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

Análisis probabilístico de algoritmos y problemas combinatorios sobre cuerpos finitos

Tesis presentada para optar al título de Doctor de la Universidad de Buenos Aires en el área Ciencias Matemáticas

Mariana Valeria Pérez

Director de tesis: Guillermo Matera
Director Asistente: Eda Cesaratto
Consejero de estudios: Pablo Solernó

Lugar de trabajo: Universidad Nacional de General Sarmiento. Instituto del Desarrollo Humano.

Buenos Aires, 2016

Análisis probabilístico de algoritmos y problemas combinatorios sobre cuerpos finitos

Resumen

En esta tesis analizamos la complejidad en promedio de dos algoritmos probabilísticos. Uno de ellos calcula puntos $\mathbb{F}_q$–racionales de hipersuperficies definidas sobre el cuerpo finito $\mathbb{F}_q$ de q elementos en base a una estrategia de "búsqueda en bandas verticales". El otro es el algoritmo clásico de factorización de polinomios univariados con coeficientes en $\mathbb{F}_q$. En este caso nos interesa su comportamiento cuando se aplica a familias de polinomios cuyos coeficientes satisfacen ciertas relaciones lineales.

A fin de realizar dichos análisis, proporcionamos estimaciones explícitas del promedio del cardinal del conjunto de valores y la distribución de patrones de factorización en familias lineales de polinomios sobre cuerpos finitos. Los resultados expuestos, que mejoran los existentes en la literatura sobre el tema, se basan en un nuevo enfoque, que reduce estas cuestiones combinatorias a la estimación del número de puntos $\mathbb{F}_q$–racionales de ciertas intersecciones completas singulares. Por tal motivo, parte de esta tesis se centra en el estudio de ciertas propiedades geométricas de dichas variedades.

Palabras clave. Complejidad en promedio, algoritmos probabilísticos, familias lineales de polinomios con coeficientes en un cuerpo finito, conjunto de valores, patrones de factorización, intersecciones completas singulares, lugar singular, puntos racionales.

Probabilistic analysis of algorithms and combinatorial problems over finite fields

Abstract

In this thesis we analyze the average-case complexity of two probabilistic algorithms. The first one computes $\mathbb{F}_q$–rational points of an hysuperface defined over the finite field $\mathbb{F}_q$ of q elements based on a search strategy in "vertical strips". The second one is the classical factorization algorithm for univariate polynomials with coefficients in $\mathbb{F}_q$. In this case we are interested in the behavior of the algorithm applied to families of polynomials whose coefficients satisfy certain linear relations.

For this purpose, we obtain explicit estimates on the average cardinality of the value set and on the distribution of factorization patterns on linear families of polynomials over a finite field. The above results, which improve the existing results in the literature, were obtained with a new approach that reduces these combinatorial issues to estimating the number of $\mathbb{F}_q$–rational points of certain singular complete intersections. For this reason, part of this thesis focuses on the study of certain geometric properties of these varieties.

Key words Average–case complexity, probabilistic algorithms, linear families of polynomials over finite fields, value sets, factorization patterns, singular complete intersections, singular locus, $\mathbb{F}_q$–rational points.

Agradecimientos

A Guillermo, por confiar en que podía emprender este camino y terminarlo, por todo lo que me enseñó, por su paciencia, por su compromiso, sin lugar a dudas sin él este momento no hubiera llegado.

A Pablo, por aceptar ser mi consejero de estudios.

A Carlos D'Andrea, Gabriela Jerónimo y María Inés Pacharoni, por aceptar ser los jurados de esta tesis.

Quiero agradecer a mi amiga, ya hermana, Melina, por las horas de estudio a mi lado, las charlas interminables, por la compañía en todo momento, el miedo al "no puedo" a su lado disminuye y por todo lo que nos falta todavía aprender juntas.

A Eda, por su paciencia, sobre todo al principio, cuando todo costaba, por sus explicaciones sobre algoritmos y complejidad, por confiar que podía lograr entender estos temas.

A Deborah, por sus explicaciones sobre cuestiones administrativas, por su paciencia y ayudarme con todos los trámites.

A Nardo, por estos años compartidos.

A tantos amigos que conocí en este camino, a Luciano, a Anita, a Ezequiel, a Martín.

A los amigos, hermanos de siempre, a Marina, a Ezequiel, a Daniela, a Jessi, por aguantar mis faltas de tiempo y apoyarme.

A mi querido Padre Max, que desde el cielo me guía y confía en mí.

A Gustavo, por sus palabras de apoyo siempre.

A mi hermana de toda la vida, Roxi, por su cariño y sus oraciones.

A mi familia, sin ellos nada hubiera sido posible, a los que están a mi lado, a mamá, a mi hermano Ale, a mi abu querida; y a los que no están, a mi querido papá.

A mi nueva familia, la que Dios me puso en el camino, Nati, Bete, Amalucha y María.

A Mauri, por acompañarme, sostenerme, amarme, elegirme, enseñarme, por tanto recorrido y por lo que todavía nos falta recorrer. Como dice siempre, lo mejor está por venir.

Índice general

Índice general **9**

1. Introducción **11**
1.1. Antecedentes 12
1.1.1. Búsqueda de puntos $\mathbb{F}_q$–racionales en hipersuperficies 12
1.1.2. Factorización de polinomios univariados 16
1.1.3. Problemas combinatorios 19
1.2. Resultados obtenidos y organización del trabajo 22

2. Preliminares **31**
2.1. Definiciones y resultados básicos de geometría algebraica 31
2.1.1. Intersecciones completas 35
2.1.2. El grado de una variedad 36
2.2. Puntos $\mathbb{F}_q$–racionales de $\mathbb{F}_q$–variedades 37
2.2.1. Algunas cotas superiores 37
2.2.2. Estimaciones del número de puntos $\mathbb{F}_q$–racionales 38

3. El lugar discriminante y variedades de incidencia asociados a familias lineales **43**
3.1. Irreducibilidad del discriminante 43
3.2. Estimaciones para variedades de incidencia 48
3.2.1. Aspectos geométricos 49
3.2.2. La geometría de la clausura proyectiva 57
3.2.3. El número de puntos $\mathbb{F}_q$–racionales 59

4. Intersecciones completas dadas por polinomios simétricos **61**
4.1. Estimaciones para intersecciones completas simétricas 61
4.1.1. Aspectos geométricos 63
4.1.2. La geometría de la clausura proyectiva 66
4.1.3. El número de puntos $\mathbb{F}_q$–racionales 69
4.2. Estimaciones para ciertas intersecciones completas asociadas a familias lineales 71
4.2.1. Aspectos geométricos 71
4.2.2. La geometría de la clausura proyectiva 78
4.2.3. El número de puntos $\mathbb{F}_q$–racionales 80

5. Conjunto de valores en familias lineales 83
5.1. El problema 83
5.2. El promedio del cardinal del conjunto de valores 85
5.2.1. Una reducción combinatoria 85
5.2.2. Un enfoque geométrico 86
5.3. Una estimación del promedio 88
5.3.1. Polinomios con coeficientes prescriptos 92
5.4. Conjunto de valores para familias no lineales 92
5.5. El segundo momento del conjunto de valores 97

6. Conjunto de valores con coeficientes prescriptos 101
6.1. El conjunto de valores en términos de ceros de polinomios simétricos . 102
6.2. Una estimación para el promedio del conjunto de valores 105

7. La distribución de patrones de factorización 111
7.1. Patrones de factorización y raíces 112
7.2. El número de polinomios con patrón de factorización dado 116
7.2.1. Polinomios con coeficientes prescriptos y aplicaciones 120

8. Búsqueda de puntos $\mathbb{F}_q$–racionales en hipersuperficies 123
8.1. Algoritmo BBV para hipersuperficies 123
8.2. Probabilidad de éxito en las primeras 2 bandas verticales 126
8.2.1. Probabilidad de éxito en la primera banda vertical 126
8.2.2. Probabilidad de éxito en la segunda banda vertical 131
8.3. Probabilidad de éxito en más bandas verticales 135
8.3.1. Imagen de la proyección que definen s bandas verticales 137
8.3.2. La probabilidad de s búsquedas en términos de cardinales de conjuntos de valores 142
8.4. Análisis probabilístico del Algoritmo BBV 148
8.4.1. Distribución de probabilidades del número de búsquedas . . . 149
8.4.2. Complejidad en promedio 151
8.5. Simulaciones sobre el número de bandas verticales 156

9. Distribución de las salidas del Algoritmo BBV 161
9.1. Sobre el número de bandas verticales 162
9.2. Una cota inferior para la entropía 167

10. Algoritmo de factorización en familias 173
10.1. El algoritmo clásico de factorización y preliminares 173
10.2. Eliminación de factores repetidos 176
10.3. Factorización en distintos grados 179
10.4. Factorización en grados iguales 185
10.5. Costo en promedio del algoritmo clásico 192

Bibliografía 195

Capítulo 1

Introducción

En todo este trabajo denotaremos con $\mathbb{F}_q$ el cuerpo finito de $q := p^l$ elementos, donde p es un número primo, y con $\overline{\mathbb{F}}_q$ su clausura algebraica. En esta tesis vamos a estudiar los siguientes problemas:

1. Diseñar y analizar la complejidad en promedio de un algoritmo probabilístico que calcula puntos $\mathbb{F}_q$–racionales de hipersuperficies definidas sobre $\mathbb{F}_q$, o equivalentemente, soluciones $\mathbb{F}_q$–racionales de polinomios multivariados con coeficientes en $\mathbb{F}_q$, basado en la estrategia de "búsqueda en bandas verticales", es decir, búsquedas sobre líneas paralelas en una dirección dada.

2. Analizar la complejidad en promedio del algoritmo clásico de factorización para familias lineales de polinomios mónicos univariados con coeficientes en $\mathbb{F}_q$.

3. Estudiar dos problemas combinatorios clásicos sobre cuerpos finitos relevantes para los dos primeros puntos: el comportamiento promedio del cardinal de la imagen o "conjunto de valores", y la distribución de los patrones de factorización, en familias lineales de polinomios mónicos univariados con coeficientes en $\mathbb{F}_q$.

Más precisamente, vamos a estudiar los problemas combinatorios mencionados desde un enfoque geométrico, es decir, traduciéndolos al de determinar la cantidad de puntos $\mathbb{F}_q$–racionales de ciertas variedades algebraicas singulares. Las características geométricas de estas variedades nos permitirán dar estimaciones explícitas de la cantidad de puntos $\mathbb{F}_q$–racionales de las mismas. De esta manera, dichos resultados nos servirán para:

(a) Obtener estimaciones explícitas del promedio del cardinal del "conjunto de valores" de una familia lineal de polinomios univariados de grado d con coeficientes en $\mathbb{F}_q$, en el caso en que d sea menor que q, con un término de error explícito en términos de d y q.

(b) Obtener estimaciones explícitas del número de elementos de una familia lineal de polinomios mónicos univariados de grado d con coeficientes en $\mathbb{F}_q$ y con

un patrón de factorización dado, en el caso en que d sea menor que q, con un término de error explícito en términos de ciertos parámetros referidos a la familia y su patrón de factorización.

Estas cotas superiores nos permitirán dar cuenta de las siguientes cuestiones referidas a los dos primeros problemas algorítmicos:

(a) Determinar el comportamiento asintótico de la distribución de probabilidad del número de "bandas verticales" que deben ser generadas hasta que dicho algoritmo encuentre un punto $\mathbb{F}_q$–racional de una hipersuperficie dada.

(b) Analizar la complejidad en promedio del algoritmo de búsquedas en bandas verticales.

(c) Analizar la distribución en promedio de la salida de tal algoritmo mediante el concepto de entropía de Shannon.

(d) Analizar el costo en promedio de las tres etapas fundamentales del algoritmo clásico de factorización en familias lineales: la eliminación de factores repetidos, la factorización en distintos grados y la factorización en igual grado.

Salvo mención explícita de lo contrario, los resultados de esta tesis son originales y se encuentran en los siguientes artículos: [CMPP14], [MPP14], [MPP16a], [CMP15b] y [MPP16b]. Los tres primeros ya se encuentran publicados, en tanto que el cuarto se encuentra en proceso de publicación y el último fue sometido a publicación. Por otro lado, los resultados del Capítulo 4, Sección 1 y del Capítulo 10 no se encuentran todavía publicados.

1.1. Antecedentes

1.1.1. Búsqueda de puntos $\mathbb{F}_q$–racionales en hipersuperficies

La búsqueda de puntos $\mathbb{F}_q$–racionales (es decir, con coordenadas en $\mathbb{F}_q$) en hipersuperficies definidas sobre $\mathbb{F}_q$ es un problema clásico de la geometría algebraica, con importantes aplicaciones en criptografía, teoría de códigos y álgebra computacional, entre otras [KY08]. Muchos problemas se reducen a encontrar puntos $\mathbb{F}_q$–racionales de hipersuperficies definidas sobre $\mathbb{F}_q$. Por ejemplo, en teoría de códigos y criptografía, podemos mencionar que encontrar puntos $\mathbb{F}_q$–racionales de hipersuperficies sirve para clasificar los códigos cíclicos y funciones casi perfectamente no lineales [HM11], o en la decodificación de máxima verosimilitud para códigos de Reed Solomon [GGL06]. También la búsqueda de puntos $\mathbb{F}_q$–racionales de hipersuperficies surge de manera natural en el problema de detectar espacios de matrices no singulares (ver por ejemplo el trabajo de L. Lovász [Lov89] o también el trabajo [BFS99]). Este problema está relacionado con el del "testeo de identidades polinomiales", es decir, determinar cuando un polinomio multivariado es idénticamente cero.

Otras motivaciones surgen de la búsqueda de puntos $\mathbb{F}_q$–racionales de sistemas polinomiales subdeterminados con coeficientes en $\mathbb{F}_q$: existen algoritmos que reducen la resolución de estos sistemas a encontrar puntos $\mathbb{F}_q$–racionales de hipersuperficies definidas sobre $\mathbb{F}_q$ (podemos citar los trabajos [HW99] o [CM06a]). Solo para mencionar la técnica utilizada en el último trabajo, podemos decir que, a partir del sistema original se obtiene una hipersuperficie birracionalmente equivalente a la variedad original, contenida en un espacio ambiente adecuado (existe un morfismo racional entre la variedad y la hipersuperficie, y dicho morfismo tiene inversa que resulta también un morfismo racional). Esta hipersuperficie se obtiene como la imagen de la variedad original por una proyección lineal genérica. Por último, se calcula un punto $\mathbb{F}_q$–racional de dicha hipersuperficie y se "levanta" a un punto de la variedad original.

Existen familias de algoritmos que calculan puntos $\mathbb{F}_q$–racionales de hipersuperficies definidas sobre $\mathbb{F}_q$ siempre que éstas posean ciertas propiedades geométricas, como por ejemplo la de ser absolutamente irreducible. Cabe mencionar que una hipersuperficie se dice absolutamente irreducible si cada polinomio de grado mínimo que la define es absolutamente irreducible, es decir, es irreducible sobre $\overline{\mathbb{F}}_q$. De esta información geométrica es posible obtener estimaciones sobre el número de puntos $\mathbb{F}_q$–racionales de dicha hipersuperficie en términos de su grado y la cantidad de elementos del cuerpo que, a su vez, puede usarse para el diseño de algoritmos probabilísticos de búsqueda de puntos $\mathbb{F}_q$–racionales (ver, por ejemplo, [Mat10]).

Una familia de algoritmos de este tipo se basa en la estrategia de "búsqueda en bandas verticales", esto es, dado un polinomio F en $\mathbb{F}_q[X_1, \ldots, X_r]$ de grado a lo sumo d, se trata de determinar elementos $a_1, \ldots, a_{r-1}$ de dicho cuerpo de modo tal que el polinomio $F(a_1, \ldots, a_{r-1}, X_r)$ tenga raíces en $\mathbb{F}_q$. Para el caso $r = 2$ esta estrategia fue analizada en profundidad por J. von zur Gathen y colaboradores en [vzGSS03]. En dicho trabajo, la determinación del elemento a_1 de $\mathbb{F}_q$ se realiza de manera aleatoria, aplicando, en el caso en que la curva definida por dicho polinomio tenga componentes absolutamente irreducibles definidas sobre $\mathbb{F}_q$, la conocida estimación de A. Weil sobre la cantidad de puntos $\mathbb{F}_q$–racionales en curvas absolutamente irreducibles sobre cuerpos finitos (ver [Wei48]). Estos autores proponen un algoritmo probabilístico de "búsqueda en bandas verticales" que calcula puntos $\mathbb{F}_q$–racionales de una curva plana de grado d definida sobre $\mathbb{F}_q$ de forma tal que cada punto $\mathbb{F}_q$–racional tenga la misma probabilidad de ser calculado (ver [vzGSS03, Algorithm 2.1]). Prueban que si $q \geq 16d^4$, el algoritmo encuentra un punto $\mathbb{F}_q$–racional con probabilidad al menos $1/2d$. De esto, se deduce que, con al menos d elecciones aleatorias es posible encontrar un punto $\mathbb{F}_q$–racional en una banda vertical con probabilidad al menos $1/2$. De todas formas, cabe destacar que en este trabajo se analiza el costo del peor caso, es decir, la cantidad máxima de operaciones aritméticas en $\mathbb{F}_q$ que realiza el algoritmo hasta encontrar un punto $\mathbb{F}_q$–racional. Más precisamente, los autores prueban que el algoritmo encuentra un punto $\mathbb{F}_q$–racional de una curva plana de grado d, con componentes absolutamente irreducibles definidas sobre $\mathbb{F}_q$ y sin líneas verticales, con $\mathcal{O}(M(d)\log(dq))$ operaciones aritméticas en $\mathbb{F}_q$, donde $M(d) := d\log d\log\log d$ es el costo de la multiplicación rápida entre dos polinomios univariados de grado a lo sumo d con coeficientes en $\mathbb{F}_q$ (ver [vzGG99, Chapter 8]).

Observamos que la notación $\log N$ indica el logaritmo de N en base 2. En el mismo artículo proponen un algoritmo probabilístico que usa $(d \log q)^{\mathcal{O}(1)}$ operaciones aritméticas en $\mathbb{F}_q$ y trata con el caso relativamente $\mathbb{F}_q$–irreducible. Cabe mencionar que un polinomio definido sobre $\mathbb{F}_q$ se dice relativamente $\mathbb{F}_q$–irreducible si ninguno de sus factores irreducibles sobre $\mathbb{F}_q$ es absolutamente irreducible. A su vez, una curva plana definida sobre $\mathbb{F}_q$ se dice relativamente $\mathbb{F}_q$–irreducible si cada polinomio de grado mínimo con coeficientes en $\mathbb{F}_q$ que la define es relativamente $\mathbb{F}_q$–irreducible. Este algoritmo determina cuando el polinomio de entrada es relativamente $\mathbb{F}_q$–irreducible y, si es así, encuentra todos los ceros $\mathbb{F}_q$–racionales del mismo.

Siguiendo estas ideas, G. Matera en [Mat10] propone un algoritmo para la búsqueda de puntos $\mathbb{F}_q$–racionales sobre curvas planas generales de grado d definidas sobre $\mathbb{F}_q$. Observa que, si la curva de entrada o alguna componente irreducible sobre $\mathbb{F}_q$ resulta absolutamente irreducible, se aplica un algoritmo similar a [vzGSS03, Algorithm 2.1] para curvas planas absolutamente irreducibles. En tal caso, prueba que si $d \geq 2$ y $q > 16d^4$, entonces el algoritmo calcula un cero $\mathbb{F}_q$–racional con una cantidad máxima de $\mathcal{O}(d\mathcal{U}(d)\log(dq))$ operaciones aritméticas en $\mathbb{F}_q$. En este caso $\mathcal{U}(d) := M(d)\log d$ representa la cantidad de operaciones máximas en $\mathbb{F}_q$ para el cálculo del máximo común divisor entre dos polinomios de grado a lo sumo d, utilizando la multiplicación rápida (ver, por ejemplo, [vzGG99, Chapter 11]). En cambio, si ninguna de las componentes irreducibles sobre $\mathbb{F}_q$ de la curva en cuestión es absolutamente irreducible, se aplica el algoritmo propuesto en [vzGSS03] para curvas relativamente $\mathbb{F}_q$–irreducibles, que requiere $\mathcal{O}(d\mathcal{U}(d^2)\log(dq))$ operaciones aritméticas en $\mathbb{F}_q$. Cabe agregar que en [vzG08] se prueba que la probabilidad de que un polinomio bivariado con coeficientes en $\mathbb{F}_q$ de grado $d \geq 2$ sea reducible es menor o igual que $q^{1-d}(1 + 6q^{-1})$, y la probabilidad de que un polinomio bivariado con coeficientes en $\mathbb{F}_q$ de grado d sea relativamente irreducible es menor o igual que $q^{-d^2/4}$. Los resultados anteriores muestran que el número esperado de operaciones aritméticas en $\mathbb{F}_q$ para calcular un cero $\mathbb{F}_q$–racional de cualquier polinomio bivariado con coeficientes en $\mathbb{F}_q$ de grado d es asintóticamente el del caso absolutamente irreducible.

Más aún, A. Cafure y G. Matera generalizan esta estrategia de búsqueda para el caso de polinomios absolutamente irreducibles en r variables. Más precisamente, en [CM06a] (ver también [Mat10, Section 3]) diseñan y analizan un algoritmo probabilístico que calcula puntos $\mathbb{F}_q$–racionales de una hipersuperficie absolutamente irreducible de grado d definida sobre $\mathbb{F}_q$. A dicho algoritmo lo llaman Algoritmo de búsqueda en secciones planas. Para calcular un punto $\mathbb{F}_q$–racional de una hipersuperficie definida sobre $\mathbb{F}_q$ reducen el problema al del cálculo de puntos $\mathbb{F}_q$–racionales de una curva plana absolutamente irreducible definida sobre dicho cuerpo. La curva plana que consideran es una sección plana de la hipersuperficie en cuestión, es decir, es la intersección de la hipersuperficie con una variedad lineal de dimensión 2. Para extender la estrategia de búsqueda en bandas verticales a hipersuperficies necesitan utilizar estimaciones explícitas del tipo Lang-Weil sobre la cantidad de puntos $\mathbb{F}_q$–racionales de hipersuperficies absolutamente irreducibles como las que se encuentran en los trabajos [GL02a], [CM06b], [CMP15a] y [MPP16a]. Los pasos más importantes del algoritmo propuesto son, primero, asegurar la existencia de una sec-

ción plana absolutamente irreducible definida sobre $\mathbb{F}_q$. Sobre este punto prueban que si $q > 16d^4$ es posible encontrar una sección plana con estas características con probabilidad al menos $\frac{7}{8}$. Otro punto importante es encontrar un cero $\mathbb{F}_q$–racional en una banda vertical adecuada de dicha sección plana. En tal sentido, los autores prueban que con al menos d elecciones aleatorias es posible encontrar un cero $\mathbb{F}_q$–racional en una banda vertical con probabilidad al menos $\frac{1}{2}$, con un análisis inspirado en el de [vzGSS03]. Finalmente, prueban que si $F \in \mathbb{F}_q[X_1, \ldots, X_r]$ es absolutamente irreducible de grado $d \geq 2$, $H := \{F = 0\} \subset \overline{\mathbb{F}}_q^r$ es la hipersuperficie definida por F y $q > 16d^4$, entonces el algoritmo de búsqueda en secciones planas calcula un punto $\mathbb{F}_q$–racional de H con $\mathcal{O}(\mathcal{L}d^2 + d\mathcal{U}(d)\log(dq))$ operaciones aritméticas en $\mathbb{F}_q$, donde $\mathcal{L}$ es el número de operaciones aritméticas en $\mathbb{F}_q$ que se requiere para evaluar F.

Para extender la búsqueda a hipersuperficies generales definidas sobre $\mathbb{F}_q$, en [Mat10] Matera propone un algoritmo probabilístico que calcula ceros $\mathbb{F}_q$–racionales de un polinomio multivariado arbitrario con coeficientes en $\mathbb{F}_q$. En este algoritmo, si el polinomio o alguno de sus factores irreducibles sobre $\mathbb{F}_q$ resulta absolutamente irreducible, se utiliza el algoritmo propuesto en [CM06a] para polinomios absolutamente irreducibles. Si no, se calcula el discriminante del polinomio y se aplica el algoritmo a él o a un factor absolutamente irreducible del mismo. Si ningún factor de dicho discriminante es absolutamente irreducible, el proceso continúa con este discriminante. Se realiza un análisis del peor caso de este algoritmo de búsqueda y se prueba que se comporta bien para el caso absolutamente irreducible. El peor caso se da cuando el polinomio no es absolutamente irreducible. Por su parte, J. von zur Gathen y colaboradores prueban en [vzGVZ13] que la probabilidad de que un polinomio multivariado sea absolutamente irreducible es alta. Así, Matera observa que la probabilidad de que el algoritmo de búsqueda en bandas verticales para hipersuperficies que propone necesite reducir la búsqueda a polinomios discriminantes es baja.

Por otro lado, cabe mencionar que el número promedio $N(F)$ de ceros $\mathbb{F}_q$–racionales de polinomios F en $\mathbb{F}_q[X_1, \ldots, X_r]$ y de grado a lo sumo d es q^{r-1} (ver, por ejemplo, [LN83, Theorem 6.16]). Además, si el polinomio en consideración es absolutamente irreducible, entonces existen cotas superiores explícitas sobre la desviación $|N(F) - q^{r-1}|$ (ver [CM06a]). Estas ideas sugieren que se puede extender la estrategia de búsqueda en bandas verticales al caso de polinomios en varias variables sin requerir hipótesis de absoluta irreducibilidad ni reducir esta búsqueda a curvas planas. Más precisamente, como el número esperado de ceros de F es igual al número de elementos de $\mathbb{F}_q^{r-1}$, dado un elemento $\boldsymbol{a}_1 \in \mathbb{F}_q^{r-1}$, se puede tratar de encontrar un cero de F de la forma $(\boldsymbol{a}_1, x_r)$, con $x_r \in \mathbb{F}_q$, o equivalentemente, un cero $\mathbb{F}_q$–racional de $F(\boldsymbol{a}_1, X_r)$. Si este polinomio no tiene ceros $\mathbb{F}_q$–racionales, dado $\boldsymbol{a}_2 \in \mathbb{F}_q^{r-1}$, se determina si el polinomio $F(\boldsymbol{a}_2, X_r)$ tiene un cero en $\mathbb{F}_q$, y así se sigue hasta encontrar un cero de F en $\mathbb{F}_q^r$.

El algoritmo correspondiente, que denominamos BBV (búsqueda en bandas verticales), funciona de la siguiente manera: dado un polinomio $F \in \mathbb{F}_q[X_1, \ldots, X_r]$ de entrada de grado a lo sumo d, genera progresivamente una sucesión $(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{q^{r-1}})$ de $\mathbb{F}_q^{r-1}$, evalúa a F sucesivamente en dichos puntos $\boldsymbol{a}_i$, busca raíces del polinomio univariado $F(\boldsymbol{a}_i, X_r)$, y termina cuando encuentra una raíz de alguno de esos

polinomios. Dado que el análisis de la complejidad del peor caso es poco significativo, como se comprueba con la experimentación numérica que realizamos, vamos a realizar un análisis de la complejidad en promedio del mismo. Este análisis se inscribe en la tradición del análisis probabilístico de algoritmos, popularizado por D. Knuth (ver [Knu98a]), que propone definir una probabilidad sobre el conjunto de entradas y considerar el parámetro a estimar como una variable aleatoria. En esta dirección, consideramos la probabilidad uniforme sobre el conjunto de todos los polinomios multivariados con coeficientes en $\mathbb{F}_q$ y de grado a lo sumo d y sobre todas las posibles elecciones de bandas verticales, y estudiamos el parámetro que determina el costo de este algoritmo de búsqueda en bandas verticales, es decir, la variable aleatoria X que cuenta la cantidad de operaciones aritméticas en $\mathbb{F}_q$ que realiza el algoritmo hasta encontrar un cero $\mathbb{F}_q$–racional, para un polinomio dado y una elección de bandas verticales. Vamos a demostrar que el costo en promedio se relaciona fuertemente con la cantidad de polinomios univariados de grado a lo sumo d, con $d < q$, y con ciertos coeficientes consecutivos prefijados, que tienen al menos un cero $\mathbb{F}_q$–racional. Cabe mencionar que esta última cantidad tiene una formulación equivalente en términos del promedio del cardinal del conjunto de valores de los polinomios univariados de grado d con ciertos coeficientes prefijados.

Asimismo, una vez que se obtiene una banda vertical con ceros $\mathbb{F}_q$–racionales del polinomio de entrada, es necesario calcular una raíz en $\mathbb{F}_q$ del correspondiente polinomio univariado, para lo cual será menester obtener algún tipo de factorización del mismo. Dado que el análisis probabilístico de los algoritmos de factorización se basa en resultados sobre la distribución de patrones de factorización, para el análisis probabilístico de esta última etapa del algoritmo necesitamos estudiar la distribución de patrones de factorización en polinomios con coeficientes prescriptos.

1.1.2. Factorización de polinomios univariados

En lo que sigue vamos a notar con $\mathbb{F}_q[T]_d$ el conjunto de todos los polinomios mónicos de grado d y con coeficientes en $\mathbb{F}_q$. Comenzamos definiendo el problema: dado un polinomio $f \in \mathbb{F}_q[T]_d$, el objetivo es encontrar la factorización completa $f = f_1^{e_1} \dots f_r^{e_r}$, donde $f_1, \dots, f_r$ son polinomios mónicos, irreducibles, distintos dos a dos y con coeficientes en $\mathbb{F}_q$, y $e_1, \dots e_r$ son números estrictamente positivos.

La factorización de polinomios univariados sobre cuerpos finitos es un problema fundamental, con aplicaciones, por ejemplo, en la factorización de polinomios sobre enteros [Zas69, Col79, LLL82, Knu98a], en la criptografía [CR88, Odl85, Len91], en la teoría de números [Buc90] y en la teoría de códigos [Ber68]. Se necesita factorizar polinomios sobre cuerpos finitos, por ejemplo, en la búsqueda de descomposiciones en fracciones parciales (ver, por ejemplo, [vzGG99, Chapter 5]), en la construcción de códigos de redundancia cíclica y de códigos BCH [Ber68, MS77, vL92], en el cálculo del número de puntos sobre curvas elípticas [Buc90] y en el diseño de sistemas criptográficos de clave pública [CR88, Odl85, Len91]. En particular, la factorización de un polinomio aleatorio sobre un cuerpo finito es necesario en el método del cálculo del índice aleatorio para calcular logaritmos discretos sobre cuerpos finitos

(ver [Odl85]). En nuestro caso, necesitamos la factorización de polinomios para el análisis de la complejidad en promedio del algoritmo BBV para hipersuperficies. Observamos que, a medida que este algoritmo genera sucesivas bandas verticales hasta encontrar un cero $\mathbb{F}_q$–racional del polinomio de entrada, se obtienen familias de polinomios univariados con cada vez más coeficientes prefijados. Con el fin de obtener ceros $\mathbb{F}_q$–racionales de estos polinomios necesitamos una factorización parcial de los mismos.

Existen numerosos algoritmos eficientes para factorizar polinomios sobre cuerpos finitos (ver, por ejemplo [vzGG99, Chapter 14], [Sho05, Chapter 21] o [Shp99, Chapter 1]). Los trabajos pioneros se deben a E.R. Berlekamp (ver [Ber67], [Ber68] y [Ber70]), H. Zassenhaus (ver [Zas69]), y D. G. Cantor y Zassenhaus [CZ81].

Muchos de estos algoritmos proceden en tres pasos: la eliminación de factores repetidos (ERF), la factorización en distintos grados (DDF) y la factorización de igual grado (EDF). Una familia de algoritmos que contienen estos tres pasos se conoce con el nombre de **algoritmo clásico de factorización**.

- **La eliminación de factores repetidos (ERF)** reemplaza el polinomio $f = f_1^{e_1} \dots f_r^{e_r}$ de entrada por un polinomio libre de cuadrados que contiene todos los factores irreducibles del polinomio original f pero con exponente 1, es decir,

$$f \longrightarrow a := \prod_{j=1}^{r} f_j.$$

- **La factorización en distintos grados (DDF)** descompone un polinomio libre de cuadrados en polinomios cuyos factores irreducibles tienen todos el mismo grado, es decir,

$$a \longrightarrow b(1) \dots b(s), \text{ donde } b(k) := \prod_{\deg(f_j)=k} f_j.$$

- **La factorización de igual grado (EDF)** descompone completamente un polinomio cuyo factores irreducibles tienen el mismo grado, es decir,

$$b(k) \longrightarrow b(k,1) \dots b(k,l_k), \text{ donde } \deg(b(k,j)) = k.$$

Los algoritmos para el primer y segundo paso son determinísticos, mientras que los algoritmos más rápidos para el tercer paso son probabilísticos (ver, por ejemplo [vzGG99, Chapter 14] o [Sho05, Chapter 21]). Cabe mencionar que la factorización en grados distintos aparece en los trabajos de Zassenhaus [Zas69], H. Kempfert [Kem69], Berlekamp [Ber70], Cantor–Zassenhaus [CZ81] y en Knuth [Knu98a], entre otros. El primer trabajo publicado sobre este algoritmo se debe a E. Galois en el año 1830; en dicho trabajo el autor se olvidó de decir que el máximo común divisor debe eliminarse después del siguiente paso y que debe contemplarse la posibilidad de que la derivada del polinomio de entrada sea cero (ver [Gal46]). Más adelante, el libro de J. Serret del año 1866 contiene una versión correcta del algoritmo de

Galois (ver [Ser66]). Por otra parte, el trabajo de A. Arwin en el año 1918 contiene muchas de las ideas modernas sobre la factorización, incluyendo este algoritmo (ver [Arw18]). En cuanto a la factorización en grados iguales, podemos citar el trabajo de Cantor–Zassenhaus [CZ81]. Otras variantes de este último algoritmo se deben a M. Ben-Or [BO81] y a von Zur Gathen y Shoup [vzGS92, Algorithm 3.6].

En [vzGG99, Exercise 14.27] los autores analizan el costo del peor caso del primer paso de este algoritmo de factorización (el algoritmo ERF) y demuestran que necesita $\mathcal{O}(\mathcal{U}(d) + d\log(q/p))$ operaciones aritméticas en $\mathbb{F}_q$.

El segundo paso, algoritmo DDF, se basa en un resultado que dice que para todo cuerpo finito $\mathbb{F}_q$ y todo entero positivo d, el producto de todos los polinomios mónicos, irreducibles cuyo grado divide a d es igual a $T^{q^d} - T$. Así el máximo común divisor $g_1 := \gcd(T^q - T, f)$ da los factores irreducibles de f de grado 1, si calculamos $g_2 := \gcd(T^{q^2} - T, f/g_1)$ obtenemos todos los factores irreducibles de f de grado 2, y sucesivos cálculos del máximo común divisor $\gcd(T^{q^k} - T, f/g_{k-1})$, con $k = 1, 2, \ldots, d$, dan la factorización en grados distintos de f. En [vzGG99, Algorithm 14.3] los autores prueban que el algoritmo DDF realiza $\mathcal{O}(sM(d)\log(dq))$ operaciones aritméticas en $\mathbb{F}_q$, donde s es el máximo grado de los factores irreducibles del polinomio de entrada.

En relación con el último paso, el algoritmo EDF, en [vzGG99, Theorem 14.11] los autores prueban que, si el polinomio de entrada tiene s factores irreducibles de grado k, su costo es de $\mathcal{O}((k\log q + \log d)M(d)\log s)$ operaciones aritméticas en $\mathbb{F}_q$. Concluimos que el algoritmo clásico de factorización calcula un cero $\mathbb{F}_q$–racional del polinomio de entrada con $\mathcal{O}(dM(d)\log(dq))$ operaciones aritméticas en $\mathbb{F}_q$ (ver, [vzGG99, Theorem 14.14]).

Por su parte, P. Flajolet y colaboradores, realizan en [FGP01] un análisis en promedio de dicho algoritmo y demuestran que la distribución de patrones de factorización sigue el denominado "modelo de permutaciones", es decir, para q suficientemente grande (con d fijo) la distribución conjunta de los grados de los factores irreducibles en un polinomio aleatorio de grado d converge a la distribución conjunta de las longitudes de los ciclos en una permutación aleatoria de tamaño d.

En cuanto al primer paso del algoritmo, en [FGP01] se prueba que el costo en promedio del algoritmo ERF está dominado por el costo del máximo común divisor entre el polinomio de entrada y su derivada, es decir, el algoritmo realiza en promedio $\mathcal{O}(\mathcal{U}(d))$ operaciones aritméticas en $\mathbb{F}_q$ para obtener la parte libre de cuadrados del polinomio de entrada.

En relación al algoritmo DDF, en [FGP01, Section 3, Theorem 5] los autores prueban que un polinomio aleatorio en $\mathbb{F}_q[T]_d$ tiene alta probabilidad de poseer factores irreducibles de grados altos, y dan estimaciones asintóticas para las distribuciones conjuntas de los dos grados más altos de los factores irreducibles del polinomio de entrada. Prueban, por ejemplo, que el grado más alto esperado tiende al número $\xi \cdot d$, donde $\xi \sim 0{,}62432\ldots$ es la constante de Golomb que representa la longitud más grande esperada entre los ciclos de una permutación aleatoria. A partir de estos resultados demuestran que el costo promedio del algoritmo DDF aplicado a polinomios libre de cuadrados de grado d es del orden de $0{,}26689(\lambda(q)\tau_1 + \tau_2)d^3$, donde $\lambda(q) \leq 2\log q$ y τ_1 y τ_2 son constantes. Observan además que la mayoría de las

factorizaciones se completan luego de la aplicación del algoritmo DDF. Más precisamente, muestran que, cuando d está fijo y q tiende a infinito, la probabilidad de que el algoritmo DDF produzca una factorización completa de un polinomio aleatorio es del orden de $e^{-\gamma} \sim 0{,}5614\ldots$, donde $\gamma \sim 0{,}577215664\ldots$ es la constante de Euler.

En relación con el algoritmo EDF, los autores combinan un proceso de refinamiento recursivo similar al de los árboles binarios (ver [Knu98b]) junto con estimaciones sobre los grados de los factores irreducibles de un polinomio aleatorio (ver [KK90a]), a fin de demostrar que el costo promedio del algoritmo EDF es comparativamente pequeño, es decir, es del orden de $\frac{3}{4}\tau_1 \frac{q^2}{q^2-1} \log q(1+\xi_d)d^2 = \mathcal{O}(d^2 \log q)$ operaciones aritméticas en $\mathbb{F}_q$, donde $|\xi_d| \leq \frac{1}{3} + o(1)$.

Asintóticamente, el algoritmo clásico de factorización no es el más rápido que existe hasta el momento (comparar con, por ejemplo, [Sho90], [Sho95] o [vzGP01]). Una de las razones para estudiarlo es que es eficiente, completo y aparece en muchos paquetes de álgebra computacional (ver [GCL92]). Asimismo, se puede obtener información importante del comportamiento del algoritmo en cada uno de sus pasos y del estado de la factorización del polinomio de entrada al finalizar cada etapa del mismo.

1.1.3. Problemas combinatorios

Motivados por el análisis de los algoritmos que describimos en las secciones anteriores, vamos a considerar dos problemas combinatorios clásicos sobre cuerpos finitos:

- El comportamiento promedio del cardinal de la imagen, o conjunto de valores, de familias lineales de polinomios univariados sobre cuerpos finitos.
- La distribución de los patrones de factorización en familias lineales de polinomios univariados sobre cuerpos finitos.

Conjunto de valores de polinomios. El estudio del cardinal del conjunto de valores de familias de polinomios univariados sobre cuerpos finitos ha sido objeto de diversas investigaciones, por sus aplicaciones a la teoría de códigos, la criptografía y a problemas de interpolación. Dado un polinomio univariado f con coeficientes en $\mathbb{F}_q$, se define el conjunto de valores de f sobre $\mathbb{F}_q$ como el conjunto imagen de la función polinomial de $\mathbb{F}_q$ en $\mathbb{F}_q$ que define f. Denotamos como $\mathcal{V}(f)$ al cardinal de dicho conjunto.

Para todo $f \in \mathbb{F}_q[T]$ se verifica fácilmente que $1 \leq \mathcal{V}(f) \leq q$. El problema de calcular $\mathcal{V}(f)$ ha sido muy estudiado (ver, por ejemplo [MP13]); sin embargo, solo se conocen fórmulas exactas de $\mathcal{V}(f)$ para polinomios especiales. Por ejemplo existen fórmulas para polinomios de grado 1, 2 y 3, para el polinomio $f := T^d$ y para los polinomios de Dickson, entre otros; en cambio, para polinomios generales sólo se conocen fórmulas asintóticas de $\mathcal{V}(f)$. En este sentido, S. Uchiyama [Uch54] prueba que si $\frac{f(x)-f(y)}{x-y}$ es absolutamente irreducible y $d \geq 4$ entonces

$$\mathcal{V}(f) = \mu_d q + \mathcal{O}(1).$$

Uchiyama prueba además que con las mismas hipótesis se tiene que $\mathcal{V}(f) > q/2$ para q grande y con la característica de $\mathbb{F}_q$ mayor a d. Por otro lado, Birch y Swinnerton-Dyer demostraron en [BS59] que, si f es un polinomio general de grado d (el grupo de Galois $\mathrm{Gal}(f(x)-y/\overline{\mathbb{F}}_q(y))$ es el grupo de permutaciones de d elementos), entonces

$$\mathcal{V}(f) = \mu_d q + \mathcal{O}(q^{1/2}),$$

donde $\mu_d := \sum_{r=1}^{d}(-1)^{r-1}/r!$ y la constante que subyace en la notación $\mathcal{O}$ depende sólo de d.

En los años 50, L. Carlitz y S. Uchiyama, utilizando técnicas de análisis combinatorio, estudiaron el comportamiento del valor promedio $\mathcal{V}(d,0)$ de $\mathcal{V}(f)$ cuando f recorre todos los polinomios en $\mathbb{F}_q[T]_d$ con $f(0)=0$, suponiendo que la característica de $\mathbb{F}_q$ es mayor que d (ver [Car55], [CU57], [Uch55a] y [Uch56]). Estos resultados fueron mejorados por S. Cohen en [Coh73], quien para todo d y todo q demostró que

$$\mathcal{V}(d,0) = \sum_{r=1}^{d}(-1)^{r-1}\binom{q}{r}q^{1-r} = \mu_d q + \mathcal{O}(1).$$

Observemos que si $d \geq q$, este resultado dice que $\mathcal{V}(d,0) = q\big(1-(1-1/q)\big)^q$. En [KK90b], A. Knopfmacher y J. Knopfmacher estudiaron la distribución del número de polinomios f de grado $d \geq q$ tal que $\mathcal{V}(f) = n$, con $1 \leq n \leq q$.

Por otro lado, si alguno de los coeficientes de los polinomios f están fijos, los resultados que se conocen sobre el promedio de $\mathcal{V}(f)$ son menos precisos. Más precisamente, Uchiyama y Cohen determinaron el comportamiento asintótico del promedio de $\mathcal{V}(f)$ cuando f recorre todos los polinomios en $\mathbb{F}_q[T]_d$ cuyos primeros s coeficientes consecutivos están prefijados. En los trabajos [Uch55b], [Coh72] y [Coh73] obtuvieron estimaciones sobre este promedio, con restricciones sobre la característica del cuerpo $\mathbb{F}_q$ y para el caso en que q sea mayor que d. Más precisamente, si $\mathcal{V}(d,s)$ denota tal promedio, demostraron que, si la característica de $\mathbb{F}_q$ es mayor que d y $1 \leq s \leq d-2$, entonces el término principal de $\mathcal{V}(d,s)$ es $\mu_d q$ y el error que se comete en dicha aproximación es del orden de $\mathcal{O}(q^{1/2})$, en el caso de Cohen, o de $\mathcal{O}(q^{\theta})$, donde θ depende del grado y de la cantidad de coeficientes que se fijan, en el caso de Uchiyama. Cabe mencionar que en todos estos trabajos ni Cohen ni Uchiyama dieron una expresión explícita del término del error en sus estimaciones.

Cohen por su parte estudia el comportamiento asintótico del cardinal de conjunto de valores promedio en familias lineales de polinomios univariados con coeficientes en $\mathbb{F}_q$. Más precisamente, en [Coh72] afirma que, para una familia lineal $\mathcal{A} \subset \mathbb{F}_q[T]_d$ que satisface ciertas condiciones técnicas, se tiene que si la característica de $\mathbb{F}_q$ es mayor a d y la codimensión de $\mathcal{A}$ es igual a $m \leq d-2$, entonces

$$\mathcal{V}(\mathcal{A}) := \frac{1}{|\mathcal{A}|}\sum_{f\in\mathcal{A}}\mathcal{V}(f) = \mu_d q + \mathcal{O}(q^{1/2}).$$

Las dificultades que presenta este resultado es que las hipótesis sobre la familia son complicadas de verificar en casos concretos e imponen fuertes restricciones sobre la característica del cuerpo, que impiden utilizar esta estimación para cuerpos de característica pequeña.

Generalmente los resultados existentes sobre el promedio del cardinal del conjunto de valores de una familia de polinomios en $\mathbb{F}_q[T]_d$ se basan en dos tipos de estrategias esencialmente distintas. Una de ellas utiliza métodos típicos de la combinatoria analítica, es decir, se consideran funciones generatrices cuyos coeficientes expresan las propiedades en cuestión, como, por ejemplo, el artículo [KK90b], donde se analiza el caso en que d es mayor que q. Cuando no es posible calcular con exactitud tales coeficientes se recurre al análisis asintótico para obtener una estimación de los mismos (ver [FS09]). Otro acercamiento posible es por medio de estimaciones de series exponenciales como se hace, por ejemplo, en [Uch55a], [Uch56], [Uch55b], [Coh72] y [Coh73], donde se estudia el problema para el caso en que q es mayor que d. Sin embargo, cuando aparecen relaciones algebraicas entre los coeficientes de los polinomios en consideración, y se buscan resultados sin restricciones sobre la característica del cuerpo finito, como en los problemas que nos interesan, estos métodos no pueden utilizarse.

Patrones de factorización de polinomios. Frecuentemente el cálculo de ceros de polinomios univariados sobre cuerpos finitos implica algún tipo de factorización de dichos polinomios. En este sentido, existen numerosos algoritmos probabilísticos que factorizan un polinomio de $\mathbb{F}_q[T]_d$ en tiempo polinomial en d y $\log q$; de hecho, es posible calcular una factorización con alrededor de $d^2 + d \log q$ operaciones en $\mathbb{F}_q$ (ver, por ejemplo, [vzGG99, Chapter 14]). El análisis de [FGP01] demuestra que el estudio de la distribución de los patrones de factorización resulta fundamental a efectos de realizar un análisis probabilístico de los algoritmos de factorización. Como ya dijimos anteriormente, la distribución de los patrones de factorización sigue el denominado modelo de permutaciones. En dicho modelo, propiedades probabilísticas de la descomposición de un polinomio en factores irreducibles sobre cuerpos finitos se relacionan con las correspondientes propiedades de la descomposición en ciclos de una permutación, cuando q es suficientemente grande. Por ejemplo, cuando d está fijo y q tiende a infinito, la probabilidad de que un polinomio aleatorio en $\mathbb{F}_q[T]_d$ admita factores irreducibles de grados distintos dos a dos tiende a la probabilidad de que una permutación de longitud d tenga todos sus ciclos de longitudes distintas dos a dos.

Si $f \in \mathbb{F}_q[T]_d$, se dice que f tiene *patrón de factorización* $\boldsymbol{\lambda} := 1^{\lambda_1} \dots d^{\lambda_d}$ si tiene λ_i factores irreducibles de grado i para $1 \leq i \leq d$. Un trabajo fundacional para el modelo de permutaciones es [Coh70], en donde Cohen relacionó la cantidad de polinomios que tienen un cierto patrón de factorización con la descomposición en ciclos de los elementos del grupo simétrico de permutaciones. Más precisamente, demostró que si $\mathbb{F}_q[T]_{d,\boldsymbol{\lambda}}$ es el conjunto formado por todos los polinomios de $\mathbb{F}_q[T]_d$ que tienen patrón de factorización $\boldsymbol{\lambda}$, para d fijo, entonces

$$|\mathbb{F}_q[T]_{d,\boldsymbol{\lambda}}| = \mathcal{T}(\boldsymbol{\lambda}) q^d + \mathcal{O}(q^{d-1/2}),$$

donde $\mathcal{T}(\boldsymbol{\lambda}) := \frac{1}{\prod_{i=1}^{d} i^{\lambda_i} \lambda_i!}$ es la proporción de permutaciones en el grupo simétrico de d elementos que tienen patrón de descomposición en ciclos $\boldsymbol{\lambda}$, es decir, permutaciones con exactamente λ_i ciclos de longitud i para $1 \leq i \leq d$. Por ejemplo, la

proporción del conjunto de permutaciones cuyo patrón es $\boldsymbol{\lambda} := 1^0 \dots (d-1)^0 d^1$, esto es, permutaciones que consisten de un ciclo de longitud d, es $\frac{1}{d}$, y por lo tanto, la cantidad de polinomios mónicos irreducibles en $\mathbb{F}_q[T]_d$ es del orden de $\frac{1}{d}q^d$.

Para una familia lineal $\mathcal{A}$ de polinomios de $\mathbb{F}_q[T]_d$, denotamos con $|\mathcal{A}_{\boldsymbol{\lambda}}|$ al número de elementos de $\mathcal{A}$ con patrón de factorización $\boldsymbol{\lambda}$. Cohen definió en [Coh72] que una familia $\mathcal{A} \subset \mathbb{F}_q[T]_d$ es *uniformemente distribuida* si la proporción de elementos en $\mathcal{A}$ con patrón de factorización $\boldsymbol{\lambda}$ es del orden de $\mathcal{T}(\boldsymbol{\lambda})$. A su vez, en [Coh72, Theorem 3] propuso un criterio para que una familia lineal resulte uniformemente distribuida. Más precisamente, demostró que si la característica de $\mathbb{F}_q$ es mayor a d y $\mathcal{A}$ cumple ciertas restricciones técnicas y tiene codimensión $m \leq d-2$, entonces

$$|\mathcal{A}_{\boldsymbol{\lambda}}| = \mathcal{T}(\boldsymbol{\lambda}) q^{d-m} + \mathcal{O}(q^{d-m-1/2}).$$

Podemos observar que, al igual que la estimación de Cohen para el promedio del cardinal $\mathcal{V}(\mathcal{A})$, este criterio se aplica con fuertes restricciones sobre la característica del cuerpo finito y las hipótesis que determinan que $\mathcal{A}$ resulta uniformemente distribuida son complicadas de verificar en casos concretos. El resultado [Coh72, Theorem 3] se aplica, en particular, al conjunto $\mathcal{A}_s$ que consiste de todos los polinomios de $\mathbb{F}_q[T]_d$ con los primeros $s \leq d-2$ coeficientes prescriptos (ver, [Coh72, Theorem 1]). Otro resultado sobre esta familia se encuentra en [Ste87, Theorem 1]. En el mismo, S. Stepanov demuestra que bajo ciertas condiciones técnicas, $|\mathcal{A}_{s,\boldsymbol{\lambda}}|$, la cantidad de elementos de $\mathcal{A}_s$ con patrón de factorización $\boldsymbol{\lambda} := d^1$, se aproxima a cq^{d-s} con un error del orden de $\mathcal{O}(q^{(d-s)/2})$, donde c depende del conjunto de todos los posibles patrones de factorización de la familia $\mathcal{A}_s$ y cumple que $1/d \leq c < 1$. Observemos que ni Cohen ni Stepanov dieron cotas explícitas del término de error en sus aproximaciones. Por último, podemos mencionar que el problema de estudiar el número de elementos de una familia de polinomios irreducibles con ciertos coeficientes prefijados posee a su vez un interés teórico, en tanto constituye un análogo en "cuerpos de funciones" de los resultados sobre la distribución de primos en intervalos pequeños (ver el trabajo [BBSR15] para una discusión de dicha analogía y [MP13, Section 3] o [Pol13] para el estudio del número de polinomios irreducibles con ciertos coeficientes prefijados).

1.2. Resultados obtenidos y organización del trabajo

Sean $f_1, \dots, f_m$ polinomios en $\mathbb{F}_q[X_1, \dots, X_n]$. Consideremos el conjunto de soluciones en el espacio afín n–dimensional $\mathbb{A}^n := \overline{\mathbb{F}}_q^n$ sobre $\overline{\mathbb{F}}_q$ del sistema que estos polinomios definen, es decir, $V := \{x \in \mathbb{A}^n : f_1(x) = \cdots = f_m(x) = 0\}$. Decimos que V es una $\mathbb{F}_q$–variedad afín. Cuando $f_1, \dots, f_m$ son polinomios homogéneos en $\mathbb{F}_q[X_0, \dots, X_n]$ y se considera el conjunto de ceros comunes de ellos en el espacio proyectivo n–dimensional $\mathbb{P}^n$ sobre $\overline{\mathbb{F}}_q$, entonces decimos que V es una $\mathbb{F}_q$–variedad proyectiva. Sea V una $\mathbb{F}_q$–variedad afín o proyectiva. Dado $x \in V$, decimos que x es un punto $\mathbb{F}_q$–racional de V si todas sus coordenadas pertenecen a $\mathbb{F}_q$. Denotamos con $V(\mathbb{F}_q)$ al conjunto de puntos $\mathbb{F}_q$–racionales de V.

El Capítulo 2 está dedicado a introducir los preliminares geométricos y fijar las notaciones necesarias para el desarrollo de los demás capítulos. También hacemos una revisión de los resultados clásicos sobre estimaciones de puntos $\mathbb{F}_q$–racionales de $\mathbb{F}_q$–variedades tanto afines como proyectivas.

Con el objetivo de mejorar los resultados existentes sobre los dos problemas combinatorios que vamos a estudiar, el cardinal del promedio del conjunto de valores y la distribución de los patrones de factorización de familias lineales en $\mathbb{F}_q[T]_d$, desarrollamos un nuevo enfoque. Nuestro enfoque reduce las cuestiones combinatorias a la estimación del número de puntos $\mathbb{F}_q$–racionales de ciertas intersecciones completas singulares. Un punto importante es el análisis del lugar singular de dichas intersecciones completas; para ello estudiamos el lugar discriminante asociado a las familias lineales de $\mathbb{F}_q[T]_d$ que consideramos. Más precisamente, la familia $\mathcal{A} \subset \mathbb{F}_q[T]_d$ a la que nos referimos se define de la siguiente manera: sean m y r enteros positivos tales que $3 \leq r \leq d-m$, sean $A_{d-1}, \ldots, A_r$ indeterminadas sobre $\overline{\mathbb{F}}_q$ y sean $L_1, \ldots, L_m \in \mathbb{F}_q[A_{d-1}, \ldots, A_r]$ formas lineales afines linealmente independientes. Si $\boldsymbol{L} := (L_1, \ldots, L_m)$, la familia lineal $\mathcal{A} := \mathcal{A}_{\boldsymbol{L}} \subset \mathbb{F}_q[T]_d$ se define por:

$$\mathcal{A} := \{T^d + a_{d-1}T^{d-1} + \cdots + a_0 \in \mathbb{F}_q[T]_d : \boldsymbol{L}(a_{d-1}, \ldots, a_r) = \mathbf{0}\}. \tag{1.1}$$

Comenzamos el Capítulo 3 estudiando el lugar discriminante de la variedad lineal $\mathcal{L} := \{L_1 = \cdots = L_m = 0\}$ suponiendo que la característica de $\mathbb{F}_q$ es al menos 3. Se define el lugar discriminante $\mathcal{D}(\mathcal{L})$ de $\mathcal{L}$ como el conjunto de todos los elementos de $\boldsymbol{a}_0 := (a_{d-1}, \ldots, a_0) \in \mathcal{L}$ tales que $\mathrm{Dis}(F(A_{d-1}, \ldots, A_0, T))|_{(A_{d-1},\ldots,A_0)=\boldsymbol{a}_0} = 0$, donde $F(A_{d-1} \ldots, A_0, T) := T^d + A_{d-1}T^{d-1} + \cdots + A_0$. Referidos a este tema, encontramos en la literatura el trabajo de M. Fried y J. Smith [FS84], donde se prueba que el lugar discriminante de familias de polinomios mónicos univariados con ciertos coeficientes prescriptos es absolutamente irreducible para cuerpos de característica suficientemente grande (ver [FS84, Proposition 3.1]). Nosotros necesitamos un resultado análogo sobre el lugar discriminante de la familia lineal $\mathcal{A}$ y sobre cuerpos de característica al menos 3. Es por ello que demostramos el siguiente resultado.

Teorema 1.2.1. *Sea la característica de $\mathbb{F}_q$ mayor a 2, $q > d$ y $3 \leq r \leq d-m$. Entonces $\mathcal{D}(\mathcal{L}) \subset \mathbb{A}^{d-m}$ es una $\mathbb{F}_q$–hipersuperficie absolutamente irreducible.*

En el mismo capítulo, usando la absoluta irreducibilidad del lugar discriminante $\mathcal{D}(\mathcal{L})$, estudiamos la geometría de ciertas variedades de incidencia $\Gamma_i^* \subset \mathbb{A}^{d+i}$ $(r+1 \leq i \leq d)$ asociadas a la familia $\mathcal{A}$ definida en (1.1). Estas variedades están definidas por las formas lineales $L_1, \ldots, L_m$ y las diferencias divididas $\Delta^{j-1}F(A_{d-1}, \ldots, A_0, T_1, \ldots, T_j)$ $(1 \leq j \leq i)$ de orden $j-1$ del polinomio $F(A_{d-1}, \ldots, A_0, T)$, es decir

$$\Delta^{j-1}F(\boldsymbol{A}_0, T_1, \ldots, T_j) := \frac{\Delta^{j-2}F(\boldsymbol{A}_0, T_1, \ldots, T_{j-1}) - \Delta^{j-2}F(\boldsymbol{A}_0, T_1, \ldots, T_{j-2}, T_j)}{(T_{j-1} - T_j)},$$

donde $\boldsymbol{A}_0 := (A_{d-1}, \ldots, A_0)$.

Suponiendo que la característica de $\mathbb{F}_q$ es mayor a 2, probamos que estas variedades resultan ser intersecciones completas definidas sobre $\mathbb{F}_q$, con buen comportamiento en el infinito y cuyo lugar singular tiene codimensión al menos 2. Gracias a

estas propiedades geométricas, podemos utilizar las estimaciones sobre el número de puntos $\mathbb{F}_q$–racionales de una intersección completa proyectiva normal de [CMP15a, Theorem 1.3] a fin de obtener estimaciones sobre la cantidad de puntos $\mathbb{F}_q$–racionales $|\Gamma_i^*(\mathbb{F}_q)|$ de las variedades Γ_i^*. Más precisamente, demostramos el siguiente resultado.

Teorema 1.2.2. *Sea la característica de $\mathbb{F}_q$ mayor a 2 y $q > d$. Sean r, d y m enteros positivos tales que $3 \leq r \leq d - m$ y sea i un entero tal que $r + 1 \leq i \leq d$. Entonces*

$$\left||\Gamma_i^*(\mathbb{F}_q)| - q^{d-m}\right| \leq (\delta_i(D_i - 2) + 2)q^{d-m-\frac{1}{2}} + (14D_i^2\delta_i^2 + 2i)q^{d-m-1},$$

donde $D_i := id - i(i+1)/2$ y $\delta_i := d!/(d-i)!$.

Proporcionamos también una estimación del número de puntos $\mathbb{F}_q$–racionales de la variedad Γ_i^* con coordenadas distintas dos a dos, cota que nos permitirá, en los capítulos siguientes, dar estimaciones explícitas sobre el número promedio del cardinal de la imagen del conjunto de valores de la familia $\mathcal{A}$ y, en particular, del conjunto de elementos de $\mathbb{F}_q[T]_d$ con los primeros s coeficientes prescriptos, cuando $1 \leq s \leq d - 3$ y la característica de $\mathbb{F}_q$ es mayor que 2.

En el Capítulo 4 nos dedicamos al problema de estimar la cantidad de puntos $\mathbb{F}_q$–racionales de variedades definidas por polinomios invariantes bajo la acción del grupo simétrico de permutaciones de sus coordenadas. Muchos problemas de la teoría de códigos, de la criptografía o de la combinatoria requieren el estudio del conjunto de puntos $\mathbb{F}_q$–racionales de este tipo de variedades. Por ejemplo, en la teoría de códigos encontramos que la existencia de deep holes en los códigos de Reed Solomon sobre $\mathbb{F}_q$ pueden expresarse en términos de la cantidad de ceros $\mathbb{F}_q$–racionales de ciertos polinomios simétricos (ver, por ejemplo, [CM07] o [CMP15a]). Además, el estudio del conjunto de ceros $\mathbb{F}_q$–racionales de cierta clase de polinomios simétricos es fundamental para el análisis del algoritmo de decodificación para el código de Reed Solomon estándar sobre $\mathbb{F}_q$ (ver [Sid94]). En criptografía, la caracterización de los monomios que definen un polinomio casi perfectamente no lineal o una función uniformemente diferenciable puede reducirse a estimar el número de ceros $\mathbb{F}_q$–racionales de ciertos polinomios simétricos (ver, por ejemplo, [Rod09] o [AR10]). Nuestro interés por esta cuestión se debe a que hemos expresado los dos problemas combinatorios que son objeto de nuestro estudio en términos de la cantidad de puntos $\mathbb{F}_q$–racionales de ciertas variedades definidas por polinomios simétricos.

Sean s, r, m enteros positivos tales que $m \leq s \leq r - m - 2$. Sean $R_1, \ldots, R_m \in \mathbb{F}_q[X_1, \ldots, X_r]$ polinomios de la forma $R_i := S_i(\Pi_1, \ldots, \Pi_s)$ para $1 \leq i \leq m$, donde $S_1, \ldots, S_m \in \mathbb{F}_q[Y_1, \ldots, Y_s]$ satisfacen ciertas hipótesis geométricas y donde $\Pi_1, \ldots, \Pi_s$ son los primeros s polinomios simétricos elementales de $\mathbb{F}_q[X_1, \ldots, X_r]$. En la Sección 4.1 probamos que la $\mathbb{F}_q$–variedad $V \subset \mathbb{A}^r$ definida por $R_1, \ldots, R_m$ es una intersección completa con buen comportamiento en el infinito y cuyo lugar singular tiene codimensión al menos 3. Estos resultados, junto con estimaciones sobre el número de puntos $\mathbb{F}_q$–racionales de intersecciones completas proyectivas de [CMP15a, Corollary 8.4], nos permiten dar la siguiente estimación del número de puntos $\mathbb{F}_q$–racionales de la variedad V.

Teorema 1.2.3. *Sean s, r, m enteros positivos tales que $m \leq s \leq r - m - 2$. Sean $R_1, \ldots, R_m \in \mathbb{F}_q[X_1, \ldots, X_r]$ los polinomios simétricos definidos arriba. Denotamos*

por $d_i := \deg R_i$ *para* $1 \leq i \leq m$, $D := \sum_{i=1}^m (d_i - 1)$ *y* $\delta := \prod_{i=1}^m d_i$. *Si* $V := V(R_1, \ldots, R_m) \subset \mathbb{A}^r$, *entonces vale la siguiente estimación:*

$$\big||V(\mathbb{F}_q)| - q^{r-m}\big| \leq 14D^3\delta^2(q+1)q^{r-m-2}.$$

Al igual que antes, también damos un resultado sobre la cantidad de puntos $\mathbb{F}_q$–racionales de V con coordenadas distintas dos a dos.

Por último en la Sección 4.2 estudiamos la geometría de otras variedades definidas por polinomios simétricos asociadas a la familia lineal $\mathcal{A}$ definida en (1.1). Sean d, s y m enteros positivos tales que $m \leq s \leq d-3$. En este caso, los polinomios simétricos que definen estas variedades son de la forma $R_i := S_i(\Pi_1, \ldots, \Pi_s)$ para $1 \leq i \leq m$, donde $S_1, \ldots, S_m \in \mathbb{F}_q[Z_1, \ldots, Z_s]$ son polinomios de grado 1. Probamos que estas variedades resultan ser intersecciones completas normales y damos una estimación de la cantidad de puntos $\mathbb{F}_q$–racionales de las mismas. Estas intersecciones completas aparecerán en el estudio de la distribución de los patrones de factorización en familias lineales sobre $\mathbb{F}_q$.

En el Capítulo 5 damos estimaciones explícitas del promedio $\mathcal{V}(\mathcal{A})$ del cardinal del conjunto de valores de la familia $\mathcal{A}$ definida en (1.1). Para obtener dichas estimaciones, traducimos este problema combinatorio en el de estimar el número de puntos $\mathbb{F}_q$–racionales con coordenadas distintas dos a dos de las familias de intersecciones completas $\Gamma_i^* \subset \mathbb{A}^{d+i}$ definidas en la Sección 3.2 y utilizamos los resultados sobre la cantidad de puntos $\mathbb{F}_q$–racionales de variedades singulares obtenidos en dicha sección. Probamos el siguiente resultado.

Teorema 1.2.4. *Si la característica de* $\mathbb{F}_q$ *es mayor que* 2, $q > d$ *y* $3 \leq r \leq d-m$, *entonces*

$$|\mathcal{V}(\mathcal{A}) - \mu_d\, q| \leq d^2 2^{d-1} q^{1/2} + 133\, d^{d+5} e^{2\sqrt{d}-d}.$$

Mejoramos así las estimaciones del trabajo de Cohen [Coh72], cuya validez impone fuertes restricciones sobre la característica de $\mathbb{F}_q$, proporcionando a su vez una estimación explícita del error.

Luego aplicamos esta estimación a un caso particular, al conjunto de polinomios de $\mathbb{F}_q[T]_d$ que tienen los primeros s coeficientes prescriptos. Más precisamente, damos la siguiente estimación explicita del valor promedio $\mathcal{V}(d, s)$ del cardinal del conjunto de valores posibles que puede tomar un polinomio f cuando f recorre todos los elementos de la familia en consideración.

Teorema 1.2.5. *Si la característica de* $\mathbb{F}_q$ *mayor que* 2, $q > d$ *y* $1 \leq s \leq d-3$, *entonces*

$$|\mathcal{V}(d, s) - \mu_d\, q| \leq d^2 2^{d-1} q^{1/2} + 133\, d^{d+5} e^{2\sqrt{d}-d}.$$

Esta estimación mejora los resultados existentes en la literatura ([Uch55b] y [Coh72]), que no proveen una expresión explícita del término de error y resultan válidas para cuerpos de característica mayor que d. Cabe mencionar que este resultado es un punto crítico para el análisis de la complejidad en promedio del algoritmo de búsqueda de puntos $\mathbb{F}_q$–racionales en hipersuperficies (ver Sección 8.3.2).

Por otro lado, en el Capítulo 6 proporcionamos otra estimación explícita del comportamiento de $\mathcal{V}(d,s)$ para el caso en que $1 \leq s \leq d/2$. De manera similar al capítulo anterior, traducimos este problema en el de determinar la cantidad de puntos $\mathbb{F}_q$–racionales con coordenadas distintas dos a dos de cierta familia de intersecciones completas definidas sobre $\mathbb{F}_q$. Los polinomios que definen tales intersecciones completas son simétricos, por lo que aplicamos las estimaciones de la Sección 4.1. Obtenemos el siguiente resultado.

Teorema 1.2.6. *Para $q > d$ y $1 \leq s \leq d/2 - 1$, tenemos que*

$$\left|\mathcal{V}(d,s) - \mu_d\, q - \frac{1}{2e}\right| \leq \frac{(d-2)^5 e^{2\sqrt{d}}}{2^{d-2}} + \frac{7}{q}.$$

Este resultado complementa el resultado del Teorema 1.2.5. En el Teorema 1.2.6 damos una estimación de $\mathcal{V}(d,s)$ para el caso en que $1 \leq s \leq d/2 - 1$ sin restricciones sobre la característica de $\mathbb{F}_q$, mientras que en el Teorema 1.2.5 damos una cota superior para este promedio para valores más grandes de s que vale cuando la característica de $\mathbb{F}_q$ es mayor a 2.

En el Capítulo 7 estudiamos la distribución de patrones de factorización en familias lineales sobre $\mathbb{F}_q$. Con un enfoque similar al de los Capítulos 5 y 6, traducimos el problema original en el de determinar la cantidad de puntos $\mathbb{F}_q$–racionales con coordenadas distintas dos a dos de ciertas intersecciones completas singulares definidas sobre $\mathbb{F}_q$. Tales intersecciones completas están definidas por polinomios simétricos, con lo cual podemos aplicar los resultados de la Sección 4.2. Más precisamente, obtenemos la siguiente estimación explícita del cardinal del conjunto $\mathcal{A}_{\boldsymbol{\lambda}}$ de elementos de la familia lineal $\mathcal{A}$ definida en (1.1) con patrón de factorización $\boldsymbol{\lambda} := 1^{\lambda_1} \ldots d^{\lambda_d}$.

Teorema 1.2.7. *Si característica de $\mathbb{F}_q$ es mayor a 2, $q > d$ y $3 \leq r \leq d - m$, entonces*

$$\big||\mathcal{A}_{\boldsymbol{\lambda}}| - \mathcal{T}(\boldsymbol{\lambda})\, q^{d-m}\big| \leq q^{d-m-1}\big(2\,\mathcal{T}(\boldsymbol{\lambda})\, D_{\boldsymbol{L}}\delta_{\boldsymbol{L}} q^{\frac{1}{2}} + 19\,\mathcal{T}(\boldsymbol{\lambda})\, D_{\boldsymbol{L}}^2\delta_{\boldsymbol{L}}^2 + d^2\big).$$

Para $q > d$ y $m + 2 \leq r \leq d - m$, tenemos que

$$\big||\mathcal{A}_{\boldsymbol{\lambda}}| - \mathcal{T}(\boldsymbol{\lambda})\, q^{d-m}\big| \leq q^{d-m-1}\big(21\,\mathcal{T}(\boldsymbol{\lambda})\, D_{\boldsymbol{L}}^3\delta_{\boldsymbol{L}}^2 + \mathcal{T}(\boldsymbol{\lambda})\, d^2\delta_{\boldsymbol{L}} + d^2\big).$$

En ambos casos, $\delta_{\boldsymbol{L}}$ y $D_{\boldsymbol{L}}$ son ciertos invariantes explícitos asociados con las formas lineales afines $\boldsymbol{L} := (L_1, \ldots, L_m)$ que definen la familia $\mathcal{A}$ y $\mathcal{T}(\boldsymbol{\lambda}) := \frac{1}{\prod_{i=1}^{d} i^{\lambda_i}\lambda_i!}$.

En el Capítulo 8 desarrollamos y analizamos la complejidad en promedio de un algoritmo probabilístico que calcula puntos $\mathbb{F}_q$–racionales de hipersuperficies en base a la estrategia de búsquedas en bandas verticales. Tal algoritmo se llama Algoritmo BBV. Para el análisis de dicho algoritmo resulta clave estudiar la cantidad de bandas verticales que deben ser generadas hasta que el algoritmo encuentra una raíz del polinomio multivariado de entrada.

En la Sección 8.1 describimos el Algoritmo BBV. Sea F el conjunto de todas las posibles elecciones de bandas verticales y sea $\mathbb{F}_q[X_1, \ldots, X_r]_{\leq d}$ el conjunto de todos

los polinomios en r variables con coeficientes en $\mathbb{F}_q$ y de grado a lo sumo d. Para un polinomio $F \in \mathbb{F}_q[X_1, \ldots, X_r]_{\leq d}$, el Algoritmo BBV genera sucesivamente una sucesión $\underline{\boldsymbol{a}} := (\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_{q^{r-1}}) \in \mathsf{F}$, y busca ceros $\mathbb{F}_q$–racionales de F en las "bandas verticales" $\{\boldsymbol{a}_i\} \times \mathbb{F}_q$ para $1 \leq i \leq q^{r-1}$, hasta que encuentra un cero de F o se terminan las bandas verticales. Observamos que el número de operaciones aritméticas en $\mathbb{F}_q$ necesarias para realizar una búsqueda en una banda vertical arbitraria es $\tau(d, r, q) := \mathcal{O}^{\sim}(D + d \log q)$, donde $D := \binom{d+r}{r}$ es la cantidad de coeficientes del polinomio F y la notación $\mathcal{O}^{\sim}$ ignora factores logarítmicos.

En las Secciones 8.2 y 8.4 analizamos la distribución de probabilidades del número de bandas verticales que deben ser generadas por el Algoritmo BBV hasta encontrar un cero $\mathbb{F}_q$–racional del polinomio de entrada. Para ello, consideramos la probabilidad uniforme P sobre el conjunto $\mathsf{F} \times \mathbb{F}_q[X_1, \ldots, X_r]_{\leq d}$, y estudiamos la variable aleatoria C que cuenta el número de bandas verticales que deben ser generadas hasta que el Algoritmo BBV encuentra un cero $\mathbb{F}_q$–racional del polinomio de entrada.

En una primera etapa demostramos que la probabilidad $P[C = 1]$ de que el algoritmo de búsqueda en bandas verticales encuentre un cero $\mathbb{F}_q$–racional en la primera banda vertical coincide con la de que un polinomio univariado aleatorio con coeficientes en $\mathbb{F}_q$ tenga ceros $\mathbb{F}_q$–racionales.

Avanzando con el análisis de la complejidad en promedio de dicho algoritmo, obtenemos estimaciones explícitas de la probabilidad del evento $[C = s]$ cuando $s > 1$. Para ello, relacionamos dicha probabilidad con el promedio $\mathcal{V}(d, s)$, aplicamos las estimaciones obtenidas en los Capítulos 5 y 6, y demostramos el siguiente resultado.

Teorema 1.2.8. *Para $s \leq \binom{d/2+r-1}{r-1}$, tenemos que*

$$P[C = s] = (1 - \mu_d)^{s-1} \mu_d + \mathcal{O}(q^{-1}).$$

Por otro lado, si la característica de $\mathbb{F}_q$ es mayor a 2 y $s \leq \binom{d+r-3}{r-1}$, entonces

$$P[C = s] = (1 - \mu_d)^{s-1} \mu_d + \mathcal{O}(q^{-1/2}).$$

Este resultado indica que la probabilidad de que el algoritmo encuentre un punto $\mathbb{F}_q$–racional en las dos primeras bandas verticales de la hipersuperficie en consideración es del orden de $\mu_d(1 - \mu_d) \approx 0{,}8646\ldots$. Esto mejora los resultados obtenidos en el trabajo [Mat10], en donde se describe un algoritmo probabilístico que aplica la estrategia de búsqueda en bandas verticales para el cálculo de puntos $\mathbb{F}_q$–racionales de hipersuperficies y se prueba que con al menos d elecciones aleatorias es posible encontrar un punto $\mathbb{F}_q$–racional con probabilidad al menos $1/2$.

Luego determinamos el comportamiento en promedio del Algoritmo BBV. Para ello, consideramos la variable aleatoria X que cuenta el número de operaciones aritméticas en $\mathbb{F}_q$ realizadas por dicho algoritmo y determinamos el comportamiento asintótico del valor esperado de dicha variable. Utilizando el Teorema 1.2.8, demostramos los siguientes resultados.

Teorema 1.2.9. *Sean $r > 2$ y $s^* := \binom{d/2+r-1}{r-1}$. Entonces la complejidad en promedio del algoritmo BBV está acotada de la siguiente manera:*

$$E[X] \leq \tau(d, r, q)\big(\mu_d^{-1} + d(1 - d^{-1})^{s^*}\big) + \mathcal{O}(q^{-1/2}).$$

Por otro lado, para $r := 2$, $s^* := d/2 + 1$ *y* $\alpha^* := 1 - 1/\sqrt{s^*}$, *tenemos la siguiente cota superior para la complejidad en promedio del algoritmo BBV:*

$$E[X] \leq \tau(d, r, q)\left(\frac{1}{\alpha^{*2}}\left(\frac{1-\mu_d}{\mu_d} + \frac{1}{(d!)^2\mu_d^2}\right) + \frac{1}{\mu_d} + \left(1 - \frac{\mu_d}{\sqrt{s^*}}\right)^{s^*+1}\right) + \mathcal{O}(q^{-1}).$$

En ambos casos, $\tau(d, r, q)$ *es el costo de la búsqueda en una banda vertical arbitraria.*

Observamos que $1/\mu_d \approx 1{,}58\ldots$. Este es el número de bandas verticales que deben ser generadas en promedio.

Terminamos el capítulo exhibiendo algunas simulaciones que obtuvimos ejecutando una implementación del Algoritmo BBV en `Maple`, que confirman los resultados asintóticos obtenidos.

En el Capítulo 9 realizamos un análisis de la distribución de las salidas del Algoritmo BBV. Observamos que cada cero $\mathbb{F}_q$–racional que devuelve este algoritmo queda determinado por ciertas elecciones aleatorias que se realizan durante la ejecución del mismo. Cabe mencionar que, para un algoritmo ideal, las salidas están equidistribuidas, es decir, cada cero en $\mathbb{F}_q^r$ del polinomio en consideración tiene la misma probabilidad de resultar la salida del algoritmo. En este capítulo analizamos la distribución promedio de las salidas del algoritmo utilizando el concepto de entropía de Shannon. Siguiendo el trabajo de C. Beltrán y M. Pardo [BP11], definimos la entropía de Shannon H_F asociada a una entrada F del algoritmo BBV como $H_F := \sum -P_{\boldsymbol{x},F}\log(P_{\boldsymbol{x},F})$, donde la suma recorre todas las raíces de F y $P_{\boldsymbol{x},F}$ denota la probabilidad puntual de que el Algoritmo BBV obtenga como salida a $\boldsymbol{x}$ para la entrada F. Esta suma representa una medida de cuán concentradas están las salidas del algoritmo para una entrada del algoritmo. En tal sentido, analizamos la entropía promedio H del Algoritmo BBV cuando F varía entre todos los elementos de $\mathbb{F}_q[X_1, \ldots, X_r]_{\leq d}$ y demostramos que es parecida a la del algoritmo ideal. Más precisamente, demostramos el siguiente resultado.

Teorema 1.2.10.

$$H \geq \frac{1}{2\mu_d}\log(q^{r-1})(1 + \mathcal{O}(q^{-1})).$$

Observemos que $1/(2\mu_d) \approx 0{,}79$ para d suficientemente grande. Teniendo en cuenta que una cota superior de la entropía promedio de un algoritmo ideal es $\log(q^{r-1})$, el Teorema 1.2.10 indica que nuestro algoritmo es al menos un 79 % tan bueno como cualquier algoritmo ideal, desde el punto de vista de las distribución de las salidas.

En el Capítulo 10 analizamos la complejidad en promedio del algoritmo clásico de factorización aplicado a la familia lineal $\mathcal{A}$ definida en (1.1). Para ello, seguimos las ideas de [FGP01], reemplazando el estudio de singularidades de ciertas funciones generatrices asociadas a la factorización de los polinomios en consideración por las estimaciones explícitas del número de elementos de $\mathcal{A}$ con cierto patrón de factorización del Capítulo 7.

En la Sección 10.2 probamos que el costo promedio del algoritmo ERF (eliminación de factores repetidos) aplicado a los elementos de $\mathcal{A}$ es asintóticamente cercano

a $\mathcal{U}(d)$, cantidad que corresponde al costo del máximo común divisor de f con su derivada.

En la Sección 10.3 analizamos el costo promedio del algoritmo DDF (factorización en distintos grados) aplicado a los elementos de $\mathcal{A}$ que resultan libres de cuadrados. Para este análisis observamos que dicho algoritmo se aplica tantas veces como el grado del factor irreducible de mayor grado que aparece en el polinomio de entrada. Así, descomponemos la familia en consideración como la unión disjunta de los elementos de $\mathcal{A}$ libres de cuadrados cuyo patrón de factorización es $\boldsymbol{\lambda} := (\lambda_1, \ldots, \lambda_i, 0 \ldots, 0)$, con $1 \leq i \leq d$. Utilizando los resultados del Capítulo 7 sobre el número de elementos de $\mathcal{A}$ con patrón de factorización $\boldsymbol{\lambda}$, y cotas superiores sobre la longitud más grande esperada de los ciclos de una permutación aleatoria de d elementos (ver [GG98]), obtenemos el siguiente resultado.

Teorema 1.2.11. *El algoritmo DDF utiliza en promedio $\mathcal{O}(\mathcal{U}(d)\log(dq))$ operaciones aritméticas en $\mathbb{F}_q$ para calcular la factorización en distintos grados de un polinomio $f \in \mathcal{A}$ libre de cuadrados.*

Demostramos también que la probabilidad de que el algoritmo DDF complete la factorización de un polinomio aleatorio de $\mathcal{A}$ es del orden de $e^{-\gamma} + o(1)$, donde γ es la constante de Euler.

Por último, en la Sección 10.4 realizamos el análisis probabilístico del algoritmo EDF (factorización en grados iguales) aplicado a los elementos de $\mathcal{A}$. Utilizamos la misma estrategia que la del artículo [FGP01], que combina un proceso recursivo que permite "aislar" los factores irreducibles del polinomio de entrada con estimaciones asintóticas de la probabilidad de que un polinomio aleatorio de grado d tenga determinados factores irreducibles de un grado dado; en vez de utilizar dichos resultados asintóticos usamos nuestras estimaciones del número de elementos de $\mathcal{A}$ con un determinado patrón de factorización. Demostramos el siguiente resultado.

Teorema 1.2.12. *El algoritmo EDF utiliza en promedio $\mathcal{O}(M(d)\log q)$ operaciones aritméticas en $\mathbb{F}_q$ para un polinomio $f \in \mathcal{A}$.*

En resumen, a lo largo de esta tesis desarrollamos un nuevo enfoque que nos permitió obtener estimaciones que mejoran significativamente los resultados existentes sobre el cardinal promedio del conjunto de valores y la distribución de patrones de factorización en familias de polinomios univariados sobre $\mathbb{F}_q$ (Capítulos 5, 6 y 7). Nuestro enfoque reduce estas cuestiones combinatorias a la estimación del número de puntos $\mathbb{F}_q$–racionales de ciertas intersecciones completas singulares. En los Capítulos 3 y 4 obtenemos estimaciones sobre el número de puntos $\mathbb{F}_q$–racionales de dichas intersecciones completas, resultados que nos permiten obtener estimaciones explícitas de los problemas combinatorios que nos interesa. A su vez, estas estimaciones nos permiten analizar la complejidad en promedio de dos algoritmos probabilísticos: el Algoritmo BBV para hipersuperficies y el algoritmo clásico de factorización aplicado a familias lineales de polinomios (Capítulos 8, 9 y 10).

Capítulo 2

Preliminares

En este capítulo damos todas las definiciones, notaciones y resultados básicos de geometría algebraica que usaremos a lo largo de esta tesis. Para la exposición de estos resultados utilizamos principalmente los textos [Eis95], [Kun85] y [Sha94]. Además enunciamos algunos resultados clásicos sobre la cantidad de puntos $\mathbb{F}_q$–racionales de $\mathbb{F}_q$–variedades.

2.1. Definiciones y resultados básicos de geometría algebraica

Denotamos con $\mathbb{A}^n$ y $\mathbb{P}^n$ al espacio afín y proyectivo de dimensión n definido sobre $\overline{\mathbb{F}}_q$ respectivamente. Ambos son espacios topológicos con la **topología de Zariski** sobre $\overline{\mathbb{F}}_q$, según la cual los cerrados son los conjuntos de ceros comunes de polinomios en $\overline{\mathbb{F}}_q[X_1, \ldots, X_n]$, o de polinomios homogéneos en $\overline{\mathbb{F}}_q[X_0, \ldots, X_n]$ en el caso proyectivo.

Los conjuntos abiertos en la topología de Zariski de $\mathbb{A}^n$ o $\mathbb{P}^n$ son densos. En tal sentido, decimos que una propiedad sobre los elementos de $\mathbb{A}^n$ o $\mathbb{P}^n$ es **genérica** si la satisfacen todos los puntos que pertenecen a un abierto Zariski de $\mathbb{A}^n$ o $\mathbb{P}^n$.

Definición 2.1.1. *Sea $K := \mathbb{F}_q$ o $K := \overline{\mathbb{F}}_q$.*

*i) Un subconjunto $V \subset \mathbb{A}^n$ es una K-*variedad afín *o una* variedad afín de $\mathbb{A}^n$ definida sobre K *si es el conjunto de ceros comunes en $\mathbb{A}^n$ de polinomios $f_1, \ldots, f_m \in K[X_1, \ldots, X_n]$. En particular, una K-*hipersuperficie afín *es el conjunto de ceros en $\mathbb{A}^n$ de un único polinomio $f \in K[X_1, \ldots, X_n]$ no nulo.*

*ii) Un subconjunto $V \subset \mathbb{P}^n$ es una K-*variedad proyectiva *o una* variedad proyectiva de $\mathbb{P}^n$ definida sobre K *si es el conjunto de ceros comunes en $\mathbb{P}^n$ de polinomios homogéneos $f_1, \ldots, f_m \in K[X_0, \ldots, X_n]$. En particular, una K-*hipersuperficie proyectiva *es el conjunto de ceros en $\mathbb{P}^n$ de un único polinomio homogéneo $f \in K[X_0, \ldots, X_n]$ no nulo.*

Observemos que una K-variedad afín (respectivamente proyectiva) resulta un espacio topológico con la topología inducida de $\mathbb{A}^n$ (respectivamente de $\mathbb{P}^n$). Vamos

a denotar por $V(f_1, \ldots, f_m)$ o $\{f_1 = \cdots = f_m = 0\}$ a la K-variedad afín o proyectiva dada por el conjunto de ceros comunes de los polinomios $f_1, \ldots, f_m$.

Sea V una K-variedad en $\mathbb{A}^n$ o $\mathbb{P}^n$. Denotamos por $I(V)$ al **ideal de la variedad**, es decir el conjunto de polinomios en $K[X_1, \ldots, X_n]$ o en $K[X_0, \ldots, X_n]$ que se anulan en todos los puntos de V. Se sabe que $I(V)$ es un ideal radical. Vamos a denotar con $K[V]$ al **anillo coordenado** de V, o sea, $K[V]$ es el anillo cociente $K[X_1, \ldots, X_n]/I(V)$ o $K[X_0, \ldots, X_n]/I(V)$.

A continuación damos una serie de definiciones y propiedades que son válidas tanto para K-variedades afines como para K-variedades proyectivas, por tal motivo vamos a llamarlas simplemente K-variedades.

Definición 2.1.2. *Sea V una K–variedad. Entonces,*

(i) V se dice irreducible *si no puede escribirse como unión finita de K-variedades propias.*

(ii) V se dice absolutamente irreducible *si es irreducible como $\overline{\mathbb{F}}_q$-variedad.*

Toda K–variedad V es irreducible si y solo si su ideal $I(V)$ es primo. Asimismo, V es irreducible si y solo si todo subconjunto abierto no vacío de V es denso en V.

Toda K-variedad V puede descomponerse como una unión irredundante de K-variedades irreducibles, es decir, $V = C_1 \cup \cdots \cup C_s$ donde C_i son K–variedades irreducibles que cumplen que $C_i \not\subset C_j$ para todo $i \neq j$. A esta descomposición se la conoce como la **descomposición en componentes irreducibles** y es única salvo reordenamiento. Cada C_i se denomina una componente K-**irreducible** de V. En particular, las componentes $\overline{\mathbb{F}}_q$–irreducibles se denominan las componentes **absolutamente irreducibles** de V.

Dada una K-variedad V, definimos la **dimensión r** de V como la longitud de la mayor cadena de K-variedades irreducibles no vacías $V_0 \subsetneq V_1 \subsetneq \cdots \subsetneq V_r \subset V$ contenida en V. Decimos que una K-variedad es **equidimensional de dimensión r** o que tiene **dimensión pura r** si toda componente K-irreducible de dicha variedad tiene dimensión r. Observemos que una K–hipersuperficie es una K–variedad de $\mathbb{A}^n$ o $\mathbb{P}^n$ de dimensión pura $n-1$.

Damos la siguiente propiedad básica entre K-variedades [Sha94, Chapter 1, Section §6.1, Theorem 1]:

Teorema 2.1.3. *Sean V y W K–variedades.*

- *Si $V \subset W$ entonces* $\dim V \leq \dim W$.
- *Si W es irreducible y $V \subset W$ tal que* $\dim V = \dim W$, *entonces $V = W$.*

Sean $V \subset \mathbb{A}^n$ y $W \subset \mathbb{A}^m$ K-variedades. Una función $f : V \to W$ es un **morfismo regular** (de K–variedades) si existen polinomios $f_1, \ldots, f_m \in K[X_1, \ldots, X_n]$ tales que para todo $x \in V$, $f(x) = (f_1(x), \ldots, f_m(x))$. Decimos que f es un **morfismo dominante** si $\overline{f(V)} = W$, donde $\overline{f(V)}$ es la clausura de $f(V)$ con respecto a la topología Zariski de W. En esta situación, f induce, por composición, una extensión de anillos $K[W] \hookrightarrow K[V]$, y decimos que este morfismo es **finito** si dicha extensión

es entera, es decir, cada elemento $\eta \in K[V]$ satisface una ecuación mónica con coeficientes en $K[W]$.

En el caso en que V y W son K-variedades proyectivas, un morfismo $f : V \to W$ se dice **regular** si para cada $x \in V$ existen entornos afines $U \subset V$ de x y $U' \subset W$ de $f(x)$ tales que $f : U \to U'$ es regular. A su vez, definimos morfismo **dominante** de manera similar a como lo hicimos al caso afín y $f : V \to W$ se dice **finito** si para todo $y \in W$ existe un abierto afín W_y tal que $U := f^{-1}(W_y)$ es afín y $f : U \to W_y$ es un morfismo finito de variedades afines. A continuación damos una propiedad importante de los morfismos finitos (ver, por ejemplo, [Dan94, §4.2, Proposition]).

Teorema 2.1.4. *Sean V y W K–variedades y sea $f : V \to W$ un morfismo finito. Si $S \subset W$ es una subvariedad irreducible entonces la preimagen $f^{-1}(S)$ es una variedad de dimensión pura* $\dim S$.

El siguiente resultado se conoce como el Teorema de la dimensión de la fibra [Sha94, Chapter 1, Section §6.3, Theorem 7].

Teorema 2.1.5. *Sea $f : V \to W$ un morfismo regular entre K–variedades irreducibles. Supongamos que f sobreyectivo, y que* $\dim V = n$ *y* $\dim W = m$. *Entonces $m \leq n$, y*

1. $\dim F \geq n - m$ *para todo $w \in W$ y para toda componente F de la fibra $f^{-1}(w)$;*
2. *Existe un subconjunto abierto no vacío $U \subset W$ tal que* $\dim f^{-1}(w) = n - m$ *para todo $w \in U$.*

Tenemos también la siguiente propiedad de morfismos regulares entre K–variedades (ver, por ejemplo, [Kun85, Chapter §III. 2, Exercise 6]).

Teorema 2.1.6. *Sea $f : V \to W$ un morfismo regular entre K–variedades. Entonces:*

- *Si Z es una subvariedad irreducible de V, entonces $\overline{f(Z)}$ es irreducible, donde $\overline{f(Z)}$ denota la clausura de $f(Z)$ con respecto a la topología Zariski de W.*
- $\dim \overline{f(V)} \leq \dim V$.

Sea $V \subset \mathbb{A}^n$ una K–variedad de dimensión pura r y sea $f \in K[V]$. Si $W := V \cap \{f = 0\}$, entonces vale una y sólo una de las siguientes afirmaciones:

- $W = \emptyset$ (esto sucede cuando f es una unidad de $K[V]$);
- W tiene dimensión r (esto sucede cuando f es divisor de cero en $K[V]$).
- W tiene dimensión pura $r - 1$ (esto sucede cuando f no es divisor de cero ni unidad en $K[V]$).

En particular, si $f_1, \ldots, f_s$ son polinomios en $K[X_1, \ldots, X_n]$ y $W := V(f_1, \ldots, f_s) \subset \mathbb{A}^n$, entonces o bien $W = \emptyset$ o bien $\dim W \geq n - s$.

En el caso proyectivo tenemos el siguiente resultado.

Teorema 2.1.7 ([Sha94, Chapter 1, Section §2.6, Corollary 2])**.** *Sea $V \subset \mathbb{P}^n$ una variedad proyectiva de dimensión r y sea $W := V(g_1, \ldots, g_s)$ una subvariedad de V. Entonces toda componente no vacía irreducible de W tiene dimensión por lo menos $r - s$.*

Sea $V \subset \mathbb{A}^n$ una K-variedad afín, sea $I(V) \subset K[X_1, \ldots, X_n]$ el ideal de V y $x \in V$. La **dimensión** $\dim_x V$ de V en x es el máximo de las dimensiones de las componentes K-irreducibles de V que contienen a x. Si $I(V) = (f_1, \ldots, f_r)$, entonces el espacio tangente $\mathcal{T}_x V$ de V en x se define como el núcleo de la matriz Jacobiana $(\partial f_i/\partial X_j)_{1\le i\le r, 1\le j\le n}(x)$ de $f_1, \ldots, f_r$ con respecto a las variables $X_1, \ldots, X_n$ en x. Se tiene que si $g_1, \ldots, g_r \in I(V)$, entonces $\mathcal{T}_x V \subset \ker((\partial g_i/\partial X_j)_{1\le i\le r, 1\le j\le n}(x))$. Se satisface la siguiente desigualdad (ver, por ejemplo, [Sha94, página 94]):

$$\dim_x V \le \dim \mathcal{T}_x V.$$

Un punto x se dice **regular** si $\dim_x V = \dim \mathcal{T}_x V$. En caso que $\dim_x V < \dim \mathcal{T}_x V$, decimos que x es un punto **singular** de V. El conjunto de puntos singulares de V se denomina el **lugar singular** de V y lo notamos con Σ; se verifica que Σ es una K-subvariedad cerrada de V. Una K-variedad se dice **no singular o regular** si el conjunto de puntos singulares es vacío. Para una K-variedad proyectiva, los conceptos de espacio tangente, punto singular y regular se definen considerando un entorno afín del punto en cuestión.

Sea V una K-variedad afín o proyectiva cuya descomposición en componentes K irreducibles es $V = \cup_{i=1}^N \mathcal{C}_i$. Se satisface que $\mathcal{C}_i \cap \mathcal{C}_j \subset \Sigma$ para todo $i \ne j$ y que Σ no contiene componentes irreducibles de V. Además, si consideramos el lugar singular Σ_i de cada componente irreducible C_i, se tiene que $\Sigma = \bigcup_{i\ne j}(C_i \cap C_j) \cup \bigcup_i \Sigma_i$.

A cada K-variedad afín $V \subset \mathbb{A}^n$ podemos asociarle una K-variedad proyectiva $\mathrm{pcl}(V) \subset \mathbb{P}^n$, que llamamos la **clausura proyectiva** de V y definimos de la siguiente manera. Consideramos la inmersión de $\mathbb{A}^n$ en $\mathbb{P}^n$ que a cada $x = (x_1, \ldots, x_n) \in \mathbb{A}^n$ le asigna el punto $(1 : x_1 : \cdots : x_n) \in \mathbb{P}^n$. La clausura proyectiva $\mathrm{pcl}(V)$ es entonces la clausura con respecto a la topología Zariski en $\mathbb{P}^n$ de la imagen de V vía esta inmersión . Así, por definición, $\mathrm{pcl}(V)$ es la menor K-variedad proyectiva que contiene a V. Los puntos de $\mathrm{pcl}(V) \setminus V$ se llaman puntos de V **en el infinito**. Se verifican además las siguientes propiedades [Kun85, §I.5, Proposition 5.17 and Exercise 6; and §II.4, Proposition 4.1].

Teorema 2.1.8. *Sea $V \subset \mathbb{A}^n$ una K–variedad afín y sea $\mathrm{pcl}(V) \subset \mathbb{P}^n$ la clausura proyectiva de V. Entonces:*

(i) V es irreducible si y sólo si $\mathrm{pcl}(V)$ lo es.

(ii) Si $V = V_1 \cup V_2 \cup \cdots \cup V_r$ es la descomposición de V en K–variedades irreducibles entonces $\mathrm{pcl}(V) = \mathrm{pcl}(V_1) \cup \mathrm{pcl}(V_2) \cup \cdots \cup \mathrm{pcl}(V_r)$ es la descomposición de $\mathrm{pcl}(V)$ en componentes K-irreducibles.

(iii) V y $\mathrm{pcl}(V)$ tienen la misma dimensión.

(iv) El ideal $I(V)^h$ *de* $\mathrm{pcl}(V)$ *es el ideal generado por la homogeneización* $f^h \in K[X_0, \ldots, X_n]$ *de todos los polinomios* $f \in I(V) \subset K[X_1, \ldots, X_n]$. *Además,* $I(V)^h$ *es radical si y sólo si* $I(V)$ *lo es.*

Sea ahora $V \subset \mathbb{P}^n$ una K-variedad proyectiva. Consideramos el morfismo

$$\theta : \mathbb{A}^{n+1} \setminus \{(0, \ldots, 0)\} \to \mathbb{P}^n$$

que a un punto con coordenadas afines $(a_0, \ldots, a_n)$ le asocia el punto proyectivo con coordenadas homogéneas $(a_0 : \cdots : a_n)$. Se define el **cono afín** de V como la variedad afín

$$C(V) = \theta^{-1}(V) \cup \{(0, \ldots, 0)\}.$$

Se satisfacen las siguientes propiedades.

(*i*) $\dim C(V) = \dim V + 1$.

(*ii*) V es irreducible si y sólo si $C(V)$ lo es.

(*iii*) V es no singular si y sólo si $C(V)$ es no singular o su único punto singular es el origen.

2.1.1. Intersecciones completas

En esta sección vamos a considerar una familia particular de K-variedades, que se denominan intersecciones completas.

Definición 2.1.9. *Sea* V *una* K*-variedad de dimensión* r.

i) Decimos que V *es una* intersección completa conjuntista *si* V *es la intersección de* $n - r$ K*-hipersuperficies.*

ii) Decimos que V *es una* intersección completa *si* $I(V)$ *puede ser generado por* $n - r$ *polinomios en* $K[X_1, \ldots, X_n]$.

Una K-variedad V es **regular en codimensión m** si el lugar singular Σ de V tiene codimensión al menos $m + 1$ en V, es decir si $\dim V - \dim \Sigma \geq m + 1$. Una intersección completa se dice **normal** si es regular en codimensión 1. Un resultado importante para intersecciones completas proyectivas es el Teorema de Conexión de Hartshorne (ver, por ejemplo, [Kun85, Theorem VI.4.2]), que enunciamos a continuación. Si $V \subset \mathbb{P}^n$ es una intersección completa definida sobre K y $W \subset V$ es una K-subvariedad de codimensión al menos 2, entonces $V \setminus W$ es conexo con la topología Zariski de $\mathbb{P}^n$ sobre K. De acuerdo a este lema, considerando $W = \Sigma$, deducimos el siguiente resultado que utilizaremos frecuentemente.

Teorema 2.1.10. *Si* $V \subset \mathbb{P}^n$ *es una intersección completa normal, entonces* V *es absolutamente irreducible.*

Cada intersección completa resulta definida por polinomios que forman una sucesión regular.

Definición 2.1.11. *Sean $f_1, \ldots, f_{n-r} \in K[X_1, \ldots, X_n]$. Decimos que $f_1, \ldots, f_{n-r}$ forman una* sucesión regular *si f_1 no es el polinomio cero, cada f_i no es divisor de cero en el anillo $K[X_1, \ldots, X_n]/(f_1, \ldots, f_{i-1})$ para $2 \leq i \leq n-r$ y $V(f_1, \ldots, f_{n-r}) \neq \emptyset$.*

Si $f_1, \ldots, f_{n-r}$ forman una sucesión regular en $K[X_1, \ldots, X_n]$ o $K[X_0, \ldots, X_n]$, entonces la K-variedad afín o proyectiva que ellos definen es una intersección completa conjuntista y es de dimensión pura r. Más aún, si el ideal $(f_1, \ldots, f_{n-r})$ es radical entonces dicha variedad es una intersección completa.

2.1.2. El grado de una variedad

Sea V una K-variedad irreducible. Se define el **grado** $\deg(V)$ de V como el número máximo de puntos en la intersección de V con una variedad lineal L de codimensión $\dim V$ para la cual dicha intersección es finita. Más generalmente, si $V = C_1 \cup C_2 \cup \cdots \cup C_r$ es la descomposición de V en componentes K-irreducibles, definimos el grado de V como $\deg V := \sum_{i=1}^{r} \deg C_i$ (ver [Hei83]). El grado de una K-hipersuperficie H es el grado de un polinomio de grado mínimo que define a H. El grado de un abierto denso contenido en una K-variedad V es igual al grado de V. A continuación enunciamos una desigualdad de Bézout que usaremos para obtener las estimaciones (ver [Hei83, Ful84, Vog84]).

Teorema 2.1.12. *Si V y W son K-variedades, entonces*

$$\deg(V \cap W) \leq \deg V \cdot \deg W. \tag{2.1}$$

También usaremos el siguiente resultado.

