FACULTAD DE MATEMÁTICAS

DEPARTAMENTO DE ÁLGEBRA

Trabajo Fin de Grado

Bases de Gröbner:
Eliminación y programación lineal entera

Alba González Parra

Dirigido por:
D. Francisco Jesús Castro Jiménez

Sevilla, Junio 2015.

Abstract

Gröbner basis is a basic concept in Computational Algebra; it was introduced by the Austrian mathematician Bruno Buchberger in 1965. A Gröbner basis of an ideal in a polinomial ring with coefficients over a field is a special generator set of the ideal. These bases have a very useful properties and applications.

In this project we deal with the theory of Gröbner bases, starting with the definition of the concept, studying their main properties and explaining Buchberger's algorithm for their computation. Then we give some applications.

The main objective of Elimination theory is the resolution of systems of polynomials equations. We use the Elimination theorem and the Extension theorem repeatedly, so at each step we only have to solve equations that depend on a finite subset of the original variables. The simplest case is when, at each step, we only need to solve equations depending on one single variable, but unfortunately this is not always posible.

We also include here a geometric interpretation of Elimination theory, the Closure theorem, and its proof, being the main result in this subject.

Finally, we give an application of Gröbner bases theory to Integer Lineal Programming. Our ultimate aim is to provide an algorithm whose input is a problem of Integer Lineal Programming, say in its standard form, and by using Gröbner basis, the algorithm returns, as an output, an optimal solution if it exists, or otherwise, the algorithm informs us that the problem has no solution.

Índice general

Introducción

El anillo de polinomios en una variable x con coeficientes en un cuerpo $\mathbb{K}$, $\mathbb{K}[x]$, es un dominio de ideales principales, de hecho es un dominio euclídeo, y por ello, todo ideal puede ser generado por un único elemento. Esto y el algoritmo de división euclídea, nos permite saber fácilmente si dado un polinomio $f \in \mathbb{K}[x]$, f está en un ideal dado o no, es decir se puede decidir la pertenencia de un polinomio a un ideal. Contamos con un orden en los monomios y un algoritmo de división en $\mathbb{K}[x]$ entre otras cosas.

La primera idea de este trabajo es adaptar estos conceptos al anillo de polinomios en varias variables $\mathbb{K}[x_1, \ldots, x_n]$. Se dedica la primera parte del *Capítulo 2* a definir un orden en los monomios de $\mathbb{K}[x_1, \ldots, x_n]$ y un algoritmo de división que de alguna manera sea consistente con lo ya conocido en $\mathbb{K}[x]$. Posteriormente llegaremos al objeto principal de estudio de este trabajo, las *bases de Gröbner*; que se puede considerar una noción básica del álgebra computacional.

El matemático austriaco Bruno Buchberger introdujo el concepto de *base de Gröbner* en su tesis doctoral en el año 1965. En este trabajo y en otros posteriores desarrolló entre otras muchas cosas el *Algoritmo de Buchberger*, que permite calcular una base de Gröbner de cualquier ideal de $\mathbb{K}[x_1, \ldots, x_n]$ a partir de un sistema finito de generadores del ideal. El nombre de base de Gröbner se debe a su director de tesis Wolfgang Gröbner.

Se introduce detalladamente el concepto de *base de Gröbner* a lo largo del *Capítulo 2*, veremos que podemos calcular una base de Gröbner de todo ideal de $\mathbb{K}[x_1, \ldots, x_n]$, y cómo se caracterizan estas bases. Parte de su importancia radica en que usando bases de Gröbner podemos solucionar el problema de la pertenencia de un polinomio a un ideal dado. Este capítulo está basado casi por completo en el capítulo 2 de [2]; aunque también me he apoyado en el capítulo 1 de [4].

Una vez familiarizados con el concepto de base de Gröbner, dedicaremos el resto del trabajo a ver algunas de sus muchas aplicaciones.

El propósito de la *Teoría de Eliminación*, *Capítulo 3*, (para el cual hemos usando principalmente las secciones 1 y 2 del capítulo 3 de [2] de la bibliografía), es simplificar sistemas de ecuaciones polinómicas con coeficientes en un cuerpo algebraicamente cerrado. Lejos de dar un algoritmo para poder resolver completamente estos sistemas, nos hemos conformado con usar bases de Gröbner para simplificarlos. Coloquialmente lo que hemos hecho es ver una manera de escalonarlos. Para ello nos servimos de los *Teoremas de Eliminación y Extensión*, que usados iteradamente, nos permiten reducir considerablemente la tarea de resolver dichos sistemas. En primer lugar se calcula un sistema de ecuaciones polinómicas equivalentes, es decir, con las mismas soluciones, que viene dado por una base de Gröbner respecto a un orden monomial que llamaremos de eliminación. A partir de ahí en cada paso se intenta resolver ecuaciones que solo dependan de algunas variables del sistema (lo ideal es que en cada paso solo tengamos que resolver polinomios de una variable pero esto desafortunadamente no siempre es posible), de manera que tenemos las soluciones de estas variables. Usaremos entonces el Teorema de Extensión para

saber cuáles de estas soluciones pueden extenderse a soluciones del sistema original y cuáles no. Para las que sí podamos extenderlas, lo que se hace es sustituir las soluciones parciales, ya obtenidas, en las ecuaciones restantes. Y volver a repetir el proceso de nuevo con el Teorema de Eliminación para encontrar las soluciones de las siguientes variables.

La prueba del *Teorema de Extensión* se realizará en el *Capítulo 4*, ya que como era necesario definir el concepto de *resultante generalizada* únicamente para desarrollar esta prueba, hemos decidido incorporarla aparte. Este capítulo está basado en la sección 6 del capítulo 3 de [2].

También se incluye, en la parte dedicada a la *Teoría de Eliminación*, una interpretación geométrica de estos resultados siendo el *Teorema de la Clausura* el resultado principal. La última sección del *Capítulo 3* está dedicada por completo a la prueba de la segunda parte del *Teorema de la Clausura*, que la hemos hecho muy detalladamente dada su complejidad y su alto contenido en conceptos de geometría algebraica. Para llevar a cabo esta demostración nos hemos basado en la sección 6 del capítulo 5 de [2].

Por último hemos incorporado a este trabajo una primera aplicación de las bases de Gröbner a la resolución de problemas de *Programación Lineal Entera*, *Capítulo 5*, eligiendo en parte este tema por su gran aplicación práctica a muchas otras disciplinas. Para ello hemos usado [4] y principalmente el capítulo 8 de [3]. El objetivo de este capítulo es el desarrollo de un algoritmo que dado un problema de Programación Lineal Entera en forma estándar, nos devuelva una solución óptima, si existe; y en caso contrario nos informe de que no existe solución del problema. Por último se ilustra con un ejemplo sencillo el desarrollo de este algoritmo.

He intentado escribir este trabajo de la forma más autocontenida posible. Por ello he incorporado el *Capítulo 1* donde están recogidos muchos resultados que se usan posteriormente, pero que no he considerado necesario probar pues son resultados con los que estoy familiarizada ya que forman parte del temario impartido en el Grado en Matemáticas. Más detalles sobre estos resultados y sus pruebas se pueden encontrar en la bibliografía.

Muchos sistemas de cálculo simbólico tienen implementados paquetes relacionados con el uso de órdenes monomiales, algoritmo de división de polinomios y bases de Gröbner. En particular he utilizado *Maple* y *Sage* para llevar a cabo los cálculos en los ejemplos tratados. De hecho considero que gracias a ello he mejorado notablemente mis capacidades en el uso de estos programas. En el trabajo he insistido sobre todo en el desarrollo teórico y solo he incluido algunos ejemplos que he considerado importantes para entender determinados conceptos y resultados. Naturalmente en el transcurso del trabajo he utilizado los sistemas de cálculo simbólico antes citados para realizar otros muchos ejemplos que no he incluido aquí.

Al final del trabajo se incluyen las referencias bibliográficas detalladas más arriba que han servido para la escritura de este texto. Así mismo, la obra más usada ha sido [2], hasta el punto de que en cierto modo, originariamente la idea de este trabajo era plasmar el esfuerzo empleado en la compresión de parte de esta obra. También he usado de manera especial mis apuntes personales de algunas asignaturas de mis estudios de Grado, que por consejo de mi tutor he incluido en la bibliografía.

Quiero hacer notar también que empecé a estudiar temas relacionados con las bases de Gröbner, durante el curso 2013-14 como alumna interna del Departamento de Álgebra tutelada por el profesor D. Francisco Jesús Castro Jiménez. Durante ese periodo me dediqué principalmente a la lectura y comprensión de los capítulos 1, 2 y 3 de [2], y los capítulos 1 y 2 de [3]. También me gustaría hacer constar que empecé a escribir mis notas sobre estos temas, antes de cursar la asignatura de Álgebra, Combinatoria y Computación, [9], que se ha impartido en el segundo cuatrimestre del presente curso 2014-15.

Capítulo 1

Resultados Previos

En este trabajo vamos a usar el símbolo $\mathbb{N}$ para denotar al conjunto de los números enteros no negativos y $\mathbb{K}$ para denotar un cuerpo. De la misma forma, usaremos la palabra anillo para referirnos a un anillo conmutativo con elemento unidad. Salvo mención expresa de lo contrario todos los resultados enunciados y no probados en este capítulo han sido consultados en [1], [2] y [11].

Definición 1.1.1. *Un polinomio f en las variables $x_1, \ldots, x_n$ con coeficientes en $\mathbb{K}$ es una suma finita*

$$f = \sum_{\alpha} a_\alpha \boldsymbol{x}^\alpha$$

con $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$, $a_\alpha \in \mathbb{K}$ y $\boldsymbol{x}^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$. Cada sumando $a_\alpha \boldsymbol{x}^\alpha$ se llamará un monomio (o un término) de f. El grado de un monomio no nulo (es decir, $a_\alpha \boldsymbol{x}^\alpha$ con a_α no nulo) es la suma $|\alpha| := \alpha_1 + \cdots + \alpha_n$. El grado total (o simplemente el grado) de un polinomio no nulo f es el máximo de los grados de los monomios que lo forman.

Notaremos $\mathbb{K}[x_1, \ldots, x_n]$ al conjunto de todos los polinomios en las variables $x_1, \ldots, x_n$ con coeficientes en $\mathbb{K}$ que tiene estructura de anillo.

Teorema 1.1.2. (Teorema Fundamental del Álgebra)
Todo polinomio $f \in \mathbb{C}[x]$ de grado n tiene a lo más n raíces distintas en $\mathbb{C}$.

En realidad este resultado es cierto para todo cuerpo $\mathbb{K}$ algebraicamente cerrado.

Definición 1.1.3. *Sea A un anillo, A se dice noetheriano si todo ideal de A es finitamente generado.*

Teorema 1.1.4. (Teorema de la Base de Hilbert)
Si A es un anillo noetheriano entonces $A[x]$ es noetheriano.

Teorema 1.1.5. (Condición de Cadena Ascendente, CCA)
Sea A un anillo. Son equivalentes:

i) *A es noetheriano.*

ii) *Todo conjunto no vacío de ideales tiene un elemento maximal (para la inclusión).*

iii) *Toda cadena $I_1 \subseteq I_2 \subseteq I_3 \subseteq \ldots$ creciente de ideales es estacionaria.*

A continuación vamos a ver algunos conceptos básicos sobre geometría algebraica que usaremos fundamentalmente en el *Capítulo 3*.

Definición 1.1.6. *Un conjunto* $\mathrm{V} \subset \mathbb{A}^n(\mathbb{K})$ *es algebraico si existe* $S \subset \mathbb{K}[x_1, \ldots, x_n]$ *tal que* $\mathrm{V} = \mathcal{V}(S)$ *donde*

$$\mathcal{V}(S) = \{(a_1, \ldots, a_n) \in \mathbb{A}^n(\mathbb{K}) \mid f(a_1, \ldots, a_n) = 0 \ \forall f \in S\}.$$

Definición 1.1.7. *Dado un conjunto cualquiera* $Z \subset \mathbb{A}^n(\mathbb{K})$ *se define el ideal de* Z *como*

$$\mathcal{I}(Z) = \{f \in \mathbb{K}[x_1, \ldots, x_n] \mid f(a_1, \ldots, a_n) = 0 \ \forall (a_1, \ldots, a_n) \in Z\}.$$

La intersección de conjuntos algebraicos es un conjunto algebraico y la unión finita también lo es. Por esta razón se define una topología en $\mathbb{A}^n(\mathbb{K})$, la *Topología de Zariski*, donde los conjuntos algebraicos son los cerrados. Dado un conjunto $\mathrm{V} \subset \mathbb{A}^n(\mathbb{K})$, llamamos clausura de Zariski de V, $\overline{\mathrm{V}}$ al menor conjunto algebraico que contiene a V.

Además tenemos que $\mathcal{V}(S) = \mathcal{V}(\langle S \rangle)$ para cualquier conjunto $S \subset \mathbb{K}[x_1, \ldots, x_n]$; luego un ideal define un conjunto algebraico, y recíprocamente, un conjunto algebraico siempre viene definido por un ideal. Pero esta correspondencia no es biyectiva, pues hay ideales distintos que dan lugar al mismo conjunto algebraico.

Definición 1.1.8. *Sea* $I \subset \mathbb{K}[x_1, \ldots, x_n]$ *un ideal. Se define el radical de* I *como*

$$\sqrt{I} = \{f \in \mathbb{K}[x_1, \ldots, x_n] \mid f^m \in I \text{ para algún } m \geq 1\}.$$

Un ideal I *se dice radical si* $I = \sqrt{I}$.

Observación 1.1.9. *Se tiene* $\mathcal{V}(I) = \mathcal{V}(\sqrt{I})$ *para todo ideal* I.

Teorema 1.1.10. (Nullstellensatz o Teorema de los Ceros de Hilbert)
Supongamos que $\mathbb{K}$ *es un cuerpo algebraicamente cerrado. Sea* I *un ideal del anillo* $\mathbb{K}[x_1, \ldots, x_n]$. *Entonces* $\mathcal{I}(\mathcal{V}(I)) = \sqrt{I}$.

Como consecuencia del Nullstellensatz, si $\mathbb{K}$ es algebraicamente cerrado, existe una correspondencia biyectiva entre los ideales radicales de $\mathbb{K}[x_1, \ldots, x_n]$ y los conjuntos algebraicos de $\mathbb{A}^n(\mathbb{K})$.

La correspondencia viene dada por

$$\begin{array}{ccc} \text{Ideales Radicales} & \longleftrightarrow & \text{Conjuntos Algebraicos} \\ \text{de } \mathbb{K}[x_1, \ldots, x_n] & & \text{de } \mathbb{A}^n(\mathbb{K}) \\ I & \longmapsto & \mathcal{V}(I) \\ \mathcal{I}(\mathrm{V}) & \longleftarrow & \mathrm{V}. \end{array}$$

Además $\mathcal{V}(\mathcal{I}(\mathrm{V})) = \overline{\mathrm{V}}$ para todo $\mathrm{V} \subset \mathbb{A}^n(\mathbb{K})$; y se invierten las inclusiones, es decir,

$$I_1 \subset I_2 \Rightarrow \mathcal{V}(I_2) \subset \mathcal{V}(I_1)$$

$$\mathrm{V}_1 \subset \mathrm{V}_2 \Rightarrow \mathcal{I}(\mathrm{V}_2) \subset \mathcal{I}(\mathrm{V}_1).$$

Si $\mathbb{K}$ es un cuerpo cualquiera, y Z_1, Z_2 son dos subconjuntos de $\mathbb{A}^n(\mathbb{K})$, entonces $\mathcal{I}(Z_1 \cup Z_2) = \mathcal{I}(Z_1) \cap \mathcal{I}(Z_2)$.

Definición 1.1.11. *Sea* $\mathrm{V} \subset \mathbb{A}^n(\mathbb{K})$ *un conjunto algebraico. Se define el anillo de coordenadas de* V *como el anillo cociente*

$$\mathcal{A}(\mathrm{V}) = \frac{\mathbb{K}[x_1, \ldots, x_n]}{\mathcal{I}(\mathrm{V})}.$$

Entonces $\mathcal{A}(\mathrm{V})$ es una $\mathbb{K}$-álgebra finitamente generada y reducida. Es decir, $\mathcal{A}(\mathrm{V})$ además de estructura de anillo tiene estructura de espacio vectorial, y es finitamente generada como anillo y es reducida (i.e. no tiene elementos nilpotentes). Ambas estructuras, la de anillo y la de espacio vectorial, son compatibles para las operaciones correspondientes.

Definición 1.1.12. *Un conjunto algebraico* $\mathrm{V} \subset \mathbb{A}^n(\mathbb{K})$ *es irreducible si* $\forall X, Y \subset \mathbb{A}^n(\mathbb{K})$ *conjuntos algebraicos tales que* $\mathrm{V} = X \cup Y$ *es* $\mathrm{V} = X$ *o* $\mathrm{V} = Y$*. En tal caso decimos que* V *es una variedad algebraica.*

Si $\mathrm{V} \subset \mathbb{A}^n(\mathbb{K})$ es una variedad algebraica, $\mathcal{I}(\mathrm{V})$ es un ideal primo. Y recíprocamente, si $I \subset \mathbb{K}[x_1, \ldots, x_n]$ es un ideal primo, $\mathcal{V}(I)$ es una variedad algebraica. Notemos que si I es un ideal primo, I es radical, es decir, $I = \sqrt{I}$.

Teorema 1.1.13. *Todo conjunto algebraico* $\mathrm{V} \subset \mathbb{A}^n(\mathbb{K})$ *se descompone de manera única (salvo reordenación) como unión finita de variedades algebraicas donde ninguna de ellas está contenida en ninguna otra. Es decir, existe una única expresión*

$$\mathrm{V} = \mathrm{V}_1 \cup \cdots \cup \mathrm{V}_m$$

con $\mathrm{V}_i \subset \mathbb{A}^n(\mathbb{K})$ *variedad algebraica para* $i = 1, \ldots, m$*; y* $\mathrm{V}_i \nsubseteq \mathrm{V}_j \ \forall i \neq j$.

Capítulo 2

Bases de Gröbner

2.1. Órdenes monomiales

En esta sección vamos a definir un orden en los monomios del anillo de polinomios $\mathbb{K}[x_1, \ldots, x_n]$ ya que ello es necesario para definir posteriormente un algoritmo de división.

Si nos fijamos en casos particulares ya conocidos como el algoritmo de división de polinomios en $\mathbb{K}[x]$ o la eliminación Gaussiana para polinomios lineales en $\mathbb{K}[x_1, \ldots, x_n]$ implícitamente estamos utilizando un orden en los monomios. Para dividir polinomios de una variable, usamos el grado de los monomios, de modo que $\ldots > x^{n+1} > x^n > \ldots > x^2 > x > 1$, mientras que cuando hacemos el método de Gauss seguimos usualmente el orden de las variables $x_1 > x_2 > \ldots > x_n$. Ahora pretendemos ampliar este concepto al anillo de polinomios $\mathbb{K}[x_1, \ldots, x_n]$.

Sea $f \in \mathbb{K}[x_1, \ldots, x_n]$, entonces f es una suma finita de la forma

$$f = \sum_{\alpha} a_\alpha \mathbf{x}^\alpha = \sum_{\alpha} a_\alpha x_1^{\alpha_1} \cdots x_n^{\alpha_n}$$

con $a_\alpha \in \mathbb{K},\ \alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$.

Podemos establecer una correspondencia biunívoca entre los monomios mónicos de $\mathbb{K}[x_1, \ldots, x_n]$ y $\mathbb{N}^n$, de modo que si tenemos un orden establecido en $\mathbb{N}^n$, tendremos un orden en los monomios.

Definamos $\mathbb{T}^n = \{\mathbf{x}^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n} \mid \alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n\}$.

De todas las formas existentes de ordenar $\mathbb{T}^n$, nosotros tenemos que considerar aquellas que sean consistentes con el algoritmo de división en $\mathbb{K}[x]$ y de eliminación Gaussiana en $\mathbb{K}[x_1, \ldots, x_n]$.

Para ordenar los términos de un polinomio necesitamos una relación de orden total (es decir, donde dos elementos cualesquiera se pueden comparar). Y por último, cuando usemos un algoritmo de división, queremos acabar en un número finito de pasos, lo que nos lleva a la noción de un buen orden.

Definición 2.1.1. *Un orden de los términos en $\mathbb{T}^n$, o equivalentemente, un orden monomial en $\mathbb{K}[x_1, \ldots, x_n]$, es una relación de orden total $\leq$ en $\mathbb{T}^n$ tal que:*

i) $\leq$ *es un buen orden,*

ii) *Si* $\boldsymbol{x}^\alpha \leq \boldsymbol{x}^\beta \ \Rightarrow\ \boldsymbol{x}^\alpha \boldsymbol{x}^\gamma \leq \boldsymbol{x}^\beta \boldsymbol{x}^\gamma,\ \forall\, \boldsymbol{x}^\gamma \in \mathbb{T}^n$.

Recordemos que una relación de orden en $\mathbb{T}^n$ es una relación binaria que es reflexiva, antisimétrica y transitiva. Con un buen orden nos referimos a que todo subconjunto no vacío de $\mathbb{T}^n$, debe tener un elemento mínimo respecto a la relación de orden $\leq$. Escribimos $\mathbf{x}^\alpha < \mathbf{x}^\beta$ si $\mathbf{x}^\alpha \leq \mathbf{x}^\beta$ y $\alpha \neq \beta$.

Observación 2.1.2. *Si $\leq$ es un orden monomial en $\mathbb{K}[x_1, \ldots, x_n]$ $\Rightarrow$*

$$1 \leq \boldsymbol{x}^\alpha, \ \forall \alpha \in \mathbb{N}^n.$$

Demostración.

Reducción al Absurdo. Supongamos $\exists \alpha \in \mathbb{N}^n$ tal que $\mathbf{x}^\alpha \leq 1$ y $\alpha \neq \underline{0}$.

Multiplicamos a ambos lados de la desigualdad por $\mathbf{x}^\alpha$ y aplicamos *ii*), con lo que obtenemos $\mathbf{x}^{2\alpha} \leq \mathbf{x}^\alpha$. Multiplicando sucesivamente por $\mathbf{x}^\alpha$ y aplicando *ii*) en cada paso,

$$\mathbf{x}^\alpha \geq \mathbf{x}^{2\alpha} \geq \mathbf{x}^{3\alpha} \geq \ldots$$

Luego el conjunto $\{\mathbf{x}^\alpha, \mathbf{x}^{2\alpha}, \mathbf{x}^{3\alpha}, \ldots\}$ no tiene primer elemento, lo que hace que $\leq$ no sea un buen orden y tenemos una contradicción. □

Veamos una caracterización de buen orden.

Lema 2.1.3. *Una relación de orden total $\leq$ en $\mathbb{T}^n$ es un buen orden si y solo si toda sucesión estrictamente decreciente en $\mathbb{T}^n$ es finita.*

Demostración.

Probaremos el contrarecíproco: *Una relación de orden total $\leq$ en $\mathbb{T}^n$ no es un buen orden si y solo si existe una sucesión infinita estrictamente decreciente en $\mathbb{T}^n$.*

$\boxed{\Rightarrow}$ Si $\leq$ en $\mathbb{T}^n$ no es un buen orden $\Rightarrow \exists S \subset \mathbb{T}^n$, $S \neq \emptyset$, que no tiene un primer elemento. Sea $\mathbf{x}^{\alpha_1} \in S$. Como $\mathbf{x}^{\alpha_1}$ no es un mínimo en S, debe existir $\mathbf{x}^{\alpha_2} \in S$ con $\mathbf{x}^{\alpha_1} > \mathbf{x}^{\alpha_2}$.

De la misma forma, debe existir $\mathbf{x}^{\alpha_3} \in S$ tal que $\mathbf{x}^{\alpha_1} > \mathbf{x}^{\alpha_2} > \mathbf{x}^{\alpha_3}$. Así sucesivamente, podemos construir una sucesión estrictamente decreciente en $\mathbb{T}^n$ con infinitos términos.

$\boxed{\Leftarrow}$ Sea $\mathbf{x}^{\alpha_1} > \mathbf{x}^{\alpha_2} > \mathbf{x}^{\alpha_3} > \ldots$ sucesión estrictamente decreciente en $\mathbb{T}^n$ con infinitos términos. Tomando S el conjunto formado por los términos de esta sucesión, es claro que S no tiene un mínimo, y por tanto, la relación $\leq$ no es un buen orden.

□

Veamos ahora algunos órdenes monomiales que serán los que usaremos próximamente.

Definición 2.1.4. (Orden Lexicográfico)
Dados dos monomios distintos $\boldsymbol{x}^\alpha, \boldsymbol{x}^\beta \in \mathbb{K}[x_1, \ldots, x_n]$ con $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ y $\beta = (\beta_1, \ldots, \beta_n) \in \mathbb{N}^n$. Diremos que $\boldsymbol{x}^\alpha <_{lex} \boldsymbol{x}^\beta$ si el primer elemento no nulo (por la izquierda) del vector $\alpha - \beta \in \mathbb{Z}^n$ es negativo.

Ejemplo 1.

- *En el caso de 3 variables* $\mathbb{K}[x_1, x_2, x_3]$,
$$x_2^2 x_3^4 <_{lex} x_1 x_2 x_3^2$$
pues $(0,2,4) - (1,1,2) = (-1,1,2)$ *tiene el primer elemento negativo.*
- *Está claro que* $x_n <_{lex} x_{n-1} <_{lex} \ldots <_{lex} x_1$.
- *Si usamos letras para denotar las variables, por ejemplo,* $\mathbb{K}[x,y,z]$*, es importante fijar primero un orden entre estas variables. En nuestro caso,* $x > y > z$.

Proposición 2.1.5. *El orden lexicográfico es un orden monomial.*

Demostración.

El orden así definido establece una relación binaria que verifica la propiedad reflexiva, antisimétrica y transitiva.

- $<_{lex}$ es un orden total. Dados $\mathbf{x}^\alpha, \mathbf{x}^\beta$:
  Si $\alpha = \beta$ entonces $\mathbf{x}^\alpha = \mathbf{x}^\beta$.
  Si no, consideramos el vector $\alpha - \beta$ que tendrá al menos un elemento no nulo. Si el primer elemento no nulo es negativo, entonces $\mathbf{x}^\alpha <_{lex} \mathbf{x}^\beta$. Si el primer elemento no nulo es positivo, entonces el primer elemento no nulo de $\beta - \alpha$ es negativo y por lo tanto es $\mathbf{x}^\beta <_{lex} \mathbf{x}^\alpha$.
- Veamos que $<_{lex}$ es un buen orden, para ello basta ver que toda sucesión estrictamente decreciente en $\mathbb{T}^n$ es finita.
  Reducción al absurdo. Supongamos que existe una sucesión estrictamente decreciente infinita $\mathbf{x}^{\alpha_1} >_{lex} \mathbf{x}^{\alpha_2} >_{lex} \ldots$ Si consideramos la primera coordenada de los vectores α_i, éstos deben formar una sucesión decreciente (no necesariamente estrictamente decreciente) de elementos en $\mathbb{N}$. Como tenemos un buen orden en $\mathbb{N}$, debe existir k tal que a partir del elemento $\mathbf{x}^{\alpha_k}$ el grado de la primera variable no cambie. A partir del elemento $\mathbf{x}^{\alpha_k}$ de la sucesión, sucede algo similar con las segundas coordenadas de los vectores α_i; y por el mismo motivo también se estabiliza el grado de la segunda variable. Repitiendo este procedimiento y teniendo en cuenta que el número de variables es finito, llegamos a la conclusión de que la sucesión completa debe estabilizarse; y por tanto a una contradicción con nuestra hipótesis. Luego $<_{lex}$ es un buen orden.
- Si $\mathbf{x}^\alpha <_{lex} \mathbf{x}^\beta$, el primer elemento no nulo por la izquierda de $\alpha - \beta$ es negativo. Pero $(\alpha + \gamma) - (\beta + \gamma) = \alpha - \beta,\ \forall \gamma \in \mathbb{N}^n$. Luego $\mathbf{x}^\alpha \mathbf{x}^\gamma <_{lex} \mathbf{x}^\beta \mathbf{x}^\gamma,\ \forall \mathbf{x}^\gamma \in \mathbb{T}^n$.

□

Veamos otros órdenes monomiales.

Definición 2.1.6. (Orden Lexicográfico Graduado)
Dados dos monomios distintos $\boldsymbol{x}^\alpha, \boldsymbol{x}^\beta \in \mathbb{K}[x_1, \ldots, x_n]$ *con* $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ *y* $\beta = (\beta_1, ..., \beta_n) \in \mathbb{N}^n$*. Diremos que* $\boldsymbol{x}^\alpha <_{grlex} \boldsymbol{x}^\beta$ *si*

$$|\alpha| = \sum_{i=1}^n \alpha_i < \sum_{i=1}^n \beta_i = |\beta|, \ \ o\, si \ \ \left(|\alpha| = \sum_{i=1}^n \alpha_i = \sum_{i=1}^n \beta_i = |\beta| \right) \ y \ \left(\boldsymbol{x}^\alpha <_{lex} \boldsymbol{x}^\beta \right).$$

Definición 2.1.7. (Orden Lexicográfico Graduado Inverso)
Dados dos monomios $\boldsymbol{x}^\alpha, \boldsymbol{x}^\beta \in \mathbb{K}[x_1, \ldots, x_n]$ *con* $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ *y* $\beta = (\beta_1, \ldots, \beta_n) \in \mathbb{N}^n$. *Diremos que* $\boldsymbol{x}^\alpha <_{grevlex} \boldsymbol{x}^\beta$ *si*

$$|\alpha| = \sum_{i=1}^{n} \alpha_i < \sum_{i=1}^{n} \beta_i = |\beta|, \ \ o\, si \ \ |\alpha| = \sum_{i=1}^{n} \alpha_i = \sum_{i=1}^{n} \beta_i = |\beta|$$

y el primer elemento no nulo por la derecha del vector $\alpha - \beta \in \mathbb{N}^n$ *es positivo.*

Ejemplo 2. *Veamos algunos casos de comparación de monomios en* $\mathbb{K}[x_1, x_2, x_3]$ *con los órdenes vistos anteriormente:*

- $x_1x_2^2x_3^2 <_{lex} x_1^2x_2x_3$ *porque* $(1,2,2)-(2,1,1) = (-1,1,1)$ *tiene el primer elemento negativo. Este orden se distingue muy fácilmente, solo hay que fijarse que la primera variable tenga menor grado en el monomio menor. Caso de que la primera variable tenga el mismo grado en ambos monomios, nos fijaremos en que la segunda variable tenga menor grado en el monomio menor y así sucesivamente.*

- $x_1^2x_2x_3 <_{grlex} x_1x_2^2x_3^2$ *y* $x_1^2x_2x_3 <_{grevlex} x_1x_2^2x_3^2$ *porque* $|(2,1,1)| = 4 < 5 = |(1,2,2)|$. *En estos dos órdenes en lo primero que hay que fijarse es en el grado total del monomio, y será menor el monomio con menor grado.*

- $x_1x_2^3x_3 <_{grlex} x_1^2x_2x_3^2$ *pues el grado total de los monomios es el mismo pero* $x_1x_2^3x_3 <_{lex} x_1^2x_2x_3^2$ *ya que* $(1,3,1)-(2,1,2) = (-1,2,-1)$ *que tiene el primer elemento negativo. En este orden, caso de que los grados totales de los monomios sean los mismos, los comparamos siguiendo el orden lexicográfico.*

- $x_1^2x_2x_3^2 <_{grevlex} x_1x_2^3x_3$ *pues el grado total de los monomios es el mismo pero* $(2,1,2) - (1,3,1) = (1,-2,1)$ *tiene el último elemento positivo. Cuando los grados totales de los monomios coinciden, los comparamos siendo el monomio menor el que tenga grado mayor en la última variable. Caso de que tengan el mismo grado en la última variable, será menor el que tenga grado mayor en la penúltima variable y así sucesivamente.*

Proposición 2.1.8. *El orden lexicográfico graduado es un orden monomial.*

Demostración.

El orden así definido establece una relación binaria que verifica la propiedad reflexiva, antisimétrica y transitiva.

- $<_{grlex}$ es un orden total. Dados $\mathbf{x}^\alpha, \mathbf{x}^\beta$:

  Si $\alpha = \beta$ entonces $\mathbf{x}^\alpha = \mathbf{x}^\beta$.

  Si $|\alpha| < |\beta|$ entonces $\mathbf{x}^\alpha <_{grlex} \mathbf{x}^\beta$.

  Si $|\alpha| > |\beta|$ entonces $\mathbf{x}^\alpha >_{grlex} \mathbf{x}^\beta$.

  Si $|\alpha| = |\beta|$ entonces aplicamos el orden lexicográfico que ya sabemos que es un orden total.

- $<_{grlex}$ es un buen orden.

  Reducción al absurdo. Supongamos que existe una sucesión estrictamente decreciente infinita $\mathbf{x}^{\alpha_1} >_{grlex} \mathbf{x}^{\alpha_2} >_{grlex} \ldots$ Se tiene $|\alpha_i| \geq |\alpha_{i+1}|$, $\forall i$. Además, la sucesión debe estabilizarse pues $|\alpha_i| \in \mathbb{N}$ donde tenemos un buen orden. Luego debe existir $\mathbf{x}^{\alpha_j}$ a partir del

cual los monomios se comparan usando el orden lexicográfico, que ya sabemos que es un buen orden. Por lo que de una forma u otra la sucesión debe estabilizarse. Lo que es una contradicción. Por tanto, $<_{grlex}$ es un buen orden.

