s}\right)^{-1}. \tag{2.3}$$

Las expresiones (2.2) *y* (2.3) *son muy similares a las fórmulas del producto de la función zeta de Riemann clásica y a la función zeta de Dedekind; de ahí que la función $Z(V, \mathbb{F}_q; x)$ se la llame función zeta del conjunto algebraico V.*

Calculamos la función zeta en dos ejemplos, que serían los más sencillos y fundamentales. En general el problema de calcular la función zeta de un conjunto algebraico afín o proyectivo no es fácil. El estudio de las propiedades de $Z(V, \mathbb{F}_q; x)$ es el punto de partida de las conjeturas de Weil.

Teorema 3.2 (Conjeturas de Weil). *Sea V un conjunto algebraico proyectivo suave sobre $\mathbb{F}_q$, geométricamente irreducible, de dimensión n.*

(1) *Racionalidad:*

$$Z(V, \mathbb{F}_q; x) \in \mathbb{Q}(x).$$

(2) *Ecuación funcional: existe un entero ϵ llamado la característica de Euler de la variedad V, tal que*

$$Z(V, \mathbb{F}_q, \frac{1}{q^n x}) = \pm q^{n\frac{\epsilon}{2}} x^\epsilon Z(V, \mathbb{F}_q, x).$$

(3) *Hipótesis de Riemann: la función zeta se factoriza como*

$$Z(V, \mathbb{F}_q; x) = \frac{\prod_{i=1}^{n-1} P_{2i-1}(x)}{\prod_{i=0}^{n} P_{2i}(x)},$$

con cada $P_i(x) \in \mathbb{Z}[x]$, con término independiente 1, *y tal que se factoriza en $\mathbb{C}$ como*

$$P_i(x) = \prod_{j=1}^{b_i} (1 - \alpha_{ij} x) \text{ con } |\alpha_{ij}| = q^{\frac{i}{2}}.$$

La cantidad $b_i = \deg(P_i)$ se llama el i-ésimo número de Betti de V

Tomando $x = q^{-s}$, notemos que lo que dice el tercer enunciado del Teorema 3.2 es que los polos de la función $Z(V, \mathbb{F}_q; x)$ deberían tener parte real en $\{0, 1, \cdots, n\}$ mientras que los ceros deberían tener parte real en $\{\frac{1}{2}, \frac{3}{2}, \cdots, \frac{2n-1}{2}\}$. Por esta razón es que el tercer enunciado se lo conoce como hipótesis de Riemann.

Observación 3.4. *Es importante notar que, si bien en el Ejemplo 3.8, vimos que la función zeta asociada a la variedad afín $\mathbb{A}^n$ era racional. En general, no es cierto que una variedad afin verifica que su función zeta asociada es racional.*

La demostración de los resultados del Teorema 3.2 requirió del trabajo de muchos matemáticos importantes como Weil, Artin, Grothendieck, Dwork y Deligne. Para esta tesis sólamente nos va a interesar el tercer enunciado del Teorema 3.2 en el caso que V sea una curva. Esto se debe a que en este caso, la hipótesis de Riemann nos permite dar una estimación para $\#V(\mathbb{F}_q)$ y por lo tanto nos permite obtener datos acerca de la distribución de V módulo $q = p$ con p primo.

La estimación de $\#V(\mathbb{F}_q)$ que necesitamos se encuentra dentro de un resultado debido a Lang y Weil de 1956 [**25**], que no enunciamos en su mayor generalidad:

Teorema 3.3 (Lang-Weil). *Sean n, d, enteros positivos. Existe una constante $A := A(n, d)$ tal que para toda $V \subseteq \mathbb{P}^{n+1}$ hipersuperficie geométricamente irreducible de grado d, definida sobre $\mathbb{F}_q$, se tiene*

$$|\#V(\mathbb{F}_{q^r}) - q^{rn}| \leq A(n, d)q^{r(n-\frac{1}{2})}.$$

En el caso de una curva algebraica proyectiva, con $n = 1$, obtenemos la estimación

$$\#V(\mathbb{F}_{q^r}) \leq A(1, d)q^{\frac{r}{2}} + q^r. \tag{2.4}$$

Tomando $q = p$ con $p \geq 2$ un primo, y $r = 1$ la estimación (2.4) nos da una noción de cómo está distribuida módulo p una curva algebraica plana. En efecto, sea C una curva algebraica plana, que vamos a suponer geométricamente irreducible, definida por la ecuación $F(x, y) = 0$ con $F \in \mathbb{Z}[x, y]$. Como siempre podemos asumir que los coeficientes de F son coprimos, el polinomio $\tilde{F}$ reducido es no nulo, de grado a lo sumo $\deg(F)$. Si proyectivizamos en $\mathbb{P}^2$, obtenemos una curva algebraica proyectiva $\tilde{C} = V$ de grado a lo sumo $\deg(F)$. Utilizando la desigualdad (2.4) obtenemos la estimación

$$\#V(\mathbb{F}_p) \leq A(1, \deg(F))p^{\frac{1}{2}} + p = O_{\deg(F)}(p^{\frac{1}{2}}) + p.$$

Puesto que al proyectivizar agregamos puntos, lo que acabamos de obtener es una estimación para los puntos racionales en C, a saber

$$\#C(\mathbb{F}_p) = O_{\deg(F)}(p^{\frac{1}{2}}) + p,$$

que dice que para p grande, dependiente del grado de F, la curva C ocupa pocas clases residuales módulo p. De aquí concluímos que las curvas algebraicas del plano son conjuntos mal distribuidos, pero la mala distribución es debida a la estructura algebraica que poseen. De la misma forma que para curvas, es posible obtener el enunciado análogo para hipersuperficies, es decir, las hipersuperficies algebraicas son conjuntos mal distribuidos.

El teorema 3.3 para curvas, es decir, para $n = 1$, puede verse que es una reformulación de la hipótesis de Riemann sobre curvas definidas en cuerpos finitos. Aún el caso para curvas de la hipótesis de Riemann no es sencillo, sin embargo hay un argumento "sencillo" de Bombieri [**4**], que apela a ideas de Stepanov [**51**], que sólamente utiliza el teorema de Riemann-Roch.

Una vez que se tiene el resultado para curvas, el Teorema 3.3 se puede probar mediante un argumento inductivo, haciendo inducción en la dimensión r de la variedad. Los detalles de la demostración pueden consultarse en [**25**].

3. La criba como una técnica analítica

El método más antiguo para estimar la cantidad de primos en un intervalo $[N]$ viene dada por la criba de Eratosthenes. Este método es bastante sencillo: sustraemos el 1 de $[N]$, que no es un número primo, luego sustraemos de los números del intervalo $[N]$ los múltiplos propios[4] de 2, luego los múltiplos propios de 3, y continuamos sustrayendo múltiplos propios de los primos menores o iguales a N. Finalizado este proceso de eliminación, los números que no fueron eliminados del intervalo $[N]$ son todos los primos menores o iguales a N. La siguiente tabla muestra este proceso para $N = 30$.

~~1~~	2	3	~~4~~	5	~~6~~	7	~~8~~	~~9~~	~~10~~
11	~~12~~	13	~~14~~	~~15~~	~~16~~	17	~~18~~	19	~~20~~
~~21~~	~~22~~	23	~~24~~	~~25~~	~~26~~	~~27~~	~~28~~	29	~~30~~

[4]Por múltiplo propio de p nos referimos a un número de la forma kp con $k > 1$.

En la tabla, los números que están marcados como ~~1~~ son múltiplos de 2, los números que están marcados como ~~9~~ son múltiplos de 3 y los números marcados como ~~25~~ son múltiplos de 5. Los números que tienen más de una marca, se debe a que son múltiplos de más de un primo. Por ejemplo, tenemos marcado ~~30~~ debido a que 30 es múltiplo de 2, 3 y 5.

El método de la criba de Eratosthenes se puede adaptar para estimar los primos del intervalo $[\sqrt{N}, N]$, sustrayendo los múltiplos de todos los primos $p \leq \sqrt{N}$. Este proceso puede verse en la siguiente tabla.

~~1~~	~~2~~	~~3~~	~~4~~	~~5~~	~~6~~	7	~~8~~	~~9~~	~~10~~
11	~~12~~	13	~~14~~	~~15~~	~~16~~	17	~~18~~	19	~~20~~
~~21~~	~~22~~	23	~~24~~	~~25~~	~~26~~	~~27~~	~~28~~	29	~~30~~

En sí, el método que describimos no es práctico para N grande, pues coincide con aplicar el principio de inclusión-exclusión. Viggo Brun [5] en el trabajo [**8**] del 1915, prueba, adaptando la criba de Eratosthenes, que existen infinitos infinitos enteros positivos n tales que n y $n+2$ tienen a lo sumo nueve factores primos y que todo entero par positivo suficientemente grande se escribe como suma de dos enteros positivos, cada con a lo sumo nueve factores primos. En [**9**], Brun obtiene:

TEOREMA 3.4 (Brun). *Sea $\mathcal{P}$ el conjunto de los primos positivos. La serie $\sum_{p:p\in\mathcal{P},p+2\in\mathcal{P}} \frac{1}{p}$ es convergente.*

Los trabajos de Brun no recibieron mucha atención hasta que Selberg, en 1947, desarrolló un método, de la misma naturaleza que el método de Brun, con el que mostró que una proporción positiva de los ceros de la zeta de Riemann se encuentran en la recta $\mathrm{Re}(s) = \frac{1}{2}$, y Linnik, en 1941, desarrolló el método de la criba grande, que posteriormente fue mejorado por Rényi (1950), Roth (1965), Bombieri (1965), Davenport y Halberstam (1966), Gallagher (1967) y otros. Explicar en profundidad las distintas técnicas que ahora reciben el nombre de "métodos de criba" escapa los objetivos de esta tesis. Existen numerosos libros y notas que explican en detalle las técnicas de criba, como [**10**], [**19**] y la exposición de Friedlander [**14**]; nosotros explicaremos brevemente la idea general detrás de estas técnicas.

Sea $\mathcal{A}$ un conjunto finito de objetos y $\mathcal{P}$ un índice de primos tales que para cada $p \in \mathcal{P}$ tenemos asociado un subconjunto $\mathcal{A}_p \subseteq \mathcal{A}$. El "problema de la criba" es dar cotas superiores e inferiores para

$$S(\mathcal{A}, \mathcal{P}) := \mathcal{A} \setminus \bigcup_{p\in\mathcal{P}} \mathcal{A}_p. \tag{3.1}$$

La razón por la cual el problema de estimar $\#S(\mathcal{A}, \mathcal{P})$ se conoce como el "problema de la criba" es la siguiente: si hacemos una tabla con los elementos de $\mathcal{A}$, lo que queremos estudiar son los elementos de $\mathcal{A}$ que no poseen elementos de $\mathcal{A}_p$ para ningún p. Consideramos, digamos, $\mathcal{A}_2$. Tachamos de la tabla que hicimos con $\mathcal{A}$ las entradas comunes con $\mathcal{A}_2$. Luego consideramos $\mathcal{A}_3$, procedemos igual, y así seguimos hasta terminar. El conjunto $S(\mathcal{A}, \mathcal{P})$ está compuesto por las entradas de la tabla que no fueron tachadas. El nombre de "criba" viene del hecho de que esto mismo es lo que se hace en la criba de Eratosthenes.

Por el principio de inclusión-exclusión , tenemos una expresión para $\#S(\mathcal{A}, \mathcal{P})$, sin embargo, dicha expresión no es práctica. Por esa razón interesa dar cotas para $S(\mathcal{A}, \mathcal{P})$ y no dar una fórmula explícita.

Un planteo típico del problema de criba es considerar $\mathcal{A}$ un conjunto finito de enteros positivos, $\mathcal{P}$ los primos o los primos mayores que algún $c > 0$, y $\mathcal{A}_p$ un subconjunto de $\mathcal{A}$ que cumple algunas condiciones de congruencia módulo p. Por ejemplo:

$$\mathcal{P}_(x) = \{p \leq \sqrt{x}\},$$
$$\mathcal{A}(x) = \{1 \leq n \leq x : n \in \mathbb{Z}\},$$
$$\mathcal{A}_p(x) = \{1 \leq n \leq x : n \equiv 0(p)\}.$$

El problema de la criba en este caso es estimar la cantidad de primos que se obtienen de extraer aquellos números divisibles por algún primo $p \leq \sqrt{x}$, es decir, estimar los números primos en el intervalo $[\sqrt{x}, x]$; esta era una de las problemáticas para las que se aplicaba la criba de Eratosthenes.

Un problema en algún sentido "inverso" al problema de la criba es el siguiente. Supongamos que partimos de un conjunto S finito y queremos obtener alguna estimación para $\#S$. Entonces miramos módulo p la imagen de S, que denotamos S_p, para $p \in \mathcal{P}$ un conjunto de primos.

[5] Es probable que Jean Merlin [**34**] haya sido el primero en tratar de modificar la criba de Eratosthenes, en 1911. Desafortunadamente, Merlin murió en la Primera Guerra Mundial y solo dos de sus manuscritos sobrevivieron. Seguramente Brun habrá leído los trabajos de Merlin y fue inspirado por los mismos, conduciéndolo a su adaptación de la criba de Eratosthenes.

PREGUNTA 3.4. *¿Puede decirse algo de $\#S$ a partir de tener estimaciones para los $\#S_p$, para todo $p \in \mathcal{P}$?*

Para estudiar la pregunta anterior se usa en general la criba grande (large sieve), nombre que hace referencia a un conjunto de técnicas analíticas que comparten una estrategia común. También se emplea la criba más grande (larger sieve). Dado que la Pregunta 3.4 es de interés para nosotros, ya que está relacionada con obtener una estimación de la cantidad de soluciones enteras de una curva algebraica a partir de tener una estimación de la cantidad de soluciones módulo p de la curva reducida módulo p (ver la Pregunta 3 y los comentarios previos), vamos a explicar un poco la criba grande y la criba más grande.

Linnik introdujo en 1941 la criba grande. La motivación original de Linnik era estudiar la hipótesis de Vinogradov, relacionada con el tamaño del menor residuo no cuadrático $n_p \pmod p$ con p un primo. Vinogradov conjeturó que para todo $\varepsilon > 0$ se verificaba $n_p = O(p^\varepsilon)$. Con la criba grande, Linnik en [**26**] prueba:

TEOREMA 3.5 (Linnik). *Dado p primo, sea $n_p \pmod p$ el menor residuo no cuadrático módulo p. Entonces, para todo $\varepsilon > 0$, el número de primos $p \leq x$ para los que $n_p > p^\varepsilon$ es $O(\log\log(x))$.*

La criba grande de Linnik es un método desarrollado en [**27**], que da una estimación para conjuntos de la forma

$$(\mathcal{M}, \mathcal{P}, \Omega) := \{m \in \mathcal{M} : m \not\equiv w_{i,p} \pmod p \notin,\ \forall 1 \leq i \leq \omega(p), \forall p \in \mathcal{P}\},$$

donde $\mathcal{M}$ es un conjunto no vacío, finito, de enteros positivos, $\mathcal{P}$ es un conjunto de primos y Ω es un conjunto de enteros positivos tal que, para todo $p \in \mathcal{P}$, si Ω_p es el conjunto de clases residuales módulo p de Ω, se verifica $\Omega_p = \{w_{1,p}, \cdots, w_{\omega(p),p}\}$. Esencialmente, $(\mathcal{M}, \mathcal{P}, \Omega)$ es el conjunto que se obtiene luego de eliminar los elementos de $\mathcal{M}$ que verifican ciertas condiciones de congruencia módulo p, para $p \in \mathcal{P}$.

La estimación de $\#(\mathcal{M}, \mathcal{P}, \Omega)$ de Linnik [**27**] se deduce de una desigualdad:

TEOREMA 3.6 (Desigualdad de la criba grande). *Sea $(a_n)_n$ una sucesión de números complejos y sean x, z enteros positivos. Entonces*

$$\sum_{d \leq z} \sum_{1 \leq a \leq d, \gcd(a,d)=1} |\sum_{n \leq x} a_n \exp(\frac{2\pi i n a}{d})|^2 \leq (z^2 + 4\pi x) \sum_{n \leq x} |a_n|^2.$$

TEOREMA 3.7 (Criba grande). *Sean $\mathcal{M}$ un conjunto no vacío, finito, de enteros positivos, $\mathcal{P}$ un conjunto de primos positivos y Ω un conjunto, no vacío, de enteros positivos tal que para todo $p \in \mathcal{P}$, si Ω_p es el conjunto de clases residuales módulo p, se tiene $\Omega_p = \{w_{1,p}, \cdots, w_{\omega(p),p}\}$. Sea $z > 0$. Sea $P(z) := \prod_{p \in \mathcal{P}: p < z} p$. Sean*

$$(\mathcal{M}, \mathcal{P}, \Omega, z) := \{m \in \mathcal{M} : m \pmod p \notin \Omega_p,\ \forall p \in \mathcal{P},\ p < z\},$$

$$L(z) := \sum_{d \leq z} \mu^2(d) \prod_{p | d} \frac{\omega(p)}{p - \omega(p)},$$

donde μ es la función de Möbius, definida como

$$\mu(d) = \begin{cases} 1 \text{ si } d = 1 \\ 0 \text{ si } d \text{ no es libre de cuadrados} \\ (-1)^k \text{ si } d = \prod_{i=1}^{k} p_i, \text{ con los enteros } p_i \text{ todos primos distintos} \end{cases}.$$

Entonces, tenemos la estimación

$$\#(\mathcal{M}, \mathcal{P}, \Omega, z) \leq \frac{z^2 + 4\pi x}{L(z)}.$$

Desde Linnik, la criba grande evolucionó hasta ser una técnica apicable a muchos problemas matemáticos diversos (consultar [**21**] para ver varias aplicaciones de la criba grande en distintos contextos). La evolución de la técnica fue debida a que la desigualdad del Teorema 3.6 se abstrajo adecuadamente.

Una adaptación particularmente útil de la criba grande, es la criba más grande de Gallaher [**15**]. Esta adaptación da una estimación de $\#(\mathcal{M}, \mathcal{P}, \Omega)$, permitiendo que el conjunto $\mathcal{P}$ contenga potencias de primos.

TEOREMA 3.8 (criba más grande). *Sean $\mathcal{M}$ un conjunto no vacío, finito, de enteros positivos y $\mathcal{T}$ un conjunto no vacío, de potencias de primos positivos. Para cada $t \in \mathcal{T}$, definimos $\mathcal{M}_t := \{x \pmod t : x \in \mathcal{M}\}$. Supongamos que para cada $t \in \mathcal{T}$ se tiene un número real positivo $u(t)$ tal que $\mathcal{M}$ ocupa a lo sumo $u(t)$ clases residuales módulo t, es decir, para todo $t \in \mathcal{T}$ se verifica $\#\mathcal{M}_t \leq u(t)$. Sea $X := \max_{m \in \mathcal{M}}(|m|)$ y Λ la función de von Mangoldt, definida como*

$$\Lambda(t) := \begin{cases} \log(p) \text{ si } n = p^a, a \in \mathbb{N} \\ 0 \text{ en otro caso} \end{cases}.$$

Si

$$\sum_{t \in \mathcal{T}} \frac{\Lambda(t)}{u(t)} - \log(2X) > 0,$$

entonces

$$\#\mathcal{M} \leq \frac{\sum_{t \in \mathcal{T}} \Lambda(t) - \log(2X)}{\sum_{t \in \mathcal{T}} \frac{\Lambda(t)}{u(t)} - \log(2X)}.$$

Como aplicación del Teorema 3.8, Gallagher en [**15**] prueba los siguientes dos resultados:

TEOREMA 3.9 (Gallagher). *Sean a, b enteros positivos tales que para todo t potencia de un primo existe ν_t con la propiedad*

$$b \equiv a^{\nu_t} \pmod{t}.$$

Entonces existe un entero ν tal que

$$b \equiv a^{\nu} \pmod{t}.$$

TEOREMA 3.10 (Gallagher). *Sea $S \subseteq [N]$ un subconjunto tal que existe $0 < \alpha < 1$ y $c > 0$ tal que S ocupa a lo sumo αp clases residuales módulo p, para todo primo $p > c$. Entonces*

$$\#S \ll_\alpha N^\alpha.$$

Sea $S \subseteq \mathbb{Z}^2$. De las técnicas de criba que mencionamos, la criba grande y la criba más grande parecen adecuadas para obtener una estimación del tamaño de S, a partir de estimaciones de la cantidad de clases residuales módulo p del conjunto S.

4. El método del determinante como una criba

Lo que vamos a hacer ahora es adaptar la criba más grande de Gallagher, tomando como base el método del determinante. Empezamos probando el último teorema de la sección anterior, que recordamos a continuación.

TEOREMA 3.11 (Gallagher). *Sea $S \subseteq [N]$ un subconjunto tal que existe $0 < \alpha < 1$ y $c > 0$ tal que S ocupa a lo sumo αp clases residuales módulo p, para todo primo $p > c$. Entonces*

$$\#S \ll_\alpha N^\alpha.$$

Antes de dar la demostración del Teorema 3.11, recordamos las siguientes estimaciones:

$$\sum_{p \leq n} \frac{\log(p)}{p} = \log(n) + O(1) \tag{4.1}$$

$$\sum_{p \leq n} \log(p) = O(n) \tag{4.2}$$

DEMOSTRACIÓN. Vamos a contar de dos maneras distintas los pares x, y y los $p > c$ primos que verifican $x \equiv y \pmod{p}$. Para esto consideramos la magnitud

$$\Delta := \prod_{x,y \in S : x \neq y} (x - y).$$

Sea $p > c$. Definimos

$$S(a; p) := \{x \in S : x \equiv a \pmod{p}\}.$$

Observamos que $S = \coprod_{a \pmod p} S(a; p)$. Notemos $A_p := \#\{x \pmod p : x \in S\}$. Luego, usando Cauchy-Schwartz:

$$(\#S)^2 = \left(\sum_{a \pmod p} \#S(a;p) \right)^2 \leq A_p \sum_{a \pmod p} \#S(a;p)^2 \leq \alpha p \sum_{a \pmod p} \#S(a;p)^2.$$

Puesto que

$$\sum_{a \pmod p} \#S(a;p)^2 = \sum_{a \pmod p} \sum_{\{(x,y)\in S^2 : x\equiv y\equiv a \pmod p\}} 1,$$

podemos particionar el conjunto $\{(x,y) \in S^2 : x \equiv y \equiv a \pmod p\}$ en la diagonal y su complemento, con lo que la anterior suma es igual a

$$\#\{(x,y) \in S^2 : x \neq y, x \equiv y(p)\} + \#S.$$

Concluímos entonces

$$\#\{(x,y) \in S^2 : x \neq y, x \equiv y(p)\} \geq \frac{\#S^2}{\alpha p} - \#S.$$

Sea $\nu_p(\Delta)$ el exponente de p en la factorización de Δ. Este exponente cuenta exactamente la cantidad de $x \neq y$ con $x \equiv y(p)$. Por lo anterior,

$$\nu_p(\Delta) \geq \frac{\#S^2}{\alpha p} - \#S.$$

Entonces, tenemos la cota

$$|\Delta| \geq \prod_{c<p\leq Q} p^{\nu_p(\Delta)} \geq \prod_{c<p\leq Q} e^{\nu_p(\Delta)\log(p)} \geq \prod_{c<p\leq Q} e^{(\frac{\#S^2}{\alpha p} - |S|)\log(p)} = e^{\frac{\#S^2}{\alpha} \sum_{c<p\leq Q} \frac{\log(p)}{p} - \#S \sum_{c<p\leq Q} \log(p)}.$$

Usando las estimaciones de (4.1) y (4.2), concluímos

$$\#\Delta \geq e^{\frac{\#S^2}{\alpha}\log(Q) - \#SO(Q) + O(1)} = e^{\frac{\#S^2}{\alpha}(\log(Q) - \frac{O(Q)}{\#S} + O_\alpha(1))}.$$

Por otro lado, tenemos la cota superior trivial

$$\#\Delta \leq \prod_{(x,y)\in S^2, x\neq y} N \leq N^{\#S^2},$$

con lo que comparando con la cota inferior y tomando logaritmos, tenemos luego de simplificar $\#S^2$, la estimación

$$\frac{1}{\alpha}(\log(Q) - \frac{O(Q)}{\#S} + O_\alpha(1)) \leq \log(N).$$

Tomando ahora $Q = \#S$ y exponenciando, llegamos a $\#S = O_\alpha(N^\alpha)$. □

En lo que sigue, vamos a adaptar la demostración del Teorema 3.11, para generalizar a conjuntos $S \subseteq [N]^2$ tal que para cada p se ocupan a lo sumo αp clases residuales. Esta generalización es:

Teorema 3.12 (Helfgott-Venkatesh). *Sea $S \subseteq [N]^2 \subseteq \mathbb{Z}^2$ con $N \geq 1$. Supongamos que existe $\alpha > 0$ tal que $\#S_p \leq \alpha p$ para todo primo p. Entonces, para todo $\varepsilon > 0$ se tiene que ocurre alguna de las proposiciones siguientes:*

- $\#S = O_{\alpha,\epsilon}(N^\varepsilon)$,
- *existe una curva algebraica plana C de grado $O_{\alpha,\varepsilon}(1)$ tal que al menos $(1-\varepsilon)\#S$ puntos de S se encuentran en C.*

En la demostración del Teorema 3.11, la función $w(x,y) = x - y$ detectaba cuándo dos puntos eran distintos. En dos variables, la idea es encontrar una función w que mida cuándo finitos puntos $\{P_i\}_i$ están en una curva algebraica plana de grado pequeño: esto va a permitir "detectar" si el conjunto de puntos $\{P_i\}_i$ en el que evaluamos w posee algún tipo de estructura algebraica. En el capítulo 1 observamos que en este tipo de problema se podía aplicar el método del determinante de Bombieri-Pila: la función w puede ser un determinante, que si vale 0, entonces los puntos $\{P_i\}$ están contenidos en una curva algebraica. Esta idea motiva las siguientes definiciones.

Sea $\mathcal{M}$ un subconjunto finito de $\mathbb{Z}[x,y]$ tal que

- es un conjunto linealmente independiente.

- el polinomio $P(x,y)=1$ se encuentra en $\mathcal{M}$.
- separa puntos, es decir, para todo $P_1, P_2 \in \mathbb{Z}^2$ existe $F \in \mathcal{M}$ tal que $F(P_1) \neq F(P_2)$.

Vamos a denotar $d_{\mathcal{M}}$ al grado total de $\mathcal{M}$, o sea, la suma de los grados de todos los polinomios de $\mathcal{M}$ y $D = \#\mathcal{M}$.

DEFINICIÓN 3.5. *Una $\mathcal{M}$-curva es una curva algebraica plana definida por un polinomio $g(x,y)=0$ con g una combinación lineal de los elementos de $\mathcal{M}$.*

Como ejemplos, tenemos los $\mathcal{M}$ de antes:

EJEMPLO 3.10. *Sea $\mathcal{M}_d := \{x^i y^j : i+j \leq d\}$. Entonces $\#\mathcal{M}_d = \frac{1}{2}(d+1)(d+2), d_{\mathcal{M}} = \frac{1}{2}d(d+1)(d+2)$. Las $\mathcal{M}$-curvas son las curvas algebraicas de grado a lo sumo d.*

EJEMPLO 3.11. *Sea $\mathcal{M}_{d,M} := \{x^i y^j : i \leq d, j \leq M\}$. Entonces se tiene $\#\mathcal{M}_{d,M} = (d+1)(M+1)$ y $d_{\mathcal{M}_{d,M}} = (d+1)(M+1)\frac{d+M}{2}$*

Vamos a probar el siguiente resultado para un $\mathcal{M}$ general, del que deduciremos el Teorema 3.12.

