Tesis de Posgrado

Un Algoritmo efectivo para el Teorema de los Ceros de Hilbert

Sabia, Juan Vicente Rafael

1992

Tesis presentada para obtener el grado de Doctor en Ciencias Matemáticas de la Universidad de Buenos Aires

Universidad de Buenos Aires

Facultad de Ciencias Exactas y Naturales

Tema de Tesis

Un Algoritmo Efectivo para el Teorema de los Ceros de Hilbert

Autor

Juan Vicente Rafael Sabia

Director

Joos Heintz

Lugar de trabajo

Departamento de Matemática

Tesis presentada para optar al título de

Doctor en Ciencias Matemáticas

1992

Agradecimiento

A Joos Heintz, Teresa Krick, Pablo Solernó y Susana Tesauri.

Resumen

En esta tesis se demuestra el siguiente teorema de los ceros de Hilbert efectivo:

Sea k un cuerpo infinito y perfecto, sean $X_1,...,X_n$ indeterminadas sobre k y sean $f_1,...,f_s$ polinomios en $k[X_1,...,X_n]$ de grado acotado por un número dado d, que satisface $d \geq n$. Entonces existe una red aritmética sobre k de tamaño $s^{O(1)} d^{O(n)}$ y profundidad $O(n^4 \log^2 sd)$ que decide si el ideal generado por $f_1,...,f_s$ en $k[X_1,...,X_n]$ es trivial y, si este es el caso, produce un cálculo de evaluación de longitud $s^{O(1)} d^{O(n)}$ y profundidad $O(n^4 \log^2 sd)$ en el cuerpo de funciones $k(X_1,...,X_n)$ que calcula polinomios $p_1,...,p_s$ de $k[X_1,...,X_n]$ de grado $d^{O(n^2)}$ que satisfacen

$$1 = \sum_{1 \leq j \leq s} p_j f_j .$$

Introducción

Unos de los teoremas básicos que conecta la Geometría Algebraica con el Algebra Conmutativa es es Teorema de los Ceros de Hilbert:

Sean $f_1,...,f_s$ polinomios en $k[X_1,...,X_n]$ y sea $(f_1,...,f_s)$ el ideal generado por ellos en $k[X_1,...,X_n]$. Si $V(f_1,...,f_s)$ nota a la variedad algebraica definida por $f_1,...,f_s$ en $\mathbb{A}^n(\bar{k})$ (donde $\bar{k}$ es una clausura algebraica de k), son equivalentes:

(i) $1 \in (f_1,...,f_s)$
(ii) $V(f_1,...,f_s) = \emptyset$.

La versión algorítmica de este teorema consiste, dados los polinomios $f_1,...,f_s$, en decidir si la variedad que ellos definen es vacía y, si éste es el caso, encontrar polinomios $p_1,...,p_s$ en $k[X_1,...,X_n]$ que satisfagan la identidad

$$1 = \sum_{1 \le j \le s} p_j f_j .$$

Existen varios trabajos al respecto (ver, por ejemplo, [Brownawell, 1987], [Caniglia-Galligo-Heintz, 1988 y 1989] y [Kollár, 1988]). Para una reseña de métodos y resultados en este tema puede remitirse a [Teissier, 1991] (o también a [Fitchas-Galligo, 1990] y a [Caniglia *et al.*, 1991] para ver técnicas elementales y aplicaciones a la computación).

El teorema principal de esta tesis fue señalado por E. Kaltofen como un problema abierto en 1988 y es motivado por el siguiente problema: la versión efectiva del Teorema de los Ceros de Hilbert de [Kollár, 1988] (ver también [Fitchas-Galligo, 1990], Théorème 1) tiene, para $d \ge 3$ y $n \ge 2$ el siguiente enunciado:

El ideal $(f_1,...,f_s)$ es trivial si y sólo si existen polinomios $p_1,...,p_s \in k[X_1,...,X_n]$ tales que $1 = \sum_{1 \le j \le s} p_j f_j$ y máx $(\deg p_j f_j;\ 1 \le j \le s) \le d^n$.

De un ejemplo muy conocido descubierto por Mora-Lazard e, independientemente,por Masser-Philippon (ver [Brownawell, 1987]) se deduce que la cota del grado, simplemente exponencial, d^n de la versión de Kollár es optimal.

Si se usa este teorema para el diseño de un algoritmo que decida la trivialidad del ideal $(f_1,...,f_s)$, se obtiene un algoritmo bien paralelizable uniforme de complejidad secuencial $s^{O(1)}d^{O(n^2)}$(ver [Dickenstein *et al.*, 1988]). Debe mencionarse que esta es aún el mejor resultado de complejidad uniforme conocido que incluye el problema de decidir si un sistema de ecuaciones polinomiales puede resolverse en un cuerpo algebraicamente cerrado.

El algoritmo mencionado anteriormente no sólo decide si el ideal $(f_1,...,f_s)$ es trivial sino que también da polinomios $p_1,...,p_s \in k[X_1,...,X_n]$ de grado acotado por d^n que satisfacen $1 = \sum_{1 \le j \le s} p_j f_j$ si tal relación existe. Estos polinomios se expresan en representación densa que es de tamaño $O(sd^{n^2})$. Este tamaño no puede ser reducido a causa del ejemplo de Mora-Lazard y Masser-Philippon. Esto significa que la aproximación al problema de trivialidad algorítmico para ideales polinomiales por medio de los teoremas de los ceros de Algebra Conmutativa conduce, necesariamente, a cotas de complejidad que son simplemente exponenciales en el tamaño de la entrada que es del orden de $O(sd^n)$. La única forma de evitar este problema es renunciar a exhibir la relación $1 = \sum_{1 \le j \le s} p_j f_j$ si el ideal es trivial o bien cambiar la estructura de datos que representa a los polinomios $p_1,...,p_s$. La primera alternativa se trata en [Giusti-Heintz, 1991 b] por medio de un algoritmo bien paralelizable con complejidad secuencial no uniforme de

orden $s^{O(1)}d^{O(n)}$ (ver sección 2.4., Proposición A). La segunda alternativa es el tema de esta tesis: se representan las salidas $p_1,...,p_s$ por medio de un cálculo de evaluación de longitud $s^{O(1)}d^{O(n)}$ en $k(X_1,...,X_n)$ (esto significa que la longitud del cálculo de evaluación que representa la salida $p_1,...,p_s$ es polinomial en el tamaño de la entrada $f_1,...,f_s$). La complejidad paralela que se obtiene aquí es del orden $O(n^4 \log^2 sd)$, y es la misma que la que se obtiene usando representación densa de polinomios y la mejor cota de d^n para $d \geq 3$ y $n \geq 2$ en el teorema de los ceros efectivo del Algebra Conmutativa (ver por ejemplo [Fitchas-Galligo, 1990], Théorème 1). Como se dijo anteriormente esta última técnica da una complejidad secuencial de $s^{O(1)}d^{O(n^2)}$ que es simplemente exponencial en el tamaño de la entrada.

Combinando el resultado principal de esta tesis con [Valiant *et al*., 1983], Theorem 2, se puede mejorar la complejidad paralela mencionada a $O(n^3 \log^2 sd)$. Sin embargo, este procedimiento hace que la complejidad secuencial simultánea aumente a $s^{O(1)}d^{O(n^2)}$.

1. Definiciones y notaciones

1.1 Definiciones y notaciones básicas.

Sea k un cuerpo infinito y perfecto. Se supone que k es *efectivo*, es decir que las operaciones aritméticas básicas de k (suma, resta, multiplicación, división y comparación de elementos) son realizables por algoritmos. Sea $\bar{k}$ una clausura algebraica de k. Se notará $\mathbf{A}^n := \mathbf{A}^n(\bar{k})$ al espacio afín n-dimensional sobre $\bar{k}$, munido de la topología de Zariski y de su anillo de coordenadas de funciones polinomiales. Cuando la característica de k sea positiva (p := car k) se supondrá que la extracción de raíces p-ésimas es efectiva, o sea realizable por un algoritmo.

Sean $X_1,...,X_n$ indeterminadas sobre k. El grado total de un polinomio f a coeficientes en k se notará deg f y el grado parcial con respecto a la variable X_i se notará $\deg_{X_i} f$. Un polinomio f se dirá *mónico* en X_i si su grado en X_i es igual a su grado total y éste a la vez es positivo.

En lo que sigue, $f_1,...,f_s$ será un sistema finito de polinomios no constantes en $k[X_1,...,X_n]$. El ideal generado por dichos polinomios en $k[X_1,...,X_n]$ se notará $I(f_1,...,f_s)$, $(f_1,...,f_s)$ o más simplemente I. En caso de quedar claro en el contexto también se usará esta notación para el ideal generado por $f_1,...,f_s$ en extensiones del anillo $k[X_1,...,X_n]$. La variedad algebraica afín, en el sentido clásico, definida por $(f_1,...,f_s)$ en $\mathbf{A}^n$ se notará $V(f_1,...,f_s)$, V(I) ó más simplemente V cuando no haya lugar a dudas. En las condiciones antedichas d siempre notará a un número natural que satisfaga $d \geq \max\{ n, \deg f_j ; 1 \leq j \leq s\}$.

También se considerarán invariantes geométricos intrínsecos a $V(f_1,...,f_s)$, esencialmente la dimensión y el grado. Sea

$$V = \bigcup_{1 \leq k \leq N} C_k$$

la descomposición de V en componentes irreducibles. La *dimensión* (dim C) de una componente irreducible $C := C_k$, $1 \leq k \leq N$, se identifica con la dimensión de Krull de su anillo de coordenadas, mientras que su *grado* (deg C) con el número de puntos que hay en la intersección cuando se corta a C con (dim C) hiperplanos genéricos. Se define entonces la dimensión de V como

$$\dim V := \max \{ \dim C_k, 1 \leq k \leq N\}$$

y su grado por

$$\deg V := \sum_{1 \leq k \leq N} \deg C_k$$

Esta noción no convencional de grado tiene la ventaja de satisfacer una desigualdad de Bezout del tipo

$$\deg V \leq \prod_{1 \leq k \leq N} \deg f_i$$

sin restricción sobre el tipo de intersección (ver [Heintz, 1983]).

1.2 Estructuras de datos, modelos algorítmicos y de complejidad.

Los algoritmos considerados admitirán como entradas (*input*) conjuntos finitos de polinomios. Es necesario describir tal conjunto por una estructura de datos, medir su longitud y precisar el tipo de algoritmos que se van a considerar.

1.2.1 Representación "ingenua" de entradas-salidas y su longitud.

Sean $f_1,...,f_s$ el conjunto de polinomios ya introducido en 1.1. Recordemos que sus grados totales están acotados superiormente por un número prefijado d que también mayora a n. Cada polinomio puede ser escrito en *representación densa*, es decir, prefijando un orden a todos los monomios de grado menor o igual que d y representando al polinomio por un vector con entradas en k, donde cada coordenada será el coeficiente del monomio correspondiente. Por ejemplo, si $d = 2$, $n = 2$ y el orden de monomios prefijado es $1, X_1, X_2, X_1^2, X_1X_2, X_2^2$, el polinomio $3 + 2X_1 + 5X_1X_2 + X_2^2 \in \mathbb{Q}[X_1, X_2]$ se representará por el vector $(3, 2, 0, 0, 5, 1)$. De esta forma, cada polinomio podrá ser escrito como un vector con coordenadas en k y su longitud dependerá de los parámetros n, d y s.

Existe un caso de particular importancia: $k := \mathbb{Q}$. En esta situación también se puede tener en cuenta la *longitud aritmética* t, es decir el supremo del número de bits necesarios para escribir cualquier coeficiente de los polinomios f_i. Dicho de otra forma, el máximo de los valores absolutos de los numeradores y denominadores de los coeficientes de todos los f_i es mayorado por 2^t.

La longitud de la entrada $f_1,...,f_s$ es la cantidad de memoria necesaria para guardarlos, es decir, la cantidad de memoria que ocuparán sus coeficientes. Se puede estimar de antemano la longitud de la entrada a partir de los parámetros d , n , s y eventualmente t en el caso k : = $\mathbb{Q}$. El número de monomios en n variables de grado a lo sumo d es el coeficiente binomial $\binom{d+n}{n}$, cantidad mayorada por e d^n, que es por lo tanto un $O(d^n)$ (n fijo, d tendiendo al infinito). Como se convino codificar a los polinomios de entrada por su coeficientes en la representación densa, harán falta $O(sd^n)$ coeficientes y, en el caso k : = $\mathbb{Q}$, demandará $O(std^n)$ bits.

Las salidas (*output*) podrán ser de diverso tipo: valores booleanos, números enteros entre -1 y n (representados por vectores de valores booleanos), matrices cuadradas de dimensión n x n a coeficientes en k o bien polinomios. Uno de los problemas centrales es decidir cómo codificar los polinomios de salida. La representación densa para los polinomios de salida es incompatible con el orden de complejidad buscada, por lo tanto se utilizará una forma de representar polinomios más ventajosa que será introducida más adelante.

1.2.2 Descripción de los modelos de algoritmos y de complejidad y codificación de polinomios de salida.

Los algoritmos que se van a utilizar serán, en principio, descriptos por una *red aritmética* con entradas en k , representada por un grafo orientado acíclico ([von zur Gathen, 1986]). A cada vértice o nodo interno le corresponde un procesador que efectúa una operación elemental del cuerpo de base k (ya sea una operación aritmética, incluyendo la eventual extracción de raíces p-ésimas; una operación booleana que corresponde a la lógica proposicional o bien selectores asociados a control de igualdad de

elementos de k) y cada arista indica el envío de una salida de un procesador como entrada de otro. Los vértices o nodos externos del grafo representan las entradas y las salidas de la red.

Todo algoritmo admite un desarrollo secuencial o paralelo. La *complejidad secuencial* (o *tiempo secuencial*) es la longitud de la red, es decir el número de procesadores o vértices del grafo. La *complejidad paralela* (o *tiempo paralelo*) es la profundidad de la red, es decir la longitud del camino más largo en el grafo orientado.

Si el cuerpo de base es el de los racionales, cada procesador aritmético se transforma en un circuito booleano donde los procesadores manipulan ahora bits. Es necesario entonces, tener en cuenta el crecimiento eventual de los coeficientes de los polinomios intermedios. De esta forma, la red aritmética se transforma naturalmente en una red booleana, a la que se le puede asignar análogamente nociones de complejidad secuencial y paralela.

Resumiendo, los algoritmos serán familias de redes aritméticas parametrizadas por las cantidades d , s y n (o de redes booleanas parametrizadas por d , n , s y t). Para una discusión más profunda sobre este modelo de complejidad, pueden verse [von zur Gathen, 1986] y [Fitchas-Galligo-Morgenstern, 1990].

Existen redes aritméticas particulares especialmente interesantes: aquellas que no hacen intervenir comparaciones ni ramas. Se llamarán *cálculos de evaluación*, *circuitos aritméticos* o *straight line programs*. Un tal cálculo de evaluación en $k(X_1,...,X_n)$ será un vector $b = (Q_1,...,Q_L)$ que cumpla las siguientes condiciones

(i) $Q_\rho \in k(X_1,...,X_n) \ \forall \ 1 \leq \rho \leq L$

(ii) Se cumplen alguna de estas tres posibilidades:

a) $Q_\rho \in (X_1,...,X_n)$

b) $Q_\rho \in k$

c) Existen $\tau, \sigma \leq \rho$ tales que $Q_\rho = Q_\tau * Q_\sigma$ dónde $* \in (+, -, ., \%)$.