Proposición 2.1.13 ([HS82, Proposition 2.3])**.** *Sean $V_1, \ldots, V_s$ K-variedades afines. Supongamos que $\dim V_1 = r$ y sea D el máximo de los grados de $V_2, \ldots, V_s$. Entonces $\deg(V_1 \cap \cdots \cap V_s) \leq \deg V_1 D^r$.*

Damos a continuación propiedades relacionadas con la noción de grado de K–variedades.

(*i*) Sean $V \subset \mathbb{A}^n$, $\mathrm{pcl}(V) \subset \mathbb{P}^n$ su clausura proyectiva y $\widehat{V} \subset \mathbb{A}^{n+1}$ el cono afín de $\mathrm{pcl}(V)$. Se satisface (ver, por ejemplo, [CGH91, Proposition 1.11]):

$$\deg V = \deg \mathrm{pcl}(V) = \deg \widehat{V}.$$

(*ii*) Sea $\phi : V \to W$ un morfismo regular lineal de K-variedades. Entonces [Hei83, Lemma 2],

$$\deg \overline{\phi(V)} \leq \deg V \tag{2.2}$$

donde $\overline{\phi(V)}$ es la clausura de Zariski de $\phi(V)$ con respecto a la topología Zariski de W.

Sea $V \subset \mathbb{P}^n$ una K-variedad intersección completa de grado δ y dimensión r, y sea $f_1, \ldots, f_{n-r}$ un conjunto de generadores homogéneos de $I(V)$. Los grados $d_1, \ldots, d_{n-r}$ dependen de V y no del sistema de generadores de $I(V)$. Sin pérdida de generalidad, podemos suponer que $d_1 \geq \cdots \geq d_{n-r}$. Definimos entonces el multigrado de V como $\boldsymbol{d} := (d_1, \ldots, d_{n-r})$. El siguiente es un resultado fundamental sobre intersecciones completas, que se denomina el **Teorema de Bézout**.

Teorema 2.1.14 ([Har92, Theorem 18.3])**.** *Sea $V \subset \mathbb{P}^n$ una intersección completa de grado δ, dimensión r y sean $f_1, \ldots, f_{n-r}$ generadores homogéneos de $I(V)$ de grados $d_1 \geq \cdots \geq d_{n-r}$ respectivamente. Entonces*

$$\delta = \prod_{i=1}^{n-r} d_i.$$

2.2. Puntos $\mathbb{F}_q$–racionales de $\mathbb{F}_q$–variedades

Sea V una $\mathbb{F}_q$-variedad afín o proyectiva. Dado $x \in V$, decimos que x es un punto $\mathbb{F}_q$–**racional** de V si todas sus coordenadas pertenecen a $\mathbb{F}_q$. Notamos por $V(\mathbb{F}_q)$ al conjunto de puntos $\mathbb{F}_q$–racionales de V. Cabe observar que el espacio afín de dimensión n sobre $\mathbb{F}_q$ tiene cardinal $|\mathbb{A}^n(\mathbb{F}_q)| = q^n$ y el espacio proyectivo de dimensión n sobre $\mathbb{F}_q$ tiene cardinal

$$p_n := |\mathbb{P}^n(\mathbb{F}_q)| = q^n + q^{n-1} + \cdots + q + 1.$$

Dada una $\mathbb{F}_q$-variedad V, estimar la cantidad de puntos $\mathbb{F}_q$-racionales de dicha variedad es un problema clásico de geometría aritmética. Dado que se conocen pocos resultados sobre la cantidad exacta de puntos $\mathbb{F}_q$-racionales, muchas veces resulta útil contar con estimaciones de dichos puntos. Observamos que las cotas superiores que se conocen del número $|V(\mathbb{F}_q)|$ de puntos $\mathbb{F}_q$–racionales de V se enuncian en términos de la dimensión y del grado de dicha variedad.

A continuación vamos a dar algunas cotas superiores conocidas.

2.2.1. Algunas cotas superiores

Proposición 2.2.1. *Sea V una variedad afín o proyectiva de dimensión r y grado δ definida en el espacio de dimensión n sobre $\overline{\mathbb{F}}_q$. Entonces la cantidad de puntos $\mathbb{F}_q$-racionales de V satisface*

(i) Si V es una variedad afín entonces $|V(\mathbb{F}_q)| \leq \delta q^r$.

(ii) Si V es una variedad proyectiva entonces $|V(\mathbb{F}_q)| \leq \delta p_r$.

En los trabajos [CM06b, Lemma 2.1] y [CM07, Proposition 3.1] los autores dan demostraciones de estos resultados que se basan en la aplicación de la desigualdad de Bézout (2.1). Para otras demostraciones podemos citar los trabajos [Lac96] y [LR15]. Encontramos también en [LR15, Proposition 4.3] un resultado análogo a la

Proposición 2.2.1 para el caso de intersecciones completas. Cabe mencionar que en [LN83, Theorems 6.13 y 6.15] se encuentran las demostraciones clásicas para el caso de hipersuperficies.

Podemos observar que, para el caso de variedades afines, la cota superior del Teorema 2.2.1 es óptima. Por ejemplo, en el caso de una $\overline{\mathbb{F}}_q$–hipersuperficie $H \subset \mathbb{A}^n$ definida por un polinomio $f \in \overline{\mathbb{F}}_q[X_1, \ldots, X_n]$ de grado δ, la cota superior $H(\mathbb{F}_q) \leq \delta \cdot q^{n-1}$ es óptima si $\delta \leq q$. En efecto, considerando el polinomio $f = (X_1 - c_1) \cdots (X_1 - c_\delta)$, siendo $c_1, \ldots, c_\delta$ elementos distintos en $\mathbb{F}_q$, la cantidad de puntos $\mathbb{F}_q$-racionales de la hipersuperficie definida por f es δq^{n-1}. Sin embargo, la cota del Teorema 2.2.1 no es óptima para el caso de variedades proyectivas. De hecho, J. P. Serre proporciona una cota más precisa para $\mathbb{F}_q$–hipersuperficies proyectivas, que enunciamos a continuación.

Proposición 2.2.2 ([Ser91]). *Sea $H \subset \mathbb{P}^n$ una $\mathbb{F}_q$–hipersuperficie de grado $\delta \leq q+1$. Entonces tenemos que $|H(\mathbb{F}_q)| \leq \delta q^{n-1} + p_{n-2}$.*

Por último, cabe mencionar que recientemente A. Couvreur obtuvo un resultado análogo al de Serre para el caso de variedades proyectivas equidimensionales.

Proposición 2.2.3 ([Cou16, Corollary 3.3]). *Sea $V \subset \mathbb{P}^n$ una variedad proyectiva de dimensión pura $d < n$ y de grado δ. Entonces,*

$$|V(\mathbb{F}_q)| \leq \delta(p_d - p_{2d-n}) + p_{2d-n}.$$

2.2.2. Estimaciones del número de puntos $\mathbb{F}_q$–racionales

En esta sección comenzamos exhibiendo fórmulas sobre el número promedio y la varianza del número de puntos $\mathbb{F}_q$–racionales de hipersuperficies (ver, por ejemplo, [LN83, Theorems 6.16 y 6.17]). Estos resultados permiten, intuir cómo estimar la cantidad de puntos $\mathbb{F}_q$-racionales de una hipersuperficie y el error que cometemos al estimar dicho número por el valor promedio. Luego mostramos estimaciones sobre la cantidad de puntos $\mathbb{F}_q$–racionales de hipersuperficies. También damos un resultado análogo al de hipersuperficies para el número promedio de puntos $\mathbb{F}_q$–racionales para $\mathbb{F}_q$–variedades y proporcionamos estimaciones sobre la cantidad de puntos $\mathbb{F}_q$–racionales de la misma.

Sea d un entero positivo y sean $X_1, \ldots, X_n$ indeterminadas sobre $\overline{\mathbb{F}}_q$. Sea $\boldsymbol{X} := (X_1, \ldots, X_n)$ y $\mathbb{F}_q[\boldsymbol{X}]$ el anillo de polinomios en $\boldsymbol{X}$ con coeficientes en $\mathbb{F}_q$. Denotamos con $\mathbb{F}_q[\boldsymbol{X}]_{\leq d}$ al conjunto de todos los polinomios en $\mathbb{F}_q[\boldsymbol{X}]$ de grado a lo sumo d. Si $N(F)$ es la cantidad de ceros $\mathbb{F}_q$-racionales de F, se satisface que

$$\frac{1}{|\mathbb{F}_q[\boldsymbol{X}]_{\leq d}|} \sum_{F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}} N(F) = q^{n-1}. \tag{2.3}$$

Con las mismas hipótesis, se tiene la siguiente fórmula para la desviación de este promedio:

$$\frac{1}{|\mathbb{F}_q[\boldsymbol{X}]_{\leq d}|} \sum_{F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}} (N(F) - q^{n-1})^2 = q^{n-1} - q^{n-2}. \tag{2.4}$$

De aquí se ve que un polinomio $F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$ tiene en promedio q^{n-1} ceros $\mathbb{F}_q$-racionales, y el error que se comete al estimar la cantidad de ceros $\mathbb{F}_q$-racionales de F por dicho promedio es del orden de $q^{\frac{n-1}{2}}$. Luego, si H es una $\mathbb{F}_q$-hipersuperficie afín, podemos estimar el número de puntos $\mathbb{F}_q$-racionales de la misma por la cantidad q^{n-1}. Sería de esperar entonces que, el error cometido $\big||H(\mathbb{F}_q)| - q^{n-1}\big|$ sea del orden de $q^{\frac{n-1}{2}}$. Desafortunadamente, esto no es cierto para una hipersuperficie afín arbitraria; por ejemplo, si H es relativamente irreducible se tiene el siguiente resultado.

Proposición 2.2.4 ([CM06b, Lemma 2.3])**.** *Sea $V \subset \mathbb{A}^n$ una $\mathbb{F}_q$-variedad relativamente irreducible de dimensión r y grado δ. Entonces $|V(\mathbb{F}_q)| \leq \delta^2\, q^{r-1}/4$.*

En el caso en que la $\mathbb{F}_q$-hipersuperficie $H \subset \mathbb{A}^n$ es absolutamente irreducible, es posible estimar en forma satisfactoria el error $||H(\mathbb{F}_q)| - q^{n-1}|$, como se expresa en el siguiente resultado.

Teorema 2.2.5 ([CM06b, Theorems 5.2 y 5.3])**.** *Sea $H \subset \mathbb{A}^n$ una $\mathbb{F}_q$-hipersuperficie absolutamente irreducible de grado δ. Entonces se satisface la siguiente estimación:*

$$\big||H(\mathbb{F}_q)| - q^{n-1}\big| \leq (\delta-1)(\delta-2)\, q^{n-3/2} + 5\delta^{13/3}\, q^{n-2}.$$

Si además $q > 15\,\delta^{13/3}$, entonces

$$\big||H(\mathbb{F}_q)| - q^{n-1}\big| \leq (\delta-1)(\delta-2)\, q^{n-3/2} + (5\delta^2+\delta+1)\, q^{n-2}.$$

Cabe aclarar de todas maneras que la absoluta irreducibilidad no es una restricción importante, dado que "casi todas" las hipersuperficies son absolutamente irreducibles (ver [vzGVZ13, Corollary 6.8]).

Por otro lado, se tiene un resultado similar al del número promedio de ceros de un polinomio con coeficientes en $\mathbb{F}_q$ para sistemas de polinomios.

Teorema 2.2.6. *Sea $\boldsymbol{d} = (d_1, \ldots, d_{n-r}) \in \mathbb{N}^{n-r}$ y $\Omega_{\boldsymbol{d}}$ el conjunto de $(n-r)$–uplas de polinomios*

$$\Omega_{\boldsymbol{d}} := \{\boldsymbol{F} := (f_1, \ldots, f_{n-r}) : \ \ f_i \in \mathbb{F}_q[\boldsymbol{X}],\ \deg f_i \leq d_i \ \text{para}\ 1 \leq i \leq n-r\}\,.$$

Se satisface entonces:

$$\frac{1}{|\Omega_{\boldsymbol{d}}|} \sum_{\boldsymbol{F} \in \Omega_{\boldsymbol{d}}} |V(\boldsymbol{F})(\mathbb{F}_q)| = q^r.$$

Demostración. Se tiene

$$\sum_{\boldsymbol{F} \in \Omega_{\boldsymbol{d}}} |V(\boldsymbol{F})(\mathbb{F}_q)| = \sum_{\boldsymbol{F} \in \Omega_{\boldsymbol{d}}} \sum_{\substack{x \in \mathbb{F}_q^n \\ \boldsymbol{F}(x) = \boldsymbol{0}}} 1 = \sum_{x \in \mathbb{F}_q^n} \sum_{\substack{\boldsymbol{F} \in \Omega_{\boldsymbol{d}} \\ \boldsymbol{F}(x) = \boldsymbol{0}}} 1 = \sum_{x \in \mathbb{F}_q^n} q^{\dim \Omega_{\boldsymbol{d}} - (n-r)} = |\Omega_{\boldsymbol{d}}| q^r.$$

El enunciado del teorema se sigue fácilmente. □

Análogamente, se puede obtener un resultado similar para la varianza, a saber:

$$\frac{1}{|\Omega_{\boldsymbol{d}}|}\sum_{\boldsymbol{F}\in\Omega_{\boldsymbol{d}}}(|V(\boldsymbol{F})(\mathbb{F}_q)|-q^r)^2 = q^r - q^{r-1}.$$

A partir de estos resultados, al igual que en el caso de hipersuperficies, podemos pensar en estimar la cantidad de puntos de una $\mathbb{F}_q$-variedad afín de dimensión r por q^r y esperar que el error cometido sea del orden $q^{\frac{r}{2}}$. En el caso en que la variedad en consideración sea absolutamente irreducible, tenemos los siguientes resultados (ver [GL02b, CM06b]).

Teorema 2.2.7 ([GL02b, Theorem 4.1])**.** *Sea $V \subset \mathbb{A}^n$ una $\mathbb{F}_q$-variedad absolutamente irreducible de dimensión r y grado δ. Entonces*

$$\big||V(\mathbb{F}_q)| - q^r\big| \leq (\delta-1)(\delta-2)q^{r-1/2} + 6\cdot 2^s(sd+3)^{n+1}q^{r-1},$$

donde s es el número de ecuaciones que definen a V y d es el grado máximo de las mismas.

Teorema 2.2.8 ([CM06b, Theorem 7.1])**.** *Sea $V \subset \mathbb{A}^n$ una $\mathbb{F}_q$-variedad absolutamente irreducible de dimensión r y grado δ. Si $q > 2(r+1)\delta^2$, entonces se satisface la siguiente estimación:*

$$\big||V(\mathbb{F}_q)| - q^r\big| \leq (\delta-1)(\delta-2)q^{r-1/2} + 5\,\delta^{13/3}q^{r-1}.$$

En el caso en que la variedad en consideración no es absolutamente irreducible, no es válido estimar la cantidad de puntos por q^r. Por ejemplo, en el trabajo [FHJ94, Proposition 3.3 (b)] (ver también [LR15, Proposition 3.8]) los autores prueban que una $\mathbb{F}_q$-variedad normal que no es absolutamente irreducible no tiene puntos $\mathbb{F}_q$-racionales.

Concluimos este capítulo dando estimaciones para intersecciones completas proyectivas definidas sobre $\mathbb{F}_q$. Un resultado conocido es la estimación de P. Deligne para intersecciones completas no singulares. Más precisamente, sea $V \subset \mathbb{P}^n$ una intersección completa no singular definida sobre $\mathbb{F}_q$, de dimensión r y multigrado $\boldsymbol{d}$. Entonces se satisface la siguiente estimación:

$$\big||V(\mathbb{F}_q)| - p_r\big| \leq b'_r(n,\boldsymbol{d})q^{r/2}, \tag{2.5}$$

donde $b'_r(n,\boldsymbol{d})$ es el r-ésimo número de Betti primitivo de V (ver, por ejemplo, [GL02a, Theorem 4.1] para una expresión explícita de $b'_r(n,\boldsymbol{d})$ en términos de n, r y $\boldsymbol{d}$).

Por su parte, Ghorpade y Lachaud dan la siguiente estimación para intersecciones completas arbitrarias (ver [GL02a, GL02b]).

Teorema 2.2.9. *Si $V \subset \mathbb{P}^n$ es una intersección completa definida sobre $\mathbb{F}_q$ de dimensión r y multigrado $\boldsymbol{d}$, cuyo lugar singular tiene dimensión a lo sumo $0 \leq s \leq r-2$, entonces*

$$\big||V(\mathbb{F}_q)| - p_r\big| \leq b'_{r-s-1}(n-s-1,\boldsymbol{d})q^{(r+s+1)/2} + C_s(V)q^{(r+s)/2}, \tag{2.6}$$

donde $C_s(V)$ es una constante independiente de q que se puede acotar por

$$C_s(V) \leq 9 \,.\, 2^{n-r}((n-r)d+3)^{n+1},$$

siendo $d := \max\{d_1, \ldots, d_{n-r}\}$ si $\boldsymbol{d} := (d_1, \ldots, d_{n-r})$.

En [CMP15a, Theorem 1.3] y [CMP15a, Corollary 8.4] los autores obtienen estimaciones que complementan las de los teoremas anteriores cuando las variedades en consideración son intersecciones completas proyectivas normales. Más precisamente, se tiene el siguiente resultado.

Teorema 2.2.10 ([CMP15a, Theorem 1.3])**.** *Sea $V \subset \mathbb{P}^n$ una intersección completa normal definida sobre $\mathbb{F}_q$ de dimensión $r \geq 2$, grado δ y multigrado $\boldsymbol{d} := (d_1, \ldots, d_{n-r})$. Entonces*

$$\big||V(\mathbb{F}_q)| - p_r\big| \leq (\delta(D-2)+2)q^{r-\frac{1}{2}} + 14D^2\delta^2 q^{r-1}, \tag{2.7}$$

donde $D := \sum_{i=1}^{n-r}(d_i - 1)$.

Por otro lado, el siguiente teorema muestra una estimación para intersecciones completas proyectivas regulares en codimensión 2.

Teorema 2.2.11 ([CMP15a, Corollary 8.4])**.** *Sea $V \subset \mathbb{P}^n$ una intersección completa definida sobre $\mathbb{F}_q$, de dimensión r y multigrado $\boldsymbol{d} := (d_1, \ldots, d_{n-r})$, la cual es regular en codimensión 2. Entonces*

$$\big||V(\mathbb{F}_q)| - p_r\big| \leq 14D^3\delta^2 q^{r-1}. \tag{2.8}$$

Para finalizar, cabe mencionar que en [MPP16a, Theorem 1.2] se muestra una estimación explícita para intersecciones completas proyectivas generales que complementa la del Teorema 2.2.9. El resultado es el siguiente.

Teorema 2.2.12. *Supongamos que $q \geq 2(s+1)D^{r-s-1}(D+r-s)\delta$ y sea $V \subset \mathbb{P}^n$ una intersección completa definida sobre $\mathbb{F}_q$, de dimensión r, multigrado $\boldsymbol{d} := (d_1, \ldots, d_{n-r})$, grado δ y lugar singular de dimensión a lo sumo s con $0 \leq s \leq r-2$. Entonces*

$$\big||V(\mathbb{F}_q)| - p_r\big| \leq \big(b'_{r-s-1}(n-s-1, \boldsymbol{d}) + 2\sqrt{\delta} + 1\big)q^{\frac{r+s+1}{2}}.$$

Capítulo 3

El lugar discriminante y variedades de incidencia asociados a familias lineales

El objetivo de este capítulo es, en primer lugar, demostrar la absoluta irreducibilidad del lugar discriminante asociado a ciertas familias lineales de polinomios univariados con coeficientes en $\mathbb{F}_q$. Luego, usando este resultado, analizamos la geometría de ciertas intersecciones completas singulares determinadas por tales familias y damos estimaciones de la cantidad de puntos $\mathbb{F}_q$–racionales de las mismas. Dichas estimaciones nos permitirán, más adelante, mejorar los resultados existentes sobre los problemas combinatorios sobre cuerpos finitos que estudiaremos: el comportamiento promedio del cardinal del conjunto de valores, y la distribución de patrones de factorización, en familias lineales de polinomios mónicos univariados con coeficientes en $\mathbb{F}_q$.

3.1. Irreducibilidad del discriminante

El lugar discriminante es un objeto de estudio clásico de la geometría algebraica (ver [GKZ94, Chapter 12] para una descripción de algunas de sus propiedades geométricas). En este capítulo vamos a estudiar el lugar discriminante de las familias de polinomios univariados asociadas a los problemas combinatorios que nos interesan. Esto nos permitirá obtener información importante sobre el lugar singular de algunas variedades algebraicas subyacentes a dichos problemas combinatorios.

Comenzamos definiendo el lugar discriminante de una familia cualquiera de polinomios univariados con coeficientes en $\overline{\mathbb{F}}_q$. Sea T una indeterminada sobre $\overline{\mathbb{F}}_q$. Sea $d > 2$ un entero positivo y sea $\overline{\mathbb{F}}_q[T]_d$ el conjunto de todos los polinomios mónicos en $\overline{\mathbb{F}}_q[T]$ de grado d. Para $\mathcal{A} \subset \overline{\mathbb{F}}_q[T]_d$, definimos el lugar discriminante $\mathcal{D}(\mathcal{A})$ de $\mathcal{A}$ como el conjunto de los elementos de $\mathcal{A}$ los cuales no son libres de cuadrados. Con un leve abuso de notación, vamos a identificar cada elemento $f_{\boldsymbol{a}_0} := T^d + a_{d-1}T^{d-1} + \cdots + a_0 \in \mathcal{A}$ con la d–upla $\boldsymbol{a}_0 := (a_{d-1}, \ldots, a_0) \in \mathbb{A}^d$, y consideramos $\mathcal{A}$ como un subconjunto de $\mathbb{A}^d$. Para $f_{\boldsymbol{a}_0} \in \mathcal{A}$, sea $\mathrm{Disc}(f_{\boldsymbol{a}_0}) := \mathrm{Res}(f_{\boldsymbol{a}_0}, f'_{\boldsymbol{a}_0})$

el discriminante de $f_{\boldsymbol{a}_0}$, esto es, la resultante de $f_{\boldsymbol{a}_0}$ y su derivada $f'_{\boldsymbol{a}_0}$. Observemos que $f_{\boldsymbol{a}_0} \in \mathcal{D}(\mathcal{A})$ si y solo si $\mathrm{Dis}(f_{\boldsymbol{a}_0}) = 0$. Sean $A_{d-1}, \ldots, A_0$ indeterminadas sobre $\overline{\mathbb{F}}_q$ y sea $\boldsymbol{A}_0 := (A_{d-1}, \ldots, A_0)$. Consideramos el polinomio $F \in \mathbb{F}_q[\boldsymbol{A}_0, T]$ definido como

$$F(\boldsymbol{A}_0, T) := T^d + A_{d-1}T^{d-1} + \cdots + A_0. \tag{3.1}$$

Dado que $f_{\boldsymbol{a}_0}$ tiene grado d, por una propiedad básica de las resultantes (ver, por ejemplo, [CLO92, §3.6, Proposition 3]) obtenemos que $\mathrm{Dis}(f_{\boldsymbol{a}_0}) = \mathrm{Dis}(F(\boldsymbol{A}_0, T))|_{\boldsymbol{A}_0=\boldsymbol{a}_0}$. Así, $\mathcal{D}(\mathcal{A})$ puede expresarse de la siguiente manera:

$$\mathcal{D}(\mathcal{A}) := \{\boldsymbol{a}_0 \in \mathcal{A} : \mathrm{Dis}(F(\boldsymbol{A}_0, T))|_{\boldsymbol{A}_0=\boldsymbol{a}_0} = 0\}. \tag{3.2}$$

Sea $\mathbb{F}_q[T]_d$ el conjunto de todos los polinomios en $\mathbb{F}_q[T]$ de grado d. En este capítulo vamos a considerar las familias lineales de polinomios en $\mathbb{F}_q[T]_d$ que describimos a continuación. Sean m y r enteros positivos tales que $3 \leq r \leq d - m$, sean $A_{d-1}, \ldots, A_r$ indeterminadas sobre $\overline{\mathbb{F}}_q$ y sean $L_1, \ldots, L_m \in \mathbb{F}_q[A_{d-1}, \ldots, A_r]$ las formas lineales afines definidas como sigue:

$$L_k := b_{k,d-1}A_{d-1} + \cdots + b_{k,r}A_r + b_{k,0} \quad (1 \leq k \leq m). \tag{3.3}$$

Sin pérdida de generalidad, podemos suponer que $L_1, \ldots, L_m$ son linealmente independientes. Sea $\boldsymbol{L} := (L_1, \ldots, L_m)$ y sea $\mathcal{A}_{\boldsymbol{L}} \subset \mathbb{F}_q[T]_d$ la familia lineal definida de la siguiente manera:

$$\mathcal{A}_{\boldsymbol{L}} := \{T^d + a_{d-1}T^{d-1} + \cdots + a_0 \in \mathbb{F}_q[T]_d : \boldsymbol{L}(a_{d-1}, \ldots, a_r) = \boldsymbol{0}\}. \tag{3.4}$$

Suponemos sin pérdida de generalidad que $M(\boldsymbol{L}) := (b_{k,d-j})_{1\leq k\leq m,\, 1\leq j\leq d-r}$ es una matriz escalonada por filas y denotamos con $1 \leq j_1 < \cdots < j_m \leq d - r$ a las posiciones de las columnas de $M(\boldsymbol{L})$ correspondientes a los pivotes. Consideramos también $\mathcal{L} \subset \mathbb{A}^d$ la variedad lineal definida por $L_1, \ldots, L_m$ y $\mathcal{D}(\mathcal{L}) \subset \mathbb{A}^d$ el lugar discriminante de $\mathcal{L}$.

En relación al lugar discriminante asociado a familias lineales de polinomios encontramos el trabajo [FS84], donde M. Fried y J. Smith demuestran el siguiente resultado sobre el lugar discriminante asociado a una familia de polinomios mónicos con ciertos coeficientes prescriptos.

Teorema 3.1.1 ([FS84, Proposición 3.1])**.** *Sean $\boldsymbol{j} := (j_1, \ldots, j_m)$ enteros no negativos tales que $1 \leq j_1 < \cdots < j_m \leq d$ y $\gcd(j_1, \ldots, j_m) = 1$ y sea $\boldsymbol{b} := (b_{j_1}, \ldots, b_{j_m}) \in \mathbb{F}_q^m$. Sea $\mathcal{A}_{\boldsymbol{j}}$ la familia lineal de polinomios de $\mathbb{F}_q[T]_d$ definida como*

$$\mathcal{A}_{\boldsymbol{j}} := \{T^d + a_{d-1}T^{d-1} + \cdots + a_0 \in \mathbb{F}_q[T]_d : a_{j_i} = b_{j_i} \quad (1 \leq i \leq m)\}.$$

Sea $\mathcal{L}_{\boldsymbol{j}} \subset \mathbb{A}^d$ la variedad lineal definida por $L_{j_i} := A_{j_i} - b_{j_i}$ para $1 \leq i \leq m$. Entonces existe $n(\boldsymbol{j}) \in \mathbb{N}$ tal que $\mathcal{D}(\mathcal{L}_{\boldsymbol{j}})$ es absolutamente irreducible si $\gcd(n(\boldsymbol{j}), p) = 1$.

Desafortunadamente, no podemos aplicar este resultado al lugar discriminante $\mathcal{D}(\mathcal{L})$ de la variedad $\mathcal{L}$ definida por las formas lineales $L_1, \ldots, L_m$ de (3.3), ya que, primero, la familia lineal $\mathcal{A}$ asociada a él consiste del conjunto de polinomios cuyos

coeficientes cumplen relaciones lineales (para los cuales el caso de coeficientes prescriptos es solo un caso particular), y segundo, necesitamos un resultado válido para cuerpos de característica pequeña. Es por esto que, en esta sección, demostramos que el lugar discriminante $\mathcal{D}(\mathcal{L})$ resulta una hipersuperficie absolutamente irreducible cuando $p > 2$, extendiendo así de forma significativa el resultado de [FS84].

Para ello, introducimos la noción de polinomios homogéneos con peso y algunos resultados básicos al respecto que serán necesarios. Sea $\mathbb{K}$ un cuerpo cualquiera. Sea $\mathbb{K}[X_1, \ldots, X_n]$ el anillo de polinomios multivariados con coeficientes en $\mathbb{K}$. Dados enteros positivos $a_1, \ldots, a_n$, definimos el peso $\mathsf{wt}(\boldsymbol{X}^{\boldsymbol{\alpha}})$ de un monomio $\boldsymbol{X}^{\boldsymbol{\alpha}} := X_1^{\alpha_1} \cdots X_n^{\alpha_n}$ como $\mathsf{wt}(\boldsymbol{X}^{\boldsymbol{\alpha}}) := \sum_{i=1}^{n} a_i \cdot \alpha_i$. De aquí se deduce que el peso de cada variable X_i es a_i. Se define el peso $\mathsf{wt}(F)$ de un elemento arbitrario $F \in \mathbb{K}[X_1, \ldots, X_n]$ como el mayor de los pesos de todos los monomios con coeficientes no nulos que aparecen en la representación densa de F.

Un elemento $F \in \mathbb{K}[X_1, \ldots, X_n]$ se dice homogéneo con peso o cuasi–homogéneo (con respecto al peso wt definido arriba) si todos sus términos tienen el mismo peso. Observamos que F es homogéneo con peso si y solo si $F(t^{a_1}X_1, \ldots, t^{a_n}X_n) = t^{\mathsf{wt}(F)}F(X_1, \ldots, X_n)$ para cada $t \in \mathbb{K} \setminus \{0\}$. Equivalentemente, F es homogéneo con peso si y solo si $F(X_1^{a_1}, \ldots, X_n^{a_n})$ es un polinomio homogéneo en los X_i de grado $\mathsf{wt}(F)$.

Todo polinomio $F \in \mathbb{K}[X_1, \ldots, X_n]$ puede escribirse de manera única como una suma de polinomios homogéneos con peso $F = \sum_i F_i$, donde cada F_i es un polinomio homogéneo con $\mathsf{wt}(F_i) = i$. Los polinomios F_i se llaman las componentes homogéneas con peso de F. En lo que sigue usaremos la propiedad que enunciamos a continuación.

Lema 3.1.2 ([HH11, Proposición 3.3.7])**.** *Sea $F \in \mathbb{K}[X_1, \ldots, X_n]$ un polinomio de grado positivo. Si la componente $F_{\mathsf{wt}(F)}$ de mayor peso de F es irreducible en $\mathbb{K}[X_1, \ldots, X_n]$, entonces F es irreducible en $\mathbb{K}[X_1, \ldots, X_n]$.*

También usaremos los siguientes criterios de irreducibilidad para polinomios multivariados.

Lema 3.1.3. *Sea $F \in \mathbb{K}[X_1, \ldots, X_n]$ un polinomio de grado positivo, $s < n$,*

$$R := \mathbb{K}[X_1, \ldots, X_s] \text{ y } Q(R) := \mathbb{K}(X_1, \ldots, X_s).$$

Si F es un polinomio primitivo de $R[X_{s+1}, \ldots, X_n]$ y es un elemento irreducible de $Q(R)[X_{s+1}, \ldots, X_n]$, entonces F es irreducible en $\mathbb{K}[X_1, \ldots, X_n]$.

Demostración. El resultado es una consecuencia inmediata del lema de Gauss. □

Lema 3.1.4 ([Gib98, Lema 3.15])**.** *Sean $F, G \in \mathbb{K}[X_1, \ldots, X_n]$ polinomios homogéneos de grado d y $d+1$ respectivamente, sin factores comunes. Entonces $F+G$ es irreducible en $\mathbb{K}[X_1, \ldots, X_n]$.*

Lema 3.1.5 ([LN83, Lema 6.54])**.** *Sea $f \in \mathbb{K}[T]$ un polinomio no constante, y $m \in \mathbb{N}$. Supongamos que f se factoriza en $\overline{\mathbb{K}}$ como $f(T) = a(T-\alpha_1)^{e_1} \ldots (T-\alpha_d)^{e_d}$ con $\alpha_i \neq \alpha_j$ si $i \neq j$. Entonces el polinomio $X^m - f(T)$ es irreducible sobre $\overline{\mathbb{K}}$ si y solo si* $\gcd(m, e_1, \ldots, e_d) = 1$

Observemos que $\boldsymbol{a}_0 \in \mathcal{D}(\mathcal{L})$ si y solo si $\boldsymbol{a}_0 \in \mathbb{A}^d$ es tal que $\mathrm{Dis}(F(\boldsymbol{A}_0, T))|_{\boldsymbol{A}_0=\boldsymbol{a}_0} = 0$ y $L_k(\boldsymbol{A}_0)|_{\boldsymbol{A}_0=\boldsymbol{a}_0} = 0$ $(1 \leq k \leq m)$, donde $F(\boldsymbol{A}_0, T)$ es el polinomio definido en (3.1) y $L_1, \ldots, L_m$ son las formas lineales afines definidas en (3.3). De la forma escalonada de las ecuaciones $L_1(\boldsymbol{A}_0) = \cdots = L_m(\boldsymbol{A}_0) = 0$ deducimos que, para $1 \leq l \leq m$ existe un polinomio $h_l \in \overline{\mathbb{F}}_q[A_k : k \in \mathcal{J}]$ de grado 1 tal que $A_{d-j_l} = h_l(A_k : k \in \mathcal{J})$, donde $\mathcal{J} := \{d-1, \ldots, 0\} \setminus \{d-j_1, \ldots, d-j_m\}$ y donde $1 \leq j_1 < \cdots < j_m \leq d-r$. Sea $\hat{\boldsymbol{A}}_0 \in \overline{\mathbb{F}}_q[A_k : k \in \mathcal{J}]^d$ el elemento que se obtiene de sustituir en la coordenada A_{d-j_l} de $\boldsymbol{A}_0$ el polinomio h_l, para $l = 1, \ldots, m$. Concluimos que $\boldsymbol{a}_0 \in \mathcal{D}(\mathcal{L})$ si y solo si $\mathrm{disc}(F(\hat{\boldsymbol{A}}_0, T))|_{\boldsymbol{A}_0=\boldsymbol{a}_0} = 0$, donde $\mathrm{disc}(F(\hat{\boldsymbol{A}}_0, T)) \in \overline{\mathbb{F}}_q[A_k : k \in \mathcal{J}]$ es el discriminante de $F(\hat{\boldsymbol{A}}_0, T) \in \overline{\mathbb{F}}_q[A_k : k \in \mathcal{J}, T]$ con respecto a la variable T.

En toda esta sección vamos a considerar el peso wt en $\overline{\mathbb{F}}_q[A_k : k \in \mathcal{J}]$ definido por $\mathsf{wt}(A_k) := d-k$ con $k \in \mathcal{J}$. Observamos que, extendiendo esta noción de peso al anillo de polinomios $\overline{\mathbb{F}}_q[A_d, \ldots, A_0]$, es decir, definiendo $\mathrm{wt}(A_k) := d-k$ para $1 \leq k \leq d$, tenemos el siguiente resultado.

Lema 3.1.6 ([FS84, Lema 2.2])**.** *El discriminante* $\mathrm{disc}(F) \in \overline{\mathbb{F}}_q[A_d, \ldots, A_0]$ *de un polinomio genérico* $F \in \overline{\mathbb{F}}_q[A_d, \ldots, A_0][T]$ *de grado* d *es un polinomio homogéneo con peso, de peso* $d(d-1)$*.*

A continuación vamos a probar el resultado más importante de esta sección, que asegura la absoluta irreducibilidad de $\mathrm{disc}(F(\hat{\boldsymbol{A}}_0, T)) \in \overline{\mathbb{F}}_q[A_k : k \in \mathcal{J}]$.

Teorema 3.1.7. *Sea* $p > 2$*,* $q > d$ *y* $3 \leq r \leq d-m$*. Entonces* $\mathrm{disc}(F(\hat{\boldsymbol{A}}_0, T))$ *es un polinomio irreducible en* $\overline{\mathbb{F}}_q[A_k : k \in \mathcal{J}]$*, donde* $\mathcal{J} := \{d-1, \ldots, 0\} \setminus \{d-j_1, \ldots, d-j_m\}$ *y* $1 \leq j_1 < \cdots < j_m \leq d-r$ *son las posiciones de las columnas correspondientes a los pivotes de la matriz* $M(\boldsymbol{L})$ *definida más arriba.*

Demostración. Supongamos primero que p no divide a $d(d-1)$. Sea $\mathsf{K}_2 := \overline{\mathbb{F}}_q(A_k : k \in \mathcal{J}_1)$, donde $\mathcal{J}_1 := \{d-1, \ldots, 2\} \setminus \{d-j_1, \ldots, d-j_m\}$. Consideramos $\mathrm{disc}(F(\hat{\boldsymbol{A}}_0, T))$ como un elemento en $\mathsf{K}_2[A_1, A_0]$ y consideramos el peso w_2 sobre $\mathsf{K}_2[A_1, A_0]$ definido por $\mathrm{w}_2(A_0) := d$ y $\mathrm{w}_2(A_1) := d-1$. Es fácil ver que la componente homogénea de mayor peso de $\mathrm{disc}(F(\hat{\boldsymbol{A}}_0, T))$ es $\Delta_2 := d^d A_0^{d-1} + (-1)^{d-1}(d-1)^{d-1} A_1^d$. Por la hipótesis sobre p tenemos que ninguno de los dos monomios de Δ_2 se anula. Más aún, por el Lema 3.1.5 deducimos que Δ_2 es irreducible en $\mathsf{K}_2[A_1, A_0]$. Por el Lema 3.1.2 concluimos que $\mathrm{disc}(F(\hat{\boldsymbol{A}}_0, T))$ es un elemento irreducible en $\mathsf{K}_2[A_1, A_0]$. Finalmente, como $\mathrm{disc}(F(\hat{\boldsymbol{A}}_0, T))$ es un polinomio primitivo de $\overline{\mathbb{F}}_q\big[A_k : k \in \mathcal{J}_1\big][A_1, A_0]$, el Lema 3.1.3 muestra que $\mathrm{disc}(F(\hat{\boldsymbol{A}}_0, T))$ es irreducible en $\overline{\mathbb{F}}_q[A_k : k \in \mathcal{J}]$.

Supongamos ahora que p divide d. Sea $\mathsf{K}_3 := \overline{\mathbb{F}}_q(A_k : k \in \mathcal{J}_2)$, donde $\mathcal{J}_2 := \{d-1, \ldots, 3\} \setminus \{d-j_1, \ldots, d-j_m\}$. Consideramos $\mathrm{disc}(F(\hat{\boldsymbol{A}}_0, T))$ como un elemento de $\mathsf{K}_3[A_2, A_1, A_0]$. Consideramos el peso w_3 sobre $\mathsf{K}_3[A_2, A_1, A_0]$ definido por $\mathrm{w}_3(A_0) = d$, $\mathrm{w}_3(A_1) := d-1$ y $\mathrm{w}_3(A_2) := d-2$.

Si $G := T^d + A_2 T^2 + A_1 T + A_0$, entonces $G' = 2A_2 T + A_1$. Por lo tanto, aplicando la fórmula de Poisson para la resultante es fácil probar que

$$\begin{aligned} \mathrm{disc}(G) = \mathrm{Res}(G, G', T) &= (-1)^d (2A_2)^d G\left(\frac{-A_1}{2A_2}\right) \\ &= A_1^d + (-1)^{d+1} 2^{d-2} A_2^{d-1} A_1^2 + (-1)^d 2^d A_2^d A_0. \end{aligned}$$

Como el discriminante de un polinomio genérico de grado d es homogéneo con peso, de peso $d(d-1)$, $\deg F(\hat{\boldsymbol{A}}_0, T) = \deg G = d$ y $\operatorname{disc}(G)$ es un término de $\operatorname{disc}(F(\hat{\boldsymbol{A}}_0, T))$, que resulta homogéneo con peso, de peso $d(d-1)$, es fácil ver que $\operatorname{disc}(G)$ es la componente de mayor peso de $\operatorname{disc}(F(\hat{\boldsymbol{A}}_0, T)) \in \mathsf{K}_3[A_2, A_1, A_0]$. Más aún, afirmamos que $\operatorname{disc}(G)$ es irreducible in $\mathsf{K}_3[A_2, A_1, A_0]$. En efecto, si consideramos $\operatorname{disc}(G)$ como un polinomio en $\mathsf{K}_3(A_0)[A_2, A_1]$, vemos que $\operatorname{disc}(G)$ es la suma de dos polinomios homogéneos de grados d y $d+1$ que no tienen factores en común, esto es, $A_1^d + (-1)^d 2^d A_2^d A_0$ y $(-1)^{d+1} 2^{d-2} A_2^{d-1} A_1^2$ respectivamente. Entonces, por el Lema 3.1.4 tenemos que $\operatorname{disc}(G)$ es irreducible in $\mathsf{K}_3(A_0)[A_2, A_1]$. Como además $\operatorname{disc}(G)$ es un polinomio primitivo en $\mathsf{K}_3[A_0][A_2, A_1]$, resulta a su vez irreducible en $\mathsf{K}_3[A_2, A_1, A_0]$ por el Lema 3.1.3. Combinando este resultado con el Lema 3.1.2 deducimos que $\operatorname{Disc}(F(\hat{\boldsymbol{A}}_0, T))$ es irreducible en $\mathsf{K}_3[A_2, A_1, A_0]$. Como $\operatorname{Disc}(F(\hat{\boldsymbol{A}}_0, T))$ es un polinomio primitivo en $\overline{\mathbb{F}}_q\big[A_k : k \in \mathcal{J}_2\big][A_2, A_1, A_0]$, aplicando nuevamente el Lema 3.1.3 concluimos que $\operatorname{Disc}(F(\hat{\boldsymbol{A}}_0, T))$ es irreducible en $\overline{\mathbb{F}}_q[A_k : k \in \mathcal{J}]$.

Finalmente, supongamos que p divide a $d-1$ y consideremos $\operatorname{disc}(F(\hat{\boldsymbol{A}}_0, T))$ como un elemento de $\mathsf{K}_3[A_2, A_1, A_0]$. Argumentando como arriba podemos concluir que el discriminante $\operatorname{disc}(G)$ del polinomio $G := T^d + A_2T^2 + A_1T + A_0$ es la componente homogénea de mayor peso de $\operatorname{disc}(F(\hat{\boldsymbol{A}}_0, T))$. Observemos que $G' = T^{d-1} + 2A_2T + A_1$. Así, utilizando propiedades elementales de la resultante, tenemos que

$$\begin{aligned}\operatorname{disc}(G) = \frac{\operatorname{Res}_T(G, TG' - G)}{\operatorname{Res}_T(G, T)} &= \frac{\operatorname{Res}_T(G, A_2T^2 - A_0)}{\operatorname{Res}_T(G, T)} \\ &= \frac{\operatorname{Res}_T(G - (A_2T^2 - A_0), A_2T^2 - A_0)}{\operatorname{Res}_T(G, T)} \\ &= \frac{\operatorname{Res}_T(T^d + A_1T + 2A_0, A_2T^2 - A_0)}{\operatorname{Res}_T(G, T)}.\end{aligned}$$

Aplicando la fórmula de Poisson para la resultante, es fácil deducir que:

$$\operatorname{disc}(G) = \begin{cases} 4A_2^dA_0 + A_0^{d-1} + 4A_0^{d/2}A_2^{d/2} - A_1^2A_2^{d-1} & \textit{para } d \textit{ par}, \\ -4A_2^dA_0 + A_0^{d-1} + 2A_1A_0^{\frac{d-1}{2}}A_2^{\frac{d-1}{2}} + A_1^2A_2^{d-1} & \textit{para } d \textit{ impar}. \end{cases}$$

Entonces $\operatorname{disc}(G)$ es irreducible en $\overline{\mathbb{F}}_q[A_0, A_2][A_1]$ por el criterio de Eisenstein (tomando como primo A_0) y así tenemos que $\operatorname{disc}(F(\hat{\boldsymbol{A}}_0, T))$ es irreducible en $\mathsf{K}_3[A_2, A_1, A_0]$ por el Lema 3.1.2. Argumentando como arriba tenemos que $\operatorname{disc}(F(\hat{\boldsymbol{A}}_0, T))$ es irreducible en $\overline{\mathbb{F}}_q[A_k : k \in \mathcal{J}]$, finalizando así la demostración del teorema. □

Del teorema anterior deducimos fácilmente el siguiente resultado sobre el lugar discriminante $\mathcal{D}(\mathcal{L})$, donde $\mathcal{L}$ es la variedad lineal definida por las formas lineales afines $L_1, \ldots, L_m$ de (3.3).

Corolario 3.1.8. *Sean $p > 2$, $q > d$ y $3 \leq r \leq d - m$. Entonces $\mathcal{D}(\mathcal{L}) \subset \mathbb{A}^{d-m}$ es una $\mathbb{F}_q$–hipersuperficie absolutamente irreducible.*

Para terminar esta sección, sea s un entero positivo tal que $1 \leq s \leq d-2$. Dado $\boldsymbol{a} := (a_{d-1}, \ldots, a_{d-s}) \in \mathbb{F}_q^s$. Consideramos la familia $\mathcal{A}_{\boldsymbol{a}}$ de polinomios en $\mathbb{F}_q[T]_d$ con los primeros s coeficientes fijos, es decir,

$$\mathcal{A}_{\boldsymbol{a}} := \{T^d + a_{d-1}T^{d-1} + \cdots + a_{d-s-1}T^{d-s-1} + \cdots + a_0 : a_{d-s-1}, \ldots, a_0 \in \mathbb{F}_q\}.$$

Observemos que esta familia es un caso particular de la familia $\mathcal{A}_{\boldsymbol{L}}$. Por lo tanto, si consideramos las formas lineales afines $L_i^{\boldsymbol{a}} := A_{d-i} - a_{d-i}$ para $1 \leq i \leq s$ y la variedad lineal $\mathcal{L}^{\boldsymbol{a}} \subset \mathbb{A}^{d-s}$ definida por $L_1^{\boldsymbol{a}}, \ldots, L_s^{\boldsymbol{a}}$, tomando $r := d-s$ y $m := s$ deducimos del Corolario 3.1.8 el siguiente resultado.

Corolario 3.1.9. *Sean $p > 2$ y $q > d$, y sea $1 \leq s \leq d-3$. Entonces $\mathcal{D}(\mathcal{L}^{\boldsymbol{a}}) \subset \mathbb{A}^{d-s}$ es una $\mathbb{F}_q$–hipersuperficie absolutamente irreducible.*

3.2. Estimaciones para variedades de incidencia

En esta sección presentamos estimaciones de la cantidad de puntos $\mathbb{F}_q$–racionales de ciertas variedades de incidencia definidas sobre $\mathbb{F}_q$ asociadas a la familia lineal $\mathcal{A}_{\boldsymbol{L}}$ dada en (3.4). Estas estimaciones nos permitirán más adelante dar cotas superiores explícitas sobre el cardinal promedio del conjunto de valores de dichas familias lineales y el número de elementos de $\mathcal{A}_{\boldsymbol{L}}$ con determinado patrón de factorización.

A continuación definimos las variedades de incidencia a la que hacemos referencia. Sean m, r y d enteros positivos tales que $3 \leq r \leq d-m$. Fijamos i con $r+1 \leq i \leq d$. Sean $A_{d-1}, \ldots, A_0$ indeterminadas sobre $\overline{\mathbb{F}}_q$ y sean $L_1, \ldots, L_m \in \mathbb{F}_q[A_{d-1}, \ldots, A_r]$ las formas lineales afines de (3.3) que definen la familia $\mathcal{A}_{\boldsymbol{L}}$. Sean $\boldsymbol{A} := (A_{d-1}, \ldots, A_1)$ y $\boldsymbol{A}_0 := (\boldsymbol{A}, A_0)$. Sean $T, T_1, \ldots, T_i$ nuevas indeterminadas sobre $\overline{\mathbb{F}}_q$ y denotemos $\boldsymbol{T} := (T_1, \ldots, T_i)$. Consideramos el polinomio $F \in \mathbb{F}_q[\boldsymbol{A}_0, T]$ definido en (3.1). Observemos que si $\boldsymbol{a}_0 \in \mathbb{F}_q^d$, entonces podemos escribir $F(\boldsymbol{a}_0, T) = f + a_0$, donde $f \in \mathbb{F}_q[T]$ es un polinomio mónico de grado d con $f(0) = 0$.

Consideremos la $\mathbb{F}_q$–cuasi–variedad afín $\Gamma_i \subset \mathbb{A}^{d+i}$ definida por el polinomio $F(\boldsymbol{A}_0, T)$ y las m formas lineales afines $L_k(\boldsymbol{A}_0)$, es decir:

$$\begin{aligned}\Gamma_i := \{(\boldsymbol{a_0}, \boldsymbol{\alpha}) \in \mathbb{A}^{d+i} : F(\boldsymbol{a_0}, \alpha_j) = 0 \ (1 \leq j \leq i),\ \alpha_j \neq \alpha_k \ (1 \leq j < k \leq i), \\ L_1(\boldsymbol{a_0}) = \cdots = L_m(\boldsymbol{a_0}) = 0\}.\end{aligned} \tag{3.5}$$

A efectos del análisis de los problemas combinatorios sobre cuerpos finitos que ya mencionamos, como veremos más adelante, vamos a estimar la cantidad de puntos $\mathbb{F}_q$–racionales de Γ_i. Para ello, vamos a considerar la clausura Zariski $\overline{\Gamma}_i$ de $\Gamma_i \subset \mathbb{A}^{d+i}$ y obtener ecuaciones que definan dicha clausura. Para este propósito, usamos la siguiente notación. Sean $X_1, \ldots, X_{l+1}$ indeterminadas sobre $\overline{\mathbb{F}}_q$ y sea $f \in \overline{\mathbb{F}}_q[T]$ un polinomio de grado a lo sumo l. Por conveniencia de notaciones, definimos la diferencia dividida $\Delta^0 f \in \overline{\mathbb{F}}_q[X_1]$ de orden 0 de f como $\Delta^0 f := f(X_1)$. Para $1 \leq j \leq l$ definimos la diferencia dividida $\Delta^j f \in \overline{\mathbb{F}}_q[X_1, \ldots, X_{j+1}]$ de orden j de f como

$$\Delta^j f(X_1, \ldots, X_{j+1}) = \frac{\Delta^{j-1} f(X_1, \ldots, X_j) - \Delta^{j-1} f(X_1, \ldots, X_{j-1}, X_{j+1})}{X_j - X_{j+1}}. \tag{3.6}$$

Con estas notaciones, definimos la siguiente $\mathbb{F}_q$–variedad $\Gamma_i^* \subset \mathbb{A}^{d+i}$:

$$\Gamma_i^* := \{(\boldsymbol{a}_0, \boldsymbol{\alpha}) \in \mathbb{A}^d \times \mathbb{A}^i : \Delta^{j-1} F(\boldsymbol{a}_0, \alpha_1, \ldots, \alpha_j) = 0 \ (1 \leq j \leq i), \\ L_k(\boldsymbol{a_0}) = 0 \, (1 \leq k \leq m)\} \tag{3.7}$$

donde $\Delta^{j-1} F(\boldsymbol{a}_0, T_1, \ldots, T_j)$ denota la diferencia dividida de orden $j-1$ de $F(\boldsymbol{a}_0, T) \in \overline{\mathbb{F}}_q[T]$.

En las próximas secciones, suponiendo que $p > 2$, probamos que las variedades de incidencia Γ_i^* resultan ser intersecciones completas definidas sobre $\mathbb{F}_q$ con buen comportamiento en el infinito y cuyo lugar singular tiene codimensión al menos 2. Esto nos permitirá aplicar las estimaciones sobre intersecciones completas proyectivas normales que se encuentran en el trabajo [CMP15a] (ver Teorema 2.2.10) y dar estimaciones sobre la cantidad de puntos $\mathbb{F}_q$–racionales de Γ_i^*.

3.2.1. Aspectos geométricos

En esta sección discutimos ciertas propiedades geométricas de las variedades de incidencia $\Gamma_i^* \subset \mathbb{A}^{d+i}$, con $r+1 \leq i \leq d$. El primer resultado muestra la relación entre la cuasi–variedad Γ_i y la variedad Γ_i^* .

Lema 3.2.1. *Sean m, r y d enteros positivos tales que $3 \leq r \leq d-m$. Sea i un entero tal que $r+1 \leq i \leq d$. Entonces tenemos la siguiente identidad:*

$$\Gamma_i = \Gamma_i^* \cap \{(\boldsymbol{a}_0, \boldsymbol{\alpha}) : \alpha_j \neq \alpha_k \ (1 \leq j < k \leq i)\}. \tag{3.8}$$

Demostración. Sea $(\boldsymbol{a}_0, \boldsymbol{\alpha})$ un punto arbitrario de Γ_i. Por la definición de la diferencia dividida de $F(\boldsymbol{a}_0, T)$ es fácil concluir que $(\boldsymbol{a}_0, \boldsymbol{\alpha}) \in \Gamma_i^*$. Por otro lado, sea $(\boldsymbol{a}_0, \boldsymbol{\alpha})$ un punto arbitrario perteneciente al conjunto del lado derecho de (3.8). Afirmamos que $F(\boldsymbol{a}_0, \alpha_j) = 0$ para $1 \leq j \leq i$. Observamos que $F(\boldsymbol{a}_0, \alpha_1) = \Delta^0 F(\boldsymbol{a}_0, \alpha_1) = 0$. Argumentando inductivamente, supongamos que $F(\boldsymbol{a}_0, \alpha_1) = \cdots = F(\boldsymbol{a}_0, \alpha_{j-1}) = 0$. De la definición concluimos que $\Delta^{j-1} F(\boldsymbol{a}_0, \alpha_1 \cdots \alpha_j)$ puede expresarse como una combinación lineal de las diferencias $F(\boldsymbol{a}_0, \alpha_{k+1}) - F(\boldsymbol{a}_0, \alpha_k)$ con $1 \leq k \leq j-1$. Por lo tanto, combinando la hipótesis inductiva con el hecho de que $\Delta^{j-1} F(\boldsymbol{a}_0, \alpha_1, \ldots, \alpha_j) = 0$, es fácil concluir que $F(\boldsymbol{a}_0, \alpha_j) = 0$. Esto finaliza la demostración de la afirmación. □

Con el objetivo de estudiar la geometría de Γ_i^*, mostramos primero que dicha variedad es una intersección completa conjuntista. Luego analizamos el lugar singular de Γ_i^*, mostrando que tiene codimensión al menos 2 en Γ_i^*.

Con este propósito recordamos que las formas lineales $\boldsymbol{L} := (L_1, \ldots, L_m)$ de (3.3), que definen la familia $\mathcal{A}_{\boldsymbol{L}}$, son linealmente independientes, y que $(\partial \boldsymbol{L}/\partial \boldsymbol{A}) := (b_{k,d-j})_{1 \leq k \leq m, \ 1 \leq j \leq d-r}$ es una matriz escalonada por filas, siendo $1 \leq j_1 < \cdots < j_m \leq d-r$ las posiciones de las columnas de $(\partial \boldsymbol{L}/\partial \boldsymbol{A})$ correspondientes a los pivotes.

Lema 3.2.2. *Γ_i^* es una intersección completa conjuntista de dimensión $d-m$.*

Demostración. Consideramos el orden lexicográfico graduado de $\overline{\mathbb{F}}_q[\boldsymbol{A}_0, \boldsymbol{T}]$ con $T_i > \cdots > T_1 > A_{d-1} > A_{d-2} > \cdots > A_0$. Es fácil ver que para cada j el polinomio $\Delta^{j-1}F(\boldsymbol{A}_0, T_1, \ldots, T_j)$ tiene grado $d-j+1$ en las variables $\boldsymbol{T}$ y el monomio T_j^{d-j+1} aparece en su representación densa con un coeficiente no nulo. Deducimos que el término principal de $\Delta^{j-1}F(\boldsymbol{A}_0, T_1, \ldots, T_j)$ en el orden monomial definido arriba es T_j^{d-j+1} para $1 \leq j \leq i$. Por otro lado, el término principal de $L_k(\boldsymbol{A}_0)$ en este orden monomial es A_{d-j_k} para $1 \leq k \leq m$. Así, los términos principales de $\Delta^{j-1}F(\boldsymbol{A}_0, T_1, \ldots, T_j)$ $(1 \leq j \leq i)$ y L_k $(1 \leq k \leq m)$ son coprimos, y por lo tanto forman una base de Gröbner del ideal J que generan (ver, por ejemplo, [CLO92, Section 2.9, Proposition 4]). El ideal inicial de J está generado por T_j^{d-j+1} $(1 \leq j \leq i), A_{d-j_k}$ $(1 \leq k \leq m)$, los cuales forman una sucesión regular en $\overline{\mathbb{F}}_q[\boldsymbol{A}_0, \boldsymbol{T}]$. Por lo tanto, por [Eis95, Proposition 15.15] tenemos que los polinomios que definen la variedad Γ_i^* también forman una sucesión regular en $\overline{\mathbb{F}}_q[\boldsymbol{A}_0, \boldsymbol{T}]$. Concluimos que Γ_i^* es una intersección completa conjuntista de dimensión $d-m$. □

Ahora demostramos que el lugar singular de Γ_i^* tiene codimensión al menos 2 en Γ_i^*. Comenzamos con el siguiente criterio de no singularidad.

Lema 3.2.3. *Sea $J_{F,L} \in \overline{\mathbb{F}}_q[\boldsymbol{A}_0, \boldsymbol{T}]^{(m+i)\times(d+i)}$ la matriz Jacobiana de los polinomios $F(\boldsymbol{A}_0, T_j)$ $(1 \leq j \leq i)$ y $L_k(\boldsymbol{A}_0)$ $(1 \leq k \leq m)$ con respecto a $\boldsymbol{A}_0$, $\boldsymbol{T}$, y sea $(\boldsymbol{a}_0, \boldsymbol{\alpha}) \in \Gamma_i^*$. Si $J_{F,\boldsymbol{L}}(\boldsymbol{a}_0, \boldsymbol{\alpha})$ es de rango completo, entonces $(\boldsymbol{a}_0, \boldsymbol{\alpha})$ es un punto no singular de Γ_i^*.*

Demostración. Considerando la forma de Newton del polinomio que interpola a $F(\boldsymbol{a}_0, T)$ en $\alpha_1, \ldots, \alpha_i$ deducimos fácilmente que $F(\boldsymbol{a}_0, \alpha_j) = 0$ para $1 \leq j \leq i$. Esto implica que $F(\boldsymbol{A}_0, T_j)$ se anula en Γ_i^* para $1 \leq j \leq i$. Por lo tanto, todo elemento del espacio tangente $\mathcal{T}_{(\boldsymbol{a}_0,\boldsymbol{\alpha})}\Gamma_i^*$ de Γ_i^* en $(\boldsymbol{a}_0, \boldsymbol{\alpha})$ pertenece al núcleo de la matriz Jacobiana $J_{F,\boldsymbol{L}}(\boldsymbol{a}_0, \boldsymbol{\alpha})$.

Por hipótesis, la matriz $J_{F,\boldsymbol{L}}(\boldsymbol{a}_0, \boldsymbol{\alpha})$ es de tamaño $(m+i) \times (d+i)$ y tiene rango $m+i$, y así su núcleo tiene dimensión $d-m$. Por lo tanto, el espacio tangente $\mathcal{T}_{(\boldsymbol{a}_0,\boldsymbol{\alpha})}\Gamma_i^*$ tiene dimensión a lo sumo $d-m$. Como Γ_i^* es de dimensión pura $d-m$, deducimos que $(\boldsymbol{a}_0, \boldsymbol{\alpha})$ es un punto no singular de Γ_i^*. □

Sea $(\boldsymbol{a}_0, \boldsymbol{\alpha})$ un punto arbitrario de Γ_i^*, con $\boldsymbol{\alpha} := (\alpha_1, \ldots, \alpha_i)$, y sea $f_{\boldsymbol{a}_0} := F(\boldsymbol{a}_0, T)$. Entonces la matriz Jacobiana $J_{F,\boldsymbol{L}}$ evaluada en $(\boldsymbol{a}_0, \boldsymbol{\alpha})$ tiene la siguiente forma:

$$J_{F,\boldsymbol{L}}(\boldsymbol{a}_0, \boldsymbol{\alpha}) := \begin{pmatrix} \dfrac{\partial \boldsymbol{L}}{\partial \boldsymbol{A_0}}(\boldsymbol{a}_0, \boldsymbol{\alpha}) & \boldsymbol{0} \\ \dfrac{\partial F}{\partial \boldsymbol{A_0}}(\boldsymbol{a}_0, \boldsymbol{\alpha}) & \dfrac{\partial F}{\partial \boldsymbol{T}}(\boldsymbol{a}_0, \boldsymbol{\alpha}) \end{pmatrix}.$$

Observemos que $(\partial F/\partial \boldsymbol{T})(\boldsymbol{a}_0, \boldsymbol{\alpha})$ es una matriz diagonal cuya j–ésima entrada diagonal es $f'_{\boldsymbol{a}_0}(\alpha_j)$. Como la matriz $(\partial \boldsymbol{L}/\partial \boldsymbol{A_0})(\boldsymbol{a}_0, \boldsymbol{\alpha})$ es de rango completo, si todas las raíces en $\overline{\mathbb{F}}_q$ del polinomio $f_{\boldsymbol{a}_0}$ son simples, la matriz $J_{F,\boldsymbol{L}}(\boldsymbol{a}_0, \boldsymbol{\alpha})$ es también de rango completo, y por lo tanto $(\boldsymbol{a}_0, \boldsymbol{\alpha})$ es un punto regular en Γ_i^*. Así, para probar que el lugar singular de Γ_i^* es una subvariedad de codimensión al menos 2 en Γ_i^*, es

suficiente considerar el conjunto de puntos $(\boldsymbol{a}_0, \boldsymbol{\alpha}) \in \Gamma_i^*$ tales que al menos una de las coordenadas de $\boldsymbol{\alpha}$ es una raíz múltiple de $f_{\boldsymbol{a}_0}$. En particular, observamos que el polinomio $f_{\boldsymbol{a}_0}$ debe tener raíces múltiples.

Comenzamos considerando el caso "extremo" donde el polinomio derivado $f'_{\boldsymbol{a}_0}$ es nulo; para ello, sea el morfismo de $\mathbb{F}_q$–variedades definido de la siguiente manera:

$$\begin{array}{rccc} \Psi_i : & \Gamma_i^* & \to & \mathcal{L} \\ & (\boldsymbol{a}_0, \boldsymbol{\alpha}) & \mapsto & \boldsymbol{a}_0, \end{array} \tag{3.9}$$

donde $\mathcal{L} := \{L_1 = 0, \ldots, L_m = 0\} \subset \mathbb{A}^d$ es el conjunto de ceros comunes de las formas lineales afines $L_1, \ldots, L_m$ de (3.3).

Recordemos que $1 \leq j_1 < \cdots < j_m \leq d - r$ representan las posiciones de las columnas correspondientes a los pivotes de la matriz escalonada por filas $(\partial \boldsymbol{L}/\partial \boldsymbol{A}) := (b_{k,d-j})_{1 \leq k \leq m,\ 1 \leq j \leq d-r}$ y $\mathcal{J} := \{d-1, \ldots, 0\} \setminus \{d - j_1, \ldots, d - j_m\}$. Observemos que el anillo de coordenadas $\overline{\mathbb{F}}_q[\mathcal{L}]$ de $\mathcal{L}$ es isomorfo al anillo de polinomios $\overline{\mathbb{F}}_q[A_k : k \in \mathcal{J}]$. Tenemos el siguiente resultado.

Lema 3.2.4. *Ψ_i es un morfismo finito.*

Demostración. Dado que es fácil ver que Ψ_i es un morfismo sobreyectivo, alcanza con mostrar que la función de coordenada t_j of $\overline{\mathbb{F}}_q[\Gamma_i^*]$ definida por T_j satisface una ecuación mónica con coeficientes en $\overline{\mathbb{F}}_q[A_k : k \in \mathcal{J}]$ for $1 \leq j \leq i$. Con este propósito, observamos que el polinomio $F(\boldsymbol{A}_0, T_j)$ se anula en Γ_i^* para $1 \leq j \leq i$ y es un elemento mónico de $\overline{\mathbb{F}}_q[\boldsymbol{A}_0][T_j]$. Teniendo en cuenta el isomorfismo entre el anillo de coordenadas $\overline{\mathbb{F}}_q[\mathcal{L}]$ y $\overline{\mathbb{F}}_q[A_k : k \in \mathcal{J}]$, es fácil concluir que existe un polinomio $G \in \overline{\mathbb{F}}_q[A_k : k \in \mathcal{J}][T_j]$ tal que $F(\boldsymbol{A}_0, T_j) = G(A_k : k \in \mathcal{J}; T_j)$ en Γ_i^* para $1 \leq j \leq i$. Deducimos así la existencia de una ecuación mónica que anula a t_j con coeficientes en $\overline{\mathbb{F}}_q[A_k : k \in \mathcal{J}]$ para $1 \leq j \leq i$. Esto concluye la demostración del lema. □

Una primera consecuencia de este lema es el siguiente resultado sobre el conjunto de puntos $(\boldsymbol{a}_0, \boldsymbol{\alpha}) \in \Gamma_i^*$ tales que $f'_{\boldsymbol{a}_0} = 0$.

Lema 3.2.5. *Si $3 \leq r \leq d - m$, entonces el conjunto $\mathcal{W}_1$ de puntos $(\boldsymbol{a}_0, \boldsymbol{\alpha}) \in \Gamma_i^*$ tal que $f'_{\boldsymbol{a}_0} = 0$ está contenido en una subvariedad de codimensión 2 en Γ_i^*.*

Demostración. Dado que $p > 2$, la condición $f'_{\boldsymbol{a}_0} = 0$ implica $a_1 = a_2 = 0$. Así tenemos que el conjunto de puntos $(\boldsymbol{a}_0, \boldsymbol{\alpha}) \in \Gamma_i^*$ con $f'_{\boldsymbol{a}_0} = 0$ es un subconjunto de $\Psi_i^{-1}(\mathcal{Z}_{1,2})$, donde $\mathcal{Z}_{1,2} \subset \mathcal{L}$ es la variedad de dimensión $d - m - 2$ definida por las ecuaciones $A_1 = A_2 = 0$. Teniendo en cuenta que Ψ_i es un morfismo finito, deducimos que $\Psi_i^{-1}(\mathcal{Z}_{1,2})$ tiene dimensión $d - m - 2$ (ver Teorema 2.1.4). □

En lo que sigue vamos a suponer que $f'_{\boldsymbol{a}_0}$ es no nulo y que $f_{\boldsymbol{a}_0}$ tiene raíces múltiples. Analizamos ahora el caso donde exactamente una de las coordenadas de $\boldsymbol{\alpha}$ es raíz múltiple de $f_{\boldsymbol{a}_0}$.

Lema 3.2.6. *Supongamos que existe una única coordenada α_j de $\boldsymbol{\alpha}$ que resulta una raíz múltiple de $f_{\boldsymbol{a}_0}$. Entonces $(\boldsymbol{a}_0, \boldsymbol{\alpha})$ es un punto regular de Γ_i^*.*

Demostración. Supongamos sin pérdida de generalidad que α_1 es la única raíz múltiple de $f_{\boldsymbol{a}_0}$ entre las coordenadas de $\boldsymbol{\alpha}$. De acuerdo al Lema 3.2.3, basta mostrar que la matriz Jacobiana $J_{F,\boldsymbol{L}}(\boldsymbol{a}_0, \boldsymbol{\alpha})$ es de rango completo. Con este objetivo, consideramos la submatriz $(\partial F/\partial(A_0, \boldsymbol{T}))(\boldsymbol{a}_0, \boldsymbol{\alpha})$ de tamaño $i \times (i+1)$ que consiste de las últimas i filas e $i+1$ columnas de $J_{F,\boldsymbol{L}}(\boldsymbol{a}_0, \boldsymbol{\alpha})$, es decir,

$$\frac{\partial F}{\partial(A_0, \boldsymbol{T})}(\boldsymbol{a}_0, \boldsymbol{\alpha}) := \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & f'_{\boldsymbol{a}_0}(\alpha_2) & 0 & \cdots & 0 \\ \vdots & \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\ 1 & 0 & 0 & \cdots & 0 & f'_{\boldsymbol{a}_0}(\alpha_i) \end{pmatrix}.$$

Dado que por hipótesis α_j es una raíz simple de $f'_{\boldsymbol{a}_0}$ para $j \geq 2$, tenemos que $f'_{\boldsymbol{a}_0}(\alpha_j) \neq 0$ para $j \geq 2$, y así deducimos que $(\partial F/\partial(A_0, \boldsymbol{T}))(\boldsymbol{a}_0, \boldsymbol{\alpha})$ tiene rango i.

Por otro lado, la matriz $(\partial \boldsymbol{L}/\partial \boldsymbol{A}_0)(\boldsymbol{a}_0, \boldsymbol{\alpha})$ tiene rango m y sus últimas columnas son nulas. Por lo tanto, denotando por $(\partial \boldsymbol{L}/\partial \boldsymbol{A})(\boldsymbol{a}_0, \boldsymbol{\alpha})$ a la submatriz de $(\partial \boldsymbol{L}/\partial \boldsymbol{A}_0)(\boldsymbol{a}_0, \boldsymbol{\alpha})$ que se obtiene eliminando las últimas columnas, podemos reescribir a $J_{F,\boldsymbol{L}}(\boldsymbol{a}_0, \boldsymbol{\alpha})$ como una matriz de bloques:

$$J_{F,\boldsymbol{L}}(\boldsymbol{a}_0, \boldsymbol{\alpha}) = \begin{pmatrix} \dfrac{\partial \boldsymbol{L}}{\partial \boldsymbol{A}}(\boldsymbol{a}_0, \boldsymbol{\alpha}) & \boldsymbol{0} \\ * & \dfrac{\partial F}{\partial(A_0, \boldsymbol{T})}(\boldsymbol{a}_0, \boldsymbol{\alpha}) \end{pmatrix}.$$

Como las matrices $(\partial \boldsymbol{L}/\partial \boldsymbol{A})(\boldsymbol{a}_0, \boldsymbol{\alpha})$ y $(\partial F/\partial(A_0, \boldsymbol{T}))(\boldsymbol{a}_0, \boldsymbol{\alpha})$ son de rango completo, concluimos que $J_{F,\boldsymbol{L}}(\boldsymbol{a}_0, \boldsymbol{\alpha})$ tiene rango $m+i$. □

El siguiente caso a considerar es cuando aparecen dos raíces múltiples distintas de $f_{\boldsymbol{a}_0}$ entre las coordenadas de $\boldsymbol{\alpha}$.

Lema 3.2.7. *Sea $\mathcal{W}_2$ el conjunto de puntos de $(\boldsymbol{a}_0, \boldsymbol{\alpha}) \in \Gamma_i^*$ tales que existen $1 \leq j < k \leq i$ para los cuales $\alpha_j \neq \alpha_k$ y α_j, α_k son raíces múltiples de $f_{\boldsymbol{a}_0}$. Entonces $\mathcal{W}_2$ está contenido en una subvariedad de codimensión 2 de Γ_i^*.*

Demostración. Sea $(\boldsymbol{a}_0, \boldsymbol{\alpha})$ un punto arbitrario de $\mathcal{W}_2$. Dado que, por hipótesis, $f_{\boldsymbol{a}_0}$ tiene al menos dos raíces múltiples distintas, el grado del máximo común divisor entre $f_{\boldsymbol{a}_0}$ y $f'_{\boldsymbol{a}_0}$ es al menos 2. Esto implica que

$$\operatorname{Res}(f_{\boldsymbol{a}_0}, f'_{\boldsymbol{a}_0}) = \operatorname{Subres}(f_{\boldsymbol{a}_0}, f'_{\boldsymbol{a}_0}) = 0,$$

donde $\operatorname{Res}(f_{\boldsymbol{a}_0}, f'_{\boldsymbol{a}_0})$ y $\operatorname{Subres}(f_{\boldsymbol{a}_0}, f'_{\boldsymbol{a}_0})$ denotan la resultante y la subresultante de primer orden de $f_{\boldsymbol{a}_0}$ y $f'_{\boldsymbol{a}_0}$ respectivamente.

Por otro lado, por el isomorfismo entre el anillo de coordenadas $\overline{\mathbb{F}}_q[\mathcal{L}]$ y el anillo de polinomios $\overline{\mathbb{F}}_q[A_k : k \in \mathcal{J}]$ deducimos que existen polinomios $h_l \in \overline{\mathbb{F}}_q[A_k : k \in \mathcal{J}]$ de grado 1 tales que se satisface $A_{d-j_l} = h_l$ en Γ_i^* para $1 \leq l \leq m$. Sea $\hat{\boldsymbol{A}}_0 \in \overline{\mathbb{F}}_q[A_k : k \in \mathcal{J}]^d$ el elemento que se obtiene al sustituir el polinomio h_l en la coordenada A_{d-j_l} de

$\boldsymbol{A}_0$, para $l = 1, \ldots, m$. Como el grado del polinomio $f_{\boldsymbol{a}_0}$ es d, por propiedades básicas de las resultantes y subresultantes (ver, por ejemplo, [CLO92, §3.6, Proposition 3]), tenemos que

$$\begin{aligned}\operatorname{Res}(f_{\boldsymbol{a}_0}, f'_{\boldsymbol{a}_0}) &= \operatorname{Res}(F(\hat{\boldsymbol{A}}_0, T), \Delta^1 F(\hat{\boldsymbol{A}}_0, T, T), T)|_{\hat{\boldsymbol{A}}_0 = \boldsymbol{a}_0},\\ \operatorname{Subres}(f_{\boldsymbol{a}_0}, f'_{\boldsymbol{a}_0}) &= \operatorname{Subres}(F(\hat{\boldsymbol{A}}_0, T), \Delta^1 F(\hat{\boldsymbol{A}}_0, T, T), T))|_{\hat{\boldsymbol{A}}_0 = \boldsymbol{a}_0},\end{aligned}$$

donde

$$\begin{aligned}\mathcal{R} &:= \operatorname{Res}(F(\hat{\boldsymbol{A}}_0, T), \Delta^1 F(\hat{\boldsymbol{A}}_0, T, T), T),\\ \mathcal{S}_1 &:= \operatorname{Subres}(F(\boldsymbol{A}_0, T), \Delta^1 F(\hat{\boldsymbol{A}}_0, T, T), T),\end{aligned} \tag{3.10}$$

son la resultante y subresultante de primer orden de $F(\hat{\boldsymbol{A}}_0, T)$ y $\Delta^1 F(\hat{\boldsymbol{A}}_0, T, T)$ con respecto a T. Por lo tanto, $\mathcal{W}_2 \subset \Psi_i^{-1}(\mathcal{Z}_2)$, donde Ψ_i es el morfismo de (3.9) y $\mathcal{Z}_2$ es la subvariedad de $\mathcal{L}$ definida por las ecuaciones

$$\mathcal{R}(\hat{\boldsymbol{A}}_0) = \mathcal{S}_1(\hat{\boldsymbol{A}}_0) = 0. \tag{3.11}$$

Observemos que $\mathcal{R}(\hat{\boldsymbol{A}}_0)$ es un polinomio no nulo ya que $F(\hat{\boldsymbol{A}}_0, T)$ es un elemento separable de $\mathbb{F}_q[A_k : k \in \mathcal{J}][T]$. Como $p > 2$, tenemos que $\mathcal{S}_1$ también es un polinomio no nulo. En efecto, si p no divide a $d(d-1)$, el monomio $d(d-1)^{d-2}A_1^{d-2}$ aparece en la representación densa de $\mathcal{S}_1$. En cambio, si p divide a $d(d-1)$, el término no nulo $2(-1)^d(d-2)^{d-2}A_2^{d-1}$ aparece en la representación densa de $\mathcal{S}_1$. Por otro lado, los polinomios $\mathcal{R}(\hat{\boldsymbol{A}}_0)$ y $\mathcal{S}_1(\hat{\boldsymbol{A}}_0)$ forman una sucesión regular de $\overline{\mathbb{F}}_q[A_k : k \in \mathcal{J}]$. En efecto, dado que $p > 2$, por el Teorema 3.1.7 tenemos que $\mathcal{R}(\hat{\boldsymbol{A}}_0)$ es un elemento irreducible de $\overline{\mathbb{F}}_q[A_k : k \in \mathcal{J}]$, y por lo tanto el anillo cociente $\overline{\mathbb{F}}_q[A_k : k \in \mathcal{J}]/(\mathcal{R}(\hat{\boldsymbol{A}}_0))$ es un dominio. Si $\mathcal{S}_1(\hat{\boldsymbol{A}}_0)$ fuera un divisor de cero en $\overline{\mathbb{F}}_q[A_k : k \in \mathcal{J}]/(\mathcal{R}(\hat{\boldsymbol{A}}_0))$, entonces debería ser un múltiplo de $\mathcal{R}(\hat{\boldsymbol{A}}_0)$ en $\overline{\mathbb{F}}_q[A_k : k \in \mathcal{J}]$. Esto último no puede ocurrir porque $\max\{\deg_{A_1} \mathcal{R}(\hat{\boldsymbol{A}}_0), \deg_{A_2} \mathcal{R}(\hat{\boldsymbol{A}}_0)\} = d$, mientras que $\max\{\deg_{A_1} \mathcal{S}_1(\hat{\boldsymbol{A}}_0), \deg_{A_2} \mathcal{S}_1(\hat{\boldsymbol{A}}_0)\} \leq d-1$. Así concluimos que $\dim \mathcal{Z}_2 = d-m-2$, y por lo tanto $\dim \Psi_r^{-1}(\mathcal{Z}_2) = d-m-2$. Por lo tanto, $\mathcal{W}_2$ está contenido en una subvariedad de Γ_i^* de codimensión 2 en Γ_i^*. □

Resta considerar el caso en donde aparece una única raíz múltiple de $f_{\boldsymbol{a}_0}$ entre las coordenadas de $\boldsymbol{\alpha}$, pero en al menos dos coordenadas distintas de $\boldsymbol{\alpha}$. En tal caso, tenemos que, o bien las restantes coordenadas de $\boldsymbol{\alpha}$ resultan ser raíces simples de $f_{\boldsymbol{a}_0}$, o bien existe al menos una tercera coordenada cuyo valor es la misma raíz múltiple. El siguiente resultado trata el primero de estos dos casos.

Lema 3.2.8. *Sea $(\boldsymbol{a}_0, \boldsymbol{\alpha}) \in \Gamma_i^*$ un punto que satisface las siguientes condiciones:*

- *existen $1 \leq j < k \leq i$ tales que $\alpha_j = \alpha_k$ y α_j es una raíz múltiple de $f_{\boldsymbol{a}_0}$;*
- *para todo $l \notin \{j, k\}$, α_l es una raíz simple de $f_{\boldsymbol{a}_0}$.*

Entonces $(\boldsymbol{a}_0, \boldsymbol{\alpha})$ es un punto regular de Γ_i^*.

Demostración. Podemos suponer sin pérdida de generalidad que $j = 1$ and $k = 2$. Observemos que los polinomios $\Delta^1 F(\boldsymbol{A}_0, T_1, T_2)$ y $F(\boldsymbol{A}_0, T_j)$ $(2 \leq j \leq i)$ se anulan en Γ_i^*. Por lo tanto, el espacio tangente $\mathcal{T}_{(\boldsymbol{a}_0,\boldsymbol{\alpha})}\Gamma_i^*$ de Γ_i^* en $(\boldsymbol{a}_0, \boldsymbol{\alpha})$ está incluido en el núcleo de la matriz Jacobiana $J_{\Delta,F,\boldsymbol{L}}(\boldsymbol{a}_0, \boldsymbol{\alpha})$ de $\Delta^1 F(\boldsymbol{A}_0, T_1, T_2)$, $F(\boldsymbol{A}_0, T_j)$ $(2 \leq j \leq i)$ y $\boldsymbol{L}$ con respecto a $\boldsymbol{A}_0, \boldsymbol{T}$. Afirmamos que $J_{\Delta,F,\boldsymbol{L}}(\boldsymbol{a}_0, \boldsymbol{\alpha})$ tiene rango $i + m$.

Ahora probamos esta afirmación. Es fácil ver que $\frac{\partial \Delta^1 F}{\partial A_0}(\boldsymbol{a}_0, \alpha_1, \alpha_1) = 0$ y que $\frac{\partial \Delta^1 F}{\partial A_j}(\boldsymbol{a}_0, \alpha_1, \alpha_1) = j\alpha_1^{j-1}$ para $j \geq 1$. Por lo tanto, podemos expresar a $J_{\Delta,F,\boldsymbol{L}}(\boldsymbol{a}_0, \boldsymbol{\alpha})$ como la siguiente matriz por bloques:

$$J_{\Delta,F,\boldsymbol{L}}(\boldsymbol{a}_0, \boldsymbol{\alpha}) = \begin{pmatrix} \dfrac{\partial \boldsymbol{L}}{\partial \boldsymbol{A}_2}(\boldsymbol{a}_0, \boldsymbol{\alpha}) & \boldsymbol{0} \\ * & \dfrac{\partial F}{\partial (A_1, A_0, \boldsymbol{T})}(\boldsymbol{a}_0, \boldsymbol{\alpha}) \end{pmatrix},$$

donde $(\partial \boldsymbol{L}/\partial \boldsymbol{A}_2)(\boldsymbol{a}_0, \boldsymbol{\alpha}) \in \mathbb{F}_q^{m\times(d-2)}$ es la matriz Jacobiana de $\boldsymbol{L}$ con respecto a $A_{d-1}, \ldots, A_2$ y $(\partial F/\partial(A_1, A_0, \boldsymbol{T}))(\boldsymbol{a}_0, \boldsymbol{\alpha}) \in \mathbb{F}_q^{i\times(i+2)}$ está definida como

$$\frac{\partial F}{\partial (A_1, A_0, \boldsymbol{T})}(\boldsymbol{a}_0, \boldsymbol{\alpha}) := \begin{pmatrix} 1 & 0 & * & * & 0 & \cdots & 0 \\ \alpha_2 & 1 & 0 & 0 & 0 & \cdots & 0 \\ \alpha_3 & 1 & 0 & 0 & f'_{\boldsymbol{a}_0}(\alpha_3) & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ \alpha_i & 1 & 0 & 0 & 0 & \cdots & f'_{\boldsymbol{a}_0}(\alpha_i) \end{pmatrix}.$$

Como α_j es una raíz simple de $f_{\boldsymbol{a}_0}$ para $j \geq 3$, se sigue que $f'_{\boldsymbol{a}_0}(\alpha_j) \neq 0$ para $j \geq 3$. Esto implica que la submatriz de $(\partial F/\partial(A_1, A_0, \boldsymbol{T}))(\boldsymbol{a}_0, \boldsymbol{\alpha})$ de tamaño $i \times i$ que consiste de eliminar la tercera y cuarta columna de $(\partial F/\partial(A_1, A_0, \boldsymbol{T}))(\boldsymbol{a}_0, \boldsymbol{\alpha})$ es de rango i. Así, concluimos que $J_{\Delta,F,\boldsymbol{L}}(\boldsymbol{a}_0, \boldsymbol{\alpha})$ tiene rango $m + i$, finalizando la demostración de la afirmación.

Ahora, como $J_{\Delta,F,\boldsymbol{L}}(\boldsymbol{a}_0, \boldsymbol{\alpha})$ tiene rango $m+i$, el núcleo de dicha matriz Jacobiana tiene dimensión $d - m$. Esto implica que $\dim \mathcal{T}_{(\boldsymbol{a}_0,\boldsymbol{\alpha})}\Gamma_i^* \leq d - m$, lo cual prueba que $(\boldsymbol{a}_0, \boldsymbol{\alpha})$ es un punto regular de Γ_i^*.

□

Finalmente analizamos el conjunto de puntos de Γ_i^* tales que al menos tres coordenadas distintas de $\boldsymbol{\alpha}$ resultan ser la misma raíz múltiple de $f_{\boldsymbol{a}_0}$.

Lema 3.2.9. *Sea $\mathcal{W}_3 \subset \Gamma_i^*$ el conjunto de puntos $(\boldsymbol{a}_0, \boldsymbol{\alpha})$ tales que existen $1 \leq j < k < l \leq i$ con $\alpha_j = \alpha_k = \alpha_l$, siendo α_j una raíz múltiple de $f_{\boldsymbol{a}_0}$. Entonces $\mathcal{W}_3$ está contenido en una subvariedad de codimensión 2 en Γ_i^*.*

Demostración. Sea $(\boldsymbol{a}_0, \boldsymbol{\alpha})$ un punto arbitrario de $\mathcal{W}_3$. Sin pérdida de generalidad podemos suponer que $\alpha_1 = \alpha_2 = \alpha_3$ es la raíz múltiple de $f_{\boldsymbol{a}_0}$ de la hipótesis del lema. Teniendo en cuenta que $(\boldsymbol{a}_0, \boldsymbol{\alpha})$ satisface las ecuaciones

$$F(\boldsymbol{A}_0, T_1) = \Delta F(\boldsymbol{A}_0, T_1, T_2) = \Delta^2 F(\boldsymbol{A}_0, T_1, T_2, T_3) = 0,$$

vemos que α_1 es una raíz común de $f_{\boldsymbol{a}_0}$, $\Delta F(\boldsymbol{a}_0, T, T)$ y $\Delta^2 F(\boldsymbol{a}_0, T, T, T)$.

Dado que $\deg_T F(\boldsymbol{A}_0, T) = \deg_T F(\boldsymbol{a}_0, T)$, por [CLO92, §3.6, Proposition 3] tenemos que

$$\operatorname{Res}(f_{\boldsymbol{a}_0}, f'_{\boldsymbol{a}_0}) = \mathcal{R}(\hat{\boldsymbol{A}}_0)|_{\hat{\boldsymbol{A}}_0 = \boldsymbol{a}_0}, \tag{3.12}$$

donde $\mathcal{R}(\hat{\boldsymbol{A}}_0)$ es la resultante de (3.10) y $\hat{\boldsymbol{A}}_0 \in \overline{\mathbb{F}}_q[A_k : k \in \mathcal{J}]^d$ es el elemento que aparece en la demostración del Lema 3.2.7.

Supongamos que $\Delta^2 F(\boldsymbol{a}_0, T, T, T) = 0$ y sea $\mathcal{W}'_3$ el conjunto de puntos $(\boldsymbol{a}_0, \boldsymbol{\alpha}) \in \Gamma_i^*$ tales que $\Delta^2 F(\boldsymbol{a}_0, T, T, T) = 0$. Entonces

$$0 = 2\Delta^2 F(\boldsymbol{a}_0, T, T, T) = d(d-1)T^{d-2} + (d-1)(d-2)a_{d-1}T^{d-3} + \cdots + 2a_2.$$

Esto implica que $2a_2 = 0$ y, como $p > 2$, tenemos que $a_2 = 0$. Como consecuencia de esta identidad y de (3.12), el conjunto $\mathcal{W}'_3$ está contenido en $\Psi_i^{-1}(\mathcal{Z}'_3)$, donde $\mathcal{Z}'_3 \subset \mathcal{L}$ es la variedad definida por las ecuaciones

$$A_2 = 0, \quad \mathcal{R}(\hat{\boldsymbol{A}}_0) = 0.$$

El Teorema 3.1.7 prueba que $\mathcal{R}(\hat{\boldsymbol{A}}_0)$ es un polinomio irreducible de $\overline{\mathbb{F}}_q[A_k : k \in \mathcal{J}]$ de grado $d-1$ en A_0. Así, $\mathcal{R}(\hat{\boldsymbol{A}}_0)$ y A_2 forman una sucesión regular en $\overline{\mathbb{F}}_q[A_k : k \in \mathcal{J}]$. Dado que Ψ_i es un morfismo finito, tenemos que $\Psi_i^{-1}(\mathcal{Z}'_3)$ tiene dimensión $d-m-2$. Por lo tanto, podemos suponer que $\Delta^2 F(\boldsymbol{a}_0, T, T, T)$ es no nulo.

Ahora supongamos que p no divide a d. Entonces $f_{\boldsymbol{a}_0}$ y $f'_{\boldsymbol{a}_0}$ son polinomios no nulos de grado d y $d-1$ respectivamente. Así, por [CLO92, §3.6, Proposition 3], tenemos que

$$\begin{aligned}
\operatorname{Res}(f_{\boldsymbol{a}_0}, f'_{\boldsymbol{a}_0}) &= \mathcal{R}(\hat{\boldsymbol{A}}_0)|_{\hat{\boldsymbol{A}}_0 = \boldsymbol{a}_0}, \\
\operatorname{Res}(f'_{\boldsymbol{a}_0}, \Delta^2 f_{\boldsymbol{a}_0}) &= \operatorname{Res}\big(\Delta^1 F(\hat{\boldsymbol{A}}_0, T, T), \Delta^2 F(\hat{\boldsymbol{A}}_0, T, T, T), T\big)\big|_{\hat{\boldsymbol{A}}_0 = \boldsymbol{a}_0}.
\end{aligned}$$

Por lo tanto, deducimos que $(\mathcal{W}_3 \setminus \mathcal{W}'_3) \cap \Gamma_i^* \subset \Psi_i^{-1}(\mathcal{Z}_3)$, donde Ψ_i es el morfismo definido en (3.9) y $\mathcal{Z}_3$ es la subvariedad $\mathcal{L}$ definida por las ecuaciones

$$\mathcal{R}(\hat{\boldsymbol{A}}_0) = 0, \quad \mathcal{R}' := \operatorname{Res}\big(\Delta^1 F(\hat{\boldsymbol{A}}_0, T, T), \Delta^2 F(\hat{\boldsymbol{A}}_0, T, T, T), T\big) = 0.$$

Como $\mathcal{R}(\hat{\boldsymbol{A}}_0)$ es un elemento irreducible de $\overline{\mathbb{F}}_q[A_k : k \in \mathcal{J}]$ de grado $d-1$ en A_0 y el polinomio no nulo $\mathcal{R}'$ tiene grado 0 en A_0, concluimos que $\mathcal{R}$ y $\mathcal{R}'$ forman una sucesión regular en $\overline{\mathbb{F}}_q[A_k : k \in \mathcal{J}]$. Esto muestra que $\mathcal{Z}_3$ tiene codimensión 2 en $\mathcal{L}$ y por lo tanto $\Psi_i^{-1}(\mathcal{Z}_3)$ es una subvariedad de codimensión 2 en Γ_i^*.

Por último, si p divide a d, a fin de demostrar en este caso que $(\mathcal{W}_3 \setminus \mathcal{W}'_3)$ está contenida en una subvariedad de codimensión 2 en Γ_i^*, vamos a considerar la noción de grado genérico $\operatorname{gendeg}(\partial F/\partial T)(\boldsymbol{A}_0, T)$ del polinomio $(\partial F/\partial T)(\boldsymbol{A}_0, T)$, que definimos como

$$\operatorname{gendeg}\left(\frac{\partial F}{\partial T}(\boldsymbol{A}_0, T)\right) := \operatorname{máx}\left\{ \deg \frac{\partial F}{\partial T}(\boldsymbol{a}_0, T) : \boldsymbol{a}_0 \in \mathcal{L} \right\}.$$

Sea $l \in \{1, \ldots, d-2\}$ el grado genérico de $(\partial F/\partial T)(\boldsymbol{A}_0, T)$. Es fácil ver que se cumplen las siguientes afirmaciones:

- p no divide a $l+1$,
- $\mathcal{L} \subset \{A_j = 0 : l+2 \leq j \leq d-1,\ p \text{ no divide } j\}$,
- $\mathcal{L} \cap \{A_{l+1} = 0\} \subsetneq \mathcal{L}$.

Si $l = d-2$, entonces $f'_{\boldsymbol{a}_0}$ y $\Delta^1 F(\hat{\boldsymbol{A}}_0, T, T)$ tienen el mismo grado en T y el argumento se sigue como arriba. En cambio, si $l \in \{1, \ldots, d-3\}$, consideramos la variedad $\mathcal{Z}_3 \subset \mathcal{L}$ de dimensión $d-m-2$ definida por las ecuaciones $\mathcal{R}(\hat{\boldsymbol{A}}_0) = A_{l+1} = 0$, concluimos que $\mathcal{W}_3 \setminus \mathcal{W}'_3 \subset \Psi_i^{-1}(\mathcal{Z}_3)$, y en consecuencia, que $\mathcal{W}_3 \setminus \mathcal{W}'_3$ tiene codimensión al menos 2 en Γ_i^*. □

Ahora podemos probar el principal resultado de esta sección. De los Lemas 3.2.5, 3.2.6, 3.2.7, 3.2.8 y 3.2.9 se concluye que el conjunto de puntos singulares de Γ_i^* está contenido en el conjunto $\mathcal{W}_1 \cup \mathcal{W}_2 \cup \mathcal{W}_3$, donde $\mathcal{W}_1$, $\mathcal{W}_2$ y $\mathcal{W}_3$ están definidos en las afirmaciones de los Lemas 3.2.5, 3.2.7 y 3.2.9. Dado que cada $\mathcal{W}_i$ está contenido en una subvariedad de codimensión 2 de Γ_i^*, obtenemos el siguiente resultado.

Teorema 3.2.10. *Sea $p > 2$ y $q > d$. Si $3 \leq r \leq d-m$, entonces el lugar singular de Γ_i^* tiene codimensión al menos 2 en Γ_i^*.*

Finalizamos esta sección con una consecuencia importante del Teorema 3.2.10.

Corolario 3.2.11. *Con las mismas hipótesis que en el Teorema 3.2.10, el ideal $J \subset \mathbb{F}_q[\boldsymbol{A}_0, \boldsymbol{T}]$ generado por $\Delta^{j-1}F(\boldsymbol{A}_0, T_1, \ldots, T_j)$ $(1 \leq j \leq i)$ y $L_k(\boldsymbol{A}_0)$ $(1 \leq k \leq m)$ es radical. Más aún, la variedad de incidencia Γ_i^* es una intersección completa de dimensión $d-m$.*

Demostración. Empezamos probando que J es un ideal radical. Denotamos con $J_{\Delta,\boldsymbol{L}}(\boldsymbol{A}_0, \boldsymbol{T})$ la matriz Jacobiana de $\Delta^{j-1}F(\boldsymbol{A}_0, T_1, \ldots, T_j)$ $(1 \leq j \leq i)$ y $L_k(\boldsymbol{A}_0)$ $(1 \leq k \leq m)$ con respecto a $\boldsymbol{A}_0, \boldsymbol{T}$. Por el Lema 3.2.2, dichos polinomios forman una sucesión regular. Así, por [Eis95, Theorem 18.15], es suficiente probar que el conjunto de puntos $(\boldsymbol{a}_0, \boldsymbol{\alpha}) \in \Gamma_i^*$ tales que $J_{\Delta,\boldsymbol{L}}(\boldsymbol{a}_0, \boldsymbol{\alpha})$ no es de rango completo está contenido en una subvariedad de Γ_i^* de codimensión al menos 1.

Observemos primero que en la demostración del Lema 3.2.3 mostramos que $F(\boldsymbol{A}_0, T_j) \in J$ para $1 \leq j \leq i$. Esto implica que cada gradiente $\nabla F(\boldsymbol{a}_0, \alpha_j)$ es una combinación lineal de los gradientes de los polinomios $\Delta^{j-1}F(\boldsymbol{a}_0, \boldsymbol{\alpha})$ $(1 \leq j \leq i)$ y $L_k(\boldsymbol{A}_0)$ $(1 \leq k \leq m)$. Concluimos así que rango $J_{F,\boldsymbol{L}}(\boldsymbol{a}_0, \boldsymbol{\alpha}) \leq$ rango $J_{\Delta,\boldsymbol{L}}(\boldsymbol{a}_0, \boldsymbol{\alpha})$.

Sea $(\boldsymbol{a}_0, \boldsymbol{\alpha})$ un punto arbitrario de Γ_i^* tal que $J_{\Delta,\boldsymbol{L}}(\boldsymbol{a}_0, \boldsymbol{\alpha})$ no es de rango completo. Entonces $J_{F,\boldsymbol{L}}(\boldsymbol{a}_0, \boldsymbol{\alpha})$ no es de rango completo y así $f_{\boldsymbol{a}_0}$ tiene raíces múltiples.

Afirmación. *El conjunto de puntos $(\boldsymbol{a}_0, \boldsymbol{\alpha}) \in \Gamma_i^*$ tales que $f_{\boldsymbol{a}_0}$ tiene raíces múltiples está contenido en una subvariedad de codimensión 1 en Γ_i^*.*

Demostración de la afirmación. Por el Lema 3.2.5, el conjunto de puntos $(\boldsymbol{a}_0, \boldsymbol{\alpha}) \in \Gamma_i^*$ tales que $f'_{\boldsymbol{a}_0} = 0$ está contenido en una subvariedad de codimensión 2 de Γ_i^*. Por otro lado, si $f_{\boldsymbol{a}_0}$ tiene raíces múltiples y $f'_{\boldsymbol{a}_0} \neq 0$ para algún $(\boldsymbol{a}_0, \boldsymbol{\alpha}) \in \Gamma_i^*$, entonces $(\boldsymbol{a}_0, \boldsymbol{\alpha}) \in \Psi_i^{-1}(\mathcal{Z})$, donde $\mathcal{Z}$ es la subvariedad de $\mathcal{L}$ definida por la ecuación $\mathcal{R}(\hat{\boldsymbol{A}}_0) := \mathrm{Res}(F(\hat{\boldsymbol{A}}_0, T), \Delta^1 F(\hat{\boldsymbol{A}}_0, T, T), T) = 0$. Por lo tanto, $\Psi_i^{-1}(\mathcal{Z})$ tiene codimensión 1 en Γ_i^*. □

Como consecuencia de esta afirmación se sigue que el conjunto de puntos $(\boldsymbol{a}_0, \boldsymbol{\alpha}) \in \Gamma_i^*$ tales que $J_{\Delta,\boldsymbol{L}}(\boldsymbol{a}_0, \boldsymbol{\alpha})$ no es de rango completo está contenido en una subvariedad de Γ_i^* de codimensión al menos 1. Así, J es un ideal radical, lo que implica que Γ_i^* es una intersección completa de dimensión $d-m$. □

3.2.2. La geometría de la clausura proyectiva

En esta sección demostramos que la clausura proyectiva de la variedad de incidencia Γ_i^* satisface ciertas propiedades geométricas necesarias para estimar la cantidad de puntos $\mathbb{F}_q$–racionales de Γ_i^*.

Consideramos la clausura proyectiva $\mathrm{pcl}(\Gamma_i^*) \subset \mathbb{P}^{d+i}$ de Γ_i^*. Recordamos que $\mathrm{pcl}(\Gamma_i^*) \subset \mathbb{P}^{d+i}$ está definida por la homogeneización $F^h \in \mathbb{F}_q[\boldsymbol{A}_0, T_0, \boldsymbol{T}]$ de cada polinomio F en el ideal $J \subset \mathbb{F}_q[\boldsymbol{A}_0, \boldsymbol{T}]$ generado por $\Delta^{j-1}F(\boldsymbol{A}_0, T_1, \ldots, T_j)$ $(1 \leq j \leq i)$ y $L_k(\boldsymbol{A}_0)$ $(1 \leq k \leq m)$. Denotamos con J^h al ideal generado por todos los polinomios F^h con $F \in J$. El Corolario 3.2.11 muestra que J es un ideal radical, por lo que también J^h es radical (ver Teorema 2.1.8). Además, $\mathrm{pcl}(\Gamma_i^*)$ es de dimensión pura $d-m$ y de grado igual a $\deg \Gamma_i^*$ (ver Teoremas 2.1.8 y 2.1.13).

Lema 3.2.12. *Los polinomios homogeneizados $\Delta^{j-1}F(\boldsymbol{A}_0, T_0, T_1, \ldots, T_j)^h$ $(1 \leq j \leq i)$ y $L_k^h(\boldsymbol{A}_0)$ $(1 \leq k \leq m)$ generan el ideal J^h. Además, $\mathrm{pcl}(\Gamma_i^*)$ es una intersección completa de dimensión $d-m$ y grado $d!/(d-i)!$.*

Demostración. En la demostración del Lema 3.2.2 se prueba que los polinomios $\Delta^{j-1}F(\boldsymbol{A}_0, T_1, \ldots, T_j)$ $(1 \leq j \leq i)$ y $L_k(\boldsymbol{A}_0)$ $(1 \leq k \leq m)$ forman una base de Gröbner del ideal J con el orden lexicográfico graduado definido por $T_i > \cdots > T_1 > A_{d-1} > \cdots > A_0$. Así, de, por ejemplo [CLO92, §8.4, Theorem 4], deducimos la primera afirmación del lema. En particular, tenemos que $\mathrm{pcl}(\Gamma_i^*)$ es una intersección completa de dimensión $d-m$. Finalmente, el Teorema 2.1.14 prueba que el grado de $\mathrm{pcl}(\Gamma_i^*)$ es $d!/(d-i)!$. □

Nuestro siguiente objetivo es estudiar el lugar singular de $\mathrm{pcl}(\Gamma_i^*)$. Empezamos con la siguiente caracterización de los puntos de $\mathrm{pcl}(\Gamma_i^*)$ en el hiperplano del infinito.