TEOREMA 3.13 (Helfgott-Venkatesh). *Sea $S \subseteq [N]^2, N \geq 1$. Supongamos que existe $\alpha > 0$ tal que $\#S_p \leq \alpha p$ para todo primo $p > c$.*

Sea $\mathcal{M} \subseteq \mathbb{Z}[x,y]$ con las condiciones que dijimos antes. Entonces, para todo $\delta \in (0,1)$ se cumple algunas de las siguientes condiciones:

(1) *existe una $\mathcal{M}$-curva que contiene al menos $\delta\#S$ puntos de S.*
(2) *se tiene $\#S \ll_{\alpha,\delta,\mathcal{M}} (N^{\frac{2\alpha d_{\mathcal{M}}}{D(D+1)}+O_{\alpha,\mathcal{M}}(\delta)})$.*

Observemos que este teorema, que en sí constituye la adaptación de la criba de Gallagher, contiene exactamente la "dicotomía" del Teorema 3.12, y esto se refleja en que de dicho resultado se deduce como corolario el Teorema 3.13. La razón de esto es que si consideramos un conjunto $S \subseteq [N]^2$, si el conjunto no es pequeño, el Teorema 3.13 dice que una buena proporción está contenida en una $\mathcal{M}$-curva. Debido a que queremos mantener un control del grado de estas curvas, vamos a terminar tomando particiones sucesivas del conjunto S, cada una cubierta por una curva algebraica plana de grado pequeño, y la curva en cuestión termina siendo la unión de todas estas curvas (en términos de polinomios, el producto de todos los polinomios que definen a cada curva). Él punto crucial del Teorema 3.13 reside en que podemos elegir $\mathcal{M}$ adecuado.

DEMOSTRACIÓN DEL TEOREMA 3.12. Vamos a considerar $\mathcal{M}$ del primer ejemplo; recordamos que $D = \frac{1}{2}(d+1)(d+2)$ y $d_{\mathcal{M}} = \frac{1}{2}d(d+1)(d+2) = \frac{2}{3}dD$. Entonces

$$\frac{2d_{\mathcal{M}}\alpha}{D(D-1)} = \frac{4d\alpha}{D-1} = \Theta(\frac{8\alpha}{d+1}),$$

de donde concluímos que para d suficientemente grande, la cantidad anterior es menor a $\frac{\varepsilon}{2}$. Para δ suficientemente chico y este d tenemos que la cantidad $O_{\alpha,\mathcal{M}}(\delta)$ del exponente del Teorema 3.13 también es menor a $\frac{\varepsilon}{2}$.

Usamos ahora el Teorema 3.13. Por la elección de estos paramétros tenemos que, o bien $\#S \ll_{\alpha,\varepsilon} N^{\varepsilon}$ o bien existe una $\mathcal{M}$-curva algebraica C que contiene al menos $\delta\#S$ puntos de S. Supongamos que ocurre lo segundo. Sea $S' = S\backslash C$. Si $\#S' \leq \varepsilon\#S$ entonces $\#(S\backslash S') \geq (1-\varepsilon)\#S$ y, como $S\backslash S'$ se encuentra en C, terminamos. Supongamos entonces que $\#S' > \varepsilon\#S$. Aplicando de nuevo el Teorema 3.13 tenemos que o bien $\#S' \ll_{\alpha,\varepsilon} N^{\varepsilon}$ (lo que implica $\#S \ll_{\alpha,\varepsilon} \frac{1}{\varepsilon}N^{\varepsilon} \ll_{\alpha,\varepsilon} N^{\varepsilon}$) o existe una $\mathcal{M}$-curva algebraica que contiene al menos $\delta\#S'$ puntos de S'. Razonamos así de manera recursiva, terminando en j cuando $\#S^{(j)} \ll_{\alpha,\varepsilon} N^{\varepsilon}$ o $\#S^{(j)} \leq \varepsilon\#S$. Dado que $\#S' \leq (1-\delta)\#S, \#S'' \leq (1-\delta)\#S'$, tenemos que $\#S^{(j)} \leq (1-\delta)^j\#S$ con lo que tenemos $\#S^{(j)} \leq \varepsilon\#S$ si $j \geq \frac{\log(\varepsilon)}{\log(1-\delta)}$. Si ocurre $|S^{(j)}| \ll_{\alpha,\varepsilon} N^{\varepsilon}$ seguimos el mismo razonamiento que antes para obtener $\#S \ll_{\alpha,\varepsilon} N^{\varepsilon}$. Si ocurre lo segundo, como

$$S\backslash S^{(j)} = (S\backslash S') \cup (S\backslash S'') \cup \cdots (S^{(j-1)}\backslash S^{(j)}),$$

tenemos que cada de los conjuntos $S^{i-1}\backslash S^i$ está contenido en una curva algebraica de grado $O_{\alpha,\varepsilon}(1)$. Sea C la unión de tales curvas, que son a lo sumo $\frac{\log(\varepsilon)}{\log(1-\delta)}$. Entonces C es una curva de grado

$$\frac{\log(\varepsilon)}{\log(1-\delta)}O_{\alpha,\varepsilon}(1) = O_{\alpha,\varepsilon}(1),$$

con lo que terminamos. □

Consideremos $f_1, \cdots, f_D$ los elementos de $\mathcal{M}$ ordenados de alguna forma. Vamos a definir la función que mide si D puntos de S están en una $\mathcal{M}$-curva. Esta función, la notaremos W, y será un determinante. Concretamente, $W : (\mathbb{R}^2)^D \to \mathbb{R}$ está dada por

$$W(P_1, \cdots, P_D) = \det(f_i(P_j))_{1 \leq i,j \leq D}.$$

Tenemos que W cumple la siguiente propiedad, que va a reemplazar a la cota $p^{\nu_p(\Delta)}$ en la demostración de la Proposición 3.11, cuya demostración dejamos a cargo del lector.

PROPOSICIÓN 3.1. *Si el conjunto $\{P_j\}_{1 \leq j \leq D}$ tiene a lo sumo k clases residuales módulo p, con p primo, entonces $p^{D-k} \mid W(P_1, \cdots, P_D)$.*

DEMOSTRACIÓN DEL TEOREMA 3.13. Denotando $\mathbf{P} = (P_1, \cdots, P_D)$, vamos a decir que $\mathbf{P}$ es *admisible* si $W(\mathbf{P}) \neq 0$, y vamos a decir que $\mathbf{P}$ es *inadmisible* si $W(\mathbf{P}) = 0$. Entonces definimos

$$\Delta = \prod_{\mathbf{P}}^{*} |W(\mathbf{P})|,$$

donde $*$ denota el producto sobre todas las D-uplas admisibles de S^D. Tenemos que para cada $\mathbf{P} \in S^D$ se verifica $W(\mathbf{P}) = O_{\mathcal{M}}(N^{d_{\mathcal{M}}})$ con lo que entonces

$$\log(\Delta) = \sum_{\mathbf{P}}^{*} \log(|W(\mathbf{P})|) \leq \#S^D[N^{d_{\mathcal{M}}} + O_{\mathcal{M}}(1)].$$

Concluímos la cota superior

$$\frac{\log(\Delta)}{\#S^D} \leq d_{\mathcal{M}} \log(N) + O_{\mathcal{M}}(1).$$

Lo que vamos a hacer ahora es obtener una cota inferior para Δ y de ahí deducir que o bien S es pequeño o que S contiene pocas D-uplas admisibles, lo que va a terminar implicando que una proporción grande de S esté contenida en una $\mathcal{M}$-curva.

Sea Q un parámetro a definir después y sea $p \leq Q$ un primo. Para cada $x \in (\frac{\mathbb{Z}}{p\mathbb{Z}})^2$ sea ρ_x la fracción de puntos de S que tiene por residuo módulo p a x. Para cada $\mathbf{P}$ sea $\kappa(\mathbf{P}) \in \{0, 1, \cdots, p-1\}$ tal que $D - \kappa(\mathbf{P})$ es la cantidad de P_i's de $\mathbf{P}$ que dan restos distintos módulo p. Sea $\nu_p(\Delta)$ el exponente de p en la factorización de Δ. Por la proposición 3.1, tenemos la cota inferior

$$\nu_p(\Delta) \geq \sum_{\mathbf{P}}^{*} \kappa(\mathbf{P}).$$

Para el análisis que vamos a hacer ahora, vamos a suponer que S no verifica la primera condición del teorema, o sea, toda $\mathcal{M}$-curva algebraica se anula en a lo sumo $\delta \# S$ puntos de S. Vamos a estimar la suma de arriba, primero considerando todos los posibles $\mathbf{P}$ luego sustrayendo los $\mathbf{P}$ inadmisibles (para esta estimación vamos a usar la consideración anterior).

Sea

$$\sum_{\mathbf{P}} \kappa(\mathbf{P}),$$

la suma sobre todas las D-uplas posibles de S^D. Vamos a pensar $\mathbf{P}$ como una variable aleatoria uniformemente distribuida en todas sus $\#S^D$ D-uplas. Resulta que $\frac{1}{\#S^D} \sum_{\mathbf{P}} D - \kappa(\mathbf{P})$ es el promedio de los posibles P_i todos distintos de las D-uplas módulo p. Si definimos

$$Y_x(\mathbf{P}) := \begin{cases} 1 \text{ si alguno de los } P_i \text{ es congruente a } x(\ mod\ p) \\ 0 \text{ si no} \end{cases},$$

entonces

$$\sum_{x} Y_x(\mathbf{P}) = D - \kappa(\mathbf{P}).$$

Se tiene que

$$\mathbb{E}(\sum_x Y_x) = \sum_x \mathbb{E}(Y_x).$$

La cantidad $\mathbb{E}(Y_x)$ no es otra cosa que la probabilidad de que alguno de los P_i's sea congruente a x módulo p. Luego, la probabilidad de que ningún P_i sea congruente a x módulo p es:

$$\prod_{1\leq i\leq D} \text{Prob}(P_i \not\equiv x(p)) = \prod_{1\leq i\leq D} 1 - \text{Prob}(P_i \equiv x(p)) = \prod_{1\leq i\leq D} (1-\rho_x) = (1-\rho_x)^D.$$

Entonces

$$\sum_x \mathbb{E}(Y_x) = \sum_x 1-(1-\rho_x)^D.$$

Como por otro lado ya sabemos que $\sum_x Y_x(\mathbf{P}) = D - \kappa(\mathbf{P})$, deducimos la igualdad

$$\frac{1}{\#S^D}\sum_{\mathbf{P}} \kappa(\mathbf{P}) = D - (\sum_x 1-(1-\rho_x)^D) = \sum_x ((1-\rho_x)^D + D\rho_x - 1),$$

que nos da una estimación para $\sum_{\mathbf{P}} \kappa(\mathbf{P})$. Ahora consideremos los $\mathbf{P}$ no admisibles; nos van a interesar de estas D-uplas aquellas que cumplan $\kappa(\mathbf{P}) > 0$ (pues los otros no aportan a la suma a estimar). Para cada uno de tales $\mathbf{P}$, debe ocurrir alguna de las siguientes dos situaciones

- Existen $i \neq j$ tales que $P_i = P_j$.
- Existen $i \neq j$ tales que $P_i \neq P_j$ pero $P_i \equiv P_j \pmod p$

La cantidad de puntos $\mathbf{P}$ que cumplen la primera condición es a lo sumo $O_D(\#S^{D-1})$; en efecto, elegimos 2 de las D coordenadas que van a ser iguales y el resto distintas. Esto da un total de

$$\binom{D}{2}\#S^{D-1} = O_D(\#S^{D-1})$$

puntos. Notamos que estamos contando con repetición, lo que no nos afecta, pues sólo necesitamos una cota superior.

Contemos ahora aquellos puntos inadmisibles que cumplen la segunda condición. Dado que $\mathcal{M}$ separa puntos, tenemos que para i, j de la condición, existe $f_k \in \mathcal{M}$ tal que $f_k(P_i) \neq f_k(P_j)$. Entonces, podemos en un principio suponer $i = 1, j = 2$ y permutar los $f_k \in \mathcal{M}$ para que $f_1 = 1, f_2(P_1) \neq f_2(P_2)$. En general podemos estudiar sólo los $\mathbf{P}$ con $i = 1, j = 2$ pero hay que compensar con el factor $\frac{1}{2}D(D-1)$, constante que termina dentro del orden ya que depende de D (en este caso este factor compensa que estamos estudiando un caso y hay $\frac{1}{2}D(D-1)$ casos similares). En definitiva, sabemos que $\det(f_i(P_j))_{1\leq i,j\leq l}$ no se anula para $l = 2$. Elijamos l maximal para que el determinante anterior no se anule. Entonces P_{l+1} se encuentra en una $\mathcal{M}$-curva algebraica definida por $P_1, \cdots, P_l$. Por la hipótesis que hicimos al principio, tenemos que hay a lo sumo $\delta\#S$ posibles valores para P_{l+1}. Denotando R a la cantidad de pares $(P, Q) \in S^2$ tales que $P \equiv Q \pmod p$, concluímos que la cantidad de $\mathbf{P}$'s inadmisibles que cumplen la segunda condición es a lo sumo $O_D(\delta|S|^{D-2}R)$. Dado que

$$R = (\#S)^2 \sum_x \rho_x^2,$$

se tiene que la cantidad de $\mathbf{P}$'s inadmisibles con $\kappa(\mathbf{P}) > 0$ es a lo sumo

$$O_D((\#S)^{D-1}) + O_D(\delta(\#S)^{D-2}((\#S)^2\sum_x \rho_x^2)) = (\#S)^D O_D(\frac{1}{\#S} + \delta\sum_x \rho_x^2).$$

Usando que $\kappa(\mathbf{P}) \leq D$, tenemos

$$\frac{\nu_p(\Delta)}{(\#S)^D} \geq \sum_x ((1-\rho_x)^D + D\rho_x - 1) - O_D(\frac{1}{\#S} + \delta\sum_x \rho_x^2).$$

Para terminar, vamos a acotar inferiormente la expresión anterior. Combinando con la cota superior, vamos a concluir el teorema. Vamos a usar la hipótesis que las clases residuales de S para cada $p > c$ son a lo sumo αp. Separaremos en dos casos:

Caso 1: supongamos que $\rho_x < \frac{\delta}{D}$. Para cada x, se tiene, desarrollano el binomio y usando que $0 < \rho_x < \frac{\delta}{D}$:

$$(1-\rho_x)^D + D\rho_x - 1 \geq (\binom{w}{2} - O_D(\delta))\rho_x^2.$$

Como siempre, la desigualdad de Cauchy-Schwartz da la desigualdad

$$\sum_x \rho_x^2 \geq \frac{1}{\alpha p}(\sum_x \rho_x)^2 = \frac{1}{\alpha p},$$

de la que concluímos

$$\frac{\nu_p(\Delta)}{(\#S)^D} \geq \left(\binom{w}{2} - O_D(\delta)\right)\frac{1}{\alpha p} + O_D(\#S^{-1}).$$

Caso 2: supongamos que existe x tal que $\rho_x \geq \dfrac{\delta}{D}$. Dado que para $0 \leq z < 1$

$$\frac{\partial}{\partial z}((1-z)^D + Dz - 1) = D(1-(1-z)^{D-1}) \geq D(1-(1-z)) = Dz,$$

tenemos, integrando de 0 a ρ_x:

$$(1-\rho_x)^D + D\rho_x - 1 \geq \frac{1}{2}\rho_x^2 \geq \frac{1}{2}D(\frac{\delta}{D})^2 = \frac{\delta^2}{2D}.$$

Dado que para $x' \neq x$ tenemos $(1-\rho_{x'})^D + D\rho_{x'} - 1 \geq 0$ y como $\sum_x \rho_x^2 \leq 1$, se tiene

$$\frac{\nu_p(\Delta)}{(\#S)^D} \geq \frac{\delta^2}{2D} + O_D\left(\delta + \frac{1}{\#S}\right).$$

Para p suficientemente grande, digamos $p > c = c_{D,\delta}$, esta cota implica la del primer caso, con lo que vamos a usar la primera cota. Ahora, promediando sobre los primos $c < p \leq Q$ con peso logarítmico y usando las estimaciones (4.1) y (4.2), tenemos

$$\frac{\log(\nu_p(\Delta))}{(\#S)^D} \geq \left(\frac{1}{2\alpha}D(D-1) + O_D(\delta)\right)\sum_{c<p\leq Q}\frac{\log(p)}{p} + O_D\left(\frac{1}{\#S}\right)\sum_{c<p\leq Q}\log(p) \geq$$
$$\left(\frac{1}{2\alpha}D(D-1) + O_D(\delta)\right)(\log(Q) - \log(c) + O(1)) + O_D\left(\frac{Q}{\#S}\right).$$

Finalmente, usando la cota superior que teníamos para $\dfrac{\log(\nu_p(\Delta))}{(\#S)^D}$ y poniendo $Q = \#S$ concluímos

$$\left(\frac{1}{2\alpha}D(D-1) + O_D(\delta)\right)(\log(\#S) - \log(c) + O(1)) \leq d_{\mathcal{M}}\log(N) + O_{\mathcal{M}}(1).$$

Tomando logaritmos y despejando, conseguimos la cota

$$\#S \ll_{\alpha,\delta,\mathcal{M}} N^{\frac{2\alpha d_{\mathcal{M}}}{D(D+1)} + O_{\alpha,\mathcal{M}}(\delta)},$$

como queríamos ver. □

Observación 3.5. *Todas las cotas del Teorema 3.13 y el Teorema 3.12 son efectivas.*

Como resaltamos al principio del capítulo, el Teorema 3.12 puede ser usado para volver a obtener la cota del Teorema 2.2. Esto puede consultarse en [**18**].

5. Generalizaciones a dimensiones superiores

Con el Teorema 3.12 pudimos probar el Teorema 3.13, que caracterizaba los conjuntos mal distribuidos del plano, y adicionalmente, comentamos que se puede demostrar el Teorema 2.2 como corolario del Teorema 3.12. Cabe preguntarse si es posible usar las mismas técnicas empleadas en las secciones anteriores, para obtener algún resultado para $S \subseteq \mathbb{Z}^3$, un conjunto mal distribuido, por ejemplo, caracterizando los conjuntos mal distribuidos como aquellos conjuntos que son pequeños o están eficientemente contenidos en hipersuperficies algebraicas (tienen algún tipo de estructura algebraica). Es posible obtener un resultado del tipo mencionado, generalizando el Teorema 3.12 a dimensiones superiores, por medio del método del determinante (usando las ideas de la sección 4 del capítulo 1), para obtener:

Teorema 3.14. *Sea $d \geq 2$. Sea $S \subseteq [N]^d$ tal que S ocupa a lo sumo αp claes residuales módulo p para todo p primo, es decir, si $S_p = \{(x_1 \pmod p), \cdots, x_d \pmod p)) \in (\mathbb{Z}/p\mathbb{Z})^d : (x_1, \cdots, x_d) \in S\}$, entonces $\#S_p \leq \alpha p$ para cada primo p. Entonces para todo $\varepsilon > 0$ se tiene alguna de las siguientes condiciones:*

- $\#S \ll_{\epsilon,d,\alpha} N^{\epsilon}$
- *existe una hipersuperficie algebraica H definida por $f(x_1, \cdots, x_d) = 0, f \in \mathbb{Z}[x_1, \cdots, x_d]$ de grado a lo sumo $O_{\varepsilon,d,\alpha}(1)$ tal que al menos $(1-\varepsilon)\#S$ puntos de S están contenidos en H.*

Notemos que el Teorema 3.14 no es una generalización del todo satisfactoria, pues la condición de ocupar a lo sumpo αp clases residuales es muy restrictiva. En el caso de dimensiones superiores, la estimación de Lang-Weil que obtenemos del Teorema 3.3 para una hipersuperficie algebraica $V \subseteq \overline{\mathbb{Q}}^n$, definida por un polinonio geométricamente irreducible $P(x_1, \cdots, x_n) \in \mathbb{Z}[x_1, \cdots, x_n]$ de grado d, nos dice que $\#V(\mathbb{F}_p) = p^{n-1} + O_d(p^{n-\frac{1}{2}})$, luego una hipersuperficie algebraica en un principio podría ocupar $O(p^{n-1})$ clases residuales, para todo primo p. Considerando $P(x_1, \cdots, x_n) = x_1 - x_2^2$, la hipersuperficie algebraica que define ocupa $O(p^{n-1})$ clases residuales módulo p pero no $O(p^{n-2})$ clases residuales módulo p para todo p primo.

Podemos entonces preguntarnos:

Pregunta. *Sea $S \subseteq [N]^d$ con $d \geq 2$, tal que S ocupa αp^i clases residuales para todo p primo, con $i < d$. ¿Qué podemos decir de S?*

Walsh estudia esta pregunta en [**59**]:

Teorema 3.15. *Sea $0 \leq k < d$ un entero y sean $\varepsilon, \alpha, \eta > 0$ números reales. Entonces existe una constante $C = O_{\varepsilon,\alpha,\nu,k}(1)$ tal que para todo $S \subseteq [N]^d$ que verifica $\#S_p \leq \alpha p^k$, para cada primo p, donde*

$$S_p := \{(x_1 \pmod p), \cdots, x_d \pmod p)) \in (\mathbb{Z}/p\mathbb{Z})^d : (x_1, \cdots, x_d) \in S\},$$

se tiene al menos alguna de las siguientes situaciones:

- $\#S \ll_{d,k,\varepsilon,\alpha} N^{k-1+\varepsilon}$.
- *Existe una hipersuperficie algebraica H definida por los ceros de $f \in \mathbb{Z}[x_1, \cdots, x_d]$ con grado a lo sumo C y los coeficientes de f acotados por N^C, que se anula en al menos $(1-\eta)\#S$ puntos de H.*

Notamos que el Teorema 3.15 es una generalización adecuada del Teorema 3.13, pues no sólo considera la situación que un conjunto S ocupe $O(p^i)$ clases residuales módulo p, con $i < d$, si no que para cada i, da información del tipo de hipersuperficie algebraica que anula al conjunto S. No vamos a explicar la demostración de Walsh del Teorema 3.15; el lector interesado puede consultar [**59**]. Sin embargo, sí vamos a explicar el ejemplo de [**59**], que muestra que en el Teorema 3.13 y en su generalización, el Teorema 3.15, no es posible tomar $\varepsilon = 0$, es decir, existen conjuntos mal distribuidos, pequeños, que no poseen una estructura algebraica natural que se les pueda atribuir.

Ejemplo 3.12 (Walsh). *Sea $0 < \eta < 1$. Supongamos que N es suficientemente grande tal que exista Q entero con $Q < \log(N) < 2Q$ tal que $R = N^{\frac{1}{4}} < \prod_{p \leq Q} p < N$. Para cada primo $p \leq Q$ elegimos $\lfloor p^{\eta} \rfloor$ clases residuales. Por el teorema chino del resto*
indexteorema!chino del resto, hay $\sim R^{\eta}$ elementos menores a R que se reducen módulo p a las clases elegidas, para $p \leq Q$. Elejimos $\left\lfloor \frac{(\log(N))^{\eta}}{2} \right\rfloor$ de estos elementos y llamamos a este conjunto X. Dado que para cada primo $p > Q$ se tiene

$$\#X \leq \frac{1}{2}(\log(N))^{\eta} < \frac{1}{2}2^{\eta}Q^{\eta} < p^{\eta}.$$

Pero por construcción, esto mismo ocurre para $p \leq Q$. Concluímos que X ocupa a lo sumo p^{η} restos módulo p para cada primo p. Notamos además que $\#X = O((\log N)^{\eta})$.

Sean d, h con $d \geq h+1$ y consideremos $h+1$ conjuntos distintos $X_1, \cdots, X_{h+1}$ construídos como el ejemplo anterior, pero con $\eta = \frac{1}{h+1}$. Definimos

$$S := \{(x_1, \cdots, x_d) \in [N]^d : x_i \in X, 1 \leq i \leq h+1\}.$$

Como hay $d-h-1$ coordenadas libres en S y las otras toman, por construcción, a lo sumo p^η restos módulo p, y son $h+1$, de la elección de η deducimos que $\#S_p \leq p^{d-h}$. Por otro lado, como $\#X = O((\log(N))^\eta)$, razonando de manera análoga tenemos que

$$\#S \gg N^{d-h-1}\log(N).$$

Esto permite exhibir un ejemplo de un conjunto de $\mathbb{Z}^n$ mal distribuido pero que posee un tamaño muy pequeño y ninguna estructura algebraica aparente[6].

Al igual que con el Teorema 3.13, es posible que los argumentos empleados en la demostración del Teorema 3.15 puedan dar una demostración del Teorema 2.2, para dimensiones superiores.

En la Pregunta 3.2, nos preguntamos cómo era un conjunto de enteros mal distribuido. Esta pregunta, como ya mencionamos, es bastante complicada. Esto se debe a que no existe una "dicotomía" clara para tales conjuntos. En efecto, sea $S \subseteq \mathbb{Z}$ un conjunto mal distribuido. Si aceptamos por noción de estructurado, que exista un polinomio $p(x) \in \mathbb{Z}[x]$ que anule eficientemente a S, con $\deg(p)$ pequeño, no tardamos en notar que, si p anula al menos $\delta\#S$ puntos de S, entonces $\deg(p) \geq \delta\#S$, con lo que el grado de p no puede ser pequeño. El problema reside en que el conjunto de ceros de un polinomio $p(x) \in \mathbb{Z}[x]$ no nulo, es finito, con lo que no permite dar una noción adecuada de estructura algebraica. Helfgott y Venkatesh [**18**] conjeturan:

Conjetura 3.1. *Sea $S \subseteq [N]$, con $N \geq 1$. Supongamos que existe $0 < \alpha < 1$ tal que el número de residuos $\{x \pmod p : x \in S\}$ es a lo sumo αp, para todo primo p. Sea $\varepsilon > 0$. Si $\#S \geq N^\varepsilon$, entonces casi todo punto de S está contenido en la imagen de un polinomio con coeficientes enteros, con grado acotado por α y ε.*

Podemos enunciar de manera precisa la conjetura anterior en un caso particular ([**59**]). Sea $S \subseteq [N]$ tal que $\#S \geq N^{0.49}$, y tal que ocupa a lo sumo $\frac{2p}{3}$ clases residuales, módulo p, para todo primo p. Entonces casi todo S debería estar contenido en un conjunto de la forma $\{an^2 + bn + c : n \in \mathbb{Z}\}$.

La Conjetura 3.1 parece complicada. Por ejemplo, Helfgott y Venkatesh [**18**] señalan que la conjetura 3.1 implica otra conjetura, que esencialmente dice que la cantidad de puntos enteros de una curva irracional Γ, es decir, una curva tal que $\Gamma(\mathbb{Z}) \cap C \cap [N]$ es "pequeño' para toda curva algebraica C racional, verifica la estimación $O_\varepsilon(N^\varepsilon)$. Esta conjetura parece difícil, con lo que es de esperar que la Conjetura 3.1 sea difícil.