Para mayor precisión sobre estas redes ver [Strassen, 1972], [von zur Gathen, 1986], [Stoss, 1989] o [Heintz, 1989].

La complejidad secuencial o longitud de un cálculo de evaluación será el número de operaciones aritméticas que contiene y, de ser necesario, se contarán también las variables o constantes introducidas. . Estos cálculos de evaluación permiten codificar polinomios de muchas variables, calculando su valor en un punto de $\mathbf{A}^n$. Esta será, en general, la forma de codificar los polinomios de salida.

Por ejemplo, sea $f \in k[X_1,...,X_n]$, $f = (X_1 + ... + X_n)^d$. Si se quiere codificar a f en forma densa, será necesario un vector de $(d + n)! / d!\, n!$ coordenadas, número que, ya se vio, es del orden de $O(d^n)$. Sin embargo, mediante un cálculo de evaluación bastará un circuito aritmético de $2n + d - 2$ nodos (n para representar las variables, n - 1 para las sumas y d - 1 para los productos) y este número es del orden $O(d)$. Entonces, los polinomios de salida serán generalmente dados por programas de evaluación.

Una última observación debe ser hecha para los cálculos de evaluación. Si algún cálculo contiene divisiones, el polinomio que represente no podrá ser evaluado en cualquier punto, porque se corre el riesgo de tener que dividir por cero. Sin embargo, la totalidad de los puntos en donde un cálculo de evaluación no puede ser evaluado siempre forman un cerrado de Zariski de $\mathbf{A}^n$ que no es todo el espacio y por lo tanto el cálculo de evaluación representará unívocamente al polinomio en cuestión.

1.3 Clases de complejidad

En la situación ya definida, diremos que un algoritmo es de complejidad secuencial *polinomial en la longitud de la entrada* si la red correspondiente de parámetros (d , n , s) (respectivamente (d , n , s , t) en el caso booleano) admite una complejidad secuencial del tipo $s^{O(1)}\ d^{O(n)}$ (respectivamente $(st)^{O(1)}\ d^{O(n)}$). Esta terminología se justifica cuando se considera la longitud $O(sd^n)$ (respectivamente $O(std^n)$) de la entrada $f_1,...,f_s$ para la estructura de datos elegida (la representación densa de polinomios).

Se dirá que el algoritmo es *bien paralelizable* si la profundidad de la red es del tipo $O(n^2 \log^2 (sd))$ (respectivamente $O(n^2 \log^2 (std))$), es decir, si la complejidad paralela del algoritmo es del orden del cuadrado del logaritmo de la complejidad secuencial, tomándose siempre logaritmo en base 2.

2. Algunas herramientas

2.1 Extensiones del cuerpo de base

Una idea fundamental consistirá en introducir parámetros auxiliares en forma de nuevas indeterminadas, por ejemplo $T_1,...,T_n$. Se reemplazará entonces provisoriamente el cuerpo de base k por $k[T_1,...,T_n]$. Los resultados intermedios representarán a los polinomios considerados como dependientes de variables principales a coeficientes en los polinomios en los parámetros $T_1,...,T_n$. Con respecto a las variables principales, los polinomios son codificados por la representación densa, pero los polinomios coeficientes son representados por cálculos de evaluación en $k(T_1,...,T_n)$.

Los algoritmos ejecutarán, entonces, operaciones aritméticas (en general sin divisiones) y comparaciones en $k[T_1,...,T_n]$. El punto esencial para la complejidad de esta nueva aritmética residirá en las comparaciones y, más precisamente, en los tests de no nulidad. Por ejemplo, si los parámetros auxiliares distintos de las variables de origen $X_1,...,X_n$ aparecen aún en los polinomios finales, deberán ser eliminados por especialización en valores apropiados de k , teniendo en cuenta de no anularlos. Así, todos los algoritmos podrán ser realizados por redes sobre el cuerpo k (ver [Heintz-Sieveking, 1981] y [Kaltofen, 1988] para la utilización de esta representacion de polinomios en cálculo formal).

A menudo será imprescindible dividir las indeterminadas $X_1,...,X_n$ en "parámetros", es decir $X_1,...,X_i$, y variables principales distinguidas, es decir $X_{i+1},...,X_n$ para un i dado , $1 \leq i \leq n$. Los algoritmos producirán elementos de $k(X_1,...,X_i)[X_{i+1},...,X_n]$ que serán considerados como polinomios en las variables principales $X_{i+1},...,X_n$ con coeficientes en el cuerpo $k(X_1,...,X_i)$. Después de una transformación adecuada de las variables principales, estos polinomios nunca contendrán más de dos de ellas. Además, el grado de estos polinomios será del orden de $d^{O(n)}$ o de $d^{O(n^2)}$, de acuerdo al contexto. Serán representados, o al menos representables por cálculos de evaluación en $k(X_1,...,X_i)[X_{i+1},...,X_n]$ de longitud $s^{O(1)}\ d^{O(n)}$, cuyas divisiones eventuales pueden ser ejecutadas en $k(X_1,...,X_i)$. La complejidad paralela de estos cálculos de evaluación será, en general, del tipo $O(n^2 \log^2(sd))$ (y en este sentido serán "bien paralelizables"). Por lo tanto, se puede pensar que los polinomios producidos por el algoritmo dados en representación densa con respecto a las dos variables que realmente aparecen. De esta forma, podrán ser dados como vectores de funciones racionales de $k(X_1,...,X_i)$ (o de $k[X_1,...,X_i]$ si no aparecen divisiones). Estas funciones racionales son representadas como resultado de un cálculo de evaluación en $k(X_1,...,X_i)$ (o en $k[X_1,...,X_i]$) generalmente bien paralelizable de longitud $s^{O(1)}\ d^{O(n)}$.

En el caso de característica $p > 0$, las extracciones de raíces p-ésimas aparecerán sólo en subrutinas sin efecto en los resultados finales. El número de iteraciones de este proceso está acotado a priori por un número entero m conocido del tipo O (n log d). Se pueden introducir nuevas indeterminadas $S_1,...,S_i$ y reemplazar los parámetros $X_1,...,X_i$ por $S_1^{p^m},...,S_i^{p^m}$. Los pasos algorítmicos que deben ser ejecutados en $k(X_1,...,X_i)$ son simulados en $k(S_1,...,S_i)$, donde las extracciones necesarias de raíces p-ésimas pueden ser efectuadas sin pérdidas en la complejidad (aquí es donde la suposición k perfecto se convierte en esencial). Como la salida final de la subrutina que requiere extracción de raíces p-ésimas será una transformacion k-lineal de las variables principales $X_{i+1},...,X_n$, la introducción momentánea de parámetros auxiliares $S_1,...,S_i$ no tiene influencia sobre el comportamiento global del algoritmo, en particular de su salida.

2.2 El teorema de Heintz-Schnorr y sus consecuencias para la complejidad

Se utilizará de manera esencial el siguiente teorema de Heintz- Schnorr ([Heintz-Schnorr, 1982], theorem 4.4):

Teorema : Se considera el conjunto $W(D,n,v)$ de polinomios de $k[T_1,...,T_n]$ de grado menor o igual que D que pueden ser evaluados por un cálculo de evaluación de longitud a lo sumo v. Sea Γ un subconjunto de k de cardinal $2v\,(1+D)^2$. Entonces existe un subconjunto $Q\,(D,n,v,\Gamma) = \{\gamma_1,...,\gamma_m\}$ de Γ^n (donde $m = 6\,(v+n)\,(v+n+1)$) que verifica la propiedad siguiente: todo polinomio de $W(D,n,v)$ que se anula sobre $\{\gamma_1,...,\gamma_m\}$ es idénticamente nulo.

Un subconjunto como en el teorema se llamara (siguiendo la terminología usada por [Henry-Merle, 1987] *cuestor* o bien *correct test sequence* .

Aplicado a la situación precedente, donde D será del tipo $s\, d^{O(n)}$ ó $d^{O(n)}$ y v, $s^{O(1)}\, d^{O(n)}$, esto posibilitará efectuar las comparaciones necesarias de elementos en $k[T_1,...,T_n]$ ejecutando solamente $s^{O(1)}\, d^{O(n)}$ operaciones aritméticas en k, ya que el cardinal de Γ y el del conjunto cuestor son del orden $s^{O(1)}\, d^{O(n)}$.

El problema de la elección del conjunto cuestor en Γ^n es esencial para la evaluación de la complejidad. Dicha elección podría hacerse paso a paso algorítmicamente, pero da una cota elevada que depende sobre todo del parámetro n. Como esta elección es independiente del problema en sí, para n, s y d fijos, se pensará que se tiene el conjunto cuestor por medio de un preprocesamiento, donde el costo no interviene en el cálculo particular. Dicho de otra forma, para cada terna d, s, n se construye una red aritmética que resuelve un cierto trabajo en tiempo secuencial $s^{O(1)}\, d^{O(n)}$ y en tiempo paralelo $O(n^2 \log^2(sd))$ pero el costo de esta construcción no se contará. Se procederá de manera análoga para los parámetros d, s, t, n y las redes boleanas en el caso $k := \mathbb{Q}$ (ver [von zur Gathen, 1986]). En este sentido, se dirá que los algoritmos no son uniformes.

En el tratamiento algorítmico completo, la elección de los cuestores puede ser hecha de forma aleatoria (ver [Heintz-Schnorr, 1982], Theorem 4.4). El tiempo secuencial de desarrollo de los algoritmos será una variable aleatoria con una esperanza del tipo $s^{O(1)}\, d^{O(n)}$. La cota superior de complejidad que se obtiene (*worst case complexity*) es del tipo $s^{O(1)}\, d^{O(n^2)}$.

2.3 Algebra lineal efectiva (Berkowitz-Mulmuley)

Las técnicas de álgebra lineal en las que se basan los algoritmos utilizan algoritmos bien paralelizables y sin divisiones. El ingrediente fundamental es

el el algoritmo de Berkowitz ([Berkowitz, 1984]) que calcula en tiempo polinomial todos los coeficientes del polinomio característico de una matriz cuadrada a coeficientes en un dominio íntegro. Estos coeficientes del polinomio característico se representan mediante un cálculo de evaluación sin divisiones.

Para calcular el rango de cualquier matriz, se combina el algoritmo de Berkowitz con un resultado de Mulmuley ([Mulmuley,1986]) que realiza el rango como la multiplicidad del cero en el polinomio característico de una matriz cuadrada auxiliar. Así, se puede decidir en tiempo polinomial por medio de un algoritmo bien paralelizable y sin efectuar divisiones, si un sistema de ecuaciones lineales admite soluciones y, en caso que la respuesta sea positiva, calcularlas efectuando una sola división por un elemento precalculado.

En la situación afín, hará falta calcular máximos comunes divisores de polinomios en una variable y a coeficientes en un anillo íntegro. En este caso, se siguen las líneas de [Giusti-Heintz, 1991 a] y la técnica de subresultante de [Brown,1971], combinándolas con el algoritmo de Berkowitz. Sólo en estas subrutinas la extracción de raíces p-ésimas (en el caso $p > 0$) es relevante.

2.4. Un resultado previo básico (Giusti-Heintz)

Los algoritmos aquí presentados se basan fundamentalmente en las técnicas y resultados de M. Giusti y J. Heintz ([Giusti-Heintz, 1991 b]); más específicamente, en las dos proposiciones siguientes.

Proposición A ([Giusti-Heintz, 1991 b], Théorème 3.5 y Théorème 3.7.2):
Existe una algoritmo bien paralelizable sin divisiones que calcula en tiempo no uniforme $s^{O(1)}\, d^{O(n)}$ los siguientes items:

(1) la dimensión $r := \dim V$ de la variedad algebraica definida por $f_1,...,f_s$ en $\mathbb{A}^n$.

(2) una matriz inversible $M \in k^{n \times n}$ tal que las variables $Y_1,...,Y_n$ que se obtienen transformando $X_1,...,X_n$ por medio de M, estan en posición de Noether con respecto a V. (Se dice que las variables $Y_1,...,Y_n$ están en posición de Noether con respecto a V si para cada $r < i \leq n$, donde $r := \dim V$, existe un polinomio de $k[Y_1,...,Y_r,Y_i]$ que es mónico en Y_i, y se anula sobre V.)

Proposición B ([Giusti-Heintz, 1991 b] Section 3.4.7 y Lemma 3.6):

Si $I = (f_1,...,f_s)$ es un ideal radical de dimensión cero (es decir que $V(I)$ es un conjunto finito no vacío), existe un algoritmo bien paralelizable sin divisiones que calcula en tiempo secuencial no uniforme $s^{O(1)} d^{O(n)}$ los siguientes items:

(1) Una matriz inversible $M \in k^{n \times n}$ que transforma las variables $X_1,...,X_n$ en nuevas variables $Y_1,...,Y_n$.

(2) Un elemento no nulo $\alpha \in k$ que es un polinomio de grado $d^{O(n)}$ en los coeficientes de $f_1,...,f_s$ en un subanillo finitamente generado de k.

(3) Polinomios $g_1,...,g_s$ en las variables $Y_1,...,Y_n$ con coeficientes en k tales que dichos coeficientes tienen las mismas propiedades que α con respecto a la entrada $f_1,...,f_s$.

El elemento α y los polinomios $g_1,...,g_s$ satisfacen las siguientes condiciones:

(i) $\max (\deg g_i ; 1 \leq i \leq n) \leq \deg V \leq d^n$.

(ii) los polinomios $g_1,...,g_s$ forman una sucesión regular de $k[X_1,...,X_n] - k[Y_1,...,Y_n]$

(iii) $g_1/\alpha ,..., g_n/\alpha$ es la base de Groebner reducida con respecto al orden monomial lexicográfico $Y_1 < ... < Y_n$.

En vista de un conocido lema (ver [Kobayashi *et al.*, 1989]) la condición (iii) puede reformularse como sigue:

(iv) existen polinomios $r_1,...,r_n$ en $k[Y_1]$ que tienen las mismas

propiedades que α con respecto a la entrada $f_1,...,f_s$, tales que $g_1 - r_1$, $g_2 - \alpha Y_2 - r_2,..., g_n - \alpha Y_n - r_n$. Más aún, $g_1,...,g_n$ generan el ideal I.

Para las aplicaciones que se harán de las proposiciones A y B, es importante observar que las entradas de la matriz M pueden elegirse de cualquier subconjunto de k prefijado Γ de cardinal suficientemente grande de orden asimptótico $s^{O(1)} d^{O(n)}$. Con la misma filosofía, el elemento α y los coeficientes de $g_1,...,g_s$ son polinomios de grado $d^{O(n)}$ en los coeficientes de $f_1,...,f_s$ sobre $\Omega[\Gamma]$, donde Ω denota el cuerpo primo de k.

3. Demostraciones

3.1. Verificación de la trivialidad del ideal I.

A lo largo de esta sección $f_1,...,f_s$ será una familia de polinomios de $k[X_1,...,X_n]$ cuyos grados estan acotados por un entero dado d que satisface $n \leq d$. Por medio de (2.4), Proposición A , es posible decidir en tiempo secuencial no uniforme $s^{O(1)} d^{O(n)}$ por medio de un algoritmo bien paralelizable si los polinomios $f_1,...,f_s$ tienen un cero en común en $\mathbb{A}^n$. Para la demostración del teorema principal enunciado en la introducción, es suficiente por lo tanto suponer que $f_1,...,f_s$ generan el ideal trivial de $k[X_1,...,X_n]$. Esta será, en lo que sigue, la suposición básica.