Lema 3.2.13. $\mathrm{pcl}(\Gamma_i^*) \cap \{T_0 = 0\} \subset \mathbb{P}^{d+i-1}$ *es una unión finita de a lo sumo $i+1$ variedades lineales de $\mathbb{P}^{d+i-1}$ de dimensión $d-m-1$.*

Demostración. Afirmamos que $\Delta^1 F(\boldsymbol{A}_0, T_0, T_j, T_k)^h \in J^h$ para $1 \leq j < k \leq i$. En efecto, tenemos la siguiente identidad $\Delta^1 F(\boldsymbol{A}_0, T_j, T_k)(T_j - T_k) = F(\boldsymbol{A}_0, T_j) - F(\boldsymbol{A}_0, T_k)$. Teniendo en cuenta que los polinomios $F(\boldsymbol{A}_0, T_l)$ se anulan en Γ_i^* para $1 \leq l \leq i$, se deduce que $\Delta^1 F(\boldsymbol{A}_0, T_j, T_k)$ se anula en el subconjunto abierto denso Zariski no vacío $\{T_j \neq T_k\} \cap \Gamma_i^*$ de Γ_i^*, y por lo tanto en Γ_i^*, lo que demuestra nuestra afirmación.

Combinando esta afirmación con el hecho de que $F(\boldsymbol{A}_0, T_0, T_j)^h \in J^h$ para $1 \leq j \leq i$, concluimos que cualquier punto $(\boldsymbol{a}_0, \boldsymbol{\alpha}) \in \mathrm{pcl}(\Gamma_i^*) \cap \{T_0 = 0\}$ satisface las

siguientes identidades:

$$F(\boldsymbol{A}_0, T_j)^h|_{T_0=0} = T_j^d + A_{d-1}T_j^{d-1} = T_j^{d-1}(T_j + A_{d-1}) = 0 \quad (1 \le j \le i), \tag{3.13}$$

$$\begin{aligned}\Delta^1 F(\boldsymbol{A}_0, T_0, T_j, T_k)^h|_{T_0=0} &= \frac{T_j^d - T_k^d}{T_j - T_k} + A_{d-1}\frac{T_j^{d-1} - T_k^{d-1}}{T_j - T_k} \\ &= \sum_{l=0}^{d-2} T_k^l T_j^{d-2-l}(T_j + A_{d-1}) + T_k^{d-1} = 0 \quad (1 \le j < k \le i).\end{aligned} \tag{3.14}$$

De (3.13) y (3.14) deducimos que $\mathrm{pcl}(\Gamma_i^*) \cap \{T_0 = 0\}$ está contenido en una unión de $\mathbb{F}_q$–variedades lineales de $\mathbb{P}^{d+i-1}$ de dimensión $d - m - 1$. Más precisamente, es fácil ver que

$$\mathrm{pcl}(\Gamma_i^*) \cap \{T_0 = 0\} \subset \bigcup_{j=0}^{i} \mathcal{L}_j,$$

donde $\mathcal{L}_0$ es la variedad definida por $T_k = 0$ $(1 \le k \le i)$ y $L_k = 0$ $(1 \le k \le m)$, y $\mathcal{L}_j$ es la variedad lineal definida por las siguientes ecuaciones para $1 \le j \le i$:

$$T_j + A_{d-1} = 0, \quad T_l = 0 \ \ (1 \le l \le i,\, l \ne j), \quad L_k = 0 \ \ (1 \le k \le m).$$

Por el Lema 3.2.12 tenemos que $\mathrm{pcl}(\Gamma_i^*)$ es de dimensión pura $d - m$. Así, por el Teorema 2.1.7 cada componente irreducible de $\mathrm{pcl}(\Gamma_i^*) \cap \{T_0 = 0\}$ tiene dimensión al menos $d - m - 1$ y está contenida en una variedad lineal $\mathcal{L}_j$ para algún $j \in \{0, \ldots, i\}$. Así, del Teorema 2.1.3 deducimos que cada componente irreducible de $\mathrm{pcl}(\Gamma_i^*) \cap \{T_0 = 0\}$ debe ser la variedad lineal $\mathcal{L}_j$. Esto finaliza la demostración del lema. □

En el siguiente resultado estudiamos la dimensión del lugar singular de $\mathrm{pcl}(\Gamma_i^*)$ en el hiperplano del infinito.

Lema 3.2.14. *El lugar singular de* $\mathrm{pcl}(\Gamma_i^*)$ *en el hiperplano del infinito tiene dimensión a lo sumo* $d - m - 2$.

Demostración. Por [GL02a, Lemma 1.1], el lugar singular de $\mathrm{pcl}(\Gamma_i^*)$ en el hiperplano del infinito está contenido en el lugar singular de $\mathrm{pcl}(\Gamma_i^*) \cap \{T_0 = 0\}$. Por otro lado, el Lema 3.2.13 prueba que $\mathrm{pcl}(\Gamma_i^*) \cap \{T_0 = 0\}$ es una unión de variedades lineales de dimensión $d-m-1$. Por lo tanto, su lugar singular es una unión finita de variedades lineales de dimensión a lo sumo $d - m - 2$ (ver, por ejemplo, [CLO92, §9.6, Exercise 11]), lo cual implica el lema. □

Finalizamos dando el principal resultado de esta sección.

Teorema 3.2.15. *Sea* $p > 2$ *y* $q > d$. *Si* $3 \le r \le d - m$, *entonces* $\mathrm{pcl}(\Gamma_i^*) \subset \mathbb{P}^{d+i}$ *es una intersección completa normal de dimensión* $d - m$ *y grado* $d!/(d-i)!$.

Demostración. El Lema 3.2.12 muestra que $\mathrm{pcl}(\Gamma_i^*)$ es una intersección completa de dimensión $d-m$ y grado $d!/(d-i)!$. Por otro lado, el Teorema 3.2.10 y el Lema 3.2.14 muestran que el lugar singular de $\mathrm{pcl}(\Gamma_i^*)$ tiene codimension al menos 2 en $\mathrm{pcl}(\Gamma_i^*)$. Esto implica que $\mathrm{pcl}(\Gamma_i^*)$ es regular en codimensión 1 y, por lo tanto, una variedad normal. □

Combinando el Teorema 3.2.15 con el Teorema 2.1.10 concluimos que $\mathrm{pcl}(\Gamma_i^*)$ es absolutamente irreducible de dimensión $d-m$ y grado $d!/(d-i)!$, y por lo tanto $\Gamma_i^* \subset \mathbb{A}^{d+i}$ resulta también absolutamente irreducible de dimensión $d-m$ y grado $d!/(d-i)!$. El Lema 3.2.1 muestra que la variedad Γ_i definida en (3.5) es un subconjunto abierto Zariski no vacío de Γ_i^*. Como Γ_i^* es absolutamente irreducible concluimos que la clausura Zariski de Γ_i es Γ_i^*.

3.2.3. El número de puntos $\mathbb{F}_q$–racionales

En esta sección damos una estimación del número de puntos $\mathbb{F}_q$–racionales de la variedad $\Gamma_i^* \subset \mathbb{A}^{d+i}$. Para esto vamos a usar la estimación sobre el número de puntos $\mathbb{F}_q$–racionales de una intersección completa proyectiva normal del Teorema 2.2.10.

Por el Teorema 3.2.15, la variedad proyectiva $\mathrm{pcl}(\Gamma_i^*) \subset \mathbb{P}^{d+i}$ es una intersección completa normal definida sobre $\mathbb{F}_q$ de dimensión $d-m$. Por lo tanto, aplicando la estimación (2.7) del Teorema 2.2.10, obtenemos que

$$\big||\mathrm{pcl}(\Gamma_i^*)(\mathbb{F}_q)| - p_{d-m}\big| \leq (\delta_i(D_i-2)+2)q^{d-m-\frac{1}{2}} + 14D_i^2\delta_i^2 q^{d-m-1},$$

donde $D_i := \sum_{j=1}^{i}(d-j) = id - i(i+1)/2$ y $\delta_i := d!/(d-i)!$.

Por otro lado, por el Lema 3.2.13 tenemos que $\mathrm{pcl}(\Gamma_i^*)^\infty := \mathrm{pcl}(\Gamma_i^*) \cap \{T_0 = 0\} \subset \mathbb{P}^{d+i-1}$ es una unión finita de a lo sumo $i+1$ variedades lineales de dimensión $d-m-1$. Deducimos que el número de puntos $\mathbb{F}_q$–racionales de $\mathrm{pcl}(\Gamma_i^*)^\infty$ es al menos p_{d-m-1} y a lo sumo $(i+1)p_{d-m-1}$. Así,

$$\begin{aligned}
\big||\Gamma_i^*(\mathbb{F}_q)| - q^{d-m}\big| =& \big||\mathrm{pcl}(\Gamma_i^*)(\mathbb{F}_q)| - |\mathrm{pcl}(\Gamma_i^*(\mathbb{F}_q))^\infty| - p_{d-m} + p_{d-m-1}\big| \\
\leq& \big||\mathrm{pcl}(\Gamma_i^*)(\mathbb{F}_q)| - p_{d-m}\big| + \big||\mathrm{pcl}(\Gamma_i^*(\mathbb{F}_q))^\infty| - p_{d-m-1}\big| \\
\leq& (\delta_i(D_i-2)+2)q^{d-m-\frac{1}{2}} + 14D_i^2\delta_i^2 q^{d-m-1} + ip_{d-m-1} \\
\leq& (\delta_i(D_i-2)+2)q^{d-m-\frac{1}{2}} + (14D_i^2\delta_i^2 + 2i)q^{d-m-1}. \qquad (3.15)
\end{aligned}$$

Por lo tanto, tenemos el siguiente resultado.

Teorema 3.2.16. *Sea $p > 2$ y $q > d$. Sean r, d y m enteros positivos tales que $3 \leq r \leq d-m$ y sea i un entero positivo tal que $r+1 \leq i \leq d$. Si $\Gamma_i^* \subset \mathbb{A}^{d+i}$ es la $\mathbb{F}_q$–variedad afín definida en* (3.7)*, entonces la cantidad de puntos $\mathbb{F}_q$–racionales de Γ_i^* satisface la siguiente estimación:*

$$\big||\Gamma_i^*(\mathbb{F}_q)| - q^{d-m}\big| \leq (\delta_i(D_i-2)+2)q^{d-m-\frac{1}{2}} + (14D_i^2\delta_i^2 + 2i)q^{d-m-1},$$

donde $D_i := \sum_{j=1}^{i}(d-j) = id - i(i+1)/2$ y $\delta_i := d!/(d-i)!$.

A continuación estimamos el número de puntos $\mathbb{F}_q$–racionales de la variedad Γ_i^* con coordenadas distintas dos a dos. Observemos que, debido al Lema 3.2.1, esta cantidad corresponde al número de puntos $\mathbb{F}_q$–racionales de la cuasi–variedad Γ_i definida en (3.5). Esta estimación la usaremos más adelante en el estudio del problema de determinar el comportamiento del conjunto de valores y la distribución de patrones de factorización en familias lineales.

Con este propósito, consideramos la siguiente $\mathbb{F}_q$–variedad afín:

$$\Gamma_i^{*,=} := \Gamma_i^* \bigcap \bigcup_{1\leq j<k\leq i} \{T_j = T_k\}.$$

Observamos que $\Gamma_i^{*,=} = \Gamma_i^* \cap \mathcal{H}_i$, donde $\mathcal{H}_i \subset \mathbb{A}^{d+i}$ es la hipersuperficie definida por el polinomio $F_i := \prod_{1\leq j<k\leq i}(T_j - T_k)$. De la desigualdad de Bézout (2.1.14) deducimos que

$$\deg \Gamma_i^{*,=} \leq \delta_i \binom{i}{2}. \tag{3.16}$$

Por otro lado, afirmamos que $\Gamma_i^{*,=}$ tiene dimensión a lo sumo $d-m-1$. En efecto, sea $(\boldsymbol{a}_0, \boldsymbol{\alpha})$ un punto arbitrario de $\Gamma_i^{*,=}$. Sin pérdida de generalidad podemos suponer que $\alpha_1 = \alpha_2$. De la definición de las diferencias divididas deducimos que $f'_{\boldsymbol{a}_0}(\alpha_1) = 0$, lo que implica que el polinomio $f_{\boldsymbol{a}_0}$ tiene raíces múltiples. Por la afirmación de la demostración del Corolario 3.2.11, el conjunto de puntos $(\boldsymbol{a}_0, \boldsymbol{\alpha})$ de Γ_i^* tales que $f_{\boldsymbol{a}_0}$ tiene raíces múltiples está contenido en una subvariedad de Γ_i^* de codimensión al menos 1. Por lo tanto, deducimos nuestra afirmación.

Combinando esta afirmación con (3.16) y la Proposición 2.2.1, obtenemos que

$$\left|\Gamma_i^{*,=}(\mathbb{F}_q)\right| \leq \delta_i \binom{i}{2} q^{d-m-1}. \tag{3.17}$$

Así, dado que $\Gamma_i(\mathbb{F}_q) = \Gamma_i^*(\mathbb{F}_q) \setminus \Gamma_i^{*,=}(\mathbb{F}_q)$, de (3.15) y (3.17) deducimos que

$$\begin{aligned}\left||\Gamma_i(\mathbb{F}_q)| - q^{d-m}\right| \leq& \left||\Gamma_i^*(\mathbb{F}_q)| - q^{d-m}\right| + |\Gamma_i^{*,=}(\mathbb{F}_q)| \\ \leq& (\delta_i(D_i-2)+2)q^{d-m-\frac{1}{2}} + \big(14D_i^2\delta_i^2 + i(i-1)\delta_i/2 + 2i\big)q^{d-m-1}.\end{aligned}$$

Por consiguiente, obtenemos el siguiente resultado.

Teorema 3.2.17. *Con las notaciones y las hipótesis del Teorema 3.2.16, la cantidad de puntos $\mathbb{F}_q$–racionales de la cuasi–variedad $\Gamma_i \subset \mathbb{A}^{d+i}$ definida en* (3.5) *satisface la siguiente estimación:*

$$\left||\Gamma_i(\mathbb{F}_q)| - q^{d-m}\right| \leq (\delta_i(D_i-2)+2)q^{d-m-\frac{1}{2}} + \big(14D_i^2\delta_i^2 + i(i-1)\delta_i/2 + 2i\big)q^{d-m-1},$$

donde $D_i := id - i(i+1)/2$ y $\delta_i := d!/(d-i)!$.

Capítulo 4

Intersecciones completas dadas por polinomios simétricos

En este capítulo estudiamos el conjunto de puntos $\mathbb{F}_q$–racionales de $\mathbb{F}_q$–variedades definidas por polinomios invariantes bajo la acción del grupo simétrico de permutaciones de sus coordenadas. Probamos que ciertas propiedades geométricas de los polinomios simétricos que definen estas variedades implican que las mismas son intersecciones completas con buen comportamiento en el infinito, cuyo lugar singular tiene codimensión al menos 3. Estos resultados, junto con las estimaciones para intersecciones completas proyectivas definidas sobre $\mathbb{F}_q$ que se encuentra en la Sección 2.2.2 (ver Teoremas 2.2.10 y 2.2.11), nos permitirán estimar el número de puntos $\mathbb{F}_q$–racionales de las correspondientes intersecciones completas. Dichas estimaciones nos servirán para estudiar, en los capítulos siguientes, el valor promedio del cardinal del conjunto de valores y la distribución de los patrones de factorización de familias de polinomios univariados con coeficientes en $\mathbb{F}_q$.

4.1. Estimaciones para intersecciones completas simétricas

En esta sección presentamos estimaciones de la cantidad de puntos $\mathbb{F}_q$–racionales de intersecciones completas definidas por polinomios simétricos con coeficientes en $\mathbb{F}_q$. Comenzamos precisando las $\mathbb{F}_q$–variedades que vamos a estudiar. Sean s, r, m enteros positivos con $m \leq s \leq r - m - 2$. Sean $Y_1, \ldots, Y_s$ indeterminadas sobre $\overline{\mathbb{F}}_q$. Sean $S_1, \ldots, S_m \in \mathbb{F}_q[Y_1, \ldots, Y_s]$ y sea $W_s \subset \mathbb{A}^s$ la $\mathbb{F}_q$–variedad afín definida por ellos. Consideramos el peso wt sobre $\mathbb{F}_q[Y_1, \ldots, Y_s]$ definido por $\mathsf{wt}(Y_j) := j$ para $1 \leq j \leq s$ y denotamos por $S_1^{\mathsf{wt}}, \ldots, S_m^{\mathsf{wt}}$ las componentes de mayor peso de $S_1, \ldots, S_m$. Sean $(\partial \boldsymbol{S}/\partial \boldsymbol{Y})$ y $(\partial \boldsymbol{S}^{\mathsf{wt}}/\partial \boldsymbol{Y})$ las matrices Jacobianas de $S_1, \ldots, S_m$ y $S_1^{\mathsf{wt}}, \ldots, S_m^{\mathsf{wt}}$ con respecto a $Y_1, \ldots, Y_s$ respectivamente. Supongamos que los polinomios $S_1, \ldots, S_m$ satisfacen las siguientes condiciones:

(H_1) $S_1, \ldots, S_m$ forman una sucesión regular de $\mathbb{F}_q[Y_1, \ldots, Y_s]$;

(H_2) $(\partial \boldsymbol{S}/\partial \boldsymbol{Y})(\boldsymbol{y})$ tiene rango máximo m para cada $\boldsymbol{y} \in W_s$;

(H_3) $S_1^{\mathsf{wt}}, \ldots, S_m^{\mathsf{wt}}$ satisfacen (H_1) y (H_2).

Como ya mencionamos en el capítulo anterior, un polinomio $F \in \mathbb{F}_q[Y_1, \ldots, Y_s]$ se dice *homogéneo con peso* (con respecto a la graduación definida por el peso wt) si todos los monomios que aparecen en su representación densa tienen el mismo peso. En este sentido, $S_1^{\mathsf{wt}}, \ldots, S_m^{\mathsf{wt}}$ son homogéneos con peso.

A continuación demostramos que polinomios $S_1, \ldots, S_m$ como los de arriba, tales que $S_1, \ldots, S_m$ y $S_1^{\mathsf{wt}}, \ldots, S_m^{\mathsf{wt}}$ satisfacen la hipótesis (H_2), necesariamente satisfacen las hipótesis (H_1) y (H_3). Sin embargo, a efectos de la exposición resulta conveniente referirnos de forma separada a las hipótesis (H_1), (H_2) y (H_3).

Observación 4.1.1. *Si $S_1^{\mathsf{wt}}, \ldots, S_m^{\mathsf{wt}}$ satisfacen (H_2), entonces $S_1^{\mathsf{wt}}, \ldots, S_m^{\mathsf{wt}}$ forman una sucesión regular de $\mathbb{F}_q[Y_1, \ldots, Y_s]$.*

Demostración. Denotamos por $W^{\mathsf{wt}} \subset \mathbb{A}^s$ la variedad afín definida por $S_1^{\mathsf{wt}}, \ldots, S_m^{\mathsf{wt}}$ y sea $\mathcal{C}$ una componente absolutamente irreducible de W^{wt}. Por el Teorema 2.1.7 tenemos que $\dim \mathcal{C} \geq s - m$. Por otro lado, si $\boldsymbol{y} \in \mathcal{C}$, entonces el hecho de que $S_1^{\mathsf{wt}}, \ldots, S_m^{\mathsf{wt}}$ satisfacen (H_2) implica que el espacio tangente $\mathcal{T}_{\boldsymbol{y}} W^{\mathsf{wt}}$ a W^{wt} en $\boldsymbol{y}$ tiene dimensión a lo sumo $s - m$. Como además $\dim \mathcal{T}_{\boldsymbol{y}} W^{\mathsf{wt}} \geq \dim_{\boldsymbol{y}} W^{\mathsf{wt}} \geq \dim \mathcal{C}$, concluimos que $\dim \mathcal{T}_{\boldsymbol{y}} W^{\mathsf{wt}} = \dim \mathcal{C} = s - m$. En otras palabras, W^{wt} es de dimensión pura $s - m$. Finalmente, como $S_1^{\mathsf{wt}}, \ldots, S_m^{\mathsf{wt}}$ son homogéneos con peso, la variedad afín $V(S_1^{\mathsf{wt}}, \ldots, S_j^{\mathsf{wt}})$ debe ser de dimensión pura $s - j$ para todo $1 \leq j \leq m$. Resulta así que $S_1^{\mathsf{wt}}, \ldots, S_m^{\mathsf{wt}}$ forman una sucesión regular de $\mathbb{F}_q[Y_1, \ldots, Y_s]$. □

Observación 4.1.2. *Si $S_1, \ldots, S_m$ y $S_1^{\mathsf{wt}}, \ldots, S_m^{\mathsf{wt}}$ satisfacen (H_2), entonces $S_1, \ldots, S_m$ forman una sucesión regular de $\mathbb{F}_q[Y_1, \ldots, Y_s]$.*

Demostración. Sea $S_j^{h_{\mathsf{wt}}} \in \mathbb{F}_q[Y_0, Y_1, \ldots, Y_s]$ la homogeneización de S_j con respecto al peso wt para $1 \leq j \leq m$. Afirmamos que la variedad afín $V(S_1^{h_{\mathsf{wt}}}, \ldots, S_m^{h_{\mathsf{wt}}}) \subset \mathbb{A}^{s+1}$ es de dimensión pura $s - m + 1$.

Para mostrar esta afirmación, sea $\mathcal{C}$ una componente absolutamente irreducible de $V(S_1^{h_{\mathsf{wt}}}, \ldots, S_m^{h_{\mathsf{wt}}})$. Es claro que $\dim \mathcal{C} \geq s - m + 1$. Sea ahora $\boldsymbol{y} := (y_0, \ldots, y_s) \in \mathcal{C}$ un punto regular de $V(S_1^{h_{\mathsf{wt}}}, \ldots, S_m^{h_{\mathsf{wt}}})$. Sin pérdida de generalidad podemos suponer que, o bien $y_0 = 1$, o bien $y_0 = 0$. Si $y_0 = 1$, entonces $\boldsymbol{y} \in W_s$, y dado que $S_1, \ldots, S_m$ satisfacen (H_2), entonces $\operatorname{rg}(\partial \boldsymbol{S}^{\boldsymbol{h}_{\mathsf{wt}}} / \partial \boldsymbol{Y})(\boldsymbol{y}) = \operatorname{rg}(\partial \boldsymbol{S} / \partial \boldsymbol{Y})(\boldsymbol{y}) = m$. Así, tenemos que $\dim \mathcal{T}_{\boldsymbol{y}} V(S_1^{h_{\mathsf{wt}}}, \ldots, S_m^{h_{\mathsf{wt}}}) \leq s - m + 1$. Por otro lado, si $y_0 = 0$, entonces $(y_1, \ldots, y_s) \in W^{\mathsf{wt}}$ y, dado que $S_1^{\mathsf{wt}}, \ldots, S_m^{\mathsf{wt}}$ satisfacen (H_2), resulta $\operatorname{rg}(\partial \boldsymbol{S}^{\boldsymbol{h}_{\mathsf{wt}}} / \partial \boldsymbol{Y})(\boldsymbol{y}) = \operatorname{rg}(\partial \boldsymbol{S}^{\mathsf{wt}} / \partial \boldsymbol{Y})(\boldsymbol{y}) = m$. Así, $\dim \mathcal{T}_{\boldsymbol{y}} V(S_1^{h_{\mathsf{wt}}}, \ldots, S_m^{h_{\mathsf{wt}}}) \leq s - m + 1$. En ambos casos, tenemos que $\dim \mathcal{T}_{\boldsymbol{y}} V(S_1^{h_{\mathsf{wt}}}, \ldots, S_m^{h_{\mathsf{wt}}}) \leq s - m + 1$. Como $\boldsymbol{y} \in \mathcal{C}$ es un punto regular de $V(S_1^{h_{\mathsf{wt}}}, \ldots, S_m^{h_{\mathsf{wt}}})$, tenemos que $\dim \mathcal{T}_{\boldsymbol{y}} W^{\mathsf{wt}} = \dim_{\boldsymbol{y}} W^{\mathsf{wt}}$. Así concluimos que $\dim \mathcal{C} = s - m + 1$, lo que finaliza la demostración de la afirmación.

Combinando la afirmación anterior con el hecho de que $S_1^{h_{\mathsf{wt}}}, \ldots, S_m^{h_{\mathsf{wt}}}$ son homogéneos con peso concluimos que $S_1^{h_{\mathsf{wt}}}, \ldots, S_m^{h_{\mathsf{wt}}}$ forman una sucesión regular de $\mathbb{F}_q[Y_0, \ldots, Y_s]$. En particular, tenemos que $V(S_1^{h_{\mathsf{wt}}}, \ldots, S_j^{h_{\mathsf{wt}}}) \subset \mathbb{A}^{s+1}$ es de dimensión pura $s - j + 1$ para cada $1 \leq j \leq m$. Por otro lado, sea $W_s^j := V(S_1, \ldots, S_j) \subset \mathbb{A}^s$ y sea $\operatorname{pcl}(W_s^j)^{\mathsf{wt}} \subset \mathbb{P}^s$ la variedad definida por la homogeneización de cada elemento del ideal $(S_1, \ldots, S_j)$ con respecto a la graduación definida por el peso wt. Observemos

que $\mathrm{pcl}(W_s^j)^{\mathsf{wt}}$ tiene dimensión $\dim W_s^j$. Así, si consideramos el cono afín $\widehat{W_s^j} \subset \mathbb{A}^{s+1}$ de $\mathrm{pcl}(W_s^j)^{\mathsf{wt}}$, observamos que tiene dimensión $\dim W_s^j + 1$ y está contenido en la variedad afín $V(S_1^{h_{\mathsf{wt}}}, \ldots, S_j^{h_{\mathsf{wt}}})$ definida por $S_1^{h_{\mathsf{wt}}}, \ldots, S_j^{h_{\mathsf{wt}}}$ para cada $1 \leq j \leq m$, por lo que concluimos que $\dim W_s^j = s - j$ para $1 \leq j \leq m$. Esto prueba que $S_1, \ldots, S_m$ forman una sucesión regular. □

Observación 4.1.3. *De la hipótesis* (H_1) *tenemos que la variedad afín* $W_s \subset \mathbb{A}^s$ *definida por* $S_1, \ldots, S_m$ *es una intersección completa conjuntista de dimensión* $s-m$. *Además, por* (H_2) *la subvariedad de* W_s *definida por el conjunto de ceros comunes de los menores maximales de la matriz Jacobiana* $(\partial \boldsymbol{S}/\partial \boldsymbol{Y})$ *tiene codimensión al menos uno. Entonces [Eis95, Theorem 18.15] prueba que* $S_1, \ldots, S_m$ *definen un ideal radical. Por lo tanto,* W_s *resulta una intersección completa.*

Sean $X_1, \ldots, X_r$ indeterminadas sobre $\overline{\mathbb{F}}_q$ y $\Pi_1, \ldots, \Pi_s$ los primeros s polinomios simétricos elementales de $\mathbb{F}_q[X_1, \ldots, X_r]$. Sean $R_1, \ldots, R_m \in \mathbb{F}_q[X_1, \ldots, X_r]$ los polinomios definidos por

$$R_i := S_i(\Pi_1, \ldots, \Pi_s) \quad (1 \leq i \leq m). \tag{4.1}$$

Sea $d_i := \deg(R_i)$ para $1 \leq i \leq m$.

4.1.1. Aspectos geométricos

En esta sección discutimos algunas propiedades de la geometría de la $\mathbb{F}_q$–variedad afín $V_r \subset \mathbb{A}^r$ definida por los polinomios $R_1, \ldots, R_m$ de (4.1). Con este objetivo, consideramos el siguiente morfismo sobreyectivo de $\mathbb{F}_q$–variedades afines:

$$\begin{aligned} \boldsymbol{\Pi}^r : \mathbb{A}^r &\to \mathbb{A}^r \\ \boldsymbol{x} &\mapsto (\Pi_1(\boldsymbol{x}), \ldots, \Pi_r(\boldsymbol{x})). \end{aligned}$$

Es fácil ver que $\boldsymbol{\Pi}^r$ es un morfismo finito (ver Teorema 2.1.4).

Considerando los polinomios $S_1, \ldots, S_m$ como elementos de $\mathbb{F}_q[Y_1, \ldots, Y_r]$, denotamos por $W_r := V(S_1, \ldots, S_m) \subset \mathbb{A}^r$ la variedad definida por $S_1, \ldots, S_m$. Observemos que $V_r = (\boldsymbol{\Pi}^r)^{-1}(W_r)$. Dado que $S_1, \ldots, S_m$ forman una sucesión regular de $\mathbb{F}_q[Y_1, \ldots, Y_r]$, la variedad W_r tiene dimensión pura $r - m$. Del Teorema 2.1.4 deducimos que V_r es de dimensión pura $r - m$. Por otro lado, la variedad $W_r^j := V(S_1, \ldots, S_j) \subset \mathbb{A}^r$ definida por los polinomios $S_1, \ldots, S_j$ tiene dimensión pura $r - j$ para cada $1 \leq j \leq m$. Esto implica que la variedad $V_r^j := (\boldsymbol{\Pi}^r)^{-1}(W_r^j)$ definida por $R_1, \ldots, R_j$ tiene dimensión pura $r - j$ para cada $1 \leq j \leq m$. Así, los polinomios $R_1, \ldots, R_m$ forman una sucesión regular de $\mathbb{F}_q[X_1, \ldots, X_r]$ y probamos el siguiente resultado.

Lema 4.1.4. *Sea* $V_r \subset \mathbb{A}^r$ *la* $\mathbb{F}_q$*–variedad afín definida por* $R_1, \ldots, R_m$. *Entonces* V_r *es una intersección completa conjuntista de dimensión* $r - m$.

A continuación estudiamos el lugar singular de V_r. Para esto, consideramos el siguiente morfismo de $\mathbb{F}_q$–variedades afines:

$$\begin{aligned} \boldsymbol{\Pi} : V_r &\to W_s \\ \boldsymbol{x} &\mapsto (\Pi_1(\boldsymbol{x}), \ldots, \Pi_s(\boldsymbol{x})). \end{aligned}$$

Para $\boldsymbol{x} \in V_r$ e $\boldsymbol{y} := \boldsymbol{\Pi}(\boldsymbol{x})$, denotamos por $\mathcal{T}_{\boldsymbol{x}} V_r$ y $\mathcal{T}_{\boldsymbol{y}} W_s$ los espacios tangentes a V_r en $\boldsymbol{x}$ y a W_s en $\boldsymbol{y}$ respectivamente. Asimismo, consideramos la diferencial de $\boldsymbol{\Pi}$ en $\boldsymbol{x}$:

$$\begin{aligned} \mathrm{d}_{\boldsymbol{x}}\boldsymbol{\Pi} : \mathcal{T}_{\boldsymbol{x}} V_r &\to \mathcal{T}_{\boldsymbol{y}} W_s \\ \boldsymbol{v} &\mapsto A(\boldsymbol{x}) \cdot \boldsymbol{v}, \end{aligned}$$

donde $A(\boldsymbol{x})$ es la siguiente matriz de tamaño $s \times r$:

$$A(\boldsymbol{x}) := \left(\frac{\partial \boldsymbol{\Pi}}{\partial \boldsymbol{X}}\right)(\boldsymbol{x}) := \left(\frac{\partial \Pi_i}{\partial X_j}(\boldsymbol{x})\right)_{1 \le i \le s,\, 1 \le j \le r}. \tag{4.2}$$

El principal resultado de esta sección es una cota superior de la dimensión del lugar singular de V_r. Para probar tal cota, comenzamos con algunas observaciones acerca de la matriz Jacobiana de los polinomios simétricos elementales. En primer lugar, es sabido que las derivadas parciales de los polinomios simétricos elementales Π_i satisfacen las siguientes igualdades para $1 \le i, j \le r$ (ver, por ejemplo, [LP02]):

$$\frac{\partial \Pi_i}{\partial X_j} = \Pi_{i-1} - X_j \Pi_{i-2} + X_j^2 \Pi_{i-3} + \cdots + (-1)^{i-1} X_j^{i-1}.$$

En consecuencia, si A_r denota la matriz de Vandermonde de tamaño $r \times r$

$$A_r := (X_j^{i-1})_{1 \le i,j \le r},$$

entonces la matriz Jacobiana $(\partial \boldsymbol{\Pi}^r / \partial \boldsymbol{X})$ de $\boldsymbol{\Pi}^r := (\Pi_1, \ldots, \Pi_r)$ con respecto a $X_1, \ldots, X_r$ se puede expresar de la siguiente manera:

$$\left(\frac{\partial \boldsymbol{\Pi}^r}{\partial \boldsymbol{X}}\right) := B_r \cdot A_r := \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ \Pi_1 & -1 & 0 & & \\ \Pi_2 & -\Pi_1 & 1 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 0 \\ \Pi_{r-1} & -\Pi_{r-2} & \Pi_{r-3} & \cdots & (-1)^{r-1} \end{pmatrix} \cdot A_r. \tag{4.3}$$

Observamos que B_r es una matriz cuadrada y triangular inferior cuyo determinante es igual a $(-1)^{(r-1)r/2}$. Esto implica que el determinante de $(\boldsymbol{\Pi}^r / \partial \boldsymbol{X})$ es igual, salvo el signo, al determinante de A_r, es decir,

$$\det\left(\frac{\partial \boldsymbol{\Pi}^r}{\partial \boldsymbol{X}}\right) = (-1)^{(r-1)r/2} \prod_{1 \le i < j \le r} (X_j - X_i).$$

Denotemos por $(\partial \boldsymbol{R} / \partial \boldsymbol{X}) := (\partial R_i / \partial X_j)_{1 \le i \le m, 1 \le j \le r}$ la matriz Jacobiana de $R_1, \ldots, R_m$ con respecto a $X_1, \ldots, X_r$.

Teorema 4.1.5. *El conjunto de puntos $\boldsymbol{x} \in V_r$ para los cuales $(\partial \boldsymbol{R} / \partial \boldsymbol{X})(\boldsymbol{x})$ no es de rango completo tiene dimensión a lo sumo $s - 1$. En particular, el lugar singular Σ_r de V_r tiene dimensión a lo sumo $s - 1$.*

Demostración. Por la regla de la cadena, las derivadas parciales de los polinomios R_i satisfacen la siguiente igualdad:

$$\left(\frac{\partial \boldsymbol{R}}{\partial \boldsymbol{X}}\right) = \left(\frac{\partial \boldsymbol{S}}{\partial \boldsymbol{Y}} \circ \boldsymbol{\Pi}\right) \cdot \left(\frac{\partial \boldsymbol{\Pi}}{\partial \boldsymbol{X}}\right).$$

Fijemos un punto arbitrario $\boldsymbol{x} \in V_r$ para el cual $(\partial \boldsymbol{R}/\partial \boldsymbol{X})(\boldsymbol{x})$ no es de rango completo. Sea $\boldsymbol{v} \in \mathbb{A}^m$ un vector no nulo en el núcleo a izquierda de $(\partial \boldsymbol{R}/\partial \boldsymbol{X})(\boldsymbol{x})$. Así,

$$\boldsymbol{0} = \boldsymbol{v} \cdot \left(\frac{\partial \boldsymbol{R}}{\partial \boldsymbol{X}}\right)(\boldsymbol{x}) = \boldsymbol{v} \cdot \left(\frac{\partial \boldsymbol{S}}{\partial \boldsymbol{Y}}\right)\big(\boldsymbol{\Pi}(\boldsymbol{x})\big) \cdot A(\boldsymbol{x}),$$

donde $A(\boldsymbol{x})$ es la matriz definida en (4.2). Por la hipótesis (H_2) la matriz Jacobiana $(\partial \boldsymbol{S}/\partial \boldsymbol{Y})\big(\boldsymbol{\Pi}(\boldsymbol{x})\big)$ es de rango completo; por lo tanto, el vector $\boldsymbol{w} := \boldsymbol{v} \cdot (\partial \boldsymbol{S}/\partial \boldsymbol{Y})\big(\Pi(\boldsymbol{x})\big) \in \mathbb{A}^s$ es no nulo y $\boldsymbol{w} \cdot A(\boldsymbol{x}) = \boldsymbol{0}$. Así, todos los menores maximales de la matriz $A(\boldsymbol{x})$ deben ser cero.

Observemos que $A(\boldsymbol{x})$ es la submatriz de $(\partial \boldsymbol{\Pi}^r/\partial \boldsymbol{X})(\boldsymbol{x})$ de tamaño $s \times r$ que se obtiene al considerar las primeras s filas de $(\partial \boldsymbol{\Pi}^r/\partial \boldsymbol{X})(\boldsymbol{x})$. Por lo tanto, de (4.3) concluimos que

$$A(\boldsymbol{x}) = B_{s,r}(\boldsymbol{x}) \cdot A_r(\boldsymbol{x}),$$

donde $B_{s,r}(\boldsymbol{x})$ es la submatriz de $B_r(\boldsymbol{x})$ de tamaño $s \times r$ que consiste de las primeras s filas de $B_r(\boldsymbol{x})$. Dado que las últimas $r-s$ columnas de $B_{s,r}(\boldsymbol{x})$ son nulas, podemos reescribir la identidad anterior de la siguiente manera:

$$A(\boldsymbol{x}) = B_s(\boldsymbol{x}) \cdot (x_j^{i-1})_{1 \le i \le s,\, 1 \le j \le r}, \tag{4.4}$$

donde $B_s(\boldsymbol{x})$ es la submatriz de $B_r(\boldsymbol{x})$ de tamaño $s \times s$ que se obtiene al considerar las primeras s filas y las primeras s columnas de $B_r(\boldsymbol{x})$.

Para $1 \le l_1 < \cdots < l_s \le r$, sea $I := (l_1, \ldots, l_s)$ y consideremos la submatriz $M_I(\boldsymbol{x})$ de $A(\boldsymbol{x})$ de tamaño $s \times s$ que se obtiene al elegir las columnas $l_1, \ldots, l_s$ de $A(\boldsymbol{x})$, es decir, $M_I(\boldsymbol{x}) := (\partial \Pi_i/\partial X_{l_j})_{1 \le i,j \le s}(\boldsymbol{x})$. De (4.3) y (4.4) deducimos que $M_I(\boldsymbol{x}) = B_s(\boldsymbol{x}) \cdot A_{s,I}(\boldsymbol{x})$, donde $A_{s,I}(\boldsymbol{x})$ es la matriz de Vandermonde $A_{s,I}(\boldsymbol{x}) := (x_{l_j}^{i-1})_{1 \le i,j \le s}$. En consecuencia, tenemos que

$$\det\big(M_I(\boldsymbol{x})\big) = (-1)^{\frac{(s-1)s}{2}} \det A_{s,I}(\boldsymbol{x}) = (-1)^{\frac{(s-1)s}{2}} \prod_{1 \le m < n \le s} (x_{l_n} - x_{l_m}) = 0. \tag{4.5}$$

Como (4.5) es válido para todo $I := (l_1, \ldots, l_s)$ elegido como arriba, concluimos que $\boldsymbol{x}$ tiene a lo sumo $s-1$ coordenadas distintas. En particular, el conjunto de puntos $\boldsymbol{x} \in V_r$ para los cuales $\operatorname{rg}(\partial \boldsymbol{R}/\partial \boldsymbol{X})(\boldsymbol{x}) < m$ está contenido en una unión finita de variedades lineales de $\mathbb{A}^r$ de dimensión $s-1$ y por lo tanto tiene dimensión a lo sumo $s-1$.

Finalmente, sea $\boldsymbol{x}$ un punto arbitrario de Σ_r. Por el Lema 4.1.4 se tiene que $\dim \mathcal{T}_{\boldsymbol{x}} V_r > r - m$. Así, el rango de $(\partial \boldsymbol{R}/\partial \boldsymbol{X})(\boldsymbol{x})$ es menor que m, pues de otra forma tendríamos que $\dim \mathcal{T}_{\boldsymbol{x}} V_r \le r - m$, lo que contradiría el hecho de que $\boldsymbol{x}$ es un punto singular de V_r. Esto finaliza la demostración del teorema. □

De la demostración del Teorema 4.1.5 concluimos que el lugar singular de V_r está contenido en una variedad simple de dimensión baja.

Observación 4.1.6. *Sean las notaciones e hipótesis como en el Teorema 4.1.5. De la demostración del Teorema 4.1.5 obtenemos la siguiente inclusión:*

$$\Sigma_r \subset \bigcup_{\mathcal{I}} \mathcal{L}_{\mathcal{I}},$$

donde $\mathcal{I} := \{I_1, \ldots, I_{s-1}\}$ *recorre todas las particiones de* $\{1, \ldots, r\}$ *en* $s-1$ *subconjuntos no vacíos* $I_j \subset \{1, \ldots, r\}$ *y* $\mathcal{L}_{\mathcal{I}} := \operatorname{span}(\boldsymbol{v}^{I_1}, \ldots, \boldsymbol{v}^{I_{s-1}})$ *es la variedad lineal generada por los vectores* $\boldsymbol{v}^{I_j} := (v_1^{I_j}, \ldots, v_r^{I_j})$ *definidos por* $v_m^{I_j} := 1$ *para* $m \in I_j$ *y* $v_m^{I_j} := 0$ *para* $m \notin I_j$.

Del Lema 4.1.4 y del Teorema 4.1.5 se obtienen más propiedades algebraicas y geométricas de los polinomios R_i y de la variedad afín V_r. De acuerdo al Teorema 4.1.5, el conjunto de puntos $\boldsymbol{x} \in V_r$ para los cuales la matriz $(\partial \boldsymbol{R}/\partial \boldsymbol{X})(\boldsymbol{x})$ no es de rango completo, tiene dimensión a lo sumo $s-1$. Como los polinomios $R_1, \ldots, R_m$ forman una sucesión regular y $s \leq r - m - 2$ concluimos, por [Eis95, Theorem 18.15], que $R_1, \ldots, R_m$ definen un ideal radical de $\mathbb{F}_q[X_1, \ldots, X_r]$, y así V_r es una intersección completa. Finalmente, por la desigualdad de Bézout (2.1.14) se tiene que $\deg V_r \leq \prod_{i=1}^m d_i$. En otras palabras, obtenemos el siguiente resultado.

Corolario 4.1.7. *Los polinomios* $R_1, \ldots, R_m$ *definen un ideal radical y la variedad* V_r *es una intersección completa de grado a lo sumo* $\deg V_r \leq \prod_{i=1}^m d_i$.

4.1.2. La geometría de la clausura proyectiva

Vamos a utilizar los resultados obtenidos en la sección anterior sobre la geometría de la $\mathbb{F}_q$–variedad $V_r \subset \mathbb{A}^r$ para estimar el número de puntos $\mathbb{F}_q$–racionales de V_r. Dado que vamos a aplicar las estimaciones para intersecciones completas proyectivas definidas sobre $\mathbb{F}_q$ de la Sección 2.2.2, en esta sección consideramos la clausura proyectiva $\operatorname{pcl}(V_r) \subset \mathbb{P}^r$ de V_r. Comenzamos estudiando el comportamiento de $\operatorname{pcl}(V_r)$ en el hiperplano del infinito, para luego dar algunas propiedades de la geometría de $\operatorname{pcl}(V_r)$.

Consideramos la descomposición de cada polinomio R_i de (4.1) en sus componentes homogéneas, es decir,

$$R_i = R_i^{d_i} + R_i^{d_i-1} + \cdots + R_i^0,$$

donde cada $R_i^j \in \mathbb{F}_q[X_1, \ldots, X_r]$ es homogéneo de grado j o cero y $R_i^{d_i}$ es no nulo para $1 \leq j \leq m$. Así, la homogeneización de cada R_i es el siguiente polinomio de $\mathbb{F}_q[X_0, \ldots, X_r]$:

$$R_i^h = R_i^{d_i} + R_i^{d_i-1} X_0 + \cdots + R_i^0 X_0^{d_i}. \tag{4.6}$$

Se deduce que $R_i^h(0, X_1, \ldots, X_r) = R_i^{d_i}$ para $1 \leq i \leq m$.

A fin de obtener una cota superior no trivial sobre la dimensión del lugar singular de $\operatorname{pcl}(V_r)$ en el hiperplano del infinito, vamos a estudiar el conjunto de ceros comunes

de los polinomios $R_1^{d_1}, \ldots, R_m^{d_m}$. Con este objetivo, en el siguiente lema relacionamos cada $R_i^{d_i}$ con la componente S_i^{wt} de mayor peso de S_i. En efecto, sea $a_{i_1,\ldots,i_s} Y_1^{i_1} \cdots Y_s^{i_s}$ un monomio arbitrario que aparece en la representación densa de S_i. Entonces su peso $\mathsf{wt}(a_{i_1,\ldots,i_s} Y_1^{i_1} \cdots Y_s^{i_s}) = \sum_{j=1}^{s} j \cdot i_j$ coincide con el grado del correspondiente monomio $a_{i_1,\ldots,i_s} \Pi_1^{i_1} \cdots \Pi_s^{i_s}$ de R_i. Así deducimos fácilmente el siguiente resultado.

Lema 4.1.8. *Sea $R_i^{d_i}$ la componente homogénea de mayor grado de R_i y sea S_i^{wt} la componente de mayor peso de S_i. Entonces $R_i^{d_i} = S_i^{\mathsf{wt}}(\Pi_1, \ldots, \Pi_s)$ para $1 \leq i \leq m$.*

Sea $\Sigma_r^\infty \subset \mathbb{P}^r$ el lugar singular de $\mathrm{pcl}(V_r)$ en el hiperplano del infinito, es decir el conjunto de puntos singulares de $\mathrm{pcl}(V_r)$ que se encuentran en $\{X_0 = 0\}$. A partir del Lema 4.1.8 obtenemos el siguiente resultado en relación a Σ_r^∞.

Lema 4.1.9. *El lugar singular $\Sigma_r^\infty \subset \mathbb{P}^r$ en el hiperplano del infinito tiene dimensión a lo sumo $s-2$.*

Demostración. Sea $\boldsymbol{x} := (0 : x_1 : \ldots : x_r)$ un punto arbitrario de Σ_r^∞. Dado que los polinomios R_i^h se anulan en $\mathrm{pcl}(V_r)$, tenemos que $R_i^h(\boldsymbol{x}) = R_i^{d_i}(x_1, \ldots, x_r) = 0$ para $1 \leq i \leq m$. Denotamos por $(\partial \boldsymbol{R^d}/\partial \boldsymbol{X}) := (\partial R_i^{d_i}/\partial X_j)_{1 \leq i \leq m, 1 \leq j \leq r}$ la matriz Jacobiana de $R_1^{d_1}, \ldots, R_m^{d_m}$ con respecto a $X_1, \ldots, X_r$. Afirmamos que $(\partial \boldsymbol{R^d}/\partial \boldsymbol{X})(\boldsymbol{x})$ no es de rango completo. En efecto, si el rango de dicha matriz fuera igual a m, tendríamos que $\dim \mathcal{T}_{\boldsymbol{x}}(\mathrm{pcl}(V_r)) \leq r - m$, lo cual implicaría que $\boldsymbol{x}$ es un punto no singular de $\mathrm{pcl}(V_r)$. Esto contradice la hipótesis sobre $\boldsymbol{x}$.

Por otro lado, el Lema 4.1.8 asegura que $R_i^{d_i} = S_i^{\mathsf{wt}}(\Pi_1, \ldots, \Pi_s)$ para $1 \leq i \leq m$. Combinando la hipótesis $(\mathsf{H_3})$ con el Lema 4.1.8 deducimos que los polinomios $R_1^{d_1}, \ldots, R_m^{d_m}$ satisfacen las hipótesis del Teorema 4.1.5. Concluimos así que el conjunto de puntos $\boldsymbol{x}_{\mathrm{aff}} := (0, x_1, \ldots, x_r) \in V(R_1^{d_1}, \ldots, R_m^{d_m}) \subset \mathbb{A}^r$ tal que $(\partial \boldsymbol{R^d}/\partial \boldsymbol{X})(\boldsymbol{x}_{\mathrm{aff}})$ no es de rango completo, es un cono afín de $\mathbb{A}^{r+1}$ de dimensión a lo sumo $s-1$. Por lo tanto, la variedad proyectiva Σ_r^∞ tiene dimensión a lo sumo $s-2$. □

A continuación damos un resultado sobre la variedad $V(R_1^{d_1}, \ldots, R_m^{d_m}) \subset \mathbb{P}^{r-1}$ que nos permitirá obtener información sobre el comportamiento de $\mathrm{pcl}(V_r)$ en el hiperplano del infinito.

Lema 4.1.10. *$V(R_1^{d_1}, \ldots, R_m^{d_m}) \subset \mathbb{P}^{r-1}$ es absolutamente irreducible de dimensión $r-m-1$, grado a lo sumo $\prod_{i=1}^m d_i$ y lugar singular de dimensión a lo sumo $s-2$.*

Demostración. Por el Lema 4.1.8 tenemos que $R_i^{d_i} = S_i^{\mathsf{wt}}(\Pi_1, \ldots, \Pi_s)$ para $1 \leq i \leq m$. Dado que los polinomios $S_1^{\mathsf{wt}}, \ldots, S_m^{\mathsf{wt}}$ satisfacen las hipótesis $(\mathsf{H_1})$ y $(\mathsf{H_2})$, por el Lema 4.1.4, el Teorema 4.1.5 y el Corolario 4.1.7 tenemos que la $\mathbb{F}_q$–variedad afín de $\mathbb{A}^r$ definida por $R_1^{d_1}, \ldots, R_m^{d_m}$ es un cono de dimensión pura $r-m$, grado a lo sumo $\prod_{i=1}^m d_i$ y lugar singular de dimensión a lo sumo $s-1$. Por lo tanto, la variedad proyectiva $V(R_1^{d_1}, \ldots, R_m^{d_m}) \subset \mathbb{P}^{r-1}$ tiene dimensión pura $r-m-1$, grado a lo sumo $\prod_{i=1}^m d_i$ y lugar singular de dimensión a lo sumo $s-2$. En particular, $V(R_1^{d_1}, \ldots, R_m^{d_m}) \subset \mathbb{P}^{r-1}$ es una intersección completa conjuntista cuyo lugar singular tiene codimensión al menos $r-m-1-s+2 \geq 3$. El Teorema 2.1.10 asegura que ésta resulta absolutamente irreducible, lo que completa la demostración del lema. □

En el siguiente teorema damos una caracterización completa sobre el comportamiento de $\mathrm{pcl}(V_r)$ en el hiperplano del infinito.

Teorema 4.1.11. $\mathrm{pcl}(V_r) \cap \{X_0 = 0\} \subset \mathbb{P}^{r-1}$ *es una intersección completa de dimensión* $r-m-1$ *y grado* $\prod_{i=1}^{m} d_i$*, que es regular en codimensión* $r-m-s \geq 2$.

Demostración. Recordemos que $\mathrm{pcl}(V_r)$ tiene dimensión pura $r-m$. Así, por el Teorema 2.1.7, cada componente irreducible de $\mathrm{pcl}(V_r) \cap \{X_0 = 0\}$ tiene dimensión al menos $r-m-1$.

De (4.6) deducimos que $\mathrm{pcl}(V_r) \cap \{X_0 = 0\} \subset V(R_1^{d_1}, \ldots, R_m^{d_m})$. Por el Lema 4.1.10 tenemos que $V(R_1^{d_1}, \ldots, R_m^{d_m})$ es absolutamente irreducible de dimensión $r-m-1$. Concluimos que $\mathrm{pcl}(V_r) \cap \{X_0 = 0\}$ es también absolutamente irreducible de dimensión $r-m-1$, y por el Teorema 2.1.3 tenemos que

$$\mathrm{pcl}(V_r) \cap \{X_0 = 0\} = V(R_1^{d_1}, \ldots, R_m^{d_m}).$$

De acuerdo al Corolario 4.1.7, los polinomios $R_1^{d_1}, \ldots, R_m^{d_m}$ definen un ideal radical. Así concluimos que $V(R_1^{d_1}, \ldots, R_m^{d_m})$ es una intersección completa, y del Teorema 2.1.14 tenemos que

$$\deg(\mathrm{pcl}(V_r) \cap \{X_0 = 0\}) = \prod_{i=1}^{m} d_i.$$

Resta demostrar la afirmación sobre la regularidad de $\mathrm{pcl}(V_r) \cap \{X_0 = 0\}$. Del Lema 4.1.10 deducimos que el lugar singular de $\mathrm{pcl}(V_r) \cap \{X_0 = 0\}$ tiene dimensión a lo sumo $s-2$. Por lo tanto, $\mathrm{pcl}(V_r) \cap \{X_0 = 0\}$ es regular en codimensión $r-m-1-(s-2)-1 = r-m-s$. □

Concluimos esta sección con un teorema que recopila todas las propiedades de la clausura proyectiva $\mathrm{pcl}(V_r)$ que necesitamos.

Teorema 4.1.12. *La variedad proyectiva* $\mathrm{pcl}(V_r) \subset \mathbb{P}^r$ *es una intersección completa de dimensión* $r-m$ *y grado* $\prod_{i=1}^{m} d_i$*, que es regular en codimensión* $r-m-s \geq 2$.

Demostración. Observamos primeramente que $\mathrm{pcl}(V_r)$ es de dimensión pura $r-m$. Por un lado, tenemos que el conjunto de puntos singulares de $\mathrm{pcl}(V_r)$ que pertenecen al abierto $\{X_0 \neq 0\}$ está contenido en el lugar singular de $\mathrm{pcl}(V_r) \cap \{X_0 \neq 0\}$, y el Teorema 4.1.5 muestra que el lugar singular de $\mathrm{pcl}(V_r) \cap \{X_0 \neq 0\}$ tiene dimensión a lo sumo $s-1$. Por otro lado, por [GL02a, Lemma 1.1] tenemos que el lugar singular de $\mathrm{pcl}(V_r)$ en el hiperplano en el infinito está contenido en el lugar singular de $\mathrm{pcl}(V_r) \cap \{X_0 = 0\}$, y el Teorema 4.1.11 muestra que el lugar singular de $\mathrm{pcl}(V_r) \cap \{X_0 = 0\}$ tiene dimensión a lo sumo $s-2$. Por lo tanto, el lugar singular de $\mathrm{pcl}(V_r)$ tiene dimensión a lo sumo $s-1$ y $\mathrm{pcl}(V_r)$ es regular en codimensión $r-m-(s-1)-1 = r-m-s$.

Observemos que $\mathrm{pcl}(V_r)$ está contenido en la variedad proyectiva $V(R_1^h, \ldots, R_m^h)$. Además, tenemos las siguientes inclusiones:

$$V(R_1^h, \ldots, R_m^h) \cap \{X_0 \neq 0\} \subset V(R_1, \ldots, R_m),$$
$$V(R_1^h, \ldots, R_m^h) \cap \{X_0 = 0\} \subset V(R_1^{d_1}, \ldots, R_m^{d_m}).$$

El Lema 4.1.10 prueba que $V(R_1^{d_1},\dots,R_m^{d_m}) \subset \mathbb{P}^{r-1}$ es absolutamente irreducible de dimensión $r-m-1$, mientras que el Lema 4.1.4 muestra que $V(R_1,\dots,R_m) \subset \mathbb{A}^r$ es de dimensión pura $r-m$. Por lo tanto, $V(R_1^h,\dots,R_m^h) \subset \mathbb{P}^r$ tiene dimensión a lo sumo $r-m$. Teniendo en cuenta que dicha variedad está definida por m polinomios, deducimos que es una intersección completa conjuntista de dimensión $r-m$. Aplicando el Teorema 4.1.5 a $R_1,\dots,R_m$, y a $R_1^{d_1},\dots,R_m^{d_m}$, vemos que los polinomios $R_1^h,\dots,R_m^h$ definen un ideal radical y el lugar singular de $V(R_1^h,\dots,R_m^h)$ tiene codimensión $r-m-(s-1) \geq 3$. Del Teorema 2.1.10 deducimos que $V(R_1^h,\dots,R_m^h)$ es absolutamente irreducible.

Dado que $\mathrm{pcl}(V_r)$ está contenido en $V(R_1^h,\dots,R_m^h)$, ambas variedades proyectivas tienen dimensión $r-m$ y $V(R_1^h,\dots,R_m^h)$ es absolutamente irreducible, del Teorema 2.1.3 deducimos que

$$\mathrm{pcl}(V_r) = V(R_1^h,\dots,R_m^h). \tag{4.7}$$

Finalmente, dado que los polinomios $R_1^h,\dots,R_m^h$ definen un ideal radical, (4.7) y el Teorema de Bézout prueban que $\mathrm{pcl}(V_r)$ es una intersección completa de grado $\prod_{i=1}^m d_i$. Esto finaliza la demostración del teorema. □

4.1.3. El número de puntos $\mathbb{F}_q$–racionales

En esta sección damos una estimación del número de puntos $\mathbb{F}_q$–racionales de la variedad afín $V_r \subset \mathbb{A}^r$ definida por los polinomios simétricos $R_1,\dots,R_m$ de (4.1). Para esto, vamos a usar una estimación sobre el número de puntos $\mathbb{F}_q$–racionales de una intersección completa proyectiva definida sobre $\mathbb{F}_q$, regular en codimensión 2, del Teorema 2.2.11.

Consideramos la clausura proyectiva $\mathrm{pcl}(V_r)$ de V_r y sea $V_{r,\infty} := \mathrm{pcl}(V_r) \cap \{X_0 = 0\}$. Combinando los Teoremas 4.1.11 y 4.1.12 con la estimación (2.8) del Teorema 2.2.11, obtenemos que

$$\begin{aligned} \big||\mathrm{pcl}(V_r)(\mathbb{F}_q)| - p_{r-m}\big| &\leq 14D^3\delta^2 q^{r-m-1}, \\ \big||V_{r,\infty}(\mathbb{F}_q)| - p_{r-m-1}\big| &\leq 14D^3\delta^2 q^{r-m-2}, \end{aligned}$$

donde $D := \sum_{i=1}^m (d_i-1)$ y $\delta := \prod_{i=1}^m d_i$. En consecuencia,

$$\begin{aligned} \big||V_r(\mathbb{F}_q)| - q^{r-m}\big| &= \big||\mathrm{pcl}(V_r)(\mathbb{F}_q)| - |V_{r,\infty}(\mathbb{F}_q)| - p_{r-m} + p_{r-m-1}\big| \\ &\leq \big||\mathrm{pcl}(V_r)(\mathbb{F}_q)| - p_{r-m}\big| + \big||V_{r,\infty}(\mathbb{F}_q)| - p_{r-m-1}\big| \\ &\leq 14D^3\delta^2(q+1)q^{r-m-2}. \end{aligned}$$

Por lo tanto tenemos el siguiente resultado.

Teorema 4.1.13. *Sean s, r, m enteros positivos tales que $m \leq s \leq r-m-2$. Sean $R_1,\dots,R_m \in \mathbb{F}_q[X_1,\dots,X_r]$ los polinomios definidos como $R_i := S_i(\Pi_1,\dots,\Pi_s)$ para $1 \leq i \leq m$, donde $S_1,\dots,S_m \in \mathbb{F}_q[Y_1,\dots,Y_s]$ satisfacen las hipótesis (H_1), (H_2) y (H_3). Denotamos por $d_i := \deg R_i$ para $1 \leq i \leq m$, $D := \sum_{i=1}^m (d_i-1)$ y $\delta := \prod_{i=1}^m d_i$. Si $V_r := V(R_1,\dots,R_m) \subset \mathbb{A}^r$, entonces*

$$\big||V_r(\mathbb{F}_q)| - q^{r-m}\big| \leq 14D^3\delta^2(q+1)q^{r-m-2}.$$

En los capítulos siguientes necesitaremos no solo estimaciones del número de puntos $\mathbb{F}_q$–racionales de una intersección completa dada, sino también del número de puntos $\mathbb{F}_q$–racionales con todas las coordenadas distintas dos a dos. En el siguiente teorema damos una cota superior de la cantidad de puntos $\mathbb{F}_q$–racionales de V_r tales que dos coordenadas toman el mismo valor y, finalmente, damos un corolario de la cantidad de puntos $\mathbb{F}_q$–racionales de V_r con coordenadas distintas dos a dos.

Teorema 4.1.14. *Con las notaciones y las hipótesis del Teorema 4.1.13, dados i y j con $1 \leq i < j \leq r$, tenemos que $V_r \cap \{X_i = X_j\}$ es de dimensión pura $r-m-1$. En particular, vale la siguiente estimación:*

$$|V_r(\mathbb{F}_q) \cap \{X_i = X_j\}| \leq \delta q^{r-m-1}.$$

Demostración. El Teorema 4.1.12 prueba que $\mathrm{pcl}(V_r)$ es una intersección completa regular en codimensión $r-m-s \geq 2$. Por lo tanto, por el Teorema 2.1.10 concluimos que es absolutamente irreducible. Esto implica que V_r también es absolutamente irreducible.

Sin pérdida de generalidad suponemos que $i = r-1$ y $j = r$. Podemos considerar a $V_{r-1,r} := V_r \cap \{X_{r-1} = X_r\}$ como la subvariedad de $\mathbb{A}^{r-1}$ definida por los polinomios $R_i := S_i(\Pi_1^*, \ldots, \Pi_s^*) \in \mathbb{F}_q[X_1, \ldots, X_{r-1}]$ para $1 \leq i \leq m$, donde $\Pi_i^* := \Pi_i(X_1, \ldots, X_{r-1}, X_{r-1})$ es el polinomio que se obtiene al sustituir X_r por X_{r-1} en el i–ésimo polinomio simétrico elemental Π_i de $\mathbb{F}_q[X_1, \ldots, X_r]$. Observemos que

$$\Pi_i^* = \Pi_i^{r-2} + 2X_{r-1} \cdot \Pi_{i-1}^{r-2} + X_{r-1}^2 \cdot \Pi_{i-2}^{r-2} \tag{4.8}$$

donde Π_j^{r-2} es el j–ésimo polinomio simétrico elemental de $\mathbb{F}_q[X_1, \ldots, X_{r-2}]$ para $1 \leq j \leq s$.

Sea $\boldsymbol{\Pi}^* := (\Pi_1^*, \ldots, \Pi_s^*)$ y denotemos por $(\partial \boldsymbol{\Pi}^* / \partial \boldsymbol{X}^*)$ la matriz Jacobiana de $\boldsymbol{\Pi}^*$ con respecto a $X_1, \ldots, X_{r-1}$. Observemos también que el conjunto de puntos $\boldsymbol{x}$ de $V_{r-1,r}$ para los cuales la matriz Jacobiana $\big(\partial(\boldsymbol{R} \circ \boldsymbol{\Pi}^*)/\partial \boldsymbol{X}^*\big)(\boldsymbol{x})$ no es de rango completo, tiene dimensión a lo sumo s. En efecto, de (4.8) concluimos que el menor no nulo de tamaño $s \times s$ de la matriz Jacobiana $(\partial \boldsymbol{\Pi}^* / \partial \boldsymbol{X}^*)$ que determina cualquier elección $i_1, \ldots, i_s$ de columnas con $1 \leq i_1 < i_2 < \cdots < i_s \leq r-2$ coincide con el correspondiente menor no nulo de $(\partial \Pi_i^{r-2} / \partial X_j)_{1 \leq i \leq s, 1 \leq j \leq r-1}$. En particular, tal menor maximal no nulo de $(\partial \Pi_i^* / \partial X_j)_{1 \leq i \leq s, 1 \leq j \leq r-2}$ es, salvo signo, un determinante de Vandermonde que depende de s de las indeterminadas $X_1, \ldots, X_{r-2}$. Argumentando como en la demostración del Teorema 4.1.5 deducimos que el conjunto de puntos $\boldsymbol{x}$ de $V_{r-1,r}$ para los cuales la matriz Jacobiana $\big(\partial(\boldsymbol{R} \circ \boldsymbol{\Pi}^*)/\partial \boldsymbol{X}^*\big)(\boldsymbol{x})$ no es de rango completo, está incluido en una unión de variedades lineales de dimensión s (ver la Observación 4.1.6), y así tiene dimensión a lo sumo s.

Sea $\mathcal{C}$ una componente irreducible de $V_{r-1,r}$. Entonces, por el Teorema 2.1.7, $\mathcal{C}$ tiene dimensión al menos $r-m-1$. Como $r-m-1-s \geq 1$, se tiene que para un punto genérico $\boldsymbol{x} \in \mathcal{C}$ la matriz Jacobiana $\big(\partial(\boldsymbol{R} \circ \boldsymbol{\Pi}^*)/\partial \boldsymbol{X}^*\big)(\boldsymbol{x})$ tiene rango m. Concluimos que el espacio tangente de $V_{r-1,r}$ en $\boldsymbol{x}$ tiene dimensión a lo sumo $r-m-1$, lo cual implica que $\mathcal{C}$ tiene dimensión $r-m-1$. Concluimos que $V_{r-1,r}$ es de dimensión pura $r-m-1$, finalizando así la demostración de la primera afirmación del teorema.

Por otro lado, de la desigualdad de Bézout (2.1.14) se sigue que $\deg V_r \cap \{X_i = X_j\} \leq \deg V_r$. La segunda afirmación se deduce inmediatamente del Teorema 2.2.1. □

Sea $\mathcal{I}$ un subconjunto de $\{(i,j) : 1 \leq i < j \leq r\}$ y sea $V_r^{=} \subset \mathbb{A}^r$ la variedad definida como

$$V_r^{=} := \bigcup_{(i,j)\in\mathcal{I}} V_r \cap \{X_i = X_j\}.$$

Asimismo, denotamos por $V_r^{\neq} := V_r \setminus V_r^{=}$. Obtenemos el siguiente resultado.

Corolario 4.1.15. *Con las notaciones y las hipótesis del Teorema 4.1.13,*

$$\left||V_r^{\neq}(\mathbb{F}_q)| - q^{r-m}\right| \leq 14D^3\delta^2(q+1)q^{r-m-2} + |\mathcal{I}|\delta q^{r-m-1}.$$

Demostración. Del Teorema 4.1.14, tenemos que

$$|V_r^{=}(\mathbb{F}_q)| \leq \sum_{(i,j)\in\mathcal{I}} \delta q^{r-m-1} = |\mathcal{I}|\delta q^{r-m-1}.$$

Por lo tanto, del Teorema 4.1.13 deducimos que

$$\begin{aligned}\left||V_r^{\neq}(\mathbb{F}_q)| - q^{r-m}\right| &\leq \left||V_r(\mathbb{F}_q)| - q^{r-m}\right| + |V_r^{=}(\mathbb{F}_q)| \\ &\leq 14D^3\delta^2(q+1)q^{r-m-2} + |\mathcal{I}|\delta q^{r-m-1}.\end{aligned}$$

Esto finaliza la demostración del corolario. □

4.2. Estimaciones para ciertas intersecciones completas asociadas a familias lineales

En esta sección estudiamos el conjunto de puntos $\mathbb{F}_q$–racionales de ciertas variedades definidas por polinomios simétricos asociadas a familias lineales de polinomios univariados. Probamos que dichas variedades resultan intersecciones completas simétricas normales. Esto nos permite aplicar la estimación para intersecciones completas normales proyectivas del Teorema 2.2.10 y proporcionar, así, una estimación de la cantidad de puntos $\mathbb{F}_q$–racionales de dichas intersecciones completas. Estas intersecciones completas normales aparecerán en el estudio de los patrones de factorización en familias lineales sobre cuerpos finitos.

4.2.1. Aspectos geométricos

Sean d, s y m enteros tales que $m \leq s \leq d-3$. Sean $Z_1, \ldots, Z_d$ indeterminadas sobre $\overline{\mathbb{F}}_q$. Sea $\boldsymbol{Z} := (Z_1, \ldots, Z_s)$ y sean $S_1, \ldots, S_m \in \mathbb{F}_q[\boldsymbol{Z}]$ los polinomios definidos de la siguiente manera:

$$S_k := \sum_{j=1}^{s} (-1)^j b_{k,d-j} Z_j + b_{k,0} \ \ (1 \leq k \leq m). \tag{4.9}$$

Observemos que $S_1, \ldots, S_m$ tienen grado 1. Podemos suponer sin pérdida de generalidad que las componentes homogéneas de grado 1 son linealmente independientes en $\mathbb{F}_q[\boldsymbol{Z}]$. Así, la matriz Jacobiana $(\partial \boldsymbol{S}/\partial \boldsymbol{Z})(\boldsymbol{z})$ de $\boldsymbol{S} := (S_1, \ldots, S_m)$ con respecto a $\boldsymbol{Z} := (Z_1, \ldots, Z_s)$ tiene rango m para todo $\boldsymbol{z} \in \mathbb{A}^s$.

Sean $Y_1, \ldots, Y_d$ indeterminadas sobre $\overline{\mathbb{F}}_q$ e $\boldsymbol{Y} := (Y_1, \ldots, Y_d)$. Sean $\Pi_1, \ldots, \Pi_s$ los primeros s polinomios simétricos elementales de $\mathbb{F}_q[\boldsymbol{Y}]$ y $R_1, \ldots, R_m \in \mathbb{F}_q[\boldsymbol{Y}]$ los polinomios definidos como:

$$R_k = S_k(\Pi_1, \ldots, \Pi_s) \ \ (1 \leq k \leq m). \tag{4.10}$$

Sea $V_d \subset \mathbb{A}^d$ la $\mathbb{F}_q$–variedad afín definida por $R_1, \ldots, R_m \in \mathbb{F}_q[\boldsymbol{Y}]$. En esta sección vamos a estudiar la geometría de dicha variedad. El principal resultado que damos es que V_d es regular en codimensión 1, de lo que concluimos que V_d es una intersección completa normal.

Consideremos $S_1, \ldots, S_m$ como elementos de $\mathbb{F}_q[Z_1, \ldots, Z_d]$. Dado que la matriz Jacobiana $(\partial \boldsymbol{S}/\partial \boldsymbol{Z})(\boldsymbol{z})$ es de rango completo para todo $\boldsymbol{z} \in \mathbb{A}^s$, la variedad lineal $W_d \subset \mathbb{A}^d$ definida por $S_1, \ldots, S_m$ tiene dimensión $d - m$. Consideramos el siguiente morfismo sobreyectivo:

$$\begin{aligned} \boldsymbol{\Pi^d} : \mathbb{A}^d &\to \mathbb{A}^d \\ \boldsymbol{y} &\mapsto (\Pi_1(\boldsymbol{y}), \ldots, \Pi_d(\boldsymbol{y})). \end{aligned}$$

De la misma manera que en la sección anterior, es fácil demostrar que $\boldsymbol{\Pi^d}$ es un morfismo finito.

Observemos que la variedad lineal $W_d^j := V(S_1, \ldots, S_j) \subset \mathbb{A}^d$ es irreducible de dimensión $d - j$. Del Teorema 2.1.4 deducimos que la variedad $(\boldsymbol{\Pi^d})^{-1}(W_d^j) = V(R_1, \ldots, R_j) \subset \mathbb{A}^d$ es de dimensión pura $d-j$. En consecuencia, $R_1, \ldots, R_m$ forman una sucesión regular de $\mathbb{F}_q[\boldsymbol{Y}]$, de donde deducimos el siguiente resultado.

Lema 4.2.1. *Sea $V_d \subset \mathbb{A}^d$ la variedad definida por $R_1, \ldots, R_m$. Entonces V_d es una intersección completa conjuntista de dimensión $d - m$.*

A continuación analizamos la dimensión del lugar singular de V_d. Suponemos sin pérdida de generalidad que $(\partial \boldsymbol{S}/\partial \boldsymbol{Z})(\boldsymbol{z})$ es una matriz escalonada por columnas. Sean $1 \leq i_1 < \cdots < i_m \leq s$ los índices correspondientes a los pivotes. Sea $\mathcal{I} := \{i_1, \ldots, i_m\}$ y $\mathcal{J} := \{j_1, \ldots, j_{s-m}\} := \{1, \ldots, s\} \setminus \mathcal{I}$. Entonces la matriz Jacobiana

$$\mathcal{M} := \big(\partial(S_1, \ldots, S_m, Z_{j_1}, \ldots, Z_{j_{s-m}})/\partial \boldsymbol{Z}\big) \tag{4.11}$$

es inversible. Sean $B_0, \ldots, B_{d-m-1}$ nuevas indeterminadas sobre $\overline{\mathbb{F}}_q$ y definamos $S_{m+k} := Z_{j_k} + B_{d-m-k}$ $(1 \leq k \leq s - m)$ y $S_k := Z_k + B_{d-k}$ $(s + 1 \leq k \leq d)$. Sean $\boldsymbol{B} := (B_{d-m-1}, \ldots, B_0)$, $\boldsymbol{S}^e := (S_1, \ldots, S_d)$ y $\boldsymbol{Z}^e := (Z_1, \ldots, Z_d)$. Observemos que la siguiente matriz Jacobiana de tamaño $d \times d$:

$$\big(\partial \boldsymbol{S}^e/\partial \boldsymbol{Z}^e\big) := \big(\partial(S_1, \ldots, S_m, Z_{j_1}, \ldots, Z_{j_{s-m}}, Z_{s+1}, \ldots, Z_d)/\partial \boldsymbol{Z}^e\big),$$

también es inversible. Consideremos el morfismo sobreyectivo

$$\begin{aligned} \boldsymbol{\Pi} : \mathbb{A}^d &\to \mathbb{A}^s \\ \boldsymbol{y} &\mapsto (\Pi_1(\boldsymbol{y}), \ldots, \Pi_s(\boldsymbol{y})). \end{aligned}$$

Finalmente, introducimos la variedad $V_d^e \subset \mathbb{A}^{2d-m}$ definida de la siguiente manera:

$$V_d^e := \{(\boldsymbol{y}, \boldsymbol{b}) \in \mathbb{A}^d \times \mathbb{A}^{d-m} : S_j(\boldsymbol{\Pi^d}(\boldsymbol{y}), \boldsymbol{b}) = 0 \ (1 \leq j \leq d)\}.$$

Vamos a establecer una relación entre las variedades V_d y V_d^e. Si $(\boldsymbol{y}, \boldsymbol{b})$ es un punto de V_d^e, entonces $S_j(\boldsymbol{\Pi^d}(\boldsymbol{y}), \boldsymbol{b}) = S_j(\boldsymbol{\Pi}(\boldsymbol{y})) = 0$ para $1 \leq j \leq m$, lo cual implica que $\boldsymbol{y} \in V_d$. Esto muestra que el siguiente morfismo regular de variedades afines está bien definido:

$$\begin{aligned} \Phi_1^e : V_d^e &\to V_d \\ (\boldsymbol{y}, \boldsymbol{b}) &\mapsto \boldsymbol{y}. \end{aligned}$$

Más aún, por la definición de V_d^e es fácil ver que Φ_1^e resulta ser un isomorfismo de variedades afines, cuya inversa es el siguiente morfismo:

$$\begin{aligned} \Psi^e : V_d &\to V_d^e \\ \boldsymbol{y} &\mapsto \big(\boldsymbol{y}, -\Pi_{j_1}(\boldsymbol{y}), \ldots, -\Pi_{j_{s-m}}(\boldsymbol{y}), -\Pi_{s+1}(\boldsymbol{y}), \ldots, -\Pi_d(\boldsymbol{y})\big). \end{aligned}$$

Concluimos así que $V_d^e \subset \mathbb{A}^{2d-m}$ es una variedad afín de dimensión pura $d-m$. Nuestro objetivo es mostrar que el lugar singular Σ de V_d tiene codimensión al menos 2 en V_d. Para este próposito, vamos a mostrar que el lugar singular Σ^e de V_d^e tiene codimensión al menos 2 en V_d^e.

Definimos $R_{m+k} := S_{m+k}(\boldsymbol{\Pi^d}, \boldsymbol{B})$ para $1 \leq k \leq d-m$. Denotamos con $(\partial \boldsymbol{R}/\partial \boldsymbol{Y})$ la matriz Jacobiana de $\boldsymbol{R} := (R_1, \ldots, R_m)$ con respecto a $\boldsymbol{Y}$ y con $(\partial \boldsymbol{R}^e/\partial(\boldsymbol{Y}, \boldsymbol{B}))$ la matriz Jacobiana de $\boldsymbol{R}^e := (R_1, \ldots, R_d)$ con respecto a $\boldsymbol{Y}$ y $\boldsymbol{B}$. La siguiente observación muestra la relación entre el lugar singular de V_d y el de V_d^e.

Observación 4.2.2. *Para $\boldsymbol{y} \in V_d$, sea $(\boldsymbol{y}, \boldsymbol{b}) := \Psi^e(\boldsymbol{y})$. Entonces $(\partial \boldsymbol{R}/\partial \boldsymbol{Y})(\boldsymbol{y})$ tiene rango máximo m si y solo si $\big(\partial \boldsymbol{R}^e/\partial(\boldsymbol{Y}, \boldsymbol{B})\big)(\boldsymbol{y}, \boldsymbol{b})$ tiene rango máximo d.*

Demostración. Sea $\boldsymbol{y} \in V_d$. Observemos que de la definición de $\boldsymbol{R}^e$ deducimos que $\big(\partial \boldsymbol{R}^e/\partial(\boldsymbol{Y}, \boldsymbol{B})\big)(\boldsymbol{y}, \boldsymbol{b})$ tiene la siguiente estructura por bloques:

$$\frac{\partial \boldsymbol{R}^e}{\partial(\boldsymbol{Y}, \boldsymbol{B})}(\boldsymbol{y}, \boldsymbol{b}) = \begin{pmatrix} \frac{\partial \boldsymbol{R}}{\partial \boldsymbol{Y}}(\boldsymbol{y}) & \boldsymbol{0} \\ \frac{\partial (\boldsymbol{R}^e \setminus \boldsymbol{R})}{\partial \boldsymbol{Y}}(\boldsymbol{y}, \boldsymbol{b}) & \frac{\partial (\boldsymbol{R}^e \setminus \boldsymbol{R})}{\partial \boldsymbol{B}}(\boldsymbol{y}, \boldsymbol{b}) \end{pmatrix} = \begin{pmatrix} \frac{\partial \boldsymbol{R}}{\partial \boldsymbol{Y}}(\boldsymbol{y}) & \boldsymbol{0} \\ \frac{\partial (\boldsymbol{R}^e \setminus \boldsymbol{R})}{\partial \boldsymbol{Y}}(\boldsymbol{y}, \boldsymbol{b}) & \boldsymbol{I} \end{pmatrix},$$

donde $\boldsymbol{0}$ denota la matriz nula de tamaño $m \times (d-m)$ e $\boldsymbol{I}$ denota la matriz identidad de tamaño $(d-m) \times (d-m)$. Se sigue fácilmente la conclusión de la observación. □

Con el objetivo de obtener una cota superior de la dimensión del lugar singular de V_d^e, consideramos la siguiente proyección:

$$\begin{aligned} \Phi_2^e : V_d^e &\to \mathbb{A}^{d-m} \\ (\boldsymbol{y}, \boldsymbol{b}) &\mapsto \boldsymbol{b}, \end{aligned}$$

y analizamos $\Phi_2^e(\Sigma^e)$, donde Σ^e es el lugar singular de V_d^e. Comenzamos con el siguiente lema.

Lema 4.2.3. *Sea $b_{i,0} = S_i(\mathbf{0})$ para $1 \leq i \leq m$ y sea $\boldsymbol{b}_0 := (b_{1,0}, \ldots, b_{m,0})$. Denotamos con $\boldsymbol{B}_J := (B_{d-m-1}, \ldots, B_{d-s})$ y sea $\boldsymbol{m}_j$ la j–ésima fila de $\mathcal{M}^{-1}$ para $1 \leq j \leq s$, donde $\mathcal{M}$ es la matriz definida en (4.11). Entonces el polinomio*

$$P_j := Y_j^d - \sum_{k=1}^{s}(-1)^k \boldsymbol{m}_k \cdot (\boldsymbol{b}_0, \boldsymbol{B}_J)^t Y_j^{d-k} - \sum_{k=s+1}^{d}(-1)^k B_{d-k} Y_j^{d-k} \tag{4.12}$$

se anula en V_d^e para $1 \leq j \leq d$.

Demostración. De la definición de la matriz $\mathcal{M}$, deducimos que

$$(R_1, \ldots, R_s)^t = \mathcal{M} \cdot \boldsymbol{\Pi}^t + (\boldsymbol{b}_0, \boldsymbol{B}_J)^t, \tag{4.13}$$

donde $\boldsymbol{\Pi} := (\Pi_1, \ldots, \Pi_s)$. Como $\mathcal{M}$ es inversible, $\boldsymbol{\Pi}(\boldsymbol{y})^t + \mathcal{M}^{-1}(\boldsymbol{b}_0, \boldsymbol{b}_J)^t = \mathbf{0}$ para todo $(\boldsymbol{y}, \boldsymbol{b}) \in V^e$. Por lo tanto,

$$\Pi_k(\boldsymbol{y}) + \boldsymbol{m}_k \cdot (\boldsymbol{b}_0, \boldsymbol{b}_J)^t = 0 \quad (1 \leq k \leq s)$$

para todo $(\boldsymbol{y}, \boldsymbol{b}) \in V_d^e$.

Fijamos j con $1 \leq j \leq d$. Por la igualdad $\prod_{j=1}^{d}(T - y_j) = T^d + \sum_{k=1}^{d}(-1)^k \Pi_k(y) T^{d-k}$, se tiene que

$$(y_j)^d + \sum_{k=1}^{d}(-1)^k \, \Pi_k(\boldsymbol{y}) \, (y_j)^{d-k} = 0 \quad (1 \leq j \leq d). \tag{4.14}$$

Combinando (4.13) y (4.14) y la definición de $\boldsymbol{S}^e$, deducimos fácilmente que el polinomio P_j de (4.12) se anula en V_d^e. □

El siguiente resultado muestra que Φ_2^e es un morfismo finito. Este resultado nos permitirá estudiar el lugar singular Σ^e de V_d^e a partir de información sobre $\Phi_2^e(\Sigma^e)$.

Lema 4.2.4. *Φ_2^e es un morfismo finito.*

Demostración. Tenemos que demostrar que Φ_2^e es dominante y que la extensión $\overline{\mathbb{F}}_q[\boldsymbol{B}] \hookrightarrow \overline{\mathbb{F}}_q[V_d^e]$ es entera.

Empezamos con la primera afirmación. Sea $\boldsymbol{b} \in \mathbb{A}^{d-m}$ un punto de la imagen de Φ_2^e y sea $\boldsymbol{y} \in V_d$ un punto tal que $(\boldsymbol{y}, \boldsymbol{b}) \in V_d^e$. De la definición de P_j deducimos que el polinomio $P_j(Y_j, \boldsymbol{b})$ tiene a lo sumo d raíces en $\overline{\mathbb{F}}_q$. El Lema 4.2.3 muestra que $P_j(y_j, \boldsymbol{b}) = 0$ para cada $\boldsymbol{y} \in V_d$, con $1 \leq j \leq d$. Se deduce que, para cada j con $1 \leq j \leq d$, existen finitas posibles elecciones para la j–ésima coordenada de un punto $\boldsymbol{y} \in V_d$ con $(\boldsymbol{y}, \boldsymbol{b}) \in V_d^e$. En otras palabras, la fibra $(\Phi_2^e)^{-1}(\boldsymbol{b})$ tiene finitos puntos, es decir, es de dimensión cero. El Teorema de la dimensión de la fibra (ver Teorema 2.1.5) asegura que $\dim V_d^e - \dim \Phi_2^e(V_d^e) \leq \dim(\Phi_2^e)^{-1}(\boldsymbol{b}) = 0$, es decir, $\dim \Phi_2^e(V_d^e) \geq d - m$. Por otro lado, del Teorema 2.1.6 tenemos que $\dim \Phi_2^e(V_d^e) \leq \dim V_d^e = d - m$. Deducimos así que Φ_2^e es dominante.

Veamos ahora la segunda afirmación del lema. Dado que $P_j(Y_j, \boldsymbol{B}) = 0$ en $\overline{\mathbb{F}}_q[V_d^e]$ para $1 \leq j \leq d$, la extensión $\overline{\mathbb{F}}_q[\boldsymbol{B}] \hookrightarrow \overline{\mathbb{F}}_q[V_d^e]$ es entera. Esto implica que Φ_2^e es un morfismo finito y finaliza así la demostración del lema. □

La siguiente proposición da una caracterización parcial del lugar singular de V_d^e. Para ello usamos el hecho de que la variedad V_d está definida por polinomios simétricos. La demostración sigue parte de las ideas de la demostración del Teorema 4.1.5.

Proposición 4.2.5. *Sea $(\boldsymbol{y}, \boldsymbol{b}) \in V_d^e$ un punto tal que $(\partial \boldsymbol{R}^e/\partial(\boldsymbol{Y}, \boldsymbol{B}))(\boldsymbol{y}, \boldsymbol{b})$ no es de rango completo. Entonces existen $i, j, k, l \in \{1, \ldots, d\}$ con $i < j$, $k < l$ y $\{i, j\} \cap \{k, l\} = \emptyset$ tales que $y_i = y_j$ y $y_k = y_l$.*

Demostración. Sea $(\boldsymbol{y}, \boldsymbol{b}) \in V_d^e$ un punto en las hipótesis de la proposición. Por la Observación 4.2.2, la matriz Jacobiana $(\partial \boldsymbol{R}/\partial \boldsymbol{Y})(\boldsymbol{y})$ no es de rango completo. Como $\boldsymbol{R} = \boldsymbol{S} \circ \boldsymbol{\Pi}$, por la regla de la cadena vemos que

$$\left(\frac{\partial \boldsymbol{R}}{\partial \boldsymbol{Y}}\right) = \left(\frac{\partial \boldsymbol{S}}{\partial \boldsymbol{Z}} \circ \boldsymbol{\Pi}\right) \cdot \left(\frac{\partial \boldsymbol{\Pi}}{\partial \boldsymbol{Y}}\right).$$

Sea $\boldsymbol{v} \in \mathbb{A}^m$ un vector no nulo en el núcleo a izquierda de $(\partial \boldsymbol{R}/\partial \boldsymbol{Y})(\boldsymbol{y})$. Entonces

$$\boldsymbol{0} = \boldsymbol{v} \cdot \left(\frac{\partial \boldsymbol{R}}{\partial \boldsymbol{Y}}\right)(\boldsymbol{y}) = \boldsymbol{v} \cdot \left(\frac{\partial \boldsymbol{S}}{\partial \boldsymbol{Z}}\right)\big(\boldsymbol{\Pi}(\boldsymbol{y})\big) \cdot \left(\frac{\partial \boldsymbol{\Pi}}{\partial \boldsymbol{Y}}\right)(\boldsymbol{y}).$$

Como la matriz Jacobiana $(\partial \boldsymbol{S}/\partial \boldsymbol{Z})\big(\boldsymbol{\Pi}(\boldsymbol{y})\big)$ es de rango completo, el vector $\boldsymbol{w} := \boldsymbol{v} \cdot (\partial \boldsymbol{S}/\partial \boldsymbol{Z})\big(\boldsymbol{\Pi}(\boldsymbol{y})\big) \in \mathbb{A}^s$ es no nulo, y satisface la identidad

$$\boldsymbol{w} \cdot \left(\frac{\partial \boldsymbol{\Pi}}{\partial \boldsymbol{Y}}\right)(\boldsymbol{y}) = \boldsymbol{0}.$$

Concluimos que todos los menores maximales de $(\partial \boldsymbol{\Pi}/\partial \boldsymbol{Y})(\boldsymbol{y})$ son nulos.

Como ya dijimos en la sección anterior, las derivadas parciales de los polinomios simétricos elementales Π_i sastifacen las siguientes igualdades para $1 \leq i \leq s$ y $1 \leq j \leq d$:

$$\frac{\partial \Pi_i}{\partial Y_j} = \Pi_{i-1} - Y_j \Pi_{i-2} + Y_j^2 \Pi_{i-3} + \cdots + (-1)^{i-1} Y_j^{i-1}.$$

En consecuencia, si $\mathcal{V}$ es la matriz de Vandermonde definida por $\mathcal{V} := (Y_j^{i-1})_{1 \leq i \leq s,\, 1 \leq j \leq d}$, entonces la matriz Jacobiana $(\partial \boldsymbol{\Pi}/\partial \boldsymbol{Y})$ puede factorizarse de la siguiente manera:

$$\left(\frac{\partial \boldsymbol{\Pi}}{\partial \boldsymbol{Y}}\right) = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ \Pi_1 & -1 & 0 & & \\ \Pi_2 & -\Pi_1 & 1 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 0 \\ \Pi_{s-1} & -\Pi_{s-2} & \Pi_{s-3} & \cdots & (-1)^{s-1} \end{pmatrix} \cdot \mathcal{V}.$$

Dado que todos los menores maximales de $(\partial \boldsymbol{\Pi}/\partial \boldsymbol{Y})(\boldsymbol{y})$ son nulos, todos los menores maximales de $\mathcal{V}(\boldsymbol{y})$ son también nulos.

Fijamos $1 \leq k_1 < \cdots < k_s \leq d$, sea $K := (k_1, \ldots, k_s)$ y sea $\mathcal{V}_K(\boldsymbol{y})$ la matriz de tamaño $s \times s$ formada por las columnas $k_1, \ldots, k_s$ de la matriz $\mathcal{V}(\boldsymbol{y})$, es decir, $\mathcal{V}_K(\boldsymbol{y}) := (y_{k_j}^{i-1})_{1 \leq i,j \leq s}$. Tenemos que

$$\det\big(\mathcal{V}_K(\boldsymbol{y})\big) = \prod_{1\le j<j'\le s}(y_{k_j} - y_{k_{j'}}) = 0.$$

Como esta identidad vale para todo $K := (k_1, \ldots, k_s)$ como arriba, concluimos que $\boldsymbol{y}$ tiene a lo sumo $s - 1 \le d - 4$ coordenadas distintas dos a dos. En particular, existen $1 \le i < j \le d - 2$ con $y_i = y_j$. Supongamos sin pérdida de generalidad que $i = 1$ y $j = 2$. Entonces existen $3 \le k < l \le d$ con $y_k = y_l$. Esto finaliza la demostración de la proposición. □

En la siguiente proposición damos una cota superior de la dimensión del lugar singular de V_d^e. Para ello, observamos que, dado $(\boldsymbol{y}, \boldsymbol{b}) \in V_d^e$, si fijamos $\boldsymbol{b} \in \mathbb{A}^{d-m}$, las raíces del polinomio $P_{\boldsymbol{b}} := P_j(Y_j, \boldsymbol{b})$ son las coordenadas del punto $\boldsymbol{y} \in V_d$. La Proposición 4.2.5 implica que $P_{\boldsymbol{b}}$ tiene dos raíces dobles o una triple, hecho que nos permitirá controlar la dimensión del lugar singular de V_d^e.

Proposición 4.2.6. *Sea $p > 2$. El conjunto $\mathcal{D}$ de puntos $(\boldsymbol{y}, \boldsymbol{b}) \in V_d^e$ tales que la matriz Jacobiana $(\partial \boldsymbol{R}^e/\partial(\boldsymbol{Y}, \boldsymbol{B}))(\boldsymbol{y}, \boldsymbol{b})$ no es de rango completo, tiene codimensión al menos 2 en V_d^e. En particular, el lugar singular de V_d^e tiene codimensión al menos 2 en V_d^e.*

Demostración. Por el Lema 4.2.3 tenemos que el polinomio

$$P_j := Y_j^d - \sum_{k=1}^{s}(-1)^k \boldsymbol{m}_k \cdot (\boldsymbol{b}_0, \boldsymbol{B}_J)^t Y_j^{d-k} - \sum_{k=s+1}^{s}(-1)^k B_{d-k} Y_j^{d-k}$$

se anula en V_d^e para $1 \le j \le d$. Sea $(\boldsymbol{y}, \boldsymbol{b})$ un punto de $\mathcal{D}$ y sea $P_{\boldsymbol{b}} \in \overline{\mathbb{F}}_q[T]$ el polinomio definido como

$$\begin{aligned} P_{\boldsymbol{b}} &:= T^d - \sum_{k=1}^{s}(-1)^k \boldsymbol{m}_k \cdot (\boldsymbol{b}_0, \boldsymbol{b}_J)^t T^{d-k} - \sum_{k=s+1}^{d}(-1)^k b_{d-k} T^{d-k} \\ &= T^d + \sum_{k=1}^{d}(-1)^k \Pi_k(\boldsymbol{y})\, T^{d-k} = \prod_{j=1}^{d}(T - y_j). \end{aligned}$$

Como las raíces de $P_{\boldsymbol{b}}$ en $\overline{\mathbb{F}}_q$ son las coordenadas del punto $\boldsymbol{y} \in V_d$, por la Proposición 4.2.5 tenemos que $P_{\boldsymbol{b}}$, o bien tiene dos raíces múltiples distintas, o bien tiene una raíz múltiple de multiplicidad al menos 3.

Observemos que si $P'_{\boldsymbol{b}}$ es nulo, entonces el Lema 3.2.5 asegura que el conjunto $\mathcal{C}_0$ de todos los elementos $\boldsymbol{b} \in \mathbb{A}^{d-m}$ tales que $P'_{\boldsymbol{b}} = 0$ está contenido en una subvariedad de codimensión 2 de $\mathbb{A}^{d-m}$. Por otro lado, por el Lema 3.2.7 deducimos que el conjunto $\mathcal{C}_1$ de todos los elementos $\boldsymbol{b} \in \mathbb{A}^{d-m}$ tales que $P_{\boldsymbol{b}}$ tiene dos raíces múltiples distintas está contenido en una subvariedad de codimensión 2 de $\mathbb{A}^{d-m}$. Finalmente, el Lema 3.2.9 implica que el conjunto $\mathcal{C}_2$ de todos los elementos $\boldsymbol{b} \in \mathbb{A}^{d-m}$ tales que $P_{\boldsymbol{b}}$ tiene una raíz de multiplicidad al menos tres está contenido en una subvariedad de codimensión 2 de $\mathbb{A}^{d-m}$. En consecuencia, dado que la imagen de $\mathcal{D}$ por Φ_2^e está contenida en $\mathcal{C}_0 \cup \mathcal{C}_1 \cup \mathcal{C}_2$, y cada $\mathcal{C}_i$ está contenida en una subvariedad de codimensión

2 de $\mathbb{A}^{n-m}$, deducimos que $\Phi_2^e(\mathcal{D})$ está contenido en una subvariedad de codimensión 2 de $\mathbb{A}^{d-m}$.

Por otro lado, el Lema 4.2.4 muestra que Φ_2^e es un morfismo finito. Por lo tanto, del hecho de que la imagen inversa por Φ_2^e de una subvariedad de $\mathbb{A}^{d-m}$ de codimensión 2 es una subvariedad de V^e de codimensión 2, deducimos la primera afirmación de la proposición.

Demostramos ahora la segunda afirmación. Sea $(\boldsymbol{y}, \boldsymbol{b})$ un punto singular de V_d^e y sea $\mathcal{T}_{\boldsymbol{y}} V_d^e$ el espacio tangente de V_d^e en $\boldsymbol{y}$. Como $V_d^e = V(R_1, \ldots, R_d)$, para cualquier punto $\boldsymbol{v} \in \mathcal{T}_{\boldsymbol{y}} V_d^e$ tenemos que $(\partial \boldsymbol{R}^e/\partial \boldsymbol{Y})(\boldsymbol{y}, \boldsymbol{b}) \cdot \boldsymbol{v} = \boldsymbol{0}$. Si la matriz Jacobiana $(\partial \boldsymbol{R}^e/\partial \boldsymbol{Y})(\boldsymbol{y}, \boldsymbol{b})$ tuviera rango máximo, entonces $\mathcal{T}_{\boldsymbol{y}} V_d^e$ debería tener dimensión a lo sumo $d - m$, lo que contradiría el hecho de que $(\boldsymbol{y}, \boldsymbol{b})$ es un punto singular de V_d^e. Esto finaliza la demostración de la segunda afirmación. □

Finalizamos dando el resultado principal de esta sección, que acota superiormente la dimensión del lugar singular de V_d.

Teorema 4.2.7. *Sea $p > 2$. El conjunto de puntos $\boldsymbol{y} \in V_d$ para los cuales $(\partial \boldsymbol{R}/\partial \boldsymbol{Y})(\boldsymbol{y})$ no es de rango completo, tiene codimensión al menos 2 en V_d. En particular, el lugar singular Σ de V_d tiene codimensión al menos 2 en V_d.*

Demostración. Sea $\mathcal{D}$ el conjunto formado por todos los puntos $(\boldsymbol{y}, \boldsymbol{b}) \in V_d^e$ para los cuales $(\partial \boldsymbol{R}^e/\partial(\boldsymbol{Y}, \boldsymbol{B}))(\boldsymbol{y}, \boldsymbol{b})$ no es de rango completo y sea $\mathcal{E}$ el conjunto de puntos $\boldsymbol{y} \in V_d$ para los cuales $(\partial \boldsymbol{R}/\partial \boldsymbol{Y})(\boldsymbol{y})$ no es de rango completo. Recordemos que la proyección $\Phi_1^e : V_d^e \to V_d$ definida por $\Phi_1^e(\boldsymbol{y}, \boldsymbol{b}) := \boldsymbol{y}$ es un isomorfismo de variedades afines. Más aún, la Observación 4.2.2 asegura que $\Phi_1^e(\mathcal{D}) = \mathcal{E}$. Por otro lado, la Proposición 4.2.6 muestra que $\mathcal{D}$ está contenido en una subvariedad de codimensión 2 en V_d^e, lo que implica que $\mathcal{E}$ está contenido en una subvariedad de codimensión 2 en V_d. Esto prueba la primera afirmación.

Sea ahora $\boldsymbol{y}$ un punto de Σ. Del Lema 4.2.1 tenemos que $\dim \mathcal{T}_{\boldsymbol{y}} V_d > d - m$. Esto implica que $\operatorname{rg}(\partial \boldsymbol{R}/\partial \boldsymbol{Y})(\boldsymbol{y}) < m$, pues de otra manera tendríamos que $\dim \mathcal{T}_{\boldsymbol{y}} V_d \leq d - m$, lo cual contradiría el hecho de que $\boldsymbol{y}$ es un punto singular de V_d. Así de la primera afirmación, que ya demostramos, se sigue la segunda afirmación. □

Del Lema 4.2.1 y el Teorema 4.2.7 obtenemos las siguientes propiedades de los polinomios $R_1, \ldots, R_m$ y de la variedad V_d. Por el Teorema 4.2.7, el conjunto de puntos $\boldsymbol{y} \in V_d$ para los cuales la matriz Jacobiana $(\partial \boldsymbol{R}/\partial \boldsymbol{Y})(\boldsymbol{y})$ no es de rango completo, tiene codimensión al menos 2 en V_d. Como los polinomios $R_1, \ldots, R_m$ forman una sucesión regular de $\mathbb{F}_q[\boldsymbol{Y}]$, de [Eis95, Theorem 18.15] concluimos que $R_1, \ldots, R_m$ definen un ideal radical de $\mathbb{F}_q[\boldsymbol{Y}]$. Así, V_d es una intersección completa.

Por otro lado, recordemos que la matriz Jacobiana $(\partial \boldsymbol{S}/\partial \boldsymbol{Z})$ es una matriz escalonada por columnas y que los índices $i_1, \ldots, i_m$ corresponden a las posiciones de los pivotes de $(\partial \boldsymbol{S}/\partial \boldsymbol{Z})$. Entonces cada polinomio R_j tiene grado i_j para $1 \leq j \leq m$. Por la desigualdad de Bézout (2.1) tenemos que $\deg V_d \leq \prod_{j=1}^m \deg R_j = i_1 \cdots i_m$. En otras palabras, obtenemos el siguiente resultado.

Corolario 4.2.8. *Sea $p > 2$. Los polinomios $R_1, \ldots, R_m$ definen un ideal radical y la variedad V_d tiene grado $\deg V_d \leq \prod_{j=1}^m \deg R_j = i_1 \cdots i_m$.*

4.2.2. La geometría de la clausura proyectiva

En esta sección estudiamos la clausura proyectiva $\mathrm{pcl}(V_d) \subset \mathbb{P}^d$ de la variedad V_d y del conjunto de puntos de dicha clausura en el hiperplano del infinito. Estas propiedades nos permitirán dar una estimación explícita de la cantidad de puntos $\mathbb{F}_q$–racionales de V_d.

Comenzamos discutiendo el comportamiento de $\mathrm{pcl}(V_d)$ en el hiperplano del infinito. De acuerdo a (4.10), cada R_k está definido por

$$R_k = S_k(\Pi_1, \ldots, \Pi_s),$$

donde $S_1, \ldots, S_m \in \mathbb{F}_q[\boldsymbol{Z}]$ están definidos en (4.9). Como antes, vamos a suponer que la matriz Jacobiana $(\partial \boldsymbol{S}/\partial \boldsymbol{Z})$ está escalonada por columnas, es decir, existen $1 \leq i_1 < i_2 < \cdots < i_m \leq s$ tales que

$$R_k = b_{k,0} + \sum_{j=1}^{i_k} (-1)^j b_{k,d-j} \Pi_j, \tag{4.15}$$

donde $b_{k,d-i_k} \neq 0$ para cada $1 \leq k \leq m$. Así, la homogeneización de cada R_k es el siguiente polinomio de $\mathbb{F}_q[Y_0, \ldots, Y_d]$:

$$R_k^h = b_{k,0} Y_0^{i_k} + \sum_{j=1}^{i_k} (-1)^j b_{k,d-j} \Pi_j Y_0^{i_k - j} \quad (1 \leq k \leq m). \tag{4.16}$$

Se deduce que $R_k^h(0, Y_1, \ldots, Y_d) = b_{k,d-i_k} \Pi_{i_k}$ para cada $1 \leq k \leq m$. Observemos que los polinomios $\Pi_{i_1}, \ldots, \Pi_{i_m}$ son una posible elección para los polinomios $R_1, \ldots, R_m$ definidos en (4.10). Por lo tanto, puede aplicarse el Lema 4.2.1, el Teorema 4.2.7 y el Corolario 4.2.8 a dichos polinomios.