[6]Como ya mencionamos, siempre es posible darle una estructura algebraica a un conjunto finito S, pues siempre existe un polinomio que anule a S. Sin embargo, nos importa que el grado del polinomio sea pequeño; el tamaño del grado del polinomio es, en algún sentido, un índice del tipo de estructura algebraica del conjunto S (ver [**60**]). El conjunto del Ejemplo 3.12 fue construído de manera "aleatoria", y no es tan pequeño como un punto; esa aleatoriedad es a la que nos referimos con que no posee estructura algebraica aparente. ¡Pero recordemos que esta "aleatoriedad" no coincide con la noción de conjunto aleatorio, pues el conjunto del Ejemplo 3.12 ocupa pocas clases residuales módulo p para todo primo p!

CAPÍTULO 4

Puntos algebraicos y especiales de una curva

En el capítulo 2 de la tesis estudiamos los puntos enteros de curvas trascendentes y probamos, mediante el método del determinante de Bombieri-Pila, el resultado:

TEOREMA (Bombieri-Pila). *Sea $f : I \to \mathbb{R}$ una función analítica trascendente sobre I, un intervalo cerrado y acotado. Sea Γ el gráfico de f. Para todo $t \geq 1$ y $\varepsilon > 0$, se tiene la estimación:*

$$\#((t\Gamma)(\mathbb{Z})) = O_{f,\epsilon}(t^{\varepsilon}).$$

Del teorema anterior, tomando $t = N$, obtenemos la misma estimación para los puntos racionales de Γ con denominador N. En general, es de interés obtener una estimación como la del teorema anterior pero para los puntos racionales de altura acotada por H. Recordamos que la altura[1] de un punto $P = (\frac{a}{b}, \frac{c}{d})$ con $\gcd(a, b) = \gcd(c, d) = 1$, se define como $H(P) := \max\{|a|, |b|, |c|, |d|\}$. Pila en [**41**] obtiene una estimación para los puntos racionales de altura acotada, de una curva trascendente analítica, por medio de una modificación del método del determinante de Bombieri-Pila:

TEOREMA (Pila). *Sea $f : I \to \mathbb{R}$ una función analítica trascendente sobre I, un intervalo cerrado y acotado. Sea Γ el gráfico de f. Para todo $N \geq 1$ entero positivo y $\varepsilon > 0$, se tiene la estimación:*

$$\#\Gamma(\mathbb{Q}, N) = O_{f,\epsilon}(N^{\varepsilon}).$$

El estudio de los puntos racionales de una curva trascendente es importante, y está relacionado con varios problemas de la teoría de trascendencia y la geometría diofántica. Por ejemplo, para el caso que Γ sea el gráfico de la función exponencial $f(x) = \exp(x)$, estudiar los puntos racionales, o en general, los puntos algebraicos de Γ da información acerca de algunos resultados de trascendencia clásicos como el teorema de las seis exponenciales o el teorema de Gelfond-Schneider.

TEOREMA (Teorema de las seis exponenciales). *Sean $x_1, x_2, x_3 \in \mathbb{R}$, linealmente independientes sobre $\mathbb{Q}$, y supongamos que $\xi_1, \xi_2 \in \mathbb{R}$ son linealmente independientes sobre $\mathbb{Q}$. Entonces al menos uno de los seis números*

$$\exp(x_i \xi_j)$$

es trascendente

TEOREMA (Gelfond-Schneider). *Sea α un número algebraico real positivo, $\alpha \neq 1$, y supongamos que β es un número algebraico real irracional. Entonces α^{β} es un número trascendente.*

Muchas condiciones algebraicas que aparecen en la geometría diofántica se pueden interpretar en términos de ecuaciones analíticas trascendentes. Por ejemplo, dada una hipersuperficie algebraica S definida por los ceros de $P(x_1, \cdots, x_n) = 0$ con $P(x_1, \cdots, x_n) \in \mathbb{C}[x_1, \cdots, x_n]$ un polinomio no nulo, resultan de interés los puntos de H cuyas coordenadas son raíces de la unidad; denominemos a estos puntos, **punto especial**. Dado que un número complejo z no nulo es una raíz de la unidad si y sólo si existe un racional q tal que $z = \exp(2\pi i q)$, los puntos de H que tienen coordenadas raíces de la unidad son las soluciones racionales de la ecuación

$$G(z_1, \cdots, z_n) := P(\exp(2\pi i z_1, \cdots, \exp(2\pi i z_n))) = 0.$$

Esta problemática, que resulta un caso particular de la conjetura de Manin-Mumford (ya probada), fue estudiada por Mann [**29**]:

TEOREMA (Mann). *Sea $n \geq 1$ un entero positivo. Sea $G(x_1, \cdots, x_l) \in \mathbb{C}[x_1, \cdots, x_l]$ un polinomio no nulo. Entonces el conjunto de puntos especiales de la hipersuperficie algebraica $\{(x_1, \cdots, x_l) \in \mathbb{C}^l : G(x_1, \cdots, x_l) = 0\}$ es una unión finita de coclases de subgrupos de $\mathbb{G}^l$.*

[1] Esta no es la altura proyectiva usual.

Pila investiga en [**39**] cómo generalizar la estimación de los puntos enteros de las dilataciones de una curva trascendente analítica a dimensiones superiores. Sin embargo, al generalizar a dimensiones superiores enseguida se encuentran complicaciones. Por ejemplo, consideremos $f : [0,1]^2 \to \mathbb{R}$ una función analítica trascendente. Sea $\Gamma = \{(x,y,z) \in \mathbb{R}^3 : f(x,y) = z, (x,y) \in [0,1]^2\}$. Siguiendo los pasos del teorema 2.1, podríamos fijar la variable y y estudiar, para cada y, el conjunto $\Gamma_y := \{(x,z) \in \mathbb{R}^2 : f(x,y) = z, x \in [0,1]\}$. El problema es que, fijada una variable, la función $f(x,y)$ puede no ser trascendente. Por ejemplo, la función $f(x,y) = x^y$ en el intervalo $[1,2]^2$ verifica, para cada $y \in [1,2] \cap \mathbb{Q}$, que $f(x,y) = x^y$ es una función algebraica. En efecto, si $y = \frac{m}{n}$, con $\gcd(m,n) = 1$ y m,n positivos, tenemos que el polinomio $P(a,b) = a^{mn} - b^n$ verifica $P(a, f(a,y)) = 0$. Para evitar esta dificultad, Pila considera la familia de conjuntos subanalíticos, que poseen un comportamiento geométrico sencillo. Sea $X \subseteq \mathbb{R}^n$ una superficie compacta subanalítica. Por medio del método del determinante de Bombieri-Pila, obtiene una cota superior para la cantidad de hipersuperficies algebraicas de grado acotado que cubren a $tX(\mathbb{Z})$, la dilatación por t de los puntos enteros de X. Dado que $X(\mathbb{Z})$ puede contener puntos que verifiquen condiciones polinomiales, como era el caso de la imagen de $f(x,y) = x^y$ (si bien el gráfico de f no es un conjunto subanalítico), es necesario considerar los puntos que no cumplen "condiciones algebraicas". Extrayendo los **puntos algebraicos** de X, que denotamos X^{alg}, obtenemos la **parte trascendente** de X. Luego, Pila da una cota superior de los puntos de intersección entre $t(X \backslash X^{\mathrm{alg}})(\mathbb{Z})$ y una hipersuperficie algebraica de grado acotado, apelando a la geometría sencilla de los conjuntos subanalíticos. El resultado que obtiene es:

Teorema (Pila). *Sea $X \subseteq \mathbb{R}^n$ una superficie compacta subanalítica. Sea $\varepsilon > 0$. Entonces, para todo $t \geq 1$ se tiene*

$$\#t(X \backslash X^{\mathrm{alg}})(\mathbb{Z}) = O_{\Omega,\varepsilon}(t^{\varepsilon}).$$

El hecho clave en la demostración del teorema anterior es la geometría sencilla de los conjuntos subanalíticos. Abstraído, este hecho puede explicarse observando que los conjuntos subanalíticos definen una estructura O-minimal sobre $\mathbb{R}$. Pila y Wilkie [**42**] dan una estimación para los puntos racionales de altura acotada de un conjunto definible en una estructura O-minimal sobre $\mathbb{R}$ que extiende la suma y el producto de $\mathbb{R}$.

Teorema (Pila-Wilkie). *Sea $X \subseteq \mathbb{R}^n$ un conjunto definible en una estructura O-minimal sobre $\mathbb{R}$ que extiende la suma y el producto de $\mathbb{R}$, y sea $\varepsilon > 0$. Entonces para todo $H > 0$, se tiene la estimación*

$$\#(X - X^{\mathrm{alg}})(\mathbb{Q}, H) = O_{X,\varepsilon}(H^{\varepsilon}).$$

Este teorema generaliza el problema de los puntos racionales de altura acotada de una curva analítica trascendente, debido a que la estructura O-minimal $\mathbb{R}_{\mathrm{an}}$ contiene como conjuntos definibles a los gráficos de funciones analíticas $f : [-1,1]^n \to \mathbb{R}$. Las técnicas utilizadas para probar el resultado de Pila-Wilkie siguen una estrategia similar al Teorema 2.1: primero cubren cierto conjunto por hipersuperficies algebraicas de grado acotado, mediante el método de Bombieri-Pila, y luego estiman los puntos de intersección de este conjunto con hipersuperficies algebraicas. Es importante notar que, la estrategia de la demostración de [**42**] del resultado de Pila-Wilkie, en algunos casos se puede hacer efectiva (calculando las constantes de manera explícita). Por ejemplo, es posible obtener el teorema de las seis exponenciales y el teorema de Gelfond-Schneider siguiendo los pasos de [**42**] con el gráfico de la función exponencial.

Para entender los argumentos empleados en el teorema de Pila-Wilkie, introducimos en la sección 1 el concepto de estructuras O-minimales y en la sección 2 desarrollamos una de las propiedades fundamentales de las estructuras O-minimales, que es la descomposición en celdas de los conjuntos definibles. Seguimos esencialmente el libro [**54**]. En las secciones 3 y 4 explicamos, no con la mayor generalidad posible, la demostración de [**42**] del teorema de Pila-Wilkie.

El resultado de Pila-Wilkie es utilizado por Pila y Zannier en [**43**] para dar una nueva demostración de la conjetura de Manin-Mumford. Posteriormente, el argumento de [**43**] recibe numerosas aplicaciones, entre ellas, una demostración incondicional (sin usar la hipótesis generalizada de Riemann) de la conjetura de André-Oort para producto arbitrario de curvas modulares [**40**]. Dado que estas aplicaciones son muy importantes, y están relacionadas con el tema de esta tesis, en la sección 5 mostramos cómo funciona el argumento de [**43**], dando una demostración del teorema de Mann, siguiendo el survey de [**46**].

1. Estructuras O-minimales

En general, no es posible "parametrizar", es decir, describir, cualquier conjunto del plano, de manera sencilla, como el caso de las curvas algebraicas, que se particionan en finitos conjuntos sencillos (gráficos de funciones), en los que se tienen parametrizaciones con pendiente acotada. . Por descripción sencilla, nos referimos a poder dar una "descomposición" finita en conjuntos sencillos. Ejemplos "sencillos" como los conjuntos G_δ o F_σ en general no admiten descripciones sencillas.

Resulta entonces natural preguntarnos por familias de conjuntos que poseen descripciones sencillas. Formalmente, para cada n, sea $\mathcal{F} = \{f_i : X_i \subseteq \mathbb{R}^n \to \mathbb{R},\ X_i \neq \emptyset, n \in \mathbb{N}\}$ una familia de funciones. Definimos los conjuntos

$$B_{f_1,\cdots,f_k,g_1,\cdots,g_l} := \{f_1(x) = \cdots = f_k(x) = 0, g_1(x) > 0, \cdots, g_l(x) > 0\},$$

con $x = (x_1, \cdots, x_n) \in \mathbb{R}^n$, y $f_i \in \mathcal{F}, g_j \in \mathcal{F}$, para todo i, j. Consideramos ahora los conjuntos que se obtienen mediante finitas operaciones conjuntistas (uniones, intersecciones, complementos, productos cartesianos, proyecciones[2]) y finitas operaciones topológicas (tomar interior, clausura) de los conjuntos $B_{f_1,\cdots,f_k,g_1,\cdots,g_l}$. Estos conjuntos definen una familia $\mathcal{A}$, que contiene conjuntos de $\mathbb{R}^n$, para todo n. Para nosotros, $\mathcal{A}$ es una familia de conjuntos que admite una *descripción sencilla* si, luego de realizar finitas operaciones como las indicadas recién, con los conjuntos $B_{f_1,\cdots,f_k,g_1,\cdots,g_l}$, llegamos a cierta "estabilización": dejamos de obtener conjuntos nuevos. Resulta natural entonces preguntarnos:

PREGUNTA. *¿Para qué familias $\mathcal{F}$ se tiene que $\mathcal{A}$, la familia obtenida por el proceso del párrafo anterior, admite una descripción sencilla?*

La pregunta anterior en general es complicada, y no vamos a estudiarla en esta tesis. Otra pregunta natural es:

PREGUNTA. *¿Qué propiedades posee una familia $\mathcal{A}$, que admite una descripción sencilla?*

El hecho clave es que si una familia $\mathcal{A}$ admite una descripción sencilla, entonces sus conjuntos poseen un comportamiento geométrico sencillo. Este comportamiento sencillo es el que les permite a Pila y Wilkie generalizar el Teorema 2.1. Por esta razón, en lo que sigue, vamos a estudiar la noción de "descripción sencilla", introduciendo el concepto de estructuras O-minimales y, posteriormente, explicando por qué los conjuntos de una estructura O-minimal tienen geometría sencilla.

Nos interesa que una familia de conjuntos $\mathcal{A}$ posea cierta estabilidad en términos de operaciones conjuntistas (uniones, interseccion, complementos). Esto motiva la definición de estructura.

DEFINICIÓN 4.1. *Sea R un conjunto no vacío. Decimos que $\delta = (\delta_m)_{m\in\mathbb{N}}$ con $\delta_m \subseteq \mathcal{P}(R)$ es una estructura[3] sobre R, si para cada m se verifican:*

1. *δ_m es un álgebra booleana de subconjuntos de $\mathbb{R}^m$*
2. *si $A \in \delta_m$, entonces $R \times A$ y $A \times R$ son elementos de δ_{m+1}*
3. *$\{(x_1, \cdots, x_m) \in R^m : x_1 = x_m\} \in \delta_m$*
4. *si $A \in \delta_{m+1}$ entonces $\pi(A) \in \delta_m$, donde $\pi : R^{m+1} \to R^m$ es la proyección en las primeras m coordenadas (estabilidad por proyecciones).*

La condición (1) nos dice que las familias de conjuntos δ_n poseen un mínimo de "estabilidad", debido a que son álgebras booleanas. Las condiciones (2) y (4) nos dicen que hay cierto tipo de "compatibilidad" entre los conjuntos de la estructura $(\delta_n)_n$. La condición (3) dice que si vemos a R^n como el conjunto $\{(x_1, \cdots, x_n, x_{n+1}) \in R^n : x_1 = x_{n+1}\}$, entonces $R^n \in \delta_{n+1}$.

Veamos algunos ejemplos de estructuras.

EJEMPLO 4.1. *Los subconjuntos semialgebraicos de $\mathbb{R}^m$ son las uniones finitas de los conjuntos de la forma*

$$\{\boldsymbol{x} = (x_1, \cdots, x_m) \in \mathbb{R}^m : f_1(\boldsymbol{x}) = \cdots = f_k(\boldsymbol{x}) = 0, g(\boldsymbol{x}) \neq 0\}, \text{ donde } f_1(\boldsymbol{x}), \cdots, f_k(\boldsymbol{x}), g(\boldsymbol{x}) \in \mathbb{R}[x_1, \cdots, x_m].$$

Los subconjuntos semialgebraicos determinan una estructura sobre $\mathbb{R}$. Notemos que, de las condiciones de la Definición 4.1, sólo la (4) no es trivial. Tarski en [**53**] *probó que los subconjuntos semialgebraicos verifican la condición (4), resultado conocido como principio de Tarski-Seidenberg.*

EJEMPLO 4.2. *Sea Ω un cuerpo algebraicamente cerrado y $K \subseteq \Omega$ un subcuerpo. Los subconjuntos K-construíbles de Ω^m son las uniones finitas de los conjuntos de la forma*

$$\{\boldsymbol{x} \in \Omega^m : f_1(\boldsymbol{x}) = \cdots = f_k(\boldsymbol{x}) = 0, g(\boldsymbol{x}) \neq 0\}, \text{ donde } f_1(x), \cdots, f_k(x), g(x) \in K[x] \text{ y } \boldsymbol{x} = (x_1, \cdots, x_m).$$

Los subconjuntos K-constructibles determinan una estructura sobre Ω. Notemos que, de las condiciones de la Definición 4.1, sólo la condición (4) no es trivial. Chevalley probó que los subconjuntos K-constructibles verifican la condición (4), primero demostrando que los cuerpos algebraicamente cerrados verifican una propiedad conocida como "eliminación de cuantificadores" (ver [**31**]*[capítulo 3, sección 3.2, pág. 84]).*

A continuación, damos propiedades elementales que verifica una estructura sobre un conjunto R no vacío.

PROPOSICIÓN 4.1. *Sea R un conjunto no vacío y δ una estructura sobre R.*

1. *Sea $A \in \delta_m, B \in \delta_n$. Entonces $A \times B \in \delta_{m+n}$.*
2. *Para $1 \leq i < j \leq m$ la diagonal $\Delta_{i,j} := \{(x_1, \cdots, x_m) \in R^m : x_i = x_j\}$ es un elemento de δ (estabilidad frente a permutación de variables).*

[2]Por proyecciones nos referimos a considerar $\pi(B)$, con $\pi : \mathbb{R}^{n+m} \to \mathbb{R}^n$ la proyección en las primeras n coordenadas.

[3]También es usual decir que (R, δ) es una estructura.

(3) *Sea $B \in \delta_n$ y sean $i(1), \cdots, i(n) \in \{1, \cdots, m\}$. Entonces el conjunto $A \subseteq R^m$ definido como $A := \{(x_1, \cdots, x_m) \in R^m : (x_{i(1)}, \cdots, x_{i(n)}) \in B\}$ es un elemento de δ (estabilidad frente a permutación e identificación de variables).*

Demostración.

(1) Sean $A \in \delta_m, B \in \delta_n$. Como $A \times R$ y $R \times B$ son elementos de δ_{m+1}, inductivamente tenemos que $A \times R^n$ y $R^m \times B$ son elementos de δ_{m+n}. Dado que $A \times B = (A \times R^n) \cap (R^m \times A)$, al ser δ_{m+n} cerrado por intersecciones, concluímos que $A \times B \in \delta_{m+n}$.
(2) Tenemos $\Delta_{ij} = R^{i-1} \times \Delta_{1j} \times R^{m-j+1}$ y como $\delta_{1j} \in \delta_{j-i}$, tenemos que $\Delta_{ij} \in \delta_m$.
(3) Para probar este punto, consideramos un caso particular; el caso general es análogo y queda a cargo del lector. Sea $n = 2, m = 3$. Luego, resulta:

$$x = (x_1, x_2, x_3) \in A \iff \exists y_1 \exists y_2 (x_{i(1)} = y_1 \wedge x_{i(2)} = y_2 \wedge y = (y_1, y_2) \in B).$$

Podemos entonces pensar en $(x, y) = (x_1, x_2, x_3, y_1, y_2) \in R^3 \times R^2$. La condición $y \in B$ se traduce $(x, y) \in R^3 \times B$. Debido a que permutar las coordenadas de un conjunto en la estructura δ sigue definiendo un conjunto en la estructura δ, las condiciones $x_{i(1)} = y_1 \wedge x_{i(2)} = y_2$ son compatibles con $(x, y) \in R^3 \times B$ y nos dicen que definen un conjunto en δ. Por ejemplo, $i(1) = 1, i(2) = 3$ el conjunto definido es $\{(x_1, x_2, x_3, y_1, y_2) \in R^3 \times B : x_1 = x_1, x_3 = y_2\}$. Basta ahora interpretar los cuantificadores de existencia. Pero estos pueden interpretarse como proyecciones. Concluímos que A se obtiene de proyectar sucesivamente las últimas coordenadas, o sea, $(x_1, x_2, x_3, y_1, y_2) \mapsto (x_2, x_3, y_1, y_2) \mapsto (x_3, y_1, y_2) \mapsto (y_1, y_2) = (x_1, x_3) \in B$. Al ser estable una estructura por proyecciones, concluímos $A \in \delta_m$.

□

La demostración de la Proposición 4.1 muestra que se puede probar que un subconjunto de R^n está en δ_n describiéndolo con "notación lógica" y luego reinterpretando como operaciones booleanas, complementos, productos cartesianos y proyecciones. Tenemos la siguiente "traducción" de lenguaje lógico a lenguaje conjuntista. Sean $\Phi(x, y, z), \Psi(x, y, z)$ "fórmulas lógicas", con $x \in X, y \in Y, z \in Z$. Sean $A := \{(x, y, z) \in X \times Y \times Z : \Phi(x, y, z)\}, B := \{(x, y, z) \in X \times Y \times Z : \Psi(x, y, z)\}$.

- La conjunción $\Phi(x, y, z) \vee \Psi(x, y, z)$ define la unión, $A \cup B$.
- La disyunción $\Phi(x, y, z) \wedge \Psi(x, y, z)$ define la intersección $A \cap B$.
- La negación $\neg\Phi(x, y, z)$ define el complemento $X \times Y \times Z \backslash A$.
- La implicación $\Phi(x, y, z) \implies \Psi(x, y, z)$, que se define como $(\neg\Phi(x, y, z)) \vee \Psi(x, y, z)$, luego define al conjunto $(X \times Y \times Z \backslash A) \cup B$.
- La doble implicación $\Phi(x, y, z) \iff \Psi(x, y, z)$, que se define como $(\Phi(x, y, z) \implies \Psi(x, y, z)) \wedge (\Psi(x, y, z) \implies \Phi(x, y, z))$ define al conjunto $[(X \times Y \times Z \backslash A) \cup B] \cap [X \times Y \times Z \backslash B] \cap A$.
- La cuantificación existencial $\exists x \Phi(x, y, z)$ define la proyección $\pi_X(A) := \{(y, z) \in X \times Z : \Phi(x, y, z)\}$. De la misma manera se definen las cuantificaciones existenciales en y, z.
- La cuantificación universal $\forall x \Phi(x, y, z)$ equivale a $\neg(\exists x \neg\Phi(x, y, z))$, luego define al conjunto $Y \times Z \backslash \Pi_X(X \times Y \times Z \backslash A)$. De la misma manera se definen las cuantificaciones universales en y, z.

La razón por la cual la "notación lógica" nos resulta útil para probar que un subconjunto de R^n está en δ_n es la siguiente afirmación:

Afirmación 1. *Un conjunto es definible en una estructura δ si se puede definir mediante una fórmula de un lenguaje de primer orden, es decir, $X \in R^n$ es definible si podemos describirlo mediante conectores lógicos*[4] $\vee, \wedge, \neg, \implies, \iff$, *cuantificadores lógicos*[5] $\forall, \exists$ *y ciertas constantes, funciones y relaciones propias de la estructura.*

Por ejemplo, los subconjuntos semialgebraicos de $\mathbb{R}$ son aquellos conjuntos definidos en un lenguaje de primer orden usando parámetros en $\mathbb{R}$, la relación de orden $<$, el símbolo $=$, las funciones $+, \cdot$ que denotan la suma y productos entre reales usuales, y las constantes 0 y 1 (es decir, los subconjuntos semialgebraicos son aquellos conjuntos definidos en un lenguaje de primer orden sobre $\mathbb{R}$ como cuerpo ordenado).

La Afirmación 1 se prueba mediante un argumento inductivo típico de la lógica proposicional (ver [**13**]); dado que para nuestros fines, nos interesa tener la Afirmación 1 como guía para entender algunas ideas, no vamos a dar una prueba de dicha afirmación.

El párrafo anterior motiva la siguiente definición.

[4]En verdad, alcanza con los conectores $\wedge, \neg$

[5]Dado que $\forall = \neg\exists$, alcanza con el cuantificador $\exists$.

DEFINICIÓN 4.2. *Sea R un conjunto no vacío, y $\delta = (\delta_n)_n$ una estructura sobre R. Decimos que $X \subseteq R^m$ es definible en δ, o simplemente que es definible si se entiende en el contexto la estructura δ considerada, si $X \subseteq \delta_n$.*

Ahora consideramos funciones que tienen algún tipo de compatibilidad con una estructura δ.

DEFINICIÓN 4.3. *Sea R un conjunto no vacío, y δ una estructura sobre R. Sea $X \in R^m$. Decimos que la función $f : X \to R^n$ es definible en δ, o simplemente que es definible si se entiende en el contexto la estructura δ considerada, si $\Gamma(f)$, el gráfico de f, es definible en δ, es decir, $\Gamma(f) \in \delta_{m+n}$.*

Las funciones definibles en una estructura verifican las siguientes propiedades:

PROPOSICIÓN 4.2. *Sea R un conjunto no vacío y δ una estructura sobre R. Sea $S \subseteq R^m$ y $f : S \to R^n$ una función definible en δ. Entonces:*

1. *$S \in \delta_m$.*
2. *Si $A \subseteq S$, $A \in \delta_m$, entonces $f(A) \in \delta_n$ y la restricción $f|_A$ es definible en δ.*
3. *Si $B \in \delta_n$, entonces $f^{-1}(B) \in \delta_m$.*
4. *Si f es biyectiva, su inversa f^{-1} es definible en δ.*
5. *Si $f(S) \subseteq T \subseteq R^n$, y $g : T \to R^p$ es una función definible en δ, entonces la composición $g \circ f : S \to R^p$ es una función definible en δ.*

DEMOSTRACIÓN.

1. Se tiene que $x \in S \iff \exists y((x, y) \in \Gamma(f))$, luego equivalentemente, $S = \pi(\Gamma(f))$ donde $\pi_1 : R^{m+n} \to R^m$ es la proyección en las primeras m coordenadas. Como $\Gamma(f) \in \delta_{n+m}$, dado que δ es una estructura, concluímos $S \in \delta_m$.
2. Se tiene que $y \in f(A) \iff \exists x(x \in A \wedge (x, y) \in \Gamma(f))$, luego equivalentemente, $f(A) = \pi(\Gamma(f))$, donde $\pi_2 : R^{m+n} \to R^n$ es la proyección en las últimas n variables. No es difícil de deducir de la proposición 4.1, que si $D \in \delta_{m+n}$ entonces $\pi_2(D) \in \delta_n$. Luego, como $\Gamma(f) \in \delta_{m+n}$, concluímos $f(A) \in \delta_n$.
   Si $g = f|_A$, dado que $\Gamma(g) = \Gamma(f) \cap R^m \times f(A)$, con $\Gamma(g)$ el gráfico de g, concluímos que $f|_A$ es definible en δ.
3. Si $B \in \delta_n$ entonces $f^{-1}(B) \in \delta_m$. En este caso, $x \in f^{-1}(B) \iff \exists y(y \in B \wedge (x, y) \in \Gamma(f))$. Interpretando las operaciones lógicas involucradas, tenemos $f^{-1}(B) = \pi(\Gamma(f) \cap R^m \times B)$ con π una proyección, con lo que $f^{-1}(B) \in \delta_m$.
4. Sea $\Gamma(f^{-1})$ el gráfico de f^{-1}. Notemos que $(y, x) \in \Gamma(f^{-1}) \iff (x, y) \in \Gamma(f)$. Por el ítem (3) de la Proposición 4.1, tenemos que f^{-1} es una función definible en δ.
5. Sea $\Gamma(g)$ el gráfico de g. Notemos que $(x, z) \in \Gamma(g \circ f) \iff \exists y((x, y) \in \Gamma(f) \wedge (y, x) \in \Gamma(g))$. Por el ítem (3) de la Proposición 4.1, tenemos que $g \circ f$ es una función definible en δ.