3.2. Preparación de los datos de entrada.

Sea $\sigma := (n + 1)\, s$. Se considera la familia de polinomios

$$g_1 := f_1 ,\ldots, g_s := f_s , g_{s+1} := x_1 f_1 ,\ldots, g_{s+n} := x_n f_1 ,\ldots$$
$$\ldots, g_{(n+1)s-n+1} := x_1 f_s ,\ldots, g_\sigma = g_{(n+1)s} := x_n f_s$$

Obsérvese que los grados de $g_1,\ldots,g_s$ están acotados por $d + 1$ y que $g_1,\ldots,g_s$ generan el ideal trivial de $k[X_1,\ldots,X_n]$.

Para $1 \le j \le s$ sea $U_j := (\mathbb{A}^n)_{f_j} = \{ x \in \mathbb{A}^n ; f_j(x) \neq 0 \}$ y $\Phi_j(x) : U_j \to \mathbb{A}^{\sigma-1}$ el morfismo de variedades afines definido por

$$\Phi_j(x) := \left(g_1(x)/f_j(x) ,\ldots, g_{j-1}(x)/f_j(x) , g_{j+1}(x)/f_j(x) ,\ldots, g_\sigma(x)/f_j(x) \right)$$

para $x \in U_j$.

Por construcción, Φ_j es no ramificada y la clausura de Zariski de su imagen es una subvariedad irreducible de $\mathbb{A}^{\sigma-1}$ de dimensión n. La familia finita $(U_j)_{1\le j\le s}$ forma un cubrimiento abierto de $\mathbb{A}^n$.

Sean $f'_1,\ldots,f'_{n+1}$ una familia de combinaciones k-lineales genéricas de $g_1,\ldots,g_\sigma$. Los grados de los polinomios $f'_1,\ldots,f'_{n+1}$ están acotados por $d + 1$ y generan el ideal trivial de $k[X_1,\ldots,X_n]$.

Fijando ahora $0 \le i \le n - 1$, se analiza el ideal $(f'_1,\ldots,f'_{n-i})$ por medio del teorema de Bertini y de las propiedades geométricas de los morfismos Φ_j, $1 \le j \le s$.

Si se fija por el momento $1 \le j \le s$, se puede observar primero que para cada $1 \le k \le n + 1$, la función racional f'_k/f_j es una combinación k-lineal afín genérica de $g_1(x)/f_j(x) ,\ldots, g_{j-1}(x)/f_j(x) , g_{j+1}(x)/f_j(x) ,\ldots, g_\sigma(x)/f_j(x)$. Por lo tanto, el teorema de Bertini ([Jouanolou, 1983], Corollaire 6.7) y el hecho de que Φ_j es no ramificada y tiene imagen de dimensión n, implican que

$f'_1/f_j,...,f'_{n-1}/f_j$ definen una subvariedad suave de U_j de dimensión i y generan un ideal radical del anillo cociente $k[X_1,...,X_n]_{f_j}$. (De hecho la variedad es irreducible y el ideal es primo para $1 \leq i \leq n-1$.) Como los U_j, $1 \leq j \leq s$, forman un cubrimiento abierto de $\mathbb{A}^n$, se concluye que la subvariedad de $\mathbb{A}^n$ definida por $f'_1,...,f'_{n-1}$ es suave y de dimensión i, y que el ideal $(f'_1,...,f'_{n-1})$ es radical para todo $0 \leq i \leq n-1$. (Se dice que una subvariedad cerrada de $\mathbb{A}^n$ eventualmente reducible es suave si es equidimensional y todos sus anillos locales son regulares.) En particular $f'_1,...,f'_n$ forman una sucesión regular de $k[X_1,...,X_n]$.

Para cada $1 \leq i \leq n+1$, se introducen nuevas indeterminadas $T_1^{(i)},...,T_\sigma^{(i)}$ que se considerarán como parámetros computacionales en el argumento que sigue (comparar con [Heintz-Giusti, 1991 b]).

Sea $R := k[T_q^{(i)};\ 1 \leq i \leq n+1,\ 1 \leq q \leq \sigma]$ y sea K el cuerpo de fracciones de R. Sea $\bar{K}$ una clausura algebraica de K. Para $1 \leq i \leq n+1$ se considera el polinomio $F_i := T_1^{(i)} g_1 + ... + T_\sigma^{(i)} g_\sigma$ de $R[x_1,...,x_n]$. El grado de F_i en las variables principales $X_1,...,X_n$ está acotado por $d+1$ y F_i es lineal en los parámetros $T_1^{(i)},...,T_\sigma^{(i)}$.

Del teorema de Bertini se deducen los hechos siguientes:

(a) Los polinomios $F_1,...,F_{n+1}$ no tienen ceros comunes en $\mathbb{A}^n(\bar{K})$
(b) Para cada $0 \leq i \leq n-1$ los polinomios $F_1,...,F_{n-i}$ generan un ideal radical de $K[x_1,...,x_n]$ y definen una subvariedad suave de $\mathbb{A}^n(\bar{K})$ de dimensión i.

Para $0 \leq i \leq n-1$, sean $r_i := (n-i)!/(n-2i)!\,i!$ y $\Delta_1^{(n-i)},...,\Delta_{r_i}^{(n-i)}$ los $n-i$ menores de la matriz jacobiana de $F_1,...,F_{n-i}$. Se tiene que $r_i \leq n^i \leq d^i$, por lo tanto r_i es de orden d^n. $\Delta_1^{(n-i)},...,\Delta_{r_i}^{(n-i)}$ son polinomios de $R[X_1,...,X_n]$ cuyos

grados en las variables principales $X_1,...,X_n$ están acotados por (n-i) d y cuyos grados en los parámetros $T_q^{(1)},...,T_q^{(i)}$, con $1 \le q \le \sigma$, no exceden a n - i.

Por medio del criterio del jacobiano para variedades suaves se puede reenunciar la propiedad (b) como sigue:

(c) Los polinomios $F_1,...,F_{n-i}$ definen una subvariedad de dimensión i de $\mathbb{A}^n(\bar{K})$ y la variedad definida por $F_1,...,F_{n-i}, \Delta_1^{(n-i)},...,\Delta_{r_i}^{(n-i)}$ es vacía.

Ahora se podrán construir los polinomios $f'_1,...,f'_{n+1}$ anteriores de forma algorítmica.

Lema 3.2.1. Si $f_1,...,f_s$ generan el ideal trivial de $k[X_1,...,X_n]$, existe un algoritmo bien paralelizable que construye de las entradas $f_1,...,f_s$ en tiempo secuencial no uniforme $s^{O(1)}\, d^{O(n)}$ polinomios $f'_1,...,f'_{n+1}$ de $k[X_1,...,X_n]$ de grado a lo sumo d + 1 que cumplen las siguientes propiedades:

(1) Los polinomios $f'_1,...,f'_{n+1}$ generan el ideal trivial de $k[X_1,...,X_n]$
(2) Para cada $0 \le i \le n - 1$, los polinomios $f'_1,...,f'_{n-i}$ generan un ideal radical de $k[X_1,...,X_n]$ y definen una subvariedad suave de $\mathbb{A}^n$ de dimensión i.
(3) Los polinomios $f'_1,...,f'_{n+1}$ pertenecen al subespacio k-lineal de $k[X_1,...,X_n]$ generado por los polinomios $f_1,...,f_s, X_1 f_1,...,X_1 f_s,...,X_n f_1,...,X_n f_s$.

Demostración:

Se mantendrán las notaciones ya establecidas. Para $y \in \mathbb{A}^{(n+1)\sigma}$ se notará con $f_1(y),...,f_{n+1}(y)$ a los polinomios que se obtienen sustituyendo en $F_1,...,F_{n+1}$ los parámetros $T_q^{(1)},...,T_q^{(n+1)}$, $1 \le q \le \sigma$, por las coordenadas de y.

Esta demostración sigue los argumentos de [Giusti-Heintz, 1991 b], Section 3.7.2.

Para cada $-1 \leq i \leq n-1$ se aplican los algoritmos de la sección 2.4, Proposición A (1) a los polinomios $F_1,\ldots,F_{n-1}$ de $R[X_1,\ldots,X_n]$ y, para cada $0 \leq i \leq n-1$, se aplica el mismo algoritmo a los polinomios $F_1,\ldots,F_{n-1}$, $\Delta_1^{(n-i)},\ldots,\Delta_{r_i}^{(n-i)}$. El procedimiento total computa durante su ejecución una familia de elementos no nulos $Q_1,\ldots,Q_N$ de R que son representados por un cálculo de evaluación sin divisiones de longitud v. Los enteros N y v son del orden de $s^{O(1)}\, d^{O(n)}$. Los resultados intermediarios $Q_1,\ldots,Q_N$ son polinomios en los parámetros $T_q^{(1)},\ldots,T_q^{(n+1)}$, $1 \leq q \leq \sigma$, con coeficientes en k y sus grados están acotados por d^{cn}, donde $c > 0$ es una constante convenientemente elegida. Los polinomios $Q_1,\ldots,Q_N$ se construyen de tal forma que satisfagan la condición siguiente:

Para cualquier $\gamma \in \mathbb{A}^{(n+1)\sigma}$ tal que $Q_1(\gamma) \neq 0,\ldots,Q_n(\gamma) \neq 0$, los polinomios $f_1^{(\gamma)},\ldots,f_{n+1}^{(\gamma)}$ tienen las propiedades (1), (2) y (3) del lema.
Esto se deduce de las observaciones hechas al principio de esta sección y de las condiciones específicas del algoritmo subyacente a la Proposición A (ver [Giusti-Heintz, 1991 b] para los detalles técnicos).

Sea $Q := Q_1 \ldots Q_N$. Sin pérdida de generalidad se puede suponer que $\deg Q \leq d^{cn}$ y que Q es computable por medio de un cálculo de evaluación de longitud a lo sumo v.

Sea Γ cualquier subconjunto de k de cardinal $\#\Gamma = 2v\,(1 + d^{cn})^2$. Obviamente, se tiene que $\#\Gamma = s^{O(1)}\, d^{O(n)}$. De acuerdo a [Heintz-Schnorr, 1982], Theorem 4.4, existe un conjunto de cuestores de $m := 6(v+(n+1)\sigma).(v+(n+1)\sigma+1)$ puntos de $\Gamma^{(n+1)\sigma}$ para la clase de polinomios a $(n+1)\sigma$ variables de grado acotado por d^{cn} que pueden ser evaluados por cálculos de evaluación de longitud a lo sumo v. Obsérvese que $m = s^{O(1)}\, d^{O(n)}$. Se elige un conjunto de cuestores $(\gamma_1,\ldots,\gamma_m)$ en $\Gamma^{(n+1)\sigma}$. (Esta elección no depende de la entrada $f_1,\ldots,f_s$; sólo depende de los parámetros d, s y n y puede hacerse por un prepocesamiento del algoritmo. Ver sección 2.2.)

Se corre el cálculo de evaluación sin divisiones que calcula $Q_1,...,Q_N$ para cada ℓ, $1 \leq \ell \leq m$. Esto puede hacerse por medio de un algoritmo bien paralelizable que tiene $s^{O(1)}\ d^{O(n)}$ operaciones aritméticas en k. Como el polinomio no nulo $Q = Q_1...Q_N$ puede ser evaluado por un cálculo de evaluación de longitud a lo sumo v y tiene grado acotado por d^{cn}, existe un punto $\gamma := \gamma_\ell$, $1 \leq \ell \leq m$ tal que $Q(\gamma) \neq 0$. Por lo tanto, se tiene que $Q_1(\gamma) \neq 0,...$ ..., $Q_N(\gamma) \neq 0$ y el punto γ se encuentra en tiempo secuencial no uniforme $s^{O(1)}\ d^{O(n)}$ por medio del algoritmo bien paralelizable indicado anteriormente. Los polinomios $f'_1 := f_1(\gamma),...,f'_{n+1} := f_{n+1}(\gamma)$ satisfacen las condiciones (1), (2) y (3) del lema. □

3.3 Notaciones y suposiciones básicas

En vista del Lema 3.2.1, de ahora en más se supondrá que las entradas son polinomios $f_1,...,f_s$ que satisfacen las siguientes condiciones:

- $s = n + 1$
- $f_1,...,f_{n+1}$ generan el ideal trivial de $k[X_1,...,X_n]$
- Para cada $0 \leq i \leq n - 1$, los polinomios $f_1,...,f_{n-i}$ generan un ideal radical de $k[X_1,...,X_n]$ y definen una subvariedad lisa $V_i := V(f_1,...,f_{n-i})$ de dimensión i.

Más aún, aplicando el algoritmo subyacente al item (2) de la Proposición A de la Sección 2.4 a los datos de entrada $f_1,...,f_{n+1}$ se supondrá de ahora en más que:

- Las variables $X_1,...,X_n$ estan en posición de Noether con respecto a todas las variedades V_i, para $0 \leq i \leq n - 1$.

Una familia de polinomios $f_1,...,f_{n+1}$ que satisfaga todas estas condiciones se llamará *entrada normalizada*.

De ahora en más se usarán las siguientes notaciones para $0 \leq i \leq n - 1$ fijo:

$A_i := k[X_1,...,X_n]$
$K_i := k(X_1,...,X_n)$
$\mathfrak{I}_i := I(f_1,...,f_{n-i})$, el ideal generado por $f_1,...,f_{n-i}$ en $k[X_1,...,X_n]$
$\mathfrak{I}'_i$ el ideal generado por $f_1,...,f_{n-i}$ en $k(X_1,...,X_i)[X_{i+1},...,X_n] = K_i[X_{i+1},...,X_n]$
$B_i := k[X_1,...,X_n]/\mathfrak{I}_i$
$B'_i := K_i[X_{i+1},...,X_n]/\mathfrak{I}'_i$.

Más aún, sea $\overline{K}_i$ una clausura algebraica de K_i. Sea $V_i := V(\mathfrak{I}_i)$ la variedad algebraica definida por $f_1,...f_{n-i}$ en $\mathbb{A}^n$ y $V'_i := V(\mathfrak{I}'_i)$ la variedad algebraica definida por $f_1,...,f_{n-i}$ en $\mathbb{A}^n(\overline{K}_i)$. Se sigue, en base a lo supuesto, que V_i es suave de dimensión i, que V'_i es de dimensión cero y que $\mathfrak{I}_i$ y $\mathfrak{I}'_i$ son ideales radicales. Por lo tanto los polinomios $f_1,...,f_{n-i}$ forman una sucesión regular en los anillos de Cohen-Macaulay $k[X_1,...,X_n]$ y $K_i[X_{i+1},...,X_n]$.

Se considera a B_i como una A_i-álgebra y a B'_i como una K_i-álgebra. Como las variables $X_1,...,X_n$ están en posición de Noether con respecto a V_i, el álgebra B_i es una A_i-módulo finito y B'_i es un K_i-espacio vectorial de dimensión finita. Los homomorfismos canónicos $A_i \to B_i$ y $K_i \to B'_i$ son extensiones enteras. Como A_i es una álgebra polinomial sobre el cuerpo k, la k-álgebra A_i es un dominio factorial y por lo tanto, íntegramente cerrado en su cuerpo de fracciones K_i. Sea $D_i := \dim_{K_i} B'_i$. Teniendo en cuenta que la variedad V_i es suave, que las variables $X_1,...,X_n$ están en posición de Noether con respecto a V_i y que $\mathfrak{I}_i$ y $\mathfrak{I}'_i$ son ideales radicales, se deduce fácilmente que $D_i \leq \deg V_i \leq d^{n-i}$ a partir de la desigualdad de Bezout (ver [Heintz, 1983] para más detalles).