Sea $\Sigma^\infty \subset \mathbb{P}^d$ el lugar singular de $\mathrm{pcl}(V_d)$ en el hiperplano del infinito, es decir, el conjunto de puntos singulares de $\mathrm{pcl}(V_d)$ en el hiperplano $\{Y_0 = 0\}$. En el siguiente lema damos una cota superior de la dimensión de Σ^∞.

Lema 4.2.9. *Si $p > 2$ y $m \leq s \leq d-3$, entonces el lugar singular $\Sigma^\infty \subset \mathbb{P}^d$ de $\mathrm{pcl}(V_d)$ en el hiperplano del infinito tiene dimensión a lo sumo $d-m-3$.*

Demostración. Sea $\boldsymbol{y} := (0 : y_1 : \cdots : y_d)$ un punto arbitrario de Σ^∞. Como los polinomios R_k^h se anulan en $\mathrm{pcl}(V_d)$, tenemos que $R_k^h(\boldsymbol{y}) = b_{k,d-i_k} \Pi_{i_k}(y_1, \ldots, y_d) = 0$ para $1 \leq k \leq m$. Denotemos por $(\partial \boldsymbol{\Pi}_{\mathcal{I}}/\partial \boldsymbol{Y})$ la matriz Jacobiana $\Pi_{i_1}, \ldots, \Pi_{i_m}$ con respecto a $Y_1, \ldots, Y_d$. Afirmamos que

$$\mathrm{rg} \left(\frac{\partial \boldsymbol{\Pi}_{\mathcal{I}}}{\partial \boldsymbol{Y}} \right) (\boldsymbol{y}) < m. \tag{4.17}$$

En efecto, si el rango de dicha matriz fuera igual a m, tendríamos $\dim \mathcal{T}_{\boldsymbol{y}}(\mathrm{pcl}(V_d)) \leq d-m$, lo cual implicaría que $\boldsymbol{y}$ sería un punto no singular de $\mathrm{pcl}(V_d)$. Esto contradice la hipótesis sobre $\boldsymbol{y}$.

Por otro lado, observemos que los polinomios $\Pi_{i_1}, \ldots, \Pi_{i_m}$ satisfacen las hipótesis del Teorema 4.2.7. Deducimos así que el conjunto de puntos $\boldsymbol{y}_{\mathrm{aff}} := (0, y_1, \ldots, y_d) \in$

$V(\Pi_1,\ldots,\Pi_{i_m})\subset\mathbb{A}^d$ tales que $(\partial\mathbf{\Pi}_{\mathcal{I}}/\partial\boldsymbol{Y})\,(\boldsymbol{y}_{\mathrm{aff}})$ no es de rango completo, resulta un cono afín equidimensional de $\mathbb{A}^{d+1}$ de dimensión a lo sumo $d-m-2$. Por lo tanto, la variedad proyectiva $\Sigma^\infty\subset\mathbb{P}^d$ tiene dimensión a lo sumo $d-m-3$. □

El siguiente resultado concierne la variedad proyectiva $V(\Pi_{i_1},\ldots,\Pi_{i_m})\subset\mathbb{P}^{d-1}$ y nos permitirá obtener información sobre el comportamiento de $\mathrm{pcl}(V_d)$ en el hiperplano del infinito.

Lema 4.2.10. *Sea $p>2$. Entonces $V(\Pi_{i_1},\ldots,\Pi_{i_m})\subset\mathbb{P}^{d-1}$ es absolutamente irreducible de dimensión $d-m-1$, grado $i_1\cdots i_m$ y lugar singular de dimensión a lo sumo $d-m-3$.*

Demostración. Por (4.15) tenemos que $\Pi_{i_1},\ldots,\Pi_{i_m}$ se anulan en $\mathrm{pcl}(V_d)\cap\{Y_0=0\}$. El Lema 4.2.1 asegura que el cono afín de $\mathbb{A}^d$ definido por los polinomios $\Pi_{i_1},\ldots,\Pi_{i_m}$ es una intersección completa conjuntista de dimensión $d-m$. Concluimos que $V(\mathbf{\Pi}_{\mathcal{I}}):=V(\Pi_{i_1},\ldots,\Pi_{i_m})\subset\mathbb{P}^{d-1}$ es una intersección completa conjuntista de dimensión $d-m-1$. Además, el Teorema 4.2.7 muestra que el lugar singular de $V(\mathbf{\Pi}_{\mathcal{I}})$ tiene codimensión al menos 2 en $V(\mathbf{\Pi}_{\mathcal{I}})$. Tenemos entonces que $V(\mathbf{\Pi}_{\mathcal{I}})$ es regular en codimensión 1, es decir, es una intersección completa normal. Del Teorema 2.1.10 se sigue que $V(\mathbf{\Pi}_{\mathcal{I}})$ es absolutamente irreducible, lo que completa la demostración del lema. □

En el siguiente teorema caracterizamos el comportamiento de $\mathrm{pcl}(V_d)$ en el hiperplano del infinito.

Teorema 4.2.11. *Si $p>2$ y $m\le s\le d-3$, entonces $\mathrm{pcl}(V_d)\cap\{Y_0=0\}\subset\mathbb{P}^{d-1}$ es una intersección completa normal de dimensión $d-m-1$ y grado $i_1\cdots i_m$.*

Demostración. Por el Lema 4.2.1 vemos que la variedad proyectiva $\mathrm{pcl}(V_d)$ es de dimensión pura $d-m$. Entonces cada componente irreducible de $\mathrm{pcl}(V_d)\cap\{Y_0=0\}$ tiene dimensión al menos $d-m-1$. Dado que $\mathrm{pcl}(V_d)\cap\{Y_0=0\}$ está contenida en la variedad proyectiva $V(\mathbf{\Pi}_{\mathcal{I}})$ y, por el Lema 4.2.10, $V(\mathbf{\Pi}_{\mathcal{I}})$ es absolutamente irreducible de dimensión $d-m-1$, $\mathrm{pcl}(V_d)\cap\{Y_0=0\}$ resulta absolutamente irreducible de dimensión $d-m-1$. Por lo tanto,

$$\mathrm{pcl}(V_d)\cap\{Y_0=0\}=V(\Pi_{i_1},\ldots,\Pi_{i_m}).$$

Por otro lado, de [Eis95, Theorem 18.15] deducimos que $\Pi_{i_1},\ldots,\Pi_{i_m}$ definen un ideal radical. Así, tenemos que $\mathrm{pcl}(V_d)\cap\{Y_0=0\}$ es una intersección completa de dimensión $d-m-1$, y por el Teorema 2.1.14 vemos que

$$\deg\big(\mathrm{pcl}(V_d)\cap\{Y_0=0\}\big)=\prod_{j=1}^{m}\deg\Pi_{i_j}=i_1\cdots i_m.$$

Finalmente, del Lema 4.2.10 deducimos que el lugar singular de $\mathrm{pcl}(V_d)\cap\{Y_0=0\}$ tiene codimensión al menos 2. Concluimos que $\mathrm{pcl}(V_d)\cap\{Y_0=0\}$ es una intersección completa normal. Esto finaliza la demostración del Teorema. □

Terminamos esta sección con un resultado sobre la clausura proyectiva $\mathrm{pcl}(V_d)$.

Teorema 4.2.12. *Si $p > 2$ y $m \leq s \leq d-3$, entonces $\mathrm{pcl}(V_d) \subset \mathbb{P}^d$ es una intersección completa normal de dimensión $d-m$ y grado $i_1 \cdots i_m$.*

Demostración. Sabemos que $\mathrm{pcl}(V_d)$ es de dimensión pura $d-m$. Por un lado, el conjunto de puntos singulares de $\mathrm{pcl}(V_d)$ que pertenecen al abierto $\{Y_0 \neq 0\}$ está contenido en el lugar singular de $\mathrm{pcl}(V_d) \cap \{Y_0 \neq 0\}$, y el Teorema 4.2.7 muestra que el lugar singular de $\mathrm{pcl}(V_d) \cap \{Y_0 \neq 0\}$ tiene dimensión a lo sumo $d-m-2$. Por otro lado, por [GL02a, Lemma 1.1] tenemos que el lugar singular de $\mathrm{pcl}(V_d)$ en el hiperplano en el infinito está contenido en el lugar singular de $\mathrm{pcl}(V_d) \cap \{Y_0 = 0\}$, y el Teorema 4.2.11 muestra que el lugar singular de $\mathrm{pcl}(V_d)$ en el hiperplano del infinito tiene dimensión a lo sumo $d-m-3$. Por lo tanto, el lugar singular de $\mathrm{pcl}(V_d)$ tiene dimensión a lo sumo $d-m-2$.

Por otro lado, observemos que $\mathrm{pcl}(V_d)$ está contenida en la variedad proyectiva $V(\boldsymbol{R}^h) := V(R_j^h : 1 \leq j \leq m) \subset \mathbb{P}^d$. Además, tenemos las siguientes inclusiones:

$$V(\boldsymbol{R}^h) \cap \{Y_0 \neq 0\} \subset V(\boldsymbol{R}), \quad V(\boldsymbol{R}^h) \cap \{Y_0 = 0\} \subset V(\boldsymbol{\Pi}_{\mathcal{I}}).$$

El Lema 4.2.10 prueba que $V(\boldsymbol{\Pi}_{\mathcal{I}}) \subset \mathbb{P}^{d-1}$ es absolutamente irreducible de dimensión pura $d-m-1$, mientras que el Lema 4.2.1 muestra que $V(\boldsymbol{R}) \subset \mathbb{A}^d$ es de dimensión pura $d-m$. Concluimos que $V(\boldsymbol{R}^h) \subset \mathbb{P}^d$ tiene dimensión a lo sumo $d-m$. Dado que $V(\boldsymbol{R}^h)$ está definida por m polinomios, resulta una intersección completa conjuntista y, por ende, equidimensional de dimensión $d-m$. Por lo tanto, ninguna de sus componentes irreducibles está contenida en el hiperplano del infinito. Esto implica que la clausura proyectiva de la restricción de $V(\boldsymbol{R}^h)$ al espacio afín $\mathbb{A}^d$ coincide con $V(\boldsymbol{R}^h)$, es decir, $\mathrm{pcl}(V(\boldsymbol{R}^h) \cap \mathbb{A}^d) = V(\boldsymbol{R}^h)$ (ver, por ejemplo, [Kun85, Proposition I.5.17]). Como $V(\boldsymbol{R}^h) \cap \mathbb{A}^d$ es la variedad afín $V_d = V(\boldsymbol{R})$, deducimos que

$$\mathrm{pcl}(V_d) = V(\boldsymbol{R}^h).$$

Como el lugar singular de $V(\boldsymbol{R}^h)$ tiene codimensión al menos 2, tenemos que $V(\boldsymbol{R}^h)$ es una intersección completa normal. De [Eis95, Theorem 18.15] deducimos que los polinomios $R_1^h, \ldots, R_m^h$ definen un ideal radical. Finalmente, el Teorema 2.1.14 asegura que $\deg \mathrm{pcl}(V) = \prod_{j=1}^m \deg R_j^h = i_1 \cdots i_m$.

□

4.2.3. El número de puntos $\mathbb{F}_q$–racionales

Sea $p > 2$ y sean m y s enteros positivos tales que $m \leq s \leq d-3$. En esta sección damos una estimación de la cantidad de puntos $\mathbb{F}_q$–racionales de la variedad afín $V_d \subset \mathbb{A}^d$ definida por los polinomios $R_1, \ldots, R_m$ de (4.10). Para esto, vamos a utilizar el Teorema 2.2.10, que nos proporciona una estimación sobre el número de puntos $\mathbb{F}_q$–racionales de una intersección completa normal proyectiva.

Los Teoremas 4.2.11 y 4.2.12 aseguran que la clausura proyectiva $\mathrm{pcl}(V_d) \subset \mathbb{P}^d$ de V_d y el conjunto de puntos $\mathrm{pcl}(V_d)^\infty \subset \mathbb{P}^{d-1}$ de $\mathrm{pcl}(V_d)$ al infinito son $\mathbb{F}_q$–variedades proyectivas que resultan intersecciones completas normales de dimensión $d-m-1$

y $d-m$ respectivamente, ambas de grado $\delta_{V_d} := i_1 \cdots i_m$. Por lo tanto, aplicando la estimación (2.7) del Teorema 2.2.10, obtenemos que

$$\begin{aligned}\big||\mathrm{pcl}(V_d)(\mathbb{F}_q)| - p_{d-m}\big| &\leq (\delta_{V_d}(D_{V_d}-2)+2)q^{d-m-\frac{1}{2}} + 14D_{V_d}^2\delta_{V_d}^2 q^{d-m-1},\\ \big||\mathrm{pcl}(V_d)^\infty(\mathbb{F}_q)| - p_{d-m-1}\big| &\leq (\delta_{V_d}(D_{V_d}-2)+2)q^{d-m-\frac{3}{2}} + 14D_{V_d}^2\delta_{V_d}^2 q^{d-m-2},\end{aligned}$$

donde $\delta_{V_d} := i_1 \cdots i_m$ y $D_{V_d} := \sum_{j=1}^m (i_j - 1)$. Así, la cantidad de puntos $\mathbb{F}_q$–racionales de V_d satisface la siguiente estimación:

$$\begin{aligned}\big||V_d(\mathbb{F}_q)| - q^{d-m}\big| &= \big||\mathrm{pcl}(V_d)(\mathbb{F}_q)| - |\mathrm{pcl}(V_d)^\infty(\mathbb{F}_q)| - p_{d-m} + p_{d-m-1}\big|\\ &\leq \big||\mathrm{pcl}(V_d)(\mathbb{F}_q)| - p_{d-m}\big| + \big||\mathrm{pcl}(V_d)^\infty(\mathbb{F}_q)| - p_{d-m-1}\big|\\ &\leq (q+1)q^{d-m-2}\big((\delta_{V_d}(D_{V_d}-2)+2)q^{1/2} + 14D_{V_d}^2\delta_{V_d}^2\big). \qquad (4.18)\end{aligned}$$

En consecuencia, tenemos el siguiente resultado.

Teorema 4.2.13. *Sea $p > 2$, sean d, s y m enteros positivos tales que $m \leq s \leq d-3$, y sean $R_1, \ldots, R_m \in \mathbb{F}_q[Y_1, \ldots, Y_d]$ los polinomios $R_k = S_k(\Pi_1, \ldots, \Pi_s)$ para $1 \leq k \leq m$, donde $S_1, \ldots, S_m \in \mathbb{F}_q[Z_1, \ldots, Z_s]$ son los polinomios de grado 1 definidos en (4.9), cuyas componentes homogéneas de grado 1 son linealmente independientes. Si $V_d := V(R_1, \ldots, R_m) \subset \mathbb{A}^d$, entonces la siguiente estimación es válida:*

$$\big||V_d(\mathbb{F}_q)| - q^{d-m}\big| \leq (q+1)q^{d-m-2}\big((\delta_{V_d}(D_{V_d}-2)+2)q^{1/2} + 14D_{V_d}^2\delta_{V_d}^2\big),$$

donde $\delta_{V_d} := i_1 \cdots i_m$ y $D_{V_d} := \sum_{j=1}^m (i_j - 1)$.

Capítulo 5

Conjunto de valores en familias lineales

Este capítulo está dedicado a estudiar un problema combinatorio clásico sobre un cuerpo finito: el comportamiento del cardinal de la imagen o "conjunto de valores" en familias lineales. Más precisamente, vamos a dar estimaciones explícitas del promedio del cardinal del conjunto de valores de familias lineales de polinomios mónicos univariados de grado d y con coeficientes en $\mathbb{F}_q$. Como un caso particular, vamos a obtener estimaciones del promedio del cardinal del conjunto de valores de familias de polinomios univariados con ciertos coeficientes consecutivos prescriptos. Para todas estas estimaciones consideramos el caso en que d es menor que q y obtenemos tanto el comportamiento asintótico como una cota superior explícita de la desviación respecto de dicho comportamiento en términos de d y q.

Para esto, traducimos el problema en un problema geométrico: el de estimar el número de puntos $\mathbb{F}_q$–racionales con coordenadas distintas dos a dos de ciertas familias de intersecciones completas singulares definidas sobre $\mathbb{F}_q$, para lo cual utilizamos los resultados del Capítulo 3.

En las últimas dos secciones de este capítulo mostramos cómo el enfoque que utilizamos nos permite abordar otros dos problemas combinatorios: el estudio del promedio del cardinal del conjunto de valores en familias no lineales y del segundo momento del conjunto de polinomios con ciertos coeficientes prescriptos.

5.1. El problema

Sea T una indeterminada sobre $\overline{\mathbb{F}}_q$. Para un polinomio $f \in \mathbb{F}_q[T]$, definimos el *conjunto de valores* de f como el conjunto imagen de la función polinomial de $\mathbb{F}_q$ en $\mathbb{F}_q$ que define f. Denotamos por $\mathcal{V}(f)$ al cardinal de dicho conjunto, es decir,

$$\mathcal{V}(f) := |\{f(c) : c \in \mathbb{F}_q\}|.$$

En [Coh72], Cohen estudia el comportamiento asintótico del cardinal del conjunto de valores promedio en familias lineales de polinomios univariados con coeficientes en $\mathbb{F}_q$. Más precisamente, afirma que, para una familia lineal $\mathcal{A} \subset \mathbb{F}_q[T]_d$ que satisface

ciertas condiciones técnicas, si $p > d$ y la codimensión de $\mathcal{A}$ es $m \leq d-2$, entonces

$$\mathcal{V}(\mathcal{A}) := \frac{1}{|\mathcal{A}|}\sum_{f\in\mathcal{A}} \mathcal{V}(f) = \mu_d q + \mathcal{O}(q^{1/2}). \tag{5.1}$$

Sin embargo, Cohen no da una expresión explícita del término de error y su resultado impone fuertes restricciones sobre la característica del cuerpo.

En este capítulo consideramos la familia lineal que describimos a continuación. Sean m y d enteros positivos tales que $q > d$ y $3 \leq r \leq d-m$, sean $A_{d-1}, \ldots, A_r$ indeterminadas sobre $\overline{\mathbb{F}}_q$ y sean $L_1, \ldots, L_m \in \mathbb{F}_q[A_{d-1}, \ldots, A_r]$ las siguientes formas lineales:

$$L_k := b_{k,d-1}A_{d-1} + \cdots + b_{k,r}A_r + b_{k,0} \ (1 \leq k \leq m). \tag{5.2}$$

Sin pérdida de generalidad podemos suponer que $L_1, \ldots, L_m$ son linealmente independientes. Sea $\boldsymbol{L} := (L_1, \ldots, L_m)$ y sea $\mathcal{A} := \mathcal{A}_{\boldsymbol{L}}$ la familia lineal definida de la siguiente manera:

$$\mathcal{A} := \{T^d + a_{d-1}T^{d-1} + \cdots + a_1T + a_0 \in \mathbb{F}_q[T]_d : \boldsymbol{L}(a_{d-1}, \ldots, a_r) = 0\}. \tag{5.3}$$

Observemos que esta familia lineal fue considerada en el Capitulo 3 (ver (3.4)). Suponemos además que la matriz $\mathcal{M}(\boldsymbol{L}) := (b_{k,d-j})_{1\leq k\leq m,\, 1\leq j\leq d-r}$ está escalonada por filas, donde $1 \leq j_1 < \cdots < j_m \leq d-r$ denotan las posiciones de las columnas de $\mathcal{M}(\boldsymbol{L})$ correspondientes a los pivotes.

En las siguientes secciones determinamos el comportamiento asintótico de $\mathcal{V}(\mathcal{A})$, dando asimismo una cota explícita de la desviación respecto de dicho comportamiento. Más precisamente, probamos que, si $p > 2$, entonces

$$|\mathcal{V}(\mathcal{A}) - \mu_d q| \leq d^2 2^{d-1} q^{1/2} + 133 d^{d+5} e^{2\sqrt{d}-d}. \tag{5.4}$$

Un caso particular de estas familias lineales es la familia $\mathcal{A}_{\boldsymbol{a}}$ que consiste de todos los polinomios en $\mathbb{F}_q[T]_d$ con los primeros $s \leq d-2$ prescriptos. Más precisamente, sea s un entero tal que $1 \leq s \leq d-2$ y sea $\boldsymbol{a} := (a_{d-1}, \cdots, a_{d-s}) \in \mathbb{F}_q^s$. Denotamos con $f_{\boldsymbol{a}} := T^d + a_{d-1}T^{d-s} + \cdots + a_{d-s}T^{d-s}$. Entonces la familia $\mathcal{A}_{\boldsymbol{a}}$ se define de la siguiente manera:

$$\mathcal{A}_{\boldsymbol{a}} := \{f_{\boldsymbol{b}} := f_{\boldsymbol{a}} + b_{d-s-1}T^{d-s-1} + \cdots + b_1T : \boldsymbol{b} := (b_{d-s-1}, \ldots, b_1) \in \mathbb{F}_q^{d-s-1}\}. \tag{5.5}$$

En la literatura encontramos a Uchiyama [Uch55b] y Cohen [Coh72], que estudian el problema de estimar el valor promedio $\mathcal{V}(d,s)$ de $\mathcal{V}(f)$ cuando f recorre todos los elementos de esta familia. Más precisamente, prueban que si $p > d$, entonces

$$\mathcal{V}(d,s) := \frac{1}{|\mathcal{A}_{\boldsymbol{a}}|}\sum_{f\in\mathcal{A}_{\boldsymbol{a}}} \mathcal{V}(f) = \mu_d q + \mathcal{O}(q^{1/2}), \tag{5.6}$$

donde la constante que aparece en la notación $\mathcal{O}$ depende solamente de d y s. Cabe mencionar que ni Uchiyama ni Cohen dan una expresión explícita de dicha constante y su resultado impone fuertes restricciones sobre la característica del cuerpo. La estimación (5.4) mejora en ambos aspectos los resultados de Cohen y Uchiyama. Dicha estimación será una herramienta fundamental en el análisis de la complejidad en promedio de los algoritmos que calculan puntos $\mathbb{F}_q$–racionales de hipersuperficies de los Capítulos 8 y 9.

5.2. El promedio del cardinal del conjunto de valores

Con el objetivo de determinar el comportamiento asintótico del promedio $\mathcal{V}(\mathcal{A})$, en la Sección 5.2.1 damos una expresión combinatoria de dicho promedio en términos del número $\mathcal{S}_i^{\mathcal{A}}$ de ciertos "conjuntos interpolantes" con $r+1 \leq i \leq d$. Luego, en la Sección 5.2.2 relacionamos el número $\mathcal{S}_i^{\mathcal{A}}$ con la cantidad de puntos $\mathbb{F}_q$–racionales con coordenadas distintas dos a dos de cierta variedad algebraica Γ_i^* para $r+1 \leq i \leq d$.

5.2.1. Una reducción combinatoria

Comenzamos mostrando cómo se puede relacionar el problema de estimar el número $\mathcal{V}(\mathcal{A})$ con un problema de interpolación. Podemos suponer sin pérdida de generalidad que $\mathcal{V}(\mathcal{A})$ es el valor promedio de $\mathcal{V}(f)$ cuando f recorre todos los elementos de $\mathcal{A}$ tales que $f(0)=0$, ya que, si tomamos f con estas características y $a_0 \in \mathbb{F}_q$, entonces $\mathcal{V}(f)$ coincide con $\mathcal{V}(f+a_0)$. Por este motivo, vamos a considerar que los elementos de $\mathcal{A}$ cumplen que $f(0)=0$.

Observamos que, si $f \in \mathcal{A}$, entonces $\mathcal{V}(f)$ es igual al número de elementos $a_0 \in \mathbb{F}_q$ para los cuales el polinomio $f+a_0$ tiene al menos un cero en $\mathbb{F}_q$. Sea $\mathbb{F}_q[T]_d$ el conjunto de polinomios en $\mathbb{F}_q[T]$ de grado d y sea $N := N_{1,d} : \mathbb{F}_q[T]_d \to \mathbb{Z}_{\geq 0}$ la función que a cada polinomio f le asigna la cantidad de ceros que éste posee en $\mathbb{F}_q$. Por último, consideramos la función característica $\mathbf{1}_{\{N>0\}} : \mathbb{F}_q[T]_d \to \{0,1\}$ del conjunto de elementos de $\mathbb{F}_q[T]_d$ que tiene al menos un cero en $\mathbb{F}_q$. De las observaciones previas deducimos la siguiente igualdad:

$$\sum_{f\in\mathcal{A}} \mathcal{V}(f) = \sum_{a_0\in\mathbb{F}_q}\sum_{f\in\mathcal{A}} \mathbf{1}_{\{N>0\}}(f+a_0) = \big|\{f+a_0 \in \mathcal{A}+\mathbb{F}_q : N(f+a_0)>0\}\big|.$$

Dado un subconjunto $\mathcal{X} \subset \mathbb{F}_q$, definimos $\mathcal{S}_{\mathcal{X}}^{\mathcal{A}} \subset \mathbb{F}_q[T]$ como el conjunto de polinomios $f+a_0 \in \mathcal{A}+\mathbb{F}_q$ que se anula en $\mathcal{X}$, es decir,

$$\mathcal{S}_{\mathcal{X}}^{\mathcal{A}} := \{f+a_0 \in \mathcal{A}+\mathbb{F}_q : \ (f+a_0)(x)=0 \text{ para cualquier } x \in \mathcal{X}\}.$$

Finalmente, dado $i \in \mathbb{N}$, notamos con $\mathcal{X}_i$ a un subconjunto de $\mathbb{F}_q$ de i elementos. El siguiente resultado nos permitirá determinar el comportamiento asintótico de $\mathcal{V}(\mathcal{A})$.

Teorema 5.2.1. *Dados $r,d,m \in \mathbb{N}$ con $d<q$ y $3 \leq r \leq d-m$, se satisface la siguiente igualdad:*

$$\mathcal{V}(\mathcal{A}) = \sum_{i=1}^{r}(-1)^{i-1}\binom{q}{i}q^{1-i} + \frac{1}{q^{d-m-1}}\sum_{i=r+1}^{d}(-1)^{i-1}\sum_{\mathcal{X}_i\subset\mathbb{F}_q}|\mathcal{S}_{\mathcal{X}_i}^{\mathcal{A}}|.$$

Demostración. Dado un subconjunto $\mathcal{X}_i := \{\alpha,\ldots,\alpha_i\} \subset \mathbb{F}_q$, consideramos el conjunto $\mathcal{S}_{\mathcal{X}_i}^{\mathcal{A}}$ definido como arriba. Es fácil ver que $\mathcal{S}_{\mathcal{X}_i}^{\mathcal{A}} = \cap_{j=1}^{i}\mathcal{S}_{\{\alpha_j\}}^{\mathcal{A}}$ y que

$$|\{f+a_0 \in \mathcal{A}+\mathbb{F}_q : \ N(f+a_0)>0\} = \Big|\bigcup_{x\in\mathbb{F}_q}\mathcal{S}_{\{x\}}^{\mathcal{A}}\Big|.$$

Por lo tanto, por el principio de inclusión–exclusión obtenemos que

$$\mathcal{V}(\mathcal{A}) = \frac{1}{q^{d-m-1}} \Big| \bigcup_{x \in \mathbb{F}_q} \mathcal{S}^{\mathcal{A}}_{\{x\}} \Big| = \frac{1}{q^{d-m-1}} \sum_{i=1}^{q} (-1)^{i-1} \sum_{\mathcal{X}_i \subset \mathbb{F}_q} |\mathcal{S}^{\mathcal{A}}_{\mathcal{X}_i}|. \tag{5.7}$$

A continuación estimamos el número $|\mathcal{S}^{\mathcal{A}}_{\mathcal{X}_i}|$ para un conjunto $\mathcal{X}_i := \{\alpha_1, \ldots, \alpha_i\} \subset \mathbb{F}_q$. Notar que si $f + a_0$ es un elemento arbitrario de $\mathcal{S}^{\mathcal{A}}_{\mathcal{X}_i}$, entonces $(f + a_0)(\alpha_j) = 0$ para $1 \leq j \leq i$ y $L_k(a_{d-1}, \ldots, a_r) = 0$ para $1 \leq k \leq m$. Estas identidades se pueden expresar en forma matricial de la siguiente manera:

$$\mathcal{M} \cdot A^T = -\Lambda^T,$$

donde $\mathcal{M} \in \mathbb{F}_q^{(m+i)\times d}$ es la siguiente matriz por bloques:

$$\mathcal{M} := \begin{pmatrix} \mathcal{M}(\boldsymbol{L}) & \mathbf{0} \\ * & \boldsymbol{V}(\mathcal{X}_i) \end{pmatrix}. \tag{5.8}$$

Aquí $V(\mathcal{X}_i) := (m_{j,k}) \in \mathbb{F}_q^{i\times r}$ es la matriz de Vandermonde definida por $m_{j,k} := \alpha_j^k$ para $1 \leq j \leq i$ y $0 \leq k \leq r-1$, $A^T := (a_{d-1}, \ldots, a_0) \in \mathbb{F}_q^{d\times 1}$ y $\Lambda^T := (b_{1,0}, \ldots, b_{m,0}, \alpha_1^d, \ldots, \alpha_i^d) \in \mathbb{F}_q^{(m+i)\times 1}$.

Como la matriz $\mathcal{M}(\boldsymbol{L})$ tiene rango m y $\operatorname{rg}(\boldsymbol{V}(\mathcal{X}_i)) = \min\{i, r\}$, concluimos que $\operatorname{rg}(\mathcal{M}) = m + i \leq d$ para $i \leq r$. Como $\mathcal{S}^{\mathcal{A}}_{\mathcal{X}_i}$ es una $\mathbb{F}_q$–variedad lineal en $\mathbb{F}_q^d$, se sigue que $\dim(\mathcal{S}^{\mathcal{A}}_{\mathcal{X}_i}) = d - m - i$, y así,

$$|\mathcal{S}^{\mathcal{A}}_{\mathcal{X}_i}| = q^{d-m-i}. \tag{5.9}$$

Por otro lado, si $i > d$ y $f + a_0 \in \mathcal{S}^{\mathcal{A}}_{\mathcal{X}_i}$, el polinomio no nulo $f + a_0$ tiene grado d y se anula en $i > d$ elementos distintos de $\mathbb{F}_q$. Esto último no es posible, por lo que deducimos que $\mathcal{S}^{\mathcal{A}}_{\mathcal{X}_i} = \emptyset$, y por lo tanto,

$$|\mathcal{S}^{\mathcal{A}}_{\mathcal{X}_i}| = 0. \tag{5.10}$$

Combinando (5.7), (5.9) y (5.10), obtenemos que

$$\mathcal{V}(\mathcal{A}) = \frac{1}{q^{d-m-1}} \sum_{i=1}^{r} (-1)^{i-1} \binom{q}{i} q^{d-m-i} + \frac{1}{q^{d-m-1}} \sum_{i=r+1}^{d} (-1)^{i-1} \sum_{\mathcal{X}_i \subset \mathbb{F}_q} |\mathcal{S}^{\mathcal{A}}_{\mathcal{X}_i}|.$$

Se sigue así la afirmación del teorema. □

5.2.2. Un enfoque geométrico

De acuerdo al Teorema 5.2.1, para determinar el comportamiento asintótico del promedio $\mathcal{V}(\mathcal{A})$ necesitamos estimar el número

$$\mathcal{S}_i^{\mathcal{A}} := \sum_{\mathcal{X}_i \subset \mathbb{F}_q} |\mathcal{S}^{\mathcal{A}}_{\mathcal{X}_i}|, \tag{5.11}$$

para cada $r+1 \leq i \leq d$ y $3 \leq r \leq d-m$. Para esto, vamos a traducir este problema en el de estimar la cantidad de puntos $\mathbb{F}_q$–racionales con coordenadas distintas dos a dos de cierta variedad de incidencia, que definimos a continuación.

Fijamos i con $r+1 \leq i \leq d$. Sean $A_{d-1}, \ldots, A_0$ indeterminadas sobre $\overline{\mathbb{F}}_q$ y sean $L_1, \ldots, L_m \in \mathbb{F}_q[A_{d-1}, \ldots, A_r]$ las formas lineales afines de (5.2). Sean $\boldsymbol{A} := (A_{d-1}, \ldots, A_1)$ y $\boldsymbol{A}_0 := (\boldsymbol{A}, A_0)$. Consideramos el polinomio $F \in \mathbb{F}_q[\boldsymbol{A}_0, T]$ definido como

$$F(\boldsymbol{A}_0, T) := T^d + A_{d-1}T^{d-1} + \cdots + A_1 T + A_0,$$

y la cuasi–$\mathbb{F}_q$–variedad afín $\Gamma_i \subset \mathbb{A}^{d+i}$ definida como sigue:

$$\begin{aligned}\Gamma_i := \{(\boldsymbol{a_0}, \boldsymbol{\alpha}) \in \mathbb{A}^d \times \mathbb{A}^i : F(\boldsymbol{a_0}, \alpha_j) = 0 \ (1 \leq j \leq i),\ \alpha_j \neq \alpha_k \ (1 \leq j < k \leq i),\\ L_1(\boldsymbol{a_0}) = \cdots = L_m(\boldsymbol{a_0}) = 0\}. \qquad (5.12)\end{aligned}$$

El siguiente resultado muestra cómo se relaciona el número $|\Gamma_i(\mathbb{F}_q)|$ con $\mathcal{S}_i^{\mathcal{A}}$.

Lema 5.2.2. *Sean i y r enteros positivos tales que $r+1 \leq i \leq d$. Entonces*

$$\frac{|\Gamma_i(\mathbb{F}_q)|}{i!} = \mathcal{S}_i^{\mathcal{A}}.$$

Demostración. Sea $(\boldsymbol{a_0}, \boldsymbol{\alpha})$ un punto de $\Gamma_i(\mathbb{F}_q)$ y sea $\sigma : \{1, \ldots, i\} \to \{1, \ldots, i\}$ una permutación. Sea $\sigma(\boldsymbol{\alpha})$ la imagen de $\boldsymbol{\alpha}$ por la función lineal que define la permutación σ. Es claro que $\big(\boldsymbol{a_0}, \sigma(\boldsymbol{\alpha})\big)$ pertenece también a $\Gamma_i(\mathbb{F}_q)$. Además, $\sigma(\boldsymbol{\alpha}) = \boldsymbol{\alpha}$ si y solo si σ es la permutación identidad. Esto muestra que $\mathbb{S}_i$, el grupo simétrico de i elementos, actúa sobre el conjunto $\Gamma_i(\mathbb{F}_q)$ vía la acción $* : \mathbb{S}_i \times \Gamma_i(\mathbb{F}_q) \hookrightarrow \Gamma_i(\mathbb{F}_q)$ definida por $*(\sigma, (\boldsymbol{a_0}, \boldsymbol{\alpha})) := (\boldsymbol{a_0}, \sigma(\boldsymbol{\alpha}))$, y cada órbita bajo esta acción tiene $i!$ elementos.

La órbita $\mathcal{O}(\boldsymbol{a_0}, \boldsymbol{\alpha})$ de un punto arbitrario $(\boldsymbol{a_0}, \boldsymbol{\alpha}) \in \Gamma_i(\mathbb{F}_q)$ determina unívocamente un polinomio $F(\boldsymbol{a_0}, T) = f + a_0$ con $f \in \mathcal{A}$ y un conjunto $\mathcal{X}_i := \{\alpha_1, \ldots, \alpha_i\} \subset \mathbb{F}_q$ con $|\mathcal{X}_i| = i$ tal que $(f + a_0)|_{\mathcal{X}_i} \equiv 0$. Por lo tanto, cada órbita determina unívocamente un conjunto $\mathcal{X}_i \subset \mathbb{F}_q$ con $|\mathcal{X}_i| = i$ y un elemento de $\mathcal{S}_{\mathcal{X}_i}^{\mathcal{A}}$. Recíprocamente, cada elemento de $\mathcal{S}_{\mathcal{X}_i}^{\mathcal{A}}$ corresponde a una única órbita de $\Gamma_i(\mathbb{F}_q)$. Esto implica que

$$\text{número de órbitas de } \Gamma_i(\mathbb{F}_q) = \sum_{\mathcal{X}_i \subseteq \mathbb{F}_q} |\mathcal{S}_{\mathcal{X}_i}^{\mathcal{A}}|. \qquad (5.13)$$

Por otro lado, como órbitas distintas resultan disjuntas y cada órbita $\mathcal{O}(\boldsymbol{a_0}, \boldsymbol{\alpha})$ tiene $i!$ elementos, tenemos que

$$|\Gamma_i(\mathbb{F}_q)| = \sum_{(\boldsymbol{a_0}, \boldsymbol{\alpha})} |\mathcal{O}(\boldsymbol{a_0}, \boldsymbol{\alpha})| = i! \cdot \text{ número de órbitas de } \Gamma_i(\mathbb{F}_q).$$

De (5.13) y (5.2.2) concluimos la demostración del lema. □

De acuerdo con el Lema 5.2.2, para dar una estimación del número $\mathcal{S}_i^{\mathcal{A}}$ basta con estimar $|\Gamma_i(\mathbb{F}_q)|$. Recordemos que estas cuasi–variedades ya fueron descriptas en la Sección 3.2. Para estimar dicha cantidad, vamos a obtener ecuaciones explícitas de

la clausura Zariski $\overline{\Gamma}_i$ de $\Gamma_i \subset \mathbb{A}^{d+i}$. Más precisamente, si $\Gamma_i^* \subset \mathbb{F}_q$ es la $\mathbb{F}_q$–variedad definida como

$$\Gamma_i^* := \{(\boldsymbol{a}_0, \boldsymbol{\alpha}) \in \mathbb{A}^d \times \mathbb{A}^i : \Delta^{j-1} F(\boldsymbol{a}_0, \alpha_1, \ldots, \alpha_j) = 0 \ (1 \leq j \leq i), \\ L_k(\boldsymbol{a_0}) = 0, \ (1 \leq k \leq m)\}, \tag{5.14}$$

donde $\Delta^{j-1} F(\boldsymbol{a}_0, T_1, \ldots, T_j)$ denota la diferencia dividida de orden $j-1$ de $F(\boldsymbol{a}_0, T) \in \overline{\mathbb{F}}_q[T]$ definida en (3.6), vamos a ver que $\overline{\Gamma}_i = \Gamma_i^*$.

Recordemos que en la Sección 3.2 estudiamos la relación que existe entre la cuasi–variedad Γ_i y la variedad $\Gamma_i^* \subset \mathbb{A}^{d+i}$. Además, suponiendo que $p > 2$, obtuvimos una serie de resultados sobre las características geométricas de la variedad Γ_i^* y su clausura proyectiva $\mathrm{pcl}(\Gamma_i^*) \subset \mathbb{P}^{d+i}$, que enunciamos a continuación.

Teorema 5.2.3. *Sea i un entero tal que $r+1 \leq i \leq d$. Entonces*

$$\Gamma_i = \Gamma_i^* \cap \{(\boldsymbol{a}_0, \boldsymbol{\alpha}) : \alpha_j \neq \alpha_k \ (1 \leq j < k \leq i)\}. \tag{5.15}$$

Si además $p > 2$, entonces

(1) $\Gamma_i^* \subset \mathbb{A}^{d+i}$ *es una intersección completa de dimensión $d-m$, cuyo lugar singular tiene codimensión al menos 2 en Γ_i^*.*

(2) $\mathrm{pcl}(\Gamma_i^*) \cap \{T_0 = 0\} \subset \mathbb{P}^{d+i-1}$ *es una unión finita de a lo sumo $i+1$ variedades lineales de $\mathbb{P}^{d+i-1}$ de dimensión $d-m-1$.*

(3) $\mathrm{pcl}(\Gamma_i^*) \subset \mathbb{P}^{d+i}$ *es una intersección completa normal de dimensión $d-m$ y grado $d!/(d-i)!$.*

Demostración. La identidad (5.15) es el contenido del Lema 3.2.1. Las propiedades de Γ_i^* se encuentran demostradas en el Teorema 3.2.10 y en el Corolario 3.2.11. Por último, las propiedades de la clausura proyectiva $\mathrm{pcl}(\Gamma_i^*)$ y de $\mathrm{pcl}(\Gamma_i^*) \cap \{T_0 = 0\}$ se encuentran demostradas en el Lema 3.2.13 y en el Teorema 3.2.15. □

Combinando el ítem (3) del Teorema 5.2.3 con el Teorema 2.1.10 concluimos que la clausura proyectiva $\mathrm{pcl}(\Gamma_i^*)$ es absolutamente irreducible de dimensión $d-m$ y grado $d!/(d-i)!$. Por el Teorema 2.1.8, Γ_i^* resulta una variedad absolutamente irreducible de dimensión $d-m$ y grado $d!/(d-i)!$. De (5.15) deducimos que Γ_i es un subconjunto abierto Zariski no vacío de Γ_i^*. Como Γ_i^* es absolutamente irreducible, la clausura Zariski $\overline{\Gamma}_i$ de Γ_i es Γ_i^*.

5.3. Una estimación del promedio

En esta sección damos una estimación del promedio del conjunto de valores $\mathcal{V}(\mathcal{A})$ de la familia lineal $\mathcal{A}$ definida en (5.3). De acuerdo al Teorema 5.2.1, tenemos que

$$\mathcal{V}(\mathcal{A}) = \sum_{i=1}^{r} (-1)^{i-1} \binom{q}{i} q^{1-i} + \frac{1}{q^{d-m-1}} \sum_{i=r+1}^{d} (-1)^{i-1} \sum_{\mathcal{X}_i \subset \mathbb{F}_q} |\mathcal{S}_{\mathcal{X}_i}^{\mathcal{A}}|,$$

donde $|\mathcal{S}^{\mathcal{A}}_{\mathcal{X}_i}|$ denota el número de polinomios de la forma $f+a_0$, con $f \in \mathcal{A}$ y $a_0 \in \mathbb{F}_q$, tales que $(f+a_0)(\alpha_j)=0$ para $1 \le j \le i$.

Sea $S_i^{\mathcal{A}} := \sum_{\mathcal{X}_i \subset \mathbb{F}_q} |S^{\mathcal{A}}_{\mathcal{X}_i}|$. De acuerdo al Lema 5.2.2 y a (5.15), tenemos, para cada $r+1 \le i \le d$, que

$$S_i^{\mathcal{A}} = \frac{|\Gamma_i(\mathbb{F}_q)|}{i!} = \frac{1}{i!}\left|\Gamma_i^*(\mathbb{F}_q) \setminus \bigcup_{j\neq k}\{T_j = T_k\}\right|, \tag{5.16}$$

donde Γ_i es la cuasi–variedad afín definida en (5.12) y Γ_i^* es la $\mathbb{F}_q$–variedad afín definida en (5.14). Como Γ_i^* y Γ_i cumplen las hipótesis de los Teoremas 3.2.16 y 3.2.17 respectivamente, deducimos la siguiente estimación:

$$\big||\Gamma_i(\mathbb{F}_q)| - q^{d-m}\big| \le (\delta_i(D_i-2)+2)q^{d-m-\frac{1}{2}} + \big(14D_i^2\delta_i^2 + i(i-1)\delta_i/2 + 2i\big)q^{d-m-1}, \tag{5.17}$$

donde $D_i := id - i(i+1)/2$ y $\delta_i := d!/(d-i)!$. De (5.16) y de (5.17) obtenemos la siguiente estimación para $\mathcal{S}_i^{\mathcal{A}}$.

Teorema 5.3.1. *Sea $p>2$ y $q>d$. Si $3 \le r \le d-m$ y $r+1 \le i \le d$, entonces*

$$\left|S_i^{\mathcal{A}} - \frac{q^{d-m}}{i!}\right| \le \frac{1}{i!}(\delta_i(D_i-2)+2)q^{d-m-\frac{1}{2}} + \frac{1}{i!}\big(14D_i^2\delta_i^2 + i(i-1)\delta_i/2 + 2i\big)q^{d-m-1},$$

donde $D_i := id - i(i+1)/2$ y $\delta_i := d!/(d-i)!$.

Combinando los Teoremas 5.2.1 y 5.3.1 obtenemos el siguiente resultado.

Corolario 5.3.2. *Con las hipótesis y las notaciones del Teorema 5.3.1,*

$$|\mathcal{V}(\mathcal{A}) - \mu_d\, q| \le d^2 2^{d-1} q^{1/2} + \frac{7}{2}\, d^4 \sum_{k=0}^{d-r-1} \binom{d}{k}^2 (d-k)!. \tag{5.18}$$

Demostración. Por el Teorema 5.2.1, tenemos que

$$\mathcal{V}(\mathcal{A}) - \mu_d\, q = \sum_{i=1}^{r}(-q)^{1-i}\left(\binom{q}{i} - \frac{q^i}{i!}\right) + \frac{1}{q^{d-m-1}} \sum_{i=r+1}^{d} (-1)^{i-1}\left(S_i^{\mathcal{A}} - \frac{q^{d-m}}{i!}\right). \tag{5.19}$$

Denotemos con $A(d,r)$ el primer término en el lado derecho de (5.19). En primer lugar, acotamos superiormente el valor absoluto de $A(d,r)$. Para ello, dados enteros positivos k, n con $k \le n$, denotamos como $\left[\begin{smallmatrix} n \\ k \end{smallmatrix}\right]$ al número de Stirling de primera clase, es decir, el número de permutaciones de n elementos con k ciclos disjuntos. Las siguientes son propiedades básicas de los números de Stirling (ver, por ejemplo, [FS09, §A.8]):

$$\begin{bmatrix} i \\ i \end{bmatrix} = 1, \ \begin{bmatrix} i \\ i-1 \end{bmatrix} = \binom{i}{2}, \ \sum_{k=0}^{i} \begin{bmatrix} i \\ k \end{bmatrix} = i!. \tag{5.20}$$

Teniendo en cuenta la identidad $\binom{q}{i} = \sum_{k=0}^{i} \frac{(-1)^{i-k}}{i!} \begin{bmatrix} i \\ k \end{bmatrix} q^k$, obtenemos

$$A(d,r) := \sum_{i=2}^{r} (-q)^{1-i} \left(\binom{q}{i} - \frac{q^i}{i!} \right) = \sum_{i=2}^{r} q^{1-i} \sum_{k=0}^{i-1} \frac{(-1)^{k+1}}{i!} \begin{bmatrix} i \\ k \end{bmatrix} q^k$$
$$= \sum_{i=0}^{r-2} \frac{(-1)^i}{2i!} + \sum_{i=2}^{r} q^{1-i} \sum_{k=0}^{i-2} \frac{(-1)^{k+1}}{i!} \begin{bmatrix} i \\ k \end{bmatrix} q^k.$$

El segundo término del lado derecho de esta expresión se puede acotar por

$$\sum_{k=0}^{i-2} \frac{1}{i!} \begin{bmatrix} i \\ k \end{bmatrix} q^k = \sum_{k=0}^{i-3} \frac{1}{i!} \begin{bmatrix} i \\ k \end{bmatrix} q^k + \frac{1}{i!} \begin{bmatrix} i \\ i-2 \end{bmatrix} q^{i-2} \leq q^{i-3} + \frac{8}{i^2} q^{i-2} \leq \left(\frac{1}{d} + \frac{8}{i^2} \right) q^{i-2}.$$

En consecuencia

$$\left| A(d,r) - \frac{1}{2e} \right| \leq \frac{1}{2 \cdot (r-1)!} + \sum_{i=2}^{r} \left(\frac{1}{d} + \frac{8}{i^2} \right) \frac{1}{q} \leq \frac{1}{2 \cdot (r-1)!} + \frac{7}{q}. \tag{5.21}$$

Consideramos ahora el valor absoluto de la segunda suma en el lado derecho de (5.19). Por el Teorema 5.3.1, tenemos que

$$B(d,r) := \frac{1}{q^{d-m-1}} \sum_{i=r+1}^{d} \left| S_i^{\mathcal{A}} - \frac{q^{d-m}}{i!} \right|$$
$$\leq q^{\frac{1}{2}} \sum_{i=r+1}^{d} \frac{\delta_i (D_i - 2) + 2}{i!} + 14 \sum_{i=r+1}^{d} \frac{D_i^2 \delta_i^2}{i!} + 2 \sum_{i=r+1}^{d} \frac{\delta_i}{2(i-2)!}.$$

Ahora bien, el primer término del lado derecho de esta última desigualdad se puede acotar de la siguiente manera:

$$\sum_{i=r+1}^{d} \frac{\delta_i (D_i - 2) + 2}{i!} \leq \sum_{i=r+1}^{d} \binom{d}{i} \frac{i(2d-1-i)}{2} \leq d^2 2^{d-1}.$$

Por otro lado,

$$\sum_{i=r+1}^{d} \frac{D_i^2 \delta_i^2}{i!} = \sum_{i=r+1}^{d} \binom{d}{i}^2 \frac{i^2 (2d-1-i)^2 \, i!}{4} \leq \frac{1}{64} (2d-1)^4 \sum_{k=0}^{d-4} \binom{d}{k}^2 (d-k)!.$$

Finalmente consideramos la última suma:

$$\sum_{i=r+1}^{d} \frac{\delta_i}{2(i-2)!} = \sum_{i=r+1}^{d} \binom{d}{i} \frac{i(i-1)}{2} = \sum_{k=0}^{d-r-1} \binom{d}{k} \frac{(d-k)!}{2\,(d-k-2)!}.$$

Por lo tanto, obtenemos que

$$B(d,r) \leq q^{1/2} d^2 2^{d-1} + \frac{1}{2} \sum_{k=0}^{d-r-1} \binom{d}{k} (d-k)! + \frac{7}{32} (2d-1)^4 \sum_{k=0}^{d-r-1} \binom{d}{k}^2 (d-k)!.$$

Combinando las cotas superiores para $|A(d,r)|$ y $B(d,r)$ se deduce la afirmación del corolario. □

Por último, analizamos el comportamiento del lado derecho de (5.18). Para ello, fijamos k con $0 \leq k \leq d-r-1$ y consideramos la función $h(k) := \binom{d}{k}^2 (d-k)!$. Analizando el signo de las diferencias $h(k+1) - h(k)$ para $0 \leq k \leq d-r-1$, mediante cálculos elementales deducimos la siguiente observación.

Observación 5.3.3. *Sea $k_0 := -1/2 + \sqrt{5+4d}/2$. Entonces h es, o bien una función creciente, o bien una función unimodal en el intervalo $[0, d-r-1]$, y alcanza su máximo en $\lfloor k_0 \rfloor$.*

A partir de la Observación 5.3.3 vemos que

$$\sum_{k=0}^{d-r-1} \binom{d}{k}^2 (d-k)! \leq (d-r) \binom{d}{\lfloor k_0 \rfloor}^2 (d-\lfloor k_0 \rfloor)! = \frac{(d-r)\,(d!)^2}{(d-\lfloor k_0 \rfloor)!\,(\lfloor k_0 \rfloor!)^2}. \tag{5.22}$$

Para obtener una cota superior del lado derecho de (5.22) utilizamos la fórmula de Stirling (ver, por ejemplo, [FS09, p. 747]): para $m \in \mathbb{N}$, existe θ con $0 \leq \theta < 1$ tal que

$$m! = (m/e)^m \sqrt{2\pi m}\, e^{\theta/12m}.$$

Aplicando esta fórmula vemos que existen θ_i $(i = 1, 2, 3)$ con $0 \leq \theta_i < 1$ tales que

$$C(d,r) := \frac{(d-r)\,(d!)^2}{(d-\lfloor k_0 \rfloor)!\,(\lfloor k_0 \rfloor!)^2} \leq \frac{(d-r)\,d^{2d+1} e^{-d+\lfloor k_0 \rfloor} e^{\frac{\theta_1}{6d} - \frac{\theta_2}{12(d-\lfloor k_0 \rfloor)} - \frac{\theta_3}{6\lfloor k_0 \rfloor}}}{\big(d-\lfloor k_0 \rfloor\big)^{d-\lfloor k_0 \rfloor} \sqrt{2\pi(d-\lfloor k_0 \rfloor)}\, \lfloor k_0 \rfloor^{2\lfloor k_0 \rfloor + 1}}.$$

De cálculos elementales obtenemos que

$$(d-\lfloor k_0 \rfloor)^{-d+\lfloor k_0 \rfloor} \leq d^{-d+\lfloor k_0 \rfloor} e^{\frac{\lfloor k_0 \rfloor (d-\lfloor k_0 \rfloor)}{d}},$$

$$\frac{d^{\lfloor k_0 \rfloor}}{\lfloor k_0 \rfloor^{2\lfloor k_0 \rfloor}} \leq e^{(d-\lfloor k_0 \rfloor^2)/\lfloor k_0 \rfloor}.$$

Luego,

$$C(d,r) \leq \frac{(d-r)\,d^{d+1} e^{2\lfloor k_0 \rfloor} e^{-\frac{\lfloor k_0 \rfloor^2}{d} + \frac{1}{6d} + \frac{d-\lfloor k_0 \rfloor^2}{\lfloor k_0 \rfloor}}}{\sqrt{2\pi} e^d \sqrt{d-\lfloor k_0 \rfloor}\, \lfloor k_0 \rfloor}.$$

De acuerdo a la definición de $\lfloor k_0 \rfloor$, es fácil ver que $d/\lfloor k_0 \rfloor \sqrt{d-\lfloor k_0 \rfloor} \leq 5/2$ y que $2\lfloor k_0 \rfloor \leq -1 + \sqrt{5+4d} \leq -1/5 + 2\sqrt{d}$. Por lo tanto, teniendo en cuenta que $d \geq 4$, concluimos que

$$C(d,r) \leq \frac{5}{2} \frac{e^{\frac{109}{30}} (d-r)\,d^d\, e^{2\sqrt{d}}}{\sqrt{2\pi} e^d}.$$

Combinando esta cota con el Corolario 5.3.2 obtenemos la siguiente estimación.

Teorema 5.3.4. *Bajo las hipótesis del Teorema 5.3.1, tenemos que*

$$|\mathcal{V}(\mathcal{A}) - \mu_d\, q| \leq d^2 2^{d-1} q^{1/2} + 133\, d^{d+5} e^{2\sqrt{d}-d}.$$

Este resultado constituye una mejora de (5.1) en varios aspectos. El primero y más importante es que las hipótesis sobre la familia lineal $\mathcal{A}$ que consideramos son relativamente generales y simples de verificar. Por otro lado, nuestro resultado es válido para $p > 2$, mientras que (5.1) requiere que p sea suficientemente grande. Finalmente, damos una expresión explícita para la constante que subyace en la notación $\mathcal{O}$ en (5.1) con buen comportamiento.

5.3.1. Polinomios con coeficientes prescriptos

En esta sección discutimos brevemente qué estimación obtenemos aplicando el Teorema 5.3.4 a la familia $\mathcal{A}_{\boldsymbol{a}}$ de elementos de $\mathbb{F}_q[T]_d$ cuyos primeros s coeficientes consecutivos están prefijados, que definimos en (5.5). Observemos que dicha familia es un caso particular de la familia $\mathcal{A}$ de (5.3), tomando para este caso $r := d - s$ y $m := s$ y las formas lineales $L_k := A_{d-k} - a_{d-k}$ con $1 \leq k \leq s$. Del Teorema 5.3.4 deducimos el siguiente resultado, que da una estimación del valor promedio $\mathcal{V}(d, s)$ definido en (5.6).

Teorema 5.3.5. *Sea $p > 2$, $q > d$ y $1 \leq s \leq d - 3$. Entonces*

$$|\mathcal{V}(d, s) - \mu_d\, q| \leq d^2 2^{d-1} q^{1/2} + 133\, d^{d+5} e^{2\sqrt{d}-d}. \tag{5.23}$$

Observemos que esta estimación mejora la de (5.6). Por un lado, nuestra estimación vale para $p > 2$, mientras que (5.6) vale para cuerpos de característica más grande. Además, proporcionamos una expresión explícita del error, cuestión de importancia para el análisis de los algoritmos que estudiaremos en los últimos capítulos.

5.4. Conjunto de valores para familias no lineales

En esta sección esbozamos de qué manera las técnicas precedentes pueden extenderse a fin de abordar el estudio del comportamiento promedio del cardinal del conjunto de valores para familias no lineales, es decir, familias de polinomios cuyos coeficientes pertenecen a una cierta variedad algebraica. Cabe mencionar que no existen resultados para tales familias. Dado que esta extensión nos desviaría del hilo argumental de esta tesis, no vamos a discutirla en detalle; una descripción completa de la misma puede encontrarse en [MPP16a].

Comenzamos precisando las familias no lineales a la que nos referimos. Sean m y d enteros no negativos tales que $q > d \geq m + 2$, sean $A_{d-1}, \ldots, A_0$ indeterminadas sobre $\overline{\mathbb{F}}_q$, $\boldsymbol{A}_0 := (A_{d-1}, \ldots, A_0)$ y sean $G_1, \ldots, G_m \in \mathbb{F}_q[A_{d-1}, \ldots, A_0]$ polinomios de grados $d_1, \ldots, d_m$ respectivamente. Sea $\boldsymbol{G} := (G_1, \ldots, G_m)$ y consideremos la familia

$$\mathcal{A}_{\boldsymbol{G}} := \left\{ T^d + \sum_{j=0}^{d-1} a_j T^j \in \mathbb{F}_q[T] : G_i(a_{d-1}, \ldots, a_0) = 0 \ (1 \leq i \leq m) \right\}.$$

Denotamos por $\mathcal{V}(\mathcal{A}_{\boldsymbol{G}})$ el promedio del cardinal del conjunto de valores posibles $\mathcal{V}(f)$ cuando f recorre todos los elementos de la familia $\mathcal{A}_{\boldsymbol{G}}$, es decir,

$$\mathcal{V}(\mathcal{A}_{\boldsymbol{G}}) := \frac{1}{|\mathcal{A}_{\boldsymbol{G}}|} \sum_{f \in \mathcal{A}_{\boldsymbol{G}}} \mathcal{V}(f).$$

El objetivo es demostrar que, bajo hipótesis generales sobre $G_1, \ldots, G_m$, el comportamiento asintótico del promedio $\mathcal{V}(\mathcal{A}_{\boldsymbol{G}})$ es el del caso lineal, es decir, $\mathcal{V}(\mathcal{A}_{\boldsymbol{G}}) = \mu_d q + \mathcal{O}(q^{1/2})$.

Al igual que para familias lineales, traducimos este problema al de determinar la cantidad de puntos $\mathbb{F}_q$-racionales de una intersección completa cuyo lugar singular tiene codimensión al menos 2. Para ello, comenzamos observando que dados m, d enteros no negativos con $q > d \geq m+2$, podemos suponer sin pérdida de generalidad que $G_1, \ldots, G_m$ son elementos de $\mathbb{F}_q[A_{d-1}, \ldots, A_1]$. En efecto, sea $\Pi : \mathcal{A}_{\boldsymbol{G}} \to \mathbb{F}_q$ definida por $\Pi(T^d + a_{d-1}T^{d-1} + \cdots + a_0) := a_0$. Denotamos con $\mathcal{A}_{\boldsymbol{G},a_0} := \Pi^{-1}(a_0)$. Tenemos que

$$\frac{1}{|\mathcal{A}_{\boldsymbol{G}}|}\sum_{f\in\mathcal{A}_{\boldsymbol{G}}}\mathcal{V}(f) - \mu_d q = \frac{1}{\sum\limits_{a_0\in\mathbb{F}_q}|\mathcal{A}_{\boldsymbol{G},a_0}|}\sum_{a_0\in\mathbb{F}_q}|\mathcal{A}_{\boldsymbol{G},a_0}|\left(\frac{1}{|\mathcal{A}_{\boldsymbol{G},a_0}|}\sum_{f\in\mathcal{A}_{\boldsymbol{G},a_0}}\mathcal{V}(f) - \mu_d q\right).$$

En consecuencia, si existe una constante $E(d_1, \ldots, d_m, d)$ tal que se satisface

$$\left|\frac{1}{|\mathcal{A}_{\boldsymbol{G},a_0}|}\sum_{f\in\mathcal{A}_{\boldsymbol{G},a_0}}\mathcal{V}(f) - \mu_d q\right| \leq E(d_1, \ldots, d_m, d)q^{\frac{1}{2}}$$

para todo $a_0 \in \mathbb{F}_q$, entonces podemos concluir que

$$\begin{aligned}\left|\frac{1}{|\mathcal{A}_{\boldsymbol{G}}|}\sum_{f\in\mathcal{A}_{\boldsymbol{G}}}\mathcal{V}(f) - \mu_d q\right| &\leq \frac{1}{\sum\limits_{a_0\in\mathbb{F}_q}|\mathcal{A}_{\boldsymbol{G},a_0}|}\sum_{a_0\in\mathbb{F}_q}|\mathcal{A}_{\boldsymbol{G},a_0}|E(d_1, \ldots, d_m, d)q^{\frac{1}{2}}\\ &\leq E(d_1, \ldots, d_m, d)q^{\frac{1}{2}}.\end{aligned}$$

Además, al igual que antes, como $\mathcal{V}(f) = \mathcal{V}(f + a_0)$ para todo $f \in \mathcal{A}_{\boldsymbol{G}}$, podemos suponer sin pérdida de generalidad que $f(0) = 0$ para todo $f \in \mathcal{A}_{\boldsymbol{G}}$.

Luego, obtenemos la siguiente expresión combinatoria para el valor promedio $\mathcal{V}(\mathcal{A}_{\boldsymbol{G}})$ en términos del número $\mathcal{S}_i^{\mathcal{A}_{\boldsymbol{G}}}$ de ciertos conjuntos interpolantes con $1 \leq i \leq d$.

Lema 5.4.1 ([MPP16a, Lemma 1.1])**.** *Dados d, m enteros no negativos tales que $q > d \geq m + 2$, tenemos que*

$$\mathcal{V}(\mathcal{A}_{\boldsymbol{G}}) = \frac{1}{|\mathcal{A}_{\boldsymbol{G}}|}\sum_{i=1}^{d}(-1)^{i-1}\mathcal{S}_i^{\mathcal{A}_{\boldsymbol{G}}},$$

donde $\mathcal{S}_i^{\mathcal{A}_{\boldsymbol{G}}} := \big|\{(\mathcal{X}_i, f) : \mathcal{X}_i \subset \mathbb{F}_q,\ f \in \mathcal{A}_{\boldsymbol{G}},\ f(x) = 0$ para todo $x \in \mathcal{X}_i\}\big|$.

Demostración. La demostración sale con argumentos similares a los del Teorema 5.2.1. □

Por lo tanto, para determinar el comportamiento asintótico de $\mathcal{V}(\mathcal{A}_{\boldsymbol{G}})$ basta con estimar el número $\mathcal{S}_i^{\mathcal{A}_{\boldsymbol{G}}}$. Para esto, traducimos este problema en el de estimar la cantidad de puntos $\mathbb{F}_q$–racionales con coordenadas distintas dos a dos de cierta variedad de incidencia. Más precisamente, fijamos i con $1 \leq i \leq d$, consideramos el polinomio $F \in \mathbb{F}_q[\boldsymbol{A}_0, T]$ definido como

$$F(\boldsymbol{A}_0, T) := T^d + A_{d-1}T^{d-1} + \cdots + A_1T + A_0,$$

y la cuasi–$\mathbb{F}_q$– variedad afín $\Gamma_i^{\boldsymbol{G}} \subset \mathbb{A}^{d+i}$ definida por

$$\Gamma_i^{\boldsymbol{G}} := \{(\boldsymbol{a}_0, \boldsymbol{\alpha}) \in \mathbb{A}^d \times \mathbb{A}^i : \ \alpha_j \neq \alpha_k \ (1 \leq j < k \leq i), \\ F(\boldsymbol{a}_0, \alpha_i) = 0 \ (1 \leq i \leq r), \ G_k(\boldsymbol{a}_0) = 0 \, (1 \leq k \leq m)\}.$$

Al igual que en el caso de familias lineales obtenemos el siguiente resultado, que muestra la relación que existe entre $|\Gamma_i^{\boldsymbol{G}}(\mathbb{F}_q)|$ y $\mathcal{S}_i^{\mathcal{A}_{\boldsymbol{G}}}$.

Lema 5.4.2. *Sean i tal que $1 \leq i \leq d$. Entonces*

$$\frac{|\Gamma_i^{\boldsymbol{G}}(\mathbb{F}_q)|}{i!} = \mathcal{S}_i^{\mathcal{A}_{\boldsymbol{G}}}.$$

Demostración. La demostración sale de la misma manera que la del Lema 5.2.2. □

De acuerdo a este lema, al igual que en el caso lineal, para estimar el número $\mathcal{S}_i^{\mathcal{A}_{\boldsymbol{G}}}$ vamos a estimar $|\Gamma_i^{\boldsymbol{G}}(\mathbb{F}_q)|$. Para ello, consideramos la clausura Zariski $\overline{\Gamma}_i^{\boldsymbol{G}}$ de $\Gamma_i^{\boldsymbol{G}} \subset \mathbb{A}^{d+i}$ y damos ecuaciones explícitas que definen dicha clausura. Más precisamente, si $\Gamma_i^{\boldsymbol{G},*} \subset \mathbb{A}^{d+i}$ es la $\mathbb{F}_q$–variedad definida como

$$\Gamma_i^{\boldsymbol{G},*} := \{(\boldsymbol{a}_0, \boldsymbol{\alpha}) \in \mathbb{A}^d \times \mathbb{A}^i : \ \Delta^{j-1}F(\boldsymbol{a}_0, \alpha_1, \ldots, \alpha_j) = 0 \ (1 \leq j \leq i), \\ G_k(\boldsymbol{a}_0) = 0 \ (1 \leq k \leq m)\},$$

donde $\Delta^{j-1}F(\boldsymbol{a}_0, T_1, \ldots, T_j)$ denota la diferencia dividida de orden $j-1$ del polinomio $F(\boldsymbol{a}_0, T) \in \overline{\mathbb{F}}_q[T]_d$ que aparece en (5.14), demostramos que $\overline{\Gamma}_i^{\boldsymbol{G}} = \Gamma_i^{\boldsymbol{G},*}$. Con este objetivo, al igual que en el caso lineal, tenemos la siguiente relación entre $\Gamma_i^{\boldsymbol{G}}$ y $\Gamma_i^{\boldsymbol{G},*}$.

Lema 5.4.3 ([MPP16a, Lemma 1.3])**.** *Sea i un entero tal que $1 \leq i \leq d$. Entonces*

$$\Gamma_i^{\boldsymbol{G}} = \Gamma_i^{\boldsymbol{G},*} \cap \{(\boldsymbol{a}_0, \boldsymbol{\alpha}) : \alpha_j \neq \alpha_k \ (1 \leq j < k \leq i)\}. \tag{5.24}$$

Demostración. La demostración sale con argumentos similares a los del Lema 3.2.1. □

El siguiente objetivo es dar una serie de características geométricas de la variedad $\Gamma_i^{\boldsymbol{G},*}$ y su clausura proyectiva $\mathrm{pcl}(\Gamma_i^{\boldsymbol{G},*}) \subset \mathbb{P}^{d+i}$. Una de las características que estudiamos es el lugar singular de $\Gamma_i^{\boldsymbol{G},*}$. Para ello, consideramos la familia $\mathcal{B}_{\boldsymbol{G}} \subset \overline{\mathbb{F}}_q[T]_d$ que consiste de todos los polinomios $f := T^d + a_{d-1}T^{d-1} + \cdots + a_0 \in \overline{\mathbb{F}}_q[T]_d$ tales que $G_k(a_{d-1}, \ldots, a_1) = 0$ para $1 \leq k \leq m$. Observamos que un punto singular de $\Gamma_i^{\boldsymbol{G},*}$ proviene de un punto singular de la variedad $V \subset \mathbb{A}^d$ definida por los polinomios $G_1, \ldots, G_m$ o de un polinomio $f \in \mathcal{B}_{\boldsymbol{G}}$ que no es libre de cuadrados. Las primeras tres condiciones que requerimos sobre los polinomios $G_1, \ldots, G_m$ garantizan que V tiene "buen comportamiento" geométrico, es decir, se trata de una intersección completa cuyo lugar singular tiene codimensión al menos 2:

(C_1) $G_1, \ldots, G_m$ forman una sucesión regular y el ideal que generan en $\mathbb{F}_q[A_{d-1}, \ldots, A_0]$ es radical.

(C_2) La variedad $V \subset \mathbb{A}^d$ definida por $G_1, \ldots, G_m$ es normal.

(C_3) Sean $G_1^{d_1}, \ldots, G_m^{d_m}$ las partes homogéneas de mayor grado de $G_1, \ldots, G_m$ respectivamente. Entonces $G_1^{d_1}, \ldots, G_m^{d_m}$ satisfacen (C_1) y(C_2).

A fin de controlar la cantidad de polinomios de $\mathcal{B}_{\boldsymbol{G}}$ que no son libres de cuadrados, establecemos una condición que asegura que el lugar discriminante y el lugar del primer subdiscriminante de la familia $\mathcal{B}_{\boldsymbol{G}}$, que definimos a continuación, cortan bien a la variedad V. Identificando cada elemento $f_{\boldsymbol{a}_0} := T^d + a_{d-1}T^{d-1} + \cdots + a_0 \in \mathcal{B}_{\boldsymbol{G}}$ con la d–upla $\boldsymbol{a}_0 := (a_{d-1}, \ldots, a_0) \in \mathbb{A}^d$, recordamos que el lugar discriminante $\mathcal{D}(\mathcal{B}_{\boldsymbol{G}})$ de $\mathcal{B}_{\boldsymbol{G}}$ es el conjunto de todos los elementos de $f_{\boldsymbol{a}_0} \in \mathcal{B}_{\boldsymbol{G}}$ para los cuales $\mathrm{Disc}(f_{\boldsymbol{a}_0}) = 0$, donde $\mathrm{Disc}(f_{\boldsymbol{a}_0}) := \mathrm{Res}(f_{\boldsymbol{a}_0}, f'_{\boldsymbol{a}_0})$ denota el discriminante de $f_{\boldsymbol{a}_0}$. Como el polinomio $f_{\boldsymbol{a}_0}$ tiene grado d, propiedades básicas de las resultantes aseguran que $\mathrm{Disc}(f_{\boldsymbol{a}_0}) = \mathrm{Disc}(F(\boldsymbol{A}_0, T))|_{\boldsymbol{A}_0=\boldsymbol{a}_0} = 0$. Así, el lugar discriminante $\mathcal{D}(\mathcal{B}_{\boldsymbol{G}})$ es el conjunto de todos los elementos de $\boldsymbol{a}_0 \in \mathcal{B}_{\boldsymbol{G}}$ tales que $\mathrm{Disc}(F(\boldsymbol{A}_0, T))|_{\boldsymbol{A}_0=\boldsymbol{a}_0} = 0$.

De la misma manera, definimos el lugar del primer subdiscriminante $\mathcal{S}_1(\mathcal{B}_{\boldsymbol{G}})$ de $\mathcal{B}_{\boldsymbol{G}}$ como el conjunto de todos los elementos $\boldsymbol{f}_{a_0} \in \mathcal{B}_{\boldsymbol{G}}$ para los cuales $\mathrm{Subdisc}(f_{\boldsymbol{a}_0}) := \mathrm{Subres}(f_{\boldsymbol{a}_0}, f'_{\boldsymbol{a}_0}) = 0$, donde $\mathrm{Subres}(f_{\boldsymbol{a}_0}, f'_{\boldsymbol{a}_0})$ denota la primera subresultante de $f_{\boldsymbol{a}_0}$ y su derivada. Como $f_{\boldsymbol{a}_0}$ tiene grado d, por propiedades básicas de las subresultantes deducimos que $\mathrm{Subdisc}(f_{\boldsymbol{a}_0}) := \mathrm{Subdisc}(F(\boldsymbol{A}_0, T))|_{\boldsymbol{A}_0=\boldsymbol{a}_0} = 0$. Por lo tanto, también podemos definir el lugar del primer subdiscriminante $\mathcal{S}_1(\mathcal{B}_{\boldsymbol{G}})$ como el conjunto de todos los elementos $\boldsymbol{a}_0 \in \mathcal{B}_{\boldsymbol{G}}$ tales que $\mathrm{Subdisc}(F(\boldsymbol{A}_0, T))|_{\boldsymbol{A}_0=\boldsymbol{a}_0} = 0$. En estos términos, vamos a requerir la siguiente condición adicional:

(C_4) $\mathcal{D}(\mathcal{B}_{\boldsymbol{G}})$ tiene codimensión uno en $\mathcal{B}_{\boldsymbol{G}}$, y $\mathcal{D}(\mathcal{B}_{\boldsymbol{G}}) \cap \mathcal{S}_1(\mathcal{B}_{\boldsymbol{G}})$ tiene codimensión dos en $\mathcal{B}_{\boldsymbol{G}}$.

Estas condiciones nos permiten dar la siguiente serie de resultados sobre las características geométricas de $\Gamma_i^{\boldsymbol{G},*}$ y su clausura proyectiva $\mathrm{pcl}(\Gamma_i^{\boldsymbol{G},*})$.

Teorema 5.4.4 ([MPP16a, Theorem 1.2, Corollary 1.2, Lemma 1.13 y Theorem 1.3])**.** *Sea i un entero tal que $1 \leq i \leq d$. Si $q > d \geq m+2$, entonces*

(1) $\Gamma_i^{\boldsymbol{G},*}$ es una intersección completa de dimensión $d-m$, cuyo lugar singular tiene codimensión al menos 2 en $\Gamma_i^{\boldsymbol{G},*}$.

(2) $\mathrm{pcl}(\Gamma_i^{\boldsymbol{G},*}) \cap \{T_0 = 0\} \subset \mathbb{P}^{d+i-1}$ está contenida en una unión de $i+1$ intersecciones completas normales definidas sobre $\mathbb{F}_q$, cada una de dimensión $d-m-1$ y grado $\prod_{i=1}^m d_i$.

(3) $\mathrm{pcl}(\Gamma_i^{\boldsymbol{G},*}) \subset \mathbb{P}^{d+i}$ es una intersección completa normal de dimensión $d-m$ y grado $\prod_{i=1}^m d_i \cdot \frac{d!}{(d-i)!}$.

Demostración. Las condiciones (C_1), (C_2) y (C_4) y argumentos similares a los del Teorema 3.2.10 y Corolario 3.2.11 prueban las propiedades geométricas de $\Gamma_i^{\boldsymbol{G},*}$. La condición ($\mathsf{C}_3$) y argumentos similares a los del Lema 3.2.13 muestran las propiedades de $\mathrm{pcl}(\Gamma_i^{\boldsymbol{G},*}) \cap \{T_0 = 0\}$. Por último, por argumentos similares a los del Teorema 3.2.15 deducimos las propiedades de $\mathrm{pcl}(\Gamma_i^{\boldsymbol{G},*})$. □

Combinando el ítem (3) del Teorema 5.4.4 con el Teorema 2.1.10 concluimos que la clausura proyectiva $\mathrm{pcl}(\Gamma_i^{\boldsymbol{G},*})$ es absolutamente irreducible de dimensión $d-m$. Por el Teorema 2.1.8, $\Gamma_i^{\boldsymbol{G},*}$ resulta una variedad absolutamente irreducible de dimensión $d-m$. De (5.24) deducimos que $\Gamma_i^{\boldsymbol{G}}$ es un subconjunto abierto Zariski no vacío de $\Gamma_i^{\boldsymbol{G},*}$. Como $\Gamma_i^{\boldsymbol{G},*}$ es absolutamente irreducible, la clausura Zariski $\overline{\Gamma}_i^{\boldsymbol{G}}$ de $\Gamma_i^{\boldsymbol{G}}$ es $\Gamma_i^{\boldsymbol{G},*}$.

Las condiciones $(\mathsf{C_1})$, $(\mathsf{C_2})$ y $(\mathsf{C_3})$, a su vez, nos permiten estimar el número de elementos de la familia no lineal $\mathcal{A}_{\boldsymbol{G}}$.

Lema 5.4.5 ([MPP16a, Lemma 1.16]). *Para $q > 16(D_V\delta_V+14D_V^2\delta_V^2q^{-\frac{1}{2}})^2$, tenemos*

$$\frac{1}{2}q^{d-m-1} < |\mathcal{A}_{\boldsymbol{G}}| \leq q^{d-m-1} + 2\big(\delta_V(D_V-2)+2+14D_V^2\delta_V^2q^{-\frac{1}{2}}\big)q^{d-m-\frac{3}{2}},$$

donde $\delta_V := \prod_{i=1}^m d_i$ y $D_V := \sum_{i=1}^m d_i - 1$.

Demostración. Las condiciones $(\mathsf{C_1})$, $(\mathsf{C_2})$ y $(\mathsf{C_3})$ implican que la clausura proyectiva $\mathrm{pcl}(V)$ de V y el conjunto $\mathrm{pcl}(V)^\infty := \mathrm{pcl}(V) \cap \{T_0 = 0\}$ de puntos en el infinito son intersecciones completas normales definidas sobre $\mathbb{F}_q$, ambas de grado $\prod_{i=1}^m d_i$, en $\mathbb{P}^{d-1}$ y $\{T_0 = 0\} \cong \mathbb{P}^{d-2}$ respectivamente. Por lo tanto, por (2.7) se sigue que

$$\begin{aligned}\big||\mathcal{A}_{\boldsymbol{G}}| - q^{d-m-1}\big| =& \big||\mathrm{pcl}(V)(\mathbb{F}_q)| - |\mathrm{pcl}(V)^\infty(\mathbb{F}_q)| - p_{d-m-1} + p_{d-m-2}\big| \\ \leq& \big||\mathrm{pcl}(V)(\mathbb{F}_q)| - p_{d-m-1}\big| + \big||\mathrm{pcl}(V(\mathbb{F}_q))^\infty| - p_{d-m-2}\big| \\ \leq& \left(\delta_V(D_V-2)+2\right)(q+1)q^{d-m-\frac{5}{2}} + 14D_V^2\delta_V^2(q+1)q^{d-m-3} \\ \leq& 2\big(\delta_V(D_V-2)+2+14D_V^2\delta_V^2q^{-\frac{1}{2}}\big)q^{d-m-\frac{3}{2}}.\end{aligned}$$

De la hipótesis sobre q, se sigue el lema. □

Por último, por el Lema 5.4.3 deducimos que, al igual que en el caso lineal, estimar $|\Gamma_i^{\boldsymbol{G}}(\mathbb{F}_q)|$ equivale a estimar la cantidad de puntos $\mathbb{F}_q$–racionales con coordenadas distintas dos a dos de $\Gamma_i^{\boldsymbol{G},*}$. Asimismo, las características geométricas de $\Gamma_i^{\boldsymbol{G},*}$ y de su clausura proyectiva, que se encuentran en el Teorema 5.4.4, nos permiten aplicar la estimación para intersecciones completas proyectivas normales del Teorema 2.2.10 a fin de estimar la cantidad de puntos $\mathbb{F}_q$–racionales de $\Gamma_i^{\boldsymbol{G},*}$ con coordenadas distintas dos a dos. De los Lemas 5.4.2 y 5.4.3 obtenemos el siguiente resultado.

Teorema 5.4.6 ([MPP16a, Theorem 1.4]). *Sea $q > d \geq m+2$. Para i tal que $1 \leq i \leq d$, tenemos que*

$$\left|S_i^{\mathcal{A}_{\boldsymbol{G}}} - \frac{q^{d-m}}{i!}\right| \leq \left(\frac{\delta_i(D_i-2)+2}{i!}\,q^{\frac{1}{2}} + \left(14\frac{D_i^2\delta_i^2}{i!} + \binom{i}{2}\frac{\delta_i}{i!} + \frac{4i}{i!}\delta_V\right)\right)q^{d-m-1},$$

donde $\delta_V := \prod_{i=1}^m d_i$, $D_V := \sum_{i=1}^m (d_i-1)$, $\delta_i := \delta_V \frac{d1}{(d-i)!}$ y $D_i := D_V + id - \frac{i(i+1)}{2}$.

Esta última estimación junto con los Lemas 5.4.1 y 5.4.5 nos permite determinar el comportamiento asintótico de $\mathcal{V}(\mathcal{A}_{\boldsymbol{G}})$, como afirmamos en el siguiente resultado.

Teorema 5.4.7 ([MPP16a, Theorem 1.5]). *Para $q > \max\{d, 16(D_V\delta_V+14D_V^2\delta_V^2q^{-\frac{1}{2}})^2\}$ y $d \geq m+2$,*

$$|\mathcal{V}(\mathcal{A}_{\boldsymbol{G}}) - \mu_d q| \leq 2^d\delta(3D_V+d^2)q^{1/2} + 67\delta^2(D_V+2)^2d^{d+5}e^{2\sqrt{d}-d}.$$

5.5. El segundo momento del conjunto de valores

En esta sección mostramos cómo el enfoque que desarrollamos en este capítulo nos permite también dar una estimación explícita del promedio del segundo momento del cardinal del conjunto de valores de la familia $\mathcal{A}_{\boldsymbol{a}}$ definida en (5.5).

Sea $p > 2$ y sean d y s enteros positivos tales que $q > d$ y $1 \leq s \leq d-3$. Fijamos $\boldsymbol{a} := (a_{d-1}, \ldots, a_{d-s}) \in \mathbb{F}_q^s$ y sea $f_{\boldsymbol{a}} := T^d + a_{d-1}T^{d-1} + \cdots + a_{d-s}T^{d-s}$. Recordemos que los elementos de $\mathcal{A}_{\boldsymbol{a}}$ son todos los polinomios de la forma $f_{\boldsymbol{b}} := f_{\boldsymbol{a}} + b_{d-s-1}T^{d-s-1} + \cdots + b_1 T$ con $\boldsymbol{b} := (b_{d-s-1}, \ldots . b_1) \in \mathbb{F}_q^{d-s-1}$.

Definimos el promedio del segundo momento de la familia $\mathcal{A}_{\boldsymbol{a}}$ como

$$\mathcal{V}_2(d,s) := \mathcal{V}_2(\mathcal{A}_{\boldsymbol{a}}) := \frac{1}{q^{d-s-1}} \sum_{\boldsymbol{b} \in \mathbb{F}_q^{d-s-1}} \mathcal{V}(f_{\boldsymbol{b}})^2.$$

En esta sección estimamos de $\mathcal{V}_2(d,s)$ y

$$\mathcal{V}_2(d,0) := q^{1-d} \sum_{\boldsymbol{b} \in \mathbb{F}_q^{d-1}} \mathcal{V}(f_{\boldsymbol{b}})^2.$$

Esta última cantidad corresponde al promedio del segundo momento del cardinal del conjunto de valores $\mathcal{V}(f_{\boldsymbol{b}})$ cuando $f_{\boldsymbol{b}}$ recorre todos los polinomios de $\mathbb{F}_q[T]_d$ tales que $f(0) = 0$. Como en la sección anterior, solo vamos a esbozar el enfoque; una descripción detallada del mismo puede verse en [MPP14].

En la literatura encontramos una expresión explícita de $\mathcal{V}_2(d,0)$ para el caso $d \geq q$ (ver [KK90b]). Por otro lado, en un trabajo de Uchiyama [Uch56] se muestra que, suponiendo la validez de la hipótesis de Riemann para funciones L, si $p > d$, entonces

$$\mathcal{V}_2(d,0) := \frac{1}{q^{d-1}} \sum \mathcal{V}(f)^2 = \mu_d^2 q^2 + \mathcal{O}(q). \tag{5.25}$$

Observemos que Uchiyama no da una expresión explícita para la constante que subyace en la notación $\mathcal{O}$. Nosotros, en cambio, damos una expresión explícita para (5.25) que vale para cuerpos de característica pequeña y es independiente de la validez de la hipótesis de Riemann mencionada.

Como en las secciones anteriores, traducimos este problema al de determinar la cantidad de puntos $\mathbb{F}_q$–racionales con coordenadas distintas dos a dos de una intersección completa cuyo lugar singular tiene codimensión al menos 2. Empezamos dando la siguiente expresión combinatoria para $\mathcal{V}_2(d,s)$ en términos del número $\mathcal{S}_{m,n}^{\boldsymbol{a}}$ de ciertos conjuntos interpolantes con $d-s+1 \leq m+n \leq 2d$. La demostración sigue argumentos similares a los del Teorema 5.2.1.

Teorema 5.5.1. *Bajo los supuestos de arriba, tenemos que*

$$\begin{aligned} \mathcal{V}_2(d,s) = \mathcal{V}(d,s) \quad &+ \sum_{\substack{1 \leq m,n \leq d \\ 2 \leq m+n \leq d-s}} (-1)^{m+n} \binom{q}{m} \binom{q}{n} q^{2-n-m} \\ &+ \frac{1}{q^{d-s-1}} \sum_{\substack{1 \leq m,n \leq d \\ d-s+1 \leq m+n \leq 2d}} (-1)^{m+n} \sum_{\substack{\Gamma_1, \Gamma_2 \subset \mathbb{F}_q \\ |\Gamma_1| = m, |\Gamma_2| = n}} \left| S_{\Gamma_1,\Gamma_2}^{\boldsymbol{a}} \right|, \end{aligned}$$

donde $\mathcal{S}^{\boldsymbol{a}}_{\Gamma_1,\Gamma_2}$ *es el conjunto que consiste de los puntos* $(\boldsymbol{b}, b_{0,1}, b_{0,2}) \in \mathbb{F}_q^{d-s+1}$ *con* $b_{0,1} \neq b_{0,2}$ *tales que se verifica* $(f_{\boldsymbol{b}} + b_{0,1})\big|_{\Gamma_1} \equiv 0$ *y* $(f_{\boldsymbol{b}} + b_{0,2})\big|_{\Gamma_2} \equiv 0$.

Fijamos s, d y $\boldsymbol{a}$ como en el Teorema 5.5.1. Para determinar el comportamiento de $\mathcal{V}_2(d,s)$ necesitamos estimar el número

$$\mathcal{S}^{\boldsymbol{a}}_{m,n} := \sum_{\substack{\Gamma_1,\,\Gamma_2 \subset \mathbb{F}_q \\ |\Gamma_1|=m, |\Gamma_2|=n}} |\mathcal{S}^{\boldsymbol{a}}_{\Gamma_1,\,\Gamma_2}| \tag{5.26}$$

para cada par (m,n) con $1 \leq m,n \leq d$ y $d-s+1 \leq m+n \leq 2d$. Con este propósito, expresamos el número $\mathcal{S}^{\boldsymbol{a}}_{m,n}$ en términos de la cantidad de puntos $\mathbb{F}_q$–racionales con coordenadas distintas dos a dos de una cierta variedad de incidencia $\Gamma^*_{m,n} \subset \mathbb{A}^{d-s+1+m+n}$. Más precisamente, fijamos enteros positivos m y n con $1 \leq m,n \leq d$ y $d-s+1 \leq m+n \leq 2d$. Introducimos nuevas variables $T, T_1, \ldots, T_m$, $U, U_1, \ldots, U_n$, $B, B_{d-s-1}, \ldots, B_1$, $B_{0,1}$, $B_{0,2}$ sobre $\overline{\mathbb{F}}_q$ y denotamos con $\boldsymbol{T} := (T_1, \ldots, T_m)$, $\boldsymbol{U} := (U_1, \ldots, U_n)$, $\boldsymbol{B} := (B_{d-s-1}, \ldots, B_1)$, $\boldsymbol{B}_1 := (\boldsymbol{B}, B_{0,1})$ y $\boldsymbol{B}_2 := (\boldsymbol{B}, B_{0,2})$. Además, consideramos el polinomio $F \in \mathbb{F}_q[\boldsymbol{B}, B, T]$ definido como

$$F := T^d + \sum_{i=d-s}^{d-1} a_i T^i + \sum_{i=1}^{d-s-1} B_i T^i + B. \tag{5.27}$$

Finalmente, sean $\Gamma_{m,n} \subset \mathbb{A}^{d-s+1+m+n}$ la cuasi–$\mathbb{F}_q$–variedad afín definida como

$$\begin{aligned}\Gamma_{m,n} := \{(\boldsymbol{b}, b_{0,1}, b_{0,2}, \boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathbb{A}^{d-s+1+m+n},\ & F(\boldsymbol{b}, b_{0,1}, \alpha_j) = 0\,(1 \leq j \leq m),\\ & \alpha_i \neq \alpha_j\,(i \neq j),\ F(\boldsymbol{b}, b_{0,2}, \beta_k) = 0\ (1 \leq k \leq n),\ \beta_i \neq \beta_j\ (i \neq j),\ b_{0,1} \neq b_{0,2}\}.\end{aligned}$$

Similarmente al Lema 5.2.2, tenemos el siguiente resultado que expresa el número $\mathcal{S}^{\boldsymbol{a}}_{m,n}$ en términos de la cantidad de puntos $\mathbb{F}_q$–racionales de $\Gamma_{m,n}$.

Lema 5.5.2. *Sean* m *y* n *enteros positivos con* $1 \leq m,n \leq d$ *y* $d-s+1 \leq m+n \leq 2d$. *Entonces*

$$\frac{|\Gamma_{m,n}(\mathbb{F}_q)|}{m!\,n!} = \mathcal{S}^{\boldsymbol{a}}_{m,n}.$$

Así, a fin de estimar $\mathcal{S}^{\boldsymbol{a}}_{m,n}$ vamos estimar la cantidad $|\Gamma_{m,n}(\mathbb{F}_q)|$. Para ello, consideramos la clausura Zariski $\overline{\Gamma}_{m,n}$ de $\Gamma_{m,n}$ en $\mathbb{A}^{d-s+1+m+n}$ y determinamos ecuaciones que definan dicha clausura. Más precisamente, si $\Gamma^*_{m,n} \subset \mathbb{A}^{d-s+1+m+n}$ es la $\mathbb{F}_q$–variedad afín

$$\begin{aligned}\Gamma^*_{m,n} := \{(\boldsymbol{b}, b_{0,1}, b_{0,2}, \boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathbb{A}^{d-s+1+m+n} :\ & \Delta^{i-1}F(\boldsymbol{b}, b_{0,1}, \alpha_1, \ldots, \alpha_i) = 0\ (1 \leq i \leq m),\\ & \Delta^{j-1}F(\boldsymbol{b}, b_{0,2}, \beta_1, \ldots, \beta_j) = 0\ (1 \leq j \leq n)\},\end{aligned}$$

donde $\Delta^{i-1}F(\boldsymbol{b}, b_{0,1}, T_1, \ldots, T_i)$ y $\Delta^{j-1}F(\boldsymbol{b}, b_{0,2}, U_1, \ldots, U_j)$ denotan las diferencias divididas de $F(\boldsymbol{b}, b_{0,1}, T) \in \overline{\mathbb{F}}_q[T]$ y $F(\boldsymbol{b}, b_{0,2}, U) \in \overline{\mathbb{F}}_q[U]$ respectivamente (ver la Definición (3.6)), vamos a ver que $\overline{\Gamma}_{m,n} = \Gamma^*_{m,n}$.

Al igual que en el estudio del promedio del conjunto de valores en familias lineales, relacionamos $\Gamma_{m,n}$ y $\Gamma^*_{m,n}$ y establecemos una serie de resultados sobre la variedad

afín $\Gamma^*_{m,n}$ y su clausura proyectiva $\mathrm{pcl}(\Gamma^*_{m,n}) \subset \mathbb{P}^{d-s+1+m+n}$, que enunciamos a continuación. Las demostraciones de estos resultados pueden verse en [MPP14, Sections 6, 7, 8, 9 y 10].

Teorema 5.5.3. *Sean m, n enteros positivos tales que $1 \leq m, n \leq d$ y $d - s + 1 \leq m + n \leq 2d$. Entonces*

$$\Gamma_{m,n} = \Gamma^*_{m,n} \cap \{\alpha_i \neq \alpha_j \ (1 \leq i < j \leq m),\ \beta_i \neq \beta_j \ (1 \leq i < j \leq n),\ b_{0,1} \neq b_{0,2}\}. \tag{5.28}$$

Si $p > 2$ y $d - s \geq 3$, entonces

- *la variedad $\Gamma^*_{m,n}$ es una intersección completa de dimensión $d - s + 1$, cuyo lugar tiene codimensión al menos 2 en $\Gamma^*_{m,n}$;*
- $\mathrm{pcl}(\Gamma^*_{m,n}) \cap \{T_0 = 0\} \subset \mathbb{P}^{d-s+m+n}$ *es una $\mathbb{F}_q$–variedad lineal de dimensión $d-s$;*
- $\mathrm{pcl}(\Gamma^*_{m,n}) \subset \mathbb{P}^{d-s+1+m+n}$ *es una intersección completa normal de dimensión $d - s + 1$ y grado $(d!)^2/(d-m)!(d-n)!$.*

Del último ítem de este teorema deducimos que $\Gamma^*_{m,n} \subset \mathbb{A}^{d-s+1+m+n}$ es absolutamente irreducible de dimensión $d - s + 1$. Recordemos que (5.28) muestra que $\Gamma_{m,n}$ coincide con el subconjunto de puntos de $\Gamma^*_{m,n}$ con $b_{0,1} \neq b_{0,2}$, $\alpha_i \neq \alpha_j$ y $\beta_k \neq \beta_l$. Así, teniendo en cuenta que $\Gamma^*_{m,n}$ es absolutamente irreducible, concluimos que $\overline{\Gamma}_{m,n} = \Gamma^*_{m,n}$.

Del Teorema 5.5.3 deducimos que estimar el número $|\Gamma_{m,n}(\mathbb{F}_q)|$ equivale a estimar la cantidad de puntos $\mathbb{F}_q$–racionales con coordenadas distintas dos a dos de $\Gamma^*_{m,n}$. Asimismo, las características geométricas de $\Gamma^*_{m,n}$ y su clausura proyectiva nos permiten utilizar la estimación para intersecciones completas proyectivas normales del Teorema 2.2.10 a fin de estimar la cantidad de puntos $\mathbb{F}_q$–racionales de $\Gamma^*_{m,n}$ con coordenadas distintas dos a dos (ver [MPP14, Section 9]). De (5.28) y el Lema 5.5.2 obtenemos el siguiente resultado.

Teorema 5.5.4 ([MPP14, Theorem 9.1])**.** *Sean $p > 2$ y d y s enteros tales que $q > d$ y $d - s \geq 3$. Para cada par (m, n) con $1 \leq m, n \leq d$ y $d - s + 1 \leq m + n \leq 2d$, tenemos que*

$$\left|\mathcal{S}^{\boldsymbol{a}}_{m,n} - \frac{q^{d-s+1}}{m!n!}\right| \leq \frac{1}{m!n!}\big(\delta_{m,n}(D_{m,n}-2)+2\big)q^{d-s+\frac{1}{2}} + \frac{1}{m!n!}(14D^2_{m,n}\delta^2_{m,n} + \xi_{m,n}\delta_{m,n})q^{d-s},$$

donde $\xi_{m,n} := \binom{m}{2} + \binom{n}{2} + 1$, $D_{m,n} := (m+n)d - \binom{m+1}{2} - \binom{n+1}{2}$ y $\delta_{m,n} := \frac{(d!)^2}{(d-m)!(d-n)!}$.

Finalmente, de los Teoremas 5.5.1 y 5.5.4, por medio de cálculos elementales del tipo del Corolario 5.3.2, obtenemos el siguiente resultado.

Corolario 5.5.5 ([MPP14, Corollary 9.2])**.** *Con las hipótesis del Teorema 5.5.4, tenemos*

$$\big|\mathcal{V}_2(d,s) - \mu_d^2\, q^2\big| \leq d^2 2^{2d+1} q^{3/2} + 14\, d^4 \left(\sum_{k=0}^{d-1} \binom{d}{k}^2 (d-k)!\right)^2 q. \tag{5.29}$$

Por último, discutimos el comportamiento del lado derecho de (5.29). Argumentando como en la demostración del Teorema 5.3.4, fijamos k con $0 \leq k \leq d-1$ y consideramos la función $h(k) := \binom{d}{k}^2 (d-k)!$. Similarmente a la Observación 5.3.3, tenemos que h es una función unimodal en el intervalo $[0, d-1]$ y alcanza su máximo en $\lfloor k_0 \rfloor$, donde $k_0 := -1/2 + \sqrt{5+4d}/2$. En consecuencia,

$$\sum_{k=0}^{d-1} \binom{d}{k}^2 (d-k)! \leq d \binom{d}{\lfloor k_0 \rfloor}^2 (d - \lfloor k_0 \rfloor)! = \frac{d\,(d!)^2}{(d-\lfloor k_0 \rfloor)!\,(\lfloor k_0 \rfloor!)^2}.$$

Argumentando como en la demostración del Teorema 5.3.4, concluimos que

$$\left(\sum_{k=0}^{d-1} \binom{d}{k}^2 (d-k)! \right)^2 \leq 8 \cdot 14^2 d^{2d+2} e^{4\sqrt{d}-2d}.$$

Así, obtenemos el siguiente resultado.

Teorema 5.5.6 ([MPP14, Theorem 9.3])**.** *Sean $p > 2$, $q > d$ y $1 \leq s \leq d-3$. Entonces*

$$\left| \mathcal{V}_2(d,s) - \mu_d^2\, q^2 \right| \leq d^2 2^{2d+1} q^{3/2} + 28^3 d^{2d+6} e^{4\sqrt{d}-2d} q.$$

Finalmente, cabe mencionar que mediante un análisis similar al precedente reducimos el problema de estudiar el comportamiento asintótico de $\mathcal{V}_2(d,0) := \frac{1}{q^{d-1}} \sum_{\boldsymbol{b} \in \mathbb{F}_q^{d-1}} \mathcal{V}(f_{\boldsymbol{b}})^2$ al de estimar la cantidad de puntos $\mathbb{F}_q$–racionales de ciertas intersecciones completas regulares en codimensión 2 con coordenadas distintas dos a dos. Así, la estimación para intersecciones completas proyectivas regulares en codimensión 2 del Teorema 2.2.11 nos permite obtener la siguiente estimación explícita para $\mathcal{V}_2(d,0)$ (para una exposición detallada de los argumentos y demostraciones, ver [MPP14, Section 10].)

Teorema 5.5.7 ([MPP14, Theorem 10.4])**.** *Sean $p > 2$, $q > d$ y $d \geq 3$. Entonces*

$$\left| \mathcal{V}_2(d,0) - \mu_d^2\, q^2 \right| \leq \left(2^{2d-2} d^2 + 28^3 d^{2d+8} e^{4\sqrt{d}-2d} \right) q.$$

Capítulo 6

Conjunto de valores de polinomios con coeficientes prescriptos

En este capítulo obtenemos una nueva estimación explícita de la desviación del comportamiento promedio "esperado" del conjunto de valores de familias de polinomios en $\mathbb{F}_q[T]_d$ con los primeros s coeficientes consecutivos prescriptos, para el caso en que $1 \leq s \leq d/2-1$, que es válida para un cuerpo finito $\mathbb{F}_q$ de característica arbitraria. Cabe recordar que en el capítulo anterior dimos una estimación de dicha desviación para el caso en que $1 \leq s \leq d-3$, válida para cuerpos de característica mayor que 2.

Más precisamente, sean d y s enteros positivos tales que $d < q$ y $1 \leq s \leq d/2-1$. Dado $\boldsymbol{a} := (a_{d-1}, \ldots, a_{d-s}) \in \mathbb{F}_q^s$, sea $f_{\boldsymbol{a}} := T^d + a_{d-1}T^{d-1} + \cdots + a_{d-s}T^{d-s}$. Sea $\mathcal{A}_{\boldsymbol{a}} \subset \mathbb{F}_q[T]_d$ la familia lineal definida en (5.5), es decir, el conjunto de todos los polinomios $f_{\boldsymbol{b}} := f_{\boldsymbol{a}} + b_{d-s-1}T^{d-s-1} + \cdots + b_1 T$ con $\boldsymbol{b} := (b_{d-s-1}, \ldots, b_1) \in \mathbb{F}_q^{d-s-1}$. Sea $\mathcal{V}(d,s)$ el promedio definido en (5.6), esto es, el valor promedio del cardinal del conjunto de valores de $\mathcal{V}(f)$ cuando f recorre todos los elementos de $\mathcal{A}_{\boldsymbol{a}}$.

En el capítulo anterior probamos que si, $p > 2$ y $1 \leq s \leq d-3$, entonces

$$|\mathcal{V}(d,s) - \mu_d\, q| \leq d^2 2^{d-1} q^{1/2} + 133\, d^{d+5} e^{2\sqrt{d}-d}. \tag{6.1}$$

En este capítulo damos una nueva estimación de $\mathcal{V}(d,s)$ para un rango de valores menos amplio de s, pero sin restricciones sobre la característica de $\mathbb{F}_q$. Más precisamente, probamos que, si $1 \leq s \leq d/2-1$ y $d < q$, entonces

$$\left|\mathcal{V}(d,s) - \mu_d\, q - \frac{e^{-1}}{2}\right| \leq \frac{(d-2)^5 e^{2\sqrt{d}}}{2^{d-2}} + \frac{7}{q}. \tag{6.2}$$

Observemos que este resultado muestra que $\mathcal{V}(d,s) = \mu_d q + \mathcal{O}(1)$, mientras que (6.1) muestra que $\mathcal{V}(d,s) = \mu_d\, q + \mathcal{O}(q^{1/2})$. Obtenemos una expresión explícita del término de error con un mejor comportamiento que la de (6.1), en el sentido de que tiende a cero cuando d tiende a infinito. De hecho, probamos que $\mathcal{V}(d,s) = \mu_d q + \frac{1}{2e} + \mathcal{O}(\rho^{-d}) + \mathcal{O}(q^{-1})$, con $1/2 < \rho < 1$.