- Sea $\mathbf{x}^\alpha <_{grlex} \mathbf{x}^\beta$. Sea $\gamma \in \mathbb{N}^n$.

  Si $|\alpha| < |\beta| \Rightarrow |\alpha+\gamma| = |\alpha| + |\gamma| < |\beta| + |\gamma| = |\beta+\gamma|$ ya que α, β, γ son vectores de enteros no negativos. Luego $\mathbf{x}^\alpha\mathbf{x}^\gamma <_{grlex} \mathbf{x}^\beta\mathbf{x}^\gamma$. Si $|\alpha| = |\beta|$, entonces $|\alpha+\gamma| = |\beta+\gamma|$. Debe ser $\mathbf{x}^\alpha <_{lex} \mathbf{x}^\beta$ y por tanto $\mathbf{x}^\alpha\mathbf{x}^\gamma <_{lex} \mathbf{x}^\beta\mathbf{x}^\gamma$, lo que implica que $\mathbf{x}^\alpha\mathbf{x}^\gamma <_{grlex} \mathbf{x}^\beta\mathbf{x}^\gamma$.

□

Nota. *Una prueba muy similar a la anterior puede hacerse en el caso del orden lexicográfico graduado inverso.*

Por último, vamos a fijar un poco más de notación para que una vez tengamos fijado un orden monomial $\leq$ en $\mathbb{K}[x_1, \ldots, x_n]$ podamos ordenar los términos de cualquier polinomio respecto a dicho orden.

Consideramos $f \in \mathbb{K}[x_1, \ldots, x_n]$ no nulo y un orden monomial $\leq$ en $\mathbb{K}[x_1, \ldots, x_n]$. Entonces escribiremos f ordenadamente de la forma:

$$f = a_1\mathbf{x}^{\alpha_1} + \cdots + a_r\mathbf{x}^{\alpha_r}$$

donde $a_i \in \mathbb{K}$, $a_i \neq 0$, $\mathbf{x}^{\alpha_i} = x_1^{\alpha_{i_1}} \cdots x_n^{\alpha_{i_n}}$ con $(\alpha_{i_1}, \ldots, \alpha_{i_n}) \in \mathbb{N}^n$ para $i = 1, \ldots, r$ tal que $\mathbf{x}^{\alpha_1} > \mathbf{x}^{\alpha_2} > \ldots > \mathbf{x}^{\alpha_r}$.

Definición 2.1.9. *Sea $f \in \mathbb{K}[x_1, \ldots, x_n]$, $f \neq 0$ y en las condiciones anteriores.*

i) El monomio líder de f (aunque no sea necesariamente mónico) es $\mathrm{ml}(f) = \boldsymbol{x}^{\alpha_1}$.

ii) El coeficiente líder de f es $\mathrm{cl}(f) = a_1$.

iii) El término líder de f es $\mathrm{tl}(f) = a_1\boldsymbol{x}^{\alpha_1}$.

Tanto $\mathrm{ml}(f)$ *como* $\mathrm{cl}(f)$ *y* $\mathrm{tl}(f)$ *dependen del orden monomial $\leq$ elegido en $\mathbb{N}^n$.*

Observación 2.1.10. *Sean $f_1, \ldots, f_m \in \mathbb{K}[x_1, \ldots, x_n]$ no nulos. Fijado un orden monomial en $\mathbb{K}[x_1, \ldots, x_n]$; si* $\mathrm{ml}(f_i) \neq \mathrm{ml}(f_j)$ *para todo $1 \leq i < j \leq m$, se verifica*

i) $f_1 + \cdots + f_m \neq 0$.

ii) $\mathrm{ml}(f_1 + \cdots + f_m) = \mathrm{máx}_i\{\mathrm{ml}(f_i)\}$.

Esto se tiene porque $\mathrm{máx}_i\{\mathrm{ml}(f_i)\}$ *no se va a anular con ningún otro término de $f_1 + \cdots + f_m$ ya que todos los monomios de los $f_1, \ldots, f_m$ son menores que* $\mathrm{máx}_i\{\mathrm{ml}(f_i)\}$.

2.2. Algoritmo de división de polinomios

Nuestro objetivo es ahora describir un algoritmo de división en $\mathbb{K}[x_1, \ldots, x_n]$ que extienda la división de polinomios en $\mathbb{K}[x]$ y la eliminación Gaussiana para polinomios lineales.

Consideremos $f, f_1, \ldots, f_s \in \mathbb{K}[x_1, \ldots, x_n]$ con $f_i \neq 0$, $1 \leq i \leq s$ y un orden monomial $\leq$ en $\mathbb{K}[x_1, \ldots, x_n]$; queremos dividir f entre la s-upla de polinomios $F = (f_1, \ldots, f_s)$. Es decir, expresar f de la forma:

$$f = h_1 f_1 + \ldots + h_s f_s + r$$

con $h_i, r \in \mathbb{K}[x_1, \ldots, x_n]$ cumpliendo las siguientes propiedades:

- $r = 0$ o $r \neq 0$ y $\mathrm{ml}(f_i)$ no divide a ningún monomio de r, donde $i = 1, \ldots, s$. (2.1)
- $\mathrm{ml}(f) = \max\{\max_{1 \leq i \leq s}\{\mathrm{ml}(h_i)\mathrm{ml}(f_i)\}, \mathrm{ml}(r)\}$. (2.2)

Algoritmo 1 Algoritmo de División

Entrada: $f, f_1, \ldots, f_s \in \mathbb{K}[x_1, \ldots, x_n]$ con cada $f_i \neq 0$.
Salida: $h_1, \ldots, h_s, r$ tales que $f = h_1 f_1 + \cdots + h_s f_s + r$ y verificando (2.1) y (2.2).
Inicialización: $h_i := 0, \ldots, h_s := 0$, $r := 0$ y definimos $p := f$
1: **mientras** $p \neq 0$ **hacer**
2: **si** $\exists i$ tal que $\mathrm{ml}(f_i)$ divide a $\mathrm{ml}(p)$ **entonces**
3: **devolver** elegimos i el menor índice cumpliendo 2 y definimos:

$$h_i := h_i + \frac{\mathrm{tl}(p)}{\mathrm{tl}(f_i)}, \qquad p := p - \frac{\mathrm{tl}(p)}{\mathrm{tl}(f_i)} f_i$$

4: **si no**
5: definimos:

$$r := r + \mathrm{tl}(p), \qquad p := p - \mathrm{tl}(p)$$

6: **fin si**
7: **fin mientras**

En el siguiente ejemplo vemos que la división depende del orden de los polinomios $f_1, ..., f_s$.

Ejemplo 3. *Consideremos $\mathbb{K}[x, y]$ y el orden lexicográfico $(x > y)$. Sean*

$$f_1 = xy - x, \quad f_2 = x - y^2, \quad f = x^2 y + xy + 1.$$

Notemos que $\mathrm{ml}(f_1) = xy$ y $\mathrm{ml}(f_2) = x$. Si dividimos f entre (f_1, f_2) obtenemos

$$f = (x + y + 2) f_1 + (x + 2) f_2 + (2y^2 + 1)$$

mientras que si dividimos f entre (f_2, f_1) obtenemos

$$f = 0 f_1 + (xy + y^3 + y) f_2 + (y^5 + y^3 + 1).$$

Teorema 2.2.1. *El algoritmo de división 1 es correcto.*

Demostración.

Veamos que el algoritmo termina. Fijémonos en que en cada etapa eliminamos el término líder de p, de forma que vamos construyendo una sucesión $p_1, p_2, \ldots$ donde p_{j+1} se obtiene de p_j sustrayendo $\mathrm{tl}(p_j)$, pues si existe $i, 1 \leq i \leq s$, tal que $\mathrm{ml}(f_i)$ divide a $\mathrm{ml}(p_j)$ entonces $p_{j+1} = p_j - \frac{\mathrm{tl}(p_j)}{\mathrm{tl}(f_i)} f_i$ (lo que elimina $\mathrm{tl}(p_j)$ de p_j y puede que algunos monomios más); y si no existe $i, 1 \leq i \leq s$, tal que $\mathrm{ml}(f_i)$ divide a $\mathrm{ml}(p_j)$, eliminamos $\mathrm{tl}(p_j)$ de p_j.

De modo que en cada etapa, $\mathrm{ml}(p_{j+1}) < \mathrm{ml}(p_j)$. Y como partimos de un orden monomial (que es un buen orden), la sucesión de los $p_j's$ cuyos términos líderes forman una sucesión de monomios estrictamente decreciente debe acabar. Por lo que nuestro algoritmo termina.

Probemos también que los polinomios h_i y r obtenidos en el algoritmo anterior verifican las propiedades (2.1) y (2.2).

Teniendo en cuenta lo anterior y que al principio del algoritmo $p = f$, en todos los pasos tenemos que $\mathrm{ml}(p) \leq \mathrm{ml}(f)$. Y como, para $i \in \{1, \ldots, s\}$ vamos obteniendo h_i añadiendo términos de la forma $\frac{\mathrm{tl}(p)}{\mathrm{tl}(f_i)}$ donde $\frac{\mathrm{tl}(p)}{\mathrm{tl}(f_i)} f_i$ cancela al término líder de p, tenemos que $\mathrm{ml}(h_i)\mathrm{ml}(f_i) \leq \mathrm{ml}(f)$.

Además r se obtiene añadiendo términos de la forma $\mathrm{tl}(p)$ en determinadas etapas, entonces $\mathrm{ml}(r) \leq \mathrm{ml}(f)$. Como $f = h_1 f_1 + \cdots + h_s f_s + r$ (que lo probaremos a continuación) debe darse la igualdad, y por tanto se tiene (2.2) y trivialmente (2.1).

Falta ver entonces que $f = h_1 f_1 + \cdots + h_s f_s + r$. Para ello basta probar que en cada paso del algoritmo tenemos $f = h_1 f_1 + \cdots + h_s f_s + p + r$ para un cierto polinomio p. Lo hacemos por *inducción.*

En el caso inicial está claro tal y como definimos los elementos en la inicialización del algoritmo. Suponemos que en una etapa intermedia del algoritmo se da la igualdad. Entonces, en el siguiente paso:

Si el término líder de algún f_i divide al término líder de p, tenemos la igualdad

$$h_i f_i + p = \left(h_i + \frac{\mathrm{tl}(p)}{\mathrm{tl}(f_i)}\right) f_i + \left(p - \frac{\mathrm{tl}(p)}{\mathrm{tl}(f_i)} f_i\right)$$

lo que muestra que $h_i f_i + p$ no ha cambiado. En caso contrario, cambian los términos de p y r, pero no de $p + r$ por lo que en cualquier caso la igualdad se mantiene del paso anterior. □

A continuación veamos una caracterización de la expresión $f = h_1 f_1 + \cdots + h_s f_s + r$ dada por el algoritmo. Sean $\mathrm{ml}(f_i) = \mathbf{x}^{\alpha(i)}$ y definimos

$$\begin{aligned}
\Delta_1 &= \alpha(1) + \mathbb{N}^n \\
\Delta_2 &= (\alpha(2) + \mathbb{N}^n) - \Delta_1 \\
\Delta_3 &= (\alpha(3) + \mathbb{N}^n) - (\Delta_1 \cup \Delta_2) \\
&\vdots \\
\Delta_s &= (\alpha(s) + \mathbb{N}^n) - \left(\textstyle\bigcup_{i=1}^{s-1} \Delta_i\right) \\
\overline{\Delta} &= \mathbb{N}^n - \left(\textstyle\bigcup_{i=1}^{s} \Delta_i\right)
\end{aligned}$$

Notemos que $\Delta_i \cap \Delta_j = \emptyset$ si $i \neq j$, $\overline{\Delta} \cap \Delta_i = \emptyset$ para $i = 1, \ldots, s$, y $\overline{\Delta} \sqcup \bigcup_{i=1}^{s} \Delta_i = \mathbb{N}^n$.

Definición 2.2.2. *Sea $f \in \mathbb{K}[x_1, \ldots, x_n]$ no nulo, $f = \sum_\alpha a_\alpha \boldsymbol{x}^\alpha$ con $a_\alpha \in \mathbb{K}$. Se define el diagrama de Newton de f como el conjunto*

$$Nw(f) = \{\alpha \in \mathbb{N}^n \mid a_\alpha \neq 0\}.$$

Teorema 2.2.3. (Caracterización de la división)
Dados $f \in \mathbb{K}[x_1,\ldots,x_n]$ *y* $F = (f_1,\ldots,f_s) \in \mathbb{K}[x_1,\ldots,x_n]^s$ *con* $f_i \neq 0$, $i = 1,\ldots,s$, *existe un único vector* $(h_1,\ldots,h_s,r) \in \mathbb{K}[x_1,\ldots,x_n]^{s+1}$ *cumpliendo las siguientes propiedades:*

i) $f = h_1f_1 + \cdots + h_sf_s + r$.

ii) *Para* $i = 1,\ldots,s$, $\alpha(i) + Nw(h_i) \subseteq \Delta_i$.

iii) *Si* $r \neq 0$, $Nw(r) \subseteq \overline{\Delta}$.

Demostración.

El vector que nos devuelve el algoritmo de división verifica *i)*. Vamos a comprobar entonces que también verifica *ii)* y *iii)* para asegurar la existencia. En la prueba del *Teorema 2.2.1* queda de manifiesto que los monomios $\mathbf{x}^\beta$ que aparecen en el cociente h_i satisfacen la condición $\beta + \alpha(i) \in \Delta_i$, dado que en cada paso del algoritmo se elige el menor índice i tal que $\mathrm{ml}(f_i)$ (que es $\mathbf{x}^{\alpha(i)}$) divide a $\mathrm{tl}(p)$. Se tiene por tanto

$$\alpha(i) + Nw(h_i) \subseteq \Delta_i.$$

Además si $r \neq 0$, sabemos que se tiene que ningún $\mathrm{ml}(f_i)$ divide a ningún monomio de r por (2.1). Luego $Nw(r) \subseteq \overline{\Delta}$. Y tenemos *ii)* y *iii)*.

Unicidad. En primer lugar observemos que:

$$\beta \in \Delta_i \;\Leftrightarrow\; \mathbf{x}^{\alpha(i)} \mid \mathbf{x}^\beta \;\; y \;\; \mathbf{x}^{\alpha(j)} \nmid \mathbf{x}^\beta \;\; \forall j < i$$

$$\gamma \in \overline{\Delta} \;\Leftrightarrow\; \mathbf{x}^{\alpha(i)} \nmid \mathbf{x}^\gamma \;\; \forall i = 1,\ldots,s.$$

Con estas observaciones probemos la unicidad. Sea

$$f = h_1f_1 + \cdots + h_sf_s + r$$

$$f = h_1'f_1 + \cdots + h_s'f_s + r'$$

por tanto tenemos que

$$0 = (h_1 - h_1')f_1 + \cdots + (h_s - h_s')f_s + (r - r').$$

Llamamos $A_i = h_i - h_i'$ y $R = r - r'$. Entonces tenemos

$$0 = A_1f_1 + \cdots + A_sf_s + R.$$

Como consecuencia de la *Observación 2.1.10* se tiene $A_1 = \cdots = A_s = R = 0$, dado que los monomios líder de los A_if_i y de R son distintos dos a dos supuesto que A_i es no nulo y R no nulo. Así probamos la unicidad.

□

Al polinomio r en la expresión de la división se le llama resto de la división de f entre F respecto del orden monomial $\leq$ y se denota $resto(f;F,\leq)$ o si no hay riesgo de confusión con el orden monomial $\leq$ se le llama forma reducida de f por F y se denota $resto(f;F)$ o incluso $\overline{f}^F$.

2.3. Bases de Gröbner

En esta sección vamos a definir el objeto central de este trabajo, que luego nos permitirá tratar algunas cuestiones referentes a geometría algebraica y optimización.

Dado un ideal $I \subset \mathbb{K}[x_1, \ldots, x_n]$, podemos asegurar la existencia de $f_1, \ldots, f_s \in \mathbb{K}[x_1, \ldots, x_n]$ tal que $I = \langle f_1, \ldots, f_s \rangle$ gracias al *Teorema 1.1.4*. El objetivo ahora es conseguir una base de polinomios, con unas propiedades especiales, que nos generen I.

Para toda la sección supondremos fijado un orden monomial $\leq$ con respecto al cual trabajaremos.

Definición 2.3.1. *Un ideal $I \subset \mathbb{K}[x_1, \ldots, x_n]$ es un ideal monomial si está generado por monomios. Lo denotaremos como $I = \langle \boldsymbol{x}^\alpha : \alpha \in A \rangle$ con $A \subset \mathbb{N}^n$.*

Definición 2.3.2. *Sea $M \subset \mathbb{K}[x_1, \ldots, x_n]$ (no necesariamente un ideal). Se define el ideal de términos líder de M como*

$$\mathrm{TL}(M) = \langle \mathrm{tl}(f) : f \in M \rangle.$$

Notemos que $\mathrm{TL}(M)$ es un ideal monomial.

Consideremos ahora $I \subset \mathbb{K}[x_1, \ldots, x_n]$ un ideal distinto de cero y su ideal de términos líder, $\mathrm{TL}(I)$. Supongamos $I = \langle f_1, \ldots, f_s \rangle$. Entonces, en general, los ideales monomiales $\langle \mathrm{tl}(f_1), \ldots, \mathrm{tl}(f_s) \rangle$ y $\mathrm{TL}(I)$ pueden ser distintos. Siempre tendremos que $\langle \mathrm{tl}(f_1), \ldots, \mathrm{tl}(f_s) \rangle \subset \mathrm{TL}(I)$; pero $\mathrm{TL}(I)$ puede ser estrictamente mayor.

Ejemplo 4. *En $\mathbb{R}[x_1, x_2]$ con $<_{lex}$ sea $I = \langle f_1, f_2 \rangle$ con $f_1 = x_1 + x_2$ y $f_2 = x_1 - x_2$. Tenemos que $\langle \mathrm{tl}(f_1), \mathrm{tl}(f_2) \rangle = \langle x_1 \rangle$, sin embargo probaremos más tarde (ver* Ejemplo 6*) que $\mathrm{TL}(I) = \langle x_1, x_2 \rangle$.*

Buscamos entonces unos generadores donde se dé la igualdad $\langle \mathrm{tl}(f_1), \ldots, \mathrm{tl}(f_s) \rangle = \mathrm{TL}(I)$.

Definición 2.3.3. *Sea $I \subset \mathbb{K}[x_1, \ldots, x_n]$ ideal no nulo. Un subconjunto de polinomios no nulos $G = \{g_1, \ldots, g_t\} \subset I$ es una base de Gröbner del ideal I respecto al orden monomial $\leq$ si*

$$\forall f \in I,\ f \neq 0,\ \exists i \in \{1, \ldots, t\} \ /\ \ \mathrm{ml}(g_i) \,|\, \mathrm{ml}(f).$$

Veamos que este conjunto G, supuesto que existe, tiene una serie de propiedades muy especiales.

Teorema 2.3.4. *Sea $I \subset \mathbb{K}[x_1, \ldots, x_n]$ un ideal distinto de cero y sea $G = \{g_1, \ldots, g_t\} \subset I$. Son equivalentes:*

i) G *es una base de Gröbner de* I.

ii) $\forall f \in \mathbb{K}[x_1, \ldots, x_n]$, $f \in I$ *si y solo si* $\overline{f}^G = 0$.

iii) $\forall f \in \mathbb{K}[x_1, \ldots, x_n]$, $f \in I$ *si y solo si* $\exists h_i \in \mathbb{K}[x_1, \ldots, x_n], i = 1, \ldots, t$ *tal que* $f = \sum_{i=1}^{t} h_i g_i$ *con* $\mathrm{ml}(f) = \mathrm{máx}_{1 \leq i \leq t}\{\mathrm{ml}(h_i)\mathrm{ml}(g_i)\}$.

iv) $\mathrm{TL}(G) = \mathrm{TL}(I)$.

Demostración.

$\boxed{i) \Rightarrow ii)}$ Supongamos que G es una base de Gröbner de I respecto al orden monomial fijado. Sea $f \in \mathbb{K}[x_1, \ldots, x_n]$. Por el algoritmo de la división, tenemos que

$$f = h_1 g_1 + \cdots + h_t g_t + r, \ \textit{donde} \ r = \overline{f}^G.$$

$\boxed{\Leftarrow}$ Si $\overline{f}^G = 0 \ \Rightarrow \ f = h_1 g_1 + \cdots + h_t g_t$, y como $G \subset I$ entonces $f \in I$.

$\boxed{\Rightarrow}$ $f - r = h_1 g_1 + \cdots + h_t g_t \in I$. Si $f \in I \ \Rightarrow \ r \in I$.
Si $r \neq 0$, por definición de base de Gröbner, $\exists i \in \{1, \cdots, t\}$ tal que $\mathrm{ml}(g_i) \,|\, \mathrm{ml}(r)$, y entonces r no sería el resto de la división por (2.1). Luego $r = 0$.

$\boxed{ii) \Rightarrow iii)}$ Por hipótesis se tiene $\forall f \in \mathbb{K}[x_1, \ldots, x_n]$, $f \in I \ \Leftrightarrow \ \overline{f}^G = 0$. Por la división se tiene para $f \in I$ que $f = h_1 g_1 + \cdots + h_t g_t$, y además $\mathrm{ml}(f) = \mathrm{máx}_{1 \leq i \leq t}\{\mathrm{ml}(h_i)\mathrm{ml}(g_i)\}$

$\boxed{iii) \Rightarrow iv)}$ Está claro que $\mathrm{TL}(G) \subset \mathrm{TL}(I)$ pues $G \subset I$. Para probar la otra contención es suficiente ver que $\forall f \in I$, $\mathrm{tl}(f) \in \mathrm{TL}(G)$ ya que $\mathrm{TL}(I)$ está generado por los $\mathrm{tl}(f)$ con $f \in I$.

Tenemos que $\mathrm{ml}(f) = \mathrm{máx}_{1 \leq i \leq t}\{\mathrm{ml}(h_i)\mathrm{ml}(g_i)\}$. Entonces $\exists j$ tal que $\mathrm{ml}(f) = \mathrm{ml}(h_j)\mathrm{ml}(g_j)$ y tenemos por tanto el resultado.

$\boxed{iv) \Rightarrow i)}$ Por hipótesis se tiene $\mathrm{TL}(G) = \mathrm{TL}(I)$. Sea $f \in I$, $f \neq 0$. Entonces

$$\mathrm{tl}(f) \in \mathrm{TL}(I) = \mathrm{TL}(G) = \langle \mathrm{tl}(g_i) : i = 1, \ldots, t \rangle.$$

Y por tanto debe existir un monomio $h \in \mathbb{K}[x_1, \ldots, x_n]$ tal que $\mathrm{tl}(f) = h \cdot \mathrm{tl}(g_i)$ para algún g_i. Con lo cual, $\mathrm{ml}(g_i) | \mathrm{ml}(f)$.

□

Observemos que de $iv)$ se obtiene que $\mathrm{TL}(I) = \langle \mathrm{tl}(g_1), \ldots, \mathrm{tl}(g_t) \rangle$ que no lo teníamos para bases en general. Además una base de Gröbner, si existe, es un conjunto generador finito de I, por el apartado $ii)$ del Teorema anterior.

Corolario 2.3.5. *Sea $f \in \mathbb{K}[x_1, ..., x_n]$ no nulo y $G = \{g_1, \ldots, g_t\}$ una base de Gröbner para un ideal $I \subset \mathbb{K}[x_1, \ldots, x_n]$. Entonces, $\overline{f}^G$ está unívocamente determinado (independientemente del orden de los elementos de G y más aún, independientemente de la base de Gröbner G si fijamos un orden monomial $\leq$).*

Demostración.

Sean $G = \{g_1, \ldots, g_t\}$ y $G' = \{g'_1, \ldots, g'_s\}$ dos bases de Gröbner del ideal I respecto del orden monomial $\leq$. Sabemos entonces que $\mathrm{TL}(I) = \mathrm{TL}(G) = \mathrm{TL}(G')$. Y por el algoritmo de división tenemos que

$$f = h_1 g_1 + \cdots + h_t g_t + r$$
$$f = h'_1 g'_1 + \cdots + h'_s g'_s + r'$$

Entonces

$$(h_1 g_1 + \cdots + h_t g_t) - (h'_1 g'_1 + \cdots + h'_s g'_s) = r' - r$$

Luego $r' - r \in I$, pero como ningún término de $r' - r$ está en $\mathrm{TL}(I)$, debe ser $r' - r = 0$ y por tanto $r = r'$. □

Teorema 2.3.6. *Todo ideal $I \subset \mathbb{K}[x_1, \ldots, x_n]$ distinto de $\{0\}$ tiene una base de Gröbner, respecto de un orden monomial fijado de antemano.*

Demostración.

Sea un ideal $I \subset \mathbb{K}[x_1, \ldots, x_n]$ distinto de $\{0\}$. Consideramos el ideal de términos líder de I, $\mathrm{TL}(I)$. $\mathrm{TL}(I) \subset \mathbb{K}[x_1, \ldots, x_n]$ y por el *Teorema 1.1.4*, $\exists g_1, \ldots, g_t \in I$ tal que $\mathrm{TL}(I) = \langle \mathrm{tl}(g_1), \ldots, \mathrm{tl}(g_t) \rangle$. Que $G = \{g_1, \ldots, g_t\}$ es una base de Gröbner del ideal I es consecuencia de la equivalencia $i)$ si y solo si $iv)$ en el *Teorema 2.3.4*. □

2.4. Algoritmo de Buchberger

Gracias al *Teorema 2.3.6* sabemos que todo ideal no nulo $I \subset \mathbb{K}[x_1, \ldots, x_n]$ tiene una base de Gröbner respecto de un orden monomial $\leq$ dado. En esta sección vamos a estudiar un algoritmo que nos permite calcularla.

Definición 2.4.1. *Sean $f, g \in \mathbb{K}[x_1, \ldots, x_n]$ no nulos. Sea $\boldsymbol{x}^\gamma = mcm(\mathrm{ml}(f), \mathrm{ml}(g))$. Se define el S-polinomio de f y g como*

$$S(f,g) = \frac{\boldsymbol{x}^\gamma}{\mathrm{tl}(f)} f - \frac{\boldsymbol{x}^\gamma}{\mathrm{tl}(g)} g.$$

Ejemplo 5. *Consideremos $<_{lex}$ en $\mathbb{R}[x, y]$ con $x > y$. Sean:*

$$f(x,y) = 2x^2y - 3x^2$$
$$g(x,y) = xy^3 + 2xy^2$$
$$\mathrm{tl}(f) = 2x^2y; \quad \mathrm{ml}(f) = x^2y$$
$$\mathrm{tl}(g) = xy^3; \quad \mathrm{ml}(g) = xy^3$$
$$\boldsymbol{x}^\gamma = mcm(x^2y, xy^3) = x^2y^3$$

$S(f,g) = \frac{x^2y3}{2x^2y} f - \frac{x^2y^3}{xy^3} g = \frac{1}{2}y^2(2x^2y - 3x^2) - x(xy^3 + 2xy^2) = x^2y^3 - \frac{3}{2}x^2y^2 - x^2y^3 - 2x^2y^2 = -\frac{7}{2}x^2y^2$

Lema 2.4.2. *Sean $f_1, \ldots, f_s \in \mathbb{K}[x_1, \ldots, x_n]$ no nulos tales que $\mathrm{ml}(f_i) = \boldsymbol{x}^\gamma$ con $i = 1, \ldots, s$. Sea $f = \sum_{i=1}^s c_i f_i$ con $c_i \in \mathbb{K}$, $i \in \{1, \ldots, s\}$. Si $\mathrm{ml}(f) < \boldsymbol{x}^\gamma \Rightarrow$*

f es una combinación lineal con coeficientes en $\mathbb{K}$ de los S-polinomios $S(f_i, f_j)$ con $1 \leq i < j \leq s$.

Además, $\mathrm{ml}(S(f_i, f_j)) < \boldsymbol{x}^\gamma$ para $1 \leq i < j \leq s$.

Demostración.

$f_i = a_i \mathbf{x}^\gamma + (\textit{términos menores})$; con $a_i \in \mathbb{K}$. Como

$$\mathrm{ml}(f) = \mathrm{ml}\left(\sum_{i=1}^s c_i f_i\right) = \mathrm{ml}\left(\sum_{i=1}^s c_i (a_i \mathbf{x}^\gamma + (\textit{términos menores}))\right) < \mathbf{x}^\gamma$$

$\Rightarrow \sum_{i=1}^{s} c_i a_i = 0$.

Por otro lado, $S(f_i, f_j) = \frac{\mathbf{x}^\gamma}{\mathrm{tl}(f_i)} f_i - \frac{\mathbf{x}^\gamma}{\mathrm{tl}(f_j)} f_j = \frac{1}{a_i} f_i - \frac{1}{a_j} f_j$.

Entonces;

$$f = c_1 f_1 + \cdots + c_s f_s = c_1 a_1 \left(\frac{1}{a_1} f_1\right) + \cdots + c_s a_s \left(\frac{1}{a_s} f_s\right) =$$

$$= c_1 a_1 \left(\frac{1}{a_1} f_1 - \frac{1}{a_2} f_2\right) + (c_1 a_1 + c_2 a_2) \left(\frac{1}{a_2} f_2 - \frac{1}{a_3} f_3\right) + (c_1 a_1 + c_2 a_2 + c_3 a_3) \left(\frac{1}{a_3} f_3 - \frac{1}{a_4} f_4\right) + \cdots$$

$$\cdots + (c_1 a_1 + \cdots + c_{s-1} a_{s-1}) \left(\frac{1}{a_{s-1}} f_{s-1} - \frac{1}{a_s} f_s\right) + (c_1 a_1 + \cdots + c_s a_s) \frac{1}{a_s} f_s =$$

$$= c_1 a_1 S(f_1, f_2) + (c_1 a_1 + c_2 a_2) S(f_2, f_3) + (c_1 a_1 + c_2 a_2 + c_3 a_3) S(f_3, f_4) + \cdots$$

$$\cdots + (c_1 a_1 + \cdots + c_{s-1} a_{s-1}) S(f_{s-1}, f_s).$$

□

Teorema 2.4.3. (Teorema de Buchberger).
Sea $G = \{g_1, \ldots, g_t\} \subset \mathbb{K}[x_1, \ldots, x_n]$ con g_i no nulo para $i \in \{1, \ldots, t\}$. Entonces, G es una base de Gröbner del ideal $I = \langle g_1, \ldots, g_t \rangle$ si y solo si $\overline{S(g_i, g_j)}^G = 0,\ \forall 1 \leq i < j \leq t$.

Demostración.

$\boxed{\Rightarrow}$ Supongamos que $G = \{g_1, \ldots, g_t\}$ es una base de Gröbner del ideal $I = \langle g_1, \ldots, g_t \rangle$. Se tiene

$$S(g_i, g_j) = \frac{\mathbf{x}^\gamma}{\mathrm{tl}(g_i)} g_i - \frac{\mathbf{x}^\gamma}{\mathrm{tl}(g_j)} g_j$$

donde $\mathbf{x}^\gamma = mcm(\mathrm{ml}(g_i), \mathrm{ml}(g_j))$. Entonces

$$\Rightarrow \ \ S(g_i, g_j) \in I \ \Rightarrow \ \overline{S(g_i, g_j)}^G = 0$$

por ser G una base de Gröbner y por el *Teorema 2.3.4*.

$\boxed{\Leftarrow}$ Supongamos $\overline{S(g_i, g_j)}^G = 0$ para $1 \leq i < j \leq t$. Para ver que G es base de Gröbner de I, vamos a probar que $\mathrm{TL}(G) = \mathrm{TL}(I)$. Probaremos que $\forall f \in I = \langle g_1, \ldots, g_t \rangle$, se tiene que $\mathrm{tl}(f) \in \langle \mathrm{tl}(g_1), \ldots, \mathrm{tl}(g_t) \rangle$.

$$f \in I = \langle g_1, \ldots, g_t \rangle \ \Rightarrow \ \exists h_i \in \mathbb{K}[x_1, \ldots, x_n] \text{ tal que } f = \sum_{i=1}^{t} h_i g_i \tag{2.3}$$

Sea $m(i) = \mathrm{ml}(h_i g_i)$ y sea $\mathbf{x}^\gamma = \mathrm{máx}_{1 \leq i \leq s}\{m(i)\}$. Por lo tanto, $\mathrm{ml}(f) \leq \mathbf{x}^\gamma$. Para cada una de las formas posibles de expresar f como en (2.3) tendremos un $\mathbf{x}^\gamma$ diferente. Como tenemos un orden monomial que es un buen orden, nos quedamos con cualquier expresión de f con menor $\mathbf{x}^\gamma$. Veamos que en este caso, se tiene que $\mathrm{ml}(f) = \mathbf{x}^\gamma$.