3.4 Un poco de álgebra conmutativa.

El siguiente es un lema conocido (compárese, por ejemplo, con [Rossi-Spangher, 1991], Proposition 3.4). Se da la demostración por carecer de referencias apropiadas.

Lema 3.4.1. Sea $0 \leq i \leq n - 1$ y sean $g_1,...,g_{n-1}$ una sucesión de polinomios de $k[X_1,...,X_n]$ que definen una variedad algebraica $V := (g_1=0,...,g_{n-1}=0)$ de dimensión i. Supóngase que $X_1,...,X_n$ estan en posición de Noether con respecto a V. Sea $A := k[X_1,...,X_i]$ y considérese la k-álgebra $B := k[X_1,...,X_n]/(g_1,...,g_{n-i})$ como A-módulo. Entonces B es libre de rango finito.

Demostración:

Se procede por inducción en i. En el caso $i = 0$ resulta $A = k$ y $B := k[X_1,...,X_n]/(g_1,...,g_n)$ es un k-espacio vectorial de dimensión finita. Luego B es un A-módulo de rango finito positivo. Basta entonces demostrar el lema para un $0 < i \leq n - 1$ fijo suponiéndolo verdadero para $i - 1$.

Como $\dim V = i$ y las variables $X_1,...,X_n$ estan en posición de Noether con respecto a la variedad V, la k-álgebra B es un A-módulo finito no trivial. Más aún, A es un anillo de polinomios sobre k. Por lo tanto, si B es un A-módulo proyectivo, entonces B es libre de rango finito positivo por el teorema de Quillen-Suslin ([Lam,1978], Chapter III, Theorem 1.8).

Por lo tanto, se demostrará que B es proyectivo.

Primero, obsérvese que $\bar{k} \otimes_k A$ es una A-álgebra fielmente playa. Por lo tanto, $\bar{k} \otimes_k B = (\bar{k} \otimes_k A) \otimes_A B$ es un $\bar{k} \otimes_k A$ -módulo proyectivo si y sólo si B es un A-módulo proyectivo ([Lam,1978], Chapter I, Proposition 2.15). Por

lo tanto, se puede suponer sin pérdida de generalidad que k es algebraicamente cerrado.

Como B es un A-módulo finito, bastará mostrar que para todo ideal maximal $\mathfrak{M}$ de A, el localizado $B_{\mathfrak{M}}$ es un $A_{\mathfrak{M}}$-módulo libre.

Sea $\mathfrak{M}$ un ideal maximal de $A = k[X_1,...,X_i]$. Como, por la reducción ya hecha, k es algebraicamente cerrado, existen elementos $\alpha_1,...,\alpha_i$ de k tales que $\mathfrak{M} = (X_1 - \alpha_1,...,X_i - \alpha_i)$. Sin pérdida de generalidad se puede suponer que $\alpha_1 = ... = \alpha_i = 0$. Entonces $\mathfrak{M} = (X_1,...,X_i)$. Como las variables $X_1,...,X_n$ están en posición de Noether con respecto a $V = (g_1 = 0,...,g_{n-i} = 0)$, la variedad $V \cap (X_i = 0) = (g_1 = 0,...,g_{n-1} = 0, X_1 = 0)$ es de dimensión i - 1 y en particular no es vacía. De esto se concluye que los polinomios $g_1,...,g_{n-1},X_i$ satisfacen las hipótesis del lema para i - 1 y que forman una sucesión regular en el anillo de Cohen-Macaulay $k[X_1,...,X_n]$ (véase [Matsumura, 1989], Theorem 17.7). Por lo tanto X_i, visto como elemento de $B = k[X_1,...,X_n]/(g_1,...,g_{n-i})$, no es divisor de cero.

Sea $\overline{A} := k[X_1,...,X_{i-1}]$ y sea $\overline{B} := k[X_1,...,X_n]/(g_1,...,g_{n-i},X_i)$. Por hipótesis inductiva $\overline{B}$ es un $\overline{A}$-módulo libre de rango finito. Sea $^{-} : B \to \overline{B}$ el homomorfismo de k-álgebras canónico. Este homomorfismo induce la identidad sobre $\overline{A} = k[X_1,...,X_{i-1}]$ y manda $\mathfrak{M}$ al ideal maximal $\overline{\mathfrak{M}} := (X_1,...,X_{i-1})$ de $\overline{A}$. Sean $e_1,...,e_N$ elementos de B tales que $e_1,...,e_N$ forman una base del $\overline{A}$-módulo libre $\overline{B}$. Por lo tanto, $1 \otimes e_1,...,1 \otimes e_N$ generan el $A/\mathfrak{M}$-espacio vectorial

$$A/\mathfrak{M} \otimes_A B \cong B/\mathfrak{M} \cong \overline{B}/\overline{\mathfrak{M}} \cong \overline{A}/\overline{\mathfrak{M}} \otimes_{\overline{A}} \overline{B}$$

Si se nota $e'_1,...,e'_N$ a las imágenes de $e_1,...,e_N$ en $B_{\mathfrak{M}}$ y $(e_1)',...,(e_N)'$ a sus imágenes en $\overline{B}_{\overline{\mathfrak{M}}}$, del lema de Nakayama se concluye que $e'_1,...,e'_N$ generan el $A_{\mathfrak{M}}$-módulo $B_{\mathfrak{M}}$. Para finalizar la demostración se probará que son linealmente independientes.

Supóngase, por el contrario, que existe una relación lineal no trivial

$$a_1 e'_1 + \ldots + a_N e'_N \qquad (1)$$

en $B_{\mathfrak{M}}$, donde $a_1,\ldots,a_N$ son elementos de $A_{\mathfrak{M}}$, no todos nulos. Sin pérdida de generalidad se puede suponer que $a_1,\ldots,a_N$ pertenecen a $A = k[X_1,\ldots,X_i]$. Dividiendo por una potencia maximal de X_i se obtienen representaciones

$a_1 = X_i^{\ell} a'_1,\ldots,a_N = X_i^{\ell} a'_N$,

donde $a'_1,\ldots,a'_N$ son elementos de A, no todos divisibles por X_i. Por lo tanto, de (1) obtenemos la relación

$$X_i^{\ell} (a'_1 e'_1 + \ldots + a'_N e'_N) = 0$$

que vale en $B_{\mathfrak{M}}$. Como X_i no es un divisor de cero en B se concluye que

$$a'_1 e'_1 + \ldots + a'_N e'_N = 0.$$

Por lo tanto se puede suponer sin pérdida de generalidad que a_1 es un elemento de A que no es divisible por X_i. Luego, la relación (1) implica que

$$a_1 e'_1 + \ldots + a_N e'_N = 0 \qquad (2)$$

vale en $B_{\mathfrak{M}}$ con $a_1,\ldots,a_N \in \overline{A}_{\mathfrak{M}}$, $a_1 \neq 0$.

Como $e_1,\ldots,e_N$ forman una base del $\overline{A}$-módulo $\overline{B}$, los elementos $(e_1)',\ldots,(e_N)'$ de $\overline{B}_{\mathfrak{M}}$ son lineamente independientes sobre $\overline{A}_{\mathfrak{M}}$. Esto contradice la relación (2). □

Si se mantienen las notaciones y las hipótesis de la Sección 3.3, del Lema 3.4.1 se obtiene el siguiente

Corolario 3.4.2. Para cada $0 \leq i \leq n-1$ la A_i-álgebra B_i es un A_i-módulo libre de rango positivo $D_i \leq d^{n-i}$.

(Recuérdense las notaciones introducidas en 3.3).

Demostración:

Sea $0 \leq i \leq n-1$. Por hipótesis (véase Sección 3.3) los polinomios $f_1,\ldots,f_{n-i}$ definen una variedad algebraica V_i de dimensión i y las variables $X_1,\ldots,X_n$ están en posición de Noether con respecto a V_i. Por lo tanto, por el Lema 3.4.1, la A_i-álgebra B_i es un A_i-módulo libre de rango finito positivo. Como K_i es el cuerpo de fracciones de A_i, este rango es $\dim_{K_i} K_i \otimes_{A_i} B_i = \dim_{K_i} B_i' = D_i \leq \deg V_i \leq d_{n-i}$. □

Lema 3.4.3. Sea $0 \leq i \leq n-1$ y sea f un polinomio de $k[X_1,\ldots,X_n]$. La multiplicación por f induce un endomorfismo lineal Φ sobre el K_i-espacio vectorial B_i' de dimensión D_i. Sea T una nueva indeterminada y sea $\chi_f \in K_i[T]$ el polinomio característico del endomorfismo lineal Φ. Sea $\chi_f = T^{Di} + \varphi_{Di-1} T^{Di-1} + \ldots + \varphi_0$ con $\varphi_{Di-1},\ldots,\varphi_0 \in K_i$. Entonces $\varphi_{Di-1},\ldots,\varphi_0$ son elementos de A_i cuyos grados están acotados por $i\,(\deg V_i)^2 \deg f \leq i\, d^{2(n-i)} \deg f$.

(Recuérdese las notaciones introducidas en 3.3).

Demostración:

Para simplificar la notación, sean $A := A_i$, $K := K_i$, $B := B_i$, $B' := B_i'$, $\mathfrak{I} := \mathfrak{I}_i$, $V := V_i$ y $D := D_i$. Entonces Φ es el endomorfismo K-lineal inducido por f.

Como, por el Corolario 3.4.2, el A-módulo B es libre, se puede considerar

una base $e_1,...,e_D$ de él. Sea $e_1',...,e_D'$ la base correspondiente de B' como K-espacio vectorial. El polinomio f induce un endomorfismo A-lineal sobre B. Por lo tanto la matriz M_f de Φ con respecto a la base $e_1',...,e_D'$ de B' tiene todos sus coeficientes en el anillo de polinomios A. Esto implica que $\varphi_{D_{i-1}},...,\varphi_0$ son elementos de A.

Ahora obsérvese que $\bar{k}\otimes_k K$ es un cuerpo (de hecho, es el cuerpo de fracciones de $\bar{k}\otimes_k A$) y que $\bar{k}\otimes_k B'$ es un $\bar{k}\otimes_k K$-espacio vectorial de dimensión D con base $1\otimes e_1',...,1\otimes e_D'$. El polinomio f induce un endomorfismo $\bar{k}\otimes_k K$-lineal sobre $\bar{k}\otimes_k B'$, cuya matriz con respecto a la base $1\otimes e_1',...,1\otimes e_D'$ es M_f. Por lo tanto el polinomio característico del endomorfismo inducido por f no varía si se extiende el cuerpo de base k a su clausura algebraica $\bar{k}$. Más aún, $\bar{k}\otimes_k B$ es un $\bar{k}\otimes_k A$-módulo libre de rango D, sobre el cual f induce un endomorfismo $\bar{k}\otimes_k A$-lineal. Se tiene que $\bar{k}\otimes_k B \cong \bar{k}[X_1,...,X_n]/(f_1,...,f_{n-i})$ y que $\bar{k}\otimes_k B' \cong \bar{k}(X_1,...,X_i)[X_{i+1},...,X_n]/(f_1,...,f_{n-i})$. Además, los polinomios $f_1,...,f_{n-i}$ generan ideales radicales de $\bar{k}[X_1,...,X_n]$ y de $\bar{k}(X_1,...,X_i)[X_{i+1},...,X_n]$. Entonces, para el resto de la demostración, se puede suponer que k es algebraicamente cerrado.

Como $\mathfrak{I} = (f_1,...,f_{n-i})$ es un ideal radical de $k[X_1,...,X_n]$ que define una variedad algebraica de dimensión i, puede ser escrito unívocamente en la forma

$$\mathfrak{I} = \mathfrak{P}_1 \cap ... \cap \mathfrak{P}_r,$$

donde $\mathfrak{P}_1,...,\mathfrak{P}_r$ son ideales primos de $k[X_1,...,X_n]$ de altura n - i. Los ideales primos $\mathfrak{P}_1,...,\mathfrak{P}_r$ definen las componentes irreducibles $C_1,...,C_r$ de la variedad algebraica V dada por $\mathfrak{I}$. Obsérvese que $\dim C_1 = ... = \dim C_r = i$.

Se demostrará que para todo $1 \leq \ell \leq r$, los ideales $\mathfrak{P}^{\ell}$ y $\bigcap_{j \neq \ell} \mathfrak{P}_j$ son coprimos. En términos geométricos esto significa que $C_1,...,C_r$ son las componentes conexas de V con respecto a la topología de Zariski.

Sea x un punto cualquiera de V y sea $\mathfrak{M}$ el ideal maximal de $k[X_1,...,X_n]$ que define a x. Como el ideal $\mathfrak{I}$ es radical y como $B = k[X_1,...,X_n]/\mathfrak{I}$, se ve que $B_{\mathfrak{M}}$ es el anillo local del punto x. Por otro lado la variedad V es suave. Entonces $B_{\mathfrak{M}}$ es un anillo regular, y por lo tanto es un dominio. Esto implica que $\mathfrak{M}$ contiene exactamente uno de los ideales primos $\mathfrak{P}_1,...,\mathfrak{P}_{\ell}$. Luego, x está en exactamente una componente de V. Como dichas componentes son conexas y el punto x fue elegido arbitrariamente se deduce lo que se quería demostrar.

Para $1 \leq \ell \leq r$, sea R_{ℓ} el anillo de coordenadas de C_{ℓ} . Entonces, $R_{\ell} \cong \cong k[X_1,...,X_n]/\mathfrak{P}_{\ell}$. Como $B = k[X_1,...,X_n]/\mathfrak{I} = k[X_1,..,X_n]/ \bigcap_{1 \leq \ell \leq r} \mathfrak{P}_{\ell}$ y todos los ideales $\mathfrak{P}_{\ell} + \bigcap_{j \neq \ell} \mathfrak{P}_j$, $1 \leq \ell \leq r$, son triviales se deduce por el teorema chino del resto que B y $\prod_{1 \leq \ell \leq r} R_{\ell}$ son A- álgebras isomorfas.

Sea $1 \leq \ell \leq r$. Como R_{ℓ} es un sumando directo de B, es un A-módulo proyectivo y, por el teorema de Quillen-Suslin, también es libre. Sea R'_{ℓ} el anillo local de la fibra genérica de la proyección canónica $\pi_{\ell} : C_{\ell} \rightarrow \mathbb{A}^i$ inducida por las variables $X_1,...,X_i$. En términos algebraicos se puede definir R'_{ℓ} como

$$R'_{\ell} := K \otimes_A R_{\ell} \cong k(X_1,...,X_i) \otimes_{k[X_1,...,X_i]} (k[X_1,...,X_n]/\mathfrak{P}_{\ell}).$$

El anillo local R'_{ℓ} es una K-álgebra de dimensión finita que se notará $m_{\ell} := \dim_K R'_{\ell}$. De la desigualdad de Bezout se deduce que $m_{\ell} \leq \deg C_{\ell}$.

Tanto el morfismo A-lineal como el K-lineal inducido por la multiplicación por f en R_ℓ y en R'_ℓ se notará ψ_ℓ. Sea T una nueva indeterminada y sea $\chi_{\psi_\ell} = T^{m_\ell} + \gamma_{m_\ell-1}^{(\ell)} T^{m_\ell-1} + \ldots + \gamma_0^{(\ell)}$ con $\gamma_{m_\ell-1}^{(\ell)}, \ldots, \gamma_0^{(\ell)}$ en A. Obsérvese que la transformación K-lineal Φ actúa sobre $B' \cong \prod_{1\le\ell\le r} R'_\ell$ diagonalmente.

Por lo tanto, $\chi_f = \prod_{1\le\ell\le r} \chi_{\psi_\ell}$.