De manera similar a lo hecho en el capítulo anterior, para obtener esta nueva estimación traducimos este problema al de determinar la cantidad de puntos

$\mathbb{F}_q$–racionales con coordenadas distintas dos a dos de una cierta familia de intersecciones completas definidas sobre $\mathbb{F}_q$. En este caso, los polinomios que definen tales intersecciones completas son simétricos; por lo tanto, podemos utilizar las estimaciones para la cantidad de puntos $\mathbb{F}_q$–racionales de intersecciones completas definidas por polinomios simétricos de la Sección 4.1.

Cabe mencionar que estas estimaciones resultarán una herramienta fundamental a fin de realizar un análisis de la complejidad en promedio del algoritmo de búsqueda en bandas verticales para hipersuperficies de los últimos capítulos.

6.1. El conjunto de valores en términos de ceros de polinomios simétricos

De la misma manera que en el Capítulo 5, expresamos el problema de estimar el promedio $\mathcal{V}(d,s)$ como una serie de problemas de interpolación.

Más precisamente, del Teorema 5.2.1 con $r := d-s$ y $m := s$ se obtiene la siguiente expresión combinatoria del promedio $\mathcal{V}(d,s)$, para $1 \leq s \leq d-3$:

$$\mathcal{V}(d,s) = \sum_{i=1}^{d-s} (-1)^{i-1} \binom{q}{i} q^{1-i} + \frac{1}{q^{d-s-1}} \sum_{i=d-s+1}^{d} (-1)^{i-1} \sum_{\mathcal{X}_i \subset \mathbb{F}_q} |\mathcal{S}^{\boldsymbol{a}}_{\mathcal{X}_i}|, \qquad (6.3)$$

donde $\mathcal{S}^{\boldsymbol{a}}_i := \sum_{\mathcal{X}_i \subset \mathbb{F}_q} |\mathcal{S}^{\boldsymbol{a}}_{\mathcal{X}_i}|$ es la cantidad de subconjuntos de i elementos $\mathcal{X}_i \subset \mathbb{F}_q$ tal que existe $g \in \mathbb{F}_q[T]_{\leq d-s-1}$ para el cual $(f_{\boldsymbol{a}} + g)|_{\mathcal{X}_i} \equiv 0$. Recordemos que $\mathbb{F}_q[T]_{\leq d-s-1}$ denota el conjunto de todos los elementos de $\mathbb{F}_q[T]$ de grado a lo sumo $d-s-1$ y que $\mathcal{S}^{\boldsymbol{a}}_{\mathcal{X}_i}$ denota el conjunto de todos los elementos $g \in \mathbb{F}_q[T]_{\leq d-s-1}$ que interpolan a $-f_{\boldsymbol{a}}$ en todos los elementos de $\mathcal{X}_i$.

Por (6.3), a fin de determinar el comportamiento de $\mathcal{V}(d,s)$ necesitamos estimar el número $\mathcal{S}^{\boldsymbol{a}}_i$ para $d-s+1 \leq i \leq d$. Para esto, seguimos un enfoque geométrico como en el Capítulo 5, es decir, expresamos a $\mathcal{S}^{\boldsymbol{a}}_i$ como la cantidad de puntos $\mathbb{F}_q$–racionales con coordenadas distintas dos a dos de cierta intersección completa definida sobre $\mathbb{F}_q$. La variedad que surge en este caso es distinta de la variedad de incidencia del Capítulo 5 y está definida por polinomios invariantes bajo la acción del grupo simétrico de permutaciones de sus coordenadas.

Dado $\boldsymbol{a} \in \mathbb{F}_q^{d-s}$ y un entero positivo i tal que $d-s+1 \leq i \leq d$, fijamos un conjunto $\mathcal{X}_i := \{x_1, \ldots, x_i\} \subset \mathbb{F}_q$ de i elementos y $g \in \mathbb{F}_q[T]_{\leq d-s-1}$. Entonces g pertenece a $\mathcal{S}^{\boldsymbol{a}}_{\mathcal{X}_i}$ si y solo si $(T-x_1)\cdots(T-x_i)$ divide a $f_{\boldsymbol{a}} + g$ en $\mathbb{F}_q[T]$. Como el grado de g es menor o igual que $d-s-1 < i$, deducimos que $-g$ es el resto de la división de $f_{\boldsymbol{a}}$ por $(T-x_1)\cdots(T-x_i)$. En otras palabras, el conjunto $\mathcal{S}^{\boldsymbol{a}}_{\mathcal{X}_i}$ es no vacío si y solo si el resto de la división de $f_{\boldsymbol{a}}$ por $(T-x_1)\cdots(T-x_i)$ tiene grado a lo sumo $d-s-1$.

Sean $X_1, \ldots, X_i$ indeterminadas sobre $\overline{\mathbb{F}}_q$, $\boldsymbol{X} := (X_1, \ldots, X_i)$ y sea

$$Q := (T-X_1)\cdots(T-X_i) \in \mathbb{F}_q[\boldsymbol{X}][T].$$

Existe un polinomio $R_{\boldsymbol{a}} \in \mathbb{F}_q[\boldsymbol{X}][T]$ de grado $\deg R_{\boldsymbol{a}} \leq i-1$ tal que satisface la

siguiente relación:

$$f_{\boldsymbol{a}} \equiv R_{\boldsymbol{a}} \mod Q.$$

Escribimos $R_{\boldsymbol{a}} = R_{i-1}^{\boldsymbol{a}}(\boldsymbol{X})T^{i-1} + \cdots + R_0^{\boldsymbol{a}}(\boldsymbol{X})$. Entonces $R_{\boldsymbol{a}}(x_1, \ldots, x_i, T) \in \mathbb{F}_q[T]$ es el resto de la división de $f_{\boldsymbol{a}}$ por $(T - x_1) \cdots (T - x_i)$. Por lo tanto, el conjunto $\mathcal{S}_{\mathcal{X}_i}^{\boldsymbol{a}}$ es no vacío si y solo si se satisfacen las siguientes igualdades:

$$R_j^{\boldsymbol{a}}(x_1, \ldots, x_i) = 0 \quad (d - s \leq j \leq i - 1). \tag{6.4}$$

Por otro lado, si existe $\boldsymbol{x} := (x_1, \ldots, x_i) \in \mathbb{F}_q^i$ con coordenadas distintas dos a dos tal que se satisface (6.4), entonces el resto de la división de $f_{\boldsymbol{a}}$ por $Q(\boldsymbol{x}, T) = (T - x_1) \cdots (T - x_i)$ es un polinomio $r_{\boldsymbol{a}} := R_{\boldsymbol{a}}(\boldsymbol{x}, T)$ de grado a lo sumo $d - s - 1$. Esto muestra que el conjunto $\mathcal{S}_{\mathcal{X}_i}^{\boldsymbol{a}}$ es no vacío, donde $\mathcal{X}_i := \{x_1, \ldots, x_i\}$. En otras palabras, obtenemos el siguiente resultado.

Lema 6.1.1. *Sean $s, d \in \mathbb{N}$ con $1 \leq s \leq d - 2$, sean $R_j^{\boldsymbol{a}}$ $(d - s \leq j \leq i - 1)$ los polinomios definidos en (6.4) y sea $\mathcal{X}_i := \{x_1, \ldots, x_i\} \subset \mathbb{F}_q$ un conjunto de i elementos. Entonces $\mathcal{S}_{\mathcal{X}_i}^{\boldsymbol{a}}$ es no vacío si y solo si se satisface (6.4).*

Por lo tanto, el número $\mathcal{S}_i^{\boldsymbol{a}}$ de conjuntos $\mathcal{X}_i \subset \mathbb{F}_q$ de i elementos tales que $\mathcal{S}_{\mathcal{X}_i}^{\boldsymbol{a}}$ es no vacío coincide con el número de puntos $\boldsymbol{x} := (x_1, \ldots, x_i) \in \mathbb{F}_q^i$ con coordenadas distintas dos a dos que satisfacen (6.4), salvo permutaciones de las coordenadas. Más precisamente, $\mathcal{S}_i^{\boldsymbol{a}} \cdot i!$ coincide con la cantidad de soluciones $\boldsymbol{x} \in \mathbb{F}_q^i$ del siguiente sistema de igualdades y desigualdades:

$$R_j^{\boldsymbol{a}}(X_1, \ldots, X_i) = 0 \quad (d - s \leq j \leq i - 1), \quad \prod_{1 \leq j < k \leq i} (X_j - X_k) \neq 0.$$

Fijamos i con $d - s + 1 \leq i \leq d$ y suponemos que $2(s + 1) \leq d$. A continuación mostramos cómo los polinomios $R_j^{\boldsymbol{a}}$ pueden expresarse en términos de los primeros s polinomios simétricos elementales $\Pi_1, \ldots, \Pi_s$ de $\mathbb{F}_q[X_1, \ldots, X_r]$. El primer paso es obtener una expresión para el resto de la división de T^j por $Q := (T - X_1) \cdots (T - X_i)$ para $i \leq j \leq d$. Por conveniencia de notaciones, denotamos con $\Pi_0 := 1$.

Lema 6.1.2. *Para $i \leq j \leq d$, se satisfacen las siguientes congruencias:*

$$T^j \equiv H_{i-1,j}T^{i-1} + H_{i-2,j}T^{i-2} + \cdots + H_{0,j} \mod Q, \tag{6.5}$$

donde cada $H_{k,j}$ es igual a cero o es un elemento homogéneo de $\mathbb{F}_q[X_1, \ldots, X_i]$ de grado $j - k$. Más aún, para $j - k \leq i$, el polinomio $H_{k,j} \in \mathbb{F}_q[\Pi_1, \ldots, \Pi_{j-k-1}][\Pi_{j-k}]$ es de grado 1 en Π_{j-k} con coeficiente principal ± 1.

Demostración. Procedemos por inducción en $j \geq i$. Teniendo en cuenta que

$$T^i \equiv \Pi_1 T^{i-1} - \Pi_2 T^{i-2} + \cdots + (-1)^{i-1}\Pi_i \mod Q, \tag{6.6}$$

deducimos inmediatamente (6.5) para $j = i$ y que $H_{0,i} = (-1)^{i-1}\Pi_i$ es mónico de grado 1 en Π_i. Supongamos ahora que (6.5) vale para j con $i \leq j$. Multiplicando

ambos lados de (6.5) por T y combinando este resultado con (6.6) deducimos que

$$\begin{aligned} T^{j+1} &\equiv H_{i-1,j}T^i + H_{i-2,j}T^{i-1} + \cdots + H_{0,j}T \mod Q \\ &\equiv (\Pi_1 H_{i-1,j} + H_{i-2,j})T^{i-1} + \cdots + ((-1)^{i-2}\Pi_{i-1}H_{i-1,j} + H_{0,j})T \\ &\quad + (-1)^{i-1}\Pi_i H_{i-1,j} \mod Q. \end{aligned}$$

Definimos

$$\begin{aligned} H_{k,j+1} &:= (-1)^{i-1-k}\Pi_{i-k}H_{i-1,j} + H_{k-1,j} \text{ para } 1 \le k \le i-1, \\ H_{0,j+1} &:= (-1)^{i-1}\Pi_i H_{i-1,j}. \end{aligned}$$

Entonces

$$T^{j+1} \equiv H_{i-1,j+1}T^{i-1} + H_{i-2,j+1}T^{i-2} + \cdots + H_{0,j+1} \mod Q.$$

Resta probar que el polinomio $H_{k,j+1}$ satisface las propiedades del enunciado del lema. Fijamos k con $1 \le k \le i-1$. Entonces $H_{k,j+1} = (-1)^{i-1-k}\Pi_{i-k}H_{i-1,j} + H_{k-1,j}$. Por la hipótesis inductiva se tiene que $H_{i-1,j}$ y $H_{k-1,j}$ son nulos u homogéneos de grados $j-i+1$ y $j-k+1$ respectivamente. Concluimos que $H_{k,j+1}$ es nulo o es homogéneo de grado $j-k+1$. Además, para $j+1-k \le i$, como $\max\{i-k, j-i+1\} \le j-k < i$, tenemos que $\Pi_{i-k}H_{i-1,j}$ es un elemento del anillo de polinomios $\mathbb{F}_q[\Pi_1, \ldots, \Pi_{j-k}]$. Por otro lado, $H_{k-1,j}$ es un elemento $\mathbb{F}_q[\Pi_1, \ldots, \Pi_{j-k}][\Pi_{j-k+1}]$ de grado 1 con coeficiente principal ± 1, lo cual implica que $H_{k,j+1}$ también lo es. Finalmente, para $k = 0$ tenemos que $H_{0,j+1} := (-1)^{i-1}\Pi_i H_{i-1,j}$, lo que muestra que $H_{0,j+1}$ es nulo o un polinomio homogéneo de $\mathbb{F}_q[X_1, \ldots, X_i]$ de grado $i + j - i + 1 = j + 1$. Esto finaliza la demostración del lema. □

Finalmente, expresamos cada polinomio $R_j^{\boldsymbol{a}}$ en términos de los polinomios $H_{k,j}$.

Proposición 6.1.3. *Sean $s, d \in \mathbb{N}$ con $1 \le s \le d-2$ y $2(s+1) \le d$. Para $d - s \le j \le i - 1$, tenemos la siguiente igualdad:*

$$R_j^{\boldsymbol{a}} = a_j + \sum_{k=i}^{d} a_k H_{j,k}, \tag{6.7}$$

donde los polinomios $H_{j,k}$ son los definidos en el Lema 6.1.2. En particular, $R_j^{\boldsymbol{a}}$ es un polinomio mónico de $\mathbb{F}_q[\Pi_1, \ldots, \Pi_{d-1-j}][\Pi_{d-j}]$, salvo una constante no nula en $\mathbb{F}_q$, de grado $d - j \le s$ para $d - s \le j \le i - 1$.

Demostración. Por el Lema 6.1.2, para $i \le j \le d$ se satisface la siguiente relación:

$$T^j \equiv H_{i-1,j}T^{i-1} + H_{i-2,j}T^{i-2} + \cdots + H_{0,j} \mod Q.$$

Así, obtenemos que

$$\begin{aligned} f_{\boldsymbol{a}} = \sum_{j=d-s}^{d} a_j T^j &= \sum_{j=d-s}^{i-1} a_j T^j + \sum_{j=i}^{d} a_j T^j \\ &\equiv \sum_{j=d-s}^{i-1} a_j T^j + \sum_{j=i}^{d} a_j \sum_{k=d-s}^{i-1} H_{k,j} T^k + \mathcal{O}(T^{d-s-1}) \mod Q \\ &\equiv \sum_{j=d-s}^{i-1} \Big(a_j + \sum_{k=i}^{d} a_k H_{j,k} \Big) T^j + \mathcal{O}(T^{d-s-1}) \mod Q, \end{aligned}$$

donde $\mathcal{O}(T^{d-s-1})$ representa una suma de los términos de $\mathbb{F}_q[X_1, \ldots, X_i][T]$ de grado a lo sumo $d-s-1$ en T. Esto muestra que los polinomios $R_j^{\boldsymbol{a}}$ satisfacen (6.7). Por otro lado, observamos que, para cada $H_{j,k}$ que aparece en la fórmula (6.7), tenemos que $k-j \leq s \leq d-s-2 \leq i$. Esto implica que $H_{j,k} \in \mathbb{F}_q[\Pi_1, \ldots, \Pi_{k-j-1}][\Pi_{k-j}]$ es de grado 1 en Π_{k-j} con coeficiente principal ± 1 y de grado $k-j$ en las variables $X_1, \ldots X_i$. En consecuencia, para cada $d-s \leq j \leq i-1$ tenemos que $R_j^{\boldsymbol{a}}$ es un elemento mónico del anillo de polinomios $\mathbb{F}_q[\Pi_1, \ldots, \Pi_{d-1-j}][\Pi_{d-j}]$ de grado $d-j$, mirado como polinomio en $X_1, \ldots, X_i$. Esto finaliza la demostración de la proposición. □

6.2. Una estimación para el promedio del conjunto de valores

Por (6.3), el comportamiento asintótico de $\mathcal{V}(d,s)$ está determinado por el del número $\mathcal{S}_i^{\boldsymbol{a}}$ para cada $d-s+1 \leq i \leq d$. Fijando i con $d-s+1 \leq i \leq d$, en la sección anterior asociamos a $f_{\boldsymbol{a}}$ ciertos polinomios $R_j^{\boldsymbol{a}} \in \mathbb{F}_q[X_1, \ldots, X_i]$ con $d-s \leq j \leq i-1$ con la propiedad de que el número de ceros comunes $\mathbb{F}_q$–racionales de $R_{d-s}^{\boldsymbol{a}}, \ldots, R_{i-1}^{\boldsymbol{a}}$ con coordenadas distintas dos a dos es igual a $i! \cdot \mathcal{S}_i^{\boldsymbol{a}}$, es decir,

$$\mathcal{S}_i^{\boldsymbol{a}} = \frac{1}{i!} \left| \left\{ \boldsymbol{x} \in \mathbb{F}_q^i : R_j^{\boldsymbol{a}}(\boldsymbol{x}) = 0 \, (d-s \leq j \leq i-1), x_k \neq x_l \, (1 \leq k < l \leq i) \right\} \right|.$$

Por la Proposición 6.1.3, podemos expresar cada polinomio $R_j^{\boldsymbol{a}}$ en términos de los polinomios simétricos elementales $\Pi_1, \ldots, \Pi_s$ de $\mathbb{F}_q[X_1, \ldots, X_i]$. Más precisamente, si $Y_1, \ldots, Y_s$ son nuevas indeterminadas sobre $\overline{\mathbb{F}}_q$, entonces podemos escribir

$$R_j^{\boldsymbol{a}} = S_j^{\boldsymbol{a}}(\Pi_1, \ldots, \Pi_{d-j}) \quad (d-s \leq j \leq i-1),$$

donde cada $S_j^{\boldsymbol{a}} \in \mathbb{F}_q[Y_1, \ldots, Y_{d-j}]$ es de grado 1 en Y_{d-j} con coeficiente principal ± 1. Consideramos el peso wt sobre $\mathbb{F}_q[Y_1, \ldots, Y_s]$ definido por $\mathsf{wt}(Y_j) := j$ para $1 \leq j \leq s$. Mediante un argumento recursivo es fácil ver que

$$\overline{\mathbb{F}}_q[Y_1, \ldots, Y_s]/(S_{d-s}^{\boldsymbol{a}}, \ldots, S_j^{\boldsymbol{a}}) \simeq \overline{\mathbb{F}}_q[Y_1, \ldots, Y_{d-j-1}] \tag{6.8}$$

para $d-s \leq j \leq i-1$. Así, $S_{d-s}^{\boldsymbol{a}}, \ldots, S_{i-1}^{\boldsymbol{a}}$ forman una sucesión regular de $\mathbb{F}_q[Y_1, \ldots, Y_s]$, es decir, estos polinomios satisfacen la hipótesis (H_1) de la Sección 4.1.

Además, teniendo en cuenta el isomorfismo (6.8) para $j = i-1$, deducimos que $S_{d-s}^{\boldsymbol{a}}, \ldots, S_{i-1}^{\boldsymbol{a}}$ forman un ideal radical de $\mathbb{F}_q[Y_1, \ldots, Y_s]$ y la $\mathbb{F}_q$–variedad afín $W_i^{\boldsymbol{a}} \subset \mathbb{A}^s$ definida por $S_{d-s}^{\boldsymbol{a}}, \ldots, S_{i-1}^{\boldsymbol{a}}$ es isomorfa al espacio afín $\mathbb{A}^{d-i}$. Concluimos así que $W_i^{\boldsymbol{a}}$ es una variedad no singular y, por lo tanto, la matriz $(\partial \boldsymbol{S}^{\boldsymbol{a}}/\partial \boldsymbol{Y})(\boldsymbol{y})$ tiene rango máximo para todo $\boldsymbol{y} \in \mathbb{A}^s$, es decir, $S_{d-s}^{\boldsymbol{a}}, \ldots, S_{i-1}^{\boldsymbol{a}}$ satisfacen la hipótesis (H_2) de la Sección 4.1.

Finalmente, vamos a mostrar que $S_{d-s}^{\boldsymbol{a}}, \ldots, S_{i-1}^{\boldsymbol{a}}$ satisfacen la hipótesis (H_3) de la Sección 4.1. Observemos que el Lema 6.1.2 y la Proposición 6.1.3 implican que la componente homogénea de mayor grado de cada $R_j^{\boldsymbol{a}}$ es $a_d H_{j,d}$ para $d-s \le j \le i-1$. Por otro lado, el Lema 4.1.8 muestra que $a_d H_{j,d} = S_j^{\boldsymbol{a},\mathrm{wt}}(\Pi_1, \ldots, \Pi_s)$, donde $S_j^{\boldsymbol{a},\mathrm{wt}}$ es la componente de mayor peso de $S_j^{\boldsymbol{a}}$. Como $H_{j,d}$ es un elemento mónico de $\mathbb{F}_q[\Pi_1, \ldots, \Pi_{d-j-1}][\Pi_{d-j}]$ de grado 1 en Π_{d-j}, se sigue que $S_j^{\boldsymbol{a},\mathrm{wt}}$ es un elemento de $\mathbb{F}_q[Y_1, \ldots, Y_{d-j-1}][Y_{d-j}]$ de grado 1 en Y_{d-j}. Por lo tanto,

$$\overline{\mathbb{F}}_q[Y_1, \ldots, Y_s]/(S_{d-s}^{\boldsymbol{a},\mathrm{wt}}, \ldots, S_j^{\boldsymbol{a},\mathrm{wt}}) \simeq \overline{\mathbb{F}}_q[Y_1, \ldots, Y_{d-j-1}]$$

para $d-s \le j \le i-1$. Argumentando como arriba concluimos que $S_{d-s}^{\boldsymbol{a},\mathrm{wt}}, \ldots, S_{i-1}^{\boldsymbol{a},\mathrm{wt}}$ satisfacen las hipótesis (H_1) y (H_2), es decir, $S_{d-s}^{\boldsymbol{a}}, \ldots, S_{i-1}^{\boldsymbol{a}}$ satisfacen (H_3).

Sea $V_i^{\boldsymbol{a}} \subset \mathbb{A}^i$ la variedad afín definida por los polinomios $R_{d-s}^{\boldsymbol{a}}, \ldots, R_{i-1}^{\boldsymbol{a}} \in \mathbb{F}_q[X_1, \ldots X_i]$. Como $i-d+s \le s \le d-s-2$ y $S_{d-s}^{\boldsymbol{a}}, \ldots, S_{i-1}^{\boldsymbol{a}}$ satisfacen las hipótesis (H_1), (H_2) y (H_3) de la Sección 4.1, podemos aplicar el Corolario 4.1.15 en este caso. Más precisamente, sea $V_{i,=}^{\boldsymbol{a}}$ el conjunto de puntos $V_i^{\boldsymbol{a}}$ con al menos dos coordenadas distintas que toman el mismo valor, es decir,

$$V_{i,=}^{\boldsymbol{a}} := \bigcup_{1 \le j < k \le i} V_i^{\boldsymbol{a}} \cap \{X_j = X_k\},$$

y sea $V_{i,\neq}^{\boldsymbol{a}} := V_i^{\boldsymbol{a}} \setminus V_{i,=}^{\boldsymbol{a}}$. Del Corolario 4.1.15 (tomando $r := i$ y $m := i-d+s$) deducimos que

$$\left| |V_{i,\neq}^{\boldsymbol{a}}(\mathbb{F}_q)| - q^{d-s} \right| \le 14 D_i^3 \delta_i^2 (q+1) q^{d-s-2} + \binom{i}{2} \delta_i \, q^{d-s-1}, \tag{6.9}$$

donde $D_i := \sum_{j=d-i+1}^{s} (j-1)$ y $\delta_i := \prod_{j=d-i+1}^{s} j = s!/(d-i)!$. De (6.9) obtenemos la siguiente estimación para $\mathcal{S}_i^{\boldsymbol{a}}$.

Teorema 6.2.1. *Sean d, i, s enteros con $1 \le s \le d-2$ y $2(s+1) \le d$. Para $d-s+1 \le i \le d$ tenemos que*

$$\left| \mathcal{S}_i^{\boldsymbol{a}} - \frac{q^{d-s}}{i!} \right| \le \frac{i(i-1)}{2i!} \, \delta_i q^{d-s-1} + \frac{14}{i!} D_i^3 \delta_i^2 (q+1) q^{d-s-2},$$

donde $D_i := \sum_{j=d-i+1}^{s} (j-1)$ y $\delta_i := \prod_{j=d-i+1}^{s} j = s!/(d-i)!$.

Finalmente, combinando (6.3) con el Teorema 6.2.1 obtenemos el siguiente resultado.

Corolario 6.2.2. *Con las hipótesis del Teorema 6.2.1, tenemos la siguiente estimación:*

$$\left|\mathcal{V}(d,s)-\mu_d\,q-\frac{1}{2e}\right|\leq\frac{s^2+1}{(d-s-1)!}+\frac{21}{8}\frac{s^6(s!)^2}{d!}\sum_{k=0}^{s-1}\binom{d}{k}\frac{1}{k!}+\frac{7}{q}. \tag{6.10}$$

Demostración. Por (6.3) obtenemos

$$\mathcal{V}(d,s)-\mu_d\,q=\sum_{i=1}^{d-s}(-q)^{1-i}\left(\binom{q}{i}-\frac{q^i}{i!}\right)+\frac{1}{q^{d-s-1}}\sum_{i=d-s+1}^{d}(-1)^{i-1}\left(\mathcal{S}_i^{\boldsymbol{a}}-\frac{q^{d-s}}{i!}\right). \tag{6.11}$$

En primer lugar, acotamos superiormente el valor absoluto $A_2(d,s)$ del primer término en el lado derecho de (6.11). Observemos que dicha suma se encuentra en (5.21). Si reemplazamos $r:=d-s$ en la cota superior de $A(d,r)$ de (5.21), tenemos que

$$\left|A_2(d,s)-\frac{1}{2e}\right|\leq\frac{1}{2\cdot(d-s-1)!}+\sum_{i=2}^{d-s}\left(\frac{1}{d}+\frac{8}{i^2}\right)\frac{1}{q}\leq\frac{1}{2\cdot(d-s-1)!}+\frac{7}{q}. \tag{6.12}$$

Consideramos ahora el valor absoluto del segundo término del lado derecho de (6.11). Por el Teorema 6.2.1 tenemos que

$$\begin{aligned}B_2(d,s)&:=\frac{1}{q^{d-s-1}}\sum_{i=d-s+1}^{d}\left|\mathcal{S}_i^{\boldsymbol{a}}-\frac{q^{d-s}}{i!}\right|\\&\leq\sum_{i=d-s+1}^{d}\frac{i(i-1)}{2i!}\,\delta_i+\sum_{i=d-s+1}^{d}\frac{14}{i!}D_i^3\delta_i^2\left(1+\frac{1}{q}\right).\end{aligned}$$

Ahora bien, el primer término del lado derecho de esta última desigualdad se puede acotar de la siguiente manera:

$$\begin{aligned}\sum_{i=d-s+1}^{d}\frac{i(i-1)}{2i!}\,\delta_i&=\frac{s!}{2(d-2)!}\sum_{i=d-s+1}^{d}\binom{d-2}{i-2}\\&\leq\frac{s\cdot s!}{2(d-2)!}\binom{d-2}{s-1}=\frac{s^2}{2(d-s-1)!}.\end{aligned}$$

Por otro lado,

$$\sum_{i=d-s+1}^{d}\frac{14}{i!}D_i^3\delta_i^2\leq\frac{7}{4}\sum_{i=d-s+1}^{d}\frac{s^3(s-1)^3(s!)^2}{i!((d-i)!)^2}=\frac{7}{4}\sum_{k=0}^{s-1}\frac{s^6(s!)^2}{(d-k)!(k!)^2}.$$

Finalmente, obtenemos

$$B_2(d,s)\leq\frac{s^2}{2(d-s-1)!}+\frac{21}{8}\frac{s^6(s!)^2}{d!}\sum_{k=0}^{s-1}\binom{d}{k}\frac{1}{k!}.$$

Combinando las cotas superiores de $A_2(d,s)$ y $B_2(d,s)$ se deduce el corolario. □

Por último, analizamos el comportamiento del lado derecho de (6.10). Esto nos permitirá mostrar que el término de error tiende a cero cuando d tiende a infinito. Fijamos k con $0 \leq k \leq s-1$ y consideramos la función $h_1(k) := \binom{d}{k}\frac{1}{k!}$. Analizando el signo de las diferencias $h_1(k+1) - h_1(k)$ para $0 \leq k \leq s-2$, deducimos el siguiente resultado.

Observación 6.2.3. *Sea $k_0 := -1/2 + \sqrt{5+4d}/2$. Entonces h_1 es una función unimodal en el intervalo de enteros $[0, s-1]$, que alcanza su máximo en $\lfloor k_0 \rfloor$.*

A partir de la Observación 6.2.3 vemos que

$$\frac{s^6(s!)^2}{d!}\sum_{k=0}^{s-1}\binom{d}{k}\frac{1}{k!} \leq \frac{s^7(s!)^2}{d!}\binom{d}{\lfloor k_0\rfloor}\frac{1}{\lfloor k_0\rfloor!} = \frac{s^7(s!)^2}{(d-\lfloor k_0\rfloor)!(\lfloor k_0\rfloor!)^2}. \tag{6.13}$$

Para obtener una cota superior del lado derecho de (6.13) utilizamos, como en el Capítulo 5, la fórmula de Stirling: para $m \in \mathbb{N}$, existe θ con $0 \leq \theta < 1$ tal que $m! = (m/e)^m\sqrt{2\pi m}\, e^{\theta/12m}$. Aplicando esta fórmula, teniendo en cuenta que $2(s+1) \leq d$, vemos que existen θ_j $(j = 1, 2, 3)$ con $0 \leq \theta_j < 1$ tales que

$$C_2(d,s) := \frac{s^7(s!)^2}{(d-\lfloor k_0\rfloor)!(\lfloor k_0\rfloor!)^2} \leq \frac{(\frac{d}{2}-1)^8(\frac{d}{2}-1)^{d-2}\, e^{2+\lfloor k_0\rfloor+\frac{\theta_1}{3d-6}-\frac{\theta_2}{12(d-\lfloor k_0\rfloor)}-\frac{\theta_3}{6\lfloor k_0\rfloor}}}{\big(d-\lfloor k_0\rfloor\big)^{d-\lfloor k_0\rfloor}\sqrt{2\pi(d-\lfloor k_0\rfloor)}\lfloor k_0\rfloor^{2\lfloor k_0\rfloor+1}}.$$

De cálculos elementales se sigue que

$$\begin{aligned}(d-\lfloor k_0\rfloor)^{-d+\lfloor k_0\rfloor} &\leq d^{-d+\lfloor k_0\rfloor}e^{\lfloor k_0\rfloor(d-\lfloor k_0\rfloor)/d},\\ \frac{d^{\lfloor k_0\rfloor}}{\lfloor k_0\rfloor^{2\lfloor k_0\rfloor}} &\leq e^{(d-\lfloor k_0\rfloor^2)/\lfloor k_0\rfloor},\\ \left(\frac{d}{2}-1\right)^{d-2} &\leq \left(\frac{d}{2}\right)^{d-2}e^{4/d-2}.\end{aligned}$$

Luego,

$$C_2(d,s) \leq \Big(\frac{d}{2}-1\Big)^8\frac{e^{\lfloor k_0\rfloor+\frac{1}{3d-6}+\frac{4}{d}+\frac{\lfloor k_0\rfloor}{d}(d-\lfloor k_0\rfloor)+\frac{1}{\lfloor k_0\rfloor}(d-\lfloor k_0\rfloor^2)}}{d^2 2^{d-2}\sqrt{2\pi(d-\lfloor k_0\rfloor)}\lfloor k_0\rfloor}.$$

De acuerdo a la definición de $\lfloor k_0\rfloor$, es fácil ver que

$$\begin{aligned}\lfloor k_0\rfloor + \frac{\lfloor k_0\rfloor}{d}(d-\lfloor k_0\rfloor) &\leq 2\lfloor k_0\rfloor - \frac{1}{5},\\ \frac{1}{\lfloor k_0\rfloor}(d-\lfloor k_0\rfloor^2) &\leq 4,\\ \frac{(\frac{d}{2}-1)^3}{d^2\lfloor k_0\rfloor\sqrt{d-\lfloor k_0\rfloor}} &\leq \frac{3}{20}.\end{aligned}$$

Por lo tanto, teniendo en cuenta que $d \geq 2$, concluimos que

$$C_2(d,s) \leq \frac{3(\frac{d}{2}-1)^5 e^{\frac{1}{3d-6}+\frac{4}{d}-\frac{1}{5}+3+\sqrt{5+4d}}}{5\sqrt{2\pi}\,2^d}. \tag{6.14}$$

Combinando estas cotas con el Corolario 6.2.2 obtenemos el siguiente resultado.

Teorema 6.2.4. *Para $q > d$ y $1 \leq s \leq d/2 - 1$, tenemos que*

$$\left| \mathcal{V}(d,s) - \mu_d\, q - \frac{1}{2e} \right| \leq \frac{(d-2)^5 e^{2\sqrt{d}}}{2^{d-2}} + \frac{7}{q}. \tag{6.15}$$

Demostración. Teniendo en cuenta (6.14) y que $\sqrt{5+4d} \leq 4/5 + 2\sqrt{d}$ para $d \geq 2$, concluimos que

$$\frac{21}{8} \frac{s^6 (s!)^2}{d!} \sum_{k=0}^{s-1} \binom{d}{k} \frac{1}{k!} \leq 3\, \frac{(d-2)^5 e^{2\sqrt{d}}}{2^d}.$$

Por otro lado, es fácil ver que

$$\frac{s^2+1}{2(d-s-1)!} \leq \frac{(d-2)^5 e^{2\sqrt{d}}}{2^d}.$$

A partir de estas desigualdades se deduce el teorema. □

Para finalizar, hacemos algunos comentarios sobre el comportamiento de la cota de (6.15)

Observación 6.2.5. *Consideremos la función $f : \mathbb{Z}_{\geq 4} \to \mathbb{R}$ definida por $f(d) := (d-2)^5 e^{2\sqrt{d}} 2^{-d}$. Se verifica que f es una función unimodal que alcanza su máximo en $d_0 := 14$ y $f(d_0) \approx 1.08 \cdot 10^5$. Es fácil ver que $\lim_{d\to+\infty} f(d) = 0$; de hecho, si $d \geq 51$ entonces $f(d) < 1$.*

Una cota superior evidente para el lado izquierdo de (6.15) es $|\mathcal{V}(d,s) - \mu_d\, q - (2e)^{-1}| \leq (1-\mu_d)q$. De cálculos directos podemos mostrar que la cota superior del Teorema 6.2.4 no es interesante para valores pequeños de q si $d \leq 44$. Por otro lado, para $1 \leq s \leq \frac{d}{2} - 3$, podemos mejorar significativamente la cota superior del Teorema 6.2.4. Más precisamente, argumentando como en la demostración del Teorema 6.2.4 obtenemos la siguiente cota superior:

$$\left| \mathcal{V}(d,s) - \mu_d\, q - \frac{1}{2e} \right| \leq \frac{9(d-6) e^{2\sqrt{d}}}{2^{d-2}} + \frac{7}{q}. \tag{6.16}$$

Sea $g := \mathbb{Z}_{\geq 7} \to \mathbb{R}$ definida por $g(d) := 9(d-6)e^{2\sqrt{d}} 2^{-d+2}$. Entonces g es una función unimodal que alcanza su máximo en $d_1 := 9$, es decir $g(d_1) := 85$. Además, $\lim_{d\to+\infty} g(d) = 0$ y $g(d) < 1$ para $d \geq 24$. En particular, (6.16) es no trivial para $d \geq 19$.

Terminamos este capítulo con la siguiente observación en donde discutimos el comportamiento asintótico del lado derecho de (6.10).

Observación 6.2.6. *Sea*

$$H(d,s) := \frac{s^6 (s!)^2}{d!} \sum_{k=0}^{s-1} \binom{d}{k} \frac{1}{k!}.$$

Sea $a_d(k) := \binom{d}{k}\frac{1}{k!}$ *para* $0 \leq k \leq d$. *En [LP81] se muestra que* a_d *es una función unimodal en el intervalo entero* $[0, d]$ *que alcanza su máximo en* $\lfloor k_0 \rfloor$, *donde* k_0 *está definido en la Observación 6.2.3. Además, para* $\epsilon > 1/4$ *se prueba que*

$$\sum_{k=0}^{d} a_d(k) \sim \sum_{k \in (k_0 - d^\epsilon, k_0 + d^\epsilon)} a_d(k) \sim \frac{1}{2\sqrt{\pi e}} d^{-1/4} e^{2\sqrt{d}},$$

donde el símbolo $\sim$ *denota el mismo comportamiento asintótico. Suponemos que* $s > \lfloor k_0 \rfloor + d^\epsilon$ *con* $\epsilon > 1/4$. *Entonces, por la fórmula de Stirling, obtenemos que*

$$H(d, s) \sim \frac{1}{\sqrt{2e}} \left(\frac{e}{d}\right)^d \left(\frac{s}{e}\right)^{2s} s^7 e^{2(s-\sqrt{d})} d^{-3/4}.$$

Finalmente, observemos que, si $s \leq \lfloor k_0 \rfloor + d^\epsilon$ *con* $\epsilon > 1/4$, *entonces el lado derecho de esta expresión es una cota superior de* $H(d, s)$ *para* d *suficientemente grande. Esto muestra que* $H(d, s)$ *converge a* 0 *con un radio doblemente exponencial* $d^{-(1-2\lambda)d}$ *para* $s \leq \lambda d$ *con* $\lambda \in [0, 1/2[$.

Capítulo 7

La distribución de patrones de factorización en familias lineales

En este capítulo vamos a estudiar otro problema combinatorio clásico sobre un cuerpo finito: la distribución de patrones de factorización en familias lineales de polinomios univariados de grado dado y con coeficientes en $\mathbb{F}_q$. Más precisamente, vamos a dar estimaciones explícitas del número de elementos de una familia lineal de polinomios mónicos univariados de grado d con coeficientes en $\mathbb{F}_q$ y con un patrón de factorización dado, en el caso que d es menor que q, en términos de parámetros sintácticos de la familia y el patrón de factorización.

Sea T una indeterminada sobre $\overline{\mathbb{F}}_q$. Sean d un entero positivo y $\mathbb{F}_q[T]_d$ el conjunto de polinomios mónicos en $\mathbb{F}_q[T]$ de grado d. Sean $\lambda_1, \ldots, \lambda_d$ enteros no negativos tales que

$$\lambda_1 + 2\lambda_2 + \cdots + d\lambda_d = d.$$

Denotamos por $\mathbb{F}_q[T]_{d,\boldsymbol{\lambda}}$ el conjunto de $f \in \mathbb{F}_q[T]_d$ con patrón de factorización $\boldsymbol{\lambda} := 1^{\lambda_1} 2^{\lambda_2} \cdots d^{\lambda_d}$, es decir, los polinomios $f \in \mathbb{F}_q[T]_d$ que tienen exactamente λ_i factores irreducibles mónicos de grado i con coeficientes en $\mathbb{F}_q$ (contados con mutiplicidad) para $1 \leq i \leq d$. En todo este capítulo usamos la notación $\mathcal{S}_{\boldsymbol{\lambda}} := \mathcal{S} \cap \mathbb{F}_q[T]_{d,\boldsymbol{\lambda}}$ para cualquier subconjunto $\mathcal{S} \subset \mathbb{F}_q[T]_d$.

S. Cohen muestra en [Coh70] que la proporción de elementos de $\mathbb{F}_q[T]_{d,\boldsymbol{\lambda}}$ en $\mathbb{F}_q[T]_d$ es del orden de $\mathcal{T}(\boldsymbol{\lambda})$, donde éste último número representa la cantidad de permutaciones cuyo patrón de descomposición en ciclos en el grupo simétrico $\mathbb{S}_d$ de d elementos es $\boldsymbol{\lambda}$. Más precisamente, Cohen prueba que

$$|\mathbb{F}_q[T]_{d,\boldsymbol{\lambda}}| = \mathcal{T}(\boldsymbol{\lambda})\, q^d + \mathcal{O}(q^{d-1/2}),$$

donde la constante que subyace a la notación $\mathcal{O}$ depende solamente de $\boldsymbol{\lambda}$. Una permutación de $\mathbb{S}_d$ se dice que tiene *patrón de descomposición en ciclos* $\boldsymbol{\lambda}$, o que es *una permutación de patrón* $\boldsymbol{\lambda}$, si se descompone en exactamente λ_i ciclos de longitud i para $1 \leq i \leq d$. Observemos que el número de permutaciones en $\mathbb{S}_d$ de patrón $\boldsymbol{\lambda}$ es $d!/w(\boldsymbol{\lambda})$, donde $w(\boldsymbol{\lambda}) := 1^{\lambda_1} 2^{\lambda_2} \ldots d^{\lambda_d} \lambda_1! \lambda_2! \ldots \lambda_d!$. En particular, $\mathcal{T}(\boldsymbol{\lambda}) := \frac{1}{w(\boldsymbol{\lambda})}$.

Posteriormente, Cohen propone en [Coh72] que un subconjunto $\mathcal{S} \subset \mathbb{F}_q[T]_{d,\boldsymbol{\lambda}}$ está *uniformemente distribuido* si la proporción $|\mathcal{S}_{\boldsymbol{\lambda}}|/|\mathcal{S}|$ es del orden de $\mathcal{T}(\boldsymbol{\lambda})$ para todo

patrón de factorización $\boldsymbol{\lambda}$. El principal resultado de este trabajo ([Coh72, Theorem 3]) provee una condición suficiente para asegurar que una familia $\mathcal{S}$ de polinomios de $\mathbb{F}_q[T]_d$ está uniformemente distribuida en el sentido de arriba. Más precisamente, si $p > d$ y $\mathcal{S}$ es una familia lineal de elementos de $\mathbb{F}_q[T]_d$ que cumple ciertas restricciones técnicas y tiene codimensión $m \leq d-2$, entonces se tiene que

$$|\mathcal{S}_{\boldsymbol{\lambda}}| = \mathcal{T}(\boldsymbol{\lambda})\, q^{d-m} + \mathcal{O}(q^{d-m-\frac{1}{2}}). \tag{7.1}$$

Una dificultad con la estimación (7.1) es que las condiciones para que una familia sea uniformemente distribuida son muy técnicas y difíciles de verificar en casos concretos. Además, el resultado es válido para cuerpos de característica p mayor que d, lo cual impide su aplicación a cuerpos de característica pequeña. Por último, queremos mejorar el comportamiento asintótico del término de error $\mathcal{O}(q^{d-m-1/2})$ y encontrar una estimación explícita del término de error subyacente en la constante $\mathcal{O}$ que aparece en (7.1).

Para esto, consideramos la familia lineal de polinomios en $\mathbb{F}_q[T]_d$ que describimos a continuación. Sean m y r enteros positivos tales que $q > d$ y $3 \leq r \leq d-m$, sean $A_{d-1}, \ldots, A_r$ indeterminadas sobre $\overline{\mathbb{F}}_q$ y sean $L_1, \ldots, L_m \in \mathbb{F}_q[A_{d-1}, \ldots, A_r]$ las formas lineales afines definidas por

$$L_k := b_{k,d-1}A_{d-1} + \cdots + b_{k,r}A_r + b_{k,0} \quad (1 \leq k \leq m). \tag{7.2}$$

Suponemos sin pérdida de generalidad que $L_1, \ldots, L_m$ son linealmente independientes. Sea $\boldsymbol{L} := (L_1, \ldots, L_m)$ y sea $\mathcal{A} := \mathcal{A}_{\boldsymbol{L}} \subset \mathbb{F}_q[T]_d$ la familia lineal definida como

$$\mathcal{A} := \left\{ T^d + a_{d-1}T^{d-1} + \cdots + a_0 \in \mathbb{F}_q[T]_d : \boldsymbol{L}(a_{d-1}, \ldots, a_r) = \mathbf{0} \right\}. \tag{7.3}$$

Observemos que esta familia lineal fue considerada en el Capítulo 3 (ver (3.4)) y en el Capítulo 5 (ver (5.3)).

Suponiendo sin pérdida de generalidad que la matriz Jacobiana $(\partial \boldsymbol{L}/\partial \boldsymbol{A})$ es escalonada por columnas, denotamos con $1 \leq i_1 < \cdots < i_m \leq d-r$ las posiciones correspondientes a los pivotes.

Dado un patrón de factorización $\boldsymbol{\lambda} := 1^{\lambda_1} \ldots d^{\lambda_d}$, el objetivo de este capítulo es demostrar que la familia $\mathcal{A}$ es uniformemente distribuida en el sentido de Cohen, es decir, $|\mathcal{A}_{\boldsymbol{\lambda}}| \approx T(\boldsymbol{\lambda})q^{d-m}$, dando una cota explícita del error $||\mathcal{A}_{\boldsymbol{\lambda}}| - T(\boldsymbol{\lambda})q^{d-m}|$. De manera similar a lo hecho en los capítulos anteriores, vamos a expresar el número $|\mathcal{A}_{\boldsymbol{\lambda}}|$ en términos de la cantidad de puntos $\mathbb{F}_q$–racionales con coordenadas distintas dos a dos de ciertas intersecciones completas singulares definidas sobre $\mathbb{F}_q$. Tales intersecciones completas están definidas por polinomios simétricos, lo que nos va a permitir utilizar los resultados del Capítulo 4.

7.1. Patrones de factorización y raíces

Sean $\mathcal{A} \subset \mathbb{F}_q[T]_d$ la familia lineal de (7.3) y $\boldsymbol{\lambda} := 1^{\lambda_1} \ldots d^{\lambda_d}$ un patrón de factorización. En esta sección mostramos que la condición de que un elemento de $\mathcal{A}$

tenga patrón de factorización $\boldsymbol{\lambda}$ puede expresarse en términos de ciertos polinomios simétricos elementales.

Sean f un elemento de $\mathbb{F}_q[T]_d$ y $g \in \mathbb{F}_q[T]$ un factor irreducible mónico de f de grado i. Entonces g es el polinomio minimal de una raíz α de f con $\mathbb{F}_q(\alpha) = \mathbb{F}_{q^i}$. Denotemos con $\mathbb{G}_i$ al grupo de Galois $\mathrm{Gal}(\mathbb{F}_{q^i}, \mathbb{F}_q)$ de $\mathbb{F}_{q^i}$ sobre $\mathbb{F}_q$. Podemos expresar al polinomio g de la siguiente manera:

$$g = \prod_{\sigma \in \mathbb{G}_i} (T - \sigma(\alpha)).$$

Así, cada factor irreducible g de f está determinado unívocamente por una raíz α de f (y su órbita bajo la acción del grupo de Galois de $\overline{\mathbb{F}}_q$ sobre $\mathbb{F}_q$), y por lo tanto, sus raíces pertenecen a una extensión de cuerpos de $\mathbb{F}_q$ cuyo grado es el de g. Para $f \in \mathbb{F}_q[T]_{d,\boldsymbol{\lambda}}$, existen λ_1 raíces de f en $\mathbb{F}_q$, digamos $\alpha_1, \ldots, \alpha_{\lambda_1}$ (contadas con multiplicidad), que están asociadas con los factores irreducibles de f en $\mathbb{F}_q[T]$ de grado 1; podemos elegir λ_2 raíces de f en $\mathbb{F}_{q^2} \setminus \mathbb{F}_q$ (contadas con multiplicidad), digamos $\alpha_{\lambda_1+1}, \ldots, \alpha_{\lambda_1+\lambda_2}$, que están asociadas con los λ_2 factores irreducibles de f de grado 2, y así sucesivamente. Vamos a suponer que se realiza una elección de $\lambda_1 + \cdots + \lambda_d$ raíces $\alpha_1, \ldots, \alpha_{\lambda_1+\cdots+\lambda_d}$ de f en $\overline{\mathbb{F}}_q$ de manera que cada factor irreducible mónico de f en $\mathbb{F}_q[T]$ está asociado con una y solo una de esas raíces. Queremos expresar la factorización de f en factores irreducibles en $\mathbb{F}_q[T]$ en términos de las coordenadas de las $\lambda_1 + \cdots + \lambda_d$ raíces elegidas de f en ciertas bases de las correspondientes extensiones $\mathbb{F}_q \hookrightarrow \mathbb{F}_{q^i}$ como $\mathbb{F}_q$–espacios vectoriales. Para este propósito, expresamos la raíz asociada con cada factor irreducible de f de grado i en una base normal Θ_i de la extensión de cuerpos $\mathbb{F}_q \hookrightarrow \mathbb{F}_{q^i}$.

Sea $\theta_i \in \mathbb{F}_{q^i}$ un elemento normal y sea Θ_i la base normal de la extensión $\mathbb{F}_q \hookrightarrow \mathbb{F}_{q^i}$ generada por θ_i, es decir,

$$\Theta_i = \left\{\theta_i, \cdots, \theta_i^{q^{i-1}}\right\}.$$

Observemos que el grupo de Galois $\mathbb{G}_i$ es cíclico y el morfismo de Frobenius $\sigma_i : \mathbb{F}_{q^i} \to \mathbb{F}_{q^i}$ definido por $\sigma_i(x) := x^q$ es un generador de $\mathbb{G}_i$. Así, las coordenadas en la base Θ_i de todos los elementos en la órbita de una raíz $\alpha_k \in \mathbb{F}_{q^i}$ de un factor irreducible de f de grado i son las permutaciones cíclicas de las coordenadas α_k en la base Θ_i.

El vector que contiene todas las coordenadas de las raíces $\alpha_1, \ldots, \alpha_{\lambda_1+\cdots+\lambda_d}$ que hemos elegido para representar los factores irreducibles de f en las bases normales $\Theta_1, \ldots, \Theta_d$ es un elemento de $\mathbb{F}_q^d$, que denotamos con $\boldsymbol{x} := (x_1, \ldots, x_d)$. Sea

$$\ell_{i,j} := \sum_{k=1}^{i-1} k\lambda_k + (j-1)\, i \tag{7.4}$$

para $1 \le j \le \lambda_i$ y $1 \le i \le d$. Observemos que el vector de coordenadas de una raíz $\alpha_{\lambda_1+\cdots+\lambda_{i-1}+j} \in \mathbb{F}_{q^i}$ es el sub–arreglo $(x_{\ell_{i,j}+1}, \ldots, x_{\ell_{i,j}+i})$ de $\boldsymbol{x}$. Con estas notaciones, los λ_i factores irreducibles de f de grado i son los polinomios

$$g_{i,j} = \prod_{\sigma \in \mathbb{G}_i} \left(T - \left(x_{\ell_{i,j}+1}\sigma(\theta_i) + \cdots + x_{\ell_{i,j}+i}\sigma(\theta_i^{q^{i-1}})\right)\right) \tag{7.5}$$

para $1 \leq j \leq \lambda_i$. En particular, tenemos que

$$f = \prod_{i=1}^{d} \prod_{j=1}^{\lambda_i} g_{i,j}. \tag{7.6}$$

Sean $X_1, \ldots, X_d$ indeterminadas sobre $\overline{\mathbb{F}}_q$, sea $\boldsymbol{X} := (X_1, \ldots, X_d)$ y consideramos el polinomio $G \in \mathbb{F}_q[\boldsymbol{X}, T]$ definido como

$$G := \prod_{i=1}^{d} \prod_{j=1}^{\lambda_i} G_{i,j}, \quad G_{i,j} := \prod_{\sigma \in \mathbb{G}_i} \Big(T - \big(X_{\ell_{i,j}+1}\sigma(\theta_i) + \cdots + X_{\ell_{i,j}+i}\sigma(\theta_i^{q^{i-1}}) \big) \Big), \tag{7.7}$$

donde los $\ell_{i,j}$ están definidos en (7.4). Por los argumentos anteriores deducimos que un polinomio $f \in \mathbb{F}_q[T]_d$ tiene patrón de factorización $\boldsymbol{\lambda}$ si y solo si existe $\boldsymbol{x} \in \mathbb{F}_q^d$ tal que $f = G(\boldsymbol{x}, T)$.

A continuación vamos a determinar cuántos elementos $\boldsymbol{x} \in \mathbb{F}_q^d$ producen un polinomio arbitrario $f = G(\boldsymbol{x}, T) \in \mathbb{F}_q[T]_{d,\boldsymbol{\lambda}}$. Para $\alpha \in \mathbb{F}_{q^i}$, tenemos que $\mathbb{F}_q(\alpha) = \mathbb{F}_{q^i}$ si y solo si su órbita bajo la acción del grupo de Galois $\mathbb{G}_i$ tiene exactamente i elementos. En particular, si α puede expresarse por su vector de coordenadas $\boldsymbol{x} \in \mathbb{F}_q^i$ en la base normal Θ_i, entonces los vectores de coordenadas de los elementos de la órbita de α forman un ciclo de longitud i, ya que el morfismo de Frobenius $\sigma_i \in \mathbb{G}_i$ permuta cíclicamente las coordenadas. En consecuencia, existe una biyección entre los ciclos de longitud i en $\mathbb{F}_q^i$ y los elementos $\alpha \in \mathbb{F}_{q^i}$ con $\mathbb{F}_q(\alpha) = \mathbb{F}_{q^i}$. Para hacer esta relación más precisa, introducimos la noción de un arreglo de tipo $\boldsymbol{\lambda}$.

Definición 7.1.1. *Sea $\ell_{i,j}$ $(1 \leq i \leq d,\ 1 \leq j \leq \lambda_i)$ definido como (7.4). Un elemento $\boldsymbol{x} = (x_1, \ldots, x_d) \in \mathbb{F}_q^d$ se dice* tipo $\boldsymbol{\lambda}$ *si y solo si cada sub–arreglo $\boldsymbol{x}_{i,j} := (x_{\ell_{i,j}+1}, \ldots, x_{\ell_{i,j}+i})$ es un ciclo de longitud i.*

Dado $\boldsymbol{x} \in \mathbb{F}_q^d$, el siguiente resultado muestra que el tipo de cada $\boldsymbol{x} \in \mathbb{F}_q^d$ determina el patrón de factorización de $G(\boldsymbol{x}, T)$.

Lema 7.1.2. *Para cualquier $\boldsymbol{x} = (x_1, \ldots, x_d) \in \mathbb{F}_q^d$, el polinomio $f := G(\boldsymbol{x}, T)$ tiene patrón de factorización $\boldsymbol{\lambda}$ si y solo si $\boldsymbol{x}$ es de tipo $\boldsymbol{\lambda}$. Más aún, para cada polinomio libre de cuadrados $f \in \mathbb{F}_q[T]_{d,\boldsymbol{\lambda}}$ hay $w(\boldsymbol{\lambda}) := \prod_{i=1}^{d} i^{\lambda_i}\lambda_i!$ diferentes $\boldsymbol{x} \in \mathbb{F}_q^d$ con $f = G(\boldsymbol{x}, T)$.*

Demostración. Sean $\Theta_1, \ldots, \Theta_d$ las bases normales introducidas anteriormente. Cada $\boldsymbol{x} \in \mathbb{F}_q^d$ está asociado con una única sucesión de elementos α_k $(1 \leq k \leq \lambda_1 + \cdots + \lambda_d)$ de la siguiente manera: $\alpha_{\lambda_1+\cdots+\lambda_{i-1}+j}$ con $1 \leq j \leq \lambda_i$ es el elemento de $\mathbb{F}_{q^i}$ cuyo vector de coordenadas en la base Θ_i es el sub–arreglo $(x_{\ell_{i,j}+1}, \ldots, x_{\ell_{i,j}+i})$ of $\boldsymbol{x}$.

Supongamos que $G(\boldsymbol{x}, T)$ tiene patrón de factorización $\boldsymbol{\lambda}$ para un $\boldsymbol{x} \in \mathbb{F}_q^d$ dado. Fijamos (i, j) con $1 \leq i \leq d$ y $1 \leq j \leq \lambda_i$. Entonces $G(\boldsymbol{x}, T)$ se factoriza como en (7.5)–(7.6), donde cada $g_{i,j} \in \mathbb{F}_q[T]$ es irreducible, y así $\mathbb{F}_q(\alpha_{\lambda_1+\cdots+\lambda_{i-1}+j}) = \mathbb{F}_{q^i}$. Concluimos que el sub–arreglo $(x_{\ell_{i,j}+1}, \ldots, x_{\ell_{i,j}+i})$ que define $\alpha_{\lambda_1+\cdots+\lambda_{i-1}+j}$ es un ciclo de longitud i. Esto prueba que $\boldsymbol{x}$ es de tipo $\boldsymbol{\lambda}$. Por otro lado, dado un $\boldsymbol{x} \in \mathbb{F}_q^d$ de

tipo $\boldsymbol{\lambda}$, fijamos (i,j) con $1 \le i \le d$ y $1 \le j \le \lambda_i$. Entonces $\mathbb{F}_q(\alpha_{\lambda_1+\cdots+\lambda_{i-1}+j}) = \mathbb{F}_{q^i}$, porque el sub–arreglo $(x_{\ell_{i,j}+1}, \ldots, x_{\ell_{i,j}+i})$ es un ciclo de longitud i y así la órbita de $\alpha_{\lambda_1+\cdots+\lambda_{i-1}+j}$ bajo la acción de $\mathbb{G}_i$ tiene i elementos. Esto implica que el factor $g_{i,j}$ de $G(\boldsymbol{x}, T)$ definido como en (7.5) es irreducible de grado i. Así deducimos que $f := G(\boldsymbol{x}, T)$ tiene patrón de factorización $\boldsymbol{\lambda}$.

Además, para $\boldsymbol{x} \in \mathbb{F}_q^d$ de tipo $\boldsymbol{\lambda}$, el polinomio $f := G(\boldsymbol{x}, T) \in \mathbb{F}_q[T]_{d,\boldsymbol{\lambda}}$ es libre de cuadrados si y solo si todas las raíces $\alpha_{\lambda_1+\cdots+\lambda_{i-1}+j}$ con $1 \le j \le \lambda_i$ son distintas dos a dos y no son elementos conjugados de $\mathbb{F}_{q^i}$. Esto implica que ninguna permutación cíclica de un sub–arreglo $(x_{\ell_{i,j}+1}, \ldots, x_{\ell_{i,j}+i})$ con $1 \le j \le \lambda_i$ coincide con otra permutación cíclica de otro sub–arreglo $(x_{\ell_{i,j'}+1}, \ldots, x_{\ell_{i,j'}+i})$. Dado que las permutaciones cíclicas de cualquiera de esos sub–arreglos y las permutaciones de esos sub–arreglos producen elementos de $\mathbb{F}_q^d$ que están asociados al mismo polinomio f, hay $w(\boldsymbol{\lambda}) := \prod_{i=1}^n i^{\lambda_i}\lambda_i!$ diferentes elementos $\boldsymbol{x} \in \mathbb{F}_q^d$ tales que $f = G(\boldsymbol{x}, T)$. □

Consideremos ahora el polinomio G de (7.7) como un elemento de $\mathbb{F}_q[\boldsymbol{X}][T]$. Vamos a expresar los coeficientes de G mediante el vector de formas lineales $\boldsymbol{Y} := (Y_1, \ldots, Y_d)$, donde $Y_i \in \overline{\mathbb{F}}_q[\boldsymbol{X}]$ se define de la siguiente manera para $1 \le i \le d$:

$$(Y_{\ell_{i,j}+1}, \ldots, Y_{\ell_{i,j}+i})^t := A_i \cdot (X_{\ell_{i,j}+1}, \ldots, X_{\ell_{i,j}+i})^t \quad (1 \le j \le \lambda_i,\ 1 \le i \le d), \quad (7.8)$$

siendo $A_i \in \mathbb{F}_{q^i}^{i\times i}$ la matriz

$$A_i := \left(\sigma(\theta_i^{q^h})\right)_{\sigma\in\mathbb{G}_i,\, 0\le h\le i-1}.$$

Por (7.7), podemos expresar al polinomio G como

$$G = \prod_{i=1}^{d}\prod_{j=1}^{\lambda_i}\prod_{k=1}^{i}(T - Y_{\ell_{i,j}+k}) = \prod_{k=1}^{d}(T - Y_k) = T^d + \sum_{k=1}^{d}(-1)^k\left(\Pi_k(\boldsymbol{Y})\right)T^{d-k},$$

donde $\Pi_1(\boldsymbol{Y}), \ldots, \Pi_d(\boldsymbol{Y})$ son los polinomios simétricos elementales de $\mathbb{F}_q[\boldsymbol{Y}]$. De acuerdo con (7.7), G pertenece al anillo de polinomios $\mathbb{F}_q[\boldsymbol{X}, T]$, lo cual implica en particular que $\Pi_k(\boldsymbol{Y})$ pertenece a $\mathbb{F}_q[\boldsymbol{X}]$ para $1 \le k \le d$. Combinando estos argumentos con el Lema 7.1.2 obtenemos el siguiente resultado.

Lema 7.1.3. *Un polinomio $f := T^d + a_{d-1}T^{d-1} + \cdots + a_0 \in \mathbb{F}_q[T]_d$ tiene patrón de factorización $\boldsymbol{\lambda}$ si y solo si existe $\boldsymbol{x} \in \mathbb{F}_q^d$ de tipo $\boldsymbol{\lambda}$ tal que*

$$a_k = (-1)^{d-k}\,\Pi_{d-k}(\boldsymbol{Y}(\boldsymbol{x})) \quad (0 \le k \le d-1). \quad (7.9)$$

En particular, si f es libre de cuadrados, entonces hay $w(\boldsymbol{\lambda})$ elementos $\boldsymbol{x}$ tales que se satisface (7.9).

En consecuencia, podemos expresar la condición de que un elemento de la familia $\mathcal{A}$ de (7.3) tiene patrón de factorización $\boldsymbol{\lambda}$ en términos de los polinomios simétricos elementales $\Pi_1, \ldots, \Pi_{d-r}$ de $\mathbb{F}_q[\boldsymbol{Y}]$.

Corolario 7.1.4. *Un polinomio* $f := T^d + a_{d-1}T^{d-1} + \cdots + a_0 \in \mathcal{A}$ *tiene patrón de factorización* $\boldsymbol{\lambda}$ *si y solo si existe* $\boldsymbol{x} \in \mathbb{F}_q^d$ *de tipo* $\boldsymbol{\lambda}$ *tal que se satisface* (7.9) *y*

$$L_k\big(-\Pi_1(\boldsymbol{Y}(\boldsymbol{x})), \ldots, (-1)^{d-r}\,\Pi_{d-r}(\boldsymbol{Y}(\boldsymbol{x}))\big) = 0 \quad (1 \le k \le m). \tag{7.10}$$

En particular, si $f := G(\boldsymbol{x}, T) \in \mathcal{A}_{\boldsymbol{\lambda}}$ *es libre de cuadrados, entonces hay* $w(\boldsymbol{\lambda})$ *elementos* $\boldsymbol{x}$ *tales que se satisface (7.10).*

7.2. El número de polinomios con patrón de factorización dado

Sean d, r y m enteros positivos tales que $q > d$ y $3 \le r \le d-m$. Sean $A_{d-1}, \ldots, A_r$ indeterminadas sobre $\overline{\mathbb{F}}_q$ y $L_1, \ldots, L_m$ las formas lineales de $\mathbb{F}_q[A_{d-1}, \ldots, A_r]$ definidas en (7.2). Sea $\mathcal{A} \subset \mathbb{F}_q[T]_d$ la familia definida en (7.3). Dado un patrón de factorización $\boldsymbol{\lambda} := 1^{\lambda_1} \cdots d^{\lambda_d}$, consideramos el conjunto $\mathcal{A}_{\boldsymbol{\lambda}}$ formado por todos los elementos de la familia $\mathcal{A} \subset \mathbb{F}_q[T]_d$ que tienen patrón de factorización $\boldsymbol{\lambda}$. En esta sección estimamos el número de elementos de $\mathcal{A}_{\boldsymbol{\lambda}}$.

Para este propósito, el Corolario 7.1.4 muestra que podemos asociar los elementos $\mathcal{A}_{\boldsymbol{\lambda}}$ con los siguientes polinomios de $\mathbb{F}_q[\boldsymbol{X}]$:

$$R_k := R_k^{\boldsymbol{\lambda}} := L_k\big(-\Pi_1(\boldsymbol{Y}(\boldsymbol{X})), \ldots, (-1)^{d-r}\Pi_{d-r}(\boldsymbol{Y}(\boldsymbol{X}))\big) \quad (1 \le k \le m). \tag{7.11}$$

Mediante el cambio de coordenadas definido por $\boldsymbol{Y} := (Y_1, \ldots, Y_d)$, donde $\boldsymbol{Y}$ es el vector de formas lineales de (7.8), podemos expresar cada R_k como un polinomio lineal en los primeros s polinomios simétricos elementales $\Pi_1, \ldots, \Pi_{d-r}$ de $\mathbb{F}_q[\boldsymbol{Y}]$. Más precisamente, sean $Z_1, \ldots, Z_{d-r}$ nuevas indeterminadas sobre $\overline{\mathbb{F}}_q$, sea $\boldsymbol{Z} := (Z_1, \ldots, Z_{d-r})$ y $\mathbb{F}_q[\boldsymbol{Z}]$ el anillo de polinomios en $Z_1, \ldots, Z_{d-r}$ con coeficientes en $\mathbb{F}_q$. Entonces podemos escribir

$$R_k = S_k(\Pi_1, \ldots, \Pi_{d-r}) \quad (1 \le k \le m), \tag{7.12}$$

donde $S_1, \ldots, S_m \in \mathbb{F}_q[\boldsymbol{Z}]$ están definidos por $S_k := L_k(-Z_1, \ldots, (-1)^{d-r}Z_{d-r})$ $(1 \le k \le m)$.

Observemos que $S_1, \ldots, S_m$ son elementos de grado 1 cuyas componentes homogéneas de grado 1 son linealmente independientes en $\mathbb{F}_q[\boldsymbol{Z}]$. Así, la matriz Jacobiana $(\partial \boldsymbol{S}/\partial \boldsymbol{Z})(\boldsymbol{z})$ de $\boldsymbol{S} := (S_1, \ldots, S_m)$ con respecto a $\boldsymbol{Z} := (Z_1, \ldots, Z_{d-r})$ tiene rango máximo m para todo $\boldsymbol{z} \in \mathbb{A}^{d-r}$. Por otro lado, podemos suponer que la matriz Jacobiana $(\partial \boldsymbol{S}/\partial \boldsymbol{Z})$ está escalonada por columnas, denotando por $1 \le i_1 < \cdots < i_m \le d-r$ las posiciones correspondientes a los pivotes.

Si $p > 2$, teniendo en cuenta las propiedades de los polinomios $S_1, \ldots, S_m$, podemos aplicar el Teorema 4.2.13 a $R_1, \ldots, R_m$, tomando en este caso $s := d-r$. En consecuencia, observando que $\deg(R_k) = i_k$ para $1 \le k \le m$, tenemos el siguiente resultado.

Teorema 7.2.1. *Sea $p > 2$. Sean d, r y m enteros positivos tales que $3 \leq r \leq d-m$ y sea $V := V(R_1, \ldots, R_m) \subset \mathbb{A}^d$ la $\mathbb{F}_q$–variedad definida por los polinomios $R_1, \ldots, R_m$ de* (7.11)*. Entonces*

$$\big||V(\mathbb{F}_q)| - q^{d-m}\big| \leq (q+1)q^{d-m-2}\big((\delta_{\boldsymbol{L}}(D_{\boldsymbol{L}}-2)+2)q^{1/2} + 14D_{\boldsymbol{L}}^2\delta_{\boldsymbol{L}}^2\big), \tag{7.13}$$

donde $\delta_{\boldsymbol{L}} := i_1 \cdots, i_m$ y donde $D_{\boldsymbol{L}} := \sum_{j=1}^m (i_j - 1)$.

Ahora, sea $V^=$ la subvariedad afín de V definida por

$$V^= := \bigcup_{\substack{1\leq i\leq d\\ 1\leq j_1<j_2\leq\lambda_i,\ 1\leq k_1<k_2\leq i}} V \cap \{Y_{\ell_{i,j_1}+k_1} = Y_{\ell_{i,j_2}+k_2}\},$$

donde $Y_{\ell_{i,j}+k}$ son las formas lineales determinadas por (7.8), es decir,

$$Y_{\ell_{i,j}+k} := X_{\ell_{i,j}+1}\sigma_{k,i}(\theta_i) + \cdots + X_{\ell_{i,j}+i}\sigma_{k,i}(\theta_i^{q^{i-1}}), \tag{7.14}$$

siendo $\mathbb{G}_i := \{\sigma_{k,i} : 1 \leq k \leq i\}$ el grupo de Galois de $\mathbb{F}_{q^i}$ sobre $\mathbb{F}_q$. Sea $V^{\neq}(\mathbb{F}_q) := V(\mathbb{F}_q)\backslash V^=(\mathbb{F}_q)$. A continuación obtenemos una cota superior para el número $|V^=(\mathbb{F}_q)|$.

El Teorema 4.2.12 asegura que la clausura proyectiva $\mathrm{pcl}(V) \subset \mathbb{P}^d$ de V es una intersección completa normal. Por lo tanto, por el Teorema 2.1.10 concluimos que V es absolutamente irreducible. Así tenemos que $V \cap \{Y_{\ell_{i,j_1}+k_1} = Y_{\ell_{i,j_2}+k_2}\}$ tiene dimensión a lo sumo $d-m-1$ para todo $1 \leq i \leq d$, $1 \leq j_1 < j_2 \leq \lambda_i$ y $1 \leq k_1 < k_2 \leq i$. Concluimos que $V^=$ tiene dimensión a lo sumo $d-m-1$. Luego, por la desigualdad de Bézout (2.1.14), observando que $\deg V \leq \delta_{\boldsymbol{L}} = i_1 \ldots i_m$, deducimos que

$$\deg V^= \leq \deg V \sum_{i=1}^d i^2\lambda_i^2 \leq d^2\delta_{\boldsymbol{L}}.$$

En consecuencia, la Proposición 2.2.1 implica

$$|V^=(\mathbb{F}_q)| \leq \deg V^= q^{d-m-1} \leq d^2\delta_{\boldsymbol{L}}q^{d-m-1}. \tag{7.15}$$

Combinando el Teorema 7.2.1 con (7.15) obtenemos el siguiente resultado.

Corolario 7.2.2. *Con las notaciones del Teorema 7.2.1, tenemos que*

$$\big|V^{\neq}(\mathbb{F}_q)| - q^{d-m}\big| \leq (q+1)q^{d-m-2}\big((\delta_{\boldsymbol{L}}(D_{\boldsymbol{L}}-2)+2)q^{1/2} + 14D_{\boldsymbol{L}}^2\delta_{\boldsymbol{L}}^2\big) + d^2\delta_{\boldsymbol{L}}q^{d-m-1},$$

donde $\delta_{\boldsymbol{L}} := i_1 \ldots i_m$ y $D_{\boldsymbol{L}} := \sum_{j=1}^m (i_j - 1)$.

Demostración. Por (7.15),

$$|V^=(\mathbb{F}_q)| \leq \deg V^= q^{d-m-1} \leq d^2\delta_{\boldsymbol{L}}q^{d-m-1}.$$

Por lo tanto, del Teorema 7.2.1 se sigue que

$$\begin{aligned}\big|V^{\neq}(\mathbb{F}_q)| - q^{d-m}\big| &\leq \big|V(\mathbb{F}_q) - q^{d-m}\big| + \big|V^=(\mathbb{F}_q)\big|\\ &\leq (q+1)q^{d-m-2}\big((\delta_{\boldsymbol{L}}(D_{\boldsymbol{L}}-2)+2)q^{1/2} + 14D_{\boldsymbol{L}}^2\delta_{\boldsymbol{L}}^2\big) + d^2\delta_{\boldsymbol{L}}q^{d-m-1}.\end{aligned}$$

Esto finaliza la demostración del corolario. □

Supongamos ahora que las formas lineales $L_1, \ldots, L_m \in \mathbb{F}_q[A_{d-1}, \ldots, A_{d-r}]$ que definen la variedad lineal $\mathcal{A}$ son ralas. Más precisamente, supongamos que $m + 2 \leq r \leq d - m$. Como los polinomios lineales $S_1, \ldots, S_m \in \mathbb{F}_q[\boldsymbol{Z}]$ definidos por $S_k := L_k(-Z_1, \ldots, (-1)^{d-r} Z_{d-r})$ $(1 \leq k \leq m)$ cumplen que la matriz Jacobiana $(\partial \boldsymbol{S}/\partial \boldsymbol{Z})(\boldsymbol{z})$ tiene rango m para todo $\boldsymbol{z} \in \mathbb{A}^{d-r}$, tenemos que $S_1, \ldots, S_m$ forman una sucesión regular de $\mathbb{F}_q[\boldsymbol{Z}]$. Por lo tanto, $S_1, \ldots, S_m$ cumplen las hipótesis (H_1) y (H_2) de la Sección 4.1, tomando $s := d-r$ y $r := d$. Si suponemos como antes que la matriz Jacobiana $(\partial \boldsymbol{S}/\partial \boldsymbol{Z})(\boldsymbol{z})$ está escalonada por columnas y $1 \leq i_1 < \cdots < i_m \leq d - r$ son las posiciones correspondientes a los pivotes, las componentes homogéneas de mayor peso de $S_1, \ldots, S_m$ son $S_1^{\mathrm{wt}} = c_1 Z_{i_1}, \ldots, S_m^{\mathrm{wt}} = c_m Z_{i_1}$ respectivamente. Así, $S_1, \ldots, S_m$ satisfacen la hipótesis (H_3) de la Sección 4.1. Finalmente, como los enteros m, d y r satisfacen la desigualdad $m + 2 \leq r \leq d - m - 2$, estamos en condiciones de aplicar el Teorema 4.1.13, tomando como antes $s := d - r$ y $r := d$. Como $\deg(R_k) = i_k$ para $1 \leq k \leq m$, por las consideraciones previas y el Corolario 4.1.15 obtenemos el siguiente resultado.

Teorema 7.2.3. *Sean r, m y d enteros positivos tales que $m + 2 \leq r \leq d - m$ y sean $R_1, \ldots, R_m \in \mathbb{F}_q[X_1, \ldots, X_d]$ los polinomios definidos en* (7.12)*. Si $V := V(R_1, \ldots, R_m) \subset \mathbb{A}^d$ es la $\mathbb{F}_q$–variedad afín definida por $R_1, \ldots, R_m$,*

$$V^{=} := \bigcup_{\substack{1 \leq i \leq d \\ 1 \leq j_1 < j_2 \leq \lambda_i,\ 1 \leq k_1 < k_2 \leq i}} V \cap \{Y_{\ell_{i,j_1}+k_1} = Y_{\ell_{i,j_2}+k_2}\},$$

y $V^{\neq} := V \setminus V^{=}$, donde $Y_{\ell_{i,j}+k}$ son las formas lineales afines definidas en (7.8)*, se tiene que*

$$\big||V^{\neq}(\mathbb{F}_q)| - q^{d-m}\big| \leq 14 D_{\boldsymbol{L}}^3 \delta_{\boldsymbol{L}}^2 (q+1) q^{d-m-2} + d^2 \delta_{\boldsymbol{L}} q^{d-m-1},$$

donde $D_{\boldsymbol{L}} := \sum_{j=1}^{m} (i_j - 1)$ y $\delta_{\boldsymbol{L}} := i_1 \cdots i_m$.

Por otro lado, el Corolario 7.1.4 relaciona el número $|V(\mathbb{F}_q)|$ de ceros comunes $\mathbb{F}_q$–racionales de $R_1, \ldots, R_m$ con la cantidad $|\mathcal{A}_{\boldsymbol{\lambda}}|$. Más precisamente, sea $\boldsymbol{x} := (\boldsymbol{x}_{i,j} : 1 \leq i \leq d, 1 \leq j \leq \lambda_i) \in \mathbb{F}_q^d$ un cero $\mathbb{F}_q$–racional de $R_1, \ldots, R_m$ de tipo $\boldsymbol{\lambda}$ (ver la Definición 7.1.1). Entonces asociamos a $\boldsymbol{x}$ con un elemento $f \in \mathcal{A}_{\boldsymbol{\lambda}}$ tal que tiene como raíz $\mathbb{F}_{q^i}$–racional a $Y_{\ell_{i,j}+k}(\boldsymbol{x}_{i,j})$ para $1 \leq i \leq d$, $1 \leq j \leq \lambda_i$ y $1 \leq k \leq i$, donde $Y_{\ell_{i,j}+k}$ es la forma lineal definida en (7.8).

Sea $\mathcal{A}_{\boldsymbol{\lambda}}^{sq} := \{f \in \mathcal{A}_{\boldsymbol{\lambda}} : f \text{ es libre de cuadrados}\}$ y $\mathcal{A}_{\boldsymbol{\lambda}}^{nsq} := \mathcal{A}_{\boldsymbol{\lambda}} \setminus \mathcal{A}_{\boldsymbol{\lambda}}^{sq}$. El Corolario 7.1.4 además asegura que todo elemento $f \in \mathcal{A}_{\boldsymbol{\lambda}}^{sq}$ está asociado con $w(\boldsymbol{\lambda}) := \prod_{i=1}^{d} i^{\lambda_i} \lambda_i!$ ceros comunes $\mathbb{F}_q$–racionales de $R_1, \ldots, R_m$ de tipo $\boldsymbol{\lambda}$. Observemos que $\boldsymbol{x} \in \mathbb{F}_q^d$ es de tipo $\boldsymbol{\lambda}$ si y solo si $Y_{\ell_{i,j}+k_1}(\boldsymbol{x}) \neq Y_{\ell_{i,j}+k_2}(\boldsymbol{x})$ para $1 \leq i \leq d$, $1 \leq j \leq \lambda_i$ y $1 \leq k_1 < k_2 \leq i$. Además, un $\boldsymbol{x} \in \mathbb{F}_q^d$ de tipo $\boldsymbol{\lambda}$ está asociado con $f \in \mathcal{A}_{\boldsymbol{\lambda}}^{sq}$ si y solo si $Y_{\ell_{i,j_1}+k_1}(\boldsymbol{x}) \neq Y_{\ell_{i,j_2}+k_2}(\boldsymbol{x})$ para $1 \leq i \leq d$, $1 \leq j_1 < j_2 \leq \lambda_i$ y $1 \leq k_1 < k_2 \leq i$. En consecuencia, vemos que $|\mathcal{A}_{\boldsymbol{\lambda}}^{sq}| = \mathcal{T}(\boldsymbol{\lambda}) \big|V^{\neq}(\mathbb{F}_q)\big|$, lo cual implica que

$$\big||\mathcal{A}_{\boldsymbol{\lambda}}^{sq}| - \mathcal{T}(\boldsymbol{\lambda})\, q^{d-m}\big| = \mathcal{T}(\boldsymbol{\lambda})\, \big||V^{\neq}(\mathbb{F}_q)| - q^{d-m}\big|. \tag{7.16}$$

Si $p > 2$ y $3 \leq r \leq d-m$, por el Corolario 7.2.2 concluimos que

$$\begin{aligned}\big||\mathcal{A}_{\boldsymbol{\lambda}}^{sq}| - \mathcal{T}(\boldsymbol{\lambda})\, q^{d-m}\big| &\leq \mathcal{T}(\boldsymbol{\lambda})\Big((q+1)q^{d-m-2}\big((\delta_{\boldsymbol{L}}(D_{\boldsymbol{L}}-2)+2)q^{1/2}+14D_{\boldsymbol{L}}^2\delta_{\boldsymbol{L}}^2\big)+d^2\delta_{\boldsymbol{L}}q^{d-m-1}\Big)\\ &\leq q^{d-m-1}\mathcal{T}(\boldsymbol{\lambda})\,\big(2\delta_{\boldsymbol{L}}D_{\boldsymbol{L}}q^{1/2}+19\,D_{\boldsymbol{L}}^2\delta_{\boldsymbol{L}}^2+d^2\delta_{\boldsymbol{L}}\big).\end{aligned}$$

Así, tenemos la siguiente estimación para el número $|\mathcal{A}_{\boldsymbol{\lambda}}|$:

$$\begin{aligned}\big||\mathcal{A}_{\boldsymbol{\lambda}}| - \mathcal{T}(\boldsymbol{\lambda})\, q^{d-m}\big| &= \big||\mathcal{A}_{\boldsymbol{\lambda}}^{sq}| + |\mathcal{A}_{\boldsymbol{\lambda}}^{nsq}| - \mathcal{T}(\boldsymbol{\lambda})q^{n-m}\big|\\ &\leq q^{d-m-1}\mathcal{T}(\boldsymbol{\lambda})\,\big(2\delta_{\boldsymbol{L}}D_{\boldsymbol{L}}q^{1/2}+19\,D_{\boldsymbol{L}}^2\delta_{\boldsymbol{L}}^2+d^2\delta_{\boldsymbol{L}}\big)+|\mathcal{A}_{\boldsymbol{\lambda}}^{nsq}|.\end{aligned} \tag{7.17}$$

Finalmente, resta obtener una cota superior para el número $|\mathcal{A}_{\boldsymbol{\lambda}}^{nsq}|$. Con este propósito, observemos que $f \in \mathcal{A}$ no es libre de cuadrados si y solo si su discriminante es igual a cero. Sea $\mathcal{A}^{nsq}$ el lugar discriminante de $\mathcal{A}$, es decir, el conjunto de elementos de la familia $\mathcal{A}$ cuyo discriminante es igual a cero. De [FS84] es fácil deducir que el lugar discriminante $\mathcal{A}^{nsq}$ es el conjunto de puntos $\mathbb{F}_q$–racionales de una hipersuperficie de grado a lo sumo $d(d-1)$ de un espacio afín adecuado de dimensión $d-m$. De (2.2.1) deducimos que

$$|\mathcal{A}_{\boldsymbol{\lambda}}^{nsq}| \leq |\mathcal{A}^{nsq}| \leq d(d-1)\, q^{d-m-1}. \tag{7.18}$$

Así, combinando (7.17) y (7.18) concluimos que

$$\begin{aligned}\big||\mathcal{A}_{\boldsymbol{\lambda}}| - \mathcal{T}(\boldsymbol{\lambda})\, q^{d-m}\big| &\leq q^{d-m-1}\mathcal{T}(\boldsymbol{\lambda})\,\big(2\delta_{\boldsymbol{L}}D_{\boldsymbol{L}}q^{1/2}+19\,D_{\boldsymbol{L}}^2\delta_{\boldsymbol{L}}^2+d^2\delta_{\boldsymbol{L}}\big)+d^2q^{d-m-1}\\ &\leq q^{d-m-1}\big(2\mathcal{T}(\boldsymbol{\lambda})D_{\boldsymbol{L}}\delta_{\boldsymbol{L}}q^{1/2}+19\,\mathcal{T}(\boldsymbol{\lambda})\,D_{\boldsymbol{L}}^2\delta_{\boldsymbol{L}}^2+d^2\big).\end{aligned}$$

En otras palabras, tenemos el siguiente resultado.

Teorema 7.2.4. *Para $p > 2$, $q > d$ y $3 \leq r \leq d-m$, tenemos que*

$$\begin{aligned}\big||\mathcal{A}_{\boldsymbol{\lambda}}^{sq}| - \mathcal{T}(\boldsymbol{\lambda})\, q^{d-m}\big| &\leq q^{d-m-1}\mathcal{T}(\boldsymbol{\lambda})\,\big(2\,D_{\boldsymbol{L}}\delta_{\boldsymbol{L}}q^{\frac{1}{2}}+19\,D_{\boldsymbol{L}}^2\delta_{\boldsymbol{L}}^2+d^2\delta_{\boldsymbol{L}}\big),\\ \big||\mathcal{A}_{\boldsymbol{\lambda}}| - \mathcal{T}(\boldsymbol{\lambda})\, q^{d-m}\big| &\leq q^{d-m-1}\big(2\,\mathcal{T}(\boldsymbol{\lambda})\,D_{\boldsymbol{L}}\delta_{\boldsymbol{L}}q^{\frac{1}{2}}+19\,\mathcal{T}(\boldsymbol{\lambda})\,D_{\boldsymbol{L}}^2\delta_{\boldsymbol{L}}^2+d^2\big).\end{aligned}$$

donde $\delta_{\boldsymbol{L}} := i_1 \cdots i_m$ y $D_{\boldsymbol{L}} := \sum_{j=1}^{m}(i_j - 1)$.

Por otro lado, si suponemos que $m+2 \leq r \leq d-m$, combinando el Teorema 7.2.3 y la estimación (7.16) deducimos que, sin restricción sobre la característica de $\mathbb{F}_q$,

$$\begin{aligned}\big||\mathcal{A}_{\boldsymbol{\lambda}}^{sq}| - \mathcal{T}(\boldsymbol{\lambda})\, q^{d-m}\big| &\leq \mathcal{T}(\boldsymbol{\lambda})\big(14\,D_{\boldsymbol{L}}^3\delta_{\boldsymbol{L}}^2(q+1)q^{d-m-2}+d^2\delta_{\boldsymbol{L}}q^{d-m-1}\big)\\ &\leq q^{d-m-1}\mathcal{T}(\boldsymbol{\lambda})\,\big(21\,D_{\boldsymbol{L}}^3\delta_{\boldsymbol{L}}^2+d^2\delta_{\boldsymbol{L}}\big).\end{aligned}$$

En este caso una estimación para el número $|\mathcal{A}_{\boldsymbol{\lambda}}|$ es:

$$\begin{aligned}\big||\mathcal{A}_{\boldsymbol{\lambda}}| - \mathcal{T}(\boldsymbol{\lambda})\, q^{d-m}\big| &= \big||\mathcal{A}_{\boldsymbol{\lambda}}^{sq}| + |\mathcal{A}_{\boldsymbol{\lambda}}^{nsq}| - \mathcal{T}(\boldsymbol{\lambda})q^{d-m}\big|\\ &\leq q^{d-m-1}\mathcal{T}(\boldsymbol{\lambda})\,\big(21\,D_{\boldsymbol{L}}^3\delta_{\boldsymbol{L}}^2+d^2\delta_{\boldsymbol{L}}\big)+|\mathcal{A}_{\boldsymbol{\lambda}}^{nsq}|.\end{aligned} \tag{7.19}$$

Combinando (7.18) y (7.19) concluimos que

$$\begin{aligned}\big||\mathcal{A}_{\boldsymbol{\lambda}}| - \mathcal{T}(\boldsymbol{\lambda})\, q^{d-m}\big| &\le q^{d-m-1}\mathcal{T}(\boldsymbol{\lambda})\left(21\, D_{\boldsymbol{L}}^3\delta_{\boldsymbol{L}}^2 + d^2\delta_{\boldsymbol{L}}\right) + |\mathcal{A}_{d,\boldsymbol{\lambda}}^{nsq}| \\ &\le q^{d-m-1}\big(21\,\mathcal{T}(\boldsymbol{\lambda})\, D_{\boldsymbol{L}}^3\delta_{\boldsymbol{L}}^2 + \mathcal{T}(\boldsymbol{\lambda})\, d^2\delta_{\boldsymbol{L}} + d^2\big).\end{aligned}$$

En otras palabras, obtenemos el siguiente resultado.

Teorema 7.2.5. *Para $q > d$ y $m+2 \le r \le d-m$, tenemos que*

$$\begin{aligned}\big||\mathcal{A}_{\boldsymbol{\lambda}}^{sq}| - \mathcal{T}(\boldsymbol{\lambda})\, q^{d-m}\big| &\le q^{d-m-1}\mathcal{T}(\boldsymbol{\lambda})\left(21\, D_{\boldsymbol{L}}^3\delta_{\boldsymbol{L}}^2 + d^2\delta_{\boldsymbol{L}}\right), \\ \big||\mathcal{A}_{\boldsymbol{\lambda}}| - \mathcal{T}(\boldsymbol{\lambda})\, q^{d-m}\big| &\le q^{d-m-1}\big(21\,\mathcal{T}(\boldsymbol{\lambda})\, D_{\boldsymbol{L}}^3\delta_{\boldsymbol{L}}^2 + \mathcal{T}(\boldsymbol{\lambda})\, d^2\delta_{\boldsymbol{L}} + d^2\big),\end{aligned}$$

donde $\delta_{\boldsymbol{L}} := i_1 \cdots i_m$ y $D_{\boldsymbol{L}} := \sum_{j=1}^m (i_j - 1)$.

Si comparamos las estimaciones de $|\mathcal{A}_{\boldsymbol{\lambda}}|$ que obtuvimos, observamos que el Teorema 7.2.5 muestra que $|\mathcal{A}_{\boldsymbol{\lambda}}| = \mathcal{T}(\boldsymbol{\lambda})q^{d-m} + \mathcal{O}(q^{d-m-1})$, mientras que el Teorema 7.2.4 muestra que $|\mathcal{A}_{\boldsymbol{\lambda}}| = \mathcal{T}(\boldsymbol{\lambda})q^{d-m} + \mathcal{O}(q^{d-m-1/2})$. Para $q \ge (11D_{\boldsymbol{L}}^2\delta_{\boldsymbol{L}})^2$, la cota superior para $\big||\mathcal{A}_{\boldsymbol{\lambda}}| - \mathcal{T}(\boldsymbol{\lambda})\, q^{d-m}\big|$ del Teorema 7.2.5 es menor que la cota superior del Teorema 7.2.4. Observamos también que el Teorema 7.2.5 vale sin restricciones sobre la característica p de $\mathbb{F}_q$, mientras que el Teorema 7.2.4 vale para $p > 2$. Por otro lado, el Teorema 7.2.4 es válido para un rango grande de valores de m, es decir, $1 \le m \le d-3$, mientras que el Teorema 7.2.5 requiere que $1 \le m \le d/2 - 1$. Podemos hacer observaciones similares para el número $\big||\mathcal{A}_{\boldsymbol{\lambda}}^{sq}| - \mathcal{T}(\boldsymbol{\lambda})\, q^{d-m}\big|$. Resumiendo, podemos decir que ambos resultados resultan complementarios.

7.2.1. Polinomios con coeficientes prescriptos y aplicaciones

En esta sección aplicamos los Teoremas 7.2.4 y 7.2.5 a familias de elementos de $\mathbb{F}_q[T]_d$ con coeficientes prescriptos. Dados $0 < i_1 < i_2 < \cdots < i_m \le d$ y $\boldsymbol{b}_0 := (b_{i_1,0}, \ldots, b_{i_m,0}) \in \mathbb{F}_q^m$, sea $\mathcal{I} := \{i_1, \ldots, i_m\}$ y consideremos el conjunto $\mathcal{A}^{\mathcal{I}}$ definido de la siguiente manera:

$$\mathcal{A}^{\mathcal{I}} := \left\{T^d + a_1T^{d-1} + \cdots + a_d \in \mathbb{F}_q[T]_d : a_{i_j} = b_{i_j,0}\ (1 \le j \le m)\right\}. \tag{7.20}$$

Además, denotamos por $\mathcal{A}^{\mathcal{I},sq}$ el conjunto de elementos $f \in \mathcal{A}^{\mathcal{I}}$ que son libres de cuadrados. Dado un patrón de factorización $\boldsymbol{\lambda}$, sea $G \in \mathbb{F}_q[\boldsymbol{X}, T]$ el polinomio definido en (7.7). Por el Lema 7.1.3, un elemento $f \in \mathcal{A}^{\mathcal{I}}$ tiene patrón de factorización $\boldsymbol{\lambda}$ si y solo si existe $\boldsymbol{x}$ de tipo $\boldsymbol{\lambda}$ tal que

$$(-1)^{i_j}\Pi_{i_j}(\boldsymbol{Y}(\boldsymbol{x})) = b_{i_j,0} \quad (1 \le j \le m).$$

Sea $\delta_{\mathcal{I}} := i_1 \cdots i_m$ y $D_{\mathcal{I}} := \sum_{j=1}^m (i_j - 1)$. De los Teoremas 7.2.4 y 7.2.5 deducimos el siguiente resultado.

Corolario 7.2.6. *Para $p > 2$, $q > d$ y $i_m \le d-3$, tenemos que*

$$\begin{aligned}\big||\mathcal{A}_{\boldsymbol{\lambda}}^{\mathcal{I},sq}| - \mathcal{T}(\boldsymbol{\lambda})\, q^{d-m}\big| &\le q^{d-m-1}\mathcal{T}(\boldsymbol{\lambda})\left(2\, D_{\mathcal{I}}\,\delta_{\mathcal{I}}\, q^{\frac{1}{2}} + 19\, D_{\mathcal{I}}^2\,\delta_{\mathcal{I}}^2 + d^2\delta_{\mathcal{I}}\right), \\ \big||\mathcal{A}_{\boldsymbol{\lambda}}^{\mathcal{I}}| - \mathcal{T}(\boldsymbol{\lambda})\, q^{d-m}\big| &\le q^{d-m-1}\big(2\,\mathcal{T}(\boldsymbol{\lambda})\, D_{\mathcal{I}}\,\delta_{\mathcal{I}}\, q^{\frac{1}{2}} + 19\,\mathcal{T}(\boldsymbol{\lambda})\, D_{\mathcal{I}}^2\,\delta_{\mathcal{I}}^2 + d^2\big).\end{aligned}$$

Por otro lado, si $q > d$ y $i_m \le d-m-2$, entonces

$$\big||\mathcal{A}_{\boldsymbol{\lambda}}^{\mathcal{I},sq}| - \mathcal{T}(\boldsymbol{\lambda})\, q^{d-m}\big| \le q^{d-m-1}\mathcal{T}(\boldsymbol{\lambda})\left(21\, D_{\mathcal{I}}^3\, \delta_{\mathcal{I}}^2 + d^2\delta_{\mathcal{I}}\right),$$
$$\big||\mathcal{A}_{\boldsymbol{\lambda}}^{\mathcal{I}}| - \mathcal{T}(\boldsymbol{\lambda})\, q^{d-m}\big| \le q^{d-m-1}\big(21\, \mathcal{T}(\boldsymbol{\lambda})\, D_{\mathcal{I}}^3\, \delta_{\mathcal{I}}^2 + d^2\big).$$

Observemos que, por simplicidad, cambiamos la enumeración de los coeficientes de los elementos de la familia $\mathcal{A}^{\mathcal{I}}$ de (7.20) con respecto a la familia $\mathcal{A}$ definida en (7.3). Por lo tanto, las condiciones $3 \le r \le d-m$ en el Teorema 7.2.4 y $m+2 \le r \le d-m$ en el Teorema 7.2.5 se expresan en este caso como $i_m \le d-3$ y $i_m \le d-m-2$ respectivamente.

Terminamos esta sección aplicando nuestras estimaciones al caso de familias de polinomios con coeficientes consecutivos prescriptos y considerando el patrón de factorización $\boldsymbol{\lambda}^* := 1^d$, es decir, consideramos los polinomios que se factorizan en factores lineales sobre $\mathbb{F}_q$. Más precisamente, para $1 \le m \le d-3$ y $\boldsymbol{b}_0 := (b_{1,0}, \ldots, b_{m,0}) \in \mathbb{F}_q^m$, sea

$$\mathcal{A}^m := \left\{1 + a_1T + \cdots + a_dT^d \in \mathbb{F}_q[T]_d : a_j = b_{j,0}\ (1 \le j \le m)\right\}.$$

Queremos obtener una estimación asintótica del número $|\mathcal{A}_{\boldsymbol{\lambda}^*}^m|$ y dar condiciones de existencia de un elemento en $f \in \mathcal{A}_{\boldsymbol{\lambda}^*}^m$. Este problema es importante en la teoría de códigos, por ejemplo, en la decodificación de códigos de Reed–Solomon (ver [CW10]) y en la determinación de la distancia mínima de un código lineal (ver [CW12]). También tiene aplicaciones en la teoría de grafos (ver [Coh94, Coh98]).

Observemos que la presentación de la familia $\mathcal{A}^m$ difiere de la familia $\mathcal{A}^{\mathcal{I}}$ definida en (7.20), ya que ahora fijamos los primeros $m+1$ coeficientes. Sin embargo, considerando los polinomios recíprocos de los elementos de $\mathcal{A}^m$ podemos aplicar el Corolario 7.2.6 a este caso.

Corolario 7.2.7. *Para $p > 2$, $q > d$ y $m \le d-3$, tenemos que*

$$\left||\mathcal{A}_{\boldsymbol{\lambda}^*}^m| - \frac{q^{d-m}}{d!}\right| \le \frac{m(m-1)\,m!}{d!}\, q^{d-m-\frac{1}{2}} + \left(5\,\frac{m^2(m-1)^2\,m!^2}{d!} + d^2\right)q^{d-m-1}.$$

Más aún, para $q > 44m^4m!^2 + 8d^2m!$ existe un elemento en $f \in \mathcal{A}_{\boldsymbol{\lambda}}^m$. Por otro lado, si $q > d$ y $2m+2 \le d$, entonces

$$\left||\mathcal{A}_{\boldsymbol{\lambda}^*}^m| - \frac{q^{d-m}}{d!}\right| \le q^{d-m-1}\left(\frac{3\,m^3(m-1)^3m!^2}{d!} + d^2\right).$$

Más aún, para $q > 3m^6m!^2 + d^2m!$ existe un elemento $f \in \mathcal{A}_{\boldsymbol{\lambda}^*}^m$.

Demostración. Observemos que un polinomio $f \in \mathcal{A}^m$ se factoriza en factores lineales sobre $\mathbb{F}_q$ si y solo si su polinomio recíproco $T^d f(T^{-1})$ también lo hace. Además, es claro que $\mathcal{T}(\boldsymbol{\lambda}^*) = 1/d!$. Por lo tanto, del Corolario 7.2.6 se deducen fácilmente las estimaciones del corolario.

Más aún, denotamos por $\mathcal{A}^{m,sq}$ el conjunto de $f \in \mathcal{A}^m$ que son libres de cuadrados. Si $p > 2$, $q > d$ y $m \le d-3$, entonces el Corolario 7.2.6 implica que

$$\left||\mathcal{A}_{\boldsymbol{\lambda}^*}^{m,sq}| - \frac{q^{d-m}}{d!}\right| \le \frac{q^{d-m-1}}{d!}\left(m(m-1)\,m!\,q^{\frac{1}{2}} + 5\,m^2(m-1)^2\,m!^2 + d^2m!\right).$$

Así, se sigue que $|\mathcal{A}^{m,sq}_{\boldsymbol{\lambda}^*}| > 0$ cuando

$$q > m(m-1)\, m!\, q^{\frac{1}{2}} + 5\, m^2(m-1)^2\, m!^2 + d^2 m!.$$

Deducimos fácilmente la primera afirmación de existencia de elementos de $\mathcal{A}^m_{d,\boldsymbol{\lambda}^*}$.

Finalmente, para $q > d$ y $2m+2 \leq d$, del Corolario 7.2.6 tenemos que

$$\left| |\mathcal{A}^{m,sq}_{d,\boldsymbol{\lambda}^*}| - \frac{q^{d-m}}{d!} \right| \leq \frac{q^{d-m-1}}{d!} \left(3m^3(m-1)^3 m!^2 + d^2 m!\right),$$

lo cual implica la segunda afirmación de existencia de elementos de $\mathcal{A}^m_{d,\boldsymbol{\lambda}^*}$. □

En [Coh98], Cohen prueba que, para $1 \leq m \leq d-2$ y $q > (d^2(d+2)!)^2$, existe un elemento $f \in \mathcal{A}^m_{\boldsymbol{\lambda}^*}$. El Corolario 7.2.7 mejora significativamente este resultado, ya que la dependencia de d en la condición de q se reemplaza por la de m. En particular, si fijamos m mostramos la existencia de un elemento de $\mathcal{A}^m_{\boldsymbol{\lambda}^*}$ para valores de q del orden de $\mathcal{O}(d^2)$.

Por otro lado, en [LW10] se obtiene la siguiente estimación:

$$\left| |\mathcal{A}^m_{\boldsymbol{\lambda}^*}| - \frac{1}{q^m} \binom{q}{d} \right| \leq \binom{q/p + (m-1)\sqrt{q} + d - 1}{d}. \tag{7.21}$$

De la estimación (7.21) los autores concluyen que, para cualquier $\varepsilon > 0$ existe una constante $c_\varepsilon > 0$ tal que, si $m < \varepsilon d^{1/2}$ y $4\varepsilon^2 \ln^2 q < d \leq c_\varepsilon q$, existe $f \in \mathcal{A}^m_{\boldsymbol{\lambda}^*}$. La estimación del Corolario 7.2.7 mejora (7.21) en varios casos importantes. En particular, es válida cuando el radio q/p es grande, es decir, para cuerpos grandes de característica pequeña.