□

Podemos entonces concluir que la clase de funciones en δ tiene "buena estabilidad": la composición se mantiene en la estructura y las funciones biyectivas tienen inversa en δ. Sin embargo, como antes, van a resultar de interés las funciones que cumplen algún otro tipo de propiedad, como ser continuas.

Notemos que, con la Definición 4.1, si δ es una estructura sobre un espacio topológico R (no vacío), no resulta claro que δ refleje propiedades geométricas dadas por la topología de R. Sin embargo, podemos pedirle a una estructura sobre R que posea conjuntos que estén relacionados con la topología de R, como hacemos en la siguiente proposición.

PROPOSICIÓN 4.3. *Consideremos R con un orden total $<$ y démosle la topología del orden. A cada R^n le damos la topología producto. Supongamos que δ es una estructura sobre R tal que $\{(x, y) \in R^2 : x < y\} \in \delta_2$. Si $A \in \delta_n$ entonces $\overline{A}, \dot{A} \in \delta_n$.*

DEMOSTRACIÓN. Vamos a probar sólo el caso de la clausura, pues el caso del interior es similar. Se tiene que x está en la clausura de A si y sólo si todo entorno que contiene a x tiene intersección no vacía con A. Esta definición es equivalente a que todo entorno de x de la forma $(a_1, b_1) \times \cdots \times (a_n, b_n)$ tiene intersección no vacía con A. Reescribimos esta condición utilizando lenguaje lógico:

$$(x_1, \cdots, x_n) \in \overline{A} \iff \forall a_1 \cdots \forall a_n \forall b_1 \cdots \forall b_n[(a_1 < x_1 < b_1 \wedge \cdots \wedge a_n < x_n < b_n) \implies \exists y_1 \cdots \exists y_m(a_1 < y_1 < b_1 \wedge \cdots \wedge a_n < y_n < b_n \wedge (y_1, \cdots, y_m) \in A)].$$

Reinterpretando los símbolos lógicos, en el sentido de la Afirmación 1, concluímos la demostración. □

Asumiendo la condición topológica de la proposición, también tenemos que el límite de funciones es definible.

PROPOSICIÓN 4.4. *Consideremos R con un orden total $<$ y démosle la topología del orden. A cada R^n le damos la topología producto. Supongamos que δ es una estructura sobre R tal que $\{(x,y) \in R^2 : x < y\} \in \delta_2$. Si $f : R^{n+1} \to R$ es definible en δ, entonces el conjunto*

$$A := \{a \in R^m : f(a,t) \text{ tiene límite } l(a) \text{ cuando } t \to +\infty\},$$

es definible en δ y la función límite $l : A \to \mathbb{R}$ es definible en δ.

DEMOSTRACIÓN. Describiendo el conjunto A y la función l mediante notación lógica, la proposición queda probada. Los detalles quedan a cargo del lector. □

De la proposición anterior, obtenemos el siguiente corolario:

COROLARIO 4.1. *Sea $R = \mathbb{R}$ y δ una estructura sobre $\mathbb{R}$ que contiene el gráfico de la suma y la multiplicación usual de $\mathbb{R}$. Entonces*

- $\{(x,y) \in \mathbb{R}^2 : x < y\} \in \delta_2$.
- $\{q\} \in \delta_1$ *para cada* $q \in \mathbb{Q}$.
- *Los polinomios racionales $f \in \mathbb{Q}[x_1, \cdots, x_n]$ son definibles. Luego, se tiene $\{r\} \in \delta_1$ para todo $r \in \mathbb{R}$ número algebraico y $(a,b) \in \delta_1$ para todo $a \in \mathbb{R}, b \in \mathbb{R}$ números algebraicos.*

En particular, δ_1 contiene las uniones de finitos puntos que son números algebraicos y finitos intervalos cuyos extremos son números algebraicos.

Además, si $f : (a,b) \to \mathbb{R}$ es derivable, entonces I', el conjunto de puntos de I donde f es derivable, es definible en δ_1, y $f' : I' \to \mathbb{R}$ es definible en δ.

De la Proposición 4.3, de la Proposición 4.4, y el Corolario 4.1, concluímos que si la estructura δ sobre R un espacio topológico, con la topología dada por un orden $<$, verifica la condición $\{(x,y) \in R^2 : x < y\} \in \delta_2$, entonces δ manifiesta información topológica de R. Sin embargo, puede ocurrir que la estructura δ sea demasiado complicada, y termine por contener demasiada información topológica, y por lo tanto, geometría muy complicada. En efecto, si R es un conjunto no vacío, y $\delta(1), \delta(2)$ son dos estructuras sobre R tales que $\delta(1)_m \subseteq \delta(2)_m$, escribimos $\delta(1) \subseteq \delta(2)$. Esto define un orden parcial sobre la colección de estructuras sobre R. Entonces, dada una familia $\{\delta(i)\}_{i \in I}$ de estructuras sobre R, siempre tiene un ínfimo δ (que es una estructura sobre R), dado por

$$\delta = \bigcap_{i \in I} \delta(i) \text{ con } \delta_m := \bigcap_{i \in I} \delta(i)_m \text{ para cada } m.$$

Consideremos entonces la menor estructura δ sobre $\mathbb{R}$ que contiene a los enteros $\mathbb{Z}$, el gráfico de las funciones $+, \cdot$ y los singletones $\{r\}$ para todo $r \in \mathbb{R}$. Se verifica [**22**][ejercicio 37.6] que δ es una estructura que contiene estrictamente a los borelianos. ¡Claramente los conjuntos de la estructura δ poseen geometría complicada!

La estructura δ del párrafo anterior es tan complicada debido a que $\mathbb{Z} \subseteq \delta_1$. En algún sentido, la complejidad de una estructura δ se ve fuertemente influenciada por la complejidad del álgebra δ_1. Esto motiva la definición de las estructuras O-minimales.

Sea R un conjunto totalmente ordenado por un orden $<$. Decimos que R es denso, si para todo par a, b de elementos de R, si $a < b$ entonces existe $c \in R$ tal que $a < c < r$. Decimos que R no tiene puntos extremos si no tiene máximo ni mínimo. Definimos dos elementos distintos $-\infty, +\infty$ que cumplen $-\infty < a < +\infty$ para todo $a \in R$. Entonces, R_∞ denota el conjunto R extendido, es decir, $R_\infty = R \cup \{\pm\infty\}$

En lo que sigue, R denotará un conjunto no vacío, totalmente ordenado por un orden $<$, denso y sin puntos extremos.

DEFINICIÓN 4.4. *Una estructura O-minimal sobre R es una estructura δ sobre R que cumple*

1. $\{(x,y) \in R^2 : x < y\} \in \delta_2$ *y* $\{r\} \in \delta_1$ *para cada $r \in R$ (compatibilidad de la estructura con el orden).*
2. *los conjuntos de δ_1 son exactamente las uniones finitas de intervalos abiertos y puntos (minimalidad de la estructura).*

La condición $\{(x,y) \in R^2 : x < y\} \in \delta_2$ y $\{r\} \in \delta_1$ para cada $r \in R$ fuerza en muchos casos a que los intervalos y puntos sean elementos de δ_1, luego sus uniones finitas son elementos de δ_1. La segunda condición resulta entonces "minimal" porque no contiene más que tales elementos. Esencialmente, la propiedad de minimalidad expresa la compatibilidad de naturaleza más sencilla entre R y su orden.

EJEMPLO 4.3. *Sea $\mathbb{R}_{\text{alg}}$, el cuerpo ordenado de números reales algebraicos, es decir, $\overline{\mathbb{Q}} \cap \mathbb{R}$. Es fácil de ver que el orden usual de $\mathbb{R}$ restringido a $\mathbb{R}_{\text{alg}}$ define una estructura de cuerpo totalmente ordenado, denso, sin puntos extremos, de $\mathbb{R}_{\text{alg}}$.*

Supongamos además que δ contiene el gráfico de la suma y producto de números algebraicos reales (con la suma y el producto de números reales usual). Por el Corolario 4.1, tenemos que δ_1 contiene las uniones finitas de puntos e intervalos. Sin embargo, δ no tiene por qué ser O-minimal; existe una estructura δ sobre $\mathbb{R}_{\text{alg}}$ para la cual $\mathbb{Z} \in \delta_1$. En efecto, sea δ' la estructura sobre $\mathbb{R}$ del párrafo anterior, que contenía a los borelianos de $\mathbb{R}^n$ para todo n. Definimos $\delta_n := \{\mathbb{R}_{\text{alg}} \cap \delta'_n\}$. No es difícil de ver que $\delta := (\delta_n)_n$ define una estructura sobre $\mathbb{R}_{\text{alg}}$, tal que δ_1 contiene (muchos) más conjuntos que las uniones finitas de intervalos y puntos.

OBSERVACIÓN 4.1. *Antes de introducir la noción de estructura O-minimalidad, observamos que era necesario no considerar en las estructuras conjuntos discretos infinitos, como $\mathbb{Z}$. Debido a la condición (2) de la definición de estructura O-minimal, todo conjunto definible en una estructura O-minimal sobre $\mathbb{R}$ con el orden usual, que es contable, es finito. En efecto, si $A \subseteq \mathbb{R}^n$ es definible contable infinito, entonces existe $1 \leq i \leq n$ tal que existen infinitos puntos en A con coordenada i-ésima todas distintas entre sí. Luego, la proyección en la coordenada i, $\pi_i(A)$ es un conjunto definible, contable de $\mathbb{R}$. Dado que los intervalos tienen cardinal la potencia del continuo, $\pi_i(A)$ no contiene intervalos. Por la condición (2) de la definición, es unión finita de puntos e intervalos; al no contener intervalos, resulta ser una unión finita de puntos, luego $\pi_i(A)$ es finito, lo que es contradictorio con la suposición inicial. Concluímos que A es finito.*

Por cuestiones de claridad, vamos a suponer de ahora en más, por el resto del capítulo, que cuando nos referimos a intervalos de R, pensamos en intervalos de la forma (a, b) con $a, b \in R_\infty$. Se tienen las siguientes consecuencias inmediatas de la definición de estructura O-minimal, que usaremos a lo largo del capítulo.

LEMA 4.1. *Sea $A \subseteq R$ un conjunto no vacío definible. Entonces:*

- *$\inf(A)$ y $\sup(A)$ existen en R_∞ (completitud).*
- *∂A es finito, y si $a_1 < \cdots < a_k$ son los puntos de ∂A ordenados, entonces cada intervalo (a_i, a_{i+1}) es o bien parte de A o es disjunto con A, con $a_0 = -\infty, a_{k+1} = +\infty$ (minimalidad de δ_1).*
- *Los conjuntos $\overline{A}, \dot{A}$ son definibles.*

Tenemos los siguientes ejemplos de estructuras O-minimales.

EJEMPLO 4.4. *Los subconjuntos semialgebraicos de $\mathbb{R}$ con su orden total usual definen una estructura O-minimal y constituyen el ejemplo más sencillo de estructuras O-minimales*

EJEMPLO 4.5. *Sea δ_n la colección de todos los subconjuntos de $\mathbb{R}^n$ que son de la forma $X = \pi(f^{-1}(0))$ donde, para $m \geq n$, $\pi : \mathbb{R}^m \to \mathbb{R}^n$ es la proyección en las primeras n coordenadas y $f : \mathbb{R}^m \to \mathbb{R}$ es un polinomio exponencial, es decir, $f(x_1, \cdots, x_m) = Q(x_1, \cdots, x_m, e^{x_1}, \cdots, e^{x_m})$ con $Q \in \mathbb{R}[X_1, \cdots, X_{2m}]$. Esta clase de conjuntos define una estructura O-minimal llamada $\mathbb{R}_{\text{exp}}$. La O-minimalidad de $\mathbb{R}_{\text{exp}}$ es un resultado de Wilkie* [**61**].

EJEMPLO 4.6. *Sea δ_n la colección de todos los subconjuntos de $\mathbb{R}^n$ que se escribe como unión de conjuntos de la forma*

$$\{\boldsymbol{x} = (x_1, \cdots, x_n) \in \mathbb{R}^n : f_1(\boldsymbol{x}) = \cdots = f_k(\boldsymbol{x}) = 0, g_1(\boldsymbol{x}) > 0, \cdots, g_l(\boldsymbol{x}) > 0\},$$

con f_i, g_j funciones analíticas restringidas. Una función analítica restringida $\tilde{f} : [0,1]^n \to \mathbb{R}$ es una función de la forma

$$\tilde{f}(\boldsymbol{x}) := \begin{cases} f(\boldsymbol{x}) \text{ si } x \in [0,1]^n \\ 0 \text{ en otro caso} \end{cases},$$

con $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función analítica en U un abierto de $\mathbb{R}^n$, que contiene al $[0,1]^n$. Esta clase de conjuntos determina una estructura O-minimal sobre $\mathbb{R}$, llamada $\mathbb{R}_{\text{an}}$. La O-minimalidad de $\mathbb{R}_{\text{an}}$ se deduce de un resultado de Gabrielov (consultar [**3**]*).*

EJEMPLO 4.7. *Sea $\delta_n = (\mathbb{R}_{\text{exp}})_n \cup (\mathbb{R}_{\text{an}})_n$, y consideremos la menor estructura O-minimal que contiene a las colecciones $(\delta_n)_n$. Esta estructura, denotada $\mathbb{R}_{\text{exp,an}}$, es O-minimal, y es el contexto más adecuado para generalizar el Teorema 2.1, pues las funciones analíticas trascendentes como la exponencial, son definibles en $\mathbb{R}_{\text{exp,an}}$. La O-minimalidad de $\mathbb{R}_{\text{exp,an}}$ es un resultado de van den Dries y Miller* [**35**].

2. Geometría de las estructuras O-minimales

Para entender por qué una estructura O-minimal refleja geometría sencilla de un espacio R, con R no vacío, totalmente ordenado, denso y sin puntos extremos, empezamos caracterizando las funciones definibles con dominio un intervalo.

Supongamos en lo que sigue, que δ es una estructura O-minimal sobre R un conjunto no vacío, totalmente ordenado, denso y sin puntos extremos.

LEMA 4.2. *Sea $f : I = (a, b) \to R$ una función definible en δ. Entonces existe un subintervalo de I sobre el que f es constante o inyectiva.*

DEMOSTRACIÓN. Si existe algún $y \in R$ tal que $f^{-1}(y)$ es infinito, por O-minimalidad contiene algún subintervalo de I con lo que $f|_I$ sería constantemente y. Asumamos entonces que $f^{-1}(y)$ es finito para todo y. Entonces $f(I)$ es infinito, con lo que por O-minimalidad contiene un intervalo J. Definamos entonces $g : J \to I$ mediante

$$g(y) := \min\{x \in I : f(x) = y\}.$$

Como g es inyectiva por definición, $g(J)$ es infinito. Además, $f(g(y)) = y$ de donde se deduce que g es una función definible. De nuevo, por O-minimalidad, $g(J)$ contiene un subintervalo de I, y f es necesariamente inyectiva en este intervalo. □

LEMA 4.3. *Sea $f : I = (a, b) \to R$ una función definible en δ. Si f es inyectiva, entonces f es estrictamente monótona en un subintervalo de I.*

DEMOSTRACIÓN. Para cada $x \in I$ el intervalo (a, x) es una unión disjunta de dos subconjuntos,

$$(a, x) = \{y \in (a, x) : f(y) < f(x)\} \cup \{y \in (a, x) : f(y) > f(x)\},$$

con lo que alguna de las partes contiene un intervalo (c, x), con $a < c < x$. El intervalo (x, b) se descompone de manera similar. Esto muestra que para cada $x \in I$ se satisface exactamente una de las cuatro fórmulas de primer orden siguientes:

$$\Phi_{++}(x) := \exists c_1, c_2 \in I\,[c_1 < x < c_2 \wedge \forall y \in (c_1, x) : f(y) > f(x) \wedge \forall y \in (x, c_2) : f(y) > f(x)],$$
$$\Phi_{+-}(x) := \exists c_1, c_2 \in I\,[c_1 < x < c_2 \wedge \forall y \in (c_1, x) : f(y) > f(x) \wedge \forall y \in (x, c_2) : f(y) < f(x)],$$

y se definen $\Phi_{-+}(x), \Phi_{--}(x)$ análogamente. Tenemos entonces que I contiene un subintervalo en el que todos sus puntos satisfacen una misma fórmula de las cuatro anteriores. Reemplazando I por este subintervalo, y f por su restricción a ese subintervalo, podemos asumir que todos los puntos satisfacen la misma fórmula. Esto nos lleva a cuatro casos.

Caso 1: $\Phi_{-+}(x)$ para todo $x \in I$.

Para cada $x \in I$ definimos $s(x) := \sup\{s \in (x, b) : f > f(x) \text{ en } (x, s]\}$. Entonces claramente $s(x) = b$, ya que si $s(x) < b$ se contradice $\Phi_{-+}(s(x))$. Entonces f es estrictamente creciente en I.

Caso 2: $\Phi_{+-}(x)$ para todo $x \in I$.

Este caso es análogo al caso 1.

Caso 3: $\Phi_{++}(x)$ para todo $x \in I$.

Sea $B := \{x \in I : \forall y \in I\ \ (y > z \implies f(y) > f(x))\}$, que es un conjunto definible. Si B es infinito, entonces por O-minimalidad, B contiene un intervalo, y en este intervalo f es estrictamente creciente, con lo que concluímos. Entonces asumamos que B es finito. Pasando a un subintervalo a la derecha de B, podemos asumir

$$\forall x \in I \exists y \in I\,(y > x \wedge f(y) < f(x))\,. \tag{2.1}$$

Sea $c \in I$. Afirmamos que para y suficientemente grande en I, se tiene $f(y) < f(c)$. En efecto, supongamos que no. Luego $f(y) > f(c)$ para todo $y \in (c, b)$ suficientemente grande. Sea $d \in [c, b)$ minimal tal que $\forall y\,(d < y < b \implies f(y) > f(c))$. Si $f(d) > f(c)$, entonces d no sería mínimo ya que $\Phi_{++}(d)$ es verdadera. Luego $f(d) < f(c)$. Pero por (2.1), existe e con $d < e < b$ y $f(e) < f(d)$, tal que $f(e) < f(c)$, lo que es una contradicción. Esto prueba que $f(y) < f(c)$ para todo $y \in I$ suficientemente grande.

Definamos $y(c)$ como el menor elemento de $[c, b)$ para el cual $f(y) < f(c)$ si $y(c) < y < b$.Notemos que $\Phi_{++}(c)$ nos dice que $c < y(c)$ y $f(y(c)) < f(c)$. La minimalidad de $y(c)$ implica que $y(c)$ satisface la siguiente fórmula de primer orden $\Psi_{+-}(v)$:

$$\Psi_{+-}(v) := \exists v_1, v_2 \in I\,[v_1 < v <_2 \wedge \forall z_1, z_2\,(v_1 < z_1 < v < z_2 < v_2 \implies f(z_1) > f(z_2))].$$

Dado que c era arbitrario, hemos probado que para todo $c \in I$ existe $v \in I$ tal que $v > c$ y $\Psi_{+-}(v)$. Entonces $\Psi_{+-}(v)$ vale para todo v en un intervalo de la forma (d, b) con $d \in I$. Achicando suficientemente I, podemos asumir que $\Psi_{+-}(v)$ vale para todo $v \in I$.

De manera análoga, definimos Ψ_{-+}. Podemos achicar más el subintervalo de manera que $\Psi_{-+}(v)$ valga para todo $v \in I$. Pero esto es una contradicción, ya que no pueden valer simultáneamente $\Psi_{+-}(v)$ y $\Psi_{-+}(v)$. Dado que la contradicción provino de suponer que B era finito, concluímos que B es infinito, y por lo tanto, que existe un subintervalo donde f es monótona.

Caso 4: $\Phi_{--}(x)$ para todo $x \in I$.
Este caso es análogo al caso 3. □

LEMA 4.4. *Sea $f : I = (a, b) \to R$ una función definible en δ. Si f es estrictamente monótona, entonces f es continua en un subintervalo de I.*

DEMOSTRACIÓN. Asumamos que f es estrictamente creciente (el caso que f es estrictamente decreciente es análogo). Dado que $f(I)$ es infinito, existe un intervalo $J \subseteq f(I)$. Sean $r, s \in J$ con $r < s$, y sean c, d sus preimágenes: $f(c) = r, f(d) = s, c < d$. Claramente f define una biyección de (c, d) sobre (r, s), que preserva el orden. Como la topología de R es la topología del orden, f resulta continua en (c, d). □

TEOREMA 4.1 (Teorema de monotonía). *Sea $f : (a, b) \to R$ una función definible sobre un intervalo. Entonces existen puntos $a_1 < \cdots < a_k$ en (a, b) tales que en cada subintervalo (a_j, a_{j+1}) con $a_0 = a, a_{k+1} = b$, la función f es, o bien constante, o estrictamente monótona y continua.*

DEMOSTRACIÓN. Sea

$$X := \{x \in (a, b) : \exists I \subseteq (a, b) \text{ intervalo}, x \in I \text{ tal que} f|_I \text{ es o bien constante, o estrictamente monótona y continua}\}.$$

Observemos que X es definible, pues f era definible, luego $(a, b) - X$ también es definible con lo que es una unión finita de intervalos y puntos. Si no fuera finito, contendría un intervalo I. Utilizando el Lema 4.2, el Lema 4.3 y el Lema 4.4, sucesivamente, podemos achicar suficientemente I de manera que $f|_I$ sea constante o estrictamente monótona y continua. Luego $I \subseteq X$, lo que entra en contradicción con la suposición.

Dado que $(a, b) - X$ es finito, podemos reducirnos a probar $X = (a, b)$ reemplazando (a, b) por cada uno de los intervalos que constituyen X. En particular, podemos asumir que f es continua. Particionando aún más el intervalo podemos reducirnos a probar uno de los siguientes casos:

(1) Para cada $x \in (a, b)$, se tiene que f es constante en un entorno de x.
(2) Para cada $x \in (a, b)$, se tiene que f es estrictamente creciente en un entorno de x.
(3) Para cada $x \in (a, b)$, se tiene que f es estrictamente decreciente en un entorno de x.

Estos hechos siguen de la propiedad de completitud que tienen las estructuras O-minimales (es decir, de la existencia de supremo y mínimo, ver Lema 4.1). Vamos a probar el caso (1); los otros dos casos se demuestran de manera similar y quedan a cargo del lector. Sea $x_0 \in (a, b)$ y consideremos

$$s := \sup\{x : x_0 < x < b : f \text{ es constante en } [x_0, x)\}.$$

Si $s < b$ entonces f es constante en algún entorno de s (pues X es abierto), lo que contradice el hecho que sea el supremo. Luego $s = b$, con lo que f es constante sobre $[x_0, b)$. Similarmente se prueba que f es constante en $(a, x_0]$. Concluímos que f es constante en (a, b). Las otras dos situaciones se estudian de manera similar. □

Se tienen los siguientes corolarios importantes.

COROLARIO 4.2. *Sea $f : (a, b) \to R$ definible. Entonces para cada $c \in (a, b)$ los límites $\lim_{x \to c^+} f(x), \lim_{x \to c^-} f(x)$ existen en R_∞. También los límites $\lim_{x \to b^-} f(x)$ y $\lim_{x \to a^+} f(x)$ en R_∞.*

COROLARIO 4.3. *Sea $f : [a, b] \to R$ continua y definible. Entonces f toma el máximo y el mínimo valor en $[a, b]$.*

El Teorema 4.1 está muy lejos de ser verdadero para funciones continuas. Esto se debe a que una función definible en una estructura O-minimal cumple condiciones bastante fuertes. Por ejemplo, si $f : (a, b) \subseteq \mathbb{R} \to \mathbb{R}$ es definible en una estructura O-minimal δ sobre $\mathbb{R}$ y $J \subseteq (a, b)$ es un intervalo, $f(I)$ es un conjunto definible de δ, luego es unión finita de intervalos y puntos. Una función continua arbitraria puede aplicar un intervalo en un conjunto muy complicado. La razón por la cuál una función definible tiene más regularidad que una función continua es que una función continua g es "regular" sobre los cerrados de $\mathbb{R}$, en el sentido que $g^{-1}(F)$ es cerrado para todo F cerrado, y los cerrados de $\mathbb{R}$ pueden ser bastante complicados.

Otro manifestación de geometría sencilla en las estructuras O-minimales está dada en el siguiente teorema.

TEOREMA 4.2 (Propiedad de finitud). *Sea $A \subseteq R^2$ definible y supongamos que para cada $x \in R$ la fibra $A_x := \{y \in R : (x, y) \in A\}$ es finita. Entonces existe un $N \in \mathbb{N}$ tal que $\#A_x \leq N$ para todo $x \in R$.*

DEMOSTRACIÓN. Vamos a decir que $(a, b) \in R^2$ es normal si existe una caja $I \times J$ alrededor de (a, b) tal que:

- $(I \times J) \cap A = \emptyset$ (luego $(a, b) \notin A$), o bien

- $(a,b) \in A$ y $(I \times J) \cap A = \Gamma(f)$ para alguna función continua $fI \to R$ (luego en particular f es única y definible).

Además, vamos a decir que un punto $(a,-\infty) \in R \times R_\infty$ es normal si existe una caja $I \times J$, disjunto de A, tal que $a \in I$ y $J = (-\infty, b)$ para algún b. Finalmente, vamos a decir que $(a,+\infty) \in R \times R_\infty$ es normal si existe una caja $I \times J$ disjunta de A tal que $a \in I$ y $J = (b,+\infty)$ para algún b. Notemos que los conjuntos

$$\begin{aligned} Y_1 &:= \{(a,b) \in R^2 : (a,b) \text{ es normal}\}, \\ Y_2 &:= \{a \in R : (a,-\infty) \text{ es normal}\}, \\ Y_3 &:= \{a \in R : (a,+\infty) \text{ es normal}\}, \end{aligned}$$

son definibles.

Consideremos ahora funciones $f_1, \cdots, f_n, \cdots$ con dominios:

$$\begin{aligned} \mathrm{Dom}(f_n) &:= \{x \in R : \#A_x \geq n\}, \\ f_n(x) &:= \text{n-ésimo elemento de } A_x. \end{aligned}$$

Las funciones f_n son definibles (si bien los dominios de definición $\mathrm{Dom}(f_n)$ pueden ser vacíos) Sea $a \in R$ y consideremos $n \geq 0$ maximal tal que $f_1, \cdots, f_n$ están definidas y son continuas en un intervalo que contiene a a. Decimos que:

- a es bueno, si $a \notin \overline{\mathrm{Dom}(f_{n+1})}$,
- a es malo, si $a \notin \overline{\mathrm{Dom}(f_{n+1})}$.