Reducción al Absurdo. Supongamos $\mathrm{ml}(f) < \mathbf{x}^\gamma \ \Rightarrow$

$$f = \sum_{m(i) = \mathbf{x}^\gamma} h_i g_i + \sum_{m(i) < \mathbf{x}^\gamma} h_i g_i = \sum_{m(i) = \mathbf{x}^\gamma} \mathrm{tl}(h_i) g_i + \sum_{m(i) = \mathbf{x}^\gamma} (h_i - \mathrm{tl}(h_i)) g_i + \sum_{m(i) < \mathbf{x}^\gamma} h_i g_i$$

Tanto los dos últimos sumandos de la expresión anterior como f tienen monomio líder $< \mathbf{x}^\gamma$ por lo que

$$\mathrm{ml}\left(\sum_{m(i)=\mathbf{x}^\gamma} \mathrm{tl}(h_i)g_i\right) < \mathbf{x}^\gamma.$$

Expresando $\mathrm{tl}(h_i) = c_i\mathbf{x}^{\alpha(i)}$ podemos escribir el primer sumando como

$$\sum_{m(i)=\mathbf{x}^\gamma} \mathrm{tl}(h_i)g_i = \sum_{m(i)=\mathbf{x}^\gamma} c_i\mathbf{x}^{\alpha(i)}g_i.$$

Llamamos $f_i = \mathbf{x}^{\alpha(i)}g_i$, y entonces tenemos que $\mathrm{ml}(f_i) = \mathrm{ml}(\mathbf{x}^{\alpha(i)}g_i) = \mathbf{x}^\gamma$ (si $m(i) = \mathbf{x}^\gamma$); aplicamos el *Lema 2.4.2* al último sumatorio para escribirlo como una combinación lineal con coeficientes en $\mathbb{K}$ de los S-polinomios $S(f_j, f_k) = S(\mathbf{x}^{\alpha(j)}g_j, \mathbf{x}^{\alpha(k)}g_k)$, con $j < k$ en un cierto subconjunto Φ no vacío de $\{1, \ldots, t\}$.

Se tiene

$$S(\mathbf{x}^{\alpha(j)}g_j, \mathbf{x}^{\alpha(k)}g_k) = \frac{\mathbf{x}^\gamma}{\mathbf{x}^{\alpha(j)}\mathrm{tl}(g_j)}\mathbf{x}^{\alpha(j)}g_j - \frac{\mathbf{x}^\gamma}{\mathbf{x}^{\alpha(k)}\mathrm{tl}(g_k)}\mathbf{x}^{\alpha(k)}g_k =$$

$$= \frac{\mathbf{x}^\gamma}{\mathrm{tl}(g_j)}g_j - \frac{\mathbf{x}^\gamma}{\mathrm{tl}(g_k)}g_k = \mathbf{x}^{\gamma-\gamma_{jk}}S(g_j, g_k)$$

donde $\mathbf{x}^{\gamma_{jk}} = mcm(\mathrm{ml}(g_j), \mathrm{ml}(g_k))$.

Luego podemos expresar:

$$\sum_{m(i)=\mathbf{x}^\gamma} \mathrm{tl}(h_i)g_i = \sum c_{jk}S(\mathbf{x}^{\alpha(j)}g_j, \mathbf{x}^{\alpha(k)}g_k) = \sum c_{jk}\mathbf{x}^{\gamma-\gamma_{j,k}}S(g_j, g_k)$$

para ciertos $c_{jk} \in \mathbb{K}$, con $c_{jk} = 0$ si $j, k \notin \Phi$.

Pero $\overline{S(g_j, g_k)}^G = 0 \;\Rightarrow\; S(g_j, g_k) = \sum_{l=1}^t a_{jkl}g_l$ con $a_{jkl} \in \mathbb{K}[x_1, \ldots, x_n]$, donde el monomio líder de $S(g_j, g_k)$ es igual a $\mathrm{máx}_l\{\mathrm{ml}(a_{jkl})\mathrm{ml}(g_l)\}$.

Multiplicamos la expresión $S(g_j, g_k)$ por $\mathbf{x}^{\gamma-\gamma_{j,k}}$

$$\mathbf{x}^{\gamma-\gamma_{jk}}S(g_j, g_k) = \mathbf{x}^{\gamma-\gamma_{j,k}}\sum_{l=1}^t a_{jkl}g_l = \sum_{l=1}^t b_{jkl}g_l$$

donde $b_{jkl} = \mathbf{x}^{\gamma-\gamma_{j,k}}a_{jkl}$.

Tenemos que

$$\mathrm{ml}(b_{jkl}g_l) \leq \mathrm{ml}(\mathbf{x}^{\gamma-\gamma_{j,k}}S(g_j, g_k)) < \mathbf{x}^\gamma$$

donde la primera desigualdad se sigue de la condición sobre el monomimo líder de $S(g_j, g_k)$ y la segunda desigualdad se sigue del *Lema 2.4.2.*

Se tiene entonces

$$\sum_{m(i)=\mathbf{x}^\gamma} \mathrm{tl}(h_i)g_i = \sum c_{jk}\mathbf{x}^{\gamma-\gamma_{j,k}}S(g_j, g_k) =$$

$$= \sum c_{jk}\left(\sum_{l=1}^t b_{jkl}g_l\right) = \sum_{l=1}^t \left(\sum c_{jk}b_{jkl}\right)g_l.$$

Denotamos $\widetilde{h}_l = \sum c_{jk}b_{jkl}$ con $\mathrm{ml}(\widetilde{h}_l g_l) = \mathrm{ml}(b_{jkl}g_l) < \mathbf{x}^\gamma$.

Ahora sustituimos en la ecuación de los tres sumandos y tenemos que

$$f = \sum_{l=1}^{t} \widetilde{h}_l g_l + \sum_{m(i)=\mathbf{x}^\gamma} (h_i - \mathrm{tl}(h_i))g_i + \sum_{m(i)<\mathbf{x}^\gamma} h_i g_i$$

donde todos los sumandos tienen monomio líder estrictamente menor que $\mathbf{x}^\gamma$ lo que nos lleva a una contradicción pues tomamos $\mathbf{x}^\gamma$ minimal.

Por lo tanto, $\mathrm{ml}(f) = \mathbf{x}^\gamma$ y entonces ya tenemos el resultado ya que $\mathrm{ml}(f) = \mathrm{ml}(h_i g_i)$ para algún $i = 1, \ldots, t$. Por lo que $\mathrm{tl}(g_i)$ divide a $tl(f)$ y $\mathrm{tl}(f) \in \langle \mathrm{tl}(g_1), \ldots, \mathrm{tl}(g_t) \rangle$.

□

Teorema 2.4.4. (Algoritmo de Buchberger)
Sea $I = \langle f_1, \ldots, f_s \rangle \subset \mathbb{K}[x_1, \ldots, x_n]$ ideal con $f_i \neq 0,\ \forall i = 1, \ldots, s$. Podemos construir una base de Gröbner $G = \{g_1, \ldots, g_t\}$ para I con el siguiente algoritmo:

Algoritmo 2 Algoritmo de Buchberger

Entrada: $F = \{f_1, \ldots, f_s\} \subset \mathbb{K}[x_1, \ldots, x_n]$ con $f_i \neq 0,\ \forall i = 1, \ldots, s$.
Salida: $G = \{g_1, \ldots, g_t\}$ base de Gröbner de I con $F \subset G$.

Inicialización: $G := F$
1: **Repetir:**
$\widetilde{G} := G$
2: **Para cada** par $\{f, g\} \subset \widetilde{G}$ con $f \neq g$ **hacer** $S = \overline{S(f,g)}^{\widetilde{G}}$
3: **Si** $S \neq 0$ **entonces hacer** $G := G \cup \{S\}$
4: **Hasta que** $G = \widetilde{G}$

Demostración.

En primer lugar veamos que el algoritmo termina: *Reducción al Absurdo.* Supongamos que no fuese así, construiríamos en cada paso un conjunto G_i estrictamente mayor que G_{i-1} de manera que

$$G_1 \subsetneq G_2 \subsetneq G_3 \subsetneq \ldots$$

donde cada G_i se obtiene añadiendo a G_{i-1} un polinomio $S \in I$ pero que está reducido con respecto a G_{i-1}. Por lo tanto, $\mathrm{tl}(S) \notin \mathrm{TL}(G_{i-1}) \Rightarrow$

$$\mathrm{TL}(G_1) \subsetneq \mathrm{TL}(G_2) \subsetneq \mathrm{TL}(G_3) \subsetneq ...$$

cadena estrictamente creciente de ideales monomiales que debe estabilizarse debido a la *Condición de Cadena Ascendente* de los anillos Noetherianos, *Teorema 1.1.5*. Por lo tanto el algoritmo debe terminar.

Además $F = \{f_1, \ldots, f_s\} \subset \{g_1, \ldots, g_t\} = G \subset I$, luego está claro que $\langle g_1, \ldots, g_t \rangle = I$; pues $I = \langle f_1, \ldots, f_s \rangle$ por hipótesis.

Por último, observemos que cuando el algoritmo termina, $\forall f, g \in G$ con $f \neq g$ es $\overline{S(f,g)}^{G} = 0$ por construcción, y gracias al *Teorema 2.4.3*, G es una base de Gröbner de I. □

Ejemplo 6. *Sean $f_1 = x + y$, $f_2 = x - y$ en el anillo de polinomios $\mathbb{Q}[x, y]$. Consideremos el orden lexicográfico con $x > y$ en los monomios del anillo. Aplicando el Algoritmo de Buchberger, una base de Gröbner del ideal $I = \langle f_1, f_2 \rangle$ respecto del orden mencionado es $G = \{g_1, g_2\}$ con $g_1 = f_1$ y $g_2 = y$. Por tanto* $\mathrm{TL}(I) = \langle x, y \rangle$.

Lema 2.4.5. *Sea $G = \{g_1, \ldots, g_t\} \subset I$ una base de Gröbner de un ideal $I \subset \mathbb{K}[x_1, \ldots, x_n]$. Si existe $g_i \in G$ tal que $\mathrm{tl}(g_i) \in \mathrm{TL}(G - \{g_i\})$. Entonces $G - \{g_i\}$ también es una base de Gröbner de I.*

Demostración.

Como G es una base de Gröbner $\Rightarrow$ $\mathrm{TL}(G) = \mathrm{TL}(I)$. Veamos que $\mathrm{TL}(G) = \mathrm{TL}(G - \{g_i\})$.

Que $\mathrm{TL}(G - \{g_i\}) \subset \mathrm{TL}(G)$ es trivial. Para la otra contención; como

$$\mathrm{tl}(g_i) \in \mathrm{TL}(G - \{g_i\}) \;\Rightarrow\; \langle \mathrm{tl}(g_i) \rangle \subset \mathrm{TL}(G - \{g_i\}) \;\Rightarrow\; \mathrm{TL}(G) \subset \mathrm{TL}(G - \{g_i\})$$

Luego $\mathrm{TL}(I) = \mathrm{TL}(G - \{g_i\}) \Rightarrow G - \{g_i\}$ es base de Gröbner de I. □

Definición 2.4.6. *Una base de Gröbner $G = \{g_1, \ldots, g_t\}$ de un ideal $I \subset \mathbb{K}[x_1, \ldots, x_n]$ respecto del orden monomial $\leq$ fijado de antemano se dice que es minimal si*

i) $\mathrm{cl}(g_i) = 1, \;\; \forall i = 1, \ldots, t.$

ii) $\mathrm{tl}(g_i) \notin \mathrm{TL}(G - \{g_i\}), \;\; \forall i = 1, \ldots, t.$

Está claro que una vez calculada una base de Gröbner con el Algoritmo de Buchberger, basta eliminar los elementos g_i tales que $\mathrm{ml}(g_i) \in \mathrm{TL}(G - \{g_i\})$ y después hacer los términos líder mónicos dividiendo cada g_i por $\mathrm{cl}(g_i)$; de esa forma ya tendremos calculada una base de Gröbner minimal.

Definición 2.4.7. *Una base de Gröbner $G = \{g_1, \ldots, g_t\}$ de un ideal $I \subset \mathbb{K}[x_1, \ldots, x_n]$ respecto del orden monomial $\leq$ fijado de antemano se dice que es reducida si*

i) $\mathrm{cl}(g_i) = 1, \;\; \forall i = 1, \ldots, t.$

ii) $\overline{g_i}^{G - \{g_i\}} = g_i, \;\; \forall i = 1, \ldots, t.$

Aquí *ii)* no solo pide que $\mathrm{tl}(g_i) \notin \mathrm{TL}(G - \{g_i\})$, sino que ningún término de g_i esté en $\mathrm{TL}(G - \{g_i\})$. Luego toda base de Gröbner reducida es minimal.

Proposición 2.4.8. *Sea un ideal $I \subset \mathbb{K}[x_1, \ldots, x_n]$ no nulo. Entonces I tiene una única base de Gröbner reducida (fijado por supuesto con anterioridad un orden monomial).*

Demostración.

Primero se calcula G una base de Gröbner minimal para I. Ahora modificamos esta base hasta conseguir que sea reducida.

Sea $g \in G$, definimos $\widetilde{g} = \overline{g}^{G - \{g\}}$ y $\widetilde{G} = (G - \{g\}) \cup \{\widetilde{g}\}$.

Notemos que $\widetilde{G}$ sigue siendo una base de Gröbner minimal de I donde el elemento $\widetilde{g}$ ya está reducido respecto de $\widetilde{G} - \{\widetilde{g}\}$. Luego hay que hacer este proceso con todos los elementos de G y obtendremos una base de Gröbner reducida.

Unicidad. Supongamos G y $\widetilde{G}$ dos bases de Gröbner reducidas para el ideal I; en particular son minimales. Veamos primero que por el hecho de ser minimales G y $\widetilde{G}$ tienen el mismo número de elementos y $\forall g \in G, \exists \widetilde{g} \in \widetilde{G}$ tal que $\mathrm{tl}(g) = \mathrm{tl}(\widetilde{g})$.

Sea $G = \{g_1, \ldots, g_t\}$, $\widetilde{G} = \{\widetilde{g}_1, \ldots, \widetilde{g}_s\}$. Como $g_1 \in I$ y $\widetilde{G}$ es base de Gröbner de I, $\exists \widetilde{g}_i \in \widetilde{G}$ tal que $\mathrm{ml}(\widetilde{g}_i) \,|\, \mathrm{ml}(g_1)$. Reordenamos y consideramos $\widetilde{g}_i = \widetilde{g}_1 \;\Rightarrow\; \mathrm{ml}(\widetilde{g}_1) \,|\, \mathrm{ml}(g_1)$.

Pero como $\widetilde{g}_1 \in I$ y G también es una base de Gröbner, $\exists g_j \in G$ tal que $\mathrm{ml}(g_j) \,|\, \mathrm{ml}(\widetilde{g}_1)$. Entonces

$$\mathrm{ml}(g_j) \,|\, \mathrm{ml}(\widetilde{g}_1) \,|\, \mathrm{ml}(g_1) \;\Rightarrow\; g_j = g_1 \, y \, \mathrm{ml}(g_1) = \mathrm{ml}(\widetilde{g}_1).$$

Pero como $\mathrm{cl}(g_1) = \mathrm{cl}(\widetilde{g}_1) = 1$ debe ser $\mathrm{tl}(g_1) = \mathrm{tl}(\widetilde{g}_1)$.

Ahora sea $g_2 \in I \;\Rightarrow \exists \widetilde{g}_i \in \widetilde{G}$ tal que $\mathrm{ml}(\widetilde{g}_i) \,|\, \mathrm{ml}(g_2)$.

Reordenando y repitiendo el procedimiento, $\mathrm{tl}(g_2) = \mathrm{tl}(\widetilde{g}_2)$. Al final, llegamos a que $s = t$ y $\mathrm{tl}(g_i) = \mathrm{tl}(\widetilde{g}_i) \;\; \forall i = 1, \ldots, t$.

Volvamos entonces a nuestras dos bases reducidas G y $\widetilde{G}$ como antes con $s = t$ y $\mathrm{tl}(g_i) = \mathrm{tl}(\widetilde{g}_i)$ para $1 \leq i \leq s$.

Dado $g \in G, \exists \widetilde{g} \in \widetilde{G}$ tal que $\mathrm{tl}(g) = \mathrm{tl}(\widetilde{g})$. Basta ver que $g = \widetilde{g}$ para probar que $G = \widetilde{G}$. Pero como $g - \widetilde{g} \in I$ y como G es base de Gröbner, $\overline{g - \widetilde{g}}^{G} = 0$, además $\mathrm{tl}(g) = \mathrm{tl}(\widetilde{g})$, así que los términos líder se cancelan en $g - \widetilde{g}$ y el resto de los términos no deben ser divisibles por $\mathrm{TL}(G) = \mathrm{TL}(\widetilde{G})$ puesto que G y $\widetilde{G}$ son reducidas. De modo que $\overline{g - \widetilde{g}}^{G} = g - \widetilde{g} = 0 \;\Rightarrow\; G = \widetilde{G}$. □

Capítulo 3

Teoría de Eliminación

El objetivo fundamental de este capítulo es describir un procedimiento que permita resolver sistemas de ecuaciones polinómicas. Sean $f_1, \ldots, f_s \in \mathbb{K}[x_1, \ldots, x_n]$, queremos calcular los puntos $(a_1, \ldots, a_n) \in \mathbb{A}^n(\mathbb{K})$ tales que anulen a los polinomios $f_1, \ldots, f_s$. Es decir, calcular el conjunto algebraico $\mathcal{V}(f_1, \ldots, f_s) = \{\mathbf{a} = (a_1, \ldots, a_n) \in \mathbb{A}^n(\mathbb{K}) \mid f_i(\mathbf{a}) = 0, i = 1, \ldots, s\}$. Además sabemos que $\mathcal{V}(f_1, \ldots, f_s) = \mathcal{V}(\langle f_1, \ldots, f_s \rangle)$.

Nos ayudaremos de las bases de Gröbner para poder resolver este problema. La idea es calcular un sistema equivalente a

$$\begin{cases} f_1 = 0 \\ \quad \vdots \\ f_s = 0 \end{cases}$$

que sea más sencillo de resolver. Más adelante explicaremos qué se entiende por un sistema más sencillo que el dado.

3.1. Teoremas de Eliminación y de Extensión

Dado $I = \langle f_1, \ldots, f_s \rangle$ con $f_1, \ldots, f_s \in \mathbb{K}[x_1, \ldots, x_n]$ no nulos, nuestro objetivo aquí es calcular una base de Gröbner $G = \{g_1, \ldots, g_t\}$, del ideal I, con buenas propiedades; de modo que como $\mathcal{V}(G) = \mathcal{V}(I)$, baste resolver el sistema

$$\begin{cases} g_1 = 0 \\ \quad \vdots \\ g_t = 0 \end{cases}$$

que buscaremos que sea sencillo de resolver, en el sentido de que consecutivamente solo tengamos que calcular las raíces de polinomios de una variable.

Definición 3.1.1. *Dado un ideal $I \subset \mathbb{K}[x_1, \ldots, x_n]$, se define el ideal de eliminación I_ℓ para $\ell \in \{1, \ldots, n-1\}$ como*

$$I_\ell = I \cap \mathbb{K}[x_{\ell+1}, \ldots, x_n].$$

Observación 3.1.2. *I_ℓ es un ideal de $\mathbb{K}[x_{\ell+1}, \ldots, x_n]$; ya que la contracción I_ℓ de un ideal I de $\mathbb{K}[x_1, \ldots, x_n]$ es un ideal del subanillo $\mathbb{K}[x_{\ell+1}, \ldots, x_n] \subseteq \mathbb{K}[x_1, \ldots, x_n]$.*

Se trata de un resultado básico de teoría de anillos. Es consecuencia de un resultado más general ([1], pag 5): $A \subset B$ anillos, $f : A \to B$ morfismo de anillos, y $J \subset B$ ideal. Entonces $f^{-1}(J)$ es un ideal de A.

En la *Definición 3.1.1* podemos tomar $\ell \in \{0, \ldots, n\}$, notemos entonces que $I_0 = I$; y abusando del lenguage $I_n = I \cap \mathbb{K}$, es decir, el ideal formado por las constantes de I, que únicamente será no nulo si $I = \mathbb{K}[x_1, \ldots, x_n]$. Es importante también fijar un orden en las variables $x_1, \ldots, x_n$ ya que el ideal I_ℓ está compuesto por los polinomios de I que solo dependen de las $n - \ell$ últimas variables.

El objetivo es ahora encontrar unos generadores de cada ideal I_ℓ partiendo de un sistema de generadores del ideal I.

Teorema 3.1.3. (Teorema de Eliminación)
Sea $I \subset \mathbb{K}[x_1, \ldots, x_n]$ un ideal y sea G una base de Gröbner del ideal I respecto del orden lexicográfico, $<_{lex}$ (con $x_1 > x_2 > \cdots > x_n$). Para $\ell \in \{1, \ldots, n-1\}$ sea $G_\ell = G \cap \mathbb{K}[x_{\ell+1}, \ldots, x_n]$. Se verifica que,

i) $G_\ell = \emptyset$ si y solo si $I_\ell = (0)$.

ii) Si $G_\ell \neq \emptyset$ entonces G_ℓ es base de Gröbner de I_ℓ respecto del orden $<_{lex}$ (restringido a los monomios de $\mathbb{K}[x_{\ell+1}, \ldots, x_n]$).

Demostración.

Sea $\ell \in \{1, \ldots, n-1\}$ fijo. Consideramos el ideal de eliminación $I_\ell = I \cap \mathbb{K}[x_{\ell+1}, \ldots, x_n]$ y el conjunto $G_\ell = G \cap \mathbb{K}[x_{\ell+1}, \ldots, x_n]$.

i) Supongamos que $G_\ell = \emptyset$.

Reducción al Absurdo. Supongamos $f \in I_\ell$ no nulo, en particular $f \in I$ y por lo tanto $\mathrm{tl}(f) \in \mathrm{TL}(I) = \mathrm{TL}(G)$ por ser G una base de Gröbner de I. Luego debe existir $g \in G$ tal que $\mathrm{tl}(g) \,|\, \mathrm{tl}(f)$. Como $\mathrm{tl}(f) \in \mathbb{K}[x_{\ell+1}, \ldots, x_n]$ entonces $\mathrm{tl}(g)$ tampoco depende de las variables $x_1, \ldots, x_\ell$. Como estamos considerando el orden lexicográfico (con $x_1 > \cdots > x_n$) ningún monomio de g puede depender de las variables $x_1, \ldots, x_\ell$. Entonces $g \in G_\ell$ lo que es una contradicción.

Si $I_\ell = (0)$ entonces $G_\ell = \emptyset$ ya que $G_\ell \subset I_\ell$ y G no contiene al polinomio 0.

ii) Supongamos que $G_\ell \neq \emptyset$. Para ver que G_ℓ es una base de Gröbner de I_ℓ, por el *Teorema 2.3.4* basta probar que $\mathrm{TL}(I_\ell) = \mathrm{TL}(G_\ell)$.

Que $\mathrm{TL}(G_\ell) \subset \mathrm{TL}(I_\ell)$ está claro pues $G_\ell \subset I_\ell$. Veamos la otra contención:

Sea $f \in I_\ell$ no nulo, hay que demostrar que $\mathrm{tl}(f) \in \mathrm{TL}(G_\ell)$. Pero como $f \in I$; y como G es una base de Gröbner de I, $\exists g \in G$ tal que $\mathrm{tl}(g) \,|\, \mathrm{tl}(f)$. Veamos entonces que $g \in G_\ell$.

Notemos que $\mathrm{tl}(f)$ y por tanto, $\mathrm{tl}(g)$ están en $\mathbb{K}[x_{\ell+1}, \ldots, x_n]$, pues $f \in I_\ell$. Y como estamos trabajando con el orden lexicográfico donde $x_1 > x_2 > \cdots > x_n$; si $\mathrm{tl}(g) \in \mathbb{K}[x_{\ell+1}, \ldots, x_n] \;\Rightarrow\; g \in \mathbb{K}[x_{\ell+1}, \ldots, x_n]$; y por tanto $g \in G_\ell$.

□

Retomemos nuestro objetivo de encontrar $\mathcal{V}(I)$ para cierto ideal $I \subset \mathbb{K}[x_1, \ldots, x_n]$. Si calculamos G_{n-1} base de Gröbner para el ideal I_{n-1}, tendremos polinomios de I que solo involucran a la variable x_n. La idea es resolver estos polinomios de una variable, es decir, hallar $\mathcal{V}(I_{n-1}) \subset \mathbb{A}^1(\mathbb{K})$; y sustituir x_n por cada valor posible (en $\mathcal{V}(I_{n-1})$) en los polinomios de G_{n-2}; y así, paso a paso ir extendiendo las soluciones de $\mathcal{V}(I_\ell)$ a las de $\mathcal{V}(I_{\ell-1})$; para al final llegar a calcular $\mathcal{V}(I_0) = \mathcal{V}(I)$; de forma que en cada paso solo tengamos que resolver polinomios de una variable.

Ejemplo 7. *Supongamos que queremos calcular las soluciones en* $\mathbb{A}^3(\mathbb{C})$ *del sistema*

$$\begin{cases} x^2 + y^2 + z^2 = 4 \\ x^2 + 2y^2 = 5 \\ xz = 1 \end{cases}$$

Consideramos el ideal $I = \langle x^2 + y^2 + z^2 - 4,\ x^2 + 2y^2 - 5,\ xz - 1\rangle \subset \mathbb{C}[x, y, z]$. *Calculamos*[1] $G = \{x + 2z^3 - 3z,\ y^2 - z^2 - 1,\ z^4 - \frac{3}{2}z^2 + \frac{1}{2}\}$, *la base de Gröbner reducida de* I *respecto del orden lexicográfico* $(x > y > z)$.

Aplicando el Teorema de Eliminación *tenemos que* $G_2 = \{z^4 - \frac{3}{2}z^2 + \frac{1}{2}\}$ *es base de Gröbner de* $I_2 = I \cap \mathbb{C}[z]$ *respecto del orden lexicográfico (que en este caso es el orden de las potencias de* z*). Resolviendo la ecuación* $z^4 - \frac{3}{2}z^2 + \frac{1}{2} = 0$ *obtenemos que* $\mathcal{V}(I_2) = \{1,\ -1,\ \frac{\sqrt{2}}{2},\ \frac{-\sqrt{2}}{2}\}$.

Aplicando de nuevo el Teorema de Eliminación $G_1 = \{y^2 - z^2 - 1,\ z^4 - \frac{3}{2}z^2 + \frac{1}{2}\}$ *es una base de Gröbner de* $I_1 = I \cap \mathbb{C}[y, z]$ *respecto del orden lexicográfico. Como ya tenemos* $\mathcal{V}(I_2)$*, para cada* $c \in \mathcal{V}(I_2)$ *sustituimos* $z = c$ *en la primera ecuación de* G_1 *y extendemos* $\mathcal{V}(I_2)$ *a* $\mathcal{V}(I_1)$ *de manera que*

$$\mathcal{V}(I_1) = \left\{ \left(\sqrt{2}, 1\right), \left(-\sqrt{2}, 1\right), \left(\sqrt{2}, -1\right), \left(-\sqrt{2}, -1\right), \right.$$
$$\left. \left(\frac{\sqrt{6}}{2}, \frac{\sqrt{2}}{2}\right), \left(-\frac{\sqrt{6}}{2}, \frac{\sqrt{2}}{2}\right), \left(\frac{\sqrt{6}}{2}, -\frac{\sqrt{2}}{2}\right), \left(-\frac{\sqrt{6}}{2}, -\frac{\sqrt{2}}{2}\right) \right\}.$$

Repetimos el proceso ahora en G *y calculamos* $\mathcal{V}(I)$ *que era nuestro propósito.*

$$\mathcal{V}(I) = \left\{ \left(1, \sqrt{2}, 1\right), \left(1, -\sqrt{2}, 1\right), \left(-1, \sqrt{2}, -1\right), \left(-1, -\sqrt{2}, -1\right), \right.$$
$$\left. \left(\sqrt{2}, \frac{\sqrt{6}}{2}, \frac{\sqrt{2}}{2}\right), \left(\sqrt{2}, -\frac{\sqrt{6}}{2}, \frac{\sqrt{2}}{2}\right), \left(-\sqrt{2}, \frac{\sqrt{6}}{2}, -\frac{\sqrt{2}}{2}\right), \left(-\sqrt{2}, -\frac{\sqrt{6}}{2}, -\frac{\sqrt{2}}{2}\right) \right\}.$$

El problema aquí radica en que no todas las soluciones de $\mathcal{V}(I_\ell)$ pueden extenderse a soluciones de $\mathcal{V}(I_{\ell-1})$ en general.

Ejemplo 8. *Supongamos que queremos calcular las soluciones en* $\mathbb{A}^3(\mathbb{C})$ *del sistema*

$$\begin{cases} xz = 1 \\ y - z = 0 \end{cases}$$

Consideramos el ideal $I = \langle xz - 1, y - z\rangle \subset \mathbb{C}[x, y, z]$. *Sea* $G = \{xz - 1, y - z\}$ *la base de Gröbner reducida de* I *respecto del orden lexicográfico* $(x > y > z)$.

[1] Para los cálculos en los ejemplos hemos usado los sistemas de cálculo simbólico Maple y Sage.

Entonces $G_2 = G \cap \mathbb{C}[z] = \emptyset$, *por lo que* $I_2 = (0)$ *y por tanto* $\mathcal{V}(I_2) = \mathbb{C}$.

$G_1 = G \cap \mathbb{C}[y, z] = \{y - z\}$, *luego* $\mathcal{V}(I_1) = \{(a, a) \,|\, a \in \mathbb{C}\}$. *Ahora calculamos* $\mathcal{V}(I)$ *resolviendo la ecuación* $xz = 1$ *para cada valor de* z *obtenido anteriormente. Pero vemos que si* $z = 0$ *no podemos encontrar* x *que verifique la ecuación, luego* $\mathcal{V}(I) = \{(\frac{1}{a}, a, a) \,|\, a \in \mathbb{C} - \{0\}\}$. *Es decir no hemos podido extender* $(0, 0) \in \mathcal{V}(I_1)$ *a una solución de* $\mathcal{V}(I)$.

Teorema 3.1.4. (Teorema de Extensión)
Sea $I = \langle f_1, \ldots, f_s \rangle$ *un ideal de* $\mathbb{C}[x_1, \ldots, x_n]$ *y sea* I_1 *el primer ideal de eliminación de* I. *Para cada* $i \in \{1, \ldots, s\}$ *escribimos* f_i *como*

$$f_i = h_i(x_2, \ldots, x_n)x_1^{N_i} + (\textit{términos con grado } < N_i \textit{ en } x_1)$$

donde $N_i \geq 0$ *y* $h_i \in \mathbb{C}[x_2, \ldots, x_n]$ *no nulo.*

Sea una solución parcial $(a_2, \ldots, a_n) \in \mathcal{V}(I_1)$. *Si* $(a_2, \ldots, a_n) \notin \mathcal{V}(h_1, \ldots, h_s)$, *entonces existe* $a_1 \in \mathbb{C}$ *tal que* $(a_1, \ldots, a_n) \in \mathcal{V}(I)$.

Dedicaremos el siguiente capítulo a la prueba de este resultado.

Hasta ahora hemos trabajado en un cuerpo arbitrario $\mathbb{K}$, pero en el *Teorema de Extensión* fijamos $\mathbb{C}$ pues necesitamos ahora un cuerpo algebraicamente cerrado para poder asegurar que todo polinomio, ya en una variable, tenga raíces en el cuerpo; es decir, para que se verifiquen las hipótesis del *Teorema Fundamental del Álgebra*.

Ejemplo 9. *Si en el* Ejemplo 7 *escribimos* $f_1 = xz - 1$ *como en el* Teorema de Extensión *tenemos que* $h_1 = z$. *Para* $(0, 0) \in \mathcal{V}(I_1)$ *se tiene* $(0, 0) \in \mathcal{V}(h_1)$, *luego no tenemos la seguridad de que podamos extender este punto a una solución total (es decir, a una solución de* $\mathcal{V}(I)$*); de hecho en el ejemplo vimos que* $(0, 0) \in \mathcal{V}(I_1)$ *no se puede extender a una solución en* $\mathcal{V}(I)$. *Sin embargo,* $\forall a \in \mathbb{C} - \{0\}$, $(a, a) \notin \mathcal{V}(h_1)$ *y por eso la solución parcial* (a, a) *se puede extender a una solución en* $\mathcal{V}(I)$.

El *Teorema de Extensión* nos dice que si una solución parcial $(a_2, \ldots, a_n) \in \mathcal{V}(I_1)$ cumple que $(a_2, \ldots, a_n) \notin \mathcal{V}(h_1, \ldots, h_s)$, es decir, que al sustituir $(a_2, \ldots, a_n)$ en $f_1, \ldots, f_s$ no se anulen todos los términos líderes simultáneamente entonces podremos extender $(a_2, \ldots, a_n)$ a una solución $(a_1, a_2, \ldots, a_n)$ de $\mathcal{V}(I)$.

Luego en el caso de que algún f_i tenga $\mathrm{tl}(f_i) = c_i x_1^{N_i}$, con $c_i \in \mathbb{C} - \{0\}$, es decir, si algún f_i tiene término líder donde solo aparezca la variable x_1, cualquier solución parcial $(a_2, \ldots, a_n) \in \mathcal{V}(I_1)$ podrá extenderse.

El teorema nos dice cómo extender soluciones de $\mathcal{V}(I_1)$ a $\mathcal{V}(I)$; pero se puede aplicar de forma similar para pasar de cualquier solución en $\mathcal{V}(I_\ell)$ a una solución en $\mathcal{V}(I_{\ell-1})$ ya que I_ℓ es el primer ideal de eliminación de $I_{\ell-1}$.

3.2. Geometría de la Eliminación

Aquí daremos una interpretación geométrica a lo visto en la sección anterior.

Definición 3.2.1. *En el espacio afín* $\mathbb{A}^n(\mathbb{C})$, *definimos la proyección sobre las últimas* $n - \ell$ *variables como*

$$\begin{array}{rccc} \pi_\ell : & \mathbb{A}^n(\mathbb{C}) & \longrightarrow & \mathbb{A}^{n-\ell}(\mathbb{C}) \\ & (a_1,\ldots,a_n) & \longmapsto & (a_{\ell+1},\ldots,a_n) \end{array}$$

para $\ell \in \{0,\ldots,n-1\}$.