Capítulo 8

Búsqueda de puntos $\mathbb{F}_q$–racionales en hipersuperficies

En este capítulo comenzamos describiendo un algoritmo probabilístico que calcula puntos $\mathbb{F}_q$–racionales de hipersuperficies, en base a la estrategia de "búsqueda en bandas verticales". Esta estrategia, que en el caso de curvas planas fue analizada en el trabajo de J. von zur Gathen y colaboradores [vzGSS03], consiste en reducir el problema original al de encontrar puntos $\mathbb{F}_q$–racionales en la intersección de la hipersuperficie en cuestión con rectas paralelas en una dirección dada. Luego analizamos el algoritmo propuesto desde un punto de vista probabilístico. Con este propósito, observamos que su comportamiento está determinado por el número de bandas verticales que deben generarse hasta hallar un punto $\mathbb{F}_q$–racional de la hipersuperficie dada. Utilizando las herramientas técnicas desarrolladas en los Capítulos 5 y 6, determinamos el comportamiento asintótico de la distribución de probabilidad del número de búsquedas que se debe realizar. Este estudio nos permitirá, al final del capítulo, obtener una cota superior de la complejidad en promedio de dicho algoritmo.

8.1. Algoritmo BBV para hipersuperficies

Fijamos enteros $r \geq 2$ y $d \geq 2$. Sean $X_1, \ldots, X_r$ indeterminadas sobre $\mathbb{F}_q$, sean $\boldsymbol{X} := (X_1, \ldots, X_r)$ y $\mathbb{F}_q[\boldsymbol{X}]$ el anillo de polinomios en $\boldsymbol{X}$ con coeficientes en $\mathbb{F}_q$. Consideramos el conjunto $\mathbb{F}_q[\boldsymbol{X}]_{\leq d} := \{F \in \mathbb{F}_q[\boldsymbol{X}] : \deg(F) \leq d\}$ y un elemento arbitrario F de $\mathbb{F}_q[\boldsymbol{X}]_{\leq d}$. El propósito de esta sección es describir un algoritmo probabilístico que calcula un cero $\mathbb{F}_q$–racional de F, es decir, un punto $\boldsymbol{x} \in \mathbb{F}_q^r$ tal que $F(\boldsymbol{x}) = 0$.

Un dato fundamental para el diseño de un algoritmo para esta tarea es el estudio del conjunto de ceros $\mathbb{F}_q$–racionales de F. Al cardinal de dicho conjunto lo denotamos con $N(F)$. El cardinal promedio $N(F)$ cuando F varía entre todos los elementos de $\mathbb{F}_q[\boldsymbol{X}]_{\leq d}$ es q^{r-1} (ver (2.3)), cantidad que coincide con el número de elementos de $\mathbb{F}_q^{r-1}$. Esto sugiere una estrategia para encontrar un cero $\mathbb{F}_q$–racional de $F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$, que extiende las ideas propuestas en los trabajos de [vzGSS03] y [Mat10] al caso de

polinomios en varias variables. Como el número esperado de ceros de F es igual al número de elementos de $\mathbb{F}_q^{r-1}$, dado $\boldsymbol{a}_1 \in \mathbb{F}_q^{r-1}$, se puede tratar de encontrar un cero $\mathbb{F}_q$–racional de F de la forma $(\boldsymbol{a}_1, x_r)$, con $x_r \in \mathbb{F}_q$, o equivalentemente, un cero $\mathbb{F}_q$–racional de $F(\boldsymbol{a}_1, X_r)$. Si este polinomio no tiene ceros en $\mathbb{F}_q$, entonces, dado $\boldsymbol{a}_2 \in \mathbb{F}_q^{r-1}$, se determina si el polinomio $F(\boldsymbol{a}_2, X_r)$ tiene un cero en $\mathbb{F}_q$. El algoritmo procede de esta manera hasta que encuentra un cero de F en $\mathbb{F}_q^r$.

Siguiendo la terminología de [vzGSS03], que considera el caso $r = 2$, cada conjunto $\{\boldsymbol{a}_i\} \times \mathbb{F}_q$ se denomina una *banda vertical*. En consecuencia, el algoritmo correspondiente, que extiende el de [vzGSS03] al caso de polinomios en r variables, se denomina *Algoritmo de búsqueda en bandas verticales*, o **Algoritmo BBV**, y procede como describimos a continuación.

Algoritmo BBV.

Entrada: un polinomio $F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$.

Output: un cero $\boldsymbol{x} \in \mathbb{F}_q^r$ *de* F*, o "fracaso".*

Sea $i := 1$ *y* $f := 1$

Mientras $1 \leq i \leq q^{r-1}$ *y* $f = 1$ *hacer*

Elegir aleatoriamente $\boldsymbol{a}_i \in \mathbb{F}_q^{r-1} \setminus \{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{i-1}\}$
Calcular $f := \gcd(F(\boldsymbol{a}_i, X_r), X_r^q - X_r)$
Si $f = 0$*, entonces elegir* $x_{r,i} \in \mathbb{F}_q$ *aleatoriamente*
Si $f \notin \{0, 1\}$*, entonces se calcula una raíz* $x_{r,i} \in \mathbb{F}_q$ *de* f
$i := i + 1$

Fin mientras

Si $f \neq 1$ *devuelve* $(\boldsymbol{a}_i, x_{r,i})$*, sino devuelve "fracaso".*

Para estimar el costo de este algoritmo, es decir, la cantidad de operaciones aritméticas en $\mathbb{F}_q$ que éste realiza, vamos a suponer que los polinomios están representados por medio de su codificación densa, esto es, cada elemento de $\mathbb{F}_q[\boldsymbol{X}]_{\leq d}$ se representa por el vector de todos los coeficientes de F (sean éstos nulos o no) con un orden prefijado de todos los monomios de grado a lo sumo d de $\mathbb{F}_q[\boldsymbol{X}]$. Dado que un elemento de $\mathbb{F}_q[\boldsymbol{X}]_{\leq d}$ tiene $D := \binom{d+r}{r}$ coeficientes, la codificación densa de elementos de $\mathbb{F}_q[\boldsymbol{X}]_{\leq d}$ tiene longitud D.

Sin considerar el costo de generar elementos aleatorios de $\mathbb{F}_q^{r-1}$, en la i–ésima iteración del Algoritmo BBV calculamos el polinomio $F(\boldsymbol{a}_i, X_r)$. Si utilizamos el método de Horner multivariado para evaluar F en $\boldsymbol{a}_i$, y tenemos en cuenta que F tiene D coeficientes, el número de operaciones aritméticas en $\mathbb{F}_q$ que se necesitan para calcular el vector de coeficientes de $F(\boldsymbol{a}_i, X_r)$ es $\mathcal{O}^\sim(D)$, donde la notación $\mathcal{O}^\sim$ ignora factores logarítmicos (en los trabajos de [BE16] y [BES13] los autores dan una cota superior de la cantidad de productos en $\mathbb{F}_q$ que se necesita para evaluar un polinomio multivariado en un punto). Luego el algoritmo calcula el máximo común

divisor f y, si $f \neq 1$, calcula un cero de f en $\mathbb{F}_q$. Esto puede realizarse con $\mathcal{O}^{\sim}(d \log q)$ operaciones aritméticas en $\mathbb{F}_q$ (ver, por ejemplo, [vzGG99, Corolario 14.16]). Así, deducimos el siguiente resultado.

Lema 8.1.1. *Sean $F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$ y $\underline{\boldsymbol{a}} := (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{q^{r-1}})$ una elección de bandas verticales. El Algoritmo BBV realiza $\mathcal{O}^{\sim}\big(C_{\underline{\boldsymbol{a}}}(F) \cdot (D + d \log q)\big)$ operaciones aritméticas en $\mathbb{F}_q$ hasta encontrar un cero $\mathbb{F}_q$–racional de F, donde $C_{\underline{\boldsymbol{a}}}(F)$ es el mínimo i tal que $F(\boldsymbol{a}_i, X_r)$ tiene un cero en $\mathbb{F}_q$.*

En tal sentido, cabe preguntarse cuántas bandas verticales se necesitan para que el Algoritmo BBV encuentre un cero $\mathbb{F}_q$–racional del polinomio en consideración.

Los algoritmos probabilísticos propuestos en [vzGSS03] para el caso de curvas, y [Mat10] para el caso de hipersuperficies, dan una respuesta a dicha pregunta. Estos algoritmos proponen como máximo d búsquedas a fin de obtener una probabilidad de éxito mayor que $1/2$. Más precisamente, tenemos el siguiente resultado.

Lema 8.1.2. *Si $F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$ es absolutamente irreducible y $q > 15d^{13/3}$ se tiene que, con a lo sumo d elecciones aleatorias, es posible encontrar un cero $\mathbb{F}_q$–racional en una banda vertical con probabilidad al menos $1/2$.*

Demostración. Denotamos con $V(F) \subset \mathbb{A}^r$ a la hipersuperficie definida por F. Como F es absolutamente irreducible y $q > 15d^{13/3}$, del Teorema 2.2.5 deducimos que

$$|V(F)(\mathbb{F}_q)| \geq q^{r-1} - d^2 q^{r-3/2}. \tag{8.1}$$

Sea $\boldsymbol{a} \in \mathbb{F}_q^{r-1}$ y sea $C_{\boldsymbol{a}}(\mathbb{F}_q) := \{x \in \mathbb{F}_q : F(\boldsymbol{a}, x) = 0\}$ el conjunto de ceros $\mathbb{F}_q$–racionales de F en la banda vertical definida por $\boldsymbol{a}$. Observemos que, para cada $\boldsymbol{a} \in \mathbb{F}_q^{r-1}$, existen a lo sumo d elementos en $C_{\boldsymbol{a}}(\mathbb{F}_q)$. Combinando este hecho con (8.1) obtenemos la estimación

$$\big|\{\boldsymbol{a} \in \mathbb{F}_q^{r-1} : C_{\boldsymbol{a}}(\mathbb{F}_q) \neq \emptyset\}\big| \geq \frac{q^{r-1} - d^2 q^{r-3/2}}{d}.$$

Así, si consideramos la probabilidad uniforme $P_{\mathbb{F}_q^{r-1}}$ sobre $\mathbb{F}_q^{r-1}$, por la desigualdad anterior obtenemos la siguiente cota inferior para la probabilidad de que exista un cero $\mathbb{F}_q$–racional de F en la banda vertical determinada por un elemento $\boldsymbol{a}$ elegido aleatoriamente:

$$P_{\mathbb{F}_q^{r-1}}[\boldsymbol{a} \in \mathbb{F}_q^{r-1} : C_{\boldsymbol{a}}(\mathbb{F}_q) \neq \emptyset] \geq \frac{1}{d}(1 - d^2 q^{-1/2}) \geq \frac{1}{2d}.$$

Se deduce así el lema. □

A partir del Lema 8.1.2, cabe preguntarse sobre la optimalidad de la cantidad de búsqueda en bandas aleatorias y la necesidad de la hipótesis de absoluta irreducibilidad.

Para contestar esto nos proponemos analizar el Algoritmo BBV desde un punto de vista probabilístico. Para este propósito, consideramos la probabilidad uniforme sobre $\mathbb{F}_q[\boldsymbol{X}]_{\leq d}$ y el parámetro que determina el costo del Algoritmo BBV, esto es, el

número de bandas verticales que deben ser generadas, como una variable aleatoriamente definida sobre $\mathbb{F}_q[\boldsymbol{X}]_{\leq d}$.

Dado $F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$ y $1 \leq s \leq q^{r-1}$, los elementos $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_s \in \mathbb{F}_q^{r-1}$ que definen las bandas verticales a considerar se eligen aleatoriamente sin repeticiones. Más precisamente, si el algoritmo ha realizado $s-1$ búsquedas sobre las bandas verticales determinadas por $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{s-1} \in \mathbb{F}_q^{r-1}$ sin tener éxito, entonces en el paso s genera aleatoriamente un elemento $\boldsymbol{a}_s \in \mathbb{F}_q^{r-1} \setminus \{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{s-1}\}$ y busca un cero $\mathbb{F}_q$–racional del polinomio univariado $F(\boldsymbol{a}_s, X_r)$.

En las siguientes secciones nos ocupamos de analizar la distribución de probabilidad del número de bandas verticales que deben ser generadas por el Algoritmo BBV hasta encontrar un cero $\mathbb{F}_q$–racional de F. Con este propósito, consideramos el conjunto F de todas las posibles elecciones de bandas verticales y la variable aleatoria $C : \mathsf{F} \times \mathbb{F}_q[\boldsymbol{X}]_{\leq d} \to \mathbb{N} \cup \{\infty\}$ que cuenta el número de bandas verticales que deben ser generadas, y determinamos el comportamiento asintótico de la probabilidad de éxito en las primeras s bandas verticales, para $s \leq \binom{d/2+r-1}{r-1}$ y para $s \leq \binom{d+r-3}{r-1}$, si $p > 2$. Este estudio nos permitirá, al final del capítulo, contestar las cuestiones anteriores y obtener una cota superior de la complejidad en promedio del Algoritmo BBV. Más precisamente, vamos a mostrar que en promedio dicho algoritmo necesita a lo sumo $1/\mu_d \approx 1{,}58$ bandas verticales para obtener un cero $\mathbb{F}_q$–racional del polinomio de entrada.

8.2. Probabilidad de éxito en las primeras 2 bandas verticales

En esta sección analizamos la probabilidad de que el Algoritmo BBV encuentre un cero $\mathbb{F}_q$–racional en una o dos bandas verticales. Los resultados muestran que existe una alta probabilidad, cercana a $0{,}865\ldots$, de que el Algoritmo BBV encuentre un cero $\mathbb{F}_q$–racional del polinomio de partida en a lo sumo dos bandas verticales. Por este motivo, una estimación precisa de estas probabilidades nos permitirá dar una mejor descripción del comportamiento del algoritmo.

8.2.1. Probabilidad de éxito en la primera banda vertical

Sea $F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$. Queremos estimar la probabilidad de que el Algoritmo BBV encuentre un cero $\mathbb{F}_q$–racional de F en la primera banda vertical. Cada posible elección de esta primera banda está representada por un elemento de $\mathbb{F}_q^{r-1}$. Por lo tanto, podemos representar esta situación mediante la variable aleatoria $C_1 := C_{1,r,d} : \mathbb{F}_q^{r-1} \times \mathbb{F}_q[\boldsymbol{X}]_{\leq d} \to \{1, \infty\}$ definida como sigue:

$$C_1(\boldsymbol{a}, F) := \begin{cases} 1 & \text{si } F(\boldsymbol{a}, X_r) \text{ tiene un cero } \mathbb{F}_q\text{–racional,} \\ \infty & \text{si no.} \end{cases}$$

Como r y d están fijos, eliminamos dichos índices de las notaciones. Consideramos la probabilidad uniforme $P_1 := P_{1,r,d}$ sobre el conjunto $\mathbb{F}_q^{r-1} \times \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$ y estudiamos

la probabilidad del evento $[C_1 = 1]$. En el siguiente resultado damos una fórmula exacta de esta probabilidad.

Teorema 8.2.1. *Para $q > d$, tenemos la siguiente identidad:*

$$P_1[C_1 = 1] = \sum_{j=1}^{d} (-1)^{j-1} \binom{q}{j} q^{-j} + (-1)^d \binom{q-1}{d} q^{-d-1}.$$

Demostración. Para $F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$, denotamos por $VS(F)$ el conjunto de las bandas verticales donde F tiene un cero $\mathbb{F}_q$–racional y por $NS(F)$ su cardinal, es decir,

$$VS(F) := \{\boldsymbol{a} \in \mathbb{F}_q^{r-1} : (\exists\, x_r \in \mathbb{F}_q)\ F(\boldsymbol{a}, x_r) = 0\}, \quad NS(F) := |VS(F)|.$$

Es fácil ver que

$$\{C_1 = 1\} = \{(\boldsymbol{a}, F) \in \mathbb{F}_q^{r-1} \times \mathbb{F}_q[\boldsymbol{X}]_{\leq d} : C_1(\boldsymbol{a}, F) = 1\} = \bigcup_{F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}} VS(F) \times \{F\}.$$

Dado que expresamos a $\{C_1 = 1\}$ como una unión de subconjuntos disjuntos de $\mathbb{F}_q^{r-1} \times \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$, se sigue que

$$P_1[C_1 = 1] = \frac{1}{q^{r-1} |\mathbb{F}_q[\boldsymbol{X}]_{\leq d}|} \sum_{F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}} NS(F). \tag{8.2}$$

Fijemos $F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$. Observemos que

$$VS(F) = \bigcup_{x \in \mathbb{F}_q} \{\boldsymbol{a} \in \mathbb{F}_q^{r-1} : F(\boldsymbol{a}, x) = 0\}.$$

Por el principio de inclusión-exclusión obtenemos que

$$\begin{aligned} NS(F) &= \left| \bigcup_{x \in \mathbb{F}_q} \{\boldsymbol{a} \in \mathbb{F}_q^{r-1} : F(\boldsymbol{a}, x) = 0\} \right| \\ &= \sum_{j=1}^{q} (-1)^{j-1} \sum_{\mathcal{X}_j \subset \mathbb{F}_q} \left| \{\boldsymbol{a} \in \mathbb{F}_q^{r-1} : (\forall x \in \mathcal{X}_j)\, F(\boldsymbol{a}, x) = 0\} \right|, \end{aligned}$$

donde $\mathcal{X}_j$ recorre todos los subconjuntos de $\mathbb{F}_q$ de cardinal j. Concluimos que

$$\begin{aligned} \sum_{F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}} NS(F) &= \sum_{F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}} \sum_{j=1}^{q} (-1)^{j-1} \sum_{\mathcal{X}_j \subset \mathbb{F}_q} \left| \{\boldsymbol{a} \in \mathbb{F}_q^{r-1} : (\forall x \in \mathcal{X}_j)\, F(\boldsymbol{a}, x) = 0\} \right| \\ &= \sum_{j=1}^{q} (-1)^{j-1} \sum_{\mathcal{X}_j \subset \mathbb{F}_q} \sum_{F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}} \left| \{\boldsymbol{a} \in \mathbb{F}_q^{r-1} : (\forall x \in \mathcal{X}_j)\, F(\boldsymbol{a}, x) = 0\} \right| \\ &= \sum_{j=1}^{q} (-1)^{j-1} \sum_{\mathcal{X}_j \subset \mathbb{F}_q} \sum_{\boldsymbol{a} \in \mathbb{F}_q^{r-1}} \left| \{F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d} : (\forall x \in \mathcal{X}_j)\, F(\boldsymbol{a}, x) = 0\} \right| \end{aligned}$$

Para $\boldsymbol{a} \in \mathbb{F}_q^{r-1}$ y un subconjunto $\mathcal{X} \subset \mathbb{F}_q$, denotamos con

$$\mathcal{S}_{\underline{\boldsymbol{a}}}(\mathcal{X}) := \{F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d} : (\forall x \in \mathcal{X})\, F(\boldsymbol{a}, x) = 0\}.$$

Se sigue así que

$$\sum_{F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}} NS(F) = \sum_{j=1}^{q} (-1)^{j-1} \sum_{\mathcal{X}_j \subset \mathbb{F}_q} \sum_{\boldsymbol{a} \in \mathbb{F}_q^{r-1}} |\mathcal{S}_{\underline{\boldsymbol{a}}}(\mathcal{X}_j)|. \tag{8.3}$$

Para cualquier j con $1 \leq j \leq q$, denotamos por

$$\mathcal{N}_j := \frac{1}{q^{r-1}|\mathbb{F}_q[\boldsymbol{X}]_{\leq d}|} \sum_{\mathcal{X}_j \subset \mathbb{F}_q} \sum_{\boldsymbol{a} \in \mathbb{F}_q^{r-1}} |\mathcal{S}_{\underline{\boldsymbol{a}}}(\mathcal{X}_j)|.$$

Fijemos un conjunto $\mathcal{X}_j \subset \mathbb{F}_q$ de j elementos. Si $j \leq d$ y $\underline{\boldsymbol{a}}$ está fijo, entonces las igualdades $F(\boldsymbol{a}, x) = 0$ para cada $x \in \mathcal{X}_j$ determinan un sistema de j ecuaciones linealmente independientes, cuyas incógnitas son los coeficientes de F en el $\mathbb{F}_q$–espacio vectorial $\mathbb{F}_q[\boldsymbol{X}]_{\leq d}$. Así tenemos $q^{\dim \mathbb{F}_q[\boldsymbol{X}]_{\leq d} - j}$ soluciones de dicho sistema, o sea $|\mathcal{S}_{\underline{\boldsymbol{a}}}(\mathcal{X}_j)| = q^{\dim \mathbb{F}_q[\boldsymbol{X}]_{\leq d} - j}$. Por lo tanto,

$$\begin{aligned} \mathcal{N}_j &= \frac{1}{q^{r-1+\dim \mathbb{F}_q[\boldsymbol{X}]_{\leq d}}} \sum_{\mathcal{X}_j \subset \mathbb{F}_q} \sum_{\boldsymbol{a} \in \mathbb{F}_q^{r-1}} |\mathcal{S}_{\underline{\boldsymbol{a}}}(\mathcal{X}_j)| \\ &= \frac{1}{q^{r-1+\dim \mathbb{F}_q[\boldsymbol{X}]_{\leq d}}} \sum_{\mathcal{X}_j \subset \mathbb{F}_q} \sum_{\boldsymbol{a} \in \mathbb{F}_q^{r-1}} q^{\dim \mathbb{F}_q[\boldsymbol{X}]_{\leq d} - j} = \binom{q}{j} q^{-j}. \end{aligned} \tag{8.4}$$

Por otro lado, si $j > d$, entonces la condición $F(\boldsymbol{a}, x) = 0$ se satisface para todo $x \in \mathcal{X}_j$ si y solo si $F(\boldsymbol{a}, X_r) = 0$. Por lo tanto, la condición $F(\boldsymbol{a}, X_r) = 0$ se puede expresar mediante $d+1$ ecuaciones lineales, linealmente independientes, cuyas incógnitas son los coeficientes de F en $\mathbb{F}_q[\boldsymbol{X}]_{\leq d}$. Así tenemos que $|\mathcal{S}_{\underline{\boldsymbol{a}}}(\mathcal{X}_j)| = q^{\dim \mathbb{F}_q[\boldsymbol{X}]_{\leq d} - (d+1)}$. Concluimos que

$$\begin{aligned} \mathcal{N}_j &= \frac{1}{q^{r-1+\dim \mathbb{F}_q[\boldsymbol{X}]_{\leq d}}} \sum_{\mathcal{X}_j \subset \mathbb{F}_q} \sum_{\boldsymbol{a} \in \mathbb{F}_q^{r-1}} |\mathcal{S}_{\underline{\boldsymbol{a}}}(\mathcal{X}_j)| \\ &= \frac{1}{q^{r-1+\dim \mathbb{F}_q[\boldsymbol{X}]_{\leq d}}} \sum_{\mathcal{X}_j \subset \mathbb{F}_q} \sum_{\boldsymbol{a} \in \mathbb{F}_q^{r-1}} q^{\dim \mathbb{F}_q[\boldsymbol{X}]_{\leq d} - (d+1)} = \binom{q}{j} q^{-d-1}. \end{aligned} \tag{8.5}$$

Combinando (8.4) y (8.5) obtenemos que

$$P_1[C_1 = 1] = \sum_{j=1}^{q} (-1)^{j-1} \mathcal{N}_j = \sum_{j=1}^{d} (-1)^{j-1} \binom{q}{j} q^{-j} + \sum_{j=d+1}^{q} (-1)^{j-1} \binom{q}{j} q^{-d-1}.$$

Finalmente, usando la siguiente igualdad [GKP94, §5.1]:

$$\sum_{j=d+1}^{q} (-1)^{j-1} \binom{q}{j} = \sum_{j=0}^{d} (-1)^{j} \binom{q}{j} = (-1)^d \binom{q-1}{d}, \tag{8.6}$$

deducimos el resultado del teorema. □

El siguiente corolario muestra el comportamiento asintótico de la probabilidad $P_1[C_1 = 1]$.

Corolario 8.2.2. *Para $q > d$, tenemos*

$$\left|P_1[C_1 = 1] - \mu_d\right| \leq \frac{2}{q},$$

donde $\mu_d := \sum_{j=1}^{d} \frac{(-1)^{j-1}}{j!}$.

Demostración. Fijamos $d \geq 2$. Veamos que

$$P_1[C_1 = 1] = \mu_d + \mathcal{O}(q^{-1}).$$

Para mostrar esto, dado enteros positivos k, j con $k \leq j$, consideramos el número de Stirling de la primera clase $\begin{bmatrix} j \\ k \end{bmatrix}$, es decir, el número de permutaciones de j elementos con k ciclos disjuntos. Recordamos que valen las siguientes propiedades (ver (5.20)):

$$\begin{bmatrix} j \\ j \end{bmatrix} = 1, \quad \begin{bmatrix} j \\ j-1 \end{bmatrix} = \binom{j}{2}, \quad \sum_{k=0}^{j} \begin{bmatrix} j \\ k \end{bmatrix} = j!.$$

Usamos también la siguiente identidad que relaciona los números combinatorios y los números de Stirling de la primera clase (ver, por ejemplo, [GKP94, (6,13)]):

$$\binom{q}{j} = \sum_{k=0}^{j} \frac{(-1)^{j-k}}{j!} \begin{bmatrix} j \\ k \end{bmatrix} q^k. \tag{8.7}$$

De acuerdo al Teorema 8.2.1 y (8.7), tenemos que

$$\begin{aligned} P_1[C_1 = 1] &= \sum_{j=1}^{d} (-1)^{j-1} \sum_{k=0}^{j} \frac{(-1)^{j-k}}{j!} \begin{bmatrix} j \\ k \end{bmatrix} q^{k-j} + (-1)^d \binom{q-1}{d} q^{-d-1} \\ &= \sum_{j=1}^{d} \frac{(-1)^{j-1}}{j!} \begin{bmatrix} j \\ j \end{bmatrix} + \sum_{j=1}^{d} \frac{(-1)^j}{j!} \begin{bmatrix} j \\ j-1 \end{bmatrix} q^{-1} \\ &\quad + \sum_{j=1}^{d} \sum_{k=0}^{j-2} \frac{(-1)^{k-1}}{j!} \begin{bmatrix} j \\ k \end{bmatrix} q^{k-j} + (-1)^d \binom{q-1}{d} q^{-d-1}. \end{aligned}$$

Se sigue que

$$P_1[C_1 = 1] = \mu_d + \frac{1}{q} \sum_{j=1}^{d} \frac{(-1)^j}{j!} \binom{j}{2} - \sum_{j=1}^{d} \sum_{k=0}^{j-2} \frac{(-1)^k}{j!} \begin{bmatrix} j \\ k \end{bmatrix} q^{k-j} + \frac{(-1)^d}{q^{d+1}} \binom{q-1}{d}.$$

En consecuencia, para $d > 2$ obtenemos

$$\begin{aligned} |P_1[C_1 = 1] - \mu_d| &\leq \frac{1}{q} \left| \sum_{j=1}^{d} \frac{(-1)^j}{j!} \binom{j}{2} \right| + \sum_{j=1}^{d} \sum_{k=0}^{j-2} \frac{1}{j!} \begin{bmatrix} j \\ k \end{bmatrix} \frac{1}{q^2} + \frac{1}{q^{d+1}} \binom{q-1}{d} \\ &\leq \frac{1}{4q} + \frac{d}{q^2} + \frac{1}{2q}. \end{aligned}$$

Para $d = 2$, esta desigualdad se obtiene mediante un cálculo directo. Deducimos así el resultado del corolario. □

Notemos que, cuando d tiende a infinito, el número $P_1[C_1 = 1]$ tiende a $1-e^{-1} = 0{,}6321\ldots$, donde e denota la base del logaritmo natural. Esto explica los resultados numéricos de la primera fila de las tablas de las simulaciones de la Sección 8.5.

Terminamos esta sección mencionando que la probabilidad $P_1[C_1 = 1]$ está relacionada con la probabilidad de que un polinomio univariado de grado a lo sumo d tenga al menos una raíz en $\mathbb{F}_q$. Más precisamente, sea el conjunto $\mathbb{F}_q[T]_{\leq d}$ de los polinomios univariados de grado a lo sumo d y con coeficientes en $\mathbb{F}_q$. Consideremos la probabilidad uniforme $p_{1,d}$ sobre $\mathbb{F}_q[T]_{\leq d}$ y sea $N := N_{1,d} : \mathbb{F}_q[T]_{\leq d} \to \mathbb{Z}_{\geq 0}$ la variable aleatoria que cuenta el número de ceros $\mathbb{F}_q$–racionales de un elemento de $\mathbb{F}_q[T]_{\leq d}$, es decir,

$$N(f) := |\{x \in \mathbb{F}_q : \ f(x) = 0\}|.$$

La variable aleatoria $N_{1,d}$ ha sido estudiada implícitamente en la literatura; por ejemplo, en [KK90b, Theorem 3] se proporciona una fórmula explícita del número total de polinomios mónicos de grado d con coeficientes en $\mathbb{F}_q$ que tienen $k \leq q$ ceros distintos en $\mathbb{F}_q$. En [Coh73, §2] se proporciona una fórmula exacta del número de polinomios mónicos de grado d con coeficientes en $\mathbb{F}_q$ que no son divisibles por un factor lineal.

Lema 8.2.3. *Para $d < q$,*

$$p_{1,d}[N > 0] = P_1[C_1 = 1].$$

Demostración. Dado un conjunto $\mathcal{X}_j := \{x_1, \ldots, x_j\} \subset \mathbb{F}_q$ con j elementos, sea

$$\mathcal{S}_{\mathcal{X}_j} := \{f \in \mathbb{F}_q[T]_{\leq d} : f(x_1) = 0, \ldots, f(x_j) = 0\}.$$

Es fácil ver que $\mathcal{S}_{\mathcal{X}_j} = \bigcap_{i=1}^{j} \mathcal{S}_{\{x_i\}}$. Por lo tanto, por el principio inclusión-exclusión obtenemos que

$$\left|\{N > 0\}\right| = \left| \bigcup_{x \in \mathbb{F}_q} \mathcal{S}_{\{x\}} \right| = \sum_{j=1}^{q} (-1)^{j-1} \sum_{\mathcal{X}_j \subset \mathbb{F}_q} \left|\mathcal{S}_{\mathcal{X}_j}\right|.$$

Denotamos $\mathcal{S}^*_{\mathcal{X}_j} := \mathcal{S}_{\mathcal{X}_j} \setminus \{0\}$. Observemos que $|\mathcal{S}^*_{\mathcal{X}_j}| = 0$ para $j > d$. En efecto, si $f \in \mathcal{S}^*_{\mathcal{X}_j}$, entonces f es un polinomio no nulo de grado d con j raíces distintas, lo cual implica que $\mathcal{S}^*_{\mathcal{X}_j} = \emptyset$. Por lo tanto, podemos reescribir esta identidad de la siguiente manera:

$$\left|\{N > 0\}\right| = 1 + \sum_{j=1}^{d} (-1)^{j-1} \sum_{\mathcal{X}_j \subset \mathbb{F}_q} \left|\mathcal{S}^*_{\mathcal{X}_j}\right|. \tag{8.8}$$

Estimamos ahora el número $|\mathcal{S}^*_{\mathcal{X}_j}|$ para un conjunto dado $\mathcal{X}_j := \{x_1, \ldots, x_j\} \subset \mathbb{F}_q$ con $1 \leq j \leq d$. Como las condiciones $f(x_1) = 0, \ldots, f(x_j) = 0$ son ecuaciones

lineales en los coeficientes de f que resultan linealmente independientes, concluimos que $|\mathcal{S}_{\mathcal{X}_j}| = q^{d+1-j}$ y $|\mathcal{S}^*_{\mathcal{X}_j}| = q^{d+1-j} - 1$. En consecuencia, por (8.8) obtenemos

$$\big|\{N > 0\}\big| = 1 + \sum_{j=1}^{d}(-1)^{j-1} \sum_{\mathcal{X}_j \subset \mathbb{F}_q} (q^{d+1-j} - 1) = 1 + \sum_{j=1}^{d}(-1)^{j-1} \binom{q}{j} (q^{d+1-j} - 1).$$

Se sigue así inmediatamente el resultado del lema. □

8.2.2. Probabilidad de éxito en la segunda banda vertical

En lo que sigue analizamos la probabilidad de que el Algoritmo BBV realice exactamente dos búsquedas hasta encontrar un cero $\mathbb{F}_q$–racional del polinomio en consideración.

Observamos que cada posible elección de las primeras dos bandas verticales es un elemento $\underline{\boldsymbol{a}} := (\boldsymbol{a}_1, \boldsymbol{a}_2) \in \mathbb{F}_q^{r-1} \times \mathbb{F}_q^{r-1}$ con $\boldsymbol{a}_1 \neq \boldsymbol{a}_2$. Por lo tanto, denotamos con F_2 el conjunto de todas las posibles elecciones para las primeras dos bandas verticales y por N_2 su cardinal, es decir,

$$\mathsf{F}_2 := \{\underline{\boldsymbol{a}} := (\boldsymbol{a}_1, \boldsymbol{a}_2) \in \mathbb{F}_q^{r-1} \times \mathbb{F}_q^{r-1} : \boldsymbol{a}_1 \neq \boldsymbol{a}_2\}, \quad N_2 = |\mathsf{F}_2| = q^{r-1}(q^{r-1} - 1).$$

Consideramos la probabilidad uniforme $P_2 := P_{2,r,d}$ sobre el conjunto $\mathsf{F}_2 \times \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$ y la variable aleatoria $C_2 := C_{2,r,d} : \mathsf{F}_2 \times \mathbb{F}_q[\boldsymbol{X}]_{\leq d} \to \{1, 2, \infty\}$ definida como

$$C_2(\underline{\boldsymbol{a}}, F) := \begin{cases} 1 & \text{si } N_{1,d}(F(\boldsymbol{a}_1, X_r)) > 0, \\ 2 & \text{si } N_{1,d}(F(\boldsymbol{a}_1, X_r)) = 0 \text{ y } N_{1,d}(F(\boldsymbol{a}_2, X_r)) > 0, \\ \infty & \text{en otro caso.} \end{cases}$$

El objetivo es analizar la probabilidad $P_2[C_2 = 2]$. Con este propósito, en el siguiente lema expresamos la probabilidad $P_2[C_2 = 2]$ en términos de las probabilidades de las variables aleatorias $C_{\underline{\boldsymbol{a}}} := C_{\underline{\boldsymbol{a}},r,d} : \mathbb{F}_q[\boldsymbol{X}]_{\leq d} \to \{1, 2, \infty\}$ que cuentan el número de búsquedas que deben realizarse sobre las bandas verticales definidas por $\underline{\boldsymbol{a}} := (\boldsymbol{a}_1, \boldsymbol{a}_2) \in \mathsf{F}_2$ hasta que el algoritmo encuentra un cero $\mathbb{F}_q$–racional del polinomio de partida. Definimos $C_{\underline{\boldsymbol{a}}}(F) := \infty$ cuando F no tiene ceros $\mathbb{F}_q$–racionales en dichas bandas. Consideramos la probabilidad uniforme $p_{r,d}$ sobre el conjunto $\mathbb{F}_q[\boldsymbol{X}]_{\leq d}$.

Lema 8.2.4. *Tenemos que*

$$P_2[C_2 = 2] = \frac{1}{N_2} \sum_{\underline{\boldsymbol{a}} \in \mathsf{F}_2} p_{r,d}[C_{\underline{\boldsymbol{a}}} = 2].$$

Demostración. Observemos que el conjunto $\{C_2 = 2\}$ puede expresarse como una unión disjunta de la siguiente manera:

$$\{C = 2\} = \bigcup_{\underline{\boldsymbol{a}} \in \mathsf{F}_2} \{\underline{\boldsymbol{a}}\} \times \{F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d} : C_{\underline{\boldsymbol{a}}}(F) = 2\}.$$

Por lo tanto,

$$P_2[C_2 = 2] = \frac{1}{N_2} \sum_{\underline{\boldsymbol{a}} \in \mathsf{F}_2} \frac{\left|\{F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d} : C_{\underline{\boldsymbol{a}}}(F) = 2\}\right|}{|\mathbb{F}_q[\boldsymbol{X}]_{\leq d}|} = \frac{1}{N_2} \sum_{\underline{\boldsymbol{a}} \in \mathsf{F}_2} p_{r,d}[C_{\underline{\boldsymbol{a}}} = 2],$$

lo que demuestra el lema. □

En la siguiente proposición estimamos la probabilidad $p_{r,d}[C_{\underline{\boldsymbol{a}}} = 2]$.

Proposición 8.2.5. *Para $q > d$ y $\underline{\boldsymbol{a}} := (\boldsymbol{a}_1, \boldsymbol{a}_2) \in \mathsf{F}_2$, tenemos*

$$\left|p_{r,d}[C_{\underline{\boldsymbol{a}}} = 2] - \mu_d(1 - \mu_d)\right| \leq \frac{3}{q}.$$

Demostración. Observemos que

$$\{C_{\underline{\boldsymbol{a}}} = 2\} = \{F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d} : N_{1,d}(F(\boldsymbol{a}_2, T)) > 0\} \setminus \{F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d} : N_{1,d}(F(\boldsymbol{a}_1, T)) > 0\}. \tag{8.9}$$

Notemos que el número de elementos de $\mathbb{F}_q[\boldsymbol{X}]_{\leq d}$ que tiene al menos un cero $\mathbb{F}_q$–racional en la banda definida por $\boldsymbol{a}_2$ está determinado en el Teorema 8.2.1. Por lo tanto, resta determinar el número $N_{\underline{\boldsymbol{a}}}$ de elementos de $\mathbb{F}_q[\boldsymbol{X}]_{\leq d}$ que tienen al menos un cero $\mathbb{F}_q$–racional en las bandas definidas por $\boldsymbol{a}_1$ y $\boldsymbol{a}_2$. Tenemos que

$$N_{\underline{\boldsymbol{a}}} = \left| \bigcup_{x \in \mathbb{F}_q} \bigcup_{y \in \mathbb{F}_q} \{F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d} : F(\boldsymbol{a}_1, x) = F(\boldsymbol{a}_2, y) = 0\} \right|.$$

Dados subconjuntos $\mathcal{X} \subset \mathbb{F}_q$ e $\mathcal{Y} \subset \mathbb{F}_q$, denotamos

$$\mathcal{S}_{\underline{\boldsymbol{a}}}(\mathcal{X}, \mathcal{Y}) := \{F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d} : F(\boldsymbol{a}_1, x) = F(\boldsymbol{a}_2, y) = 0 \text{ para todo } x \in \mathcal{X} \text{ y } y \in \mathcal{Y}\}.$$

Por el principio de inclusión–exclusión tenemos que

$$N_{\underline{\boldsymbol{a}}} = \sum_{j=1}^{q} \sum_{k=1}^{q} (-1)^{j+k} \sum_{\mathcal{X}_j \subset \mathbb{F}_q} \sum_{\mathcal{Y}_k \subset \mathbb{F}_q} \left|\mathcal{S}_{\underline{\boldsymbol{a}}}(\mathcal{X}_j, \mathcal{Y}_k)\right|, \tag{8.10}$$

donde la suma recorre todos los subconjuntos $\mathcal{X}_j \subset \mathbb{F}_q$ e $\mathcal{Y}_k \subset \mathbb{F}_q$ de j y k elementos respectivamente.

Afirmación. $\dfrac{N_{\underline{\boldsymbol{a}}}}{|\mathbb{F}_q[\boldsymbol{X}]_{\leq d}|} = \big(P_1[C_1 = 1]\big)^2 + \frac{q-1}{q^{2d+2}} = \big(P_1[C_1 = 1]\big)^2 + \mathcal{O}(q^{-1})$.

Demostración de la afirmación. Para $1 \leq j, k \leq q$, sea

$$\mathcal{N}_{j,k} := \sum_{\mathcal{X}_j \subset \mathbb{F}_q} \sum_{\mathcal{Y}_k \subset \mathbb{F}_q} |\mathcal{S}_{\underline{\boldsymbol{a}}}(\mathcal{X}_j, \mathcal{Y}_k)|.$$

Calculemos $\mathcal{N}_{j,k}$ para los siguientes cuatros casos.

Supongamos primero que $j, k \leq d$. Como $\boldsymbol{a}_1 \neq \boldsymbol{a}_2$, las igualdades $F(\boldsymbol{a}_1, x) = 0, F(\boldsymbol{a}_2, y) = 0$ para todo $x \in \mathcal{X}_j$ e $y \in \mathcal{Y}_k$ determinan un sistema de $j+k$ ecuaciones linealmente independientes cuyas incógnitas son los coeficientes de $F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$. Por lo tanto, $|\mathcal{S}_{\underline{\boldsymbol{a}}}(\mathcal{X}_j, \mathcal{Y}_k)| = q^{\dim \mathbb{F}_q[\boldsymbol{X}]_{\leq d} - j - k}$, lo cual implica que

$$\mathcal{N}_{j,k} = \sum_{\mathcal{X}_j \subset \mathbb{F}_q} \sum_{\mathcal{Y}_k \subset \mathbb{F}_q} q^{\dim \mathbb{F}_q[\boldsymbol{X}]_{\leq d} - j - k} = \binom{q}{j}\binom{q}{k} q^{\dim \mathbb{F}_q[\boldsymbol{X}]_{\leq d} - j - k}.$$

Supongamos ahora que $j > d$ y $k \leq d$. Por un lado, si $j > d$ y $\mathcal{X}_j \subset \mathbb{F}_q$ es un subconjunto de cardinal j, entonces la condición $F(\boldsymbol{a}_1, x) = 0$ se satisface para todo $x \in \mathcal{X}_j$ si y solo si $F(\boldsymbol{a}_1, X_r) = 0$, condición que se puede expresar mediante $d+1$ ecuaciones lineales en los coeficientes de $F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$ que resultan linealmente independientes. Por otro lado, las igualdades $F(\boldsymbol{a}_2, y) = 0$ para todo $y \in \mathcal{Y}_k$ imponen k condiciones adicionales sobre los coeficientes de F que resultan linealmente independientes. Así tenemos que $|\mathcal{S}_{\underline{\boldsymbol{a}}}(\mathcal{X}_j, \mathcal{Y}_k)| = q^{\dim \mathbb{F}_q[\boldsymbol{X}]_{\leq d} - (d+1) - k}$. Esto implica que

$$\mathcal{N}_{j,k} = \sum_{\mathcal{X}_j, \mathcal{Y}_k \subset \mathbb{F}_q} q^{\dim \mathbb{F}_q[\boldsymbol{X}]_{\leq d} - (d+1) - k} = \binom{q}{j}\binom{q}{k} q^{\dim \mathbb{F}_q[\boldsymbol{X}]_{\leq d} - (d+1) - k}.$$

El caso $j \leq d$ y $k > d$ resulta completamente análogo al segundo caso. Finalmente, cuando $j > d$ y $k > d$, las condiciones $F(\boldsymbol{a}_1, x) = 0$ y $F(\boldsymbol{a}_2, y) = 0$ para todo $x \in \mathcal{X}_j$ e $y \in \mathcal{Y}_k$ implican que $F(\boldsymbol{a}_1, X_r) = F(\boldsymbol{a}_2, X_r) = 0$. En este caso tenemos que $|\mathcal{S}_{\underline{\boldsymbol{a}}}(\mathcal{X}_j, \mathcal{Y}_k)| = q^{\dim \mathbb{F}_q[\boldsymbol{X}]_{\leq d} - 2d - 1}$, por lo que concluimos que

$$\mathcal{N}_{j,k} = \binom{q}{j}\binom{q}{k} q^{\dim \mathbb{F}_q[\boldsymbol{X}]_{\leq d} - 2d - 1}.$$

De la expresión para $\mathcal{N}_{j,k}$ en los cuatro casos considerados, concluimos que

$$\begin{aligned}
\frac{N_{\underline{\boldsymbol{a}}}}{|\mathbb{F}_q[\boldsymbol{X}]_{\leq d}|} &= \frac{1}{q^{\dim \mathbb{F}_q[\boldsymbol{X}]_{\leq d}}} \sum_{j=1}^{q}\sum_{k=1}^{q} (-1)^{j+k} \mathcal{N}_{j,k} \\
&= \sum_{j=1}^{d}\sum_{k=1}^{d} (-1)^{j+k} \binom{q}{j}\binom{q}{k} q^{-j-k} + 2\sum_{j=1}^{d}\sum_{k=d+1}^{q} (-1)^{j+k} \binom{q}{j}\binom{q}{k} q^{-j-(d+1)} \\
&\quad + \sum_{j=d+1}^{q}\sum_{k=d+1}^{q} (-1)^{j+k} \binom{q}{j}\binom{q}{k} q^{-2d-1}.
\end{aligned}$$

Por (8.6), el Teorema 8.2.1 y cálculos elementales tenemos que

$$\begin{aligned}
\frac{N_{\underline{\boldsymbol{a}}}}{|\mathbb{F}_q[\boldsymbol{X}]_{\leq d}|} &= \left(\sum_{j=1}^{d} (-1)^j \binom{q}{j} q^{-j}\right)^2 - 2\left(\sum_{j=1}^{d} (-1)^j \binom{q}{j} q^{-j}\right)(-1)^d \binom{q-1}{d} q^{-d-1} \\
&\quad + \binom{q-1}{d}^2 q^{-2d-1} = \big(P_1[C_1 = 1]\big)^2 + \frac{q-1}{q^{2d+2}} \binom{q-1}{d}^2.
\end{aligned}$$

□

Combinando esta afirmación con (8.9), deducimos que

$$\begin{aligned} p_{r,d}[C_{\underline{\boldsymbol{a}}} = 2] = \frac{[C_{\underline{\boldsymbol{a}}} = 2]}{|\mathbb{F}_q[\boldsymbol{X}]_{\leq d}|} &= \frac{[C_1 = 1]}{|\mathbb{F}_q[\boldsymbol{X}]_{\leq d}|} - \frac{N_{\underline{\boldsymbol{a}}}}{|\mathbb{F}_q[\boldsymbol{X}]_{\leq d}|} \\ &= P_1[C_1 = 1] - \big(P_1[C_1 = 1]\big)^2 - \frac{q-1}{q^{2d+2}}\binom{q-1}{d}^2 \\ &= \big(1 - P_1[C_1 = 1]\big)P_1[C_1 = 1] - \frac{q-1}{q^{2d+2}}\binom{q-1}{d}^2. \end{aligned}$$

Sea $f : \mathbb{R} \to \mathbb{R}$ la función definida por $f(x) := (1-x)x$. El Teorema del Valor Medio muestra que existe $\xi \in (0,1)$ tal que

$$\big(1 - P_1[C_1 = 1]\big)P_1[C_1 = 1] - (1-\mu_d)\mu_d = f'(\xi)\,\big(P_1[C_1 = 1] - \mu_d\big).$$

Como $-1 \leq f'(x) \leq 1$ para todo $x \in [0,1]$, deducimos que $|f'(\xi)| \leq 1$. Por lo tanto, del Corolario 8.2.2 se sigue que

$$\big|(1 - P_1[C_1 = 1])P_1[C_1 = 1] - (1-\mu_d)\mu_d\big| \leq \big|P_1[C_1 = 1] - \mu_d\big| \leq \frac{2}{q}.$$

Por otro lado, es fácil deducir que $\frac{q-1}{q^{2d+2}}\binom{q-1}{d}^2 \leq \frac{1}{q}$. Esto implica la afirmación de la proposición. □

La Proposición 8.2.5 es el paso fundamental para analizar el comportamiento de la probabilidad $P_2[C_2 = 2]$. El siguiente resultado proporciona una estimación de dicha probabilidad.

Corolario 8.2.6. *Para $q > d$,*

$$|P_2[C_2 = 2] - (1-\mu_d)\mu_d| \leq \frac{3}{q}.$$

Demostración. Por el Lema 8.2.4 y la Proposición 8.2.5 obtenemos que

$$|P_2[C_2 = 2] - (1-\mu_d)\mu_d| \leq \frac{1}{N_2}\sum_{\underline{\boldsymbol{a}} \in \mathsf{F}_2} |p_{r,d}[C_{\underline{\boldsymbol{a}}} = 2] - (1-\mu_d)\mu_d| \leq \frac{3}{q},$$

lo que demuestra el corolario. □

El análisis del Algoritmo BBV desde un punto de vista probabilístico prueba que la probabilidad de encontrar un cero $\mathbb{F}_q$–racional de F en a lo sumo 2 bandas verticales es del orden de $(2-\mu_d)\mu_d \approx 0{,}8646\ldots$. Esto mejora los resultados de [Mat10], donde se describe una versión del Algoritmo BBV para polinomios $F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$ absolutamente irreducibles y se demuestra que con al menos d elecciones aleatorias es posible encontrar un cero $\mathbb{F}_q$–racional de F con probabilidad al menos $1/2$. Más adelante vamos a obtener una cota superior de la complejidad en promedio del Algoritmo BBV.

Finalizamos esta sección con un comentario sobre los espacios muestrales que hemos considerado hasta el momento. Para el análisis de la probabilidad de que el algoritmo termine en la primera banda vertical consideramos el espacio muestral $\mathsf{F}_1 := \mathbb{F}_q^{r-1}$ y la variable aleatoria $C_1 : \mathsf{F}_1 \times \mathbb{F}_q[\boldsymbol{X}]_{\leq d} \to \{1, \infty\}$. En cambio, para el análisis de la probabilidad de que el algoritmo necesite dos bandas verticales el espacio muestral es $\mathsf{F}_2 \times \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$ y la variable aleatoria $C_2 : \mathsf{F}_2 \times \mathbb{F}_q[\boldsymbol{X}]_{\leq d} \to \{1, 2, \infty\}$. Para vincular ambos análisis, en el Lema 8.4.1 probaremos que

$$P_2[C_2 = 1] = P_1[C_1 = 1].$$

Esta igualdad muestra la "consistencia" de los espacios de probabilidad del Teorema 8.2.1 y del Corolario 8.2.6.

8.3. Probabilidad de éxito en más bandas verticales

El paso más importante para el análisis probabilístico del Algoritmo BBV es determinar la probabilidad de que realicen búsquedas en s bandas verticales, para una determinada elección de s bandas verticales distintas dos a dos. Los casos $s = 1$ y $s = 2$ fueron discutidos en la sección anterior. En esta sección analizamos el caso general.

Fijamos $3 \leq s \leq \min\{\binom{d+r-1}{r-1}, q^{r-1}\}$ y $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_s \in \mathbb{F}_q^{r-1}$. Supongamos que $\boldsymbol{a}_i \neq \boldsymbol{a}_j$ para $i \neq j$ y denotemos por $\underline{\boldsymbol{a}} := (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_s)$. En esta sección estudiamos la probabilidad de que el algoritmo realice s búsquedas hasta que encuentra una banda vertical con un cero $\mathbb{F}_q$–racional del polinomio de partida, suponiendo que $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_s$ son las elecciones que consideramos para las primeras s bandas verticales.

Para este propósito, consideramos la probabilidad uniforme $p_{r,d}$ sobre el conjunto $\mathbb{F}_q[\boldsymbol{X}]_{\leq d}$ y la variable aleatoria $C_{\underline{\boldsymbol{a}}} := C_{\underline{\boldsymbol{a}},r,d} : \mathbb{F}_q[\boldsymbol{X}]_{\leq d} \to \{1, 2, \ldots, s, \infty\}$ que cuenta el número de búsquedas que el algoritmo realiza sobre las bandas verticales determinadas por $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_s$, donde $C_{\underline{\boldsymbol{a}}}(F) := \infty$ significa que $F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$ no tiene ceros $\mathbb{F}_q$–racionales sobre esas s bandas verticales.

Comenzamos con el siguiente resultado elemental.

Lema 8.3.1. *Sean $\mathbb{V}$ y $\mathbb{W}$ dos espacios $\mathbb{F}_q$–vectoriales de dimensión finita y sea $\Phi : \mathbb{V} \to \mathbb{W}$ cualquier función $\mathbb{F}_q$–lineal. Consideramos las probabilidades uniformes $P_{\mathbb{V}}$ y $P_{\mathbb{W}}$ sobre $\mathbb{V}$ y $\mathbb{W}$ respectivamente. Entonces, para cualquier conjunto $A \subset \mathbb{W}$ tenemos que*

$$P_{\mathbb{V}}(\Phi^{-1}(A)) = \frac{|A \cap \mathrm{Im}(\Phi)|}{|\mathrm{Im}(\Phi)|} = \frac{P_{\mathbb{W}}(A \cap \mathrm{Im}(\Phi))}{P_{\mathbb{W}}(\mathrm{Im}(\Phi))} =: P_{\mathrm{Im}\Phi}(A).$$

Demostración. Tenemos que

$$\frac{1}{|\mathbb{V}|}|\Phi^{-1}(A)| = \frac{1}{|\mathbb{V}|}\sum_{\boldsymbol{w} \in A} |\Phi^{-1}(\boldsymbol{w})| = \frac{1}{|\mathbb{V}|}|\mathrm{Ker}(\Phi)|\,|A \cap \mathrm{Im}(\Phi)|.$$

Por el Teorema de la dimensión y la igualdad $|\mathbb{S}| = q^{\dim \mathbb{S}}$, resulta que

$$\frac{1}{|\mathbb{V}|}|\Phi^{-1}(A)| = \frac{|A \cap \mathrm{Im}(\Phi)|}{|\mathrm{Im}(\Phi)|} = \frac{P_{\mathbb{W}}(A \cap \mathrm{Im}(\Phi))}{P_{\mathbb{W}}(\mathrm{Im}(\Phi))}.$$

Esto finaliza la demostración del lema. □

Por simplicidad, vamos a reemplazar la variable X_r por una nueva indeterminada T, y vamos a encontrar una "buena" descripción de la imagen de la proyección lineal que describe los polinomios resultantes $F(\boldsymbol{a}_i, T)$ de las primeras s búsquedas. Consideremos la función $\mathbb{F}_q$–lineal $\Phi := \Phi_{\underline{\boldsymbol{a}}} : \mathbb{F}_q[\boldsymbol{X}]_{\leq d} \to \mathbb{F}_q[T]^s_{\leq d}$ definida como

$$\Phi(F) := \big(F(\boldsymbol{a}_1, T), \ldots, F(\boldsymbol{a}_s, T)\big). \tag{8.11}$$

Como $\mathrm{Im}(\Phi)$ es un espacio $\mathbb{F}_q$–vectorial, aplicando el Lema 8.3.1 obtenemos que

$$p_{r,d}[C_{\underline{\boldsymbol{a}}} = s] = \frac{\big|(\{N = 0\}^{s-1} \times \{N > 0\}) \cap \mathrm{Im}(\Phi)\big|}{|\mathrm{Im}(\Phi)|}, \tag{8.12}$$

donde $N := N_{1,d}$ denota la variable aleatoria que cuenta el número de ceros en $\mathbb{F}_q$ de los elementos de $\mathbb{F}_q[T]_{\leq d}$. Así, a fin de estimar la probabilidad $p_{r,d}[C_{\underline{\boldsymbol{a}}} = s]$ vamos a estimar la siguiente cantidad:

$$R_s := \big|\big(\{N = 0\}^{s-1} \times \{N > 0\}\big) \cap \mathrm{Im}(\Phi)\big|.$$

En la siguiente sección obtenemos una caracterización de la imagen de Φ para una elección general de $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_s$. Esta caracterización nos permitirá expresar R_s en términos del promedio del cardinal del conjunto de valores de ciertas familias lineales de polinomios univariados con coeficientes prescriptos, y utilizar así los resultados de los Capítulos 5 y 6.

Como explicamos más adelante, existe un único entero positivo $\kappa_s \leq d$ tal que

$$\binom{\kappa_s + r - 2}{r-1} < s \leq \binom{\kappa_s + r - 1}{r-1}.$$

Vamos a pedir que los puntos $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_s$ que consideramos satisfagan la condición que enunciamos a continuación. Para $1 \leq j \leq \kappa_s$, sea $D_j := \binom{j+r-1}{r-1}$ y sea $\Omega_j := \{\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_{D_j}\} \subset (\mathbb{Z}_{\geq 0})^{r-1}$ el conjunto de las $(r-1)$–uplas $\boldsymbol{\omega}_k := (\omega_{k,1}, \ldots, \omega_{k,r-1})$ con $|\boldsymbol{\omega}_k| := \omega_{k,1} + \cdots + \omega_{k,r-1} \leq j$. Sea $\boldsymbol{a}_i^{\boldsymbol{\omega}_k} := a_{i,1}^{\omega_{k,1}} \cdots a_{i,r-1}^{\omega_{k,r-1}}$ para $1 \leq i \leq s$ y sea $1 \leq k \leq D_j$. Entonces necesitamos que la matriz de Vandermonde multivariada

$$\mathcal{M}_j := \begin{pmatrix} \boldsymbol{a}_1^{\boldsymbol{\omega}_1} & \cdots & \boldsymbol{a}_1^{\boldsymbol{\omega}_{D_j}} \\ \vdots & & \vdots \\ \boldsymbol{a}_s^{\boldsymbol{\omega}_1} & \cdots & \boldsymbol{a}_s^{\boldsymbol{\omega}_{D_j}} \end{pmatrix} \in \mathbb{F}_q^{s \times D_j} \tag{8.13}$$

tenga rango máximo $\min\{D_j, s\}$ para $1 \leq j \leq \kappa_s$.

Esta condición es un requisito débil, que probablemente se cumpla para cualquier elección "razonable" de los elementos $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_s \in \mathbb{F}_q^{r-1}$. Sean $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_s$

$(r-1)$–uplas de indeterminadas sobre $\overline{\mathbb{F}}_q$, esto es, $\boldsymbol{A}_i := (A_{i,1}, \ldots, A_{i,r-1})$ para $1 \leq i \leq s$, y denotemos por $\mathcal{V}_j$ la siguiente matriz de Vandermonde de tamaño $\text{mín}\{D_j, s\} \times \text{mín}\{D_j, s\}$ con entradas en $\mathbb{F}_q[\boldsymbol{A}_1, \ldots, \boldsymbol{A}_s]$:

$$\mathcal{V}_j := \begin{pmatrix} \boldsymbol{A}_1^{\boldsymbol{\omega}_1} & \cdots & \boldsymbol{A}_1^{\boldsymbol{\omega}_{\text{mín}\{D_j,s\}}} \\ \vdots & & \vdots \\ \boldsymbol{A}_{\text{mín}\{D_j,s\}}^{\boldsymbol{\omega}_1} & \cdots & \boldsymbol{A}_{\text{mín}\{D_j,s\}}^{\boldsymbol{\omega}_{\text{mín}\{D_j,s\}}} \end{pmatrix}.$$

Supongamos que enumeramos los elementos de $\Omega_j := \{\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_{D_j}\} \subset (\mathbb{Z}_{\geq 0})^{r-1}$ de manera graduada, es decir, $|\boldsymbol{\omega}_k| \leq |\boldsymbol{\omega}_l|$ siempre que $k \leq l$. En particular, $\boldsymbol{\omega}_1 = (0, \ldots, 0)$. Por [DT09, Teorema 1.5] se sigue que $\det \mathcal{V}_j$ es absolutamente irreducible para $1 \leq j \leq \kappa_s$. Sea δ_j el grado del polinomio $\det \mathcal{V}_j$. Tenemos que $\delta_j \leq jD_j$. Por el Teorema 2.2.5 obtenemos que el número $\mathcal{N}_j$ de $(r-1)$–uplas $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_s \in \mathbb{F}_q^{r-1}$ que anulan al polinomio $\det \mathcal{V}_j$ satisface la siguiente estimación:

$$|\mathcal{N}_j - q^{s(r-1)-1}| \leq (\delta_j - 1)(\delta_j - 2) q^{s(r-1)-\frac{3}{2}} + 5\delta_j^{\frac{13}{3}} q^{s(r-1)-2}. \tag{8.14}$$

Así, si evitamos cualquier elección de $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_s$ en esas $\mathcal{N}_j = \mathcal{O}(q^{s(r-1)-1})$ uplas para $1 \leq j \leq \kappa_s$, obtenemos puntos $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_s$ que cumplirán nuestros requerimientos. Más aún, muchas "malas" elecciones $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_s$ que anulan al polinomio $\det \mathcal{V}_j$ para un j dado también funcionarán, dado que otros menores de la matriz de Vandermonde $\mathcal{M}_j$ definida en (8.13) pueden ser no singulares. En particular, esta condición se cumplirá, en el caso en que $s \leq r$, si $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_s$ son afínmente independientes en $\mathbb{F}_q^{r-1}$.

Resumiendo, denotemos $\mathcal{V}^s := \prod_{j=1}^{\kappa_s} \det \mathcal{V}_j \in \mathbb{F}_q[\boldsymbol{A}_1, \ldots, \boldsymbol{A}_s]$ y sea

$$\mathtt{B}_s := \{\underline{\boldsymbol{a}} := (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_s) \in \mathbb{F}_q^{s(r-1)} : \mathcal{V}^s(\underline{\boldsymbol{a}}) = 0\}. \tag{8.15}$$

Entonces $|\mathtt{B}_s| = \mathcal{O}(q^{s(r-1)-1})$. Todos los resultados de esta sección son válidos siempre que $\underline{\boldsymbol{a}} \in \mathbb{F}_q^{s(r-1)} \setminus \mathtt{B}_s$.

8.3.1. Imagen de la proyección que definen *s* bandas verticales

En esta sección caracterizamos la imagen $\text{Im}(\Phi)$ de Φ. Para este propósito, expresamos cada elemento del espacio $\mathbb{F}_q$–lineal $\mathbb{F}_q[\boldsymbol{X}]_{\leq d}$ por medio de sus coordenadas en la base monomial usual $\mathcal{B}$ de $\mathbb{F}_q[\boldsymbol{X}]_{\leq d}$, considerando el orden monomial que definimos ahora. Denotamos por $\mathcal{B}_i$ el conjunto de los monomios de $\mathbb{F}_q[X_1, \ldots, X_{r-1}]$ de grado a lo sumo i para $0 \leq i \leq d$, con el orden lexicográfico definido por $X_1 < X_2 < \cdots < X_{r-1}$. Entonces la base $\mathcal{B}$ se considera con el orden $\mathcal{B} = \{X_r^d, X_r^{d-1}\mathcal{B}_1, \ldots, X_r\mathcal{B}_{d-1}, \mathcal{B}_d\}$, donde cada conjunto $X_r^{d-i}\mathcal{B}_i$ se ordena siguiendo el orden inducido por el de $\mathcal{B}_i$. En otras palabras, si expresamos cada $F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$ de manera única de la forma

$$F = \sum_{i=0}^{d} F_i(X_1, \ldots, X_{r-1}) X_r^i,$$

donde cada F_i tiene grado a lo sumo $d-i$ para $0 \leq i \leq d$, entonces el vector de coeficientes $(F)_{\mathcal{B}}$ de F en la base $\mathcal{B}$ es $(F)_{\mathcal{B}} = \big((F_d)_{\mathcal{B}_0}, \ldots, (F_0)_{\mathcal{B}_d}\big)$. Por otro lado, expresamos los elementos de $\mathbb{F}_q[T]^s_{\leq d}$ en la base $\mathcal{B}' := \{T^d, \ldots, T, 1\}^s$.

Sean

$$D_i := \binom{i+r-1}{r-1} = |\mathcal{B}_i| \quad (0 \leq i \leq d), \quad D := \binom{d+r}{r} = |\mathcal{B}| = \sum_{i=0}^{d} |\mathcal{B}_i|.$$

Definimos $D_{-1} := 0$. Observemos que la sucesión $(D_i)_{i \geq -1}$ es estrictamente creciente, y por lo tanto, para cada i con $1 \leq i \leq s$, existe un único $\kappa_i \in \mathbb{N}$ tal que

$$D_{\kappa_i - 1} < s \leq D_{\kappa_i}.$$

Observemos que por definición se tiene fácilmente la siguiente observación.

Observación 8.3.2.

- $\kappa_i \leq j$ *si y solo si* $i \leq D_j$.
- $\kappa_1 = 0$, $\kappa_s \leq d$.

La matriz $\mathsf{M}_\Phi \in \mathbb{F}_q^{s(d+1)\times D}$ de la función $\mathbb{F}_q$– lineal Φ con respecto a las bases $\mathcal{B}$ y $\mathcal{B}'$ puede escribirse como una matriz de bloques, de la siguiente manera:

$$\mathsf{M}_\Phi = \begin{pmatrix} \mathsf{M}_1 \\ \vdots \\ \mathsf{M}_s \end{pmatrix},$$

donde $\mathsf{M}_i \in \mathbb{F}_q^{(d+1)\times D}$ es la matriz diagonal por bloques

$$\mathsf{M}_i := \begin{pmatrix} \mathsf{M}_{i,0} & & & \\ & \mathsf{M}_{i,1} & & \\ & & \ddots & \\ & & & \mathsf{M}_{i,d} \end{pmatrix}, \quad \mathsf{M}_{i,j} := \big(\boldsymbol{a}_i^{\boldsymbol{\alpha}} : |\boldsymbol{\alpha}| \leq j\big) \in \mathbb{F}_q^{1\times D_j}.$$

En el Lema 8.3.4 determinamos la dimensión de $\mathrm{Im}(\Phi)$. Para ello, utilizamos la siguiente identidad combinatoria.

Observación 8.3.3. *Dados enteros positivos R, K, tenemos que*

$$\sum_{j=0}^{K} j\binom{j+R}{R} = (R+1)\binom{R+1+K}{R+2}. \tag{8.16}$$

Demostración. A partir de cálculos de combinatoria elemental, tenemos que

$$\sum_{j=0}^{K} j\binom{j+R}{R} = \sum_{j=1}^{K} \frac{(j+R)!}{R!(j-1)!} = (R+1)\sum_{j=0}^{K-1}\binom{j+R+1}{R+1} = (R+1)\binom{R+1+K}{R+2}.$$

Esto muestra (8.16). □

Lema 8.3.4. *Para $s \leq \min\{D_d, q^{r-1}\}$, tenemos que*

$$\dim \mathrm{Im}(\Phi) = \binom{\kappa_s - 1 + r}{r} + s(d - \kappa_s + 1) = \sum_{i=1}^{s}(d + 1 - \kappa_i).$$

Demostración. Sea $\boldsymbol{h} := (h_1, \ldots, h_s)$ un elemento de $\mathrm{Im}(\Phi)$. Entonces existe un polinomio $F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$ tal que $\boldsymbol{h} = \Phi(F)$. Si denotamos por $(F)_{\mathcal{B}} = \big((F_d)_{\mathcal{B}_0}, \ldots, (F_0)_{\mathcal{B}_d}\big)$ las coordenadas de F en la base monomial $\mathcal{B}$, por la estructura en bloques de la matriz M_Φ tenemos que

$$\Phi(F) = \sum_{j=0}^{d} \begin{pmatrix} \mathsf{M}_{1,j} \\ \vdots \\ \mathsf{M}_{s,j} \end{pmatrix} (F_{d-j})_{\mathcal{B}_j} T^{d-j}. \tag{8.17}$$

Como $\underline{\boldsymbol{a}} \in \mathbb{F}_q^{s(r-1)} \setminus \mathtt{B}_s$, tenemos que

$$\mathrm{rg} \begin{pmatrix} \mathsf{M}_{1,j} \\ \vdots \\ \mathsf{M}_{s,j} \end{pmatrix} = \min\{D_j, s\} = \begin{cases} D_j & \text{for } 0 \leq j \leq \kappa_{s-1}, \\ s & \text{for } \kappa_s \leq j \leq d. \end{cases}$$

Por lo tanto,

$$\dim \mathrm{Im}(\Phi) = \sum_{j=0}^{\kappa_{s-1}} D_j + s(d - \kappa_s + 1) = \binom{\kappa_s - 1 + r}{r} + s(d - \kappa_s + 1).$$

Esto prueba la primera afirmación del lema. Para demostrar la segunda afirmación, vemos que

$$\begin{aligned} \sum_{i=1}^{s}(d + 1 - \kappa_i) &= \sum_{j=0}^{\kappa_s} \sum_{i=D_{j-1}+1}^{\min\{D_j, s\}} (d + 1 - j) \\ &= \sum_{j=0}^{\kappa_s - 1} (d + 1 - j)(D_j - D_{j-1}) + (d + 1 - \kappa_s)(s - D_{\kappa_s - 1}). \end{aligned}$$

Observemos que

$$\sum_{j=0}^{\kappa_{s-1}}(D_j - D_{j-1}) = \sum_{j=0}^{\kappa_{s-1}} \binom{j + r - 2}{r - 2} = D_{\kappa_{s-1}}.$$

Así, concluimos que

$$\sum_{i=1}^{s}(d + 1 - \kappa_i) = -\sum_{j=0}^{\kappa_s - 1} j(D_j - D_{j-1}) + (d + 1 - \kappa_s)s + \kappa_s D_{\kappa_s - 1}.$$

Teniendo en cuenta (8.16), obtenemos que

$$\sum_{j=0}^{\kappa_s-1} j(D_j - D_{j-1}) = \sum_{j=0}^{\kappa_s-1} j\binom{j+r-2}{r-2} = (r-1)\binom{\kappa_s+r-2}{r}.$$

Por lo tanto,

$$\sum_{i=1}^{s}(d+1-\kappa_i) = -(r-1)\binom{\kappa_s+r-2}{r} + (d+1-\kappa_s)s + \kappa_s D_{\kappa_s-1}.$$

Por un cálculo elemental tenemos que

$$-(r-1)\binom{\kappa_s+r-2}{r} + \kappa_s D_{\kappa_s-1} = \binom{\kappa_s+r-1}{r}.$$

Así, concluimos que

$$\sum_{i=1}^{s}(d+1-\kappa_i) = \binom{\kappa_s+r-1}{r} + (d+1-\kappa_s)s.$$

Esto finaliza la demostración del lema. □

De los argumentos del Lema 8.3.4 obtenemos la siguiente observación sobre los rangos de las matrices M_i para cada $1 \leq i \leq s$.

Observación 8.3.5.

$$\operatorname{rg}(\mathsf{M}_i) = d+1-\kappa_i \quad (1 \leq i \leq s). \tag{8.18}$$

Sea $\boldsymbol{h} := (h_1, \ldots, h_s)$ un elemento arbitrario de $\operatorname{Im}(\Phi)$. Por (8.18) tenemos que la i–ésima coordenada h_i es un polinomio de grado d que tiene los últimos $d+1-k_i$ coeficientes consecutivos libres para $1 \leq i \leq s$. A continuación presentamos una parametrización de $\operatorname{Im}(\Phi)$ que nos servirá para estimar la probabilidad deseada. Sea $\Phi^* : \operatorname{Im}(\Phi) \to \mathbb{F}_q^{\dim \operatorname{Im}(\Phi)}$ la función $\mathbb{F}_q$–lineal definida por

$$\Phi^*(\boldsymbol{h}) := \boldsymbol{h}^*,$$

donde $\boldsymbol{h} := (h_1, \ldots, h_s)$, $h_i := (h_{d,i}, \ldots, h_{0,i}) \in \mathbb{F}_q^{d+1}$ para $1 \leq i \leq s$ y

$$\boldsymbol{h}^* := (h_1^*, \ldots, h_s^*), \quad h_i^* := (h_{d-\kappa_i,i}, \ldots, h_{0,i}) \ \ (1 \leq i \leq s). \tag{8.19}$$

El Lema 8.3.4 muestra que Φ^* está bien definida. Más aún, tenemos el siguiente resultado.

Lema 8.3.6. Φ^* *es un isomorfismo.*

Demostración. Como Φ^* es una transformación lineal entre espacios $\mathbb{F}_q$–vectoriales de la misma dimensión, es suficiente mostrar que Φ^* monomorfismo. Fijemos $\boldsymbol{h} := \Phi(F) \in \mathrm{Im}(\Phi)$ con $\boldsymbol{h}^* = \mathbf{0}$. De (8.17) deducimos que

$$\begin{pmatrix} \mathsf{M}_{1,j} \\ \vdots \\ \mathsf{M}_{s,j} \end{pmatrix} (F_{d-j})_{\mathcal{B}_j} = \begin{pmatrix} h_{d-j,1} \\ \vdots \\ h_{d-j,s} \end{pmatrix}. \tag{8.20}$$

Fijemos j con $0 \leq j \leq \kappa_s - 1$. Entonces los elementos $h_{d-j,i}$ se encuentran incluidos en la definición de h_i^* si y solo si $i \leq D_j$ (ver Observación 8.3.2). Por hipótesis tenemos que $\boldsymbol{h}^* = \mathbf{0}$; se sigue que $h_{d-j,i} = 0$ para $0 \leq i \leq D_j$. Así, tenemos la siguiente identidad:

$$\begin{pmatrix} \mathsf{M}_{1,j} \\ \vdots \\ \mathsf{M}_{D_j,j} \\ \mathsf{M}_{D_j+1,j} \\ \vdots \\ \mathsf{M}_{s,j} \end{pmatrix} (F_{d-j})_{\mathcal{B}_j} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ h_{d-j,D_j+1} \\ \vdots \\ h_{d-j,s} \end{pmatrix}.$$

Como por hipótesis $\underline{\boldsymbol{a}} \in \mathbb{F}_q^{s(r-1)} \setminus \mathtt{B}_s$, la submatriz superior de tamaño $D_j \times D_j$ de la matriz del lado izquierdo de la igualdad anterior es inversible. Concluimos que $(F_{d-j})_{\mathcal{B}_j} = \mathbf{0}$. Esto implica que $h_{d-j,D_j+1} = \cdots = h_{d-j,s} = 0$. Por otro lado, para $j \geq \kappa_s$ los elementos $h_{d-j,i}$ están incluidos en la definición de h_i^* para $1 \leq i \leq s$ y por lo tanto $h_{d-j,i} = 0$ para $1 \leq i \leq s$. Esto muestra que $\boldsymbol{h} = \mathbf{0}$, y demuestra el lema. □

Denotamos por $\Psi := (\psi_1, \ldots, \psi_s) : \mathbb{F}_q^{\dim \mathrm{Im}(\Phi)} \to \mathrm{Im}(\Phi)$ la función inversa de Φ^*. En el siguiente lema damos información sobre las funciones coordenadas ψ_i de Ψ. Este resultado muestra que, para $1 \leq i \leq s$, los coeficientes $h_{d,i}, \ldots, h_{d-\kappa_i+1,i}$ del polinomio h_i quedan unívocamente determinados por los coeficientes $h_{d-\kappa_j,j}, \ldots, h_{0,j}$ con $1 \leq j \leq i-1$ de los polinomios $h_1, \ldots, h_{i-1}$.

Lema 8.3.7. *Sea* $h_i^* := (h_{d-\kappa_i,i}, \ldots, h_{0,i}) \in \mathbb{F}_q^{d+1-\kappa_i}$ *para* $1 \leq i \leq s$. *Sea* $\boldsymbol{h}^* := (h_1^*, \ldots, h_s^*) \in \mathbb{F}_q^{\dim \mathrm{Im}(\Phi)}$ *y* $\boldsymbol{h} := \Psi(\boldsymbol{h}^*)$. *Denotamos por*

$$h_i := \psi_i(\boldsymbol{h}^*) := h_{d,i}\, T^d + \cdots + h_{d+1-\kappa_i,i}\, T^{d+1-\kappa_i} + h_{d-\kappa_i,i}\, T^{d-\kappa_i} + \cdots + h_{0,i}.$$

Entonces $h_{d,i}, \ldots, h_{d+1-\kappa_i,i}$ *están unívocamente determinados por* $h_1^*, \ldots, h_{i-1}^*$.

Demostración. Fijemos k con $0 \leq k \leq \kappa_i - 1$. Como $\boldsymbol{h} \in \mathrm{Im}(\Phi)$, tenemos que $\boldsymbol{h} := \Phi(F)$ para algún $F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$. De (8.17) deducimos que

$$\begin{pmatrix} \mathsf{M}_{1,k} \\ \vdots \\ \mathsf{M}_{D_k,k} \end{pmatrix} (F_{d-k})_{\mathcal{B}_k} = \begin{pmatrix} h_{d-k,1} \\ \vdots \\ h_{d-k,D_k} \end{pmatrix},$$

donde la matriz de tamaño $D_k \times D_k$ del lado izquierdo es inversible. Los elementos $h_{d-k,l}$ están incluidos en la definición de h_l^* si y solo si $l \leq D_k$. Además, tenemos que $k \leq \kappa_i - 1 \leq \kappa_{i-1}$. Concluimos que el vector columna de la igualdad anterior está unívocamente determinado por $h_1^*, \ldots, h_{i-1}^*$, y $(F_{d-k})_{\mathcal{B}_k}$ está también unívocamente determinado por dichos elementos. Por lo tanto, la identidad

$$\begin{pmatrix} \mathsf{M}_{1,k} \\ \vdots \\ \mathsf{M}_{i,k} \end{pmatrix} (F_{d-k})_{\mathcal{B}_k} = \begin{pmatrix} h_{d-k,1} \\ \vdots \\ h_{d-k,i} \end{pmatrix},$$

muestra que los elementos $h_{d-k,i}$ para $1 \leq i \leq s$ están unívocamente determinados por $h_1^*, \ldots, h_{i-1}^*$. □

Terminamos esta sección con la siguiente observación referida a los coeficientes principales de cada coordenada h_i de cualquier elemento $\boldsymbol{h} := (h_1, \ldots, h_s) \in \mathrm{Im}(\Phi)$.

Observación 8.3.8. *Para cada $\boldsymbol{h} := (h_1, \ldots, h_s) \in \mathrm{Im}(\Phi)$, tenemos que $h_{d,1} = \ldots = h_{d,s}$. En efecto, de (8.17) deducimos que*

$$\begin{pmatrix} \mathsf{M}_{1,0} \\ \vdots \\ \mathsf{M}_{s,0} \end{pmatrix} (F_d)_{\mathcal{B}_0} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} (F_d)_{\mathcal{B}_0} = \begin{pmatrix} h_{d,1} \\ \vdots \\ h_{d,s} \end{pmatrix}.$$

Esto implica que $h_{d,1} = \ldots = h_{d,s} = (F_d)_{\mathcal{B}_0}$. En particular, el coeficiente $h_{d,1}$ del monomio T^d en el polinomio h_1 determina unívocamente los coeficientes $h_{d,j}$ del monomio T^d en h_j para $2 \leq j \leq s$. □

8.3.2. La probabilidad de s búsquedas en términos de cardinales de conjuntos de valores

Sea $\underline{\boldsymbol{a}} := (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_s) \in \mathbb{F}_q^{s(r-1)} \setminus \mathtt{B}_s$, donde $\mathtt{B}_s$ está definido en (8.15). El objetivo de esta sección es estimar la probabilidad $p_{r,d}[C_{\underline{\boldsymbol{a}}} = s]$ del conjunto de polinomios F de $\mathbb{F}_q[\boldsymbol{X}]_{\leq d}$ para los cuales el Algoritmo BBV realiza s búsquedas sobre las bandas verticales determinadas por $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_s$ hasta encontrar una raíz en $\mathbb{F}_q$ del polinomio F de entrada. De acuerdo con (8.12), vamos a estimar la cantidad

$$R_s := \big|\big(\{N = 0\}^{s-1} \times \{N > 0\}\big) \cap \mathrm{Im}(\Phi)\big|.$$

A fin de obtener estimaciones explícitas de R_s, vamos a relacionar dicha cantidad con el promedio del cardinal del conjunto de valores de ciertas familias de polinomios mónicos univariados de grado a lo sumo d con coeficientes prescriptos.

Por el Lema 8.3.6 tenemos que cada elemento $\boldsymbol{h} \in \mathrm{Im}(\Phi)$ puede expresarse de manera única como $\boldsymbol{h} = \Psi(\boldsymbol{h}^*)$, donde $\boldsymbol{h}^*$ está definida en (8.19). Se sigue que

$$R_s = \sum_{\boldsymbol{h}^* \in \mathbb{F}_q^{\dim \mathrm{Im}(\Phi)}} \mathbf{1}_{\{N=0\}^{s-1} \times \{N>0\}}\big(\Psi(\boldsymbol{h}^*)\big), \tag{8.21}$$

donde $\mathbf{1}_{\{N=0\}^{s-1}\times\{N>0\}} : \mathbb{F}_q[T]^s_{\leq d} \to \{0,1\}$ denota la función característica del conjunto $\{N=0\}^{s-1}\times\{N>0\}$. Observemos que, por el Lema 8.3.7, la coordenada $\psi_i(\boldsymbol{h}^*)$ depende solo de $\boldsymbol{h}^*_i := (h^*_1,\dots,h^*_i)$ para $1\leq i\leq s$, por lo que, con un leve abuso de notaciones, escribiremos $\psi_i(\boldsymbol{h}^*)$ como $\psi_i(\boldsymbol{h}^*_i)$ para $1\leq i\leq s$.

A continuación reescribimos la expresión (8.21) para R_s de una manera adecuada para nuestros propósitos.

Lema 8.3.9. *Sea* $\boldsymbol{h} := (\sum_{j=0}^d h_{j,1}T^j,\dots,\sum_{j=0}^d h_{j,s}T^j)$ *un elemento arbitrario de* $\mathrm{Im}(\Phi)$ *y sea* $\boldsymbol{h}^* := \Phi^*(\boldsymbol{h}) := (h^*_1,\dots,h^*_s) \in \mathbb{F}_q^{\dim \mathrm{Im}(\Phi)}$ *definido como en* (8.19). *Para* $s\leq \min\{D_d, q^{r-1}\}$, *vale la siguiente igualdad:*

$$R_s = \sum_{\substack{h^*_1\in\mathbb{F}_q^{d+1}\\ N(\psi_1(\boldsymbol{h}^*_1))=0}} \cdots \sum_{\substack{h^*_{s-1}\in\mathbb{F}_q^{d-\kappa_{s-1}+1}\\ N(\psi_{s-1}(\boldsymbol{h}^*_{s-1}))=0}} \sum_{h^*_s\in\mathbb{F}_q^{d-\kappa_s+1}} \mathbf{1}_{\{N>0\}}\big(\psi_s(\boldsymbol{h}^*_s)\big).$$

Demostración. Podemos reescribir (8.21) de la siguiente manera:

$$R_s = \sum_{h^*_1\in\mathbb{F}_q^{d+1}} \cdots \sum_{h^*_s\in\mathbb{F}_q^{d-\kappa_s+1}} \mathbf{1}_{\{N=0\}^{s-1}\times\{N>0\}}\big(\Psi(\boldsymbol{h}^*)\big).$$

Observamos además que, como consecuencia de las observaciones antes del Lema 8.3.9, tenemos que

$$\begin{aligned}\mathbf{1}_{\{N=0\}^{s-1}\times\{N>0\}}\big(\Psi(\boldsymbol{h}^*)\big) &= \prod_{i=1}^{s-1} \mathbf{1}_{\{N=0\}}\big(\psi_i(\boldsymbol{h}^*)\big)\cdot \mathbf{1}_{\{N>0\}}\big(\psi_s(\boldsymbol{h}^*)\big)\\ &= \prod_{i=1}^{s-1} \mathbf{1}_{\{N=0\}}\big(\psi_i(\boldsymbol{h}^*_i)\big)\cdot \mathbf{1}_{\{N>0\}}\big(\psi_s(\boldsymbol{h}^*_s)\big).\end{aligned}$$

Por lo tanto, podemos reescribir la expresión (8.21) de R_s de la siguiente manera:

$$R_s = \sum_{h^*_1\in\mathbb{F}_q^{d+1}} \mathbf{1}_{\{N=0\}}\big(\psi_1(\boldsymbol{h}^*_1)\big) \cdots \sum_{h^*_{s-1}\in\mathbb{F}_q^{d-\kappa_{s-1}+1}} \mathbf{1}_{\{N=0\}}\big(\psi_{s-1}(\boldsymbol{h}^*_{s-1})\big) \sum_{h^*_s\in\mathbb{F}_q^{d-\kappa_s+1}} \mathbf{1}_{\{N>0\}}\big(\psi_s(\boldsymbol{h}^*_s)\big).$$

Obtenemos así el enunciado del lema. □

Para $1\leq i\leq s-1$, fijamos el elemento $h^*_i \in \mathbb{F}_q^{d+1-\kappa_i}$. Para cada $h^*_s := (h_{d-\kappa_s,s},\dots,h_{0,s}) \in \mathbb{F}_q^{d+1-\kappa_s}$, denotamos por $f_{h^*_s}$ el polinomio univariado

$$f_{h^*_s} := \psi_s(h^*_1,\dots,h^*_s) := h_{d,s}T^d + \cdots + h_{d-\kappa_s+1,s}T^{d-\kappa_s+1} + h_{d-\kappa_s,s}T^{d-\kappa_s} + \cdots + h_{0,s}.$$

Del Lema 8.3.9 vemos que, a fin de estimar R_s, alcanza con estimar la suma

$$\sum_{h^*_s\in\mathbb{F}_q^{d-\kappa_s+1}} \mathbf{1}_{\{N>0\}}(f_{h^*_s}). \tag{8.22}$$

Para cada $h^*_s := (h_{d-\kappa_s,s},\dots,h_{0,s}) \in \mathbb{F}_q^{d+1-\kappa_s}$, denotamos con $\widehat{h}^*_s := (h_{d-\kappa_s,s},\dots,h_{1,s}) \in \mathbb{F}_q^{d-\kappa_s}$ y $f_{\widehat{h}^*_s} := \sum_{j=1}^d h_{j,s}T^j = f_{h^*_s} - f_{h^*_s}(0)$. Recordemos que el cardinal del conjunto

de valores $\mathcal{V}(f)$ de $f \in \mathbb{F}_q[T]$ se define como $\mathcal{V}(f) := |\{f(c) : c \in \mathbb{F}_q\}|$. Como observamos en la Sección 5.2.1, el cardinal $\mathcal{V}(f_{\widehat{h}_s^*})$ del conjunto de valores de $f_{\widehat{h}_s^*}$ es igual al número de $h_{0,s} \in \mathbb{F}_q$ para los cuales el polinomio $f_{\widehat{h}_s^*} + h_{0,s}$ tiene al menos una raíz en $\mathbb{F}_q$. Por lo tanto,

$$\sum_{h_s^* \in \mathbb{F}_q^{d+1-\kappa_s}} \mathbf{1}_{\{N>0\}}(f_{h_s^*}) = \sum_{\widehat{h}_s^* \in \mathbb{F}_q^{d-\kappa_s}} \sum_{h_{0,s} \in \mathbb{F}_q} \mathbf{1}_{\{N>0\}}(f_{h_s^*}) = \sum_{\widehat{h}_s^* \in \mathbb{F}_q^{d-\kappa_s}} \mathcal{V}(f_{\widehat{h}_s^*})$$
$$= \frac{1}{q} \sum_{h_s^* \in \mathbb{F}_q^{d+1-\kappa_s}} \mathcal{V}(f_{h_s^*}), \quad (8.23)$$

En consecuencia, podemos describir la suma (8.22) en términos de la suma del cardinal del conjunto de valores de los elementos de la familia $\{f_{h_s^*} : h_s^* \in \mathbb{F}_q^{d+1-\kappa_s}\}$. El Lema 8.3.7 prueba que $h_{d,s}, \ldots, h_{d+1-\kappa_s,s}$ están unívocamente determinados por $\boldsymbol{h}_{s-1}^* := (h_1^*, \ldots, h_{s-1}^*)$. Así, la suma del lado derecho de (8.23) toma como argumento el cardinal del conjunto de valores de todos los elementos de $\mathbb{F}_q[T]_{\leq d}$ con los primeros κ_s coeficientes $(h_{d,s}, \ldots, h_{d+1-\kappa_s,s})$ prescriptos. Denotemos con $\psi_s^{\mathrm{fix}}(\boldsymbol{h}_{s-1}^*) := (h_{d,s}, \ldots, h_{d+1-\kappa_s,s})$ y con $\mathcal{V}_d(\kappa_s, \psi_s^{\mathrm{fix}}(\boldsymbol{h}_{s-1}^*))$ el valor promedio del cardinal del conjunto de valores de la familia $\{f_{h_s^*} : h_s^* \in \mathbb{F}_q^{d+1-\kappa_s}\}$, es decir,

$$\mathcal{V}_d(\kappa_s, \psi_s^{\mathrm{fix}}(\boldsymbol{h}_{s-1}^*)) := \frac{1}{q^{d+1-\kappa_s}} \sum_{h_s^* \in \mathbb{F}_q^{d+1-\kappa_s}} \mathcal{V}(f_{h_s^*}). \quad (8.24)$$

En el siguiente lema expresamos la probabilidad $p_{r,d}[C_{\underline{\boldsymbol{a}}} = s]$ en términos del valor promedio $\mathcal{V}_d(\kappa_s, \psi_s^{\mathrm{fix}}(\boldsymbol{h}_{s-1}^*))$.

Lema 8.3.10. *Para $s \leq \min\{D_d, q^{r-1}\}$, vale la siguiente identidad:*

$$p_{r,d}[C_{\underline{\boldsymbol{a}}} = s] = \frac{1}{q^{\sum_{i=1}^{s-1}(d+1-\kappa_i)}} \sum_{\substack{h_1^* \in \mathbb{F}_q^{d+1} \\ N(\psi_1(\boldsymbol{h}_1^*))=0}} \cdots \sum_{\substack{h_{s-1}^* \in \mathbb{F}_q^{d+1-\kappa_{s-1}} \\ N(\psi_{s-1}(\boldsymbol{h}_{s-1}^*))=0}} \frac{\mathcal{V}_d(\kappa_s, \psi_s^{\mathrm{fix}}(\boldsymbol{h}_{s-1}^*))}{q}.$$

Demostración. Del Lema 8.3.4 tenemos que $\dim \mathrm{Im}(\Phi) = \sum_{i=1}^{s}(d+1-\kappa_i)$. Combinando esta identidad con (8.12) y el Lema 8.3.9, obtenemos que

$$p_{r,d}[C_{\underline{\boldsymbol{a}}} = s] =$$
$$\frac{1}{q^{\sum_{i=1}^{s-1}(d+1-\kappa_i)}} \sum_{\substack{h_1^* \in \mathbb{F}_q^{d+1} \\ N(\psi_1(\boldsymbol{h}_1^*))=0}} \cdots \sum_{\substack{h_{s-1}^* \in \mathbb{F}_q^{d+1-\kappa_{s-1}} \\ N(\psi_{s-1}(\boldsymbol{h}_{s-1}^*))=0}} \frac{1}{q^{d+1-\kappa_s}} \sum_{h_s^* \in \mathbb{F}_q^{d+1-\kappa_s}} \mathbf{1}_{\{N>0\}}\big(\psi_s(\boldsymbol{h}_s^*)\big).$$

De (8.23) y de (8.24) deducimos el enunciado del lema. □

A continuación estimamos la probabilidad $p_{r,d}[C_{\underline{\boldsymbol{a}}} = s]$ utilizando las estimaciones explícitas sobre el promedio del cardinal del conjunto de valores en familias lineales de los Capítulos 5 y 6. Más precisamente, supongamos que $s \leq \min\{D_{d-2}, q^{r-1}\}$. Por los Teoremas 5.3.5 y 6.2.4, tenemos estimaciones explícitas del promedio $\mathcal{V}_d(\kappa_s, \psi_s^{\mathrm{fix}}(\boldsymbol{h}_{s-1}^*))$

de (8.24) para cualquier $\boldsymbol{h}^*_{s-1}$ tal que $f_{h^*_s}$ es de grado d. Estas estimaciones, junto con el Lema 8.3.10, permiten expresar la probabilidad $p_{r,d}[C_{\underline{\boldsymbol{a}}} = s]$ en términos del promedio del cardinal del conjunto de valores de las familias de polinomios que introducimos a continuación. Para $1 \leq i \leq s-1$ y $1 \leq j \leq i-1$, fijemos $h^*_j := (h_{d-\kappa_j,j}, \ldots, h_{0,j}) \in \mathbb{F}_q^{d+1-\kappa_j}$. Para cada $h^*_i := (h_{d-\kappa_i,i}, \ldots, h_{0,i}) \in \mathbb{F}_q^{d+1-\kappa_i}$, denotamos por $f_{h^*_i}$ el polinomio

$$f_{h^*_i} := \psi_i(h^*_1, \ldots, h^*_i) := h_{d,i}T^d + \cdots + h_{d+1-\kappa_i,i}T^{d+1-\kappa_i} + h_{d-\kappa_i,i}T^{d-\kappa_i} + \cdots + h_{0,i}.$$

De acuerdo al Lema 8.3.7, los coeficientes $h_{d,i}, \ldots, h_{d+1-\kappa_i,i}$ están unívocamente determinados por $\boldsymbol{h}^*_{i-1} := (h^*_1, \ldots, h^*_{i-1})$. En consecuencia, consideramos $\psi_i^{\mathrm{fix}}(\boldsymbol{h}^*_{i-1}) := (h_{d,i}, \ldots, h_{d+1-\kappa_i,i})$ y el promedio del cardinal $\mathcal{V}(d, \kappa_i, \psi_i^{\mathrm{fix}}(\boldsymbol{h}^*_{i-1}))$ del conjunto de valores de la familia $\{f_{h^*_i} : h^*_i \in \mathbb{F}_q^{d+1-\kappa_i}\}$, es decir,

$$\mathcal{V}_d(\kappa_i, \psi_i^{\mathrm{fix}}(\boldsymbol{h}^*_{i-1})) := \frac{1}{q^{d+1-\kappa_i}} \sum_{h^*_i \in \mathbb{F}_q^{d+1-\kappa_i}} \mathcal{V}(f_{h^*_i}). \tag{8.25}$$

Nuevamente, los Teoremas 5.3.5 y 6.2.4 proveen estimaciones explícitas del promedio $\mathcal{V}_d(\kappa_i, \psi_i^{\mathrm{fix}}(\boldsymbol{h}^*_{i-1}))$ para cualquier $\boldsymbol{h}^*_{i-1}$ tal que $f_{h^*_i}$ es de grado d, para $1 \leq i \leq s-1$. Observemos que dichas estimaciones valen para familias de polinomios con ciertos coeficientes prescriptos y término independiente nulo, y que los polinomios $f_{h^*_i}$ que consideramos tienen término independiente no necesariamente nulo. Sin embargo, podemos aplicar tales resultados a estas familias. Más precisamente, de la igualdad

$$\frac{1}{q^{d-\kappa_i}} \sum_{\widehat{h}^*_i \in \mathbb{F}_q^{d-\kappa_i}} \mathcal{V}(f_{\widehat{h}^*_i}) = \frac{1}{q^{d+1-\kappa_i}} \sum_{h^*_i \in \mathbb{F}_q^{d+1-\kappa_i}} \mathcal{V}(f_{h^*_i})$$

y las estimaciones de los Teoremas 5.3.5 y 6.2.4 deducimos el siguiente resultado.

Observación 8.3.11. *Si $1 \leq i \leq s$ y $1 \leq \kappa_i \leq d/2$, entonces*

$$|\mathcal{V}_d(\kappa_i, \psi_i^{\mathrm{fix}}(\boldsymbol{h}^*_{i-1})) - \mu_d\, q| \leq \frac{e^{-1}}{2} + \frac{(d-2)^5 e^{2\sqrt{d}}}{2^{d-2}} + \frac{7}{q}. \tag{8.26}$$

Si $p > 2$, y $1 \leq \kappa_i \leq d-2$ para $1 \leq i \leq s$, tenemos que

$$|\mathcal{V}_d(\kappa_i, \psi_i^{\mathrm{fix}}(\boldsymbol{h}^*_{i-1})) - \mu_d\, q| \leq d^2\, 2^{d-1} q^{\frac{1}{2}} + 133\, d^{d+5} e^{2\sqrt{d}-d}. \tag{8.27}$$

El siguiente resultado expresa la probabilidad $p_{r,d}[C_{\underline{\boldsymbol{a}}} = s]$ en términos del promedio $\mathcal{V}_d(\kappa_i, \psi_i^{\mathrm{fix}}(\boldsymbol{h}^*_{i-1}))$ para $1 \leq i \leq s$.

Teorema 8.3.12. *Para $s \leq \min\{D_d, q^{r-1}\}$, tenemos*

$$p_{r,d}[C_{\underline{\boldsymbol{a}}} = s] = (1-\mu_d)^{s-1}\mu_d\, \frac{q-1}{q} + \sum_{i=0}^{s} \mathcal{T}_i,$$

donde $|\mathcal{T}_0| \leq 1/q$,

$$\mathcal{T}_i := (1-\mu_d)^{s-i-1}\mu_d \frac{q-1}{q^{\sum_{j=1}^{i-1}(d+1-\kappa_j)}} \sum_{\substack{h_1^* \in \mathbb{F}_q^{d+1} \\ N(\psi_1(\boldsymbol{h}_1^*))=0 \\ h_{d,1}=1}} \cdots \sum_{\substack{h_{i-1}^* \in \mathbb{F}_q^{d+1-\kappa_{i-1}} \\ N(\psi_{i-1}(\boldsymbol{h}_{i-1}^*))=0}} \left(\mu_d - \frac{\mathcal{V}_d(\kappa_i, \psi_i^{\mathrm{fix}}(\boldsymbol{h}_{i-1}^*))}{q}\right) \tag{8.28}$$

para $1 \leq i \leq s-1$, *y*

$$\mathcal{T}_s := \frac{q-1}{q^{\sum_{i=1}^{s-1}(d+1-\kappa_i)}} \sum_{\substack{h_1^* \in \mathbb{F}_q^{d+1} \\ N(\psi_1(\boldsymbol{h}_1^*))=0 \\ h_{d,1}=1}} \cdots \sum_{\substack{h_{s-1}^* \in \mathbb{F}_q^{d+1-\kappa_{s-1}} \\ N(\psi_{s-1}(\boldsymbol{h}_{s-1}^*))=0}} \left(\frac{\mathcal{V}_d(\kappa_s, \psi_s^{\mathrm{fix}}(\boldsymbol{h}_{s-1}^*))}{q} - \mu_d\right). \tag{8.29}$$

Demostración. Denotemos con $C := C_{\underline{\boldsymbol{a}}}$. Utilizando la Observación 8.3.8, vamos a descomponer la expresión para $p_{r,d}[C = s]$ de la Proposición 8.3.10 en dos sumas, dependiendo de si $h_{d,1} = 0$ o no. Más precisamente, escribimos

$$p_{r,d}[C = s] = p_{r,d}[C = s, F_d = 0] + p_{r,d}[C = s, F_d \neq 0],$$

donde

$$p_{r,d}[C = s, F_d = 0] = \frac{1}{q^{\sum_{i=1}^{s-1}(d+1-\kappa_i)}} \sum_{\substack{h_1^* \in \mathbb{F}_q^{d+1} \\ N(\psi_1(\boldsymbol{h}_1^*))=0 \\ h_{d,1}=0}} \cdots \sum_{\substack{h_{s-1}^* \in \mathbb{F}_q^{d+1-\kappa_{s-1}} \\ N(\psi_{s-1}(\boldsymbol{h}_{s-1}^*))=0}} \frac{\mathcal{V}_d(\kappa_s, \psi_s^{\mathrm{fix}}(\boldsymbol{h}_{s-1}^*))}{q},$$

$$\begin{aligned} p_{r,d}[C = s, F_d \neq 0] &= \frac{1}{q^{\sum_{i=1}^{s-1}(d+1-\kappa_i)}} \sum_{\substack{h_1^* \in \mathbb{F}_q^{d+1} \\ N(\psi_1(\boldsymbol{h}_1^*))=0 \\ h_{d,1}\neq 0}} \cdots \sum_{\substack{h_{s-1}^* \in \mathbb{F}_q^{d+1-\kappa_{s-1}} \\ N(\psi_{s-1}(\boldsymbol{h}_{s-1}^*))=0}} \frac{\mathcal{V}_d(\kappa_s, \psi_s^{\mathrm{fix}}(\boldsymbol{h}_{s-1}^*))}{q}, \\ &= \frac{q-1}{q^{\sum_{i=1}^{s-1}(d+1-\kappa_i)}} \sum_{\substack{h_1^* \in \mathbb{F}_q^{d+1} \\ N(\psi_1(\boldsymbol{h}_1^*))=0 \\ h_{d,1}=1}} \cdots \sum_{\substack{h_{s-1}^* \in \mathbb{F}_q^{d+1-\kappa_{s-1}} \\ N(\psi_{s-1}(\boldsymbol{h}_{s-1}^*))=0}} \frac{\mathcal{V}_d(\kappa_s, \psi_s^{\mathrm{fix}}(\boldsymbol{h}_{s-1}^*))}{q}. \end{aligned}$$

A fin de estimar el primer término, como $\mathrm{Im}(\Phi)$ no está contenida en $\mathbb{F}_q[T]^s_{\leq d-1}$, la intersección $\mathrm{Im}(\Phi) \cap \mathbb{F}_q[T]^s_{\leq d-1}$ tiene codimensión al menos 1 en $\mathrm{Im}(\Phi)$. Combinando esta observación con el Lema 8.3.1, vemos que

$$\mathcal{T}_0 := p_{r,d}[C = s, F_d = 0] \leq \frac{|\mathrm{Im}(\Phi) \cap \mathcal{F}^s_{1,d-1}|}{|\mathrm{Im}(\Phi)|} \leq \frac{q^{\dim \mathrm{Im}(\Phi)-1}}{q^{\dim \mathrm{Im}(\Phi)}} = \frac{1}{q}.$$

Por otro lado, es fácil ver que la expresión para $p_{r,d}[C = s, F_d \neq 0]$ se puede reescribir de la siguiente manera:

$$p_{r,d}[C = s, F_d \neq 0] = \mu_d \frac{q-1}{q^{\sum_{i=1}^{s-1}(d+1-\kappa_i)}} \sum_{\substack{h_1^* \in \mathbb{F}_q^{d+1} \\ N(\psi_1(\boldsymbol{h}_1^*))=0 \\ h_{d,1}=1}} \cdots \sum_{\substack{h_{s-1}^* \in \mathbb{F}_q^{d+1-\kappa_{s-1}} \\ N(\psi_{s-1}(\boldsymbol{h}_{s-1}^*))=0}} 1 + \mathcal{T}_s,$$

donde $\mathcal{T}_s$ está definido en (8.29).