Se demostrará que los grados de los polinomios $\gamma_{m_\ell-1}^{(\ell)}, \ldots, \gamma_0^{(\ell)}$ de $A = k[X_1,\ldots,X_i]$ están acotados por $i\,(\deg C_\ell)^2 \deg f$. Entonces, $\deg V = \sum_{1\le\ell\le r} \deg C_\ell$ implica que los grados de los coeficientes $\varphi_{D-1},\ldots,\varphi_0$ del polinomio característico χ_f están acotados por $i\,(\deg V)^2 \deg f$, lo que termina con la demostración del Lema.

Para el resto de la demostración se fija $1 \le \ell \le r$. Para simplificar las notaciones sean $\mathfrak{P} := \mathfrak{P}_\ell$; $R := R_\ell$; $R' := R_\ell'$; $C := C_\ell$; $\psi := \psi_\ell$; $m := m_\ell$; $\gamma_{m-1} := \gamma_{m_\ell-1}^{(\ell)}, \ldots, \gamma_0 := \gamma_0^{(\ell)}$ y $\chi_\psi := \chi_{\psi_\ell}$.

Obsérvese que la variedad algebraica irreducible C es de dimensión i y que las variables $X_1,\ldots,X_n$ estan en posición de Noether con respecto a C. Las variables $X_1,\ldots,X_i$ inducen una proyeccion π: $C \to \mathbb{A}^i$ que es un morfismo finito y suryectivo de variedades algebraicas. Por lo tanto, el anillo de coordenadas R de C es una extensión entera de $A = k[X_1,\ldots,X_i]$. Más aún, R' es el cuerpo de fracciones de R y es una extensión finita de $K = k(X_1,\ldots,X_i)$. Entonces se tiene el siguiente diagrama de cuerpos y dominios:

$$\begin{array}{ccc} A & \subset & R \\ \cap & & \cap \\ K & \subset & R' \end{array}$$

Como R es isomorfo a $k[X_1,...,X_n]/\mathfrak{P}$, al polinomio f le corresponde una imagen en R, por ejemplo g, que induce el morfismo A-lineal (y K-lineal) ψ. El polinomio minimal $\mu \in K[T]$ del elemento g y el del morfismo K-lineal ψ es el mismo. Como A es íntegramente cerrado en su cuerpo de fracciones y $g \in R$ es entero sobre A se tiene que $\mu \in A[T]$. Más aún, μ es primitivo e irreducible con respecto a K[T]. Por lo tanto, μ es un polinomio irreducible de $k[X_1,...,X_i,T] = A[T]$.

El polinomio minimal μ es de la forma

$$\mu = T^q + \mu_{q-1} T^{q-1} + ... + \mu_0$$

con μ_{q-1} ..... $\mu_0 \in A = k[X_1,...,X_i]$. Como R'/K es una extensión finita de cuerpos de grado m, el polinomio característico de g, que coincide con χ_ψ, es una potencia de μ. Entonces se tiene que $\chi_\psi = \mu^h$ con $h \leq m \leq \deg C$.

Se demostrará ahora que los grados de μ_{q-1} ..... μ_0 están acotados por (i deg C deg f). Esto implica que los grados de los coeficientes γ_{m-1} ..... γ_0 de χ_ψ están acotados por i $(\deg C)^2$ deg f, que es lo que se quiere probar.

Sea Y cualquiera de las variables $X_1,...,X_i$ que aparecen en el polinomio $\mu \in k[X_1,...,X_i,T]$. Se probará que el grado parcial $\deg_Y \mu$ de μ en Y esta acotado por deg f deg C. Sin pérdida de generalidad se puede suponer que μ realmente depende de Y (es decir que Y aparece por lo menos en uno de los monomios de μ). Para simplificar la notación supongamos $Y = X_1$. El isomorfismo $R \cong k[X_1,...,X_n]/\mathfrak{P}$ induce un morfismo canónico de k-álgebras

$\bar{\ }: k[X_1,...,X_n] \to R$. Se tiene que $g = \bar{f}$ y $\mu(Y,X_2,...,X_i,f) = \mu(X_1,...,X_i,g) = 0$. Sea Z una nueva indeterminada. Como el polinomio μ es irreducible y depende de X_1, los elementos $X_2,...,X_i,f$ de R son algebraicamente independientes sobre k y $\mu(Z,X_2,...,X_i,f)$ es el polinomio minimal de Y sobre $k[X_2,...,X_i,f]$. Por lo tanto, $\deg_Y \mu$ esta acotado por $[R' : k(X_2,...,X_i,f)]$, el grado de la extensión finita de cuerpos $k(X_2,...,X_i,f) \subset R'$. (Obsérvese que i - dim C es el grado de trascendencia del cuerpo R' sobre k.)

Considérese el morfismo de variedades afines $\theta : C \to \mathbb{A}^i$ definido por las secciones globales $X_2,...,X_i,f$ de C. Como estos elementos son algebraicamente independientes sobre k, el morfismo θ es dominante y su fibra típica es de dimensión cero. Elíjase un punto $\lambda = (\lambda_1 , ... , \lambda_i)$ de $\mathbb{A}^i$ que satisfaga $\dim \theta^{-1}(\lambda) = 0$. Luego $X_2 - \lambda_1, ... ,X_i - \lambda_{i-1}, f - \lambda_i$ generan un ideal propio $\mathfrak{B}$ de R cuyo anillo cociente $R/\mathfrak{B}$ tiene dimensión de Krull cero. Teniendo en cuenta que $k(X_2 - \lambda_1, ... ,X_i - \lambda_{i-1}, f - \lambda_i) = k(X_2,...,X_i,f)$ se concluye (usando [Caniglia-Galligo-Heintz, 1989], Proposition 5) que $[R' : k(X_2,...,X_i,f)] \leq \deg C \deg f$. Entonces se obtiene que $\deg_Y \mu \leq \deg C \deg f$.

Como Y era cualquiera de las variables $X_1,...,X_i$ que aparecían en μ, el grado parcial de μ en $X_1,...,X_i$ está acotado por (i deg C deg f). Esto significa que los grados de $\mu_{q-1},...,\mu_0$ están acotados por (i deg C deg f). □

3.5. Cálculos de evaluación y algebra lineal sobre $k[X_1,...,X_n]$.

A lo largo de esta Sección se mantendra fijo $1 \leq i \leq n - 1$, y también se mantendrán las notaciones y convenciones introducidas en 3.1, 3.2 y 3.3.

Sea $V := V_i$ la subvariedad suave de $\mathbb{A}^n$ de dimensión i definida por $f_1,...,f_{n-i}$ y $\mathfrak{I} = \mathfrak{I}_i$ y $\mathfrak{I} = \mathfrak{I}_i'$ los ideales generados por $f_1,...,f_{n-i}$ en $k[X_1,...,X_n]$

y en $k(X_1,...,X_i)[X_{i+1},...,X_n]$ respectivamente. Por hipótesis las variables $X_1,...,X_n$ estan en posición de Noether con respecto a V y los ideales $\mathfrak{I}$ y $\mathfrak{I}'$ son radicales. Sean $A := A_i = k[X_1,...,X_i]$; $K := K_i = k(X_1,...,X_i)$; $B := B_i = k[X_1,...,X_n]/\mathfrak{I}_i$; $B' := B_i' = k(X_1,...,X_i)[X_{i+1},...,X_n)/(f_1,...,f_{n-i})$ y $D := D_i$. Del Corolario 3.4.2 se deduce que B es un A-módulo libre de rango finito positivo D. Por otra parte, B es un K-espacio vectoria de dimensión D y se tiene que $D \leq \deg V \leq d^{n-i}$. Sean $\bar{}: k[X_1,...,X_n] \to B$ y $\bar{}: K[X_{i+1},...,X_n] \to B'$ los morfismos canónicos.

Considérese un polinomio g de $k[X_1,...,X_n]$ o de $K[X_{i+1},...,X_n]$. Se notará $\bar{g}$ a las imágenes de g en B y en B' así como también al endomorfismo A-lineal de B y al endomorfismo K-lineal de B' inducidos por la multiplicación por g. Sea T una nueva indeterminada. Se notara χ_g al polinomio característico del endomorfismo K-lineal $\bar{g}$ de B'. Se considera χ_g como un elemento de K[T]. Del Lema 3.4.3 se deduce que de hecho χ_g es un polinomio de $A[T] = k[X_1,...,X_i,T]$ cuyo grado total no supera a $(D + i\, d^{2(n-i)} \deg g)$. Más precisamente, los coeficientes de χ_g, que son elementos de $k[X_1,...,X_i]$ tienen grado a lo sumo $(i\, d^{2(n-i)} \deg g)$.

Si M es una matriz en $K^{D\times D}$, se notará χ_M a su polinomio característico, que pertenece a K[T]. Se dirá que $M \in K^{D\times D}$ está dada (o representada) por un cálculo de evaluación β en K si β calcula todas las entradas de M. Sea $M \in A^{D\times D}$. Se define deg M, el grado de M, como el máximo de los grados de todas las entradas de M. El polinomio característico χ_M pertenece a A[T] en este caso.

Lema 3.5.1. Sea $M \in K^{D\times D}$ una matriz dada por un cálculo de evaluación β de longitud L y profundidad ℓ en K. Entonces los coeficientes de χ_M pueden computarse por medio de un cálculo de evaluación β' de longitud $L + cD^c$ y profundidad $\ell + c \log^2 D$ en K, donde $c > 0$ es una constante adecuada que no

depende del cuerpo K. Más aún, β' puede ser calculado a partir de β en tiempo secuencial (uniforme) $c\,D^c$ por medio de una red aritmética en k sin divisiones de profundidad $c \log^2 D$.

Demostración:

El algoritmo de Berkowitz ([Berkowitz, 1984]) calcula los coeficientes de χ_M a partir de las entradas de M por medio de un cálculo de evaluación sin divisiones de longitud $c\,D^c$ y produndidad $c \log^2 D$, donde $c > 0$ es una constante que no depende de K. Esto implica el lema. □

Lema 3.5.2. Existe una algoritmo sin divisiones bien paralelizable que calcula en tiempo secuencial no uniforme $d^{O(n)}$ a partir de las entradas $f_1,...,f_{n-i}$ los siguientes items:

- un polinomio no nulo α de $A = k[X_1,...,X_i]$
- polinomios $h_1,...,h_D$ de $k[X_1,...,X_n]$
- matrices $M_1,...,M_n$ de $A^{D \times D}$

tales que los grados de α, $h_1,...,h_D$ y $M_1,...,M_n$ son de orden $d^{O(n)}$, las clases $h_1,...,h_D$ forman una base del K-espacio vectorial B' y $M_1,...,M_n$ son las matrices de los endomorfismos K-lineales de B' $\alpha X_1,...,\alpha X_n$ con respecto a la base $h_1,...,h_D$. Dicho algoritmo representa al polinomio α y a las entradas de $M_1,...,M_n$ por medio de un cálculo de evaluación bien paralelizable y sin divisiones de longitud $d^{O(n)}$.

Demostración:

Se consideran las indeterminadas $X_1,...,X_i$ como parámetros y $X_{i+1},...,X_n$ como variables principales (ver Sección 2.1). Por lo tanto se trabajará en el álgebra de polinomios $A[X_{i+1},...,X_n]$ sobre el anillo de base $A = k[X_1,...,X_i]$.

Recuérdese que el cuerpo de fracciones de A es $K = k(X_1,...,X_i)$ y que el ideal $\mathfrak{I}'$, generado por los polinomios $f_1,...,f_{n-i}$ en $K[X_{i+1},...,X_n]$ es radical y de dimensión cero. Aplicando la Proposición B de la Sección 2.4 a esta situación se obtienen un polinomio no nulo α de $A = k[X_1,...,X_i]$, polinomios $g_{i+1},...,g_n$ de $k[X_1,...,X_n]$ y $n - i$ formas lineales independientes $y_{i+1},...,y_n$ de $K[X_{i+1},...,X_n]$ tales que $g_{i+1}/\alpha,...,g_n/\alpha$ es la base de Groebner reducida de $\mathfrak{I}'$ con respecto al orden monomial lexicográfico $y_{i+1} < ... < y_n$ de $K[y_{i+1},...,y_n] = K[X_{i+1},...,X_n]$. El algoritmo subyacente en la Proposición B de la Sección 2.4 computa a los polinomios α y $g_{i+1},...,g_n$ y a las formas lineales $y_{i+1},...,y_n$ a partir de la entrada $f_1,...,f_{n-i}$ por medio de una red aritmética sobre A (sin divisiones) de tamaño d^{cn} y produndidad $(c\ n^2 \log^2 d)$, donde $c > 0$ es una constante apropiada que no depende de A. (Obsérvese que el eventual chequeo de identidades entre elementos de A que puede aparecer durante la ejecución de este algoritmo puede realizarse por medio de una red aritmética de tamaño d^{cn} y profundidad $c\ n^2 \log^2 d$ sobre k usando [Heintz-Schnorr, 1982], Theorem 4.4. Véase [Giusti-Heintz, 1991 b] para detalles sobre esta técnica.)

Se consideran $f_1,...,f_{n-i}$ como polinomios en las variables principales $X_{i+1},...,X_n$ y $g_{i+1},...,g_n$ como polinomios en $y_{i+1},...,y_n$ con coeficientes en $A = k[X_1,...,X_i]$.

El algoritmo presenta al elemento α de A y a los coeficientes de $g_{i+1},...,g_n$ (que también son elementos de A) por medio de un cálculo de evaluación en $A = k[X_1,...,X_i]$ sin divisiones de longitud d^{cn} y profundidad $c\ n^2 \log^2 d$. A partir de estos datos, es fácil ahora encontrar por medio de [Heintz-Schnorr, 1982], Theorem 4.4 un conjunto de monomios $h_1,...,h_D$ en $y_{i+1},...,y_n$ tales que $h_1,...,h_D$ formen una base del K-espacio vectorial $B' = K[X_{i+1},...,X_n]/\mathfrak{I}'$ de dimension D así como también las matrices $M_1,...,M_n$ de los endomorfismos K-lineales $\alpha.X_1,...,\alpha.X_n$ de B' con respecto a esta base. Se verifica inmediatamente que $M_1,...,M_n$ son elementos de A^{DxD} cuyas entradas están representadas en el algoritmo por medio de un cálculo de evaluación en A =

$k[X_1,...,X_i)$ sin divisiones de longitud d^{cn} y produndidad c $n^2 \log^2 d$. □

Para el resto de esta sección se pensarán fijos a los polinomios $h_1,...,h_D$. Por el Lema 3.5.2 pueden encontrarse en tiempo secuencial $s^{O(1)}$ $d^{O(n)}$ por medio de un algoritmo bien paralelizable sin divisiones, sus grados son del orden $d^{O(n)}$ y $h_1,...,h_D$ forman una base del K-espacio vectorial B'.