Notemos que π_0 es la identidad sobre $\mathbb{A}^n(\mathbb{C})$.

Sea $\mathrm{V} = \mathcal{V}(f_1,\ldots,f_s) \subset \mathbb{A}^n(\mathbb{C})$ un conjunto algebraico. Consideremos el conjunto $\pi_\ell(\mathrm{V})$ obtenido aplicando la proyección π_ℓ a V. Es decir,

$$\pi_\ell(\mathrm{V}) = \{(a_{\ell+1},\ldots,a_n) \in \mathbb{A}^{n-\ell}(\mathbb{C}) : \exists a_1,\ldots,a_\ell \in \mathbb{C} \text{ con } (a_1,\ldots,a_\ell,a_{\ell+1},\ldots,a_n) \in \mathrm{V}\}.$$

Queremos en primer lugar ver qué relación existe entre $\pi_\ell(\mathrm{V})$ y el ideal de eliminación I_ℓ de $I = \langle f_1,\ldots,f_s\rangle$.

Lema 3.2.2. *Sea* $I_\ell = I \cap \mathbb{C}[x_{\ell+1},\ldots,x_n]$ *el* ℓ*-ésimo ideal de eliminación de* $I = \langle f_1,\ldots,f_s\rangle$. *Entonces,* $\pi_\ell(\mathrm{V}) \subset \mathcal{V}(I_\ell)$ *y en general la inclusión es estricta.*

Demostración.

Sea $(a_{\ell+1},\ldots,a_n) \in \pi_\ell(\mathrm{V}) \;\Rightarrow\; \exists a_1,\ldots,a_\ell \in \mathbb{C}$ tal que $(a_1,\ldots,a_\ell,a_{\ell+1},\ldots,a_n) \in \mathrm{V}$. Entonces $f(a_1,\ldots,a_n) = 0$, $\forall f \in I$.

En particular $\forall f \in I_\ell$, $f(a_1,\ldots,a_n) = f(a_{\ell+1},\ldots,a_n) = 0$ y por tanto $(a_{\ell+1},\ldots,a_n) \in \mathcal{V}(I_\ell)$. Que la inclusión es estricta en general se verá en el siguiente ejemplo. □

Ejemplo 10. *Volvemos de nuevo a la situación de los* Ejemplos 7 y 8*, donde teníamos que*

$$\mathrm{V} = \mathcal{V}(I) = \Big\{ \left(\frac{1}{a}, a, a\right) \mid a \in \mathbb{C} - \{0\} \Big\},$$

$$\mathcal{V}(I_1) = \{(a,a) \mid a \in \mathbb{C}\}.$$

Luego tenemos entonces que

$$\pi_1(\mathrm{V}) = \{(a,a) \mid a \in \mathbb{C} - \{0\}\}.$$

Por lo tanto en este caso $\pi_1(\mathrm{V}) \subsetneq \mathcal{V}(I_1)$ *pues* $(0,0) \in \mathcal{V}(I_1) - \pi_1(\mathrm{V})$.

De la definición del conjunto $\pi_\ell(\mathrm{V})$, se obtiene fácilmente que $\pi_\ell(\mathrm{V})$ es el conjunto de soluciones parciales de $\mathcal{V}(I_\ell)$ que se pueden a extender a soluciones de $\mathcal{V}(I)$.

Teorema 3.2.3. (Teorema de Extensión Geométrico)
Sea $I = \langle f_1,\ldots,f_s\rangle$ *un ideal de* $\mathbb{C}[x_1,\ldots,x_n]$*, sea* I_1 *el primer ideal de eliminación de* I*, y sea* $\mathrm{V} = \mathcal{V}(I) \subset \mathbb{A}^n(\mathbb{C})$*. Para cada* $i \in \{1,\ldots,s\}$ *escribimos* f_i *como*

$$f_i = h_i(x_2,\ldots,x_n)x_1^{N_i} + (\textit{términos con grado } < N_i \textit{ en } x_1)$$

donde $N_i \geq 0$ *y* $h_i \in \mathbb{C}[x_2,\ldots,x_n]$ *no nulo.*

Entonces,

$$\mathcal{V}(I_1) = \pi_1(\mathrm{V}) \cup (\mathcal{V}(h_1,\ldots,h_s) \cap \mathcal{V}(I_1))$$

donde $\pi_1 : \mathbb{A}^n(\mathbb{C}) \longrightarrow \mathbb{A}^{n-1}(\mathbb{C})$ *es la proyección en las últimas* $n-1$ *variables.*

Demostración.

Vamos a probar que $\mathcal{V}(I_1) = \pi_1(\mathrm{V}) \cup (\mathcal{V}(h_1, \ldots, h_s) \cap \mathcal{V}(I_1))$.

$\boxed{\subset}$ Sea $(a_2, \ldots, a_n) \in \mathcal{V}(I_1)$.

Si $(a_2, \ldots, a_n) \in \mathcal{V}(h_1, \ldots, h_s)$ entonces $(a_2, \ldots, a_n) \in \pi_1(\mathrm{V}) \cup (\mathcal{V}(h_1, \ldots, h_s) \cap \mathcal{V}(I_1))$.

Si $(a_2, \ldots, a_n) \notin \mathcal{V}(h_1, \ldots, h_s)$, aplicando el *Teorema de Extensión*, $(a_2, \ldots, a_n)$ puede extenderse a una solución de $\mathcal{V}(I)$; luego $(a_2, \ldots, a_n) \in \pi_1(\mathrm{V})$.

$\boxed{\supset}$ Se tiene directamente porque $\pi_1(\mathrm{V}) \subseteq \mathcal{V}(I_1)$ y $\mathcal{V}(h_1, \ldots, h_s) \cap \mathcal{V}(I_1) \subseteq \mathcal{V}(I_1)$.

□

Nota. *Si en las condiciones anteriores $\mathcal{V}(h_1, \ldots, h_s) \cap \mathcal{V}(I_1) = \emptyset$ entonces $\mathcal{V}(I_1) = \pi_1(\mathrm{V})$.*

Además tenemos que $\mathcal{V}(I_1) = \pi_1(\mathrm{V})$ si y solo si $\mathcal{V}(h_1, \ldots, h_s) \cap \mathcal{V}(I_1) \subseteq \pi_1(\mathrm{V})$. Osea que toda solución de $\mathcal{V}(I_1)$ se extiende a una solución total de $\mathcal{V}(I)$ si y solo si toda solución de $\mathcal{V}(h_1, \ldots, h_s) \cap \mathcal{V}(I_1)$ se extiende a $\mathcal{V}(I)$.

Pero ésta no es la única relación existente entre $\pi_\ell(\mathrm{V})$ y $\mathcal{V}(I_\ell)$.

Observemos que el conjunto $\pi_\ell(\mathrm{V})$ no tiene en principio por qué ser un subconjunto algebraico de $\mathbb{A}^{n-\ell}(\mathbb{C})$.

Ejemplo 11. *Vimos en el* Ejemplo 9 *que $\pi_1(\mathrm{V}) = \{(a, a) \in \mathbb{A}^2(\mathbb{C}) \,|\, a \in \mathbb{C} - \{0\}\}$. Veamos que no es un cerrado para la Topología de Zariski, es decir, que no es un conjunto algebraico.*

Reducción al Absurdo. *Supongamos que si lo es. Entonces $\pi_1(\mathrm{V}) = \mathcal{V}(f_1, \ldots, f_s)$ para ciertos polinomios $f_i \in \mathbb{C}[x, y]$ no nulos*[2]*. Escribimos*

$$f_i(x, y) = q_i(x, y)(y - x) + r_i(x)$$

(división por $y - x$) con $r_i(x) \in \mathbb{C}[x]$. Como $f_i(a, a) = 0 \ \forall a \in \mathbb{C} - \{0\}$, se tiene que $f_i(a, a) = r_i(a) = 0 \ \forall a \in \mathbb{C} - \{0\}$, lo que implica que $r_i(x) = 0$. Tenemos entonces que $f_i = q_i(y - x)$ y por tanto $\langle f_1, \ldots, f_s \rangle \subseteq \langle y - x \rangle \subseteq \mathbb{C}[x, y]$. Entonces $\mathcal{V}(y - x) = \{(a, a) \in \mathbb{A}^2(\mathbb{C}) \,|\, a \in \mathbb{C}\} \subseteq \pi_1(\mathrm{V})$ lo que es una contradicción.

El siguiente resultado nos dice que $\mathcal{V}(I_\ell)$ es la clausura de Zariski de $\pi_\ell(\mathrm{V})$.

Teorema 3.2.4. (Teorema de la Clausura)
Sea $\mathrm{V} = \mathcal{V}(f_1, \ldots, f_s) \subset \mathbb{A}^n(\mathbb{C})$ y sea I_ℓ el ℓ-ésimo ideal de eliminación del ideal $I = \langle f_1, \ldots, f_s \rangle$ con $\ell \in \{1, \ldots, n - 1\}$. Entonces se verifica

i) $\mathcal{V}(I_\ell)$ es el menor conjunto algebraico que contiene a $\pi_\ell(\mathrm{V})$ en el espacio afín $\mathbb{A}^{n-\ell}(\mathbb{C})$.

ii) Si $\mathrm{V} \neq \emptyset$, existe un conjunto algebraico $W \subsetneq \mathcal{V}(I_\ell)$ tal que $\mathcal{V}(I_\ell) - W \subset \pi_\ell(\mathrm{V})$.

Demostración.

Veamos de momento solo la prueba de *i*), la demostración de *ii*) la veremos en la siguiente sección.

[2]En los *Ejemplos 7, 8* y *9* hemos trabajado en el anillo $\mathbb{C}[y, z]$. Por comodidad hacemos aquí el cambio de z por x.

La clausura de Zariski de $\pi_\ell(\mathrm{V})$, o lo que es lo mismo, el menor conjunto algebraico que contiene a $\pi_\ell(\mathrm{V})$, sabemos que es $\mathcal{V}(\mathcal{I}(\pi_\ell(\mathrm{V})))$ (ver Resultados Previos, *Capítulo 1*). Luego lo que tenemos que probar es que $\mathcal{V}(\mathcal{I}(\pi_\ell(\mathrm{V}))) = \mathcal{V}(I_\ell)$.

$\boxed{\subset}$ Como por el *Lema 3.2.2* sabemos que $\pi_\ell(\mathrm{V}) \subset \mathcal{V}(I_\ell)$, se tiene que

$$\mathcal{V}(\mathcal{I}(\pi_\ell(\mathrm{V}))) = \overline{\pi_\ell(\mathrm{V})} \subseteq \overline{\mathcal{V}(I_\ell)} = \mathcal{V}(I_\ell).$$

$\boxed{\supset}$ Vamos a ver que $\mathcal{I}(\pi_\ell(\mathrm{V})) \subset \sqrt{I_\ell}$ (una vez probado esto, se tiene que $\mathcal{V}(\sqrt{I_\ell}) = \mathcal{V}(I_\ell) \subseteq \mathcal{V}(\mathcal{I}(\pi_\ell(\mathrm{V})))$.

Sea $f \in \mathcal{I}(\pi_\ell(\mathrm{V}))$. Se tiene

$$f(a_{\ell+1}, \ldots, a_n) = 0, \ \ \forall (a_{\ell+1}, \ldots, a_n) \in \pi_\ell(\mathrm{V})$$

donde $f \in \mathbb{C}[x_{\ell+1}, \ldots, x_n] \subseteq \mathbb{C}[x_1, \ldots, x_n]$. Luego

$$f(a_1, \ldots, a_n) = 0, \ \ \forall a_1, \ldots, a_\ell \in \mathbb{C} \text{ y } \forall (a_{\ell+1}, \ldots, a_n) \in \pi_\ell(\mathrm{V}).$$

En particular,

$$f(a_1, \ldots, a_n) = 0 \ \ \forall (a_1, \ldots, a_n) \in \mathrm{V} \ \ \Rightarrow \ \ f \in \mathcal{I}(\mathrm{V}) = \mathcal{I}(\mathcal{V}(I)) = \sqrt{I}$$

(hemos usado aquí el *Nullstellensatz*). Luego existe $N \in \mathbb{N}$ tal que $f^N \in I$. Pero como $f \in \mathbb{C}[x_{\ell+1}, \ldots, x_n] \ \ \Rightarrow \ \ f^N \in \mathbb{C}[x_{\ell+1}, \ldots, x_n]$. Luego $f^N \in I \cap \mathbb{C}[x_{\ell+1}, \ldots, x_n] = I_\ell$, por lo que $f \in \sqrt{I_\ell}$.

□

Nota. *Como ya vimos cuando enunciamos el* Teorema de Extensión*, si algún polinomio f_i verifica que* $\mathrm{tl}(f_i) = c_i x_1^{N_i}$ *con* $c_i \in \mathbb{C} - \{0\}$*; toda solución parcial de $\mathcal{V}(I_1)$ se puede extender a una solución de $\mathcal{V}(I)$, luego en ese caso, $\pi_1(\mathrm{V}) = \mathcal{V}(I_1)$.*

En este capítulo hemos trabajado con el cuerpo de los números complejos, $\mathbb{C}$. En realidad, todos estos resultados son válidos para cualquier cuerpo $\mathbb{K}$ algebraicamente cerrado dado que solo hemos usado el *Nullstellensatz* que es válido sobre cualquier cuerpo algebraicamente cerrado.

3.3. Continuación de la demostración del Teorema de la Clausura.

Vamos a dedicar esta última sección a demostrar la segunda parte *ii*) del *Teorema de la Clausura*. En primer lugar lo haremos en el caso particular donde $\ell = 1$, aunque realmente no necesitemos esta prueba para la demostración del caso general que haremos posteriormente, aun así es interesante ver este caso particular.

Teorema 3.3.1. *Sea* $\mathrm{V} = \mathcal{V}(f_1, \ldots, f_s) \subset \mathbb{A}^n(\mathbb{C})$ *y sea I_1 el primer ideal de eliminación del ideal $I = \langle f_1, \ldots, f_s \rangle$. Entonces si $\mathrm{V} \neq \emptyset$, existe un conjunto algebraico $W \subset \mathbb{A}^{n-1}(\mathbb{C})$, $W \subsetneq \mathcal{V}(I_1)$ tal que $\mathcal{V}(I_1) - W \subset \pi_1(\mathrm{V})$.*

Demostración.

Por el *Teorema de Extensión Geométrico*, tenemos que (con las notaciones del *Teorema 3.2.3*)

$$\mathcal{V}(I_1) = \pi_1(\mathrm{V}) \cup (\mathcal{V}(h_1, \ldots, h_s) \cap \mathcal{V}(I_1)) .$$

Sea $W = \mathcal{V}(h_1, \ldots, h_s) \cap \mathcal{V}(I_1)$ que es un subconjunto algebraico de $\mathbb{A}^{n-1}(\mathbb{C})$ por ser intersección de otros dos conjuntos algebraicos.

Si $W \neq \mathcal{V}(I_1)$, entonces ya tenemos el resultado.

Si $W = \mathcal{V}(I_1)$, probaremos al final que podemos expresar

$$\mathrm{V} = \mathcal{V}(f_1, \ldots, f_s, h_1, \ldots, h_s). \tag{3.1}$$

Consideremos $I = \langle f_1, \ldots, f_s \rangle$ nuestro ideal original e $\widetilde{I} = \langle f_1, \ldots, f_s, h_1, \ldots, h_s \rangle$. En principio estos dos ideales no tienen por qué ser iguales, aunque sí definen el mismo conjunto algebraico por (3.1). Como no son iguales, en general tampoco lo son sus ideales de eliminación I_1 y $\widetilde{I}_1$. Sin embargo, notemos que $\mathcal{V}(I_1) = \mathcal{V}(\widetilde{I}_1)$ pues

$$\pi_1(\mathrm{V}) = \pi_1(\mathcal{V}(f_1, \ldots, f_s)) = \pi_1(\mathcal{V}(I)) = \pi_1(\mathcal{V}(\widetilde{I})),$$

y como $\mathcal{V}(I_1)$ es el menor conjunto algebraico que contiene a $\pi_1(\mathcal{V}(I))$, del mismo modo que $\mathcal{V}(\widetilde{I}_1)$ es el menor conjunto algebraico que contiene a $\pi_1(\mathcal{V}(\widetilde{I}))$, tenemos que $\mathcal{V}(I_1) = \mathcal{V}(\widetilde{I}_1)$.

Recordemos que tenemos, para $i = 1, \ldots, s$,

$$f_i = h_i(x_2, \ldots, x_n) x_1^{N_i} + (\text{términos con grado } < N_i \ en \ x_1)$$

donde $N_i \geq 0$ y $h_i \in \mathbb{C}[x_2, \ldots, x_n]$ no nulo.

Definimos entonces $\widetilde{f}_i = f_i - h_i x_1^{N_i}$. Notemos que para cada $i = 1, \ldots, s$; el polinomio $\widetilde{f}_i$ es nulo o tiene en x_1 grado estrictamente menor que N_i. Tenemos además que

$$\widetilde{I} = \langle f_1, \ldots, f_s, h_1, \ldots, h_s \rangle = \langle \widetilde{f}_1, \ldots, \widetilde{f}_s, h_1, \ldots, h_s \rangle \tag{3.2}$$

y por tanto

$$\mathrm{V} = \mathcal{V}(f_1, \ldots, f_s) = \mathcal{V}(f_1, \ldots, f_s, h_1, \ldots, h_s) = \mathcal{V}(\widetilde{f}_1, \ldots, \widetilde{f}_s, h_1, \ldots, h_s).$$

Aplicamos el *Teorema de Extensión Geométrico*, a $\mathcal{V}(\widetilde{f}_1, \ldots, \widetilde{f}_s, h_1, \ldots, h_s)$.

$$\mathcal{V}(I_1) = \mathcal{V}(\widetilde{I}_1) = \pi_1(\mathrm{V}) \cup \widetilde{W}$$

donde $\widetilde{W}$ son aquellas soluciones parciales, es decir, soluciones de $\mathcal{V}(\widetilde{I}_1)$, donde los coeficientes líderes como polinomios en $\mathbb{C}[x_2, \ldots, x_n][x_1]$ de $\widetilde{f}_1, \ldots, \widetilde{f}_s, h_1, \ldots, h_s$ se anulan simultáneamente. Además notemos que esta descomposición es diferente de la que teníamos al principio pues los coeficientes líderes de los generadores son ahora en general distintos.

Si ahora $\widetilde{W} \neq \mathcal{V}(I_1)$; ya tenemos el resultado probado. Caso contrario, habría que repetir todo el proceso. Y así sucesivamente hasta que en alguna etapa demos en esta descomposición con un conjunto algebraico estrictamente menor que $\mathcal{V}(I_1)$ y ya habríamos terminado. La clave está en notar que cada vez que repetimos el proceso, los grados N_i en x_1 bajan o se mantienen en 0; por lo que al final si en ninguna etapa conseguimos que $\widetilde{W} \neq \mathcal{V}(I_1)$ todos los generadores tendrían grado 0 en x_1. En este caso, si $(a_2, \ldots, a_n)$ es una solución parcial siempre se puede extender

a $(a_1, a_2, \ldots, a_n) \in \mathrm{V}$, $\forall a_1 \in \mathbb{C}$. Lo cual quiere decir que $\pi_1(\mathrm{V}) = \mathcal{V}(I_1)$ y basta tomar como conjunto algebraico W el vacío.

Para completar la prueba falta demostrar (3.1).

$$\mathrm{V} = \mathcal{V}(f_1, \ldots, f_s, h_1, \ldots, h_s)$$

$\boxed{\supset}$ Como $\langle f_1, \ldots, f_s\rangle \subseteq \langle f_1, \ldots, f_s, h_1, \ldots, h_s\rangle$ tenemos que $\mathcal{V}(f_1, \ldots, f_s, h_1, \ldots, h_s) \subseteq \mathcal{V}(f_1, \ldots, f_s) = \mathrm{V}$.

$\boxed{\subset}$ Sea $(a_1, \ldots, a_n) \in \mathrm{V}$. Entonces

$$(a_2, \ldots, a_n) \in \pi_1(\mathrm{V}) \subset \mathcal{V}(I_1) = W = \mathcal{V}(h_1, \ldots, h_s) \cap \mathcal{V}(I_1) \;\Rightarrow$$

$$(a_2, \ldots, a_n) \in \mathcal{V}(h_1, \ldots, h_s) \;\Rightarrow\; h_i(a_2, \ldots, a_n) = 0,\ \forall i = 1, \ldots, s.$$

Luego, $(a_1, \ldots, a_n) \in \mathcal{V}(f_1, \ldots, f_s, h_1, \ldots, h_s)$.

□

En lo que queda de sección nos dedicaremos a terminar poco a poco la prueba de *ii*) del *Teorema de la Clausura* para $\ell > 1$. Lo que haremos será probarlo antes en el supuesto de que $\mathrm{V} = \mathcal{V}(f_1, \ldots, f_s)$ sea un conjunto algebraico irreducible, es decir, una variedad algebraica. Y luego extenderlo al caso general. Vamos a ver a continuación algunas observaciones relativas a este caso particular.

Observación 3.3.2. *Si el ideal $I = \langle f_1, \ldots, f_s\rangle \subset \mathbb{C}[x_1, \ldots, x_n]$ es primo entonces para cada $\ell = \{1, \ldots, n-1\}$ el ideal $I_l = I \cap \mathbb{C}[x_{\ell+1}, \ldots, x_n]$ es primo (La contracción de un ideal primo es un ideal primo. [1], pag 4.), y por tanto $\mathcal{V}(I_\ell)$ es una variedad algebraica (ver Resultados Previos,* Capítulo 1*). Además si $\mathrm{V} = \mathcal{V}(I)$, por el apartado i) del* Teorema de la Clausura, 3.2.4*, el conjunto algebraico $\overline{\pi_\ell(\mathrm{V})} = \mathcal{V}(I_\ell)$ es irreducible si V lo es.*

Nuestro objetivo es ahora probar el apartado *ii*) del *Teorema de la Clausura, 3.2.4* en el caso de que tengamos un conjunto algebraico irreducible, es decir:

Sea $\mathrm{V} = \mathcal{V}(f_1, \ldots, f_s) \subset \mathbb{A}^n(\mathbb{C})$, $\mathrm{V} \neq \emptyset$ e irreducible. Sea I_ℓ el ℓ-ésimo ideal de eliminación del ideal $I = \langle f_1, \ldots, f_s\rangle$ con $\ell \in \{1, \ldots, n-1\}$. Entonces existe un conjunto algebraico $W \subsetneq \mathcal{V}(I_\ell)$ tal que $\mathcal{V}(I_\ell) - W \subset \pi_\ell(\mathrm{V})$.

De hecho vamos a probar un teorema más fuerte, del cual se obtiene directamente el resultado anterior tomando $W_0 = \emptyset$.

Teorema 3.3.3. *Sea $\mathrm{V} = \mathcal{V}(f_1, \ldots, f_s) \subset \mathbb{A}^n(\mathbb{C})$, $\mathrm{V} \neq \emptyset$ e irreducible. Sea I_ℓ el ℓ-ésimo ideal de eliminación del ideal $I = \langle f_1, \ldots, f_s\rangle$ con $\ell \in \{1, \ldots, n-1\}$. Dado un conjunto algebraico $W_0 \subsetneq \mathrm{V}$, existe un conjunto algebraico $W_\ell \subsetneq \mathcal{V}(I_\ell)$ tal que*

$$\mathcal{V}(I_\ell) - W_\ell \subset \pi_\ell(\mathrm{V} - W_0).$$

Demostración.

En primer lugar por el apartado *i*) del *Teorema de la Clausura, 3.2.4* se tiene $\overline{\pi_\ell(\mathrm{V})} = \mathcal{V}(I_\ell)$. En la igualdad anterior, el conjunto de la izquierda solo depende de $\mathrm{V} = \mathcal{V}(\mathcal{I}(\mathrm{V}))$. Por tanto podemos tomar $I = \mathcal{I}(\mathrm{V})$ que es un ideal primo dado que V es irreducible (ver Resultados Previos, *Capítulo 1*).

Lo demostramos por inducción sobre $\ell \in \{1, \ldots, n-1\}$. Sea $W_0 \subsetneq \mathrm{V}$.

En primer lugar lo probamos para $\ell = 1$. Como $W_0 \subsetneq \mathrm{V}$, existirá $(a_1, \ldots, a_n) \in V - W_0$. E igualmente debe existir $f \in \mathcal{I}(W_0)$ tal que $f(a_1, \ldots, a_n) \neq 0$, pues si todo polinomio de $\mathcal{I}(W_0)$ se anula en $(a_1, \ldots, a_n)$ sería $(a_1, \ldots, a_n) \in \mathcal{V}(\mathcal{I}(W_0)) = W_0$ y tendríamos una contradicción. Distinguimos dos casos:

Caso 1 Supongamos que $\forall (b_2, \ldots, b_n) \in \mathcal{V}(I_1)$ y $\forall b_1 \in \mathbb{C}$ se tiene que $(b_1, b_2, \ldots, b_n) \in \mathrm{V}$.

Vamos a escribir f como un polinomio en x_1 con coeficientes en $\mathbb{C}[x_2, \ldots, x_n]$:

$$f = \sum_{i=0}^{m} g_i(x_2, \ldots, x_n) x_1^i.$$

Y definimos $W_1 = \mathcal{V}(I_1) \cap \mathcal{V}(g_0, \ldots, g_m)$, veamos que el conjunto W_1 es el que buscamos.

En primer lugar por la elección que hicimos de $(a_1, \ldots, a_n)$, como $(a_1, \ldots, a_n) \in \mathrm{V}$ en particular $(a_2, \ldots, a_n) \in \pi_1(\mathrm{V}) \subset \mathcal{V}(I_1)$. Pero por otro lado como $f(a_1, \ldots, a_n) \neq 0$ debe ser $g_i(a_2, \ldots, a_n) \neq 0$ para algún i de modo que $(a_2, \ldots, a_n) \notin \mathcal{V}(g_0, \ldots, g_m)$ y por tanto $(a_2, \ldots, a_n) \notin W_1$. Luego tenemos que $W_1 \subsetneq \mathcal{V}(I_1)$.

Ahora probemos que $\mathcal{V}(I_1) - W_1 \subset \pi_1(\mathrm{V} - W_0)$. Sea entonces $(c_2, \ldots, c_n) \in \mathcal{V}(I_1) - W_1$, lo que significa que $g_i(c_2, \ldots, c_n) \neq 0$ para algún i y en ese caso $f(x_1, c_2, \ldots, c_n)$ no es el polinomio nulo. Un polinomio no nulo con coeficientes en $\mathbb{C}$ tiene un número finito de raíces en $\mathbb{C}$ y como $\mathbb{C}$ es infinito, debe existir $c_1 \in \mathbb{C}$ tal que $f(c_1, c_2, \ldots, c_n) \neq 0$ lo que implica que $(c_1, c_2, \ldots, c_n) \notin W_0$ pues en su momento tomamos $f \in \mathcal{I}(W_0)$. Pero como estamos en el **Caso 1** se tiene entonces que $(c_1, c_2, \ldots, c_n) \in \mathrm{V}$. De modo que entonces $(c_1, c_2, \ldots, c_n) \in \mathrm{V} - W_0$ y $(c_2, \ldots, c_n) \in \pi_1(\mathrm{V} - W_0)$.

Luego ya tenemos en este caso asegurada la existencia de un conjunto $W_1 \subsetneq \mathcal{V}(I_1)$ tal que $\mathcal{V}(I_1) - W_1 \subset \pi_1(\mathrm{V} - W_0)$.

Caso 2 Supongamos que $\exists (b_2, \ldots, b_n) \in \mathcal{V}(I_1)$ y $\exists b_1 \in \mathbb{C}$ tal que $(b_1, b_2, \ldots, b_n) \notin \mathrm{V}$.

Como $\mathcal{V}(I) = \mathrm{V}$ es irreducible, $\mathcal{I}(\mathrm{V}) = \sqrt{I} = I$ pues es un ideal primo. Luego debe existir $h \in I$ tal que $h(b_1, b_2, \ldots, b_n) \neq 0$. Escribimos h como un polinomio en x_1 con coeficientes en $\mathbb{C}[x_2, \ldots, x_n]$:

$$h = \sum_{i=0}^{r} u_i(x_2, \ldots, x_n) x_1^i.$$

Como $h(b_1, b_2, \ldots, b_n) \neq 0$, para algún i debe ser $u_i(b_2, \ldots, b_n) \neq 0$ por lo que $u_i \notin I_1$ ya que $(b_2, \ldots, b_n) \in \mathcal{V}(I_1)$. Además si $u_r \in I_1 \Rightarrow u_r(b_2, \ldots, b_n) = 0 \Rightarrow (h - u_r x_1^r)(b_1, b_2, \ldots, b_n) = h(b_1, \ldots, b_n) = 0$. Y por otro lado $u_r \in I_1 \subset I$ luego, $h - u_r x_1^r \in I$ lo que implica que podemos reemplazar h por $h - u_r x_1^r$. De manera que repitiendo el proceso y renombrando podemos suponer que en el h original $u_r \notin I_1$.

El siguiente objetivo es probar que

$$\exists v_i \in \mathbb{C}[x_2, \ldots, x_n] \text{ tal que } \sum_{i=0}^{r} v_i f^i \in I \text{ y } v_0 \notin I_1. \tag{3.3}$$

Para ello veamos f y $h = \sum_{i=0}^{r} u_i(x_2, \ldots, x_n) x_1^i$ como polinomios en x_1. Si dividimos f entre h (como polinomios en x_1) es posible que algunos coeficientes del cociente y del resto no sean polinomios en $\mathbb{C}[x_2, \ldots, x_n]$; así que consideramos $u_r^{N_1} f$ en lugar de f, tomando

N_1 suficientemente grande como para que no aparezcan denominadores al dividir $u_r^{N_1} f$ entre h, obteniendo entonces que

$$u_r^{N_1} f = q_1 h + v_{10} + v_{11}x_1 + \cdots + v_{1,r-1}x_1^{r-1}$$

donde $q_1 \in \mathbb{C}[x_1, \ldots, x_n]$ y $v_{1i} \in \mathbb{C}[x_2, \ldots, x_n]$ para $i = 0, \ldots, r-1$.

Ahora hacemos el mismo procedimiento para f^j con $j \in \{0, \ldots, r\}$.

$$\begin{array}{ccccccccccc}
u_r^{N_0} & = & q_0 h & + & v_{00} & + & v_{01}x_1 & + & \cdots & + & v_{0,r-1}x_1^{r-1} \\
u_r^{N_1} f & = & q_1 h & + & v_{10} & + & v_{11}x_1 & + & \cdots & + & v_{1,r-1}x_1^{r-1} \\
u_r^{N_2} f^2 & = & q_2 h & + & v_{20} & + & v_{21}x_1 & + & \cdots & + & v_{2,r-1}x_1^{r-1} \\
\vdots & & \vdots & & \vdots & & \vdots & & \ddots & & \vdots \\
u_r^{N_r} f^r & = & q_r h & + & v_{r0} & + & v_{r1}x_1 & + & \cdots & + & v_{r,r-1}x_1^{r-1}
\end{array}$$

En general,

$$u_r^{N_j} f^j = q_j h + v_{j0} + v_{j1}x_1 + \cdots + v_{j,r-1}x_1^{r-1} \text{ con } j = 0, \ldots, r. \tag{3.4}$$

Usamos ahora el anillo de coordenadas de $\mathcal{V}(I_1)$, $\mathcal{A}(\mathcal{V}(I_1))$, al que para abreviar notaremos $\mathcal{A}$. Sabemos que $\mathcal{A} \cong \frac{\mathbb{C}[x_2,\ldots,x_n]}{I_1}$ y que además $\mathcal{A}$ es un dominio de integridad pues I_1 es un ideal primo ya que es la restricción a $\mathbb{C}[x_2, \ldots, x_n]$ del ideal primo I. Consideramos las clases de equivalencia de los elementos $v_{ji} \in \mathbb{C}[x_2, \ldots, x_n]$ en el anillo de coordenadas, es decir,

$$[v_{ji}] = v_{ji} + \langle I_1 \rangle \in \mathcal{A}, \text{ con } j = 0, \ldots, r, \ i = 0, \ldots, r-1.$$

$$\text{Sea } M = \begin{pmatrix} [v_{00}] & \cdots & [v_{0,r-1}] \\ \vdots & \ddots & \vdots \\ \lceil v_{r,0} \rfloor & \cdots & [v_{r,r-1}] \end{pmatrix}$$

la matriz de dimensión $(r+1) \times r$ con entradas en el anillo $\mathcal{A}$. De manera que podemos ver las $r+1$ filas de la matriz M como $r+1$ vectores de $Q(\mathcal{A})^r$, que deben ser obligatoriamente linealmente dependientes (donde $Q(\mathcal{A})$ es el cuerpo de fracciones de $\mathcal{A}$). Luego (después de reducir a un común denominador) existen $[w_0], \ldots, [w_r] \in \mathcal{A}$ no todos nulos tal que

$$\sum_{j=0}^{r} [w_j][v_{ji}] = [0] \text{ para } i = 0, \ldots, r-1,$$

con $w_j \in \mathbb{C}[x_2, \ldots, x_n]$ para $j = 0, \ldots, r$. O lo que es lo mismo, $\exists w_0, \ldots, w_r \in \mathbb{C}[x_2, \ldots, x_n]$ tal que $\sum_{j=0}^{r} w_j v_{ji} \in I_1$ para $i = 0, \ldots, r-1$, y $w_j \notin I_1$ para algún $j \in \{0, \ldots, r\}$.