Sea $\mathcal{G}$ el conjunto de puntos buenos y $\mathcal{B}$ el conjunto de puntos malos. Notemos que si $a \in \mathcal{G}$ entonces (con n como recién) el dominio de f_{n+1} es disjunto con un intervalo alrededor de a, en el que $f_1, \cdots, f_n$ están definidas y son continuas. Esto muestra que para $a \in \mathcal{G}$ se tiene:

(1) $\#A_x$ es constante en el intervalo alrededor de a.
(2) (a,b) es normal para todo $b \in R_\infty$.

Lo que vamos a ver ahora es que los conjuntos $\mathcal{B}$ y $\mathcal{G}$ son definibles. Para cada $a \in \mathcal{B}$ y n como antes, sean:

$$\lambda(a,-) := \begin{cases} \lim_{x \to a^-} f_{n+1}(x) \text{ si } f_{n+1} \text{ está definida en algún intervalo } (t,a) \\ +\infty \text{ en otro caso} \end{cases},$$

$$\lambda(a,0) := \begin{cases} f_{n+1}(a) \text{ si } a \in \mathrm{Dom}(f_{n+1}) \\ +\infty \text{ en otro caso} \end{cases},$$

$$\lambda(a,-) := \begin{cases} \lim_{x \to a^+} f_{n+1}(x) \text{ si } f_{n+1} \text{ está definida en algún intervalo } (t,a) \\ +\infty \text{ en otro caso} \end{cases},$$

Sea $\beta(a) := \min\{\lambda(a,-), \lambda(a,0), \lambda(a,+)\}$. No es difícil de ver que $\beta(a)$ es el menor elemento $b \in R_\infty$ tal que (a,b) no es normal. Esto prueba que si $a \in \mathcal{B}$, entonces existe un menor elemento $b \in R_\infty$ tal que (a,b) no es normal. Este hecho, combinado con que los conjuntos Y_1, Y_2, Y_3 son definibles, y el hecho que (a,b) es normal para todo $b \in R_\infty$, nos permiten concluir que $\mathcal{B}$ y $\mathcal{G}$ son conjuntos definibles.

Supongamos ahora que $\mathcal{B}$ es finito, digamos $\mathcal{B} = \{a_1, \cdots, a_k\}$, con $-\infty = a_0 < a_1 < \cdots < a_k < a_{k+1} = +\infty$, entonces $\#A_x$ es constante en cada intervalo (a_i, a_{i+1}). En efecto, para cualquier a en este intervalo, sea $n = \#A_a$. Por (1), el conjunto $\{x \in (a_i, a_{i+1}) : \#A_x = n\}$ es abierto, y por la misma razón, el conjunto $\{x \in (a_i, a_{i+1}) : \#A_x \neq n\}$ es abierto. Dado que ambos conjuntos son definibles, el segundo debe ser vacío.

Supongamos ahora que $\mathcal{B}$ no es finito. Veamos que llegamos a una contradicción, con lo que concluímos la demostración del teorema. Sea $\beta(a)$ el menor $b \in R_\infty$ tal que (a,b) no es normal. Definimos los conjuntos

$$\begin{aligned} \mathcal{B}_- &:= \{a \in \mathcal{B} : \exists y\, (y < \beta(a) \wedge (a,y) \in A)\}, \\ \mathcal{B}_+ &:= \{a \in \mathcal{B} : \exists y\, (y > \beta(a) \wedge (a,y) \in A)\}, \end{aligned}$$

y las funciones $\beta_- : \mathcal{B}_- \to R$ y $\beta_+ : \mathcal{B}_+ \to R$ definidas por

$$\begin{aligned} \beta_-(a) &:= \max\{y : y < \beta(a) \wedge (a,y) \in A\}, \\ \beta_+(a) &:= \min\{y : y > \beta(a) \wedge (a,y) \in A\}. \end{aligned}$$

Dado que $\mathcal{B}$ estamos suponiendo que es infinito, uno de los conjuntos $\mathcal{B}_- \cap \mathcal{B}_+, \mathcal{B}_- \setminus \mathcal{B}_+, \mathcal{B}_+ \setminus \mathcal{B}_-, \mathcal{B} \setminus (\mathcal{B}_- \cup \mathcal{B}_+)$ es infinito, y cada uno de estos conduce a una contradicción. Dado que cada una de estas situaciones se trata de manera similar, sólo vamos a estudiar el caso que $\mathcal{B}_- \cap \mathcal{B}_+$ es infinito. Dado que β_-, β y β_+ son funciones definibles, por el teorema de monotonicidad 4.1, existe un intervalo $I \subseteq \mathcal{B}_- \cap \mathcal{B}_+$ en el cual las funciones β_-, β y β_+ son continuas. Notemos que $\beta_- < \beta < \beta_+$ en I. Tenemos que I se particiona en los subconjuntos $\{x \in I : (x, \beta(x)) \in A\}$ y $\{x \in I : (x, \beta(x)) \notin A\}$, y uno de estos subconjuntos contiene un intervalo. Reemplazando I por este subintervalo, podemos asumir que, o bien, $\Gamma(\beta|_I) \subseteq A$, o $\Gamma(\beta|_I) \cap A = \emptyset$, donde $\Gamma(\beta|_I)$ es el gráfico de $\beta|_I$. En cualquier caso, concluímos que $\Gamma(\beta|_I)$ posee sólo puntos normales, ya que β_-, β y β_+ don continuas en I. Dado que $(a, \beta(a))$ nunca es normal, llegamos a una contradicción. □

Combinando el Teorema 4.1 y el Teorema 4.2, tenemos una descripción sencilla de los conjuntos definibles de R^2.

Teorema 4.3. *Sea $A \subseteq R^2$ un conjunto definible tal que A_x es finito para cada $x \in R$. Entonces existen puntos $a_1 < \cdots < a_k$ en R tales que las intersecciones de A con los segmentos $(a_i, a_{i+1}) \times R$ son uniones finitas de gráficos $\Gamma(f_{ij})$ con $f_{ij} : (a_i, a_{i+1}) \to R$ funciones continuas definibles con $f_{i1}(x) < f_{i2}(x) < \cdots f_{in(i)}(x)$ para $x \in (a_i, a_{i+1})$. Hemos asumido $a_0 = -\infty$ y $a_{k+1} = +\infty$*

El teorema 4.1, el teorema 4.2 y el teorema 4.3 pueden generalizarse a dimensiones superiores. Esto es posible particionando en finitas "celdas" (conjuntos definibles muy sencillos) y tal que cada función definible en un subconjunto de R^m sea continua en cada celda. En el caso $m = 1$, las celdas tentativamente serían los puntos y los intervalos.

Sea $X \in R^m$ un conjunto definible. Definimos

$$C(X) := \{f : X \to R : f \text{ es definible y continua}\},$$
$$C_\infty(X) := C(X) \cup \{-\infty, +\infty\},$$

donde pensamos a $-\infty, +\infty$ como funcions constantes. Para $f, g \in C_\infty(X)$ escribimos $f < g$ si $f(x) < g(x)$ para todo $x \in X$, y en este caso, denotamos:

$$(f, g)_X := \{(x, r) \in X \times R : f(x) < r < g(x)\}.$$

Luego $(f, g)_X$ es un subconjunto definible de R^{m+1}, que usualmente denotaremos (f, g) si X es claro del contexto.

Definición 4.5. *Sea $(i_1, \cdots, i_m)$ una secuencia de ceros y unos de longitud m. Definimos una $(i_1, \cdots, i_m)$-celda de manera inductiva en m como sigue:*

1. *una (0)-celda es un singleton $\{r\} \subseteq R$, una (1)-celda es un intervalo $(a, b) \subseteq R$.*
2. *supongamos que tenemos definidas las $(i_1, \cdots, i_m)$-celdas. Entonces una $(i_1, \cdots, i_m, 0)$-celda es el gráfico $\Gamma(f)$ de una función $f \in C(X)$, con X una $(i_1, \cdots, i_m)$-celda; una $(i_1, \cdots, i_m, 1)$-celda es un conjunto $(f, g)_X$ con X una $(i_1, \cdots, i_m)$-celda y $f, g \in C_\infty(X), f < g$.*

Luego una $(0, 0)$-celda es un "singletón" $\{(r, s)\} \subseteq R^2$, una $(0, 1)$-celda es un intervalo en la línea vertical $\{a\} \times R$, y una $(1, 0)$-celda es el gráfico de una función definable definida en un intervalo. Adicionalmente, se puede observar que un paralelepípedo de R^m es una $(1, \cdots, 1)$-celda. En general, *una celda* en R^m es una $(i_1, \cdots, i_m))$-celda para alguna (necesariamente única) sucesión $(i_1, \cdots, i_m)$. Dado que las $(1, \cdots, 1)$-celdas son las únicas celdas abiertas, vamos a llamarlas *celdas abiertas* Por conveniencia, R^0, el espacio de un punto, es una celda, pensándola como ()-celda, donde () es la secuencia de longitud 0.

Observación 4.2. *Las celdas de R^m que no son abiertas tienen interior vacío. Este hecho puede intuirse geométricamente. Luego, se tiene que la unión de finitas celdas no abiertas tienen interior vacío.*

Observación 4.3. *Cada celda es localmente cerrada, es decir, es abierta en su clausura. Esto se ve haciendo inducción en la estructura. Si $C \subseteq R^{m+1}$ es una celda, entonces $\pi(C)$ es una celda en R^m, con $\pi : R^{m+1} \to R^m$ es la proyección en las primeras m-coordenadas. Poniendo $B = \pi(C)$, por hipótesis inductiva se tiene que B es abierto en $\overline{B}$. Entonces se usa que C como celda es el gráfico de $f : B \to R$ definible o es un conjunto de la forma $(f, g)_B$ con f, g definibles.*

Observación 4.4. *Cada celda es homeomorfa (mediante una función definible) a una celda abierta. Sea $i = (i_1, \cdots, i_m)$ una secuencia de ceros y unos. Definimos $p_i : R^m \to R^k$ de la siguiente manera: sean $\lambda(1) < \cdots < \lambda(k)$ los índices $\lambda \in [m]$ para los cuales $i_\lambda = 1$, con lo que $k = i_1 + \cdots + i_m$. Sea $p_i(x_1, \cdots, x_m) := (x_{\lambda(1)}, \cdots, x_{\lambda(k)})$. Es fácil de ver por inducción en m, que cada p_i aplicada cada i-celda A homeomórficamente, de manera definible, a una celda abierta $p_i(A)$ en R^k. Escribiendo $p(A) := p_i(A)$ y $p_A = p_i|_A : A \to p(A)$, tenemos que p_A es un homeomorfismo con la celda abierta $p(A)$, y $p_A = id_A$ si A es una celda abierta.*

Observación 4.5. *Cada celda abierta de R^l es homeomorfa a $(0,1)^l$. Para probar este hecho, procedemos por inducción en l. Notemos que una celda abierta de R es un intervalo (a,b), luego es homeo definiblemente al intervalo $(0,1)$, mediante la aplicación $\varphi(z) = \frac{1}{b-a}z$. En general, si consideramos C una celda abierta de R^n, tenemos que existe X una celda abierta de R^{n-1} y $f, g : X \to R$ funciones definibles continuas tales que $f(x) < g(x)$ para todo $x \in X$, que verifican $C = (f,g)_X$. Por hipótesis inductiva, existe $\varphi : X \to (0,1)^{n-1}$ homemorfismo definible. Entonces, la función $\psi : C \to (0,1)^n$ dada por $\psi(x_1, \cdots, x_{n-1}, x_n) := (\varphi(x_1, \cdots, x_{n-1}), \frac{1}{g(x_1,\cdots,x_{n-1}) - f(x_1,\cdots,x_{n-1})} x_{n+1})$ resulta un homeomorfismo definible entre C y $(0,1)^n$.*

Observación 4.6. *Combinando la Observación 4.4 y la Observación 4.5, deducimos que toda celda es homeomorfa, mediante un homeomorfismo definible, al $(0,1)^l$ para algún l.*

Ahora definimos lo que es una *descomposición* de R^m; como antes, esta definición es recursiva y tiene la particularidad de relacionar todos los niveles de la estructura δ sobre R, por medio de particiones.

Definición 4.6. *Sea R^m. Definimos una descomposición de R^m:*

- *Si $m = 1$, es una colección $\{(-\infty, a_1), (a_1, a_2), \cdots, (a_k, +\infty), \{a_1\}, \cdots \{a_k\}\}$ con $a_1 < \cdots < a_k$ puntos de R. O sea, es una partición finita en celdas de R.*
- *Si estamos en R^{m+1}, una descomposición es una partición finita de R^{m+1} en celdas $\{A\}$ tal que las proyecciones $\{\pi(A)\}$ son una descomposición de R^m (π es la proyección en las primeras m coordenadas).*

Observemos que es bastante fácil mostrar el aspecto de una descomposición de R^m, conociendo una partición para algunos de los m. En efecto, si $\mathcal{D} = \{A_1, \cdots, A_k\}$ es una descomposición de R^m (todas las celdas son distintas entre sí), obtenemos una descomposición de R^{m-1} proyectando. Ahora, para cada $i \in \{1, \cdots, k\}$, consideremos funciones $f_{i1} < \cdots < f_{in(i)}$ en $C(A_i)$ dadas. Entonces

$$\mathcal{D}_i := \{(-\infty, f_{i1}), (f_{i1}, f_{i2}), \cdots, (f_{in(i)}, +\infty), \Gamma(f_{i1}), \cdots, \Gamma(f_{in(i)})\},$$

es una partición de $A_i \times R$ que termina definiendo $\mathcal{D}^* = \bigcup_{i=1}^k \mathcal{D}_i$ una descomposición de R^{m+1}.

Una descomposición $\mathcal{D}$ de R^m se dice que *particiona* un conjunto $S \subseteq R^m$ si es unión de celdas en $\mathcal{D}$. Estamos finalmente en condiciones de enunciar el *teorema de descomposición en celdas*.

Teorema 4.4 (Descomposición en celdas).

(1) *Dados $A_1, \cdots, A_k \subseteq R^m$ definibles existe una descomposición de R^m que particiona a cada $A_1, \cdots, A_k$.*
(2) *Para cada función $f : A \to R, A \subseteq R^m$, existe una descomposición $\mathcal{D}$ de R^m particionando A tal que la restricción $f|_B : B \to R$ en cada celda $B \in \mathcal{D}$ con $B \subseteq A$ es continua.*

El Teorema 4.4 es la razón fundamental por la cual las estructuras O-minimales reflejan geometría sencilla. Un conjunto definible en δ una estructura O-minimal sobre R admite una descomposición en conjuntos relativamente sencillos (las celdas), en donde las funciones definibles en δ son sumamente sencillas. En particular, si la estructura O-minimal que consideramos es sobre $\mathbb{R}$, el cuerpo de los números reales con el orden usual, dado que el n-cubo $(0,1)^n$ es conexo, por la Tbservación 4.6, tenemos que las celdas son conexas. Concluímos:

Corolario 4.4. *Sea $X \subseteq \mathbb{R}^n$ un conjunto definible en una estructura O-minimal sobre $\mathbb{R}$. Entonces X tiene finitas componentes conexas.*

Destaquemos que la suposición que la estructura O-minimal sea sobre $\mathbb{R}$ es necesaria. Existen cuerpos totalmente ordenados, densos, sin puntos extremos, que verifican que $(0,1)^l$ no es conexo, con lo que hay estructuras O-minimales en las que las celdas no son conexas. Estos ejemplos pueden consultarse en [**7**]

Si bien el Teorema 4.4 da una descripción sencilla de los conjuntos y funciones definibles en una estructura O-minimal, para la geometría es fundamental el concepto de "suavidad", que utiliza esencialmente la estructura de cuerpo ordenado de $\mathbb{R}$ o la estructura de cuerpo de $\mathbb{C}$. Por esta razón, en lo que sigue, el conjunto R totalmente ordenado con $<$, sobre el que definimos una estructura O-minimal, vamos a asumir que es un cuerpo, con cierta compatibilidad con el orden $<$.

En lo que sigue, supondremos que R es un cuerpo ordenado, con orden $<$, suma $+$ y neutro aditivo 0, producto $\cdot$ y neutro multiplicativo 1. Consideramos en R la topología del orden dada por $<$, y en R^m la topología producto usual. Dado $x \in R$ y $\mathbf{x} = (x_1, \cdots, x_n) \in R^n$, definimos

$$|x| := \begin{cases} x \text{ si } x > 0 \\ -x \text{ si } x \leq 0 \end{cases}.$$
$$||(\mathbf{x})||_\infty = \max_{1 \leq i \leq n}\{|x_i|\}.$$

Dado que R es un cuerpo ordenado, podemos definir la derivada de una función $f : I \to R^n$ con $I \subseteq R$ un abierto.

DEFINICIÓN 4.7. *Sea $I \subseteq R$ un abierto. Una función $f : I \to R^n$ se dice diferenciable en un punto $x \in I$ con derivada $a \in R^n$ si*

$$\lim_{t\to 0} t^{-1}[f(x+t) - f(x)] = a.$$

La derivación de funciones definidas en un cuerpo ordenado verifica las mismas reglas de derivación usuales de la suma, producto, división y composición.

PROPOSICIÓN 4.5. *Sean $f, g : I \to R^n$ funciones diferenciables en x.*

- *Se tiene que $f + g : I \to R^n$ es diferenciable en x y $(f+g)'(x) = f'(x) + g'(x)$.*
- *Se tiene que $f \cdot g : I \to R^n$ es diferenciable en x y $(f \cdot g)'(x) = f'(x)g(x) + f(x)g'(x)$.*
- *Si $n = 1$ y g no se anula en I entonces $\frac{f}{g}$ es diferenciable en x y $(\frac{f}{g})'(x) = \frac{(f'(x)g(x) - f(x)g'(x))}{g^2(x)}$.*

PROPOSICIÓN 4.6. *Sean $I, J \subseteq R$ abiertos no vacíos. Sea $f : I \to R$ una función diferenciable en x y $g : J \to R$ una función diferenciable en $f(x) \in J$. Entonces $g \circ f : I \cap f^{-1}(J) \to R$ es diferenciable en x, y se tiene $(g \circ f)'(x) = g'(f(x)) \cdot f'(x)$.*

Definimos ahora la diferenciabilidad de una función $f : U \to R^n$ con $U \subseteq R^m$ un abierto.

DEFINICIÓN 4.8. *Sea $f : U \to R^n$ con $U \subseteq R^m$ abierto. Sea $x \in U$ y $v \in R^m$. Decimos que f es diferenciable en x en la dirección v, con derivada $a \in^n$, si se tiene*

$$\lim_{t\to 0} t^{-1}\left[f(x+tv) - f(x)\right] = a.$$

En tal caso, escribimos $\frac{\partial f}{\partial v}(x) = a$. Si definimos e^i como el vector de R^m que tiene un 1 en la coordenada i-ésima y 0's en el resto de coordenadas, denotamos $\frac{\partial f}{\partial e^i}(x) := \frac{\partial f}{\partial x_i}(x)$ y llamamos a este valor la derivada parcial i-ésima de f en x.

DEFINICIÓN 4.9. *Sea $f : U \to R^n$ con $U \subseteq R^m$ abierto. Sea $x \in U$ y $T : R^m \to R^n$ una función lineal. Decimos que f es diferenciable en x con diferencial T si para cada $\varepsilon > 0$ se verifica*

$$||f(x+v) - f(x) - T(v)||_\infty < \varepsilon ||v||_\infty,$$

para todo $v \in R^m$ suficientemente chico (en norma $||\cdot||_\infty$). La función lineal T la denotamos $T := D_x f$.

Como ocurre en el análisis clásico, se tienen las siguientes propiedades:

PROPOSICIÓN 4.7. *Sea $f = (f_1, \cdots, _m) : U \to R^n$ con $U \subseteq R^n$. Sea $x \in R^m$.*

- *Si f es diferenciable en x, para todo $v \in R^m$ se verifica $D_x f(v) = \frac{\partial f}{\partial v}(x)$.*
- *La función $f = (f_1, \cdots, f_m)$ es diferenciable en x si y sólo si f_i es diferenciable en x, para todo $1 \leq i \leq m$.*

PROPOSICIÓN 4.8.

- *Sean $f, g : U \to R^n$ funciones diferenciables en $x \in U$, con $U \subseteq R^m$ abierto. Entonces $f + g : U \to R^n$ es diferenciable en x, y se tiene $D_x(f+g) = D_x f + D_x g$.*
- *Sea $f : U \to R^n$ una función diferenciable en $x \in U$, con $U \subseteq R^m$ abierto. Entonces, para cada $c \in R$,se tiene que $c \cdot f : U \to R^n$ es diferenciable en x, y se tiene $D_x(cf) = cD_x f$.*
- *Sean $f : U \to R^n$, $h : V \to R^p$ con $U \subseteq R^m, V \subseteq R^n$ abiertos. Sea $x \in U$ tal que $f(x) \in V$. Si f es diferenciable en x y h es diferenciable en $f(x)$, entonces $h \circ f : U \cap f^{-1}(V) \to R^p$ es diferencialbe en x, y $D_x(h \circ f) = (D_{f(x)}h) \circ D_x f$.*

Sea de ahora en adelante δ, una estructura O-minimal sobre R, un cuerpo totalmente ordenado, denso, sin extremos tal que el gráfico de la suma $+$ y el producto $\cdot$ de R son definibles en δ. Observemos que, por el Corolario 4.1, tenemos:

PROPOSICIÓN 4.9. *Sea $f : I \to R$ una función definible en un intervalo I. Sea I' el subconjunto de puntos de I, donde f es derivable. Entonces I' es definible y $f' : I' \to R$ es una función definible.*

Dados $x = (x_1, \cdots, x_n), y = (y_1, \cdots, y_n) \in R^n$, definimos

$$x \cdot y := \sum_{i=1}^{n} x_i \cdot y_i$$

Sea $||x||_2 := (x \cdot x)^{\frac{1}{2}}$. Nuestro objetivo es probar:

TEOREMA 4.5. *Sea $f : I = (a, b) \to R$ una función definible en δ. Entonces f es diferenciable en (a, b) salvo en finitos puntos.*

Al igual que en el caso del Teorema 4.1, vamos a requerir varios lemas previos, para comprender la relación entre las funciones definibles en δ y la diferenciabilidad de las mismas.

LEMA 4.5 (Rolle). *Sean $a, b \in R$ tales que $a < b$. Sea $f : [a, b] \to R$ definible, continua, diferenciable en (a, b) y que verifica $f(a) = f(b)$. Entonces existe $c \in (a, b)$ tal que $f'(c) = 0$.*

DEMOSTRACIÓN. Por el Corolario 4.3, sea $a < c < b$, tal que (c) es máximo o mínimo. Al igual que en el análisis clásico, $f'(c) = 0$. □

Deducimos del Lema 4.5 los siguientes corolarios:

COROLARIO 4.5. *Sean $a, b \in R$ tales que $a < b$. Sea $f : [a, b] \to R$ definible, continua en $[a, b]$, y diferenciable en (a, b).*

- *Se tiene que existe $c \in (a, b)$ tal que $f(b) - f(a) = (b - a)f'(c)$ [teorema del valor medio].*
- *Supongamos que $f'(x) = 0$ para todo $x \in (a, b)$. Entonces f es constante.*

LEMA 4.6. *Sea $f : I = (a, b) \to R$ definible. Entonces, para todo $x \in I$, los límites*

$$f'(x^+) := \lim_{t\to 0^+} t^{-1} [f(x + t) - f(x)],$$
$$f'(x^-) := \lim_{t\to 0^-} t^{-1} [f(x + t) - f(x)],$$

existen en R_∞. Más aún, si f es continua y $f'(x^+) > 0$ para todo x, entonces f es estrictamente creciente y su inversa $f^{-1} : f(I) \to R$ satisface $(f^{-1})'(y^+) = \frac{1}{f'(x^+)}$, para todo $x \in I$ y $f(x) = y \in f(I)$. (Estamos asumiendo $\frac{1}{+\infty} = 0$)

DEMOSTRACIÓN. Fijado x, la función $g : (0, \epsilon) \to R, g(t) = t^{-1} [f(x + t) - f(x)]$ resulta definible, y además el límite $f'(x^+) = \lim_{t\to 0}$ existe por el corolario 4.3. Concluímos la existencia de $f'(x^+)$ de manera análoga.

Supongamos ahora que $f'(x^+) > 0$ para todo x, y que f es continua. Si f no fuera estrictamente creciente, entonces por el teorema de monotonía 4.1, f sería constante o estrictamente decreciente en algún subintervalo, contradiciendo $f'(x^+) > 0$ en tal subintervalo.

La expresión $(f^{-1})'(y^+) = \frac{1}{f'(x^+)}$ se obtiene de derivar la expresión $f^{-1}f(x) = x$. □

LEMA 4.7. *Sea $f : I \to R$ definible y continua, y supongamos que las funciones $x \mapsto f'(x^+), x \mapsto f'(x^-)$ son finitas y continuas en I. Entonces I es diferenciable en cada punto de I, y $f' : I \to R$ es continua.*

DEMOSTRACIÓN. Basta probar que $f'(a^+) = f'(a^-)$ para todo $a \in I$. Supongamos que no, luego existe $a \in I$ tal que $f'(a^+) > f'(a^-)$. Luego, por continuidad, existe $c \in R$ y un subintervalo J de I, alrededor de a, tal que $f'(x^+) > c > f'(x^-)$ en J. Luego, la función $g : J \to R$ dada por $g(x) := f(x) - cx$ es definible y verifica que $g'(x^+) > 0, g'(x^-) < 0$, para todo $x \in J$, luego g sería estrictamente decreciente y creciente en J, lo que es una contradicción. □

LEMA 4.8. *Sea $f : I \to R$ definible. Entonces existen finitos $x \in I$ tales que $f'(x) \in \{-\infty, +\infty\}$.*

DEMOSTRACIÓN. Consideremos el conjunto $\{x \in I : f'(x^+) = +\infty\}$, que resulta definible. Supongamos que es infinito. Por O-minimalidad, este conjunto contiene un intervalo J. Por el teorema de monotonía, podemos achicar el intervalo J de manera que f sea continua sobre J y $f'(x^+) = +\infty$. Por el Lema 4.6 esto implica que f es estrictamente creciente, luego $f'(x^-) \geq 0$ para todo $x \in J$.

Achicando más el intervalo J, podemos asumir alguna de las siguientes situaciones

(1) $f'(x^-) = +\infty$ para todo $x \in I$,
(2) $f'(x^-) \in R$ para todo $x \in I$, y la función $f'(x^-)$ es continua en J.

En el caso (1), la inversa de f satisface $(f^{-1})'(y^-) = (f^{-1})'(y^+)$ para todo $y \in f(I)$, con lo que por el Corolario 4.5, deducimos f^{-1} es constante, contradiciendo la inyectividad de f^{-1}. En el caso (2), podemos aplicar el mismo argumento que en el Lema 4.7 para obtener una contradicción. □

Ahora podemos dar la demostración del Teorema 4.5

DEMOSTRACIÓN DEL TEOREMA 4.5. Por el Teorema de monotonicidad 4.1 y el Lema 4.8, podemos reducirnos al caso en que f sea continua y $f'(x^+), f'(x^-)$ sean finitas y continuas en $I = (a, b)$. Aplicando el Lema 4.7, concluímos la proposición. □

Usando que f' es definible (por la Proposición 4.9), y el teorema de monotonía 4.1, del Teorema 4.5 obtenemos:

COROLARIO 4.6. *Sea $f : I = (a, b) \to R$ una función definible. Entonces la derivada f' es continua salvo en finitos puntos del intervalo I.*

El Corolario 4.6 puede expresarse en términos de la clase de funciones $\mathcal{C}^1$, que se define de la misma manera que en el análisis clásico.