Afirmamos que, para $1 \le j \le s$,

$$p_{r,d}[C = s, F_d \neq 0] = (1-\mu_d)^{s-j}\mu_d \frac{q-1}{q^{\sum_{i=1}^{j-1}(d+1-\kappa_i)}} \sum_{\substack{h_1^* \in \mathbb{F}_q^{d+1} \\ N(\psi_1(\boldsymbol{h}_1^*))=0 \\ h_{d,1}=1}} \cdots \sum_{\substack{h_{j-1}^* \in \mathbb{F}_q^{d+1-\kappa_{j-1}} \\ N(\psi_{j-1}(\boldsymbol{h}_{j-1}^*))=0}} 1 + \sum_{i=j}^{s} \mathcal{T}_i, \tag{8.30}$$

donde $\mathcal{T}_i$ está definido en (8.28). Observemos que, si $j = 1$, entonces (8.30) es la afirmación del teorema.

Demostramos (8.30) por inducción "regresiva" en j, comenzando en $j = s$ y terminando en $j = 1$. Comenzamos observando que ya demostramos anteriormente el caso $j = s$. Para $j < s$, supongamos que (8.30) es válida para $j+1$; queremos ver que es válida también para j. Tenemos que

$$\frac{1}{q^{d+1-\kappa_j}} \sum_{\substack{h_j^* \in \mathbb{F}_q^{d+1-\kappa_j} \\ N(\psi_j(\boldsymbol{h}_j^*))=0}} 1 = 1 - \frac{1}{q^{d+1-\kappa_j}} \sum_{\substack{h_j^* \in \mathbb{F}_q^{d+1-\kappa_j} \\ N(\psi_j(\boldsymbol{h}_j^*))>0}} 1 = 1 - \frac{\mathcal{V}_d(\kappa_j, \psi_j^{\mathrm{fix}}(\boldsymbol{h}_{j-1}^*))}{q}.$$

Así, reemplazando esta identidad en la expresión para $p_{r,d}[C = s, F_d \neq 0]$ correspondiente a la afirmación para $j+1$, deducimos la afirmación (8.30) para el caso j. Esto concluye la demostración del teorema. □

En el siguiente teorema, el más importante de esta sección, exhibimos estimaciones explícitas de la probabilidad del evento $\{C_{\underline{\boldsymbol{a}}} = s\}$.

Teorema 8.3.13. *Sea* $\underline{\boldsymbol{a}} := (\boldsymbol{a}_1, \dots, \boldsymbol{a}_s) \in \mathbb{F}_q^{s(r-1)} \setminus \mathtt{B}_s$, *siendo* $\mathtt{B}_s$ *el conjunto de* (8.15). *Para* $s \le \min\left\{\binom{d/2+r-1}{r-1}, q^{r-1}\right\}$, *tenemos que*

$$\left|p_{r,d}[C_{\underline{\boldsymbol{a}}} = s] - (1-\mu_d)^{s-1}\mu_d\right| \le \left(e^{-1} + \frac{(d-2)^5 e^{2\sqrt{d}}}{2^{d-1}} + 1\right) q^{-1} + 14q^{-2}.$$

Por otro lado, si $p > 2$ *y* $s \le \min\left\{\binom{d+r-3}{r-1}, q^{r-1}\right\}$, *entonces*

$$\left|p_{r,d}[C_{\underline{\boldsymbol{a}}} = s] - (1-\mu_d)^{s-1}\mu_d\right| \le d^2\, 2^d q^{-\frac{1}{2}} + (266\, d^{d+5} e^{2\sqrt{d}-d} + 1) q^{-1}.$$

Demostración. Si $s \le \min\left\{\binom{d/2+r-1}{r-1}, q^{r-1}\right\}$, entonces $\kappa_s \le d/2$. Para $1 \le i \le s$, por la definición de κ_i tenemos que $1 \le \kappa_i \le \kappa_s \le d/2$, y por lo tanto, de (8.26) deducimos que

$$\left|\frac{\mathcal{V}_d(\kappa_i, \psi_i^{\mathrm{fix}}(\boldsymbol{h}_{i-1}^*))}{q} - \mu_d\right| \le \left(\frac{e^{-1}}{2} + \frac{(d-2)^5 e^{2\sqrt{d}}}{2^{d-2}}\right) q^{-1} + 7q^{-2}$$

para todo $1 \le i \le s$. Así, tenemos las siguientes estimaciones para las expresiones $\mathcal{T}_i$ con $1 \le i \le s$ de (8.28) y (8.29):

$$|\mathcal{T}_i| \le (1-\mu_d)^{s-i-1}\mu_d \left(\left(\frac{e^{-1}}{2} + \frac{(d-2)^5 e^{2\sqrt{d}}}{2^{d-2}}\right) q^{-1} + 7q^{-2}\right) \quad (1 \le i \le s-1),$$

$$|\mathcal{T}_s| \le \left(\frac{e^{-1}}{2} + \frac{(d-2)^5 e^{2\sqrt{d}}}{2^{d-2}}\right) q^{-1} + 7q^{-2}.$$

Combinando estas estimaciones con el Teorema 8.3.12 obtenemos la primera parte del teorema.

Por otro lado, si $p > 2$ y $s \leq \min\left\{\binom{d+r-3}{r-1}, q^{r-1}\right\}$, entonces $\kappa_s \leq d-2$. Para $1 \leq i \leq s$, por la definición de κ_i tenemos que $1 \leq \kappa_i \leq \kappa_s \leq d-2$, y de (8.27) concluimos que

$$\left|\frac{\mathcal{V}_d(\kappa_i, \psi_i^{\mathrm{fix}}(\boldsymbol{h}_{i-1}^*))}{q} - \mu_d\right| \leq d^2\, 2^{d-1} q^{-\frac{1}{2}} + 133\, d^{d+5} e^{2\sqrt{d}-d} q^{-1}$$

para $1 \leq i \leq s$. Así, tenemos las siguientes estimaciones para $\mathcal{T}_i$:

$$|\mathcal{T}_i| \leq (1-\mu_d)^{s-i-1}\mu_d\big(d^2\, 2^{d-1} q^{-\frac{1}{2}} + 133\, d^{d+5} e^{2\sqrt{d}-d} q^{-1}\big) \quad (1 \leq i \leq s-1),$$
$$|\mathcal{T}_s| \leq d^2\, 2^{d-1} q^{-\frac{1}{2}} + 133\, d^{d+5} e^{2\sqrt{d}-d} q^{-1}.$$

De estas estimaciones y el Teorema 8.3.12 deducimos la segunda parte del teorema. □

Para terminar esta sección, observamos que el esquema de demostración del Teorema 8.3.13 no puede aplicarse para estimar la probabilidad de que se realicen $s > s^* := \binom{d+r-3}{r-1}$ búsquedas en bandas verticales hasta que el algoritmo encuentra un cero $\mathbb{F}_q$–racional del polinomio en consideración, ya que el comportamiento de la función $\Phi := \Phi_{\underline{\boldsymbol{a}}} : \mathbb{F}_q[\boldsymbol{X}]_{\leq d} \to \mathbb{F}_q[T]^s_{\leq d}$ de (8.11) puede cambiar significativamente en este caso. Sin embargo, para valores grandes de s, del Teorema 8.3.13 y de la igualdad $p_{r,d}[C_{\underline{\boldsymbol{a}}} > s^*] = 1 - \sum_{i=1}^{s^*} p_{r,d}[C_{\underline{\boldsymbol{a}}} = i]$ se deduce el siguiente resultado.

Corolario 8.3.14. *Con las notaciones del Teorema 8.3.13, para* $s^* := \min\left\{\binom{\frac{d}{2}+r-1}{r-1}, q^{r-1}\right\}$ *tenemos*

$$p_{r,d}[C_{\underline{\boldsymbol{a}}} > s^*] = (1-\mu_d)^{s^*} + \mathcal{O}(q^{-1}).$$

Por otro lado, si $p > 2$ *y* $s^* := \min\left\{\binom{d+r-3}{r-1}, q^{r-1}\right\}$, *entonces*

$$p_{r,d}[C_{\underline{\boldsymbol{a}}} > s^*] = (1-\mu_d)^{s^*} + \mathcal{O}(q^{-1/2}).$$

Como $|1-\mu_d| \leq 1/2$, de la expresión para s^* se sigue que el término principal de esta probabilidad decrece exponencialmente con r y d.

8.4. Análisis probabilístico del Algoritmo BBV

En esta sección determinamos la complejidad en promedio del Algoritmo BBV. Para realizar dicho análisis, vamos a estudiar la distribución de probabilidades del número de búsquedas que dicho algoritmo realiza hasta encontrar una raíz del polinomio de partida.

8.4.1. Distribución de probabilidades del número de búsquedas

Similarmente a la sección 8.2, para $s \geq 3$ denotamos por

$$\mathsf{F}_s := \{(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_s) \in \mathbb{F}_q^{r-1} \times \cdots \times \mathbb{F}_q^{r-1} : \boldsymbol{a}_i \neq \boldsymbol{a}_j \text{ for } i \neq j\}, \quad N_s := |\mathsf{F}_s|,$$

y consideramos la variable aleatoria $C_s := C_{s,r,d} : \mathsf{F}_s \times \mathbb{F}_q[\boldsymbol{X}]_{\leq d} \to \{1, \ldots, s, \infty\}$ definida por $\underline{\boldsymbol{a}} := (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_s) \in \mathsf{F}_s$ y $F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$ de la siguiente manera:

$$C_s(\underline{\boldsymbol{a}}, F) := \begin{cases} \text{mín}\{j : N_{1,d}(F(\boldsymbol{a}_j, X_r)) > 0\} & \text{si } \exists j \text{ con } N_{1,d}(F(\boldsymbol{a}_j, X_r)) > 0, \\ \infty & \text{en otro caso.} \end{cases}$$

Consideramos la probabilidad uniforme $P_s := P_{s,r,d}$ sobre el espacio muestral $\mathsf{F}_s \times \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$ y analizamos la probabilidad $P_s[C_s = s]$.

A continuación damos un resultado que vincula los espacios de probabilidad determinados por $\mathsf{F}_s \times \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$ y P_s para todo $1 \leq s \leq q^{r-1}$.

Lema 8.4.1. *Sea $s > 1$ y sea $\pi_s : \mathsf{F}_s \times \mathbb{F}_q[\boldsymbol{X}]_{\leq d} \to \mathsf{F}_{s-1} \times \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$ la función que induce la proyección $\mathsf{F}_s \to \mathsf{F}_{s-1}$ sobre las primeras $s-1$ coordenadas. Si $\mathcal{S} \subset \mathsf{F}_{s-1} \times \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$, entonces $P_s[\pi_s^{-1}(\mathcal{S})] = P_{s-1}[\mathcal{S}]$.*

Demostración. Dado $\mathcal{S} \subset \mathsf{F}_{s-1} \times \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$, tenemos que

$$\begin{aligned} \pi_s^{-1}(\mathcal{S}) &= \bigcup_{F \in \mathcal{F}_{r,d}} \{(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_s) \in \mathsf{F}_s : (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{s-1}, F) \in \mathcal{S}\} \times \{F\} \\ &= \bigcup_{F \in \mathcal{F}_{r,d}} \bigcup_{\substack{(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{s-1}) \in \mathsf{F}_{s-1}: \\ (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{s-1}, F) \in \mathcal{S}}} \{(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{s-1})\} \times (\mathbb{F}_q^{r-1} \setminus \{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{s-1}\}) \times \{F\}. \end{aligned}$$

Se sigue que

$$\begin{aligned} P_s[\pi_s^{-1}(\mathcal{S})] &= \frac{1}{N_s |\mathbb{F}_q[\boldsymbol{X}]_{\leq d}|} \sum_{F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}} \sum_{\underline{\boldsymbol{a}} \in \mathsf{F}_{s-1} : (\underline{\boldsymbol{a}}, F) \in \mathcal{S}} (q^{r-1} - s + 1) \\ &= \frac{q^{r-1} - s + 1}{N_{s-1} |\mathbb{F}_q[\boldsymbol{X}]_{\leq d}| (q^{r-1} - s + 1)} \sum_{F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}} \big|\{\underline{\boldsymbol{a}} \in \mathsf{F}_{s-1} : (\underline{\boldsymbol{a}}, F) \in \mathcal{S}\}\big| \\ &= \frac{1}{N_{s-1} |\mathbb{F}_q[\boldsymbol{X}]_{\leq d}|} \sum_{F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}} \big|\{\underline{\boldsymbol{a}} \in \mathsf{F}_{s-1} : (\underline{\boldsymbol{a}}, F) \in \mathcal{S}\}\big| = P_{s-1}[\mathcal{S}]. \end{aligned}$$

Esto prueba el lema. □

Por el teorema de extensión de Kolmogorov (ver, por ejemplo, [Fel91, Chapter IV, Section 5, Extension Theorem]), la condición de consistencia del Lema 8.4.1 implica que las probabilidades P_s con $1 \leq s \leq q^{r-1}$ pueden considerarse en un marco unificado. Más precisamente, definimos $\mathsf{F} := \mathsf{F}_{q^{r-1}}$ y $P := P_{q^{r-1}}$. La medida de probabilidad P definida sobre el espacio muestral F nos permite interpretar consistentemente los

resultados de esta sección. De la misma manera, las variables C_s $(1 \leq s \leq q^{r-1})$ se extienden naturalmente a la variable aleatoria $C : \mathsf{F} \times \mathbb{F}_q[\boldsymbol{X}]_{\leq d} \to \mathbb{N} \cup \{\infty\}$. En particular, las variables aleatorias C_1 y C_2 analizadas en la Sección 8.2 resultan la restricción de C a F_1 y F_2 respectivamente. En lo que sigue eliminamos el subíndice s de las notaciones de P_s y C_s.

Para el análisis de la distribución de probabilidad del número de búsquedas, expresamos la probabilidad $P[C = s]$ en términos de las probabilidades de los eventos $\{C_{\underline{\boldsymbol{a}}} = s\}$, donde $C_{\underline{\boldsymbol{a}}} : \mathbb{F}_q[\boldsymbol{X}]_{\leq d} \to \mathbb{N}$ es la variable aleatoria que cuenta el número de búsquedas que el algoritmo realiza para una entrada F sobre las bandas verticales que determina $\underline{\boldsymbol{a}} \in \mathsf{F}_s$. El resultado correspondiente es una generalización del Lema 8.2.4.

Lema 8.4.2. *Tenemos que*

$$P[C = s] = \frac{1}{N_s} \sum_{\underline{\boldsymbol{a}} \in \mathsf{F}_s} p_{r,d}[C_{\underline{\boldsymbol{a}}} = s].$$

Demostración. El conjunto $[C = s]$ se expresa como una unión disjunta de conjuntos de la siguiente manera:

$$\{C = s\} = \bigcup_{\underline{\boldsymbol{a}} \in \mathsf{F}_s} \{\underline{\boldsymbol{a}}\} \times \{F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d} : C_{\underline{\boldsymbol{a}}}(F) = s\}.$$

Por lo tanto,

$$P[C = s] = \frac{1}{N_s} \sum_{\underline{\boldsymbol{a}} \in \mathsf{F}_s} \frac{\left|\{F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d} : C_{\underline{\boldsymbol{a}}}(F) = s\}\right|}{|\mathbb{F}_q[\boldsymbol{X}]_{\leq d}|} = \frac{1}{N_s} \sum_{\underline{\boldsymbol{a}} \in \mathsf{F}_s} p_{r,d}[C_{\underline{\boldsymbol{a}}} = s],$$

Esto demuestra la afirmación del lema. □

En el Teorema 8.3.13 determinamos el comportamiento asintótico de la probabilidad $p_{r,d}[C_{\underline{\boldsymbol{a}}} = s]$ para $\underline{\boldsymbol{a}} \notin \mathsf{B}_s$, donde B_s es el conjunto definido en (8.15). Notemos que $\mathsf{B}_s \subset \mathsf{F}_s$; así, el Teorema 8.3.13 se cumple para cada $\underline{\boldsymbol{a}} \in \mathsf{F}_s \setminus \mathsf{B}_s$. Por (8.14) tenemos que $|\mathsf{B}_s| = \mathcal{O}(q^{s(r-1)-1})$, donde la constante que aparece en la notación $\mathcal{O}$ depende de s, d y r, pero es independiente de q.

Terminamos esta sección, dando una estimación de la probabilidad $P[C = s]$. El Lema 8.4.2 implica que

$$\begin{aligned} P[C = s] &= \frac{1}{N_s} \sum_{\underline{\boldsymbol{a}} \in \mathsf{F}_s} p_{r,d}[C_{\underline{\boldsymbol{a}}} = s] \\ &= \frac{1}{N_s} \sum_{\underline{\boldsymbol{a}} \in \mathsf{F}_s \setminus \mathsf{B}_s} p_{r,d}[C_{\underline{\boldsymbol{a}}} = s] + \frac{1}{N_s} \sum_{\underline{\boldsymbol{a}} \in \mathsf{B}_s} p_{r,d}[C_{\underline{\boldsymbol{a}}} = s] \\ &= \frac{1}{N_s} \sum_{\underline{\boldsymbol{a}} \in \mathsf{F}_s \setminus \mathsf{B}_s} p_{r,d}[C_{\underline{\boldsymbol{a}}} = s] + \mathcal{O}(q^{-1}). \end{aligned}$$

Sea $\underline{\boldsymbol{a}} \in \mathsf{F}_s \setminus \mathtt{B}_s$. Por el Teorema 8.3.13, si $s \leq \binom{d/2+r-1}{r-1}$, entonces $p_{r,d}[C_{\underline{\boldsymbol{a}}} = s] = (1-\mu_d)^{s-1}\mu_d + \mathcal{O}(q^{-1})$. Por otro lado, si $p > 2$ y $s \leq \binom{d+r-3}{r-1}$, entonces $p_{r,d}[C_{\underline{\boldsymbol{a}}} = s] = (1-\mu_d)^{s-1}\mu_d + \mathcal{O}(q^{-1/2})$. En consecuencia, tenemos el siguiente resultado.

Teorema 8.4.3. *Para $s \leq \binom{d/2+r-1}{r-1}$, tenemos que*

$$P[C = s] = (1-\mu_d)^{s-1}\mu_d + \mathcal{O}(q^{-1}).$$

Por otro lado, si $p > 2$ y $s \leq \binom{d+r-3}{r-1}$, entonces

$$P[C = s] = (1-\mu_d)^{s-1}\mu_d + \mathcal{O}(q^{-1/2}).$$

8.4.2. Complejidad en promedio

En esta sección obtenemos una cota superior de la complejidad en promedio del Algoritmo BBV. Recordemos que, dado $F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$, el Algoritmo BBV genera sucesivamente una sucesión $\underline{\boldsymbol{a}} := (\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_{q^{r-1}}) \in \mathsf{F}_{q^{r-1}}$, y busca ceros $\mathbb{F}_q$–racionales de F en las bandas $\{\boldsymbol{a}_i\} \times \mathbb{F}_q$ para $1 \leq i \leq q^{r-1}$, hasta que encuentra un cero de F o se terminan las bandas verticales. En el Lema 8.1.1 probamos que el Algoritmo BBV realiza $\tau(d,r,q)\, C_{\underline{\boldsymbol{a}}}(F)$ operaciones aritméticas en $\mathbb{F}_q$, donde $\tau(d,r,q) := \mathcal{O}^{\sim}(D + d\log q)$ es el número máximo de operaciones aritméticas en $\mathbb{F}_q$ necesarias para realizar una búsqueda en una banda vertical arbitraria y $C_{\underline{\boldsymbol{a}}}(F)$ es el número de búsquedas que el algoritmo realiza en las bandas verticales determinadas por $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{q^{r-1}}$ hasta encontrar un cero $\mathbb{F}_q$–racional de F.

El Algoritmo BBV tiene una rutina probabilística que busca ceros $\mathbb{F}_q$–racionales de los elementos de $\mathbb{F}_q[T]_{\leq d}$, que realiza r_d elecciones aleatorias de elementos de $\mathbb{F}_q$, para un $r_d \in \mathbb{N}$ adecuado. Denotamos por $\Omega_d := \mathbb{F}_q^{r_d}$ el conjunto de todas esas posibles elecciones aleatorias; consideramos a Ω_d con la probabilidad uniforme, el espacio muestral $\mathsf{F} \times \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$ con la probabilidad uniforme P definida en la Sección 8.4, y el espacio muestral $\mathsf{F} \times \mathbb{F}_q[\boldsymbol{X}]_{\leq d} \times \Omega_d$ con la probabilidad producto. A fin de analizar el costo del Algoritmo BBV, consideramos la variable aleatoria $X := X_{r,d} : \mathsf{F} \times \mathbb{F}_q[\boldsymbol{X}]_{\leq d} \times \Omega_d :\to \mathbb{N}_{\geq 0}$ que cuenta el número $X(\underline{\boldsymbol{a}}, F, \omega)$ de operaciones aritméticas en $\mathbb{F}_q$ que realiza el Algoritmo BBV sobre la entrada $F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$, cuando se consideran las bandas verticales determinadas por $\underline{\boldsymbol{a}}$ y la elección ω para los parámetros de la rutina que encuentra ceros en $\mathbb{F}_q$ de los elementos de $\mathbb{F}_q[T]_{\leq d}$.

Nuestro objetivo es determinar el comportamiento asintótico de la esperanza de la variable aleatoria X, es decir,

$$E[X] := \frac{1}{|\mathsf{F}||\mathcal{F}_{r,d}||\Omega_d|} \sum_{(\underline{\boldsymbol{a}},F,\omega)} X(\underline{\boldsymbol{a}}, F, \omega) \leq \frac{\tau(d,r,q)}{|\mathsf{F}||\mathcal{F}_{r,d}|} \sum_{F \in \mathcal{F}_{r,d}} \sum_{\underline{\boldsymbol{a}} \in \mathsf{F}} C(\underline{\boldsymbol{a}}, F).$$

Empezamos estudiando el caso $r > 2$, para el cual tenemos el siguiente resultado.

Teorema 8.4.4. *Sean $r > 2$ y $s^* := \binom{d/2+r-1}{r-1}$. Entonces la complejidad en promedio del Algoritmo BBV está acotado superiormente de la siguiente manera:*

$$E[X] \leq \tau(d,r,q)\big(\mu_d^{-1} + d(1-d^{-1})^{s^*}\big) + \mathcal{O}(q^{-1/2}), \tag{8.31}$$

donde $\tau(d,r,q)$ es el costo de la búsqueda en una banda vertical arbitraria.

Demostración. Recordemos que un elemento de $\mathbb{F}_q[\boldsymbol{X}]_{\leq d}$ se dice relativamente $\mathbb{F}_q$–irreducible si ninguno de sus factores irreducibles en $\mathbb{F}_q$ es absolutamente irreducible. Consideremos los conjuntos

$$A := \{F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d} : F \text{ es relativamente } \mathbb{F}_q\text{–irreducible}\}, \quad B := \mathbb{F}_q[\boldsymbol{X}]_{\leq d} \setminus A.$$

Tenemos que

$$\sum_{F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}} \sum_{\underline{\boldsymbol{a}} \in \mathsf{F}} C(\underline{\boldsymbol{a}}, F) = \sum_{F \in A} \sum_{\underline{\boldsymbol{a}} \in \mathsf{F}} C(\underline{\boldsymbol{a}}, F) + \sum_{F \in B} \sum_{\underline{\boldsymbol{a}} \in \mathsf{F}} C(\underline{\boldsymbol{a}}, F). \tag{8.32}$$

De [vzGVZ13, Corollary 6.7] se sigue que $|A|/|\mathbb{F}_q[\boldsymbol{X}]_{\leq d}| = \mathcal{O}\big(q^{\frac{-r(r-1)}{2}}\big)$. Por lo tanto,

$$\frac{1}{|\mathsf{F}||\mathbb{F}_q[\boldsymbol{X}]_{\leq d}|} \sum_{F \in A} \sum_{\underline{\boldsymbol{a}} \in \mathsf{F}} C(\underline{\boldsymbol{a}}, F) \leq \frac{q^{r-1}}{|\mathbb{F}_q[\boldsymbol{X}]_{\leq d}|} |A| = \mathcal{O}\big(q^{\frac{(r-1)(2-r)}{2}}\big) = \mathcal{O}(q^{-1}). \tag{8.33}$$

Ahora estudiamos el segundo término del lado derecho de (8.32). Tenemos

$$\frac{1}{|\mathsf{F}||\mathbb{F}_q[\boldsymbol{X}]_{\leq d}|} \sum_{F \in B} \sum_{\underline{\boldsymbol{a}} \in \mathsf{F}} C(\underline{\boldsymbol{a}}, F) = \frac{1}{|\mathbb{F}_q[\boldsymbol{X}]_{\leq d}|} \sum_{F \in B} \sum_{s=1}^{q^{r-1}} s \frac{|\{\underline{\boldsymbol{a}} \in \mathsf{F} : C(\underline{\boldsymbol{a}}, F) = s\}|}{|\mathsf{F}|}.$$

De las condiciones de consistencia del Lema 8.4.1 se sigue que

$$\begin{aligned}\frac{1}{|\mathsf{F}||\mathbb{F}_q[\boldsymbol{X}]_{\leq d}|} \sum_{F \in B} \sum_{\underline{\boldsymbol{a}} \in \mathsf{F}} C(\underline{\boldsymbol{a}}, F) &= \frac{|B|}{|\mathbb{F}_q[\boldsymbol{X}]_{\leq d}|} \sum_{s=1}^{q^{r-1}} s \frac{1}{|B|} \sum_{F \in B} \frac{|\{\underline{\boldsymbol{a}} \in \mathsf{F}_s : C(\underline{\boldsymbol{a}}, F) = s\}|}{|\mathsf{F}_s|} \\ &= \frac{|B|}{|\mathbb{F}_q[\boldsymbol{X}]_{\leq d}|} \sum_{s=1}^{q^{r-1}} s P_{\mathsf{F} \times B}[C = s],\end{aligned}$$

donde $P_{\mathsf{F} \times B}$ denota la probabilidad uniforme en $\mathsf{F} \times B$.

Para $s \leq s^*$, el Teorema 8.4.3 nos permite estimar la probabilidad del evento $[C = s]$. Por lo tanto, descomponemos la última suma de la siguiente manera:

$$\begin{aligned}\sum_{s=1}^{q^{r-1}} s P_{\mathsf{F} \times B}[C = s] &= \sum_{s=1}^{s^*} s P_{\mathsf{F} \times B}[C = s] + (s^* + 1) \sum_{s=s^*+1}^{q^{r-1}} P_{\mathsf{F} \times B}[C = s] \\ &\qquad + \sum_{s=s^*+2}^{q^{r-1}} (s - s^* - 1) P_{\mathsf{F} \times B}[C = s] \\ &= \sum_{s=1}^{s^*} s P_{\mathsf{F} \times B}[C = s] + (s^*+1) P_{\mathsf{F} \times B}[C \geq s^*+1] + \sum_{s=s^*+2}^{q^{r-1}} P_{\mathsf{F} \times B}[C \geq s].\end{aligned} \tag{8.34}$$

Primero estimamos la suma S_1 de los dos primeros términos del lado derecho de (8.34). Argumentando como en el Lema 8.4.2, vemos que

$$P_{\mathsf{F} \times B}[C = s] = \frac{1}{|\mathsf{F_s}|} \sum_{\underline{\boldsymbol{a}} \in \mathsf{F_s}} p_B[C_{\underline{\boldsymbol{a}}} = s].$$

Por el Teorema 8.4.3 y el Corolario 8.3.14 tenemos que

$$\begin{aligned} S_1 &= \sum_{s=1}^{s^*} s(\mu_d(1-\mu_d)^{s-1} + \mathcal{O}(q^{-1})) + (s^*+1)(1-\mu_d)^{s^*} + \mathcal{O}(q^{-1}) \\ &= \mu_d \sum_{s=1}^{s^*} s(1-\mu_d)^{s-1} + (s^*+1)(1-\mu_d)^{s^*} + \mathcal{O}(q^{-1}). \end{aligned}$$

Teniendo en cuenta que $\sum_{n\geq 1} nz^{n-1} = 1/(1-z)^2$ para todo $|z| \leq 1$, obtenemos

$$S_1 = \frac{1}{\mu_d} - \mu_d \sum_{s\geq s^*+1} s(1-\mu_d)^{s-1} + (s^*+1)(1-\mu_d)^{s^*} + \mathcal{O}(q^{-1}) = \frac{1}{\mu_d} + \mathcal{O}(q^{-1}), \tag{8.35}$$

donde la última igualdad se sigue de la identidad $\sum_{s\geq s^*+1} sz^{s-1} = z^{s^*}(s^*+1-zs^*)/(1-z)^2$, que vale para todo $|z| < 1$ (ver, por ejemplo, [GKP94, §2.3]).

Luego, estimamos la suma S_2 del lado derecho de (8.34). Observemos que

$$p_B[C_{\underline{\boldsymbol{a}}} \geq s] = p_B[F \in B : N_{1,d}(F(\boldsymbol{a}_i, X_r)) = 0 \ (1 \leq i \leq s-1)].$$

Así,

$$\begin{aligned} S_2 &\leq \frac{1}{|B|} \sum_{s=s^*+2}^{q^{r-1}} \frac{1}{|\mathsf{F_s}|} \sum_{(\underline{\boldsymbol{a}},\boldsymbol{a}_s)\in \mathsf{F}_{s-1}\times\mathbb{F}_q^{r-1}} |\{F \in B : N_{1,d}(F(\boldsymbol{a}_i, X_r)) = 0 \ (1 \leq i \leq s-1)\}| \\ &\leq \frac{q^{r-1}}{|B|} \sum_{s=s^*+2}^{q^{r-1}} \frac{1}{q^{r-1}-(s-1)} \sum_{\underline{\boldsymbol{a}}\in \mathsf{F}_{s-1}} \sum_{\substack{F\in B \\ N_{1,d}(F(\boldsymbol{a}_i,X_r))=0 \ (1\leq i\leq s-1)}} \frac{1}{|\mathsf{F_{s-1}}|} \\ &\leq \frac{q^{r-1}}{|B|} \sum_{s=s^*+2}^{q^{r-1}} \frac{1}{q^{r-1}-(s-1)} \sum_{F\in B} P_{\mathsf{F}_{s-1}}[N_{1,d}=0], \end{aligned}$$

donde $P_{\mathsf{F}_{s-1}}[N_{1,d}=0] := P_{\mathsf{F}_{s-1}}[\{\underline{\boldsymbol{a}} \in \mathsf{F}_{s-1} \ : N_{1,d}(F(\boldsymbol{a}_i, X_r)) = 0, \ 1 \leq i \leq s-1\}]$. Como $N_{1,d} = 0$ sigue una distribución hipergeométrica, podemos expresar la probabilidad $P_{\mathsf{F}_{s-1}}[N_{1,d}=0]$ de la siguiente manera (ver, por ejemplo, [Fel68, Chapter 6]):

$$P_{\mathsf{F}_{s-1}}[N_{1,d}=0] = \frac{\binom{q^{r-1}-NS(F)}{s-1}}{\binom{q^{r-1}}{s-1}}.$$

Deducimos que

$$S_2 \leq \frac{1}{|B|} \sum_{s=s^*+2}^{q^{r-1}} \sum_{F\in B} \left(1 - \frac{NS(F)-1}{q^{r-1}-1}\right)^{s-1}. \tag{8.36}$$

Fijamos $F \in B$. Luego, F tiene al menos un factor absolutamente irreducible definido sobre $\mathbb{F}_q$. Así, para $q > d^4$, del Teorema 2.2.5 se sigue que $NS(F) \geq \frac{q^{r-1}}{d}(1-$

α), donde $\alpha := d^2 q^{-1/2}$. Esto implica que

$$1 - \frac{NS(F)-1}{q^{r-1}-1} \leq 1 - \frac{1-\alpha}{d} + \mathcal{O}\big(q^{1-r}\big).$$

Combinando esta desigualdad con (8.36), concluimos que

$$\begin{aligned} S_2 &\leq \frac{1}{|B|} \sum_{s=s^*+2}^{q^{r-1}} \sum_{F\in B} \big(1-(1-\alpha)d^{-1} + \mathcal{O}(q^{1-r})\big)^{s-1} \\ &= \sum_{s=s^*+2}^{q^{r-1}} \big(1-(1-\alpha)d^{-1} + \mathcal{O}(q^{1-r})\big)^{s-1} \\ &= \frac{\big(1-(1-\alpha)d^{-1}\big)^{s^*+1}}{(1-\alpha)d^{-1}} + \mathcal{O}(q^{1-r}) = d(1-d^{-1})^{s^*+1} + \mathcal{O}(q^{-1/2}). \end{aligned}$$

Combinando (8.32), (8.33) y (8.35) con esta desigualdad, deducimos (8.31). □

Como $s^* > d^2/4$, el término $d(1-d^{-1})^{s^*+1}$ tiende a cero a medida que d y r crecen, y por lo tanto el lado derecho de (8.31) se comporta como ${\mu_d}^{-1}\tau(d,r,q)$. Esto indica que, en promedio, el Algoritmo BBV realiza a lo sumo ${\mu_d}^{-1} \approx 1{,}58\ldots$ búsquedas en bandas verticales hasta encontrar un cero $\mathbb{F}_q$–racional del polinomio de entrada. Esto mejora los resultados obtenidos en [vzGSS03] (para el caso de polinomios bivariados) y [CM06a] y [Mat10] (para el caso de polinomios en r variables), en donde se demuestra que con d búsquedas es posible encontrar un cero $\mathbb{F}_q$–racional en una banda vertical con probabilidad al menos $1/2$.

Resta analizar la complejidad en promedio $E[X]$ para el caso $r = 2$, es decir,

$$E[X] := \frac{1}{|\mathsf{F}||\mathbb{F}_q[X_1,X_2]_{\leq d}||\Omega_d|} \sum_{(\underline{\boldsymbol{a}},F,\omega)} X(\underline{\boldsymbol{a}},F,\omega) \leq \frac{\tau(d,r,q)}{|\mathsf{F}||\mathbb{F}_q[X_1,X_2]_{\leq d}|} \sum_{F\in\mathbb{F}_q[X_1,X_2]_{\leq d}} \sum_{\underline{\boldsymbol{a}}\in F} C(\underline{\boldsymbol{a}},F).$$

Para un número real $0 < \alpha < 1$ a determinar, consideramos los conjuntos

$$\begin{aligned} A &:= \{F \in \mathbb{F}_q[X_1,X_2]_{\leq d} : NS(F) \leq (1-\alpha)NS(2,d)\}, \\ B &:= \{F \in \mathbb{F}_q[X_1,X_2]_{\leq d} : NS(F) > (1-\alpha)NS(2,d)\}, \end{aligned}$$

donde $NS(F)$ es el número de bandas verticales donde F tiene un cero $\mathbb{F}_q$–racional, y $NS(2,d)$ es el número promedio de tales bandas verticales. Tenemos que

$$\sum_{F\in\mathbb{F}_q[X_1,X_2]_{\leq d}} \sum_{\underline{\boldsymbol{a}}\in\mathsf{F}} C(\underline{\boldsymbol{a}},F) = \sum_{F\in A} \sum_{\underline{\boldsymbol{a}}\in\mathsf{F}} C(\underline{\boldsymbol{a}},F) + \sum_{F\in B} \sum_{\underline{\boldsymbol{a}}\in\mathsf{F}} C(\underline{\boldsymbol{a}},F). \tag{8.37}$$

Para estimar el primer término del lado derecho de (8.37), comenzamos dando una estimación de $|A|$. Para esto, necesitamos dos resultados que demostramos en el capítulo siguiente, y que dan información sobre el comportamiento asintótico del promedio $NS(2,d)$ y la varianza $NS_2(2,d)$ de $NS(\cdot)$. Más precisamente, del Lema 9.1.1 y de la Proposición 9.1.2, que enunciamos más adelante, tenemos que

$NS(2,d) = \mu_d q + \mathcal{O}(1)$ y $NS_2(2,d) = ((d!)^{-2} + \mu_d(1-\mu_d))q + \mathcal{O}(1)$ respectivamente. Luego, de la desigualdad de Chebychev (ver Corolario 9.1.3 del capítulo siguiente) deducimos que

$$|A| \leq \left(\frac{1}{(\alpha\,\mu_d\,d!)^2} + \frac{1-\mu_d}{\alpha^2\mu_d}\right) q^{\dim \mathbb{F}_q[X_1,X_2]_{\leq d}-1} + \mathcal{O}(q^{\dim \mathbb{F}_q[X_1,X_2]_{\leq d}-2}).$$

Se sigue que

$$\frac{1}{|\mathsf{F}||\mathbb{F}_q[X_1,X_2]_{\leq d}|}\sum_{F\in A}\sum_{\underline{\boldsymbol{a}}\in\mathsf{F}} C(\underline{\boldsymbol{a}},F) \leq \frac{|A|q}{|\mathbb{F}_q[X_1,X_2]_{\leq d|}} \leq \left(\frac{1}{(\alpha\,\mu_d\,d!)^2}+\frac{1-\mu_d}{\alpha^2\mu_d}\right)+\mathcal{O}(q^{-1}). \tag{8.38}$$

Ahora estudiamos el segundo término del lado derecho de (8.37). Argumentando como en el caso $r > 2$, para $s^* := d/2 + 1$ tenemos que

$$\frac{1}{|\mathsf{F}||\mathbb{F}_q[X_1,X_2]_{\leq d}|}\sum_{F\in B}\sum_{\underline{\boldsymbol{a}}\in\mathsf{F}} C(\underline{\boldsymbol{a}},F) \leq \frac{1}{\mu_d} + \frac{1}{|B|}\sum_{s=s^*+2}^{q}\sum_{F\in B}\left(1-\frac{NS(F)-1}{q-1}\right)^{s-1} + \mathcal{O}(q^{-1}).$$

Fijamos $F \in B$. Por definición, $NS(F) > (1-\alpha)NS(2,d)$ y, del Lema 9.1.1 que damos en el siguiente capítulo, tenemos que $NS(2,d) = \mu_d\, q + \mathcal{O}(1)$. Así,

$$1 - \frac{NS(F)-1}{q-1} \leq 1 - (1-\alpha)\mu_d + \mathcal{O}(q^{-1}).$$

Por lo tanto,

$$\frac{1}{|B|}\sum_{s=s^*+2}^{q}\sum_{F\in B}\left(1-\frac{NS(F)-1}{q-1}\right)^{s-1} \leq \frac{(1-(1-\alpha)\mu_d)^{s^*+1}}{(1-\alpha)\mu_d} + \mathcal{O}(q^{-1}).$$

Combinando (8.37) con (8.38) y esta desigualdad, concluimos que

$$E[X] \leq \tau(d,r,q)\left(\frac{1}{\alpha^2}\left(\frac{1-\mu_d}{\mu_d}+\frac{1}{(d!)^2\mu_d^2}\right)+\frac{1}{\mu_d}+\big(1-(1-\alpha)\mu_d\big)^{s^*+1}\right)+\mathcal{O}(q^{-1}).$$

Fijando $\alpha^* := 1 - 1/\sqrt{s^*}$, obtenemos el siguiente resultado, que completa el análisis de la complejidad en promedio del Algoritmo BBV.

Teorema 8.4.5. *Sean $r := 2$, $s^* := d/2+1$ y $\alpha^* := 1-1/\sqrt{s^*}$. Tenemos la siguiente cota superior para la complejidad en promedio del Algoritmo BBV:*

$$E[X] \leq \tau(d,r,q)\left(\frac{1}{\alpha^{*2}}\left(\frac{1-\mu_d}{\mu_d}+\frac{1}{(d!)^2\mu_d^2}\right)+\frac{1}{\mu_d}+\left(1-\tfrac{\mu_d}{\sqrt{s^*}}\right)^{s^*+1}\right)+\mathcal{O}(q^{-1}), \tag{8.39}$$

donde $\tau(d,r,q)$ es el costo de la búsqueda en una banda vertical arbitraria.

A medida que d crece, la cantidad s^* tiende a infinito y la expresión del lado derecho de (8.39) tiende a $(2-\mu_d)/\mu_d \approx 2{,}16\ldots$. Esto es una cota superior para el número promedio de búsquedas en bandas verticales que deben realizarse en el caso $r = 2$.

8.5. Simulaciones sobre el número de bandas verticales

Terminamos este capítulo describiendo los resultados sobre la distribución del número de búsquedas que obtuvimos ejecutando una implementación del Algoritmo BBV en `Maple` sobre una muestra aleatoria de elementos de $\mathbb{F}_q[\boldsymbol{X}]_{\leq d}$, para valores de d, r y q dados. Recordamos que $C : \mathsf{F} \times \mathbb{F}_q[\boldsymbol{X}]_{\leq d} \mapsto \mathbb{N} \cup \{\infty\}$ es la variable aleatoria que cuenta el número de búsquedas que realiza el algoritmo sobre todas las posibles elecciones de bandas verticales. El Teorema 8.4.3 muestra que $P[C = s] \approx (1-\mu_d)^{s-1}\mu_d$. La experimentación numérica que realizamos está dirigida a mostrar que $(1-\mu_d)^{s-1}\mu_d$ se aproxima a la probabilidad $P[C = s]$.

Con este fin, dada una muestra aleatoria $\mathcal{S} \subset \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$ y un elemento $\underline{\boldsymbol{a}} \in \mathsf{F}_s$ usamos las siguientes notaciones:

$$p_{\underline{\boldsymbol{a}}} := p_{r,d}[\mathcal{S} \cap C_{\underline{\boldsymbol{a}},r,d} = s], \quad \widehat{p}_s := (1-\mu_d)^{s-1}\mu_d.$$

Tomamos una muestra de $N := 30$ elecciones aleatorias de $\underline{\boldsymbol{a}} \in \mathsf{F}_s$ y calculamos la media muestral

$$\overline{p}_s := \sum_{i=1}^{N} \frac{p_{\underline{\boldsymbol{a}}_i}}{N}.$$

Asimismo, consideramos el correspondiente error relativo:

$$\epsilon_s := \frac{|\overline{p}_s - \widehat{p}_s|}{\widehat{p}_s}.$$

Finalmente, comparamos el número promedio $\overline{N}^q_{r,d}$ de bandas verticales que deben ser consideradas hasta que el algoritmo encuentra un cero $\mathbb{F}_q$–racional con la cota superior teórica del Teorema 8.4.4, es decir, $1/\mu_d$.

Consideramos valores de s no tan grandes porque sino, la probabilidad $p_{\underline{\boldsymbol{a}}}$ es tan pequeña que nos es imposible compararla con el estimador teórico $\widehat{p}_s$.

La primera columna de las tablas a continuación corresponde al número de búsquedas que realiza el Algoritmo BBV hasta encontrar un cero $\mathbb{F}_q$–racional del polinomio en consideración; la columna siguiente contiene la correspondiente media muestral de las probabilidades $p_{\underline{\boldsymbol{a}}_i}$ que se obtienen tomando una muestra de 30 elecciones aleatorias de $\underline{\boldsymbol{a}}$ y polinomios aleatorios de grado a lo sumo d; en la tercera columna consignamos el estimador teórico $\widehat{p}_s$ y en la última columna incluimos el error relativo ϵ_s.

Ejemplos con $q := 67$ y $r := 2$. Consideramos una muestra aleatoria de polinomios bivariados con coeficientes en el cuerpo finito $\mathbb{F}_{67}$. En el Cuadro 8.1 consideramos una muestra aleatoria de 1000000 polinomios de $\mathbb{F}_{67}[X_1, X_2]$ de grado a lo sumo 30 y analizamos cuántas búsquedas realiza el Algoritmo BBV hasta que encuentra un cero $\mathbb{F}_q$–racional del polinomio de entrada. En este caso tenemos que $\widehat{p}_s := (1-\mu_{30})^{s-1}\mu_{30}$, donde $\mu_{30} := 0{,}6321205588\ldots$. El número promedio de búsquedas que realiza el algoritmo es $\overline{N}^{67}_{2,30} = 1{,}574924\ldots$, mientras que la cota teórica es $1/\mu_{30} = 1{,}581977\ldots$.

Cuadro 8.1: Ejemplo aleatorio con $q = 67$, $r = 2$ y $d = 30$.

s	$\overline{p}_s$	$\widehat{p}_s$	ϵ_s
1	0,635031	0,632121	0,004583
2	0,231664	0,232544	0,003799
3	0,084627	0,085548	0,010889
4	0,030921	0,031471	0,017789
5	0,011279	0,011578	0,026473
6	0,004101	0,004259	0,038575
7	0,001509	0,001567	0,038166
8	0,000553	0,000576	0,042349
9	0,000199	0,000212	0,067918
10	0,000076	0,000078	0,030513
11	0,000025	0,000029	0,161872
12	0,000010	0,000011	0,038441
13	0,000038	0,000003	0,022074
14	0,000011	0,000001	0,339501
15	0,000001	0,000001	0,051253

Consideramos también otro ejemplo de 1000000 polinomios de $\mathbb{F}_{67}[X_1, X_2]$ de grado a lo sumo $d := 5$. Por lo tanto, $\widehat{p}_s := (1-\mu_5)^{s-1}\mu_5$, donde $\mu_5 := 0{,}6333333\ldots$. Observemos que $\overline{N}_{2,5}^{67} = 1{,}572816\ldots$, mientras que la correspondiente cota teórica es $1/\mu_5 = 1{,}578947\ldots$. Los resultados de la simulación se encuentran en el Cuadro 8.2.

Cuadro 8.2: Muestra aleatoria con $q = 67$, $r = 2$ y $d = 5$.

s	$\overline{p}_s$	$\widehat{p}_s$	ϵ_s
1	0,635885	0,633333	0,004012
2	0,231459	0,232222	0,003298
3	0,084318	0,085148	0,009844
4	0,030727	0,031221	0,016085
5	0,011188	0,011448	0,023224
6	0,004091	0,004197	0,025996
7	0,001481	0,001539	0,039029
8	0,000543	0,000564	0,040109
9	0,000195	0,000207	0,056976
10	0,000069	0,000076	0,085938
11	0,000029	0,000028	0,030685
12	0,000009	0,000010	0,129198
13	0,000003	0,000003	0,133380
14	0,000002	0,000001	0,085740
15	0,000001	0,000001	0,057169

Ejemplos con $r = 3$ y $q := 11$, $q := 67$ y $q := 8$. Consideramos también dos ejemplos de 1000000 polinomios aleatorios de $\mathbb{F}_q[X_1, X_2, X_3]$. El primero contiene polinomios de grado a lo sumo $d := 5$ con coeficientes en $\mathbb{F}_{11}$, mientras que el segundo contiene polinomios de grado a lo sumo $d := 5$ con coeficientes en $\mathbb{F}_{67}$. Los números promedio de búsquedas en bandas verticales que realizó el algoritmo son $\overline{N}_{3,5}^{11} = 1{,}539646\ldots$ y $\overline{N}_{3,5}^{67} = 1{,}572975\ldots$; ambos deben ser comparados con $1/\mu_5 = 1{,}578947\ldots$. Los resultados de la experimentación numérica se encuentran en los Cuadros 8.3 y 8.4.

Cuadro 8.3: Muestra aleatoria con $q = 11$, $r = 3$ y $d = 5$.

s	$\overline{p}_s$	$\widehat{p}_s$	ϵ_s
1	0,649494	0,633333	0,024881
2	0,227637	0,232222	0,020145
3	0,079769	0,085148	0,067430
4	0,027999	0,031221	0,115075
5	0,009822	0,011448	0,165519
6	0,003419	0,004198	0,227683
7	0,001213	0,001539	0,269344
8	0,000421	0,000564	0,340555
9	0,000149	0,000207	0,382851
10	0,000050	0,000076	0,504379
11	0,000017	0,000028	0,662509
12	0,000002	0,000010	0,500062
13	0,000002	0,000004	0,726225
14	0,000001	0,000001	0,523767
15	0,000000	0,000001	2,017058

Cuadro 8.4: Muestra aleatoria con $q = 67$, $r = 3$ y $d = 5$.

s	$\overline{p}_s$	$\widehat{p}_s$	ϵ_s
1	0,635802	0,633333	0,003883
2	0,231571	0,232222	0,002810
3	0,084285	0,085148	0,010237
4	0,030732	0,031221	0,015898
5	0,011192	0,011447	0,022809
6	0,004081	0,004197	0,028645
7	0,001482	0,001539	0,038865
8	0,000541	0,000564	0,042865
9	0,000199	0,000207	0,039628
10	0,000071	0,000076	0,062618
11	0,000027	0,000028	0,017780
12	0,000010	0,000010	0,003320
13	0,000003	0,000004	0,078891
14	0,000001	0,000001	0,111938
15	0,000000	0,000001	0,257107

Por último, consideramos una muestra aleatoria de 100000 polinomios de grado a lo sumo $d := 3$ con coeficientes en un cuerpo finito no primo, esto es, $\mathbb{F}_8[X_1, X_2, X_3]$. Tenemos $\overline{N}^8_{3,3} = 1{,}504509\ldots$, a comparar con $1/\mu_3 = 1{,}5$.

Cuadro 8.5: Muestra aleatoria con $q = 8$, $r = 3$ y $d = 3$.

s	$\overline{p}_s$	$\widehat{p}_s$	ϵ_s
1	0,663161	0,666666	0,005259
2	0,222801	0,222222	0,002605
3	0,075617	0,074074	0,014151
4	0,025319	0,024691	0,020831
5	0,008725	0,008230	0,060146
6	0,002859	0,002743	0,042289
7	0,000808	0,000914	0,115974
8	0,000148	0,000304	0,513158

En los Cuadros $1 - 7$ observamos que, para los distintos valores de q, r, d y s considerados, las medias muestrales $\overline{p}_s$ se aproximan al estimador teórico $\widehat{p}_s$. Así, los ejemplos que consideramos confirman el comportamiento predicho por la estimación asintótica de $P[C = s]$ en el Teorema 8.4.3. Sin embargo, como el costo del Algoritmo BBV crece exponencialmente con el número r de variables de los polinomios considerados, para nuestros experimentos sólo consideramos los casos $r = 2$ y $r = 3$.

Capítulo 9

Distribución de las salidas del Algoritmo BBV

En este capítulo, continuando con el análisis del Algoritmo BBV del capítulo anterior, vamos a analizar la distribución de las salidas del mismo. Dado un polinomio de entrada, el algoritmo calcula un cero $\mathbb{F}_q$–racional, que queda determinado por ciertas elecciones aleatorias que se realizan durante la ejecución del Algoritmo BBV. Un punto importante, entonces, es que la respuesta del algoritmo no sea "sesgada", es decir, no haya ceros que aparezcan con más frecuencia que otros como salida del algoritmo.

Un algoritmo "ideal", desde el punto de vista de la "calidad" de la salida, es aquel para el cual todos los ceros $\mathbb{F}_q$–racionales del polinomio en consideración tienen la misma probabilidad de resultar la salida del algoritmo. Por esta razón, en [vzGSS03] la estrategia de búsqueda en bandas verticales para polinomios bivariados se modifica de manera que todos los ceros $\mathbb{F}_q$–racionales del polinomio de entrada tengan la misma probabilidad de ser calculados. Dicha modificación también puede aplicarse al Algoritmo BBV; sin embargo, como esta modificación implica una "ralentización" del algoritmo, nos proponemos un curso de acción diferente. En este capítulo vamos a analizar la distribución de las salidas utilizando un concepto de la teoría de la información: la *entropía de Shannon.* Si la salida para un entrada dada tiende a estar concentrada en pocos ceros $\mathbb{F}_q$–racionales, entonces la información que obtenemos es poca. Si, por el contrario, todos los ceros son salidas igualmente probables del algoritmo, la información que obtenemos es mayor, ya que no existen soluciones que no sean detectadas. El concepto de entropía de Shannon permite medir los estados intermedios entre estos dos extremos, es decir, se trata de una medida de cuán concentradas están las salidas del Algoritmo BBV para una entrada dada.

Dado $F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$, denotamos $Z(F) := \{\boldsymbol{x} \in \mathbb{F}_q^r : F(\boldsymbol{x}) = \boldsymbol{0}\}$ y $N(F) := |Z(F)|$. Siguiendo el trabajo de C. Beltrán y M. Pardo [BP11], definimos la *entropía de Shannon* H_F asociada a F y al Algoritmo BBV, como

$$H_F = \sum_{\boldsymbol{x} \in Z(F)} -P_{\boldsymbol{x},F} \log(P_{\boldsymbol{x},F}), \tag{9.1}$$

donde $P_{\boldsymbol{x},F}$ es la probabilidad puntual de que el algoritmo BBV obtenga como salida

a $\boldsymbol{x}$ para la entrada F y, a diferencia de los demás capítulos, en este capítulo, log denota el logaritmo natural.

Es sabido que $H_F \leq \log N(F)$, y vale la igualdad si y solo si $P_{\boldsymbol{x},F} = 1/N(F)$ para todo $\boldsymbol{x} \in Z(F)$. Así, es equivalente decir que la entropía H_F alcanza el valor máximo a decir que las salidas están equidistribuidas, esto es, todos los ceros $\mathbb{F}_q$–racionales del polinomio F de entrada tienen la misma probabilidad de ser obtenidos como salida.

En esta sección, vamos a analizar la entropía promedio del Algoritmo BBV cuando F recorre todos los elementos de $\mathbb{F}_q[\boldsymbol{X}]_{\leq d}$, es decir,

$$H := \frac{1}{|\mathbb{F}_q[\boldsymbol{X}]_{\leq d}|} \sum_{F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}} H_F. \tag{9.2}$$

Sea $F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$ y sea $\mathbf{A}^{\text{ideal}}$ un algoritmo "ideal" de búsqueda de ceros $\mathbb{F}_q$-racionales de F. Desde el punto de vista de la distribución de las salidas tenemos que la probabilidad $P^{\text{ideal}}_{\boldsymbol{x},F}$ de que $\boldsymbol{x} \in Z(F)$ resulte la salida es igual a $1/N(F)$. Así, de la definición (9.1), tenemos que la correspondiente entropía es

$$H^{\text{ideal}}_F := \sum_{\boldsymbol{x} \in Z(F)} -P^{\text{ideal}}_{\boldsymbol{x},F} \log(P^{\text{ideal}}_{\boldsymbol{x},F}) = \sum_{\boldsymbol{x} \in Z(F)} \frac{\log N(F)}{N(F)} = \log N(F).$$

Por la concavidad de la función $x \mapsto \log x$ y (2.3) se sigue el siguiente resultado.

Observación 9.0.1. *Si $\boldsymbol{A}^{\text{ideal}}$ es un algoritmo ideal de búsqueda de ceros $\mathbb{F}_q$–racionales de $F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$, entonces la entropía promedio para $\boldsymbol{A}^{\text{ideal}}$ es*

$$H^{\text{ideal}} := \frac{1}{|\mathbb{F}_q[\boldsymbol{X}]_{\leq d}|} \sum_{F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}} H^{\text{ideal}}_F \leq \log \left(\frac{\sum_{F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}} N(F)}{|\mathbb{F}_q[\boldsymbol{X}]_{\leq d}|} \right) = \log(q^{r-1}). \tag{9.3}$$

Vamos a exhibir una cota inferior para la entropía promedio H del Algoritmo BBV, que muestra que dicha estimación se acerca a la cota superior (9.3) de la entropía promedio del algoritmo "ideal". Esto indica que la entropía promedio del Algoritmo BBV será parecida a la del "algoritmo ideal", o sea, las salidas estarán "cerca" de resultar equidistribuidas.

9.1. Sobre el número de bandas verticales

Para obtener una cota inferior de la entropía promedio H del Algoritmo BBV, necesitamos analizar la distribución de probabilidades de la variable aleatoria $NS : \mathbb{F}_q[\boldsymbol{X}]_{\leq d} \to \mathbb{Z}_{\geq 0}$ que cuenta, para cada $F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$, el número de $\boldsymbol{a} \in \mathbb{F}_q^{r-1}$ para los cuales el polinomio $F(\boldsymbol{a}, X_r)$ tiene un cero $\mathbb{F}_q$–racional.

Recordemos que, para $F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$, denotamos por $VS(F)$ el conjunto de bandas verticales donde F tiene un cero $\mathbb{F}_q$–racional y por $NS(F)$ su cardinal, es decir,

$$VS(F) := \{\boldsymbol{a} \in \mathbb{F}_q^{r-1} : (\exists\, x_r \in \mathbb{F}_q)\ F(\boldsymbol{a}, x_r) = 0\}, \quad NS(F) := |VS(F)|.$$

En esta sección vamos a estudiar el promedio y la varianza de la variable aleatoria NS. Empezamos con el siguiente lema sobre el número promedio $NS(r,d)$ de bandas verticales en $\mathbb{F}_q[\boldsymbol{X}]_{\leq d}$, es decir,

$$NS(r,d) := \frac{1}{|\mathbb{F}_q[\boldsymbol{X}]_{\leq d}|} \sum_{F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}} NS(F).$$

Lema 9.1.1. *El número promedio $NS(r,d)$ satisface*

$$\begin{aligned} NS(r,d) &= \sum_{k=1}^{d} (-1)^{k-1} \binom{q}{k} q^{r-1-k} + (-1)^d \binom{q-1}{d} q^{r-d-2} \\ &= \mu_d \, q^{r-1} + \mathcal{O}(q^{r-2}). \end{aligned}$$

Demostración. De acuerdo a (8.2), tenemos que $NS(r,d) = q^{r-1} P[C=1]$. Por lo tanto, la primera afirmación se sigue inmediatamente del Teorema 8.2.1. La segunda afirmación se sigue del Corolario 8.2.2. □

En la siguiente proposición estimamos la varianza $NS_2(r,d)$ de la variable aleatoria NS, es decir,

$$\begin{aligned} NS_2(r,d) &:= \frac{1}{|\mathbb{F}_q[\boldsymbol{X}]_{\leq d}|} \sum_{F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}} \big(NS(F) - NS(r,d)\big)^2 \\ &= \frac{1}{|\mathbb{F}_q[\boldsymbol{X}]_{\leq d}|} \sum_{F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}} NS(F)^2 - NS(r,d)^2. \end{aligned}$$

Proposición 9.1.2. *La varianza $NS_2(r,d)$ satisface*

$$NS_2(r,d) = \frac{1}{(d!)^2} q^{2r-3} + \mu_d(1-\mu_d)\, q^{r-1} + \mathcal{O}(q^{2r-4}).$$

Demostración. Recordemos las notaciones $\mathsf{F}_2 := (\mathbb{F}_q^{r-1})^2 \setminus \{(\boldsymbol{a},\boldsymbol{a}) : \boldsymbol{a} \in \mathbb{F}_q^{r-1}\}$ y $N_2 := |\mathsf{F}_2|$. Dado un elemento $F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$ fijo, tenemos que

$$NS(F)^2 = \Big| \bigcup_{x,y \in \mathbb{F}_q} \{(\boldsymbol{a}_1,\boldsymbol{a}_2) \in (\mathbb{F}_q^{r-1})^2 : F(\boldsymbol{a}_1,x) = F(\boldsymbol{a}_2,y) = 0\} \Big|.$$

Por el principio de inclusión–exclusión resulta

$$\begin{aligned} \sum_{F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}} NS(F)^2 &= \sum_{F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}} \sum_{j=1}^{q} \sum_{k=1}^{q} (-1)^{j+k} \sum_{\mathcal{X}_j \subset \mathbb{F}_q} \sum_{\mathcal{Y}_k \subset \mathbb{F}_q} \mathcal{S}(\mathcal{X}_j, \mathcal{Y}_k) \\ &= \sum_{j=1}^{q} \sum_{k=1}^{q} (-1)^{j+k} \sum_{\mathcal{X}_j \subset \mathbb{F}_q} \sum_{\mathcal{Y}_k \subset \mathbb{F}_q} \sum_{F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}} \mathcal{S}(\mathcal{X}_j, \mathcal{Y}_k), \end{aligned}$$

donde $\mathcal{X}_j$ e $\mathcal{Y}_k$ recorren todos los subconjuntos de $\mathbb{F}_q$ de cardinal j y k respectivamente y, para subconjuntos cualesquiera $\mathcal{X} \subset \mathbb{F}_q$ e $\mathcal{Y} \subset \mathbb{F}_q$,

$$\mathcal{S}(\mathcal{X},\mathcal{Y}) := \big|\{(\boldsymbol{a}_1,\boldsymbol{a}_2) \in (\mathbb{F}_q^{r-1})^2 : (\forall x \in \mathcal{X})(\forall x \in \mathcal{Y})\, F(\boldsymbol{a}_1,x) = 0, F(\boldsymbol{a}_2,y) = 0\}\big|.$$

Para $\underline{\boldsymbol{a}} := (\boldsymbol{a}_1, \boldsymbol{a}_2) \in (\mathbb{F}_q^{r-1})^2$ y subconjuntos $\mathcal{X} \subset \mathbb{F}_q$ e $\mathcal{Y} \subset \mathbb{F}_q$, denotamos con

$$\mathcal{S}_{\underline{\boldsymbol{a}}}(\mathcal{X}, \mathcal{Y}) := \{F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d} : (\forall x \in \mathcal{X})(\forall x \in \mathcal{Y})\, F(\boldsymbol{a}_1, x) = 0, F(\boldsymbol{a}_2, y) = 0\}.$$

Se sigue que

$$\begin{aligned}\sum_{F \in \mathcal{F}_{r,d}} NS(F)^2 &= \sum_{j=1}^{q}\sum_{k=1}^{q}(-1)^{j+k} \sum_{\mathcal{X}_j \subset \mathbb{F}_q} \sum_{\mathcal{Y}_k \subset \mathbb{F}_q} \sum_{\underline{\boldsymbol{a}} \in (\mathbb{F}_q^{r-1})^2} |\mathcal{S}_{\underline{\boldsymbol{a}}}(\mathcal{X}_j, \mathcal{Y}_k)| \\ &= \sum_{\underline{\boldsymbol{a}} \in (\mathbb{F}_q^{r-1})^2} \sum_{j=1}^{q}\sum_{k=1}^{q}(-1)^{j+k} \sum_{\mathcal{X}_j \subset \mathbb{F}_q} \sum_{\mathcal{Y}_k \subset \mathbb{F}_q} |\mathcal{S}_{\underline{\boldsymbol{a}}}(\mathcal{X}_j, \mathcal{Y}_k)| =: \sum_{\underline{\boldsymbol{a}} \in (\mathbb{F}_q^{r-1})^2} N_{\underline{\boldsymbol{a}}},\end{aligned}$$

donde la definición $N_{\underline{\boldsymbol{a}}}$ coincide con la de (8.10).

Si $\underline{\boldsymbol{a}} \in \mathsf{F}_2$, entonces (8.2) y la afirmación en la demostración de la Proposición 8.2.5 aseguran que

$$\frac{N_{\underline{\boldsymbol{a}}}}{|\mathbb{F}_q[\boldsymbol{X}]_{\leq d}|} = \big(P[C = 1]\big)^2 + \frac{q-1}{q^{2d+2}}\binom{q-1}{d}^2. \tag{9.4}$$

Por otro lado, para $(\boldsymbol{a}, \boldsymbol{a}) \in (\mathbb{F}_q^{r-1})^2 \setminus \mathsf{F}_2$ tenemos que

$$\begin{aligned}N_{(\boldsymbol{a},\boldsymbol{a})} &:= \sum_{j,\,k=1}^{q} (-1)^{j+k} \sum_{\mathcal{X}_j, \mathcal{Y}_k \subset \mathbb{F}_q} |\mathcal{S}_{(\boldsymbol{a},\boldsymbol{a})}(\mathcal{X}_j, \mathcal{Y}_k)| \\ &= \sum_{l=1}^{q} \sum_{\mathcal{Z}_l \subset \mathbb{F}_q} \Bigg(\sum_{\substack{\mathcal{X}_j, \mathcal{Y}_k \subset \mathbb{F}_q \\ \mathcal{X}_j \cup \mathcal{Y}_k = \mathcal{Z}_l}} (-1)^{|\mathcal{X}_j| + |\mathcal{Y}_k|} \Bigg) |\mathcal{S}_{\boldsymbol{a}}(\mathcal{Z}_l)|,\end{aligned}$$

donde $\mathcal{Z}_l \subset \mathbb{F}_q$ recorre todos los subconjuntos de $\mathbb{F}_q$ de cardinal l y $\mathcal{S}_{\boldsymbol{a}}(\mathcal{Z}) := \{F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d} : (\forall z \in \mathcal{Z})\, F(\boldsymbol{a}, z) = 0\}$ para cualquier subconjunto $\mathcal{Z} \subset \mathbb{F}_q$.

Fijando $1 \leq l \leq q$ y un subconjunto $\mathcal{Z}_l \subset \mathbb{F}_q$ de cardinal l, hacemos la siguiente afirmación:

Afirmación.

$$\sum_{\substack{\mathcal{X}_j, \mathcal{Y}_k \subset \mathbb{F}_q \\ \mathcal{X}_j \cup \mathcal{Y}_k = \mathcal{Z}_l}} (-1)^{|\mathcal{X}_j| + |\mathcal{Y}_k|} = (-1)^{|\mathcal{Z}_l|+1}. \tag{9.5}$$

Demostración. Usando la igualdad $|\mathcal{X}_j| + |\mathcal{Y}_k| = |\mathcal{Z}_l| + |\mathcal{X}_j \cap \mathcal{Y}_k|$, basta probar que

$$\mathcal{S}_{j,k} := \sum_{\substack{\mathcal{X}_j, \mathcal{Y}_k \subset \mathbb{F}_q \\ \mathcal{X}_j \cup \mathcal{Y}_k = \mathcal{Z}_l}} (-1)^{|\mathcal{X}_j \cap \mathcal{Y}_k|} = -1. \tag{9.6}$$

Expresemos $\mathcal{S}_{j,k}$ en dos sumas, dependiendo de si $\mathcal{X}_j$ e $\mathcal{Y}_k$ son disjuntos o no. Más precisamente, escribamos

$$\mathcal{S}_{j,k} = \mathcal{S}_{j,k}^{\mathrm{D}} + \mathcal{S}_{j,k}^{\mathrm{ND}},$$

donde

$$\mathcal{S}^{\mathrm{D}}_{j,k} := \sum_{\substack{\mathcal{X}_j,\mathcal{Y}_k\subset\mathbb{F}_q\ \mathcal{X}_j\cap\mathcal{Y}_k=\emptyset\\ \mathcal{X}_j\cup\mathcal{Y}_k=\mathcal{Z}_l}} (-1)^{|\mathcal{X}_j\cap\mathcal{Y}_k|}, \quad \mathcal{S}^{\mathrm{ND}}_{j,k} := \sum_{\substack{\mathcal{X}_j,\mathcal{Y}_k\subset\mathbb{F}_q\ \mathcal{X}_j\cap\mathcal{Y}_k\neq\emptyset\\ \mathcal{X}_j\cup\mathcal{Y}_k=\mathcal{Z}_l}} (-1)^{|\mathcal{X}_j\cap\mathcal{Y}_k|}.$$

Por un cálculo de combinatoria elemental, de los 3^l-2 pares de subconjuntos no vacíos $\mathcal{X}_j$ e $\mathcal{Y}_k$ de $\mathcal{Z}_l$ tales que $\mathcal{X}_j\cup\mathcal{Y}_k=\mathcal{Z}_l$, hay exactamente 2^l-2 que son disjuntos. Por lo tanto, es fácil ver que

$$\mathcal{S}^{\mathrm{D}}_{j,k} = (2^l-2)(-1)^0 = 2^l-2. \tag{9.7}$$

Resta considerar la suma $\mathcal{S}^{\mathrm{ND}}_{j,k}$. Supongamos que $\mathcal{Z}_l := \{a_1,\dots,a_l\}$; podemos reescribir a $\mathcal{S}^{\mathrm{ND}}_{j,k}$ de la siguiente manera:

$$\mathcal{S}^{\mathrm{ND}}_{j,k} = \sum_{i_1=1}^{l} A_{i_1}(-1)^1 + \sum_{1\le i_1<i_2\le l} A_{i_1i_2}(-1)^2 + \cdots + \sum_{1\le i_1<\cdots<i_{l-1}\le l} A_{i_1\dots i_{l-1}}(-1)^{l-1} + (-1)^l, \tag{9.8}$$

donde $A_{i_1\dots i_m} := \big|\{\mathcal{X}_j\cap\mathcal{Y}_k:\ \ \mathcal{X}_j\cap\mathcal{Y}_k=\{a_{i_1},\dots,a_{i_m}\}\big|$ para $1\le i_1<\cdots<i_m\le l$.

Fijemos $1\le m\le l-1$. Observamos que $A_{i_1\dots i_m}=A_{j_1\dots j_m}$ para cada par de m–uplas ordenadas $(i_1,\dots,i_m)$ y $(j_1,\dots j_m)$. Por otro lado, hay $\binom{l}{m}$ subconjuntos no vacíos de $\mathcal{Z}_l$ formados por m elementos. Por lo tanto, a fin de obtener una expresión explícita para la suma $\sum_{1\le i_1<\cdots<i_m\le l} A_{i_1\dots i_m}(-1)^m$, basta calcular $\binom{l}{m}\cdot A_{1\dots m}$. Más precisamente, podemos reescribir la expresión de $\mathcal{S}^{\mathrm{ND}}_{j,k}$ de (9.8) de la siguiente manera:

$$\mathcal{S}^{\mathrm{ND}}_{j,k} = \binom{l}{1}A_1(-1)^1 + \binom{l}{2}A_{12}(-1)^2 + \dots\binom{l}{l-1}A_{12\dots l-1}(-1)^{l-1} + (-1)^l.$$

Notemos que el número $A_{1\dots m}$ coincide con la cantidad de pares de subconjuntos disjuntos de $\mathcal{Z}_l\setminus\{a_1,\dots,a_m\}$ tales que su unión es $\mathcal{Z}_l\setminus\{a_1,\dots,a_m\}$ para $1\le m\le l-1$. Dado que $|\mathcal{Z}_l\setminus\{a_1,\dots,a_m\}|=l-m$, vemos que hay exactamente 2^{l-m} pares de subconjuntos de $\mathcal{Z}_l$ con estas características, es decir, $A_{1\dots m}=2^{l-m}$ para cada $1\le m\le l-1$. Por lo tanto,

$$\mathcal{S}^{\mathrm{ND}}_{j,k} = \sum_{m=1}^{l}\binom{l}{m}2^{l-m}(-1)^m = 1-2^l. \tag{9.9}$$

Combinando (9.7) y (9.9) demostramos la afirmación (9.6), ya que $\mathcal{S}_{j,k}=2^l-2+1-2^l=-1$. □

De la afirmación (9.6) obtenemos que

$$N_{(\boldsymbol{a},\boldsymbol{a})} := \sum_{j=1}^{q}\sum_{k=1}^{q}(-1)^{j+k}\sum_{\mathcal{X}_j\subset\mathbb{F}_q}\sum_{\mathcal{Y}_k\subset\mathbb{F}_q}|\mathcal{S}_{(\boldsymbol{a},\boldsymbol{a})}(\mathcal{X}_j,\mathcal{Y}_k)| = \sum_{j=1}^{q}(-1)^{j-1}\sum_{\mathcal{X}_j\subset\mathbb{F}_q}|\mathcal{S}_{\boldsymbol{a}}(\mathcal{X}_j)|. \tag{9.10}$$

Por lo tanto, de (9.4), (9.10) y (8.3) tenemos que

$$\begin{aligned}\frac{1}{|\mathbb{F}_q[\boldsymbol{X}]_{\le d}|}\sum_{F\in\mathbb{F}_q[\boldsymbol{X}]_{\le d}} NS(F)^2 &= \sum_{\underline{\boldsymbol{a}}\in\mathsf{F}_2}\frac{N_{\underline{\boldsymbol{a}}}}{|\mathbb{F}_q[\boldsymbol{X}]_{\le d}|}+\frac{1}{|\mathbb{F}_q[\boldsymbol{X}]_{\le d}|}\sum_{\boldsymbol{a}\in\mathbb{F}_q^{r-1}}\sum_{j=1}^{q}(-1)^{j-1}\sum_{\mathcal{X}_j\subset\mathbb{F}_q}|\mathcal{S}_{\boldsymbol{a}}(\mathcal{X}_j)|\\ &= N_2\left(\left(q^{1-r}NS(r,d)\right)^2+\frac{q-1}{q^{2d+2}}\binom{q-1}{d}^2\right)+\frac{\sum_{F\in\mathbb{F}_q[\boldsymbol{X}]_{\le d}}NS(F)}{|\mathbb{F}_q[\boldsymbol{X}]_{\le d}|}\\ &= N_2\big(q^{1-r}NS(r,d)\big)^2+NS(r,d)+N_2\frac{q-1}{q^{2d+2}}\binom{q-1}{d}^2.\end{aligned}$$

Así vemos que

$$\begin{aligned}NS_2(r,d) &= \big((q^{2(r-1)}-q^{r-1})q^{2(1-r)}-1\big)NS(r,d)^2+NS(r,d)+N_2\frac{q-1}{q^{2d+2}}\binom{q-1}{d}^2\\ &= -q^{1-r}NS(r,d)^2+NS(r,d)+\frac{q^{2r-3}}{d!^2}+\mathcal{O}(q^{2d-2}).\end{aligned}$$

De esta igualdad y del Lema 9.1.1 se sigue fácilmente la afirmación de la proposición. □

El siguiente corolario es una consecuencia de la desigualdad de Chebyshev, y muestra una cota superior del cardinal del conjunto de elementos $F\in\mathbb{F}_q[\boldsymbol{X}]_{\le d}$ tales que el número de bandas verticales $NS(F)$ en las que F tiene un cero $\mathbb{F}_q$–racional difiere en cierta proporción del valor esperado $NS(r,d)$.

Corolario 9.1.3. *Para $0<\alpha<1$, el número $A(\alpha)$ de $F\in\mathbb{F}_q[\boldsymbol{X}]_{\le d}$ tales que $NS(F)\le(1-\alpha)NS(r,d)$ está acotado de la siguiente manera:*

$$A(\alpha)\le\frac{1}{(\alpha\,\mu_d\,d!)^2}q^{\dim\mathbb{F}_q[\boldsymbol{X}]_{\le d}-1}+\frac{1}{\alpha^2}\,\frac{1-\mu_d}{\mu_d}q^{\dim\mathbb{F}_q[\boldsymbol{X}]_{\le d}-r+1}+\mathcal{O}(q^{\dim\mathbb{F}_q[\boldsymbol{X}]_{\le d}-2}).$$

Demostración. Por el Lema 9.1.1 y la Proposición 9.1.2, la desigualdad de Chebyshev implica que

$$p_{r,d}\left(|NS(F)-NS(r,d)|\ge\alpha NS(r,d)\right)\le\frac{NS_2(r,d)}{\alpha^2NS(r,d)^2}.$$

Teniendo en cuenta que

$$\frac{NS_2(r,d)}{\alpha^2NS(r,d)^2}=\frac{1}{(\alpha\,\mu_d\,d!)^2}q^{-1}+\frac{1-\mu_d}{\alpha^2\mu_d}q^{1-r}+\mathcal{O}(q^{-2}),$$

obtenemos la afirmación del corolario.

□

9.2. Una cota inferior para la entropía

Con el objetivo de analizar la entropía promedio de Shannon definida en (9.2), vamos a determinar la probabilidad $P_{\boldsymbol{x},F}$ de que un elemento $\boldsymbol{x} := (\boldsymbol{a}, x) \in \mathbb{F}_q^r$ sea la salida del Algoritmo BBV, con $F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$ como entrada.

Dado una entrada $F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$ y la banda vertical definida por un elemento $\boldsymbol{a} \in \mathbb{F}_q^{r-1}$, el Algoritmo BBV busca un cero $\mathbb{F}_q$–racional del polinomio univariado $f := \gcd(F(\boldsymbol{a}, T), T^q - T)$. Si para realizar esta búsqueda utilizamos el algoritmo probabilístico de Cantor y Zassenhaus [CZ81], entonces todos los ceros $\mathbb{F}_q$–racionales de f tienen la misma probabilidad de salir. Este algoritmo divide al polinomio f en dos factores; uno de ellos es $\gcd(T^{(q-1)/2} - 1, f(T + b))$, para $b \in \mathbb{F}_q$ aleatorio. Si necesitamos conseguir un factor irreducible de f podemos aplicar recursivamente este proceso al factor de grado más pequeño. Así obtenemos todos los factores irreducibles de f aplicando el algoritmo de Cantor-Zassenhaus de manera recursiva. Por lo tanto, sin pérdida de generalidad podemos suponer que la búsqueda de raíces en $\mathbb{F}_q$ de los elementos de $f \in \mathbb{F}_q[T]_{\leq d}$ se realiza utilizando un algoritmo probabilístico del tipo de Cantor-Zassenhaus, de tal manera que todas las raíces tengan la misma probabilidad de salir.

Recordamos el modelo probabilístico de la Sección 8.4.2, que también vamos a utilizar en el análisis de la distribución de las salidas. Para $r_d \in \mathbb{N}$ adecuado, denotamos por $\Omega_d := \mathbb{F}_q^{r_d}$ al conjunto de todas las posibles elecciones aleatorias de elementos de $\mathbb{F}_q$ que realiza la rutina del Algoritmo BBV que busca ceros $\mathbb{F}_q$–racionales de los elementos de $\mathbb{F}_q[T]_{\leq d}$. Consideramos Ω_d con la probabilidad uniforme, el espacio muestral $\mathsf{F} \times \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$ con la probabilidad uniforme P, y el espacio muestral $\mathsf{F} \times \mathbb{F}_q[\boldsymbol{X}]_{\leq d} \times \Omega_d$ con la probabilidad producto $P \times P_{\Omega_d}$. Finalmente, consideramos la variable aleatoria $C_{\text{out}} : \mathsf{F} \times \mathbb{F}_q[\boldsymbol{X}]_{\leq d} \times \Omega_d \to \mathbb{F}_q^r \cup \{\emptyset\}$ definida de la siguiente manera: para un elemento $(\underline{\boldsymbol{a}}, F, \gamma) \in \mathsf{F} \times \mathbb{F}_q[\boldsymbol{X}]_{\leq d} \times \Omega_d$, si F tiene un cero $\mathbb{F}_q$–racional sobre cualquiera de las bandas verticales definidas por $\underline{\boldsymbol{a}}$, y $\boldsymbol{a}_j$ es la primera banda vertical con esa propiedad, entonces $C_{\text{out}}(\underline{\boldsymbol{a}}, F, \gamma) := (\boldsymbol{a}_j, x)$, donde $x \in \mathbb{F}_q$ es el cero del polinomio univariado $F(\boldsymbol{a}_j, T)$ que calcula la rutina correspondiente en el Algoritmo BBV determinado por la elección aleatoria γ. En caso contrario, definimos $C_{\text{out}}^{\text{var}}(\underline{\boldsymbol{a}}, F, \gamma) := \emptyset$. Por lo tanto, podemos expresar la probabilidad $P_{\boldsymbol{x},F}$ de que un elemento $\boldsymbol{x} := (\boldsymbol{a}, x) \in \mathbb{F}_q^r$ sea la salida del Algoritmo BBV para la entrada $F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$ como la probabilidad condicional $P \times P_{\Omega_d}\big[C_{\text{out}} = \boldsymbol{x} | F\big]$, es decir,

$$P_{\boldsymbol{x},F} := P \times P_{\Omega_d}\big[C_{\text{out}} = \boldsymbol{x} | F\big] := \frac{P \times P_{\Omega_d}\big[\{C_{\text{out}} = \boldsymbol{x}\} \cap (\mathsf{F} \times \{F\} \times \Omega_d)\big]}{P \times P_{\Omega_d}\big[\mathsf{F} \times \{F\} \times \Omega_d\big]}.$$

Con estas definiciones y observaciones podemos determinar la probabilidad $P_{\boldsymbol{x},F}$. Para este propósito, denotamos con $N_{\boldsymbol{a}}(F)$ al número de ceros $\mathbb{F}_q$–racionales de F en la banda vertical definida por $\boldsymbol{a}$, es decir,

$$N_{\boldsymbol{a}}(F) := |\{x \in \mathbb{F}_q : F(\boldsymbol{a}, x) = 0\}|.$$

En el siguiente lema determinamos la probabilidad $P_{\boldsymbol{x},F}$ de que un elemento arbitrario $\boldsymbol{x} \in \mathbb{F}_q^r$ resulte ser la salida del Algoritmo BBV para la entrada F.

Lema 9.2.1. *Sea $F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$ y $\boldsymbol{x} := (\boldsymbol{a}, x) \in Z(F)$. Entonces*

$$P_{\boldsymbol{x},F} = \frac{1}{NS(F)\, N_{\boldsymbol{a}}(F)}.$$

Demostración. Si $\boldsymbol{x}$ es la salida que se produce en el paso j, entonces el Algoritmo BBV elige elementos $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{j-1}$ para las primeras $j-1$ búsquedas tales que $N_{\boldsymbol{a}_k}(F) = 0$ para $k = 1, \ldots, j-1$, y el elemento $\boldsymbol{a}$ para la búsqueda j. Además, la rutina de búsqueda de raíces en $\mathbb{F}_q$ para polinomios univariados obtiene la raíz $x \in \mathbb{F}_q$ del polinomio $F(\boldsymbol{a}, T)$, lo cual ocurre, según nuestras observaciones y suposiciones anteriores, con probabilidad $1/N_{\boldsymbol{a}}(F)$.

Recordemos que los elementos $\boldsymbol{a}_j \in \mathbb{F}_q^{r-1}$ para la búsqueda j son elegidos aleatoriamente entre los elementos de $\mathbb{F}_q^{r-1} \setminus \{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{j-1}\}$ con equiprobabilidad. Asimismo, si $\boldsymbol{a}$ surge como la elección para el paso j, entonces el Algoritmo BBV debe elegir elementos $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{j-1} \in \mathbb{F}_q^{r-1} \setminus NS(F)$ distintos dos a a dos para las primeras $j-1$ búsquedas. Así, la probabilidad de dichas elecciones es

$$\begin{aligned} P(N_{\boldsymbol{a}_1}(F) = 0, \ldots, N_{\boldsymbol{a}_{j-1}}(F) = 0, \boldsymbol{a}_j = \boldsymbol{a} | F) &= \prod_{k=0}^{j-2} \left(1 - \frac{NS(F)}{q^{r-1} - k}\right) \cdot \frac{1}{q^{r-1} - j + 1} \\ &= \frac{1}{q^{r-1}} \frac{\binom{q^{r-1} - NS(F)}{j-1}}{\binom{q^{r-1}-1}{j-1}}. \end{aligned}$$

Como hay $q^{r-1} - NS(F)$ elementos $\boldsymbol{b} \in \mathbb{F}_q^{r-1}$ tales que $N_{\boldsymbol{b}}(F) = 0$, el Algoritmo BBV realiza a lo sumo $q^{r-1} - NS(F) + 1$ búsquedas. Así, tenemos que

$$\begin{aligned} P_{\boldsymbol{x},F} &= \sum_{j=1}^{q^{r-1} - NS(F) + 1} P(N_{\boldsymbol{a}_1}(F) = 0, \ldots, N_{\boldsymbol{a}_{j-1}}(F) = 0, \boldsymbol{a}_j = \boldsymbol{a} | F) \cdot \frac{1}{N_{\boldsymbol{a}}(F)} \\ &= \frac{1}{q^{r-1} N_{\boldsymbol{a}}(F)} \sum_{j=0}^{q^{r-1} - NS(F)} \frac{\binom{q^{r-1} - NS(F)}{j}}{\binom{q^{r-1}-1}{j}}. \end{aligned}$$

Utilizando, por ejemplo, [GKP94, §5.2, Problema 1], vemos que

$$\sum_{j=0}^{q^{r-1} - NS(F)} \frac{\binom{q^{r-1} - NS(F)}{j}}{\binom{q^{r-1}-1}{j}} = \frac{q^{r-1}}{NS(F)}.$$

Concluimos que

$$P_{\boldsymbol{x},F} = \frac{1}{q^{r-1} N_{\boldsymbol{a}}(F)} \frac{q^{r-1}}{NS(F)} = \frac{1}{NS(F)\, N_{\boldsymbol{a}}(F)},$$

lo que completa la demostración del lema. □

Para $F \in \mathbb{F}_q[\boldsymbol{X}]_{\leq d}$, consideramos la entropía

$$H_F = \sum_{(\boldsymbol{a},x) \in Z(F)} \frac{\log\big(NS(F)\, N_{\boldsymbol{a}}(F)\big)}{NS(F)\, N_{\boldsymbol{a}}(F)}. \tag{9.11}$$

Nuestro objetivo es determinar el comportamiento asintótico de la entropía promedio

$$H := \frac{1}{|\mathbb{F}_q[\boldsymbol{X}]_{\le d}|} \sum_{F\in\mathbb{F}_q[\boldsymbol{X}]_{\le d}} H_F = \frac{1}{|\mathbb{F}_q[\boldsymbol{X}]_{\le d}|} \sum_{F\in\mathbb{F}_q[\boldsymbol{X}]_{\le d}} \sum_{(\boldsymbol{a},x)\in Z(F)} \frac{\log\big(NS(F)\,N_{\boldsymbol{a}}(F)\big)}{NS(F)\,N_{\boldsymbol{a}}(F)}.$$

Observemos que

$$\begin{aligned}\sum_{F\in\mathbb{F}_q[\boldsymbol{X}]_{\le d}} \sum_{(\boldsymbol{a},x)\in Z(F)} 1 &= \sum_{(\boldsymbol{a},x)\in\mathbb{F}_q^r} |\{F\in\mathbb{F}_q[\boldsymbol{X}]_{\le d} : F(\boldsymbol{a},x)=0\}| \\ &= \sum_{F\in\mathbb{F}_q[\boldsymbol{X}]_{\le d}} N(F) = q^{r-1}|\mathbb{F}_q[\boldsymbol{X}]_{\le d}| = q^{\dim\mathbb{F}_q[\boldsymbol{X}]_{\le d}+r-1}. \qquad (9.12)\end{aligned}$$

Por otro lado, la función $f : (0,+\infty)\to\mathbb{R}$, $f(x) := \log x/x$ es decreciente en el intervalo $[e,+\infty)$ y es convexa en el intervalo $[e^{3/2},+\infty)$. Por el Corolario 9.1.3, la probabilidad del conjunto formado por los elementos $F\in\mathbb{F}_q[\boldsymbol{X}]_{\le d}$ que tienen hasta $e^{3/2}=4{,}48\ldots$ bandas verticales es $\mathcal{O}(q^{-1})$. Por lo tanto, tenemos que

$$\begin{aligned}H &= \frac{\sum_{F\in\mathbb{F}_q[\boldsymbol{X}]_{\le d}}\sum_{(\boldsymbol{a},x)\in Z(F)} 1}{|\mathbb{F}_q[\boldsymbol{X}]_{\le d}|} \frac{\sum_{F\in\mathbb{F}_q[\boldsymbol{X}]_{\le d}}\sum_{(\boldsymbol{a},x)\in Z(F)} \frac{\log(NS(F)\,N_{\boldsymbol{a}}(F))}{NS(F)\,N_{\boldsymbol{a}}(F)}}{\sum_{F\in\mathbb{F}_q[\boldsymbol{X}]_{\le d}}\sum_{(\boldsymbol{a},x)\in Z(F)} 1} \\ &\ge q^{r-1}\, f\left(\frac{\sum_{F\in\mathbb{F}_q[\boldsymbol{X}]_{\le d}}\sum_{(\boldsymbol{a},x)\in Z(F)} NS(F)\,N_{\boldsymbol{a}}(F)}{\sum_{F\in\mathbb{F}_q[\boldsymbol{X}]_{\le d}}\sum_{(\boldsymbol{a},x)\in Z(F)} 1}\right)(1+\mathcal{O}(q^{-1})). \qquad (9.13)\end{aligned}$$

Ahora analizamos el numerador

$$\mathcal{N} := \sum_{F\in\mathbb{F}_q[\boldsymbol{X}]_{\le d}} \sum_{(\boldsymbol{a},x)\in Z(F)} NS(F)\,N_{\boldsymbol{a}}(F)$$

del argumento de la función f en la última expresión.

Lema 9.2.2. *Tenemos que* $\mathcal{N} = 2\,\mu_d\, q^{2r-2+\dim\mathbb{F}_q[\boldsymbol{X}]_{\le d}}(1+\mathcal{O}(q^{-1}))$.

Demostración. Para $F\in\mathbb{F}_q[\boldsymbol{X}]_{\le d}$ y $\boldsymbol{a}\in VS(F)$, tenemos que

$$NS(F) = \left|\bigcup_{x\in\mathbb{F}_q}\{\boldsymbol{a}\in\mathbb{F}_q^{r-1} : F(\boldsymbol{a},x)=0\}\right|,\ N_{\boldsymbol{a}}(F) = |\{x\in\mathbb{F}_q : F(\boldsymbol{a},x)=0\}|.$$

En consecuencia,

$$\begin{aligned}\mathcal{N} &= \sum_{F\in\mathbb{F}_q[\boldsymbol{X}]_{\le d}} \sum_{\substack{(\boldsymbol{a},x)\in\mathbb{F}_q^r\\ F(\boldsymbol{a},x)=0}} \sum_{\substack{y\in\mathbb{F}_q\\ F(\boldsymbol{a},y)=0}} \left|\bigcup_{z\in\mathbb{F}_q}\{\boldsymbol{b}\in\mathbb{F}_q^{r-1} : F(\boldsymbol{b},z)=0\}\right| \\ &= \sum_{F\in\mathbb{F}_q[\boldsymbol{X}]_{\le d}} \sum_{\substack{(\boldsymbol{a},x)\in\mathbb{F}_q^r\\ F(\boldsymbol{a},x)=0}} \sum_{\substack{y\in\mathbb{F}_q\\ F(\boldsymbol{a},y)=0}} \sum_{k=1}^{q}(-1)^{k-1} \sum_{\substack{\mathcal{Z}_k\subset\mathbb{F}_q\\ |\mathcal{Z}_k|=k}} \big|\{\boldsymbol{b}\in\mathbb{F}_q^{r-1} : F(\boldsymbol{b},T)|_{\mathcal{Z}_k}\equiv 0\}\big| \\ &= \sum_{k=1}^{q}(-1)^{k-1} \sum_{\boldsymbol{a}\in\mathbb{F}_q^{r-1}} \sum_{x\in\mathbb{F}_q}\sum_{y\in\mathbb{F}_q} \sum_{\substack{\mathcal{Z}_k\subset\mathbb{F}_q\\ |\mathcal{Z}_k|=k}} \mathcal{N}_{\boldsymbol{a},x,y,\mathcal{Z}_k},\end{aligned}$$

donde

$$\begin{aligned}\mathcal{N}_{\boldsymbol{a},x,y,\mathcal{Z}_k} &:= \sum_{\substack{F\in\mathbb{F}_q[\boldsymbol{X}]_{\le d}\\ F(\boldsymbol{a},x)=F(\boldsymbol{a},y)=0}} \big|\{\boldsymbol{b}\in\mathbb{F}_q^{r-1} : F(\boldsymbol{b},T)|_{\mathcal{Z}_k}\equiv 0\}\big| \\ &= \sum_{\boldsymbol{b}\in\mathbb{F}_q^{r-1}} \big|\{F\in\mathbb{F}_q[\boldsymbol{X}]_{\le d} : F(\boldsymbol{a},x)=0, F(\boldsymbol{a},y)=0, F(\boldsymbol{b},T)|_{\mathcal{Z}_k}\equiv 0\}\big|.\end{aligned}$$

Supongamos que $k\le d$. Para el caso en que $\boldsymbol{b}\neq\boldsymbol{a}$ y $x\neq y$, tenemos que las igualdades $F(\boldsymbol{a},x)=0, F(\boldsymbol{a},y)=0, F(\boldsymbol{b},T)|_{\mathcal{Z}_k}\equiv 0$ forman un sistema de $k+2$ ecuaciones linealmente independientes cuyas incógnitas son los coeficientes de F en el espacio vectorial $\mathbb{F}_q[\boldsymbol{X}]_{\le d}$. Si en cambio $\boldsymbol{b}\neq\boldsymbol{a}$ y $x=y$, entonces las igualdades $F(\boldsymbol{a},x)=0, F(\boldsymbol{b},T)|_{\mathcal{Z}_k}\equiv 0$ forman un sistema de $k+1$ ecuaciones linealmente independientes. Finalmente, para el caso en que $\boldsymbol{b}=\boldsymbol{a}$, el número de condiciones linealmente independientes depende del cardinal de la intersección $\{x,y\}\cap\mathcal{Z}_k$. En efecto, si $x=y$ y $x\notin\mathcal{Z}_k$, entonces tenemos $k+1$ condiciones linealmente independientes; si, en cambio, $x\in\mathcal{Z}_k$, tenemos k condiciones linealmente independientes. Por otro lado, si $x\neq y$, y $\mathcal{Z}_k\cap\{x,y\}=\emptyset$, entonces tenemos $\min\{d+1,k+2\}$ condiciones linealmente independientes. En cambio, si $\mathcal{Z}_k\cap\{x,y\}\neq\emptyset$, tenemos $k+1$ o k condiciones linealmente independientes, dependiendo de si este cardinal es uno o dos. Resumiendo todos estos casos, vemos que

$$\mathcal{N}_{\boldsymbol{a},x,y,\mathcal{Z}_k} = (q^{r-1}-1)\,q^{\dim\mathbb{F}_q[\boldsymbol{X}]_{\le d}-k-|\{x,y\}|} + q^{\dim\mathbb{F}_q[\boldsymbol{X}]_{\le d}-\min\{d+1,|\{x,y\}\cup\mathcal{Z}_k|\}}.$$

Supongamos que $x\neq y$. Si $\mathcal{Z}_k\cap\{x,y\}=\emptyset$, entonces tenemos $\binom{q-2}{k}$ posibles elecciones del conjunto $\mathcal{Z}_k$. Si, en cambio, $|\mathcal{Z}_k\cap\{x,y\}|=1$, tenemos $2\binom{q-2}{k-1}$ posibles elecciones de dicho conjunto; por último, si $|\mathcal{Z}_k\cap\{x,y\}|=2$, tenemos $\binom{q-2}{k-2}$ posibles elecciones. Así, por estas observaciones y cálculos elementales, obtenemos

$$\begin{aligned}\sum_{\substack{\mathcal{Z}_k\subset\mathbb{F}_q\\ |\mathcal{Z}_k|=k}} \mathcal{N}_{\boldsymbol{a},x,y,\mathcal{Z}_k} &= \binom{q-2}{k}\Big((q^{r-1}-1)q^{\dim\mathbb{F}_q[\boldsymbol{X}]_{\le d}-(k+2)} + q^{\dim\mathbb{F}_q[\boldsymbol{X}]_{\le d}-\min\{d+1,k+2\}}\Big) \\ &\quad + 2\binom{q-2}{k-1}\Big((q^{r-1}-1)q^{\dim\mathbb{F}_q[\boldsymbol{X}]_{\le d}-(k+2)} + q^{\dim\mathbb{F}_q[\boldsymbol{X}]_{\le d}-(k+1)}\Big) \\ &\quad + \binom{q-2}{k-2}\Big((q^{r-1}-1)q^{\dim\mathbb{F}_q[\boldsymbol{X}]_{\le d}-(k+2)} + q^{\dim\mathbb{F}_q[\boldsymbol{X}]_{\le d}-k}\Big) \\ &= q^{\dim\mathbb{F}_q[\boldsymbol{X}]_{\le d}-k}(q^{r-1}-1)\binom{q}{k}\Big(\frac{1}{q^2} + \frac{(q-k)(q-k-1)q^{-3}+k(k-1)q^{-1}}{(q-1)(q^{r-1}-1)} \\ &\quad + \frac{2k(q-k)q^{-2}}{(q-1)(q^{r-1}-1)}\Big) \\ &= q^{\dim\mathbb{F}_q[\boldsymbol{X}]_{\le d}-k}(q^{r-1}-1)\binom{q}{k}\Big(\frac{1}{q^2}+\mathcal{O}(q^{-1-r})\Big). \qquad (9.14)\end{aligned}$$

De la misma manera, por cálculos elementales tenemos que, si $x=y$,

$$\begin{aligned}\sum_{\substack{\mathcal{Z}_k\subset\mathbb{F}_q\\|\mathcal{Z}_k|=k}}\mathcal{N}_{\boldsymbol{a},x,x,\mathcal{Z}_k} &= \binom{q-1}{k}\Big((q^{r-1}-1)q^{\dim\mathbb{F}_q[\boldsymbol{X}]_{\le d}-(k+1)}+q^{\dim\mathbb{F}_q[\boldsymbol{X}]_{\le d}-(k+1)}\Big)\\ &\quad+\binom{q-1}{k-1}\Big((q^{r-1}-1)q^{\dim\mathbb{F}_q[\boldsymbol{X}]_{\le d}-(k+1)}+q^{\dim\mathbb{F}_q[\boldsymbol{X}]_{\le d}-k}\Big)\\ &= q^{\dim\mathbb{F}_q[\boldsymbol{X}]_{\le d}-k}(q^{r-1}-1)\binom{q}{k}\left(\frac{1}{q}+\frac{(q-k)q^{-1}+k}{q(q^{r-1}-1)}\right)\\ &= (q^{r-1}-1)q^{\dim\mathbb{F}_q[\boldsymbol{X}]_{\le d}-k}\binom{q}{k}\Big(\frac{1}{q}+\mathcal{O}(q^{-r})\Big).\end{aligned}\tag{9.15}$$

Por lo tanto, de (9.14) y (9.15) concluimos que

$$\begin{aligned}\sum_{x\in\mathbb{F}_q}\sum_{y\in\mathbb{F}_q}\sum_{\substack{\mathcal{Z}_k\subset\mathbb{F}_q\\|\mathcal{Z}_k|=k}}\mathcal{N}_{\boldsymbol{a},x,y,\mathcal{Z}_k} &= (q^{r-1}-1)\binom{q}{k}q^{\dim\mathbb{F}_q[\boldsymbol{X}]_{\le d}-k}\left(\frac{q^2-q}{q^2}+\frac{q}{q}\right)(1+\mathcal{O}(q^{1-r}))\\ &= \frac{2q-1}{q}(q^{r-1}-1)\binom{q}{k}q^{\dim\mathbb{F}_q[\boldsymbol{X}]_{\le d}-k}(1+\mathcal{O}(q^{1-r})).\end{aligned}$$

Supongamos ahora que $k>d$. Entonces la condición $F(\boldsymbol{b},T)|_{\mathcal{Z}_k}\equiv 0$ es equivalente a que $F(\boldsymbol{b},T)=0$, es decir hay $d+1$ condiciones linealmente independiente sobre los coeficientes de F. Con argumentos similares a los del caso anterior, tenemos que

$$\sum_{x\in\mathbb{F}_q}\sum_{y\in\mathbb{F}_q}\sum_{\substack{\mathcal{Z}_k\subset\mathbb{F}_q\\|\mathcal{Z}_k|=k}}\mathcal{N}_{\boldsymbol{a},x,y,\mathcal{Z}_k}=\frac{2q-1}{q}(q^{r-1}-1)\binom{q}{k}q^{\dim\mathbb{F}_q[\boldsymbol{X}]_{\le d}-(d+1)}(1+\mathcal{O}(q^{1-r})).$$

Por estas dos últimas igualdades, el Teorema 8.2.1 y el hecho de que $P[C=1]=\mu_d+\mathcal{O}(q^{-1})$, tenemos que

$$\begin{aligned}\mathcal{N} =&2q^{2r-2+\dim\mathbb{F}_q[\boldsymbol{X}]_{\le d}}\frac{2q-1}{2q}(1-q^{1-r})\\ &\left(\sum_{k=1}^{d}(-1)^{k-1}\binom{q}{k}q^{-k}+\sum_{k=d+1}^{q}(-1)^{k-1}\binom{q}{k}q^{-d-1}\right)(1+\mathcal{O}(q^{1-r}))\\ =&2q^{2r-2+\dim\mathbb{F}_q[\boldsymbol{X}]_{\le d}}\frac{2q-1}{2q}(1-q^{1-r})P[C=1]\big(1+\mathcal{O}(q^{1-r})\big)\\ =&2\,\mu_d\,q^{2r-2+\dim\mathbb{F}_q[\boldsymbol{X}]_{\le d}}(1+\mathcal{O}(q^{-1})).\end{aligned}$$

Esto finaliza la demostración del teorema. □

Combinando (9.12) con (9.13) y el Lemma 9.2.2 obtenemos

$$H \geq q^{r-1} f\left(\frac{2\,\mu_d\, q^{2r-2+\dim \mathbb{F}_q[\boldsymbol{X}]_{\leq d}}(1+\mathcal{O}(q^{-1}))}{q^{r-1+\dim \mathbb{F}_q[\boldsymbol{X}]_{\leq d}}}\right)(1+\mathcal{O}(q^{1-r}))$$
$$= \frac{\log\left(2\,\mu_d q^{r-1}\big(1+\mathcal{O}(q^{-1})\big)\right)}{2\,\mu_d(1+\mathcal{O}(q^{-1}))}(1+\mathcal{O}(q^{-1})).$$

Por lo tanto, tenemos el siguiente resultado

Teorema 9.2.3. *Si H denota la entropía promedio del Algoritmo BBV, entonces*

$$H \geq \frac{1}{2\mu_d}\log(q^{r-1})(1+\mathcal{O}(q^{-1})).$$

Recordemos que, de acuerdo a (9.3), para un algoritmo en el cual las salidas están equidistribuidas, la entropía promedio H está acotada superiormente por $\log(q^{r-1})$. Observemos también que, para d grande,

$$\frac{1}{2\mu_d} \approx \frac{1}{2(1-e^{-1})} \approx 0{,}79.$$

Por lo tanto, del Teorema 9.2.3 deducimos el Algoritmo BBV, desde el punto de vista de la distribución de las salidas, es al menos 79 % tan bueno como un algoritmo "ideal".