Multiplicando (3.4) por w_j, obtenemos

$$w_j u_r^{N_j} f^j = w_j q_j h + w_j v_{j0} + w_j v_{j1}x_1 + \cdots + w_j v_{j,r-1}x_1^{r-1} \text{ con } j = 0, \ldots, r.$$

Si los sumamos, $\sum_{j=0}^{r} w_j u_r^{N_j} f^j \in I$ pues $h \in I$ y $\sum_{j=0}^{r} w_j v_{ji} \in I_1 \subset I$ para $i = 0, \ldots, r-1$.

Sea entonces $v_j = w_j u_r^{N_j}$ de manera que $\sum_{j=0}^{r} v_j f^j \in I$. Falta ver que $v_0 \notin I_1$.

Como $u_r \notin I_1$, $w_j \notin I_1$ para algún $j \in \{0, \ldots, r\}$ e I_1 es un ideal primo, $v_j \notin I_1$ para algún $j \in \{0, \ldots, r\}$. Supongamos $v_t \notin I_1$ y $v_0, v_1, \ldots, v_{t-1} \in I_1$.

$$\sum_{j=0}^{r} v_j f^j = \sum_{j=0}^{t-1} v_j f^j + \sum_{j=t}^{r} v_j f^j \in I \ \Rightarrow$$

$$\sum_{j=0}^{t-1} v_j f^j + f^t \sum_{j=t}^{r} v_j f^{j-t} \in I \;\Rightarrow$$

Como $\sum_{j=0}^{t-1} v_j f^j \in I$ pues $v_0, \ldots, v_{t-1} \in I_1$ se tiene que

$$f^t \sum_{j=t}^{r} v_j f^{j-t} \in I$$

y ahora como $f \notin I$ e I es primo,

$$\sum_{j=t}^{r} v_j f^{j-t} \in I.$$

Luego renombrando podemos suponer que $v_0 \notin I_1$.

Partiendo de que se verifica (3.3) vamos a ver que entonces se tiene

$$\pi_1(\mathrm{V}) \cap \left(\mathbb{A}^{n-1}(\mathbb{C}) - \mathcal{V}(v_0)\right) \subset \pi_1(\mathrm{V} - W_0). \tag{3.5}$$

Sea $(c_2, \ldots, c_n) \in \pi_1(\mathrm{V}) \cap \left(\mathbb{A}^{n-1}(\mathbb{C}) - \mathcal{V}(v_0)\right)$, entonces tenemos que debe existir $c_1 \in \mathbb{C}$ tal que $(c_1, c_2, \ldots, c_n) \in \mathrm{V}$ (lo que implica que todo polinomio de $I = \mathcal{I}(\mathrm{V})$ debe anularse en el punto $(c_1, c_2, \ldots, c_n)$) y que $(c_2, \ldots, c_n) \notin \mathcal{V}(v_0)$. Ahora usamos (3.3) y entonces se tiene que

$$0 = \left(\sum_{i=0}^{r} v_i f^i\right)(c_1, c_2, \ldots, c_n) =$$

$$= v_0(c_2, \ldots, c_n) + f(c_1, c_2, \ldots, c_n)\left(\sum_{i=1}^{r} v_i(c_2, \ldots, c_n)\left(f(c_1, c_2, \ldots, c_n)\right)^{i-1}\right)$$

Como $v_0(c_2, \ldots, c_n) \neq 0$, se tiene que $f(c_1, c_2, \ldots, c_n) \neq 0$ y teníamos que $f \in \mathcal{I}(W_0)$ por lo que $(c_1, c_2, \ldots, c_n) \notin W_0 \Rightarrow (c_1, c_2, \ldots, c_n) \in \mathrm{V} - W_0$, y por tanto, $(c_2, \ldots, c_n) \in \pi_1(\mathrm{V} - W_0)$ lo que prueba (3.5).

Falta ver para acabar con el **Caso 2** que $\exists W_1 \subsetneq \mathcal{V}(I_1)$ tal que $\mathcal{V}(I_1) - W_1 \subset \pi_1(\mathrm{V} - W_0)$.

Como $u_r, v_0 \notin I_1$ y I_1 es un ideal primo, se tiene que $g = u_r v_0 \notin I_1$. Sea entonces el conjunto $W_1 = \mathcal{V}(g) \cap \mathcal{V}(I_1) \subset \mathcal{V}(I_1)$. Además no se tiene $W_1 = \mathcal{V}(I_1)$ pues en ese caso tendríamos que $\mathcal{V}(I_1) \subseteq \mathcal{V}(g)$ y por tanto $g \in \sqrt{I_1} = I_1$ lo cual es una contradicción. Luego se tiene que $W_1 = \mathcal{V}(g) \cap \mathcal{V}(I_1) \subsetneq \mathcal{V}(I_1)$.

Sea $(c_2, \ldots, c_n) \in \mathcal{V}(I_1) - W_1$, lo que implica que $g(c_2, \ldots, c_n) \neq 0$, luego ni u_r ni v_0 se anulan en $(c_2, \ldots, c_n)$.

Como $v_0(c_2, \ldots, c_n) \neq 0$ se tiene que $(c_2, \ldots, c_n) \in \mathbb{A}^n(\mathbb{C}) - \mathcal{V}(v_0)$.

Como $h = \sum_{i=0}^{r} u_i(x_2, \ldots, x_n) x_1^i \in I$ podemos considerar h otro generador del ideal I, es decir, $I = \langle f_1, \ldots, f_s, h\rangle$ donde $u_r(c_2, \ldots, c_n) \neq 0$.

Entonces tenemos que $(c_2, \ldots, c_n) \in \mathcal{V}(I_1)$ y $u_r(c_2, \ldots, c_n) \neq 0$, basta aplicar el *Teorema de Extensión* para saber que la solución parcial $(c_2, \ldots, c_n)$ se puede extender, es decir, $\exists c_1 \in \mathbb{C}$ tal que $(c_1, c_2, \ldots, c_n) \in \mathrm{V}$; y por tanto $(c_2, \ldots, c_n) \in \pi_1(\mathrm{V})$.

Luego $(c_2, \ldots, c_n) \in \pi_1(\mathrm{V}) \cap (\mathbb{A}^n(\mathbb{C}) - \mathcal{V}(v_0))$ y aplicando (3.5), $(c_2, \ldots, c_n) \in \pi_1(\mathrm{V} - W_0)$, lo que demuestra que $\mathcal{V}(I_1) - W_1 \subset \pi_1(\mathrm{V} - W_0)$.

De manera que obtenemos que existe un conjunto algebraico $W_1 \subsetneq \mathcal{V}(I_1)$ tal que $\mathcal{V}(I_1) - W_1 \subset \pi_1(\mathrm{V} - W_0)$.

Hipótesis de Inducción. Suponemos el resultado cierto para $\ell - 1$, es decir; existe un conjunto algebraico $W_{\ell-1} \subsetneq \mathcal{V}(I_{\ell-1})$ tal que $\mathcal{V}(I_{\ell-1}) - W_{\ell-1} \subset \pi_{\ell-1}(\mathrm{V} - W_0)$ y suponemos $\ell > 1$.

Vamos a ver ahora la prueba para ℓ. Fijémonos en que el ideal I_ℓ es el $(\ell-1)$-ésimo ideal de eliminación de I_1, y que por la *Observacion 3.3.2* tenemos que $\mathcal{V}(I_1)$ es irreducible. Por la primera parte de la inducción, teníamos que existe $W_1 \subsetneq \mathcal{V}(I_1)$ tal que $\mathcal{V}(I_1) - W_1 \subset \pi_1(\mathrm{V} - W_0)$. Luego podemos aplicar la *Hipótesis de Inducción* a este conjunto $W_1 \subsetneq \mathcal{V}(I_1)$. De manera que tenemos que existe $W_\ell \subsetneq \mathcal{V}(I_\ell)$ tal que

$$\mathcal{V}(I_\ell) - W_\ell \subset \widetilde{\pi}_{\ell-1}(\mathcal{V}(I_1) - W_1)$$

donde $\widetilde{\pi}_{\ell-1} : \mathbb{A}^{n-1}(\mathbb{C}) \longrightarrow \mathbb{A}^{n-\ell}(\mathbb{C})$ es la proyección en las últimas $n - \ell$ variables.

Pero observemos que $\pi_\ell = \widetilde{\pi}_{\ell-1} \circ \pi_1$. Y como por el caso $\ell = 1$ teníamos que $\mathcal{V}(I_1) - W_1 \subset \pi_1(\mathrm{V} - W_0)$ obtenemos ya el resultado pues

$$\mathcal{V}(I_\ell) - W_\ell \subset \widetilde{\pi}_{\ell-1}(\mathcal{V}(I_1) - W_1) \subset \widetilde{\pi}_{\ell-1}(\pi_1(\mathrm{V} - W_0)) = \pi_\ell(\mathrm{V} - W_0).$$

□

Necesitamos por último un lema más, que añadimos a continuación, para poder probar el resultado general.

Lema 3.3.4. *Sean* $\mathrm{V}, \mathrm{V}_1, \ldots, \mathrm{V}_m \subset \mathbb{A}^n(\mathbb{C})$ *conjuntos algebraicos con* $m \geq 2$ *y* V *irreducible. Si* $\mathrm{V} \subset \mathrm{V}_1 \cup \cdots \cup \mathrm{V}_m$*, entonces* $\mathrm{V} \subset \mathrm{V}_i$ *para algún* $i \in \{1, \ldots, m\}$.

Demostración.

La hacemos por inducción sobre m. Supongamos $m = 2$.

Si $\mathrm{V} \subset \mathrm{V}_1 \cup \mathrm{V}_2$ se tiene que (ver Resultados Previos, *Capítulo 1*) $\mathcal{I}(\mathrm{V}_1) \cap \mathcal{I}(\mathrm{V}_2) = \mathcal{I}(\mathrm{V}_1 \cup \mathrm{V}_2) \subset \mathcal{I}(\mathrm{V})$ donde $\mathcal{I}(\mathrm{V})$ es un ideal primo. Por la Proposición 1.11 de [1] se tiene que o bien $\mathcal{I}(\mathrm{V}_1) \subset \mathcal{I}(\mathrm{V})$ o $\mathcal{I}(\mathrm{V}_2) \subset \mathcal{I}(\mathrm{V})$ y por tanto $\mathrm{V} \subset \mathrm{V}_1$ o $\mathrm{V} \subset \mathrm{V}_2$.

Hipótesis de Inducción. Suponemos el resultado cierto para $m - 1$.

Sea entonces $\mathrm{V} \subset \mathrm{V}_1 \cup \cdots \cup \mathrm{V}_m$ con V irreducible. Sea $\mathrm{V}' = \mathrm{V}_2 \cup \cdots \cup \mathrm{V}_m$, de modo que tenemos entonces $\mathrm{V} \subset \mathrm{V}_1 \cup \mathrm{V}'$. Aplicamos el caso para $m = 2$ y obtenemos que $\mathrm{V} \subset \mathrm{V}_1$ o $\mathrm{V} \subset \mathrm{V}'$. Si se tuviese $\mathrm{V} \subset \mathrm{V}'$ aplicamos la *Hipótesis de Inducción* a V'. Luego se puede asegurar que $\mathrm{V} \subset \mathrm{V}_i$ para algún $i \in \{1, \ldots, m\}$. □

Teorema 3.3.5. *Sea* $\mathrm{V} = \mathcal{V}(f_1, \ldots, f_s) \subset \mathbb{A}^n(\mathbb{C})$ *y sea* I_ℓ *el* ℓ*-ésimo ideal de eliminación del ideal* $I = \langle f_1, \ldots, f_s \rangle$ *con* $\ell \in \{1, \ldots, n-1\}$*. Entonces si* $\mathrm{V} \neq \emptyset$*, existe un conjunto algebraico* $W \subsetneq \mathcal{V}(I_\ell)$ *tal que* $\mathcal{V}(I_\ell) - W \subset \pi_\ell(\mathrm{V})$.

El resultado anterior para $\ell = 1$ es el *Teorema 3.3.1.*

Demostración.

Podemos escribir el conjunto algebraico V como unión de sus componentes irreducibles, de manera que tendremos que $\mathrm{V} = \mathrm{V}_1 \cup \cdots \cup \mathrm{V}_m$ donde cada V_i es una variedad algebraica. Consideramos entonces para cada $i \in \{1, \ldots, m\}$ los conjuntos

$$\pi_\ell(\mathrm{V}_i),\ \mathrm{V}'_i = \overline{\pi_\ell(\mathrm{V}_i)} \subset \mathbb{A}^{n-\ell}(\mathbb{C}).$$

Probaremos al final que

$$\mathcal{V}(I_\ell) = \mathrm{V}'_1 \cup \cdots \cup \mathrm{V}'_m \tag{3.6}$$

donde además sabemos que cada V'_i es irreducible por la *Observación 3.3.2* ya que como V_i es irreducible, también lo es $\overline{\pi_\ell(\mathrm{V}_i)} = \mathrm{V}'_i$. Por lo tanto tenemos una descomposición de $\mathcal{V}(I_\ell)$ en componentes irreducibles que no tiene por qué ser la descomposición minimal; pero dónde si podemos suponer que la primera de esas componentes no está contenida en la unión de las demás (pues en caso contrario, podemos quitarla). Es decir, en particular podemos suponer, $\mathrm{V}'_1 \not\subset \mathrm{V}'_i$ para todo $i = 2, \ldots, m$.

Aplicamos ahora el *Teorema 3.3.3* a V_1 (tomando $W_0 = \emptyset$), luego existe un conjunto algebraico $W_{1\ell} \subsetneq \mathrm{V}'_1$ tal que $\mathrm{V}'_1 - W_{1\ell} \subset \pi_\ell(\mathrm{V}_1)$ ya que $\mathrm{V}'_1 = \overline{\pi_l(\mathrm{V}_1)}$.

Sea entonces $W = W_{1\ell} \cup \mathrm{V}'_2 \cup \cdots \cup \mathrm{V}'_m$. Se tiene que $W \subset \mathcal{V}(I_\ell)$ y por tanto es fácil ver que

$$\begin{aligned}\mathcal{V}(I_\ell) - W &= \mathrm{V}'_1 \cup \cdots \cup \mathrm{V}'_m - (W_{1\ell} \cup \mathrm{V}'_2 \cup \cdots \cup \mathrm{V}'_m) \\ &\subset \mathrm{V}'_1 - (W_{1\ell} \cup \mathrm{V}'_2 \cup \cdots \cup \mathrm{V}'_m) \subset \mathrm{V}'_1 - W_{1\ell} \subset \pi_\ell(\mathrm{V}_1) \subset \pi_\ell(\mathrm{V}).\end{aligned}$$

Nos queda ver que $W \neq \mathcal{V}(I_\ell)$.

Reducción al Absurdo. Supongamos $W = \mathcal{V}(I_\ell)$. Entonces tendríamos que

$$W_{1\ell} \cup \mathrm{V}'_2 \cup \cdots \cup \mathrm{V}'_m = \mathrm{V}'_1 \cup \cdots \cup \mathrm{V}'_m$$

y en particular

$$\mathrm{V}'_1 \subset W_{1\ell} \cup \mathrm{V}'_2 \cup \cdots \cup \mathrm{V}'_m.$$

Como V'_1 es irreducible, aplicando el *Lema 3.3.4* se tiene que V'_1 debería estar contenido en uno de los $W_{1\ell}$, V'_2, $\cdots$, V'_m, lo cual es imposible por la elección de V'_1 y de $W_{1\ell}$ y tenemos la contradicción.

Queda por último probar (3.6).

$$\mathcal{V}(I_\ell) = \mathrm{V}'_1 \cup \cdots \cup \mathrm{V}'_m$$

$\boxed{\subset}$ Sea $(a_{\ell+1}, \ldots, a_n) \in \pi_\ell(\mathrm{V}) \ \Rightarrow$

$$\exists a_1, \ldots, a_\ell \in \mathbb{C} \text{ tal que } (a_1, \ldots, a_\ell, a_{\ell+1}, \ldots, a_n) \in \mathrm{V} = \mathrm{V}_1 \cup \cdots \cup \mathrm{V}_m \ \Rightarrow$$

$$(a_1, \ldots, a_\ell, a_{\ell+1}, \ldots, a_n) \in \mathrm{V}_j \text{ para algún } j \ \Rightarrow$$

$$(a_{\ell+1}, \ldots, a_n) \in \pi_\ell(\mathrm{V}_j) \subset \mathrm{V}'_j \subset \mathrm{V}'_1 \cup \cdots \cup \mathrm{V}'_m$$

Luego $\pi_\ell(\mathrm{V}) \subset \mathrm{V}'_1 \cup \cdots \cup \mathrm{V}'_m$. Como $\mathcal{V}(I_\ell)$ es el menor conjunto algebraico que contiene a $\pi_\ell(\mathrm{V})$ y $\mathrm{V}'_1 \cup \cdots$, $\cup \mathrm{V}'_m$ es un conjunto algebraico, se tiene que

$$\mathcal{V}(I_\ell) \subset \mathrm{V}'_1 \cup \cdots \cup \mathrm{V}'_m.$$

$\boxed{\supset}$ $\pi_\ell(\mathrm{V}_i) \subset \pi_\ell(\mathrm{V}) \subset \mathcal{V}(I_\ell),\ \ \forall i = 1, \ldots, m.$

Tomando clausura, $\mathrm{V}'_i \subset \mathcal{V}(I_\ell),\ \ \forall i = 1, \ldots, m$. Luego $\mathrm{V}'_1 \cup \cdots \cup \mathrm{V}'_m \subset \mathcal{V}(I_\ell)$.

□

Capítulo 4

Resultantes y prueba del Teorema de Extensión

El propósito de este capítulo es probar el Teorema de Extensión que enunciamos en el capítulo anterior, para ello necesitamos introducir el concepto de resultante. Este concepto surge al preguntarnos cuándo dos polinomios de una variable tiene un factor en común.

Consideremos el problema siguiente: dados $f, g \in \mathbb{K}[x]$, queremos saber si $\exists h \in \mathbb{K}[x]$ tal que $h \mid f,\ h \mid g$ y $gr(h) > 0$.

Podemos resolver esta cuestión calculando el máximo común divisor de f y g, o factorizándolos en productos de polinomios irreducibles, pero queremos evitar estos dos métodos y buscar un método que resuelva el problema usando solo álgebra lineal.

Proposición 4.1.1. *Sean $f, g \in \mathbb{K}[x]$ donde $gr(f) = \ell > 0$ y $gr(g) = m > 0$. Entonces f y g tienen un factor común si y solo si existen polinomios $A, B \in \mathbb{K}[x]$ no nulos tales que:*

i) $gr(A) \leq m - 1$ *y* $gr(B) \leq \ell - 1$.

ii) $Af + Bg = 0$.

Demostración.

Supongamos que existe $h \in \mathbb{K}[x]$ tal que $h \mid f$, $h \mid g$ y $gr(h) \geq 1$. Entonces podremos escribir

$$f = hf_1 \text{ con } gr(f_1) \leq \ell - 1$$

$$g = hg_1 \text{ con } gr(g_1) \leq m - 1$$

Entonces

$$g_1 f + (-f_1)g = g_1 h f_1 - f_1 h g_1 = 0.$$

Por lo tanto basta tomar $A = g_1$ y $B = -f_1$.

Supongamos que existen $A, B \in \mathbb{K}[x]$ no ambos nulos que verifican $i)$ e $ii)$. *Reducción al Absurdo.* Supongamos que f y g no tienen ningún factor en común, entonces $m.c.d.(f, g) = 1$ y por la identidad de Bézout existirán $\widetilde{A}, \widetilde{B} \in \mathbb{K}[x]$ tal que $\widetilde{A}f + \widetilde{B}g = 1$.

Como $Bg = -Af$, tenemos que

$$B = (\widetilde{A}f + \widetilde{B}g)B = \widetilde{A}Bf + \widetilde{B}Bg = \widetilde{A}Bf - \widetilde{B}Af = (\widetilde{A}B - \widetilde{B}A)f.$$

Como B es no nulo, debe ser $gr(B) \geq \ell$, y ya tenemos la contradicción. $\square$

Luego gracias a este resultado basta saber si existen los polinomios A y B para decidir si f y g tienen un factor común o no. Veamos que podemos trasladar la existencia de A y B a un problema de álgebra lineal.

Sean $f, g \in \mathbb{K}[x]$ donde $gr(f) = \ell > 0$ y $gr(g) = m > 0$. Buscamos $A, B \in \mathbb{K}[x]$ no nulos

$$A = c_0 + c_1 x + \cdots + c_{m-1} x^{m-1}$$

$$B = d_0 + d_1 x + \cdots + d_{\ell-1} x^{\ell-1}$$

con $c_i, d_j \in \mathbb{K}$ no todos nulos y tales que $Af + Bg = 0$.

Nuestro objetivo entonces es encontrar los coeficientes c_i, d_j. Para ello expresamos f y g como

$$f = a_0 + a_1 x + \cdots + a_\ell x^\ell \quad \text{con } a_\ell \neq 0$$

$$g = b_0 + b_1 x + \cdots + b_m x^m \quad \text{con } b_m \neq 0$$

Luego

$$\begin{gathered} 0 = Af + Bg = (c_0 + c_1 x + \cdots + c_{m-1} x^{m-1})(a_0 + a_1 x + \cdots + a_\ell x^\ell) + \\ + (d_0 + d_1 x + \cdots + d_{\ell-1} x^{\ell-1})(b_0 + b_1 x + \cdots + b_m x^m) = \\ = (a_0 c_0 + b_0 d_0) + (a_0 c_1 + a_1 c_0 + b_0 d_1 + b_1 d_0) x + \cdots \\ \cdots + (a_\ell c_{m-2} + a_{\ell-1} c_{m-1} + b_m d_{\ell-2} + b_{m-1} d_{\ell-1}) x^{m+l-2} + (a_\ell c_{m-1} + b_m d_{\ell-1}) x^{m+\ell-1}. \end{gathered}$$

Por lo que debe verificarse el sistema de ecuaciones lineales

$$\left\{ \begin{array}{ccccccccccccc} a_0 c_0 & & \cdots & & & + & b_0 d_0 & & \cdots & & & = & 0 \\ a_1 c_0 & + & a_0 c_1 & & \cdots & + & b_1 d_0 & + & b_0 d_1 & & \cdots & = & 0 \\ \ddots & & \ddots & & & & \ddots & & \ddots & & & \vdots & \\ \cdots & & a_\ell c_{m-2} & + & a_{\ell-1} c_{m-1} & \cdots & + & b_m d_{\ell-2} & + & b_{m-1} d_{\ell-1} & = & 0 \\ & & \cdots & & a_\ell c_{m-1} & & & \cdots & + & b_m d_{\ell-1} & = & 0 \end{array} \right. \tag{4.1}$$

donde tenemos $\ell + m$ ecuaciones, una por cada potencia de x y $\ell + m$ incógnitas. Luego nuestro sistema tendrá solución no nula si y solo si la matriz de los coeficientes tiene determinante nulo por el *Teorema de Rouché-Frobenius*. A la matriz de dimensión $(\ell + m) \times (\ell + m)$ con coefientes en $\mathbb{K}$ que define nuestro sistema le llamaremos matriz de Sylvester de f y g respecto de x.

$$Syl(f, g, x) = \begin{pmatrix} a_0 & & & & b_0 & & & \\ a_1 & a_0 & & & b_1 & b_0 & & \\ & a_1 & \ddots & & & b_1 & \ddots & \\ \vdots & & \ddots & a_0 & \vdots & & \ddots & b_0 \\ & \vdots & & a_1 & & \vdots & & b_1 \\ a_\ell & & & & b_m & & & \\ & a_\ell & & \vdots & & b_m & & \vdots \\ & & \ddots & & & & \ddots & \\ & & & a_\ell & & & & b_m \end{pmatrix}$$

Llamaremos Resultante de f y g respecto de x al determinante de la matriz de Sylvester de f y g respecto de x,

$$Res(f, g, x) = |\, Syl(f, g, s)\,|.$$

Por lo tanto obtenemos directamente de la *Proposición 4.1.1* que dados $f, g \in \mathbb{K}[x]$ con $gr(f) = \ell > 0$ y $gr(g) = m > 0$, f y g tendrán un factor en común si y solo si $Res(f, g, x) = 0$.

Ejemplo 12. *Sean los polinomios $f = x^2 - 1$ y $g = x^3 - 2x^2 - 3x$. Veamos si f y g tienen algún factor en común usando la resultante (ya que en este caso es evidente que -1 es raíz común de f y g), para ello buscamos $A = c_0 + c_1x + c_2x^2$ y $B = d_0 + d_1x$ en $\mathbb{C}[x]$ no nulos tales que*

$$0 = Af + Bg = (c_0 + c_1x + c_2x^2)(x^2 - 1) + (d_0 + d_1x)(x^3 - 2x^2 - 3x) =$$

$$= -c_0 + (-c_1 - 3d_0)x + (c_0 - c_2 - 2d_0 - 3d_1)x^2 + (c_1 + d_0 - 2d_1)x^3 + (c_2 + d_1)x^4 = 0$$

$$\begin{cases} -c_0 & & & & & = & 0 \\ & -c_1 & & -3d_0 & & = & 0 \\ c_0 & & -c_2 & -2d_0 & -3d_1 & = & 0 \\ & c_1 & & +d_0 & -2d_1 & = & 0 \\ & & c_2 & & +d_1 & = & 0 \end{cases}$$

Luego tenemos que

$$Syl(f, g, x) = \begin{pmatrix} -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & -3 & 0 \\ 1 & 0 & -1 & -2 & -3 \\ 0 & 1 & 0 & 1 & -2 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix}$$

Y por tanto $Res(f, g, x) = |\, Syl(f, g, x)\,| = 0$. Luego f y g tienen algún factor en común.

Proposición 4.1.2. *Dados $f, g \in \mathbb{K}[x]$ con $gr(f) = \ell > 0$ y $gr(g) = m > 0$, existen $A, B \in \mathbb{K}[x]$ tales que $Af + Bg = Res(f, g, x)$.*

Demostración.

Ya hemos visto que en el caso de que f y g tengan algún factor en común, se tiene que $Af+Bg = 0$ para ciertos polinomios $A, B \in \mathbb{K}[x]$ no nulos.

Supongamos entonces que f y g no tienen ningún factor en común, entonces $m.c.d.(f, g) = 1$ y por la identidad de Bézout existirán $\widetilde{A}, \widetilde{B} \in \mathbb{K}[x]$ tales que $\widetilde{A}f + \widetilde{B}g = 1$. Si expresamos

$$f = a_0 + a_1x + \cdots + a_\ell x^\ell \quad \text{con } a_\ell \neq 0$$

$$g = b_0 + b_1x + \cdots + b_m x^m \quad \text{con } b_m \neq 0$$

Y suponemos $\widetilde{A}, \widetilde{B}$ de la forma

$$\widetilde{A} = c_0 + c_1x + \cdots + c_{m-1}x^{m-1}$$

$$\widetilde{B} = d_0 + d_1x + \cdots + d_{l-1}x^{l-1}$$

donde desconocemos los coeficientes $c_0, \ldots, c_{m-1}, d_0, \ldots, d_{\ell-1} \in \mathbb{K}$. Imponemos entonces que se verifique $\widetilde{A}f + \widetilde{B}g = 1$ y obtenemos el siguiente sistema de ecuaciones lineales:

$$\begin{cases} a_0c_0 & \cdots & & + & b_0d_0 & \cdots & & & = & 1 \\ a_1c_0 + & a_0c_1 & \cdots & + & b_1d_0 + & b_0d_1 & \cdots & & = & 0 \\ \ddots & \ddots & & & \ddots & \ddots & & & & \vdots \\ \cdots & a_\ell c_{m-2} + & a_{\ell-1}c_{m-1} & \cdots & + & b_m d_{\ell-2} + & b_{m-1}d_{\ell-1} & & = & 0 \\ & \cdots & a_\ell c_{m-1} & & & \cdots & + & b_m d_{\ell-1} & = & 0 \end{cases}$$

Este sistema tiene la misma matriz de coeficientes que el sistema (4.1), es decir, la matriz de coeficientes coincide con $Syl(f,g,x)$ y además tenemos que $|\,Syl(f,g,x)\,| = Res(f,g,x) \neq 0$; luego existe una única solución de este sistema. Si usamos la regla de Cramer para obtenerla:

$$\forall i,\ \ c_i = \frac{1}{Res(f,g,x)} \text{ (determinante de cierta submatriz de la matriz de Sylvester)}$$

$$\forall j,\ \ d_j = \frac{1}{Res(f,g,x)} \text{ (determinante de cierta submatriz de la matriz de Sylvester)}$$

Luego podemos expresar:

$$\widetilde{A} = \frac{1}{Res(f,g,x)}A \quad \text{donde } A \in \mathbb{K}[x]$$

$$\widetilde{B} = \frac{1}{Res(f,g,x)}B \quad \text{donde } B \in \mathbb{K}[x].$$

Por lo que obtenemos

$$1 = \widetilde{A}f + \widetilde{B}g = \frac{A}{Res(f,g,x)}f + \frac{B}{Res(f,g,x)} \ \Rightarrow\ Af + Bg = Res(f,g,x).$$

□

Trabajemos ahora con polinomios en varias variables. Sean $f, g \in \mathbb{K}[x_1,\ldots,x_n]$ con grado positivo en x_1. Supongamos

$$f = a_0 + a_1x_1 + \cdots + a_\ell x_1^\ell \quad \text{con } a_\ell \neq 0,\ \ell > 0$$

$$g = b_0 + b_1x_1 + \cdots + b_m x_1^m \quad \text{con } b_m \neq 0,\ m > 0$$

donde $a_i, b_j \in \mathbb{K}[x_2,\ldots,x_n]$. Entonces definimos la matriz de Sylvester y la resultante de f y g respecto a x_1 análogamente a como lo hicimos antes, con la diferencia de que tanto las entradas de $Syl(f,g,x_1)$ como $Res(f,g,x_1)$ están en $\mathbb{K}[x_2,\ldots,x_n]$.

Veamos a continuación un resultado que nos aporta una primera relación de la resultante con la Teoría de Eliminación. Para ello necesitamos un lema previo.

Lema 4.1.3. *Sean $f, g \in \mathbb{K}[x_1,\ldots,x_n]$ ambos con grado positivo en x_1. Entonces f y g tienen un factor común en $\mathbb{K}[x_1,\ldots,x_n]$ con grado positivo en x_1 si y solo si tienen un factor común en $\mathbb{K}(x_2,\ldots,x_n)[x_1]$ (y por tanto con grado positivo en x_1).*

Demostración.

Si $\exists h \in \mathbb{K}[x_1,\ldots,x_n]$ con grado positivo en x_1 factor común de f y g, está claro que $h \in \mathbb{K}(x_2,\ldots,x_n)[x_1]$.

Veamos el recíproco. Supongamos que $\exists \widetilde{h} \in \mathbb{K}(x_2,\ldots,x_n)[x_1]$ con grado positivo en x_1 tal que $f = \widetilde{h}\widetilde{f_1}$, $g = \widetilde{h}\widetilde{g_1}$ con $\widetilde{f_1}, \widetilde{g_1} \in \mathbb{K}(x_2,\ldots,x_n)[x_1]$; pero $\widetilde{h}, \widetilde{f_1}, \widetilde{g_1}$ pueden tener coeficientes con denominadores que sean polinomios en $\mathbb{K}[x_2,\ldots,x_n]$. Sea $d \in \mathbb{K}[x_2,\ldots,x_n]$ un múltiplo común a los denominadores de $\widetilde{h}, \widetilde{f_1}, \widetilde{g_1}$ y sea $h = \widetilde{h}d$, $f_1 = \widetilde{f_1}d$, $g_1 = \widetilde{g_1}d \in \mathbb{K}[x_2,\ldots,x_n]$.

$$f = \widetilde{h}\widetilde{f_1} \ \Rightarrow\ d^2f = \widetilde{h}\widetilde{f_1}d^2 = hf_1 \in \mathbb{K}[x_1,\ldots,x_n]$$

$$g = \widetilde{h}\widetilde{g_1} \ \Rightarrow\ d^2g = \widetilde{h}\widetilde{g_1}d^2 = hg_1 \in \mathbb{K}[x_1,\ldots,x_n]$$

Como $\widetilde{h} = \frac{h}{d}$ tiene grado positivo en x_1, debe existir h_1 un factor irreducible de h con grado positivo en x_1. Luego $h_1 \,|\, hf_1 \ \Rightarrow\ h_1 \,|\, d^2f$. Como h_1 es irreducible y además h_1 depende de x_1 y d no, debe ser $h_1 \,|\, f$. Por el mismo motivo $h_1 \,|\, g$. Luego $h_1 \in \mathbb{K}[x_1,\ldots,x_n]$ con grado positivo en x_1 es un factor común de f y g. □

Proposición 4.1.4. *Sean $f, g \in \mathbb{K}[x_1, \ldots, x_n]$ ambos con grado positivo en x_1. Sea I el ideal $\langle f, g\rangle \subset \mathbb{K}[x_1, \ldots, x_n]$ y sea I_1 el primer ideal de eliminación de I. Entonces se verifica:*

i) $Res(f, g, x_1) \in I_1$.

ii) $Res(f, g, x_1) = 0$ *si y solo si* f *y* g *tienen un factor común en* $\mathbb{K}[x_1, \ldots, x_n]$ *con grado positivo en* x_1.