DEFINICIÓN 4.10. *Sea $f = (f_1, \cdots, f_m) : U \to R^n$ con $U \subseteq R^m$ abierto. Decimos que f es de clase $\mathcal{C}^1(U)$ si las derivadas parciales $(\frac{\partial f_i}{\partial x_j})_{i,j}$ son funciones finitas en U y continuas.*

De la definición anterior, se sigue que si $f : U \to R^n$ es de clase $C^1(U)$ para algún $k \geq 1$ entonces f es continua.
Usando la Definición 4.10, podemos enunciar el Corolario 4.6 de la siguiente manera:

COROLARIO 4.7. *Sea $f : I = (a, b) \to R$ una función definible. Entonces existe un conjunto finito $A \subseteq I$ tal que f es de clase $\mathcal{C}^1(I \backslash A)$.*

También tenemos un resultado análogo al Teorema 4.4, que vincula la diferenciabilidad. Las celdas que se consideran para armar una descomposición son similares, salvo que las funciones que las definen son de clase $\mathcal{C}^1$.

DEFINICIÓN 4.11. *Sea $X \in R^m$ es un conjunto definible, decimos que la función $f : X \to R^n$, definible, es de clase $\mathcal{C}^1(X)$ si existe un abierto definible $U \subseteq R^m$, conteniendo al conjunto X, y una función $F : U \to R^n$ de clase $\mathcal{C}^1(U)$ tal que $F|_A = f$.*

Definimos

$$C^1(X) := \{f : X \to R : f \text{ es definible y de clase } \mathcal{C}^1(X)\},$$
$$C^1_\infty(X) := C^1(X) \cup \{-\infty, +\infty\},$$

Para $f, g \in C^1_\infty(X)$, denotamos:

$$(f, g)_X := \{(x, r) \in X \times R : f(x) < r < g(x)\}.$$

Definimos ahora las C^1-celdas:

DEFINICIÓN 4.12. *Sea $(i_1, \cdots, i_m)$ una secuencia de ceros y unos de longitud m. Definimos una C^1-$(i_1, \cdots, i_m)$-celda de manera inductiva en m como sigue:*

1. *una C^1-(0)-celda es un singleton $\{r\} \subseteq R$, una C^1-(1)-celda es un intervalo $(a, b) \subseteq R$.*
2. *supongamos que tenemos definidas las C^1-$(i_1, \cdots, i_m)$-celdas. Entonces una C^1-$(i_1, \cdots, i_m, 0)$-celda es el gráfico $\Gamma(f)$ de una función $f \in C^1(X)$, con X una C^1-$(i_1, \cdots, i_m)$-celda; una C^1-$(i_1, \cdots, i_m, 1)$-celda es un conjunto $(f, g)_X$ con X una C^1-$(i_1, \cdots, i_m)$-celda y $f, g \in C^1_\infty(X), f < g$.*

Los comentarios posteriores a la Definición 4.5 permanecen siendo válidos para la definición de C^1-celdas. Particularmente, la Observación 4.6 ahora tiene una versión $\mathcal{C}^1$:

PROPOSICIÓN 4.10. *Toda $\mathcal{C}^1$-celda es homeomorfa, mediante un homeomorfismo definible de clase $\mathcal{C}^1$, al $(0, 1)^l$ para algún l.*

La definición de descomposición 4.6 se generaliza para descomposición en C^1-celdas. Si $\mathcal{D}$ es una descomposición de R^m en C^1-celdas, decimos que *particiona* un conjunto $S \subseteq R^m$ si es unión de celdas en $\mathcal{D}$.
Enunciamos ahora la generalización del Teorema 4.4.

TEOREMA 4.6 (Descomposición en C^1-celdas).

1. *Dados $A_1, \cdots, A_k \subseteq R^m$ definibles existe una descomposición de R^m en C^1-celdas, que particiona a cada $A_1, \cdots, A_k$.*
2. *Para cada función $f : A \to R, A \subseteq R^m$, existe una descomposición $\mathcal{D}$ de R^m en C^1-celdas, particionando A, tal que la restricción $f|_B : B \to R$, en cada C^1-celda $B \in \mathcal{D}$ con $B \subseteq A$, es de clase $\mathcal{C}^1(B)$.*
3. *Para cada función $f : A \to R, A \subseteq R^m$, si p es un punto interior de A, definimos $\nabla f(p) := (\frac{\partial f}{\partial x_1}(p), \cdots, \frac{\partial f}{\partial x_m}(p))$, siempre que todas las derivadas parciales existan en p. Si alguna derivada parcial on está definida en p, entonces ∇f no está definida en p. Si $A' := \{p \in A : p$ es un punto interior de A en el que ∇f está definido$\}$, se tiene que $A \backslash A'$ tiene interior vacío.*

Es posible obtener una versión más general del Teorema 4.6, igual de útil, que pide más condiciones de diferenciabilidad. En efecto, como así definimos la clase $\mathcal{C}^1(X)$ de funciones definibles, con $X \subseteq R^m$ definible, podemos definir la clase $\mathcal{C}^k(X)$ con $k \geq 1$. Teniendo una noción de diferenciabilidad para derivadas superiores, la Proposición 4.10, el Corolario 4.7 y el Teorema 4.6 se pueden generalizar:

PROPOSICIÓN 4.11. *Toda $\mathcal{C}^k$-celda es difeomorfa,, mediante un difeomorfismo definible de clase $\mathcal{C}^k$, al $(0, 1)^l$ para algún l.*

TEOREMA 4.7 (Descomposición en C^k-celdas). *Sea $f : I = (a, b) \to R$ una función definible. Entonces, dado $k \geq 1$ un entero positivo, existe un conjunto finito $A_k \subseteq I$ tal que f es de clase $\mathcal{C}^k(I \backslash A_k)$.*

TEOREMA 4.8.

(1) *Dados $A_1, \cdots, A_k \subseteq R^m$ definibles existe una descomposición de R^m en C^k-celdas, que particiona a cada $A_1, \cdots, A_k$.*
(2) *Para cada función $f : A \to R, A \subseteq R^m$, existe una descomposición $\mathcal{D}$ de R^m en C^k-celdas, particionando A, tal que la restricción $f|_B : B \to R$, en cada C^k-celda $B \in \mathcal{D}$ con $B \subseteq A$, es de clase $\mathcal{C}^1(B)$.*

OBSERVACIÓN 4.7. *Una función definible $f : I = (a, b) \to R$ no tiene por qué ser analítica, o al menos ser de clase $\mathcal{C}^\infty$, en un subintervalo de I; en* [**45**] *se construyen estructuras O-minimales sobre $\mathbb{R}$ que contienen el gráfico de la suma y el producto de números reales, que no admiten descomposición en celdas analíticas.*

Antes de terminar esta sección, vamos a mencionar la noción de *dimensión* de un conjunto definible.

DEFINICIÓN 4.13. *Sea $C \subseteq R^n$ una $\mathcal{C}^k$-celda. Por la Proposición 4.11, tenemos que C es difeomorfa definiblemente al $(0, 1)^d$, para algún d entero positivo. Definimos la dimensión de C como* $\dim(C) = d$.

DEFINICIÓN 4.14. *Sea $X \subseteq R^n$ un conjunto definible. Supongamos que X admite una descomposición en $\mathcal{C}^k$-celdas $C_1, \cdots, C_r$. Definimos la dimensión de X como el máximo de las dimensiones de las celdas que aparecen en su descomposición, es decir,* $\dim(X) = \max_{1 \leq i \leq n}(\dim_i(C_i))$.

La Definición 4.14 en un principio podría no ser adecuada, ya que depende de la descomposición en $\mathcal{C}^k$-celdas, en particular, depende del k. Otro problema es que podría ocurrir que la dimensión no fuera invariante frente a homeomorfismos/difeomorfismos definibles. Estas posibles complicaciones no ocurren; dado que no requerimos en las próximas secciones propiedades sobre la dimensión de los conjuntos definibles, asumiremos ad hoc que la Definición 4.14 es invariante por homeomorfismos/difeomorfismos definibles y que no depende de la descomposición de la descomposición en celdas, ni de la regularidad de las celdas. La buena definición y buen comportamiento de la Definición 4.14 puede consultarse en el capítulo 4 de [**54**].

Como caso particular al concepto de dimensión de un conjunto definible, tenemos la dimensión de conjuntos semialgebraicos, noción que utilizaremos en la próxima sección. Sea $X \subseteq \mathbb{R}^n$ un subconjunto semialgebraico. Dado que los subconjuntos semialgebraicos reales son una estructura O-minimal, tenemos que X admite una descomposición en celdas semialgebraicas $C_1, \cdots, C_n$. Si definimos la dimensión de una celda semialgebraica de la misma manera que una celda general, tenemos que la dimensión de X es la dimensión máxima de todas las celdas semialgebraicas de su descomposición, es decir, $\dim(X) = \max_{1 \leq i \leq n}(\dim(C_i))$.

Antes de concluir esta sección, mencionamos el siguiente resultado, que refuerza la idea que los conjuntos definibles en estructuras O-minimales son geométricamente sencillos.

TEOREMA 4.9. *Sean $A \subseteq R^m, B \subseteq R^n$ conjuntos definibles. Decimos que A y B son definiblemente homeomórficos si existe $f : A \to B$ un homeomorfismo definible. La noción de definiblemente homeomórfico define una clase de equivalencia en los conjuntos de definibles de la estructura. Se tiene que existen a lo sumo numerables clases de homeomorfismos definibles.*

El teorema anterior puede consultarse en el capítulo 8 de [**54**].

3. Existencia de parametrizaciones buenas

El interés de las estructuras O-minimales viene de buscar una familia de conjuntos y funciones suficientemente amplia, que refleje propiedades geométricas de un espacio, sin tener una geometría demasiado complicada. El interés detrás de este objetivo provenía de buscar una generalización del teorema:

TEOREMA. *Sea $f : [0, 1] \to \mathbb{R}$ una función analítica trascendente. Sea Γ el gráfico de f. Para todo $N \geq 1$ y $\varepsilon > 0$, se tiene:*

$$\#\Gamma(\mathbb{Q}, N) = O_{f,\varepsilon}(t^\varepsilon).$$

La razón por la cuál las estructuras O-minimales dan un lenguaje adecuado para generalizar el teorema anterior, es que los conjuntos definibles en una estructura O-minimal sobre $\mathbb{R}$, o un cuerpo R totalmente ordenado, denso, sin puntos extremos, admiten "parametrizaciones buenas". Para entender qué entendemos por "parametrizaciones buenas", introducimos las siguientes definiciones. A lo largo de esta sección, R denotará un cuerpo totalmente ordenado, denso, sin puntos extremos, y δ una estructura O-minimal sobre R, que contiene el gráfico de la suma y el producto de R. Dado que un cuerpo con las

hipótesis dadas no puede tener característica p, se verifica que R tiene a $\mathbb{Q}$ como cuerpo primo de R; identificamos $\mathbb{Q}$ en R y, por lo tanto, identificamos $\mathbb{N}$ en R.

DEFINICIÓN 4.15.

- *Decimos que un elemento $a \in R$ es finito si $|a| \leq c$ para algún $c \in \mathbb{N}$. Un elemento finito de R también decimos que está fuertemente acotado.*
- *Decimos que $\boldsymbol{a} = (a_1, \cdots, a_n) \in R^n$ está fuertemente acotado si todas sus coordenadas están fuertemente acotadas.*
- *Decimos que X, un conjunto definible de R^n, está fuertemente acotado si existe $c \in \mathbb{N}$ tal que $|a| \leq c$ para todo $a \in X$.*
- *Decimos que una función definible $f : X \to R^n$ está fuertemente acotada si su gráfico está fuertemente acotado (equivalentemente, si sus dominios y rangos lo están).*

OBSERVACIÓN 4.8. *Las definiciones anteriores son equivalentes a la noción usual de acotado, si el cuerpo R es arquimediano.*

Introducimos ahora la definición de parametrización de un conjunto definible.

DEFINICIÓN 4.16. *Sea $X \subseteq R^n$ definible y sea $d := \dim(X)$, la dimensión de X como conjunto definible.*

- *Una función definible $\varphi : (0,1)^d \to X$ se dice una parametrización parcial de X.*
- *Una función definible $\varphi : (0,1)^d \to X$ de clase $\mathcal{C}^k(X)$ se dice una $\mathcal{C}^k$-parametrización parcial de X.*
- *Un conjunto finito de parametrizaciones parciales de X se dice una parametrización de X si $\bigcup_{\varphi \in S} \operatorname{rank}(\varphi) = S$.*
- *Un conjunto finito de $\mathcal{C}^k$-parametrizaciones parciales de X se dice una $\mathcal{C}^k$-parametrización de X si $\bigcup_{\varphi \in S} \operatorname{rank}(\varphi) = S$.*

Notemos que decir que un conjunto X definible admite una parametrización es, en algún sentido, equivalente a que admita una descripción sencilla. Puesto que los conjuntos definibles admiten una descomposición en celdas, siempre existen parametrizaciones.

PROPOSICIÓN 4.12. *Sea $X \subseteq R^n$ un conjunto definible, de dimensión d. Para todo $k \geq 1$, se tiene que existe una $\mathcal{C}^k$-parametrización de X.*

DEMOSTRACIÓN. Sea $k \geq 1$. Por el Teorema 4.6, tenemos que existe una descomposición en C^k-celdas de R^n que particiona al conjunto X, es decir, existen $\mathcal{C}^k$-celdas $C_1, \cdots, C_r$ tales que $C_i \cap C_j = \emptyset$, $R^n = \bigcup_{i=1}^r C_i$ y $X = \bigcup_{1 \leq i \leq r, C_i \cap X \neq \emptyset} C_i$. Por la Proposición 4.11, cada celda C_i es difeomorfa definiblemente a $(0,1)^{l_i}$, con difeomorfismo $\varphi_i : (0,1)^{l_i} \to C_i$ de clase $\mathcal{C}^k$. Tenemos tres posibilidades:

- Si $l_i = d$, entonces $\psi_i := \varphi_i$ define una parametrización parcial de X.
- Si $l_i = 0$, entonces $C_i = \{x\}$ con $x \in X$, luego $\psi_i : (0,1)^d \to X$, dada por $\psi_i(y) := x$ define una parametrización parcial de X.
- Si $0 < l_i < d$, entonces definimos $\psi_i : (0,1)^d \to X$ como $\psi_i := \varphi_i \circ \pi_{l_i}$, con $\pi_{l_i} : R^d \to R^{l_i}$ la proyección en las primeras l_i-coordenadas.

En cualquier caso, el conjunto finito $S := \{\psi_i : 1 \leq i \leq r\}$ es una $\mathcal{C}^k$-parametrización de X. □

OBSERVACIÓN 4.9. *Sea $X \subseteq \mathbb{R}^n$ una variedad diferencial de dimensión n. Si X es compacto, resulta que existen finitas cartas $\{(U_i, \varphi_i)\}_i$ que permiten parametrizar a la variedad X. Luego, podemos decir que la condición de ser X un conjunto definible en una estructura O-minimal, sumada a la condición de ser fuertemente acotado, permiten pensar a X, en algunos aspectos, como una variedad compacta.*

La Proposición 4.12 expresa, de una manera más geométrica, el contenido del Teorema 4.6, y muestra un comportamiento geométrico sencillo de los conjuntos definibles en una estructura O-minimal. Sin embargo, las parametrizaciones que describen a un conjunto definible, se pueden elegir de manera que sean todavía más regulares.

DEFINICIÓN 4.17. *Una parametrización $S = \{\varphi_i\}_{1 \leq i \leq n}$ de un conjunto definible X se dice una r-parametrización si cada $\varphi \in S$ es de clase $\mathcal{C}^r$ y tiene la propiedad de que $\varphi^{(\alpha)}$ está fuertemente acotada para cada $\alpha \in \mathbb{N}^{\dim(X)}$ con $||\alpha||_1 \leq r$, donde $||\alpha||_1$ es la suma de las coordenadas de α.*

Una r-parametrización es una parametrización muy regular; no es obvio en general que una parametrización así exista para una variedad diferencial compacta.

Ahora, sea $X \subseteq R^m$ definible, y consideremos una función $F : X \to R^n$ definible. Nos interesa expresar F en coordenadas, es decir, si S es una parametrización de X, las coordenadas de F son $F \circ \varphi$ para toda $\varphi \in S$.

DEFINICIÓN 4.18. *Supongamos que S es una r-parametrización de un conjunto definible $X \subseteq R^m$. Sea $F : x \to M^n$ una función definible. Decimos que S es una r-reparametrización de F si, para cada $\varphi \in S$, se tiene que $F \circ \varphi$ es de clase $\mathcal{C}^r$ y $(F \circ \varphi)^{(\alpha)}$ está fuertemente acotada para todo $\alpha \in \mathbb{N}^{\dim(X)}$ con $||\alpha||_1 \leq r$, donde $||\alpha||_1$ es la suma de las coordenadas de α.*

El hecho increíble, probado por Pila y Wilkie en [**42**], que verifican los conjuntos y funciones definibles en una estructura O-minimal es que siempre admiten r-parametrizaciones y r-reparametrizaciones

TEOREMA 4.10 (Gromov[6]-Pila-Wilkie). *Sea $r \geq 1$ un entero positivo. Si $X \subseteq R^n$ es definible y fuertemente acotado, entonces existe una r-parametrización de X*

TEOREMA 4.11 (Pila-Wilkie). *Sea $r \geq 1$ un entero positivo y $X \subseteq R^p$ definible. Si $F : X \to R^n$ es una función definible, entonces existe una r-reparametrización.*

La idea/estrategia que se suele aplicar para estudiar un conjunto $X \subseteq R^n$ definible en una estructura O-minimal, es la siguiente:

(1) Particionamos el conjunto X en celdas.
(2) Dado que hay, para cada R^n, esencialmente dos tipos de celdas, que son o bien de la forma $(f, g)_Y$ o $\Gamma(h)$ con $f, g, h : Y \to R$ funciones definibles, Y un conjunto definible, estudiamos estos dos tipos de celdas.
(3) Combinamos la partición en celdas y el estudio de las celdas particulares, para obtener información del conjunto X.

En el caso de un conjunto definible $X \subseteq R$, la descomposición en celdas da una unión finita de intervalos y puntos. Debido a que en este caso las celdas son sumamente sencillas, tenemos que el Teorema 4.10, para conjuntos $X \subseteq R$ definible, es inmediata.

LEMA 4.9 (Teorema 4.10 para $X \subseteq R$). *Sea $r \geq 1$ un entero positivo. Si $X \subseteq R$ es definible y fuertemente acotado, entonces existe una r-parametrización de X*

DEMOSTRACIÓN. Sea $X \subseteq R$ definible. Por O-minimalidad, se tiene que X es unión disjunta de finitos intervalos $(a_1, b_1), \cdots, (a_n, b_n)$ y finitos puntos $\{x_1\}, \cdots, \{x_m\}$. Como X es fuertemente acotado, los intervalos y los puntos son fuertemente acotados.

- Dado un intervalo (a_i, b_i), la asignación $\varphi_i(x) := (b_i - a_i)x + a_i$ es una función definible, fuertemente acotada (porque el intervalo lo es), de clase $\mathcal{C}^r$, que aplica suryectivamente el intervalo $(0, 1)$ al intervalo (a_i, b_i).
- Dado $\{x_j\}$, la asignación $\psi_j(x) := x_j$ es definible, fuertemente acotada (porque el punto x_j es fuertemente acotado), de clase $\mathcal{C}^r$, que aplica suryectivamente el intervalo $(0, 1)$ al punto $\{x_j\}$.

Entonces $S = \{\varphi_i, \psi_j : i \in [n], j \in [m]\}$ es una r-parametrización de X.

□

Para estudiar funciones definibles $F : X \to R^n$ con $X \subseteq R^m$ definible, la idea/estrategia es similar a la idea/estrategia para estudiar conjuntos definibles:

(1) Particionamos el conjunto X en celdas de manera que F, restringida a cada celda, sea muy regular.
(2) Estudiamos funciones definibles $F : C \to R^n$ cuyo dominio de definición es una celda.
(3) Combinamos la partición en celdas y el estudio de las restricciones de la función F a celdas particulares, para obtener información del conjunto X.

En el caso de una función definible $F : I = (a, b) \to R$, la descomposición en celdas da el teorema de monotonía 4.1 y el Teorema 4.7, es decir, la función F es muy regular, salvo en finitos puntos. Por medio del Teorema 4.7, podemos dar una demostración del Teorema 4.11 para el caso de un conjunto definible $X \subseteq R$.

LEMA 4.10. *Sea $r \geq 2$ y supongamos que $f : (0, 1) \to R$ es una función definible de clase $\mathcal{C}^r$, con $f^{(j)}$ fuertemente acotado para $0 \leq j \leq r - 1$. Supongamos adicionalmente que $|f^{(r)}|$ es monótona decreciente. Sea $g : (0, 1) \to M$ definida por*

$$g(x) := f(x^2).$$

Entonces $g^{(j)}$ está fuertemente acotada para $0 \leq j \leq r$.

[6]El primero en obtener un resultado similar fue Gromov [**16**], que mostró que los conjuntos semialgebraicos fuertemente acotados admiten r-parametrizaciones

DEMOSTRACIÓN. Por la regla de la cadena, $g^{(i)}(x) = \sum_{j=0}^{i} \rho_{i,j}(x) f^{(j)}(x^2)$ para cada $0 \leq i \leq r$ y $x \in (0,1)$, donde cada $\rho_{i,j}$ es un polinomio con coeficientes enteros de grado j.

Notemos que por las hipótesis sobre f, todos los sumandos de $g^{(i)}$ están fuertemente acotados salvo, quizás, el que tiene $i = j = r$. Es fácil de comprobar que este sumando es $2^r x^r f^{(r)}(x^2)$. Sea c un entero positivo que acota fuertemente a $f^{(r-1)}$ y supongamos que existe $x_0 \in (0,1)$ tal que $|f^{(r)}(x_0)| > \frac{4c}{x_0}$. Por el teorema del valor medio 4.5, existe $\xi \in (\frac{x_0}{2}, x_0)$ tal que

$$f^{(r-1)}(x_0) - f^{(r-1)}(\frac{x_0}{2}) = f^{(r)}(\xi)(x_0 - \frac{x_0}{2}).$$

Como $|f^{(r)}|$ es monótona decreciente, se tiene $|f^{(r)}(\xi)| \geq |f^{(r)}(x_0)| > \dfrac{4c}{x_0}$. Entonces:

$$2c \geq |f^{(r-1)}(x_0) - f^{(r-1)}(\frac{x_0}{2})| > \frac{4c}{x_0}(x_0 - \frac{x_0}{2}) = 2c,$$

lo que es una contradicción. Concluímos entonces que

$$|2^r x^r f^{(r)}(x^2)| \leq 2^r x^r \frac{4c}{x^2} \leq 2^{r+2} c,$$

donde en la cota final usamos $r \geq 2$ y $x \in (0,1)$. Luego $g^{(i)}$ está fuertemente acotada para $0 \leq i \leq r$, como queríamos ver. □

LEMA 4.11. *Sea $F : (0,1) \to R$ una función definible, fuertemente acotada. Entonces F admite una 1-reparametrización S con la propiedad adicional de que para cada $\varphi \in S$, se tiene φ o $F \circ \varphi$ es un polinomio (en el $(0,1)$) con coeficientes fuertemente acotados.*

DEMOSTRACIÓN. Por el Corolario 4.7, sabemos que existen $a_0 = 0 < a_1 < \cdots < a_p < a_{p-1} = 1$ elementos de R tales que $F|_{(a_i,a_{i+1})}$ es de clase $\mathcal{C}^1$ y se verifica o bien $|F|'_{(a_i,a_{i+1})}| \leq 1$ o $|F|'_{(a_i,a_{i+1})}| \geq 1$ (en cuyo caso la función restringida es estrictamente monótona).

En el primer caso, $|F'| \leq 1$, definimos $\varphi_i : (0,1) \to R$ dada por $x \mapsto (a_{i+1} - a_i)x + a_i$.

En el segundo caso, sea $b_i := \lim_{x \to a_i^+} F(x), b_{i+1} = \lim_{x \to a_i^-} F(x)$ (bien definidos por ser F estrictamente monótona en los intervalos) y definimos $\varphi_i : (0,1) \to R$ dada por $x \mapsto F^{-1}((b_{i+1} - b_i)x + b_i)$.

En cualquier caso, $\operatorname{rank}(\varphi_i) = (a_i, a_{i+1})$ y tanto φ_i como $F \circ \varphi_i$ son de clase $\mathcal{C}^1$ en $(0,1)$ con derivadas fuertemente acotadas. Adicionalmente, al menos una de tales funciones es lineal con coeficientes en $[-1,1]$. Agregando los puntos que faltan (en donde eventualmente F no era diferenciable) como funciones constantes, tenemos que $S = \{\varphi_0, \cdots, \varphi_p, \hat{a}_1, \cdots, \hat{a}_p\}$ es una 1-reparametrización de F con la propiedad requerida, donde $\hat{a}_i \equiv a_i$ (la función constantemente a_i). □

LEMA 4.12. *Sea $r \geq 1$ y supongamos que $F : (0,1) \to R$ es una función definible, fuertemente acotada. Entonces F tiene una r-reparametrización para la que, o bien φ o bien $F \circ \varphi$ es un polinomio (restringido al $(0,1)$) con coeficientes fuertemente acotados.*

PROOF. La demostración es por inducción en r. El caso $r = 1$ es el Lema 4.11. Supongamos entonces $r \geq 2$ y que S es una $(r-1)$-reparametrización de F con la propiedad del lema. Sea $\varphi \in S$ y escribamos $\{\varphi, F \circ \varphi\} = \{g, h\}$ donde g es un polinomio en el $(0,1)$ con coeficientes fuertemente acotados. Entonces, en particular, $g^{(i)}$ existe y es fuertemente acotada para todo i. Sin embargo, de $h^{(i)}$ sólo sabemos que existe, es continua, y es fuertemente acotada para $0 \leq i \leq r-1$. Por el Teorema 4.7, podemos particionar el $(0,1)$, con $0 = a_0 < a_1 < \cdots < a_{p_\varphi} < a_{p_\varphi+1} = 1$ de manera que para cada $0 \leq i \leq p_\varphi$, la función h es de clase $\mathcal{C}^r$ en (a_i, a_{i+1}) y $|h^{(r)}|$ es monótona (posiblemente constante) en (a_i, a_{i+1}).