Sea g un polinomio en $k[X_1,...,X_n]$ o en $K[X_{i+1},...,X_n]$. Se notará M_g a la matriz del endomorfismo K-lineal g de B' con respecto a la base $h_1,...,h_D$. La matriz M_g pertenece a $K^{D \times D}$ y puede contener funciones racionales (elementos de K-A) entre sus coeficientes aún cuando g sea un polinomio en $k[X_1,...,X_n]$. Sin embargo, si g pertenece a $k[X_1,...,X_n]$, el polinomio característico χ_{M_g} es un elemento de $A[T] = k[X_1,...,X_n,T]$, porque $\chi_{M_g} = \chi_g$ no depende de la base particular $h_1,...,h_D$ de B' y porque g es un endomorfismo A-lineal del A-módulo libre B. Los coeficientes de χ_{M_g}, que son polinomios en $k[X_1,...,X_i]$, no tienen grados que excedan la cota $i\, d^{2(n-i)}$ deg g. Por el Lema 3.5.2 las matrices $M_{X_1},...,M_{X_n}$ pueden encontrarse en tiempo secuencial no uniforme $d^{O(n)}$ por medio de un algoritmo bien paralelizable. Ellas son representadas por medio de un cálculo de evaluación en $K = k(X_1,...,X_i)$ bien paralelizable de longitud $d^{O(n)}$. La funcion $K[X_{i+1},...,X_n] \to K^{D \times D}$ que asocia a cada polinomio g de $K[X_{i+1},...,X_n]$ su matriz M_g es un homomorfismo de K-álgebras (esta observación sera crucial en el próximo lema).

Definición 3.5.3. Sea β un cálculo de evaluación en $k(X_1,...,X_i)[X_{i+1},...,X_n] = K[X_{i+1},...,X_n]$. Se llama complejidad no escalar de β al número de multiplicaciones y divisiones esenciales de β, es decir que no se cuentan las operaciones K-lineales (sumas, restas, multiplicaciones y divisiones por elementos de K son "gratis"). En este sentido β tendrá una longitud no escalar y una profundidad no escalar.

Sean L, ℓ, Λ, λ números naturales y sea g un polinomio en $K[X_{i+1},...,X_n]$. Se

dice que g tiene (simultáneamente) complejidad secuencial L, complejidad secuencial no escalar Λ, complejidad paralela ℓ y complejidad no escalar paralela λ, si existe un cálculo de evaluación β en $K[X_{i+1},...,X_n]$ que tiene sólo divisiones por elementos de K y que tiene longitud L, longitud no escalar Λ, profundidad ℓ y profundidad no escalar λ, que calcula a g (ver [von zur Gathen,1986] y [Strassen,1972] para más detalles).

Lema 3.5.4. Sea g un polinomio en $k[X_1,...,X_n]$ que puede ser evaluado por un cálculo de evaluación β en $k(X_1,...,X_i)[X_{i+1},...,X_n]$ de longitud L, longitud no escalar Λ, profundidad ℓ y produndidad no escalar λ (se supone que β tiene sólo divisiones por elementos de $k(X_1,...,X_i)$). Entonces la matriz M_g, puede representarse por medio de un cálculo de evaluación β' en $K = k(X_1,...,X_i)$ de longitud $L + c\,\Lambda^2 D^3 + d^{cn}$ y profundidad $\ell + c(\lambda \log D + n^2 \log^2 d)$, donde $c > 0$ es una constante que no depende de K. Más aún, β' puede producirse a partir de β en tiempo secuencial no uniforme $L + c\,\Lambda^2 D^3 + d^{cn}$ y tiempo paralelo $\ell + c(\lambda \log D + n^2 \log^2 d)$ por medio de una red aritmética sobre k.

Demostración:

Sea $\beta = (q_1,...,q_{L'})$ dónde $L \leq L'$ y $(q_1,...,q_{L'})$ son elementos de $k(X_1,...,X_i)[X_{i+1},...,X_n]$. Sin pérdida de generalidad se puede suponer que $g = q_{L'}$. Se considera la siguiente sucesión de matrices en $K^{D\times D}$:

$$\beta^* := (M_{q_1},...,M_{q_{L'}}).$$

Se interpretará a β^* como un cálculo de M_g a partir de las entradas $M_{X_1},...,M_{X_n}$ en el álgebra de matrices $K^{D\times D}$. En este sentido β^* tiene longitud L, longitud no escalar Λ, produndidad ℓ y profundidad no escalar λ.

A la concatenación de β con la sucesión de matrices β^* le corresponde un cálculo de evaluación β_1 en $k(X_1,...,X_i)$ que calcula los coeficientes de M_g a

partir de $X_1,...,X_i$ y los coeficientes de $M_{X_1},...,M_{X_n}$ como entradas. Aplicando el algoritmo más obvio para sumar y multiplicar matrices, se ve que la longitud de β_1 está acotada por $L + c\,A^2\,D^3$ y que la profundidad de β_1 está acotada por $\ell + c\,\lambda \log D$, donde $c > 0$ es una constante apropiada que no depende de K. (Se sigue de un argumento standard en teoría de complejidad algebraica que la computación de la complejidad no escalar A puede ser normalizada de tal forma que la nueva computación tenga longitud A^2 aproximadamente. Aquí se consideran todas las operaciones, inclusive las lineales. Sin embargo las operaciones constantes que requieren acceso a elementos del cuerpo de base son "gratis". En este caso el cuerpo de base es $K = k(X_1,...,X_i)$ y las constantes pueden ser precomputadas por un cálculo de evaluación de longitud L. Obviamente todo el procedimiento puede ser paralelizado. Para los detalles, véase [Stoss,1989].)

Ahora se concatena β_1 con el cálculo de evaluación β_2 en $k(X_1,...,X_i)$ que computa las matrices $M_{X_1},...,M_{X_n}$ y que se obtiene como consecuencia del Lema 3.4.2. El cálculo de evaluación β_2 puede obtenerse en tiempo secuencial d^{cn} y tiempo paralelo $c\,n^2 \log^2 d$ a partir de la entrada $f_1,...,f_{n-i}$ y tiene longitud d^{cn} y profundidad $c\,n^2 \log^2 d$, donde $c > 0$ es una constante suficientemente grande.

Sea β' el cálculo de evaluación en $k(X_1,...,X_i)$ que se obtiene concatenando β_1 y β_2. Luego β' calcula los coeficientes de la matriz M_g a partir de las entradas $X_1,...,X_i$ en tiempo secuencial $L + c\,A^2\,D^3 + d^{cn}$ y tiempo paralelo $\ell + c\,(\lambda \log D + n^2 \log^2 d)$. En las mismas cotas de tiempo (aunque no uniformes), β' puede obtenerse a partir de β por medio de una red aritmética sobre k. □

La proposición siguiente y su corolario son una herramienta esencial para la demostración del teorema principal.

Proposición 3.5.5. Sean a , f , r polinomios en $k[X_1,...,X_n]$ tales que r = a f vale en B. Supóngase que f no es divisor de cero en la A-álgebra B. Supóngase también que se tiene un cálculo de evaluación β en $k(X_1,...,X_i)[X_{i+1},...,X_n)$ que calcula f y r de longitud L, longitud no escalar A, profundidad ℓ y profundidad no escalar λ, con la condición adicional que β sólo contiene divisiones por elementos de $k(X_1,...,X_i)$. Sea T una nueva indeterminada y sea $\chi_a = T^D + \alpha_{D-1} T^{D-1} + ... + \alpha_0$ el polinomio característico del endomorfismo K-lineal a de B. Entonces $\alpha_{D-1},...,\alpha_0$ son polinomios de A = $k[X_1,...,X_i]$ de grado a lo sumo $i\, d^{2(n-i)}$ deg a que pueden ser computados por un cálculo de evaluación β' en $k(X_1,...,X_i)$ de longitud $L + c\, A^2\, D^3 + d^{cn}$ y profundidad $\ell + c\,(\lambda \log D + n^2 \log^2 d)$, con c > 0 una constante adecuada que no depende de $K = k(X_1,...,X_i)$. Además, β' puede obtenerse a partir de β en tiempo secuencial no uniforme $L + c\, A^2\, D^3 + d^{cn}$ y tiempo paralelo $\ell + c\,(\lambda \log D + n^2 \log^2 d)$ por medio de una red aritmética sobre k.

Demostración:

Obsérvese primero que por el Lema 3.4.3 las funciones racionales $\alpha_{D-1},...,\alpha_0$ de $K = k(X_1,...,X_i)$ son de hecho polinomios de $k[X_1,...,X_i]$ de grado a lo sumo $i\, d^{2(n-i)}$ deg a.

Aplicando el Lema 3.5.4 al cálculo β se obtiene otro calculo β_1 en $k(X_1,...,X_i)$ de longitud $L + c\, A^2\, D^3 + d^{cn}$ y profundidad $\ell + c\,(\lambda \log D + n^2 \log^2 d)$ que evalúa las entradas de las matrices M_r y M_f. (Aquí y en lo que sigue c > 0 será una constante adecuada que no depende del cuerpo $K = k(X_1,...,X_i)$.) Como f no es un divisor de cero en la A-álgebra B, el K-endomorfismo de B' correspondiente es inversible. Esto significa que la matriz $M_f \in K^{D \times D}$ es regular.

Primero se computan los coeficientes de $M_{f^{-1}}$ y luego los de $M_a = M_r\, M_{f^{-1}}$ a partir de los coeficientes de M_r y M_f por medio de un cálculo de evaluación

β_2 en $k(X_1,...,X_i)$ de longitud c D^5 y profundidad c $\log^2 D$ (esto es una consecuencia del algoritmo de Berkowitz. Compárese también con [von zur Gathen, 1986], Theorem 5.2).

Finalmente, se computan los polinomios $\alpha_{D-1},...,\alpha_0 \in k[X_1,...,X_i]$, que son los coeficientes del polinomio característico $\chi_a - \chi_{M_a}$ a partir de las entradas de M_a por medio del algoritmo de Berkowitz ([Berkowitz, 1984]). Esto se hace por medio de un cálculo de evaluación β_3 en $k(X_1,...,X_i)$ de longitud c D^5 y profundidad c $\log^2 D$.

Concatenando los cálculos de evaluación β_1, β_2 y β_3 y teniendo en cuenta que $D \leq d^n$, se obtiene un cálculo de evaluación β' en $k(X_1,...,X_i)$ de longitud $L + c\,A^2\,D^3 + d^{cn}$ y profundidad $\ell + c\,(\lambda \log D + n^2 \log^2 d)$ que calcula los polinomios $\alpha_{D-1},...,\alpha_0$.

Por construcción es evidente que β' puede calcularse a partir de β en tiempo secuencial no uniforme $L + cA^2 D^3 + d^{cn}$ y tiempo paralelo $\ell + c(\lambda \log D + n^2 \log^2 d)$ por medio de una red aritmética sobre k. □

Combinando la Proposición 3.5.5 con una versión local del Nullstellensatz efectivo de Caniglia (ver [Caniglia,1989] o [Caniglia *et al.*, 1991] se obtiene el siguiente

Corolario 3.5.6. Sean f y r polinomios de $k[X_1,...,X_n]$ tales que r es divisible por f en B. Supóngase que deg f ≤ d y que f no es divisor de cero en B. Supóngase también que se tiene un cálculo de evaluación β en $k(X_1,...,X_i)[X_{i+1},...,X_n]$ que computa al polinomio r de longitud L, longitud no escalar A, profundidad ℓ y profundidad no escalar λ y además que β contiene divisiones sólo por elementos de $k(X_1,...,X_i)$. Entonces existe un polinomio a de $k[X_1,...,X_n]$ tal que su grado satisface

$$\deg a \leq i\,(d^{2(n-i)} \deg r + d^{3(n-i)+1}) + D\,(d + \deg r)$$

y tal que, en B, vale la igualdad

$$r^D = a\,f.$$

Más aún, el polinomio a puede computarse por medio de un cálculo de evaluación β' en $k(X_1,...,X_i)[X_{i+1},...,X_n]$ de longitud $2L + c\,A^2\,D^3 + d^{cn}$, longitud no escalar $L + 2D + d^{c(n-i)}$, profundidad $2\ell + c\,(\lambda \log D + n^2 \log^2 d)$ y profundidad no escalar $\lambda + 2\log D + c\,(n-i)\log d$, donde $c > 0$ es una constante adecuada que no depende de $k(X_1,...,X_i)[X_{i+1},...,X_n]$.

El cálculo de evaluación β' contiene divisiones sólo por elementos de $k(X_1,...,X_i)$ y β' puede obtenerse a partir de β en tiempo secuencial no uniforme $2L + c\,A^2\,D^3 + d^{cn}$ y tiempo paralelo $2\ell + c\,(\lambda \log D + n^2 \log^2 d)$ por medio de una red aritmética sobre k.

Demostración:

Antes de comenzar con el argumento principal de la demostración obsérvese lo siguiente:

Como f es un polinomio de $k[X_1,...,X_n]$ de grado a lo sumo d, tiene en representación densa a lo sumo ed^n coeficientes. Por lo tanto f puede ser evaluado por medio de un cálculo de evaluación sin divisiones β_1 en $k[X_1,...,X_n]$ (y por lo tanto en $k(X_1,...,X_i)[X_{i+1},...,X_n]$) de longitud d^{cn}, longitud no escalar $d^{c(n-i)}$, profundidad cn log d y profundidad no escalar c(n-i) log d , donde $c > 0$ es una constante adecuada que no depende de $k[X_1,...,X_n]$.

Supóngase primero que f es una unidad de B. Sea T una nueva indeterminada y sea $\chi_f = T^D + \varphi_{D-1}\,T^{D-1} + ... + \varphi_0$ el polinomio característico del endomorfismo K-lineal f de B'. Del Lema 3.4.3 se deduce que $\varphi_{D-1},...,\varphi_0$ son polinomios en $A = k[X_1,...,X_i]$ de grado a lo sumo $i\,d^{2(n-i)+1}$. En particular

el polinomio φ_0 es el determinante del endomorfismo inversible A-lineal f del A-módulo libre B (compárese con el Corolario 3.3.2). Por lo tanto φ_0 es una unidad de $k[X_1,...,X_i]$. Esto significa que φ_0 es un elemento no nulo de k. Sea

$$q := (-\varphi_0)^{-1}(f^{D-1} + \varphi_{D-1} f^{D-2} + ... + \varphi_1).$$

Se ve que q es un polinomio en $k[X_1,...,X_n]$ de grado a lo sumo i $d^{2(n-i)+1} + D\ d$. De la proposición 3.5.5 y el hecho de que $D \leq d^n$ se deduce inmediatamente que $\varphi_{D-1},...,\varphi_0$ pueden ser evaluados por medio de un cálculo de evaluación en $k(X_1,...,X_i)$ de longitud d^{cn} y profundidad $c\ n^2 \log^2 d$, donde $c > 0$ es una constante adecuada que no depende de $k(X_1,...,X_i)$.

Esto implica que q puede computarse por medio de un cálculo de evaluación β^* en $k(X_1,...,X_i)[X_{i+1},...,X_n]$ de (asintóticamente) la misma longitud y profundidad (adaptando la constante $c > 0$). El cálculo de evaluación β^* sólo contiene divisiones por elementos de $k(X_1,...,X_i)$. El teorema de Hamilton-Cayley asegura que

$$(1) \quad f^D + \varphi_{D-1} f^{D-1} + ... + \varphi_0 = 0$$

vale en B. De (1) se deduce que la igualdad

$$(2) \quad f\ q = 1$$

vale en B.

Sea a el polinomio de $k[X_1,...,X_n]$ dado por $a := r^D\ q$. De (2) se infiere inmediatamente que la igualdad

$$r^D = a\ f$$

vale en B y que el grado de a satisface

$$\deg a \leq i\, d^{2(n-i)+1} + D\,(d + \deg r) \leq i\,(d^{2(n-i)} \deg r + d^{3(n-i)+1}) + D\,(d + \deg r).$$

El polinomio a de $k[X_1,...,X_n]$ puede ser evaluado por medio de un cálculo de evaluación β' en $k(X_1,...,X_i)[X_{i+1},...,X_n]$ de longitud $L + d^{cn} \leq 2L + c\,A^2\,D^3 + d^{cn}$, longitud no escalar $L + 2D + d^{c(n-i)}$, profundidad $\ell + c\,n^2 \log^2 d \leq 2\ell + c\,(\lambda \log D + n^2 \log^2 d)$ y profundidad no escalar $\lambda + 2\log D + c\,(n-i) \log d$, donde $c > 0$ es una constante adecuada que no depende de $k(X_1,...,X_i)[X_{i+1},...,X_n]$.