Demostración.

Tenemos

$$f = a_0 + a_1 x_1 + \cdots + a_\ell x_1^\ell \quad \text{con } a_\ell \neq 0,\ \ell > 0$$
$$g = b_0 + b_1 x_1 + \cdots + b_m x_1^m \quad \text{con } b_m \neq 0,\ m > 0$$

donde $a_i, b_j, Res(f, g, x_1) \in \mathbb{K}[x_2, \ldots, x_n]$. Sabemos entonces (ver *Proposición 4.1.2*) que existen $A, B \in \mathbb{K}[x_2, \ldots, x_n][x_1]$ tales que $Af + Bg = Res(f, g, x_1)$. En principio $A, B \in \mathbb{K}(x_2, \ldots, x_n)[x_1]$ pero están en $\mathbb{K}[x_2, \ldots, x_n][x_1]$ gracias a la prueba de la *Proposición 4.1.2*. Luego se tiene que $Res(f, g, x_1) \in \langle f, g\rangle \cap \mathbb{K}[x_2, \ldots, x_n] = I_1$.

Probemos *ii*). Hemos visto que f, g como polinomios en $\mathbb{K}[x_2, \ldots, x_n][x_1]$, luego en particular los coeficientes a_i, b_j de f y g viven en el cuerpo $\mathbb{K}(x_2, \ldots, x_n)$. Aplicamos entonces que $Res(f, g, x_1) = 0$ si y solo si f y g tienen un factor común en $\mathbb{K}(x_2, \ldots, x_n)[x_1]$ con grado positivo en x_1; que por el *Lema 4.1.3* es equivalente a que f y g tengan un factor común en $\mathbb{K}[x_1, \ldots, x_n]$ con grado positivo en x_1. □

Ahora veamos que también podemos usar las resultantes para ver cuándo se pueden extender las soluciones de $\mathcal{V}(I_1)$ a soluciones de $\mathcal{V}(I)$ cuando $I = \langle f, g\rangle$.

Proposición 4.1.5. *Sean $f, g \in \mathbb{C}[x_1, \ldots, x_n]$ ambos de grado positivo en x_1 expresados como*

$$f = a_0 + a_1 x_1 + \cdots + a_\ell x_1^\ell \quad \textit{con } a_\ell \neq 0,\ \ell > 0$$
$$g = b_0 + b_1 x_1 + \cdots + b_m x_1^m \quad \textit{con } b_m \neq 0,\ m > 0$$

donde $a_i, b_j \in \mathbb{C}[x_2, \ldots, x_n]$; y sea $I = \langle f, g\rangle$.

Sea $(c_2, \ldots, c_n) \in \mathbb{A}^{n-1}(\mathbb{C})$ tal que $Res(f, g, x_1)(c_2, \ldots, c_n) = 0$. Entonces se verifica i) ó ii):

i) $a_\ell(c_2, \ldots, c_n) = 0$ *ó* $b_m(c_2, \ldots, c_n) = 0$.

ii) Existe $c_1 \in \mathbb{C}$ *tal que* $(c_1, \ldots, c_n) \in \mathcal{V}(I)$.

Demostración.

Notemos $\mathbf{c} = (c_2, \ldots, c_n)$. Para hacer la prueba supondremos que $a_\ell(\mathbf{c}) \neq 0$ y $b_m(\mathbf{c}) \neq 0$, y tenemos que probar que se verifica *ii*).

Se tiene

$$f(x_1, \mathbf{c}) = f(x_1, c_2, \ldots, c_n)$$
$$g(x_1, \mathbf{c}) = g(x_1, c_2, \ldots, c_n).$$

Luego basta probar que $f(x_1, \mathbf{c}), g(x_1, \mathbf{c}) \in \mathbb{C}[x_1]$ tienen una raíz en común, o lo que es lo mismo, tienen un factor común con grado positivo en x_1 (hemos usado aquí que $\mathbb{C}$ es algebraicamente cerrado).

Usando la hipótesis de que $Res(f,g,x_1)$ se anula en $\mathbf{c}$, tenemos que

$$0 = Res(f,g,x_1)(\mathbf{c}) =$$

$$= det \begin{pmatrix} a_0(\mathbf{c}) & & & & b_0(\mathbf{c}) & & & \\ a_1(\mathbf{c}) & a_0(\mathbf{c}) & & & b_1(\mathbf{c}) & b_0(\mathbf{c}) & & \\ & a_1(\mathbf{c}) & \ddots & & & b_1(\mathbf{c}) & \ddots & \\ \vdots & & \ddots & a_0(\mathbf{c}) & \vdots & & \ddots & b_0(\mathbf{c}) \\ & \vdots & & a_1(\mathbf{c}) & & \vdots & & b_1(\mathbf{c}) \\ a_\ell(\mathbf{c}) & & & & b_m(\mathbf{c}) & & & \\ & a_\ell(\mathbf{c}) & & \vdots & & b_m(\mathbf{c}) & & \vdots \\ & & \ddots & & & & \ddots & \\ & & & a_\ell(\mathbf{c}) & & & & b_m(\mathbf{c}) \end{pmatrix} =$$

$$= Res(f(x_1,\mathbf{c}), g(x_1,\mathbf{c}), x_1).$$

Luego $f(x_1,\mathbf{c})$ y $g(x_1,\mathbf{c})$ tienen un factor común con grado positivo en x_1. □

Ya estamos en condiciones de probar el *Teorema de Extensión* en el caso de que el ideal del que partamos esté generado únicamente por dos polinomios usando para ello los conceptos y resultados vistos hasta ahora.

Teorema 4.1.6. (Teorema de Extensión para un ideal generado por dos polinomios.)
Sea $I = \langle f, g\rangle$ (que es un ideal de $\mathbb{C}[x_1,\ldots,x_n]$) y sea I_1 el primer ideal de eliminación de I. Escribimos

$$f = a_0 + a_1x_1 + \cdots + a_\ell x_1^\ell \quad \textit{con } a_\ell \neq 0,\ \ell \geq 0$$

$$g = b_0 + b_1x_1 + \cdots + b_m x_1^m \quad \textit{con } b_m \neq 0,\ m \geq 0$$

donde $a_i, b_j \in \mathbb{C}[x_2,\ldots,x_n]$.

Sea una solución parcial $(c_2,\ldots,c_n) \in \mathcal{V}(I_1)$. Si $(c_2,\ldots,c_n) \notin \mathcal{V}(a_\ell, b_m)$, entonces existe $c_1 \in \mathbb{C}$ tal que $(c_1,c_2,\ldots,c_n) \in \mathcal{V}(I)$.

Demostración.

Notemos $\mathbf{c} = (c_2,\ldots,c_n)$.

Por la *Proposición 4.1.4* tenemos que $Res(f,g,x_1) \in I_1$, y como $\mathbf{c} \in \mathcal{V}(I_1)$, se tiene que $Res(f,g,x_1)$ se anula en $\mathbf{c}$.

Sabemos que $\mathbf{c} \notin \mathcal{V}(a_\ell, b_m)$. Si $a_\ell(\mathbf{c}) \neq 0$ y $b_m(\mathbf{c}) \neq 0$, entones obtenemos directamente el resultado por la *Proposición 4.1.5*.

Falta estudiar el caso donde $a_\ell(\mathbf{c}) = 0$ y $b_m(\mathbf{c}) \neq 0$ (o $a_\ell(\mathbf{c}) \neq 0$ y $b_m(\mathbf{c}) = 0$, que se hace de manera análoga). Vamos a buscar entonces unos nuevos generadores del ideal I. Veamos a continuación que

$$I = \langle f, g\rangle = \langle f + x_1^N g, g\rangle \quad \forall N \in \mathbb{N}. \tag{4.2}$$

$\boxed{\subset}$ Si $h \in \langle f, g\rangle$ entonces

$$h = h_1 f + h_2 g = h_1(f + x_1^N g) - h_1 x_1^N g + h_2 g \in \langle f + x_1^N g, g\rangle.$$

$\boxed{\supset}$ Si $h \in \langle f + x_1^N g, g\rangle$ entonces

$$h = h_1(f + x_1^N g) + h_2 g = h_1 f + (h_1 x_1^N + h_2)g \in \langle f, g\rangle.$$

Elegimos entonces $N \in \mathbb{N}$ lo suficientemente grande como para que $m + N > \ell$. Luego ahora si escribimos los polinomios $f + x_1^N g$ y g como polinomios en $\mathbb{C}[x_2, \ldots, x_n][x_1]$, sus coeficientes líderes no se anulan en $\mathbf{c}$ ya que $b_m(\mathbf{c}) \neq 0$. Por lo que si volvemos a aplicar la *Proposición 4.1.5* obtenemos que existe $c_1 \in \mathbb{C}$ tal que $(c_1, c_2, \ldots, c_n) \in \mathcal{V}(I)$. □

Nuestro objetivo es acabar este capítulo con la demostración del *Teorema de Extensión* para un ideal $I = \langle f_1, \ldots, f_s\rangle \subset \mathbb{C}[x_1, \ldots, x_n]$; pero para ello tenemos que extender el concepto de resultante para un número arbitrario de polinomios.

Supongamos entonces que tenemos $f_1, \ldots, f_s \in \mathbb{C}[x_1, \ldots, x_n]$. Introducimos $s - 1$ variables nuevas $u_2, \ldots, u_s$ y consideramos el polinomio

$$u_2 f_2 + \cdots + u_s f_s \in \mathbb{C}[u_2, \ldots, u_s, x_1, \ldots, x_n].$$

Además debemos ver f_1 como un polinomio en $\mathbb{C}[u_2, \ldots, u_s, x_1, \ldots, x_n]$.

Sabemos que $Res(f_1, u_2 f_2 + \cdots + u_s f_s, x_1) \in \mathbb{C}[u_2, \ldots, u_s, x_2, \ldots, x_n]$, lo expresamos entonces de la siguiente forma:

$$Res(f_1, u_2 f_2 + \cdots + u_s f_s, x_1) = \sum_{\alpha} h_\alpha(x_2, \ldots, x_n)\mathbf{u}^\alpha$$

donde $\mathbf{u}^\alpha = u_2^{\alpha_2} \cdots u_s^{\alpha_s}$ y $h_\alpha \in \mathbb{C}[x_2, \ldots, x_n]$.

Llamaremos a los polinomios h_α las *resultantes generalizadas* de $f_1, \ldots, f_s$ respecto de x_1.

Nota. *Las resultantes generalizadas de un conjunto de polinomios depende de qué polinomio del conjunto consideremos en primer lugar (en nuestro caso f_1).*

Teorema 4.1.7. (Teorema de Extensión)
Sea $I = \langle f_1, \ldots, f_s\rangle$ un ideal de $\mathbb{C}[x_1, \ldots, x_n]$ y sea I_1 el primer ideal de eliminación de I. Para cada $i \in \{1, \ldots, s\}$ escribimos f_i como

$$f_i = g_i(x_2, \ldots, x_n)x_1^{N_i} + (\textit{términos con grado } < N_i \textit{ en } x_1)$$

donde $N_i \geq 0$ y $g_i \in \mathbb{C}[x_2, \ldots, x_n]$ no nulo.

Sea una solución parcial $(a_2, \ldots, a_n) \in \mathcal{V}(I_1)$. Si $(a_2, \ldots, a_n) \notin \mathcal{V}(g_1, \ldots, g_s)$, entonces existe $a_1 \in \mathbb{C}$ tal que $(a_1, \ldots, a_n) \in \mathcal{V}(I)$.

Demostración.

En la demostración suponemos que $s \geq 3$ pues en el caso de que $s = 1$ es trivial y $s = 2$ ya está probado en el *Teorema 4.1.6*.

Notemos $\mathbf{a} = (a_2, \ldots, a_n) \in \mathcal{V}(I_1)$ y supongamos $\mathbf{a} \notin \mathcal{V}(g_1, \ldots, g_s)$. Para cada $i = 1, \ldots, s$ consideremos el polinomio $f_i(x_1, \mathbf{a})$. En realidad queremos probar que existe una raíz en $\mathbb{C}$ común a los polinomios $f_1(x_1, \mathbf{a}), \ldots, f_s(x_1, \mathbf{a})$.

Como $\mathbf{a} \notin \mathcal{V}(g_1, \ldots, g_s)$, reordenando los polinomios $f_1, \ldots, f_s$ podemos suponer que $g_1(\mathbf{a}) \neq 0$.

Consideremos las resultantes generalizadas de $f_1, \ldots, f_s$,

$$h = Res(f_1, u_2 f_2 + \cdots + u_s f_s, x_1) = \sum_{\alpha} h_\alpha \mathbf{u}^\alpha.$$

Vamos a probar en primer lugar que $h_\alpha \in I_1$ para todo α.

Aplicando la *Proposición 4.1.2* se tiene que existen $A, B \in \mathbb{C}[u_2, \ldots, u_s, x_1, \ldots, x_n]$ tales que

$$Af_1 + B(u_2 f_2 + \cdots + u_s f_s) = Res(f_1, u_2 f_2 + \cdots + u_s f_s, x_1) = \sum_{\alpha} h_\alpha \mathbf{u}^\alpha \in \mathbb{C}[u_2, \ldots, u_s, x_2, \ldots, x_n].$$

Expresamos A y B de la siguiente forma:

$$A = \sum_{\alpha} A_\alpha(x_1, \ldots, x_n)\mathbf{u}^\alpha$$

$$B = \sum_{\beta} B_\beta(x_1, \ldots, x_n)\mathbf{u}^\beta.$$

Notemos también que para todo $i \in \{2, \ldots, s\}$ se tiene que $u_i = \mathbf{u}^{e_i}$ donde $e_i = (0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{N}^{s-1}$ donde el elemento no nulo está en la posición $i-1$. Ahora tenemos lo siguiente:

$$Af_1 + B(u_2 f_2 + \cdots + u_s f_s) = Res(f_1, u_2 f_2 + \cdots + u_s f_s, x_1)$$

$$\left(\sum_{\alpha} A_\alpha \mathbf{u}^\alpha\right) f_1 + \left(\sum_{\beta} B_\beta \mathbf{u}^\beta\right)(\mathbf{u}^{e_2} f_2 + \cdots + \mathbf{u}^{e_s} f_s) = \sum_{\alpha} h_\alpha \mathbf{u}^\alpha$$

$$\sum_{\alpha} h_\alpha \mathbf{u}^\alpha = \left(\sum_{\alpha} A_\alpha \mathbf{u}^\alpha\right) f_1 + \left(\sum_{\beta} B_\beta \mathbf{u}^\beta\right)\left(\sum_{i=2}^{s} \mathbf{u}^{e_i} f_i\right) =$$

$$= \sum_{\alpha} (A_\alpha f_1)\mathbf{u}^\alpha + \sum_{\beta} \sum_{i=2}^{s} (B_\beta f_i)\mathbf{u}^{\beta + e_i} =$$

$$= \sum_{\alpha} (A_\alpha f_1)\mathbf{u}^\alpha + \sum_{\alpha} \left(\sum_{\substack{2 \le i \le s \\ \beta + e_i = \alpha}} B_\beta f_i\right)\mathbf{u}^\alpha =$$

$$\sum_{\alpha} \left(A_\alpha f_1 + \sum_{\substack{2 \le i \le s \\ \beta + e_i = \alpha}} B_\beta f_i\right)\mathbf{u}^\alpha.$$

Igualando los términos obtenemos que para todo α es

$$h_\alpha = A_\alpha f_1 + \sum_{\substack{2 \le i \le s \\ \beta + e_i = \alpha}} B_\beta f_i.$$

Luego se tiene entonces que $h_\alpha \in I$; y por tanto $h_\alpha \in I \cap \mathbb{C}[x_2, \ldots, x_n] = I_1$.

Tenemos entonces que $h_\alpha \in I_1$ para todo α y de aquí se deduce que $h_\alpha(\mathbf{a}) = 0$ ya que $\mathbf{a} \in \mathcal{V}(I_1)$. Recordemos que

$$h = \sum_{\alpha} h_\alpha(x_2, \ldots, x_n)\mathbf{u}^\alpha \in \mathbb{C}[u_2, \ldots, u_s, x_2, \ldots, x_n]$$

por lo que tenemos que $h(u_2, \ldots, u_s, a_2, \ldots, a_n) \in \mathbb{C}[u_2, \ldots, u_s]$ es el polinomio nulo.

Supongamos ahora que se da

$$g_2(\mathbf{a}) \neq 0 \text{ y } N_2 > N_i \ \forall i = 3, \ldots, s. \tag{4.3}$$

En este caso se tiene que

$$0 = h(u_2, \ldots, u_s, \mathbf{a}) = Res(f_1, u_2 f_2 + \cdots + u_s f_s, x_1)(u_2, \ldots, u_s, \mathbf{a}) =$$

$$= Res(f_1(x_1, \mathbf{a}), u_2 f_2(x_1, \mathbf{a}) + \cdots + u_s f_s(x_1, \mathbf{a}), x_1)$$

donde $f_1(x_1, \mathbf{a}), u_2 f_2(x_1, \mathbf{a}) + \cdots + u_s f_s(x_1, \mathbf{a}) \in \mathbb{C}[u_2, \ldots, u_s, x_1]$. Luego si su resultante se anula quiere decir por la *Proposición 4.1.4* que deben tener un factor común $F \in \mathbb{C}[u_2, \ldots, u_s, x_1]$ de grado positivo en x_1. Pero como $F \mid f_1(x_1, \mathbf{a})$ debe ser $F \in \mathbb{C}[x_1]$; y en ese caso como $F \mid u_2 f_2(x_1, \mathbf{a}) + \cdots + u_s f_s(x_1, \mathbf{a})$ se tiene que $F \mid f_i(x_1, \mathbf{a})$ para todo $i = 1, \ldots, s$. Por esta razón si todos los polinomios $f_i(x_1, \mathbf{a})$ comparten un factor común $F \in \mathbb{C}[x_1]$ de grado positivo, debe existir una raíz $a_1 \in \mathbb{C}$ común a todos los $f_i(x_1, \mathbf{a})$ y por tanto el resultado queda demostrado si se da (4.3).

Veamos ahora el caso donde no se verifique (4.3). Recordemos que $g_1(\mathbf{a}) \neq 0$, luego se verifica la siguiente igualdad:

$$I = \langle f_1, f_2 + x_1^N f_1, f_3, \ldots, f_s \rangle \ \forall N \in \mathbb{N}.$$

La demostración de esta igualdad es análoga a la que se hizo en el Teorema de Extensión para dos polinomios (4.2). Tomamos entonces N lo suficientemente grande para que el coeficiente líder de $f_2 + x_1^N f_1$ sea g_1, que no se anula en $\mathbf{a}$, y para que el grado de $f_2 + x_1^N f_1$ en x_1 sea mayor que N_i para todo $i = 3, \ldots, s$. De esta manera ya se cumple (4.3) y por lo tanto existe una raíz común $a_1 \in \mathbb{C}[x_1]$ a los polinomios $f_1(x_1, \mathbf{a}), f_2(x_1, \mathbf{a}) + x_1^N f_1(x_1, \mathbf{a}), f_3(x_1, \mathbf{a}), \ldots, f_s(x_1, \mathbf{a})$. Además si a_1 es raíz de $(f_2 + x_1^N f_1)(x_1, \mathbf{a})$ y de $f_1(x_1, \mathbf{a})$ también debe ser raíz de $f_2(x_1, \mathbf{a})$ y hemos acabado. □

Capítulo 5

Programación Lineal Entera

En el campo de la Programación Matemática merecen gran interés los problemas de Programación Lineal Entera por su aplicación práctica a muchas otras ciencias. En este capítulo nos centraremos en estudiar y presentar una aplicación de las bases de Gröbner a la resolución de problemas de Programación Lineal Entera.

5.1. Resolución de un problema PLE

Vamos a introducir a continuación el contexto donde trabajaremos. Consideramos pues un problema de Programación Lineal Entera puro, es decir, donde todas variables que aparezcan representen números enteros.

$$\begin{array}{rl} \text{mín} \,/\, \text{máx} & \ell(\mathbf{A}) \\ s.a.: & a_{11}A_1 + a_{12}A_2 + \cdots + a_{1n}A_n \leq b_1 \\ & a_{21}A_1 + a_{22}A_2 + \cdots + a_{2n}A_n \leq b_2 \\ & \vdots \quad\quad \vdots \quad\quad \ddots \quad \vdots \quad\quad \vdots \\ & a_{m1}A_1 + a_{m2}A_2 + \cdots + a_{mn}A_n \leq b_m \\ & \mathbf{A} = (A_1, \ldots, A_n) \in \mathbb{Z}^n \end{array} \tag{5.1}$$

Donde ℓ es una función lineal con coeficientes números reales, $\mathbf{A} = (A_1, \ldots, A_n)$ son las variables y los coeficientes a_{ij}, $b_i \in \mathbb{Z}$. Nos restringiremos al caso donde las variables representen números enteros no negativos, es decir, $\mathbf{A} \in \mathbb{N}^n$.

Observemos también que podemos restringirnos al caso de minimizar funciones, pues $\text{máx}\, \ell(\mathbf{A}) = -\,\text{mín}\,(-\ell(\mathbf{A}))$. Además basta considerar restricciones del tipo $\leq$ pues

$$a_{i1}A_1 + a_{i2}A_2 + \cdots + a_{in}A_n \geq b_i$$

es equivalente a

$$-a_{i1}A_1 - a_{i2}A_2 - \cdots - a_{in}A_n \leq -b_i.$$

De esta forma los problemas que consideraremos pueden ser expresados como:

$$\begin{array}{rl} \text{mín} & c_1 A_1 + c_2 A_2 + \cdots + c_n A_n \\ s.a.: & a_{11}A_1 + a_{12}A_2 + \cdots + a_{1n}A_n \leq b_1 \\ & a_{21}A_1 + a_{22}A_2 + \cdots + a_{2n}A_n \leq b_2 \\ & \vdots \quad\quad \vdots \quad\quad \ddots \quad \vdots \quad\quad \vdots \\ & a_{m1}A_1 + a_{m2}A_2 + \cdots + a_{mn}A_n \leq b_m \\ & \mathbf{A} = (A_1, \ldots, A_n) \in \mathbb{N}^n \end{array} \tag{5.2}$$

donde $a_{ij}, b_i \in \mathbb{Z}$ y $c_j \in \mathbb{R}$.

Llamaremos región factible del problema (5.2) a todos los puntos $(A_1, \ldots, A_n) \in \mathbb{N}^n$ que verifican las restricciones, y región factible relajada del problema (5.2) a los puntos $(A_1, \ldots, A_n) \in \mathbb{R}^n$ que verifican todas las desigualdades.

Por último, para escribir un problema en forma estándar solo consideraremos restricciones de igualdad. Para ellos introducimos una variable de holgura por cada restricción. Si tenemos

$$a_{i1}A_1 + a_{i2}A_2 + \cdots + a_{in}A_n \leq b_i$$

consideramos una nueva variable β_i de modo que tengamos

$$a_{i1}A_1 + a_{i2}A_2 + \cdots + a_{in}A_n + \beta_i = b_i$$

con $\beta_i = b_i - (a_{i1}A_1 + a_{i2}A_2 + \cdots + a_{in}A_n)$, por lo que tendremos que $\beta_i \in \mathbb{N}$. Esta variable de holgura aparece en la función objetivo con coeficiente nulo. Luego todo problema de Programación Lineal Entera en $\mathbb{N}^n$ puede ser escrito en forma estándar como:

$$\begin{array}{rl} \text{mín} & c_1 A_1 + c_2 A_2 + \cdots + c_n A_n \\ s.a.: & a_{11}A_1 + a_{12}A_2 + \cdots + a_{1n}A_n = b_1 \\ & a_{21}A_1 + a_{22}A_2 + \cdots + a_{2n}A_n = b_2 \\ & \vdots \quad\quad \vdots \quad\quad \ddots \quad \vdots \quad\quad \vdots \\ & a_{m1}A_1 + a_{m2}A_2 + \cdots + a_{mn}A_n = b_m \\ & \mathbf{A} = (A_1, \ldots, A_n) \in \mathbb{N}^n \end{array} \tag{5.3}$$

donde $a_{ij}, b_i \in \mathbb{Z}$ y $c_j \in \mathbb{R}$.

Notemos que podemos suponer que todas las varibles $A_1, \ldots, A_n$ aparecen en al menos una de las desigualdades, es decir, que para todo j no son los $a_{1j}, \ldots, a_{mj}$ todos nulos simultáneamente. Pues en caso contrario la variable A_j no estaría sujeta a ninguna restricción salvo a la de integridad; entonces si el c_j correspondiente es positivo basta tomar $A_j = 0$ para minimizar la función y resolver el problema con una variable menos; y si $c_j < 0$ el problema no tendrá, en general, solución pues la función $\ell(\mathbf{A})$ no tiene mínimo cuando $\mathbf{A}$ pertenece a la región factible.

A partir de ahora nos centraremos en convertir este problema de Programación Lineal Entera estándar (5.3) en un problema sobre ecuaciones polinómicas que trataremos de resolver usando bases de Gröbner.

En primer lugar vamos a considerar el caso donde todos los coeficientes $a_{ij}, b_i \in \mathbb{N}$. Posteriormente extenderemos el proceso al caso general.

Para empezar vamos a introducir una variable z_i por cada restricción de (5.3) de modo que ésta se escriba de forma equivalente como:

$$z_i^{a_{i1}A_1 + a_{i2}A_2 + \cdots + a_{in}A_n} = z_i^{b_i} \quad \text{con } i \in \{1, \ldots, m\}$$

que podemos expresar abreviadamente como

$$\prod_{j=1}^{n} z_i^{a_{ij}A_j} = z_i^{b_i} \quad \text{con } i \in \{1,\dots,m\}. \tag{5.4}$$

Estas igualdades deben verificarse para cualquier valor de cada z_i si $(A_1,\dots,A_n)$ está en la región factible (de hecho en la región factible relajada) del problema (5.3). Multiplicando entre sí las m expresiones de (5.4) correspondientes a cada restricción obtenemos:

$$\prod_{i=1}^{m}\left(\prod_{j=1}^{n} z_i^{a_{ij}A_j}\right) = \prod_{j=1}^{n}\left(\prod_{i=1}^{m} z_i^{a_{ij}}\right)^{A_j} = \prod_{i=1}^{m} z_i^{b_i} \tag{5.5}$$

De aquí podemos deducir el siguiente resultado.

Proposición 5.1.1. *Sea* $\mathbb{K}$ *un cuerpo y supongamos que estamos en las condiciones del problema* (5.3)*. Definimos el morfismo de* $\mathbb{K}$*-álgebras*

$$\varphi : \mathbb{K}[w_1,\dots,w_n] \longrightarrow \mathbb{K}[z_1,\dots,z_m]$$

por

$$\varphi(w_j) = \prod_{i=1}^{m} z_i^{a_{ij}}, \quad \textit{con } j \in \{1,\dots,n\},$$

$$\varphi(g(w_1,\dots,w_n)) = g(\varphi(w_1),\dots,\varphi(w_n)).$$

Entonces $(A_1,\dots,A_n)$ *pertenece a la región factible de* (5.3) *si y solo si*

$$\varphi(w_1^{A_1}\cdots w_n^{A_n}) = z_1^{b_1}\cdots z_m^{b_m}.$$

Demostración.

$(A_1,\dots,A_n)$ está en la región factible de (5.3) si y solo si

$$\prod_{j=1}^{n}\left(\prod_{i=1}^{m} z_i^{a_{ij}}\right)^{A_j} = \prod_{i=1}^{m} z_i^{b_i}$$

donde la igualdad se da en el anillo $\mathbb{K}[z_1,\dots,z_m]$. Pero como $\varphi(w_j) = \prod_{i=1}^{m} z_i^{a_{ij}} \in \mathbb{K}[z_1,\dots,z_m]$; esto último es equivalente a

$$\prod_{j=1}^{n} (\varphi(w_j))^{A_j} = \prod_{i=1}^{m} z_i^{b_i}$$

$$\varphi(w_1)^{A_1}\cdots\varphi(w_n)^{A_n} = z_1^{b_1}\cdots z_m^{b_m}$$

$$\varphi(w_1^{A_1}\cdots w_n^{A_n}) = z_1^{b_1}\cdots z_m^{b_m}$$

□

En general el morfismo φ no es sobreyectivo, veámoslo en el siguiente ejemplo.

Ejemplo 13. *Supongamos que tenemos el siguiente problema de Programación Lineal Entera.*

$$\begin{array}{rl} \text{mín} & cA \\ s.a.: & aA = b \\ & A \in \mathbb{N} \end{array}$$

donde $a, b \in \mathbb{N}$ *y* $c \in \mathbb{R}$. *Aquí solo estamos considerando una variable y una restricción, y estamos en las condiciones del problema* (5.3). *Definimos entonces el morfismo de* $\mathbb{K}$*-álgebras de la* Proposición 5.1.1*:*

$$\begin{array}{rcl} \varphi : \mathbb{K}[w] & \longrightarrow & \mathbb{K}[z] \\ w & \longmapsto & z^a \end{array}$$

Observemos que si $a \neq 1$ *el morfismo* φ *no es sobreyectivo.*

Veremos a continuación un resultado que nos dice cómo caracterizar a los polinomios de $\mathbb{K}[z_1, \ldots, z_m]$ que son imagen por φ de algún polinomio en $\mathbb{K}[w_1, \ldots, w_n]$. Observemos que la imagen de φ es la subálgebra $\mathbb{K}[f_1, \ldots, f_n]$ de $\mathbb{K}[z_1, \ldots, z_m]$ donde $f_j = \prod_{i=1}^{m} z_i^{a_{ij}} = \varphi(w_j)$.

Consideremos de manera análoga el morfismo de $\mathbb{K}$-álgebras

$$\psi : \mathbb{K}[z_1, \ldots, z_m, w_1, \ldots, w_n] \longrightarrow \mathbb{K}[z_1, \ldots, z_m]$$

donde

$$\begin{array}{c} \psi(z_i) = z_i \\ \psi(w_j) = \varphi(w_j) = f_j \end{array}$$

$$\psi(g(z_1, \ldots, z_m, w_1, \ldots, w_n)) = g(\psi(z_1), \ldots, \psi(z_m), \psi(w_1), \ldots, \psi(w_n)).$$

que extiende al morfismo φ, considerado para polinomios arbitrarios $f_1(\mathbf{z}), \ldots, f_n(\mathbf{z})$ en $\mathbb{K}[\mathbf{z}]$. Para simplificar denotaremos por $\mathbb{K}[\mathbf{z}, \mathbf{w}]$ a $\mathbb{K}[z_1, \ldots, z_m, w_1, \ldots, w_n]$ y por $\mathbb{K}[\mathbf{z}]$ a $\mathbb{K}[z_1, \ldots, z_m]$.

Vamos a introducir un lema que usaremos posteriormente.

Lema 5.1.2. *Sean polinomios arbitrarios* $f_1(\boldsymbol{z}), \ldots, f_n(\boldsymbol{z})$ *en* $\mathbb{K}[\boldsymbol{z}]$. *En las condiciones anteriores sea* $I = \langle f_1 - w_1, \ldots, f_n - w_n \rangle \subset \mathbb{K}[\boldsymbol{z}, \boldsymbol{w}]$. *Entonces se tiene que* $ker(\psi) = I$.

Demostración.

Consideremos por tanto el morfismo de $\mathbb{K}$-álgebras

$$\begin{array}{rcl} \psi : \mathbb{K}[z_1, \ldots, z_m, w_1, \ldots, w_n] & \longrightarrow & \mathbb{K}[z_1, \ldots, z_m] \\ z_i & \longmapsto & z_i \\ w_j & \longmapsto & f_j \end{array}$$

$\boxed{I \subseteq ker(\psi)}$ Sea $h \in I$, entonces podemos expresar

$$h = h_1(f_1 - w_1) + \ldots + h_n(f_n - w_n), \quad \text{con } h_j \in \mathbb{K}[\mathbf{z}, \mathbf{w}].$$

Luego

$$\psi(h) = \psi(h_1)\,[\psi(f_1) - \psi(w_1)] + \cdots + \psi(h_n)\,[\psi(f_n) - \psi(w_n)] = 0$$

ya que $\psi(f_j) = f_j$ pues $f_j \in \mathbb{K}[\mathbf{z}]$ y $\psi(w_j) = f_j$.