Sea $\theta_{\varphi,i} : (0,1) \to (0,1)$ definido por

$$\theta_{\varphi,i}(x) := \begin{cases} (a_{i+1} - a_i)x + a_i \text{ si } |h^{(r)}| \text{ es decreciente o constante} \\ (a_i - a_{i+1})x + a_{i+1} \text{ si } |h^{(r)}| \text{ es creciente pero no constante} \end{cases}.$$

Resulta entonces que $h \circ \theta_{\varphi,i} : (0,1) \to R$ es de clase $\mathcal{C}^r$ y $(h \circ \theta_{\varphi,i})^{(i)}$ está fuertemente acotada para $0 \leq i \leq r-1$. Adicionalmente, por construcción se tiene $|(h \circ \theta_{\varphi,i})^{(r)}|$ es monótona decreciente. Llamemos $\rho : (0,1) \to (0,1)$ a la función (de clase $\mathcal{C}^\infty$) que asigna $x \mapsto x^2$. Por el Lema 4.10, la función $h \circ \theta_{\varphi,i} \circ \rho : (0,1) \to R$ tiene derivada i-ésima fuertemente acotada para $0 \leq i \leq r$. Adicionalmente, la función $g \circ \theta_{\varphi,i} \circ \rho$ es un polinomio con coeficientes fuertemente acotados y $\{h \circ \theta_{\varphi,i} \circ \rho, g \circ \theta_{\varphi,i} \circ \rho\} = \{\varphi \circ \theta_{\varphi,i} \circ \rho, F \circ (\varphi \circ \theta_{\varphi,i} \circ \rho)\}$. Notemos que i varía de 0 a p_φ, luego $\operatorname{rank}(\varphi \circ \theta_{\varphi,i} \circ \rho)$ cubre $\operatorname{rank}(\varphi)$,

salvo finitos puntos (aquellos puntos donde h no era de clase $\mathcal{C}^r$, dados posiblemente por los extremos de los intervalos). Luego sólo necesitamos agregar finitas funciones constantes (con valor en $(0,1)$) al conjunto $\varphi \circ \theta_{\varphi,i} \circ \rho : \varphi \in S\}$ para que se vuelva una r-reparametrización de F con la propiedad del lema. Esto completa la inducción. □

COROLARIO 4.8. *Sea $X \subseteq R$ fuertemente acotado y $F : X \to R$ una función fuertemente acotada. Entonces para cada $r \geq 1$ se tiene que F posee una r-reparametrización.*

DEMOSTRACIÓN. Ya vimos que si $X \subseteq R$ es definible, entonces tiene una r-parametrización S. Usando el Lema 4.8, podemos r-reparametrizar cada función $F \circ \varphi : (0,1) \to R$ para $\varphi \in S$ y tomar la unión de todas estas r-reparametrizaciones como r-reparametrización. □

LEMA 4.13. *Sean $m, r \geq 1$ enteros positivos. Supongamos que toda función $f : X \subseteq R^l \to R$ definible, fuertemente acotada, con $l \leq m$, tiene una r-reparametrización. Entonces para cada $n \geq 1$, se tiene que toda función $F : X \subseteq R^l \to R^n$ definible, fuertemente acotada, con $l \leq m$, tiene una r-reparametrización.*

DEMOSTRACIÓN. Por medio de un argumento inductivo, alcanza con probar que si $n \geq 2$ y $F : X \to R^{n-1}, f : X \to R$ son definibles y fuertemente acotadas y F tiene una r-reparametrización, entonces la función $(F, f) : X \to R^n$, definida como $(F, f)(x) = (F(x), f(x))$, tiene una r-reparametrización. Por medio de la inclusión $R^l \subseteq R^m$ (que es definible, recordar los comentarios posteriores a la Definición 4.1), podemos suponer que X es un subconjunto definible, fuertemente acotado, de R^m.

Sea S una r-reparametrización de F, y sea $\varphi \in S$ con $\varphi : (0,1)^l \to X$ y $l = \dim(X) \leq m$. Aplicando la hipótesis del lema a la función $f \circ \varphi : (0,1)^l \to R$ obtenemos una r-reparametrización T_φ. Entonces, para cada $\psi \in T_\varphi$ se tiene que su dominio de definición es $(0,1)^l$ y se sigue, por uso seguido de la regla de la cadena, que cada $(\varphi \circ \psi)^{(\alpha)} : (0,1)^l \to R^m$ con $||\alpha||_1 \leq r$, es fuertemente acotada. Entonces, tenemos que $(\{\varphi \circ \psi : \varphi \in S, \psi \in T_\varphi\})$ es una r-reparametrización de (F, f). □

TEOREMA 4.12 (Teorema 4.11 para el caso $X \subseteq R$). *Sea $r \geq 1$ un entero positivo y $X \subseteq R$ definible. Si $F : X \to R^n$ es una función definible, entonces existe una r-reparametrización.*

DEMOSTRACIÓN. El resultado a probar se deduce del Corolario 4.8 y el caso $m = 1$ del Lema 4.13. □

Con el Teorema 4.11 para el caso $X \subseteq R$, podemos probar el Teorema 4.10 para $X \subseteq R^2$.

TEOREMA 4.13 (Teorema 4.10 para el caso $X \subseteq R^2$). *Sea $r \geq 1$ un entero positivo. Si $X \subseteq R^2$ es definible y fuertemente acotado, entonces existe una r-parametrización de X*

DEMOSTRACIÓN. Sea $r \geq 1$. Por medio del teorema de descomposición en celdas 4.6, si mostramos que las $\mathcal{C}^r$-celdas admiten r-parametrizaciones, uniendo todas las r-parametrizaciones de las $\mathcal{C}^r$-celdas que aparecen en la partición de X en $\mathcal{C}^r$-celdas, obtenemos que X admite una r-parametrización. Supongamos entonces que X es una $\mathcal{C}^r$-celda de R^2.

- Supongamos que $X = \Gamma(f)$ con $f : Y \to R$ una función definible, y $Y \subseteq R$ una celda, necesariamente fuertemente acotada, al serlo X. Por el lema 4.9, tenemos que existe una r-parametrización de Y. Luego, para cada $\varphi \in S$, consideramos la función $G_\varphi = (\mathrm{Id} \circ \varphi, f \circ \varphi) : (0,1) \to R^2$ definida mediante $(\mathrm{Id} \circ \varphi, f \circ \varphi)(x) := (\mathrm{Id} \circ \varphi(x), f \circ \varphi(x))$. Por el teorema 4.12, obtenemos una r-reparametrización S_φ de G_φ. La unión de todas las r-reparametrización S_φ define una r-parametrización de X.
- Sea $X = (f, g)_Y$ con $f, g : Y \to R$ funciones definibles, y $Y \subseteq R$ una celda, fuertemente acotada. Por el lema 4.9, tenemos que existe S una r-parametrización de Y. Para cada $\varphi \in S$, consideramos la función $G_\varphi = (f \circ \varphi, g \circ \varphi) : (0,1) \to R^2$ definida como $(f \circ \varphi, g \circ \varphi)(x) := (f \circ \varphi(x), g \circ \varphi(x))$. Por medio del teorema 4.12, tenemos que existe una r-reparametrización S_φ de G_φ. Para cada $\psi \in S_\varphi$, definimos $\theta_{\varphi,\psi} : (0,1)^2 \to X$ mediante la expresión:
$$\theta_{\varphi,\psi}(x_1, x_2) := (\varphi \circ \psi(x_1), (1 - x_2) f \circ \varphi \circ \psi(x_1) + x_2 g \circ \varphi \circ \psi(x_1).$$
El conjunto finito $\{\theta_{\varphi,\psi} : \varphi \in S, \psi \in S_\varphi\}$ es fácil de ver que es una r-parametrización de X.

□

Observemos que el uso del Teorema 4.12 no es necesario. Esto se debe a que las celdas $Y \subseteq R$ son, o bien, intervalos (a, b) o puntos $\{x\}$, con lo que en el Teorema 4.13 es posible reemplazar el empleo del Teorema 4.12 por el teorema de descomposición en celdas 4.6. La razón por la cuál utilizamos el Teorema 4.12 es para mostrar la estrategia general de la demostración del Teorema 4.10 y el teorema 4.11. Dado $m \geq 1$ un entero positivo, escribimos las proposiciones:

$(I)_m$ Para cada $r, n \geq 1$ y toda función $F : (0,1)^m \to R^n$ definible, fuertemente acotada, existe una r-reparametrización de F.

$(II)_m$ Para cada $r \geq 1$ y todo subconjunto $X \subseteq R^{m+1}$ definible, fuertemente acotado, existe una r-parametrización de X.

La demostración del teorema 4.10 y el Teorema 4.11 se realiza mediante el siguiente argumento inductivo. Sabemos que las proposiciones $(I)_1$ y $(II)_0$ valen (por el Teorema 4.12 y el Lema 4.9, respectivamente). Luego, si $m \geq 1$, asumimos que $(I)_m$ vale para todo $l \leq m$ y que $(II)_m$ vale para todo $l < m$.

- Si X tiene dimensión $l \leq m$, por hipótesis inductiva $(II)_{m-1}$ y $(I)_k$ para todo $k \leq m$ son verdaderas, luego deducimos, mediante un argumento análogo al empleado en la demostración del teorema 4.13, que $(II)_m$ es verdadero.
- Usando que asumimos $(I)_m$ es verdadera, y usando que $(II)_m$ es verdadera bajo las hipótesis de inducción, se concluye que $(I)_{m+1}$ es verdadero.

El segundo ítem de la "estrategia" anterior es el único paso complicado del Teorema 4.10 y el Teorema 4.11. Para los detalles, consultar el trabajo original de Pila y Wilkie [**42**].

Del Teorema 4.10 y el teorema 4.11 se obtiene fácilmente el siguiente corolario:

Corolario 4.9. *Sean $m, r \geq 1$, y supongamos que $X \subseteq (0,1)^m$ es un conjunto definible. Entonces existe un conjunto finito de funciones S, tal que para toda $\varphi \in S$, se tiene $\varphi : (0,1)^{\dim(X)} \to X$ es de clase $\mathcal{C}^r$, tal que*

(1) $\bigcup_{\varphi \in S} \operatorname{rank}(\varphi) = X$,
(2) $|\varphi^{(\alpha)}(x)| \leq 1$ *para toda* $\varphi \in S, \alpha \in \mathbb{N}^{\dim(X)}$ *con* $||\alpha||_1 \leq r$ *y todo* $x \in (0,1)^{\dim(X)}$.

Demostración. Sea S^* una r-parametrización de X, dada por el Teorema 4.10. Entonces (1) vale para S^*, y (2) vale con c reemplazado por 1 para algún $c \in \mathbb{N}$. Cubrimos $(0,1)^{\dim(X)}$ con $(2c)^{\dim(X)}$ cubos de lado $\frac{1}{c}$, y para cada cubo K de estos, sea $\lambda_K : (0,1)^{\dim(X)} \to K$ una función linear biyectiva. Entonces

$$S := \{\varphi \circ \lambda_K : \varphi \in S^*, K \text{ cubo del cubrimiento}\},$$

verifica las condiciones requeridas. □

4. El Teorema de Pila-Wilkie

En esta sección, todas las estructuras O-minimales que consideramos son estructuras O-minimales sobre $\mathbb{R}$, que contienen la suma y el producto de números reales. En particular, estas estructuras O-minimales contienen los subconjuntos semialgebraicos de $\mathbb{R}$.

El objetivo de esta sección es generalizar el Teorema 2.2, que ya recordamos en la sección anterior. Vimos en el Ejemplo 2.1 y el Ejemplo 2.2, que en general no es posible obtener una estimación mucho mejor que la del Teorema 2.2. Adicionalmente, la hipótesis de trascendencia del Teorema 2.2 era importante, pues en caso contrario, la curva podía tener demasiados puntos racionales "algebraicos", en el sentido, que estén contenidos en conjuntos semialgebraicos de dimensión positiva. En dimensiones superiores, por ejemplo, si I es un intervalo cerrado y acotado, y consideramos $f : I^2 \to \mathbb{R}$, una función analítica trascendente, puede ocurrir que el gráfico de f, que denotamos X, contenga subconjuntos semialgebraicos de dimensión positiva, y por lo tanto, demasiados puntos racionales. Por ejemplo, si $X = \{(x,y,z) : z = x^y : x, y \in [1,2]\}$, para cada y racional fijo, el conjunto $X_y = \{(x,z) : z - x^y,\ x \in [1,2]\}$ determina una curva semialgebraica[7] en X. Por esta razón, para generalizar el Teorema 2.2 se excluyen los puntos que están contenidos en subconjuntos semialgebraicos "grandes".

Definición 4.19. *Sea $X \subseteq \mathbb{R}^n$. La parte algebraica de X, que denotamos X^{alg}, es la unión de todas los subconjuntos conexos semialgebraicos de X, de dimensión positiva. La parte trascendente de X es el complemento $X \backslash X^{\text{alg}}$.*

Si consideramos la parte trascendente de X, tenemos una estimación similar al Teorema 2.2. Pila y Wilkie en [**42**] observan que, de hecho, la estimación se sigue manteniendo para X definible en una estructura O-minimal.

Teorema 4.14 (Pila-Wilkie). *Sea $X \subseteq \mathbb{R}^n$ un conjunto definible en una estructura O-minimal sobre $\mathbb{R}$, que contiene el gráfico de la suma y el producto, y sea $\varepsilon > 0$. Entonces para todo $H > 0$, se tiene la estimación*

$$\#((X - X^{\text{alg}})(\mathbb{Q}), H) = O_{X,\varepsilon}(H^{\varepsilon}).$$

[7]Una curva semialgebraica es la imagen de una función $\sigma : (0,1) \to R^n$ semialgebraica.

Para entender por qué el teorema de Pila-Wilkie 4.14 generaliza adecuadamente el Teorema 2.2, consideremos la estructura O-minimal $\mathbb{R}_{\text{an}}$. Si X es el gráfico de una función analítica trascendente $f : [0, N] \to \mathbb{R}$, tenemos que X es definible en $\mathbb{R}_{\text{an}}$. Por el Teorema 2.2, ya tenemos la estimación del Teorema 4.14. Por otro lado, en la Observación 4.7, hay estructuras O-minimales que poseen funciones definibles que no son analíticas en ningún punto. Luego, si consideramos como X el gráfico de una tal función, este conjunto es definible en alguna estructura O-minimal sobre $\mathbb{R}$ que contiene la suma y el producto de reales, y por lo tanto, verifica el Teorema 4.14.

Para probar el Teorema 4.14, se podría proceder como hicimos para el Teorema 2.1. Sea $X \subseteq \mathbb{R}^n$ definible en una estructura O-minimal.

(1) Dado $\varepsilon > 0$, determinamos la cantidad de hipersuperficies algebraicas de grado $d(\varepsilon)$ que cubren los puntos $(X - X^{\text{alg}})(\mathbb{Q}, H)$, por medio del método del determinante de Bombieri-Pila.
(2) Acotamos $\#\left[((X\backslash X^{\text{alg}})(\mathbb{Q}), H) \cap V\right]$ para $V \subseteq \mathbb{R}^n$ una hipersuperficie algebraica.
(3) Usando las dos estimaciones anteriores, obtenemos una cota superior para $\#((X\backslash X^{\text{alg}})(\mathbb{Q}), H)$.

El principal problema con la estrategia anterior es el paso (2), debido a que el conjunto X^{alg} puede ser muy complicado. En efecto, supongamos que el paso (1) pudimos realizarlo, con un conjunto $X \subseteq \mathbb{R}^n$ definible en una estructura O-minimal que contiene el gráfico de la suma y el producto de $\mathbb{R}$. En la demostración del Teorema 2.1 usamos que toda curva trascendente analítica se interseca en finitos puntos con una curva algebraica. Si consideramos $S \subseteq \mathbb{R}^n$ una hipersuperficie algebraica, quisiéramos que $(X\backslash X^{\text{alg}}) \cap S$ fuera finito. Sin la hipótesis de trascendencia, podríamos intentar usar que $X\backslash X^{\text{alg}}$ sea semialgebraico. En tal caso $(X\backslash X^{\text{alg}}) \cap S$ sería semialgebraico, con lo que si es infinito, tendría dimensión positiva y por lo tanto $(X\backslash X^{\text{alg}}) \cap S \subseteq X^{\text{alg}}$. El problema principal es que X^{alg} no tiene ni siquiera por qué ser definible en alguna estructura O-minimal. En efecto:

Ejemplo 4.8. *Si $X = \{(x, y, z) : z = x^y : x, y \in [1, 2]\}$ observamos que X contiene infinitas componentes conexas, una por cada $y \in \mathbb{Q} \cap [1, 2]$: estas componentes conexas son los conjuntos $X_y = \{(x, z) : z = x^y, x \in [1, 2]\}$. Debido a que un conjunto definible en una estructura O-minimal sobre $\mathbb{R}$, tiene finitas componentes conexas (Corolario 4.4), concluímos que X^{alg} no es definible en ninguna estructura O-minimal sobre $\mathbb{R}$ con lo que, en particular, la parte trascendente de X no es definible en ninguna estructura O-minimal sobre $\mathbb{R}$.*

Observación 4.10. *La complejidad de la parte algebraica de un conjunto definible puede verse en un caso bien particular. Sea $f : (0, 1) \to \mathbb{R}$ definible. Consideramos entonces el conjunto $\Gamma(f)$. Puede verse (consultar [**54**]) que si existe un polinomio $P(x, y) \in \mathbb{R}[x, y]$ no nulo tal que $P(x, f(x)) = 0$ para todo $x \in (0, 1)$, entonces f es semialgebraica, luego $\Gamma(f) = \Gamma(f)^{\text{alg}}$ es un conjunto semialgebraico. Sin embargo, puede ocurrir que exista un polinomio no nulo $Q(x, y) \in \mathbb{R}[x, y]$ tal que $Q(x, f(x)) = 0$ para todo $x \in I'$ con $I' \subseteq (0, 1)$ un intervalo propio de $(0, 1)$, con lo que $\Gamma(f)$ contiene un tramo, que es el gráfico de una función semialgebraica. Si $g = f|_{I'}$, entonces $\Gamma(g) \subseteq \Gamma(f)^{\text{alg}}$. Y así, podemos seguir encontrando tramos semialgebraicos en $\Gamma(f)$. El problema es, que en un principio, este proceso no tiene por qué terminar, con lo que, de nuevo, puede ocurrir que $\Gamma(f)^{\text{alg}}$ no sea, ni siquiera definible.*

No vamos a probar el teorema de Pila-Wilkie, sin embargo, explicamos cómo podemos adaptar la estrategia anterior bosquejada.

Pila Y Wilkie observan que estudiar los puntos de intersección de varias hipersuperficies algebraicas con un conjunto definible $X \subseteq \mathbb{R}^n$ puede interpretarse en términos del concepto de familia definible

Definición 4.20. *Sea $Z \subseteq \mathbb{R}^n \times \mathbb{R}^m$, y sean π_1, π_2 las proyecciones en $\mathbb{R}^n$ y $\mathbb{R}^m$ respectivamente. Sea $Y := Y_Z = \pi_2(Z)$, y para cada $y \in Y$, sea $Z_y := \pi_2^{-1}(\{y\})$. Sea $X_y := X_{Z,y} := \pi_1(Z_y)$. Una familia de conjuntos $Z \subseteq \mathbb{R}^n \times \mathbb{R}^m$ significa una colección de fibras $\{X_y : y \in Y_Z\}$. Una familia Z se dice definible si el conjunto Z es definible.*

Definición 4.21. *Sea $Z \subseteq \mathbb{R}^n \times \mathbb{R}^m$ una familia definible. La dimensión de bira de Z es el máximo de las dimensiones de todas las fibras de Z.*

De ahora en más, las notaciones Y_Z, X_y de la Definición 4.20 van a ser usadas en el resto de la sección.

En términos de la Definición 4.20, el teorema que se puede intentar probar es:

Teorema 4.15 (Pila-Wilkie). *Sea $Z \subseteq \mathbb{R}^n \times \mathbb{R}^m$ una familia definible, y sea $\varepsilon > 0$. Entonces existe una constante $c(Z, \varepsilon)$ con la siguiente propiedad. Sea X una fibra de Z. Entonces para todo $H > 0$ se tiene*

$$\#(X\backslash X^{\text{alg}})(\mathbb{Q}, H) \leq c(Z, \varepsilon) H^{\varepsilon}.$$

En el Ejemplo 4.8 observamos que para $X \subseteq \mathbb{R}^n$ definible, no tiene por qué ser X^{alg} siquiera definible. Sin embargo, podría ocurrir que para todo X y $\varepsilon > 0$, existe un subconjunto semialgebraico $X_\varepsilon \subseteq X$ tal que $\#(X \backslash X_\varepsilon)(\mathbb{Q}, H) = O_{X,\varepsilon}(H^\varepsilon)$. Este no es el caso:

EJEMPLO 4.9. *Sea $X := \{(x, y) : 0 < x < 1, 0 < y < \exp(x)\}$. Entonces $X^{\text{alg}} = X$, pues X es abierto, luego todo punto $p \in X$ está contenido en un disco alrededor de p, y es claro que los discos son conjuntos semialgebraicos, de dimensión positiva. Pero X no es semialgebraico, pues en tal caso, el borde de X, el gráfico de la función exponencial en $[0, 1]$, sería semialgebraica. No es muy difícil de ver que esto no es posible; por ejemplo, vale que toda función semialgebraica está polinomialmente acotada (ver por ejemplo [**7**][capítulo 2, prop. 2.6.1, pág. 42]).*

Sin embargo, dado X y $\varepsilon > 0$, prodríamos esperar que exista un conjunto definible $X_\varepsilon \subseteq X^{\text{alg}}$ tal que $\#(X \backslash X_\varepsilon)(\mathbb{Q}, H) = O_{X,\varepsilon}(H^\varepsilon)$. Para una familia definible Z, los conjuntos X_ε podrían ser elegidos como las fibras de una familia definible $W(Z, \varepsilon) \subseteq Z$. Esto es lo que prueban Pila y Wilkie en [**42**]

TEOREMA 4.16 (Pila-Wilkie). *Sea $Z \subseteq \mathbb{R}^n \times \mathbb{R}^m$ una familia definible. Sean π_1, π_2 las proyecciones en $\mathbb{R}^n$ y $\mathbb{R}^m$ respectivamente. Sea $Y := Y_Z = \pi_2(Z)$, y para cada $y \in Y$, sea $Z_y := \pi_2^{-1}(\{y\})$. Sea $X_y := X_{Z,y} := \pi_1(Z_y)$. Dado $\varepsilon > 0$, se tiene que existe una familia definible $W := W(Z, \varepsilon)$ y una constante $c(Z, \varepsilon)$ con la siguiente propiedad: dado $y \in Y$, sea $X = X_{Z,y}$ y $X_\varepsilon = X_{W,y}$. Entonces $X_\varepsilon \subseteq X^{\text{alg}}$, y, para todo $H > 0$, se tiene*

$$(X \backslash X_\varepsilon)(\mathbb{Q}, H) \leq c(Z, \varepsilon) H^\varepsilon.$$

No es difícil de ver que el Teorema 4.16 implica al Teorema 4.15, y que a su vez, este resultado implica el Teorema 4.14.

En términos del concepto de familia definible, adaptamos el método del determinante Bombieri-Pila para estimar la cantidad de hipersuperficies algebraicas que cubren los puntos racionales de altura acotada de las fibras de una familia definible. Para ello, recordamos y reenunciamos el Teorema 1.2 que probamos en el capítulo 1:

PROPOSICIÓN 4.13. *Sean k, n enteros positivos con $k < n$. Entonces, para cada $d \geq 1$ entero positivo, existe $r = r(k, n, d)$ un entero positivo, y una constante positiva $\varepsilon(k, n, d)$, con la siguiente propiedad: supongamos que $\varphi : (0, 1)^k \to \mathbb{R}^n$ es una función de clase $\mathcal{C}^r$ con $|\phi^{(\alpha)}(x)| \leq 1$ para todo $x \in (0, 1)^k$ y todo $\alpha \in \mathbb{N}^k$ con $||\alpha||_1 \leq r$. Sea $X = \phi((0, 1)^k)$, y sea $H \geq 1$. Entonces $X(\mathbb{Q}, H)$ está contenido en la unión de no más de $O_{k,n,d}(H^{\varepsilon(k,n,d)})$ hipersuperficies algebraicas de grado a lo sumo d. Además, $\varepsilon(k, n, d) \to 0$ cuando $d \to +\infty$.*

La Proposición 4.13 se deduce fácilmente del Teorema 1.2. Podemos observar la utilidad del Teorema 4.10, concretamente, su Corolario 4.9: dado $X \subseteq \mathbb{R}^n$, por el Corolario 4.9, existe una r-parametrización S de X que verifica que cada $\varphi \in S$ satisface las hipótesis de la Proposición 4.13. Podemos entonces concluir que $X(\mathbb{Q}, H)$ está contenido en la unión de no más de $\#SO_{k,n,d}(H^{\varepsilon(k,n,d)}) = O_{X,n,d}(H^{\varepsilon(k,n,d)})$ hipersuperficies algebraicas de grado a lo sumo d.

Dado que el Teorema 4.16 está enunciado en términos de una familia definible, se debe adaptar la Proposición 4.13 a una familia definible. Para ello, se recurre a una versión del Corolario 4.9 para familias definibles:

PROPOSICIÓN 4.14 (versión del Corolario 4.9 para familias definibles). *Sean $n, m, r \geq 1$, y supongamos que $X \subseteq (0, 1)^n \times \mathbb{R}^m$ es una familia definible. Entonces existe un $N \geq 1$ entero positivo y, para cada $y \in \mathbb{R}^m$, existe un conjunto S_y de N funciones, tales que si $\varphi \in S_y$ entonces $\varphi : (0, 1)^{\dim(X_y)} \to X_y$ es de clase $\mathcal{C}^r$, y además:*

1. $\bigcup_{\varphi \in S_y} \text{rank}(\varphi) = X_y$,
2. $|\varphi^{(\alpha)}(x)| \leq 1$, *para cada $\varphi \in S_y, \alpha \in \mathbb{N}^{\dim(X_y)}$, con $||\alpha||_1 \leq r$ y para todo $x \in (0, 1)^{\dim(X_y)}$.*
3. *Las funciones que componen S_y dependen de manera definible en y.*

No vamos a probar la Proposición 4.14; destacamos que hay cierto tipo de técnicas que permiten pasar, por ejemplo, de conjuntos definibles a familias definibles (en este caso, para deducir la Proposición 4.14 del Corolario 4.9), pero no vamos a desarrollar estas técnicas. Referirse a [**42**] para los detalles.

Con la Proposición 4.14 podemos obtener una estimación para un cubrimiento de hipersuperficies algebraicas, en el contexto de familias definibles.

LEMA 4.14. *Sea $Z \subseteq (0, 1)^n \times \mathbb{R}^m$ una familia definible, de dimensión de fibra $k < n$. Sea $\varepsilon > 0$. Entonces existe $d = d(\varepsilon, k, n)$ un entero positivo, y una constante $K(Z, \varepsilon)$ con la siguiente propiedad: para todo $y \in Y$ y $H \geq 1$, el conjunto $X(\mathbb{Q}, H)$, con $X = X_y$, está contenido en la unión de no más de $K(Z, \varepsilon) H^\varepsilon$ hipersuperficies algebraicas de grado a lo sumo d.*

Por medio del Lema 4.14, lo único que faltaría para poder completar el esquema de la demostración del Teorema 4.16, es tener una cota suficientemente uniforme entre la intersección de hipersuperficies algebraicas y los conjuntos $X(\mathbb{Q}, H)$ del Lema 4.14. Como explicamos antes del Ejemplo 4.8, la complicación proviene de la complejidad de la parte algebraica de un conjunto definible. Por un momento, consideremos el Teorema 4.14, para $X \subseteq \mathbb{R}^2$ definible, fuertemente acotado. Por el teorema de descomposición en celdas 4.6, tenemos que X es unión finita de:

- puntos,
- intervalos de la forma $(a, b) \times \{c\}$,
- intervalos de la forma $\{f\} \times (d, e)$,
- conjuntos $\Gamma(f)$, que son gráficos de funciones definibles $f : I \to \mathbb{R}$, con $I \subseteq \mathbb{R}$ un intervalo,
- conjuntos de la forma $(g, h)_I$, con $g, h : I' \to \mathbb{R}$ funciones definibles en un intervalo $I' \subseteq \mathbb{R}$.