El cálculo de evaluación β' sólo contiene divisiones en $k(X_1,...,X_i)$ y β' puede obtenerse a partir de β por un circuito aritmético sobre k en tiempo secuencial no uniforme $L + d^{cn} \leq 2L + c\,A^2\,D^3 + d^{cn}$ y tiempo paralelo $\ell + c\,n^2 \log^2 d \leq 2\ell + c\,(\lambda \log D + n^2 \log^2 d)$.

Supóngase ahora que f no es una unidad en B. Luego la variedad algebraica $W := V \cap (f=0)$ no es vacía. Como f no es divisor de cero en B la variedad W es de dimensión $i - 1 \geq 0$. Por otro lado W es la subvariedad de $\mathbb{A}^n$ definida por los $n - i + 1$ polinomios $f_1,...,f_{n-i},f$. Teniendo en cuenta que $\dim W = i - 1$ se deduce del teorema de Macaulay (ver [Matsumura, 1989], Theorem 17.7) que $f_1,...,f_{n-i},f$ forman una sucesión regular en $k[X_1,...,X_n]$. Como $B = k[X_1,...,X_n]/(f_1,...,f_{n-i})$ y f divide a r en B, se concluye que el polinomio f pertenece al ideal $(f_1,...,f_{n-i},f)$ de $k[X_1,...,X_n]$.

Luego existen polinomios $q_1,...,q_{n-i}$ y q de $k[X_1,...,X_n]$ que satisfacen la igualdad

$$(3) \qquad r = q_1 f_1 + ... + q_{n-i} f_{n-i} + q f$$

y el grado la acotación

$$\max (\deg (q_1 f_1), \ldots , \deg (q_{n-i} f_{n-i}), \deg (q f)) \leq \deg r + (\max (3,d))^{n-i+1}$$

(Esto es una consecuencia de la versión local del Nullstellensatz mencionado anteriormente. Ver [Caniglia *et al.*, 1991], Theorem 1.3 y Corollary 3.3, o [Dickenstein *et al.*, 1988], Theorem 5.1.)

Teniendo en cuenta la suposición $d \geq n$ (ver Sección 1.1) y tratando el caso $n = 2$ separadamente se ve de inmediato que se puede suponer sin pérdida de generalidad $\deg (q f) \leq \deg r + d^{n-i+1}$.

Por otro lado se deduce de la igualdad (3) que

$$(4) \qquad r = q f$$

vale en B. Sea T una nueva variable y sea

$$\chi_q = T^D + \mu_{D-1} T^{D-1} + \ldots + \mu_0$$

el polinomio característico del endomorfismo K-lineal q de B'.

De la Proposición 3.5.5 se deduce que $\mu_{D-1},\ldots,\mu_0$ son polinomios de $k[X_1,\ldots,X_i]$ de grados a lo sumo $i\,(d^{2(n-i)} \deg r + d^{3(n-i)+1})$ que pueden ser evaluados por medio de un cálculo de evaluación β_2 en $k(X_1,\ldots,X_i)$ de longitud $L + c\,A^2\,D^3 + d^{cn}$ y profundidad $\ell + c\,(\lambda \log D + n^2 \log^2 d)$, donde $c > 0$ es una constante adecuada que no depende de $k(X_1,\ldots,X_i)$.

Del Teorema de Hamilton-Cayley se concluye que la igualdad

$$(5) \qquad q^D + \mu_{D-1} q^{D-1} + \ldots + \mu_0 = 0$$

vale en B. Combinando (4) y (5) se obtiene la igualdad

$$(6) \qquad r^D + \mu_{D-1} f\, r^{D-1} + \ldots + \mu_0 f^D = 0.$$

Se considera el polinomio $a := -(\mu_{D-1} r^{D-1} + \mu_{D-2} f\, r^{D-2} + \ldots + \mu_0 f^{D-1})$ en $k[X_1,\ldots,X_n]$. De (6) se deduce que la igualdad

$$r^D = a f$$

vale en B.

Más aún el grado del polinomio a está acotado por $i(d^{2(n-i)} \deg r + d^{3(n-i)+1}) + D(d + \deg r)$ y a puede computarse a partir de las entradas f, r y $\mu_{D-1},\ldots,\mu_0$ por medio de un cálculo de evaluación β_3 en $k[X_1,\ldots,X_n]$ sin divisiones de longitud 3D, longitud no escalar 2D, profundidad $2 + 3 \log D$ y profundidad no escalar $1 + 2 \log D$.

Concatenando los cálculos de evaluación β, β_1, β_2 y β_3 se obtiene un cálculo de evaluación β' en $k(X_1,\ldots,X_i)[X_{i+1},\ldots,X_n]$ de longitud $2L + c A^2 D^3 + d^{cn}$, longitud no escalar $L + 2D + d^{c(n-i)}$, profundidad $2\ell + c(\lambda \log D + n^2 \log^2 d)$ y profundidad no escalar $\lambda + 2\log D + c(n-i) \log d$ que computa al polinomio a de $k[X_1,\ldots,X_n]$. (Aquí, $c > 0$ es nuevamente una constante adecuada que no depende de $k(X_1,\ldots,X_i)[X_{i+1},\ldots,X_n]$.) Más aún, β' sólo contiene divisiones por elementos de $k(X_1,\ldots,X_i)$ y β' puede obtenerse a partir de β por una red aritmética sobre k de tamaño $2L + c A^2 D^3 + d^{cn}$ y profundidad $2\ell + c(\lambda \log D + n^2 \log^2 d)$. □

3.6. Demostración del resultado principal.

En esta sección se mantendrán las siguientes notaciones y convenciones

ya introducidas en 3.2 y 3.3:

Sean $f_1,...,f_{n+1}$ una familia de polinomios de $k[X_1,...,X_n]$ que representan una entrada normalizada para el algoritmo que se va a diseñar. Supóngase que los grados de $f_1,...,f_{n+1}$ están acotados por un entero d que satisface también $d \leq n$ y que $f_1,...,f_{n+1}$ generan el ideal trivial de $k[X_1,...,X_n]$.

Para cada $0 \leq i \leq n-1$, se considerarán los items siguientes:

$A_i := k[X_1,...,X_i]$; $K_i := k(X_1,...,X_i)$; $\mathfrak{I}_i$ y $\mathfrak{I}_i'$ los ideales generados por $f_1,...,f_{n-i}$ en $k[X_1,...,X_n]$ y en $K[X_{i+1},...,X_n]$ respectivamente; $Bi := k[X_1,...,X_n]/\mathfrak{I}_i$, $B_i' := K[X_{i+1},...,X_n]/\mathfrak{I}_i'$; $Di := \dim_{K_i} B_i'$ y $V_i := V(f_1,...,f_{n-i}) = V(\mathfrak{I}_i)$, la subvariedad algebraica cerrada de $\mathbb{A}^n$ definida por los polinomios $f_1,...,f_{n-i}$.

Además supóngase que para $0 \leq i \leq n-1$ se satisfacen las siguientes condiciones:

- $\mathfrak{I}_i$ y $\mathfrak{I}_i'$ son ideales radicales
- V_i es una variedad algebraica suave de dimensión i
- las variables $X_1,...,X_n$ están en posición de Noether con respecto a V_i.

Para cada $0 \leq i \leq n-1$ y para cada polinomio g de $k[X_1,...,X_n]$ se consideran los homomorfismos canónicos de k-álgebras $\bar{}\ : k[X_1,...,X_n] \rightarrow B_i$ y $\bar{}\ : K[X_{i+1},...,X_n] \rightarrow B_i'$ y se nota g a la imagen de g en B_i y en B_i' así como también el endomorfismo A_i-lineal de B_i y el endomorfismo K_i-lineal de B_i' inducido por la multiplicación por g.

Con estas notaciones y suposiciones se ha demostrado que el resultado principal, es decir el teorema de la introducción, es una consecuencia directa de la proposición siguiente:

Proposición 3.6.1. Existe un cálculo de evaluación β en $k(X_1,...,X_n)$ de longitud $d^{O(n)}$ y profundidad $O(n^4 \log^2 d)$ que calcula polinomios $p_1,...,p_{n+1}$ de $k[X_1,...,X_n]$ de grado $d^{O(n^2)}$ tales que

$$1 = p_1 f_1 + ... + p_{n+1} f_{n+1}.$$

El cálculo de evaluación β puede obtenerse en tiempo secuencial no uniforme $d^{O(n)}$ y en tiempo paralelo $O(n^4 \log^2 d)$ por medio de una red aritmética sobre k.

Demostración:

Esta demostración se basa en una aplicación iterada del Corolario 3.4.6 y sigue la línea de [Caniglia-Galligo-Heintz, 1988, 1989].

Para cada $1 \le i \le n$ se construye recursivamente un cálculo de evaluación $\beta^{(i)}$ en $k(X_1,...,X_{i-1})[X_i,...,X_n]$ que sólo contiene divisiones por elementos de $k(X_1,...,X_{i-1})$ y que calcula polinomios $a_{n+1}^{(i)},...,a_{n-i+2}^{(i)}$ de $k[X_1,...,X_n]$ tales que

$$r_i := a_{n+1}^{(i)} f_{n+1} + ... + a_{n-i+2}^{(i)} f_{n-i+2} - 1$$

pertenece al ideal $\mathfrak{I}_{i-1} = (f_1,...,f_{n-i+1})$. El cálculo de evaluación $\beta^{(i)}$ sera realizado por una red aritmética sobre k de tamaño L_i y profundidad ℓ_i. Sin pérdida de generalidad se puede suponer que L_i y ℓ_i acotan también la longitud y la profundidad del cálculo de evaluación $\beta^{(i)}$. Además sea A_i la longitud no escalar y λ_i la profundidad no escalar de $\beta^{(i)}$ y sea $d_i := \max(\deg a_{n+1}^{(i)},...,\deg a_{n-i+2}^{(i)})$. Entonces para $1 \le i < n$ las cantidades L_i, A_i, ℓ_i, λ_i y d_i satisfacen las siguientes fórmulas de recursión:

(1) $L_1 \le d^{cn}$, $L_{i+1} \le 3L_i + c\, A_i^2\, D_i^3 + d^{cn}$

(2) $A_1 \le 2D_1 + 1$, $A_{i+1} \le 2A_i + (3i + 2)\, D_i + d^{c(n-i)}$

(3) $\ell_1 \le c\, n^2 \log^2 d$, $\ell_{i+1} \le \ell_i + c\,(\lambda_i \log D_i + n^2 \log^2 d)$

(4) $\lambda_1 \le 2 \log D_1$, $\lambda_{i+1} \le \lambda_i + 6 \log D_i + c\, n \log d$

(5) $d_1 \le (D_1 + 1)\, d$, $d_{i+1} \le i\,(d^{2(n-i)}\, d^i + d^{3(n-i)+1}) + D_i\,(d + d_i)$

(aquí $c > 0$ es una constante adecuada que no depende de $k(X_1,\ldots,X_i)[X_{i+1},\ldots,X_n]$ y $1 \le i < n$). Teniendo en cuenta que $d \ge n$ y que $D_i \le d^{n-i} \le d^n$ vale para todo $0 \le i \le n$, se pueden hacer las siguientes conclusiones.

De (2) y (4) se infiere que para $1 \le i \le n$:

$$(6) \quad \Lambda_i \le 2^n\,(2d^{n-1} + 1) + \sum_{1 \le i \le n} (3i + 7)\, d^{n-i} + \sum_{1 \le i \le n} d^{c(n-i)} \qquad y$$

$$\lambda_i \le c\, n^2 \log d + 8 \sum_{1 \le i \le n} \log D_i \le c'\, n^2 \log d,$$

donde $c' > 0$ es una constante adecuada.

De (1), (3) y (6) se infiere que para $1 \le i < n$:

$$(7) \quad L_{i+1} \le 3\, L_i + c\, d^{2c'n} D_i^3 + d^{cn} \le 3\, L_i + d^{c''n} \qquad y$$

$$\ell_{i+1} \le \ell_i + c(c'\, n^2 \log d \log D_i + n^2 \log^2 d) \le \ell_i + c''\, n^3 \log^2 d,$$

donde $c'' > 0$ es una constante adecuada.

De (1), (3) y (7) se deduce que para $1 \le i \le n$:

$$(8) \quad L_i \le 3^n\, d^{cn} + n\, d^{c''n} \le d^{c'''n} \qquad y$$

$$\ell_i \le c\, n^2 \log^2 d + c''\, n^4 \log^2 d \le c'''\, n^4 \log^2 d,$$

donde $c''' > 0$ es una constante adecuada.

Luego (8) implica

$$(9) \qquad L_n = d^{O(n)} \quad y$$
$$\ell_n = O(n^4 \log^2 d).$$

De forma similar se concluye que

$$(10) \qquad d_n = d^{O(n^2)}$$

Se considerará primero el caso i = 1:

Se tiene que $\mathfrak{I}_{i-1} = \mathfrak{I}_0 = (f_1,...,f_n)$. Luego $\mathfrak{I}_{i-1}$ es un ideal de $k[X_1,...,X_n]$ de altura n que define una subvariedad de $\mathbb{A}^n$ de dimensión cero $V_{i-1} = V_0 = V(f_1,...,f_n)$. Más aún, f_{n+1} no se anula en todos los puntos de V_{i-1}. Esto significa que f_{n+1} es una unidad de $B_1 = B_1 = k[X_1,...,X_n]/(f_1,...,f_n)$. En particular f_{n+1} no es un divisor de cero de B_1 y divide al elemento 1 de B_1. Entonces por el Corolario3.5.6 existe un polinomio de $k[X_1,...,X_n]$ $a_{n+1}^{(1)}$ de grado $d_1 \leq (D_1 + 1)\, d$ tal que $1 = a_{n+1}^{(1)} f_{n+1}$ vale en B_1. Esto significa que $r_1 := a_{n+1}^{(1)} f_{n+1} - 1$ pertenece a $\mathfrak{I}_0 = (f_1,...,f_n)$. Más aún, el polinomio $a_{n+1}^{(1)}$ puede evaluarse por medio de un cálculo de evaluación $\beta^{(1)}$ en $k[X_1,...,X_n]$ de longitud $L_1 \leq d^{cn}$, longitud no escalar $A_1 \leq 2\, D_1 + 1$, profundidad $\ell_1 \leq c\, n^2 \log^2 d$ y profundidad no escalar $\lambda_1 \leq 2 \log D_1$, donde $c > 0$ es una constante adecuada que no depende de $k[X_1,...,X_n]$. El cálculo de evaluación $\beta^{(1)}$ puede obtenerse en tiempo secuencial no uniforme L_1 y en tiempo paralelo ℓ_1 por medio de una red aritmética sobre k.