$\boxed{ker(\psi) \subseteq I}$ Sea $f \in \mathbb{K}[\mathbf{z}, \mathbf{w}]$ tal que $\psi(f) = 0$. Obtenemos

$$f(z_1, \ldots, z_m, w_1, \ldots, w_n) = \sum_{j=1}^{n} h_j(f_j(\mathbf{z}) - w_j) + R(\mathbf{z})$$

de la división de f entre $\{f_1 - w_1, \ldots, f_n - w_n\}$ tomando $w_1, \ldots, w_n$ como monomios líderes, y por tanto se tiene que el resto $R(\mathbf{z})$ es un polinomio que no depende de $\mathbf{w}$. Luego,

$$0 = \psi(f) = \sum_{j=1}^{n} \psi(h_j)\,[\psi(f_j(\mathbf{z})) - \psi(w_j)] + \psi(R(\mathbf{z})).$$

Y debe ser entonces $\psi(R(\mathbf{z})) = 0$, pero como $\psi(R(\mathbf{z})) = R(\mathbf{z})$ tenemos que $R(\mathbf{z}) = 0$ y $f \in I$. □

Proposición 5.1.3. *Consideremos $f_1, \ldots, f_n \in \mathbb{K}[z_1, \ldots, z_m]$ polinomios arbitrarios y $\leq$ un orden monomial en $\mathbb{K}[\boldsymbol{z}, \boldsymbol{w}]$ con la Propiedad de Eliminación respecto del par $(\boldsymbol{z}, \boldsymbol{w})$ (es decir, donde todo monomio que contenga a algún z_i es mayor que aquellos monomios que solo involucren a los w_j). Sea G una base de Gröbner para el ideal*

$$I = \langle f_1 - w_1, \ldots, f_n - w_n \rangle \subset \mathbb{K}[z_1, \ldots, z_m, w_1, \ldots, w_n]$$

y sea $f \in \mathbb{K}[z_1, \ldots, z_m]$. Si denotamos $g = \overline{f}^G$ entonces,

i) $f \in \mathbb{K}[f_1, \ldots, f_n]$ si y solo si $g = \overline{f}^G \in \mathbb{K}[w_1, \ldots, w_n]$.

ii) Si $f \in \mathbb{K}[f_1, \ldots, f_n]$, y por tanto, $g = \overline{f}^G \in \mathbb{K}[w_1, \ldots, w_n]$; entonces $f = g(f_1, \ldots, f_n)$ nos da una expresión de f como polinomio en $\mathbb{K}[f_1, \ldots, f_n]$.

iii) Si f y f_j son monomios para $j = 1, \ldots, n$, y $f \in \mathbb{K}[f_1, \ldots, f_n]$ entonces g también es un monomio de $\mathbb{K}[w_1, \ldots, w_n]$.

Demostración.

Supongamos $G = \{g_1, \ldots, g_t\}$. Al dividir f entre G obtenemos la expresión

$$f = h_1 g_1 + \cdots + h_t g_t + g$$

donde $h_1, \ldots, h_t, g \in \mathbb{K}[\mathbf{z}, \mathbf{w}]$ verificando las propiedades del *Teorema 2.2.3* y $g = \overline{f}^G$. Como G es una base de Gröbner de I tenemos también que

$$I = \langle f_1 - w_1, \ldots, f_n - w_n \rangle = \langle g_1, \ldots, g_t \rangle.$$

Luego como $h_1 g_1 + \cdots + h_t g_t \in I$, podemos escribir

$$f = h'_1(f_1 - w_1) + \cdots + h'_n(f_n - w_n) + g$$

donde $h'_j \in \mathbb{K}[\mathbf{z}, \mathbf{w}]$.

i) Vamos a considerar el morfismo de $\mathbb{K}$-álgebras

$$\begin{array}{ccc} \psi : \mathbb{K}[z_1,\ldots,z_m,w_1,\ldots,w_n] & \longrightarrow & \mathbb{K}[z_1,\ldots,z_m] \\ z_i & \longmapsto & z_i \\ w_j & \longmapsto & f_j \end{array}$$

$\boxed{\Leftarrow}$ Supongamos que $g = \overline{f}^G \in \mathbb{K}[w_1,\ldots,w_n]$.

$$\psi(f) = \psi(h'_1)\,[\psi(f_1) - \psi(w_1)] + \cdots + \psi(h'_n)\,[\psi(f_n) - \psi(w_n)] + \psi(g) =$$

$$= \psi(g) = g(f_1,\ldots,f_n)$$

pero como además tenemos que $\psi(f) = f$ pues $f \in \mathbb{K}[z_1,\ldots,z_m]$ es entonces $f = g(f_1,\ldots,f_n)$ y por tanto se tiene que $f \in \mathbb{K}[f_1,\ldots,f_n]$.

$\boxed{\Rightarrow}$ Supongamos $f = g'(f_1,\ldots,f_n) \in \mathbb{K}[f_1,\ldots,f_n]$ para cierto polinomio $g'(w_1,\ldots,w_n) \in \mathbb{K}[w_1,\ldots,w_n]$. Veamos que $g = \overline{f}^G \in \mathbb{K}[w_1,\ldots,w_n]$. En $\mathbb{K}[z_1,\ldots,z_m,w_1,\ldots,w_n]$ un monomio en $f_1,\ldots,f_n \in \mathbb{K}[z_1,\ldots,z_m]$ puede ser escrito como

$$f_1^{A_1}\cdots f_n^{A_n} = (w_1 + (f_1 - w_1))^{A_1}\cdots(w_n + (f_n - w_n))^{A_n} =$$

$$= w_1^{A_1}\cdots w_n^{A_n} + B_1(f_1 - w_1) + \cdots + B_n(f_n - w_n)$$

para algunos $B_1,\ldots,B_n \in \mathbb{K}[z_1,\ldots,z_m,w_1,\ldots,w_n]$. Por lo que

$$f = g'(f_1,\ldots,f_n) = g'(w_1,\ldots,w_n) + C_1(f_1 - w_1) + \cdots + C_n(f_n - w_n)$$

con $C_j \in \mathbb{K}[z_1,\ldots,z_m,w_1,\ldots,w_n]$. Tenemos entonces que

$$\overline{f}^G = \overline{g'(w_1,\ldots,w_n)}^G + \overline{\sum_{j=1}^{n} C_j(f_j - w_j)}^G .$$

Además es $\overline{\sum_{j=1}^{n} C_j(f_j - w_j)}^G = 0$ por ser G una base de Gröbner de I. Entonces, $\overline{f}^G = \overline{g'(w_1,\ldots,w_n)}^G \in \mathbb{K}[w_1,\ldots,w_n]$ porque G es una base de Gröbner de $I = \langle f_1 - w_1,\ldots,f_n - w_n\rangle$ respecto de un orden monomial con la *Propiedad de Eliminación* respecto del par $(\mathbf{z},\mathbf{w})$. En efecto $g'(w_1,\ldots,w_n)$ solo se reduce respecto de los $g_k \in G$ tales que $\mathrm{tl}(g_k) \in \mathbb{K}[\mathbf{w}]$. Pero por el orden de eliminación elegido los tales g_k debe estar en $\mathbb{K}[\mathbf{w}]$. Así, $g'(w_1,\ldots,w_n)$ solo se reduce respecto de los g_k que están en $\mathbb{K}[\mathbf{w}]$ y el resto de la división está por tanto en $\mathbb{K}[\mathbf{w}]$.

ii) Queda automáticamente probado por la primera parte de la demostración de *i)*, pues $f(\mathbf{z}) = g(f_1,\ldots,f_n)$.

iii) Si ahora f_j, f son monomios, entonces el ideal I está generado por polinomios que son diferencias de dos monomios o equivalentemente, generado por binomios. Cuando calculamos S-polinomios con polinomios de esta forma seguimos obteniendo binomios y al dividir binomios por binomios también obtendremos como restos polinomios que estarán formados como mucho por dos términos. Luego al aplicar el algoritmo de Buchberger para calcular una base de Gröbner G de I, se sigue que todo elemento de G también será o un monomio o un binomio. Al dividir f que es un monomio por una base de Gröbner de estas características, el resto que obtenemos de aplicar el algoritmo de división debe ser un monomio, ya que en cada paso restamos a un monomio un polinomio formado por dos monomios, donde uno de ellos se cancela y solo queda a lo más un monomio. Luego $g = \overline{f}^G$ es un

monomio de $\mathbb{K}[w_1, \ldots, w_n]$. En el caso de que la base de Gröbner que tengamos no se haya obtenido aplicando el algoritmo de Buchberger no podemos asegurar que tenga esa forma, sin embargo el resto de dividir f por una base de Gröbner ya vimos que era el mismo, independientemente de la base de Gröbner que tengamos, si el orden monomial siempre es el mismo. Luego se verifica la tesis del apartado *iii*).

□

Vamos a recapitular resumidamente todo lo que hemos visto hasta ahora. Tenemos entonces por la *Proposición 5.1.1* que $(A_1, \ldots, A_n)$ es factible si y solo si

$$\varphi(w_1^{A_1} \cdots w_n^{A_n}) = z_1^{b_1} \cdots z_n^{b_n}$$

donde $\varphi(w_j) = \prod_{i=1}^m z_i^{a_{ij}}$ con $j = 1, \ldots, n$. Si ahora en este caso particular aplicamos la *Proposición 5.1.3* tenemos lo siguiente:

$$\text{Para } j = 1, \ldots, n, \;\; f_j = \prod_{i=1}^m z_i^{a_{ij}} \text{ es un monomio en } \mathbb{K}[z_1, \ldots, z_m],$$

$$I = \langle f_1 - w_1, \ldots, f_n - w_n \rangle \subset \mathbb{K}[z_1, \ldots, z_m, w_1, \ldots, w_n]$$

y G una base de Gröbner de I respecto a un orden monomial con la *Propiedad de Eliminación*. Sea $f = z_1^{b_1} \cdots z_m^{b_m} \in \mathbb{K}[z_1, \ldots, z_m]$. Luego tenemos como resultado que si $f \in \mathbb{K}[f_1, \ldots, f_n]$ entonces $g = \overline{f}^G \in \mathbb{K}[w_1, \ldots, w_n]$. Además $f = g(f_1, \ldots, f_n)$ nos da la expresión de f como polinomio en f_j y g es un monomio. De aquí obtenemos que si $z_1^{b_1} \cdots z_m^{b_m}$ está en la imagen de φ, entonces debe ser imagen de algún monomio $w_1^{A_1} \cdots w_n^{A_n}$ por *iii*) de la *Proposición 5.1.3*.

A continuación vamos a definir los órdenes monomiales que nos permitirán trabajar satisfactoriamente en la resolución de problemas de Programación Lineal Entera usando bases de Gröbner.

Pongamos $\boldsymbol{x} = (\mathbf{z}, \mathbf{w}) = (z_1, \ldots, z_m, w_1, \ldots, w_n)$.

Definición 5.1.4. *Un orden monomial $\leq$ en $\mathbb{K}[\boldsymbol{z}, \boldsymbol{w}]$ se dice que está adaptado al problema de Programación Lineal Entera* (5.3) *si verifica:*

i) Propiedad de eliminación. *Para todo monomio $\boldsymbol{x}^\alpha$ que contenga a alguna de las variables z_i y para todo monomio $\boldsymbol{x}^\beta$ que solo contenga a las variables w_j debe ser $\boldsymbol{x}^\alpha > \boldsymbol{x}^\beta$.*

ii) Compatibilidad con ℓ. *Sean $\mathbf{A} = (A_1, \ldots, A_n) \in \mathbb{N}^n$ y $\mathbf{A'} = (A'_1, \ldots, A'_n) \in \mathbb{N}^n$. Si los monomios $\mathbf{w}^{\mathbf{A}} = w_1^{A_1} \cdots w_n^{A_n}$ y $\mathbf{w}^{\mathbf{A'}} = w_1^{A'_1} \cdots w_n^{A'_n}$ satisfacen que $\varphi(\mathbf{w}^{\mathbf{A}}) = \varphi(\mathbf{w}^{\mathbf{A'}})$ y $\ell(\mathbf{A}) > \ell(\mathbf{A'})$ debe ser $\mathbf{w}^{\mathbf{A}} > \mathbf{w}^{\mathbf{A'}}$.*

Veremos en la *Sección 5.3* que todo problema de Programación Lineal Entera (5.3) tiene al menos un orden monomial adaptado.

Teorema 5.1.5. *Consideremos un problema de Programación Lineal Entera en forma estándar como en* (5.3) *y sea $\leq$ un orden monomial adaptado para dicho problema. Supongamos que todos los coeficientes $a_{ij}, b_i \in \mathbb{N}$. Sea $f_j = \prod_{i=1}^m z_i^{a_{ij}}$ para $j = 1, \ldots, n$ y sea el ideal*

$$I = \langle f_1 - w_1, \ldots, f_n - w_n \rangle \subset \mathbb{K}[z_1, \ldots, z_m, w_1, \ldots, w_n]$$

con G una base de Gröbner de I con respecto al orden monomial $\leq$.

Entonces si $f = z_1^{b_1} \cdots z_m^{b_m} \in \mathbb{K}[f_1, \ldots, f_n]$, el resto $\overline{f}^G \in \mathbb{K}[w_1, \ldots, w_n]$ (que es un monomio en $\mathbb{K}[w_1, \ldots, w_n]$) nos proporciona una solución de (5.3).

En el caso en el que la solución de (5.3) no sea única nos devolverá un punto óptimo factible entre todos los que haya.

Demostración.

Está claro que si $f = z_1^{b_1} \cdots z_m^{b_m} \in \mathbb{K}[f_1, \ldots, f_n]$ entonces por lo visto anteriormente existe $(A_1, \ldots, A_n) \in \mathbb{N}^n$ tal que

$$f = z_1^{b_1} \cdots z_m^{b_m} = \varphi(w_1^{A_1} \cdots w_n^{A_n}) = \varphi(w_1)^{A_1} \cdots \varphi(w_n)^{A_n} = f_1^{A_1} \cdots f_n^{A_n} \in \mathbb{K}[f_1, \ldots, f_n],$$

y por tanto $g = \overline{f}^G = w_1^{A_1} \cdots w_n^{A_n} \in \mathbb{K}[w_1, \ldots, w_n]$ nos da un punto $(A_1, \ldots, A_n)$ factible por la *Proposición 5.1.1 y 5.1.3*.

Lo que tenemos que probar ahora es que ese punto $(A_1, \ldots, A_n)$ es óptimo.

Reducción al Absurdo. Supongamos que $\overline{f}^G = g = w_1^{A_1} \cdots w_n^{A_n} = \mathbf{w}^{\mathbf{A}}$ con $\varphi(\mathbf{w}^{\mathbf{A}}) = f$, pero que $\mathbf{A} = (A_1, \ldots, A_n)$ no es óptimo para (5.3). Entonces debe existir $\mathbf{A'} = (A'_1, \ldots, A'_n) \neq \mathbf{A}$ tal que $\varphi(\mathbf{w}^{\mathbf{A'}}) = f$ y $\ell(\mathbf{A}) > \ell(\mathbf{A'})$. Observemos que como el orden monomial es adaptado, se verifica que $\mathbf{w}^{\mathbf{A}} > \mathbf{w}^{\mathbf{A'}}$.

Sea entonces $h = \mathbf{w}^{\mathbf{A}} - \mathbf{w}^{\mathbf{A'}} \in \mathbb{K}[w_1, \ldots, w_n]$ y por tanto $\varphi(h) = \psi(h) = \varphi(\mathbf{w}^{\mathbf{A}} - \mathbf{w}^{\mathbf{A'}}) = \varphi(\mathbf{w}^{\mathbf{A}}) - \varphi(\mathbf{w}^{\mathbf{A'}}) = f - f = 0$. Luego tenemos que $h \in I$ por el *Lema 5.1.2*. Entonces $\mathrm{tl}(h) = \mathbf{w}^{\mathbf{A}} \in \mathrm{TL}(I)$ (ya que $\mathbf{w}^{\mathbf{A}} > \mathbf{w}^{\mathbf{A'}}$). Por otra parte $\mathrm{tl}(h) = \mathbf{w}^{\mathbf{A}} = \overline{f}^G$, y por tanto $\mathrm{tl}(h)$ no está en $\mathrm{TL}(I)$ lo cual es una contradicción.

Así que $\mathbf{A}$ es una solución del problema (5.3). □

Este último resultado nos proporciona un algoritmo para resolver problemas de Programación Lineal Entera como (5.3) donde todos los coeficientes $a_{ij}, b_i \in \mathbb{N}$.

Algoritmo 3 Algoritmo de resolución de problemas de Programación Lineal Entera usando bases de Gröbner.

Entrada: a_{ij}, b_i de (5.3) y un orden monomial $\leq$ adaptado a dicho problema.
Salida: Una solución $\mathbf{A} = (A_1, \ldots, A_n)$ de (5.3) si existe.

Inicialización: Definimos

$$f_j := \prod_{i=1}^{m} z_i^{a_{ij}} \quad \text{con } j = 1, \ldots, n$$

$$I := \langle f_1 - w_1, \ldots, f_n - w_n \rangle$$

$$f := \prod_{i=1}^{m} z_i^{b_i}$$

Calculamos G una base de Gröbner de I respecto a $\leq$ y hacemos $g := \overline{f}^G$.

1: **Si** $g \in \mathbb{K}[w_1, \ldots, w_n]$
Devolver: el vector formado por los exponentes de g.

2: **Si** $g \notin \mathbb{K}[w_1, \ldots, w_n]$
Devolver: el problema (5.3) no tiene solución.

5.2. Resolución del Problema General

Vamos a dedicar esta sección a generalizar el algoritmo para los problemas de Programación Lineal Entera (5.3)

$$\begin{array}{rl} \text{mín} & c_1 A_1 + c_2 A_2 + \cdots + c_n A_n \\ s.a.: & a_{11}A_1 + a_{12}A_2 + \cdots + a_{1n}A_n = b_1 \\ & a_{21}A_1 + a_{22}A_2 + \cdots + a_{2n}A_n = b_2 \\ & \vdots \quad\quad \vdots \quad\quad \ddots \quad \vdots \quad\quad \vdots \\ & a_{m1}A_1 + a_{m2}A_2 + \cdots + a_{mn}A_n = b_m \\ & \mathbf{A} = (A_1, \ldots, A_n) \in \mathbb{N}^n \end{array}$$

donde $a_{ij}, b_i \in \mathbb{Z}$ y $c_j \in \mathbb{R}$.

Hasta ahora hemos conseguido resolver el problema usando bases de Gröbner en el caso de que los coeficientes $a_{ij}, b_i \in \mathbb{N}$. Pero ahora vamos a ver como generalizamos estos resultados para $a_{ij}, b_i \in \mathbb{Z}$.

Como en el caso anterior volvemos a introducir una variable z_i por cada restricción de nuestro problema y de nuevo obtenemos las siguientes igualdades:

$$z_i^{a_{i1}A_1 + a_{i2}A_2 + \cdots + a_{in}A_n} = \prod_{j=1}^{n} z_i^{a_{ij}A_j} = z_i^{b_i} \quad \text{con } i \in \{1, \ldots, m\}.$$

Multiplicando entre sí las m expresiones obtenemos (5.5):

$$\prod_{i=1}^{m} \left(\prod_{j=1}^{n} z_i^{a_{ij}A_j} \right) = \prod_{j=1}^{n} \left(\prod_{i=1}^{m} z_i^{a_{ij}} \right)^{A_j} = \prod_{i=1}^{m} z_i^{b_i}$$

El problema ahora radica en que estas expresiones no tienen porqué ser polinomios del anillo $\mathbb{K}[z_1, \ldots, z_m]$.

Lo que vamos a considerar entonces es una nueva variable t, el anillo de polinomios $\mathbb{K}[z_1, \ldots, z_m, t]$ y el ideal $J = \langle t z_1 \cdots z_m - 1 \rangle$ de este anillo.

A continuación vamos a adaptar todos los resultados de la sección anterior donde trabajamos con el anillo de polinomios $\mathbb{K}[z_1, \ldots, z_m]$ a este caso, donde trabajaremos con el anillo cociente:

$$\frac{\mathbb{K}[z_1, \ldots, z_m, t]}{\langle t z_1 \cdots z_m - 1 \rangle}. \tag{5.6}$$

Podemos entender t como el producto de las variables z_i inversas, es decir, $t = z_1^{-1} \cdots z_m^{-1}$.

Ahora siempre podremos escribir

$$(a_{1j}, \ldots, a_{mj}) = (a'_{1j}, \ldots, a'_{mj}) + \alpha_j(-1, \ldots, -1) \quad \text{para } j = 1, \ldots, n$$

donde $a'_{ij}, \alpha_j \in \mathbb{N}$ para todo j y para todo i. De manera que obtenemos la siguiente expresión en el anillo cociente (5.6) para $j = 1, \ldots, n$:

$$\prod_{i=1}^{m} z_i^{a_{ij}} = \prod_{i=1}^{m} z_i^{a'_{ij} - \alpha_j} = \prod_{i=1}^{m} z_i^{a'_{ij}} z_i^{-\alpha_j} =$$

$$= \left(\prod_{i=1}^{m} z_i^{-\alpha_j}\right)\left(\prod_{i=1}^{m} z_i^{a'_{ij}}\right) = (z_1 \cdots z_m)^{-\alpha_j}\left(\prod_{i=1}^{m} z_i^{a'_{ij}}\right) = t^{\alpha_j}\prod_{i=1}^{m} z_i^{a'_{ij}}.$$

Análogamente siempre podremos expresar

$$(b_1, \ldots, b_m) = (b'_1, \ldots, b'_m) + \beta(-1, \ldots, -1)$$

donde $b'_i, \beta \in \mathbb{N}$ para todo i. Y de nuevo en el anillo cociente (5.6) obtenemos la expresión:

$$\prod_{i=1}^{m} z_i^{b_i} = \prod_{i=1}^{m} z_i^{b'_i-\beta} = \prod_{i=1}^{m} z_i^{b'_i} z_i^{-\beta} =$$

$$= \left(\prod_{i=1}^{m} z_i^{-\beta}\right)\left(\prod_{i=1}^{m} z_i^{b'_i}\right) = (z_1 \cdots z_m)^{-\beta}\left(\prod_{i=1}^{m} z_i^{b'_i}\right) = t^{\beta}\prod_{i=1}^{m} z_i^{b'_i}.$$

Viendo la ecuación (5.5) en el anillo cociente (5.6) y aplicando los desarrollos que acabamos de ver obtenemos lo siguiente:

$$\prod_{j=1}^{n}\left(t^{\alpha_j}\prod_{i=1}^{m} z_i^{a'_{ij}}\right)^{A_j} = t^{\beta}\prod_{i=1}^{m} z_i^{b'_i}.$$

Hay que entender las dos últimas igualdades en el anillo cociente (5.6), es decir, módulo la relación $tz_1 \cdots z_m - 1 = 0$.

Vamos a adaptar la *Proposicion 5.1.1* a este caso.

Proposición 5.2.1. *Sea* $\mathbb{K}$ *un cuerpo y supongamos que estamos en las condiciones que hemos visto hasta ahora en esta sección. Definimos el morfismo de* $\mathbb{K}$*-álgebras*

$$\widetilde{\varphi} : \mathbb{K}[w_1, \ldots, w_n] \longrightarrow \frac{\mathbb{K}[z_1, \ldots, z_m, t]}{J}$$

por

$$\widetilde{\varphi}(w_j) = \left(t^{\alpha_j}\prod_{i=1}^{m} z_i^{a'_{ij}}\right) + J$$

$$\widetilde{\varphi}(g(w_1, \ldots, w_n)) = g(\varphi(w_1), \ldots, \varphi(w_n)) + J.$$

Siendo $J = \langle tz_1 \cdots z_m - 1\rangle \subset \mathbb{K}[\boldsymbol{z}, t]$ *y* φ *el morfismo* $\varphi : \mathbb{K}[\boldsymbol{w}] \longrightarrow \mathbb{K}[\boldsymbol{z}, t]$ *equivalente a* $\widetilde{\varphi}$ *pero sin tomar cociente módulo* J*.*

Entonces $(A_1, \ldots, A_n)$ *pertenece a la región factible de* (5.3) *si y solo si*

$$\varphi(w_1^{A_1} \cdots w_n^{A_n}) + J = t^{\beta} z_1^{b'_1} \cdots z_m^{b'_m} + J.$$

Demostración.

$(A_1, \ldots, A_n)$ está en la región factible de (5.3) si y solo si

$$\prod_{j=1}^{n}\left(\prod_{i=1}^{m} z_i^{a_{ij}}\right)^{A_j} = \prod_{i=1}^{m} z_i^{b_i}$$

que hemos visto que se corresponde en el anillo cociente con

$$\prod_{j=1}^{n}\left(t^{\alpha_j}\prod_{i=1}^{m}z_i^{a'_{ij}}\right)^{A_j}=t^{\beta}\prod_{i=1}^{m}z_i^{b'_i}.$$

Pero como $\widetilde{\varphi}(w_j)=\left(t^{\alpha_j}\prod_{i=1}^{m}z_i^{a'_{ij}}\right)+J$, tomando clases de equivalencia, esto último es equivalente a

$$\left(\prod_{j=1}^{n}(\varphi(w_j))^{A_j}\right)+J=\left(t^{\beta}\prod_{i=1}^{m}z_i^{b'_i}\right)+J$$

$$\left(\varphi(w_1^{A_1}\cdots w_n^{A_n})\right)+J=\left(t^{\beta}z_1^{b'_1}\cdots z_m^{b'_m}\right)+J$$

□

De nuevo en este caso el morfismo φ no tiene por qué ser sobreyectivo. Volvemos a definir los polinomios

$$f_j=t^{\alpha_j}\prod_{i=1}^{m}z_i^{a'_{ij}}\in\mathbb{K}[z_1,\ldots,z_m,t]$$

de modo que la imagen por φ de $\mathbb{K}[w_1,\ldots,w_n]$ será el conjunto de polinomios de (5.6) que podamos expresar como polinomios en $f_1,\ldots,f_n$.

En la proposición siguiente los polinomios $f_1,\ldots,f_n$ en el anillo $\mathbb{K}[\mathbf{z},t]$ pueden ser arbitrarios en cuyo caso consideraremos $\widetilde{\varphi}$ el morfismo análogo al definido en la *Proposición 5.2.1* tomando $\widetilde{\varphi}(w_j)=f_j+J$. De todos modos, aplicaremos dicha proposición en el caso de que los f_j sean los monomios descritos más arriba.

Proposición 5.2.2. *Consideremos $f_1,\ldots,f_n\in\mathbb{K}[z_1,\ldots,z_m,t]$ polinomios arbitrarios. Y sea $\leq$ un orden monomial en $\mathbb{K}[\boldsymbol{z},t,\boldsymbol{w}]$ con la Propiedad de Eliminación respecto del par $((\boldsymbol{z},t),\boldsymbol{w})$. Sea G una base de Gröbner para el ideal*

$$I=\langle tz_1\cdots z_m-1,f_1-w_1,\ldots,f_n-w_n\rangle\subset\mathbb{K}[z_1,\ldots,z_m,t,w_1,\ldots,w_n]$$

y sea $f\in\mathbb{K}[z_1,\ldots,z_m,t]$. Si denotamos $g=\overline{f}^{G}$ entonces,

i) Existe $f'\in\mathbb{K}[w_1,\ldots,w_n]$ tal que $\widetilde{\varphi}(f')=[f]\in\frac{\mathbb{K}[z_1,\ldots,z_m,t]}{J}$ si y solo si $g=\overline{f}^{G}\in\mathbb{K}[w_1,\ldots,w_n]$.

ii) Si se cumple i), $f=g(f_1,\ldots,f_n)$ nos da una expresión de f como polinomio en los f_j (en el anillo $\mathbb{K}[z_1,\ldots,z_m,t]$).

iii) Si f y f_j son monomios para $j=1,\ldots,n$; y $[f]$ está en la imagen por $\widetilde{\varphi}$, es decir se cumple cualquiera de las dos condiciones equivalentes de i), entonces g también es un monomio de $\mathbb{K}[w_1,\ldots,w_n]$.

Demostración.

La demostración es muy similar a la de la *Proposición 5.1.3*, pero hay que hacer algunas modificaciones al trabajar con el anillo cociente. Vamos a incluir un pequeño esbozo de la misma.

i) $\boxed{\Leftarrow}$ Supongamos $[f] \in Im(\widetilde{\varphi})$. Entonces existe $f'(\mathbf{w}) \in \mathbb{K}[\mathbf{w}]$ tal que $[f] = \widetilde{\varphi}(f')$, que es equivalente a que

$$f'(f_1, \ldots, f_n) - f \in J \subset \mathbb{K}[\mathbf{z}, t].$$

Luego se tiene que $f'(f_1, \ldots, f_n) = f + \widetilde{f}(\mathbf{z}, t)(tz_1 \cdots z_m - 1)$. Hacemos la división euclídea de $f'(w_1, \ldots, w_n)$ por los elementos de $w_i - f_i$ de manera que obtenemos:

$$f'(\mathbf{w}) = \sum_{i=1}^{n} q_i(w_i - f_i) + r(\mathbf{z}, t) \;\Rightarrow$$

$$f'(f_1, \ldots, f_n) = r(\mathbf{z}, t).$$

Entonces tenemos

$$f'(\mathbf{w}) = \sum_{i=1}^{n} q_i(w_i - f_i) + f + \widetilde{f}(\mathbf{z}, t)(tz_1 \cdots z_m - 1) \;\Rightarrow$$

$$\overline{f'(\mathbf{w})}^{G} = \overline{f}^{G} \in \mathbb{K}[\mathbf{w}]$$

debido a que estamos trabajando con un orden de eliminación respecto del par $((\mathbf{z}, t), \mathbf{w})$.

$\boxed{\Rightarrow}$ Supongamos que $g = \overline{f}^{G} \in \mathbb{K}[\mathbf{w}]$. Al dividir $f \in \mathbb{K}[\mathbf{z}, t]$ entre $G = \{g_1, \ldots, g_t\}$ obtenemos la expresión $f = h_1 g_1 + \cdots + h_t g_t + g$ donde $h_1, \ldots, h_t \in \mathbb{K}[\mathbf{z}, t, \mathbf{w}]$ y $g \in \mathbb{K}[\mathbf{w}]$.

Como f no depende de $\mathbf{w}$ es

$$f(\mathbf{z}, t) = f(\mathbf{z}, t, \mathbf{w}) = f(\mathbf{z}, t, f_1, \ldots, f_n) = g(f_1, \ldots, f_n)$$

pues $g_i \in I$ y por tanto $g_i(\mathbf{z}, t, f_1, \ldots, f_n) = 0$.

Luego tenemos que

$$f(\mathbf{z}, t) = g(f_1, \ldots, f_n) \in Im(\varphi)$$

$$[f] = [g(f_1, \ldots, f_n)] \in Im(\widetilde{\varphi}).$$

ii) Se obtiene directamente de la prueba de *i)*.

iii) El razonamiento es análogo al usado en *iii)* de la *Proposicion 5.1.3.*

□

Tenemos entonces un resultado análogo al del *Teorema 5.1.5* que nos da la resolución del problema (5.3); es decir, si contamos con un orden monomial adaptado al problema (5.3) tenemos que:

Si $[f] = \left[z_1^{b'_1} \cdots z_m^{b'_m}\right] \in \widetilde{\varphi}(\mathbb{K}[w_1, \ldots, w_n])$, *entonces* $\overline{f}^{G} \in \mathbb{K}[w_1, \ldots, w_n]$ *(que es un monomio en* $\mathbb{K}[w_1, \ldots, w_n]$*) nos proporciona una solución del problema* (5.3)*. En caso contrario el problema no tiene solución.*

5.3. Órdenes monomiales adaptados

En esta sección vamos a denotar por $\mathbb{R}_+$ al conjunto de los números reales no negativos.

Definición 5.3.1. *Sea $\boldsymbol{u} = (u_1, \ldots, u_n) \in \mathbb{R}_+^n$ no nulo y sea $<$ un orden monomial en $\mathbb{K}[x_1, \ldots, x_n]$. Entonces se define la relación de orden $<_{\boldsymbol{u}}$ en los monomios de $\mathbb{K}[x_1, \ldots, x_n]$ de la siguiente forma:*

Dados dos monomios distintos $\boldsymbol{x}^\alpha$, $\boldsymbol{x}^\beta \in \mathbb{K}[x_1, \ldots, x_n]$ con $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ y $\beta = (\beta_1, \ldots, \beta_n) \in \mathbb{N}^n$, diremos que $\boldsymbol{x}^\alpha <_{\boldsymbol{u}} \boldsymbol{x}^\beta$ si $\boldsymbol{u} \cdot \alpha < \boldsymbol{u} \cdot \beta$ o si $\boldsymbol{u} \cdot \alpha = \boldsymbol{u} \cdot \beta$ y $\boldsymbol{x}^\alpha < \boldsymbol{x}^\beta$.

Al vector $\boldsymbol{u}$ lo denotaremos vector de pesos de la relación de orden $<_{\boldsymbol{u}}$.

Proposición 5.3.2. *La relación de orden $<_{\boldsymbol{u}}$ definida anteriormente es un orden monomial.*

Demostración.

Está claro que la relación así definida establece una relación binaria que verifica las propiedades reflexiva, antisimétrica y transitiva.

- $<_{\boldsymbol{u}}$ es un orden total. Dados $\boldsymbol{x}^\alpha$, $\boldsymbol{x}^\beta$ con $\alpha \neq \beta$, se tiene:

  Si $\boldsymbol{u} \cdot \alpha < \boldsymbol{u} \cdot \beta$ entonces $\boldsymbol{x}^\alpha <_{\boldsymbol{u}} \boldsymbol{x}^\beta$.

  Si $\boldsymbol{u} \cdot \alpha > \boldsymbol{u} \cdot \beta$ entonces $\boldsymbol{x}^\alpha >_{\boldsymbol{u}} \boldsymbol{x}^\beta$.

  Si $\boldsymbol{u} \cdot \alpha = \boldsymbol{u} \cdot \beta$ entonces $\boldsymbol{x}^\alpha$ y $\boldsymbol{x}^\beta$ son comparables usando el orden $<$ que es un orden total.

- Sea $\boldsymbol{x}^\alpha <_{\boldsymbol{u}} \boldsymbol{x}^\beta$, y sea $\boldsymbol{x}^\gamma$.

  Si $\boldsymbol{u} \cdot \alpha < \boldsymbol{u} \cdot \beta$ entonces $\boldsymbol{u} \cdot (\alpha + \gamma) = \boldsymbol{u} \cdot \alpha + \boldsymbol{u} \cdot \beta < \boldsymbol{u} \cdot (\beta + \gamma)$ y por tanto $\boldsymbol{x}^\alpha \boldsymbol{x}^\gamma <_{\boldsymbol{u}} \boldsymbol{x}^\beta \boldsymbol{x}^\gamma$.