De la lista de conjuntos anteriores, no es difícil de ver que, salvo $\Gamma(f)$, el resto están contenidos en la parte semialgebraica de X. Como mencionamos en el Ejemplo 4.10, el conjunto $\Gamma(f)^{\text{alg}}$ puede ser bastante complicado, con lo que es de esperar que los conjuntos definibles de $\mathbb{R}^3$, y $\mathbb{R}^n$ en general, tenga parte algebraica muy complicada (ver Ejemplo 4.8).

Consideremos de nuevo $X \subseteq \mathbb{R}^2$ definible. En la descomposición en celdas, las celdas "difíciles" de estudiar resultaron ser las celdas de dimensión 1. En general, lo que va a ocurrir para $X \subseteq \mathbb{R}^n$ definible es que, los "puntos complicados" de X van a estar contenidos en celdas de dimensión $k < n$. Esto motiva a estudiar:

Definición 4.22. *Sea $X \subseteq \mathbb{R}^n$ un conjunto definible, y sea $p \geq 1$ un entero positivo. Definimos $\text{Reg}^p(X)$ como el conjunto de todos los $x \in X$ para los que existe un entorno U de x tal que $U \cap X$ es una $\mathcal{C}^p$-subvariedad embebida de X. Para cada $k \leq n$, definimos $\text{Reg}^p_k(X)$, como el conjunto de puntos $x \in X$ tales que existe un entorno U de x tal que $U \cap X$ es una $\mathcal{C}^p$-subvariedad embebida de X, de dimensión k.*

Se tiene el siguiente resultado, probado en [**36**][sección 1.8, pág. 4]:

Teorema 4.17. *Para cada $X \subseteq \mathbb{R}^n$ definible, y k, p enteros positivos, se tiene que $\text{Reg}^p_k(X)$ es un conjunto definible.*

Pila y Wilkie utilizan el Teorema 4.17, combinado con un argumento inductivo, para poder probar el Teorema 4.16. El Teorema 4.17 se utiliza para reducir el Teorema 4.14, esencialmente, al estudio de las celdas de dimensión $k < n$, o en el contexto de familias definibles del Teorema 4.16, para reducirnos a estudiar familias definibles $Z \subseteq \mathbb{R}^n \times \mathbb{R}^m$ con dimensión de fibra $k < n$. El resto de la demostración se basa en cubrir por hipersuperficies algebraicas de grado acotado a un conjunto $X \subseteq \mathbb{R}^n$ definible, o en el caso de una familia definible, en cubrir por hipersuperficies algebraicas de grado acotado las fibras de una familia definible Z, utilizando el Teorema 1.2 o el Lema 4.14, y luego se usa un argumento de compacidad, esencialmente, que las hipersuperficies algebraicas de $\mathbb{R}^k$ de grado d definen un espacio compacto, concretamente, son un plano proyectivo $\mathbb{P}^{\nu(k,d)}$. Para los detalles, consultar [**42**].

5. Argumento de Pila-Zannier

El teorema de Pila-Wilkie 4.14 da una estimación para los puntos racionales de altura a lo sumo H, de la parte trascendente de un conjunto definible en una estructura O-minimal sobre $\mathbb{R}$, que contiene a la suma y al producto de $\mathbb{R}$. La parte trascendente de un conjunto X, $X \backslash X^{\text{alg}}$, posee cierto parecido con los *puntos especiales* de una variedad algebraica. Esta relación entre puntos trascendentes y puntos especiales es utilizada por Pila y Zannier [**43**] para dar otra demostración de la conjetura de Manin-Mumford, por Masser y Zannier [**32**] para dar un resultado relacionado con los puntos de torsión de curvas elípticas, y posteriormente por Pila [**40**] para dar una demostración incondicional (es decir, independiente de la hipótesis de Riemann generalizada) de la conjetura de André-Oort para producto de curvas modulares. Estos resultados son bastante profundos, y son sumamente importantes para la geometría diofántica, pues dan información geométrica del comportamiento de los conjuntos algebraicos. En esta sección vamos a explicar un caso particular de [**43**], siguiendo el survey de Scanlon [**46**].

Sea $\mathbb{G} := \mathbb{C}^\times$ el grupo de unidades de los números complejos, con operación $\cdot$, el producto usual entre complejos.

Definición 4.23. *Sea $l \geq 0$. Decimos que $\boldsymbol{g} \in \mathbb{G}^l$ es un punto especial si $\boldsymbol{g} = (g_1, \cdots, g_l)$ tiene orden finito, es decir, cada g_i es una raíz de la unidad. Por convención, el punto $\mathbb{G}^0$ es especial.*

Sea $g \in \mathbb{G}$. Si g es un punto especial, entonces existe un entero positivo n tal que $g^n = 1$. En particular, $g = \exp(2\pi i z)$ para algún $z \in \mathbb{Q}$. En general, si $\mathbf{g} = (g_1, \cdots, g_l) \in \mathbb{G}^l$ es un punto especial, entonces existen $(z_1, \cdots, z_n) \in \mathbb{Q}^l$ tales que $\mathbf{g} = (\exp(2\pi i z_1, \cdots, \exp(2\pi i z_n)))$. Si definimos $E : \mathbb{C}^l \to \mathbb{G}^l$ mediante $E(z_1, \cdots, z_l) := (\exp(2\pi i z_1), \cdots, \exp(2\pi i z_l))$, tenemos:

PROPOSICIÓN 4.15. *Sea $l \geq 0$. El punto $\boldsymbol{g} \in \mathbb{G}^l$ es un punto especial si y sólo si existe $\xi = (\xi_1, \cdots, \xi_l) \in \mathbb{Q}^l$ tal que $E(\xi) = \boldsymbol{g}$.*

La Proposición 4.15 muestra que estudiar los puntos especiales de $\mathbb{G}^l$ es equivalente a estudiar los puntos racionales de la función analítica $E : \mathbb{C}^l \to \mathbb{G}^l$. Dado que tenemos una descripción explícita de los puntos especiales de $\mathbb{G}^l$, la Proposición 4.15 no parece aportarnos nada. El problema de interés para la conjetura de Manin-Mumford son los puntos especiales de variedades algebraicas.

TEOREMA 4.18 (Mann, 1965). *Sea $n \geq 1$ un entero positivo. Sea $G(x_1, \cdots, x_l) \in \mathbb{C}[x_1, \cdots, x_l]$ un polinomio no nulo. Entonces el conjunto de puntos especiales de la hipersuperficie algebraica $\{(x_1, \cdots, x_l) \in \mathbb{C}^l : G(x_1, \cdots, x_l) = 0\}$ es una unión finita de coclases de subgrupos de $\mathbb{G}^l$.*

El Teorema 4.18, que es un caso particular de la conjetura de Manin-Mumford, puede interpretarse como estudiar el conjunto de soluciones racionales de la ecuación $G(E(z)) = 0$, con G el polinomio del Teorema 4.18 y $E((z_1, \cdots, z_l) = (\exp(2\pi i z_1), \cdots, \exp(2\pi i z_l))$. Identificando $\mathbb{C}$ con $\mathbb{R}^2$ tomando parte real e imaginaria, podemos pensar la ecuación $G(E(z)) = 0$ como una ecuación analítica real. Por el teorema de Pila-Wilkie, podemos obtener información de los puntos racionales de altura a lo sumo H del conjunto $X := \{z = (\mathrm{Re}(z), \mathrm{Im}(z)) \in \mathbb{R}^{2l} : G(E(z)) = 0\}$ si este fuera definible en alguna estructura O-minimal, por ejemplo, $\mathbb{R}_{\exp}$ o $\mathbb{R}_{\mathrm{an}}$ (estas estructuras O-minimales fueron introducidas al final de la sección 1 de este capítulo). Sin embargo, si nó restringimos el dominio de la función E, el conjunto X podría no resultar definible, debido a que la función $E : \mathbb{R}^{2l} \to \mathbb{G}^l$. En efecto, si la función $E : \mathbb{C}^l \to \mathbb{G}^l$ fuera definible en una estructura O-minimal δ sobre $\mathbb{R}$, entonces las preimágenes deberían ser definibles en δ (ver Propiedad 4.2). Luego, $E^{-1}(\{1, \cdots, 1\}) = \mathbb{Z}^l$ (visto en $\mathbb{R}^{2l}$), que es un conjunto que tiene infinitas componentes conexas, por lo tanto, no resulta definible en ninguna estructura O-minimal (Corolario 4.4). Sin embargo, si consideramos $D := \{z = (z_1, \cdots, z_l) \in \mathbb{C}^l : 0 \leq \mathrm{Re}(z_i) < 1, \ \forall i\}$, identificado como subespacio de $\mathbb{R}^l$, no es difícil de ver que $\tilde{E} := E|_D$ es una función definible en $\mathbb{R}_{\exp}$, y por lo tanto, como $G|_{E(D)}$ es definible en $\mathbb{R}_{\exp}$, resulta que $G \circ E$ es definible en $\mathbb{R}_{\exp}$ y por lo tanto $\tilde{X} = \tilde{E}^{-1}(G^{-1}(\{0\}))$ es un conjunto definible en $\mathbb{R}_{\exp}$.

En lo que sigue, vamos a mantener la notación del párrafo anterior para los conjuntos $D, \tilde{X}$ y las funciones $\tilde{E}, G$. Podemos describir brevemente el argumento de Pila-Zannier, aplicado a la demostración del Teorema 4.18.

- Buscamos estudiar información sobre la distribución de los puntos especiales de una variedad algebraica V. Para ello, encontramos una función definible F en una estructura O-minimal sobre $\mathbb{R}$, como $\mathbb{R}_{\exp}, \mathbb{R}_{\mathrm{an}}$ o $\mathbb{R}_{\exp,\mathrm{an}}$ tal que los puntos especiales de V son los puntos racionales del conjunto $\tilde{X} = F^{-1}(\{0\})$.
- Estudiamos la parte algebraica de $\tilde{X}$ y utilizamos el teorema de Pila-Wilkie 4.14 para obtener una cota superior de la cantidad $\#(\tilde{X} \backslash \tilde{X}^{\mathrm{alg}})(\mathbb{Q}, H)$.
- Obtenemos una cota inferior para $\#\tilde{X}(\mathbb{Q}, H)$ y comparamos con la cota del Teorema 4.14.

De la estrategia recién bosquejada, sólo hemos realizado el primer paso. Para estudiar la parte algebraica del conjunto $\tilde{X}$, vamos a usar el siguiente resultado, que es una versión de la conjetura de Schanuel.

TEOREMA 4.19 (Conjetura de Schanuel para cuerpos funcionales, [**1**]). *Sean $\gamma_1(t), \cdots, \gamma_l(t) \in t\mathbb{C}[[t]]$ series de potencias sobre los números complejos, sin término constante, linealmente independientes sobre $\mathbb{Q}$. Entonces el grado de trascendencia sobre $\mathbb{C}(t)$ de la extensión de cuerpos $\mathbb{C}(t, \gamma_1(t), \cdots, \gamma_l(t), \exp(\gamma_1(t)), \cdots, \exp(\gamma_l(t)))$ es al menos n.*

OBSERVACIÓN 4.11. *Usando la notación posterior al Teorema 4.18, recordamos que el conjunto $\tilde{X} = \{z \in D : G(\tilde{E}(z)) = 0\}$ es definible en $\mathbb{R}_{\exp}$. Sea $\gamma = (\gamma_1, \cdots, \gamma_l) : (0, 1) \to \tilde{X}$ una curva semialgebraica. Puesto que $G(x_1, \cdots, x_l) \in \mathbb{C}[x_1, \cdots, x_l]$ es un polinomio no nulo que verifica $G(\tilde{E}(\gamma_1(t), \cdots, \gamma_l(t))) = 0$, por el Teorema 4.19, concluímos que $\gamma_1, \cdots, \gamma_l$ son linealmente dependientes sobre $\mathbb{Q}$.*

FINAL DEL ESQUEMA DE LA DEMOSTRACIÓN DEL TEOREMA 4.18. Para cada entero positivo $k \leq l$, consideramos el conjunto M_k de espacios afines V (traslaciones de subespacios) de dimensión k para los que $\dim(V \cap \tilde{X}) = k$, pero para los cuales existe un punto $a \in V \cap \tilde{X}$ para el que no hay un espacio afín W de dimensión $k + 1$ tal que $a \in W \cap \tilde{X}$ pero $\dim(W \cap \tilde{X}) < k + 1$. Dado que los espacios afines se describen por medio de ecuaciones afines, no es difícil de ver que cada M_k es definible en $\mathbb{R}_{\exp}$ (de hecho, son semialgebraicos). Usando la definición de la función $\tilde{E}$ (consultar [**46**], [**43**]), se concluye que M_k está definido en $\mathbb{Q}$. Luego, cada M_k es contable y definible en $\mathbb{R}_{\exp}$. Por la Propiedad 4.1, M_k resulta finito. Usando este argumento y la Observación 4.11, deducimos que $\tilde{X}^{\mathrm{alg}} = \bigcup_{k=1}^{l} \bigcup_{H \in M_k} H$. Concluímos que $\tilde{X}^{\mathrm{alg}}(\mathbb{Q})$ es la unión finita de coclases de grupos, intersecados con el $[0, 1)^{2l}$ (debido a que estamos restringiendo el dominio de definición de $\tilde{E}$ al conjunto D).

Para concluir el Teorema 4.18, resta estudiar el conjunto $(\tilde{X}\backslash\tilde{X}^{\mathrm{alg}})(\mathbb{Q})$. Si mostramos que este conjunto es finito, entonces va a resultar que es una unión finita de coclases de grupos (trasladados del $\{0\}$). Por el teorema de Pila-Wilkie, tenemos una cota superior para $(\tilde{X}\backslash\tilde{X}^{\mathrm{alg}})(\mathbb{Q},H)$. Lo que hacemos a continuación es obtener una cota inferior. Puesto que estamos ilustrando la estrategia de la demostración del Teorema 4.18, vamos a asumir que el polinomio $G(x_1,\cdots,x_l)$ está definido sobre un cuerpo de números K y es irreducible sobre este cuerpo [8]. Sea $\mathbf{z}\in\tilde{X}(\mathbb{Q},H)$ con $H>0$. Podemos escribir $\mathbf{z}=(\frac{a_1}{b_1},\cdots,\frac{a_{2l}}{b_{2l}})$ con a_i,b_i enteros, $0\leq a_i<b_i\leq H$ y $\gcd(a_i,b_i)=1$. El punto $\mathbf{z}$ verifica $G(\exp(2\pi i\frac{a_1}{b_1}),\cdots,\exp(2\pi i\frac{a_{2l}}{b_{2l}}))=0$. Consideramos la extensión de K dada por $L:=K(\exp(2\pi i\frac{a_1}{b_1}),\cdots,\exp(2\pi i\frac{a_{2l}}{b_{2l}}))$. Como G tiene coeficientes en K, para cada automorfismo $\sigma:L/K\to L/K$, se verifica $G(\sigma(\tilde{E}(z)))=0$. Dado que $\exp(2\pi i\frac{a_i}{b_i})$ es una raíz b_i-ésima primitiva de la unidad, tenemos[9] $\sigma(\exp(2\pi i\frac{a_i}{b_i}))=\exp(2\pi i\frac{a'}{b_j})$ para algún a' con $0\leq a'<b_i$ y $\gcd(a',b_i)=1$. De la teoría de Galois clásica, sabemos que la órbita de $\exp(2\pi i\frac{a_j}{b_j})$ bajo el grupo de Galois de la extensión L/K tiene cardinal al menos $\frac{\varphi(b_j)}{[K:\mathbb{Q}]}$, con φ la función de Euler, definida como $\varphi(n):=\#(\frac{\mathbb{Z}}{n\mathbb{Z}})^{\times}$. Dado que la función φ verifica la igualdad

$$\varphi(n)=n\prod_{p|n,p\text{ primo}}(1-\frac{1}{p}),$$

tenemos que para todo $\varepsilon<1$, para $n\gg_\varepsilon 1$ se verifica que $\varphi(n)>Cn^{\varepsilon}$. Sea ahora $H>0$ un entero positivo. Si $\#(\tilde{X}\backslash\tilde{X}^{\mathrm{alg}})(\mathbb{Q},H)>\#(\tilde{X}\backslash\tilde{X}^{\mathrm{alg}})(\mathbb{Q},H-1)$, entonces $\#\tilde{X}(\mathbb{Q},H)\geq\frac{1}{[K:\mathbb{Q}]}\varphi(t)$, con lo que para t suficientemente grande, obtendríamos una contradicción con el teorema de Pila Wilkie 4.14 para $\varepsilon<1$. Concluímos que existe H un entero positivo para el que cualquier elemento de $(\tilde{X}\backslash\tilde{X}^{\mathrm{alg}})(\mathbb{Q})$ tiene altura a lo sumo H, es decir, dicho conjunto es finito. □

[8] Se puede ver que el caso general de $G(x_1,\cdots,x_l)\in\mathbb{C}[x_1,\cdots,x_l]$ se reduce al caso que estamos asumiendo. Consultar [**46**]

[9] Estamos usando el siguiente hecho: dado $\alpha\in\overline{\mathbb{Q}}$, si $\sigma:\overline{\mathbb{Q}}/\mathbb{Q}\to\overline{\mathbb{Q}}/\mathbb{Q}$ es un automorfismo de cuerpos, entonces $\sigma(\alpha)$ es una raíz del minimal de α.

Bibliografía

[1] J. Ax, *On Schanuel's conjectures*, Ann. of Math. **93**, 1998, 203-208.

[2] A. Baker, G. Wustholz,*Logarithmic forms and Diophantine geometry*, New mathematical monographs **9**, Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge 2008.

[3] E. Bierstone, P. D. Milman, *Semianalytic and subanalytic sets*, Inst. Hautes Études Sci. Publ. Math. **67** (1988), 5-42.

[4] E. Bombieri, *Counting points and curves on finite fields*, Séminaire Bourbaki, exp. 430, 1972-1973, 234-241.

[5] E. Bombieri, *Le grand crible dans la théorie analytique des nombres*, Second ed., Astérisque **18**, Soc. Math. France, Paris, 1987.

[6] E. Bombieri y J. Pila, *The number of integral points on arcs and ovals*, Duke Math. J. **59** (1989), 337-357.

[7] J. Bochnak, M. Coste, M. Roy, *Real Algebraic Geometry*, Results in Mathematics and Related Areas (3), **36**. Springer-Verlag, Berlin, 1998.

[8] V. Brun, *Über das Goldbasche Gesetz und die Anzahl der Primzahlpaare*, Archiv for Math. og Naturvid., **B34**: 8 (1915), 19 páginas.

[9] V. Brun, *Le crible d'Eratostène et le théorème de Goldbach*, Videnskaps. Skr., Mat.-Naturv. Kl. Kristiana, no. **3** (1920), 36 páginas.

[10] A. Cojocaru, M. R. Murty, *To Sieve Methods and Their Applications*, London Mathematical Society Student Texts **66**, Cambridge University Press, 2005.

[11] H. Davenport, *Analytic Methods for Diophantine Equations and Diophantine Inequalities*, Second ed., Cambridge Mathematical Library, Cambridge University Press, 1963-2005.

[12] J. Denef, L. van den Dries, *p-adic and real subanalytic sets*, Ann. of Math **128**, 1988, 79-138.

[13] H. B. Enderton, *A Mathematical Introduction to Logic*, Hartcourt Science and Technology Company, Academic Press, 2001.

[14] J. B. Friedlander, D. R. Heath-Brown, H. Iwaniec, J. Kaczorocoski. *Analytic Number Theory*, lectures given at the C.I.M.E. Summer School held in Cetrara, Italy, July 11-18-2002, Springer, Foundazione. C.I.M.E.

[15] P. X. Gallagher, *A larger sieve*, Acta Arith. **18** (1971), 77-81.

[16] M. Gromov, *Entropy, homology and semialgebraic geometry*, AStérisque **145-146** (1987), 225-240, Séminarie Bourbaki 1985/86, no. 663.

[17] D. R. Heath-Brown, *The density of rational points on curves and surfaces*, Ann. of Math. **155** (2002), 553-595.

[18] H. A. Helfgott, A. Venkatesh, *How small must ill-distributed sets be?*, Number theory, Essays in honour of Klaus Roth, Cambridge University Press (2009), 224-234.

[19] H. Iwaniec, *Sieve Methods*, Rutgers, spring 1996

[20] V. Jarník, *Über die Gitterpunkte auf konvexen Kurven*, Math. Z. **24** (1926), 500-518.

[21] E. Kowalski, *The large sieve and its applications: arithmetic geometry, random walks and discrete groups*, Cambridge Tracts in Math. **175**, Cambridge University Press, 2008.

[22] A. Kechris, *Classical descriptive set theory*, Springer-Verlag, 1995

[23] L. Kuipers, H. Niederreiter, *Uniform distribution of sequences*, A Wiley-intersciencie publication, 1974.

[24] S. Lang, *Introduction to Trascendental Numbers*, Addison-Wesley Series in Mathematics, Addison-Wesley Publishing company, 1966.

[25] S. Lang, A. Weil, *Numbers of Points of Varieties in Finite Fields*, Amer. J. of Math. vol. **76**, 1954, 819-827.

[26] Y. V. Linnik, *Asymptotic formula in a additive problem of Hardy-Littlewood*, Izv. Akad. Nauk SSSR, Ser. Math. **24** (1960), 629-706.

[27] Y. V. Linnik, *The Large Sieve*, Dokl. Akad. Nauk. SSSR, **30** (1941), 292-4.

[28] A. Macintyre, A. J. Wilkie, *On the decidability of the real exponential field*, in Kreiseliana: About and Around Georg Kreisel, (P. Odifreddi) A. K. Peters. (1996), 441-467

[29] H. Mann, *On linear relations between roots of unity*, Mathematika, **12**:107-117, 1965.

[30] D. A. Marcus, *Number Fields*, Springer, 1977.

[31] D. Marker, *Model Theory: an Introduction*, Graduate texts in Math. **217**, Sringer, 2000

[32] D. Masser, U. Zannier, *Torsion anomalous points and families of elliptci curves*, C. R. Math. Acad. Sci. Paris **346**, 2008, 491-494.

[33] J. V. Matiyasevich, *The Diophantineness of enumerable sets*, Dokl. Akad. Nauk. SSSR **191** (1970), 279-282; English translation in Soviet Math. Dokl. **11** (1970), 354-358.

[34] J. Merlin, *Sur Quelques théorèmes d'Arithmetique et un énoncé qui les contient*, C. R. Acad. Sci. Paris, **153** (1911), 516-518.

[35] C. Miller, L. van den Dries, *On the real exponential field with restricted analytic functions*, Israel J. Math. **85**, 1994, 19-56.

[36] C. Miller, L. van den Dries, *Geometric categories an O-minimal structures*, Duke Math. J. **84** (1996), 497-540.

[37] K. Ireland, M. Rosen, *A Classical Introduction to Modern Number Theory*, Graduate Texts in Maths. **84**, 1982

[38] J. Pila, *Geometric and arithmetic postulation of the exponential function*, J. Austral. Math. Soc. Ser. A **54** (1993), 111-127

[39] J. Pila, *Integer points on the dilation of a subanalytic surface*, Q. J. Math. **55** (2004), 207-223.

[40] J. Pila, *O-minimality and the André-Oort conjecture for* $\mathbb{C}^n$, Ann. of Math. **173** (2011), 1779-1840.

[41] J. Pila, *Geometric postulation of a smooth function and the number of the number of rational points* , Duke Math. J. **63** (1991), 449-463.

[42] J. Pila y A. J. Wilkie, *The rational points of a definable set*, Duke Math. J. **133** (2006), 591-616.

[43] J. Pila y U. Zannier, *Rational points in periodic analytic sets and the Manin-Mumford conjecture*, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl. 19 (2) (2008), 149-162.

[44] G. Pólya, *On the Mean Value Theorem corresponding to a given Linear Homogeneous Differential Equation*, Trans. Amer. Math. Soc. **24** (1922), 312-324.
[45] J. -P. Rolin, P. Speissegger, A. J. Wilkie, *Quasianalytic Denjoy-Carleman clases and o-minimality*, J. Amer. Math. Soc. **16** (2003), 751-777.
[46] T. Scanlon, *Counting special points: logic, diophantine geometry and trascendence theory*, Current Events Bulletin, 8 January 2011
[47] T. Scanlon, *A proof of the André-Oort conjecture via mathematical logic*, Séminaire Bourbaki, 2010-2011, no. 63.
[48] W. M. Schmidt, *Integer Points on Curves and Surfaces*, Monatsh. Math. **99** (1985), 45-72.
[49] E. Selmer, *The diophantine equation* $ax^3 + by^3 + cz^3 = 0$, Acta Mathematica **85** (1951) 203-362,
[50] J. H. Silverman, *The Arithmetic of Elliptic Curves*, Second ed., Graduate Texts in Mathematics **106**, Springer, 2009.
[51] S. A. Stepanov, *On the numbers of points of a hiperelliptic curve over a finite prime field*, Izv. Akad. Nauk SSSR, Ser. Mat. **33**, 1969, 1103-1114.
[52] H. P. F. Swinnerton-Dyer, *The number of lattice points on a convex curve*, J. Number Theory **6** (1974), 128-135.
[53] A. Tarski, *A Decision Method for Elementary Algebra and Geometry*, RAND Corporation, Santa Monica, California, 1948.
[54] L. van den Dries , *Remarks on Tarski's problem concerning* $(\mathbb{R}, +, \cdot, \exp)$, Logic Colloquium **82**, G. Lolli, G. Longo and A. Marcja, eds. ,Nort-Holland, 1984, 97-121.
[55] L. van den Dries, *Tame topology and o-minimal structures*, London Mathematical Society Lecture Note Series **248**, Cambridge University Press, Cambridge, 1998.
[56] A. J. Van den Poorten, *Transcendental entire functions mapping every algebraica number field into itself*, J. Austral. Math. Soc. **8** (1968), 192-193.
[57] R. C. Vaughan, *The Hardy-Littlewood method*, Second ed., Cambridge Tracts in Mathematics **125**, Cambridge University Press, 1997.
[58] M. Waldschmidt, *On the transcendence methods of Gelfond and Schneider in Several Variables*, New Advances in Transcendence Theory, Cambridge University Press, 1988.
[59] M. Walsh, *The inverse sieve problem in high dimensions*, Duke Math. J. **161** (2010) 2001-2022.
[60] M. Walsh, *The algebraicity of ill-distributed sets*, prepint on arxiv.
[61] A. J. Wilkie, *Model completeness results for expansions of the ordered field of real numbers by restricted Pfaffian functions and the exponential function*, J. Amer. Soc. **9**, 1996, 1051-1094.

Índex