Supóngase ahora que $1 \leq i < n$ y que se tiene un cálculo de evaluación $\beta^{(i)}$ en $k(X_1,...,X_{i-1})[X_i,...,X_n]$ que sólo contiene divisiones por elementos de

$k(X_1,...,X_{i-1})$ y que calcula polinomios $a_{n+1}^{(i)},...,a_{n-i+2}^{(i)}$ de $k[X_1,...,X_n]$ tales que

$$r_i := a_{n+1}^{(i)} f_{n+1} + ... + a_{n-i+2}^{(i)} f_{n-i+2} - 1$$

pertenece al ideal $\mathfrak{I}_{i-1} = (f_1,...,f_{n-i+1})$

Sea L_i la longitud, A_i la longitud no escalar, ℓ_i la profundidad y λ_i la profundidad no escalar de $\beta^{(i)}$ y sea $d_i := \max(\deg a_{n+1}^{(i)},...,\deg a_{n-i+2}^{(i)})$. Supóngase que $\beta^{(i)}$ puede obtenerse por medio de una red aritmética sobre k de tamaño L_i y profundidad ℓ_i. Primero se observa que r_i puede calcularse por medio de un cálculo de evaluación β' en $k(X_1,...,X_{i-1})[X_i,...,X_n]$ de longitud $L_i + 2i + d^{cn}$, longitud no escalar $A_i + i + d^{c(n-i)}$, profundidad $\ell_i + \log i + c\,n \log d$ y profundidad no escalar $\lambda_i + c\,(n - i) \log d$. El cálculo de evaluación β' sólo contiene divisiones por elementos de $k(X_1,...,X_{i-1})$.

El polinomio r_i pertenece al ideal $\mathfrak{I}_{i-1} = (f_1,...,f_{n-i},f_{n-i+1})$. Por lo tanto f_{n-i+1} divide a r_i en la A_i-álgebra $B_i = k[X_1,...,X_n]/(f_1,...,f_{n-i})$. Por otro lado, la variedad $V_{i-1} = V(\mathfrak{I}_{i-1}) = V((f_1,...,f_{n-i+1}))$ es un subconjunto cerrado de dimensión i-1 del espacio afín $\mathbb{A}^n$. Luego, por el Teorema de Macaulay ([Matsumura, 1989], Theorem 17.7), la familia $f_1,...,f_{n-i},f_{n-i+1}$ forma una sucesión regular del álgebra de polinomios $k[X_1,...,X_n]$. Esto implica que f_{n-i+1} no es un divisor de cero de B_i. Entonces, de nuevo por el Corolario 3.5.6, se concluye que existe un polinomio a de $k[X_1,...,X_n]$ tal que

$$\deg a \le i\,(d^{2(n-i)}\,d_i + d^{3(n-i)+1} + D_i\,(d + d_i)$$

que satisface

$$(11) \qquad r_i^{D_i} - a\,f_{n-i+1}$$

en B_i. Más aun, el polinomio a puede evaluarse por medio de un cálculo de

evaluación β'' en $k(X_1,...,X_i)[X_{i+1},...,X_n]$ de longitud $2L_i + c\,\Lambda_i^2 D_i^3 + d^{cn}$, longitud no escalar $\Lambda_i + 2D_i + d^{c(n-i)}$, profundidad $2\,\ell_i + c\,(\lambda_i \log D_i + n^2 \log^2 d)$ y profundidad no escalar $\lambda_i + 2 \log D_i + c\,(n - i) \log d$.

El cálculo de evaluación β'' sólo contiene divisiones por elementos de $k(X_1,...,X_i)$ y puede obtenerse a partir de $\beta^{(i)}$ por medio de una red aritmética sobre k de tamaño $2L_i + c\Lambda_i^2 D_i^3 + d^{cn}$ y profundidad $2\,\ell_i + c\,(\lambda_i \log D_i + n^2 \log^2 d)$. (Aquí nuevamente $c > 0$ es una constante adecuada que no depende de $k(X_1,...,X_i)[X_{i+1},...,X_n]$ y $1 \le i < n$.)

De (11) se concluye que el polinomio

$$r := r_i^{D_i} - a\, f_{n-i+1}$$

pertenece al ideal $\mathfrak{I}_i = (f_1,...,f_{n-i})$.

Como $r_i := a_{n+1}^{(i)} f_{n+1} + ... + a_{n-i+2}^{(i)} f_{n-i+2} - 1$, donde $a_{n+1}^{(i)},...,a_{n-i+2}^{(i)}$ son polinomios en $k[X_1,...,X_n]$ de grado a lo sumo d_i que pueden ser evaluados por el cálculo de evaluación $\beta^{(i)}$, se concluye que $r := r_i^{D_i} - a\, f_{n-i+1}$ tiene la forma

$$r := a_{n+1} f_{n+1} + ... + a_{n-i+2} f_{n-i+2} - a\, f_{n-i+1} + (-1)^{D_i},$$

donde $a_{n+1},...,a_{n-i+2}$ son polinomios en $k[X_1,...,X_n]$ de grado a lo sumo $D_i\, d_i$. Más aún, es posible calcularlos a partir de $f_{n+1},..., f_{n-i+2}$ y $a_{n+1}^{(i)},...,a_{n-i+2}^{(i)}$ por medio de un cálculo de evaluación β''' sin divisiones en $k[X_1,...,X_n]$ de longitud $4i\, D_i$, longitud no escalar $3i\, D_i$, profundidad $5 \log D_i + 2$ y profundidad no escalar $4 \log D_i + 1$. Si D_i es impar sea

$$a_{n+1}^{(i+1)} := a_{n+1},..., a_{n-i+2}^{(i+1)} := a_{n-i+2},\ a_{n-i+1}^{(i+1)} := -a\ ;$$

y si D_i es par sea

$$a_{n+1}^{(i+1)} := -a_{n+1}, \ldots, a_{n-i+2}^{(i+1)} := -a_{n-i+2}, a_{n-i+1}^{(i+1)} := a.$$

En ambos casos

$$r_{i+1} := a_{n+1}^{(i+1)} f_{n+1} + \ldots + a_{n-i+1}^{(i+1)} f_{n-i+1} - 1$$

pertenece a $\mathfrak{I}_i = (f_1,\ldots,f_{n-i})$ y los grados de $a_{n+1}^{(i+1)},\ldots,a_{n-i+1}^{(i+1)}$ están acotados por la cantidad

$$i\,(d^{2(n-i)}\,d_i + d^{3(n-i)+1} + D_i\,(d + d_i).$$

Por lo tanto se tiene que $d_{i+1} \leq i\,(d^{2(n-i)}\,d_i + d^{3(n-i)+1} + D_i\,(d + d_i)$. Concatenando $\beta^{(i)}$, β', β'' y β''' se obtiene un cálculo de evaluación $\beta^{(i+1)}$ en $k(X_1,\ldots,X_i)[X_{i+1},\ldots,X_n]$ que sólo contiene divisiones por elementos de $k(X_1,\ldots,X_i)$ y que computa los polinomios $a_{n+1}^{(i+1)},\ldots,a_{n-i+1}^{(i+1)}$ de $k[X_1,\ldots,X_n]$. El cálculo de evaluación $\beta^{(i+1)}$ tiene longitud $L_{i+1} \leq 3L_i + c\Lambda_i^2 D_i^3 + d^{cn}$, longitud no escalar $\Lambda_{i+1} \leq 2\Lambda_i + (3i+2)\,D_i + d^{c(n-i)}$, profundidad $\ell_{i+1} \leq 2\,\ell_i + c\,(\lambda_i \log D_i + n^2 \log^2 d)$ y profundidad no escalar $\lambda_{i+1} \leq \lambda_i + 6 \log D_i + c\,n \log d$, donde $c > 0$ es una constante adecuada que no depende de $k(X_1,\ldots,X_i)[X_{i+1},\ldots,X_n]$ y $1 \leq i < n$. Más aún, $\beta^{(i+1)}$ puede obtenerse por una red aritmética sobre k de tamaño L_{i+1} y profundidad ℓ_{i+1}.

En particular, $\beta^{(n)}$ es un cálculo de evaluación en $k(X_1,\ldots,X_{n-1})[X_n]$ (y por lo tanto en $k(X_1,\ldots,X_n)$) de longitud $L_n = d^{O(n)}$ y profundidad $\ell_n = O(n^4 \log^2 d)$ que calcula polinomios $a_{n+1}^{(n)},\ldots,a_2^{(n)}$ en $k[X_1,\ldots,X_n]$ de grado $d^{O(n^2)}$ tales que $r_n := a_{n+1}^{(n)} f_{n+1} + \ldots + a_2^{(n)} f_2 - 1$ es divisible por f_1 en $k[X_1,\ldots,X_n]$.

Sean $p_1 := r_n/f_1$, $p_2 := a_2^{(n)}$,..., $p_{n+1} := a_{n+1}^{(n)}$. Luego $p_1,\ldots,p_{n+1}$ son

polinomios en $k[X_1,...,X_n]$ de grado $d^{O(n^2)}$ que pueden calcularse por medio de un cálculo de evaluación en $k(X_1,...,X_n)$ de longitud $d^{O(n)}$ y profundidad $O(n^4 \log^2 d)$. Más aún, $p_1,...,p_{n+1}$ satisfacen la condición

$$1 = p_1 f_1 + ... + p_{n+1} f_{n+1}. \quad \square$$

El resultado principal (el teorema de la introducción) es ahora una consecuencia inmediata de la Proposición 3.6.1 y del Lema 3.3.1.

Referencias

S. J. BERKOWITZ, On computing the determinant in small parallel time using a small number of processors, Inform. Process. Lett. 18 (1984) 147-150.

W. S. BROWN, On Euclid's algorithm and the computation of polynomial greatest common divisors, J. ACM 18 (1971) 478-504.

D. BROWNAWELL, Bounds for the degrees in the Nullstellensatz, Ann. of Math. Second Series, Vol 126 Nº 3 (1987) 577-591.

L. CANIGLIA, Complejidad de algoritmos en geometría computacional, Tesis de doctorado, Universidad de Buenos Aires, (1989).

L. CANIGLIA, A. GALLIGO y J. HEINTZ, Borne simple exponentielle pour les degrés dans le théorème des zéros sur un corps de charactéristique quelconque, C. R. Acad. Sci. Paris, Série I Math. t. 307 (1988) 255-258.

L. CANIGLIA, A. GALLIGO y J. HEINTZ, Some new effectivity bounds in computational geometry, en Applied Algebra, Algebraic Algorithms and Error-Correcting Codes, Proc. 6th Intern. Conf. AAECC-6, Rome 1988, T. Mora, ed., Springer LN Comput. Sci. 357 (1989) 131-151.

L. CANIGLIA, J. A. GUCCIONE y J. J. GUCCIONE, Local membership problems for polynomial ideals, en Effective Methods in Algebraic Geometry, Proc. Intern. Conf. MEGA 90, Castiglioncello 1990, T. Mora y C. Traverso, eds., Progress in Mathematics Vol 94, Birkhäuser(1991) 35-41.

A. DICKENSTEIN, N. FITCHAS, M. GIUSTI y C. SESSA, The membership problem for unmixed ideals is solvable in single exponential time, a aparecer en Discrete Appl. Math., Special Issue 7th Intern. Conf. Applied Algebra, Algebraic Algorithms and Error-Correcting Codes AAECC-7, Tolouse 1989.

N. FITCHAS y A. GALLIGO, Nullstellensatz effectif et Conjecture de Serre (Théorème de Quillen-Suslim) pour le Calcul Formel, Math. Nachr. 149 (1990) 231-253.

N. FITCHAS, A. GALLIGO y J. MORGENSTERN, Algotirhmes rapides en

séquentiel et parallèle pour l'élimination des quantificateurs en géométrie élémentaire, Séminaire sur les structures algébriques ordonnées 1984-1987 vol. I, Publications de l'Université Paris VII (F. Delon, M. Dickmann y D. Gondard eds.), 32 (1990), 103-145.

J. VON ZUR GATHEN, Parallel arithmetic computations: a survey, en Proc. 13th Symp. MFCS 1986, Springer LN Comput. Sci. 233 (1986) 93-112.

M. GIUSTI y J. HEINTZ, Algorithmes -disons rapides- pour la décomposition d'une variété algébrique en composantes irréductibles et équidimensionelles, en Effective Methods in Algebraic Geometry, Proc. Intern. Conf. MEGA 90, Castiglioncello 1990, T. Mora y C. Traverso, eds., Progress in Mathematics Vol 94, Birkhäuser(1991 a) 169-193.

M. GIUSTI y J. HEINTZ, La détermination des points isolés et de la dimension d'une variété algebrique peut se faire en temps polynomial, Manuscript Centre de Mathematiques, Ecole Polytechnique, Palaiseau (1991 b).

J. HEINTZ, Definability and fast quantifier elimination over algebraically closed fields, Theoret. Comput. Sci. 24 (1983) 239-277.

J. HEINTZ, On the computational complexity of polynomials and bilinear mappings. A survey, en Applied Algebra, Algebraic Algorithms and Error-Correcting Codes, Proc. 5th Intern. Conf. AAECC-5, Menorca 1987, L Huguet y A. Poli, eds., Springer LN Comput. Sci. 356 (1989) 269-300.

J. HEINTZ y P. SCHNORR, Testing polynomials which are easy to compute, en Logic and Algorithmic. Un simposio en honor de Ernst Specker., Monographie de l'Enseignement Mathématique, Genève 1982, 237-254; también publicado en: 12th Annual ACM Symp, Theory of Computing (1980) 262-268.

J. HEINTZ y M. SIEVEKING, Absolute primality of polynomials is decidable in random polynomial time in the number of variables, Proc. 8th Int. Coll. Automata, Languages and Programming ICALP '81, Springer LNCS 115 (1981), 16-28.

J. P. HENRY y M. MERLE, Conditions de régularité et éclatements, Annales de l'Institut Fourier, Grenoble, 37, 3 (1987), 159-190.

J. P. JOUANOLOU, Théorèmes de Bertini et applications, Progress in Mathematics, Vol. 42. Birkhäuser 1983.

E. KALTOFEN, Greatest common divisors of polynomials given by straight line programms, Journal ACM 35 Nº 1 (1988), 231-264.

H. KOBAYASHI, S. MORITSUGU y R. W. HOGAN, On solving systems of algebraic equations, en Proc. Intern. Symp. on Symbolic and Algebraic Computation, ISSAC 88, Rome 1988, P. Gianni, ed., Springer LN Comput. Sci. 358 (1989).

J. KOLLAR, Sharp effective Nullstellensatz, J. AMS 1 (1988) 963-975.

T. Y. LAM, Serre's Conjecture, Springer LN Math. 635 (1978).

H. MATSUMURA, Commutative ring theory, Cambridge Studies in Advanced Mathematics Vol. 8, Cambridge University Press, Cambridge 1989.

K. MULMULEY, A fast parallel algorithm to compute the rank of a matrix over an arbitrary field, Proc. 18th Ann. Symp. Theory of Computing (1986) 338-339.

F. ROSSI y W. SPANGHER, Some effective methods in the opennes of loci for Cohen-Macaulay and Gorenstein properties, en Effective Methods in Algebraic Geometry, Proc. Intern. Conf. MEGA 90, Castiglioncello 1990, T. Mora y C. Traverso, eds., Progress in Mathematics Vol 94, Birkhäuser (1990) 441-455.

H.-J. STOSS, On the representation of rational functions of bounded complexity, Theoret. Comput. Sci. 64 (1989) 1-13.

V. STRASSEN, Berechnung und Programm I, Acta Inform. 1 (1972) 320-334.

B. TEISSIER, Résultats récents d'algebre commutative effective, en Séminaire Bourbaki Volumen 1989/90, Exposé 718, novembre 1989, Asterisque Vol. 189-190 (1991)107-131.

L. G. VALIANT, S. SKYUM, S. BERKOWITZ y C. RACKOFF, Fast parallel computation of polynomials using few processors, SIAM J. Comput., Vol. 12, Nº 4 (1983) 641-644.