  Si $\boldsymbol{u} \cdot \alpha = \boldsymbol{u} \cdot \beta$ entonces $\boldsymbol{u} \cdot (\alpha + \gamma) = \boldsymbol{u} \cdot (\beta + \gamma)$ y por tanto $\boldsymbol{x}^\alpha \boldsymbol{x}^\gamma <_{\boldsymbol{u}} \boldsymbol{x}^\beta \boldsymbol{x}^\gamma$ ya que $\boldsymbol{x}^\alpha \boldsymbol{x}^\gamma < \boldsymbol{x}^\beta \boldsymbol{x}^\gamma$ por ser $<$ un orden monomial.

- Para ver que $<_{\boldsymbol{u}}$ es un buen orden vamos a usar un resultado que no hemos incluido en este trabajo pero que podemos encontrar en la biliografía, *(Corolario 6, Sección 4, Capítulo 2, [2])*. Gracias al cual solo tenemos que probar que para todo $\alpha \in \mathbb{N}^n$, $\alpha \neq (0, \ldots, 0)$ se tiene que $1 <_{\boldsymbol{u}} \boldsymbol{x}^\alpha$; lo cual es obvio pues si $\alpha \neq (0, \ldots, 0)$ es $\boldsymbol{u} \cdot \alpha > 0$.

□

Definición 5.3.3. *Sea $\boldsymbol{u}_1, \boldsymbol{u}_2 \in \mathbb{R}_+^n$ no nulos y sea $<$ un orden monomial en $\mathbb{K}[x_1, \ldots, x_n]$. Entonces se define la relación de orden $<_{\boldsymbol{u}_1, \boldsymbol{u}_2}$ en los monomios de $\mathbb{K}[x_1, \ldots, x_n]$ de la siguiente forma:*

Dados dos monomios distintos $\boldsymbol{x}^\alpha$, $\boldsymbol{x}^\beta \in \mathbb{K}[x_1, \ldots, x_n]$ con $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ y $\beta = (\beta_1, \ldots, \beta_n) \in \mathbb{N}^n$, diremos que $\boldsymbol{x}^\alpha <_{\boldsymbol{u}_1, \boldsymbol{u}_2} \boldsymbol{x}^\beta$ si

i) $\boldsymbol{u}_1 \cdot \alpha < \boldsymbol{u}_1 \cdot \beta$, *o bien*

ii) $\boldsymbol{u}_1 \cdot \alpha = \boldsymbol{u}_1 \cdot \beta$ *y* $\boldsymbol{u}_2 \cdot \alpha < \boldsymbol{u}_2 \cdot \beta$, *o bien*

iii) $\boldsymbol{u}_1 \cdot \alpha = \boldsymbol{u}_1 \cdot \beta$, $\boldsymbol{u}_2 \cdot \alpha = \boldsymbol{u}_2 \cdot \beta$ *y* $\boldsymbol{x}^\alpha < \boldsymbol{x}^\beta$.

Proposición 5.3.4. *La relación de orden $<_{\boldsymbol{u}_1, \boldsymbol{u}_2}$ definida anteriormente es un orden monomial.*

Demostración.

La demostración es análoga a la anterior. □

Ejemplo 14. (Orden de Eliminación, [10])

Consideremos el vector $\boldsymbol{u} = (1, \ldots, 1, 0, \ldots, 0) \in \mathbb{R}^n_+$ *donde* $u_k = 1$ *para* $k = 1, \ldots, i$ *y* $u_k = 0$ *para* $k = i+1, \ldots, n$*; y consideremos cualquier orden monomial en* $\mathbb{K}[x_1, \ldots, x_n]$, $<$.

Entonces el orden monomial $<_{\boldsymbol{u}}$ *en* $\mathbb{K}[x_1, \ldots, x_n]$ *tiene la Propiedad de Eliminación pues para todo monomio* $\boldsymbol{x}^\alpha$ *que contenga a alguna de las i primeras variables, y para todo monomio* $\boldsymbol{x}^\beta$ *que solo envuelva a las últimas* $n - i$ *variables es* $\boldsymbol{x}^\alpha >_{\boldsymbol{u}} \boldsymbol{x}^\beta$.

Observemos que nuestra función objetivo es de la forma $\ell(\mathbf{A}) = c_1 A_1 + \cdots + c_n A_n$ y suponiendo que los coeficientes $c_j \in \mathbb{R}_+$, el vector $\mathbf{c} = (c_1, \ldots, c_n)$ define un vector de pesos para el orden monomial $<_{\mathbf{c}}$ en $\mathbb{K}[w_1, \ldots, w_n]$ para cualquier orden monomial $<$ en $\mathbb{K}[w_1, \ldots, w_n]$ fijado de antemano. Además se verifica que $\mathbf{w}^{A'} <_{\mathbf{c}} \mathbf{w}^A$ si $\ell(\mathbf{A}') < \ell(\mathbf{A})$. De manera que podemos incorporar el vector de pesos $\mathbf{c}$ con un orden $<$ con la *Propiedad de Eliminación* en $\mathbb{K}[z_1, \ldots, z_m, w_1, \ldots, w_n]$ para conseguir un orden monomial adaptado.

Resumiendo entonces, sea el problema de Programación Lineal Entera (5.3)

$$\begin{array}{ll} \text{mín} & c_1 A_1 + c_2 A_2 + \cdots + c_n A_n \\ s.a.: & a_{11}A_1 + a_{12}A_2 + \cdots + a_{1n}A_n = b_1 \\ & a_{21}A_1 + a_{22}A_2 + \cdots + a_{2n}A_n = b_2 \\ & \vdots \quad\quad \vdots \quad\quad \ddots \quad \vdots \quad\quad \vdots \\ & a_{m1}A_1 + a_{m2}A_2 + \cdots + a_{mn}A_n = b_m \\ & \mathbf{A} = (A_1, \ldots, A_n) \in \mathbb{N}^n \end{array}$$

donde $a_{ij}, b_i \in \mathbb{N}$ y $c_j \in \mathbb{R}_+$.

Supongamos que estamos trabajando en las condiciones establecidas en todo este capítulo para la resolución de un problema de Programación Lineal Entera. Sea $<$ un orden monomial en $\mathbb{K}[z_1, \ldots, z_m, w_1, \ldots, w_n]$. Sean los vectores de pesos $\boldsymbol{u} = (1, \ldots, 1, 0, \ldots, 0) \in \mathbb{R}^{m+n}_+$ con $u_k = 1$ para $k = 1, \ldots, m$ y el resto de coordenadas nulas; y $\boldsymbol{c} = (0, \ldots, 0, c_1, \ldots, c_n) \in \mathbb{R}^{m+n}_+$. Entonces el orden monomial $<_{\boldsymbol{u},\boldsymbol{c}}$ es un orden monomial adaptado para nuestro problema de Programación Lineal Entera.

Esto último no nos sirve en el caso de que algún $c_j < 0$ pues en ese caso no tenemos un orden monomial de pesos con vector de pesos $\mathbf{c}$ ya que $<_{\boldsymbol{u},\boldsymbol{c}}$ no verificaría la propiedad de ser un buen orden.

Definimos el quasi-grado [1] de las variables de $\mathbb{K}[z_1, \ldots, z_m, w_1, \ldots, w_n]$ de la siguiente forma:

$$gr(z_i) = 1, \text{ con } i = 1, \ldots, m$$

$$gr(w_j) = d_j := \sum_{i=1}^{m} a_{ij}, \text{ con } j = 1, \ldots, n$$

y el quasi-grado total de un monomio como la suma de los quasi-grados de cada una de las variables que aparecen en ese monomio, es decir,

$$gr(\mathbf{z}^\alpha \mathbf{w}^\beta) = \sum_{i=1}^{m} \alpha_i gr(z_i) + \sum_{j=1}^{n} \beta_j gr(w_j) = |\alpha| + \sum_{j=1}^{n} d_j \beta_j.$$

[1] Otros autores llaman peso a lo que aquí llamamos quasi-grado. Por eso usaremos aquí peso como un sinónimo de quasi-grado.

Fijémonos en que podemos suponer cada $d_j > 0$ ya que en caso contrario, si algún $d_j = 0$ la variable A_j no estaría sujeta a ninguna restricción y vimos en su momento (sección 5.1) que podíamos suponer sin pérdida de la generalidad que esto no ocurría.

Vamos a definir a continuación algunos conceptos relacionados con el quasi-grado definido anteriormente.

Definición 5.3.5. *Sea $f \in \mathbb{K}[z_1, \ldots, z_m, w_1, \ldots, w_n]$. Diremos que f es quasi-homogéneo si $f = 0$ o si todos los monomios $\boldsymbol{z}^\alpha \boldsymbol{w}^\beta$ que aparecen en f tienen el mismo quasi-grado total.*

Nota. *Dado un sistema de pesos para las variables $(\boldsymbol{z}, \boldsymbol{w})$ como más arriba, cada polinomio no nulo $f \in \mathbb{K}[\boldsymbol{z}, \boldsymbol{w}]$ puede escribirse de forma única como suma finita de polinomios quasi-homogéneos no nulos. Es decir, $f = \sum_k f_k$, suma finita donde cada f_k es un polinomio quasi-homogéneo no nulo. Esta suma, que es única para $f \neq 0$, se llama la expresión de f como suma de sus componentes quasi-homogéneas.*

Definición 5.3.6. *Un ideal $I \subset \mathbb{K}[z_1, \ldots, z_m, w_1, \ldots, w_n]$ es quasi-homogéneo si $I = (0)$, o si $I \neq (0)$, entonces para todo $f \in I$, escribiendo f como $f = \sum_k f_k$, donde cada componente f_k de f es la suma de los términos de f con quasi-grado total k, se tiene que $f_k \in I$ para todo k.*

Teorema 5.3.7. *Sea $I \subset \mathbb{K}[z_1, \ldots, z_m, w_1, \ldots, w_n]$ ideal. Son equivalentes:*

i) I es un ideal quasi-homogéneo.

ii) Existe un sistema de generadores de I formada por polinomios quasi-homogéneos, es decir, existen $f_1, \ldots, f_s \in \mathbb{K}[x_1, \ldots, x_n]$ quasi-homogéneos tal que $I = \langle f_1, \ldots, f_s \rangle$.

iii) Toda base de Gröbner reducida de I (por base de Gröbner reducida nos referimos a la única base de Gröbner reducida respecto a un orden monomial fijado) está formada por polinomios quasi-homogéneos.

Demostración.

Podemos suponer $I \neq (0)$.

$\boxed{i) \Rightarrow ii)}$ Sea I un ideal quasi-homogéneo. Por el *Teorema de la Base de Hilbert* tenemos que $I = \langle F_1, \ldots, F_t \rangle$ donde los polinomios $F_j \in \mathbb{K}[z_1, \ldots, z_m, w_1, \ldots, w_n]$ no tienen por qué ser quasi-homogéneos. Escribimos cada F_j como la suma de sus componentes quasi-homogéneas, es decir, $F_j = \sum_i F_{ji}$ donde todo polinomio F_{ji} es quasi-homogéneo y pertenece al ideal I por hipótesis.

Sea entonces J el ideal generado por estas componentes quasi-homogéneas F_{ji}. Veamos que $I = J$

$\boxed{I \subset J}$ Pues $F_j \in J$ para todo $j = 1, \ldots, t$ ya que cada F_j es suma de sus componentes F_{ji} que son generadores de J.

$\boxed{J \subset I}$ Ya que todas las componentes $F_{ji} \in I$.

Por lo tanto $I = J$ y el ideal I tiene un sistema de generadores quasi-homogéneos.

$\boxed{ii) \Rightarrow i)}$ Notemos en primer lugar que dados dos polinomios $f, g \in \mathbb{K}[z_1, \ldots, z_m, w_1, \ldots, w_n]$, si los expresamos ambos como sumas de sus componentes quasi-homogéneas, $f = \sum_i f_i$ y $g = \sum_i g_i$, está claro que $f = g$ si y solo si $f_i = g_i$ para todo i ya que $f - g = \sum_i (f_i - g_i)$.

Sean de nuevo $f, g \in \mathbb{K}[z_1, \ldots, z_m, w_1, \ldots, w_n]$ expresados como acabamos de ver $f = \sum_i f_i$ y $g = \sum_j g_j$. Sea

$$h = f \cdot g = \left(\sum_i f_i \right) \cdot \left(\sum_j g_j \right),$$

entonces podemos escribir h como suma de sus componentes quasi-homogéneas $h = \sum_k h_k$ donde $h_k = \sum_{i+j=k} f_i g_j$.

Sea ahora $f \in I = \langle f_1, \ldots, f_s \rangle$ donde los f_i son quasi-homogéneos. Entonces $f = h_1 f_1 + \cdots + h_s f_s$, si escribimos cada h_i como suma de sus componentes quasi-homogéneas, $h_i = \sum_j h_{ij}$ donde cada h_{ij} es quasi-homogeneo no nulo de grado j, obtenemos

$$f = \left(\sum_j h_{1j} \right) f_1 + \cdots + \left(\sum_j h_{sj} \right) f_s =$$

$$= \sum_j (f_1 h_{1j} + \cdots + f_s h_{sj})$$

donde cada $f_i h_{ij}$ es quasi-homogeneo y está en I, luego f tiene todas sus componentes quasi-homogéneas en I por lo que I es un ideal quasi-homogéneo.

$\boxed{ii) \Rightarrow iii)}$ En primer lugar vamos a hacer dos observaciones.

a) Sea $f, f_1, \ldots, f_s$ polinomios quasi-homogéneos. Aplicamos en Algoritmo de División para dividir f entre $F = (f_1, \ldots, f_s)$ y obtenemos la expresión $f = h_1 f_1 + \cdots + h_s f_s + r$ donde r será quasi-homogéneo del mismo quasi-grado total que f pues se forma añadiendo términos de f o términos de p (que era el polinomio que iba apareciendo en cada paso de la división) que también se irá formando por términos con el mismo quasi-grado total que f; y cada h_i también será quasi-homogéneo pues se forma añadiendo términos de la forma $\frac{\mathrm{tl}(p)}{\mathrm{tl}(f_i)}$ que siempre tendrán el mismo quasi-grado total.

b) Sean f, g polinomios quasi-homogéneos y sea

$$S(f,g) = \frac{\mathbf{x}^\gamma}{\mathrm{tl}(f)} f - \frac{\mathbf{x}^\gamma}{\mathrm{tl}(g)} g$$

donde $\mathbf{x}^\gamma = mcm(\mathrm{ml}(f), \mathrm{ml}(g))$. Es fácil observar que $S(f,g)$ también es un polinomio quasi-homogéneo pues todos los términos de f (que tienen el mismo quasi-grado) se multiplicarán por un monomio; análogamente con g. Además si $S(f,g) \neq 0$ todos sus términos tienen el mismo quasi-grado del monomio $\mathbf{x}^\gamma$.

Si $I = \langle f_1, \ldots, f_s \rangle$ con todos los generadores quasi-homogéneos, cuando apliquemos el algoritmo de Buchberger, añadiremos S-polinomios que también serán quasi-homogéneos por b) y así obtendremos una base de Gröbner con polinomios quasi-homogéneos. Cuando a continuación hagamos dicha base de Gröbner una base de Gröbner reducida solo quitaremos algunos elementos y dividiremos otros donde nos quedaremos con el resto que por a) también será quasi-homogéneo. Luego la base de Gröbner reducida que obtengamos siempre estará formada por polinomios quasi-homogéneos independientemente del orden monomial.

$\boxed{iii) \Rightarrow ii)}$ Trivial.

□

Nota. *No toda base de Gröbner de un ideal quasi-homogéneo tiene que tener todos sus polinomios quasi-homogéneos. En concreto, si añadimos a una base de Gröbner del ideal I cualquier polinomio no quasi-homogéneo de I, el nuevo conjunto sigue siendo una base de Gröbner.*

Ahora vamos a incorporar estos resultados a nuestro contexto de orden monomial adaptado para problemas de Programación Lineal Entera. En el siguiente resultado seguiremos considerando el mismo quasi-grado que definimos en los monomios de $\mathbb{K}[\mathbf{z}, \mathbf{w}]$ al principio.

Corolario 5.3.8. *En $\mathbb{K}[z_1, \ldots, z_m, w_1, \ldots, w_n]$ se tiene que:*

i) El ideal $I = \langle f_1 - w_1, \ldots, f_n - w_n \rangle$ es quasi-homogéneo.

ii) Toda base de Gröbner reducida de I (con respecto a un orden monomial arbitrario) está formada por polinomios quasi-homogéneos.

Demostración.

$$gr(f_j) = gr\left(\prod_{i=1}^{m} z_i^{a_{ij}}\right) = \sum_{i=1}^{m} a_{ij} = gr(w_j)$$

Luego todos los generadores de I son polinomios quasi-homogéneos y el resto de la demostración se obtiene directamente del teorema anterior. □

Para encontrar un orden monomial adaptado a nuestro problema de Programación Lineal Entera (5.3) si algún $c_j < 0$ hacemos lo siguiente:

Como $d_j > 0$ para todo j, dado el vector $\mathbf{c} = (c_1, \ldots, c_n)$ procedente de la función objetivo ℓ, siempre existe $\mu \in \mathbb{N}$ suficientemente grande para que el vector

$$(c_1, \ldots, c_n) + \mu(d_1, \ldots, d_n)$$

tenga todas sus coordenadas positivas. Fijamos entonces μ para que esto ocurra y consideramos el vetor $\boldsymbol{u}_1 = (1, \ldots, 1, 0, \ldots, 0) \in \mathbb{N}^{m+n}$ que tiene las primeras m coordenadas 1 y el resto nulas; y el vector $\boldsymbol{u}_2 = (0, \ldots, 0, c_1, \ldots, c_n) + \mu(0, \ldots, 0, d_1, \ldots, d_n) \in \mathbb{R}_+^{m+n}$.

De este modo tenemos $\boldsymbol{u}_1, \boldsymbol{u}_2 \in \mathbb{R}_+^{m+n}$ y podemos definir por tanto un orden monomial de pesos $\boldsymbol{u}_1, \boldsymbol{u}_2$, $<_{\boldsymbol{u}_1, \boldsymbol{u}_2}$ en $\mathbb{K}[z_1, \ldots, z_m, w_1, \ldots, w_n]$ para cualquier $<$ orden monomial en $\mathbb{K}[z_1, \ldots, z_m, w_1, \ldots, w_n]$.

Veamos entonces que $<_{\boldsymbol{u}_1, \boldsymbol{u}_2}$ es un orden monomial adaptado al problema de Programación Lineal Entera (5.3).

Observemos en primer lugar que se cumple la *Propiedad de Eliminación* respecto de $(\mathbf{z}, \mathbf{w})$ pues si tenemos un monomio $\mathbf{z}^{\mathbf{0}}\mathbf{w}^{\beta}$ que solo depende de las variables $w_1, \ldots, w_n$ será $\boldsymbol{u}_1 \cdot (0, \ldots, 0, \beta_1, \ldots, \beta_n) = 0$. Cualquier otro monomio $\mathbf{z}^{\alpha}\mathbf{w}^{\beta}$ que dependa de al menos alguna variable z_i cumplirá que $\boldsymbol{u}_1 \cdot (\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_n) = |\alpha| > 0$ y por tanto será $\mathbf{z}^{\alpha}\mathbf{w}^{\beta} >_{\boldsymbol{u}_1, \boldsymbol{u}_2} \mathbf{z}^{\mathbf{0}}\mathbf{w}^{\beta}$.

Veamos ahora que $<_{\boldsymbol{u}_1, \boldsymbol{u}_2}$ es *compatible con* ℓ. Sean dos monomios $\mathbf{w}^{\mathbf{A}}, \mathbf{w}^{\mathbf{A}'}$ tales que $\varphi(\mathbf{w}^{\mathbf{A}}) = \varphi(\mathbf{w}^{\mathbf{A}'})$ y $\ell(\mathbf{A}) > \ell(\mathbf{A}')$.

Como $\varphi(\mathbf{w}^{\mathbf{A}}) = \varphi(\mathbf{w}^{\mathbf{A}'})$ se tiene

$$\varphi(w_1)^{A_1} \cdots \varphi(w_n)^{A_n} = \varphi(w_1)^{A'_1} \cdots \varphi(w_n)^{A'_n} \quad \Rightarrow$$

$$\left(\prod_{i=1}^{m} z_i^{a_{i1}}\right)^{A_1} \cdots \left(\prod_{i=1}^{m} z_i^{a_{in}}\right)^{A_n} = \left(\prod_{i=1}^{m} z_i^{a_{i1}}\right)^{A'_1} \cdots \left(\prod_{i=1}^{m} z_i^{a_{in}}\right)^{A'_n} \Rightarrow$$

$$\prod_{i=1}^{m} z_i^{\sum_{j=1}^{n} a_{ij}A_j} = \prod_{i=1}^{m} z_i^{\sum_{j=1}^{n} a_{ij}A'_j} \Rightarrow$$

$$\sum_{j=1}^{n} a_{ij}A_j = \sum_{j=1}^{n} a_{ij}A'_j \quad \text{para } i = 1, \ldots, m.$$

Entonces $\boldsymbol{u}_1 \cdot (0, \ldots, 0, A_1, \ldots, A_n) = \boldsymbol{u}_1 \cdot (0, \ldots, 0, A'_1, \ldots, A'_n) = 0$ pues ambos monomios solo dependen de las variables w_j. Comparamos ahora con el vector $\boldsymbol{u}_2$.

$$\mathbf{u}_2 \cdot (0, \ldots, 0, A_1, \ldots, A_n) = (0, \ldots, 0, c_1 + \mu d_1, \ldots, c_n + \mu d_n) \cdot (0, \ldots, 0, A_1, \ldots, A_n) =$$

$$= \ell(\mathbf{A}) + \mu \sum_{j=1}^{n} d_j A_j$$

Análogamente,

$$\mathbf{u}_2 \cdot (0, \ldots, 0, A'_1, \ldots, A'_n) = \ell(\mathbf{A}') + \mu \sum_{j=1}^{n} d_j A'_j$$

Pero como $d_j = \sum_{i=1}^{m} a_{ij}$ tenemos que

$$\mathbf{u}_2 \cdot (0, \ldots, 0, A_1, \ldots, A_n) = \ell(\mathbf{A}) + \mu \sum_{j=1}^{n} A_j \sum_{i=1}^{m} a_{ij} =$$

$$= \ell(\mathbf{A}) + \mu \sum_{i=1}^{m} \sum_{j=1}^{n} A_j a_{ij}$$

Análogamente

$$\mathbf{u}_2 \cdot (0, \ldots, 0, A'_1, \ldots, A'_n) = \ell(\mathbf{A}') + \mu \sum_{i=1}^{m} \sum_{j=1}^{n} A'_j a_{ij}$$

Como por hipótesis tenemos que $\sum_{j=1}^{n} a_{ij}A_j = \sum_{j=1}^{n} a_{ij}A'_j$ para todo $i = 1, \ldots, m$ y $\ell(\mathbf{A}) > \ell(\mathbf{A}')$ entonces es $\mathbf{u}_2 \cdot (0, \ldots, 0, A_1, \ldots, A_n) > \mathbf{u}_2 \cdot (0, \ldots, 0, A'_1, \ldots, A'_n)$ y por tanto $\mathbf{w}^{\mathbf{A}} >_{\boldsymbol{u}_1, \boldsymbol{u}_2} \mathbf{w}^{\mathbf{A}'}$.

Luego ya hemos encontrado por tanto órdenes monomiales adaptados para nuestro problema (5.3) para cualquier función objetivo lineal. Esto termina el desarrollo teórico de este capítulo.

A continuación vamos a ver un ejemplo muy sencillo, que carece de uso práctico, de un problema de Programación Lineal Entera resuelto utilizando Bases de Gröbner que nos permitirá entender mejor el *Algoritmo 3*.

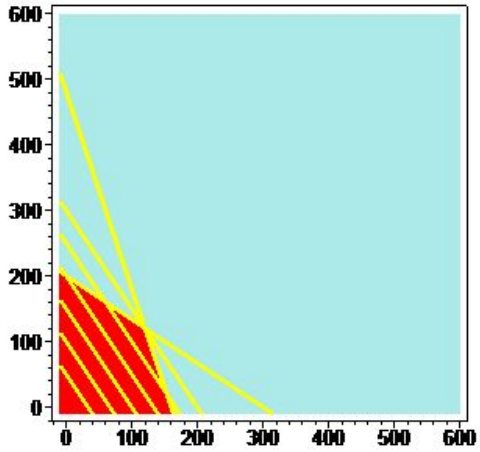
Ejemplo 15. Una determinada empresa fabrica dos tipos de productos P_1 y P_2. El beneficio que se obtiene por cada unidad del producto P_1 es de 3€, y del producto P_2 de 2€. Para producir cada unidad de P_1 se necesitan 2h de trabajo manual y 3h de trabajo con maquinaria, mientras que P_2 requiere 3h de trabajo manual y solo 1h de maquinaria. La empresa solo puede permitirse un total de 600h de trabajo manual semanales y sus máquinas solo están encendidas 480h a la semana. Queremos calcular el número de unidades que debe producirse en la empresa de cada tipo de producto para maximizar los beneficios.

Nuestro objetivo es entonces resolver el problema de Programación Lineal Entera:

$$\begin{array}{rl} \text{máx} & 3n_1 + 2n_2 \\ s.a.: & 2n_1 + 3n_2 \leq 600 \\ & 3n_1 + \ n_2 \leq 480 \\ & n_1, n_2 \in \mathbb{N} \end{array} \tag{5.7}$$

donde n_1 y n_2 representan el número de unidades que se producen de P_1 y P_2 a la semana respectivamente.

En primer lugar veamos una representación gráfica del problema.



Ya en la gráfica podemos observar que todos los vértices de la región factible relajada son puntos con coordenadas enteras, y por lo tanto la solución del problema (5.7) se encuentra en uno de estos vértices. Está claro que la función objetivo alcanza su valor máximo en el punto $(120, 120)$. Luego se deben producir por tanto 120 unidades de cada producto para maximizar los beneficios.

Seguidamente vamos a resolver este mismo problema (5.7) usando bases de Gröbner, pero antes necesitamos expresarlo en forma estándar. Para ello tenemos que escribirlo como el mínimo de cierta función objetivo y donde las restricciones sean de igualdad, luego debemos introducir dos nuevas variables de holgura, una por cada restricción. Obtenemos entonces que el problema (5.7) es equivalente al siguiente.

$$\begin{array}{rl} -\,\text{mín} & (-3n_1 - 2n_2) \\ s.a.: & 2n_1 + 3n_2 + n_3 \qquad = 600 \\ & 3n_1 + \ n_2 + \quad + n_4 = 480 \\ & n_1, n_2, n_3, n_4 \in \mathbb{N} \end{array} \tag{5.8}$$

Ahora si estamos en condiciones de resolver nuestro problema usando para ello el *Algoritmo 3*.

Introducimos las variables z_1 y z_2, una por cada restricción y obtenemos:

$$z_1^{2n_1+3n_2+n_3} = z_1^{600}$$

$$z_2^{3n_1+n_2+n_4} = z_2^{480}.$$

Vamos a trabajar en este ejemplo con el cuerpo de los números racionales $\mathbb{Q}$. Definimos entonces el morfismo de $\mathbb{Q}$-álgebras de la *Proposición 5.1.1.*

$$\begin{gathered}\varphi : \mathbb{Q}[w_1, w_2, w_3, w_4] \longrightarrow \mathbb{Q}[z_1, z_2] \\ \varphi(w_1) = z_1^2 z_2^3 \\ \varphi(w_2) = z_1^3 z_2 \\ \varphi(w_3) = z_1 \\ \varphi(w_4) = z_2 \\ \varphi(g(w_1, w_2, w_3, w_4)) = g(\varphi(w_1), \varphi(w_2), \varphi(w_3), \varphi(w_4))\end{gathered}$$

Entonces tenemos que un punto (n_1, n_2, n_3, n_4) es factible para (5.8) si

$$\varphi(w_1^{n_1} w_2^{n_2} w_3^{n_3} w_4^{n_4}) = z_1^{600} z_2^{480}.$$

Consideremos ahora los polinomios

$$f_1 = z_1^2 z_2^3,\ f_2 = z_1^3 z_2,\ f_3 = z_1,\ f_4 = z_2,\ f = z_1^{600} z_2^{480} \in \mathbb{Q}[z_1, z_2]$$

y el ideal

$$I = \langle f_1 - w_1, f_2 - w_2, f_3 - w_3, f_4 - w_4 \rangle \subset \mathbb{Q}[z_1, z_2, w_1, w_2, w_3, w_4].$$

Necesitamos también contar con un orden monomial adaptado al problema. Buscamos un orden monomial en $\mathbb{Q}[z_1, z_2, w_1, w_2, w_3, w_4]$ adaptado al problema de Programación Lineal Entera (5.8). Como los coeficientes de nuestra función objetivo son negativos, definimos los quasi-grados de las variables como hemos visto anteriormente.

$$\begin{gathered}gr(z_1) = gr(z_2) = 1 \\ gr(w_1) = 2 + 3 = 5,\ gr(w_2) = 3 + 1 = 4,\ gr(w_3) = 1,\ gr(w_4) = 1\end{gathered}$$

El vector de costes de la función objetivo es $\mathbf{c} = (-3, -2, 0, 0)$, luego se trata de buscar $\mu > 0$ tal que

$$\mathbf{c} + \mu(5, 4, 1, 1) \in \mathbb{N}^4.$$

Basta por tanto tomar $\mu = 1$ y obtendríamos $(2, 2, 1, 1) \in \mathbb{N}^4$, y considerar entonces los vectores de pesos

$$\mathbf{u}_1 = (1, 1, 0, 0, 0, 0),\ \ \mathbf{u}_2 = (0, 0, 2, 2, 1, 1).$$

Entonces tenemos que para cualquier orden monomial $<$ en $\mathbb{Q}[z_1, z_2, w_1, w_2, w_3, w_4]$, el orden $<_{\mathbf{u}_1, \mathbf{u}_2}$ es un orden monomial en $\mathbb{Q}[z_1, z_2, w_1, w_2, w_3, w_4]$ adaptado a (5.8). Nosotros vamos a considerar el orden $<_{grevlex, \mathbf{u}_1, \mathbf{u}_2}$.

Aplicamos ahora el *Algoritmo 3.* Para ello calculamos una base de Gröbner, G, del ideal

$$I = \langle f_1 - w_1, f_2 - w_2, f_3 - w_3, f_4 - w_4 \rangle \subset \mathbb{Q}[z_1, z_2, w_1, w_2, w_3, w_4]$$

respecto del orden $<_{grevlex, \mathbf{u}_1, \mathbf{u}_2}$:

$$G = \{-w_2 + w_3^3 w_4, w_4^2 w_2 - w_1 w_3, -w_1 + w_3^2 w_4^3, -w_4 w_2^2 + w_3^4 w1, z_1 - w_3, z_2 - w_4\}.$$

Hemos usado la función `gbasis` de Maple para el cálculo de esta base.

Teníamos que $f = z_1^{600} z_2^{480}$ y calculamos $\overline{f}^G = w_1^{120} w_2^{120} = w_1^{120} w_2^{120} w_3^0 w_4^0$. Luego el vector $(120, 120, 0, 0)$ nos da una solución al problema (5.8).

Por lo tanto tenemos que para maximizar los beneficios hay que producir 120 unidades de cada producto y obtendríamos así unos beneficios de 600 €.

Bibliografía

[1] Atiyah, M. F., Macdonald, I. G., *Introducción al Álgebra Conmutativa*. Reverté, Febrero de 2005.

[2] Cox, D., Little, J., O'Shea, D., *Ideals, Varieties, and Algorithms. An Introdution to Computational Algebraic Geometry and Commutative Algebra*. Segunda Edición, New York : Springer, 2007.

[3] Cox, D., Little, J., O'Shea, D., *Using Algebraic Geometry*. New York, NY : Springer Science+Business Media, Inc., 2005.

[4] W. Adams, W., Loustauau, P., *An Introduction to Gröbner Bases*. Graduate Studies in Mathematics, Volume 3. American Mathematical Society, 1937.

[5] Buchberger, B., Winkler, F., *Gröbner Bases and Applications*. London Mathematical Society. Lecture Note Series 251. Cambridge University Press, 1998.

[6] Apuntes de Programación Matemática. Tercer curso de Grado en Matemáticas, 2013-14. Universidad de Sevilla. (Profesores D. Emilio Carrizosa Priego y D. Justo Puerto Albandoz)

[7] Apuntes de Estructuras Algebraicas. Tercer curso de Grado en Matemáticas, 2012-13. Universidad de Sevilla. (Profesores D. Francisco Javier Herrera Govantes y D. Alberto Castaño Domínguez)

[8] Apuntes de Algebra Conmutativa y Geometría Algebraica. Cuarto curso de Grado en Matemáticas, 2014-15. Universidad de Sevilla. (Profesores D. Antonio Rojas León y D. Alberto Castaño Domínguez)

[9] Apuntes de Algebra, Combinatoria y Computación. Cuarto curso de Grado en Matemáticas, 2014-15. Universidad de Sevilla. (Profesores D. José María Tornero Sánchez y D. Miguel Ángel Olalla Acosta)

[10] Bayer, D., Stillman, M., *A theorem on refining division orders by the reverse lexicographic order*. Duke Math. J. Volume 55, Number 2 (1987), 321-328.

[11] Hartshorne, R., *Algebraic Geometry*. GTM 52. Springer-Verlag, 1977.