Capítulo 10

Algoritmo de factorización de polinomios univariados en familias lineales

En los dos capítulos anteriores analizamos un algoritmo probabilístico de búsqueda en bandas verticales de puntos $\mathbb{F}_q$–racionales de hipersuperficies definidas sobre $\mathbb{F}_q$. En el mismo, una vez que se obtiene una banda vertical con raíces del polinomio multivariado que define la hipersuperficie de entrada, es necesario calcular una raíz en $\mathbb{F}_q$ del correspondiente polinomio univariado. Así, si queremos calcular ceros de polinomios univariados con coeficientes en $\mathbb{F}_q$, necesitamos factorizar parcialmente dichos polinomios. Existen diversos algoritmos eficientes de factorización de polinomios univariados sobre $\mathbb{F}_q$. En el trabajo [FGP01] se muestra que la distribución de los patrones de factorización del conjunto de polinomios de $\mathbb{F}_q[T]$ de grado dado determina el comportamiento en promedio del método de factorización más utilizado. Este algoritmo se basa en tres etapas fundamentales: eliminación de factores repetidos, factorización en distintos grados y factorización en igual grado, y se lo denomina el *algoritmo clásico de factorización*. Nos interesa analizar el comportamiento de este algoritmo de factorización aplicado a las familias lineales de polinomios que fueron objeto de estudio a lo largo de esta tesis. Por este motivo, en este capítulo, siguiendo las ideas de [FGP01] y utilizando las estimaciones explícitas sobre la distribución de patrones de factorización en familias lineales del Capítulo 7, analizamos el comportamiento en promedio del algoritmo clásico de factorización aplicado a estas familias.

10.1. El algoritmo clásico de factorización y preliminares

El algoritmo clásico de factorización funciona de la siguiente manera: dado un polinomio $f \in \mathbb{F}_q[T]_d$, encuentra la factorización completa $f = f_1^{e_1} \dots f_r^{e_r}$, donde $f_1, \dots, f_r \in \mathbb{F}_q[T]$ son polinomios mónicos, irreducibles y distintos dos a dos y $e_1, \dots e_r$ son números estrictamente positivos. Para ello, opera en tres pasos: reem-

plaza el polinomio de entrada por un polinomio libre de cuadrados que contiene todos los factores irreducibles del mismo pero con multiplicidad 1, descompone a este polinomio libre de cuadrados en polinomios cuyos factores irreducibles tienen todos el mismo grado y luego descompone completamente estos polinomios. A continuación describimos el algoritmo clásico de factorización.

Algoritmo clásico de factorización.

Entrada: Un polinomio mónico $f \in \mathbb{F}_q[T]_d$.

Salida: La factorización completa de f *en polinomios irreducibles sobre* $\mathbb{F}_q$.

procedimiento factor(f: $\boldsymbol{f \in \mathbb{F}_q[T]_d}$***).***

$a_f := ERF(f)$ *[*a_f *es libre de cuadrados]*

$\boldsymbol{b}_f := DDF(a_f)$ *[*$\boldsymbol{b}_f$ *es una factorización parcial en grados distintos]*

Sea $F := 1$

Para k *desde* 1 *hasta* s $(s \leq d)$ *hacer*

$F := F \cdot EDF(b_f[k], k)$ *[refinar la factorización en grados distintos para polinomios de grado* k*]*

Fin Para

$c := \text{factor}(f/a_f)$

Devolver $F \cdot c$.

En [FGP01] los autores analizan la complejidad en promedio del algoritmo clásico de factorización. Asintóticamente, este algoritmo no es el más rápido que existe hasta el momento (comparar con, por ejemplo, [Sho90], [Sho95], [vzGP01]). Sin embargo, es importante estudiarlo ya que es eficiente, completo, y aparece en muchos paquetes de álgebra computacional (ver [GCL92]). Asimismo, se puede obtener información importante del comportamiento del algoritmo en cada uno de sus pasos y del estado de la factorización del polinomio de entrada al finalizar cada etapa del algoritmo. Por ejemplo, se puede obtener la probabilidad de que el paso DDF complete la factorización de f. En efecto, en [FGP01, Theorem 6] se prueba que dicha probabilidad tiende a $e^{-\gamma} = 0{,}5614\ldots$, donde γ es la constante de Euler, cuando q es suficientemente grande. Asimismo, al finalizar la etapa DDF, podemos obtener la factorización en grados distintos de f, gracias a los sucesivos cálculos de $g_k := \gcd(T^{q^k} - T, f/g_{k-1})$ con $k = 1, 2, \ldots$, donde $g_1 := \gcd(T^q - T, f)$ da los factores irreducibles de f de grado 1 (usamos aquí el resultado que dice que, para todo cuerpo finito $\mathbb{F}_q$ y todo entero positivo d, el producto de todos los polinomios mónicos e irreducibles cuyo grado divide a d es igual a $T^{q^d} - T$).

En este capítulo vamos a analizar el comportamiento en promedio del algoritmo clásico de factorización cuando las entradas son elementos de la familia lineal $\mathcal{A} \subset \mathbb{F}_q[T]_d$ definida en la Sección 7.3 (o también en la Sección 3.1). Cabe observar que no podemos aplicar directamente los resultados de [FGP01] para este análisis, ya que

las entradas son elementos de una familia especial de polinomios de $\mathbb{F}_q[T]_d$, es decir, los coeficientes de los mismos cumplen relaciones lineales. Recordamos la definición de dicha familia: sean m y r enteros positivos tales que $q > d$ y $3 \leq r \leq d-m$, sean $A_{d-1}, \ldots, A_r$ indeterminadas sobre $\overline{\mathbb{F}}_q$ y sean $L_1, \ldots, L_m \in \mathbb{F}_q[A_{d-1}, \ldots, A_r]$ las formas lineales afines, linealmente independientes, definidas por

$$L_k := b_{k,d-1}A_{d-1} + \cdots + b_{k,r}A_r + b_{k,0} \ (1 \leq k \leq m). \tag{10.1}$$

Sea $\boldsymbol{L} := (L_1, \ldots, L_m)$. Supongamos que la matriz $M(\boldsymbol{L}) := (b_{k,d-j})_{1\leq k\leq m,\ 1\leq j\leq d-r}$ está escalonada por filas y denotamos con $1 \leq j_1 < \cdots < j_m \leq d-r$ a las posiciones de las columnas de $M(\boldsymbol{L})$ correspondientes a los pivotes. Consideramos la familia lineal $\mathcal{A} := \mathcal{A}_{\boldsymbol{L}} \subset \mathbb{F}_q[T]_d$ definida como

$$\mathcal{A} := \left\{T^d + a_{d-1}T^{d-1} + \cdots + a_0 \in \mathbb{F}_q[T]_d : \boldsymbol{L}(a_{d-1}, \ldots, a_r) = \boldsymbol{0}\right\}.$$

A fin de analizar la complejidad en promedio del algoritmo clásico de factorización aplicado a $\mathcal{A}$, como ya dijimos, seguimos las ideas de [FGP01], utilizando las estimaciones sobre la distribución de patrones de factorización en familias lineales del Capítulo 7.

Para este propósito, consideramos la probabilidad uniforme sobre $\mathcal{A}$ y tomamos como variable aleatoria la función $\mathcal{X} : \mathcal{A} \to \mathbb{N}$ que cuenta para cada elemento $f \in \mathcal{A}$ la cantidad $\mathcal{X}(f)$ de operaciones aritméticas en $\mathbb{F}_q$ que el algoritmo clásico de factorización necesita realizar hasta lograr la factorización completa de f en $\mathbb{F}_q$. Como este algoritmo consta de cuatro etapas de factorización, podemos descomponer a la variable aleatoria $\mathcal{X}$ como la suma de las variables aleatorias que cuentan el costo de cada paso del algoritmo. Más precisamente, sea $\mathcal{X}_1 : \mathcal{A} \to \mathbb{N}$ la variable aleatoria que cuenta el número de operaciones aritméticas en $\mathbb{F}_q$ del paso ERF, es decir,

$$\mathcal{X}_1(f) := \text{Costo}(\text{ERF}(f)), \text{ con } f \in \mathcal{A}. \tag{10.2}$$

Notemos con $a_f := ERF(f)$ el polinomio libre de cuadrados que se obtiene luego de aplicar el paso ERF al polinomio de entrada f. A su vez, $\mathcal{X}_2 : \mathcal{A} \to \mathbb{N}$ es la variable aleatoria que cuenta la cantidad de operaciones aritméticas en $\mathbb{F}_q$ del paso DDF, es decir,

$$\mathcal{X}_2(f) := \text{Costo}(\text{DDF}(a_f)), \text{ con } f \in \mathcal{A}. \tag{10.3}$$

Notemos con $\boldsymbol{b}_f := \text{DDF}(a_f) = (b_f(1), \ldots, b_f(s))$ al vector de polinomios que se obtiene aplicando el paso DDF al polinomio mónico libre de cuadrados $a_f := ERF(f)$ (aquí s es el grado del factor irreducible más grande de a_f). Cada $b_f(k)$ denota el producto de los polinomios mónicos irreducibles de grado k que dividen a f, contados con multiplicidad. Sea $\mathcal{X}_3 : \mathcal{A} \to \mathbb{N}$ la variable aleatoria que cuenta la cantidad de operaciones aritméticas del paso EDF, es decir,

$$\mathcal{X}_3(f) := \sum_{k=1}^{d} \mathcal{X}_{3,k}(f), \tag{10.4}$$

donde, para cada k, la variable aleatoria $\mathcal{X}_{3,k}(f)$ se define como $\mathcal{X}_{3,k}(f)$:=Costo$(EDF(b_f(k)))$ y $b_f(k)$ es la k–ésima coordenada del vector $\boldsymbol{b}_f$. Notemos que EDF aplicado a $b_f(k)$

factoriza dicho polinomio en factores irreducibles de grado k. Por último, denotamos por $\mathcal{X}_4 : \mathcal{A} \to \mathbb{N}$ a la variable aleatoria que cuenta la cantidad de operaciones aritméticas en $\mathbb{F}_q$ del algoritmo clásico de factorización aplicado al polinomio $f/ERF(f)$.

En este capítulo vamos a estudiar la esperanza de la variable aleatoria $\mathcal{X}$, es decir, el número

$$E[\mathcal{X}] := \frac{1}{|\mathcal{A}|} \sum_{f \in \mathcal{A}} \mathcal{X}(f) = \frac{1}{|\mathcal{A}|} \sum_{k=1}^{4} \sum_{f \in \mathcal{A}} \mathcal{X}_k(f). \tag{10.5}$$

A lo largo de este capítulo, como en el Capítulo 8, vamos a suponer que la multiplicación de polinomios se realiza con el algoritmo de multiplicación rápida (ver [vzGG99, Chapter 8]). Por lo tanto, la multiplicación y la división con resto de dos polinomios de grado a lo sumo d se pueden efectuar usando a lo sumo $\tau_1 M(d)$ operaciones aritméticas en $\mathbb{F}_q$, y el cálculo del máximo común divisor entre dos polinomios de grado a lo sumo d se puede efectuar con a lo sumo $\tau_2 \mathcal{U}(d)$ operaciones aritméticas en $\mathbb{F}_q$, donde τ_1 y τ_2 son constantes (ver [vzGG99, §8.4 and 11.1]). Recordemos que $\mathcal{U}(d) := M(d) \log d$ y $M(d) := d \log d \log \log d$.

10.2. Eliminación de factores repetidos

Como hemos dicho, el primer paso del algoritmo clásico de factorización es la eliminación de factores repetidos. Sea $f := f_1^{e_1} \dots f_r^{e_r} = \prod_{p|e_i} f_i^{e_i} \prod_{p \nmid e_i} f_i^{e_i}$ la factorización en polinomios mónicos irreducibles en $\mathbb{F}_q[T]$ de un elemento $f \in \mathcal{A}$, donde $f_1, \dots, f_r$ son distintos dos a dos, $e_1, \dots, e_r \in \mathbb{N}$, y p es la característica de $\mathbb{F}_q$. Sabemos que f es libre de cuadrados si y solo si el máximo común divisor $\gcd(f, f')$ entre f y su derivada es igual a 1 (ver [vzGG99, Corollary 14.25]). Supongamos entonces que f no es libre de cuadrados. Así tenemos que $u := \gcd(f, f') \neq 1$. Si $\gcd(f, f') = f$ (o sea, $f' = 0$), entonces $f = g^p$ para algún polinomio $g \in \mathbb{F}_q[T]$. En consecuencia, $v := f/u = \prod_{p \nmid e_i} f_i$ es la parte libre de cuadrados del producto $\prod_{p \nmid e_i} f_i^{e_i}$ (ver [Sho05, Theorem 20.4]). Como cada e_i es menor o igual que $d := \deg(f)$, tenemos que $\gcd(u, v^d) = \prod_{p \nmid e_i} f_i^{e_i - 1}$. Así, $w := \frac{u}{\gcd(u, v^d)} = \prod_{p|e_i} f_i^{e_i}$ es la parte de f que es una potencia de p.

A continuación describimos el algoritmo de eliminación de factores repetidos.

Algoritmo ERF.

Entrada: $f \in \mathcal{A}$.

Salida: Parte libre de cuadrados de f, o sea, el producto de todos los factores irreducibles distintos de f en $\mathbb{F}_q[T]$.

procedimiento ERF(f: polinomio)

Calcular $u := gcd(f, f')$

Calcular $v := \frac{f}{u}$ *[parte libre de cuadrados de* $\prod_{p \nmid e_i} f_i^{e_i}$*]*
Calcular $w := \frac{u}{\gcd(u, v^d)}$ *[parte de* f *que es potencia de* p*]*
Devolver $v \cdot ERF(w^{1/p})$.

En [vzGG99, Exercise 14.27] los autores demuestran que, para un polinomio $f \in \mathbb{F}_q[T]_d$, la cantidad de operaciones aritméticas en $\mathbb{F}_q$ que el algoritmo EDF debe realizar hasta obtener la parte libre de cuadrados de f es $\mathcal{O}(M(d) \log d + d \log(q/p))$.

En esta sección analizamos la complejidad en promedio del algoritmo ERF aplicado a elementos de $\mathcal{A}$. Más precisamente, consideramos la variable aleatoria $\mathcal{X}_1$ definida en (10.2) y damos una estimación para la esperanza $E[\mathcal{X}_1]$ de dicha variable aleatoria sobre la familia de polinomios $\mathcal{A}$, es decir,

$$E[\mathcal{X}_1] := \frac{1}{|\mathcal{A}|} \sum_{f \in \mathcal{A}} \mathcal{X}_1(f). \tag{10.6}$$

Observemos primeramente que, si $q \geq 2d^2$, la probabilidad de que un polinomio aleatorio de $\mathcal{A}$ sea libre de cuadrados es al menos $1/2$. En efecto, consideremos el conjunto $\mathcal{A}^{sq}$ de todos los polinomios $f \in \mathcal{A}$ que son libres de cuadrados y $\mathcal{A}^{nsq} := \mathcal{A} \setminus \mathcal{A}^{sq}$. La probabilidad de que un polinomio aleatorio de $\mathcal{A}$ sea libre de cuadrados es

$$P[\mathcal{A}^{sq}] = \frac{|\mathcal{A}^{sq}|}{|\mathcal{A}|} = 1 - \frac{|\mathcal{A}^{nsq}|}{|\mathcal{A}|}.$$

En (7.18) dimos la siguiente cota superior del cardinal de $\mathcal{A}^{nsq}$:

$$|\mathcal{A}^{nsq}| \leq d(d-1)q^{d-m-1}. \tag{10.7}$$

Utilizando (10.7) y la condición sobre q, tenemos que

$$P[\mathcal{A}^{sq}] \geq 1 - \frac{d^2 q^{d-m-1}}{q^{d-m}} = 1 - \frac{d^2}{q} \geq \frac{1}{2}.$$

En consecuencia, obtenemos el siguiente resultado.

Teorema 10.2.1. *Para* $q \geq 2d^2$*, la probabilidad de que un polinomio aleatorio de* $\mathcal{A}$ *sea libre de cuadrados es mayor o igual que* $1 - d^2/q \geq 1/2$.

Observamos que un polinomio aleatorio de $\mathbb{F}_q[T]_d$ con $d \geq 2$ tiene probabilidad $1 - 1/q$ de ser libre de cuadrados (ver, por ejemplo, [FS09, Theorem 2.1]).

Para estimar la esperanza (10.6) descomponemos a la familia lineal $\mathcal{A}$ en los subconjuntos disjuntos $\mathcal{A}^{sq}$ y $\mathcal{A}^{nsq}$. Así, tenemos que

$$E[\mathcal{X}_1] := \frac{1}{|\mathcal{A}|} \sum_{f \in \mathcal{A}} \mathcal{X}_1(f) = \frac{1}{|\mathcal{A}|} \sum_{f \in \mathcal{A}^{sq}} \mathcal{X}_1(f) + \frac{1}{|\mathcal{A}|} \sum_{f \in \mathcal{A}^{nsq}} \mathcal{X}_1(f). \tag{10.8}$$

El Teorema 10.2.1 sugiere que la esperanza $E[\mathcal{X}_1]$ estará dominada por la primera suma del lado derecho de esta última expresión.

Empezamos dando una cota superior para la primera suma S_1^{sq} del lado derecho de la suma de (10.8). Observamos que el algoritmo ERF, en este caso, realiza los primeros tres pasos. Como $u := \gcd(f, f') = 1$ y $\gcd(u, v^d) = 1$, el costo del algoritmo en este caso está dominado por el costo de calcular u, que es de a lo sumo $\tau_2\, \mathcal{U}(d)$ operaciones aritméticas en $\mathbb{F}_q$, y el costo de calcular v^d, que es de a lo sumo $\tau_1\, \mathcal{U}(d)$ operaciones aritméticas en $\mathbb{F}_q$. Concluimos que, si $f \in \mathcal{A}^{sq}$, entonces

$$\mathcal{X}_1(f) \leq (\tau_1 + \tau_2)\, \mathcal{U}(d).$$

Así, tenemos que

$$S_1^{sq} := \frac{1}{|\mathcal{A}|} \sum_{f \in \mathcal{A}^{sq}} \mathcal{X}_1(f) \leq (\tau_1 + \tau_2)\, \mathcal{U}(d) \frac{|\mathcal{A}^{sq}|}{|\mathcal{A}|}. \tag{10.9}$$

Por otro lado, damos una cota superior para la segunda suma S_1^{nsq} de (10.8). Sea $f \in \mathcal{A}^{nsq}$. En [vzGG99, Exercise 14.27] los autores muestran que la cantidad de operaciones aritméticas en $\mathbb{F}_q$ que el algoritmo necesita hasta encontrar la parte libre de cuadrados de f está acotada por $\mathcal{X}_1(f) \leq c_1\big(\mathcal{U}(d) + d \log\left(\frac{q}{p}\right)\big)$, donde c_1 es una constante independiente de q. Así, tenemos que

$$S_1^{nsq} := \frac{1}{|\mathcal{A}|} \sum_{f \in \mathcal{A}^{nsq}} \mathcal{X}_1(f) \leq c_1\big(\mathcal{U}(d) + d \log\left(\frac{q}{p}\right)\big) \frac{|\mathcal{A}^{nsq}|}{|\mathcal{A}|}. \tag{10.10}$$

Combinando (10.9), (10.10) y (10.7) con (10.8) concluimos que

$$\begin{aligned} E[\mathcal{X}_1] leq & (\tau_1 + \tau_2)\, \mathcal{U}(d) \frac{|\mathcal{A}^{sq}|}{|\mathcal{A}|} + c_1\, \mathcal{U}(d) \frac{|\mathcal{A}^{nsq}|}{|\mathcal{A}|} + c_1 d \log\left(\frac{q}{p}\right) \frac{|\mathcal{A}^{nsq}|}{|\mathcal{A}|} \\ &\leq c_2\, \mathcal{U}(d) + c_1 d \log\left(\frac{q}{p}\right) \frac{|\mathcal{A}^{nsq}|}{|\mathcal{A}|} \\ &\leq c_2\, \mathcal{U}(d) + c_1 d^3 \log\left(\frac{q}{p}\right) \frac{1}{q}, \end{aligned}$$

donde $c_2 := \max\{\tau_1 + \tau_2, c_1\}$. Obtenemos así el siguiente resultado.

Teorema 10.2.2. *El costo en promedio $E[\mathcal{X}_1]$ del algoritmo ERF aplicado a los elementos de $\mathcal{A}$ está acotado superiormente por $E[\mathcal{X}_1] \leq c_2\, \mathcal{U}(d) + c_1 \log\left(\frac{q}{p}\right) \frac{d^3}{q}$, donde $\mathcal{X}_1$ es la variable aleatoria definida en* (10.2) *y c_1 y c_2 son constantes independientes de d y q.*

Observamos que el costo en promedio del algoritmo ERF aplicado a los elementos de $\mathcal{A}$ es asintóticamente del orden de $M(d) \log d$, que corresponde al costo de calcular el máximo común divisor $u := \gcd(f, f')$. Esto generaliza los resultados de [FGP01, Section 2],

10.3. Factorización en distintos grados

El segundo paso de la factorización clásica es la factorización en distintos grados, esto es, se tata de descomponer un polinomio libre de cuadrados en polinomios cuyos factores irreducibles tienen el mismo grado. Esto significa expresar al polinomio $a_f := ERF(f)$ en la forma $b(1)\cdots b(s)$, donde $b(k)$ es el producto de todos los factores irreducibles de grado k que aparecen en a_f.

A continuación describimos el algoritmo de factorización en distintos grados (algoritmo DDF). Este algoritmo utiliza la siguiente propiedad (ver, por ejemplo, [LN83, Theorem 3.20]): Para $k \geq 1$, el polinomio $T^{q^k} - T \in \mathbb{F}_q[T]$ es el producto de todos los polinomios mónicos irreducibles en $\mathbb{F}_q[T]$ cuyo grado divide a k. Así, si calculamos el máximo común divisor $g_1 := \gcd(T^q - T, f)$ obtenemos todos los factores irreducibles en $\mathbb{F}_q[T]$ de f de grado 1. Si removemos todos los factores lineales de f y calculamos $g_2 := \gcd(T^{q^2} - T, f/g_1)$, obtenemos todos los factores irreducibles en $\mathbb{F}_q[T]$ de f de grado 2. De esta manera, calculando el máximo común divisor $g_k := \gcd(T^{q^k} - T, f/g_{k-1})$ obtenemos todos los factores irreducibles en $\mathbb{F}_q[T]$ de f de grado k, para $1 \leq k \leq d$.

Algoritmo DDF.

Entrada: un polinomio libre de cuadrados a de grado d.

Salida: El vector de polinomios $\boldsymbol{b} := (b(1), \ldots, b(s))$, donde cada $b(k)$ es el producto de todos los factores irreducibles de a en $\mathbb{F}_q[T]$ de grado k.

Sean $g := a, \quad h := T$

Mientras $g \neq 1$ *hacer*

Calcular $h := h^q \mod g$

Calcular $b(k) := \gcd(h - T, g)$ *[producto de factores irreducibles de grado k en común con g]*

Calcular $g := \frac{g}{b(k)}$ *[a sin factores de grado menor o igual que k]*

$k := k + 1$

Fin Mientras

Devolver $\boldsymbol{b}$.

El vector $(b(1), \ldots, b(s))$ que se obtiene como salida del algoritmo DDF se denomina la *descomposición en grados distintos* de a.

En [vzGG99, Algorithm 14.3] los autores prueban que el algoritmo DDF realiza $\mathcal{O}(sM(d)\log(dq))$ operaciones aritméticas en $\mathbb{F}_q$, donde s es el máximo de los factores irreducibles del polinomio a de entrada. En esta sección analizamos la complejidad en promedio del algoritmo DDF aplicado a los elementos de la familia lineal $\mathcal{A}$. Más

precisamente, consideramos la variable aleatoria $\mathcal{X}_2$ definida en (10.3) y damos una estimación de la esperanza $E[\mathcal{X}_2]$, es decir,

$$E[\mathcal{X}_2] := \frac{1}{|\mathcal{A}|} \sum_{f \in \mathcal{A}} \mathcal{X}_2(f).$$

Descomponiendo el conjunto de entradas $\mathcal{A}$ como antes, en los subconjuntos disjuntos $\mathcal{A}^{sq}$ (elementos de $\mathcal{A}$ que son libres de cuadrados) y $\mathcal{A}^{nsq} := \mathcal{A} \setminus \mathcal{A}^{sq}$, tenemos que

$$E[\mathcal{X}_2] = \frac{1}{|\mathcal{A}|} \sum_{f \in \mathcal{A}^{sq}} \mathcal{X}_2(f) + \frac{1}{|\mathcal{A}|} \sum_{f \in \mathcal{A}^{nsq}} \mathcal{X}_2(f). \tag{10.11}$$

En primer lugar acotamos superiormente la primera suma S_2^{sq} de (10.11). Para ello expresamos el conjunto $\mathcal{A}^{sq}$ como una unión disjunta de la siguiente manera:

$$\mathcal{A}^{sq} = \bigcup_{i=1}^{d} \mathcal{A}_i^{sq},$$

donde $\mathcal{A}_i^{sq}$ es el conjunto de los elementos $f \in \mathcal{A}^{sq}$ cuyo factor irreducible de mayor grado es de grado i. Así, podemos reescribir a S_2^{sq} en la forma

$$S_2^{sq} = \frac{1}{|\mathcal{A}|} \sum_{i=1}^{d} \sum_{f \in \mathcal{A}_i^{sq}} \mathcal{X}_2(f). \tag{10.12}$$

Más aún, para $1 \leq i \leq d$ podemos expresar cada $\mathcal{A}_i^{sq}$ como la siguiente unión disjunta:

$$\mathcal{A}_i^{sq} = \bigcup_{\boldsymbol{\lambda} \in \mathcal{P}_i} \mathcal{A}_{\boldsymbol{\lambda}}^{sq},$$

donde $\mathcal{P}_i$ es el conjunto de todos los vectores $\boldsymbol{\lambda} := (\lambda_1, \ldots, \lambda_i, 0, \ldots, 0) \in (\mathbb{Z}_{>0})^d$ tales que $1 \cdot \lambda_1 + \cdots + i \cdot \lambda_i = d$ y $\lambda_i \neq 0$, y $\mathcal{A}_{\boldsymbol{\lambda}}^{sq}$ es el conjunto de todos los elementos $f \in \mathcal{A}_i^{sq}$ que tienen patrón de factorización $\boldsymbol{\lambda}$. Por lo tanto, podemos escribir la suma (10.12) de la siguiente manera:

$$S_2^{sq} = \frac{1}{|\mathcal{A}|} \sum_{i=1}^{d} \sum_{\boldsymbol{\lambda} \in \mathcal{P}_i} \sum_{f \in \mathcal{A}_{\boldsymbol{\lambda}}^{sq}} \mathcal{X}_2(f). \tag{10.13}$$

Fijemos i con $1 \leq i \leq d$, sean $\boldsymbol{\lambda} \in \mathcal{P}_i$ y $f \in \mathcal{A}_{\boldsymbol{\lambda}}^{sq}$. A fin de calcular el costo $\mathcal{X}_2(f)$, observamos que el algoritmo realiza i iteraciones del loop principal, ya que el polinomio irreducible de mayor grado que aparece en f es de grado i. Fijemos l con $1 \leq l \leq i$ y consideremos la l–ésima iteración del algoritmo DDF. Notamos con $\lambda(q)$ al número de productos necesarios para calcular $h^q \mod g$. Si usamos el método de exponenciación binaria (ver [Knu98a]), y denotamos por $\nu(q)$ el número de unos en la representación binaria de q, entonces el número de productos necesarios para calcular $h^q \mod g$ es

$$\lambda(q) := \lfloor \log q \rfloor + \nu(q) - 1.$$

Así, el primer paso de la iteración l del algoritmo DDF requiere a lo sumo $\tau_1\lambda(q)M(d_l)$ operaciones aritméticas en $\mathbb{F}_q$, donde d_l es el grado del polinomio g (notemos que $d_1 = d$ y $d_l \leq d$). El cálculo del máximo común divisor $b(k) := \gcd(h - t, g)$ requiere a lo sumo $\tau_2 M(d_l)\log d_l$ operaciones aritméticas en $\mathbb{F}_q$. El tercer paso, o sea, la división $g/b(k)$, requiere a lo sumo $\tau_1 M(d_l)$ operaciones aritméticas en $\mathbb{F}_q$. Como realizamos i iteraciones en total, vemos que $\mathcal{X}_2(f)$ está acotado superiormente por

$$\mathcal{X}_2(f) \leq \sum_{l=1}^{i}(\tau_1\lambda(q) + \tau_2\log d_l + \tau_1)\cdot M(d_l).$$

Si $a \leq b$, entonces $M(a) \leq M(b)$ (ver, por ejemplo, [vzGG99, §14.8]). Así, concluimos que

$$\mathcal{X}_2(f) \leq i\cdot(\tau_1\lambda(q) + \tau_2\log d + \tau_1)\cdot M(d).$$

En consecuencia, el costo del algoritmo DDF aplicado a $f \in \mathcal{A}_{\boldsymbol{\lambda}}^{sq}$ es

$$\mathcal{X}_2(f) \leq i\cdot c_{d,q}, \tag{10.14}$$

donde $c_{d,q} := M(d)(2\tau_1\lambda(q) + \tau_2\log d)$.

Reemplazando (10.14) en la suma (10.13), obtenemos

$$S_2^{sq} := \frac{1}{|\mathcal{A}|}\sum_{i=1}^{d}\sum_{\boldsymbol{\lambda}\in\mathcal{P}_i}\sum_{f\in\mathcal{A}_{\boldsymbol{\lambda}}^{sq}}\mathcal{X}_2(f) \leq \frac{c_{d,q}}{|\mathcal{A}|}\sum_{i=1}^{d}\sum_{\boldsymbol{\lambda}\in\mathcal{P}_i}\sum_{f\in\mathcal{A}_{\boldsymbol{\lambda}}^{sq}} i = \frac{c_{d,q}}{|\mathcal{A}|}\sum_{i=1}^{d} i\sum_{\boldsymbol{\lambda}\in\mathcal{P}_i}|\mathcal{A}_{\boldsymbol{\lambda}}^{sq}|. \tag{10.15}$$

Más aún, tenemos el siguiente resultado.

Lema 10.3.1. *La suma S_2^{sq} está acotada superiormente por*

$$S_2^{sq} \leq c_{d,q}\Big(1 + \frac{Z_{\boldsymbol{L},d}}{q}\Big)\xi(d+1), \tag{10.16}$$

donde, si $p > 2$, $q > d$ y $3 \leq r \leq d-m$, entonces $Z_{\boldsymbol{L},d} := 2\,D_{\boldsymbol{L}}\delta_{\boldsymbol{L}}q^{\frac{1}{2}} + 19\,D_{\boldsymbol{L}}^2\delta_{\boldsymbol{L}}^2 + d^2\delta_{\boldsymbol{L}}$, en tanto que $Z_{\boldsymbol{L},d} := 21\,D_{\boldsymbol{L}}^3\delta_{\boldsymbol{L}}^2 + d^2\delta_{\boldsymbol{L}}$ para $q > d$ y $m+2 \leq r \leq d-m$. En ambos casos, $\delta_{\boldsymbol{L}} := j_1\cdots j_m$ y $D_{\boldsymbol{L}} := \sum_{k=1}^{m}(j_k - 1)$, donde $1 \leq j_1 < \ldots, j_m \leq d-r$ son las posiciones de los pivotes de la matriz $M(\boldsymbol{L})$ de las formas lineales $L_1, \ldots, L_m$ que definen la familia lineal $\mathcal{A}$. A su vez, el número $T(\boldsymbol{\lambda})$ es la cantidad de permutaciones con patrón de descomposición en ciclos $\boldsymbol{\lambda}$ en el grupo simétrico $\mathbb{S}_d$ de d elementos y $\xi \sim 0{,}62432945\ldots$ es la constante de Golomb.

Demostración. Recordemos que los Teoremas 7.2.4 y 7.2.5 proporcionan una estimación del cardinal $|\mathcal{A}_{\boldsymbol{\lambda}}^{sq}|$ de la suma (10.15). Más precisamente, si $p > 2$ y $3 \leq r \leq d-m$, entonces

$$|\mathcal{A}_{\boldsymbol{\lambda}}^{sq}| \leq q^{d-m}T(\boldsymbol{\lambda})\Big(1 + \frac{M_{\boldsymbol{L},d}}{q}\Big),$$

donde $M_{\boldsymbol{L},d} := 2\,D_{\boldsymbol{L}}\delta_{\boldsymbol{L}}q^{\frac{1}{2}} + 19\,D_{\boldsymbol{L}}^2\delta_{\boldsymbol{L}}^2 + d^2\delta_{\boldsymbol{L}}$. Por otro lado, para $m+2 \leq r \leq d-m$ tenemos que

$$|\mathcal{A}_{\boldsymbol{\lambda}}^{sq}| \leq q^{d-m}T(\boldsymbol{\lambda})\Big(1 + \frac{N_{\boldsymbol{L},d}}{q}\Big),$$

donde $N_{\boldsymbol{L},d} := 21\, D_{\boldsymbol{L}}^3 \delta_{\boldsymbol{L}}^2 + d^2 \delta_{\boldsymbol{L}}$. Así, obtenemos la siguiente cota superior para la suma (10.15):

$$S_2^{sq} \leq \frac{c_{d,q}}{|\mathcal{A}|} q^{d-m} \Big(1 + \frac{Z_{\boldsymbol{L},d}}{q}\Big) \sum_{i=1}^{d} i \sum_{\boldsymbol{\lambda} \in \mathcal{P}_i} T(\boldsymbol{\lambda}), \tag{10.17}$$

donde la expresión de $Z_{\boldsymbol{L},d}$ es $M_{\boldsymbol{L},d}$ ó $N_{\boldsymbol{L},d}$, según las hipótesis sobre p, q, d, m y r.

Ahora analizamos la suma $E_d := \sum_{i=1}^{d} i \sum_{\boldsymbol{\lambda} \in \mathcal{P}_i} T(\boldsymbol{\lambda})$. Observemos que la suma $\sum_{\boldsymbol{\lambda} \in \mathcal{P}_i} T(\boldsymbol{\lambda})$ expresa la probabilidad de que el ciclo de mayor longitud de una permutación aleatoria en $\mathbb{S}_d$ sea i. Así, concluimos que E_d es la mayor longitud esperada entre los ciclos de una permutación aleatoria en $\mathbb{S}_d$. En el trabajo [GG98] los autores prueban que E_d cumple la siguiente estimación:

$$\frac{E_d}{d+1} \leq \xi, \tag{10.18}$$

donde $\xi \sim 0{,}62432945\ldots$ es la constante de Golomb (ver [Fin03] o [Knu98a]). Por lo tanto, combinando (10.18) con el hecho de que $|\mathcal{A}| = q^{d-m}$ en la suma (10.17), deducimos que

$$S_2^{sq} \leq c_{d,q} \Big(1 + \frac{Z_{\boldsymbol{L},d}}{q}\Big) \xi (d+1).$$

Así finaliza la demostración del lema. □

Ahora determinamos una cota superior para la segunda suma S_2^{nsq} de (10.11), esto es,

$$S_2^{nsq} := \frac{1}{|\mathcal{A}|} \sum_{f \in \mathcal{A}^{nsq}} \mathcal{X}_2(f). \tag{10.19}$$

Dado $f \in \mathcal{A}^{nsq}$, al igual que antes, queremos estimar $\mathcal{X}_2(f) := \text{Costo}(DDF(a_f))$, donde $a_f := ERF(f)$ es el polinomio libre de cuadrados que resulta de aplicar el algoritmo ERF al polinomio de entrada f. Por (10.14), tenemos que

$$\mathcal{X}_2(f) \leq c_{N,q} \cdot s_a, \tag{10.20}$$

donde $c_{N,q} := M(N)\big(2\tau_1 \lambda(q) + \tau_2 \log d\big)$, $N := \deg(a_f)$ y s_a es el mayor grado de los factores irreducibles de a_f. Como $f \in \mathcal{A}^{nsq}$, observamos que $\deg(a_f) \leq d-1$ y que el mayor de los grados de los factores irreducibles de a_f se acota por $d-2$; por lo tanto, tenemos que $N \leq d-1$ y $s_a \leq d-2$. Más aún, estas cotas superiores son óptimas. En efecto, si $f := f_1^2 \cdot f_{d-2}$, donde f_1 es un polinomio irreducible de grado 1 y f_{d-2} es un polinomio irreducible de grado $d-2$, entonces $N := d-1$ y $s_a := d-2$. En consecuencia, una cota superior para (10.20) es

$$\mathcal{X}_2(f) \leq c_{d-1,q} \cdot (d-2). \tag{10.21}$$

Combinando (10.21) y (10.7) con la suma (10.19), obtenemos que

$$S_2^{nsq} \leq c_{d-1,q}(d-2) \frac{|A^{nsq}|}{|\mathcal{A}|} \leq c_{d-1,q}(d-2) \frac{d^2 q^{d-m-1}}{q^{d-m}} \leq c_{d-1,q} \frac{d^3}{q}. \tag{10.22}$$

Por las cotas superiores (10.16) y (10.22) para las dos sumas en la esperanza $E[\mathcal{X}_2]$, concluimos que

$$E[\mathcal{X}_2] = \frac{1}{|\mathcal{A}|}\sum_{f\in\mathcal{A}^{sq}}\mathcal{X}_2(f) + \frac{1}{|\mathcal{A}|}\sum_{f\in\mathcal{A}^{nsq}}\mathcal{X}_2(f) \leq c_{d,q}\Big(1+\frac{Z_{\boldsymbol{L},d}}{q}\Big)\xi(d+1) + c_{d-1,q}\frac{d^3}{q}, \tag{10.23}$$

donde $c_{N,q} := M(N)(2\tau_1\lambda(q) + \tau_2\log N)$. Observemos que $c_{d-1,q} \leq c_{d,q}$ y estimemos número $c_{d,q}$. Notemos primeramente que $\lambda(q) \leq 2\log q$. En efecto, el número $\nu(q)$ cumple que $1 \leq \nu(m) \leq 1 + \log m$ para todo m. Así, $\nu(q) - 1 \leq \log q$, de lo que se deduce fácilmente la cota superior para $\lambda(q)$. Por lo tanto, $c_{d,q} \leq (4\tau_1\log q + \tau_2\log d)M(d)$ y deducimos el siguiente resultado.

Teorema 10.3.2. *El costo en promedio del algoritmo DDF aplicado a los elementos de $\mathcal{A}$ está acotado superiormente por*

$$E[X_2] \leq (4\tau_1\log q + \tau_2\log d)M(d)\Big(\xi(d+1) + (2\xi d Z_{\boldsymbol{L},d} + d^3)/q\Big),$$

donde $\mathcal{X}_2$ es la variable aleatoria definida en (10.3)*, ξ es la constante de Golomb y $Z_{\boldsymbol{L},d}$ es la constante que aparece en el Lema 10.3.1.*

Observemos que en [FGP01, Section 3, Theorem 5], los autores prueban que un polinomio aleatorio en $\mathbb{F}_q[T]_d$ con alta probabilidad tiene factores irreducibles de grados altos y dan estimaciones asintóticas de las distribuciones conjuntas de los dos grados más altos de los factores irredudibles del polinomio de entrada. Prueban, por ejemplo, que el grado más grande esperado tiende al número $\xi \cdot d$, donde $\xi \sim 0{,}62432\ldots$ es la constante de Golomb (que representa la longitud más grande esperada entre los ciclos de una permutación aleatoria). Esta información les permite probar que el costo en promedio de dicho algoritmo aplicado a un polinomio $f \in \mathbb{F}_q[T]_d$ es del orden de $0{,}26689(\lambda(q)\tau_1 + \tau_2)d^3$, donde $\lambda(q) \leq 2\log q$. Nosotros, en cambio, observamos que el algoritmo DDF realiza tantas iteraciones en el loop principal como el mayor grado de los factores irreducibles del polinomio de entrada y expresamos el conjunto de todas las entradas libres de cuadrados como la unión disjunta de todos los elementos con determinado patrón de factorización. Agrupando estos conjuntos de acuerdo al máximo grado de los factores irreducibles en consideración y utilizando las estimaciones sobre la distribución de patrones en familias lineales del Capítulo 7, probamos que el costo en promedio del algoritmo DDF aplicado a los elementos de $\mathcal{A}$ es del orden de $dM(d)\log(dq)$ operaciones aritméticas en $\mathbb{F}_q$, mejorando así la estimación dada en [FGP01].

El algoritmo DDF no factoriza completamente un polinomio $f \in \mathcal{A}$ que tiene distintos factores irreducibles del mismo grado. Concluimos esta sección con un estudio de la probablidad de que el algoritmo DDF complete la factorización de un polinomio $f \in \mathcal{A}$. Para ello, observamos que, para que el algoritmo clásico de factorización termine en este paso, debe ocurrir que el patrón de factorización de f sea $\boldsymbol{\lambda} := (\lambda_1, \ldots, \lambda_d) \in \{0,1\}^d$.

En [FS09] los autores prueban que la mayoría de las factorizaciones se completan luego de la aplicación del algoritmo DDF. Más precisamente, prueban que, cuando

d está fijo y q tiende a infinito, la probabilidad de que el algoritmo DDF produzca una factorización completa de un polinomio aleatorio en $\mathbb{F}_q[T]_d$ es del orden de $e^{-\gamma} \sim 0{,}5614\ldots$, donde $\gamma \sim 0{,}577215664\ldots$ es la constante de Euler (ver [FS09, Theorem 6]). En el resultado a continuación probamos un resultado similar para la familia $\mathcal{A}$.

Teorema 10.3.3. *La probabilidad de que el algoritmo DDF complete la factorización de un polinomio aleatorio de $\mathcal{A}$ es del orden de $e^{-\gamma} + o(1)$, donde γ es la constante de Euler.*

Demostración. Consideremos la familia $\mathcal{A}_1$ que consiste de todos los elementos de $\mathcal{A}$ cuyos factores irreducibles son todos de grados distintos. La probabilidad de que el algoritmo DDF complete la factorización de un polinomio aleatorio de $\mathcal{A}$ coincide con la probabilidad de que un polinomio aleatorio pertenezca a $\mathcal{A}_1$. A fin de estimar la probabilidad $P[\mathcal{A}_1]$, comenzamos observando que podemos expresar a $\mathcal{A}_1$ como la unión disjunta

$$\mathcal{A}_1 = \bigcup_{\boldsymbol{\lambda} \in \mathcal{P}_d} \mathcal{A}_{1,\boldsymbol{\lambda}},$$

donde $\mathcal{P}_d$ es el conjunto de todos los vectores $\boldsymbol{\lambda} := (\lambda_1, \ldots, \lambda_d) \in \{0,1\}^d$ tales que $1 \cdot \lambda_1 + \cdots + d \cdot \lambda_d = d$ y $\mathcal{A}_{1,\boldsymbol{\lambda}}$ es el conjunto de todos los elementos de $\mathcal{A}_1$ que tienen patrón de factorización $\boldsymbol{\lambda}$. Así, la probabilidad de que el algoritmo DDF complete la factorización de un polinomio aleatorio de $\mathcal{A}$ se puede expresar como la suma

$$P[A_1] = \sum_{\boldsymbol{\lambda} \in \mathcal{P}_d} P[\mathcal{A}_{1,\boldsymbol{\lambda}}] = \frac{1}{|\mathcal{A}|} \sum_{\boldsymbol{\lambda} \in \mathcal{P}_d} |\mathcal{A}_{1,\boldsymbol{\lambda}}|. \tag{10.24}$$

Si $f \in \mathcal{A}_1$, entonces f es libre de cuadrados, y los Teoremas 7.2.4 y 7.2.5 implican que

$$|\mathcal{A}_{1,\boldsymbol{\lambda}}| \leq q^{d-m} T(\boldsymbol{\lambda}) \Big(1 + \frac{Z_{\boldsymbol{L},d}}{q}\Big),$$

donde $Z_{\boldsymbol{L},d}$ son las constantes que aparecen en el Lema 10.3.1 y $\boldsymbol{\lambda} := (\lambda_1, \ldots, \lambda_d) \in \{0,1\}^d$. Así, tenemos que

$$P[\mathcal{A}_1] \leq \Big(1 + \frac{Z_{\boldsymbol{L},d}}{q}\Big) \sum_{\boldsymbol{\lambda} \in \mathcal{P}_d} T(\boldsymbol{\lambda}). \tag{10.25}$$

Ahora bien, $\sum_{\boldsymbol{\lambda} \in \mathcal{P}_d} T(\boldsymbol{\lambda})$ expresa la probabilidad de que una permutación aleatoria de $\mathbb{S}_d$ tenga ciclos de longitud distinta. En [FS09, Theorem 6] se prueba el siguiente resultado sobre dicha probabilidad:

$$\sum_{\boldsymbol{\lambda} \in \mathcal{P}_d} T(\boldsymbol{\lambda}) = e^{-\gamma} + o(1).$$

Así, de (10.25) y la probabilidad anterior, deducimos que

$$P[\mathcal{A}_1] \leq \big(1 + \frac{Z_{\boldsymbol{L},d}}{q}\big)(e^{-\gamma} + o(1)). \tag{10.26}$$

Esto concluye la demostración del teorema. □

10.4. Factorización en grados iguales

Luego de los dos primeros pasos del algoritmo clásico de factorización, el problema general de factorización se reduce a factorizar una colección de polinomios mónicos libres de cuadrados $b(k)$, cuyos factores irreducibles tienen el mismo grado k. Más precisamente, luego de aplicar los primeros dos pasos de dicho algoritmo tenemos un vector $\boldsymbol{b}_f := DDF(a_f) = (b_f(1), \ldots, b_f(s))$, donde cada $b_f(k)$ es el producto de los factores irreducibles de grado k del polinomio a_f y $a_f := ERF(f)$ es la parte libre de cuadrados de f. El siguiente paso en este proceso se conoce como el algoritmo de factorización en grados iguales (algoritmo EDF), y su objetivo es factorizar cada $b_f(k)$ como un producto de polinomios irreducibles $b_f(k,1) \ldots b_f(k,l)$. El algoritmo probabilístico que presentamos aquí está basado en el algoritmo de Cantor–Zassenhaus [CZ81].

Algoritmo EDF.

Entrada: c es un polinomio libre de cuadrados cuyos factores irreducibles tienen grado k.

Salida: La factorización completa de c.

procedimiento EDF(c: polinomio libre de cuadrados, k: entero)

Si $\deg c = k$, *entonces devolver* c
Fin si
Elegir un polinomio aleatorio $h := \text{randpoly}(\deg(c)-1)$ *de grado* $\deg c - 1$.
Calcular $g := h^{(q^k-1)/2} - 1 \mod c$
Calcular $r := \gcd(g, c)$
Devolver $\text{EDF}(r,k) \cdot \text{EDF}(c/r, k)$.

El algoritmo EDF se basa en un principio que nos permitirá "aislar" los distintos factores irreducibles del polinomio de entrada. Supongamos que el polinomio de entrada c es un producto de j factores irreducibles $f_1, \ldots, f_j$, cada uno de grado k. El Teorema Chino del resto implica que

$$\mathbb{F}_q[T]/(c) \cong \mathbb{F}_q[T]/(f_1) \times \cdots \times \mathbb{F}_q[T]/(f_j).$$

Por este isomorfismo, un elemento aleatorio h de $\mathbb{F}_q[T]/(c)$ está asociado con una j–upla $(h_1, \ldots, h_j)$, donde cada h_i es un elemento aleatorio de $\mathbb{F}_q[T]/(f_i)$. Como cada f_i es irreducible, tenemos que el anillo cociente $\mathbb{F}_q[T]/(f_i)$ es un cuerpo finito, isomorfo a $\mathbb{F}_{q^k}$. Como el grupo multiplicativo $\mathbb{F}_{q^k}^* := \mathbb{F}_{q^k} \setminus \{0\}$ es cíclico, contiene $\frac{q^k-1}{2}$ cuadrados y $\frac{q^k-1}{2}$ que no lo son (ver [vzGG99, Lemma 14.7]). Cabe recordar que un elemento $m \in \mathbb{F}_{q^k}^*$ es un cuadrado si y solo si $m^{(q^k-1)/2} = 1$. Por lo tanto, chequear si $h_i^{(q^k-1)/2} = 1$ discrimina los cuadrados en $\mathbb{F}_{q^k}^*$. Así, eligiendo un polinomio

h aleatoriamente y calculando $g := h^{(q^k-1)/2}-1 \mod c$, vemos que el máximo común divisor $\gcd(g,c)$ "extrae" el producto de todos los factores irreducibles f_i de c para los cuales h es un cuadrado en $\mathbb{F}_q[T]/(f_i)$. Finalmente, el algoritmo EDF se aplica recursivamente a los polinomios $r := \gcd(g,c)$ y c/r.

Desde un punto de vista probabilístico, cada componente h_i es un elemento aleatorio de $\mathbb{F}_q[T]/(f_i)$, que tiene probabilidad $\alpha := \frac{1}{2} - \frac{1}{2q^k}$ de ser un cuadrado (es decir, de ser discriminado por el cálculo del máximo común divisor $\gcd(g,c)$), y probabilidad dual, $\beta := \frac{1}{2} + \frac{1}{2q^k}$, de no ser un cuadrado.

De esta manera, el algoritmo EDF calcula recursivamente los factores irreducibles de cada factor $c := b(k)$. En el trabajo de [FS09, Section 5] se analiza en promedio este algoritmo aplicado a $\mathbb{F}_q[T]_d$, para lo cual se utiliza una estimación de la probabilidad de que existan j factores irreducibles de grado k en un polinomio aleatorio de grado d (ver [KK90b]).

Siguiendo estas ideas, en esta sección analizamos la complejidad en promedio del algoritmo EDF aplicado a la familia lineal $\mathcal{A}$. Más precisamente, obtenemos una cota superior para la esperanza $E[\mathcal{X}_3]$ de la variable aleatoria $\mathcal{X}_3$ de (10.4), es decir,

$$E[\mathcal{X}_3] := \frac{1}{|\mathcal{A}|}\sum_{f\in\mathcal{A}} \mathcal{X}_3(f).$$

La variable aleatoria $\mathcal{X}_3$ se puede descomponer en la suma de las siguientes variables aleatorias (ver (10.4)):

$$\mathcal{X}_3(f) := \sum_{k=1}^{d} \mathcal{X}_{3,k}(f),$$

donde, para cada k, la variable aleatoria $\mathcal{X}_{3,k}(f)$ se define como

$$\mathcal{X}_{3,k}(f) := \text{Costo}(EDF(b_f(k))),$$

siendo $b_f(k)$ la k–ésima coordenada del vector $\boldsymbol{b}_f := DDF(a_f) = (b_f(1), \dots b_f(s))$. Así, tenemos que

$$E[\mathcal{X}_3] = \frac{1}{|\mathcal{A}|}\sum_{f\in\mathcal{A}}\sum_{k=1}^{\lceil d/2\rceil} \mathcal{X}_{3,k}(f) = \frac{1}{|\mathcal{A}|}\sum_{k=1}^{\lceil d/2\rceil}\sum_{f\in\mathcal{A}} \mathcal{X}_{3,k}(f) = \sum_{k=1}^{\lceil d/2\rceil} E[\mathcal{X}_{3,k}]. \tag{10.27}$$

Fijamos k con $1 \le k \le \lceil d/2\rceil$ y estimamos la esperanza $E[\mathcal{X}_{3,k}]$. Tenemos que

$$E[\mathcal{X}_{3,k}] = \frac{1}{|\mathcal{A}|}\sum_{f\in\mathcal{A}} \mathcal{X}_{3,k}(f) = \frac{1}{|\mathcal{A}|}\sum_{f\in\mathcal{A}^{sq}} \mathcal{X}_{3,k}(f) + \frac{1}{|\mathcal{A}|}\sum_{f\in\mathcal{A}^{nsq}} \mathcal{X}_{3,k}(f). \tag{10.28}$$

Empezamos estimando la primera suma $S_{3,k}^{sq}$ del término de la derecha en (10.28). Para ello, dado $f \in \mathcal{A}^{sq}$, nuestro propósito es dar una cota superior del costo $\mathcal{X}_{3,k}(f) := \text{Costo}(EDF(b_f(k)))$ de aplicar el algoritmo EDF al polinomio $b_f(k)$. Podemos expresar a $\mathcal{A}^{sq}$ como la unión disjunta

$$\mathcal{A}^{sq} = \bigcup_{j=0}^{\lfloor d/k\rfloor} \mathcal{A}_{j,k}^{sq},$$

donde $\mathcal{A}_{j,k}^{sq}$ es el conjunto de todos los elementos $f \in \mathcal{A}^{sq}$ que tienen j factores irreducibles de grado k. En consecuencia,

$$S_{3,k}^{sq} = \frac{1}{|\mathcal{A}|} \sum_{j=0}^{\lfloor d/k \rfloor} \sum_{f \in \mathcal{A}_{j,k}^{sq}} \mathcal{X}_{3,k}(f). \tag{10.29}$$

El análisis del costo $\mathcal{X}_{3,k}(f)$ de aplicar el algoritmo EDF a cualquier $f \in \mathcal{A}_{j,k}^{sq}$ se describe en [FS09, Section 5]. Cabe observar que el costo de una llamada recursiva del algoritmo EDF aplicado a f está determinado por el costo de calcular $h^{(q^k-1)/2}$ mód f, donde h es un elemento aleatorio de $\mathbb{F}_q[T]/(f_i)$, del producto de dos polinomios de grado a lo sumo jk y del cálculo del máximo común divisor de f con un polinomio de grado a lo sumo jk. El número de productos necesarios para calcular $h^{(q^k-1)/2}$ mód f, usando el proceso de exponenciación binaria, es

$$\mu_k := \lambda\Big(\frac{q^k-1}{2}\Big) := \Big\lfloor \log\Big(\frac{q^k-1}{2}\Big)\Big\rfloor + \nu\Big(\frac{q^k-1}{2}\Big) - 1.$$

Así, la cantidad de operaciones en $\mathbb{F}_q$ necesarias para calcular $h^{(q^k-1)/2}$ mód f es a lo sumo $\tau_2\mu_k M(jk)$. Además, el costo del producto de dos polinomios de grado a lo sumo jk es de a lo sumo de $\tau_1 M(jk)$ operaciones aritméticas en $\mathbb{F}_q$, en tanto que el costo del máximo común divisor es de a lo sumo $\tau_2\mathcal{U}(jk)$ operaciones aritméticas en $\mathbb{F}_q$. El siguiente lema proporciona una cota superior del costo $\mathcal{X}_{3,k}(f)$.

Lema 10.4.1 ([FS09, Lemma 4])**.** *El costo $\mathcal{X}_{3,k}(f)$ del algoritmo EDF aplicado a cualquier $f \in \mathcal{A}_{j,k}^{sq}$ se acota por*

$$\mathcal{X}_{3,k}(f) \le \left(\frac{j(j-1)}{2\alpha\beta} + j\sum_{m=0}^{\infty}\sum_{l=0}^{m}\binom{m}{l}\alpha^{m-l}\beta^l(1-(1-\alpha^{m-l}\beta^l)^{j-1})\right)\cdot(\mu_k\tilde{\tau}_1+\tilde{\tau}_2)k^2,$$

donde $\tilde{\tau}_1 := 2\frac{\tau_1 M(d)}{kd}$, $\tilde{\tau}_2 := \frac{\tau_2\mathcal{U}(d)}{kd}$ *y* $\mu_k := \lfloor\log(\frac{q^k-1}{2})\rfloor + \nu(\frac{q^k-1}{2}) - 1$.

A continuación obtenemos una cota superior simple del costo $\mathcal{X}_{3,k}(f)$. Aplicando la desigualdad $1-(1-u)^{j-1} \le (j-1)u$ para $j \ge 2$ y $0 \le u \le 1$ en el Lema 10.4.1, vemos que

$$\begin{aligned}\mathcal{X}_{3,k}(f) &\le \left(\frac{j(j-1)}{2\alpha\beta} + j\sum_{m=0}^{\infty}\sum_{l=0}^{m}\binom{m}{l}(j-1)\alpha^{2(m-l)}\beta^{2l}\right)\cdot(\mu_k\tilde{\tau}_1+\tilde{\tau}_2)k^2 \qquad (10.30)\\ &= \frac{j(j-1)}{\alpha\beta}\cdot(\mu_k\tilde{\tau}_1+\tilde{\tau}_2)k^2,\end{aligned}$$

donde la última igualdad se sigue de la identidad

$$\sum_{m=0}^{\infty}\sum_{l=0}^{m}\binom{m}{l}\alpha^{2(m-l)}\beta^{2l} = \sum_{m=0}^{\infty}(\alpha^2+\beta^2)^m = \frac{1}{2\alpha\beta}.$$

Así, por (10.30) tenemos que

$$S_{3,k}^{sq} := \frac{1}{|\mathcal{A}|}\sum_{j=0}^{\lfloor d/k\rfloor}\sum_{f\in\mathcal{A}_{j,k}^{sq}}\mathcal{X}_{3,k}(f) \leq \sum_{j=0}^{\lfloor d/k\rfloor}\frac{j(j-1)}{\alpha\beta}\cdot(\mu_k\tilde{\tau}_1+\tilde{\tau}_2)k^2\cdot\frac{|\mathcal{A}_{j,k}^{sq}|}{|\mathcal{A}|}. \tag{10.31}$$

A continuación estimamos la probabilidad $P_{j,k}^{\mathcal{A}}[\mathcal{A}_{j,k}^{sq}]$ de que un polinomio aleatorio $f\in\mathcal{A}$ sea libre de cuadrados y tenga j factores irreducibles de grado k. En la literatura observamos que, en [KK90b], se prueba que si q es suficientemente grande, entonces la probabilidad de que un polinomio aleatorio $f\in\mathbb{F}_q[T]_d$ tenga j factores irreducibles distintos de grado k se aproxima al número $e^{-1/k}\frac{k^{-j}}{j!}$.

Partimos el conjunto $\mathcal{A}_{j,k}^{sq}$ en la siguiente unión disjunta:

$$\mathcal{A}_{j,k}^{sq} = \bigcup_{\boldsymbol{\lambda}\in\mathcal{P}_d^{j,k}}\mathcal{A}_{j,\boldsymbol{\lambda}}^{sq},$$

donde $\mathcal{P}_d^{j,k}$ es el conjunto de todas las d–uplas $\boldsymbol{\lambda} := (\lambda_1,\ldots,\lambda_d)\in\mathbb{Z}_{\geq 0}^d$ tales que $1\cdot\lambda_1+\cdots+d\cdot\lambda_d = d$ y $\lambda_k = j$. Así, tenemos que

$$P_{j,k}^{\mathcal{A}}[\mathcal{A}_{j,k}^{sq}] = \frac{1}{|\mathcal{A}|}\sum_{\boldsymbol{\lambda}\in\mathcal{P}_d^{j,k}}|\mathcal{A}_{j,\boldsymbol{\lambda}}^{sq}|. \tag{10.32}$$

De los Teoremas 7.2.4 y 7.2.5 deducimos que

$$|\mathcal{A}_{j,\boldsymbol{\lambda}}^{sq}| \leq q^{d-m}T(\boldsymbol{\lambda})\big(1+\frac{Z_{\boldsymbol{L},d}}{q}\big), \tag{10.33}$$

donde $Z_{\boldsymbol{L},d}$ denota alguna de las constantes que aparecen en Lema 10.3.1. En consecuencia, tenemos que

$$P_{j,k}^{\mathcal{A}}[\mathcal{A}_{j,k}^{sq}] = \frac{1}{|\mathcal{A}|}\sum_{\boldsymbol{\lambda}\in\mathcal{P}_d^{j,k}}|\mathcal{A}_{j,\boldsymbol{\lambda}}^{sq}| \leq \big(1+\frac{Z_{\boldsymbol{L},d}}{q}\big)\sum_{\boldsymbol{\lambda}\in\mathcal{P}_d^{j,k}}T(\boldsymbol{\lambda}). \tag{10.34}$$

Observemos que la suma del término de la derecha en (10.34) expresa la probabilidad de que una permutación aleatoria en $\mathbb{S}_d$ tenga exactamente j ciclos de longitud k. En [SL66] se muestra que dicha probabilidad es

$$\sum_{\boldsymbol{\lambda}\in\mathcal{P}_d^{j,k}}T(\boldsymbol{\lambda}) = \frac{1}{j!k^j}\sum_{i=0}^{\lfloor d/k-j\rfloor}(-1)^i\frac{1}{i!k^i}. \tag{10.35}$$

Se deduce la siguiente igualdad, que corresponde al hecho de que la suma de las probabilidades es igual a 1:

$$\sum_{j=0}^{\lfloor d/k\rfloor}\frac{1}{j!k^j}\sum_{i=0}^{\lfloor d/k-j\rfloor}(-1)^i\frac{1}{i!k^i} = 1. \tag{10.36}$$

Combinando (10.34), (10.35), (10.36) con (10.31) deducimos que

$$\begin{aligned} S_{3,k}^{sq} &\leq \sum_{j=0}^{\lfloor d/k \rfloor} \frac{j(j-1)}{\alpha\beta} \cdot (\mu_k \tilde{\tau}_1 + \tilde{\tau}_2) k^2 \cdot \big(1 + \frac{Z_{\boldsymbol{L},d}}{q}\big) \sum_{\boldsymbol{\lambda}_j \in \mathcal{P}_d^{j,k}} T(\boldsymbol{\lambda}) \\ &\leq \sum_{j=2}^{\lfloor d/k \rfloor} \frac{j(j-1)}{\alpha\beta} \cdot (\mu_k \tilde{\tau}_1 + \tilde{\tau}_2) k^2 \cdot \big(1 + \frac{Z_{\boldsymbol{L},d}}{q}\big) \frac{1}{j! k^j} \sum_{i=0}^{\lfloor d/k-j \rfloor} (-1)^i \frac{1}{i! k^i} \\ &\leq \frac{(\mu_k \tilde{\tau}_1 + \tilde{\tau}_2)}{\alpha\beta} \big(1 + \frac{Z_{\boldsymbol{L},d}}{q}\big) \sum_{j=2}^{\lfloor d/k \rfloor} \frac{1}{(j-2)! k^{j-2}} \sum_{i=0}^{\lfloor d/k-j \rfloor} (-1)^i \frac{1}{i! k^i} \\ &\leq \frac{(\mu_k \tilde{\tau}_1 + \tilde{\tau}_2)}{\alpha\beta} \big(1 + \frac{Z_{\boldsymbol{L},d}}{q}\big). \end{aligned} \tag{10.37}$$

Por otro lado, estimamos la segunda suma $S_{3,k}^{nsq}$ de (10.28):

$$S_{3,k}^{nsq} := \frac{1}{|\mathcal{A}|} \sum_{f \in \mathcal{A}^{nsq}} \mathcal{X}_{3,k}(f). \tag{10.38}$$

Fijemos $f \in \mathcal{A}^{nsq}$. El objetivo ahora es acotar superiormente el costo $\mathcal{X}_{3,k}(f) := \text{Costo}(EDF(b_f(k)))$, donde $b_f(k)$ es la k–ésima coordenada del vector $\boldsymbol{b}_f := DDF(a_f) = (b_f(1), \ldots, b_f(s))$. Notemos que $\deg(a_f) < \deg(f)$.

Supongamos que $\deg(b_f(k)) = m_k$. En [vzGG99, Theorem 14.11], los autores dan la siguiente cota superior del costo $\mathcal{X}_{3,k}(f)$:

$$\mathcal{X}_{3,k}(f) \leq c \cdot (k \log q + \log m_k) M(m_k) \log\Big(\frac{m_k}{k}\Big),$$

donde c es una constante independiente de k y q. Por esta desigualdad y la estimación para $|\mathcal{A}^{nsq}|$ de (10.7), obtenemos la siguiente cota superior:

$$\begin{aligned} S_{3,k}^{nsq} &\leq c \cdot (k \log q + \log m_k) M(m_k) \log\Big(\frac{m_k}{k}\Big) \frac{|\mathcal{A}^{nsq}|}{|\mathcal{A}|} \\ &\leq c \cdot (k \log q + \log m_k) M(m_k) \log\Big(\frac{m_k}{k}\Big) \frac{d^2}{q}. \end{aligned} \tag{10.39}$$

Por lo tanto, de (10.37) y (10.39) deducimos la siguiente cota superior para la esperanza $E[\mathcal{X}_3]$ de (10.27):

$$\begin{aligned} E[\mathcal{X}_3] &= \sum_{k=1}^{\lceil d/2 \rceil} \frac{1}{|\mathcal{A}|} \sum_{f \in \mathcal{A}^{sq}} \mathcal{X}_{3,k}(f) + \sum_{k=1}^{\lceil d/2 \rceil} \frac{1}{|\mathcal{A}|} \sum_{f \in \mathcal{A}^{nsq}} \mathcal{X}_{3,k}(f) \\ &\leq \Big(1 + \frac{Z_{\boldsymbol{L},d}}{q}\Big) \sum_{k=1}^{\lceil d/2 \rceil} \frac{(\mu_k \tilde{\tau}_1 + \tilde{\tau}_2)}{\alpha\beta} + c \frac{d^2}{q} \sum_{k=1}^{\lceil d/2 \rceil} (k \log q + \log m_k) M(m_k) \log\Big(\frac{m_k}{k}\Big). \end{aligned} \tag{10.40}$$

Más aún, tenemos el siguiente resultado.

Teorema 10.4.2. *El costo en promedio del algoritmo EDF aplicado a los elementos de $\mathcal{A}$ está acotado superiormente por*

$$E[X_3] \leq \tau M(d) \log q(1 + \log d) + \frac{\tau M(d) \log q(2Z_{\boldsymbol{L},d} \log d + d^3)}{q},$$

donde $\mathcal{X}_3$ es la variable aleatoria de (10.4)*, τes una constante independiente de q y d y $Z_{\boldsymbol{L},d}$ es la constante del enunciado del Lema 10.3.1.*

Demostración. A fin de estimar la esperanza $E[\mathcal{X}_3]$, estimamos las dos sumas de (10.40). Empecemos con la primera suma, esto es,

$$S_1 := \sum_{k=1}^{\lceil d/2 \rceil} \frac{(\mu_k \tilde{\tau}_1 + \tilde{\tau}_2)}{\alpha\beta}, \tag{10.41}$$

donde $\tilde{\tau}_1 := 2\frac{\tau_1 M(d)}{kd}$ y $\tilde{\tau}_2 := \frac{\tau_2 \mathcal{U}(d)}{kd}$ son las constantes del enunciado del Lema 10.4.1. Como $\alpha := \frac{1}{2} - \frac{1}{2q^k}$ y $\beta := \frac{1}{2} + \frac{1}{2q^k}$, tenemos que

$$\frac{1}{\alpha\beta} \leq 4\frac{q^2}{q^2 - 1} \leq \frac{16}{3}.$$

A su vez, por la definición de $\mu_k := \lfloor \log(\frac{q^k - 1}{2}) \rfloor + \nu(\frac{q^k - 1}{2}) - 1$, vemos que

$$\mu_k \leq 2k \log q.$$

Por lo tanto, reemplazando estas cotas superiores en (10.41) deducimos que

$$\begin{aligned} S_1 &\leq \sum_{k=1}^{\lceil d/2 \rceil} \frac{\mu_k \tilde{\tau}_1}{\alpha\beta} + \sum_{k=1}^{\lceil d/2 \rceil} \frac{\tilde{\tau}_2}{\alpha\beta} \qquad (10.42) \\ &\leq \frac{64\tau_1}{3} \frac{M(d)\lceil d/2 \rceil \log q}{d} + \frac{16\tau_2}{3} \frac{\mathcal{U}(d)}{d} \sum_{k=1}^{\lceil d/2 \rceil} \frac{1}{k} \\ &\leq \frac{64\tau_1}{3} M(d) \log q + \frac{16\tau_2}{3} \log q \ \frac{H(\lceil d/2 \rceil)\mathcal{U}(d)}{d} \\ &= M(d) \log q \Big(\frac{64\tau_1}{3} + \frac{16\tau_2}{3} \frac{\log d \ H(\lceil d/2 \rceil)}{d} \Big), \end{aligned}$$

donde $H(\lceil d/2 \rceil)$ es el número armónico de $\lceil d/2 \rceil$. Teniendo en cuenta que $\log(N + 1) \leq H(N) \leq 1 + \log N$ (ver, por ejemplo, [GKP94, §6.3]), es fácil deducir que, si $d \geq 2$, entonces

$$0 \leq \frac{\log d \ H(\lceil d/2 \rceil)}{d} \leq 1.$$

Concluimos que

$$S_1 \leq M(d) \log q \big(\frac{64\tau_1}{3} + \frac{16\tau_2}{3} \log d \big). \tag{10.43}$$

Ahora estimamos la segunda suma de (10.40), es decir,

$$S_2 := \sum_{k=1}^{\lceil d/2 \rceil} (k \log q + \log m_k) M(m_k) \log\Big(\frac{m_k}{k}\Big). \tag{10.44}$$

Tenemos las siguientes desigualdades:

$$\sum_{k=1}^{\lceil d/2 \rceil} kM(m_k) \log\Big(\frac{m_k}{k}\Big) \leq M(d) \sum_{k=1}^{\lceil d/2 \rceil} m_k \frac{\log\left(\frac{m_k}{k}\right)}{\frac{m_k}{k}} \leq M(d) \sum_{k=1}^{\lceil d/2 \rceil} m_k \leq dM(d),$$

$$\sum_{k=1}^{\lceil d/2 \rceil} M(m_k) \log(m_k) \log\Big(\frac{m_k}{k}\Big) \leq M(d) \sum_{k=1}^{\lceil d/2 \rceil} \log^2(m_k) \leq M(d) \sum_{k=1}^{\lceil d/2 \rceil} m_k \leq dM(d).$$

Así deducimos que

$$S_2 \leq 2dM(d) \log q. \tag{10.45}$$

Reemplazando las cotas superiores de (10.43) y (10.45) en (10.40), obtenemos la siguiente cota superior para la esperanza $E[\mathcal{X}_3]$:

$$E[\mathcal{X}_3] \leq \Big(1+\frac{Z_{\boldsymbol{L},d}}{q}\Big) S_1 + c\frac{d^2}{q} S_2 \leq M(d) \log q \bigg(\Big(1+\frac{Z_{\boldsymbol{L},d}}{q}\Big)\Big(\frac{64\tau_1}{3} + \frac{16\tau_2}{3} \log d\Big) + \frac{2cd^3}{q} \bigg). \tag{10.46}$$

Tomando $\tau := \max\{\frac{64\tau_1}{3}, \frac{16\tau_2}{3}, 2c\}$, concluimos la demostración del teorema. □

Observemos que, en [FS09, Theorem 9], utilizando la multiplicación clásica de polinomios, los autores prueban que el costo promedio de algoritmo EDF en $\mathbb{F}_q[T]_d$ es del orden de $\frac{3\tau_1}{4} \frac{q^2}{q^2-1} \log q (1+\xi_d) d^2$, donde $|\xi_d| \leq \frac{1}{3} + o(1)$, es decir, el algoritmo EDF utiliza en promedio $\mathcal{O}(d^2 \log q)$ operaciones aritméticas en $\mathbb{F}_q$. Nosotros, en cambio, utilizamos la multiplicación rápida y los resultados sobre la distribución de patrones de factorización en $\mathcal{A}$ a fin de demostrar que el algoritmo EDF aplicado a dicha familia realiza en promedio $\mathcal{O}(M(d) \log q)$ operaciones aritméticas en $\mathbb{F}_q$.

Por otro lado, cabe destacar que este análisis mejora el del costo del peor caso de [vzGG99, Theorem 14.11]. En dicho trabajo, los autores aseguran que el costo del algoritmo EDF aplicado a un polinomio de grado a lo sumo d que tiene j factores irreducibles de grado k es del orden de $\mathcal{O}((k \log q + \log d) M(d) \log j)$ operaciones aritméticas en $\mathbb{F}_q$, o sea, $\mathcal{O}^{\sim}(d^2 \log q)$ operaciones aritméticas en $\mathbb{F}_q$.

10.5. Costo en promedio del algoritmo clásico

En esta sección concluimos el análisis del costo en promedio del algoritmo de factorización en $\mathcal{A}$. Para esto, comenzamos con el análisis de la complejidad en promedio del último paso del algoritmo clásico de factorización, es decir, la esperanza de la variable aleatoria $\mathcal{X}_4$ que cuenta la cantidad de operaciones aritméticas en $\mathbb{F}_q$ que realiza el algoritmo clásico aplicado al polinomio $f/ERF(f)$, cuando f recorre los elementos de $\mathcal{A}$. Más precisamente, estudiamos

$$E[\mathcal{X}_4] := \frac{1}{|\mathcal{A}|}\sum_{f\in\mathcal{A}} \mathcal{X}_4(f).$$

Podemos reescribir la esperanza $E[\mathcal{X}_4]$ de la siguiente manera:

$$E[\mathcal{X}_4] = \frac{1}{|\mathcal{A}|}\sum_{f\in\mathcal{A}^{sq}} \mathcal{X}_4(f) + \frac{1}{|\mathcal{A}|}\sum_{f\in\mathcal{A}^{nsq}} \mathcal{X}_4(f). \tag{10.47}$$

Estimamos la primera suma S_4^{sq} de (10.47). Si $f \in \mathcal{A}^{sq}$, entonces $f/ERF(f) = 1$. Por lo tanto, el costo de este paso es de a lo sumo $\tau_1 M(d)$ operaciones aritméticas en $\mathbb{F}_q$, esto es, el costo de dividir dos polinomios de grado a lo sumo d. En consecuencia,

$$S_4^{sq} := \frac{1}{|\mathcal{A}|}\sum_{f\in\mathcal{A}^{sq}} \mathcal{X}_4(f) \leq \tau_1 M(d). \tag{10.48}$$

Por otro lado, acotamos la segunda suma S_4^{nsq} de (10.47). Para ello, descomponemos el conjunto $\mathcal{A}^{nsq}$ como la unión disjunta del conjunto $\mathcal{A}_{=2}^{nsq}$ cuyos elementos tienen todos los factores irreducibles de multiplicidad a lo sumo 2 y $\mathcal{A}_{\geq 2}^{nsq} := \mathcal{A}^{nsq} \setminus \mathcal{A}_{=2}^{nsq}$.

Así, si $f \in \mathcal{A}_{=2}^{nsq}$ entonces f es de la forma $f = \prod f_i \prod f_j^2$. Por lo tanto, tenemos que $f/ERF(f) = \prod f_j$. En consecuencia, en este caso solo se ejecutan los tres primeros pasos del algoritmo de factorización. Por lo tanto, del análisis del peor caso del algoritmo clásico de factorización de [vzGG99, Theorem 14.14] concluimos que $\mathcal{X}_4(f) \leq c_3 dM(d)\log(dq)$, donde c_3 es una constante independiente de d y q. En cambio, si $f \in \mathcal{A}_{\geq 2}^{nsq}$, entonces, además de ejecutarse los tres primeros pasos del algoritmo, se ejecuta el cuarto tantas veces como la máxima multiplicidad de los factores irreducibles de $f/ERF(f)$. Así, el análisis del peor caso de [vzGG99, Theorem 14.14] implica que $\mathcal{X}_4(f) \leq c_4 d^2 M(d)\log(dq)$, donde c_4 es una constante independiente de d y q. De estas observaciones se sigue que

$$\begin{aligned} S_4^{nsq} &:= \frac{1}{|\mathcal{A}|}\sum_{f\in\mathcal{A}_{=2}^{nsq}} \mathcal{X}_4(f) + \frac{1}{|\mathcal{A}|}\sum_{f\in\mathcal{A}_{\geq 2}^{nsq}} \mathcal{X}_4(f) \\ &\leq c_3 dM(d)\log(dq)\frac{|\mathcal{A}_{=2}^{nsq}|}{|\mathcal{A}|} + c_4 d^2 M(d)\log(dq)\frac{|\mathcal{A}_{\geq 2}^{nsq}|}{|\mathcal{A}|} \end{aligned} \tag{10.49}$$

Dado que $\mathcal{A}_{=2}^{nsq}$ es una subfamilia de $\mathcal{A}^{nsq}$, por (10.7) tenemos que

$$|\mathcal{A}_{=2}^{nsq}| \leq d(d-1)q^{d-m-1} \leq d^2 q^{d-m-1}. \tag{10.50}$$

Por otro lado, si $f \in \mathcal{A}_{\geq 2}^{nsq}$, entonces el grado del máximo común divisor entre f y su derivada es al menos 2. Deducimos que $\mathrm{Res}(f, f') = \mathrm{Subres}(f, f') = 0$. Por lo tanto, la familia $\mathcal{A}_{\geq 2}^{nsq}$ está incluida en $\mathcal{D}(V(\boldsymbol{L})) \cap \mathcal{S}_1(V(\boldsymbol{L}))$, donde $V(\boldsymbol{L}) \subset \mathbb{A}^d$ es la variedad definida por las formas lineales $L_1, \ldots, L_m$ de (10.1), $\mathcal{D}(V(\boldsymbol{L}))$ es el lugar discriminante de $V(\boldsymbol{L})$ y $\mathcal{S}_1(V(\boldsymbol{L}))$ es el lugar del primer subdisciminante de $V(\boldsymbol{L})$. Esto implica que

$$|\mathcal{A}_{\geq 2}^{nsq}| \leq d(d-1)^2(d-2)q^{d-m-2} \leq d^4 q^{d-m-2}. \tag{10.51}$$

En consecuencia, reemplazando (10.50), (10.51) en (10.49) deducimos la siguiente cota superior para la suma $\mathcal{S}_4^{nsq}$:

$$\mathcal{S}_4^{nsq} \leq c_3 M(d)\log(dq)\frac{d^3}{q} + c_4 M(d)\log(dq)\frac{d^6}{q^2}. \tag{10.52}$$

Combinando (10.48) y (10.52) obtenemos el siguiente resultado

Teorema 10.5.1. *Sea $q \geq d^4$. El costo en promedio del último paso del algoritmo clásico de factorización en $\mathcal{A}$ está acotado superiormente por*

$$E[X_4] \leq \tau_1 M(d) + \frac{cd^3 M(d)\log(dq)}{q},$$

donde c es una constante independiente de d y q.

El Teorema 10.5.1 muestra que el costo en promedio del último paso del algoritmo clásico de factorización aplicado a elementos de $\mathcal{A}$ es de $\mathcal{O}(M(d))$ operaciones aritméticas en $\mathbb{F}_q$, esto es, el costo de dividir dos polinomios de grado a lo sumo d.

Para finalizar este capítulo mostramos el costo en promedio del algoritmo clásico de factorización para familias lineales de polinomios. En las secciones anteriores analizamos la complejidad en promedio de cada paso de este algoritmo. En la siguiente tabla resumimos el costo de los tres pasos fundamentales. En la primera columna indicamos cada etapa del algoritmo, la segunda columna describe el costo del peor caso de la factorización de un polinomio en $\mathbb{F}_q[T]_d$ de acuerdo a [vzGG99], la tercera columna corresponde al costo en promedio del algoritmo clásico de factorización según el análisis de [FGP01] y la cuarta columna muestra los resultados de nuestro análisis en promedio del algoritmo clásico de factorización restringido a la familia $\mathcal{A}$ (ver Teoremas 10.2.2, 10.3.2 y 10.4.2).

Recordamos que $M(d) := d\log d \log\log d$ y $\mathcal{U}(d) := M(d)\log d$. Notamos también que, en (1), s es el máximo grado de los factores irreducibles del polinomio de entrada y que, en (2), los números k y j indican que el polinomio de entrada tiene j factores irreducibles de grado k. Observando este cuadro podemos concluir que las estimaciones del costo de los tres pasos fundamentales del algoritmo clásico de factorización en $\mathcal{A}$ mejoran las de [FGP01]. También cabe mencionar que nuestras técnicas nos permitieron dar cotas superiores explícitas para los costos de cada uno de los pasos del algoritmo. Por último, al igual que en el peor caso (ver [vzGG99, Theorem 14.14]), el costo en promedio del algoritmo clásico de factorización aplicado a los elementos de $\mathcal{A}$ es del orden de $\mathcal{O}(dM(d)\log(dq))$ operaciones aritméticas en $\mathbb{F}_q$, lo cual coincide con el del algoritmo DDF.

Cuadro 10.1: Comparación del costo del peor caso y del costo en promedio para los tres pasos fundamentales del algoritmo clásico de factorización

Pasos	Peor caso en $\mathbb{F}_q[T]_d$	Caso promedio en $\mathbb{F}_q[T]_d$	Caso promedio en $\mathcal{A}$
ERF	$\mathcal{O}(\mathcal{U}(d) + d\log(q/p))$	$\mathcal{O}(d^2)$	$\mathcal{O}(\mathcal{U}(d))$
DDF	$\mathcal{O}(sM(d)\log(dq))$ (1)	$\mathcal{O}(d^3 \log q)$	$\mathcal{O}(dM(d)\log(dq))$
EDF	$\mathcal{O}((k\log q + \log d)M(d)\log j)$ (2)	$\mathcal{O}(d^2 \log q)$	$\mathcal{O}(M(d)\log q)$

Bibliografía

[AR10] Y. Aubry y F. Rodier, *Differentially 4-uniform functions*, Arithmetic, geometry, cryptography and coding theory 2009, Contemp. Math., vol. 521, Amer. Math. Soc., Providence, RI, 2010, 1–8.

[Arw18] A. Arwin, *Ueber kongruenzen von dem funften und hoheren graden nach einem primzahl-modulus*, Ark. Mat. **14** (1918), 1–46.

[BBSR15] E. Bank, L. Bary-Soroker y L. Rosenzweig, *Prime polynomials in short intervals and in arithmetic progressions*, Duke Math. J. **164** (2015), no. 2, 277–295.

[BE16] E. Ballico y M. Elia, *On evaluating multivariate polynomials over finite fields*, Quaestiones Mathematicae **39** (2016), no. 1, 1–8.

[Ber67] E. R. Berlekamp, *Factoring polynomials over finite fields*, Bell System Tech. J. **46** (1967), 1853–1859.

[Ber68] E. R. Berlekamp, *Algebraic coding theory*, McGraw-Hill Book Co., New York, 1968.

[Ber70] E. R. Berlekamp, *Factoring polynomials over large finite fields*, Math. Comp. **24** (1970), 713–735.

[BES13] E. Ballico, M. Elia y M. Sala, *On the evaluation of multivariate polynomials over finite fields*, J. Symbolic Comput. **50** (2013), 255–262.

[BFS99] J. F. Buss, G. S. Frandsen y J. O. Shallit, *The computational complexity of some problems of linear algebra*, J. Comput. System Sci. **58** (1999), no. 3, 572–596.

[BO81] M. Ben-Or, *Probabilistic algorithms in finite fields*, Proceedings of the 22Nd Annual Symposium on Foundations of Computer Science (Washington, DC, USA), SFCS '81, IEEE Computer Society, 1981, 394–398.

[BP11] C. Beltrán y L. M. Pardo, *Fast linear homotopy to find approximate zeros of polynomial systems*, Found. Comput. Math. **11** (2011), no. 1, 95–129.

[BS59] B. Birch y H. Swinnerton-Dyer, *Note on a problem of Chowla*, Acta Arith. **5** (1959), no. 4, 417–423.

[Buc90] J. Buchmann, *Complexity of algorithms in algebraic number theory*, Number theory (Banff, AB, 1988), de Gruyter, Berlin, 1990, 37–53.

[Car55] L. Carlitz, *On the number of distinct values of a polynomial with coefficients in a finite field*, Proc. Japan Acad. **31** (1955), 119–120.

[CGH91] L. Caniglia, A. Galligo y J. Heintz, *Equations for the projective closure and effective Nullstellensatz*, Discrete Appl. Math. **33** (1991), 11–23.

[CLO92] D. Cox, J. Little y D. O'Shea, *Ideals, varieties, and algorithms: an introduction to computational algebraic geometry and commutative algebra*, Undergrad. Texts Math., Springer, New York, 1992.

[CM06a] A. Cafure y G. Matera, *Fast computation of a rational point of a variety over a finite field*, Math. Comp. **75** (2006), no. 256, 2049–2085.

[CM06b] A. Cafure y G. Matera, *Improved explicit estimates on the number of solutions of equations over a finite field*, Finite Fields Appl. **12** (2006), no. 2, 155–185.

[CM07] ———, *An effective Bertini theorem and the number of rational points of a normal complete intersection over a finite field*, Acta Arith. **130** (2007), no. 1, 19–35.

[CMP15a] A. Cafure, G. Matera y M. Privitelli, *Polar varieties, Bertini's theorems and number of points of singular complete intersections over a finite field*, Finite Fields Appl. **31** (2015), 42–83.

[CMP15b] E. Cesaratto, G. Matera y M. Pérez, *The distribution of factorization patterns on linear families of polynomials over a finite field*, Preprint `arXiv:1408.7014 [math.NT]`, to appear in Combinatorica, 2015.

[CMPP14] E. Cesaratto, G. Matera, M. Pérez y M. Privitelli, *On the value set of small families of polynomials over a finite field, I*, J. Combin. Theory Ser. A **124** (2014), 203–227.

[Coh70] S. D. Cohen, *The distribution of polynomials over finite fields*, Acta Arith. **17** (1970), 255–271.

[Coh72] ———, *Uniform distribution of polynomials over finite fields*, J. London Math. Soc. (2) **6** (1972), 93–102.

[Coh73] S. D. Cohen, *The values of a polynomial over a finite field*, Glasgow Math. J. **14** (1973), 205–208.

[Coh94] S. D. Cohen, *Polynomial factorisation, graphs, designs and codes*, Finite fields: theory, applications, and algorithms (Las Vegas, NV, 1993), Contemp. Math., vol. 168, Amer. Math. Soc., Providence, RI, 1994, 23–32.

[Coh98] ———, *Polynomial factorisation and an application to regular directed graphs*, Finite Fields Appl. **4** (1998), no. 4, 316–346.

[Col79] G. E. Collins, *Factoring univariate integral polynomials in polynomial average time*, Symbolic and algebraic computation (EUROSAM '79, Internat. Sympos., Marseille, 1979), Lecture Notes in Comput. Sci., vol. 72, Springer, Berlin-New York, 1979, 317–329.

[Cou16] A. Couvreur, *An upper bound on the number of rational points of arbitrary projective varieties over finite fields*, Proc. Amer. Math. Soc. **144** (2016), no. 9, 3671–3685.

[CR88] B. Chor y R. L. Rivest, *A knapsack-type public key cryptosystem based on arithmetic in finite fields*, IEEE Trans. Inform. Theory **34** (1988), no. 5, 901–909.

[CU57] L. Carlitz y S. Uchiyama, *Bounds for exponential sums*, Duke Math. J. **24** (1957), 37–41.

[CW10] Q. Cheng y D. Wan, *Complexity of decoding positive-rate primitive Reed-Solomon codes*, IEEE Trans. Inform. Theory **56** (2010), no. 10, 5217–5222.

[CW12] ———, *A deterministic reduction for the gap minimum distance problem*, IEEE Trans. Inform. Theory **58** (2012), no. 11, 6935–6941.

[CZ81] D. G. Cantor y H. Zassenhaus, *A new algorithm for factoring polynomials over finite fields*, Math. Comp. **36** (1981), no. 154, 587–592.

[Dan94] V. Danilov, *Algebraic varieties and schemes*, Algebraic Geometry I (I. Shafarevich, ed.), Encyclopaedia of Mathematical Sciences, vol. 23, Springer, Berlin Heidelberg New York, 1994, 167–307.

[DT09] C. D'Andrea y L. F. Tabera, *Tropicalization and irreducibility of generalized Vandermonde determinants*, Proc. Amer. Math. Soc. **137** (2009), no. 11, 3647–3656.

[Eis95] D. Eisenbud, *Commutative algebra with a view toward algebraic geometry*, Grad. Texts in Math., vol. 150, Springer, New York, 1995.

[Fel68] W. Feller, *An introduction to probability theory and its applications. Vol. I*, 3rd ed., John Wiley & Sons, Inc., New York, 1968.

[Fel91] ———, *An introduction to probability theory and its applications*, vol. 2, Wiley, 1991.

[FGP01] P. Flajolet, X. Gourdon y D. Panario, *The complete analysis of a polynomial factorization algorithm over finite fields*, J. Algorithms **40** (2001), no. 1, 37–81.

[FHJ94] M. D. Fried, D. Haran y M. Jarden, *Effective counting of the points of definable sets over finite fields*, Israel J. Math. **85** (1994), no. 1-3, 103–133.

[Fin03] S. R. Finch, *Mathematical constants*, Encyclopedia of Mathematics and its Applications, vol. 94, Cambridge Univ. Press, Cambridge, 2003.

[FS84] M. Fried y J. Smith, *Irreducible discriminant components of coefficient spaces*, Acta Arith. **44** (1984), no. 1, 59–72.

[FS09] P. Flajolet y R. Sedgewick, *Analytic combinatorics*, Cambridge Univ. Press, Cambridge, 2009.

[Ful84] W. Fulton, *Intersection theory*, Springer, Berlin Heidelberg New York, 1984.

[Gal46] E. Galois, *Sur la théorie des nombres*, Jour. Math. Pures Appl. **11** (1846), 398–407.

[GCL92] K. O. Geddes, S. R. Czapor y G. Labahn, *Algorithms for computer algebra*, Kluwer Academic Publishers, Boston, MA, 1992.

[GG98] S. W. Golomb y P. Gaal, *On the number of permutations of n objects with greatest cycle length k*, Adv. Appl. Math. **20** (1998), no. 1, 98–107.

[GGL06] P. Gopalan, V. Guruswami y R. J. Lipton, *Algorithms for modular counting of roots of multivariate polynomials*, LATIN 2006: Theoretical informatics, Lecture Notes in Comput. Sci., vol. 3887, Springer, Berlin, 2006, 544–555.

[Gib98] C. G. Gibson, *Elementary geometry of algebraic curves: an undergraduate introduction*, Cambridge Univ. Press, Cambridge, 1998.

[GKP94] R. Graham, D. Knuth y O. Patashnik, *Concrete mathematics: a foundation for computer science*, 2nd ed., Addison–Wesley, Reading, Massachusetts, 1994.

[GKZ94] I. M. Gel′fand, M. M. Kapranov y A. V. Zelevinsky, *Discriminants, resultants, and multidimensional determinants*, Mathematics: Theory & Applications, Birkhäuser Boston, Inc., Boston, MA, 1994.

[GL02a] S. Ghorpade y G. Lachaud, *Étale cohomology, Lefschetz theorems and number of points of singular varieties over finite fields*, Mosc. Math. J. **2** (2002), no. 3, 589–631.

[GL02b] S. R. Ghorpade y G. Lachaud, *Number of solutions of equations over finite fields and a conjecture of Lang and Weil*, Number theory and discrete mathematics (Chandigarh, 2000), Trends Math., Birkhäuser, Basel, 2002, 269–291.

[Har92] J. Harris, *Algebraic geometry: a first course*, Grad. Texts in Math., vol. 133, Springer, New York Berlin Heidelberg, 1992.

[Hei83] J. Heintz, *Definability and fast quantifier elimination in algebraically closed fields*, Theoret. Comput. Sci. **24** (1983), no. 3, 239–277.

[HH11] J. Herzog y T. Hibi, *Monomial ideals*, Grad. Texts in Math., vol. 260, Springer, London, 2011.

[HM11] F. Hernando y G. McGuire, *Proof of a conjecture on the sequence of exceptional numbers, classifying cyclic codes and APN functions*, J. Algebra **343** (2011), 78–92.

[HS82] J. Heintz y C.-P. Schnorr, *Testing polynomials which are easy to compute*, Logic and algorithmic (Zurich, 1980), Monograph. Enseign. Math., vol. 30, Univ. Genève, Geneva, 1982, 237–254.

[HW99] M.-D. Huang y Y.-C. Wong, *Solvability of systems of polynomial congruences modulo a large prime*, Comput. Complexity **8** (1999), no. 3, 227–257.

[Kem69] H. Kempfert, *On the factorization of polynomials*, J. Number Theory **1** (1969), 116–120.

[KK90a] A. Knopfmacher y J. Knopfmacher, *Counting polynomials with a given number of zeros in a finite field*, Linear Multilinear Algebra **26** (1990), no. 4, 287–292.

[KK90b] A. Knopfmacher y J. Knopfmacher, *The distribution of values of polynomials over a finite field*, Linear Algebra Appl. **134** (1990), 145–151.

[Knu98a] D. E. Knuth, *The art of computer programming. Vol. 2*, Addison-Wesley, Reading, MA, 1998, Seminumerical algorithms, Third edition.

[Knu98b] ———, *The art of computer programming. Vol. 3*, Addison-Wesley, Reading, MA, 1998, Sorting and searching, Second edition.

[Kun85] E. Kunz, *Introduction to commutative algebra and algebraic geometry*, Birkhäuser, Boston, 1985.

[KY08] S. Kopparty y S. Yekhanin, *Detecting rational points on hypersurfaces over finite fields*, Proceedings of the 23rd Annual IEEE Conference on Computational Complexity, CCC 2008, 23-26 June 2008, College Park, Maryland, USA, 2008, 311–320.

[Lac96] G. Lachaud, *Number of points of plane sections and linear codes defined on algebraic varieties*, Arithmetic, geometry and coding theory (Luminy, 1993), de Gruyter, Berlin, 1996, 77–104.

[Len91] H. W. Lenstra, Jr., *On the Chor-Rivest knapsack cryptosystem*, J. Cryptology **3** (1991), no. 3, 149–155.

[LLL82] A. K. Lenstra, H. W. Lenstra, Jr. y L. Lovász, *Factoring polynomials with rational coefficients*, Math. Ann. **261** (1982), no. 4, 515–534.

[LN83] R. Lidl y H. Niederreiter, *Finite fields*, Addison–Wesley, Reading, Massachusetts, 1983.

[Lov89] L. Lovász, *Singular spaces of matrices and their application in combinatorics*, Bol. Soc. Brasil. Mat. (N.S.) **20** (1989), no. 1, 87–99.

[LP81] V. Lifschitz y B. Pittel, *The number of increasing subsequences of the random permutation*, J. Combin. Theory Ser. A **31** (1981), no. 1, 1–20.

[LP02] A. Lascoux y P. Pragacz, *Jacobians of symmetric polynomials*, Ann. Comb. **6** (2002), no. 2, 169–172.

[LR15] G. Lachaud y R. Rolland, *On the number of points of algebraic sets over finite fields*, J. Pure Appl. Algebra **219** (2015), no. 11, 5117–5136.

[LW10] J. Li y D. Wan, *A new sieve for distinct coordinate counting*, Sci. China Math. **53** (2010), no. 9, 2351–2362.

[Mat10] G. Matera, *The computation of rational solutions of polinomial systems over a finite field*, LVII Jornadas de Matemática Discreta y Algorítmica (Santander, Spain) (D. S. et al., ed.), 2010, 9–33.

[MP13] G. Mullen y D. Panario, *Handbook of finite fields*, CRC Press, Boca Raton, FL, 2013.

[MPP14] G. Matera, M. Pérez y M. Privitelli, *On the value set of small families of polynomials over a finite field, II*, Acta Arith. **165** (2014), no. 2, 141–179.

[MPP16a] ———, *On the value set of small families of polynomials over a finite field,III*, Contemporary Developments in Finite Fields and Their Applications, 217–243, World Scientific Press, 2016.

[MPP16b] G. Matera, M. Pérez y M. Privitelli, *On the computation of rational points of a hypersurface over a finite field*, 2016, Disponible en `http://arxiv.org/abs/1504.06512`.

[MS77] F. J. MacWilliams y N. J. A. Sloane, *The theory of error-correcting codes. I*, North-Holland Publishing Co., Amsterdam-New York-Oxford, 1977, North-Holland Mathematical Library, Vol. 16.

[Odl85] A. M. Odlyzko, *Discrete logarithms in finite fields and their cryptographic significance*, Advances in cryptology (Paris, 1984), Lecture Notes in Comput. Sci., vol. 209, Springer, Berlin, 1985, 224–314.

[Pol13] P. Pollack, *Irreducible polynomials with several prescribed coefficients*, Finite Fields Appl. **22** (2013), 70–78.

[Rod09] F. Rodier, *Borne sur le degré des polynômes presque parfaitement nonlinéaires*, Arithmetic, geometry, cryptography and coding theory, Contemp. Math., vol. 487, Amer. Math. Soc., Providence, RI, 2009, 169–181.

[Ser66] J.-A. Serret, *Cours d'algèbre supérieure*, vol. 1512, Gauthier-Villars, París, 1866.

[Ser91] J.-P. Serre, *Lettre à M. Tsfasman*, Astérisque **198-200** (1991), 351–353.

[Sha94] I. Shafarevich, *Basic algebraic geometry: Varieties in projective space*, Springer, Berlin Heidelberg New York, 1994.

[Sho90] V. Shoup, *On the deterministic complexity of factoring polynomials over finite fields*, Inform. Process. Lett. **33** (1990), no. 5, 261–267.

[Sho95] ———, *A new polynomial factorization algorithm and its implementation*, J. Symbolic Comput. **20** (1995), no. 4, 363–397.

[Sho05] ———, *A computational introduction to number theory and algebra*, Cambridge Univ. Press, Cambridge, 2005.

[Shp99] I. E. Shparlinski, *Finite fields: Theory and computation*, Mathematics and its Applications, vol. 477, Kluwer Academic Publishers, Dordrecht, 1999.

[Sid94] V. Sidelnikov, *Decoding Reed–Solomon codes beyond* $(d-1)/2$ *and zeros of multivariate polynomials*, Probl. Inf. Transm. **30** (1994), no. 1, 44–59.

[SL66] L. A. Shepp y S. P. Lloyd, *Ordered cycle lengths in a random permutation*, Trans. Amer. Math. Soc. **121** (1966), 340–357.

[Ste87] S. Stepanov, *The number of irreducible polynomials of a given form over a finite field*, Math. Notes **41** (1987), 165–169.

[Uch54] S. Uchiyama, *Sur le nombre des valeurs distinctes d'un polynôme à coefficients dans un corps fini*, Proc. Japan Acad. **30** (1954), 930–933.

[Uch55a] ———, *Note on the mean value of* $V(f)$, Proc. Japan Acad. **31** (1955), 199–201.

[Uch55b] ———, *Note on the mean value of* $V(f)$. *II*, Proc. Japan Acad. **31** (1955), 321–323.

[Uch56] ———, *Note on the mean value of* $V(f)$. *III*, Proc. Japan Acad. **32** (1956), 97–98.

[vL92] J. H. van Lint, *Introduction to coding theory*, second ed., Springer-Verlag, Berlin, 1992.

[Vog84] W. Vogel, *Results on Bézout's theorem*, Tata Inst. Fundam. Res. Lect. Math., vol. 74, Tata Inst. Fund. Res., Bombay, 1984.

[vzG08] J. von zur Gathen, *Counting reducible and singular bivariate polynomials*, Finite Fields Appl. **14** (2008), no. 4, 944–978.

[vzGG99] J. von zur Gathen y J. Gerhard, *Modern computer algebra*, Cambridge University Press, New York, 1999.

[vzGP01] J. von zur Gathen y D. Panario, *Factoring polynomials over finite fields: a survey*, J. Symbolic Comput. **31** (2001), no. 1-2, 3–17.

[vzGS92] J. von zur Gathen y V. Shoup, *Computing Frobenius maps and factoring polynomials*, Comput. Complexity **2** (1992), no. 3, 187–224.

[vzGSS03] J. von zur Gathen, I. Shparlinski y A. Sinclair, *Finding points on curves over finite fields*, SIAM J. Comput. **32** (2003), no. 6, 1436–1448.

[vzGVZ13] J. von zur Gathen, A. Viola y K. Ziegler, *Counting reducible, powerful, and relatively irreducible multivariate polynomials over finite fields*, SIAM J. Discrete Math. **27** (2013), no. 2, 855–891.

[Wei48] A. Weil, *Sur les courbes algébriques et les varietés qui s'en déduisent*, Hermann, Paris, 1948.

[Zas69] H. Zassenhaus, *On Hensel factorization. I*, J. Number Theory **1** (1969), 291–311.